Forest Ecology and Management 405 (2017) 69–80

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Model selection changes the spatial heterogeneity and total potential carbon in a tropical dry forest

R.O. Corona-Núñez[a,*], A. Mendoza-Ponce[b], R. López-Martínez[a]

[a] Procesos y Sistemas de Información en Geomática, SA de CV (PSIG), Calle 5, Viveros de Petén 18, Col. Viveros del Valle, 54060 Tlalnepantla, Mexico
[b] International Institute for Applied Systems Analysis (IIASA) – Schlossplatz, 2361 Laxenburg, Austria

ARTICLE INFO

Keywords:
Above ground biomass
Potential carbon stocks
Tropical dry forest
Reconstruction
Mexico

ABSTRACT

Understanding how aboveground biomass (AGB) is spatially distributed in the landscape and what factors are involved is critical to identify the ecological constraints limiting the magnitude and the allocation of carbon (C) stocks. Yet these factors remain poorly quantified for much of the world. The aim of this study is to identify the factors that influence the reconstruction of potential AGB and its spatial heterogeneity under current climate. A range of statistical approaches is used here to reconstruct the spatial distribution of AGB found in a tropical dry forest in Mexico. This is one of the first studies to directly quantify the predictive performance of various techniques within a common framework applied to AGB estimates from field observations and biophysical variables. The results suggest that general linear model (GLM) and the general additive model (GAM) performed similarly and outperformed other more complex approaches, such as automated neural networks, generalized linear mixed models via penalized quasi-likelihood, MaxEnt and random forest. GLM and GAM approaches also showed good performance in comparison to independent field observations over different spatial resolutions. MaxEnt performed poorly against independent validation data. The GLM, GAM, neural networks and regression tree models returned comparable mean AGB, suggesting that the potential AGB in the studied area is $\sim 132\ \mathrm{Mg\ ha^{-1}}$. The biomass spatial distribution is represented differently by the different models. Neural networks and regression tree approaches tend to cluster similar AGB estimates with a large range of the spatial autocorrelation, while the GLM is capable of reproducing the spatial distribution of the biomass.

1. Introduction

Deforestation in the tropics is an important source of CO_2 and thus a major driver of climate change (CC) (Houghton, 2005; Pan et al., 2011). However, the uncertainty of CO_2 estimations is large, mainly as result of the following major elements: (1) limited knowledge on the current state of biomass of tropical forests (Houghton, 1999; Eva et al., 2003; Fearnside and Laurance, 2003; Ometto et al., 2015), and of the underlying processes of the natural regeneration (Tucker and Townshend, 2000; DeFries et al., 2002); (2) Understanding of the deforestation impacts, forest degradation and fires, on C stocks (Laurance et al., 1998; Laurance et al., 2000; Barlow et al., 2003); (3) the scale of analysis (Hansen and DeFries, 2004); and (4) the diverse methodologies used to measure, simulate and/or predict any of the above.

Recently there have been efforts to mitigate CC through a reduction of CO_2 emissions through projects as Reducing Emissions from Deforestation and Forest Degradation (REDD+) and, Clean Development Mechanism (Santilli et al., 2005). These projects are aimed to promote forest regeneration and reduce CO_2 emissions due to deforestation and forest degradation (Harris et al., 2012). Consequently, the understanding of the potential above ground biomass (AGB_{pot}) and the potential C sequestration from forests becomes crucial to quantify the total emissions due to anthropogenic activities, which in turn help to prioritize restoration programs and/or land management.

Forests are spatially heterogeneous due to a complex combination of environmental and topographic conditions, and human disturbance (Houghton et al., 2001; Houghton, 2005; Ryan et al., 2012; Woollen et al., 2012). Temperature and precipitation, solar irradiation and soil nutrients are the primary drivers of plant development (Holmgren et al., 1997; Berdanier and Klein, 2011; Medeiros and Drezner, 2012; Peterson, 2012), biomass accumulation and therefore, C sequestration. However, depending on the scale of analysis and the ecosystem these drivers perform differently to AGB. On the one hand, the availability of a suitable microclimate leading into water availability (Holmgren et al., 1997; Berdanier and Klein, 2011) and soil nutrients may play a decisive role (Allen and Hoekstra, 1990; Turner, 2005; Currie, 2011) for the individual development of trees. At a landscape scale variables such as solar irradiation, slope, aspect, soil texture and concavity of the terrain

