A land cover map from Ellesmere Island to Tierra del Fuego: the combination of two continental land cover mapping projects and comparison to global maps

René Roland Colditz [1]
Ricardo Magdaleno Llamas Barba [1]

[1] National Commission for the Knowledge and Use of Biodiversity (CONABIO)
Avenida Liga Periférico-Insurgentes Sur 4903, Col. Parques del Pedregal, Del. Tlapan, 14010, Mexico, D.F.
rcolditz@conabio.gob.mx, ricardo.llamas@conabio.gob.mx

Abstract. Land cover affects and is affected by climate change and therefore is a critical input to climate, biogeochemical, hydrological, and species distribution models. There are multiple global land cover datasets such as GLC2000, the MODIS land cover product, and GlobCover. Although all are useful for modeling at the global scale, the spatial resolution, thematic detail, and accuracy is often insufficient for regional studies. The North American Land Change Monitoring System (NALCMS) was founded in 2006 with the goal to provide annually updated land cover information for the North American continent, which is Canada, the United States, and Mexico. As a first effort a continental map of the land cover in 2005 based on 250m MODIS data was completed. The Latin American network to monitor and study natural resources (SERENA) was founded in 2008 with the aim to monitor, study, and distribute information associated with changes due to burnt biomass as well as land cover of Latin America and the Caribbean. The network is formed by 52 participants from 11 countries and 18 institutions distributed over the whole region. The recently completed 2008 land cover map is based on 500m MODIS data. This study combines both products and presents a land cover map from Ellesmere Island to Tierra del Fuego. Critical steps in map combination included legend harmonization, adequate selection of projection and resolution parameters, and adjustments for differences among the years. The new map was compared to an existing land cover product, the MODIS land cover map of the year 2005. For important classes in terms of their total area the spatially-explicit correspondence among the maps is moderate (approximately 70%). Other classes such as mixed forest and grassland indicate poor agreements due to their intrinsic heterogeneity. These dissimilarities may be attributed to the methods (supervised or hybrid) and data but also to remaining differences in legends and processing issues to enable map comparison.

Keywords: Land cover, Map comparison, MCD12Q1, NALCMS, SERENA, MODIS, America.

1. Introduction

Land cover and land use maps are probably the most common derivatives of satellite images. Land cover maps are required by the modeling community for climate, biodiversity, and partly also hydrology and landscape analysis. Another field of people interested in land cover maps the global change community. Land cover is defined at the physical material at the Earth surface. Typical materials are grass, trees, barren ground, or water. In this sense it is different from land use that is the description of the human use of land. Most maps are in reality a mix of both and thus termed a land use and land cover map (Fisher et al. 2005). Prominent examples are the mapping of land use classes urban area and agricultural land while other classes describe typical land covers, e.g. grassland and water. Land cover maps can be derived from satellite images in a fairly automatic manner and general land uses as described above may be inferred with reasonable accuracy.

Maps can be derived from various spatial resolutions, map extents and thematic detail that is the number of classes. Usually there is proportional relationship with increasing spatial resolution, decreasing extent and more detailed thematic classes. This is directly liked to feasibility. At least until nowadays there is no global map at 30m resolution but there exist multiple maps between 300 and 1000m. On the other hand it makes little sense to map a small city at that coarse resolution. The thematic detail not exclusively refers to the number of classes but also to the detail that the classes characterize.

This paper will focus on the combination of two existing land cover – and in the strict sense used above – also land use maps for the American continent. The maps originate from two existing mapping projects in which the authors participate. The North American Land Change Monitoring System (NALCMS) is a tri-national initiative of Canada (Canada Centre for Remote Sensing / Natural Resources Canada – CCRS/NRCan), the United States of America (United States Geological Survey – USGS), and Mexico (National Institute for Statistics and Geography – INEGI, National Commission for the Knowledge and Use of Biodiversity – CONABIO, National Forestry Commission – CONAFOR). The work is supported by the Commission for Environmental Cooperation (CEC) that is a tri-national institution to support environmental studies for the continent. The aims of NALCMS are to develop a land cover monitoring system at the continental scale using satellite images. It mainly employs images of the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor. Using algorithms developed at CCRS seven reflective bands important for land cover monitoring are filtered for clouds and errors, downscaled to 250m spatial resolution and aggregated to monthly composites (Latifovic et al, 2012). In 2010 a baseline map, the land cover of North America for the year 2005, was completed and made publically available as wall map and in digital format (CEC 2010). The map has 19 thematic classes that characterize different types of forest, shrubland, grassland, lichen and moss, wetland, cropland, barren land, urban area, water, and snow / ice. The overall accuracy of the continental map is 68% for not considering ambiguity in the labeling of reference data and 82% if an alternative label is considered as correctly classified (Latifovic et al, 2012). Current activities of this on-going mapping initiative are the detection of change area and the characterization of changes for map updating. The long-term goal is an annually-generated map product. However, also continuous products and higher spatial resolution maps are anticipated.

