A QGIS plugin for BONDS project: integrating field data with geographical, remote sensing and health information

Rogerio Flores Júnior[1,2], **Vitor Conrado Faria Gomes**[1,3],
Karine Reis Ferreira[1], **Gilberto Ribeiro de Queiroz**[1], **Marie-Paule Bonnet**[2]

[1]Image Processing Division, National Institute for Space Research (INPE)
São José dos Campos – SP – Brazil

[2]Environment and Resources Department – Institut de recherche pour le développement (IRD)
Montpellier – Occitanie – France

[3]C4ISR Division – Institute for Advanced Studies (IEAv)
São José dos Campos – SP – Brazil

`rogerio.junior@inpe.br, vitor@ieav.cta.br`

`gilberto.queiroz@inpe.br, karine.ferreira@inpe.br`

Abstract. *Open source Geographic Information System (GIS) have been fostering spatial data research such as Earth observation and environmental monitoring for more than 30 years. More recently, globally available geospatial information combined with web technologies are providing new environments and tools for data handling. Thus, binding the mapping and processing capabilities of traditional GIS to the accessibility and reliability of web-based data providers can bring new opportunities for research. In this paper, we built a QGIS plugin to explore the integration of different public data providers in Brazil along with field data produced by the BONDS project. The biOdiversity conservatioN with Development in Amazon wetlandS project (BONDS) proposes to develop biodiversity scenarios for the Amazonian floodplains aiming to support solutions to preserve biodiversity and ecosystem services. The use of web services enabled dynamic and fast access to several products ranging from remote sensing images, land use and land cover, territorial cartography, water quality, to COVID-19 health data, and more.*

1. Introduction

Within the technological revolution of the last decades, the advancement of the internet and mobile communication changed how information is produced, stored, and consumed [Tole et al. 2013]. With these new technologies and the possibility of a fast and reliable Global Navigation Satellite System (GNSS), data is now produced on the fly and mostly with Geographical Information (GI) attached to it [Maguire and Longley 2005]. Although this massive amount of data can be portrayed as big data, its geographical features bring new paradigms [Lee and Kang 2015]. Nowadays, GI data has become a paramount framework to sample and represent phenomenons and patterns of the real world in its digital twin. In this matter, Geographical Information Systems (GIS) comes into the picture to meet the necessity to understand, represent and forecast data by providing a geographic ecosystem with tools and techniques to visualize, analyze and model the GI data [Raju 2006].

In the internet era, all areas of knowledge including GIS have been influenced to a large extent by internet standards and technologies, resulting in the increasing availability of online applications using distributed web services. Ready-to-use services are shaping how we interact with data by providing standards and tools to discover, access and acquire data in an organized, fast and reliable way. [Lemmens 2006]. There is apparent migration of traditional GIS tools to web based solutions, however, these approaches still suffer with paradigms such as the need for uninterrupted and reliable internet, poor mapping and visualization features, lack of advanced processing tools, and lack of computational power for everyone. Therefore, the use of desktop or client-side GIS tools are very likely to continue to be needed [Michaelis and Ames 2012].

In spite of its benefits, traditional desktop based GIS are burdened with data duplicity, outdated data sources, manual data input, and the inability to easily share data. Therefore, combining client-side GIS with web based data tools may enhance the user's experience. The use of well established practices such as the Open Geospatial Consortium (OGC) standards could be useful for promoting better data access, normalization and shareability. These new services can be very helpful to the scientific community as research projects become more globalized. An example is the *Balancing biOdiversity conservatioN with Development in Amazon wetlandS* (BONDS) an international research project comprising seven countries and funded by the joint call of the Belmont Forum and BiodivERsA, within the scope of the BiodivScen ERA-Net COFUND program. This project proposes to develop biodiversity scenarios for the Amazonian floodplains, combining conservation and development issues. Thus, it aims to find, along with stakeholders, solutions to preserve biodiversity and ecosystem services in the context of socioeconomics and climate change.

BONDS research groups are actively generating Earth Observation (EO) derived products based on multi-disciplinary GI data which, despite being open and freely available, are often decentralized and hard to reach. In addition, there is no single place of management and distribution for data produced by the project. With the project's research areas diversity, some collaborators may not have the access or computational ability to search and effectively consume the data in the present conditions. Therefore, this work aims to explore data integration methods and evaluate the benefits and drawbacks of current technologies. This paper presents a QGIS plugin to provide access to data from both BONDS project and Brazil's public and freely available datasets. The plugin was designed to abstract the computational technologies to aid BONDS researchers and external users to reach as much relevant data as possible.

