Swiss Data Cube & SDG – Hackathon (11-13 April 2018)

Organizers: UN Environment/GIRD-Geneva & University of Geneva/ISE

The Swiss Data Cube

Pressures on natural resources are increasing and a number of challenges need to be overcome to meet the needs of a growing population in a period of environmental variability. Some of these environmental issues can be monitored using remotely-sensed Earth Observations (EO) data that are increasingly available from a number of freely and openly accessible repositories. However, the full information potential of EO data has not been yet realized. They remain still underutilized mainly because of their complexity, increasing volume, and the lack of efficient processing capabilities.

With each passing year, new generations of EO satellites are creating increasingly significant volumes of data with such comprehensive global coverage that for many applications, the lack of data is no longer a limiting factor. Extensive research and development activity has delivered new data applications that offer significant potential to deliver great impact to important environmental, economic and social challenges, including at the local, regional and global scales. Such applications highlight the value of EO, though the challenge is in providing the proper connections between data, applications and users. Even today, much archived EO satellite data is underutilized despite modern computing and analysis infrastructures.

EO Data Cubes (DC) are a new paradigm aiming to realize the full potential of EO data by lowering the barriers caused by these Big data challenges and providing access to large spatio-temporal data in an analysis ready form.

The Swiss Data Cube contains 33 years of Landsat 5,7,8 (1984-2017) and 2.5 years (2015-2017) of Sentinel-2 data. A prototype platform is running and allows testing and visualizing several algorithms.

The SDC is powered by the Open Data Cube architecture. This is an open-source analytical framework developed by GeoScience Australia & NASA that helps users organize and analyze large, standardized satellite data collections.

The main objectives of the Swiss Data Cube (SDC) is to support the Swiss government and the Cantons for environmental monitoring and reporting and enable Swiss scientific institutions (e.g., Universities) to facilitate new insights and research using the SDC and to improve the knowledge on the Swiss environment using EO data. In particular it can help: (1) Improving efficiency and effectiveness of government investment and improve management of natural resources; (2) Supporting innovation and growth in the digital economy era; and (3) Generating new information products and services that transforms everyday life.

The Sustainable Development Goals

The Sustainable Development Goals (SDGs) are a collection of 17 global goals set by the United Nations. The broad goals are interrelated though each has its own targets to achieve. The total number of targets is 169. The SDGs cover a broad range of social and economic development issues. These include poverty, hunger, health, education, climate change, gender equality, water, sanitation, energy, environment and social justice. The SDGs are also known as "Transforming our World: the 2030 Agenda for Sustainable Development" or Agenda 2030 in short. The goals were developed to replace the Millennium Development

Goals (MDGs) which ended in 2015. Unlike the MDGs, the SDG framework does not distinguish between "developed" and "developing" nations. Instead, the goals apply to all countries.

Earth observations, geospatial data, and derived information play insightful roles in monitoring targets, planning, tracking progress, and helping nations and stakeholders make informed decisions, plans, and on-going adjustments that will contribute toward achieving the SDGs. Combined with demographic and statistical data, these sources enable nations to analyze and model conditions, create maps and other visualizations, evaluate impacts across sectors and regions, monitor change over time in a consistent and standardized manner, and improve accountability.

The Challenge

Switzerland is acknowledged as the water reservoir of Europe. While its territory represents four thousandths of the continent's total area, 262 billion m^3 of water, or 6% of Europe's freshwater reserves, are stored in Switzerland. 57% of those reserves are stored in natural lakes.

The SDG 6 is aiming to *"Ensure availability and sustainable management of water and sanitation for all"* because access to safe water and sanitation and sound management of freshwater ecosystems are essential to human health and to environmental sustainability and economic prosperity. One of the target (6.3) clear states that by 2030 water quality should be improved by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials. This objective can be monitored using the indicator 6.3.2:
Proportion of bodies of water with good ambient water quality.

This indicator can be measured using remotely-sensed Earth Observations data and therefore thanks to its large spatio-temporal database the Swiss Data Cube can be extremely useful to follow progresses towards the objective of SDG6.

The Objective

The general objective of this hackathon will be to develop a prototype application for monitoring Sustainable Development Goals (SDGs) at the country scale using EO data stored in the SDC.

A team from NASA (who is leading the development of the Data Cube software) will be there to support participants and help them in developing the prototype.

More specifically, our aim is to develop an interactive, web-based cartographic graphical user interface allowing users to track the water quality (Total Suspend Matter & Chlorophyll contents) of the 20 largest lakes of Switzerland
(https://fr.wikipedia.org/wiki/Liste_des_lacs_de_Suisse)

The GUI should enable users to:

- Select one or more lakes
- Launch a water quality algorithm according to some requirements
- Visualize results under the form of maps and/or graphs
- Query each pixel and see the evolution through time.

