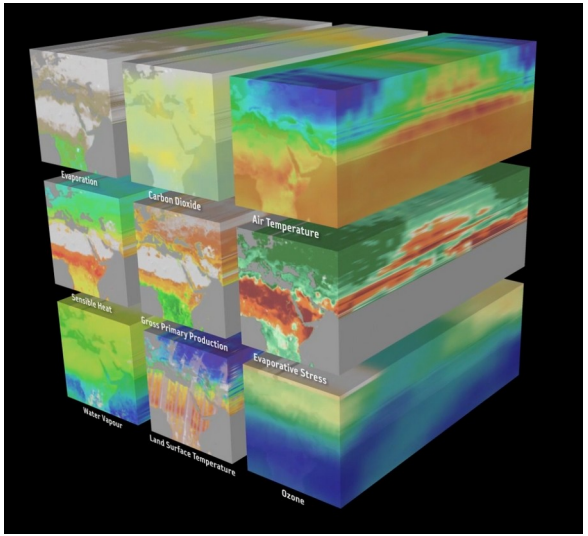
Summer of Code: Minicuber

Selene - Aug. 2nd 2023

1. Minicubes

What?

2D arrays (images, rasters...) “stacked” in time



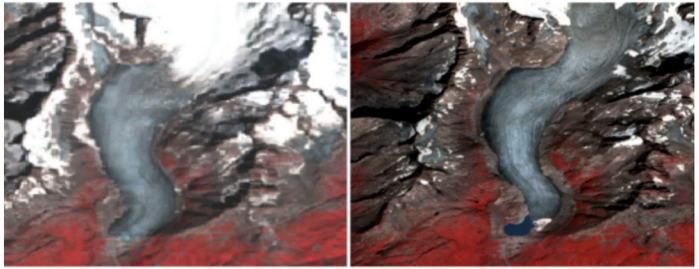
1. Minicubes

Why?

- We are interested in spatial AND temporal information.
- A format that is increasingly being used in AI with EO data (ex: Swiss Data Cube)
- Can be deployed on different spatial scales.

The Swiss Data Cube – Insights for action

The SDC will **improve our understanding of Switzerland's changing landscape**, providing much needed insights, knowledge and analysis for more informed, strategic and inclusive decision making across the country. This information will b**enefit policy makers and public officials**, enabling them to make better decisions, and will **increase commercial efficiency and economic growth** for businesses and entrepreneurs across the country.



The Rhône Glacier in 1985 (left – Landsat) and 2020 (right – Sentinel-2)

1. Minicubes

Challenges

- Spatial resolution
- Temporal resolution
- Coordinate systems
- Missing data
- Clouds
- ...

1. Minicubes

Currently:

- There exists “minicuber” python packages, but mainly for Sentinel-2

My contribution:

- Integrating ERA5 reanalysis data
- Integrating custom data

2. Creating minicubes - overview

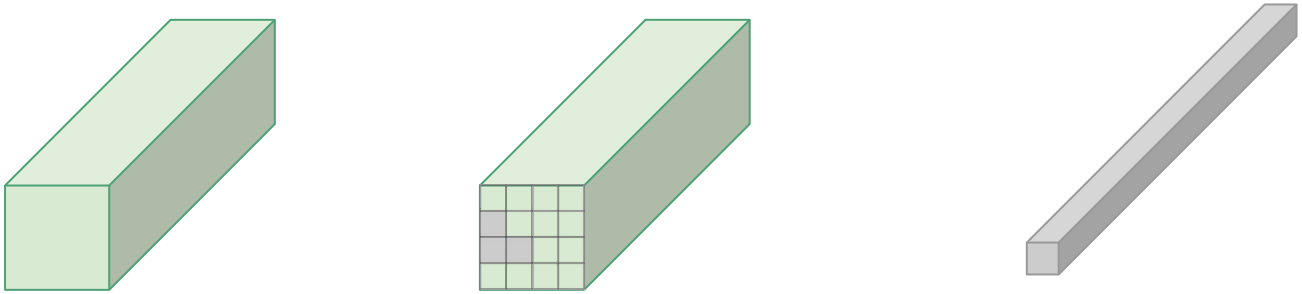
Specify parameters (dates, location, bands, resolution, aggregation…)

Retrieve data via AWS (filter for time and place)

Post-processing (match pixels, aggregate data…)

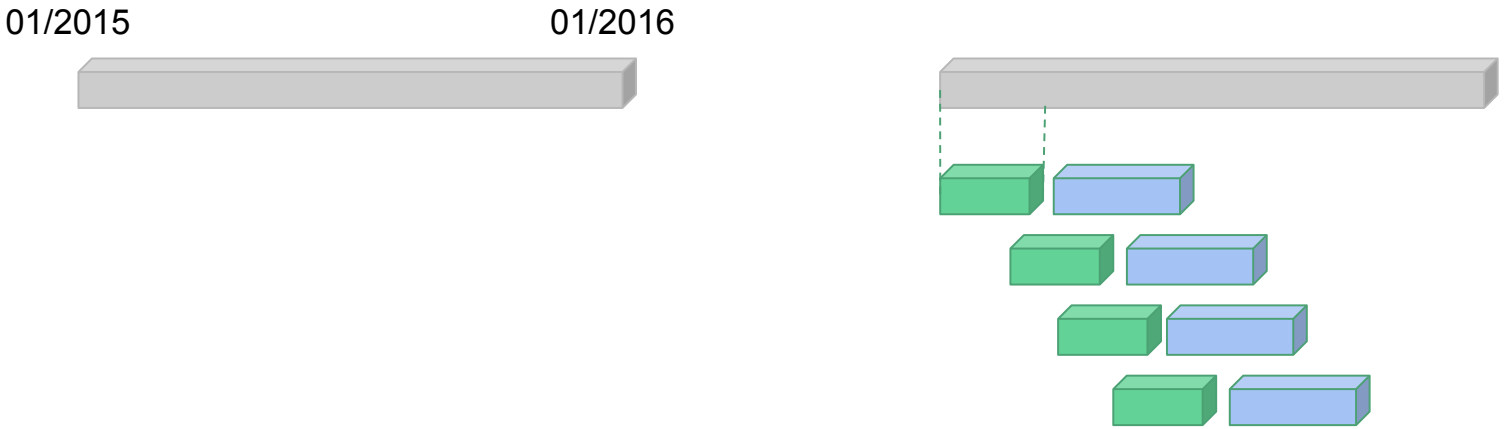
3. Pixel timeseries

Cube timeseries + masks → pixel



3. Pixel timeseries

pixel timeseries → pixel timeseries



3. Generating pixel datasets

1. Download cube(s)
2. Add your data
3. Perform cloud removal
4. Deal with NaNs/missing data
5. Extract pixels sub-timeseries

3. Generating pixel datasets

Clouds block the signal of the band → values of pixel are at 0

Removal strategy:

- Remove lower 5% per week of year
- If not too many missing values, interpolate & smooth
  - Linear interpolation
  - Then LOESS smoothing