

A PROPER LAND COVER AND FOREST TYPE CLASSIFICATION SCHEME FOR MEXICO

S. Gebhardt^{a,*}, P. Maeda^a, T. Wehrmann^a, Jesús Argumedo Espinoza^b, M. Schmidt^a

^a National Commission for the Knowledge and Use of Biodiversity (CONABIO), Liga Periférico-Insurgentes Sur 4903, Parques del Pedregal, 14010 Tlalpan, Mexico City, Mexico - steffen.gebhardt@conabio.gob.mx, pedro.diaz@conabio.gob.mx, thilo.wehrmann@conabio.gob.mx, michael.schmidt@conabio.gob.mx

^b National Institute for Statistics and Geography (INEGI), Héroe de Nacozari 2301 Sur, Jardines del Parque, 20270 Aguascalientes, Aguascalientes, Mexico - jesus.argumedo@inegi.org.mx

KEY WORDS: Land cover, Forest type, Landsat time series, Automatic classification, Wall-2-wall, REDD+

ABSTRACT:

The imminent implementation of a REDD+ MRV system in Mexico in 2015, demanding operational annual land cover change reporting, requires highly accurate, annual and high resolution forest type maps; not only for monitoring but also to establish the historical baseline from the 1990s onwards. The employment of any supervised classifier demands exhaustive definition of land cover classes and the representation of all classes in the training stage. This paper reports the process of a data driven class separability analysis and the definition and application of a national land cover classification scheme. All Landsat data recorded over Mexico in the year 2000 with cloud coverage below 10 percent and a national digital elevation model have been used. Automatic wall-2-wall image classification has been performed trained by national reference data on land use and vegetation types with 66 classes. Validation has been performed against field plots of the national forest inventory. Groups of non-separable classes have subsequently been discerned by automatic iterative class aggregation. Class aggregations have finally been manually revised and modified towards a proposed national land cover classification scheme at 4 levels with 35 classes at the highest level including 13 classes for primary temperate and tropical forests, 2 classes for secondary temperate and tropical forest, 1 for induced or cultivated forest, as also 8 different scrubland classes. The remaining 11 classes cover agriculture, grassland, wetland, water bodies, urban and other vegetation land cover classes. The remaining 3 levels provide further hierachic aggregations with 14, 10, and 8 classes, respectively. Trained by the relabeled training dataset wall-2-wall classification towards the 35 classes has been performed. The final national land cover dataset has been validated against more than 200,000 polygons randomly distributed all over the country with class labels derived by manual interpretation. The agreement for all 35 classes at level 4 was 71%. Primary forest classes have been identified with accuracies between 60% and 83%. Secondary forest classes rated only 50% finding major confusion with the primary forest classes. Accuracies over the scrubland classes have been calculated between 60% and 90%. Agreements for aggregated temperate and tropical forest classes was 85% and 80%, respectively. Separation of forest and non-forest has been achieved with an agreement of 87%.

1. MANUSCRIPT

Mexico has sported for the last 25 years national cartography featuring extremely high thematic resolution with more than 200 classes of vegetation types and land use types. The so-called INEGI (Instituto Nacional de Estadística y Geografía) series I-V have been giving the user an excellent overview of Mexico's main vegetation types at 5 to 10 years intervals (INEGI, 2005, 2008). While INEGI series are each in itself highly accurate in representing the vegetation and land use of Mexico, as a series they show serious limitations due to frequent switches in production methods, class schemes and data used (from aerial photos over Landsat to SPOT5). Moreover, INEGI series at 1:250,000 with a minimum mapping unit (MMU) of 25-50 ha are not suitable for modern planning, decision making and also not for reporting greenhouse gas (GHG) emission for instance. The temporal resolution of 5 to 10 years is also not suitable to enable Mexico's national commission for forestry (CONAFOR) maintaining it's objective to reduce emissions from land use cover change (LUCC) in forests by 50% compared to the 1990 level until 2020. Considering the additional demand for highly accurate frequent and hi resolution maps due to the current implementation of the REDD+ mechanism in Mexico, demanding operational annual

LUCC reporting by 2015, Mexico's national commissions for forestry and biodiversity (CONAFOR and CONABIO) supported by INEGI have started in 2011 developing an automated method to map at high accuracies, high resolution and frequent intervals (annually) the land cover changes over the whole federal territory using a highly efficient processing system called MAD-Mex (Monitoring Activity Data for the Mexican REDD+ program), which allows employing a suit of different optical sensors to render maps at 1:20,000 and 1:100,000 annually in full compliance with the national mapping standards established by INEGI and with a classification scheme fully compatible to INEGI's hierarchical scheme.

