Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 112 (2008) 2145–2159
Remote Sensing of Environment
www.elsevier.com/locate/rse

# Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments

Steven E. Sesnie [a,b,*], Paul E. Gessler [a], Bryan Finegan [b], Sirpa Thessler [c]

[a] University of Idaho, College of Natural Resources, Forest Resources Department, PO Box 1133, Moscow, Idaho 83844-1133, USA
[b] Departament of Natural Resources and Environment, Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Apdo. 93-7170, Turrialba, Costa Rica
[c] Finnish Forest Research Institute (Metla), Unioninkatu 40 A, 00170 Helsinki, Finland

Received 1 November 2006; received in revised form 19 July 2007; accepted 18 August 2007

## Abstract

Remotely sensed images and processing techniques are a primary tool for mapping changes in tropical forest types important to biodiversity and environmental assessment. Detailed land cover data are lacking for most wet tropical areas that present special challenges for data collection. For this study, we utilize decision tree (DT) classifiers to map 32 land cover types of varying ecological and economic importance over an 8000 $km^2$ study area and biological corridor in Costa Rica. We assess multivariate QUEST DTs with unbiased classification rules and linear discriminant node models for integrated vegetation mapping and change detection. Predictor variables essential to accurate land cover classification were selected using importance indices statistically derived with classification trees. A set of 35 variables from SRTM-DEM terrain variables, WorldClim grids, and Landsat TM bands were assessed.

Of the techniques examined, QUEST trees were most accurate by integrating a set of 12 spectral and geospatial predictor variables for image subsets with an overall cross-validation accuracy of 93%±3.3%. Accuracy with spectral variables alone was low (69%±3.3%). A random selection of training and test set pixels for the entire landscape yielded lower classification accuracy (81%) demonstrating a positive effect of image subsets on accuracy. A post-classification change comparison between 1986 and 2001 reveals that two lowland forest types of differing tree species composition are vulnerable to agricultural conversion. Tree plantations and successional vegetation added forest cover over the 15-year time period, but sometimes replaced native forest types, reducing floristic diversity. Decision tree classifiers, capable of combining data from multiple sources, are highly adaptable for mapping and monitoring land cover changes important to biodiversity and other ecosystem services in complex wet tropical environments.
© 2008 Elsevier Inc. All rights reserved.

*Keywords:* Decision trees; Biodiversity; Costa Rica; Land cover change; SRTM-DEM; Landsat TM

## 1. Introduction

Validated methods to discriminate tropical rain forest vegetation types with remotely sensed data have become progressively important following the 1992 International Convention for Biological Diversity (CBD) (Nagendra and Gadgil, 1999; UNEP, 1992). Accelerated deforestation since the 1950s has left dispersed forest remnants and regrowth areas amid agricultural lands over much of Latin America (Houghton, 1994; Mayaux et al., 2005). The Framework Convention on Climate Change targets aforestation and reforestation areas as carbon sinks that can potentially restore a level of forest connectivity within fragmented landscapes (Lamb et al., 2005). Multitemporal information regarding the extent of habitat types, regeneration areas and land cover dynamics is needed for tropical regions to achieve diverse conservation goals (Castro et al., 2003).

Creating suitable land cover information for prioritizing conservation activities will require methods to integrate data from a

* Corresponding author. Current contact information: Northern Arizona University. Environmental Sciences and Education Department. PO Box 5694, Flagstaff, Arizona, USA. 86011-5694. Tel.: +1 928 523 7889; fax: +1 928 523 7423.

*E-mail addresses:* steven.sesnie@nau.edu (S.E. Sesnie), sirpa.thessler@mtt.fi (S. Thessler) Paulg@uidaho.edu (P.E. Gessler), bfinegan@catie.ac.cr (B. Finegan).

0034-4257/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.08.025

variety of Earth observation systems (Kerr and Ostrovsky, 2003). In this study, we examine the potential for discriminating tropical rain forest and other land cover types with a decision tree (DT) classifier. Non-parametric DTs have been shown to be more capable than other methods for integrating diverse data to solve complex classification problems (De'ath and Fabricius, 2000; Franklin, 2003), but are relatively unexplored for mapping tropical vegetation types and changing landscape conditions.

Differentiating forest categories based on species composition and structural characteristics, in addition to land cover related to agricultural development is a primary objective for this study. Land cover, as it relates to differences between natural vegetation types, is a common element of biodiversity assessment, in addition to assessing vertebrate species distributions (Scott and Jennings, 1998). Spectral features from Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) images have been shown capable of discriminating some rain forest types and secondary successional stages (Foody and Hill, 1996; Lobo and Gullison, 1998; Lu et al., 2003a; Nagendra and Gadgil, 1999). However, vegetation classification can be limited by satellite image based approaches when ecologically important differences in forest structure and composition are spectrally similar (Castro et al., 2003; Helmer et al., 2002; Lu et al., 2003a; Pedroni, 2003; Sader et al., 1989; Sesnie et al., in review). Integrating biophysical factors related to natural vegetation and land use patterns with satellite data can potentially improve land cover discrimination for vegetation of different structure and composition features (Ferrier et al., 2002; Pedroni, 2003).