2. Materials and methods

2.1. BONDS Study area

The study areas for the BONDS project are floodplain areas along the Brazilian Amazon rain forest. Those areas encompass the high, medium and low Amazon river regions, and are located in both Amazonas and Pará states. The Curuai lake floodplains are the most important BONDS site and also the focus of this paper (Figure 1). The Lake is located at the lower Amazon river region ($-55.489°W$, $-2.111°S$), next to the cities of Santarém and Óbidos, being around $900km$ from the Amazon river delta. Figure 1 presents a few

---

[1] *http://www.bonds-amazonia.org/*

QGIS layers obtained with the BONDS plugin. These data are the CREN Land Use and Land Cover (LULC) for Pará state (2018), CGEO federal water bodies, BONDS water quality field data, and a true color Landsat-8 image from a 16 day (08/28-09/12/2020) datacube from Brazil Data Cube (BDC) project.

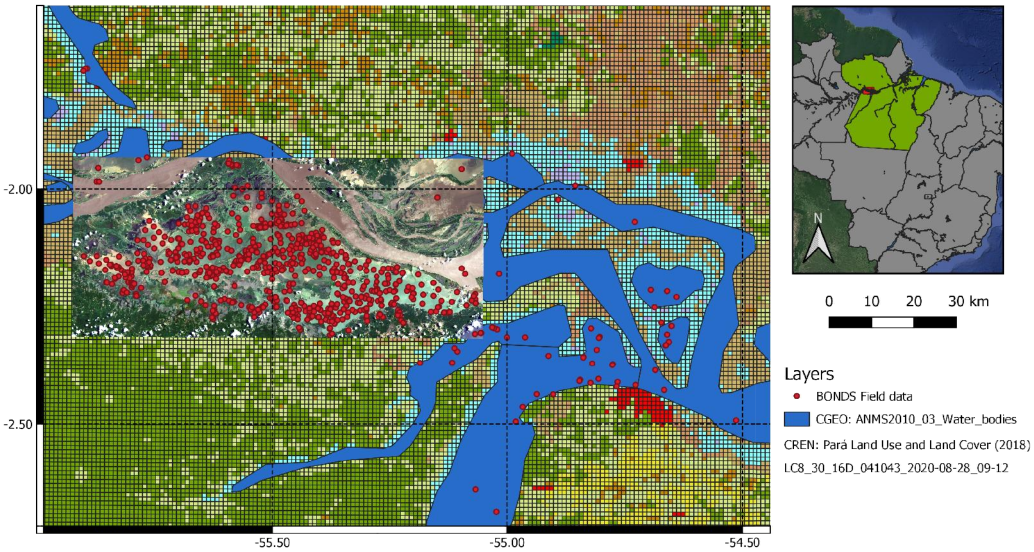


Figure 1. BONDS main study area with layers obtained through BONDS plugin. In the background CREN LULC for Pará state (2018). As vectors the CGEO federal water masses and BONDS water quality field data. On top a true color RGB Landsat-8 image from a 16 day datacube from BDC.

2.2. Brazilian public open data

Although policies regarding the use and access of public data have been discussed since the 1980s, the topic started to grow in Brazil in the early 2010s with initiatives for boosting government's transparency. Brazil's federal open data policy Decree 8.777/2016 established rules and guidelines for public agencies to comply with the Open Government Data plan (OGD) for public data sharing [Silva and Pinheiro 2019]. Surveying public agencies identifying compliance to OGD [Silva and Pinheiro 2019] conclude that although progress is slow, agencies are moving towards open datasets through Application Programming Interface (API) services.

2.2.1. Brazilian National Data Infrastructure (INDE)

The INDE proposes an integrated set of technologies, standards, policies, monitoring mechanisms and procedures to centralize data information of several governmental providers, ranging from universities and public agencies to research facilities. It is composed of 60 institutions at academic, municipal, state or national levels such as National Institute for Space Research (INPE), Amazon Protection System (SIPAM), Brazilian Institute of the Environment and Renewable Natural Resources (IBAMA), National Water

Agency (ANA), and Brazilian Institute of Geography and Statistics (IBGE). INDE's catalog comprises data from territorial cartography, LULC, biodiversity, water related products and many more.

INDE promotes common standards and agreements to facilitate and standardize data generation, storage, access, sharing, and dissemination. Thus, any partner institutions must comply with OGC standards and guidelines to be able to provide geoservices through INDE's infrastructure [INDE 2022]. Geoservices are web services specific to the geospatial domain, constituting a powerful set of functionalities to collect, store, retrieve, transform and display spatial data. These functionalities can be accessed through web browsers or any other application as programming languages and GIS.