In digital image analysis land cover classification describes the process of transforming image data (satellite imagery, auxiliary information) into discrete maps of land cover units on the earth's surface. While manual image interpretation might still produce most confident and accurate land cover maps it is an exhaustive, time-consuming but also subjective process and it is by far not applicable for larger regions or even at a national scale. Algorithms supporting automatic image classification can be generally divided into supervised and unsupervised algorithms. While unsupervised algorithms require the definition of land cover classes after clustering similarity or

* Corresponding author

dissimilarity in image data, supervised approaches require definition of classes in all possible combinations beforehand the classification (Di Gregorio, 2005). The correct use of any supervised classifier requires that all classes that occur in the study area be included in the training stage of the analysis (Congalton and Green, 2008; Foody et al., 2006). Inappropriate class definition or missing input variables for discriminating classes will inevitably lead to inaccurate and blurred classification results (Heinl et al., 2009). The failure to exhaustively define classes can result in substantial errors which may also remain undetected in a subsequent classification accuracy assessment (Foody, 2002). The exclusion classes in the training stage will typically result in cases of untrained classes being commissioned into the set of classes upon which the classifier was trained (Foody et al., 2006). The definition and selection of land cover classes, however, has shown to be crucial and not to be simply adaptable from existing land cover class schemes and a stronger focus must be put towards discriminating land cover classes by their typical spectral, topographic or seasonal properties (Heinl et al., 2009).

The incorporation of a preliminary discrimination analysis, based on the actual image data to be used in the final classification process, requires knowledge of all present and potentially spectrally discernible classes in the study area and the availability of independent training and reference data for each class. Whilst for manual class identification over a single scene or small mosaicked scenes this seems feasible, it is utterly beyond reach for large area mapping endeavors. For instance, a wall-to-wall land cover classification over Mexico or countries of similar size requires about 130 distinct Landsat scenes or about 4000 individual RapidEye scenes. It is pretty obvious that scene based class separability analysis based on manually discerned training data is not an option. The same holds true for reference data generation. All this under the assumption that the classification has a reasonable production time.

In a previous paper the MAD-Mex system has been introduced (Gebhardt et al., 2014). It is one aim of the system to automatically produce a national land cover dataset in a standardized, transparent and transferable way. Amongst others, one key application of MAD-Mex is an automatic wall-to-wall land cover classification using Landsat time series towards a national reference land cover dataset. In the before-mentioned publication the classification scheme defined 9 and 12 classes at two hierarchical levels. Overall accuracies achieved were up to 76%. Tropical and temperate forests have been classified with accuracies of 78% and 82%, respectively. However, the thematic depth especially in the forest classes was rather shallow and it is now apparent that the full potential of the system to discern various land cover classes based on the given data and methodology had not been fully explored.

This paper presents the methodology and application of a data driven class separability analysis towards a land cover scheme definition for national mapping in operational fashion. Emphasis is on the discrimination of forest vegetation types due to its main application, REDD+ MRV. It employs four different and completely independent data sources. These are i) all Landsat 7 and Landsat 5 images available in the current USGS archive over Mexico of the year 2000 featuring a cloud cover percentage of less than 10% plus a national digital elevation model with a 30m resolution (INEGI, 2010); ii) national vegetation type and land use maps from Mexican national statistics bureau (INEGI) at a scale of 1:250,000 representing about 70 different thematic classes; iii) about 14,000 field plots extracted of the national forest inventory, cycle 2004 to 2007 (CONAFOR, 2007); and iv) more than 240,000 polygons derived from Landsat 2000 imagery randomly distributed all

over the country with class labels derived by manual interpretation.

2. MATERIALS

2.1 Landsat data and digital elevation model

A full Landsat coverage for Mexico is achieved with 135 distinct path/rows (tiles). All Landsat TM and ETM+ Level 1T data from the year 2000 with maximum cloud coverage (metadata based) of less than 10 percent have been used resulting in a total of 2518 Landsat scenes. A national Digital Elevation Model (DEM) with 30 m resolution, provided by the National Institute of Statistics and Geography of Mexico (INEGI, 2010), of which aspect and slope images have been additionally processed.

2.2 Training data

INEGI creates and publishes national vegetation type maps (Uso de Suelo y Vegetación, USV) at a scale of 1:250,000 using a 25 to 50 hectare minimum mapping unit. Those maps differentiate 73 thematic classes including different vegetation types and areas devoted to agriculture, livestock, and forestry. From these 73 classes 46 classes define primary forest and shrub types. Furthermore, the applied classification scheme also enables separation of primary and three secondary forest vegetation classes. Thus, INEGI discerns in its standard classification scheme (USV) 212 land use and vegetation types. However, the maximum number of classes assigned in all so far published land use and vegetation series is 176. From the five map series published so far, series II (1993-1999), III (2002-2004), and IV (2005-2007) have been used to generate a national training dataset (INEGI, 2009, 2008, 2005). In order to account for eventual land cover change and map inconsistencies the three discrete maps have been filtered for persistency by conserving only those objects which land use attributes, not including secondary forest attributes, have not changed in all three map series. By doing so only 66 out of 73 classes could be conserved. The dataset, denominated “persistent areas dataset”, has been rasterized to a national training dataset with a 125 m pixel size designating the 66 different land use and vegetation type classes as listed in Table 1 of the appendix. This dataset served as sole source for classifier training.

