(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Remote Sensing of Environment 132 (2013) 13–31

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A land cover map of Latin America and the Caribbean in the framework of the SERENA project☆

Paula D. Blanco [a,*], Rene R. Colditz [b], Gerardo López Saldaña [c], Leonardo A. Hardtke [a], Ricardo M. Llamas [b], Nicolás A. Mari [d], Angeles Fischer [d], Constanza Caride [d], Pablo G. Aceñolaza [e,f], Héctor F. del Valle [a], Mario Lillo-Saavedra [g], Fernando Coronato [a], Sergio A. Opazo [h], Fabiano Morelli [i], Jesús A. Anaya [j], Walter F. Sione [f,k], Pamela Zamboni [f], Victor Barrena Arroyo [l]

[a] National Patagonian Center-Argentinean National Research Council, Terrestrial Ecology Unit, U9120ACD Puerto Madryn, Chubut, Argentina
[b] National Commission for the Knowledge and Use of Biodiversity (CONABIO), Av. Liga Periférico-Insurgentes Sur 4903, Parques del Pedregal, Tlalpan 14010, Mexico City, D.F., Mexico
[c] Department of Forestry, School of Agronomy, Technical University of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal
[d] Climate and Water Institute-INTA, B1712WAA Castelar, Buenos Aires, Argentina
[e] Center of Scientific Research and Technological Transfer to Productivity-Argentinean National Research Council, CP 3105 Diamante, Entre Ríos, Argentina
[f] Autonomous University of Entre Ríos, CP 3100 Paraná, Entre Ríos, Argentina
[g] University of Concepción, Department of Electronic Engineering, 4070386 Concepción, Chile
[h] University of Magallanes, Science and Agricultural Technology School, 621-0427 Punta Arenas, Chile
[i] Center for Weather Forecasting and Climate Studies, INPE, SP 12227-010 Sao José dos Campos, Brazil
[j] University of Medellín, Faculty of Engineering, Medellín, Colombia
[k] National University of Luján, 6700 Luján, Buenos Aires, Argentina
[l] National Agrarian University — La Molina (UNALM), Lima, Peru

ARTICLE INFO

Article history:
Received 1 August 2012
Received in revised form 27 December 2012
Accepted 29 December 2012
Available online xxxx

Keywords:
Land cover classification
Class memberships
Decision trees
MODIS
SERENA project
Latin America
RedLaTIF network
GOFC-GOLD

ABSTRACT

Land cover maps at different resolutions and mapping extents contribute to modeling and support decision making processes. Because land cover affects and is affected by climate change, it is listed among the 13 terrestrial essential climate variables. This paper describes the generation of a land cover map for Latin America and the Caribbean (LAC) for the year 2008. It was developed in the framework of the project Latin American Network for Monitoring and Studying of Natural Resources (SERENA), which has been developed within the GOFC-GOLD Latin American network of remote sensing and forest fires (RedLaTIF). The SERENA land cover map for LAC integrates: 1) the local expertise of SERENA network members to generate the training and validation data, 2) a methodology for land cover mapping based on decision trees using MODIS time series, and 3) class membership estimates to account for pixel heterogeneity issues. The discrete SERENA land cover product, derived from class memberships, yields an overall accuracy of 84% and includes an additional layer representing the estimated per-pixel confidence. The study demonstrates in detail the use of class memberships to better estimate the area of scarce classes with a scattered spatial distribution. The land cover map is already available as a printed wall map and will be released in digital format in the near future. The SERENA land cover map was produced with a legend and classification strategy similar to that used by the North American Land Change Monitoring System (NALCMS) to generate a land cover map of the North American continent, that will allow to combine both maps to generate consistent data across America facilitating continental monitoring and modeling.

1. Introduction

During the past decades Latin America and the Caribbean (LAC) have undergone unprecedented land cover and land use changes. Cropland expansion and forest conversion, accelerated by economic globalization and climate change, are the dominant land-use trends in the region (Grau & Aide, 2008). LAC have lost more forests since 1990 than any other major world region (FAO, 2005), and the rapid conversion of forests to agriculture has been especially evident in areas such as South America's Amazon region and the lowland forests of Central America. Latin America is responsible for 4.3% of global greenhouse gas emissions (Magrin et al., 2007), and of these, 48.3% result from deforestation and land use changes (UNEP, 2000). In this context land use and land cover (LULC) maps are vital for monitoring, understanding, and predicting the effects of complex human–nature interactions. Information on LULC changes needs to be as accurate and timely as possible if it is to be incorporated into management

☆ This article is dedicated to the memory of our beloved friend and colleague Sergio Opazo (November 26, 1978–July 13, 2012).

* Corresponding author. Tel.: +54 2804504143.
E-mail address: blanco@cenpat.edu.ar (P.D. Blanco).

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2012.12.025

and policy decisions. To meet these requirements it is necessary to develop cost-effective ways for automated processing of satellite images and production of LULC maps with high temporal resolution (DeFries & Belward, 2000; Skole et al., 1997).

Traditionally, data from medium to high spatial resolution sensors (10–60 m), e.g. Landsat, Système Pour l'Observation de la Terre (SPOT) and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), have been used for mapping LULC at local to regional scales (Rogan & Chen, 2004). This level of spatial resolution is generally adequate to detect fine-scale (circa 1:100,000) land use patterns. However, potential data costs, the small image extent, haze and cloudiness, and sporadic acquisitions can make data from medium spatial resolution sensors impractical for macro-regional, continental, and global mapping (Asner, 2001; Hansen et al., 2008). Therefore, data with lower spatial resolution (250–1000 m) but high temporal frequency of image acquisition are in many cases selected as an alternative to medium resolution data for large-area land cover mapping, because they are affordable, require less effort for data manipulation and processing, and their near-to-daily coverage increases the probability of cloud-free composites (Arino et al., 2008; Bartholomé & Belward, 2005; Friedl et al., 2002; Hansen et al., 2002; Loveland et al., 2000). Also many classes can only be mapped with high accuracy using time series that characterize vegetation phenology rather than single-date spectral response (Coppin et al., 2003; Friedl et al., 2002, 2010; Hansen et al., 2000).

Various global land cover maps have been produced from low resolution satellites, e.g. 1.1 km Advanced Very High Resolution Radiometer (AVHRR; IGBP DISCover, Loveland et al., 2000; UMD GLCC, Hansen et al., 2000), 1 km SPOT-Vegetation (GLC2000, Bartholomé & Belward, 2005), 500 m and 1000 m Moderate Resolution Imaging Spectroradiometer (MODIS; MOD12Q1, Friedl et al., 2002; MCD12Q1, Friedl et al., 2010), and 300 m Medium Resolution Imaging Spectrometer (MERIS; Globcover, Bicheron et al., 2008). Specifically for the LAC region, a vegetation map for South America (SA) for the year 2000 with twenty-two classes was developed using 1 km SPOT-Vegetation as part of a project to map Global Land Cover for the year 2000 (GLC 2000, Eva et al., 2004), and more recently a land cover map of SA with nine classes was derived from the 300 m MERIS for the years 2008 and 2010 (Hojas Gascon et al., 2012). A study by Clark et al. (2012) mapped annual land-cover data with 8 classes from 2001 to 2010 using 250 m MODIS data, and tracks change patterns of three generalized classes including woody, mix woody/plantation, and agriculture/herbaceous vegetation by linear regression at the municipality level. However, a discrete land cover map of coarse spatial resolution has several disadvantages (DeFries et al., 1995a, 1999; Fernandes et al., 2004). Discrete classes cannot represent spatially complex areas (Hansen et al., 2002), because spatial complexity increases with spatial resolution (Moody & Woodcock, 1994). As a result, homogeneous landscapes e.g. large stands of evergreen broadleaf forest in the Amazon tend to be well classified, but accuracy is poor for mixed pixels that represent small-patch landscapes and transitional zones with various spectral and temporal signals of trees, shrubs, and herbaceous vegetation (e.g. shrub-grass steppes in southern Patagonia; Herold et al., 2008).

The Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD) program is a coordinated international effort, working to provide ongoing space-based and in situ observations of forests and other vegetation cover for sustainable management of terrestrial resources and to obtain an accurate and reliable quantitative understanding of the terrestrial carbon budget (http://www.fao.org/gtos/gofc-gold/overview.html). The creation of regional networks has been encouraged, which provide a mechanism for sharing resources and expertise. Within this framework, the Latin American Remote Sensing and Forest Fires Network (*Red Latinoamericana de Teledetección e Incendios Forestales* — RedLaTIF) was created in 2002. This paper presents the production and analysis of a land cover map for Latin America and the Caribbean for the year 2008 that was part of the Latin American Network for Monitoring and Studying of Natural Resources (*Red Latinoamericana de Seguimiento y Estudio de los Recursos Naturales* — SERENA, web page http://www.proyectoserena.com.ar/). The SERENA project, developed within the RedLaTIF network, aims to monitor, study, and disseminate information associated with biomass burning (wildfires) and changes in land use and cover for LAC utilizing satellite data. The network brings together 18 institutions from 10 countries with 53 researchers. The network envisages the development of common methodologies for developing homogeneous products that build upon and are validated with local data.

The SERENA land cover map for LAC incorporates: 1) the local expertise of SERENA network members to generate training and validation data, 2) a land cover mapping approach based on decision trees using MODIS time series, and 3) estimates of class memberships to partly overcome pixel heterogeneity issues of coarse resolution land cover maps. The suite of SERENA land cover products also includes a discrete map with an additional layer that represents the estimated confidence. The land cover map was recently finished and is available as a wall map. The digital version will be released in the near future and is hosted on the SERENA web page. The SERENA land cover mapwas produced using a legend and classification strategy similar to the North American Land Change Monitoring System (NALCMS) that generated a land cover map for the North American continent (Colditz et al., 2012; Latifovic et al., 2012). The similarity between both maps will facilitate legend harmonization and map combination to form a consistent land cover map from Ellesmere Island, Canada to Tierra del Fuego, Argentina that will be of high value for monitoring and modeling across the American continent.

2. Methodology

The requirements of supervised image classification can be grouped into four broad categories: legend definition, input data generation, sample data preparation, and classifier development. Often post processing steps are needed and accuracy assessment should be an integral part of each mapping exercise. The optimal choice depends on carefully considering the characteristics of each component with respect to the others, the aim of the study and the characteristics and diversity of the study area. The modules of each component are described below and are shown in Fig. 1.

The general framework of the mapping approach to derive class-membership estimates has been described in previous studies, mainly by Colditz et al. (2011) for South Africa and Germany and by Colditz et al. (2012) for Mexico within the NALCMS initiative. Other studies for mapping land cover over broad regions focused on input data transformations and sample data analysis (e.g. Conrad et al., 2011; DeFries et al., 1998; Hansen et al., 2000). For accuracy assessment we followed the guidelines of Strahler et al. (2006) for global maps. Therefore this method description will be brief and focuses on important modifications of established protocols. Due to data processing reasons LAC were divided in two mapping frames, Central America and the Caribbean (CAC) together and South America (SA), that were combined at the Isthmus of Darien. CAC includes all countries from Mexico to Panama and islands in the Caribbean Sea. Generally map generation followed the same guidelines in both mapping frames, and differences are outlined below.

2.1. Legend definition

A legend with three levels was defined based on the FAO/UNEP Land Cover Classification System (LCCS) (Table 1). The SERENA land cover legend is similar to the North American Land Change Monitoring System (NALCMS; Latifovic et al., 2012), that should facilitate the harmonization of both datasets to obtain a continental map. The main difference is the separation of tropical from sub-tropical classes

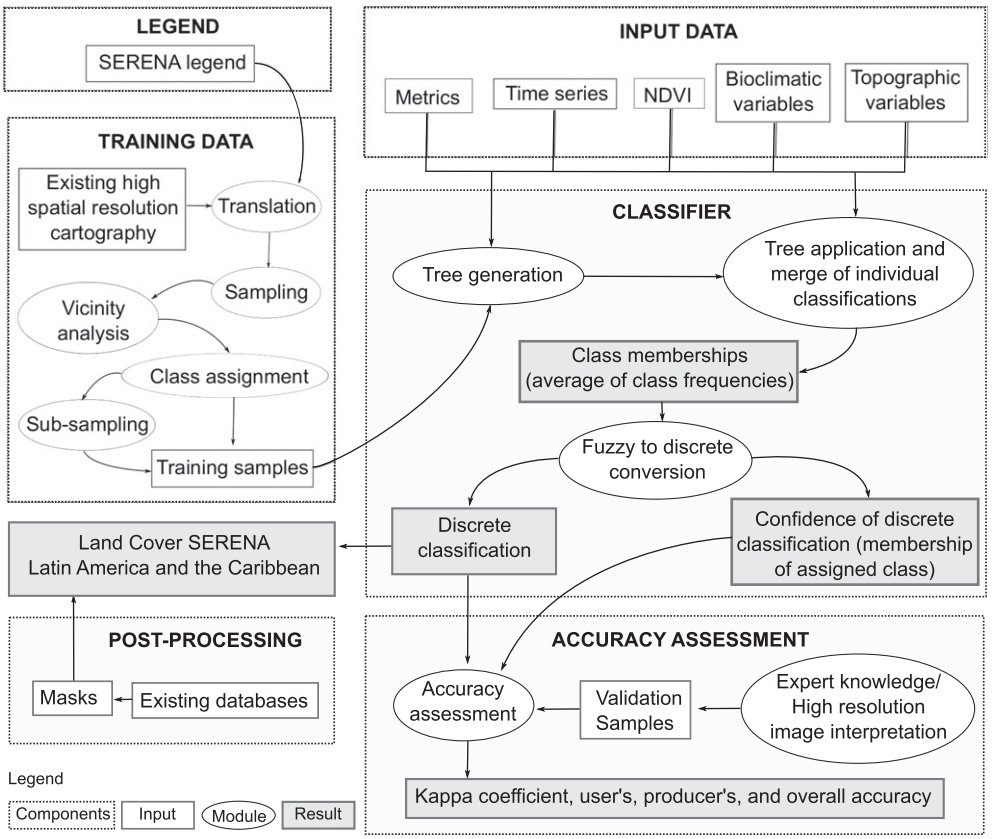


Fig. 1. Process for image classification with classification components and modules.

that were merged in NALCMS. Whereas only a small portion of NALCMS (less than 5%) is located in the tropics and sub-tropics, the study area of SERENA is mainly located in tropical and sub-tropical areas that require separation.

