CLASSIFYING THE LAND COVER OF MEXICO IN THE FRAMEWORK OF THE NORTH AMERICAN LAND CHANGE MONITORING SYSTEM

René R. Colditz, Remote Sensing Expert
Pedro Maeda, GIS Expert
Gerardo López, Remote Sensing Expert
Isabel Cruz, Remote Sensing Group Leader
Rainer Ressl, Director of Geomatics
National Commission for the Knowledge and Use of the Biodiversity (CONABIO)
Liga Periférico-Insurgentes Sur 4903, Col. Parques del Pedregal, Del. Tlalpan,
CP: 14010, Mexico City, D.F.
Tel: +52-55-50045020, FAX: +52-55-5004-4931
rene.colditz@conabio.gob.mx
pedro.diaz@conabio.gob.mx
gerardo.loez@conabio.gob.mx
isabel.cruz@conabio.gob.mx
rainer.ressl@conabio.gob.mx

ABSTRACT

Automatic derivation of land cover information from optical satellite data is one of the main research topics in remote sensing. Accurate land cover products are needed for biogeochemical modelling and biodiversity analysis. Land transformation processes are analyzed in the framework of global change studies. While many global land cover products fulfil the needs for driving global models, their accuracy is insufficient for regional to continental applications. The North American Land Change Monitoring System is a tri-national initiative to provide accurate, automatically generated annual land cover and land change products. Its first product, the land cover map of North America for 2005, has recently been completed and is published by the intergovernmental Commission for Environmental Cooperation. The paper presents the classification for the Mexican section. Although the smallest, it is without doubt the most complex portion of the study area due to the transition of temperate and tropical ecosystems and a very heterogeneous small-patch land cover assemblage. The classification is based on monthly MODIS composites and ancillary data. A large sample data base was built to train a multitude of boosted decision trees. The combination of multiple classifications assures map consistency, an important step to mitigate false change detection in the future. Class memberships of the decision trees were transformed to a discrete map, which is accompanied by a confidence layer. The discrete map assessment yielded 82 %.

INTRODUCTION

Land cover information is an important parameter for climate and biogeochemical modeling, hydrological and biodiversity analysis, and land change mapping. The distribution of land cover types and their dynamics help to understand climate change impacts, and combined with other data a distinction between natural and human change processes might be possible.

Since the mid 1990ies several global land cover maps based on NOAA-AVHRR or SPOT VEGETATCION data have been produced (IGBP: Loveland et al. 2000; GLC: Bartholomé and Belward 2005). The MODIS MOD12 product attempts to automatically map land cover on an annual basis (Friedl et al. 2002). Most recently the GLOBCOVER land cover map based on MERIS data was released (Bicheron et al. 2006). Although these land cover maps provide suitable information on a coarse mapping scale and for modeling global processes, their accuracy is insufficient for a more detailed regional analysis and for monitoring of change.

Level I	Level II
1. Needle leaved forest	1. Temperate or sub-polar needleleaf evergreen forest
	2. Sub-polar taiga needleleaf forest
2. Broadleaved forest	3. Tropical or sub-tropical broadleaf evergreen forest
	4. Tropical or sub-tropical broadleaf deciduous forest
	5. Temperate or sub-polar broadleaf deciduous forest
3. Mixed forest	6. Mixed forest
4. Shrubland	7. Tropical or sub-tropical shrubland
	8. Temperate or sub-polar shrubland
5. Herbaceous	9. Tropical or sub-tropical grassland
	10. Temperate or sub-polar grassland
6. Lichens/moss	*11. Sub-polar or polar shrubland-lichen-moss*
	12. Sub-polar or polar grassland-lichen-moss
	13. Sub-polar or polar barren-lichen-moss
7. Wetland	14. Wetland
8. Cropland	15. Cropland
9. Barren land	16. Barren land
10. Urban and built-up	17. Urban and built-up
11. Water	18. Water
12. Snow and ice	19. Snow and ice

Figure 1. Legend of NALCMS for Level I and II. Legend Level III is country-specific and linked to LCCS. Italic classes do not exist in Mexico.