2.3 National forest inventory data and manual interpreted reference polygons

Validation of the classification results and subsequent class aggregation has been based on independent field inventory samples. Forest samples have been available through the National Forest Inventory (INFyS, 2004-2007) provided by the National Forestry Commission (Comisión Nacional Forestal, CONAFOR) (CONAFOR, 2007, 2004). CONAFOR has started a national forest inventory (Inventario Nacional Forestal y de Suelos, INFyS) in 2004 as a set of techniques and procedures to obtain quantitative and qualitative information on forest resources, associated vegetation, components and abiotic parameters of the forest habitat. These plots have been designed with varying distances according to the predominant land cover. Since these inventories aim to be consistent and comparable over time, defined locations of inventory sites are being revisited in defined 5 years time intervals. Amongst other parameters, 57 different forest and vegetation types according to the INEGI USV thematic scheme (Appendix Table 1) are reported. A total of 17,749 inventory sites exist for whole Mexico. Each inventory site is composed of 4 sub-sites hence 70,996 individual sample sites are in place. Until now, INFyS field data from 2 sampling cycles are available. For the period

from 2004 until 2007 data from 17,716 inventory sites and a total of 67,253 individual sample sites are available. In 2010 3,277 inventory sites have been revisited in which a total of 12,630 sample points could be measured. To account for scale differences between individual sub-plots and classified Landsat polygons from the 17,749 inventory sites the 13,923 inventory sites have been selected where all subplots feature the same land use or vegetation type.

A set of 5,125 points derived by random stratified sampling distributed over the Mexican territory has been generated. As strata the land cover classes from the previously published land cover dataset for the year 2000 have been used (Gebhardt et al., 2014). Based on these nodal points image chips of 1 sqkm have been subset from cloud free Landsat images and subsequently been segmented. This resulted in a total of 242,170 polygons, which have been labeled manually assigning the land use and vegetation class, including the state, i.e secondary or primary, disturbed and undisturbed forest.

3. METHODS

3.1 MAD-Mex automatic image classification system

Automatic image processing and classification has been performed by the “Monitoring Activity Data in Mexico” (MAD-Mex) system. The MAD-Mex system has been introduced and described in detail (Gebhardt et al., 2014), so only a short summary is given here. It is the aim of the MAD-Mex system to deliver wall-to-wall land cover and – change information by automatic image classification within a few days for the Mexican territory. Amongst others it features fully automatic land cover classification of Landsat time series.

In a tile based approach all available Landsat imagery is extracted and pre-processed to surface and top-of-atmosphere reflectance using the LEDAPS preprocessing algorithms (Masek, J.G. et al., 2012). In addition to LEDAPS, FMASK is utilized for masking pixels obscured by clouds and cloud shadows (Zhu and Woodcock, 2012). From each individual pre-processed Landsat scene Tasseled Cap components, NDVI, EVI, SR and ARVI vegetation indices are computed. All of those indices and the original spectral bands are then combined to time-series from which selected descriptive statistics for each pixel are calculated. In this process, pixels labeled by FMASK are omitted, thus producing cloud and cloud shadow free image metrics for each calculated feature. i.e. the Tasseled Caps, the vegetation indices, and the spectral bands. As additional feature the gradient image from the mean of the NDVI metrics is calculated using the Sobel operator. Subsequently, image segmentation (*Cloud And Cloud Shadow Assessment*, n.d.) is done on the NDVI metrics and feature extraction is performed on all time-series metrics, the gradient image and on auxiliary elevation, slope and aspect datasets. Segmented objects are then labeled against a training dataset by conserving only those objects, which are spatially intersected by only one class in the training dataset. In order to create a clean training dataset a class based outlier removal (Radoux et al., 2014; Radoux and Defourny, 2010) is then performed based on the three primary bands derived from principal component transformation of the spectral bands metrics. Based on the samples that passed the outlier removal a 10-folded C5 decision tree classifier is trained and applied to all image objects. Classified objects are then transformed to raster representation. Finally, post-processing is applied to minimize effects in the Landsat scene overlapping and to generate a national classification dataset. In the previous methodology (Gebhardt et al., 2014) classification was done on a reduced feature space derived by the ranked principal components where the cumulative explained variability reached

a threshold of 95%. In contrast, the current approach only uses the first three ranked principal components (calculated only from the spectral band metrics) for outlier elimination, while all non-transformed features source the classification, as described above.

Only a given training dataset and a time range of image acquisitions, e.g. one full calendar year, parameterize the MAD-Mex Landsat based classification workflow. MAD-Mex has been developed and is maintained by CONABIO and is currently capable of processing a national Landsat based land cover dataset within half a day.