Advantages of DTs are their versatility to integrate numerical and categorical variables into vegetation classifications free of distributional assumptions (Breiman et al., 1984; Guisan and Zimmerman, 2000). Parametric statistical models are often unsuitable for data integration because of non-linear relationships and correlation among variables (Friedl and Brodley, 1997; Franklin, 2003; Gislason et al., 2006). DT classifiers utilize conditional relationships between vegetation types, spectral and geospatial information such as elevation, slope, aspect and other environmental variables to enhance forest classifications (Gislason et al., 2006). DTs also typically require less training time compared to other machine learning techniques, such as artificial neural networks and support vector machines, while attaining similar accuracies (Pal and Mather, 2003; Pal, 2005).

Limitations include the instability of trees to outliers or small changes in the training data (Miller and Franklin, 2002), in addition to requiring a large number of training samples for tree construction (Joy et al., 2003). However, improved classification accuracy over other methods has been achieved for mapping national and global-scale land cover with the advanced very high resolution radiometer (AVHRR) and moderate resolution imaging spectroradiometer (MODIS) (Friedl and Brodley, 1997; Friedl et al., 2002; Defries et al., 1998; Hansen et al., 2000; Muchoney et al., 2000). Brown de Colstoun et al. (2003) attained accuracies >80% for classifying temperate broad-leaf, coniferous and mixed forest types using Landsat ETM+ bands. Joy et al. (2003) demonstrate that incorporating topographic variables derived from a digital elevation model (DEM) with TM bands enhanced accuracy for classifying coniferous forest types.

The availability of low-cost multispectral and geographic information of known quality improves the potential for integrating data to map and monitor floristically defined forest categories (Gislason et al., 2006). Landsat TM's >20 year image archive is a primary source of repeated Earth observations at a nominal 30 m pixel resolution, important for obtaining relatively cloud free images in wet tropical areas. Recently available digital elevation data at a 3-arcsecond spatial resolution from the Space Shuttle Radar and Topography Mission (SRTM) dramatically improves data resources for all tropical regions (Hofton et al., 2006). A DEM is often the primary data source for deriving terrain variables related to vegetation composition (Franklin, 1995; Miller and Franklin, 2002; Moore et al., 1991). The WorldClim data set, used for this study, combines SRTM-DEM elevations with global meteorological records to improve monthly temperature and precipitation surfaces at an ~1 km grid cell size (Hijmans et al., 2005).

Classical "niche theory" and gradient analysis is central to DT applications for predictive vegetation modeling and pattern recognition (Franklin, 1995). Niche theory posits that ecological communities and constituent vegetation are structured along environmental gradients that are important to making predictions beyond sampled locations (Miller and Franklin, 2002; Whittaker, 1967). Similar vegetation composition is expected to occur on sites with comparable climate, soil, and terrain conditions (Franklin, 1995; Tuomisto et al., 2003a). Random disturbances and dispersal limitations can also influence tropical tree species composition (Condit, 1996; Condit et al., 2002). Consequently, geographic distance will be evaluated in this study, as it has shown success for predicting floristic distributions in temperate and tropical regions as a substitute for potentially unmeasured dispersal variables (Chust et al., 2006; Franklin, 1998).

For the present study, we consider a framework to integrate multitemporal satellite data with geographic information and ground vegetation data for enhanced land cover mapping and change detection. Our objective is to first determine which spectral and geospatial predictor variables enhance classification accuracy for mapping priority floristic alliance categories. Alliance is defined here as a floristic category based on existing vegetation of similar physiognomy, characterized by dominant tree species or other diagnostic vegetation (Jennings et al., 2004). Random Forest decision trees (Breiman, 2001) discussed below, are used to evaluate and select predictor variables leading to accurate land cover classification. Statistical indices from the Random Forest trees also enable interpretation of ecological factors driving vegetation composition and distribution patterns (Gislason et al., 2006). A second objective is to compare multivariate DT classifiers via cross-validation classification accuracy (CV-accuracy) using subsets of predictors and Landsat TM imagery. Prior studies suggest that more comprehensive land cover classifications are often implemented using smaller image subsets (Pedroni, 2003). Image subsets (2000 km$^2$) versus large-area (8000 km$^2$) classifications are assessed for accuracy to evaluate their application for heterogeneous land cover conditions. The above procedures enable the objective selection of an approach for detailed land cover classification and change analysis in a priority conservation area of Costa Rica.

Download English Version:

https://daneshyari.com/en/article/4460587

Download Persian Version:

https://daneshyari.com/article/4460587

Daneshyari.com