The North American Land Change Monitoring System (NALCMS) attempts to provide homogeneous land cover and land cover change products at an annual interval for the North American Continent. It is a trilateral effort of institutions in Canada (Canadian Centre for Remote Sensing, CCRS), the United States (United States Geological Survey, USGS), and Mexico (National Institute for Statistics and Geography, INEGI; National Commission for Forestry, CONAFOR; National Commission for the Knowledge and Use of Biodiversity, CONABIO) and is united by the intergovernmental Commission for Environmental Cooperation (CEC). The attempted products range from medium-resolution land cover maps and fractional vegetation products to higher resolution change products. The current research focuses on a 2005 baseline map of important land cover classes at 250m resolution. In this respect a common legend at three levels was defined (Figure 1 depicts level I and II); the third level is country-specific and linked to the Land Cover Classification System (LCCS). The map was recently finished and is available in printed format at the CEC. The digital data will be released in the near future and hosted by CEC.

This paper describes the map production process for the Mexican portion of the study area. Besides the INEGI vegetation maps (INEGI 2005) there is only a small number of studies mapping Mexico’s land cover, e.g. Giri and Jenkins (2005) or Turcotte et al. (1993). In this sense this is the first attempt to build a robust land cover monitoring system for Mexico. Therefore several sources of data had to be combined to allow for high map accuracy and map consistency among different years. Land cover mapping in Mexico is particularly challenging because of the spatial heterogeneity of surface types. Mexico is among the countries with the highest biodiversity (Mittermeier and Mittermeier, 1992). Its great number of ecoregions ranges from sub-tropical desert lands in the North to tropical shrublands and forests in the South. In addition to the unique location, the orography with a high plateau in the central section and ranges in the East and West leads to temperate ecosystems. The flat and karstic Yucatan Peninsula exposes tropical forests and shrublands.

DATA

This study employed seven reflective bands (visible blue to short-wave infrared) of MODIS-Terra top-of-atmosphere surface radiance data and the NDVI. The Canadian Centre for Remote Sensing (CCRS) processed high-quality continental datasets of monthly composites for 2005 and 2006 referenced to the Lambert Azimuthal Equal

Area (LAEA) projection (Figure 2a, 2b). A fusion method developed by Luo et al. (2008) downscaled all 500m channels to a spatial resolution of 250m. This homogeneous dataset was employed by all participating institutions to produce their country land cover map.

Several studies indicated the use of metrics (DeFries et al. 1995, Hansen et al. 2000), which are simple statistics such as mean, standard deviation, minimum, maximum, and range for a period of time (year, half-year, four-month, and quarter-year periods). In addition, ancillary information accompanied the satellite data for better class discrimination, using data such as digital elevation model (Figure 2c) and derived slope and aspect, mean, maximum, and minimum temperature, total of precipitation and days with precipitation, and the number of green days. The final map of Mexico combines classifications based on the monthly composites with classifications based on metrics.

A sample data base was built to train and test the land cover classification. Several sources of data were combined to provide a sufficient sample size for each class. Besides the integration of several field-based sample sets, preliminary analysis indicated the need for additional samples for some classes or underrepresented regions, which were digitized from GoogleEarth. In total a base of approximately 121,000 samples was compiled (Figure 2d).

As a reference the INEGI Series-III vegetation map (INEGI 2005) was recoded to the NALCMS legend. Although it is not the goal of this study to resemble this recoded vegetation map, the dataset served as an orientation of an approximate spatial distribution of land cover classes. Other sources were masks to *a posteriori* correct for spatial overestimations of some classes as well as masks for water and classes with insufficient spatial distribution to be mapped in the classification process (e.g. snow and ice).

