

Territorial relations between deforestation and Rural Environmental Registry (CAR) in the Amazon biome using free software QGIS/PostgreSQL/PostGIS and Data Warehouse Structure

Mayra Abboudi Brasco¹, Carlos Alberto de Carvalho¹

¹Embrapa Territorial

Avenida Soldado Passarinho, 303, CEP 13070-115 - Campinas - SP

mayra.brasco@colaborador.embrapa.br, carlos-alberto.carvalho@embrapa.br

Abstract.

This work presents the use and integration of free software to organize a spatial database (part of a data warehouse) capable of performing spatial and numerical analyses on a set of spatial data from the Amazon biome (deforestation and Rural Environmental Registry - CAR). This work is part of a larger research project (Embrapa/Censipam) that addresses the use of strategic territorial intelligence to analyze deforestation processes in the Amazon biome. More specifically, this article deals with a part of the Data Warehouse construction that serves as the basis for the larger research project to work. The focus is on the spatial part of the data warehouse, proposing a free software structure to perform spatial analysis. As preliminary results, quantifications of deforestation by state and by municipalities are presented.

Keywords: Amazon biome; Rural Environmental Registry; Data Warehouse; Quantum GIS; deforestation.

Resumo.

Este trabalho apresenta o uso e integração de software livre para organizar um banco de dados espacial (parte de um data warehouse) capaz de realizar análises espaciais e numéricas sobre um conjunto de dados espaciais do bioma Amazônia (desmatamentos e o Cadastro Ambiental Rural - CAR). Este trabalho faz parte de um projeto de pesquisa maior (Embrapa/Censipam) que aborda o uso de inteligência territorial estratégica para analisar os processos de desmatamento no bioma Amazônia. Mais especificamente, este artigo lida com uma parte da construção do Data Warehouse que serve de base para o funcionamento do projeto de pesquisa maior. O enfoque está na parte espacial do data warehouse, propondo uma estrutura de software livre para realizar as análises espaciais. Como resultados preliminares, são apresentadas quantificações de desmatamentos por estado e por municípios.

Palavras-chave: Amazônia; Cadastro Ambiental Rural; Data Warehouse; Quantum GIS; desmatamento.

1. Introduction

This article is one of the results of the project “Strategic territorial intelligence applied to the identification, quantification and qualification of deforestation processes and support and guidance of public and private policies in the Amazon biome”, a partnership between Embrapa Territorial and Centro Gestor e Operacional do Sistema de Proteção da Amazônia (Censipam). It seeks to broaden the analysis of deforestation processes in the Amazon biome through a detailed study of territorial and temporal bases, of deforestation patterns, such as the number of deforested polygons and their spatial and temporal evolution, concentration by state, microregion, legal attribution, proximity to

urban centers, logistics infrastructure, etc., and their territorial distribution according to the legal allocation of land.

The database for the project was built using the Kimball method and the scripts generated using ETL (Extract, Transform and Load) methods. The creation and structuring of the Data Warehouse were addressed in a previous work (DEUS; ALMEIDA; CARVALHO, 2021). In this paper, we will continue the analysis that supports the Data Warehouse created.

The present work consists of describing both the operational process of accessing the database created using the Data Warehouse concept, built on PostgreSQL/PostGIS and Quantum GIS, all free softwares, and the analysis of deforestation scenario between 2018 and 2020 in the Rural Environmental Registry (CAR in the Portuguese acronym) (BRASIL, 2012). However, in this analysis the legal reserve areas within the CAR's polygons and the permanent protection areas around the rivers are not considered.

2. Free software: Quantum GIS (QGIS) and PostgreSQL/PostGIS

The use of free software and the movements for their democratization began a long time ago. In 1983, Richard Stallman created a movement in opposition to the Unix world, which created Linux operating systems, called GNU¹.