This application can be inspired from the Digital Earth Australia portal:
http://eos.ga.gov.au
especially for time-series analysis.

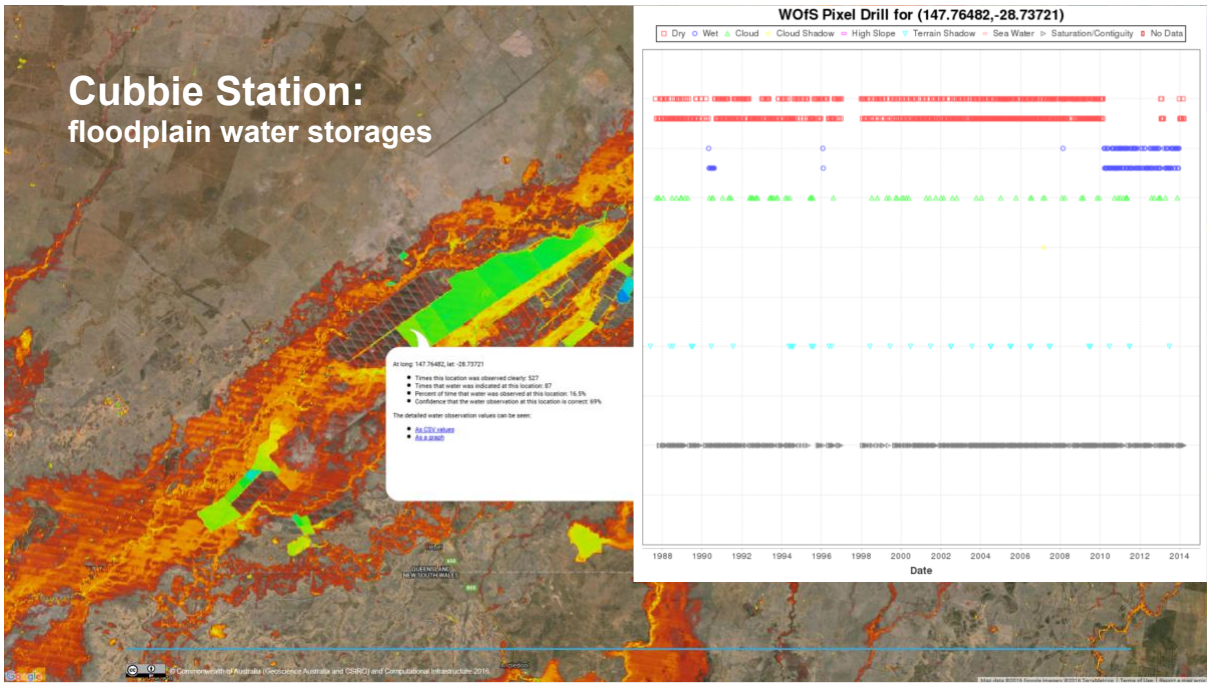

Figure 1: Query result on a specific pixel & related time-series graph

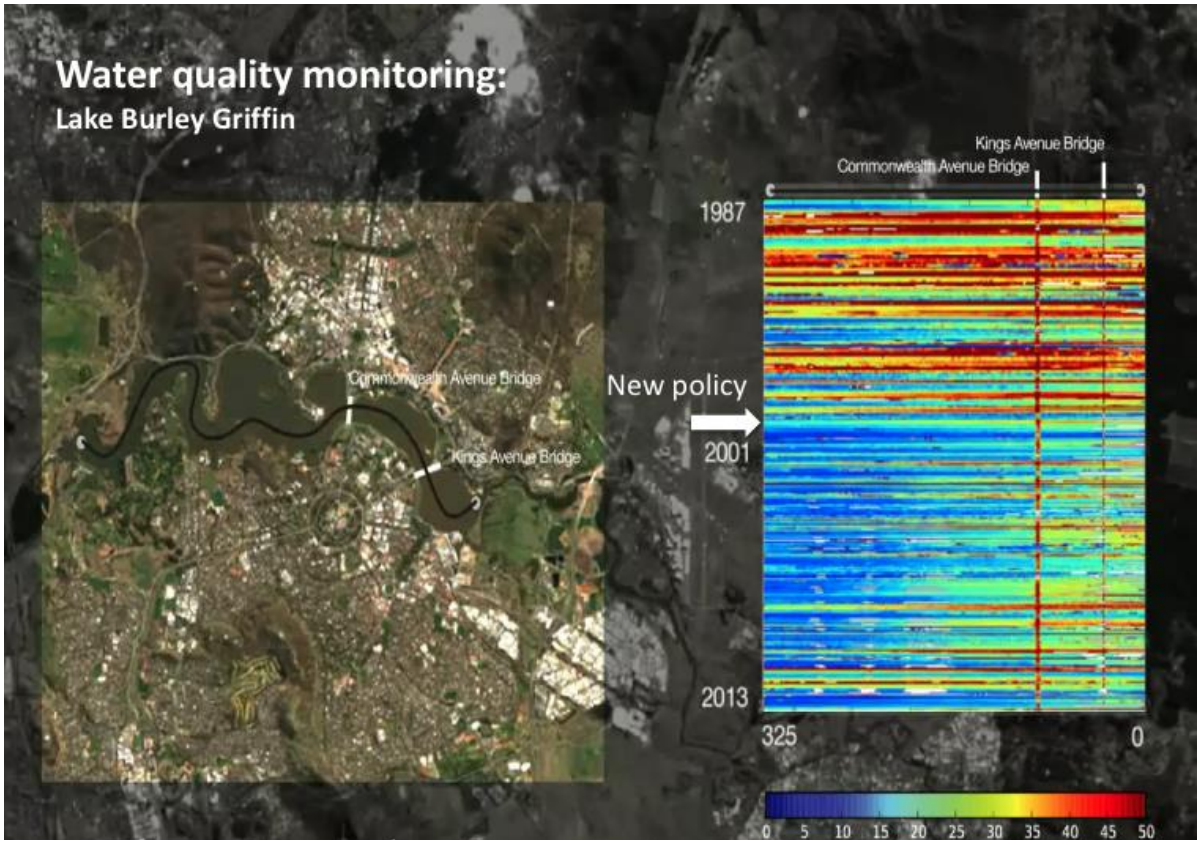

Figure 2: Transect and spatio-temportal analysis of TSM in a lake

You can be inspired by the following report:
https://www.earthobservations.org/documents/publications/201703_geo_eo_for_2030_agenda.pdf
(pp.9-10) where the Australians have used their Data Cube to assess the water quality using the Total Suspended Mater Algorithm that is also available in the Swiss Data Cube.

The following useful resource on the UN-SPIDER Knowledge Portal can give you a good view on water quality monitoring using Earth Observations:
http://un-spider.org/links-and-resources/data-sources/daotm-water-quality?utm_content=buffer1104e&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

Data

You will use a part of the Landsat 5,7,8 data stored in the Swiss Data Cube.
Landsat data are at 30m resolution, 11 bands, and available from 1984 to 2017.
More information at: https://landsat.usgs.gov/

(Suggested) technical solutions to use

- Swiss Data Cube: http://www.swissdatacube.org
- Python & Jupyter notebook to interact with the SDC API
- GeoServer (http://geoserver.org) for publishing results with OGC standards
- OpenLayers (http://openlayers.org), Leaflet (http://leafletjs.com), D3.js (https://d3js.org) or any other JavaScript library for developing the graphical user interface (Bootstrap, …)