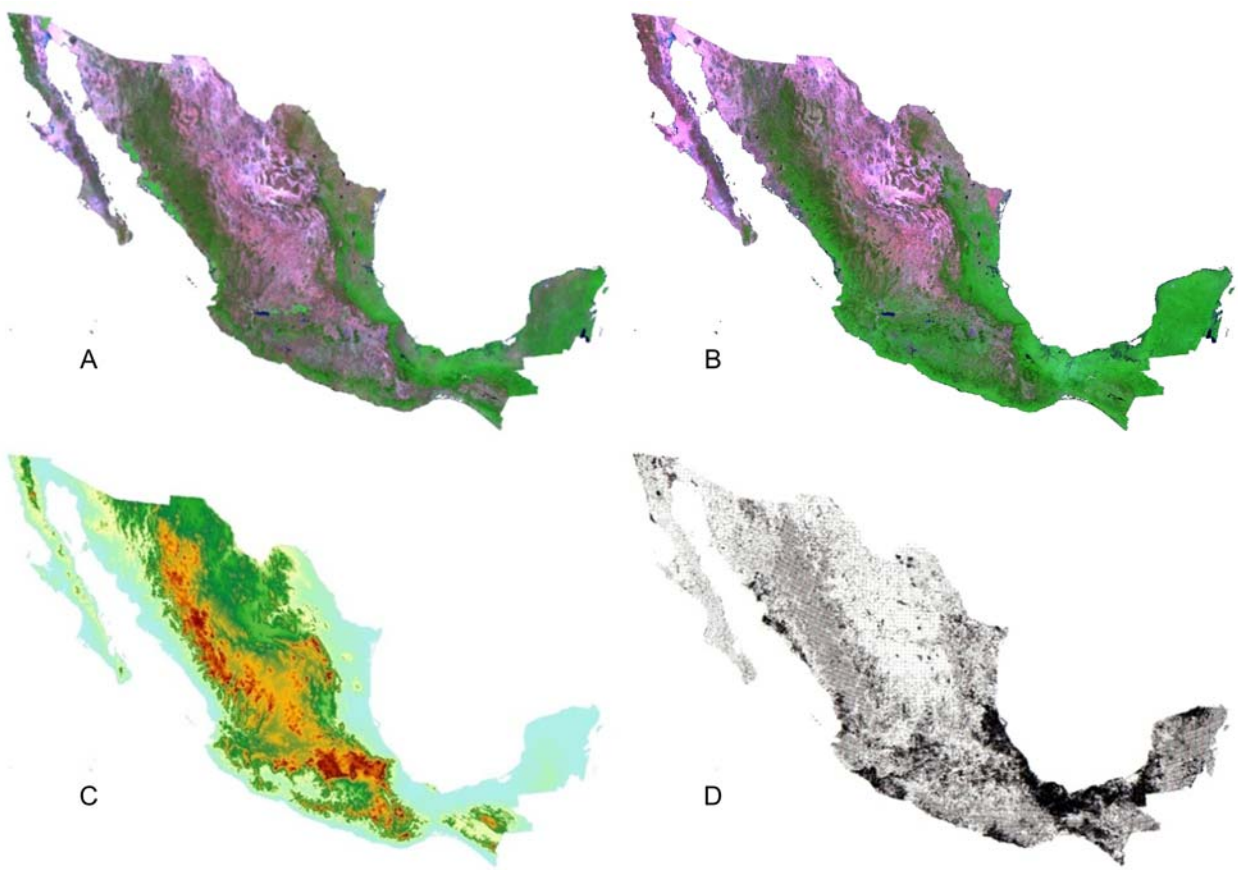

Figure 2. Selected datasets for study. A: MODIS composite of March 2005, band combination 7, 2, 1 (RGB). B: MODIS composite of October 2005, band combination 7, 2, 1 (RGB). B: Elevation. C: Sample data density.

METHODS

In a first step the sample data were split per class into 80% for training and 20 % for validation. It was ensured that sample data are equally distributed in space. Following, each set was superimposed by the reference map and samples not coinciding with the land cover patch or closer than 500m to another class were eliminated. On average this reduced the samples by 50%, however, available samples of some fragmented classes such as urban were diminished to only 5 %. Preliminary testing, however showed that classification accuracy significantly improved when samples were cleaned.

