Forest Ecology and Management 327 (2014) 240–250

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

The development of a forestry geodatabase for natural forest management plans in Costa Rica

Sienna Svob, J. Pablo Arroyo-Mora *, Margaret Kalacska

Department of Geography, McGill University, Burnside Hall Building, Montreal, QC H3A 0B9, Canada

ARTICLE INFO

Article history:
Received 30 November 2013
Received in revised form 8 May 2014
Accepted 15 May 2014
Available online 11 June 2014

Keywords:
Data management
Logging inventories
Tropical forest
Ecoinformatics
Quality assurance

ABSTRACT

Forest management data is available for many tropical countries, representing a large and spatially rich source of tree level data. Over the past decade, we have collected and digitized over 1000 Costa Rican natural forest management plans, spanning 30 years and spread out over approximately 26,700 km^2 along the country's Atlantic coast, northern lowlands, and southern Pacific coast. In order to analyze this unique collection of tree inventory data we developed a system to ensure the standardization, quality control, and reliable management of the dataset. We developed a relational geodatabase, forming logical associations between and within the spatial and tabular components of the forest management data. In this study, we outline the data standardization procedures established to permit the comparison of data across time and space. Further, we describe quality checks built-into the database's functionality to identify and reduce the presence of errors. The final customized forest management geodatabase efficiently stores a large and diverse dataset through the integration of logical relations, quality checks, and flexible data access across 32 tables and over 250,000 tree records. Through the use of quality tests, the database provides a means to improve overall data integrity and illustrates the magnitude, sources, and types of errors present in the initial dataset. Finally, the value of a comprehensive database for the management of forest data is demonstrated through an exploratory analysis of standardized tree taxonomic information. With this analysis, we begin to explore the potential strengths, weaknesses, and applications of forest management data for future ecological studies (e.g. species diversity assessment). For instance, although most of the forest management data (tree inventories and censuses) is collected using common names in the field, our standardization process has allowed us to depict trends similar to those found in ecological studies (e.g. dominant species for different ecosystems). Overall, our forestry geodatabase represents the most complete record of natural forest management practices in Costa Rica to date.

1. Introduction

While tropical forests roughly represent only 7% of global land cover (Bradshaw et al., 2009), they are home to an estimated two-thirds of the globe's terrestrial biodiversity (Gardner et al., 2009) and contain nearly 40% of the earth's carbon biomass (Brown and Lugo, 1992). Due to a rapidly changing climate combined with high rates of deforestation and forest degradation, these globally important carbon and biodiversity stocks are under threat (Malhi and Grace, 2000). As these ecosystems decline, a renewed interest in quantifying and understanding spatial and temporal variations in biodiversity and carbon distribution across tropical forests has arisen (Petrokofsky et al., 2011). However, a high level of uncertainty in the assessment and understanding of environmental factors influencing biomass and biodiversity across spatial scales persists (Houghton et al., 2001; Chave, 2008; Gardner et al., 2009). This uncertainty is primarily rooted in the inadequate spatial distribution and sampling intensity of traditional ecological data sets (Clark et al., 1998, 1999). While landscape scale tropical forest inventories may offer a solution to this uncertainty, they are rare among ecological studies due to their high cost and complex logistics (Greig-Smith, 1983). Moreover, long-term datasets require strong database management approaches to ensure reproducible and valid data analyses (Le Duc et al., 2007; Condit et al.,

Abbreviations: ACAHN, Arenal Huetar Norte conservation area; ACCVC, Central Volcanic conservation area; ACLA-C, Caribbean La Amistad conservation area; ACOSA, Osa conservation area; ACTO, Tortuguero conservation area; DBH, diameter at breast height; FGIS, forestry geographic information system; NFMP, natural forest management plant.

* Corresponding author. Tel.: +1 5143986052.
E-mail addresses: sienna.svob@mail.mcgill.ca (S. Svob), pablo.arroyo@mcgill.ca (J.P. Arroyo-Mora), margaret.kalacska@mcgill.ca (M. Kalacska).

http://dx.doi.org/10.1016/j.foreco.2014.05.024
0378-1127/

2014). Existing pre-felling inventories from selective logging (i.e. forest management), on the other hand, may represent a solution due to their spatial distribution and sampling intensity (Couteron et al., 2003; Arroyo-Mora et al., 2009).

Logging inventories are common to tropical countries. They represent a valuable and abundant source of data on forest composition and structure, permitting a variety of ecological questions to be addressed at spatial scales larger than those in traditional ecological inventories (Couteron et al., 2003; Putz et al., 2001; Réjou-Méchain et al., 2011a). Due to their unique scale, logging inventories have allowed researchers to observe the environmental factors controlling species distribution and community composition at regional and landscape scales (Couteron et al., 2003; Réjou-Méchain et al., 2008; Gourlet-Fleury et al., 2011; Fayolle et al., 2012). Research conducted at these large scales provides a link to local scale studies and therefore, insight into the importance of certain environmental factors and processes controlling community composition at different spatial scales (Réjou-Méchain et al., 2011b). The large area sampled by logging inventories also allows for studies to successfully capture the heterogeneity of tropical forest biomass distribution, and in turn forest carbon (Gourlet-Fleury et al., 2011; Maniatis et al., 2011). These studies can help reveal the impact of particular environmental factors (e.g. elevation, soil type, and land-use history) on tropical forest biomass, providing crucial information for future studies attempting to predict the regional and landscape patterns of biomass. Finally, large scale inventories aid in the development of detailed wall-to-wall vegetation and land use maps by offering a rich source of ground level data for the validation and calibration of remotely sensed data (Gond et al., 2013).