2.2. Preparation of input data

This study employed daily MODIS-Terra surface reflectance data (MOD09GA of collection 5) for the year 2008, downloaded free-of-charge from the NASA Earth Observing System Data Gateway. The MOD09GA product is an estimate of the surface spectral reflectance for seven bands at 500 m spatial resolution, as it would have been measured at ground level if there were no atmospheric scattering or absorption (Vermote et al., 2002). In total 20,069 images (47 tiles, 365 images of 2008 as well as 31 images of December 2007 and January 2009 to improve temporal interpolation) were processed to generate monthly composites for reflective bands (red: 620–670 nm, NIR: 841–876 nm, blue: 459–479 nm, green: 545–565 nm, SWIR1: 1230–1250 nm, SWIR2: 1628–1652 nm, and SWIR3: 2105–2155 nm). Based on the quality of the daily observations determined by the quality assessment science data set (QA-SDS), only pixels that met settings of Table 2 and with a view-zenith angle lower than 45° were kept for further processing. To create monthly composites, for each pixel the day corresponding to the NIR median value was selected for all bands. In the case of no valid information for the entire month, gaps were filled by linear temporal interpolation between other monthly composites. Satellite data preparation is different to NALCMS (Latifovic et al., 2012) with respect to 1) input data selection (MOD02 radiance versus MOD09 surface reflectance), 2) a simpler filtering and compositing approach (compare to Luo et al., 2008), and 3) no downscaling to 250 m (Trishchenko et al., 2006).

The Normalized Difference Vegetation Index (NDVI) was computed for each month from the red and NIR band composites. Additional ancillary information included elevation, slope, and aspect, derived from 90 m SRTM data (USGS, 2004) and bioclimatic variables, downloaded from WorldClim-Global Climate data (Hijmans et al., 2005). WorldClim-Global Climate data were generated through interpolation of average monthly climate data (1950–2000 period) from weather stations on a 30 arc-second resolution grid. Specifically bioclimatic variables include the annual mean temperature (Bio 1), mean temperature diurnal range (Bio 2), temperature seasonality (Bio 4), maximum temperature of warmest month (Bio 5), minimum temperature of coldest month (Bio 6), temperature annual range (Bio 7), annual precipitation (Bio 12), precipitation of wettest month (Bio 13), precipitation of driest month (Bio 14), and precipitation seasonality (Bio 15). The set of ancillary information used in SERENA was the same for CAC and SA, which is different to NALCMS where each country had distinct sets depending on their availability (Latifovic et al., 2012). All satellite data and ancillary layers were referenced to the Lambert Azimuthal Equal Area (LAEA) projection with center at 70°W and 10°S and Sphere datum with a radius of 6370.997 km.

Temporal metrics, i.e. simple statistics calculated for defined periods from the satellite time series, capture the seasonal spectral differences among classes that are useful for class discrimination (Clark et al., 2012; Conrad et al., 2011; DeFries et al., 1995b; Hansen et al., 2000). The mean, standard deviation, minimum value, maximum value, and range between maximum and minimum were computed for each band and the NDVI over the full year, two six-month, three four-month and four three-month periods. Multiple classification runs employed different feature sets of input data that consisted of the monthly composites of spectral bands, NDVI, and ancillary data (96 + 13 variables) or temporal metrics and ancillary data (400 + 13).

Table 1
Legend of SERENA land cover. Class names in italic correspond to stable masks.

Code	Level 1	Level 2	Level 3	Definition
1	Broadleaf forest	Tropical broadleaf forest	Tropical broadleaf evergreen forest	Forests generally taller than 5 m and more than 20% of total vegetation cover. These forests have greater than 75% of tree crown cover represented by evergreen species. This type occupies most of the Amazonian basin and also exists at the coasts of Guyana and Colombia and along the Atlantic shores of Central America.
2			Tropical broadleaf deciduous forest	Forests generally taller than 5 m and more than 20% of total vegetation cover. These forests have greater than 75% of tree crown cover represented by deciduous species. This type occurs on the western flank of Sierra Madre Occidental in Mexico and in isolated spots in Mato Grosso in Brazil.
3		Sub-tropical broadleaf forest	Sub-tropical broadleaf evergreen forest	Same as class 1, but it covers the so-called *"ceja de selva"* along the eastern flank of the Andes from Venezuela to Argentina and the plateaus in southern Brazil where it is called *"mata atlantica"*.
4			Sub-tropical broadleaf deciduous forest	Same as class 2, but this type occurs mostly in the Chaco plains in Paraguay and northern Argentina.
5		Temperate broadleaf forest	Temperate broadleaf evergreen forest	Same as class 1 but with 3 m height. It covers the western flank of southern Andes in Chile and some very isolated spots in the Andean range in Peru.
6			Temperate broadleaf deciduous forest	Same as class 2 but with 3 m height. This type is sparsely present in Sierra Madre Occidental and in the Andean piedmont in northern Argentina.
7	Needleleaf forest	Sub-tropical needleleaf forest	Sub-tropical needleleaf forest	Forests generally taller than 5 m and more than 20% of total vegetation cover. The tree crown cover contains at least 75% of needleleaf species. This type mainly occurs in highlands in Honduras and northwestern Nicaragua.
8		Temperate needleleaf forest	Temperate needleleaf forest	Same as class 7 but with 3 m height. Exists in the Sierra Madre Occidental in Mexico and in the drier sections of southern Andes in Patagonia.
9	Mixed forest	Mixed forest	Mixed forest	Forests generally taller than 3 m and more than 20% of total vegetation cover. Neither needleleaf nor broadleaf tree species occupy more than 75% of total tree cover, but are co-dominant. This type spreads over the Sierra Madre Occidental from the US border to Guatemala and Honduras.
10	Shrubland	Tropical shrubland	Tropical shrubland	Areas dominated by woody perennial plants with persistent woody stems less than 5 m tall and typically greater than 20% of total vegetation. This type is well represented by the "*sertao*" in northeastern Brazil but is also present further inland as well as in the Piura region, the Peru-Ecuador border on the Pacific.
11		Sub-tropical shrubland	Sub-tropical shrubland	Same as class 10 but this type occupies the northern Mexican plateau where is called *"chaparral"* and plains in central Argentina where is known as *"monte-espinal"*.
12		Temperate shrubland	Temperate shrubland	Same as class 10 but with woody stems less than 3 m of height. This type spreads over the dry plains of *Pampa seca* and patchily in the plateaus of Patagonia.
13	Grassland	Tropical grassland	Tropical grassland	Areas dominated by graminoid or herbaceous vegetation generally accounting for greater than 80% of total vegetation cover. These areas are not subject to intensive management such as tilling, but can be utilized for grazing. This type mainly appears in the Orinoco left basin in central Venezuela, covering what locally is known as *"llanos"*.
14		Sub-tropical grassland	Sub-tropical grassland	Same as class 13, but this type exists in the hilly *"cuchillas"* in Uruguay and in Entre-Rios in Argentina.
15	Grassland	Temperate grassland	Temperate grassland	Same as class 14, but this type covers most of Patagonian plateaus and the Andean highlands of the wet Puna in Peru and Bolivia and *"la sierra"* in Peru.
16	*Water*	*Water*	*Water*	Areas of open water, generally with less than 25% cover of non-water cover types. This class refers to areas that are consistently covered by water.
17	*Urban area*	*Urban area*	*Urban area*	Areas that contain at least 30% or greater urban constructed materials for human activities (cities, towns, transportation etc.).
18	*Permanent ice and snow*	*Permanent ice and snow*	*Permanent ice and snow*	Areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover. The two main icefields are located in the southern Patagonian Andes.
19	Barren land	Barren land	Barren land	Areas characterized by bare rock, gravel, sand, silt, clay, or other earthen material, with little or no "green" vegetation present regardless of its inherent ability to support life. Generally, vegetation accounts for less than 10% of total cover. This type occurs in driest areas such in Atacama desert in northern Chile, the coast of Peru, the Puna in the Andean altiplano in high ranges in western Argentina and some areas in central Patagonia.
20	Cropland	Cropland	Cropland	Areas dominated by intensively managed crops. These areas typically require human activities for their maintenance. This includes areas used for the production of annual crops, such as corn, soybeans, wheat, maize, vegetables, tobacco, cotton etc., and also perennial and annual grasses for grazing and woody crops such as orchards, plantations, and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class does not represent natural grasslands used for light to moderate grazing. It spreads all over Cuba, most of Central America from the Mexican lowlands to Panama, in upper Parana basin in Central Brazil and Pampa plains in central-eastern Argentina.
21	Wetland	Wetland	Wetland	Areas dominated by perennial herbaceous and woody wetland vegetation which is influenced by the water table at or near surface over extensive periods of time. This includes marshes, swamps, bogs, mangroves etc. either coastal or inland where water is present for a substantial period annually. The main examples are the upper Paraguay basin in Mato Grosso, locally called *"Pantanal"*, and the Ibera region, in Corrientes in north-eastern Argentina.
22	*Salt flat*	*Salt flat*	*Salt flat*	Flat-floored bottoms of interior desert basins which do not qualify as Wetland. Most of them are in the dry Puna, Atacama desert and north-western Argentina.
23	*Sea water*	*Sea water*	*Sea water*	Zone seaward of the lowest tide limit.
24	*No data*	*No data*	*No data*	Pixels with no satellite data.

2.3. Preparation of sample data for training

A sample data base was built to train the land cover classification algorithm. The methodological steps to obtain the sample data are described below. It should be noted that an extensive sample data base already existed for Mexico (Colditz et al., 2012). That dataset together with additional samples for Central American countries formed the sample set for the CAC map frame.

2.3.1. Reference data sources

National and regional land cover maps at the scale of 1:250,000, mainly based on Landsat 5TM and 7ETM + imagery and aerial

Table 2
Quality flags and settings for high data quality observations.

Quality flags	Quality setting
Cloud state	Clear
Cloud shadow	No
Land/water mask	Land
Aerosol quantity	Climatology/low/average
Cirrus detected	None
Internal cloud algorithm	No cloud
Internal fire algorithm	No fire
MOD35 snow/ice	No
Pixel is adjacent to cloud	No
BRDF correction performed	No
Internal snow algorithm	No

photographs between 1999 and 2006, served as reference data. Most original maps were digitized by experienced local experts and did not undergo formal accuracy assessment, but in the few cases 80 to 85% overall accuracy was estimated. Specifically, national land cover classifications of Perú (INRENA, 2000; 31 classes), Brazil (PROBIO, 2004; between 28 and 81 classes depending on the region), Bolivia (Navarro & Ferreira, 2007; 1903 classes), Central America (World Bank & CCAD, 2000; 197 classes), México (INEGI, 2005; 175 classes), Colombia (IDEAM et al., 2007; 26 classes), Chile (CONAMA, 2002; 127 classes), and regional land cover classifications of Argentina (Ayesa et al., 2002; Guerschman et al., 2003; CRPS INTA, 2004; SADS, 2005; between 5 and 62 classes depending on the region) were used. Regional experts, all members of the SERENA network, translated the maps to the SERENA legend. We merged all translated reference maps, reprojected to the LAEA projection, and gridded to a cell size of 500 m matching the MODIS data. The area covered by this reference map represented 78.8% of the total land to be classified. In the following reference data were needed for three tasks: 1) to obtain training samples distant from land cover boundaries, 2) as an a priori estimate of the expected area of each class, and 3) for comparison to class membership estimates.

2.3.2. Sample selection and class assignment

For SA and Central American countries a random sample strategy was applied with samples at least 5000 m apart and initial class assignment corresponding to the reference map. For Mexico such a sample set already existed from the NALCMS mapping effort (Colditz et al., 2012). Each sample was gridded to 500 m matching all other data sources. Although recent studies for small regions and limited number of classes also showed the potentials of using mixed pixels for classifier training (Foody & Mathur, 2006; Hansen, 2012), most supervised classifications over large regions with many classes build upon homogenous and high-quality training data (Colditz et al., 2012; Friedl et al., 2010). Therefore, all samples that did not coincide with the reference map within a 3×3 kernel (vicinity analysis) were excluded. This approach ensures sampling away from land cover boundaries and increases the likelihood of sampling in homogeneous areas. The required number of at least 100 samples per class could not be extracted in SA for both needleleaf forest classes and temperate broadleaf deciduous forest; in their case heterogeneous samples were also used. On average, this vicinity analysis for sample data reduced the total number of potential samples by approximately 50%. At the same time this automated selection procedure ensures that samples are well-distributed spatially with an appropriate representation of multi-modal frequency distributions of spectrally and temporally diverse classes (Colditz et al., 2011).

For all samples outside Mexico, interpreters (all members of the SERENA network) assigned the final class for classifier training. The local expertise and the availability of high spatial resolution imagery via Google Earth© and other locally available and temporally corresponding higher spatial resolution image sources such Landsat 5TM and 7ETM+, SPOT 5 HRG and the China–Brazil Earth Resources Satellite program (CBERS) were extremely valuable at this step for the correct interpretation of the sample set. It should be noted that this last step in final class assignment makes the training set independent from potential errors in the reference data which were only used for initial class assignment. For Mexico the sample data are based on actual field observations or were digitized from high spatial resolution images and Google Earth© (Colditz et al., 2012) and were not revisited. Table 3 (columns for full set of training samples) shows for each mapping frame (CAC and SA) the number of training samples for every class.

2.3.3. Sub-sampling strategies

Differences in sample size among classes are a controversial issue in supervised image classification. Some studies recommend equalized sample sizes among all classes and others recommend larger sizes for classes with larger area or wider spectral distribution in feature space (Borak & Strahler, 1999; Colditz et al., 2011, 2012; Friedl et al., 2002; Hansen et al., 2000). However, it remains unclear how the sample sizes should be distributed among classes and what objective criteria

Table 3
Number of training samples for each land cover class in CAC and SA, generated by vicinity analysis (full set) and sub-sampling according to expected area.

Code	Class	Full set of training samples		Sub-sampled set of training samples	
		CAC	SA	CAC	SA
1	Tropical broadleaf evergreen forest	2387	8626	875	2020
2	Tropical broadleaf deciduous forest	996	1055	392	200
3	Sub-tropical broadleaf evergreen forest	NA	1094	NA	141
4	Sub-tropical broadleaf deciduous forest	367	650	78	42
5	Temperate broadleaf evergreen forest	NA	470	NA	79
6	Temperate broadleaf deciduous forest	1091	85	272	85
7	Sub-tropical needleleaf forest	150	38	80	10
8	Temperate needleleaf forest	1269	21	221	7
9	Mixed forest	1408	NA	490	NA
10	Tropical shrubland	503	2232	485	610
11	Sub-tropical shrubland	2270	1430	2270	72
12	Temperate shrubland	1354	1249	220	318
13	Tropical grassland	365	530	155	464
14	Sub-tropical grassland	250	441	110	117
15	Temperate grassland	719	534	461	192
19	Barren land	432	1490	40	189
20	Cropland	30,513	2466	2498	468
21	Wetland	622	336	136	83
Total		44,696	22,746	8783	5097

Note: NA…not mapped in this mapping frame.

can be used. For instance, Colditz et al. (2011) suggested relating the sample size of each class to the expected area proportion obtained from a reference map.