Within the NALCMS initiative there are common guidelines for product development such as the baseline dataset, the legend that follows the Land Cover Classification System (LCCS, di Gregorio 2005) standards, and partly the algorithm. However, each country develops its own dataset also considering the national needs by not losing the overarching international requirements. Specific national requirements can be a thematically more detailed map that is aggregated for the continental map or a particular mapping approach to represent the diversity of the land cover. The latter was necessary for land cover mapping in Mexico where in many cases exists more than one label for a pixel of 6.25ha (Colditz et al. 2012). Using a multitude of discrete classifiers, C5.0 decision trees (Quinlan et al. 1993) the proportion of each land cover class was estimated. These class memberships represent commonly known ecological transition zones (ecotones) due to elevation changes or scarcity of water and heterogeneous areas with many mixed pixels, because the patchiness of the actual land cover is well beyond the resolution of 250m pixels (Colditz et al. 2011, Colditz et al 2012). Membership maps can be transformed to discrete land cover maps by assigning to each pixel the class with the highest proportion.

The same relatively new but computationally expensive approach was employed in the project “Latin American network to monitor and study natural resources” (SERENA). Founded in 2008 and embedded in the regional Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD) program RedLaTIF, the network brings together 52 researchers from 10 countries distributed over the whole Latin American continent. The aim of this initiative is to monitor, study, and distribute information associated with changes due to burnt biomass and land cover from Mexico to Tierra del Fuego. The analysis is based on a set of monthly composites of MODIS images with a spatial resolution of 500m filtered for clouds and processing issues. The land cover sub-group recently finished the land cover map of Latin America for 2008 (Blanco et al. 2012). The wall map was released in spring 2012, and the digital release is anticipated for the end of 2012. The overall map accuracy was 84% on the continental scale and major confusions in the error matrix are related to ecological gradients and less to climate zones.

Both, NALCMS and SERENA maps are comparable because a similar legend was used in both projects. The map of 19 classes in the continental NALCMS was amplified to 22 classes in SERENA. The main difference is explicitly mapping sub-tropical classes in SERENA. In NALCMS there are combined tropical and sub-tropical classes, because their distinction was not imperative due to the small area proportion of their extent. In SERENA, however the majority of the mapping area is located within the tropical and sub-tropical climate zones.

The following chapter will describe how the maps were combined and which adjustments were necessary to link classes thematically. The results chapter will present the final map for the American continent and will compare this map quantitatively to an existing global land cover map. The final chapter outlines the conclusions.

2. Methods for map combination and analysis

The NALCMS as well as the SERENA maps were generated in the Lambert Azimuthal Equal Area (LAEA) projection but with two different projection origins. Also, the NALCMS map has a spatial resolution of 250m in contrast to SERENA with 500m. For the generation of the America map both maps were reprojected to a new projection center (10°S, 90°W) and a spatial resolution of 500m.

A second step concerned legend merge of the 19 classes NALCMS and 22 classes in SERENA. The main difference between legends of both maps concerns the sub-tropical classes that were discerned in SERENA but not in NALCMS. However, the issue can be solved, because tropical classes in the NALCMS map extent only exist in Mexico, and Mexico was also covered by SERENA. Therefore the new America map employed data of SERENA for Mexico. NALCMS data were only used for Canada and the United States where all tropical and sub-tropical classes were labeled "sub-tropical". Other classes of lichens and mosses in NALCMS only exist in northern Canada and Alaska and not at all in South America.

