Volunteered Geographic Information and the Future of Geospatial Data

Cláudio Elízio Calazans Campelo
Federal University of Campina Grande, Brazil

Michela Bertolotto
University College Dublin, Ireland

Padraig Corcoran
Cardiff University, UK

A volume in the Advances in Geospatial Technologies (AGT) Book Series

IGI GLOBAL
DISSEMINATOR OF KNOWLEDGE
www.igi-global.com

Published in the United States of America by
IGI Global
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

CIP Data Pending
ISBN: 978-1-5225-2446-5
eISBN: 978-1-5225-2447-2

This book is published in the IGI Global book series Advances in Geospatial Technologies (AGT) (ISSN: 2327-5715; eISSN: 2327-5723)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 8
Mapping Regional Landscape by Using OpenstreetMap (OSM):
A Case Study to Understand Forest Patterns in Maya Zone, Mexico

Di Yang
University of Florida, USA

ABSTRACT

A forest patterns map over a large extent at high spatial resolution is a heavily computation task but is critical to most regions. There are two major difficulties in generating the classification maps at regional scale: large training points sets and expensive computation cost in classifier modelling. As one of the most well-known Volunteered Geographic Information (VGI) initiatives, OpenstreetMap contributes not only on road network distributions, but the potential of justify land cover and land use. Google Earth Engine is a platform designed for cloud-based mapping with a strong computing power. In this study, we proposed a new approach to generating forest cover map and quantifying road-caused forest fragmentations by using OpenstreetMap in conjunction with remote sensing dataset stored in Google Earth Engine. Additionally, the landscape metrics produced after incorporating OpenStreetMap (OSM) with the forest spatial pattern layers from our output indicated significant levels of forest fragmentation in Yucatan peninsula.

INTRODUCTION

Incorporating Road Networks into Forest Landscape Mapping

Forest ecosystems are one of the most important ecosystems. Forests cover one-third of the surface land and contain nearly 80% of terrestrial biodiversity. Forests also provide a broad range of ecosystem goods and services for human well-being (Aerts & Honnay 2011; Costanza et al., 2006). Forests contribute significantly to the terrestrial carbon cycle and sequester massive amounts of anthropogenic CO_2 emissions. In landscape ecology perspective, forest landscape is defined as a spatial mosaic of arbitrary

DOI: 10.4018/978-1-5225-2446-5.ch008

boundaries containing forest patches that are able to interact functionally (Turner 1989). Also, forests are amongst the most biologically rich terrestrial systems and provide broad ranges of goods and services for human well-being (Costanza et al., 2006). However, human activities are negatively transforming and altering the global forest patterns and structures. Human activities, such as road-building and resource extraction (e.g., timber harvesting and timber production), alter forest patterns and landscapes, thereby fragmenting continuous forests into smaller and more isolated patches (Jaeger et al., 2007). Understanding forest landscape patterns is critical in ecological research and forest land management. Remote sensing and GIS science are powerful tools that make monitoring forest dynamics over decades at regional scale possible. However, interpreting remote-sensing-derived images is not a new technology since the first aerial photographs (Carls, 1947) and satellite images (Estes et al., 1980 and Landsat 1). The recent increase of satellite sensing prompts the rapid development of open geoscience and helps us see the world in a more complete way.

Moreover, the updates for most land cover map projects are far from frequent, particularly at regional scale due to the uneven distribution of data resources. Most developed countries have a well-established infrastructure to provide high quality, large scale and comprehensive dataset on land cover/land-use for a variety of purposes. North America and Europe have well-done QA/QC assessments, such as COoRdination of Information on the Environment (CORINE), National Land Cover Database (NLCD) and Coast Change Analysis Program (C-CAP). Although land cover mapping techniques are well studied and have been applied to map most of the developed areas; however, there are still difficulties mapping land cover and land changes at a large scale with high in spatial and temporal resolution, especially for areas with high heterogeneity. In the developing world, such high-quality datasets are limited, and most developing countries lack the resources to collect extensive ground data.

Mapping forest fragmentation by incorporating road network is a fundamental basis of forest landscape ecological analysis (Haddad et al., 2015, Abdullah et al., 2007). A variety of forest fragmentation results have been characterized from the regional to global scale (Espírito et al., 2014, Haddad et al., 2015). McGarigal et al., investigated the landscape structure changes of the San Juan Mountains from 1950-1993 and found that roads had a more significant impact on landscape structure than logging (McGarigal et al. 2001). Wickham examined forest changes in cumulative distribution based on patch sizes for the Chesapeake Bay, Virginia (Wickham et al. 2007). However, there is persistent difficulty in delineating ecologically relevant patches based only on remote sensing images. In landscape ecology perspective, the significant impacts of road networks are continuously disregarded, not to mention the ignorance of small roads. In other cases, the systematic mapping of roads excludes small roads (e.g. logging roads and trails). Moreover, the effects of small secondary roads remain largely unknown. We therefore urgently need a more comprehensive road network database to evaluate the effects of road-caused fragmentation. As human activities make forestland more accessible, road building itself becomes a domino effect increasing secondary roads.

Roads inherently cut the existing forest patches into pieces, which is one of the most obvious reasons for forest fragmentations. Road networks are also treated as one of the most important human factors that cause forest fragmentation (add more citations here). Roads and the adjacent area contribute upwards of 20% of American total geographical land, as Forman et al., analyzed in 2000 (Forman et al. 2000). In some heavily forested areas, road and water features are hiding under canopy, so we have difficulty discovering and mapping them using regular classification techniques, especially when we want to explore the forest patterns. Also, there is persistent difficulty in delineating ecologically relevant patches based only on remote sensing images. Besides, road networks have significant secondary impacts on ecosystems and the distributions of species. The assessment of road databases is, therefore, imperative/crucial, and roads must be mapped as accurately as possible. Otherwise, the impacts end ecological effects of roads on the landscape might be underestimated. OpenStreetMap created the best map when considering temporal change trajectories by providing "up-to-date" data.

Volunteered Geographic Information (VGI) and OpenstreetMap

Volunteered Geographic Information (VGI) opens a new era in mapping and visualizing our world (Elwood et al., 2008, Haklay et al., 2008). OpenStreetMap (OSM) is one of the well-supported VGIs, and it has been studied and applied in multiple disciplines, but there are still large amounts of information needing to be investigated. OSM has its unique advantages, which might fit a wider range of applications over the official geographic databases. From the geographic perspective, scaling has always been a prominent and prevailing issue. Because of the growing coverage of spatial big data and cloud computing, OSM attracted our attention to its suitability in regional and continental scale mapping (Zook et al., 2010). Recent studies focused on land cover and land-use change accuracies by incorporating OSM (Estima et al, 2013 and 2015; Johnson 2016, Jokar et al., 2013, See et al., 2013 and 2015). In this study, we aim to incorporate OSM into regional land cover databases to investigate and quantify landscape fragmentation. Spatializing the distribution of the forest patches is a concrete way to measure the landscape structure and understand the forest patterns. The accuracy of VGI and OSM data are studied (Hakley et al, 2010). Moreover, the combining of remote-sensing data and OSM constitutes a powerful tool to monitor, characterize and quantify land cover and forest-patch distributions. The product will be a major source of information for researchers and policy-makers for investigating land parcels.