QGIS is open source licensed under the GNU General Public License and is a project of the Open Source Geospatial Foundation (OSGeo). Given this important feature, QGIS is constantly being updated, receiving contributions from communities through bug reports, plugins (or add-ons) that are created by software scholars and made available for free through the system's repositories.

PostgreSQL is an open-source, widely used relational database. For compatibility with other systems, the most stable version, 13.5, was used. PostGIS is the spatial extension to the PostgreSQL relational database for working with geospatial fields and functions. PostGIS - Spatial and Geographic Objects for SQL is a spatial extension for the PostgreSQL free database, the result of the same project as QGIS, OSGeo. It can be installed on any operating system, as well as QGIS, and there is an interface and direct connection between these software (Figure 1).

Free software from the OSGeo project, QGIS and PostGIS are tools that have a strong connection and are in constant evolution and adaptability. However, there is still resistance regarding their use for several reasons.

¹ At the time, Unix did not give freedom to its users. Currently, we see many versions of Linux that are more restrictive, with better interfaces and also those that allow users to access their source code at any time and make changes.

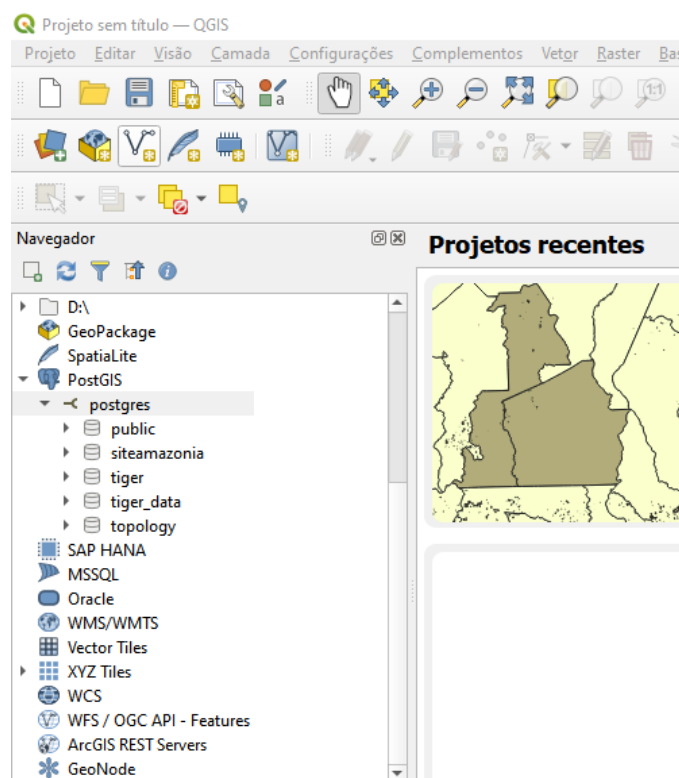


Figure 1. Screenshot of QGIS showing its direct connection with PostGIS.

Given the context of the project, we developed a method using these technologies. It gives anyone who wants to use the project's data the opportunity to analyze deforestation using free software. Therefore, with the continuity and the dissemination of free software methods, we aim to identify and quantify the numbers and dimension of deforestation that occurs annually in the Amazon biome, its states and municipalities, with specific analysis in areas registered in CAR.

3. Material and methods

This research is based on the Amazon biome and a GIS will be structured with it. This system will be the integration of software, tools and techniques with collections of numerical, iconographic and cartographic data, in GIS, supported by spatial databases.

The equipment used in this project have high storage capacity and performance: 5 HP Intel Core i7 3.4GHz Stations with 16GB of RAM, 500GB SSD HD and 2TB HD; 1 SY480Gen10 Blade Server with 10-core Intel Xeon Silver processors, 192 GB of RAM and 2 600GB HDDs and a large-capacity data storage system Storage HP MSA-2050 with a capacity of 192TB.

The software used in the integration and interconnection with the Data Warehouse structure of the project were Quantum GIS 3.24.1 Tisler, PostgreSQL 13.5 and PostGIS 3.1.