3.2 Initial wall-to-wall land cover classification, iterative class aggregation, and proposal of national land cover classification scheme

The initial national land cover classification over all Landsat 2000 data has been trained using INEGI data based persistent areas representing 66 distinct land use and vegetation type classes. All 13,923 samples from the INFyS forest inventory representing 57 different vegetation type classes have been used for product validation, which is visualized in a confusion matrix of absolute and relative classification errors. Initially, an iterative class aggregation algorithm had been applied on the confusion matrix. In each iteration, the algorithm identifies the class with the lowest producer's accuracy and searches for the class having major absolute omission and commission errors with it. Those two classes are then aggregated to one class. All respective samples are relabeled to the new class and the confusion matrix is calculated again. The procedure stops when a certain accuracy threshold is met. This has been defined as 60% in this study. With each iteration, the information of which classes have been combined along with their respective accuracies is logged to protocol for subsequent analysis. The applied approach is very much depending on the representation of all classes in both, the reference points and the classified map. This is, however, not always the case as the forest inventory lists classes which had been omitted in the persistent area dataset or which have actually never been available in any of the distinct land use and vegetation type map series, for instance “Medium Tropical Evergreen Forest” (SMP). Moreover, the unequal distribution of the number of inventory samples available per class (as a function of actual area) will influence the applied automated aggregation. This would inevitably result in some gibberish, grouping errors that make no ecological sense. The class aggregation protocols have therefore been analyzed manually in terms of compatibility of aggregated classes. Lastly, a hierarchical scheme of land cover classes had been defined including the corresponding transition rules from the original land use and vegetation types to the corresponding land cover class.

3.3 Wall-2-Wall land cover classification and validation

According to the proposed land cover classification scheme the persistent area training dataset has been relabeled and fed into the MAD-Mex classification workflow. The 242,170 samples from the reference polygons have been relabeled accordingly and used for the validation of the classification product. The validation procedure followed the protocol described by Olofsson et al. (Olofsson et al., 2013). Class based producers and users accuracies were derived from the error matrix as also the error-adjusted area estimates along with their confidence intervals. Finally, classification results and samples have been aggregated up to the higher classification scheme levels and validated accordingly.

4. RESULTS

Initial wall-to-wall classification against the national training dataset comprising 73 distinct classes was computed with the MAD-Mex system implemented at CONABIO. Validation of the final mosaic against all samples from the national forest inventory sourced the iterative class aggregation. The proposed groupings from this process corresponded to the major types of temperate forest (those with codes B*), Tropical Forest (with codes S*), and Scrubland (codes M*), as also Other Vegetation types (V*). One could observe the abrupt increase in producer's accuracy with each aggregation iteration as for example with the first grouping of oak Forest (BQ, classified with 52% accuracy) with oak-pine Mixed Forest (with only 33% accuracy) to the joined accuracy of 60.4%. It also reflected that some actual vegetation type classes could already be classified at very promising accuracies for example the Coastal Rosetophyllous Scrub (MRC) with 85%, the tall Tropical Evergreen Forest (SAP) at 82% accuracy or the Medium Tropical Deciduous Forest (SMS) with 71% accuracy. It also showed, however, that some proposed aggregations did not make any sense, for instance Induced Savanna (VSI) and the Medium Tropical Drought Sub-deciduous Forest (SMS). Class aggregations have therefore finally been manually revised and modified towards a proposed national land cover classification scheme at 4 levels with 35 classes at the highest level including 13 classes for primary temperate and tropical forests, 1 for induced or cultivated forest, 2 classes for secondary temperate and tropical forest, and 8 different scrubland classes. The remaining 11 classes cover agriculture, grassland, wetland, water bodies, urban, and other vegetation or land cover classes. The remaining 3 levels provide further hierachic aggregations with 14, 10, and 8 classes, respectively. The final classification scheme is presented in Table 2 of the appendix.

With the training dataset relabeled to the 35 classes at level 4 a new wall-to-wall classification was processed and validated against 242,170 points extracted from the respective reference polygons. Tables 3 of the appendix shows the validation results for the respective class scheme level 4. The first question to answer is the representativeness of the validation points used, i.e. does their quantitative distribution in fact represent the actual area distribution of the reference class. Therefore, in columns %Ai and %ni the percentages of area proportion of a class (calculated based on the INEGI USV series IV map) and the percentage of points available for this class are given, respectively. One can find strongest disagreements of those numbers for classes 2 and 3, however they are below 2.5%. For all classes the difference calculates lower than 1%. From this, one may conclude that the sample data very much represent the spatial distribution of every class and are therefore well suited for a stratified area weighted validation.

Table 3, showing validation results for the 35 classes at level 4, depicts an overall accuracy of 71.28%. The area errors calculated from the estimated area proportions for producers and users accuracies sum up to 26.0 thousand sqkm (1.3% of the country size) and 31.4 thousand sqkm (1.6% of the country size), respectively. With the previous published land cover dataset comprising 12 land cover classes at its second level similar overall accuracy was achieved with 71% (Gebhardt et al., 2014). The respective area errors for producers and users accuracy, however, rated 78.8 and 87.5 thousand sqkm, respectively, which is more than 4% of the country size. With the new proposed classification scheme this error has already been minimized by a factor of about 2.7, but with 35 land cover classes assigned vs. only 12 before.

The primary temperate forest classes feature producer's accuracies between 59.54% and 73.54%, user's accuracies are computed to be 61.48% to 75.28%. Little agreement was found for the secondary temperate forest (class 100) with only 45% producers accuracy. There were no validation data available for the class 1 representing Spruce and Cypress Forest (and this is the reason why it is not displayed in Table 3). According to INEGI's USV reference map this class sums up to 3000 sqkm, which is about 0.015% of the national territory. One may safely assume that exclusion of this class will not seriously effect validation.

