Copernicus - eoSC AnaLytics Engine

Christian Briese, EODC

christian.briese@eodc.eu

OpenAIRE-Nexus Launch Event | 10.03.2021

Copernicus | Problem Statement

- EU Copernicus programme has established itself globally as the predominant spatial data provider.

- There is no single European processing back-end that serves (or will any time soon serve) all datasets of interest to the scientific community.
  - hampers the integration of these data sources in science and monitoring applications, particularly when expanding the applications to regional, continental and global scales.

- Big (Copernicus) Data Analytics require a federated infrastructure with a core cloud computing and storage architecture optimised for very large data handling and fast user query response.

Copernicus – eoSC AnaLytics Engine

C-SCALE

Vision

To empower European researchers, institutions and initiatives to easily discover, access, process, analyse and share Copernicus data, tools, resources and services in a way that can be seamlessly integrated into their processes and research practices.

Mission

The C-SCALE (Copernicus eoSC AnaLytics Engine) project will deliver a pan-European federated data and computing infrastructure for Copernicus to the EOSC Portal. The open platform will integrate cross-disciplinary EOSC services, ensuring interoperability between distributed data catalogs, computational tools and infrastructure. In doing so, the federation will increase the service offer of the EOSC Portal providing state-of-the-art research enabling services to its users.

High level objectives

- The project will **integrate EOSC services to federate pan-European computing and data resources for Copernicus**, including resources from, inter alia, the Copernicus DIAS, the Collaborative Ground Segments and the European Open Science Cloud.

- **Users will contribute to the functional design** of the federation by deploying mature services (TRL>8) and providing feedback on the user experience via the C-SCALE User Forum.

- The project will **provide an architecture blueprint and onboarding process**, including training material, **to scale the federation beyond the project partners.**

The C-SCALE ambition

Federate European EO (DIASes, CollG nodes, etc.) **and e-Infrastructure services** to support Copernicus research and operations.

The C-SCALE consortium brings together expertise from:

EO sector:

EODC, Deltares, VITO, CloudFerro, TU Wien, CESNET, GRNET

e-Infrastructure:

EGI, CESNET, INFN, SURFsara, GRNET, INCD

C-SCALE services in EOSC

Suite of three services

- C-SCALE EO **Data** Archive
  - provide access to a large C-SCALE EO data archive
- C-SCALE **Compute** services
  - compute services integrated with the C-SCALE EO Data archive
- C-SCALE EO **Analytics** platforms
  - set of analytics platforms that can be deployed on top of the C-SCALE EO data archive and compute services

C-SCALE federation services will be available in the EOSC Portal around mid 2022.

Use cases

#	Use Case	Responsibility
1	Aquamonitor	Deltares
2	WaterWatch	Deltares
3	HiSea	Deltares
4	High-resolution Land Surface (Drought) Analysis	Deltares
5	RETURN – monitoring tropical forest recovery capacity using satellite data	WUR
6	Virtual European Sentinel Data Cubes	TUW

Role:
- Validate and optimise C-SCALE

Benefit for use cases:
- Cloud agnostic
- Independence from commercial, closed, non-EU providers
- Cross- / inter-disciplinary exposure
- FAIRer

Use case #1: Aquamonitor

C-SCALE

Objectives

- Mapping of surface water changes globally

- Web app to explore surface water mask and changes

- Replicate functionality implemented using Google Earth Engine using EOSC infrastructure (or use hybrid approach)


