USING REMOTE SENSING IMAGES AND CLOUD SERVICES ON AWS TO IMPROVE LAND USE AND COVER MONITORING

K. R. Ferreira[a,*], G. R. Queiroz[a], G. Camara[a], R. C. M. Souza[a], L. Vinhas[a], R. F. B. Marujo[a], R. E. O. Simoes[a], C. A. F. Noronha[a], R. W. Costa[a], J. S. Arcanjo[a], V. C. F. Gomes[a] and M. C. Zaglia[a]

[a] INPE, National Institute for Space Research, São José dos Campos, Brazil - (karine.ferreira, gilberto.queiroz, gilberto.camara, ricardo.cartaxo, lubia.vinhas, rennan.marujo, rolf.simoes, carlos.noronha, raphael.costa, jeferson.arcanjo, vitor.gomes, matheus.zaglia)@inpe.br

KEY WORDS: Cloud Computing, Remote sensing images, land use and cover monitoring, Amazon Web Service (AWS)

ABSTRACT:

The Brazilian National Institute for Space Research (INPE) produces official information about deforestation as well as land use and cover in the country, based on remote sensing images. The current open data policy adopted by many space agencies and governments worldwide provided access to petabytes of remote sensing images. To properly deal with this vast amount of images, novel technologies have been proposed and developed based on cloud computing and big data systems. This paper describes the INPE's initiatives in using remote sensing images and cloud services of the Amazon Web Services (AWS) infrastructure to improve land use and cover monitoring.

1. INTRODUCTION

The Brazilian National Institute for Space Research (INPE) is the main government research organization working in Earth Observation and Space Science areas. Since 1988, INPE has led projects to produce official information about deforestation and Land Use and Cover Change (LUCC).

Since 2004, the DETER project aims at producing deforestation alerts for the Brazilian Amazon and other biomes in Brazil (Diniz et al., 2015). The PRODES project monitors shallow-cut deforestation in the Brazilian Amazon since 1988 and is responsible for providing accurate official deforestation rates for all biomes in Brazil (INPE, 2019). The data sets generated by DETER and PRODES are disseminated in a web platform called TerraBrasilis (Assis et al., 2019). The TerraClass project investigates what the deforested areas detected by PRODES have become, understanding and explaining LUCC processes in the Brazilian Amazon and Cerrado biome. Based on remote sensing data analysis and geoinformation techniques, interpreters classify deforested areas into different land use and cover classes and evaluate the spatiotemporal and semantic dynamics of these areas (Almeida et al., 2016).

The open data policy adopted by many space agencies and governments worldwide has made accessible petabytes of remote sensing images of different spatial, spectral and temporal resolutions. For effectively managing and analyzing these vast amounts of images, novel technologies based on cloud computing and big data methods have been proposed as a viable option (Wang et al., 2013).

Currently, several cloud computing environments exist. The Microsoft Azure Cloud Services and Amazon Web Services (AWS) are, arguably, the most well known. Microsoft Azure was launched in 2010 by Microsoft and offers services that can be categorized as Platform as a Service (PaaS), software as a service (SaaS) and Infrastructure as a service (IaaS). AWS was launched in 2006 by Amazon and it also offers PaaS, SaaS and IaaS services.

Both platforms are highly scalable and offer developer tools, storage, database, and networking services. However, AWS contains several of the most used remote sensing image collections, including Landsat-8, Sentinel-2 and CBERS-4 images. This is an advantage when compared to Azure.

In the era of big Earth observation data and cloud computing environments, it is crucial to improve the land use and cover monitoring projects of INPE to take advantage of the big amount of remote sensing imagery freely available. In this context, this paper describes INPE's initiatives and experiences in using remote sensing images and cloud services on the AWS to improve its projects for land use and cover monitoring. To execute this work, the INPE's team is using the credits earned under the GEO (Group on Earth Observations) and AWS Earth Observation Cloud Credits Programme.

The general objective of this work is to use and evaluate the AWS cloud computing environment, developing applications and services to improve the land use and cover monitoring projects of INPE. The current development on cloud computing focuses on three fronts: (1) Forest Monitor application, described in Section 2; (2) Brazil Data Cube generation, described in Section 3; and (3) LUCC classification, described in Section 4. Some final remarks are presented in Section 5.