In this study the ratio between the samples for each class i and the expected area proportion of that class was calculated (Eq. (1)).

$$ratio_i = \frac{samples_i}{area_i}. \quad (1)$$

The smallest ratio obtained with (Eq. (1)) will ensure sufficient samples for each class. To obtain the number of samples proportional to the expected area, the class with the smallest ratio is multiplied by the area of each class (Eq. (2))

$$samples(proportional)_i = \min(ratio) * area_i. \quad (2)$$

Applying Eqs. (1) and (2), the total number of samples was reduced to 5097 for SA and 8783 for CAC. Table 3 (column for sub-sampled set of training samples) shows the number of samples proportional to the expected area for each class.

For CAC there is an extreme bias of samples for cropland compared to all other samples, and in Colditz et al. (2012) sub-sampling by sample source was deemed useful to limit the cropland extent in the final map. 28,276 cropland samples originate from two sources (COLPOS, 2005; PROCEDE, 2006), focusing on very different regions and types of cropland, and excluding the one or the other limited the samples to 24,921 and 7829, respectively.

2.4. Classifier

The SERENA land cover product was produced using a supervised ensemble classification algorithm. The base algorithm is a decision tree (C5.0; Quinlan, 1993), invoked with 10-folded boosting to improve the accuracy (Freund & Schapire, 1996; Friedl et al., 1999; McIver & Friedl, 2001; Quinlan, 1996; Schapire et al., 1998). Boosting iteratively improves the classifications by assigning weights to classes incorrectly classified in the previous iteration. For the SA mapping frame four runs with different input variables (time series and ancillary data or metrics and ancillary data) and samples sets (vicinity analysis or sub-sampled proportional to expected area, see Table 3) were generated. For CAC eight runs with different feature sets (same as for SA) and sample sets (vicinity analysis, sub-sampled proportional to expected area, exclusion of cropland samples by sources COLPOS or PROCEDE) were produced. For each run, i.e. a specific feature-sample data combination, C5.0 decision trees with 10-folded boosting were generated and applied to the feature set. Classically, the class of the dominating number of samples is assigned to each leaf, but most leaves are not pure, and in particular stopping and pruning were applied to prevent overfitting. Therefore each leaf contains class frequencies (Quinlan, 1996), which can be interpreted as proportional class estimates (Colditz et al., 2012). In a first instance class frequencies of each leaf were stored, and for all boosted trees the proportional estimate of each class was calculated. The class membership estimates for each pixel the proportion of every class. In a second step, membership estimates of all runs (four for SA, eight for CAC) were averaged. Finally, class memberships of all runs were transformed into a discrete map by assigning the pixel to the class with the highest class membership (majority rule). A confidence estimate by pixel for the discrete map is provided, which is defined as the membership (proportion of class) associated with the pixel's assigned majority class in the discrete map.

2.5. Masks generation and map finalization

Classes Water, Permanent ice and snow, Urban area, and Salt flat were generated as masks and superimposed on to the final classification. The water mask was obtained from the Global Raster Water Mask (Carroll et al., 2009) that uses the SRTM Water Body Data (SWBD) in combination with MODIS 250 m data. The permanent ice and snow mask was derived from the Normalized Difference Snow Index (NDSI) (Salomonson & Appel, 2006), which is defined as the difference of reflectance observed in a visible band (MODIS band 4) and a short-wave infrared band (MODIS band 6) divided by the sum of both bands. The urban area mask was derived from the Defense Meteorological Satellite Program–Operational Linescan System (DMSP-OLS) Version 4 Nighttime Lights Time Series of 2008 with 30 arc sec (approximately 1 km) of spatial resolution (NGDC, 2012). We applied a threshold of >86%, which according to Small et al. (2005) is useful to detect lighted built-up areas. Class Pan, Brackish/Saline Wetland was extracted from the Global Lakes and Wetlands Database (GLWD-3, Lehner & Döll, 2004) and labeled Salt flat in the SERENA map.

A minimum mapping unit (MMU) of 100 ha (four connected 500 m MODIS pixels using 8-neighbor rule, i.e. all surrounding cells) was applied to the final discrete map to remove remaining noise patterns. The MMU does not affect the spatial resolution (it remains at 500 m) but defines a minimum area of the smallest patch to be mapped.

For approximately 0.7% of the study area we had no cloud-free observation during the entire period of 2008. Those areas are mainly located at the windward side of the Andes cordilleras in tropical regions of Colombia, Ecuador, and Peru as well as the Patagonian portion of Chile. Still, we were able to classify these pixels quite well, i.e. the class coincided with surrounding pixels with valid satellite observations, despite only the ancillary data provided actual observations (all spectral data contained a fill value). An additional “No data” mask will also be made available to the user community when the digital data will be released.

2.6. Accuracy assessment and membership comparison protocol

The accuracy assessment of the final discrete land cover map was accomplished with an independent set of validation sites. Protocols for assessing the accuracy of discrete maps, computed from the confusion matrix, e.g. user's, producer's, and overall accuracy, have been well established and described in the literature (Congalton & Green, 2009; Foody, 2002; Strahler et al., 2006).

The first step in the generation of the validation sample set was the determination of the sample size per class. In practice, the number of samples is limited by the operational constraints of a study and often represents a compromise between the need to obtain a precise measurement and the requirement to remain efficient and able to process all samples properly.

In this study the sample size was estimated with the multinomial distribution function (Congalton & Green, 2009) resulting in a total of at least 830 samples for CAC and 877 samples for SA (95% confidence, precision of 5%, 16 classes and 39.08% area coverage of largest class in CAC, 17 classes and 45.6% area coverage of largest class in SA). Based on these theoretical estimates 50 sample pixels per class were randomly distributed based on the classified map to assess the accuracy of discrete maps for CAC and SA separately. Fig. 2 shows the spatial distribution of validation samples per class. For each sample a polygon of the size of the corresponding 500 m MODIS pixel was generated. Each sample was analyzed using fine spatial resolution data (30 m Landsat 5TM and 7ETM+, 10 m SPOT 5 HRG, and 1–30 m Google Earth) as close as possible to 2008, and SERENA network experts assigned a primary label. In many cases the heterogeneity of the landscape complicates or impedes the assignation of only one reference label, and in this case the analyst also assigned an alternative label (Latifovic et al., 2012; Sarmento et al., 2009; Zhang & Foody, 1998; Zhu et al., 2000). The error matrix with raw counts was transformed to area-weighted estimates (Card, 1982) from which user's, producer's and overall accuracy were calculated.

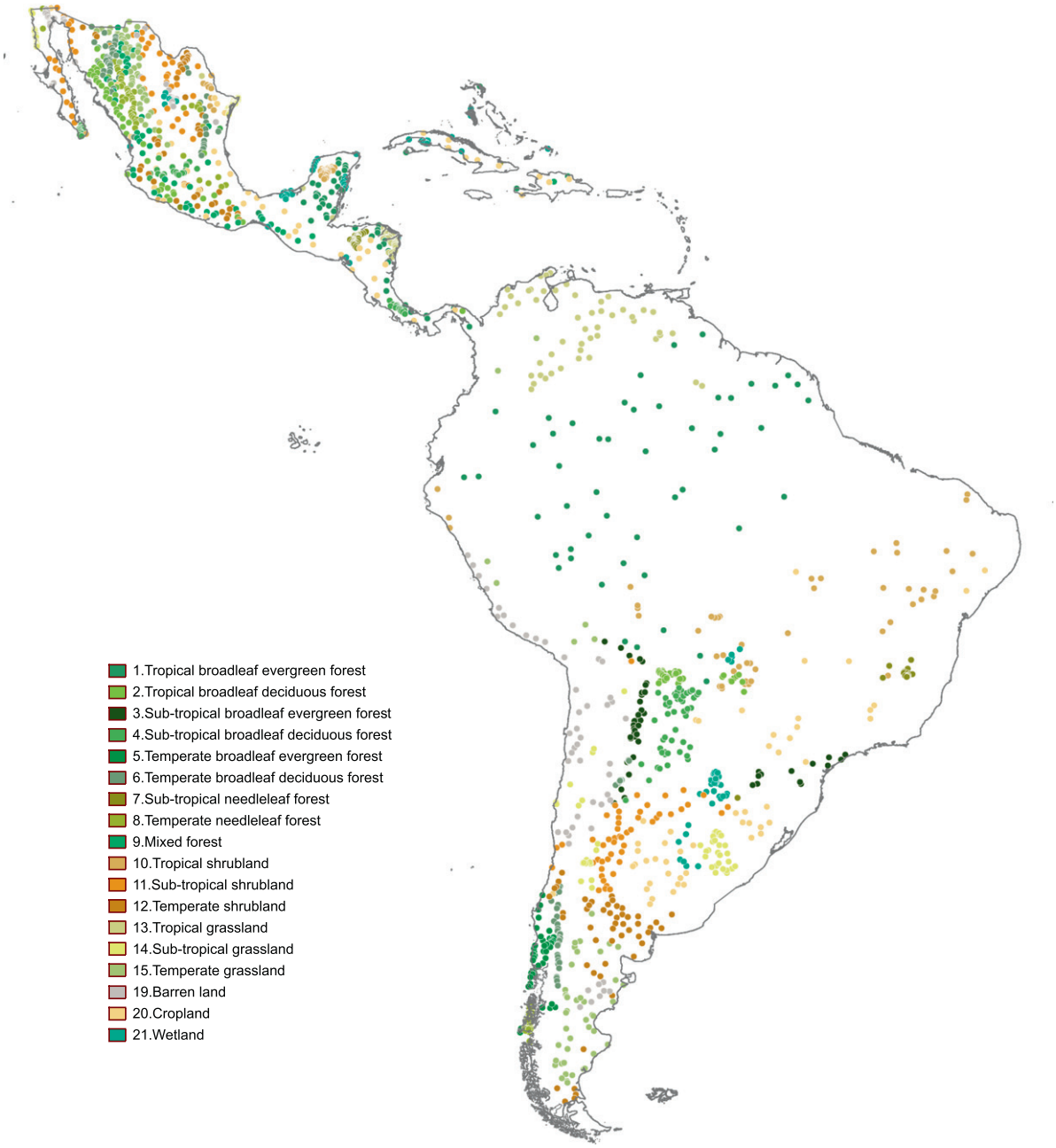


Fig. 2. Spatial distribution of independent validation samples used for accuracy assessment.

For six sites the classification membership estimations were compared to reference memberships. As a baseline, vector maps at the scale of 1:250,000 were rasterized to 25 m cell size matching the grid of the land cover data. Next, for each corresponding 500 m cell the proportion of each class was calculated from the 25 m reference data. This data set served for comparison with the memberships obtained from the classification process. As a difference measure we computed for each pixel the Root Mean Square Error (RMSE) where R is the reference membership, C is the classification membership and K is the number of classes.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{K} (R_i - C_i)^2}{K}}. \quad (3)$$

3. Results and analysis

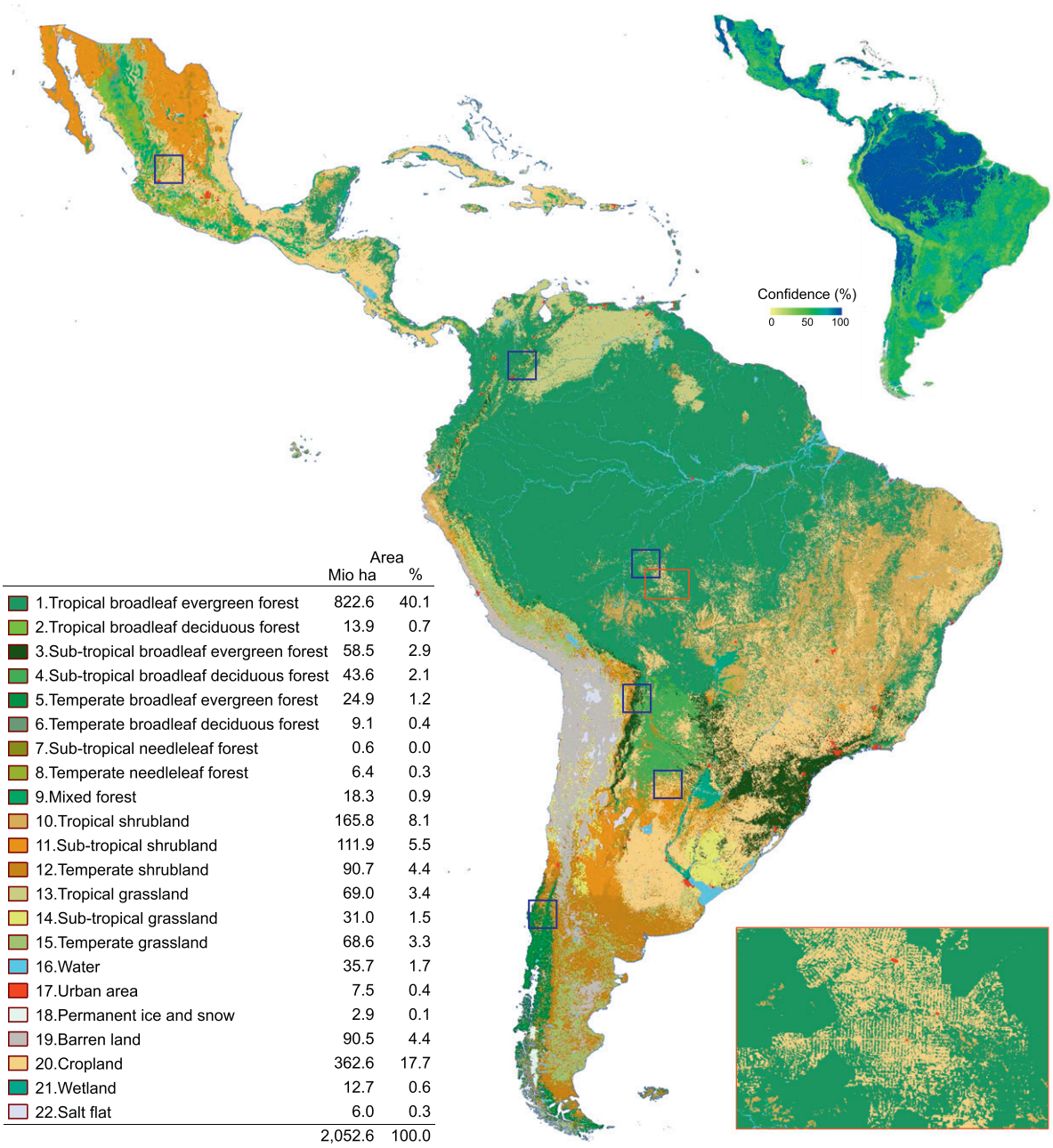
3.1. Discrete land cover map analysis

The classification as a discrete map (majority rule of class memberships, superimposed masks for classes Water, Urban area, Permanent ice and snow, and Salt flats, and MMU of 100 ha) is depicted in Fig. 3. The inset in Fig. 3 shows a close up of deforestation patterns in the southern Brazilian Amazon region that is still visible at 500 m spatial resolution. The site is within the arc of deforestation and indicates the transformation of forest into cropland and rangeland for cattle (Pacheco, 2012).

