

Acquisition and Analysis of Times Series of Satellite Data in the Cloud – Lessons from the Field

Date: May 11, 2022

Presented by: Marisol Garcia-Reyes (Farallon Institute)

(The slides are available under “Materials from the Webinar” in the above link. The tutorial is available at https://github.com/marisolgr/python_sat_tutorials)

Q. Some of the data collections in the cloud seem to be managed by Farrallon. What are some of the challenges of managing the data in the cloud?

A. Funding to store the data in different formats for accessibility is challenging and will be difficult in the future. Data access is limited because many users prefer, and only know or use, the online servers. This makes it hard to justify exploring new formats that require resources and time.

Q. What about a wishlist for tools? You mentioned many great tools in this domain, but are there any types of tools that are missing and would make this type of work easier?

A. I am still learning all the capabilities of xarray, so still in the honey-moon phase with this tool. But I think a tool that allows standardizing access to data (in different formats or achieving fashion) would be ideal.

Q. Does your python tutorial only relate to earth science?

A. Yes in the sense that only Earth data is exemplified, but No because the tools shown can be customized to other data with no more challenge than would be choosing a different earth dataset.

Q. Once you get the data, what is the biggest challenge, managing the data size, e.g. need enough RAM in the system, or is it fetching and feeding the data to keep the system busy?

A. Manipulating and saving the data is pretty easy since you could do some synthesis before obtaining the data. That's the advantage of using the cloud and xarray/python. The analysis would be as complicated as the user needs them to be.