Quantitative comparison of the America map to the MCD12Q1 map to another existing map was based on the comparison matrix. First, both maps were transformed to a common generalized legend of basic land cover classes. Cells of the diagonal of the matrix indicate the agreement (A) between both maps (Equation 1). In contrast, disagreement (D) can be calculated for both maps (X, Y) and is the sum of the column or row totals (+k, k+) minus the diagonal cell (Equations 2 and 3). An additional division by 2 is necessary for not double counting commission and omission among different classes.

$$A(X,Y)_k = n_{kk} \quad (1)$$

$$D(X)_k = \frac{n_{+k} - n_{kk}}{2} \quad (2)$$

$$D(Y)_k = \frac{n_{k+} - n_{kk}}{2} \quad (3)$$

3. Results and Analysis

This section will present the results of this study. First, the land cover map for the entire American continent from Ellesmere Island to Tierra del Fuego will be presented and described. Another section will focus on a quantitative analysis of this map and a comparison to another existing map product for the American continent.

3.1. The America map

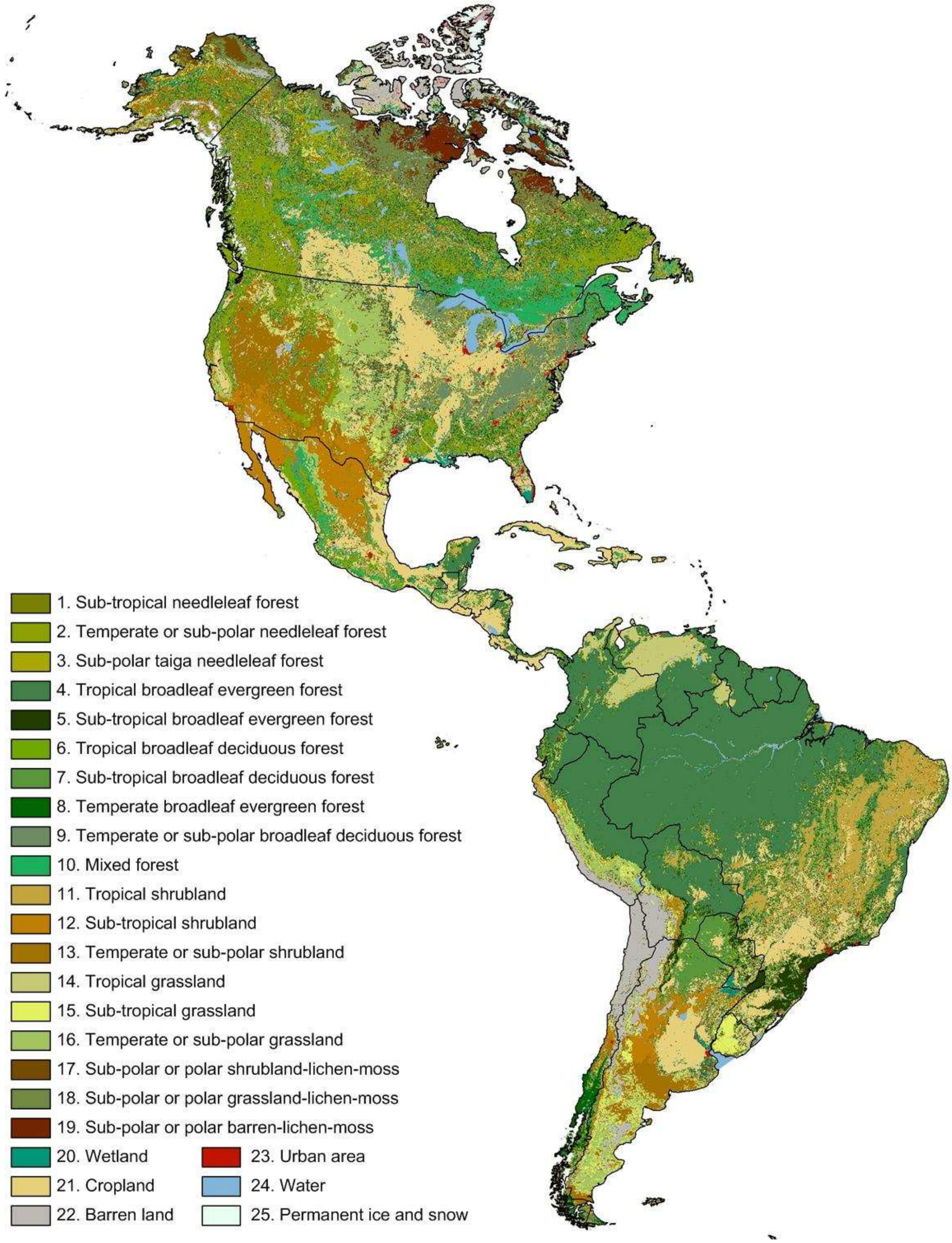

Figure 1. Land cover map of the American continent.