As classification algorithm the C5.0 classifier (Quinlan 1993) was employed and executed in boosting mode. Individual classifications were combined using the confidence estimate equation (Quinlan 1996) resulting in class membership estimations per pixel. Several classifications of varying input data and sample sets were averaged. Class memberships were transformed to discrete classes and accompanied by a confidence estimate. *A posteriori* corrections for overestimated classes urban and wetland in semi-arid areas were applied, for which the runners-up classes of the class membership-to-discrete map transformation process were assigned. Finally, stable masks for water and snow/ice were superimposed.

RESULTS

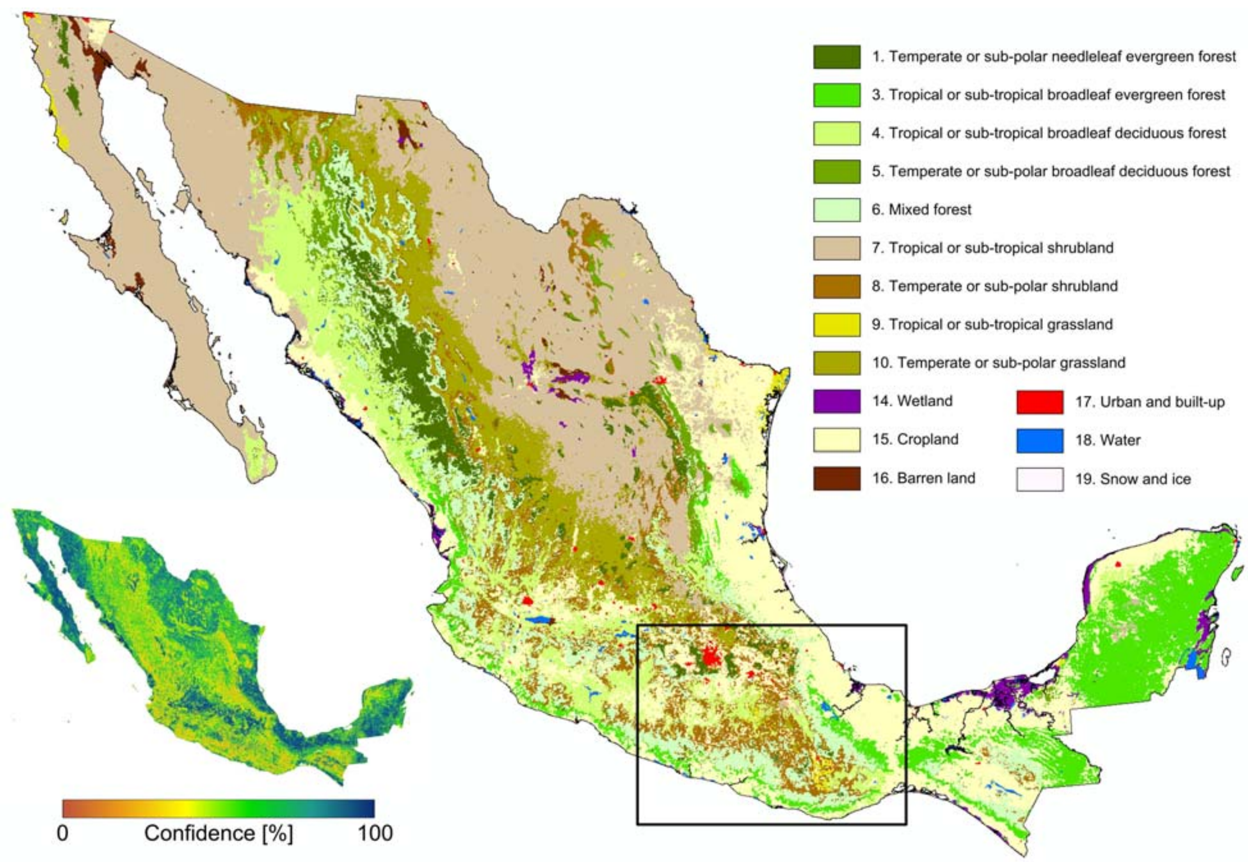

Figure 3. Land cover map of Mexico for 2005. Corrections for classes urban and wetland were applied and stable masks for classes “Water” and “Snow and Ice” were superimposed. The confidence layer estimates the probability of the discrete class. The box locates the close-up of Figure 4.