The data warehouse was structured using PostgreSQL and made available on an exclusive workstation with more computational resources and intended for the approval of the applications developed in the project. The hardware used has the following settings:

- Manufacturer: Lenovo;
- Brand: Thinkstation;
- Model: P520;
- Intel Xeon W-2255 Processor;
- Memory 64GB DDR4 ECC 2666MHz (4x 16GB);
- NVIDIA Quadro P2000 graphics card;
- 2TB 7200 RPM HD;
- SSD 512GB;
- Windows 10 Pro for Workstations 64 Operating System.

Windows 10 operating system contains the Windows Subsystem for Linux or WSL (Windows Subsystem for Linux) application [15], which “allows developers to run a GNU/Linux environment, including most command-line tools, utilities and applications, directly in Windows, without modification and without the overhead of a traditional virtual machine or dualboot installation”. Thus, we were able to migrate the Data Warehouse using the same Linux operating system where it was developed.

The limits of municipalities and states were obtained from the Brazilian Institute of Geography and Statistics’ (IBGE) territorial divisions as a shapefile (IBGE, 2014). The limits of the Amazon biome are the official ones of the Ministry of the Environment, cut from the shapefile of Brazilian biomes. The geocoded boundaries of Conservation Units, Indigenous Lands, Military Areas, Agrarian Settlements and *Quilombola* Communities were obtained from the National Spatial Data Infrastructure (INDE, 2022), digitally cut and structured using projection and nomenclature adjustments in the SIRGAS 2000 reference system (IBGE, 2010), and the projection adopted by IBGE (IBGE, 2014) for area calculations.

Deforestation by clear cutting in the Amazon biome has been monitored since 1988 by the Project for Monitoring Deforestation in the Legal Amazon by Satellite (PRODES), of the National Institute for Space Research (INPE) (INPE, 2022). Annual rates are estimated based on the deforestation increments identified in LANDSAT satellite images (20 to 30 meters of spatial resolution and revisited every 16 days). The minimum area mapped by PRODES is 6.25 hectares.

The CAR database used was the SICAR 2021 version. This database underwent a process of selection of polygons, and those contained in areas of integral conservation units, indigenous lands and military areas were removed. Then, the polygons of SICAR 2021 were intersected with data on deforestation between 2018 and 2020.

3.1. Database connection

As described at the beginning of the report, QGIS and PostgreSQL/PostGIS have an easy integration, because they are from the same developers. Therefore, their connection is easy and does not require any specific configuration.

To import a layer in QGIS, after fixing geometries and defining projection, it is necessary to click on the upper tab “Database” and import a layer using the “DB Manager”, then check the “primary key”, “geometry column”, “source SRID”, “target SRID” and “create spatial indexes” options. When importing it is necessary to define projection. For this database created using the Data Warehouse, we registered the

IBGE_South_Albers_Equal_Area_Conic projection with the SRID code 1054, which is the IBGE's standard projection for calculating areas. Thus, when importing the layer, it is necessary to reaffirm to the database which projection is adopted using a command in the SQL language (Figure 2):

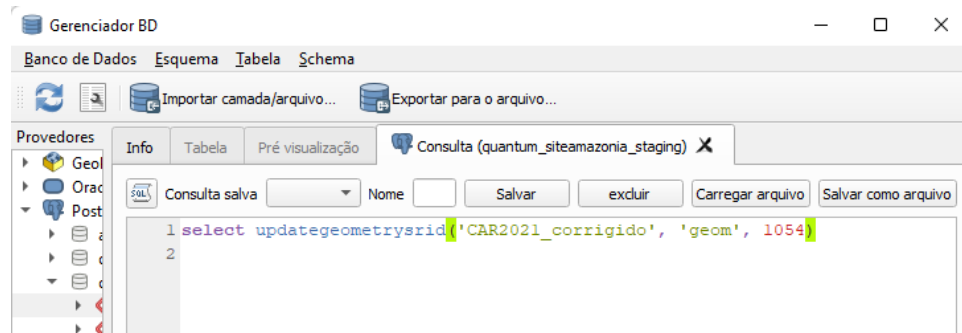


Figure 2. SQL command window in QGIS DB Manager.