https://aqua-monitor.appspot.com/
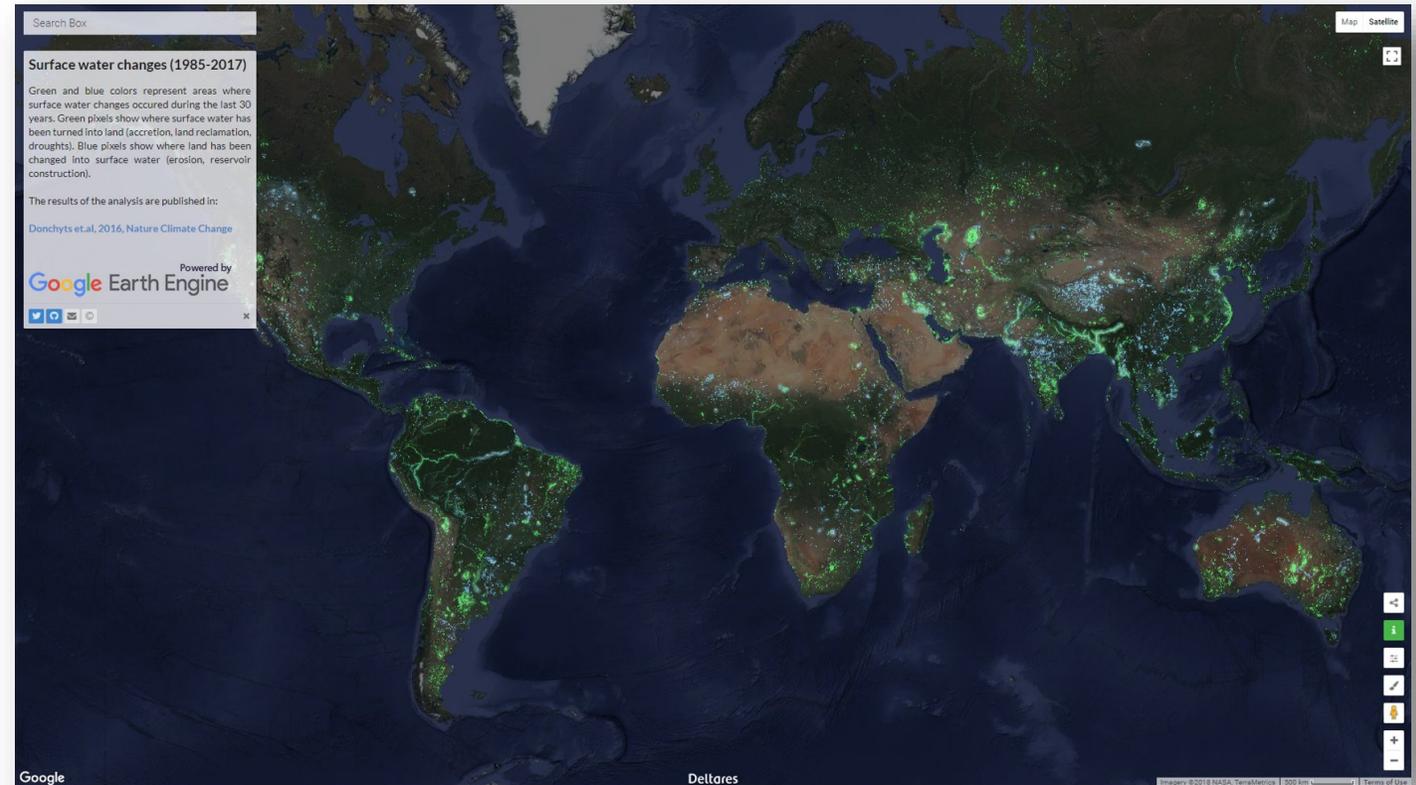
Use case #2: WaterWatch

Near-real-time updates of the availability of surface water resources foor 100k – 1M reservoirs (lakes, rivers, wetlands)
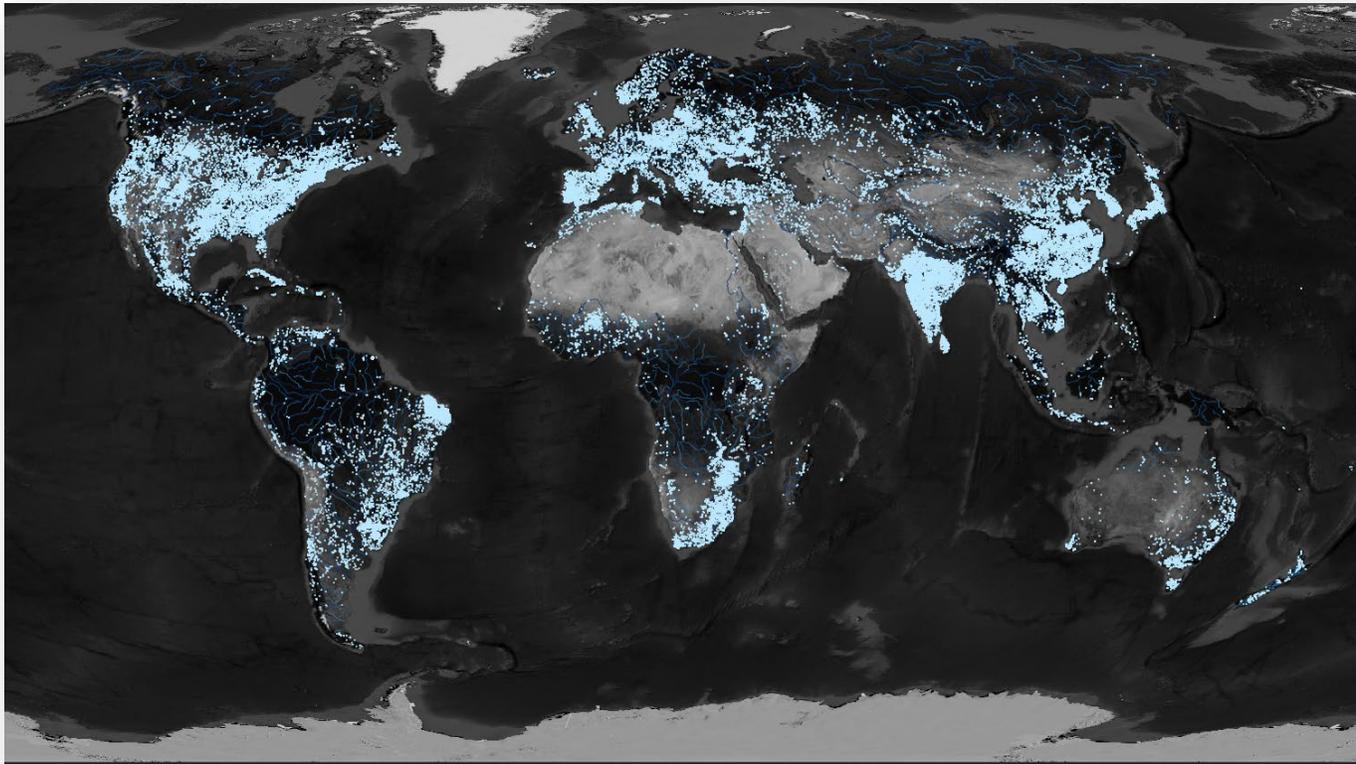
Users: water managers, agriculture, hydropower, droughts/floods, insurances