Over the 7 primary tropical forest classes producer's accuracies ranged in this study from 52.28% to 82.97%, while user's accuracies were between 63% and 88.63%. The secondary tropical forest class was correctly classified to only 52.34%. The 8 classes dedicated to scrubland rated between 63.11% and 89.55% in producer's accuracy with 63.11% found for the pachycaulous scrub class. The remaining 7 classes showed very high accuracies above 75%. Likewise, very high user's accuracies were found ranging between 73.21% and 88.7%. The non-forest and non-scrubland classes showed producer's accuracies between 59% and 90% and user's accuracies ranging from 51.5% to 90.1% with the gypsophilous and halophilous vegetation being the exception at only 4.76% producer's and 11.76% user's accuracy.

Class aggregation to 12 classes at level 3 provided an increase of overall accuracy to 75.06% and the estimated producer's and user's area errors further decreased to 17.73 and 16.95 thousand sqkm. The 2 remaining primary temperate deciduous and coniferous forest classes (classes 1 and 2) showed 74.21% and 74.73% producer's accuracies while user's accuracies were at 67.09% and 73.47%, respectively. Primary evergreen and deciduous tropical forest classes (3 and 4) rated producer's accuracies of 79.54% and 70.12% with user's accuracies of 79.22% and 73.99%. The aggregated scrubland class (class 5) rated 91.35% and 88.55% in producer's and user's accuracy, respectively. The group of wetland vegetation (class 8) had accuracies of 67.53% and 74%, the other vegetation class (9) calculated 86.29% and 78.20% in producer's and user's accuracy. Except of the 2 secondary forest classes, which naturally preserved their accuracies on level 3, no class existed with producer's accuracies smaller 65%.

At level 2 forest classes are further grouped to only 2 classes representing temperate (class 1) and tropical (class 2) forest. Overall accuracy further increased to 80.65%, estimated area errors decreased to 13 thousand sqkm. Temperate forest calculated 84.86% producer's and 85.41% user's accuracy. The respective numbers for tropical forest were 79.15 and 83.78%. The final level 1 condenses only one forest class, which rated 87.14% and 89.52% in user's and producer's accuracy with an overall classification accuracy that increased to 82.5% and estimated area errors of 11 thousand sqkm.

5. DISCUSSIONS

Direct comparison of the present land cover classification results with the one published previously exhibits a strong increase in thematic resolution but also in classification accuracy. Moreover, the new products shows further decreased errors, especially with respect to estimated area proportions. The availability of a record of 25 years national cartography featuring extremely high thematic resolution of vegetation and land use types has been used as basis for thematic class separability analysis. Because this analysis has been driven by the same satellite data and methods used for automatic land cover classification the resulting land cover scheme definition is

especially adapted to these input data and features that can be extracted from them i.e. Landsat time series.

The classification results at the highest or 4th most detailed hierarchical level of the derived classification scheme defined 35 classes. Validation results exhibited overall an accuracy of 71%. The classes showing highest errors especially with respect to their area have been identified as secondary temperate and tropical forest but also as grassland. These 3 classes already sum up to 513 thousand sqkm, which is 26 percent of the country size.

Grassland, representing 16% of the country, has been identified correctly at a rate of 67%. This class itself is a combination of different classes including natural and induced, managed and unmanaged grasslands but also savannahs, all potentially showing different patterns in a multi-temporal feature space. 10% of misclassifications are found towards agriculture (which is also a combination over different agriculture types potentially having similar feature patterns and spectral properties as some grassland types). Confusions of grassland towards agriculture are therefore very likely. Another 9% are misclassifications to scrubland often showing similar coverage to grassland and with that spectrally difficult to discriminate. However, about 12% are falsely classified towards the different forest classes but especially to the secondary temperate and tropical forest classes. This appears to be likely for secondary forests showing only small trees and sparse tree density.

The two secondary forest classes constitute an area representing 171 thousand sqkm (9% of the country). In the accuracy assessment both rate only about 50%. Both, however, find their major confusions towards their primary forest classes. More than 20% of secondary tropical forest has been classified as a primary tropical forest; over 30% of secondary temperate forest was classified as primary temperate forest.

However, it should be stressed, that MAD-Mex is designed to classify land cover, while the INEGI products and also the expert interpreted image chips feature land use. Therefore, we often encounter a grassland type assigned to agriculture due to its primary use as pasture. Unfortunately, we do not yet possess the tools to quantify this in our products.

Except from the discussed erroneous classes the results show very good results. With respect to vegetation classes and focusing forest types, this is especially true for the different primary forest type classes over both, tropical and temperate forest, as also for the scrubland classes. The MAD-Mex land cover classification benefits from a very regionalized approach. For Landsat every path/row is individually classified based on the respective multi-temporal image record for a given year. With that the classifier can be perfectly trained for land cover classes over that region. Thus, the training does not include non-present land cover types in the respective path/row. The latter occurs when training over a harmonized, mosaicked image data set to obtain a wall-to-wall classification. Therefore, high thematic class diversity can be kept, especially over scrublands and forests. However, we finally state that for Mexico the 35 proposed land cover classes seem to define the maximum feasible with automated Landsat image classification under the conditions we presented here. The year 2000 features the best temporal resolution in Landsat data availability over Mexico with combined acquisitions from the Thematic Mapper (TM) and the then still intact Enhanced Thematic Mapper Plus (ETM+). One should expect lower qualities for classification results during the 90s after 2003 with the ETM+ scanline defect and almost no TM acquisitions over Mexico.

