INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 5)
Available online at: www.ijariit.com

Open Data Cube: Future of EO satellites

Parul Mittal
parul241995@gmail.com
Devi Ahilya Vishwavidyalaya, Indore, Madhya Pradesh

ABSTRACT

The Open Data Cube (ODC) is a network of individuals and associations building ability for working with earth perception information. At the center, the ODC is a product library and set of order line devices. What's more, keeping the center together we have a gathering of associations working together on technique and who may keep up particular usage of the ODC. To study the Earth Observation data, ODC is playing a great role which will help scientists to make a confident conclusion.

Keywords— *Open Data Cube, AGDC, satellites, Earth observation, big data*

1. INTRODUCTION

With the launching of satellites every year, petabytes of satellite data are generating every year. High volume data of different varieties is generating with a high velocity. The challenge is to manage such data which is geospatial. Better management of data not only helps in the efficient memory utilization but it also helps researchers in the perfect analysis. Satellite Earth observation data collection is 'Big data'. The intervals of observation reduce from weeks to days, days to hours, hours to minutes in case of geostationary satellites.

Remotely sensed Earth Observations (EO) data are increasingly available from a number of freely and openly accessible repositories. These data are highly valuable because of their unique and globally consistent information that they include (Lewis et al., 2016). Indeed, global observations together with scientific expertise and appropriate tools provide substantial benefit supporting economic development, decision-making, and policy implementation for all countries. However, the full information potential of EO data has not been yet realized. They remain still underutilized and stored in electronic silos of data (Gore, 1998; Lewis et al., 2016).

This is because of a few reasons:

- Expanding volumes of information produced by EO satellites
- Absence of mastery, framework, or web transmission capacity to proficiently and successfully access, process, and use EO information
- The specific sort of exceedingly organized information that EO information speak to presenting challenges when attempting to coordinate or investigate them
- The generous exertion and cost required to store and process information restrain the proficient utilization of this information (CEOS, 2017; Lewis et al., 2016; Press et al., 2015).

Hence, EO information can be considered as Big Data, information that is too substantial, quick lived, heterogeneous, or complex to get comprehended and misused (Baumann, Rossi, et al., 2016). Therefore, we require new ways to deal with completely advantage from EO information and:

- Open the data intensity of EO information
- Expand the utilization of EO information to a more extensive scope of networks
- Bolster choices producers with the learning they require by methodically breaking down every accessible perception and convert them into significant geophysical factors.

There is an exploration continuing for the administration of information, there is such an idea named Data Cube which is as of late utilizing in satellite symbolism.

As per the Open Data Cube site, "The goal of Open Data Cube is to build the effect of satellite information by giving an open and uninhibitedly available misuse instrument, and to encourage a network to create, manage and develop the broadness and profundity of use".

The essential issues for clients are information get to, information readiness, and proficient examinations to help client applications. The two first issues are fundamental difficulties to handle while building a DC. Without a doubt, these means concern the age of Analysis Ready Data (ARD). Presidents characterize ARD as "satellite information that has been prepared to a base arrangement of prerequisites and sorted out into a shape that permits quick examination without extra client exertion"

The fundamental test is to get to the information, setting up the information and its proficient investigation to help client application. There are a few wants that are related to satellite information:

- Time and logical learning prerequisite in getting to and planning information ought to be limited.
- Free and open Earth perception satellite information and application calculations.
- Efficient time arrangement investigation to help arrive change applications.
- Uses of numerous datasets together.
- Use of regular GIS instruments.
- A local and provincial arrangement that stay away from business and web reliance.

2. HISTORY

Quite a while back, satellite symbolism was dispersed on huge moves of tape. A less lengthy time-frame back, in 2011, Geoscience Australia worked with various different associations to duplicate the Landsat information they had put away on tapes and in different areas onto turning circles at the National Computational Infrastructure as a major aspect of the Unlocking the Landsat Archive venture. Sometime after this, the Australian Geoscience Data Cube (AGDC) was produced, a Landsat particular device ready to improve access to these Landsat documents.

Prior, satellite pictures were put away on moves of tape, however, with the expanding satellites and their information, there is required in the progression of innovation. After at that point, Geoscience Australia begins putting away the satellite information on turning disks 2017, AGDC v2 was renamed as Open Data Cube.