Once the database is connected, it is possible to work on the layers from any computer that has access to QGIS, PostGIS and connection to the Data Warehouse built. This tool made it possible for team members of the project to connect and work remotely.

For the analyses depicted in this article we used QGIS tools to work on the data, more precisely the "dissolve" and "clip" tools. Then, using the "basic statistics" tool, we calculated the areas and quantities of polygons from the intersection of CAR 2021 data and deforestation polygons between 2018 and 2020. These were transferred to Microsoft Excel, where we made the tables and graphs presented below.

4. Analysis

Crossing the deforestation areas within CAR polygons, their relationship with the total deforested areas and the numbers of deforestation polygons that coincide with the polygons of SICAR 2021, we observed a direct relation between the CAR polygons and deforestation (Figure 3).

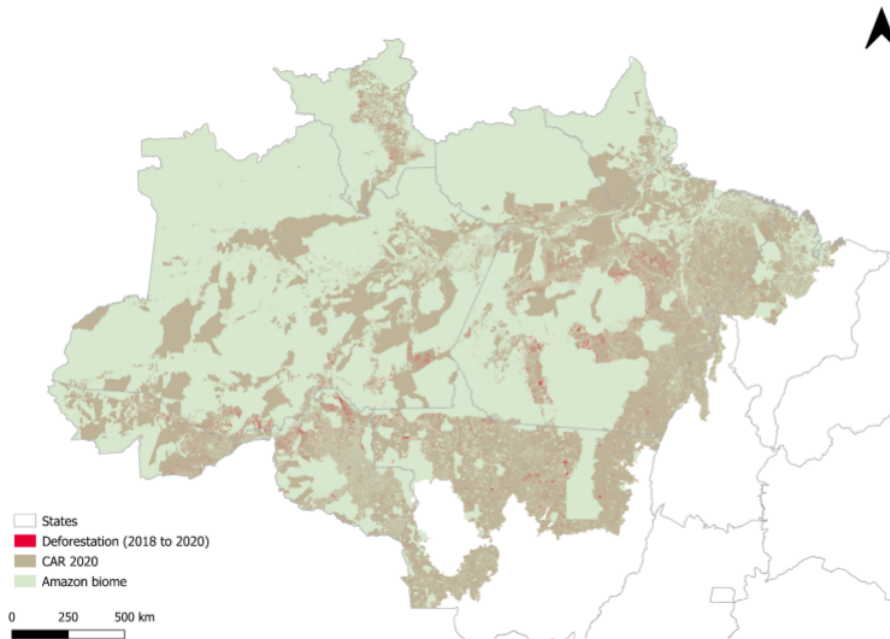


Figure 3. Amazon biome with CAR 2021 and deforestation (2018 to 2020)

Table 1 shows the general results: for each year, the total area deforested in hectares and the number of deforestation polygons, as well as the total area occupied by CAR 2021 and its number of polygons and, finally, the deforested area that coincides with the area of CAR, the percentage of deforestation in these areas that corresponds to the total area deforested in the year and the number of corresponding polygons.

Table 1. General results of the analysis of deforestation and SICAR 2021 (between 2018 and 2020)

Year	total area deforested in hectares	number of deforestation polygons	total area occupied by CAR 2021	number of polygons of CAR 2021	deforested area that coincides with the area of CAR 2021	percentage of how much deforestation of these areas occupied by CAR corresponds to the total area deforested	number of corresponding polygons that coincides with the area of CAR 2021
2020	1.031.851	41.129	156.886.040	131.616	691.048	66,97%	6.646
2019	1.066.581	41.526	156.886.040	131.616	727.541	68,21%	6.750
2018	703.710	32.661	156.886.040	131.616	481.334	68,40%	6.011

We noticed that there was an increase of more than 300 thousand hectares deforested from 2018 to 2019, but a softening of the advance of deforestation between

2019 and 2020. However, there are still more than 1 million hectares deforested in each of these years.