It remains to be seen, how the here presented results can be replicated in the near future with combined Landsat-8 and Sentinel-2 acquisitions (hopefully) available.

REFERENCES

Cloud And Cloud Shadow Assessment, n.d. . Geospatial Data Analysis Corporation.

CONAFOR, C.N.F. (CONAFOR), 2007. *Inventario Nacional Forestal y de Suelos (INFyS 2004–2007)*.

CONAFOR, C.N.F. (CONAFOR), 2004. *Inventario Nacional Forestal y de Suelos (INFyS 2004–2007)*. Manual y procedimientos para el muestreo de campo.

Congalton, R.G., Green, K., 2008. *Assessing the accuracy of remotely sensed data: principles and practices*. CRC press.

Di Gregorio, A., 2005. Land cover classification system: classification concepts and user manual: LCCS. Food & Agriculture Org.

Foody, G.M., 2002. Hard and soft classifications by a neural network with a non-exhaustively defined set of classes. *Int. J. Remote Sens.* 23, 3853–3864.

Foody, G.M., Mathur, A., Sanchez-Hernandez, C., Boyd, D.S., 2006. Training set size requirements for the classification of a specific class. *Remote Sens. Environ.* 104, 1–14. doi:10.1016/j.rse.2006.03.004

Gebhardt, S., Wehrmann, T., Ruiz, M.A.M., Maeda, P., Bishop, J., Schramm, M., Kopeinig, R., Cartus, O., Kellndorfer, J., Ressl, R., Santos, L.A., Schmidt, M., 2014. MAD-MEX: Automatic Wall-to-Wall Land Cover Monitoring for the Mexican REDD-MRV Program Using All Landsat Data. *Remote Sens.* 6, 3923–3943. doi:10.3390/rs6053923

Heinl, M., Walde, J., Tappeiner, G., Tappeiner, U., 2009. Classifiers vs. input variables - The drivers in image classification for land cover mapping. *Int. J. Appl. Earth Obs. Geoinformation* 11, 423–430. doi:10.1016/j.jag.2009.08.002

INEGI, I.N. de E. y G. (INEGI), 2010. Continuo de Elevaciones Mexicano 2.0 (CEM 2.0).

INEGI, I.N. de E. y G. (INEGI), 2009. Guía para la interpretación de cartografía uso del suelo y vegetación: Escala 1:250 000 Serie III. Instituto Nacional de Estadística, Geografía e Informática (Mexico).

INEGI, I.N. de E. y G. (INEGI), 2008. Conjunto de datos vectoriales de la carta de uso del suelo y vegetación, Escala 1:250,000, Serie IV (CONTINUO NACIONAL).

INEGI, I.N. de E. y G. (INEGI), 2005. Conjunto de datos vectoriales de la carta de uso del suelo y vegetación, Escala 1:250,000, Serie III (CONTINUO NACIONAL).

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, J., Lim, T.K., 2012. LEDAPS Landsat Calibration, Reflectance, Atmospheric Correction Preprocessing Code. Model product. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.

Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* 129, 122–131. doi:10.1016/j.rse.2012.10.031

Radoux, J., Defourny, P., 2010. Automated Image-to-Map Discrepancy Detection using Iterative Trimming. *Photogramm. Eng. Remote Sens.* 76, 173–181.

Radoux, J., Lamarche, C., Van Bogaert, E., Bontemps, S., Brockmann, C., Defourny, P., 2014. Automated Training Sample Extraction for Global Land Cover Mapping. *Remote Sens.* 6, 3965–3987. doi:10.3390/rs6053965

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* 118, 83–94. doi:10.1016/j.rse.2011.10.028

APPENDIX

Table 1: INEGI land use and vegetation type classes. Classes labeled with * could not be conserved in the persistent training dataset.