For the objective of optimizing our framework, we chose the Maya Zone of the Yucatan Peninsula, Mexico as a case study. The Maya Zone (Zona Maya) is a geographic area extending from north Quintana Roo to the borders of Belize, and it is known for its large forest coverage area and wide diversity of civilizations (Scarborough 2010). The Maya Zone of the Yucatan Peninsula is home to the most densely forested region in Mexico, which has a massive amount of forest cover and is also well known as native Mayan land. Current road networks have strongly influenced the distribution of forest patches in the Yucatan Peninsula. In the Maya Zone, local communities manage a large amount of forest. Fieldwork included in this study explores nine different sites covering a large part of the Quintana Roo state and includes at least three different types of land tenures (*Ejido*, Private land and conservation land). Spatializing the distribution of the forest patches is a concrete way to measure the landscape structure and understand the forest patterns.

In this study, we proposed a framework to improve the potential of land cover mapping based on selected features in OpenstreetMap in Google Earth Engine. Also, we were able to build a database demonstrating the near real-time road effects on forest-pattern spatial distributions. In addition, we validated and assessed the potentials of mapping land cover and land change at a regional scale with high heterogeneity.

STUDY AREA AND SOCIAL BACKGROUNDS

The Maya Zone (Zona Maya) in our study is a geographic area extending from North Yucatán state to the borders of Belize locating on the Yucatan Peninsula, and it is known for its wide diversity of civilizations (Bray et al., 2004). Yucatan Peninsula is a region dominated by semi-deciduous, tropical dry forest with distinct dry seasons (October - May) and wet seasons (June – September). Average forest height of the dry forest is 10-20m at maturity status and a lot of waterbodies are hiding under canopy because of its distinctive karstic topography (Flores et al., 1994, White and Hood. 2004). Our study area contains two states: Quintana Roo and Yucatan (Figure 1). These two adjacent states of interest were selected for comparison since their social backgrounds and different forest management practices. Also, the locations of these two regions were chosen because of the dominant land covers. In Yucatán state, the land cover is dominated by cropland and smallholder agriculture and Quintana Roo state is dominated by forest and secondary forest.

Yucatan Peninsula (the Maya Zone) has an organized combination of pre-colonial indigenous Mayan people and post-colonial development (Turner et al., 2001). In forest management and land-use practice perspective, the structure of land tenure and resource use rights essentially obligates a community-based approach in Mexico. At least three types of land tenure exist in the study area: *ejido* land (largely usufruct), private land, and conservation land. An *ejido* is a uniquely Mexican form of communal land tenure that covers large portions of the rural landscape in many states. The areas that covered the land tenure types including *ejido* land, private land and conservation land.

This uniquely Mexican form of communal land tenure covers large portions of the rural landscape in many states, and such land property is called an "*ejido*". *Edijos* and the local communities have maintained their traditional methods of managing the forest sustainably for thousands of years.

For cropland cover, local Mayan people keep the tradition of shifting cultivation known as milpa managed by *ejidos* (Klepeis and Vance 2003). Milpa croplands include the crop types of maize, squash, yoka and beans. In 1992, residents were given rights for the land based on revised Article 27 of the Constitution (Geoghegan et al., 2001). So we hypothesized that the privatization of the *ejidos* might change the land-use and forest management practice strategy that cause the land cover change which can be observed by remotely sensed satellite images. This study seeks to determine a way in which the forest cover patterns in the Quintana Roo state can be understood and the how the characteristics of this heavily forested land can be evaluated at a regional scale. This type of broader view enables the identification of the interactions between local scale (communities) and regional scale (forest management).

METHODS

In this section, we introduce the framework of producing a land cover dataset by embedding the features from OSM using Random Forest (RF) machine learning techniques. Also, a landscape ecology analysis will be performed by incorporating road networks into regional land cover maps to investigate the landscape pattern matrix. In this case study, we divided our methods into two parts: 1) Mapping land cover by exploring the feature points from OSM by using machine learning classification techniques. 2) Incorporate OSM road networks into the classification map, which produced from part 1, to quantify road-caused forest fragmentations.

Conceptual Model

To examine how road networks affect the forest pattern distributions in the Maya Zone, we proposed a measurement to refine the forest map by OpenStreetMap. Road-caused forest fragmentation was measured

Figure 1. Study region of the Maya Zone in Yucatan Peninsula, Mexico

on the land cover maps after superimposing full national road and street maps from open source OSM roads database. We incorporated updated roads data into our forest land cover map so that the maps include additional classes of roads such as private access roads and driveways in rural areas, small service roads or alleys in urban areas, and forest access roads. When superimposing road maps, we converted all the forest pixels that contained at least one road segment to non-forest pixels (30-m spatial resolution based). No distinctions were drawn between the types of road, traffic volume, or other factors. By superimposing our resulting land cover maps with road networks data, the refined forest patches map will be compared and the spatial distribution will be analyzed. Set our classification strategy with appropriate numbers of training points and their weights.

We chose random forest (RF) as our classier because it is now a widely used algorithm for remote sensing image classification. It has an excellent ability to handle high dimensional raster database, which perfectly fit our research framework. Random Forest is a classifier that produces many classification and regression-like trees, where each tree is grown according to the split training data. RF also provides an accuracy assessment called "out-of-bag" (OOB) error to avoid bias during the validations.

As our computing platform, Google Earth Engine (GEE), anewly developed data analysis environment, enables large-scale spatial analysis by its special infrastructure and automatically parallelized computation techniques (Padarian et al., 2015, Patel et al., 2015). By storing and bringing together large amounts of earth observations data on Earth Engine's server, GEE is able to analysis big data "on-the-fly" (Yu et al., 2012). We can also upload our private and personal database into the server as assets to achieve time series generating, zonal statistics, spectral analysis and many machine learning classification techniques. All the data we used in this study as inputs for the classifier are completely embedded in GEE and we concluded java program codes for any repeatable analysis of the other regions. With the embedding of GEE, the regional land cover mapping can be done in a much simpler way, as many spatial and temporal covariates are already available in GEE server and the calculation itself is parallelized automatically to optimize the computing time cost. By uploading our own training sample from OSM, we need to generate a fusion table representation the geometry including class types of our training sample extracted from OSM.

