SPECIES DISTRIBUTION AND FOREST TYPE MAPPING IN MEXICO

Anna Cord[ab], René R. Colditz[c], Michael Schmidt[ab], Stefan Dech[ab]

a German Aerospace Center – German Remote Sensing Data Center (DLR-DFD), Münchner Str. 20, 82234 Wessling, Germany
b Department of Remote Sensing, Institute of Geography – University of Wuerzburg, Am Hubland, 97074 Wuerzburg, Germany
c National Commission for the Knowledge and Use of Biodiversity (CONABIO), Avenida Liga Periférico – Insurgentes Sur No 4903, Col. Parques del Pedregal, Del. Tlalpan 14010 México, D.F., Mexico

ABSTRACT

The study analyzed the potential of multi-temporal satellite remote sensing data (Terra-MODIS) for phenological studies in Mexico. Within two conceptually and methodologically independent approaches, we analyzed vegetation index data to continuously predict species' occurrence and linked these predictions to a satellite data based land cover map of Mexico. Phenological metrics derived from interpolated MODIS EVI time series together with topographic data of the SRTM mission and bioclimatic variables (WorldClim) were used as environmental predictors for Maxent distribution modeling of six *Quercus* species. We evaluated model performance based on 20 % independent occurrence data using model sensitivity, model specificity, and AUC statistics. Forest classes were extracted from a recently accomplished national land cover map of Mexico (NALCMS) with approximate accuracy of 82 %. We found a remarkable spatial agreement of modeled *Quercus* distribution and forest classes. The value-added combination of both approaches allows for improved vegetation maps and the methodology has a great potential of transferability.

***Index Terms* —** Multi-temporal analysis, time series, MODIS, phenology, land cover classification, species distribution modeling, Maxent.

1. INTRODUCTION

Multi-temporal satellite data are nowadays of standard use in land cover classification approaches from regional to global scales [1], since classification of single-date imagery may fail at discerning land cover types that temporarily share similar spectral reflectance characteristics. Several more or less recently launched sensors, such as the *Moderate Resolution Imaging Spectroradiometer* (MODIS) used for this study, exhibit enhanced spectral and radiometric resolutions, wide geographical coverage and improved atmospheric corrections, while preserving a high temporal resolution [2]. The analysis of ecological attributes based on multi-temporal data is essential for effectively understanding ecosystem dynamics and allows for the extraction of phenological, seasonal, and latitudinal variations in vegetation cover. Since most natural and semi-natural vegetation associations feature distinct and in many cases unique phenological cycles [3], multi-temporal satellite data can be used to assign vegetation type memberships. Often, the observed phenology is directly related to vegetation type diversity and small-scale heterogeneity of topo-climatic conditions. In the context of decreasing environmental stability, climate change, accelerating anthropogenic impact, and biodiversity loss, quantitative assessments of vegetation type and distribution have become very important.

The availability of multi-temporal remote sensing data together with an increasing number of geo-referenced species' occurrence data (museum / herbarium collections or field observations) has created the opportunity to monitor and model species' geographic distribution and richness. Especially phenological characteristics (such as the starting date or length of the growing season) may provide useful ecosystem characteristics that explain the actual spatial distribution of a given species rather than its potential range, which is mainly determined by topo-climatic variables [4]. The aim of this study was to explore the potential of multi-temporal Terra-MODIS vegetation index data as environmental predictors for modeling the geographic distributions of selected plant species and to link these predictions to a satellite data based land cover map of Mexico. The focus was set on *Quercus* (white oak) species and forest types / land cover categories where these species are likely to occur. The assumption was that model predictions of several representative key species could be combined and further interpreted as proxy indicators of the distribution of the respective forest types enabling the extraction of land cover information with higher levels of thematic detail.

2. STUDY AREA

Mexico exposes a variety of ecosystems ranging from the arid northern interior to deciduous and evergreen tropical forests mainly located in the southern parts. At smaller scales, this pronounced North-South gradient is significantly modified by altitudinal effects. The marked environmental gradients throughout Mexico provide a unique opportunity for phenological studies aiming at the discrimination of different vegetation types based on time series of satellite remote sensing data. As a result of the above-mentioned physiographic and climatic variations, the mega-diverse country maintains around 10% of the world's biota [5] and has an enormous floristic richness.

3. DATA AND METHODS

We used the MODIS-Terra L3 global standard product (MOD13A2, 16-days composites, Version 5) designed for vegetation studies and the extraction of canopy biophysical parameters [6]. For this study, the Enhanced Vegetation Index (EVI) was used which is less susceptible to background soil effects or atmospheric disturbances and does not saturate in high biomass regions [7], an important issue in the tropical forests in southern and western Mexico.