2. FOREST MONITOR APPLICATION

Forest Monitor is a web-based platform to support the detection of deforestation alerts by accessing and visualizing the remote sensing images stored in the AWS buckets of data. The system allows the visualization of Sentinel-2A/MSI, Sentinel-2B/MSI, Landsat-8/OLI and CBERS-4/AWFI image collections from the moment they are published on AWS. These image collections are used by experts to map the deforestation alerts. The platform offers a set of basic functionalities such as the image contrast enhancement, spatio-temporal visualization using a temporal slider and the delineation of the alerts that are saved as vector layer. Figure 1 shows the prototype of this platform.

The Forest Monitor platform integrates all remote sensing images of medium spatial resolution available in the AWS with deforestation polygons detected in PRODES and DETER in a desktop environment. The platform is being used for the interpreters to improve the analysis and evaluation of deforested areas images from

*Corresponding author.

https://doi.org/10.5194/isprs-archives-XLII-3-W12-2020-207-2020 |

Figure 1: Forest Monitor Application.

different sensors, in particular to compare deforestation areas previously detected with the most recently satellite image available in the AWS.

Figure 2 presents the Forest Monitor architecture. The application is based on the services provided by the RemotePixel.ca, we have created a Tile Map Services (TMS) using AWS Lambda service to access Sentinel, Landsat and CBERS images from the AWS buckets of data. For image collection metadata access, we used the STAC (SpatioTemporal Asset Catalog) API provided by the Development Seed `https://sat-api.developmentseed.org/search/stac` to access Sentinel and Landsat image collections, and provided by AMS Kepler `https://cbers.stac.cloud/` to access CBERS images. The components of the Forest Monitor architecture are part of the project Brazil Data Cube (BDC) that will be described in Section 3.

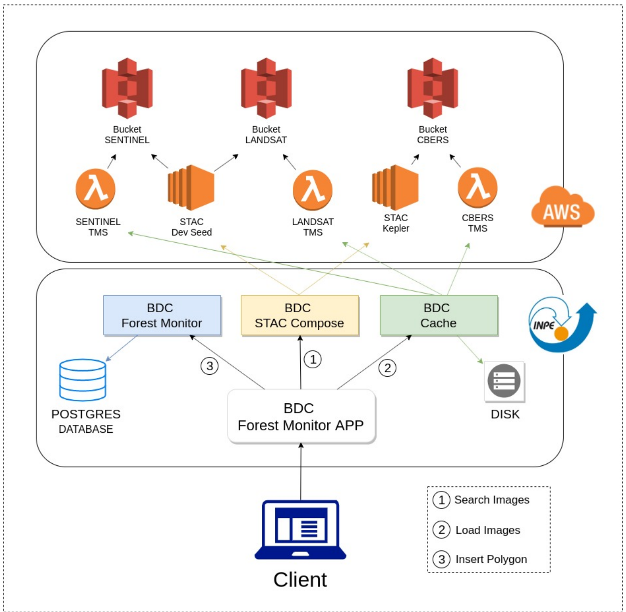

Figure 2: BDC Forest Monitor platform architecture.

3. BRAZIL DATA CUBE GENERATION

Since the beginning of 2019, the INPE's team is working in a project called Brazil Data Cube (BDC) (see `http://brazildatacube.org/`). This project has four main objectives: (1) create analysis-ready data (ARD) from remote sensing images of medium spatial resolution (10 to 30 meters) for the Brazilian territory, including images from Landsat, CBERS and Sentinel; (2) model the ARD data sets as multidimensional cubes with three or more dimensions, including space, time and spectral properties; (3) propose and develop novel methods and techniques to store, process and analyze data EO cubes using satellite image time series analysis, image processing and machine learning methods; and (4) use the data cubes and methods to improve the generation of LUCC information for Brazil. The Brazil Data Cube project follows a worldwide trend in creating multidimensional data cubes from remote sensing images for a specific country, such as the Australian Data Cube (Lewis et al., 2017) and the Swiss Data Cube (Giuliani et al., 2017) projects.

Figure 3 illustrates the process of data acquisition, preprocessing and cube generation. To generate a data cube, a grid is used to search all available images of a given image collection, these images are merged, reprojected, resampled and gridded to a common spatial reference. Than, a temporal compositing function is used to build regular intervals (16 days or monthly) and reduce the data dimensionality according to an aggregation function (such as median or best quality pixel).