Figure 1 presents the land cover map for the American continent. Zonal patterns dominate the northern parts of the North American continent. Meridianal mountainous systems like the Appalachians and the Western Cordilleras interrupt this pattern with broadleaf deciduous forests and shrubland in the semi-arid regions of the interior basins and Colorado plateau. Also, land use patterns become visible at a broad scale with cropland extending from Alberta and Saskatchewan to the southern United States East of the Mississippi/Missouri system. West of this system dominate temperate grasslands. Towards Mexico those classes change towards their tropical-subtropical counterparts with temperate classes continuing towards the South at higher elevations. Other Central American countries are dominated by cropland that intermingles with tropical and sub-tropical forests. South America is dominated by the tropical broadleaf evergreen forests in the Amazon basin that extent towards the Atlantic coast with transitions to tropical shrubland and sub-tropical forests. Climatic gradients are clearly visible for South America's lowlands, changing from tropical to sub-tropical and eventually to temperate classes in Patagonia. Also the hypsometric gradient can be easily observed for land cover classes in the Andean ranges. Latifovic et al. (2012) and Blanco et al. (2012) describe in more detail and quantitative fashion the land cover for North America and Latin America, respectively, and Colditz et al. (2012) with even more detail for Mexico.

3.2. Quantitative map comparison

For quantitative map comparison the North American map was aggregated to 12 basic land cover classes. The specific aggregation scheme is presented in Table 1. This generalized map was compared to the MODIS land cover product for the year 2005 with 17 classes following the IGBP legend (Friedl et al. 2002, Friedl et al. 2010). The MODIS product was obtained in tiles that were mosaicked for entire America and projected to the same LAEA parameters and grid as the America map. The aggregation rules of the MCD12Q1 product with 500m spatial resolution was not straight-forward and followed specific semantic and spatial rules that are also summarized in Table 1.

Figure 2 presents the quantitative class-wise comparison between the generalized America and MCD12 IGBP map. Coniferous forest that mainly exists in Canada, the United States and higher elevations for the rest of the study area shows a good agreement of 62%. There is a proportion of 25% that was only mapped in MOD12 IGBP but not in America and a proportion of 13% *vice versa*. The agreement was even better for broadleaf forest with 94%. Even more important is that this class makes up a large part of the land surface with approximately 28% of the total area. In this light one has to see the poor agreement of mixed forest (35%) because less than 4% of the total land corresponds to this diverse class that is hard to define and only exists in North and Central America. Another spatially important class is shrubland with approximately 16% of the total land surface. This class shows good correspondences above 70%. Grassland indicates lower agreements with only slightly above 50%. The larger disagreement occurs for the MOD12 IGBP map with 33% because this map classified more grassland than the America map. Lichen and moss only exists in the northern zones of North America and shows good general agreements above 70%. Although having poor agreements (23%) this number for wetland has only a limited importance because of its small area proportion over the entire study area (approximately 2%). In contrast, cropland with a share between 11 and 15% of the area is rather important. The agreement of 70% is moderate due to more mapped cropland in America than in the MOD12 IGBP map. Other classes (barren land, urban and built-up, water, snow and ice) show moderate to poor agreements but their importance in terms of total area is low.

Table 1. Aggregation of the America map legend and the MCD12 IGBP legend to a generalized 12-class legend.