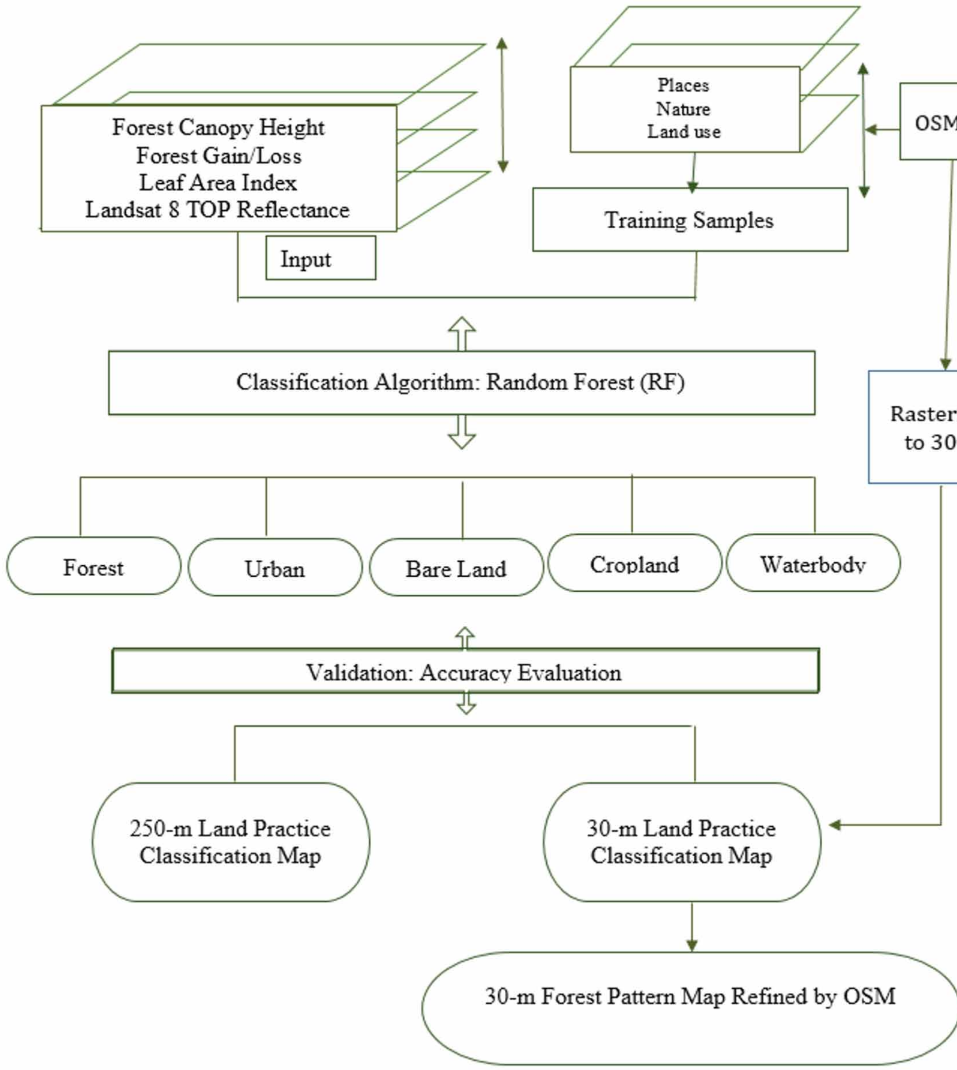
The four-step procedure for mapping forest patterns and qualifying road-caused fragmentation is summarized below, and the flow chart of the process is shown in Figure. 2:

1. Extract the datasets of "places", "land-use", and "nature" from OSM, set the land cover classification strategy based on research objective, reclassify and generate a training dataset. (In this study, we reclassify the points into urban, waterbody, forest, cropland and cropland and bare land five classes.)
2. Validate with Google Earth both on time and space for the training dataset.
3. Collaborate the spatial and temporal covariates as inputs for Random Forest classifier, and run Random Forest classifier by using Google Earth Engine API, and analyzed the classification accuracy for each classification type.
4. Extract the forest land from the resulting map at step 3, superimpose rasterized OSM road networks and run cluster analysis to generate forest pattern map.1

Mapping Land Cover Using GEE and OSM

OpenStreetMap (OSM) is one of the most popular and well-supported Volunteered Geographic Information (VGI) (Mooney et al., 2010). All OSM data was downloaded from the Geofabrik website http://download.geofabrik.de. The main reason why we used OSM was to create the fully free and openly accessible map of road networks. Community volunteers collect geographic information and submit it to the global OSM database (Ciepluch et al., 2009). The impact of the OSM has been assessed and is significant. Girres and Touya (2010) performed a spatial analysis of the quality of OSM street network representations in the UK and France respectively through a comparison to ground-truth data obtained

Figure 2. Flowchart illustrating the procedure of mapping forest land and quantify forest patterns



from the corresponding national mapping agency. Both studies found that on average the quality of the data was reasonably good but exhibited significant spatial heterogeneity. Neis et al 2012 analyzed how the quality of the OSM street network in different regions of Germany changed between the years 2007–2011.

Our study area of the Maya Zone is composed of cropland and forestland covering around 95% of the whole Yucatan Peninsula. The OSM database has a full coverage of our study area with the following sub-datasets: places, points, railways, roads, waterways, buildings, land-use and nature areas. It's hard to obtain accurate ground information that covers a broad study area from the field. We have to rely on the visual interpretation of high spatial resolution remotely sensed data for ground information; however, it is time consuming and labor intensive. To facilitate the virtual interpretation, we selected two publicly available datasets, OpenStreetMap features and Google Earth virtual global, and integrated them as our training and validation data resources. The OpenStreetMap feature classes we applied for the thematic map including “land-use”, “natural”, “places”, and “waterway”. The “land-use” is a polygon class identified forests, residential areas, industrial areas, etc. In the class “natural”, all the water body, parks, and recreation areas are represented by polygons. The point class “places” includes urban and suburb landmarks and attractions. The “waterway” is different with water body, it is a polyline feature indicated rivers, canals, streams, ditches, etc. All of these classes were converted to point features and we used a stratified randomly distributed sample design to extract training and validation points. Then we integrated the selected points with Google Earth virtual globe for visual interpretation. Google Earth virtual globe offers substantially high-resolution imagery and enabled us to assess the five land cover classes in all scenes. The reason we used Google Earth is its higher spatial resolution with a time span of more than 20 years. So we will get more spatial information, what types of land practice occur and the land cover type during this spatiotemporal validation. More details about the training sample setting strategy will be represented as a flow chart. We converted the randomly selected points into a Keyhole Markup Language (KML) file. Each selected point centered at a 30×30-meter polygon, one-pixel image size. The selected points and associated polygons overlaid in Google Earth virtual globe for visualization. We identified and assigned the classes of selected points based on the OSM feature types and their surrounding areas. The selected points were split into training and validation data sets. One-third of points were randomly selected as validation data. With a stratified random sampling design, the points were selected from five strata. The main focus of the thematic map was the forest and forest is the dominant domain.

The training samples we extracted from OSM were re-classified into five classes: Urban and built-up, waterbody, forest, cropland and bare land. Figure 3 showed the spatial distribution of our training samples of each class. Feature points density is virtualized in Figure.3. It represented the spatial distribution of the training points geographically and also illustrated the degree of urbanization.

The spatial and temporal covariates used in the land cover classification include NDVI-based reflectance, forest canopy height, forest gain and loss, and leaf area index (LAI). Those spatial and temporal covariates are significant indicators of forest ecosystems, and those remotely sensed inputs are consistently and accurately derived. All the data we used for our land cover map were up-to-date and subset to the boundaries of our study area of Yucatan peninsula and resampled to 30-m spatial resolution. In Table 1, we list our existing data for mapping task. All the data we used as spatial and temporal covariates cover our study regions and can be extended to a global scale. Therefore, it is repeatable for the other study regions if there are similar research objectives.