3. CHALLENGE

In any case, separating data from EOS information property represents a huge innovative test because of the volume and assortment of information. Satellite Earth perception information accumulations are 'Enormous Data'. The Oxford Dictionary characterizes 'Enormous Data' as "Datasets that are too vast and complex to control or grill with standard strategies or devices". The term was first depicted by Laney (2001) as "the three V's: Volume, Velocity, and Variety". Others have stretched out this rundown to incorporate legitimacy, veracity, esteem, and permeability.

New Earth perception information examination strategies are rising which request ongoing access to full documents of moderate goals Earth perception information, these include:

- The utilization of time-arrangement investigation strategies to precisely recognize the change.
- The orderly portrayal of a specific cover compose over various decades
- The utilization of time arrangement of Earth perception information and determined items as Essential Climate Variables (ECV's) and Climate Data Records.

A further test is that substantial accumulations of satellite Earth perception information are regularly unique, with new information being ceaselessly included and additionally existing information being refreshed. Information frameworks are required that can oversee such adaptable accumulations.

4. OPEN DATA CUBE

The Open Data Cube (ODC) is a network of individuals and associations building ability for working with earth perception information. At the center, the ODC is a product library and set of order line devices. What's more, keeping the center together we have a gathering of associations working together on technique and who may keep up particular usage of the ODC.

The Data Cube was initially executed as standard, non-covering, 'tiles' of gridded sensor information, 'stacked' as indicated by the season of information catch (perception), prompting the visual metaphor of a 'solid shape' of information. For Version 1 of the AGDC, we picked a 4000 ∗ 4000 cluster of pixels to be an entire tile. Communicated in geographic directions, the tiles were 1 degree by 1 degree and every pixel 0.00025 degrees on a side. In Version 2 of the AGDC, the tile measure and the pixel estimate are parameters set at the season of information ingest. Ingestion designs the information for capacity - viably making an AGDCv2 organized dataset. The capacity to design ingest parameters implies that AGDCv2 has expanded adaptability to help gridded perceptions of shifting goals from different sources.

AGDCv2 additionally can file, as opposed to instead of ingesting, a dataset. The ordering procedure adds data to the Data Cube database which empowers AGDCv2 capacities to work specifically on the local information. This information would thus be able to stay in their unique design and be custom fitted amid access to the client's need. The ordering approach improves expansion of new sensor information while permitting those datasets to be gotten to flawlessly through the API. Ordering is helpful where it is unfortunate to rebuild the information.

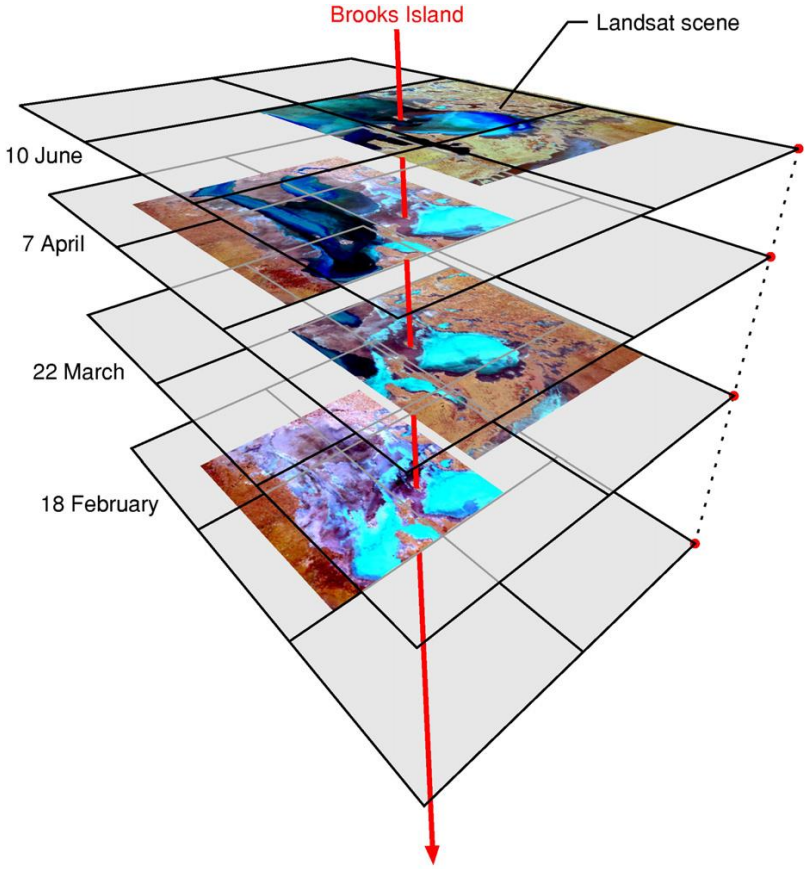

Fig. 1: Data Cube Concept: Landsat scenes change over time

4.1 Architecture

It includes four layers from base to top as pursues:

(1) Information Acquisition and Inflow- Observations are gathered and preprocessed to an 'investigation prepared' level by different overseers.

