

Using Open-Source Data and Software to Analyse Land-Use Changes and Deforestation in Marqués de Comillas, Chiapas, Mexico (Work in Progress)

Iskar Jasmani Waluyo Moreno¹, José García Hernández², Agustín Bolóm Gómez¹ and Julio Cesar García Sampayo¹

¹Centro de investigación en Ciencias de Información Geoespacial, México

²Universidad de Chalcatongo, México

Abstract

This project aimed to exploit existing open data and open-source software to allow visualising and finding spatial coincidences between three groups of indicators: 1) agricultural activity, 2) land use and forest cover changes, and 3) spatial distribution patterns of vegetation loss. This paper presents preliminary results of developing an interactive map using R, Leaflet, and QGIS to perform a simple overlay analysis of layers containing better policy and decision-making indicators. The map uses open data available through the Mexican government platforms, and all processing, storage and publication are done using free open-source software tools. The data used in this project was limited to the Marqués de Comillas municipality because of its importance in Mexico's forest ecosystem. The maps are a work in progress and a continuation of research done for a geomatics graduate program in Centro de Investigación en Ciencias de Información Geoespacial.

Keywords: overlay analysis, Chiapas deforestation, open source geospatial data, open source geospatial software

1 Introduction

Agricultural expansion continues to be one of the main drivers of deforestation and forest fragmentation, and the associated loss of forest biodiversity. The deforestation rate in the 1990s was 16 million hectares per year and 10 million hectares per year from 2015 to 2020. (FAO & UN, 2020). Decisions on land use have prioritized the benefits of agriculture over forest conservation, which in large part explains continued deforestation in the tropics. In the short term, deforestation allows greater productivity, but long-term effects may be inverse due to soil degradation (Benhin, 2006). Although simplistic, analysing links between agricultural activity and deforestation can provide preliminary information regarding the drivers of land-use changes and deforestation (Benhin, 2006).

Deforestation drivers in Mexico is strongly influenced by contradictory public policies that, on the one hand, promote conservation and, on the other, promote productive activities in forest

ecosystems. Most recently, the federal program “Sembrando Vida” provides economic support to farmers who have access to at least 2.5 ha of workable land for agroforestry projects and are located in municipalities with social lag. The broad goals of the program are: promoting food auto sufficiency, increasing employment levels and improving forest cover. A review of newspaper notes and research papers regarding the program revealed mixed results. While the program has had positive results, it has also contributed to the deforestation of natural and endemic forest cover (Olvera, 2019; Forbes Staff, 2016; Acosta and Vera-Herrera, 2019; Arturo, 2020; Enciso, 2020; Quadri, 2020; CONEVAL, 2020). However, state-level data regarding the program is clear, at municipal levels the information about the “Sembrando Vida” program is not (Colter et al., 2020). Furthermore, no current evaluations or recommendations were found that analyse the effects of the program on forest coverage. In this regard, this article presents preliminary results of a interactive map prototype intended to use open data and open source software to facilitate monitoring land use changes and agricultural activities at a sub-municipal level that could provide useful information for better evaluation and monitoring of programs such as these.

2 Proposal

There is a large amount of data regarding forest cover, land use and agricultural activities in Mexico. Additionally, satellite imagery from several different sources is also freely and widely available. Although easily accessed, the data used is often stored in different repositories and in different formats and dimensions, making combining data of related phenomena, such as agriculture and deforestation, complex. The following table compares three platforms that allow visualisation of this type of data in Mexico.

Table 1: Comparison of existing portals with georeferenced data in Mexico

	GAIA INEGI	Geoportal CONABIO	Geoweb Chiapas
Datasets	20	>13000	>200
Allows data download	Not all layers	Yes	Yes
Code is available	No	No	No
Allows overlays	No	No	No
Regional or National	National	Both	Regional
Can filter data by region	Not all layers	No	NA

Albeit not exhaustive, none of the platforms reviewed allowed overlaying layers of data at state or municipal levels, and none of them provided code or development information. Exploiting the full potential of all of this data requires it to be easily visualised and compared, which we consider as important as developing new models and generating new data. In this sense, this paper presents preliminary results of a prototype of a simple GIS that allows basic data visualisation by homologating data from different open data sources and open-source software. The map is intended to allow simple overlay analysis of the data that contributes to

finding likely or best locations for a specific phenomenon (ArcGIS, 2020). The system is designed to help to answer three general but essential questions at a local scale:

1. Where are the changes in land use and forest cover clustering?
2. What are the main agricultural activities where these clusters are occurring?
3. What may the different combinations of these clusters/activities indicate?

