



The Assimila data cube vision – geospatial data for environmental applications.

Gerardo López Saldaña, Jonathan Styles, and Taylor Day
Assimila Ltd, United Kingdom (gerardo.lopezsaldana@assimila.eu)

Addressing environmental issues at a large scale, e.g. country or continental level, requires the use of a vast amount of Earth Observation data and geospatial ancillary datasets. The challenges when working with EO data can be grouped in three broad categories: access, pre-processing and analysis. The former describes the difficulty associated with initially downloading or getting access to particular datasets, whilst pre-processing challenges capture the complexity of analysis-ready data creation using multiple data sources, each with varying file formats, resolutions and projections. The final challenge relates to a simple way to address this vast wealth of data such that the scientific analysis can take place. Even when all the required data can be stored into a single multidimensional structure, accessing array elements in terms of location and time can be challenging.

The Assimila Data Cube (ADC) provides a solution to these issues. It removes the data download difficulties through the provision of a set of tools that access multiple data portals to either directly download or access data in a cloud environment. The acquired data are then used to create seamless and homogenous analysis-ready datasets using user-specified resolutions, projections and formats. These datasets then support smart spatio-temporal querying using latitude, longitude and time in natural language, as well as traditional indexing using rows, columns and layers.

Facilitating data access will ensure that this data can be better exploited by both advanced and entry level users, particularly if the latter find the initial uptake of EO data to be overwhelming for their environmental applications.

The ADC solution is one approach to improve access the use of geospatial data but there are many more complex issues that the data cube paradigm is facing. We want to contribute to a fruitful discussion between all actors involved to truly facilitate the exploitation of EO data to provide key information to a multitude of environmental applications.