The final land cover map of Mexico is depicted in Figure 3. From a qualitative point of few there are no significant misclassifications in the automatically generated map. The accompanying confidence layer estimates the probability of the discrete class. It indicates the heterogeneity of land cover types in the central zone and areas of the eastern, western, and southern mountain ranges. The central zone of the country is intensely used and coincides also with the highly biodiverse Mexican transition zone mixing temperate and tropical species.

A close-up of the southern portion with the Mexico City in the North and Oaxaca in the southern portion is shown in Figure 4 and superimposed by the ecoregion boundaries of North America. The ecoregion polygons coincide with the class composition of certain areas. For instance, tropical evergreen broadleaf forest is only located in the coastal lowlands along the Pacific and the mountain ranges between the states of Veracruz and Oaxaca. Forests at higher elevation is tropical broadleaf deciduous, temperate, or mixed forest. Forests at even higher elevations change to needleleaf in a temperate climate, e.g. the mountains of the volcanic belt with the Nevado de Toluca West of the city of Toluca, the mountains West and South of Mexico City as well as the peaks of Iztaccíhuatl and Popocatépetl on the eastern side, La Malinche north of Puebla, and Pico de Orizaba. The gulf coastal plains were correctly identified as cropland, which in this classification also includes rangeland. Important wetlands, e.g. the Lagoons of Alvarado, South of the city of Veracruz, was identified correctly. The complex interior with its basins and ranges as mainly identified and temperate shrubland and grassland. The depression of the Río Balsas in the center-west, however, is covered by a complex pattern of tropical deciduous forest and cropland, intermingles with some shrublands.