The OGC is an international non-profit organization for creating spatial data sharing standards allowing government agencies, NGOs, companies, academic and research entities to make their geospatial data openly available [Baumann 2010, Giuliani et al. 2013]. Today, OGC is composed of more than 500 organizations and has 58 common open source standards [OGC 2022]. All specifications intend to support data interoperability by following a common architecture [Baumann 2010]. The main services at OGC for EO are: a) Web Map Service (WMS); b) Web Feature Service (WFS); c) Web Coverage Service (WCS); d) Web Processing Service (WPS); e) Web Coverage Processing Service (WCPS); and f) Catalog Service for Web (CSW).

The INDE specification is based on the OGC WMS, WFS, and WCS geoservices [CONCAR 2010]. The WMS is a protocol to serve georeferenced maps optimized for visualization, allowing users to view and consume its information but not meant for modeling or processing [Baumann 2010]. The WFS protocol, on the other hand, defines a set of interfaces for accessing geographic information as feature and feature properties. In short, WFS allows users to access a geographic phenomenon represented in vector format [Giuliani et al. 2013]. Also, WFS provides query capabilities that can be based on spatial and non-spatial constraints. Similar to WFS, the WCS provides datasets in matrix or raster format, however, it is used to represent phenomena with continuous spatial-temporal variations called *coverages*.

2.2.2. Remote sensing

In terms of remote sensing images, Brazil is a historical reference of open public data being one of the first nations to freely open its satellite catalog [Santos 2005]. More recently, the BDC project from INPE is producing and distributing data cubes of satellite images. It is a research, development and technological innovation project to produce, store, manage and analyze cubes of medium resolution satellite data [Ferreira et al. 2020]. The BDC image catalog follows the Spatio Temporal Asset Catalogs (STAC) specification providing a standard interface that allows users and applications to query and discover all image collections and data cubes stored in the project databases. To support and abstract the STAC services provided, the BDC team developed packages in python (stac.py) and R (rstack). The project also provides a web portal where users can access, visualize, and download images and data cubes [Ferreira et al. 2020].

---

[2] *https://brazildatacube.dpi.inpe.br/*

The project is also currently developing computational environments and tools with cutting-edge technologies for dealing with EO data. The project provides geoservices using common standards such as Tile Map Service (TMS) by the Open Source Geospatial Foundation (OSGeo) and WMS, WFS, and WCS by the OGC and ISO [Ferreira et al. 2020]. Also, at the assets level (image bands) the raster files are distributed in a Cloud Optimized GeoTIFF (COG), allowing for dynamic access to image subsets.

There are several public providers producing and sharing EO derived products, such as the Modular system for continuous monitoring of inland waters (MAPAQUALI) and Monitoring of Deforestation of the Brazilian Amazon Forest by Satellite (PRODES) projects. MAPAQUALI is an experimental platform for monitoring water quality parameters through satellite images [de Lucia Lobo et al. 2017]. PRODES carries out satellite monitoring of clear-cut deforestation in the Legal Amazon and, since 1988, has produced annual deforestation rates in the region, aiming to establish public policies.

2.2.3. Health COVID-19

The COVID-19 global pandemic was catastrophic for mankind, resulting in millions of lives lost and a global economic recession. Its fast and easy spread worldwide brought to light the need for disease monitoring and control in a globalized society. One of the first data portals to collect and monitor COVID-19 was the Johns Hopkins Coronavirus Resource Center (CRC) [Dong et al. 2020]. It was considered the *"go-to data source"* for COVID-19 by TIME magazine in 2020. The CRC aim is to collect data available on cases, deaths, tests, hospitalizations, and vaccines worldwide.

Brazil has a robust universal healthcare system the Unified Health System (SUS) which comprises several services at the federal, state and municipal levels. SUS provides essential, universal, and free access to health-related services for the whole country. More than 80% of Brazil's population rely on SUS [Stopa et al. 2017]. To handle its massive digital infrastructure the SUS Information Technology Department (DataSUS) was created in 1991. DataSUS developed over 200 systems and is responsible for organizing and providing free access to the data generated by SUS. At the OpendataSUS portal several public health datasets are available [BRASIL 2022b, BRASIL 2022a]. This data is being organized following the *Comprehensive Knowledge Archive Network* (CKAN) standards allowing users to send HTTP requisitions to retrieve the data. However, not every dataset in the portal is already ready for the CKAN API search. Currently, at the portal there are two datasets through API search, the COVID-19 vaccination status and flu syndrome of suspected COVID-19 cases, other datasets are available in Comma Separated Values (CSV) for direct download.