Fig. 3 presents next to the class names the area proportions in Mio ha (million hectares) and percentage including classes that were superimposed as masks. Dominating land cover classes in LAC are tropical broadleaf evergreen forest (822.6 Mio ha, 40.1%) and cropland

(362.6 Mio ha, 17.7%), followed by tropical and sub-tropical shrubland (165.8 Mio ha, 8.1% and 111.9 Mio ha, 5.5%, respectively). Tropical and temperate broadleaf deciduous forest, needleleaf forests, and wetland are the automatically mapped classes with the smallest representation in the map, occupying less than 1% of LAC. Classes Urban area, Permanent snow and ice and Salt flat, that were superimposed as stable masks, make up less than 0.8% of the total area. The confidence map of the discrete land cover classification (Fig. 3) shows the highest values for the Amazon region, in the coastal plain along the Gulf of Mexico and in northwestern Mexico. Intermediate confidence values were found for the shrubland–grassland–cropland region of eastern Brazil to Patagonia as well as the interior of Mexico and Central America. Lowest confidences are located in the northern and central Andes.

Table 4 (columns for discrete classes) shows the area and percentage of each automatically classified class for the full map and pixels with at least 50% and 75% confidence. The area and percentage are different from Fig. 3, because we only analyze the 18 classes that were mapped by the classification process and without applying a MMU. Fig. 4 depicts for each class the proportional decrease of area with increasing classification confidence. The decrease is expressed in percent with respect to the area of each class at confidence 0 (no confidence threshold). The total area with map confidences equal or higher than 50% and 75% is 79.8% and 51.0%, respectively, indicating that a



	Area Mio ha	%
1.Tropical broadleaf evergreen forest	822.6	40.1
2.Tropical broadleaf deciduous forest	13.9	0.7
3.Sub-tropical broadleaf evergreen forest	58.5	2.9
4.Sub-tropical broadleaf deciduous forest	43.6	2.1
5.Temperate broadleaf evergreen forest	24.9	1.2
6.Temperate broadleaf deciduous forest	9.1	0.4
7.Sub-tropical needleleaf forest	0.6	0.0
8.Temperate needleleaf forest	6.4	0.3
9.Mixed forest	18.3	0.9
10.Tropical shrubland	165.8	8.1
11.Sub-tropical shrubland	111.9	5.5
12.Temperate shrubland	90.7	4.4
13.Tropical grassland	69.0	3.4
14.Sub-tropical grassland	31.0	1.5
15.Temperate grassland	68.6	3.3
16.Water	35.7	1.7
17.Urban area	7.5	0.4
18.Permanent ice and snow	2.9	0.1
19.Barren land	90.5	4.4
20.Cropland	362.6	17.7
21.Wetland	12.7	0.6
22.Salt flat	6.0	0.3
	2,052.6	100.0

Fig. 3. Discrete map of Latin America and the Caribbean (LAC) and map confidence. Note: Classes Water, Urban area, Permanent ice and snow, and Salt flat were superimposed using masks. The area in Mio ha and % were calculated from the LAC map with a minimum mapping unit of 100 ha (4 connected MODIS pixels, 8-neighbor rule). The close-up depicts a zone of forest to agriculture transformation in the arc of deforestation in the southern Amazon of Brazil (see red box). Blue boxes indicate the sites for membership analysis (see Fig. 6). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Area for each land cover class calculated from memberships and discrete map for entire map and pixels with at least 50% and 75% membership or confidence.

Code	Class	Memberships						Discrete classes					
		Entire map		50% membership		75% membership		Entire map		50% confidence		75% confidence	
		Mio ha	%	Mio ha	%	Mio ha	%	Mio ha	%	Mio ha	%	Mio ha	%
1	Tr. broad. ever. f.	765.40	38.25	680.00	52.45	609.74	64.94	819.62	40.96	753.27	47.17	640.99	62.86
2	Tr. broad. decid. f.	60.39	3.02	7.04	0.54	4.24	0.45	14.97	0.75	9.45	0.59	4.95	0.49
3	Sub-tr. broad. ever. f.	63.48	3.17	28.72	2.22	15.95	1.70	59.02	2.95	39.78	2.49	19.20	1.88
4	Sub-tr. broad. decid. f.	44.60	2.23	20.76	1.60	6.89	0.73	43.93	2.20	30.64	1.92	8.56	0.84
5	Te. broad. ever. f.	19.40	0.97	12.60	0.97	6.67	0.71	24.32	1.22	17.41	1.09	8.11	0.80
6	Te broad. decid. f.	15.89	0.79	2.51	0.19	0.27	0.03	9.61	0.48	4.10	0.26	0.34	0.03
7	Sub-tr. needle. f.	4.69	0.23	0.13	0.01	0.00	0.00	0.71	0.04	0.22	0.01	0.00	0.00
8	Te. needle. f.	9.15	0.46	3.53	0.27	1.80	0.19	6.37	0.32	4.94	0.31	2.19	0.21
9	Mixed forest	15.95	0.80	8.26	0.64	2.80	0.30	18.13	0.91	12.19	0.76	3.40	0.33
10	Tr. shrubland	163.05	8.15	73.60	5.68	14.88	1.59	166.01	8.30	112.98	7.08	18.69	1.83
11	Sub-tr. shrubland	119.95	5.99	59.47	4.59	38.65	4.12	112.90	5.64	77.82	4.87	43.43	4.26
12	Te. shrubland	92.27	4.61	37.74	2.91	9.66	1.03	90.87	4.54	58.26	3.65	12.04	1.18
13	Tr. grassland	80.17	4.01	52.27	4.03	44.24	4.71	70.41	3.52	61.66	3.86	48.91	4.80
14	Sub-tr. grassland	46.19	2.31	9.30	0.72	2.83	0.30	31.75	1.59	14.17	0.89	3.50	0.34
15	Te. Grassland	69.42	3.47	20.12	1.55	1.92	0.20	68.09	3.40	33.82	2.12	2.34	0.23
19	Barren land	82.68	4.13	64.52	4.98	53.18	5.66	90.46	4.52	76.30	4.78	58.17	5.70
20	Cropland	316.60	15.82	211.50	16.31	123.66	13.17	360.42	18.01	284.08	17.79	143.28	14.05
21	Wetland	31.84	1.59	4.42	0.34	1.50	0.16	13.52	0.68	5.74	0.36	1.67	0.16
Total		2001.12	100	1296.49	100	938.90	100	2001.12	100	1596.85	100	1019.78	100

Note: tr…tropical, sub-tr…sub-tropical, te…temperate, needle…needleleaf, broad…broadleaf, ever…evergreen, decid…deciduous, f.…forest.

fifth part of the study area is not or only just dominated by the assigned class (i.e., confidence <50%). Some classes show extreme cases of decrease in area for higher confidences (see Fig. 4), including "Temperate broadleaf deciduous forest" (96.4% for the 75% confidence), "Sub-tropical needleleaf forest" (99.9%), and "Temperate grassland" (96.6%). Most of these classes also have a very small total area (Table 4). This suggests that classes with a small area proportion likely also have a lower confidence and reflects the mixed pixel issue. On the other hand, plots in Fig. 4 also show that classes with a moderate area proportion, e.g. barren and tropical grassland, have confidences above 75% for more than 64.3% and 69.5% of their area, respectively, compared to class cropland (three times larger area in the map) with only 39.8%.

3.2. Class membership map analysis

Class memberships of the classification for LAC are displayed in Fig. 5. Areas where the membership is 0 are shown in white, and memberships greater 0 to 100% are displayed in a color bar from yellow to blue. These membership maps show spatially important transitions among classes and zones of heterogeneous landscapes where several land cover classes occur at the same place. These patterns can hardly be identified with a discrete map and even the confidence may only help to some extent. For instance, the memberships in Brazil indicate the transition from tropical broadleaf evergreen forest to tropical broadleaf deciduous forest, that hardly appears in the discrete map, and to tropical shrubland at the Atlantic coast. In the same manner transitions are highlighted towards sub-tropical forest classes in southern Brazil that continue into sub-tropical shrubland in Paraguay, northern Argentina, and Uruguay and eventually temperate shrubland and grassland in Patagonia. Cropland intermingles in many ways with all classes, shown by the widespread spatial distribution of moderate to high memberships, especially in sub-tropical and temperate zones.

Fig. 6 depicts for six sites (for spatial location of sites see blue boxes in Fig. 3) the RMSE between the reference and classification memberships. White areas indicate surfaces like cities or water where no analysis was made. Summary statistics for each site are presented in Table 5. Site Mexico, the region around Guadalajara, was thematically the most

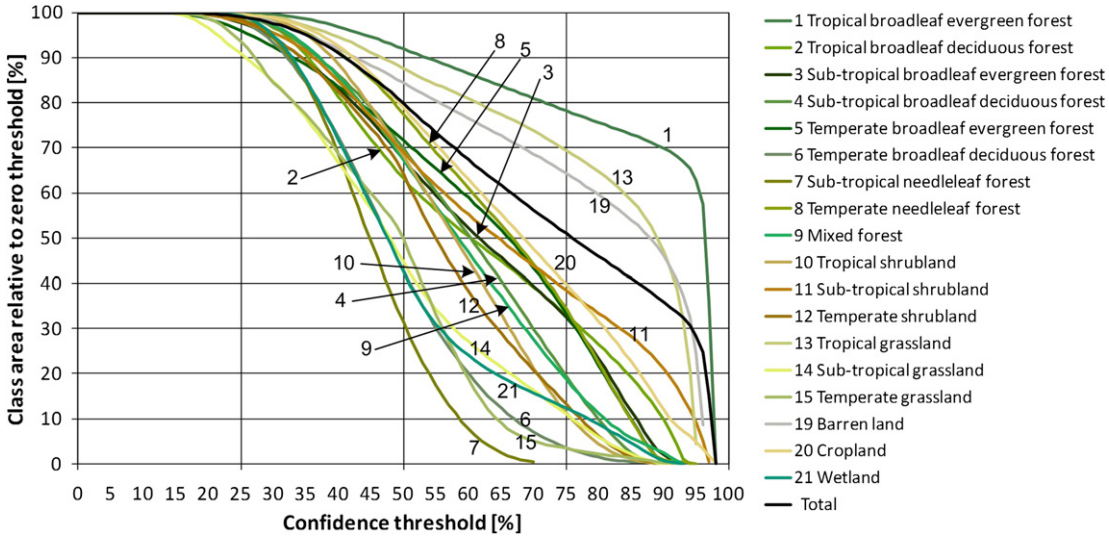


Fig. 4. Area of each land cover class in LAC as a function of the classification confidence. Note: The results are computed for 18 automatically classified land cover classes and area was calculated from the discrete map. The decrease is expressed in percent with respect to the area of each class at confidence 0 (no confidence threshold).

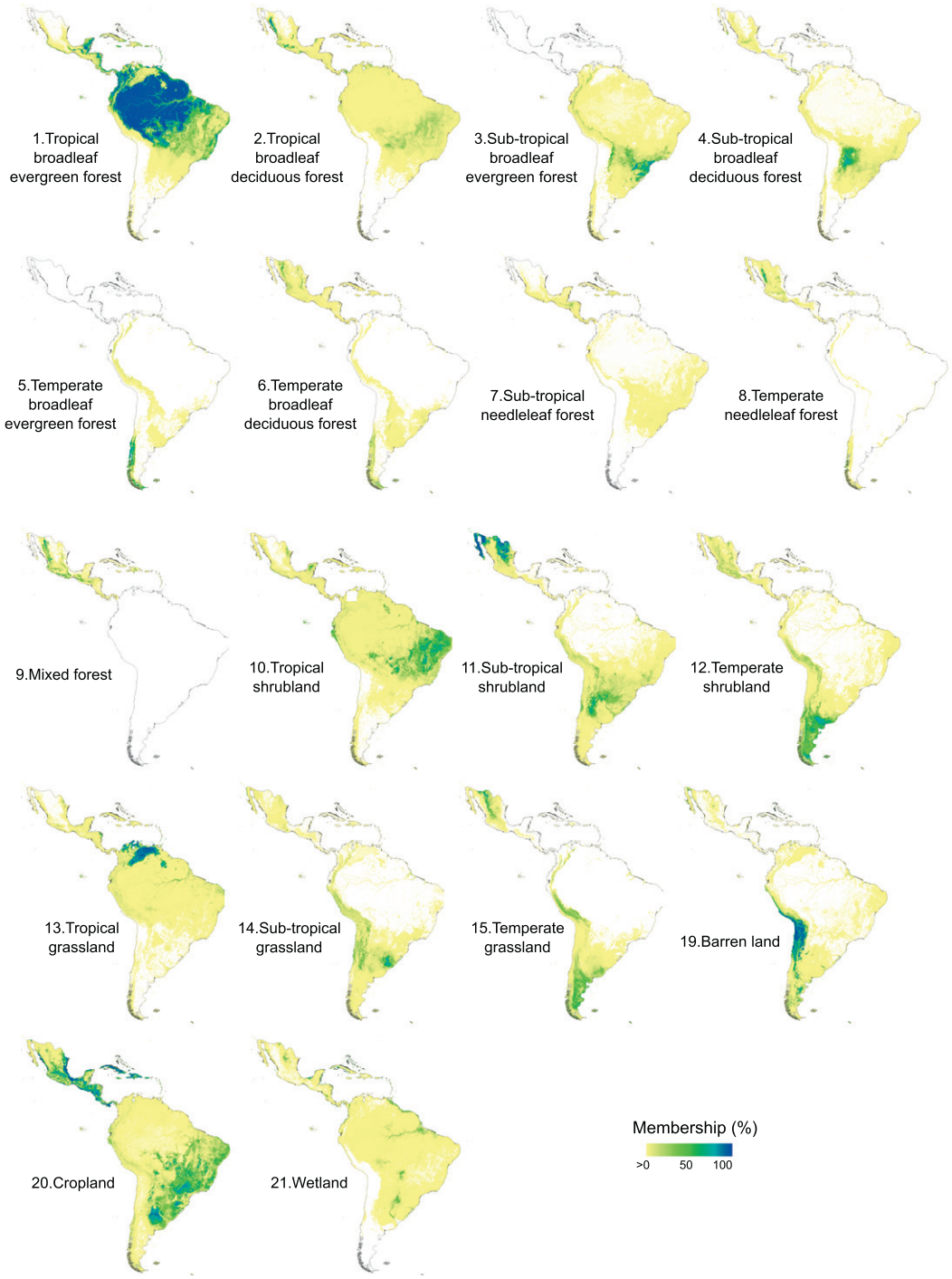


Fig. 5. Membership estimates for each land cover class in LAC. Note: The results show the 18 automatically classified land cover classes. Areas in white indicate membership value 0. Transitions from colors yellow to blue show membership values of greater than 0 to 100%. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

diverse site with 12 land cover classes, and topographically most complex was site Bolivia at the front range of the Andes. Both sites show the highest mean RMSE with values above 15%. The site in northern Argentina was spatially complex with many small patches of different land use. The site in Chile around the city of Concepción indicates that the difference depends on the actual land cover class with higher disagreements for agricultural and pasture land use of the Central Valley and lower differences in the mountainous forests. Similar patterns are shown for the site in Colombia (Northeast of Bogotá) with higher differences in the Eastern Andean range. The lowest differences (mean RMSE of 6.6%) were found for the site in Brazil, where lines of higher differences indicate the deforestation patterns of the tropical broadleaf forest. This analysis of case studies indicates that the classification membership estimates to a high degree the spatial heterogeneity on the ground and can therefore be seen as a measure to describe mixed pixels. As expected, the magnitude of the difference between classification and reference memberships is a function of thematic and spatial complexity of the landscape.