1. Landsat 8 Top-of-atmosphere (TOA) (Chander et al., 2009) of 2016: This composite image reflects the greenest pixel during the whole time period based on NDVI. (Google earth engine ID: LANDSAT/LC8_L1T_ANNUAL_GREENEST_TOA)
2. Global forest canopy height version of 2005 (Simard et al): This dataset represents the canopy height on the global level by incorporating Geoscience Laser Altimeter System (GLAS) and ancillary data. (Google earth engine ID: NASA/JPL/global_forest_canopy_height_2005)
3. Hansen et al., global forest change dataset V1.2 (Version of 2015): Hansen et al., global forest change database is the first database describing and monitoring the forest changes at a global-scale. We extracted 3 layers of “treecover2000” “forestloss” and “forestgain”. Here the layer of “treecover2000” is defined as canopy closure for all the vegetation taller than 5m and covers a percentage

Figure 3. Distribution of the OSM-derived training data for creation of the forest pattern map in Yucatan Peninsula, Mexico

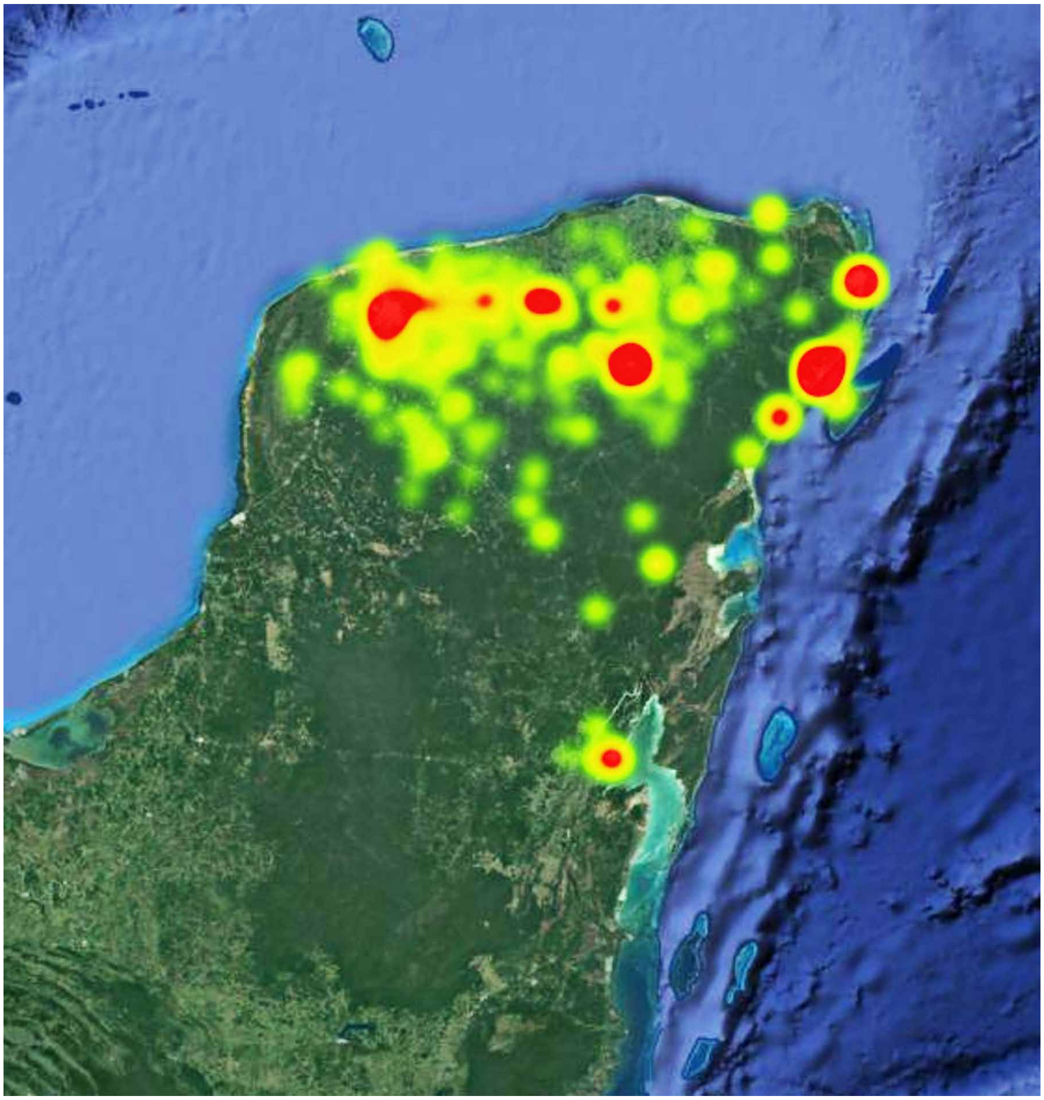


Table 1. Spatial and temporal covariates for random forest classifier

Database Name	*Objective*	*Google Earth Engine ID*	*Spatial resolution*	*Temporal resolution*
Landsat 8 TOP Reflectance	Reflectance	LANDSAT/LC8_L1T_ANNUAL_ GREENEST_TOA	30m	16-day
Forest Canopy Height	Canopy Height	NASA/JPL/global_forest_canopy_ height_2005	1km	Composite
Hansen et al., global forest change dataset V1.2	Forestland changes	UMD/hansen/global_forest_ change_2015	30m	Composite
MODIS 15 Leaf Area Index	LAI	MODIS/006/MCD15A3H	500m	4-day

Source: Google Earth Engine

in the analysis unit. "Forestloss" during the period of 2000-2014 is defined as a stand-replacement disturbance, or a change from a forest to non-forest, represented as either 1 (loss) or 0 (no loss). "Forestgain" during the period 2000-2014 is defined as the inverse of loss, or a non-forest to forest change (Hansen et al., 2013). (Google earth engine ID: UMD/hansen/global_forest_change_2015)

4. MODIS 15 Leaf Area Index global 500m: The MCD15A3H version 6 MODIS Level 4, Combined Fraction of Photosynthetically Active Radiation (FPAR), and Leaf Area Index (LAI) product is a 4-day composite database with 500-meter pixel size. (Google Earth Engine ID: MODIS/006/MCD15A3H)

Quantifying Road-Caused Forest Fragmentation

Uncertainty assessment is a critical component of any earth observation derived classification-mapping techniques. Usually, the road feature is hiding due to the dense forest canopy, whereas we were able to incorporate the road layer into the land cover map as a way of achieving our forest pattern map's assessments. The combining of remote sensing and Geographic Information Systems (GIS) is a powerful tool to monitor, characterize and quantify forest fragmentation in both time and space (Ritters et al., 2002, Heilman et al., 2002). The modeled land cover map cover those two states was re-classified into forest/nonforest dummy variable map to quantify road-caused fragmentation effects on forest patches. The non-forest class includes cropland, bare land, residential and water body. All data sets were projected into the coordinate system Albers Conical Equal Area to follow the North America mapping criteria. To characterize forest fragmentation, the OSM database was superimposed onto a 30-m spatial resolution raster vision. Comparing with our analytical extent, we assumed the interaction between urban forest and rural forest would be minimal. Additionaly, we also generated near-real time "Openroad density" map from OSM, which included all types of roads.

Choosing the unit and scale of analysis is important in quantifying forest fragmentation, as smaller landscapes are more sensitive to finer-scale patterns (Riitters et al. 2002). The effects of both spatial and temporal scale should be considered in forest fragmentation, because forestland is heterogeneous, and the structures, functions and ecological processes are themselves scale-dependent. We calculated the patch size-distributions at both the local and regional scales. The spatial distribution represented the diversity and heterogeneity of road-caused forest fragmentation; and the patch-size class map elucidated the landscape composition and configurations. Therefore, in this study, we produced the dataset at local and regional scale at a spatial resolution of 30m (Landsat spatial resolution) and 250m (MODIS spatial resolution), respectively.

RESULTS

In this section, we present the results obtained by applying the conceptual model and the methodology described in Methodology section.

