Accessing Raster Data organized in Data Cubes via R using a Shiny App and the CubeR Package

Mattia Rossi (1,2), Daniel Frisinghelli (2), and Alexander Jacob (2)

(1) Free University of Bolzano, Faculty of Science and Technology, Italy (mattia.rossi@eurac.edu), (2) Institute for Earth Observation, Eurac Research, Bolzano, Italy

The emerging topics of big data processing, machine- or deep learning techniques as well as time-series analysis require handling many datasets. Access, storage and processing often depend on the preferences of the user (or user group) and therefore lack in standardization. This leads to very time and resource consuming processes and reduces the reproducibility of a research. In geosciences, this problem comes up when handling raster datasets e.g. from remote sensing platforms. Often search, storage, denomination and pre-processing steps for each dataset are repeatedly done varying from one user to the other. Geo-spatial-temporal data where the x- and y-axes represent the coordinate system and the z-axis the time of acquisition, recently has also often been referred to as data cubes, although there is no strict or standardized definition of this term. The Open Geospatial Consortium (OGC) has also defined a standard for representing this data in so called coverages via the Coverage Implementation Schema (CIS) and for accessing those coverages the Web Coverage Service (WCS). For processing the latter has been extended with the Web Coverage Processing Service (WCPS), which allows for mathematical operations being performed on the data directly in the database, potentially decreasing IO requirements, since only results of those operations need to be transferred.

Following those standards, we developed both a Shiny frontend and the CubeR package for easy access of data cubes fed with Sentinel 1 and 2 raster images. Both have implemented functions that allow direct communication with the data cubes by generating WCPS queries and handle the response automatically. The CubeR package and the Shiny frontend provide three central functionalities: (i) return properties, (ii) return metrics, (iii) perform mathematical operations on the data cubes. Functions returning properties deliver information of the data stored in the cube such as the coverage, coordinate system, resolution etc. The metrics comprise the value extraction of a single point, a pixel history showing the changes on this pixel over time and the download of raster image(s) delimited by a bounding box. Mathematical operations such as the computation of the normalized difference index are also possible with every band combination. The benefit of implementing a Shiny app is that the properties can be immediately seen and that the data in a data cube can be visually explored. The package is useful for uncomplicated extraction of the metrics needed from a data cube.

For development and testing the Eurac Research Sentinel Alpine Observatory infrastructure was used, where Rasdaman and its extension Petascope have been set up to expose data organized in multi-dimensional arrays via OGC web services. Those multidimensional arrays allow for very efficient data access and processing of geo-spatial-temporal datasets, when configured according to the tasks foreseen.