Code	Vegetation type	Code	Vegetation type	Code	Vegetation type
BS	Douglas-fir and spruce forest	MST	Subtropical shrubland / Subtropical scrub	SMS	Medium tropical drought subdeciduous forest
BB	Cypressus forest	MK	Thorn forest	SMQ	Medium tropical seasonal evergreen rain forest
BQ	Oak forest	MKX	Desertic thorn forest	DV	Without vegetation
BQP	Oak-Pine forest	MKE	Tropical thorn forest	VT	Cattail marsh
BG	Gallery forest	VP	Induced Palm association	VD	Sand desert's vegetation
BMK*	Thorn forest	VPN	Natural Palm association	VU	Coastal dune's vegetation
BA(AO)	Boreal forest	PY	Gypsophilous grassland	VG	Gallery vegetation
BP	Pine forest	PH	Halophilous grassland	PT	Peten vegetation
BPQ	Pine-Oak forest	PI	Induced grassland	VY	Gypsophilous vegetation
BJ	Juniperus forest	PN	Natural grassland	VH	Halophilous vegetation
BI	Induced forest	VA	Arrowroot marsh	VHH	Halophilous hydrophilous vegetation
BM	Cloud forest	VW	Alpine meadow	VHX*	Halophilous dry vegetation
BBA*	Low-open forest	VS	Savanna	Code	Agricultural usage
ML	Chaparral	VSI	Induced savanna	TA	Rainfed cropland / Seasonal agriculture
VM	Mangrove	SAP	Tall tropical evergreen rain forest	BC	Cultivated forest
MC	Pachycaulous shrubland / Pachycaulous scrub	SAQ	Tall tropical seasonal evergreen rain forest	RA	Irrigated cropland / Irrigated agriculture
MJ	Coniferous shrubland / Coniferous scrub	SBC	Low tropical drought deciduous forest	PC	Cultivated grassland
MDM	Desertic microphyllous shrubland / Desertic microphyllous scrub	SBK	Low tropical thorn drought deciduous forest	ACUI*	Aquaculture
MDR	Desertic rosetophyllous shrubland / Desertic rosetophyllous scrub	SBEQ*	Low tropical thorn seasonal evergreen forest	HA	Moisture agriculture
MET	Tamaulipan thorn shrubland / Tamaulipan thorn scrub	SBP	Low tropical evergreen forest	Code	Others
MRC	Coastal rosetophyllous shrubland / Coastal rosetophyllous scrub	SBS	Low tropical drought subdeciduous forest	ZU	Urban area
MSC	Sarcocaulous shrubland / Sarcocaulous scrub	SBQ	Low tropical seasonal evergreen forest	AH*	Human settlements
MSCC	Sarco-pachycaulous shrubland / Sarco-pachycaulous scrub	SG	Tropical gallery forest	NA*	No applicable
MSN	Mist's sarco-pachycaulous shrubland / Mist's sarco-pachycaulous scrub	SMC	Medium tropical drought deciduous forest	ADV	Devoid of vegetation
MSM	Piedmont shrubland / Piedmont scrub	SMP	Medium tropical evergreen rain forest	H2O	Water body

Table 2: Derived national land cover classification scheme at four levels

C1	Level 1	C2	Level 2	C3	Level 3	C4	Level 4	Level keys	USV
1	Forest	1	Temperate forest	1	Coniferous temperate forest	1	Spruce and cypress forest	BS, BB	
							Pine, juniperous and fir forest	BA, BP, BJ	
							Pine-oak mixed forest and coniferous scrub	BPQ, MJ	
				2	Deciduous temperate forest	2	Oak and oak-pine mixed forest	BQ, BQP	
							(Tropical) thorn forest, piedmont and subtropical scrub	MK, MSM, MST, MKE	
							Induced and cultivated forest	BC, BI	
		100	Secondary temperate forest	6	Chaparral	100	Secondary temperate forest	ML	
								B/Vsa	
		2	Tropical forest	3	Evergreen tropical forest	8	Cloud forest and low tropical evergreen forest	BM, SBP	
							Low and medium tropical seasonal evergreen forest, Tropical gallery forest and Palms	BG, VP, VPN, SBQ, SG, SMQ	
							Tall tropical seasonal evergreen forest	SAQ	
						12	Tall and medium tropical evergreen forest	SAP, SMP	
				4	Deciduous tropical forest		Mangrove and peten vegetation	VM, PT	
					9	Low tropical deciduous forest and induced palm	SBC, SBK, SBS		
						Medium tropical deciduous forest	SMC, SMS		
		200	Secondary tropical forest	200	Secondary tropical forest	200	Secondary tropical forest	S/Vsa	
3	Agriculture	4	Agriculture	6	Agriculture	28	Rainfed, irrigated and moisture agriculture	HA, RA, TA	
4	Grassland	5	Grassland	7	Grassland	27	Grassland, meadow and savanna	PC, PY, PH, PI, PN, VW, VS, VSI	
5	Wetland	6	Wetland	8	Wetland	14	Cattail marsh	VT	
						15	Arrowroot marsh	VA	
6	Other vegetation	7	Other vegetation	9	Other vegetation	20	Coastal dune's vegetation	VU	
						24	Sand desert's vegetation	VD	
						25	Halophilous Hydrophilous vegetation	VHH	
						26	Gypsophilous and halophilous vegetation	VY, VH	
7	Unvegetated	8	Bare soil	10	Bare soil	30	Devored and no vegetation	ADV, DV	
		9	Human settlements	11	Human settlements	31	Urban areas and human settlements	ZU, AH	
8	Water	10	Water	12	Water	29	Water and aquaculture	H2O, ACUI	