When correlating the deforested areas with the areas occupied by rural properties, we also see that the amount of deforestation decreased within these areas in the period: in 2018, 68.4% of deforestation occurred within CAR areas; in 2019, 68.2%; and in 2020, 66.97%. Therefore, we see that most deforestation is carried out in these areas, but the increase in deforestation is inside and outside the CAR areas, as the increase is not proportional.

When we started to analyze the states, we noticed discrepancies between the deforestation numbers. In Figure 4, we see that some states have many deforested hectares, while others do not appear on the graph.

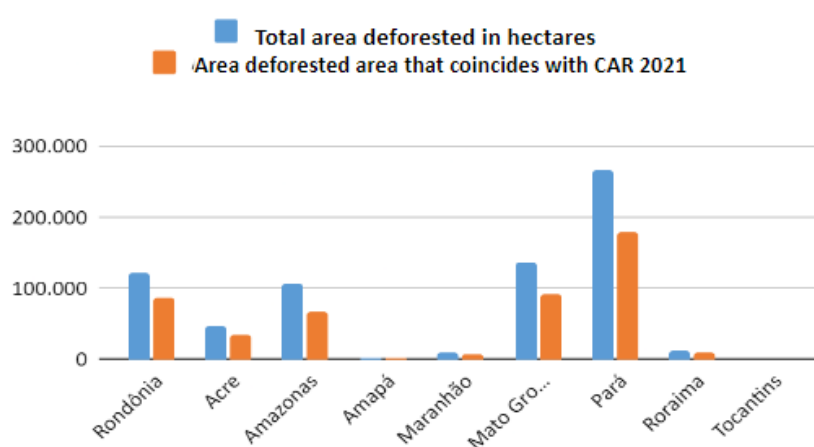


Figure 4. Graph - Deforestation that occurred in 2018 and its incidence within the 2021 CAR by states

The state of Pará had more of its area deforested in 2018, followed by Mato Grosso, Rondônia and Amazonas. Tocantins had less than a thousand hectares deforested in the year and does not appear in the graph. Most of the deforestation in the states occur within CAR areas. However, in some states, we have noticed that almost all deforestation occurs in CAR, such as Maranhão and Roraima.

When we analyze the 2019 data, presented in Figure 5, we see that Tocantins remains unimpressive, while Pará, Mato Grosso, Amazonas and Rondônia continue to be the states with the highest deforestation rates. But this time, Amazonas had a more deforested area than Rondônia. Most of the deforested areas are inside SICAR 2021, but we are beginning to notice a discrepancy in deforestation outside these areas in Pará.

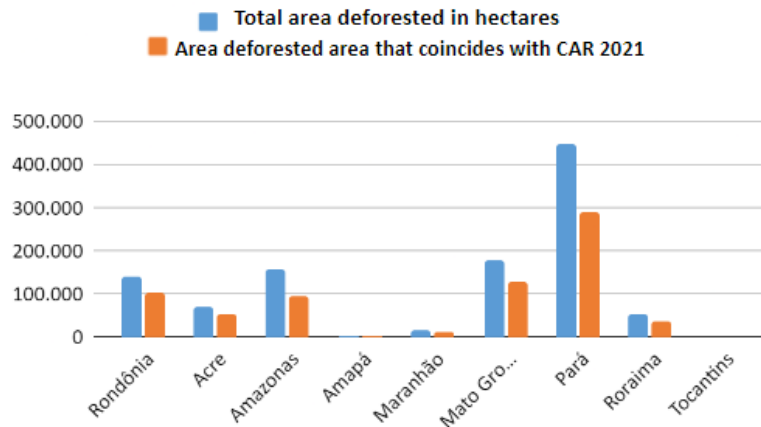


Figure 5. Graph - Percentage of deforestation carried out within the 2021 CAR areas in relation to the total deforestation carried out in the states in 2019

In 2020 (Figure 6), we notice few differences in these numbers. However, we see that the pattern of deforested areas remains the same, except for Maranhão, which starts to appear more in the graph for having a few thousand hectares more than in the other two years.