2.2.4. BONDS field data

The field data used in this work is composed by Instrumentation Laboratory for Aquatic Systems (LabISA)[3] Amazon datasets for water quality parameters at BONDS scope. This data is in a PostgreSQL database with PostGIS extension as spatial tabular structures

---

[3] *http://www.dpi.inpe.br/labisa/*

due to their relational dependence. An access at LabISA's database allowed for direct communication with PostgreSQL

2.3. Plugin architecture

In order to develop the plugin we applied some technologies and standards of web services. Web services are a way to expose the functionality of an information system and make it available through standard web technologies [Gottschalk et al. 2002]. These definitions may vary as these services are not single straightforward concepts but a complex pile of technological layers with common communication languages. To develop the plugin, three layers are necessary Data, Service and Client.

The Data layer is related to the data storage, organization and management. This layer is the farthest from the users and is totally abstracted from them, especially in the case of cloud storage solutions. It is tied to each provider (data owner or manager) and their choice of standards and technologies. At this layer, we can have data management with a traditional Database Management System (DBMS) (e.g. PostgreSQL [Momjian 2001]). On other hand, data can also be stored as plain files and distributed with a server manager, such as the Geoserver.

The Service layer is composed of all the software necessary to provide a functional environment to access the data and its information. This layer allows users to consume data without knowledge of internal data storage infrastructures through a set of APIs. These APIs can be standardized with communication standard protocols to ensure interoperability between applications. This allows the services to be language and technology independent.

The Client layer is built to consume the services and provide means of abstraction from the code and is the closest to the end-user. Clients are developed to use the API provided at the server layer, allowing for several kinds of applications to be built ranging from back-end to front-end, being independent of the programming language. For example, languages such as Python, R, PHP, and C++ can consume services for back-end operations and JavaScript can be used for front-end applications. For geospatial data, these clients often provide a Graphical User Interface (GUI) for mapping and several tools for data analyses, for example, the QGIS and ArcGIS. More recently, with the maturing of the internet, web-based applications consuming web services and providing interfaces and tools are becoming widespread.

Figure 2 presents the BONDS QGIS plugin service architecture in layers related to each of the layers presented before. At the bottom are the data providers by area of interest, in the middle are the services used for each, and on top is the BONDS QGIS plugin as a client.

2.4. QGIS Plugin development

The development of the plugin interface and code was carried out using a conjunction of three technologies: Python, QGIS, and QT Designer. This project uses, Python version 3.9.0 as available within QGIS components. QGIS is a free and open-source cross-platform desktop GIS application that supports viewing, editing, printing, and analysis of geospatial data [Lawhead 2017]. For QGIS, the selected version was the 3.24.3 Tisler.

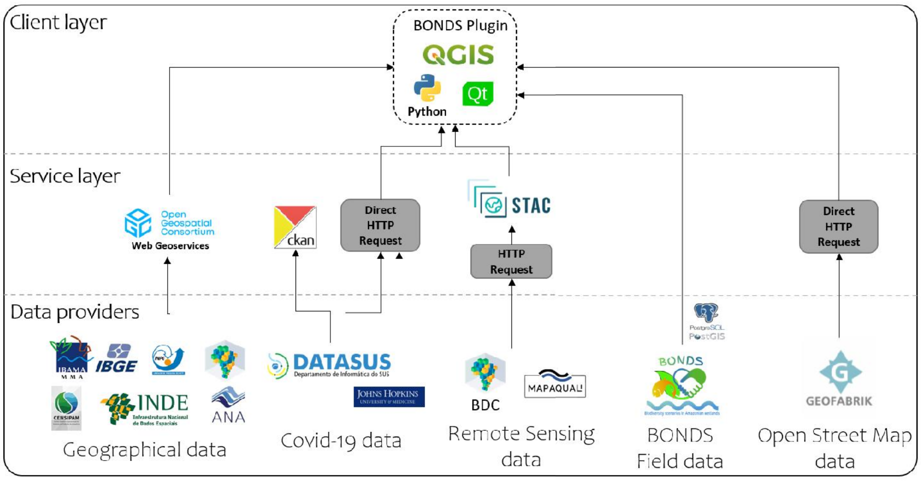


Figure 2. BONDS plugin architecture.

QGIS comes with a distribution of QT Designer by default. QT Designer is an integrated development environment that simplifies GUI application development.