In Costa Rica, selective logging inventories are developed under a standardized Natural Forest Management Plan (NFMP) framework (Arroyo-Mora et al., 2009) that has been implemented on private lands since 1996 (Forestry Law 7575). As Costa Rica's forest resources are managed under a national system of conservation areas, NFMPs are handled in local sub regional offices. In total, the country is divided into 11 conservation areas (Boza, 1993). Within each conservation area, sub-regional offices are responsible for the reception, revision, approval or rejection and follow up of the forest management plans. Due to the limited resources of sub-regional offices, this rich information is often inadequately stored, organized, and in some cases even destroyed (e.g. floods, fire).

In order to recover these archives and evaluate their capacity to support forest conservation, Arroyo-Mora (2008) developed a forestry geographic information system (FGIS) in a case study of a conservation area in Northern Costa Rica. Following that original study, this present work encompasses the complete FGIS of Costa Rica, including the five conservation areas where selective logging has been most heavily practiced. Specifically, selective logging has focused in lowland areas in the Atlantic coast, the northern lowlands, and the southern Pacific coast. The main objective of this study is to develop a countrywide tree database from data extracted from natural forest management plans (NFMPs). Our second objective is to describe the broad taxonomic patterns occurring in our study region based on the FGIS database in order to assess the utility of the NFMP dataset.

With modern ecological datasets combining spatial and temporal scales, powerful data infrastructures are a necessity for efficient data storage and analysis (McIntosh et al., 2007). The relational database model, developed by IBM in the 1970s, has become an international standard (Codd, 1979). Relational databases following this model have been used extensively to efficiently store, mine, and analyze ecological data, all while minimizing redundancy (McIntosh et al., 2007). The RAINFOR database is one such model, composed of long-term ecological information from plots established across the Amazonian rainforest (Peacock et al., 2007). More recently, Condit et al. (2014) developed the CTFS database to reduce the presence of integrity errors within a global repeated-measurements forest plot dataset. A newer extension of the relational database model is the geodatabase, which also contains spatial information that may be displayed in a geographic information system (GIS) (Zeiler, 1999). For instance, geographic information databases have been employed in Mexico to study land cover change (Mas et al., 2004) and by the US Forestry Service to model locations for forest management (Loh et al., 1994). This study presents the development and structure of the FGIS geodatabase along with the forest management, taxonomic, and spatial data that it stores. Additionally, it highlights the importance of establishing data standardization and quality assurance procedures when managing large ecological datasets. Finally, it describes the limitations of such a database before providing an example of its potential application in the study of Costa Rica's diverse forested ecosystems.

2. Natural forest management plans in Costa Rica

In Costa Rica, a natural forest management plan is a document containing a collection of technical standards combined with data, developed to govern the management activities of a privately owned area of natural forest (management unit or forest unit). The data applied in this study was extracted from NFMPs conducted by certified foresters between 1983 and 2011. Below, we provide a summary of the standard methodology followed and data collected in these NFMPs.

In general, a NFMP document produced by a certified forester contains information on the structure and composition of a forest stand (inventory and census), official land tenure information, the protected and productive areas within the stand, proposed logging roads and timber patios (MINAE, 2008). The document also contains general maps of land holdings including, among other features, the shape of the forest unit, the trees to be logged, and the trees to be preserved as progeny trees. In addition, protected areas in the management unit are defined as areas near water bodies (e.g. streams, springs) and on steep slopes where selective logging is prohibited (MINAE, 2008). The forested area outside of the protected area within the bounds of each management unit is the productive area.

Costa Rican natural forest management plans generally follow a systematic field sampling design. The design consists of transects placed across the forest stand perpendicular to a pre-established baseline. Depending on the terrain and total forest unit area, adjacent transects are separated by 50–100 m and extend the length of the forest stand. Transects are located spatially by at least one point in the field and later mapped. Inventory plots of 0.3 ha (30 by 100 m) are mounted randomly along these transects and all trees with a DBH $\geqslant$ 30 cm are recorded within the plots. Despite these standards, however, inventory plots are sometimes biased towards species rich and biomass dense areas of the forest stand (Arroyo-Mora pers. obs.) The number of plots mounted through the inventory is determined so that the sampling error is less than 20% (95% confidence level) for the basal area per hectare of trees with a DBH $\geqslant$ 30 cm (MINAE, 2008). Of the data included in this study, the number of inventory plots per NFMP ranged from one to 53. A more complete characterization of the tree population in the forest stand is given by the census. In the census, all trees with a DBH $\geqslant$ 60 cm are recorded along the previously mounted transects (50–100 m apart extending the length of the forest stand). Trees recorded in the census are spatially located in the forest stand and classified as trees for harvest or as remnants. The harvest intensity is determined in proportion to the abundance of each tree species. In Costa Rica, the logging intensity must be less than 60%