Analysis of Refined Land Cover Map

The map of the dominant land cover types in Yucatan Peninsula is shown in Figure 4, and the error matrix is Table 2. The landscape of Yucatan Peninsula, which consisted 2,222km^2, is comprised of 45%

forestland, 45% cropland, 12% urban areas. In the random forest classifier, we used to estimate the classification error without bias. In our study, the OOB error is 12% yielding a total classification accuracy of 91% of our random forest classifier building

In Table 2, we made the comparison between the forest cover and cropland cover between those two states. Quintana Roo is about 31.4% higher Yucatan at forest cover and 31% lower at cropland cover. Yucatán state has a higher average road density of 4.53km/km^2 than Quintana Roo state 2.55 km/km^2 derived from OSM road feature. This huge difference is not only just because the degree of urbanized, but also there is a big difference between the land tenure, tourism policies, and timber rights in those two states.

We constructed an error matrix from the results of our land practice map, and calculated the commission, omission and the overall errors (Table.3). The error matrix is shown in Table. 2. Rows represent the reference data and columns represent the classification results. Producer's accuracy and user's accuracy are listed in the last row and column, respectively. The overall accuracy of the thematic map was 92%, and the kappa coefficient was 77%. Overall 96% of the forest pixels were correctly identified as forest, and 95% of the forest area are successfully classified as forest. Bare land shows the lowest user's accuracy of 72%, which is because of the light weight in the training samples, and VGI contributors define bare land differently. Moreover, the training samples cannot fully represent the surrounding land cover types, for example, a small conservation park in urban regions that was marked "forest".

Forest Patterns Map

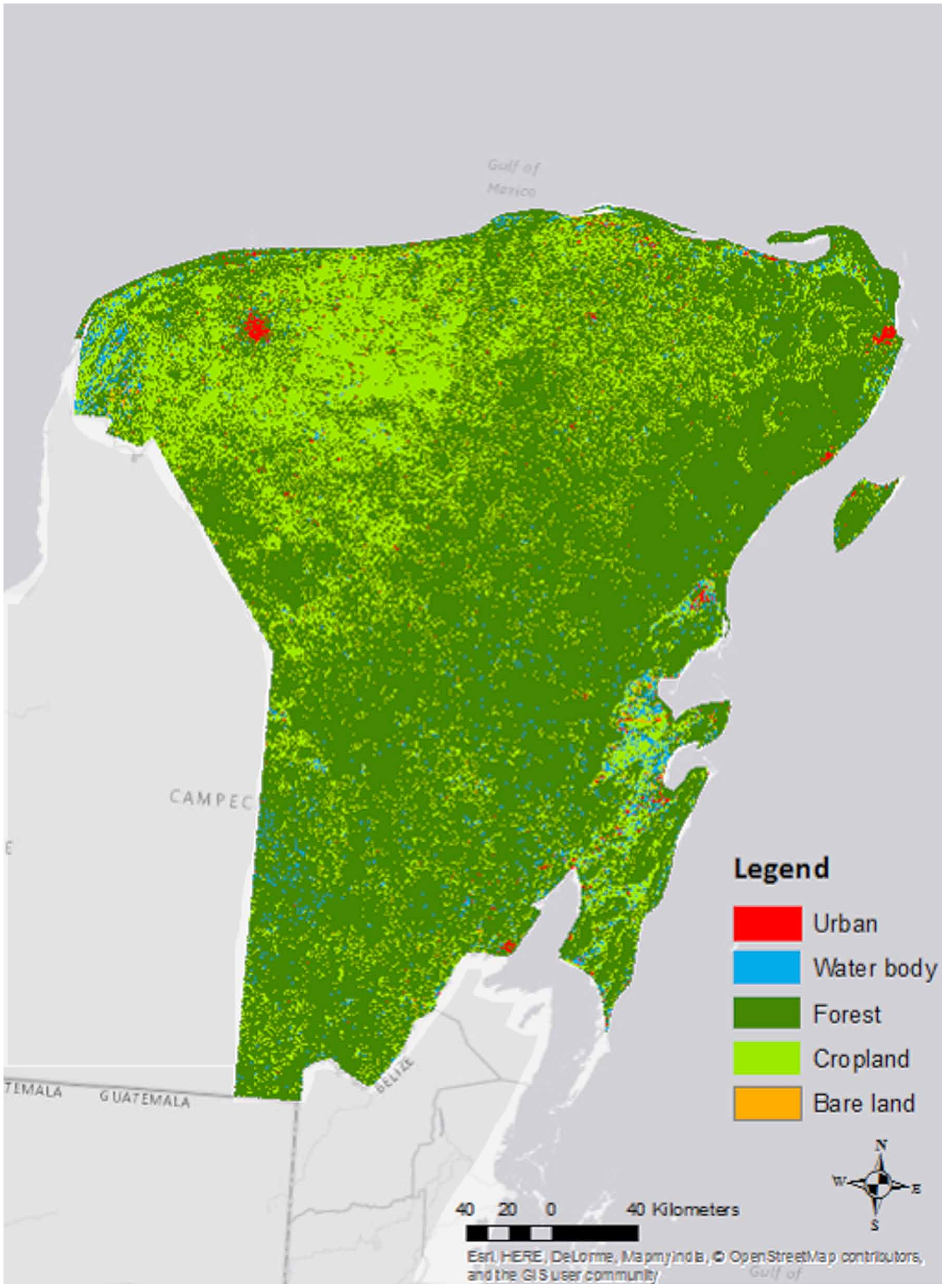
By proposing the framework by incorporating OSM road networks, we have the capability to spatially access the potential of forest fragmentation according to forest patch sizes and its surrounding land covers and road densities. Figure.5 shows the refined spatial distribution of forest patterns based on patch size. We can see clearly that there is the big difference between Yucatan and Quintana Roo state. In Yucatan forest pattern map, the visualization result tells us a large proportion of forest patches ranges from 6-100 ha which is relatively small and has a high risk for further deforestation and fragmentation. Moreover, in Quintana Roo state, the forest is less fragmented than Yucatán state, but we can still easily see the forest pattern was separated and divided by roads. In general, as the road networks increase, the forest fragmentation will increase by a wide margin. After incorporating the road network, there are 603 forest patches in the state of Quintana Roo. A total number of 477 forest patches are under the area of 500 Ha.

To account for the complexity spatial patterns at regional scale, frequency distributions are used heavily in quantifying the total number of patches and their patch sizes. Combining with the forest pattern map from Figure 6, a size-frequency distribution has the capability to represent locations where has a high risk of deforestation and forest fragmentation. As shown in Figure. 6, we calculated the size-frequency distribution of forest patches over the whole Yucatan and Quintana Roo state. Those total 15,000 forest

Table 2. Comparison of Forest and Cropland Cover between Quintana Roo and Yucatán states

	Forest Cover	Cropland Cover	Average Road Density (km/km^2)
Quintana Roo	86.20%	11.10%	2.55
Yucatan	44.80%	53.00%	4.53