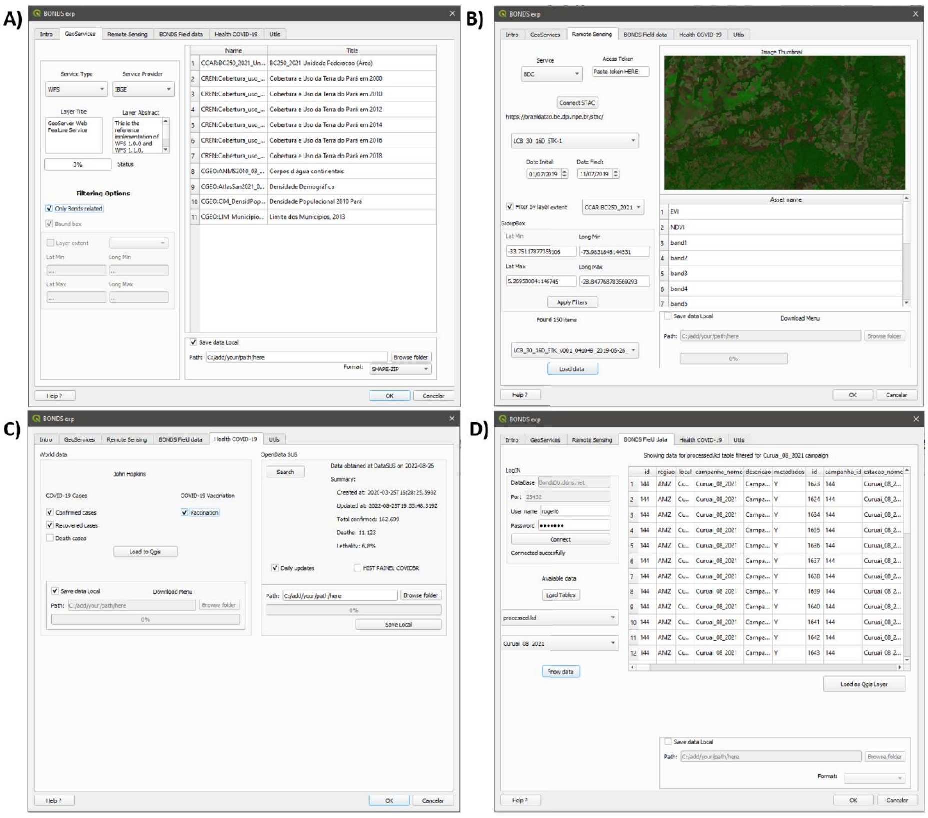
3. Results and discussion

To understand the data available in the context of BONDS, it was conducted a survey of possible data providers and products on five topics of interest: remote sensing imagery and products, field remote sensing, laboratory data, territorial cartography, and health data related to COVID-19. Table 1 presents the number of providers, relevant products and achievable products found. The relevant products were selected by assessing BONDS' objectives to develop studies related to biodiversity, aquatic ecosystems, climate change, and health. The achievable products are the ones that during the plugin implementation were possible to be accessed and consumed. Data providers and products will be presented in the following sections.

Topics of interest	Number of providers	Relevant products	Achievable products
Remote Sensing Imagery	2	10	4
Laboratory data	1	7	7
Field remote sensing	1	3	3
Territorial Cartography	7	25	25
Health COVID-19	3	7+	6

Table 1. Data availability by area of interest, BONDS relevant products and achievable products.

The next step was to catalog the maximum information of each provider and product to identify each group of technology and build the plugin tools. Therefore, the results obtained in this paper will be presented following the BONDS plugin development, as each tab represents a different data as well as a technology. This aims to simplify the access of different kinds of data and to make possible to evaluate the technologies applied to store, catalog and distribute it. The available tabs are: Intro, Geoservices, Remote sensing, BONDS field data, Health COVID-19, and Utils.



A)

Service Type: WFS
Service Provider: IBGE
Layer Title: GeoServer Web Feature Service
Layer Abstract: This is the reference implementation of WFS 1.0.0 and WFS 1.1.0,
0% Status
☑ Only Bonds related
☑ Bound box
☐ Layer extent
Lat Min: ...
Long Min: ...
Lat Max: ...
Long Max: ...

	Name	Title
1	CCAR:BC250_2021_Un...	BC250_2021 Unidade Federacao (Área)
2	CREN:Cobertura_uso_...	Cobertura e Uso da Terra do Pará em 2000
3	CREN:Cobertura_uso_...	Cobertura e Uso da Terra do Pará em 2010
4	CREN:Cobertura_uso_...	Cobertura e Uso da Terra do Pará em 2012
5	CREN:Cobertura_uso_...	Cobertura e Uso da Terra do Pará em 2014
6	CREN:Cobertura_uso_...	Cobertura e Uso da Terra do Pará em 2016
7	CREN:Cobertura_uso_...	Cobertura e Uso da Terra do Pará em 2018
8	CGEO:ANMS2010_03_...	Corpos d'água continentais
9	CGEO:AtlasSan2021_0...	Densidade Demográfica
10	CGEO:C04_DensidPop...	Densidade Populacional 2010 Pará
11	CGEO:LIM_Municipio...	Limite dos Municípios, 2013