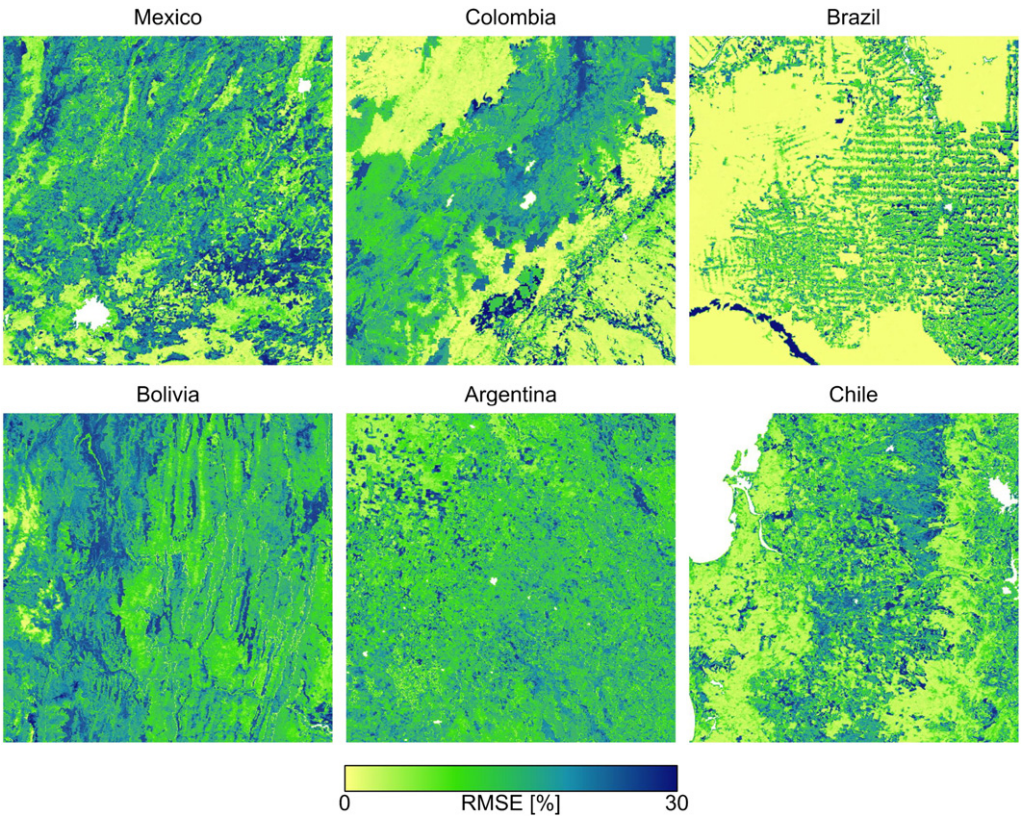


Fig. 6. Root mean square error of each pixel between reference and classification membership estimates. Areas in white indicate pixels with no analysis. For site location see blue boxes in Fig. 3.

Based on the analysis above, it appears that membership is related to the area of each pixel occupied by that cover type. Following this logic, we assumed that for each land cover class the total area is the sum of the membership values times the area of a pixel (Table 4, columns for Memberships; Chen et al., 2010; Fisher, 2010; Fonte & Lodwick, 2004; Leyk & Zimmermann, 2007; Woodcock & Gopal, 2000). There are notable differences when comparing the estimated class area derived from the discrete map and the membership approach (class percentage times pixel area). For instance, tropical broadleaf evergreen forest occupies an area of 819.62 Mio ha according to the discrete classification and only 765.40 Mio ha when calculated from the memberships. The reverse pattern is shown for temperate broadleaf deciduous forest with 9.61 Mio ha in the discrete map and 15.89 Mio ha from the memberships. In general, dominating classes such as tropical broadleaf evergreen forest and cropland present smaller total areas when calculated from memberships. On the other hand, many scarce classes (e.g. temperate broadleaf deciduous forest, both needleleaf forest classes, and wetland) show substantial increments, with class mixed forest as the only notable exception. Considering only class memberships above 50% and 75%, respectively, areas of scarce classes decrease dramatically, which subsequently results in a relative increase in area proportion of dominant classes with relatively more pixels with higher memberships, in particular tropical broadleaf evergreen forest.

Table 5
Statistics of the root mean square error between reference and classification membership estimates for pixels within each test site. For site location see blue boxes in Fig. 3.

	Minimum	Maximum	Mean	Standard deviation
Mexico	0.32	29.70	15.04	7.38
Colombia	0.38	30.48	11.20	7.98
Brazil	0.21	30.51	6.61	7.75
Bolivia	0.61	29.99	15.73	5.26
Argentina	1.01	32.60	13.95	4.86
Chile	0.05	29.21	11.84	7.20

Fig. 7 presents the area as a function of class memberships, thus the largest membership will have the smallest area and the smallest membership the largest. This representation is similar to Fonte and Lodwick (2004) and Leyk and Zimmermann (2007) and is also used in Fisher (2010), but we show the graphs with reversed axes, because the alpha cut (threshold level) is made at the memberships (x-axis, independent variable) that yields area (y-axis, dependent variable). The graphs are similar to Fig. 4 (area as a function of classification confidence) only that a minimum membership is required, and the decrease is expressed in percent with respect to the area of each class at membership 0 (no membership threshold). The graphs show that classes with high memberships such as tropical broadleaf evergreen forest and barren have relatively small decreases in area for higher alpha cuts. On the other extreme there are classes that almost never have high memberships, e.g. sub-tropical needleleaf forest and tropical broadleaf deciduous forest. In this respect the widespread spatial distribution of small memberships for class tropical broadleaf evergreen forest (Fig. 5) does not substantially impact the reduction in (membership) area in Fig. 7, because the high memberships in the Amazon region make up the most of the area calculation. On the other hand, a similar widespread spatial distribution for class tropical broadleaf deciduous forest results in an enormous reduction because even lower alpha cuts reduce a noticeable part of the area thus the total area is substantially impacted by the vast area with low memberships relative to few pixels with high membership estimates. Also, the total area decreased more rapidly when using the area of memberships (64.8% of the study area have a membership above 50%, Fig. 7) compared to confidences (79.8% of the area with a confidence above 50%, Fig. 4).

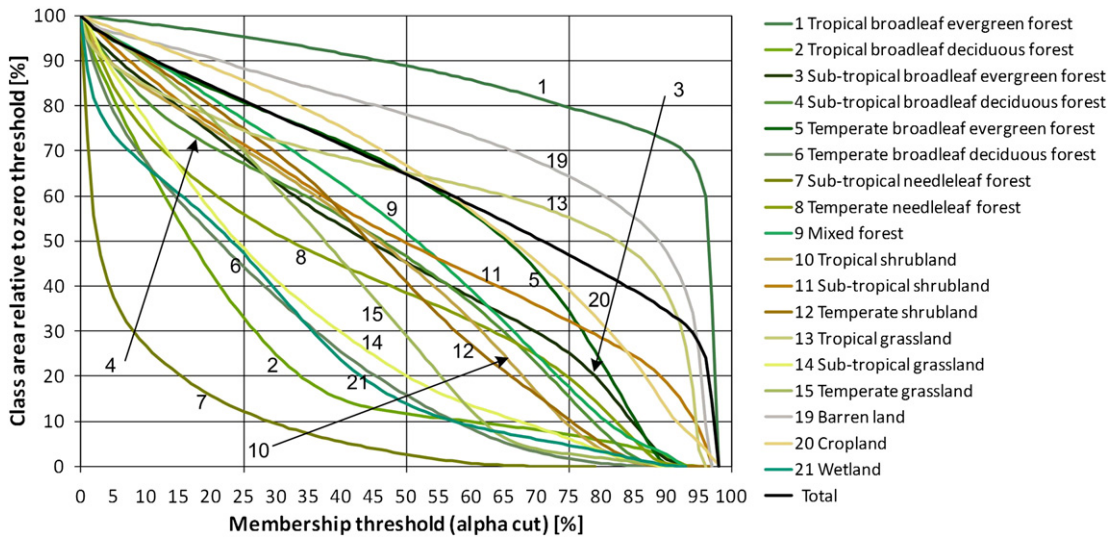


Fig. 7. Area of each land cover class in LAC as a function of membership. Note: The results are computed for 18 automatically classified land cover classes and area was calculated from memberships. The graphs are similar to Fig. 4 (area as a function of classification confidence) only that a minimum membership is required, and the decrease is expressed in percent with respect to the area of each class at membership 0 (no membership threshold).

3.3. Accuracy assessment results

The accuracy assessment of the final discrete land cover map was accomplished with an independent set of validation sites. Table 6 reports the area-weighted error matrix for the primary reference labels and user's, producer's and overall accuracies for all automatically generated classes (classes that were not superimposed as masks in post processing) of the discrete maps for CAC, SA, and LAC.

Accuracies for LAC were derived by combining the raw validation sets of CAC and SA and calculating an area-weighted matrix for the LAC mapping frame. Accuracies from that error matrix for the primary reference labels yield for LAC an overall accuracy (OA) of 84.2% and a Kappa coefficient of 0.80. Broadleaf forest classes depict confusion among each other, in particular between evergreen and deciduous vegetation, and additional confusion with mixed forest, tropical shrublands, and cropland. The confusion with cropland is augmented by the high area proportion of that class, which increases weights of its misclassified samples. This is particularly evident for sub-tropical coniferous forest with a producer's accuracy of only 5% although only three samples were misclassified as cropland. Shrublands and grasslands show class-specific accuracies above 70%. Confusion is mainly among each other with errors between shrubland-grassland for the same climatic region (tropical, sub-tropical, or temperate) and cropland. Due to the different spectral responses, barren was mapped with accuracies above 84%. Cropland also shows reasonable class-accuracies above 77%. Wetlands depict a low producer's accuracy of 24% due to some misclassifications with tropical broadleaf evergreen forest in the Amazon.

The overall accuracies (Kappa coefficients) of the primary reference labels for the individual mapping frames of the SA and CAC are 81.4% (0.76) and 81.1 (0.77), respectively. In the confusion matrix of SA many of the patterns described above were repeated, including the confusion among tropical broadleaf forests and with tropical shrubland and wetland, as well as the confusion of shrubland and grassland classes within their climatic zone. Besides general patterns described for LAC, in CAC there are notable confusions between tropical and sub-tropical deciduous broadleaf forests and between temperate deciduous broadleaf forest and mixed forest. Cropland shows confusion with all higher biomass classes (tropical broadleaf forests, subtropical needleleaf forests, and tropical and sub-tropical shrubland and grassland). In case of ambiguity in reference label assignment the alternative label was also considered for correctly classified samples. In this case the overall accuracies increased to 85.1% (CAC), 91.8% (SA), and 91.2% (LAC).

Fig. 8 illustrates the OA for the primary reference label as function of map confidence for validation sites. For LAC, OA starts to increase for map confidences higher than 15%. For a map confidence of at least 50% the OA reaches 85.4% (81.9% of validation data) and for the 75% confidence threshold 88.8% (39.5% of validation data). Similar plots for the individual mapping frames of CAC and SA shows 81.4% and 82.2% overall accuracy for 50% the confidence threshold and 86.5% and 83.5% for at least 75% confidence. The unsteady curve for higher confidences is due to the smaller validation samples size, for that a few incorrectly classified but high confidence samples have a higher impact and cause variability of the OA.

4. Discussion

4.1. Comparison to other available products

The map presented in this paper reveals the status of land cover in Latin America and the Caribbean for the year 2008. It provides detailed information with 22 thematic classes that were mapped at 500 m spatial resolution and took into account spectral, temporal, and ancillary information. It therefore can be seen in a line with historic and on-going global mapping projects such as GLC2000 and the South America mapping region (Bartholomé & Belward, 2005; Eva et al., 2004) with 22 classes based on 1 km SPOT VGT, the annual MODIS land cover maps (MOD12, Friedl et al., 2010) with 17 classes (IGBP legend) at 500 m, and Globcover of 2005/06 and 2009 (Bicheron et al., 2008) with 22 classes based on 300 m MERIS data. Other recent land cover studies at the extent of South America with 300 m MERIS (9 classes; Hojas Gascon et al., 2012), Central America with 500 m MODIS (9 classes; Giri & Jenkins, 2005), and Latin America and the Caribbean with 250 m MODIS (8 classes; Clark et al., 2012) derived maps at the same or finer spatial resolution but do not share the same thematic detail. Besides the specific aims of these studies, a reason for less classes may be also related to less spectral information by only using the EVI (Enhanced Vegetation Index), red, near infrared, and one short-wave infrared band of MOD13 data (Clark et al., 2012) and the FAPAR (Fraction of Absorbed Photosynthetically Active Radiation), and four spectral bands in the red, red edge and near infrared

Table 6
Area-proportional error matrices derived from the primary label for mapping frames Central America and the Caribbean (CAC, upper row), South America (SA, middle row), and Latin America and the Caribbean (LAC, bottom row).