This work is not focused on designing or implementing a complex model but rather on attempting to configure, process and compare readily available data from different sources in a simple interactive map that allows overlay analysis of the following groups of data (see table 2):

Group 1: Agricultural activities

Group 2: Land cover change

Group 3: Spatial autocorrelation of vegetation loss

End users will be able to select different combinations of layers data and overlay analysis of these combinations as well as access raw and processed data used.

3 Case Study: Marques de Comillas Chiapas, Mexico

This pilot project was limited to the Marqués de Comillas municipality because of its relevance Mexico's forest; it is located in the last redoubts of high evergreen forest in the country, an ecosystem that went from 10 million hectares to just over 1 million that contains remnants of the forest that allow the mobility of different species and also bordered the Montes Azules Biosphere Reserve which preserves the greatest diversity of species in Mexico (Carabias, de la Maza & Cadena, 2015; Flores, 2019). Productive activities in this region are in constant conflict with forest conservation; finding ways to guarantee the population's well-being should balance productive capacity with conserving the ecosystem. It is widely accepted that agriculture is a driver of tropical deforestation and its large scale environmental consequences. Thus through proxies such as plot size, it may be possible to analyse details that provide accurate data regarding forest cover changes (Dang, 2019).

4 Data Processing and Map Development

Data sources: Data from Censo Agropecuario 2007¹ (INEGI, 2007), AMCA 2016² (INEGI, 2016), USV series III and VI³, was pre-processed in R in order to generate data frames that could be easily concatenated with maps and compared with land cover change information. The results were then stored as .Rdata and .geojson files in a Github account.

¹ Censo Agrícola, Ganadero y Forestal 2007 – Agriculture, Livestock and Forestry Census 2007

² Actualización del Marco Censal Agropecuario 2016 – Update of the Agriculture, Livestock and Forestry Census 2016

³ Uso de suelo y vegetación serie III and serie VI – Land use series III and series IV are nationwide maps of landuse published in 2005 and 2016 respectively.

Data Processing: Data processing was divided into three groups, as shown in the following table.

Table 2: Groups and indicators

* The census was taken in 2007 and 2016 do not allow comparing data directly. Only data from 2016 data was used for creating thematic maps for the time being.

Group 1: Agricultural activities (2016)*	Group 2: Land cover change (2007-2016)	Group 2: Spatial autocorrelation of vegetation loss (2007-2016)
Parcels with some kind agricultural activity.	Continued anthropic activity (% of ha)	Spatial autocorrelation of changes in primary vegetation (no scale)
Agricultural activities (% of parcels)	Conserved forest cover (% of ha)	Spatial autocorrelation of changes in secondary vegetation (no scale)
Productive forestry activities (% of parcels)	Increase of agricultural activity (% of ha)	Spatial autocorrelation of changes in vegetation (includes primary and secondary vegetation) (no scale)
Livestock activities (% of parcels)	Increase of primary vegetation (% ha)	
	Increase of secondary vegetation (% ha)	
	Reduction of primary vegetation (% ha)	
	Reduction of secondary vegetation (% ha)	
	Increase in urbanized area (% ha)	

Interactive Map Development: Processed data was uploaded to Github and then processed again using R and its Leaflet library to create thematic map layers of the aforementioned data groups. Afterwards, the Shiny library was used to deploy the maps in an interactive map in shinyapps.io (Shiny, 2020). Initially, data was read directly from Github through shinyapp.io; this made the process slow and surpassed the shinyapp.io server free account limits. However, this was solved using the .Rdata file because it allowed loading “pre-processed” data. Data and code are stored in Github⁴, and an experimental version of the map is available on a free Shiny Apps⁵ account.