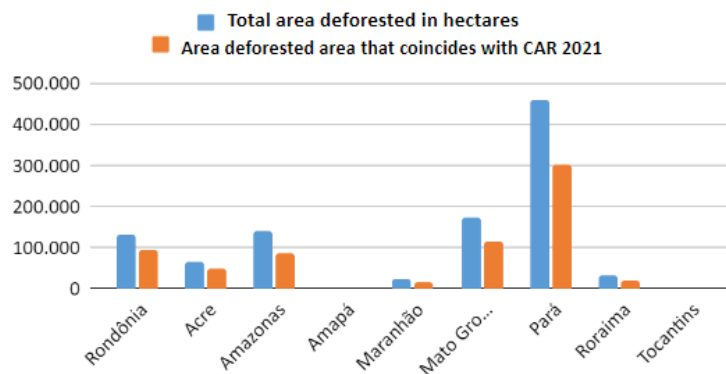


Figure 6. Graph - Percentage of deforestation carried out within the SICAR 2021 areas in relation to the total deforestation carried out in the states in 2020

When we start to analyze the municipalities, some other dimensions stand out. In Figure 7, there is a graph showing ten municipalities that deforested the most in 2018, but the percentages of deforestation that occurred within CAR areas are shown. Therefore, the calculation of the ratio between deforestation within SICAR 2021 and total deforestation produced the comparison in percentage.

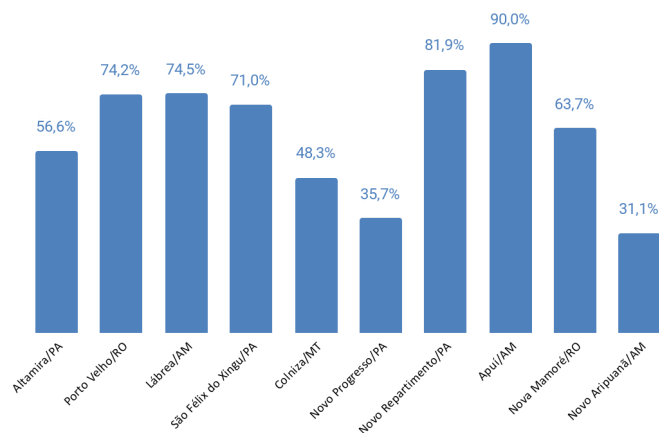


Figure 7. Graph - Ten municipalities with the highest deforestation rates in 2018 and how much of it (percentage) occurred in SICAR 2021 areas

Leaving the general analysis (states) and entering the specific analysis (municipalities), it stands out that not all municipalities have most of their deforestation taking place within the CAR polygons. By studying the municipalities that deforest the most, case by case, we may think about local public policies.

We see that most of the deforestation occur within CAR areas, with the exception of Colniza/MT, Novo Progresso/PA and Novo Aripuanã/AM occur in areas not registered in the Brazilian Rural Environmental Registry. A first analysis indicates the need for a focus on the inspection of areas outside the Registry, but when we start to analyze the municipalities that are deforested the most, we notice some other dimensions in this new scale.

In Figure 8, we see that the municipalities that stand out the most remain almost the same, with the exception of Pacajá and Portel in Pará. And most of the deforestation occur inside CAR areas with the exception of Altamira and Novo Aripuanã.