Figure 4. Spatial distribution of dominant land cover types of Yucatan Peninsula

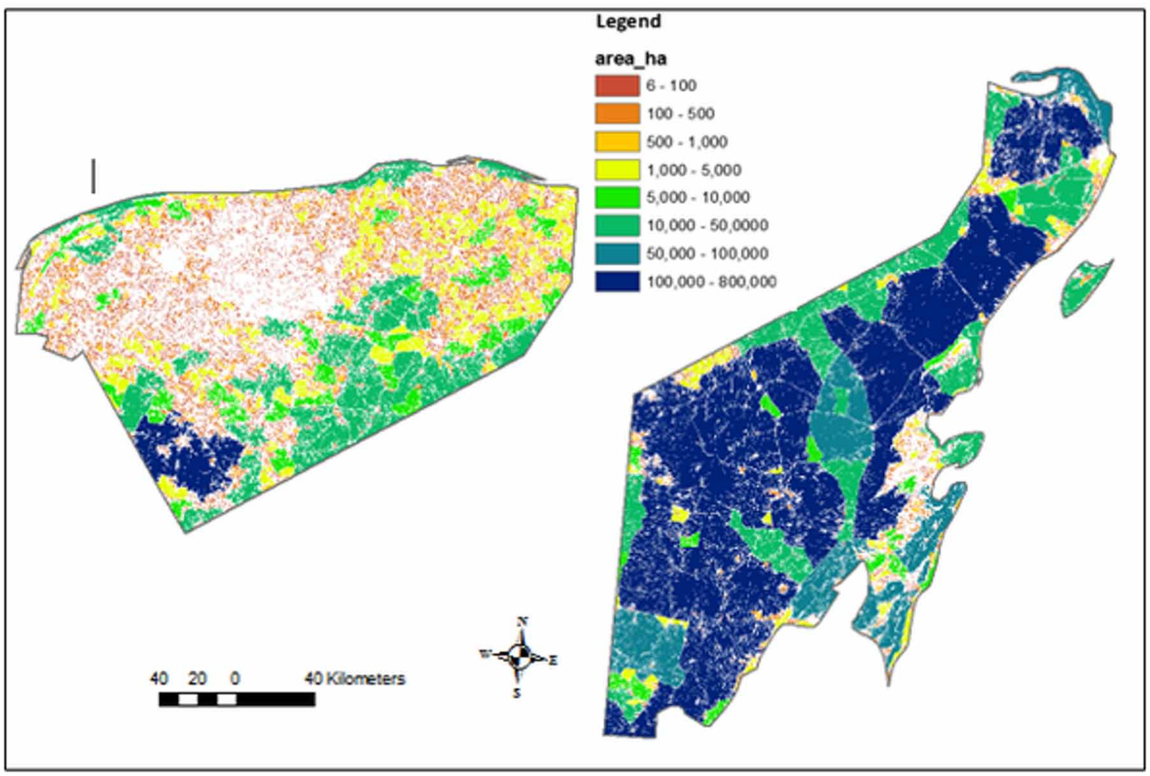


patches in Yucatán state just cover 3% of the forest cover, which can be understood with a high potential to get further fragmented. And the spatial patterns map tells us where those patch located and we can provide some advises to the land managers or decision makers. In Quintana Roo state, no more major road building will be recommended within the large continuous patches with size more than 10,000 ha.

Table 3. Classification accuracy from random forest classifier OOB of each land cover type

	Urban	Water	Forest	Cropland	Bare land	Points	User's acc. (%)
Urban	36	0	7	2	0	45	80
Water	0	56	10	7	0	73	77
Forest	4	15	1280	40	12	1351	95
Cropland	0	4	28	147	1	180	82
Bare land	0	0	4	10	36	50	72
Total	40	75	1329	206	49	1699	
Producer's acc. (%)	90	75	96	71	73		

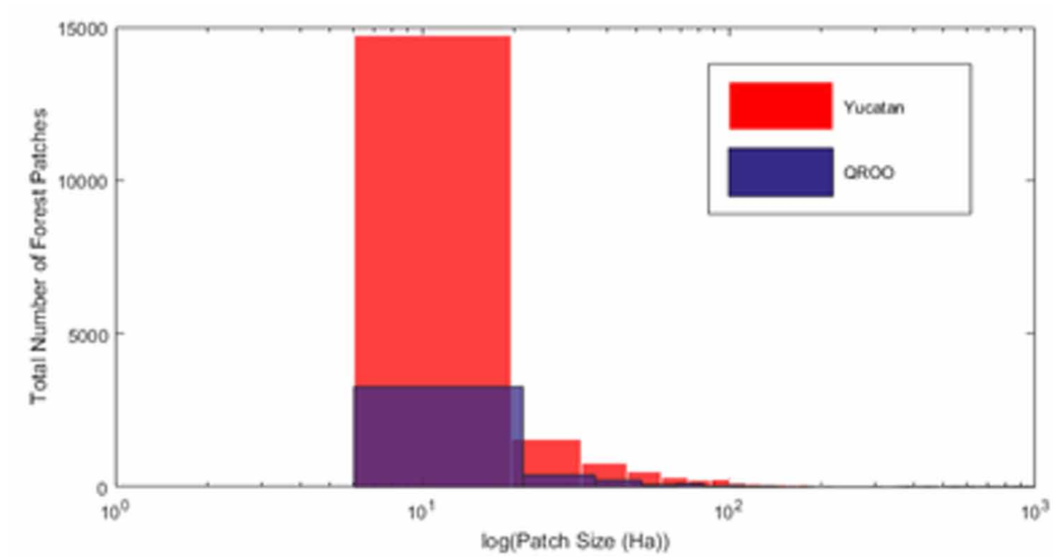
Figure 5. Spatial patterns classified by patches area of Yucatán state and Quintana Roo state



DISCUSSIONS

This study described the potential that OSM's ability to mapping forest patterns, and its critical rules in setting training samples. We incorporated the OSM into remote sensing satellite layers to classify land and forest practices. We then compared the dominant land cover types between Quintana Roo and Yucatán states.

Figure 6. Forest Patches Size-frequency Distribution of Yucatan and Quintana Roo state



OpenstreetMap Road Density

At the regional scale, road density is an overall term used to describe the road networks. It has been proposed as a broad index of roads' ecological effects in a landscape. We calculated OpenStreetMap derived road density of our study area based on our land cover map including the land transition matrix.

Ownership Affecting Forest Patterns Distributions

To support our map, to prove the by combining the satellite images and VGI data have the potential to tell us what is going on over the ground and the landscape. We also need people in situ testimonies to monitor and explain the ground truth because the land and forest management practices on Maya Zone is unique and cannot be repeatable by the others according to the historical and climate environments (Corbera et al., 2011). We took nine places in Quintana Roo state and the state border between Yucatan and Quintana Roo state as the sub-case study. First of all, we concluded those places and classified them based on the management ownership types. In this research, land ownership in the state of Quintana Roo was separated into three types: *Ejido* and local communities land; private land; and conservation areas.

- ***Ejido* Land**: Chan Santa Cruz, Petcacab, Betania, Nuevo Durango
- **Private Land**: Don's land, and agriculture land in Jose Maria Morelos
- **Conservation Land**: Siijil Nol Ha, Punta Laguna, and Coba

Forest ownership is a vital link between human and environmental factors, particularly at the re. At regional scale, forest ownership patterns explain different types of land management practices and trajectories of land cover change (Turner 1996). Turner et al., also similarly showed a correlation between land ownership categories and land cover change patterns in North Carolina and Washington (Turner et al., 1996). Albers and Ando found that federal ownership of conservation lands created crowding-in and crowding-out effects of private ownership (depending on the states), thereby also indirectly influencing fragmentation (Albers et al, 2008). But in Mexico, not many researches have been done.