Mapped class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	19	20	21	Users
1 Tr. broad. ever. f.	10.9		NA		NA														100
	37.4	0.9							NA	1.8							1.8	3.6	82.0
	37.3	0.4								0.8							0.8	1.6	91.0
2 Tr. broad. decid. f.		3.3	NA	0.7	NA	0.1				0.2	0.2								74.0
	0.0	0.1							NA	0.0							0.0	0.0	62.0
	0.1	0.5		0.1		0.0				0.0							0.0	0.0	68.0
3 Sub-tr. broadl. ever. f.	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
			2.3	0.9					NA		0.1						0.1		68.0
			2.0	0.8							0.1						0.1		68.0
4 Sub-tr. broad. decid. f.		0.0	NA	0.7	NA	0.0			0.0	0.0							0.0		90.0
			0.1	2.2					NA								0.0		92.0
		0.0	0.1	2.0		0.0			0.0	0.0							0.0		91.0
5 Te. broad. ever. f.	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
					1.1	0.0		0.1	NA			0.2					0.0		80.0
					1.0	0.0		0.0				0.1					0.0		80.0
6 Te. broad. decid.f.			NA	0.2	NA	2.1								0.0					90.0
					0.0	0.1			NA			0.0					0.0		54.0
				0.0	0.1	0.3						0.0		0.0			0.0		72.0
7 Sub-tr. needle.f.			NA		NA		0.3												100
				0.0			0.0		NA		0.0						0.0		50.0
				0.0			0.0				0.0						0.0		87.9
8 Te. needle.f.			NA		NA			2.2			0.1				0.0				94.0
								0.0	NA			0.0			0.0				52.0
								0.2			0.0	0.1			0.0				73.0
9 Mixed f.			NA	0.1	NA	1.2		0.5	4.5		0.1	0.1	0.1						66.0
	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
				0.0		0.2		0.1	0.6		0.0	0.0	0.0						66.0
10 Tr. shrubland	0.3		NA		NA					0.5			0.0						64.0
	0.4	0.8							NA	7.7			0.2			0.2	0.2		82.0
	1.6	0.3								6.1			0.2			0.1	0.1		73.0
11 Sub-tr. shrubland			NA		NA						21.3			0.5		0.9			94.0
				0.1					NA		2.8			0.1			0.1		92.0
				0.1							5.2			0.1		0.1	0.1		93.0
12 Te. shrubland			NA	0.1	NA	0.2					0.3	2.0							80.0
						0.2			NA			3.8			0.4	0.4	0.1		78.0
				0.1		0.2					0.2	3.6			0.2	0.2	0.0		79.0
13 Tr. grassland			NA		NA						0.0		0.4						98.0
									NA	0.7			2.9			0.1	0.2	0.1	72.0
										0.3	0.0		3.0			0.0	0.1	0.0	85.0
14 Sub-tr. grassland			NA		NA						0.0			0.1				0.0	76.0
									NA		0.3			1.4		0.1			80.0
											0.2			1.2		0.0		0.1	78.0
15 Te. grassland			NA	0.1	NA						0.3	0.2		0.1	3.2				82.0
					0.1				NA			0.9			1.9	0.4			58.0
				0.0	0.0						0.1	0.6		0.0	2.4	0.2			70.0
19 Barren land			NA		NA											0.6		0.0	98.0
									NA	0.2	0.1					4.7	0.1		92.0
										0.1	0.0					4.3	0.0	0.0	95.0
20 Cropland	2.3	0.8	NA		NA	0.8	2.3		0.8	0.8	1.6		1.6		0.8		27.4		70.0
			0.9	0.3	0.3				NA		0.3	0.3				0.3	12.4		84.0
	0.5	0.2	0.5	0.2	0.2	0.2	0.5		0.2	0.2	0.5	0.2	0.4		0.2	0.2	13.9		77.0
21 Wetland	0.0		NA		NA						0.1		0.0		0.0	0.1	0.0	1.5	84.0
			0.0						NA		0.0			0.0				0.4	86.0
	0.0		0.0								0.0		0.0	0.0	0.0	0.0	0.0	0.6	85.0
Producers	80.4	80.8	NA	35.4	NA	48.7	10.2	80.4	85.0	35.3	88.9	84.4	20.4	19.1	78.7	37.3	99.8	97.8	**81.1**
	98.9	5.9	68.3	63.1	73.4	31.4	100	3.1	NA	73.7	79.3	72.6	93.8	95.3	83.2	76.8	81.9	10.4	**81.4**
	94.4	35.0	76.0	60.9	78.2	35.7	5.5	65.8	74.7	80.5	79.7	77.6	84.4	88.7	86.3	83.6	90.9	24.0	**84.2**

Note: Abbreviations: tr…tropical, sub-tr…sub-tropical, te…temperate, needle…needleleaf, broad…broadleaf, ever…evergreen, decid…deciduous, f…forest. Cells with no samples have zero value. Cells with 0.0 in the matrix indicate a very small proportion. NA indicates classes that were not mapped in this mapping frame. The overall accuracies for CAC, SA, and LAC are shown in bold in the lower right corner.

(Hojas Gascon et al., 2012). Giri and Jenkins (2005) only used three dates of images that may not be sufficient to capture the variety of phenological patterns in this diverse region.

The use of ancillary information (elevation, precipitation, and temperature) helped to improve the final results. Tests from the NALCMS study over Mexico (Colditz et al., 2011) and also initial tests for South America showed substantial improvements of the accuracy in the range between 10 and 15%. This is different from Clark et al. (2012) who explored elevation and slope auxiliary variables in classifications and found no improvement. We also found that ancillary data result in locally more stable results and mapping small areas, where no satellite data were available (remaining clouds over entire compositing period), to be spatially consistent with the surrounding area. We are aware of the potential risk that ancillary data imply for change detection, because these variables are used by the classifier but do not change over the years. We recommend a

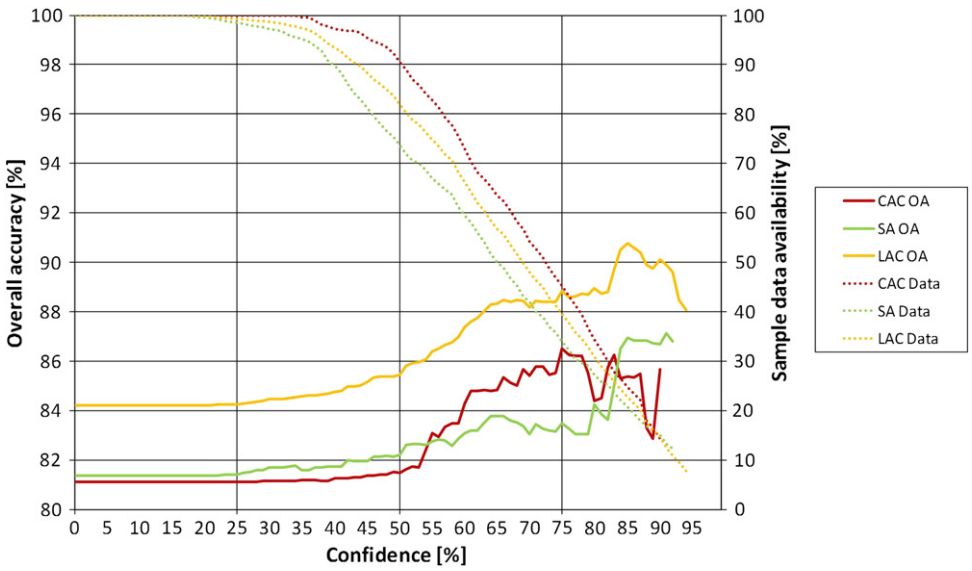


Fig. 8. Overall accuracy as a function of the classification confidence for mapping frames CAC, SA, and LAC. Note: OA…Overall accuracy, Data…Available sample data, CAC…Central America and the Caribbean, SA…South America, LAC…Latin America and the Caribbean.

careful evaluation if ancillary data help to improve the map in terms of accuracy and spatial coherence and their potential negative impacts on detecting change.

A difference to many other mapping projects (Bicheron et al., 2008; Clark et al., 2012) is also that we did not use a regionalized approach. Mapping zones are frequently used to maximize spectral discrimination between areas with comparatively uniform ecological characteristics. That helps the classifier to obtain better results for that region and this, in turn, should result in a more accurate map (Bauer et al., 1994; Homer et al., 1997, 2004; Lillesand, 1996; Reese et al., 2002). Regionalization, however, may also complicate mapping, because some classes may not be mapped in all regions (although they exist) due to insufficient training data. Additional difficulties arise in land-cover mismatch along the boundaries between mapping zones, producing false edges in the final land cover map (Clark et al., 2012; Colditz et al., 2012). Experiences from the NALCMS project for mapping Mexico (Colditz et al., 2012), one of the ecologically most diverse countries of the world, showed little improvement obtained from regionalized classifications. In fact, mapping was more difficult due to the issues described above, and initial tests for mapping CAC resulted in similar problems. The decision to divide the study area into two map frames (CAC and SA) was owed to practical reasons, because the individual map frames are smaller, and thus more manageable and faster to process. The workload was divided between the available computing facilities in Mexico (National Commission for the Knowledge and Use of Biodiversity, CONABIO) and Argentina (National Institute of Agricultural Technology, INTA and the Patagonian National Research Center, CENPAT). In this manner the project also took advantage of local expert knowledge in each sub-region and limited traveling costs by still ensuring a consistent continental product following common mapping guidelines. Merging the land area was applied at the Isthmus of Darien along the border between Panama and Colombia. In this shortest-possible transect the area is mostly tropical evergreen broadleaf forest, which is easily mapped, and the landscape has less class heterogeneity, resulting in no visible false land cover boundaries.

It should be noted that class mixed forest was only mapped for CAC and classes sub-tropical broadleaf evergreen forest and temperate broadleaf evergreen forest only for SA. The restriction of these classes to only a portion of the study area results from their too small spatial extent to collect sample data with sufficient confidence. Similar issues were noted in SA for class sub-tropical needleleaf forest with 21 training samples (see Table 3) that resulted in only 425 ha and temperate needleleaf forest with 38 training samples and 48,925 ha. Albeit these classes are rather scarce in SA clearly more training sites are needed to appropriately represent those classes in feature space (Hansen et al., 2000). Other classes with little spatial extent (urban area, permanent ice and snow, salt flats) were superimposed with masks that were derived from other data sets. This is a common approach for large-area mapping (Friedl et al., 2010) and limits confusion with spectrally similar classes such as barren or sparse grassland.

4.2. Discrete map and map confidence

At first sight the distribution of thematic classes shows the expected spatial patterns with tropical forests in the Amazon, tropical grasslands in the Orinoco basin north of the Amazon, and tropical shrublands in the Brazilian Cerrado. At increasing latitudes, the proportion of sub-tropical classes increases and finally temperate forests appear in southern Chile and higher altitudes in northern Mexico as well as temperate shrubland and grassland or mixture of both in Patagonia. At the same time, the map shows the hypsometric gradient (strongly correlated with rainfall gradients), at the wind-ward side of the central Andes in Peru and Bolivia, for instance, or in the Sierra Madre Oriental and Occidental in Mexico. Also, some specific land use patterns become evident as the fish-bone pattern of deforestation areas in Rondonia in western Brazil.

Despite good accuracies the area proportion and spatial distribution of two classes, cropland in CAC and barren in SA, need further discussion. Cropland in CAC makes up 38% of the area and is the dominating class in this map frame. Large portions of the Mexican Gulf coast, the pacific and central parts of Central American countries as well as the islands of Cuba and Hispaniola were mapped as cropland. A regional map of Mesoamerica (Mexico to Panama, Giri & Jenkins, 2005) classified 21.8% of that area as cropland and in the NALCMS map of Mexico (Colditz et al., 2012) 20.2% were mapped as cropland, compared to 34.4% and 28.9% for the respective areas in this study. One reason for an 8.7% higher estimate of cropland area for Mexico is likely related to the increased spatial resolution (250 m in NALCMS) that favors the classification of a diverse class with multiple distributions in feature space which was well described with many training data. Thus biased sampling was noted in many studies as a

potential reason for overestimation of one class (Colditz et al., 2011, 2012; Hansen et al., 2000). For CAC, eight runs were combined of which only two showed a balance between the number of samples and the expected area. In particular these balanced runs also showed a smaller area proportion of cropland with 30.2% for CAC. However, the bias of samples among classes towards cropland is only one part of the explanation. The other reason is the understanding of cropland samples that originated from the NALCMS data base (the vast majority of samples for CAC). Cropland comprises managed areas that need human activities for maintenance including all forms of herbaceous cover used for grazing, including perennial and annual grasses (Table 1). In particular, the latter included rangelands into class cropland that in other maps is usually assigned to grassland and open shrubland. In this respect, results approximately match for comparisons of this map to other maps for Central America that combine cropland, grassland and partly also open shrubland. It is also interesting to note specific spatial patterns, for instance, cropland and grassland of the Mesoamerica map (Giri & Jenkins, 2005) and cropland of CAC widely coincide for Central America and MOD12 IGBP and GLC2000 classify large parts of Cuba as cropland.

The other visible overestimation is class barren in South America, namely in southern Peru, western Bolivia, northwestern Argentina, and northern Chile. While barren was correctly classified in the coastal desert zones in southern Peru and the Atacama Desert of Chile and is also correct for many other mountainous regions in the Andes, the class extends far into the Altiplano of Bolivia and northwestern Argentina. However, in these high-plateaus, vegetation above tree line and below snow line (the highest in the world) is dominated by very open and sparse grassland known as Puna vegetation. Even though the northern part around Lake Titicaca in the Central Andean Wet Puna (Olson et al., 2001) was correctly classified as temperate grassland, regions further south, known as Central Andean Dry Puna, High Monte, Southern Andean Steppe, and patches of the Central Andean Puna (Olson et al., 2001) were classified as barren. The reason for the confusion could be mainly attributed to the significant soil background signal of open grasslands that even increases for medium resolution data, because locally denser grassland patches are often smaller than the spatial resolution of 500 m. It should be noted that other global maps also classify the Puna region largely as barren (GLC2000), or as a mixture between sparse vegetation and barren (Globcover) and open shrublands and barren (MOD12 IGBP).

The discrete map reflects the class of the dominant membership estimate; hence it is easily interpretable but cannot represent land cover diversity as compared to memberships (see below). We present evidence that classifier confidence indicates spatially the level of dominance of the assigned class. If for one pixel all classes have nearly the same membership a value of slightly above 5.5% (18 classes) could be sufficient to assign a pixel to the dominant class (majority rule) that thus will have a confidence of slightly above 5.5%. There are no pixels with 100% confidence that indicates that the classifier never mapped one pixel with pure leafs in all runs. The graphical display of area as a function of confidence (Fig. 4) illustrates important patterns of area decrease that begins for most classes for confidences between 15 and 25%. Classes with many high confidence pixels show concave curves while classes with many low confidence pixels show convex curves. There are also many classes that have a nearly linear inverse relationship with constant decreases in proportional area between confidences around 20% to close to 100%. The analysis also showed that scattered classes and classes with small area proportions are generally less confident and decrease faster in area for higher confidences.

4.3. Class memberships of Latin America and the Caribbean

Although the discrete map is the usually requested result by decision makers and planners at the federal level and by international organizations as well as other users in the modeling community, the memberships, in fact, contain a lot more information but are also more difficult to interpret. Still, it is advantageous to derive membership estimates, also because it is rather simple to transform those to a discrete map. In the simplest form (as implemented in this study) the majority rule is used, assigning the class with the highest membership, but also other class-specific or regionalized rules could be developed, e.g. multiplying the memberships with a-priori probabilities. In this sense, membership estimates provide the most flexible way to derive a discrete, and even a customized map.