The image acquisition and preprocessing scripts obtain Landsat-8 OLI and Sentinel (2A and 2B) MSI sensor images, store their metadata in an internal database catalog and process them locally to generate the surface reflectance products using LaSRC (Vermote et al., 2016) and Sen2cor (Louis et al., 2016) atmospheric correction, respectively. These products are ingested in the AWS Simple Storage Service (S3). The surface reflectance images from CBERS-4 AWFI are already available in AWS. Figure 4 illustrates the set of AWS services used to implement the generation of the cubes in this environment. The STAC service is a catalog. The AWS Lambda allows code to run without provisioning or managing servers, up to 1000 lambdas in parallel. They are used to implement image processing operations. AWS lambdas are scaled by the AWS Simple Queue Service (SQS) service. The AWS Dynamo DB service is a high performance key-value database. In the context of data cube generation it is used in the activities orchestration. AWS Relational Database Service (RDS) enables the configuration, operation, and scalability of relational databases in the cloud, using an instance of MariaDB/MySQL that stores products metadata. The Kinesis service serializes messages from lambdas, so that enables other lambdas

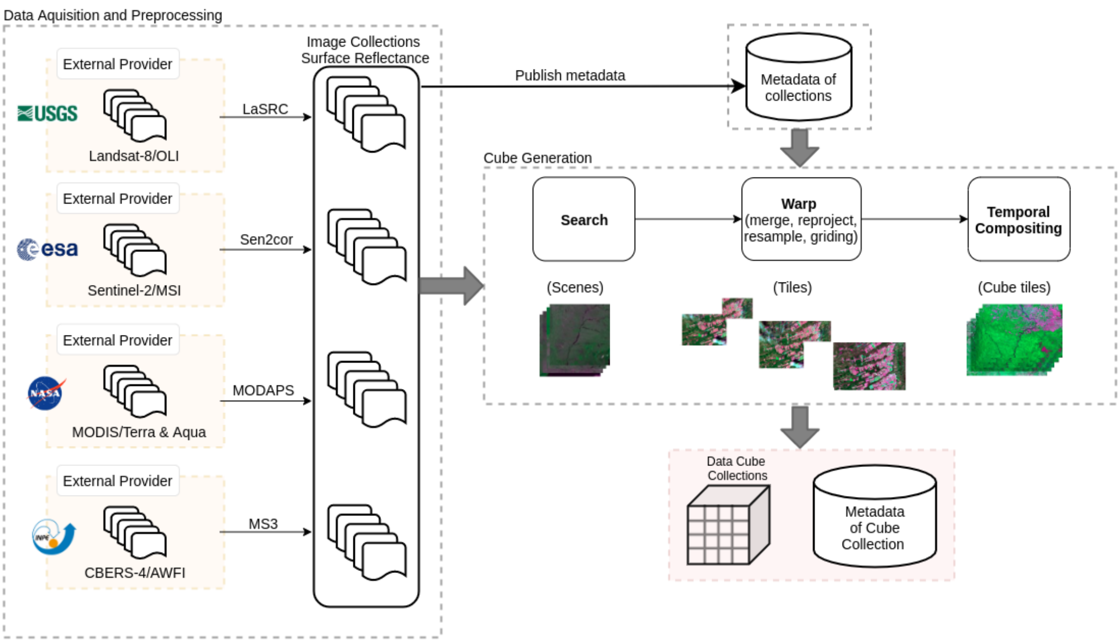

Figure 3: Data Cube generation process.

that can process these messages in an orderly manner, for example access to DynamoDB and RDS. At the end of the cube generation process, the images are stored in a S3 bucket and their metadata are organized in RDS to be accessed by users.

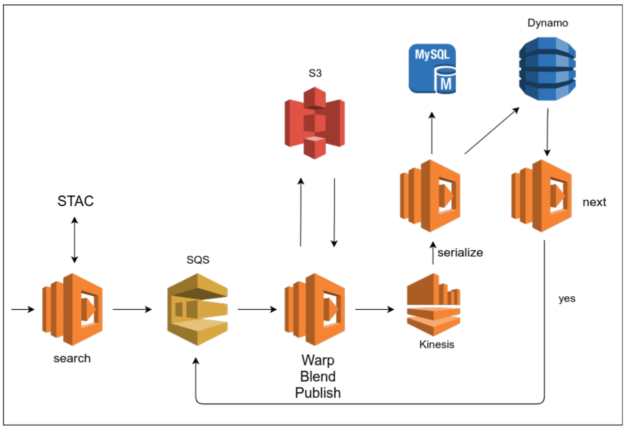

Figure 4: Brazil Data Cube generation service workflow on AWS.

The scripts developed in the context of the project can be accessed on the project code repository at `https://github.com/brazil-data-cube`, as free and open source software.