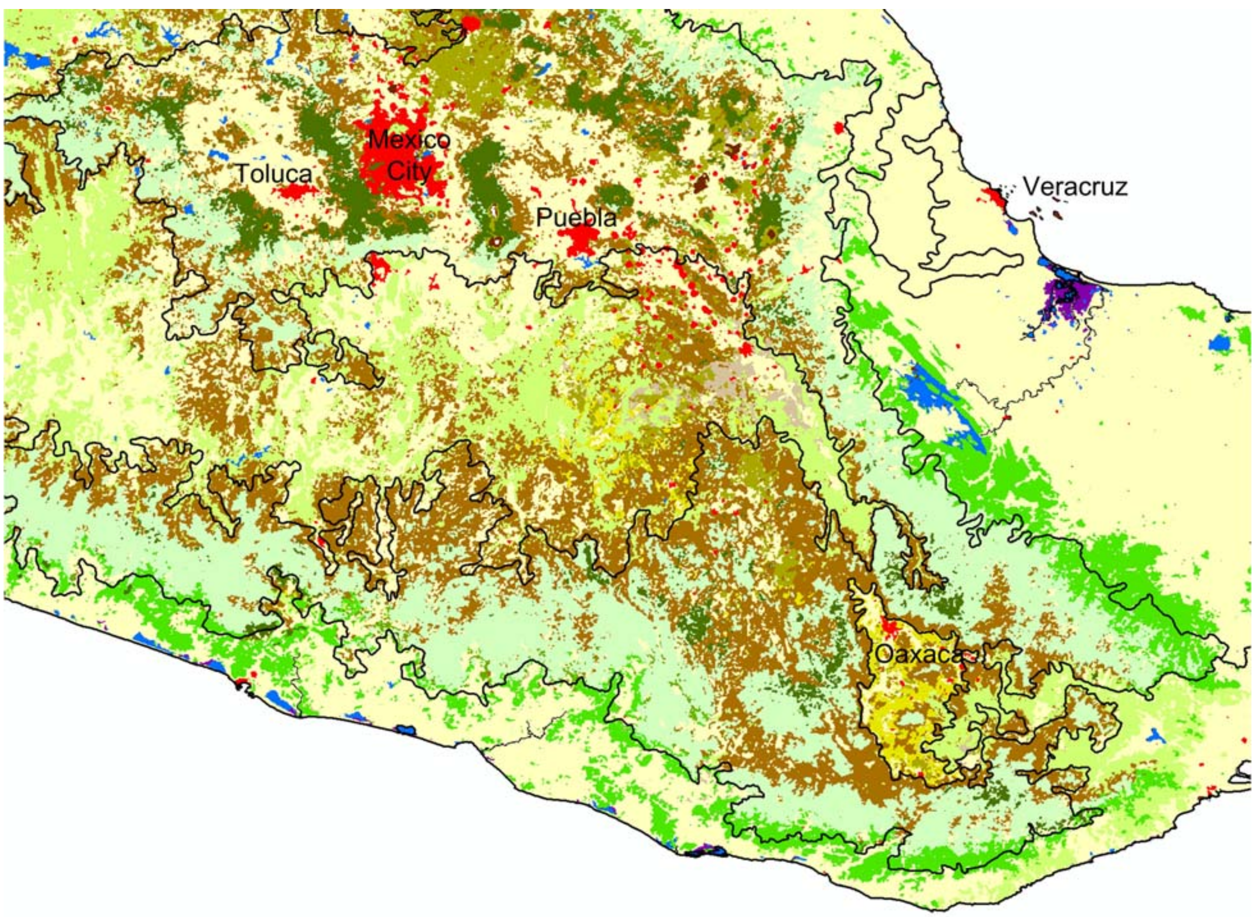

Figure 4. Close-up of land cover map 2005 for central Mexico. For location see box in Figure 3.

Table 1. Producer's and Users accuracies for mapped classes during the automated classification.

ID	Class	Prod.	Users
1	Temperate or sub-polar needleleaf evergreen forest	60.4	73.9
3	Tropical or sub-tropical broadleaf evergreen forest	88.9	74.6
4	Tropical or sub-tropical broadleaf deciduous forest	84.3	67.1
5	Temperate or sub-polar broadleaf deciduous forest	66.7	74.6
6	Mixed forest	80.1	62.9
7	Tropical or sub-tropical shrubland	94.9	79.7
8	Temperate or sub-polar shrubland	34.7	49.4
9	Tropical or sub-tropical grassland	63.1	81.9
10	Temperate or sub-polar grassland	59.9	36.1
14	Wetland	62.2	96.4
15	Cropland	73.5	97.2
16	Barren land	50.0	94.6
17	Urban and built-up	58.2	89.1

As a global error measure the normalized overall accuracy, estimated to be 82% for the 2005 map, deemed useful because it balances between unequal sample size per class (Congalton and Green 1999). Table 1 reports the users and producer's accuracies for all automatically generated classes of the product. While most classes have an acceptable accuracy between 80 and 85%, there are misclassifications for mixed forest which was confused with needleleaf and broadleaf forest classes. Also the subdivision into climatic groups tropical and temperate for shrubland and grassland caused confusion. Only dense, large urban centers were mapped successfully which caused low Producer's accuracies. Although field-based sample data existed, Mexican settlement structures in the countryside are intermingled with various other land surface types (small-patch cropland, forest, shrubland or grassland) and cannot be detected with medium resolution data.