☑ Save data Local
Path: C:/add/your/path/here
Format: SHAPE-ZIP

B)

Service: BDC
Access Token: Paste token HERE
https://brazildatacube.dpi.inpe.br/stac/
LCB_30_16D_STK-1
Date Inital: 01/07/2019
Date Final: 11/07/2019
☑ Filter by layer extent: CCAR:BC250_2021
GroupBox
Lat Min: -33.75117877355106
Long Min: -73.98318481144531
Lat Max: 5.269580041146745
Long Max: -28.847768783569293
Found 150 items
LCB_30_16D_STK_v001_041049_2019-05-26_

	Asset name
1	EVI
2	NDVI
3	band1
4	band2
5	band3
6	band4
7	band5

☐ Save data Local
Path: C:/add/your/path/here
0%

C)

World data
COVID-19 Cases
☑ Confirmed cases
☑ Recovered cases
☐ Death cases
COVID-19 Vaccination
☑ Vaccination
☑ Save data Local
Path: C:/add/your/path/here
0%

OpenData SUS
Data obtained at DataSUS on 2022-08-25
Summary:
Created at: 2020-03-25T15:28:25.593Z
Updated at: 2022-08-25T19:33:48.319Z
Total confirmed: 162.609
Deaths: 11.123
Lethality: 6,8%
☑ Daily updates
☐ HIST PAINEL COVIDBR
Path: C:/add/your/path/here
0%

D)

LogIN
DataBase: BondsDb.ddns.net
Port: 25432
User name: rogerio
Password: •••••••
Connected succesfully
Available data
processed.kd
Curuai_08_2021

Showing data for processed.kd table filtered for Curua_08_2021 campaign

	id	regiao	local	campanha_nome	descricao	metadados	id	campanha_id	estacao_nome
1	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1623	144	Curuai_08_2...
2	141	AMZ	Cu...	Curua_08_2021	Campa...	Y	1624	144	Curuai_08_2...
3	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1634	144	Curuai_08_2...
4	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1635	144	Curuai_08_2...
5	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1636	144	Curuai_08_2...
6	141	AMZ	Cu...	Curua_08_2021	Campa...	Y	1637	141	Curuai_08_2...
7	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1638	144	Curuai_08_2...
8	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1639	144	Curuai_08_2...
9	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1640	144	Curuai_08_2...
10	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1641	144	Curuai_08_2...
11	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1642	144	Curuai_08_2...
12	144	AMZ	Cu...	Curua_08_2021	Campa...	Y	1643	144	Curuai_08_2...

☐ Save data Local
Path: C:/add/your/path/here
Format:

Figure 3. BONDS QGIS plugin main tabs: a) Geoservices; b) Remote sensing; c) BONDS field data; d) Health COVID-19.

3.1. Geoservices

The main Geoservice explored in this paper was the WFS. In total, there are 11 possible operations to WFS requisitions. At the plugin we have focused mainly on the discovery and query operations. Discovery allows the service to be interrogated to determine its capabilities and to retrieve the application schema that defines the feature types that the service offers. Query allows features or values of feature properties to be retrieved from the underlying data store based upon constraints, defined by the client, on feature properties [Michaelis and Ames 2012].

Through the data survey, it was possible to identify seven providers with Geoservice enabled datasets. These providers are: INDE, BDC, INPE, SIPAM, IBAMA, ANA, IBGE. We selected a list of 25 products (layers) that could be of BONDS direct interest. However, at the GUI the user is capable of removing this filter and browsing all the provider's data. It also has the option to filter by bounding box using the WFS capabilities to send parameters along the requisitions (Figure 3a). After the filters choice, WFS vector layers are displayed and made available to be loaded to QGIS or downloaded.

With the WFS capabilities we were able to consume data from several providers in a single standardized format. This brings source and data reliability to the plugin as more and more providers join the standards. Also, allied with newer client technologies, the access and consumption of data were very fast when compared with traditional downloads. However, on the negative side, we experienced some hiccups with the layer's dynamic load. First, with instability and latency both from local or server, the WFS requisitions often retrieved ”timeout error” as the time limit for a response was reached. Also, with large datasets such as the LULC data, the loading was slower and more susceptible to interruptions, thus breaking the data already loaded. The use of resilient layers with microservices can offer solutions to reduce these problems [Nascimento et al. 2020].

From the data point of view, challenges for data integration such as temporal differences, geodetic datum, data format, and semantic inconsistencies were expected [Li et al. 2004, Li 2008]. In our dataset, temporal differences could be seen in data acquired or produced with different time spans, requiring the users to be aware of the date information. Datum problems were observed with data in different coordinate systems and projections, bringing spatial compatibility issues. Semantic problems were mainly related to different nomenclature or class systems among products. Data format was less of a problem as GDAL and QGIS offers many tools to ingest the data, however, data heterogeneity among providers is still an issue.