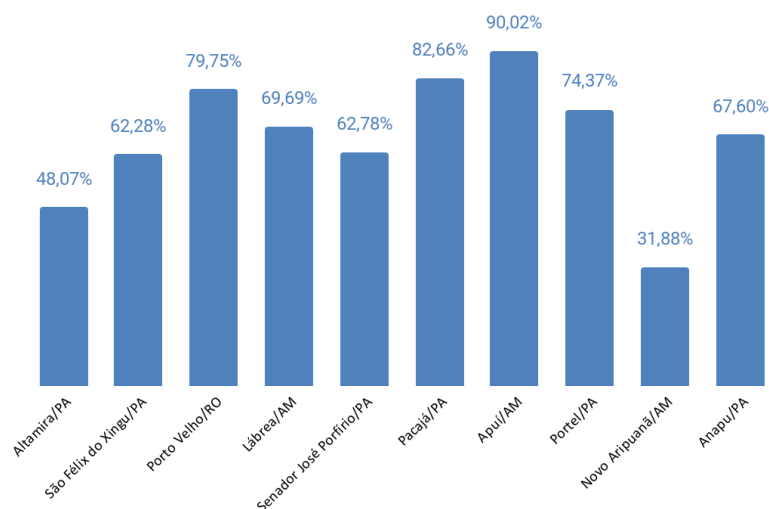


Figure 8. Graph - Ten municipalities with the highest rates (in 2019) of deforestation and how much of it occurred in SICAR 2021 areas

When looking at the 2020 numbers (Figure 9), we notice an emergence of municipalities that were not yet on the list, such as: Novo Progresso and Itaituba, both in Pará. Most deforestation took place in CAR areas, with the exception of Altamira and Itaituba. We perceive the need for greater control and inspection in these municipalities, in order to identify where the deforested areas are and make appropriate decisions to avoid illegal and unauthorized deforestation.

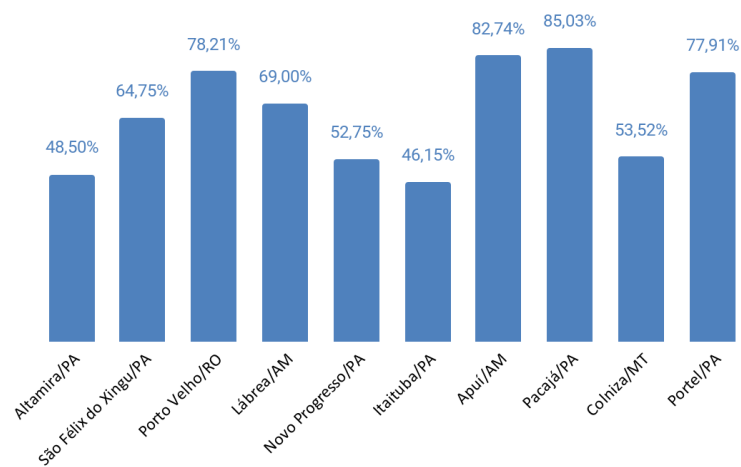


Figure 9. Graph - Ten municipalities with the highest rates (in 2020) of deforestation and how much of these occurred in SICAR 2021 areas

When spatializing the municipalities and deforestation polygons in SICAR 2021 (Figure 10), we see that they remain in the same range above the “arc of deforestation”. We also noticed an increase in polygons that coincide with the SICAR 2021 areas.