Much of Quintana Roo land is assigned to the ejidal system, which is managed by community-based planting and logging. In many *ejidos*, mahogany-a lucrative forest product, is used as the guide species to design forest management regimes. In the *ejido* of Petcacab, the total forested area is divided into five-year blocks with these in turn divided into annual harvesting blocks. In a given year, members of the Petcacab *ejido* are authorized to harvest a specified volume of wood. Within each harvesting block, only trees with a diameter at breast height (DBH) larger than 55cms are supposed to be harvested. In contrast, few people are involved in forest management practice per management unit in privately owned land. Within private land ownership systems, there is more freedom in forest management practices but less freedom regarding planting and harvesting. Conservation areas, on the other hand, are a form of passive management. These forested lands are "area where natural processes and natural disturbance regimes can develop without management intervention and where ecological and societal goals are given primacy" (Duncker et al., 2012). The characteristics of forest management practices in the Quintana Roo state vary by land ownership type, which in turn may have implications for the level of forest fragmentation associated with each of these land tenure systems. To examine the different implications of the land ownership types on forest fragmentation and hence biodiversity conservation, zonal statistics were calculated for each of the land tenures types and their forest management systems.

CONCLUSION

By using the OSM trained forest pattern map as the forest reference data, we produced the OSM refined forest parcels and analyzed the spatial size distribution of forest patterns. We successfully demonstrated the application of OSM as training sample and assessments layer embedding in land cover maps at a regional scale. The framework got a satisfactory performance of generating the maps with a total accuracy of 91% and Kappa coefficient of 77% and also did an excellent job incorporating OSM road features in characterizing forest patterns. In addition, by incorporating the OSM, the secondary and logging roads are shown to play an important role in causing fragmentation of the remnant forestlands. Landscape metrics for the Quintana Roo state were produced. Spatial frequency distributions of forest patches were calculated and analyzed. The results of the analysis showed significant levels of forest fragmentation throughout the both Quintana Roo and Yucatán state.

All platforms in this study are in active developments, GEE is building more machine learning classification techniques into the program, and increasing the computational limit for the goal of large-scale mapping. Moreover, more researchers and land managers are digitalizing their rural area maps into OpenstreetMap, which fills the gaps of VGI in rural areas.

OSM is an exciting new platform, which could be implemented in land cover mapping and assessing, because of its geo-located features and marks. The combination of Google Earth Engine with OpenstreetMap features will be a tool for large scale mapping and will present significant opportunities in the creation ideas because of its unique workflow to fulfill specific mapping needs.

FUTURE RESEARCH DIRECTIONS

Conserving forest biodiversity is a prerequisite for the long-term flow of forest ecosystem services. Forest management has evolved considerably in the last two decades, demonstrating significant positive impacts for conserving biodiversity and delivering economic benefits to local communities.

According to the temporal availability of our covariates, these up-to-date forest patterns maps can be extended to a multi-temporal changes database. It also has the capability to make predictions based on the spatiotemporal forest pattern changes and patch size distributions. There is still a strong need for a continental coverage land cover map with a higher temporal resolution, which will reflect the land cover transition processes at shorter intervals. This issue is urgent; in order to abate the rapidly increasing construction of road networks and the associated ecological impacts, we need better road network designs that combine both top-down and bottom-up processes, which would retain/preserve large forest patches and decelerate the parcelization rate. The use of open data source will be open our mind of exploring the area that doesn't have an official classification database. More information on the effects of land ownership and forest management, combined with the detailed road network and a continental coverage land cover maps can aid in thwarting further forest fragmentation by promoting more thoughtful road planning by land planners and decision makers.

REFERENCES

Abdullah, S. A., & Nakagoshi, N. (2007). Forest fragmentation and its correlation to human land-use change in the state of Selangor, peninsular Malaysia. *Forest Ecology and Management*, *241*(1), 39–48. doi:10.1016/j.foreco.2006.12.016

Aerts, R., & Honnay, O. (2011). Forest restoration, biodiversity and ecosystem functioning. *BMC Ecology*, *11*(1), 29. doi:10.1186/1472-6785-11-29 PMID:22115365

Albers, H. J., Ando, A. W., & Batz, M. (2008). Patterns of multi-agent land conservation: Crowding in/out, agglomeration, and policy. *Resource and Energy Economics*, *30*(4), 492–508. doi:10.1016/j.reseneeco.2008.04.001

Arsanjani, J. J., Mooney, P., & Zipf, A. (2015). Quality assessment of the contributed land-use information from OpenStreetMap versus authoritative datasets. OpenStreetMap in GIScience, 37-58.

Baskent, E. Z., & Jordan, G. A. (1995). Characterizing spatial structure of forest landscape. *Canadian Journal of Forest Research*, *25*(11), 1830–1849. doi:10.1139/x95-198

Bray, D. B., Ellis, E. A., Armijo-Canto, N., & Beck, C. T. (2004). The institutional drivers of sustainable landscapes: A case study of the 'Mayan Zonein Quintana Roo, Mexico. *Land Use Policy*, *21*(4), 333–346. doi:10.1016/j.landusepol.2003.11.001

Carls, N. (1947). *How to read aerial photographs for census work*. US Government Printing Office.

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, *113*(5), 893–903. doi:10.1016/j.rse.2009.01.007

Ciepluch, B., Mooney, P., Jacob, R., & Winstanley, A. C. (2009). Using openstreetmap to deliver location-based environmental information in Ireland. *SIGSPATIAL Special*, *1*(3), 17–22. doi:10.1145/1645424.1645428

Cihlar, J. (2000). Land cover mapping of large areas from satellites: Status and research priorities. *International Journal of Remote Sensing*, *21*(6-7), 1093–1114. doi:10.1080/014311600210092

Corbera, E., Estrada, M., May, P., Navarro, G., & Pacheco, P. (2011). Rights to land, forests and carbon in REDD+: Insights from Mexico, Brazil and Costa Rica. *Forests*, *2*(1), 301–342. doi:10.3390/f2010301

Corcoran, P., Mooney, P., & Bertolotto, M. (2013). Analysing the growth of OpenStreetMap networks. S*patial*. *Statistics*, *3*, 21–32.

Costanza, R., Fisher, B., Ali, S., Beer, C., Bond, L., Boumans, R., & Snapp, R. et al. (2007). Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological Economics*, *61*(2), 267–276. doi:10.1016/j.ecolecon.2006.02.023

Duncker, P. S., Raulund-Rasmussen, K., Gundersen, P., Katzensteiner, K., De Jong, J., Ravn, H. P., & Spiecker, H. (2012). How forest management affects ecosystem services, including timber production and economic return: Synergies and trade-offs. *Ecology and Society*, *17*(4), 50. doi:10.5751/ES-05066-170450

EEA Corine Land Cover. (n.d.). Available online: http://www.eea.europa.eu/publications/COR0-landcover

Ellis, E., & Pontius, R. (2007). Land-use and land cover change. Encyclopedia of Earth.

Elwood, S. (2008). Volunteered geographic information: Key questions, concepts and methods to guide emerging research and practice. *GeoJournal*, *72*(3), 133–135. doi:10.1007/s10708-008-9187-z

Espírito-Santo, F. D., Gloor, M., Keller, M., Malhi, Y., Saatchi, S., Nelson, B., & Palace, M. et al. (2014). Size and frequency of natural forest disturbances and the Amazon forest carbon balance. *Nature Communications*, 5. PMID:24643258

Estima, J., & Painho, M. (2013). Exploratory analysis of OpenStreetMap for land-use classification. *Proceedings of the second ACM SIGSPATIAL international workshop on crowdsourced and volunteered geographic information*, 39-46. doi:10.1145/2534732.2534734

Estima, J., & Painho, M. (2015). Investigating the potential of OpenStreetMap for land-use/Land cover production: a case study for Continental Portugal. In *OpenStreetMap in GIScience* (pp. 273–293). Springer International Publishing. doi:10.1007/978-3-319-14280-7_14

Flores Guido, J. S., & Espejel Carvajal, I. (1994). *Tipos de vegetación de la Península de Yucatán* (No. Y/581.61 E8/3). Academic Press.