⁴ https://github.com/iskarwaluyo/mapa_agricultura_masaforestal/

⁵ https://iskarwaluyo.shinyapps.io/mapa_agricultura_masaforestal_comillas/

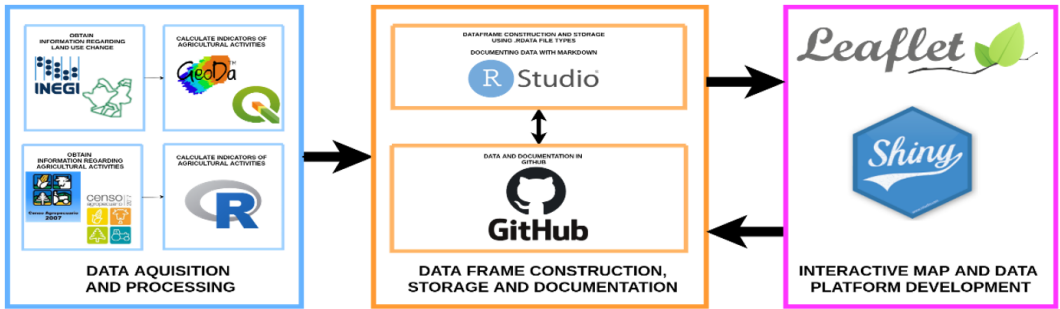


Figure 1: Project development scheme

5 Preliminary Results

The following figures show overlays of layers of autocorrelation of vegetation loss and percentage of plots with agricultural, livestock and forest activities. Although rudimentary, these types of visualisations allow quick identification of combinations of clusters of vegetation loss and main agricultural activities, theoretically providing quick contextual information for decision takers and policymakers.

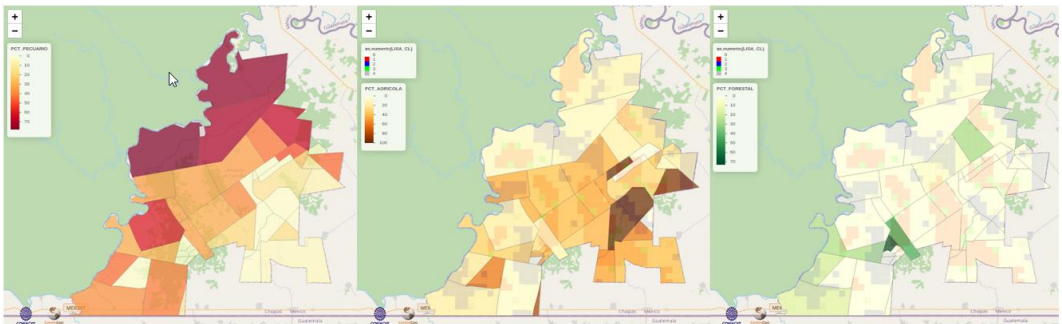


Figure 2: Overlays with autocorrelation of vegetation loss from left to right a) livestock activity, b) agricultural activity and c) productive forestry activity

A visual analysis of the layers may illustrate overlaps between data may better explain land-use changes in the region. For example, it was found that AC 07116010-1007 occupies 46% of plots with forestry activities. However, it also had a positive autocorrelation of total vegetation loss which seems contradictory; a closer look allows noticing that the net vegetation loss is negative due to secondary vegetation loss, but primary vegetation is increasing. A hypothesis that arises from this pattern is that overlaps of clusters with secondary vegetation loss and primary vegetation gain may indicate reforestation or forest stand maturity.