We compared the memberships as a result of the classification against estimates of the membership from fine spatial resolution data. The difference between memberships of all classes had an average RMSE between 6 and 15% and depended mainly on the complexity of the study site. Although this result is satisfying, the analysis has to be interpreted carefully, because there are deficiencies in the reference data. Our sources were digitized at a scale of 1:250,000 which limit the spatial detail in terms of minimum patch size and accuracy of polygon boundaries. This limitation has a direct effect on the rasterized 25 m spatial resolution data that cannot represent the actual detail on the ground. This issue cannot be further quantified with our data but a qualitative discussion for transitional zones has been presented by Colditz et al. (2011). Unfortunately we had no other data sources available to compare classification memberships to more adequate reference data and suggest future studies that should emphasize this scale-related issue. Membership map interpretation can highlight spatial patterns that cannot be revealed in a discrete map. Under particular circumstances the confidence may be an indicator for transitional zones (Keil et al., 2010). Membership maps do not only spatially depict transitional zones that are often related to broad-scale ecotones, such as the transition from forest to shrubland, but also indicate landscapes that are dominated by several classes. They may also help to better characterize mixed pixels in landscapes with patch sizes below the spatial resolution of the imagery (Colditz et al., 2011, 2012). In this mapping area the Patagonian steppe is an excellent example for a heterogeneous landscape where temperate grassland and shrubland coexist (Jobbágy et al., 1996). The discrete map shows a dominance of shrubland in the northern part and grassland in the South. Membership maps of both classes show the broad-scale co-existence of shrubs and grass throughout the region and their respective areas of relative dominance, reflecting the constraints imposed by geomorphological and edaphic factors and their effects on water availability (Paruelo et al., 2004, 2007).

A small-patch landscape, where several patches may occur within one pixel, is typical for agricultural frontiers. Cropland shows moderate memberships and thus intermingles with many other classes including the region of the arc of deforestation in southern Amazonia, where small-scale farming, and more importantly clearing of forests for pasture, have historically played the most significant role in the clearing of Amazonian forest (Kirby et al., 2006). The patterns of class dominance were not only reflected by membership but also remained in the discrete map showing deforestation patterns in that region. Other examples of representing mixed landscape by class memberships with similar mapping approaches were presented in Colditz et al. (2011), Colditz et al. (2012), and by vegetation continuous fields globally (DeFries et al., 1999) or in many parts of the world (e.g. Hansen et al., 2008).

The analysis could suggest that the total area-estimates from memberships are more accurate than from the discrete map. We cannot provide a quantitative proof because of missing reference data for comparison, but in a qualitative sense reason is given, in particular for scarce classes with generally low memberships that hardly ever dominate a pixel. Similar patterns were found by others, e.g. Fisher (2010) and Chen et al. (2010). For instance the rare class sub-tropical needleleaf forest with only 0.71 Mio ha in the discrete map yields 4.69 Mio ha when computed from memberships. The gain in area

when using memberships can be enormous such as for tropical broadleaf deciduous forest (300%), sub-tropical needleleaf forest (560%), and wetland (135%). In fact, all classes with less than 150 Mio ha (except for temperate broadleaf evergreen forest and mixed forest) showed an increase in area from memberships. The reason is that with the fuzzy to discrete class conversion the fractional detail is lost, and only the class with the highest membership is considered ("the winner takes it all"). This transformation affects to a higher degree classes that mainly occur in co-existence with others and show little dominance. The exception of temperate broadleaf evergreen forest is owed to its local occurrence in southern Chile and highest elevations in the tropical Andes, where it dominates over other classes which results in many discrete pixels and few areas where it was overruled. Mixed forest was only mapped in CAC and by its definition represents a mixed class that occurs in co-existence with others and often barely dominated the pixel.

Plots of the decrease of fuzzy area as a function of membership show three broad categories of curves. Concave curves are shown for classes that have many pixels with high memberships. That may be achieved by a large area of high confidence pixels and the widespread existence of pixels with low memberships does not impact the reduction, such as for tropical broadleaf deciduous forest, or a class with local occurrence and clear dominance as shown for barren. On the other hand there are convex curves for all classes with many pixels of low membership (e.g. sub-tropical needleleaf forest, tropical broadleaf deciduous forest, and wetland). As noted above these classes can likely only be properly accounted for in their spatial extent and total area using class membership estimates. The third category form classes with a more or less linear trend of decreasing area with increasing membership (e.g. sub-tropical broadleaf evergreen forest, mixed forest, tropical and temperate shrubland). The slope may vary for different alpha cuts of the membership as shown for class temperate needleleaf forest. An initially convex curve, showing that many pixels of low memberships impact the area of this spatially limited class, turns into a concave curve for higher memberships, which indicates the relative impact of the remaining pixels with high memberships in the Sierra Madre Occidental. In this respect the plots of membership against area illustrate graphically class-specific patterns of potential existence in a discrete map. In general terms the patterns of classes among the three groups also correspond to the plots of map area as a function of confidence. The only notable exception is tropical broadleaf deciduous forest with a convex curve for memberships but a linear decrease in confidence, which indicates that if this class was mapped in the discrete map then the dominance was substantial.

4.4. Classification accuracy

From the analysis of Table 6 it is clear that the main source of the error in the map arises from confusion among ecologically similar classes. For instance, a substantial part of wetlands was classified as tropical broadleaf evergreen forest. Both classes occur in close vicinity to each other, e.g. mangroves, riparian areas along tropical rivers and deltas, and are therefore difficult to discriminate. Also interesting are confusions between shrublands and grasslands within their respective climatic zone but with little confusion among the climatic regions. That result is particularly notable, because it is commonly expected that the climate gradient introduced at level 2 of the classification scheme causes confusion among related classes (Colditz et al., 2012). Instead, confusion in this study was higher among ecological gradients (shrubland versus grassland or tropical shrubland versus tropical forests) than climate. In this respect (although not tested), the ancillary data that describe climate gradients such as precipitation and temperature and elevation helped to discern climatic regions. On the other hand, subtle ecological differences for shrubland–grassland transitions and for high biomass tropical forest versus shrubland were predicted less accurately despite the use of spectral and temporal features. These patterns demonstrate that classification errors are largely concentrated among classes that encompass ecological and biophysical gradients, and that are quite similar both functionally and in terms of their spectral–temporal properties. They also suggest that, depending on user needs, it may make sense to improve map quality by aggregating classes in the areas with higher errors (e.g., classes temperate shrubland and temperate grassland in the Patagonian Steppe).

Another source of error of the classification map could be due to the number of available samples used to train the classifier. For example, classes sub-tropical needleleaf forest and for SA temperate needleleaf forest show especially low accuracies and also present a small and fragmented distributional pattern in the map that is related to the few training samples as mentioned earlier.

On the other hand, classes such as tropical broadleaf evergreen forest and tropical and sub-tropical shrublands that encompass large areas were mapped with accuracies above 80%. It has been proven that the spatial pattern of the landscape influences the occurrence of land cover classes at varying resolution and also the area estimates derived from coarse resolution maps (Mayaux & Lambin, 1995; Moody & Woodcock, 1994; Woodcock & Strahler, 1987; Zhao et al., 2009). In this study, the distribution in large homogeneous patches of tropical broadleaf evergreen forest and tropical and sub-tropical shrublands improved the chances of pixels of being classified successfully. In addition, tropical broadleaf evergreen forest, that dominates the Amazonia region, has a distinct reflectance signal, wide-spread pattern with relatively steady ground features, and low seasonality. These characteristics discern this class from others in satellite images and facilitate its characterization in image classification. In a similar way, a highly different spectral signal helped to classify barren with high classification accuracies. The spectral signal of this land cover type is very different from vegetated surfaces and can be particularly well mapped for a large homogeneous area such as the Atacama Desert.

Ambiguity in reference labeling was accounted for by allowing for an alternative assignment. This is a common procedure, suggested in various modifications also by Sarmento et al. (2009), Zhang and Foody (1998), and Zhu et al. (2000) and used in a similar form also to validate the continental NALCMS map (Latifovic et al., 2012). We found our approach to be effective, because the polygon that delineates the pixel could easily be superimposed over several high resolution data sources, and at the same time indicated the area that the analyst had to assess. In this sense assessment was carried out at the same approximate scale (1:2500) for all sites, and heterogeneity was easy to interpret as it did not require confidence or other measures that are subjective to the analyst. The overall accuracies increased moderately by 4.0% for CAC, 10.4% for SA, and 7% for LAC.

As thematic detail and spatial resolution affect the land cover heterogeneity it also defines the maximum achievable classification accuracy (Latifovic & Olthof, 2004). Only 8.1% of the reference samples required an alternative call, which implies significant ambiguity in assignment and therefore is an indicator of the mixed pixel issue. This sample-based estimate would set a theoretically achievable overall accuracy to 91.9% that is 7.7% higher than the estimated overall accuracy.

5. Conclusions

This paper described the production process of a land cover map for Latin America and the Caribbean for the year 2008 and discussed in detail the results. It was an international effort in the framework of the "Latin American Network for Monitoring and Studying of Natural Resources" (SERENA) that is part of the regional RedLaTIF network, integrated in the global GOFC-GOLD initiative. This map with 22 classes represents an alternative land cover map to several available products with more general or different legends (Bicheron

et al., 2008; Clark et al., 2012; Eva et al., 2004; Hojas Gascon et al., 2012).

The SERENA land cover for LAC has the following characteristics:

i. Use the local expertise of SERENA network members. The involvement of local experts to generate training and validation data not only ensured local relevance of the map and helped to produce the high degree of confidence of the final product, but it also leads to greater acceptance of the map by local communities and provides important elements of capacity building. However, it was also a considerable effort and required time and monetary resources to coordinate a larger group and to find consensus among the researchers from different fields such as engineering, geography, and biology.
ii. Apply a methodology for land cover mapping based on decision trees using MODIS time series data. Although the resolution is coarse, MODIS data have the advantage of high temporal resolution, and the surface reflectance product employed in this study also provides the necessary spectral information for detailed thematic land cover mapping. This allowed us to calculate predictor variables that describe vegetation phenology and spectral differences among classes. Indeed the decision tree classifier was successful in discriminating LULC classes using this spectral-temporal variable set, without the operator having to select optimal predictor variables or classifier parameters. The decision tree algorithm is non-parametric and can handle classes with multi-modal frequency distributions which is particularly important when mapping at broad scales, because class variance increases across environmental and anthropogenic gradients (Clark et al., 2012).
iii. Produce class membership estimations of land cover classes and a discrete map with an additional layer estimating the per-pixel confidence. Highly mixed landscapes are difficult to classify with coarse spatial resolution data, but fractional class estimates partly overcome these difficulties. It has been shown that membership estimations represent the mixed pixel issue. The comparison to reference data suggests that the difference is a function of thematic and spatial complexity. Particularly, class membership estimates are deemed useful for mapping small and scattered classes in their spatial extent and may improve the estimate of their total area. Although membership estimates would be most appropriate to represent the intrinsic class heterogeneity of the Earth surface, which even increases for coarse resolution data, the user community usually requests a discrete map for analysis, modeling, and decision making. In this conversion a part of the information is lost. That affects small and scattered classes in heterogeneous areas more than dominant classes in homogeneous regions. The discrete map of this study will be provided to the user community for free-of-charge and is accompanied by a confidence layer that indicates the potential use of the pixel assignment to the interested user. The discrete map was assessed with satisfying results using common accuracy measures.

SERENA land cover products for LAC have high scientific relevance and will have several applications. Available land cover information on a broad scale is critical for understanding land surface processes which can be related to economic, social, and environmental aspects for sustainability. Several international organizations including IGBP and FAO are particularly interested in products and methodologies to automatically classify land cover and to estimate land cover change using remote sensing data. Indeed, the implementation of carbon accounting schemes, currently arduously discussed and in several Latin American countries on-going (e.g. Panama, Paraguay and Bolivia UN-REDD programs, REDD-PAC in Brazil), will require land cover information. In this context, this new land cover map of Latin America and the Caribbean, developed in the framework of the SERENA project, is a contribution to these goals but will surely need refinements to achieve the requested detail in spatial and thematic terms as well as the necessary accuracy. In addition, it needs to be shown that the map, and in particular the class membership estimates, are useful to detect change and degradation.

The methodology employed in this paper to derive training data and to generate land cover products on the continental scale could be applied to any other year of the MODIS dataset (2000–2012). Comparing results from different years would allow checking, 1) the response of vegetation to interannual climate variability in areas where higher sensitivity of vegetation could be expected, e.g. semi-arid regions, and 2) the robustness of the procedure in generating similar results for areas in which no changes in land cover are expected, e.g. core forest areas. The role of classification confidence and membership estimates for these tasks remains to be shown at the continental scale (for country-specific analysis see Colditz et al., 2011) but may have the potential to indicate areas of change.

Acknowledgments

This research was performed in the framework of CYTED-funded SERENA project coordinated by Dr. Carlos Di Bella (Programme of Science and Technology for Development CYTED 508AC0352). The authors would like to thank Carlos Di Bella and Julieta Straschnoy for their support and invaluable assistance. We also thank three anonymous reviewers for their detailed and helpful comments that significantly improved the manuscript.

References

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., & Weber, J. (2008). The most detailed portrait of earth. *ESA Bulletin*, *136*, 24–31.
Asner, G. P. (2001). Cloud cover in Landsat observations of the Brazilian Amazon. *International Journal of Remote Sensing*, *22*, 3855–3862.
Ayesa, J. A., López, C. R., Bran, D. E., Umaña, F. J., & Lagorio, P. A. (2002). Cartografía biofísica de la Patagonia Norte. *Relevamiento integrado*, *76*, Argentina: Instituto Nacional de Tecnología Agropecuario-Estación Experimental Bariloche.
Bartholomé, E., & Belward, A. S. (2005). GLC2000: A new approach to global land cover mapping from Earth Observation data. *International Journal of Remote Sensing*, *26*, 1959–1977.
Bauer, M. E., Burk, T. E., Ek, A. R., Coppin, P. R., Lime, S. D., Walsh, T. A., et al. (1994). Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing*, *60*, 287–298.
Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., et al. (2008). GlobCover 2005 — Products description and validation report, Version 2.1. Available online on the ESA IONIA website. http://ionia1.esrin.esa.int/
Borak, J. S., & Strahler, A. H. (1999). Feature selection and land cover classification of a MODIS-like data set for a semiarid environment. *International Journal of Remote Sensing*, *20*, 919–938.
Card, D. H. (1982). Using known map categorical marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, *48*, 431–439.
Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, *2*, 291–308.
Chen, C., Li, M., Huang, Q., Chen, Z., & Mao, K. (2010). Mapping land cover types as fuzzy sets. *Proceedings of International Conference on Geoinformatics* (pp. 1–6).
Clark, M. L., Aide, T. M., & Riner, G. (2012). Land change for all municipalities in Latin America and the Caribbean assessed from 250-m MODIS imagery (2001–2010). *Remote Sensing of Environment*, *126*, 84–103.
Colditz, R. R., López Saldaña, G., Maeda, P., Espinoza, J. A., Meneses Tovar, C., Victoria Hernández, A., et al. (2012). Generation and analysis of the 2005 land cover map for Mexico using 250 m MODIS data. *Remote Sensing of Environment*, *123*, 541–552.
Colditz, R. R., Schmidt, M., Conrad, C., Hansen, M. C., & Dech, S. (2011). Land cover classification with coarse spatial resolution data to derive continuous and discrete maps for complex regions. *Remote Sensing of Environment*, *115*, 3264–3275.
COLPOS (2005). *Classification of agricultural lands of Mexico.* Colegio de Posgraduados (COLPOS): Segretaría de agricultura, ganadería, desarrollo rural, pesca y alimentación de Mexico (SAGARPA).
CONAMA (2002). *Cobertura nacional de pisos de vegetación de Chile. Escala 1:250.000.* Sistema de Información Ambiental Geográfica, Ministerio del Medio Ambiente, Gobierno de Chile.
Congalton, R., & Green, K. (2009). Assessing the accuracy of remotely sensed data: Principles and practices (2nd ed.). Boca Raton, Fl: CRC/Taylor & Francis.
Conrad, C., Colditz, R., Dech, S., Klein, D., & Vlek, P. (2011). Improved irrigated crop classification in Central Asia using temporal segmentation and MODIS time series. *International Journal of Remote Sensing*, *32*, 8763–8778.
Coppin, P. R., Jonckheere, I., & Nachaerts, K. (2003). Digital change detection in ecosystem monitoring: A review. *International Journal of Remote Sensing*, *24*, 1–33.