Foody, G. M., & Boyd, D. S. (2013). Using volunteered data in land cover map validation: Mapping West African forests. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(3), 1305–1312. doi:10.1109/JSTARS.2013.2250257

Forman, R. T. (2000). Estimate of the area affected ecologically by the road system in the United States. *Conservation Biology*, *14*(1), 31–35. doi:10.1046/j.1523-1739.2000.99299.x

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., & Obersteiner, M. et al. (2009). Geo-Wiki. Org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, *1*(3), 345–354. doi:10.3390/rs1030345

Geoghegan, J., Villar, S. C., Klepeis, P., Mendoza, P. M., Ogneva-Himmelberger, Y., Chowdhury, R. R., & Vance, C. et al. (2001). Modeling tropical deforestation in the southern Yucatan peninsula region: Comparing survey and satellite data. *Agriculture, Ecosystems & Environment*, *85*(1), 25–46. doi:10.1016/S0167-8809(01)00201-8

Girres, J. F., & Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, *14*(4), 435–459. doi:10.1111/j.1467-9671.2010.01203.x

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., & Cook, W. M. et al. (2015). Habitat fragmentation and its lasting impact on Earths ecosystems. *Science Advances*, *1*(2), e1500052. doi:10.1126/sciadv.1500052 PMID:26601154

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning. B, Planning & Design*, *37*(4), 682–703. doi:10.1068/b35097

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning. B, Planning & Design*, *37*(4), 682–703. doi:10.1068/b35097

Haklay, M., Singleton, A., & Parker, C. (2008). Web mapping 2.0: The neogeography of the GeoWeb. *Geography Compass*, *2*(6), 2011–2039. doi:10.1111/j.1749-8198.2008.00167.x

Heilman, G. E., Strittholt, J. R., Slosser, N. C., & Dellasala, D. A. (2002). Forest fragmentation of the conterminous United States: Assessing forest intactness through road density and spatial characteristics. *Bioscience*, *5241*, 1–22.

Homer, C. G., Dewitz, J. A., Yang, L., Jin, S., Danielson, P., Xian, G., & Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United states-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, *81*(5), 345–354.

Jaeger, J. A., Schwarz-von Raumer, H. G., Esswein, H., Müller, M., & Schmidt-Lüttmann, M. (2007). Time series of landscape fragmentation caused by transportation infrastructure and urban development: A case study from Baden-Württemberg, Germany. *Ecology and Society*, *12*(1), 22. doi:10.5751/ES-01983-120122

Johnson, B. A., & Iizuka, K. (2016). Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid /land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. *Applied Geography (Sevenoaks, England)*, *67*, 140–149. doi:10.1016/j.apgeog.2015.12.006

Jokar Arsanjani, J., Helbich, M., & Bakillah, M. (2013). Exploiting volunteered geographic information to ease mapping of an urban landscape. *ISPRS-International Archives of the Photogrammetry*. *Remote Sensing and Spatial Information Sciences*, *1*(1), 51–55.

Jokar Arsanjani, J., Helbich, M., Bakillah, M., Hagenauer, J., & Zipf, A. (2013). Toward mapping land-use patterns from volunteered geographic information. *International Journal of Geographical Information Science*, *27*(12), 2264–2278. doi:10.1080/13658816.2013.800871

Klepeis, P., & Vance, C. (2003). Neoliberal policy and deforestation in southeastern Mexico: An assessment of the PROCAMPO program. *Economic Geography*, *79*(3), 221–240. doi:10.1111/j.1944-8287.2003.tb00210.x

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., & George, P. et al. (2001). The causes of land-use and change: Moving beyond the myths. *Global Environmental Change*, *11*(4), 261–269. doi:10.1016/S0959-3780(01)00007-3

McGarigal, K., Romme, W. H., Crist, M., & Roworth, E. (2001). Cumulative effects of roads and logging on landscape structure in the San Juan Mountains, Colorado (USA). *Landscape Ecology*, *16*(4), 327–349. doi:10.1023/A:1011185409347

Mooney, P., Corcoran, P., & Winstanley, A. C. (2010). A study of data representation of natural features in openstreetmap. Proceedings of GIScience, 150, 150-156.

Neis, P., Zielstra, D., & Zipf, A. (2012). The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet*, *4*(4), 1–21. doi:10.3390/fi4010001

OpenStreetMap in GIScience. (2015). *Experiences, Research, and Applications*. Springer.

OpenStreetMap Map Features. (2016). Retrieved from http://wiki.openstreetmap.org/wiki/Map_Features

Padarian, J., Minasny, B., & McBratney, A. B. (2015). Using Googles cloud-based platform for digital soil mapping. *Computers & Geosciences*, *83*, 80–88. doi:10.1016/j.cageo.2015.06.023

Patel, N. N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F. R., & Trianni, G. et al. (2015). Multitemporal settlement and population mapping from Landsat using Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation*, *35*, 199–208. doi:10.1016/j.jag.2014.09.005

Scarborough, V. L., & Burnside, W. R. (2010). Complexity and sustainability: Perspectives from the ancient Maya and the modern Balinese. *American Antiquity*, *75*(2), 327–363. doi:10.7183/0002-7316.75.2.327

See, L., Comber, A., Salk, C., Fritz, S., van der Velde, M., Perger, C., & Obersteiner, M. et al. (2013). Comparing the quality of crowdsourced data contributed by expert and non-experts. *PLoS ONE*, *8*(7), e69958. doi:10.1371/journal.pone.0069958 PMID:23936126

See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., & Siraj, M. A. et al. (2015). Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 48–56. doi:10.1016/j.isprsjprs.2014.06.016

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, *116*(G4), G04021. doi:10.1029/2011JG001708

Turner, B. L., Villar, S. C., Foster, D., Geoghegan, J., Keys, E., Klepeis, P., & Plotkin, A. B. et al. (2001). Deforestation in the southern Yucatán peninsular region: An integrative approach. *Forest Ecology and Management*, *154*(3), 353–370. doi:10.1016/S0378-1127(01)00508-4

Turner, M. G. (1989). Landscape ecology: The effect of pattern on process. *Annual Review of Ecology and Systematics*, *20*(1), 171–197. doi:10.1146/annurev.es.20.110189.001131

White, D. A., & Hood, C. S. (2004). Vegetation patterns and environmental gradients in tropical dry forests of the northern Yucatan Peninsula. *Journal of Vegetation Science*, *15*(2), 151–160. doi:10.1111/j.1654-1103.2004.tb02250.x

Wickham, J. D., Riitters, K. H., Wade, T. G., & Vogt, P. (2010). A national assessment of green infrastructure and change for the conterminous United States using morphological image processing. *Landscape and Urban Planning*, *94*(3), 186–195. doi:10.1016/j.landurbplan.2009.10.003

Yu, L., & Gong, P. (2012). Google Earth as a virtual globe tool for Earth science applications at the global scale: Progress and perspectives. *International Journal of Remote Sensing*, *33*(12), 3966–3986. doi:10.1080/01431161.2011.636081

Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake. *World Medical & Health Policy*, *2*(2), 7–33. doi:10.2202/1948-4682.1069