Useful links

- Swiss Data Cube: http://www.swissdatacube.org
- Open Data Cube software: https://www.opendatacube.org
- Open Data Cube GitHub: https://github.com/opendatacube/datacube-core
- Jupyter notebook recipes: https://github.com/opendatacube/datacube-core/tree/develop/examples/notebooks
- Zonal Stats Notebook: https://github.com/opendatacube/datacube-core/blob/develop/contrib/notebooks/zonal-stats-example.ipynb
- Sustainable Development Goals: https://sustainabledevelopment.un.org
- SDG 6: https://sustainabledevelopment.un.org/sdg6
- Indicator 6.3.2: http://www.sdg6monitoring.org/indicators/target-63/indicators632/

Further readings

Anderson, K., Ryan, B., Sonntag, W., Kavvada, A., & Friedl, L. (2017). Earth observation in service of the 2030 agenda for sustainable development. *Geo-spatial Information Science, 20*, 77–96.
http://www.tandfonline.com/doi/full/10.1080/10095020.2017.1333230

Evans, B., Wyborn, L., Pugh, T., Allen, C., Antony, J., Gohar, K., ... Bell, G. (2015). The NCI high performance computing and high performance data platform to support the analysis of petascale environmental data collections. In Ralf Denzer, Robert M Argent, Gerald Schimak, & Jiří Hřebíček (Eds.), *Environmental software systems. Infrastructures, services*

and applications: 11th IFIP WG 5.11 International Symposium, ISESS 2015 (pp. 569–577), Melbourne, March 25–27, 2015.
https://hal.inria.fr/hal-01328609