(2) Information Cube Infrastructure- investigation prepared information are filed into the ODC including ingestion into multi-dimensional datasets, at present netCDF4-CF1.6, with a suite of instruments for undertaking execution, disclosure, perception etc.

(3) Information and Application Platform- Platforms and condition that permit routine age of items, and, investigation of new items in a 'virtual research center' condition.

(4) UI and Application Layer- An assorted arrangement of utilization is empowered by the hidden foundation.

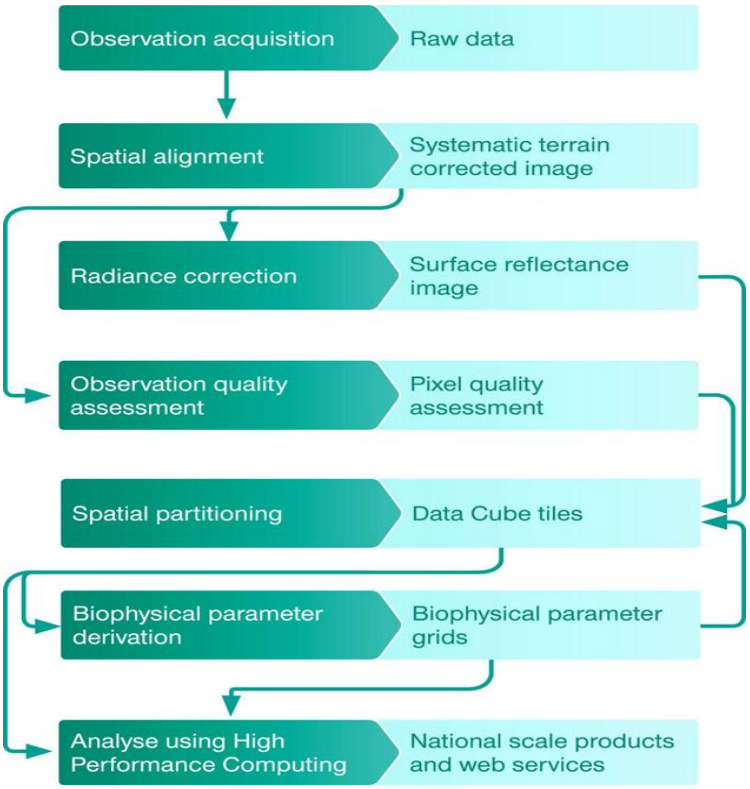

Fig. 2: Flow chart of data cube

The ODC code as given on github is composed in Python (N2.6, N3.3) and freely accessible under the Apache License, Version 2.0. The code vigorously use open source libraries including pyyaml, sqlalchemy, and jsonschema for record property and database get to; netCDF4 for Data Cube dataset creation and access; Luigi (https://github.com/spotify/luigi) for overseeing work processes, numpy, xarray and dask for exhibit dealing with, rasterio/gdal for geospatial information get to, scipy, pandas, scikit-picture, and scikit-learn for examination.

4.2 Structure of data cube

It includes five tables in Postgresql:
(1) Metadata_type: It comprises of the kind of metadata eg. Either information can be of earth perception (eo) or telemetry.
(2) Dataset_type: It comprises of the item definition the dataset.
(3) Dataset_location
(4) Dataset_source
(5) Dataset: It comprises the first information and the ingested information.

The accompanying factors might be designed while ingesting information into the Data Cube:

- Coordinate Reference System
- Spatial goals
- Chunk measurements: x, y, t
- Compression
- Temporal profundity: hour, day, month year/s
- Tile measurement: x, y
- Data compose: CF tradition