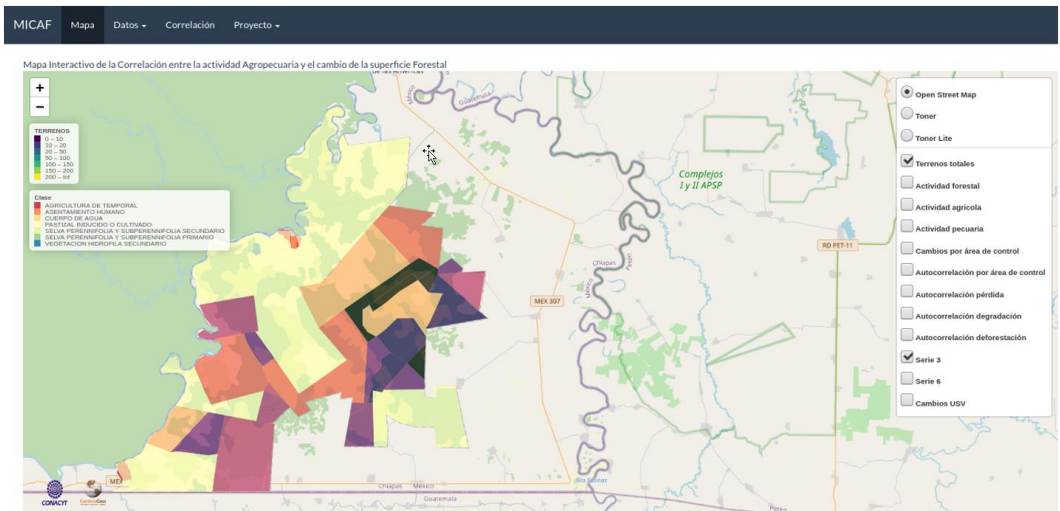


Figure 3: Screenshot of interactive map deployed in shinyapps.io

Another quick pattern that stands out is that most livestock production activity occurs near the northern natural border of the municipality, a river that borders the conservation area. Thus agroforestry programs such as the ones being promoted currently in Mexico could be beneficial in these specific areas of the municipality.

6 Conclusions

After reviewing a number of platforms, we sustain that although there is a large amount of data available, there is a need to develop lighter, less robust platforms designed for end-users. Current platforms are excellent data repositories but require a lot of data manipulation by specialists to exploit the data. Given the urgency many institutions, organisations, leaders and decision-makers have to decide sometimes; we believe that current platforms do not allow easy access to the data they contain. The prototype we present is practical and allows relatively quick overlays and data queries. Although our results are not detailed, they allow an overview that we consider helpful for end-users who often need to make general data consultations and comparisons. Additionally, the cost and time of development are relatively low, which is advantageous for areas where the cost of data analysis and software development are limiting factors for progressing towards the UNSDGs. Nonetheless, our project is only a prototype and its scalability, and real-world use is unknown at this point. We recognize that further research regarding the needs of end-users and computational costs of expanding our work is needed in order to determine the potential of using open-source data and software to develop quicker, lighter and cheaper speciality platforms such as the one we present in this paper. Additionally, it is important that our results be validated; although we are confident that the data used is reliable, further developments could include validating our results with other processes such as NDVI.