3.2. Remote sensing

For the remote sensing tab, two providers were supported in the final version, the BDC and MAPAQUALI projects. Both providers were selected due to their relevance to the BONDS project as they provide Amazon related data. These providers are being developed in the same infrastructure as they share common design patterns. We used *pystac-client* package to create a connection to the providers to browse their STAC and discover images and data cubes. The images are delivered by a Geoserver as COG, allowing GDAL to dynamically load image subsets, thus reducing the charge on the local QGIS instance.

Search can be refined through filters by collection, date, and spatial bounding box. Also, the user has the option to use the bounding box of a previously loaded QGIS layer (Figure 3b). Once the users select an item, its thumbnail (preview) and assets (bands) will be displayed and ready to be loaded to QGIS or downloaded. For BDC it is mandatory to have an API key to access the files. The key is available for registered users and can be generated at BDC web portal under the user's menu. Standardized access and discovery of data, through STAC allows to extend the plugin to reach other providers. GDAL's dynamic load proved to be an important feature to reduce the data load at the QGIS application. However, the size of remote sensing data is still a problem for data analysis and information extraction especially when you need all the data within the GIS (e.g. time series) [Ferreira et al. 2020].

3.3. BONDS field data

BONDS' field data are diverse and complex due to their data structures and the preprocessing steps needed to make data ready for use. With this heterogeneity, having a centralized common structure between datasets is very difficult. In this manner, the use of LabISA's integrated water quality database helps us to better understand and manage field data.

LabISA's tabular structure was composed of two main tables with field works and sample station information and several tables with water quality parameters linked to the main tables. QT and QGIS tools allow users to create a connection and perform queries within PosgreSQL given a valid username and password. In the GUI users have the possibility to filter the data by the desired quality parameter and also by fieldwork (time-related) (Figure 3c). After selection, the filtered data is displayed in a tabular structure and can be loaded to QGIS as a georeferenced point layer or downloaded as CSV. Using LabISA's PostgreSQL database with QGIS and QT tools was successful, however, publishing this data as OGC standards presents a more standardized and reliable way to access this field data in the future.

3.4. Health COVID-19

For COVID-19 health data, two providers and 6 products were implemented in the plugin. First, the global COVID-19 data bulletins from CRC, products are: confirmed, recovered, and death cases and also vaccination status (Figure 3d). Regarding Brazil's data, the OpendataSUS daily bulletins and the historical resumed panels were selected. These summarized data are already screened and have been available since the beginning of the pandemic.

3.5. Utils

During the plugin development, some tools were identified and added to the Utils tab to improve user's experience with the QGIS environment. This tab has currently three functionalities implemented: a) Basemaps loading; b) CSV to point layer converter; c) Brazil Open Street Map (OSM) data download. The Basemaps used in the BONDS plugin is distributed as a Tiled Web Map (TWM) which supports the loading of several tiles in parts, making it seemingly fast. Also, a simple CSV to layer converter was added to help users load local data. The OSM data is a bundle of several geospatial data from Brazil that is maintained and distributed by Geofabrik [4].

4. Conclusion

In this paper we presented a QGIS plugin for discovering, browsing and acquiring (download) datasets from several data providers with different sharing technologies. The use of QT Designer within QGIS application was helpful to design the BONDS plugin GUI as a non-programming view. It is important to note that the GDAL within QGIS was essential to the project development as it presents a reliable and widely used Toolkit for geospatial handling. Also, the QGIS Python API presented a useful set of tools for data import, export, and management within the QGIS application.

The use of geoservices and cloud databases with standardized infrastructures and protocols has proven to be a great way to deal with complex geospatial data. These services often not only provide an easy access to data but also several server-side tools to aid users. Also, in modern technologies, the possibility of requesting only the necessary data can greatly improve the user's experience. Although geoservices offer great tools it is important to note that there are also some drawbacks. The first is a dependency on good internet connections from both servers and users. Second, the users are prone to changes in the service standards or data distribution policies. Lastly, as those technologies are still maturing their support and documentation can be sometimes absent.