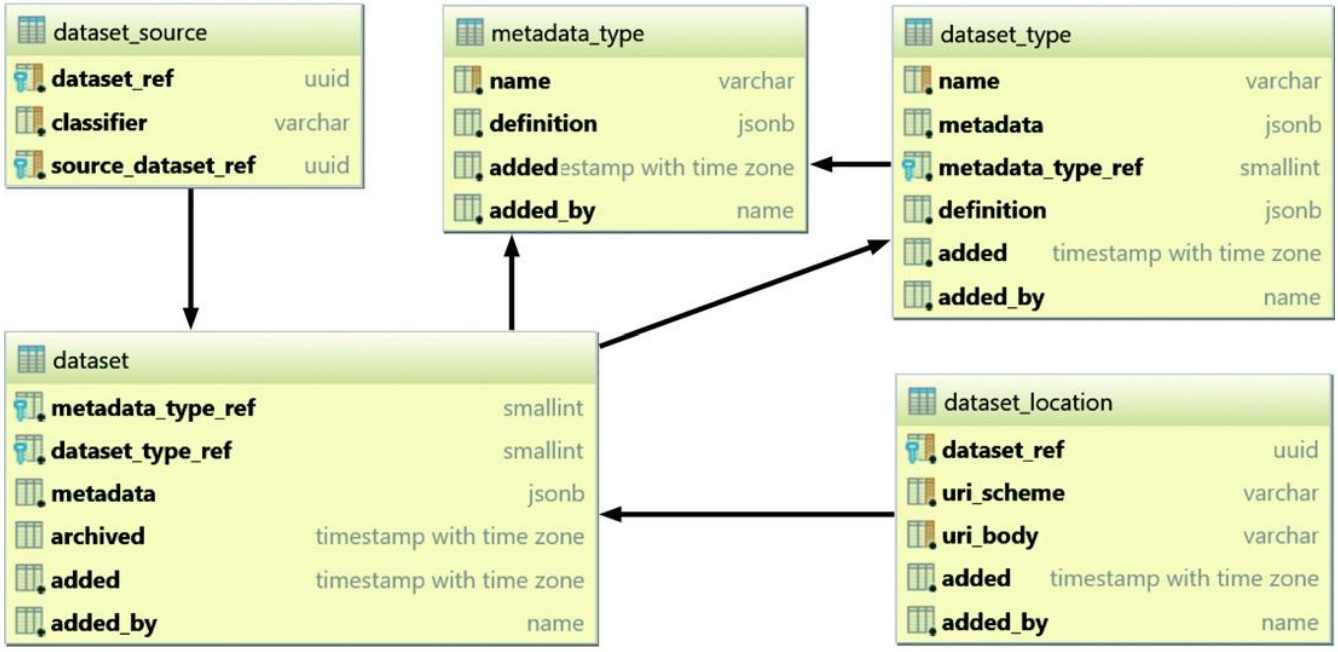

Fig. 3: Structure of ODC in PostgreSQL

5. APPLICATIONS

• Observation of water from space: ODC shows a model that aides in the investigation of the water from the space.
• Best accessible pixel composites and worldly measurements: ODC holds all perceptions the 'best pixel' idea can be reached out to incorporate fleeting insights, with the client picking run of the mill (or atypical) estimations of the pixel through middle, mean or percentile esteems or different measurements of the time arrangement; for instance 'the median surface reflectance in winter'.
• Time arrangement water quality perception: A critical preferred standpoint of the ODC over customary remote detecting picture examination approaches is that all information focuses are held, giving a more extravagant stream of information.

6. CONCLUSION AND FUTURE WORK

Information Cube frameworks are probably going to end up progressively essential to support the investigation of Earth frameworks science, empowering new worldwide change applications. These frameworks can give the helpful functionalities to investigate and extricate data from the present furthermore, (more essentially) the future long (satellite based) time arrangement that is portrayed by a spatial and fleeting goal never observed. Enormous Data Analytics requires Analysis Prepared Data, and multidimensional information stockrooms turned out to be ready to fulfill such request. Potential clients need to keep decision smoothness and use whatever number Data-Cube benefits as could reasonably be expected to pay little mind to the stage and association that control them.

Future work will incorporate the advancement of Data 3D shape for Indian information, creating calculations that can bolster Indian satellite information. Time arrangement examination of Indian information will turn out to be more successful for investigation and close outcomes on it.

7. REFERENCES

[1] https://www.opendatacube.org/
[2] https://www.tandfonline.com/doi/pdf/10.1080/20964471.2017.1398903?needAccess=true
[3] https://www.tandfonline.com/doi/pdf/10.1080/20964471.2017.1404232?needAccess=true
[4] http://www.copernicus.eu/events/open-source-%E2%80%9Copen-data-cube%E2%80%9D-platf
[5] https://medium.com/opendatacube/what-is-open-data-cube-805af60820d7
[6] http://ceos.org/document_management/Meetings/SIT/SIT-32/Side%20Meeting%20Materials/ODC_WhitePaper_v2a.pdf