* Corresponding author.
E-mail address: rogelio.corona@sigeomatica.com (R.O. Corona-Núñez).

http://dx.doi.org/10.1016/j.foreco.2017.09.018
Received 31 May 2017; Received in revised form 5 September 2017; Accepted 5 September 2017
Available online 28 September 2017

are recognized to be the major influences on water availability for plants, mainly in dry ecosystems (Leitner, 1987; Berdanier and Klein, 2011; Peterson, 2012), and thus consequently on AGB densities. Other studies have revealed that species composition plays importantly in explaining AGB distribution (Becknell et al., 2012). On the other hand, at global scales, climatic variables are main factors to explain AGB and eclipsing local factors such as topography and soil properties (Snyder and Tartowski, 2006). For example, different authors found that annual precipitation explains over 50% of the variation in AGB with a negative correlation (Brown and Lugo, 1982; Eaton and Lawrence, 2009; Becknell et al., 2012).

Tropical forests represent a large reservoir of C, nearly 50% of C stored in vegetation (Houghton, 2005) and a sink in tropical forest regrowth of 1.6 ± 0.5 PgC $year^{-1}$ (Pan et al., 2011). However, land-use and land-cover change (LULCC) has become a focus of substantial research due to large C fluxes associated with LULCC, particularly from tropical forests (Carpenter et al., 2006). A majority of research effort has focused on tropical rain forests, neglecting tropical dry forests (Trejo and Dirzo, 2000; Skutsch et al., 2009). Although they once accounted for over 40% of all tropical forests (Cao et al., 2016). Currently, information on the state and disturbance of TDF are sparse and uncertain.

To date, different approaches, with differing advantages and disadvantages have been applied to estimate current forest C stocks and rates of change. Some authors have estimated the loss of biomass by considering a mean deforestation rate and a mean biomass value (Achard et al., 2002; DeFries et al., 2002; Achard et al., 2004; Houghton, 2005) which resulted in divergent estimations due to the spatial heterogeneity of the AGB distribution (Slik et al., 2010; Slik et al., 2013; Ometto et al., 2015), and the discrepancy in deforestation rates (Ewers et al., 2008; Corona, 2012; Ometto et al., 2015). While these studies track changes over time, none of them is capable to estimate and locate the total AGB that has been depleted. Moreover, they can mislead the understanding of the location of the major AGB losses occurred and the total C emissions that can be related to such changes. Nevertheless, these methodologies are useful for implementing conservation strategies to reduce the impacts of deforestation on C stocks, but fail to prioritize sites for potential C sequestration. Therefore, to mitigate CC and to reduce the impacts on ecological processes, it is important to take into consideration the AGB_{pot} and its spatial heterogeneity. Moreover, the assessment of AGB_{pot} allows to estimate the total C emissions that took place in the past which cannot be derived by other means, for example, agricultural fields (Exbrayat and Williams, 2015) or cities (Doko et al., 2014; Lentz et al., 2014). Thus, in the light of current knowledge, this study aims to understand how the model selection influences the potential AGB predictions over different spatial scales and also to evaluate their uncertainty.

2. Methods

2.1. Study region

The study region is located on the Southern Pacific coast in the state of Oaxaca, Mexico (Fig. 1) with a total area of 215,687 ha. The altitude ranges from 0 to 1200 m asl and the local climate is classified as a dry sub-humid, Aw(w), Köppen modified by García (2004). The main dry season lasts seven months with a mean annual precipitation of less than 1600 mm, > 75% of which falls between June and September. The annual temperature ranges between 19 and 33 °C with a mean annual evaporation of ~1700 mm with maximum values that exceed 170 mm during March to May (Hijmans et al., 2005; SMN, 2014). The dominant vegetation cover (~80%) is Tropical Dry Forest (TDF) (Corona, 2012) with different levels of degradation processes (Corona, 2009; Lira and Ceballos, 2010; Corona, 2012; Mendoza, 2015). The agriculture is dominated by slash burn practices (Corona et al., 2016).