3.1. Time series generation

MOD13A2 time series were generated with the TiSeG software package [8] which makes use of the MODIS product-specific and pixel-level data quality indicators. Low-quality observations, e.g. due to cloud cover, adverse viewing angles, or significant atmospheric contaminations, were excluded while high-quality observations were used as vertices for pixel-level linear temporal interpolation.

3.2. Species distribution modeling

Phenological metrics (Table 1) were derived from the interpolated MOD13A2 time series and averaged over the seven years of the study period (2001-2007). They were used as environmental predictors in combination with topographic data of the SRTM mission including elevation, slope, and aspect, and a series of bioclimatic variables (WorldClim, [9]) derived from monthly temperature and rainfall values. We applied the Maxent algorithm (*Maximum Entropy*, Software version 3.2.1, [10]) for distribution modeling of *Quercus peninsularis* (22 occurrence samples), *Q. tuberculata* (69), *Q. pringlei* (77), *Q. durifolia* (88), *Q. polymorpha* (164), and *Q. resinosa* (168). For all models, we converted the continuous predictions into binary maps using (maximum test sensitivity and specificity) threshold probabilities indicative of species' presence or absence. The predicted distributions were evaluated based on 20 % independent species occurrence data withheld from model training. We judged model usefulness as the successful prediction of species' presence in a high proportion of test localities whilst not predicting an excessively large part of the study area as suitable. We used three quality indicators, namely model sensitivity, model specificity, and AUC statistics (area under curve of the receiver operating characteristic, Table 3).

Table 1: Time series metrics and respective ecological / bio-physiological interpretation.

MOD13A2(EVI)	**Ecological interpretation**
Annual maximum	Maximum amount and density of green biomass at the peak of the growing season
Annual minimum	Minimum amount and density of green biomass; for detecting irrigated agricultural areas and urban settlements
Annual mean	Integrated measure of annually produced green biomass
Annual range	Annual increase of green vegetation
Standard deviation	Annual variability
Date of maximum	Time of the peak of the growing season
Date of minimum	Time of minimum vegetation density and greenness throughout the year

3.3. Forest type mapping

Forest classes were extracted from the recently accomplished national land cover map of Mexico for 2005 (Table 2), [11, 12]. This product is based on monthly composites of 250 m MODIS radiance data, annual and seasonal metrics as well as ancillary data (topography, temperature, and precipitation).

Table 2: Land cover classes of Mexico [11, 12].

Level I	Level II	ID
Needle leaved forest	Temperate / sub-polar needle evergreen	1
Broadleaved forest	Tropical / sub-tropical broad evergreen	3
	Tropical / sub-tropical broad deciduous	4
	Temperate / sub-polar broad deciduous	5
Mixed forest	Mixed Forest (broad / needle/ evergreen / deciduous)	7
Shrubland	Tropical / sub-tropical	8
	Temperate / sub-polar	9
Herbaceous	Tropical / sub-tropical	11
	Temperate / sub-polar	*12*
Wetland	Wetland	15
Cropland	Cropland	16
Barren lands	Barren lands	17
Urban and built-up	Urban and built-up	18
Water	Water	19

The sample data origin from various existing datasets complemented with digitized additional samples for sparsely populated classes. Various boosted classification trees were combined to obtain class-memberships for each pixel. The discrete map is accompanied by a confidence map. Set aside validation data yielded an approximate area-normalized accuracy of 82 %.

4. RESULTS AND DISCUSSION

4.1. Species distribution modeling

The phenological analysis of the EVI time series at the known occurrence points of *Quercus* spp. revealed that the vegetation communities at these localities differed significantly in their phenological cycles, especially in the annual range and maximum values (Fig. 1).

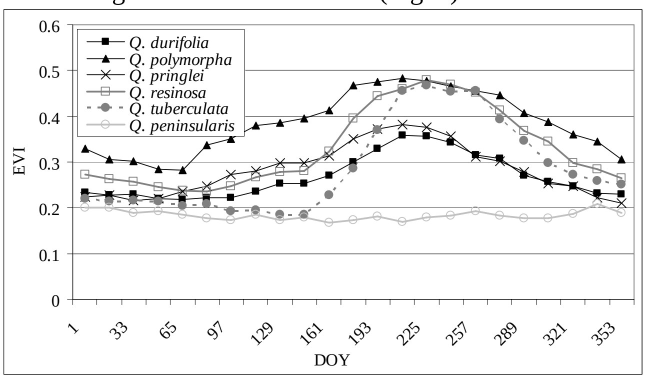

Fig. 1. Phenological cycles (year 2005) of *Quercus* study species based on MOD13A2 EVI data.