Lewis, A., Lymburner, L., Purss, M. B. J., Brooke, B., Evans, B., Ip, A., Oliver, S. (2016). Rapid, high- resolution detection of environmental change over continental scales from satellite data – The Earth Observation Data Cube. *International Journal of Digital Earth, 9*(1), 106–111.
http://www.tandfonline.com/doi/full/10.1080/17538947.2015.1111952

Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., ... Wu, W. (2017). The Australian Geoscience Data Cube – Foundations and Lessons Learned. *Remote Sensing of Environment,*.
https://www.sciencedirect.com/science/article/pii/S0034425717301086

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., ... Ip, A. (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment, 174*(March), 341–352.
https://www.sciencedirect.com/science/article/pii/S0034425715301929

Purss, M. B. J., Lewis, A., Oliver, S., Ip, A., Sixsmith, J., Evans, B., ... Chan, T. (2015). Unlocking the Australian Landsat archive – From dark data to high performance data infrastructures. *GeoResJ, Rescuing Legacy Data for Future Science, 6*(June), 135–140.
https://www.sciencedirect.com/science/article/pii/S2214242815000182

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment, 122*, 2–10.
https://www.sciencedirect.com/science/article/pii/S003442571200034X

Dhu, T., Dunn B., Lewis B., Lymburner L., Mueller N., Telfer E., McIntyre A., Minchin S., Philipps C. (2017) Digital earth Australia – unlocking new value from earth observation data. *Big Earth Data, 1(1-2)*
http://www.tandfonline.com/doi/full/10.1080/20964471.2017.1402490?scroll=top&needAccess=true

Nativi S., Mazzetti P., Craglia M. (2017) A view-based model of data-cube to support big earth data systems interoperability. *Big Earth Data, 1(1-2)*
http://www.tandfonline.com/doi/full/10.1080/20964471.2017.1404232

Giuliani G., Chatenoux B., De Bono A., Rodila D., Richard J.-P., Allenbach K., Dao H., Peduzzi P. (2017) Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data, 1(1-2)*
http://www.tandfonline.com/doi/full/10.1080/20964471.2017.1398903

Information

Venue: Centre Universitaire d'Informatique, Battelle – building A, 7 route de Drize, 1227 Carouge. http://cui.unige.ch/en/contact/planaccessglobal/
Date: 11-13 April 2018
Material:

- Bring your own laptop
- Wi-Fi access will be provided
- Beamer
- Screen
- A Virtual Machine VM with a part of the SDC will be provided
- GitHub repository

Objectives & Expected outcomes:

- Deliver a prototype for an innovative & user-friendly solution for monitoring an SDG using EO data
- Select and implement a simple algorithm to analyze Chlorophyll in Lakes
- Generate time-series analysis for Total Suspend Matter (TSM) and Chlorophyll in lakes
- Develop a nice interactive web-interface to visualize and query the results from the previous analysis
- Test the approach on a selected area & if time allows scale-up to the 20 largest lakes in Switzerland.

Principles/Rules:

- We are here in a collaborative spirit
- Learn from each other
- Have Fun
- Short development cycle (60hours)
- Provide insights on user needs for the SDC
- Scalability/Sustainability of the solution
- Link with the SDGs
- Be sure that they understand the science behind
- Advance the vision of maximizing the societal impact of EO
- Agility rather perfection is key to the success
- Each sub-team should write every day a short blog report on:
    - What did you get done today?
    - What do you plan on doing next day?
    - Are you blocked on anything?

Team composition & tasks

Algorithm team

- Identify in the literature a simple and efficient algorithm/index to derive Chlorophyll concentration in lakes
- Provide a comparison matrix of different candidate algorithm
- Select an algorithm
- Implement the selected algorithm

- Test the implemented algorithm

Analysis team

- Run a spatio-temporal analysis (according different sets of requirements) of TSM in a given lake
- Same with Chlorophyll (once the algorithm is available)
- Deliver results to the Web UI team in a useful format for visualization.

Web UI team

- Design a light-responsive web-based interactive application (mock-up)
- Select the required JavaScript libraries (front-end; map & graphs; …)
- Implement a prototype of the Web UI

Organization

Day 1: Wednesday April 11

- Keynote by Dr. Brian Killough (NASA) on the Open Data Cube Imitative
- Presentation by Dr. Gregory Giuliani (UNIGE) on the proposed challenge
- Quick presentation (NASA developers) on how to develop/interact with the SDC – Python API/Jupyter Notebook
- Set-up of sub-task team
- Start to hack
- End of the day: wrap-up session

Day 2: Thursday April 12

- Short session on the objectives of the day
- Continue to hack
- End of the day: wrap-up session

Day 3: Friday April 13

- Short session on the objectives of the day
- Continue to hack
- End of the day: final wrap-up session & follow-up