2.2. Sampling design and AGB estimates

A stratified random sampling procedure was implemented to characterize the AGB landscape heterogeneity. The stratified sampling was based on topographical features of the landscape, altitudinal gradient (every 200 m asl), and slope orientation (North, East, West, South and flat land).. With this approach, our field campaigns recorded the landscape variability by sampling the most common biophysical regions and the rare elements of the landscape. We collected 60 plots (10 m × 30 m). In each plot, all the plants ≥1 cm in diameter at height breast (DBH; defined at ≥1.3 m in height) were measured. Our data collection was complemented with field samples of the Mexican National Forest Inventory (NFI) (CONAFOR, 2007, 2012). The NFI is based on 128 plots (10 m × 40 m), where all trees ≥7.5 cm in DBH were recorded (DBH and height). For trees with DBH < 7.5 cm the NFI included a sampling in subplots of 3.54 m × 3.54 m. The number of plants in the subplot was extrapolated to the forest plot. Finally, to understand the efficiency of the models to predict AGB_{pot} over different spatial scales we conducted a sampling of eight plots of 100 m × 100 m (1 ha) in mature forest. Within each of the one-hectare plots, a nested sample was established. All the trees ≥30 cm in DBH were measured and geotagged; trees with DBH ≥20 cm were sampled in a sub-plot of 50 m × 50 m; trees with DBH ≥10 cm were recorded in a subplot of 25 m × 25 m; and trees with DBH ≥1 cm were collected in a subplot of 300 m^2. The information was up-scaled into 8 ha (n = 1), 4 ha (n = 2), 1 ha (n = 8), and 0.25 ha (n = 32) pseudo-replicated plots.

Out of the 188 samples, 83 plots were selected because they were at least 50 years old and considered as mature forest (Fig. 1). For this study, the AGB of mature forest and the AGB_{pot} were assumed to be equivalent. To ensure that only mature forests were included different approaches were considered: (1) The absence of human tracks, invasive species, logging or any kind of wood extraction or burned areas, with no evidence of cattle inside or around the plot. (2) Areas plots located over 2 km from agricultural fields. (3) Remote sensing time series data to ensure that the forest w mature (aerial photographs for 1985 and 1995, and Google Earth-DigitalGlobe imagery for 2004–2014, and when available LandSat imagery for 2013 and 2014). This process was based on a similar approach followed by Powers et al., (2009) in which the aerial photographs (Shoshany, 2000, 2002) are visually analyzed to ensure that since 1985 they showed mature forest surrounded by a matrix of natural forest.

The estimation of the AGB for each sample plot was done by using allometric equations applicable to TDF (Table 1). In total seven of the most common allometric equations were used to mitigate uncertainties of AGB estimates: from them, two equations are specifically calibrated on Mexican TDF (Martínez-Yrízar et al., 1992; Návar, 2009), two on pantropic forests (Chave et al., 2005) and three globally applicable (Brown et al., 1989; Brown, 1997). The allometric equation developed by Brown et al., (1989) is not recommended to be used in trees under 5 cm in DBH, so trees that fell into that category were adjusted by using the mean value from the remaining six allometric equations. This study did not include the recent equation reported by Chave et al., (2014) because it shows similar performance to those published in 2005. For each plot the upper and lower estimates from across the 7 allometric equations were rejected as outliers, thus the AGB estimates used of this study is the mean across the remaining five equations. These biomass estimates are then used to provide the training and validation dataset for the remainder of the analysis. This approach has suggested producing similar AGB estimates than those derived from field observations (van Breugel et al., 2011) and/or the best allometric equation (in concordance to the calculations derived from Ketterings et al., 2001, Djombo et al., 2010; Rutishauser et al., 2013).

A stand-level wood density average was derived, as recommended by Baker et al. (2004). Wood density was constructed based on a random collection of 28 plots during the dry season (N = 121 trees). For C content, 6 mixed samples from different stems and plots were