General land cover classes	America map	MCD12 IGBP
1. Needleleaf forest	1. Sub-tropical needleleaf forest, 2. Temperate or sub-polar needleleaf forest, 3. Sub-polar taiga needleleaf forest	1. Evergreen needleleaf forest, 3. Deciduous needleleaf forest, 8. Woody savanna
2. Broadleaf forest	4. Tropical broadleaf evergreen forest, 5. Sub-tropical broadleaf evergreen forest, 6. Tropical broadleaf deciduous forest, 7. Sub-tropical broadleaf deciduous forest, 8. Temperate broadleaf evergreen forest, 9. Temperate or sub-polar broadleaf deciduous forest	2. Evergreen broadleaf forest, 4. Deciduous broadleaf forest, 8. Woody savanna, 9. Savanna, 14. Cropland / natural vegetation mosaic
3. Mixed forest	10. Mixed forest	5. Mixed forest
4. Shrubland	11. Tropical shrubland, 12. Sub-tropical shrubland, 13. Temperate or sub-polar shrubland	6. Closed shrublands, 7. Open shrublands, 8. Woody savanna, 9. Savanna, 14. Cropland / natural vegetation mosaic
5. Grassland	14. Tropical grassland, 15. Sub-tropical grassland, 16. Temperate or sub-polar grassland	8. Woody savanna, 9. Savanna, 10. Grasslands, 14. Cropland / natural vegetation mosaic
6. Lichen and moss	17. Sub-polar or polar shrubland-lichen-moss, 18. Sub-polar or polar grassland-lichen-moss, 19. Sub-polar or polar barren-lichen-moss	7. Open shrublands
7. Wetland	20. Wetland	11. Permanent wetlands
8. Cropland	21. Cropland	12. Croplands, 14. Cropland / natural vegetation mosaic
9. Barren land	22. Barren land	7. Open shrublands, 16. Barren or sparsely vegetated
10. Urban and built-up	23. Urban area	13. Urban and built-up
11. Water	24. Water	0. Water
12. Snow and ice	25. Permanent ice and snow	15. Snow and ice

Note: Some classes of the MCD12 IGBP map correspond to multiple land cover classes in the generalized legend. In this case spatial rules were applied, i.e. this class was converted to a specific general land cover class if it corresponded to the label.

An interesting number can be presented if the generalized 12 classes are combined to 3 basic classes that are forest (coniferous, broadleaf and mixed forest), natural vegetation (shrubland, grassland, lichen and moss, and wetland) and other (remaining classes). For forest the agreement is 90%, for natural vegetation 81%, and for all other classes combined 76%. For the map overall, the agreement is 72%.