The observed species-specific phenological characteristics allowed for the discrimination and distribution modeling of the given species with excellent model performance and low omission and commission (FPA) rates (Table 3). The three most important explanatory variables of the time series-derived phenological metrics (Table 3) could be related to the observed phenologies, e.g. the range of the annual EVI values being an indicator for presence of *Q. tuberculata*.

4.2. Forest type mapping

All mean EVI curves of forest classes followed an expected pattern with a maximum during the rainy summer season. Even evergreen classes depict a subtle phenology due to the green-up in the understory. However, the range was most prominent for tropical / sub-tropical deciduous broadleaved forest (Fig. 2).

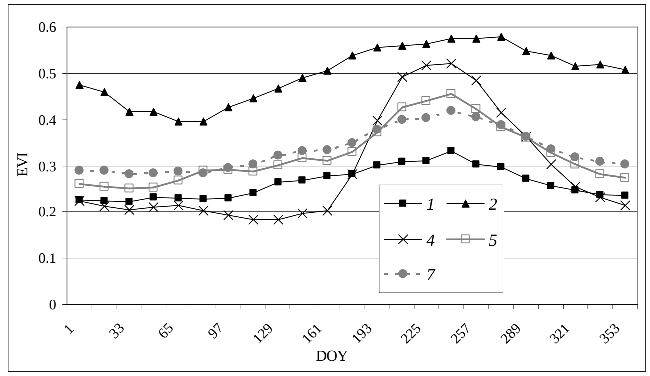

Fig. 2. Phenological cycles of forest classes (year 2005). For the meaning of land cover class numbers refer to Table 2.

The comparison of modeled *Quercus* distributions and land cover classes revealed that the predicted habitats largely coincided with areas classified as forest (Fig. 3, Table 4). Land cover and species' distribution maps thus mutually affirmed their reliability even though produced within two independent, methodologically and conceptually different approaches. Smaller fractions of the modeled *Quercus* habitats were classified as land cover categories shrubland, grassland, or cropland (Table 4) since the species of interest are non-obligate forest species but rather limited by the general environmental conditions. Moreover, due to different spatial resolutions of the input data, pixel were more often assigned to other categories especially in forest / non-forest mosaic landscapes at 250 m than at 1 km resolution.

Table 4. Agreement statistics of land cover classes with modeled species' distributions.

Land cover classes (%)	Forest	Shrub-land	Grass-land	Crop-land	Other
Q. peninsularis	80.5	18.8	0.0	0.4	0.3
Q. tuberculata	85.7	8.0	1.3	4.9	0.2
Q. pringlei	59.2	32.5	4.5	3.6	0.2
Q. durifolia	81.8	10.4	6.6	1.2	0.1
Q. polymorpha	51.8	23.8	1.6	21.7	1.1
Q. resinosa	47.2	25.2	10.8	15.9	0.8

Table 3. Results and parameters of the Maxent distribution model for each study species. Note: 'Threshold' = maximum test sensitivity and specificity threshold, 'AUC' = Area under curve of the receiver operating characteristic, 'FPA' = Fraction of predicted area referring to the land surface.

Species	**Explanatory variables**	**Threshold**	**Training AUC**	**Test AUC**	**Test omission rate**	**FPA**
Q. peninsularis	Stddev., max, range	0.829	1.000	1.000	0.000	0.0001224
Q. tuberculata	date of min, range, max	0.589	0.990	0.985	0.091	0.0034649
Q. pringlei	date of min, mean, min	0.332	0.997	0.995	0.000	0.0047756
Q. durifolia	date of min, range, max	0.399	0.989	0.912	0.119	0.0078367
Q. polymorpha	min, range, date of min	0.237	0.997	0.984	0.087	0.0106500
Q. resinosa	min, range, mean	0.134	0.993	0.987	0.000	0.0213675