CRPS INTA ( ). *Guía Geográfica Interactiva de Santa Cruz. Instituto Nacional de Teconología Agropecuaria.* Argentina: Centro Regional Patagonia Sur-Estación Experimental Santa Cruz.

DeFries, R. S., & Belward, A. S. (2000). Global and regional land cover characterization from satellite data: An introduction to the special issue. *International Journal of Remote Sensing*, *21*, 1083–1092.

DeFries, R. S., Field, C. B., Fung, I., Justice, C. O., Los, S., Matson, P. A., et al. (1995a). Mapping the land surface for global atmosphere-biosphere models: Toward continuous distributions of vegetation's functional properties. *Journal of Geophysical Research*, *100*, 20867–20882.

DeFries, R., Hansen, M., & Townshend, J. (1995b). Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sensing of Environment*, *54*, 209–222.

DeFries, R. S., Hansen, M. C., Townshend, J. R. G., & Sohlberg, R. A. (1998). Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, *19*, 3141–3168.

DeFries, R., Townshend, J. R. G., & Hansen, M. (1999). Continuous fields of vegetation characteristics at the global scale at 1 km resolution. *Journal of Geophysical Research-Atmospheres*, *104*, 16.911–16.925.

Eva, H., Belward, A., De Miranda, E., Di Bella, C., Gond, V., Huber, O., et al. (2004). A land cover map of South America. *Global Change Biology*, *10*, 731–744.

Fernandes, R., Fraser, R. H., Latifovic, R., Cihlar, J., Beaubien, J., & Du, Y. (2004). Approaches to fractional land cover and continuous field mapping: A comparative assessment over the BOREAS study region. *Remote Sensing of Environment*, *89*, 234–251.

Fisher, P. F. (2010). Remote Sensing of land cover classes as type 2 fuzzy sets. *Remote Sensing of Environment*, *114*, 309–321.

Fonte, C. C., & Lodwick, W. A. (2004). Areas of fuzzy geographical entities. *International Journal of Geographical Information Science*, *18*, 127–150.

Food and Agriculture Organization of the United Nations (FAO) (2005). FAO Stat. FAO statistical databases. Available online at. http://faostat.fao.org/

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185–201.

Foody, G. M., & Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, *103*, 179–189.

Freund, Y., & Schapire, R. E. (1996). Experiments with a new boosting algorithm. *Machine Learning: Proceedings of the Thirteenth International Conference, Bari, Italy July 3-6 1996* (pp. 148–156).

Friedl, M. A., Brodley, C. E., & Strahler, A. H. (1999). Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 969–977.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*, 287–302.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., et al. (2010). MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*, 168–182.

Giri, C., & Jenkins, J. C. (2005). Land cover mapping of Greater Mesoamerica using MODIS data. *Canadian Journal of Remote Sensing*, *31*, 274–282.

Grau, H. R., & Aide, T. M. (2008). Globalization and land use transitions in Latin America. *Ecology and Society*, *13*(2), 16 Available online at http://www.ecologyandsociety.org/vol13/iss2/art16/

Guerschman, J. P., Paruelo, J. M., DiBella, C., Giallorenzi, M. C., & Pacín, F. (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing*, *24*, 3381–3402.

Hansen, M. C. (2012). Classification trees and mixed pixel training data. In C. P. Giri (Ed.), *Remote sensing of land Use and land cover: Principles and applications*. Boca Raton, Fl: CRC/Taylor & Francis.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *216–7*, 1331–1364.

Hansen, M. C., DeFries, R. S., 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.

Hansen, M., Roy, D. P., Lindquist, E., Justice, C., & Alstatt, A. (2008). A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in Central Africa. *Remote Sensing of Environment*, *112*, 2495–2513.

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, *112*, 2538–2556.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Hojas Gascon, L. H., Eva, H. D., Gobron, N., Simonetti, D., & Fritz, S. (2012). The application of medium-resolution MERIS satellite data for continental land-cover mapping over south America—Results and caveats. In C. P. Giri (Ed.), *Remote sensing of land Use and land cover: Principles and applications*. Boca Raton, Fl: CRC/Taylor & Francis.

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 National Land-Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, *70*, 829–840.

Homer, C. G., Ramsey, R. D., Edwards, T. C., Jr., & Falconer, A. (1997). Landscape cover-type modeling using a multi-scene Thematic Mapper mosaic. *Photogrammetric Engineering and Remote Sensing*, *63*, 59–67.

IDEAM, IGAC, IAvH, Invemar, Sinchi, I., & IIAP (2007). *Ecosistemas continentales, costeros y marinos de Colombia.* Bogotá, D.C.: Instituto de Hidrología, Meteorología y Estudios Ambientales, Instituto Geográfico Agustín Codazzi, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Instituto de Investigaciones Ambientales del Pacífico Jhon von Neumann, Instituto de Investigaciones Marinas y Costeras José Benito Vives De Andréis e Instituto Amazónico de Investigaciones Científicas Sinchi (276 pp. + 37 hojas cartográficas).

INEGI (2005). Dirección General de Geografía. In INEGI (Ed.), *Conjunto de Datos Vectoriales de la Carta de Uso del Suelo y Vegetación, Escala* 1:250,000, Serie III (CONTINUO NACIONAL). Instituto Nacional de Estadística, Geografía e Informática - INEGI. Aguascalientes, Ags., México.

INRENA (2000). Mapa de Cobertura Vegetal y Uso de la Tierra del Perú a escala 1: 250,000. Base de Datos de Recursos Naturales e Infraestructura.

Jobbágy, E. G., Paruelo, J. M., & León, R. J. C. (1996). Vegetation heterogeneity and diversity in flat and mountain landscapes of Patagonia (Argentina). *Journal of Vegetation Science*, *7*, 599–608.

Keil, M., Gessner, U., Hüttich, C., & Colditz, R. R. (2010). Large-scale vegetation assessments in southern Africa: Concepts and applications using multi-source remote sensing data. In U. Schmiedel, & N. Jürgens (Eds.), *Biodiversity in southern Africa. Patterns and processes at regional scale*, *2*. (pp. 32–45) Göttingen & Windhoek: Klaus Hess Publishers.

Kirby, K. R., Laurance, W. F., Albernazc, A. K., Schroth, G., Fearnside, P. M., Bergen, S., et al. (2006). The future of deforestation in the Brazilian Amazon. *Futures*, *38*, 432–453.

Latifovic, R., Homer, C., Ressl, R., Pouliot, D., Hossain, S. N., Colditz, R. R., et al. (2012). North American land-change monitoring system. In C. P. Giri (Ed.), *Remote sensing of land Use and land cover: Principles and applications* (pp. 303–324) Boca Raton, Fl: CRC/Taylor & Francis.

Latifovic, R., & Olthof, I. (2004). Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. *Remote Sensing of Environment*, *90*, 153–165.

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, *296*, 1–22.

Leyk, S., & Zimmermann, N. E. (2007). Improving land change detection based on uncertain survey maps using fuzzy sets. *Landscape Ecology*, *22*, 257–272.

Lillesand, T. M. (1996). A protocol for satellite-based land cover classification in the Upper Midwest. In J. M. Scott, T. H. Tear, & F. W. Davis (Eds.), *Gap analysis: A landscape approach to biodiversity planning* (pp. 103–118) Bethesda, MD: ASPRS.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., et al. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *216–7*, 1303–1330.

Luo, Y., Trishchenko, A. P., & Khlopenkov, K. V. (2008). Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment*, *112*, 4167–4185.

Magrin, G., Gay García, C., Cruz Choque, D., Giménez, J. C., Moreno, A. R., Nagy, G. J., et al. (2007). Latin America. Climate change 2007: Impacts, adaptation and vulnerability. In M. L. Parry, O. F. Canziani, J. P. Palutikof, P. J. van der Linden, & C. E. Hanson (Eds.), *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 581–615) Cambridge, UK: Cambridge University Press.

Mayaux, P., & Lambin, E. F. (1995). Estimation of tropical forest area from coarse spatial resolution data: A two-step correction function for proportional errors due to spatial aggregation. *Remote Sensing of Environment*, *53*, 1–16.

McIver, D. K., & Friedl, M. A. (2001). Estimating pixel-scale land cover classification confidence using non-parametric machine learning methods. *IEEE Transactions on Geoscience and Remote Sensing*, *39*, 1959–1968.

Moody, A., & Woodcock, C. E. (1994). Scale-dependent errors in the estimation of land-cover proportions: Implications for global land-cover datasets. *Photogrammetric Engineering and Remote Sensing*, *60*, 585–594.

Navarro, G. S., & Ferreira, W. (2007). *Leyenda explicativa de las unidades del mapa de vegetación de Bolivia a escala 1:250 000,* (pp. 65). Cochabamba, Bolivia: The Nature Conservancy-Rumbol SRL.

NGDC (National Geophysical Data Center) (2012). Nighttime Lights data available at: http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience*, *51*, 933–938.

Pacheco, P. (2012). Actor and frontier types in the Brazilian Amazon: Assessing interactions and outcomes associated with frontier expansion. *Geoforum*, *43*, 864–874.

Paruelo, J. M., Golluscio, R., Guerschman, J., Cesa, A., Jouve, V., & Garbulsky, M. (2004). Regional scale relationships between ecosystem structure and functioning: The case of the Patagonian steppes. *Global Ecology and Biogeography*, *13*, 385–395.

Paruelo, J. M., Jobbágy, E. G., Oesterheld, M., Golluscio, R. A., & Aguiar, M. R. (2007). The grasslands and steppes of Patagonia and the Rio de la Plata plains. In T. Veblen, K. Young, & A. Orme (Eds.), *The physical geography of South America* (pp. 232–248) Oxford: Oxford University Press (Chapter 14).

PROBIO (2004). Mapeamento dos biomas brasileiros a escala 1:250,000.

PROCEDE (2006). *Programa de Certificacion de Derechos Ejidales y Solares Urbanos (PROCEDE).* Secretaría de la Reforma Agraria (SRA), Procuraduría Agraria (PA), Instituto Nacional de Estadística y Geografía (INEGI); Registro Agrario Nacional (RAN) (Data from 1992 to 2006).

Quinlan, J. R. (1993). C4.5: Programs for machine learning. San Mateo, CA:Morgan Kaufmann.

Quinlan, J. R. (1996). Bagging, boosting, and C4.5. *Machine Learning: Proceedings of the Thirteenth International Conference, Bari, Italy July 3–6 1996* (pp. 725–730).

Reese, H. M., Lillesand, T. M., Nagel, D. E., Stewart, J. S., Goldmann, R. A., Simmons, T. E., et al. (2002). Statewide land cover derived from multiseasonal Landsat TM data:

A retrospective of the WISCLAND project. *Remote Sensing of Environment*, *82*, 224–237.

Rogan, J., & Chen, D. (2004). Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progress in Planning*, *61*, 301–325.

SADS (2005). *Primer Inventario Nacional de Bosques Nativos a escala 1:250,000. Proyecto Bosques Nativos y Áreas Protegidas.* Buenos Aires, Argentina: Secretaria de Ambiente y Desarrollo Sostenible.

Salomonson, V. V., & Appel, I. (2006). Development of the Aqua MODIS NDSI fractional snow cover algorithm and validation results. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1747–1756.

Sarmento, P., Carrão, H., Caetano, M., & Stehman, S. V. (2009). Incorporating reference classification uncertainty into the analysis of land cover accuracy. *International Journal of Remote Sensing*, *30*, 5309–5321.

Schapire, R. E., Freund, Y., Bartlett, P., & Lee, W. S. (1998). Boosting the margin: A new explanation for the effectiveness of voting methods. *The Annals of Statistics*, *26*, 1651–1686.

Skole, D., Justice, C. O., Janetos, A. C., & Townshend, J. (1997). A land cover change monitoring program: A strategy for international effort. *Mitigation and Adaptation Stategies for Global Change*, *2*, 157–175.

Small, C., Pozzi, F., & Elvidge, C. D. (2005). Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sensing of Environment*, *96*, 277–291.

Strahler, A. H., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., et al. (2006). Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps. *GOFC/GOLD Report No 25*. Ispra, It: JRC.

Trishchenko, A. P., Luo, Y., & Khlopenkov, K. V. (2006). A method for downscaling MODIS land channels to 250 m spatial resolution using adaptive regression and normalization. *Proceedings of SPIE—The International Society for Optical Engineering*, *vol. 6366*, (8 pp. art. no. 636607).

UNEP (2000). GEO Latin America and the Caribbean. Mexico: Environmental Outlook 2000, UNEP Regional Office for Latin America and the CaribbeanMexico.

USGS (2004). *Shuttle radar topography mission, global land cover facility*. College Park, Maryland: University of Maryland (February, 2000).

Vermote, E. F., El Saleous, N. Z., & Justice, C. O. (2002). Atmospheric correction of MODIS data in the visible to middle infrared: First results. *Remote Sensing of Environment*, *83*, 97–111.

Woodcock, C. E., & Gopal, S. (2000). Fuzzy set theory and thematic maps: Accuracy assessment and area estimation. *International Journal of Geographical Information Science*, *14*, 153–172.

Woodcock, C. E., & Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*, 311–325.

World Bank, & CCAD (). *Ecosystems of Central America (Arcview map files at 1:250,000).* Washington, D.C.: World Bank, Comisión Centroamerica de Ambiente y Desarrollo (CCAD), World Institute for Conservation and Environment (WICE), and the Centro Agronómico Tropical de Investigación y Enseñanza (CIAT).

Zhang, J., & Foody, G. M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing*, *19*, 2721–2738.

Zhao, T., Bergen, K. M., Brown, D. G., & Shugart, H. H. (2009). Scale dependence in quantification of land-cover and biomass change over Siberian boreal forest landscapes. *Landscape Ecology*, *24*, 1299–1313.

Zhu, Z., Yang, L., Stehman, S. V., & Czaplewski, R. L. (2000). Accuracy assessment for the U.S. Geological Survey regional land-cover mapping programme: New York and New Jersey region. *Photogrammetric Engineering and Remote Sensing*, *66*, 1425–1435.