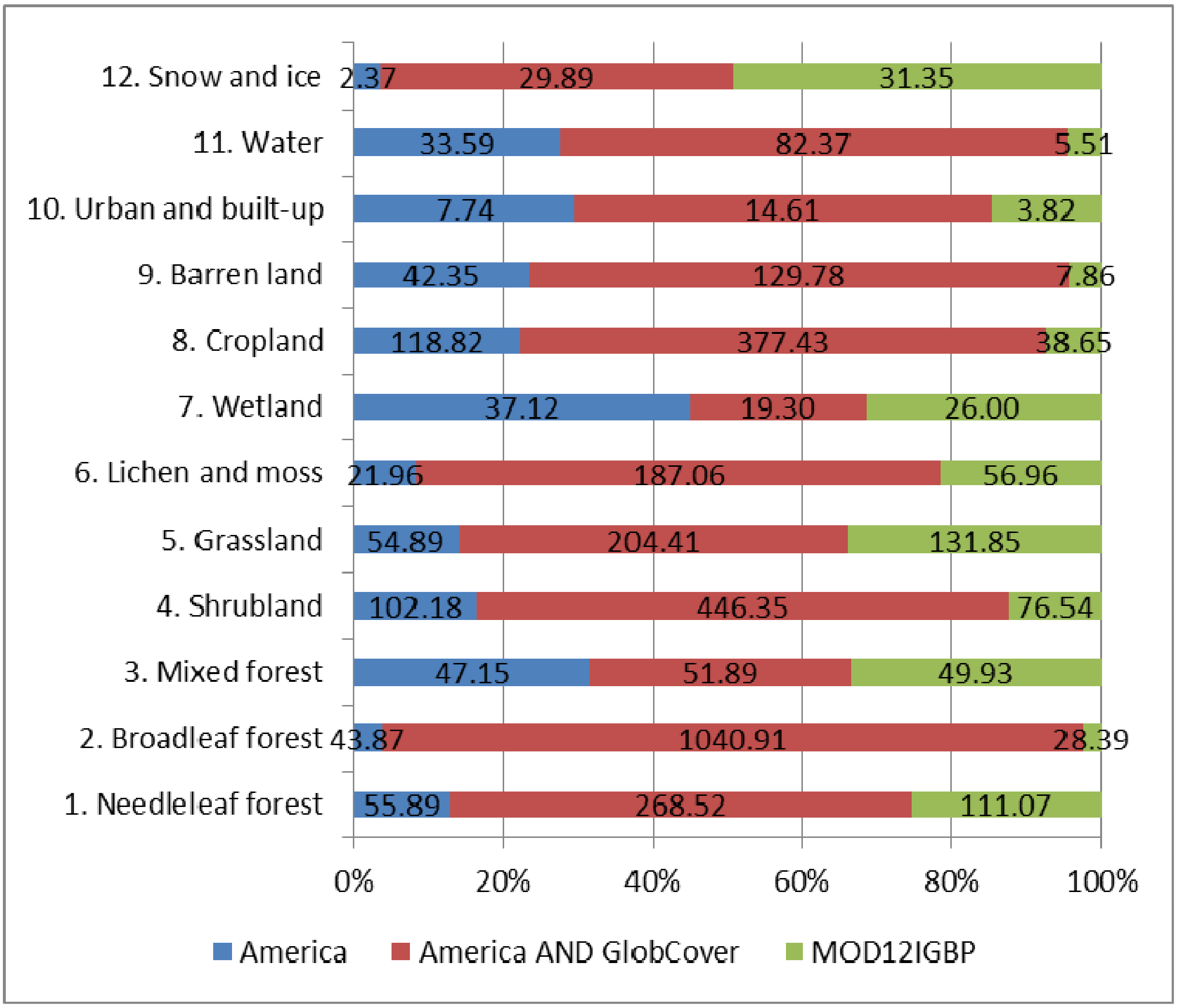

Figure 2. Class-specific agreement and disagreement (in per cent and Mio ha) between the America and MDD12 IGBP maps with generalized legends (see Table 1).

4. Conclusions

The paper presented the generation of a land cover map for the entire American continent. Due to the similarities between the NALMCS and SERENA project such a map can be accomplished by only a few processing steps. The map is important for projection on the continental scale such as climate modeling and species distribution and migration studies. The map is similar to other existing products with the advantage of presenting more thematic classes than several other products.

References

Blanco, P.D.; Colditz, R.R.; López Saldaña, G.; Hardtke, L.A.; Llamas, R.M.; Mari, N.A.; de los Angeles Fischer, M.; Caride, C.; Aceñolaza, P.G.; del Valle, H.F.; Lillo-Saavedra, M.; Coronato, F.R.; Opazo, S.A.; Morelli, F.; Anaya, J.A.; Sione, W.F.; Zamboni, P.; Barrena Arroyo, V. A new land cover map of Latin America and the Caribbean in the framework of SERENA Project (RedLaTIF - GOFC-GOLD). **Remote Sensing of Environment**, submitted, 2012.

CEC. http://www.cec.org/ 2010.

Colditz, R.R.; Schmidt, M.; Conrad, C.; Hansen, M.C.; Dech, S. Land cover classification with coarse spatial resolution data to derive continuous and discrete maps for complex regions. **Remote Sensing of Environment**, vol. 115, p. 3264-3275, 2011.

Colditz, R. R.; López Saldaña, G.; Maeda, P.; Argumedo Espinoza, J.; Meneses Tovar, C.; Victoria Hernández, A.; Zermeño Benítez, C.; Cruz López, I.; Ressl, R. Generation and analysis of the 2005 land cover map for Mexico using 250 m MODIS data. **Remote Sensing of Environment**, vol. 123, p. 541–552, 2012.

di Gregorio, **A. Land cover classification system—classification concepts and user manual for software version 2**. FAO Environment and Natural Resources Service Series, No. 8, Rome, 2005, 208 p.

Fisher, P.; Comber, A.; Wadsworth, R. Land use and land cover: contradiction or complement. In Fisher, P.; Unwin, D.J. (Org.). **Re-Presenting GIS**, Chichester, John Wiley and Sons, 2005, p. 85-98.

Friedl, M.A.; McIver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; Baccini, A.; Gao, F.; Schaaf, C.B. Global land cover mapping from MODIS: algorithms and early results. **Remote Sensing of Environment**, vol. 83, no.1-2, p. 287-302, 2002.

Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. **Remote Sensing of Environment**, vol. 114, p. 168-182, 2010.

Latifovic, R.; Homer, C.; Ressl, R.; Pouliot, D.; Hossain, S.N.; Colditz, R.R.; Olthof, I.; Giri, C.; Victoria, A. North American land-change monitoring system. In C. P. Giri (Org.). **Remote Sensing of Land Use and Land Cover: Principles and Applications**, Boca Raton, Fl: CRC/Taylor & Francis 2012, p. 303-324.

Quinlan, J. R. **C4.5: Programs for Machine Learning**. San Mateo, CA, Morgan Kaufmann, 1993.