4. LUCC CLASSIFICATION

The INPE's team is also researching methods to extract LUCC information from big Earth observation data sets, using satellite image time series analysis, machine learning algorithms and image processing procedures (Picoli et al., 2018) (Santos et al., 2019). These methods are being developed in an open source R package called Satellite Image Time Series (SITS) (`https://github.com/e-sensing/sits`). The SITS package provides a set of tools for working with analyses, clustering and classification of satellite image time series.

Figure 5 illustrates LUCC classification in the AWS environment, for the Mato Grosso state, Brazil. The SITS package runs in an R Studio Server in the Amazon Elastic Compute Cloud (EC2) service to process MOD13Q1 product files stored in AWS S3 buckets. To process the classification in an EC2 service, we choose a AWS machine called r4.8xlarge with 32 CPUs and 240 GB of RAM memory. In this environment we took 1,5 days to classify 18 years of image time series from MOD13Q1 product The LUCC map resulting from this classification is also shown in Figure 5.

The Mato Grosso classification processed around 660 thousands of 480x480 pixels images stored in AWS S3 buckets. These files composed the time series used in the classification. The images metadata were indexing in catalog, that is being developed as a web service, a central component of the data cube technology. It enables the search and retrieval of images stored in buckets and facilitates the interoperability between other services and clients software.

5. FINAL REMARKS

This paper presents three INPE initiates in using and evaluating the remote sensing images and cloud services on AWS: (1) Forest Monitor application; (2) Brazil Data Cube generation; and (3) LUCC classification. In this work, we describe the software architecture of each initiative and the AWS services used.

The AWS services tested in this work are very fast and useful, specially the AWS Lambda service that allows users to run code without provisioning or managing servers. Table 1 shows the cost to generate a data cube of Sentinel 2 images with spatial resolution of 10 meters and temporal resolution of 16 days. It compares the cost to generate the cubes using on premise hardware (1 machine with 32 CPUs of 128 GB of RAM) and using AWS lambdas. As can be seen, there is an expressive gain using the AWS environment.

	Sentinel Cube 1	Sentinel Cube 2
Spatial extent	1 tile of the BDC grid	Brazilian territory (560 tiles of BDC grid)
Temporal extent	1 year	1 year
Time to run on premise hardware	1242 min = 20 hours	11592 hours = 483 days
Time to run on AWS using 260 lambdas	5 minutes	2800 minutes = 46 hours
AWS cost using 260 Lambdas	US$ 3,00 (without S3 cost)	US$ 840,00 (without S3 cost)

Table 1: Cost to generate a Sentinel-2 data cube.

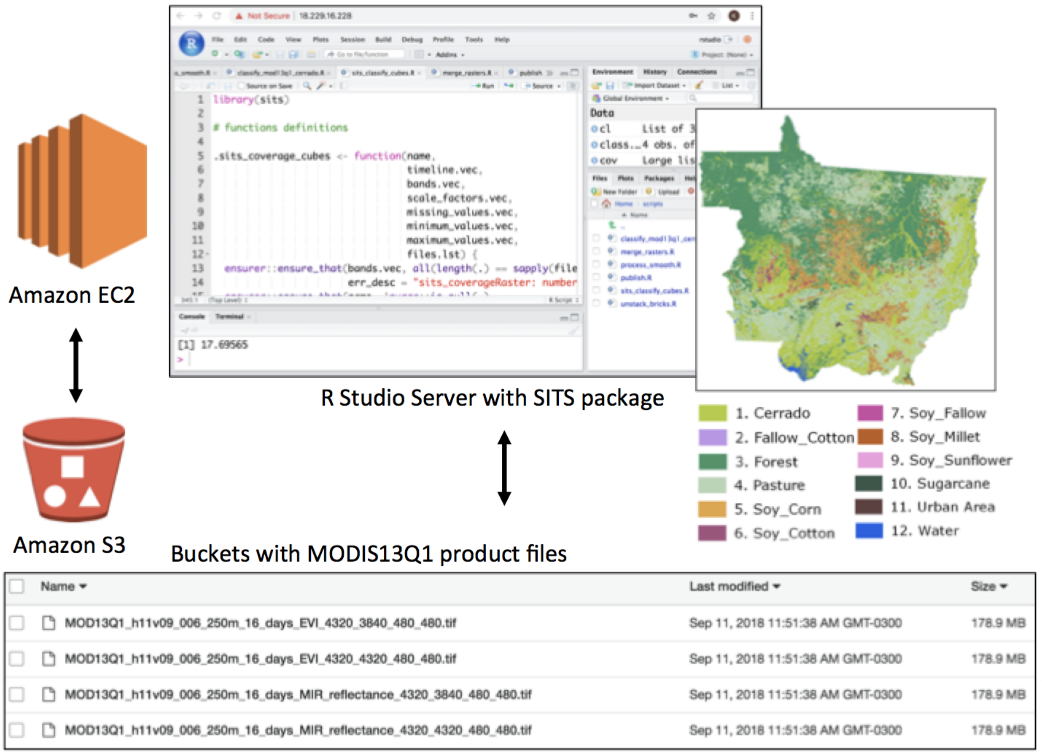

Figure 5: LUCC classification on AWS.