Table 3: Validation results at level 4

Class	Ai	%Ai	ni	%ni	PA	PA AREA	UA	UA AREA	OA
3	62.25	3.20%	10443	4.31%	71.75 ± 1.42	67.64 ± 0.96	69.48 ± 1.41	64.29 ± 0.91	71.28%
124	66.19	3.41%	9776	4.04%	61.31 ± 1.59	67.06 ± 1.07	61.62 ± 1.58	65.86 ± 1.04	
2	106.85	5.50%	19187	7.92%	73.54 ± 1.23	101.82 ± 1.25	61.48 ± 1.06	127.81 ± 1.35	
4	37.47	1.93%	5628	2.32%	72.75 ± 1.83	43.65 ± 0.80	75.28 ± 1.87	36.21 ± 0.68	
123	0.42	0.02%	93	0.04%	59.54 ± 15.66	0.77 ± 0.12	69.58 ± 18.76	0.36 ± 0.07	
6	17.98	0.93%	2700	1.11%	65.61 ± 2.66	23.56 ± 0.63	74.00 ± 2.88	15.94 ± 0.46	
100	98.43	5.07%	11123	4.59%	45.03 ± 1.33	114.93 ± 1.53	58.62 ± 1.62	75.61 ± 1.23	
8	13.50	0.69%	1674	0.69%	58.28 ± 3.72	13.77 ± 0.51	63.00 ± 3.88	12.49 ± 0.49	
10	57.72	2.97%	6340	2.62%	82.97 ± 1.51	52.17 ± 0.79	82.00 ± 1.34	58.40 ± 0.78	
11	1.47	0.08%	295	0.12%	75.54 ± 7.00	2.01 ± 0.14	88.63 ± 7.67	1.25 ± 0.10	
12	23.94	1.23%	2941	1.21%	78.46 ± 2.49	22.13 ± 0.55	75.91 ± 2.28	24.74 ± 0.56	
16	9.15	0.47%	1377	0.57%	74.17 ± 4.12	8.93 ± 0.37	70.62 ± 3.75	9.61 ± 0.36	
9	89.17	4.59%	11124	4.59%	69.58 ± 1.38	86.80 ± 1.19	70.99 ± 1.32	87.40 ± 1.15	
13	37.76	1.94%	4119	1.70%	66.43 ± 2.03	40.15 ± 0.82	75.77 ± 2.16	33.11 ± 0.72	
200	104.79	5.39%	13089	5.40%	52.34 ± 1.30	124.80 ± 1.62	57.97 ± 1.47	94.61 ± 1.39	
5	238.24	12.26%	27340	11.29%	87.06 ± 0.74	218.89 ± 1.62	77.41 ± 1.60	267.94 ± 4.29	
7	15.47	0.80%	1307	0.54%	63.11 ± 3.12	17.10 ± 0.53	81.28 ± 3.82	12.01 ± 0.46	
22	34.02	1.75%	4373	1.81%	76.36 ± 1.85	36.77 ± 0.68	79.50 ± 1.94	32.68 ± 0.63	
17	23.14	1.19%	2183	0.90%	75.84 ± 2.22	26.77 ± 0.59	88.77 ± 2.46	19.77 ± 0.49	
19	52.88	2.72%	5687	2.35%	86.07 ± 1.62	49.33 ± 0.80	82.46 ± 1.37	55.20 ± 0.76	
18	5.66	0.29%	557	0.23%	84.64 ± 4.52	5.49 ± 0.25	84.96 ± 4.36	5.64 ± 0.25	
23	4.72	0.24%	542	0.22%	89.55 ± 5.25	4.29 ± 0.23	84.22 ± 4.08	5.02 ± 0.21	
21	106.69	5.49%	12125	5.01%	83.23 ± 1.11	99.72 ± 1.11	73.21 ± 3.40	121.30 ± 4.12	
28	321.90	16.57%	39149	16.17%	75.22 ± 0.70	312.03 ± 2.19	74.47 ± 0.65	325.15 ± 2.11	
27	309.88	15.95%	37913	15.66%	66.77 ± 0.75	320.95 ± 2.41	68.95 ± 0.74	300.11 ± 2.22	
14	9.13	0.47%	1629	0.67%	67.82 ± 3.55	12.43 ± 0.44	72.74 ± 3.86	8.51 ± 0.33	
15	1.31	0.07%	148	0.06%	58.77 ± 9.38	1.77 ± 0.17	82.34 ± 12.04	0.93 ± 0.11	
20	1.54	0.08%	284	0.12%	59.26 ± 11.27	1.74 ± 0.20	51.50 ± 10.24	1.77 ± 0.18	
24	21.58	1.11%	1814	0.75%	89.83 ± 2.71	21.68 ± 0.59	90.10 ± 1.93	21.51 ± 0.42	
25	3.80	0.20%	584	0.24%	74.08 ± 6.75	3.24 ± 0.22	63.14 ± 6.03	4.46 ± 0.27	
26	25.50	1.31%	1021	0.42%	4.76 ± 32.81	0.40 ± 0.13	11.76 ± 17.33	10.33 ± 1.79	
30	9.99	0.51%	1238	0.51%	66.91 ± 5.14	9.34 ± 0.48	62.69 ± 4.28	10.66 ± 0.46	
31	15.98	0.82%	1847	0.76%	73.50 ± 3.70	13.45 ± 0.50	66.82 ± 3.23	17.57 ± 0.57	
29	14.65	0.75%	2520	1.04%	71.36 ± 2.95	17.54 ± 0.52	70.09 ± 3.05	14.92 ± 0.45	
					1943.13 ± 25.99		1943.13 ± 31.38		