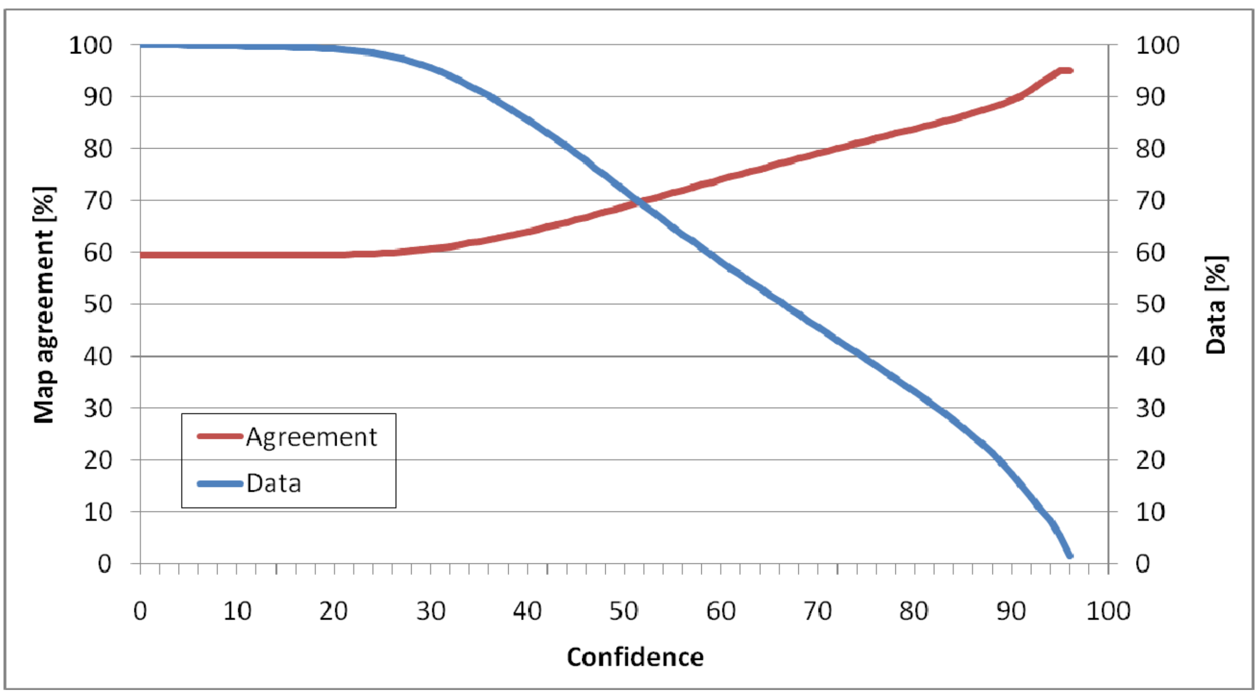

Figure 5. Map agreement between reference map and classification as a function of map confidence.

The comparison of the reference, the reclassification of the INEGI vegetation map to the NALCMS legend, with the classification as a function of map confidence is shown in Figure 5. The map agreement for all pixels is 60%. However, the agreement increases to 70% for pixels of 50% and more confidence (70% of data) and to 82% for pixels of more than 75% confidence, corresponding to 40 % of the data. It should be noted that it was never attempted to achieve the same map as the reference map, which can only partly be considered as land cover dataset. This also explains the low map agreement for comparing all pixels.

Ongoing studies of the project focus on the map consistency and change detection comparing the classification based on data of 2005 to 2006. First results indicate a high map agreement of discrete classifications among both years, which increases substantially to approximately 95% for higher confidence assignments. This encouraging result provides possibilities of producing change alarm masks, for instance when contiguous high-confidence pixels have changed. In a second step the class membership maps will be directly employed for change alarm mask generation. Detailed studies with high spatial resolution data can focus on the indicated areas and provide the necessary detail of changes by vector and magnitude (Lambin and Strahler 1994).

CONCLUSIONS

The North American Land Change Monitoring System (NALCMS) is an international effort to provide land cover products of this continent at an annual basis. The methodology of map production for the Mexican territory with data of 2005 was described. The unique product provides a high map accuracy in qualitative as well as quantitative terms and makes this map useful for land cover studies at a continental to regional scale. In fact, it already has been compared to results of recent species distribution studies for Oak forests in Mexico which yielded a good agreement (Cord et al. 2009). The Mexican map has been integrated into the continental map and this product is being released by the Commission of Environmental Cooperation (CEC). Preliminary results comparing the map of 2005 and 2006 also suggest a very high map coincidence. Ongoing work focuses on robust change alarm methodologies to mitigate false change detection.