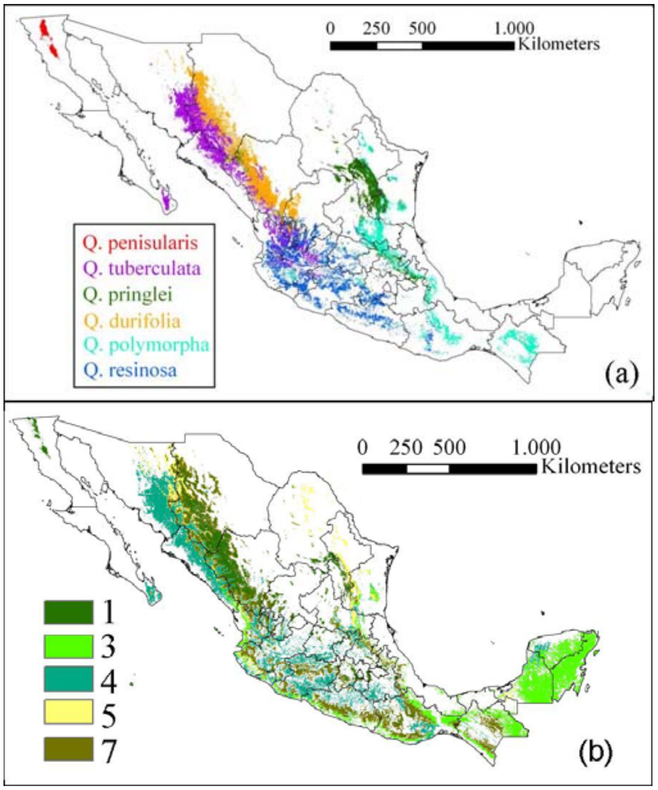

Fig. 3 (a) Predicted distribution of *Quercus* spp. and (b) distribution of forest types (refer to Table 2).

5. CONCLUSIONS

The spatial agreement of modeled *Quercus* spp. distributions and forest classes according to the NALCMS land cover product was remarkable. Both approaches used multi-temporal MODIS vegetation index data to either continuously predict species' occurrence or assign categorical land cover class memberships. Results showed that the inherent phenological characteristics could successfully be extracted from the time series of satellite remote sensing data using a set of statistical features and exploited for the discrimination of different vegetation types. We therefore suggest modeling the distributions of representative plant species and using the results as proxy indicators to derive vegetation maps. The value-added combination of both approaches will allow for the extraction of reliable vegetation distribution data at higher levels of thematic detail. By additionally modeling distributions of species with key ecological functionalities, land cover and vegetation maps can be attributed with further information related to bio- and ecological diversity. Even though developed for Mexico with its marked environmental and altitudinal gradients, the presented methods can be transferred to other study regions on the globe being dependent only on the availability of reliable geo-referenced species occurrence data.

6. ACKNOWLEDGEMENTS

The 2005 land cover map is a preliminary result of a tri-national initiative of the North American Land Change Monitoring System ([12]). The employed monthly MODIS composites were processed by the Canadian Centre for Remote Sensing (CCRS). The sample data were compiled and corrected by the National Institute for Statistics and Geography (INEGI), the National Institute for Forestry (CONAFOR), and the National Commission for the Knowledge and Use of Biodiversity (CONABIO) of Mexico.

7. REFERENCES

[1] Cihlar, J. (2000): Land cover mapping of large areas from satellites: Status and research priorities. International Journal of Remote Sensing, 21, 1093-1114.

[2] Carrao, H., Gonçalves, P., and Caetano, M. (2007): Contribution of multispectral and multitemporal information from MODIS images to land cover classification, Remote Sensing of Environment, 112(3), 986-997.

[3] Geerken, R., Zaitchik, B., and Evans, J.P. (2005): Classifying rangeland vegetation type and coverage from NDVI time series using Fourier Filtered Cycle Similarity, International Journal of Remote Sensing, 26(24), 5535-5554.

[4] Bradley, A. B. and Fleishman, E. (2008): Can remote sensing of land cover improve species distribution modelling?, Journal of Biogeography, 35(7), 1158-1159.

[5] McNeely, J. A., Miller, K. R., Reid, W. V., Mittermeier, R. A. & Werner, T. B. (1990): Conserving the world's biological diversity, International Union for Conservation of Nature and Natural Resources (IUCN), Gland, Switzerland, 200 pp.

[6] Justice, C.O., Townshend, J.R.G., Vermote, E.F., Masuoka, E., Wolfe R.E., Saleous N., Roy D.P., Morisette J.T. (2002): An overview of MODIS Land data processing and product status. Remote Sensing of Environment, 83(1-2), 3-15.

[7] Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X. and Ferreira, L. G. (2002): Overview of the radiometric and biophysical performance of the MODIS Vegetation indices, Remote Sensing of Environment 83, 195-213.

[8] Colditz, R.R., Conrad, C., Wehrmann, T., Schmidt, M. & Dech, S. (2008): TiSeG: A flexible Software Tool for Time-Series Generation of MODIS Data Utilizing the Quality Assessment Science Data Set. IEEE Transactions on Geoscience and Remote Sensing, 46(10), 3296-3308.

[9] Hijmans et al. (2005): Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25, pp. 1965-1978.

[10] Phillips, S.J., Anderson, R.P., Schapire, R.E. (2006): Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190, 231-259.

[11] Colditz, R.R., López, G., Meneses, C., Maeda, P., Crúz, I., Ressl, R., 2009. A satellite-data based land cover monitoring system for the North American Continent – the contribution of Mexico. In: 33rd International Symposium on Remote Sensing of Environment, May 4th – 8th 2009, Stresa, Italy, CD-ROM.

[12] NALCMS (2009): http://www.cec.org/naatlas/nalcms.cfm# Contacts (June 12th 2009).