References

- Anselin, L. (2004). GeoDa 0.95i Release Notes. Spatial Analysis Laboratory (SAL). Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign, IL. http://geoda-center.github.io/workbook/6a_local_auto/lab6a.html
- ArcGIS. (2020). Comprender El Análisis De Superposición—Ayuda | Arcgis for Desktop. Retrieved June 30, 2020, from <https://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-analyst-toolbox/understanding-overlay-analysis.htm>.
- Arturo, P. (2020, June 13). Sembrando Vida, la reforestación más grande del mundo. Retrieved January 30, 2021, from <https://www.excelsior.com.mx/nacional/sembrando-vida-la-reforestacion-mas-grande-del-mundo/1387883>
- Benhin, J., (2006). "Agriculture and Deforestation in the Tropics: A Critical Theoretical and Empirical Review," *AMBIO: A Journal of the Human Environment* 35(1), 9-16. <https://doi.org/10.1579/0044-7447-35.1.9>
- Quadri de la Torre, G. (2020). Sembrando Vida, la otra catástrofe. Retrieved January 30, 2021, from <https://www.eleconomista.com.mx/opinion/Sembrando-Vida-la-otra-catastrofe--20200820-0130.html>
- Enciso, A. L. (2020, October 12). Sembrando Vida no rescata bosques: ONG. Retrieved January 30, 2021, from <https://www.jornada.com.mx/ultimas/sociedad/2020/10/12/sembrando-vida-no-rescata-bosques-ong-3923.html>
- Carabias, J., de la Maza, J., Cadena, R. (2015). Conservación y desarrollo sustentable en la Selva Lacandona 25 años de Actividades y experiencias. *Natura y Ecosistemas Mexicanos*, A.C.
- Cheng, G. (2020). Leaflet. (Version 2.0.4.1) [Software library for R]. <https://rstudio.github.io/leaflet/>
- Chang, W. (2020). Shiny. (Version 1.6.0) [Software library for R]. <https://cran.r-project.org/web/packages/shiny/index.html>
- CONEVAL (Consejo Nacional de Evaluación de la Política de Desarrollo Social). (2020). Evaluación de Diseño con Trabajo de Campo del Programa Sembrando Vida 2019-2020. Mexico City.
- Cotler, H., Manson, R., Nava, J.D. (2020). Evaluación de la focalización del Programa Sembrando Vida. Retrieved January 30, 2021, from https://www.centrogeo.org.mx/cgeo_archivo/200518_evaluacion-de-la-focalizacion-del-programa-sembrando-vida.pdf
- Dang DKD, Patterson AC, Carrasco LR (2019). An analysis of the spatial association between deforestation and agricultural field sizes in the tropics and subtropics. *PLoS ONE* 14(1): e0209918. <https://doi.org/10.1371/journal.pone.0209918>
- FAO and UNEP. 2020. The State of the World's Forests 2020. Forests, biodiversity and people. Rome. <https://doi.org/10.4060/ca8642en>
- Flores, M.L. (2019). Los alcances en la producción agrícola chiapaneca. Una reflexión sobre la soberanía alimentaria en la región. *región y sociedad*, 31, e1177. doi: 10.22198/rys2019/31/1177
- Forbes Staff. (2019, July 11). Talan árboles en Veracruz para entrar al programa 'Sembrando Vida' • Actualidad • Forbes México. Retrieved January 30, 2021, from <https://www.forbes.com.mx/pobladores-talan-arboles-para-entrar-al-programa-sembrando-vida/>
- Gentleman, R., Ihaka, R. (2019). R. (Version 3.6.1) [Computer software]. <https://www.r-project.org/>
- INEGI. (2005). Uso del suelo y vegetación, escala 1:250000, serie III (continuo nacional). Retrieved June 30, 2020, from http://www.conabio.gob.mx/informacion/gis/?vns=gis_root/usv/inegi/usv250ks3gw
- INEGI. (2007). Censo Agrícola, Ganadero y Forestal 2007. Retrieved June 30, 2020, from <https://www.inegi.org.mx/programas/cagf/2007/#Documentacion>
- INEGI. (2016). Actualización del Marco Censal Agropecuario 2016. Retrieved June 30, 2020, from <https://www.inegi.org.mx/programas/amca/2016/>

- INEGI. (2016). Uso del suelo y vegetación, escala 1:250000, serie VI (continuo nacional). Retrieved June 30, 2020, from http://www.conabio.gob.mx/informacion/gis/?vns=gis_root/usv/inegi/usv250ks3gw
- Olvera, D. (2020, March 06). No se debió sembrar en primer año del programa Sembrando Vida, se saltaron la planeación: Expertos. Retrieved January 30, 2021, from <https://www.economiahoy.mx/nacional-eAm-mx/noticias/10400116/03/20/No-se-debio-sembrar-en-primer-ano-del-programa-Sembrando-Vida-se-saltaron-la-planeacion-expertos.html>
- Pontius, R. G. Jr., E. Shusas and M. McEachern (2004), “Detecting important categorical land changes while accounting for persistence”, *Agriculture, Ecosystems and Environment*, no. 101, pp. 251–268.
- Sherman, G. (2020). Quantum GIS. (Version 3.16) [Computer software]. <https://www.qgis.org/>