ACKNOWLEDGEMENTS

We thank Rasim Latifovic and Darren Pouliot from the Canadian Centre for Remote Sensing (CCRS) for providing the monthly composites of MODIS radiance data at 250m spatial resolution. In addition we are thankful to all NALCMS team members for their fruitful inputs during our discussions. Details of the NALCMS project as well as the map in printed and digital format can be obtained from the Commission of Environmental Cooperation (CEC) webpage: (http://www.cec.org/naatlas/nalcms.cfm).

REFERENCES

Bartholomé, E. and A.S. Belward, 2005. GLC2000: A new approach to global land cover mapping from Earth observation data< *International Journal of Remote Sensing*, 26(9), pp. 1959-1977.

Congalton, R.G. and K. Green, 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, 160 p.

Cord, A., R.R. Colditz, M. Schmidt, and S. Dech, 2009. Species distribution and forest type mapping in Mexico, *IEEE International Conference on Geoscience and Remote Sensing, 2009, IGARSS 2009*, July 13^{th} - 17^{th} 2009, Cape Town, South Africa.

Bicheron, P., M. Leroy, C. Brockmann, U. Krämer, B. Miras, M. Huc, F. Ninõ, P. Defourny, C. Vancutsem, O. Arino, F. Ranéra, D. Petit, V. Amberg, B. Berthelot, and D. Gross, 2006. GLOBCOVER : A 300m global land cover product for 2005 using ENVISAT/MERIS time series, in *Proceedings of the Recent Advances in Quantitative Remote Sensing Symposium*, Valencia, September 2006.

DeFries, R.S., M.C. Hansen, and J.R.G. Townshend, 1995. Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data, *Remote Sensing of Environment*, 54(3), pp. 209-222.

Friedl, M.A., D.K. McIver, J.C.F. Hodges, X.Y. Zhang, D. Muchoney, A.H. Strahler, C.E. Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C.B. Schaaf, 2002. Global land cover mapping from MODIS: Algorithms and early results, *Remote Sensing of Environment*, 83(1-2), pp. 287-302.

Giri, C. and C. Jenkins, 2005. Land cover mapping of Greater Mesoamerica using MODIS data, *Canadian Journal of Remote Sensing*, 31(4), pp. 274-282.

Hansen, M.C., R.S. DeFries, J.R.G. Townshend, and R.A. Sohlberg, 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach, *International Journal of Remote Sensing*, 21(6-7), pp. 1331-1364.

INEGI, 2005. Dirección General de Geografía - 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.

Lambin, E.F. and A.H. Strahler, 1994. Change vector analysis in multitemporal space: A tool to detect and categorize land-cover change processes using high temporal-resolution satellite data, *Remote Sensing of Environment*, 48(2), pp. 231-244.

Loveland, T.R., B.C. Reed, J.F. Brown, D.O. Ohlen, Z. Zhu, L. Yang, and J.W. Merchant, 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *International Journal of Remote Sensing*, 21(6), pp. 1303-1330.

Luo, Y., A.P. Trishchenko, and K.V. Khlopenkov, 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(12), p.p. 4167-4185.

Mittermeier, R.A. and C. Mittermeier, 1992. La importancia de la diversidad biológica de México, In: Sarukhán, J. and Dirzo, R. (Eds.). *México ante los retos de la biodiversidad*. CONABIO, México, pp. 63-73.

Quinlan, J.R., 1993. *C4.5 Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA, 302 p.

Quinlan, J.R., 1996. Bagging, boosting, and C4.5, in *Proceedings of the 14th National Conference on Artificial Intelligence*, Menlo Park, CA, pp. 725-730.

Turcotte, K.M., K. Lulla, and G. Venugopal, 1993. Mapping small-scale vegetation changes of Mexico, *Geocarto International*, 8 (4), pp. 73–85.