The main drawback faced in this work is that the repository *Earth on AWS* does not offer surface reflectance products (ARD data) from Sentinel 2, Landsat 8 and CBERS 4 images. To generate the data cubes described in Section 3, we downloaded the original images from USGS, ESA and INPE to locally to generate the surface reflectance images and upload them to AWS.

As future work, we intend to use and evaluate the AWS SageMaker service to build, train, and deploy machine learning models.

ACKNOWLEDGEMENTS

The authors would like to thank the Fundo Amazônia, Amazon AWS and GEO (Group on Earth Observations) for the support in this project.

REFERENCES

Almeida, C. A. d., Coutinho, A. C., Esquerdo, J. C. D. M., Adami, M., venturieri, A., Diniz, C. G., Dessay, N., Durieux, L. and Gomes, A. R., 2016. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amazonica 46, pp. 291 – 302.

Assis, L. F. F. G., Ferreira, K. R., Vinhas, L., Maurano, L., Carvalho, A., Rodrigues, J., Maciel, A. and Camargo, C., 2019. Terrabrasilis: A hybrid spatial data analytics infrastructure using micro-services for large-scale thematic mapping projects in brazil. pp. 1–26.

Diniz, C. G., d. A. Souza, A. A., Santos, D. C., Dias, M. C., d. Luz, N. C., d. Moraes, D. R. V., Maia, J. S., Gomes, A. R., d. S. Narvaes, I., Valeriano, D. M., Maurano, L. E. P. and Adami, M., 2015. Deter-b: The new amazon near real-time deforestation detection system. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8(7), pp. 3619–3628.

Giuliani, G., Chatenoux, B., Bono, A. D., Rodila, D., Richard, J.-P., Allenbach, K., Dao, H. and Peduzzi, P., 2017. Building an earth observations data cube: lessons learned from the swiss data cube (sdc) on generating analysis ready data (ard). Big Earth Data 1(1-2), pp. 100–117.

INPE, 2019. Prodes project: Brazilian amazon forest monitoring by satellite. http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes. (28 Dec. 2019).

Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., Raevksi, G., Hooke, J., Woodcock, R., Sixsmith, J., Wu, W., Tan, P., Li, F., Killough, B., Minchin, S., Roberts, D., Ayers, D., Bala, B., Dwyer, J., Dekker, A., Dhu, T., Hicks, A., Ip, A., Purss, M., Richards, C., Sagar, S., Trenham, C., Wang, P. and Wang, L. W., 2017. The Australian Geoscience Data Cube — Foundations and lessons learned. Remote Sensing of Environment 202, pp. 276–292.

Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E. and Gascon, F., 2016. Sentinel-2 sen2cor: L2a processor for users. European Space Agency, (Special Publication) ESA SP SP-740(August), pp. 9–13.

Picoli, M. C. A., Camara, G., Sanches, I., Simões, R., Carvalho, A., Maciel, A., Coutinho, A., Esquerdo, J., Antunes, J., Begotti, R. A., Arvor, D. and Almeida, C., 2018. Big earth observation time series analysis for monitoring Brazilian agriculture. ISPRS Journal of Photogrammetry and Remote Sensing 145(August), pp. 328–339.

Santos, L., Ferreira, K. R., Picoli, M. and Camara, G., 2019. Self-organizing maps in earth observation data cubes analysis. In: International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM+), Vol. 13, Barcelona, Spain, pp. 70–79.

Vermote, E., Justice, C., Claverie, M. and Franch, B., 2016. Preliminary analysis of the performance of the landsat 8/oli land surface reflectance product. Remote Sensing of Environment 185, pp. 46–56.

Wang, P., Wang, J., Chen, Y. and Ni, G., 2013. Rapid processing of remote sensing images based on cloud computing. Future Generation Computer Systems 29(8), pp. 1963–1968.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLII-3-W12-2020-207-2020 | © Authors 2020. CC BY 4.0 License.
Primary publication at IEEE Xplore: https://doi.org/10.1109/LAGIRS48042.2020.9165649