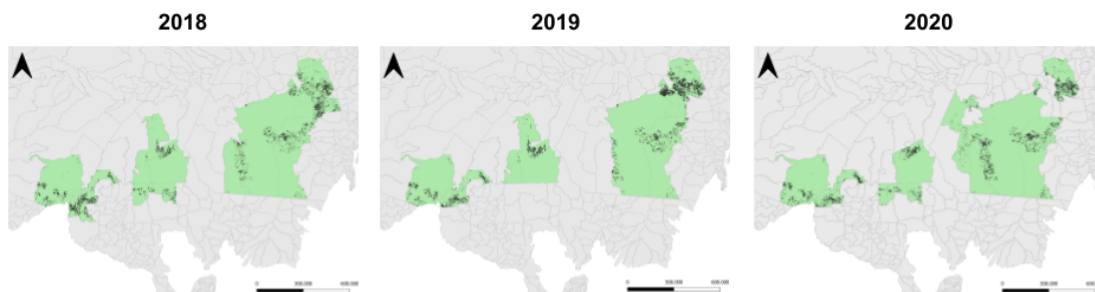


Figure 10. Map - The ten most deforested municipalities in the Amazon biome in 2018, 2019 and 2020 and deforestation polygons that coincide with the SICAR 2021 areas

This data was generated in this short period of research and approach to the theme, and there is a need to deepen concepts, analyses and other tools to deal with this complexity.

6. Conclusions

The use of free software in the project was an efficient and low cost solution. The possibility of creating a Data Warehouse according to the needs of the project and the generated data, its connection to Quantum GIS using the PostGIS extension and the great interoperability and ease of connection and processing of data offers what is necessary for the project to continue. With this structure in place, we were able to analyze deforestation in CAR areas over the temporal and spatial clipping described.

Approaching this research between deforestation and CAR using a territorial cutout detailed by municipalities opens a possibility of discussing public policies aimed directly at city managers, because it is where deforestation events actually occur. It also makes it possible to highlight cuts at the state level, enabling joint actions by different levels of public administration. By highlighting municipalities that concentrate deforestation and qualifying them in terms of their distribution in the rural world as portrayed by CAR, action may be taken to combat illegal deforestation.

In addition to the results, performing data processing and analysis using free software is also important, because when we work with information of public use, it is essential to disseminate results and methods that may be used and replicated, in order to foster the evolution of scientific knowledge and the elaboration of public policies by public institutions or companies.

References

BRASIL (2012). Lei n. 12.651, de 25 de maio de 2012. Dispõe sobre a proteção da vegetação nativa; altera as Leis n°s 6.938, de 31 de agosto de 1981, 9.393, de 19 de dezembro de 1996, e 11.428, de 22 de dezembro de 2006; revoga as Leis n°s 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a Medida Provisória n° 2.166-67, de 24 de agosto de 2001; e dá outras providências.
http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112651.htm, August 10th, 2022.

DE DEUS, R.; ALMEIDA, M.; CARVALHO, C. (2021) Estruturação de um Data Warehouse para repartição territorial de desmatamentos no bioma Amazônia e o Cadastro Ambiental Rural. *In*: CONGRESSO BRASILEIRO DE

AGROINFORMÁTICA (SBIAGRO), 13. , 2021, Evento Online. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021 . p. 116-125.

GNU (2021). What is GNU?. <https://www.gnu.org/>, August 10th, 2022.

IBGE (2014). “Estrutura Territorial”.
<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial.html>, August.

IBGE (2010). “Projeto SIRGAS: Sistema de Referência Geocêntrico para as Américas”.
<https://www.ibge.gov.br/geociencias/informacoes-sobre-posicionamento-geodesico/sirgas/16691-projeto-mudanca-do-referencial-geodesico-pmrg.html?=&t=o-que-e>, August.

IBGE. Cadastro Nacional de Endereços para Fins Estatísticos – CNEFE. Rio de Janeiro: IBGE, 2020. Disponível em
<https://censoagro2017.ibge.gov.br/resultados-censo-agro-2017/cnefe.html>, August 10th, 2022.

INCRA (2022). “Instituto Nacional de Colonização e Reforma Agrária”.
<http://www.incra.gov.br/pt/>, August 10th, 2022.

INDE (2022). “Infraestrutura Nacional de Dados Espaciais”. <https://inde.gov.br/>, August.

INPE (2022). “Projeto PRODES: Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite”.
<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>, August 10th, 2022.