---

[4] *https://www.geofabrik.de/*

We believe that the use of standardized web services and geoservices are already the main option to deal with data analysis. However, at the Client layer, the necessity to download this data locally in the GIS presents a bottleneck for large datasets or time series. On this matter, some approaches are exploring the concepts of server-side processing, such as the WPS and WCPS. Following the moving code paradigm, these approaches aim to bring new levels of abstraction for users to perform data analysis directly on the server [Müller et al. 2010, Müller 2016]. However, at the current state of data availability for heterogeneous datasets and providers, server-side approaches are still not feasible. As technologies and standards for geoservices mature, their uncertainties are constantly being reduced, and new solutions are proposed. At the current state, the BONDS plugin was assertive in centralizing some of the available services and providing tools for BONDS researchers to discover, access and download the datasets.

5. Acknowledgements

The authors would like to thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 for funding his Ph.D scholarship. Also, thanks to CAMPUS FRANCE Brazil and SABERES project (BNP Paribas) for co-funding this research. Thanks to LabISA and BONDS (Biodiversa - Belmont Forum) for providing access to data.

References

Baumann, P. (2010). The ogc web coverage processing service (wcps) standard. *Geoinformatica*, 14(4):447–479.

BRASIL (2022a). Campanha nacional de vacinação contra covid-19 [national vaccination program against covid-19]. `https://opendatasus.saude.gov.br/dataset/covid-19-vacinacao/`.

BRASIL (2022b). Opendatasus. `https://opendatasus.saude.gov.br/`.

CONCAR (2010). Plano de ação para a implantação da inde.

de Lucia Lobo, F., Barbosa, C. C. F., de Carvalho, L. S., Menino, F., Carlos, E. M. L. d. M., and Novo, E. F. d. S. (2017). Proposta de um sistema de monitoramento integrado da qualidade das águas continentais com imagens de satélites: Mapaquali.

Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534.

Ferreira, K. R., Queiroz, G. R., Vinhas, L., Marujo, R. F., Simoes, R. E., Picoli, M. C., Camara, G., Cartaxo, R., Gomes, V. C., Santos, L. A., et al. (2020). Earth observation data cubes for brazil: Requirements, methodology and products. *Remote Sensing*, 12(24):4033.

Giuliani, G., Dubois, A., and Lacroix, P. (2013). Testing ogc web feature and coverage service performance: Towards efficient delivery of geospatial data. *Journal of Spatial Information Science*, 1(7):1–23.

Gottschalk, K., Graham, S., Kreger, H., and Snell, J. (2002). Introduction to web services architecture. *IBM systems Journal*, 41(2):170–177.

INDE (2022). Catálogo de geoserviços. `https://inde.gov.br/CatalogoGeoservicos`.

Lawhead, J. (2017). *QGIS Python Programming Cookbook*. Packt Publishing Ltd.

Lee, J.-G. and Kang, M. (2015). Geospatial big data: challenges and opportunities. *Big Data Research*, 2(2):74–81.

Lemmens, R. (2006). Semantic interoperability of distributed geo-services. ITC.

Li, D. (2008). On generalised and specialised spatial information grids: are geo-services ready? *International Journal of Digital Earth*, 1(4):315–325.

Li, D., Shao, Z., Zhu, X., and Zhu, Y. (2004). From digital map to spatial information multi-grid. In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, volume 5, pages 2933–2936. IEEE.

Maguire, D. J. and Longley, P. A. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, environment and urban systems*, 29(1):3–14.

Michaelis, C. D. and Ames, D. P. (2012). Considerations for implementing ogc wms and wfs specifications in a desktop gis.

Momjian, B. (2001). *PostgreSQL: introduction and concepts*, volume 192. Addison-Wesley New York.

Müller, M. (2016). Service-oriented geoprocessing in spatial data infrastructures.

Müller, M., Bernard, L., and Brauner, J. (2010). Moving code in spatial data infrastructures–web service based deployment of geoprocessing algorithms. *Transactions in GIS*, 14:101–118.

Nascimento, H. L., de Souza Baptista, C., de Andrade, F. G., and Santos, L. C. (2020). Towards a resilient spatial data infrastructure. In *GeoInfo*, pages 34–45.

OGC (2022). Open geospatial consortium. `https://www.ogc.org/about`.

Raju, P. (2006). Fundamentals of geographical information system. *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology*, page 103.

Santos, A. F. D. (2005). The policy for commercializing cbers data. In *56th International Astronautical Congress of the International Astronautical Federation, the International Academy of Astronautics, and the International Institute of Space Law*, pages E6–1.

Silva, P. N. and Pinheiro, M. M. K. (2019). Execução da política de dados abertos no brasil. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 12(2).

Stopa, S. R., Malta, D. C., Monteiro, C. N., Szwarcwald, C. L., Goldbaum, M., and Cesar, C. L. G. (2017). Use of and access to health services in brazil, 2013 national health survey. *Revista de saude publica*, 51.

Tole, A. A. et al. (2013). Big data challenges. *Database systems journal*, 4(3):31–40.