Use case #3: HiSea

C-SCALE
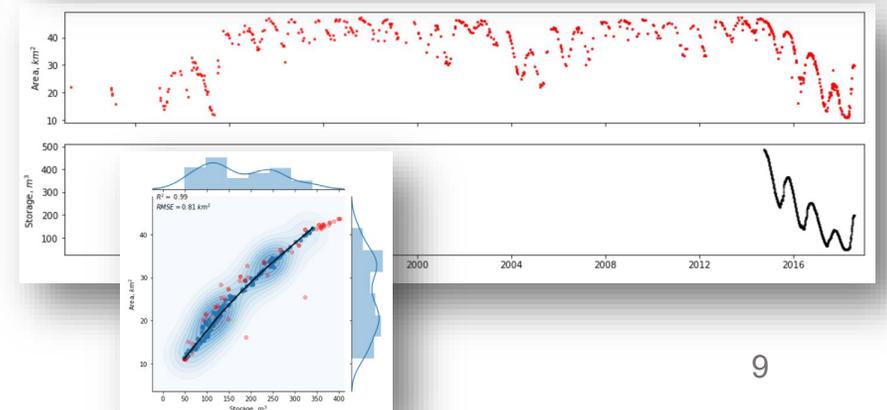
Deliver accurate and reliable information, readily available, easily understandable and with high resolution

HiSea solution

- Co-designed with users
- Provides high resolution data of water quality at sea
- Develops operational Copernicus-based downstream information services
- Improves operation, planning and management of marine activities

HiSea Users

- Targets port and the aquaculture sectors

Further use cases

C-SCALE

- Use case #4: High Resolution Land Surface (Drought) Analysis (HiResLSDA)

- Use case #5: RETURN (Monitoring tropical forest recovery capacity using Multispectral and Radar satellite data)

- Use case #6: Virtual European Sentinel Data Cubes

C-SCALE Virtual Access resources

Installations (short name)	Service Category	Unit of access	Units
EODC EO Storage	Storage Service	TB year	150
EODC EO Cloud	Cloud Service	CPU year	620
EODC HPC	HPC/HTC Service	CPU Hour	600,000
CESNET MetaCentrumCloud - CPU	Cloud Service	CPU Month	2,850
INFN Bari Storage	Storage Service	TB Month	1,500
INFN Bari CPU	Cloud Service	CPU Month	3,840
SURFsara dCache Fro			
SURFsara Spider (HTC			
SURFsara Data Proce			00
SURFsara MS4 Storag			
SURFsara MS4 SSD St			
SURFsara MS4 Compu			
VITO CVB (Storage)			
VITO CVB (Compute)			
GRNET KNS-Storage (			
GRNET KNC Cloud			,00
GRNET ARIS-Storage HPC	Storage Service	TB Hour	219,000
GRNET ARIS-Compute HPC	HPC/HTC Service	CPU Hour	1,000,000
CREODIAS – Storage	Storage Service	TB Month	2,762
CREODIAS - Compute	Cloud Service	vCore hour	1,500,000
CREODIAS – GPU	GPU service	GPU hour	6,000
INCD Lisbon (NCG) (Storage)	Storage Service	TB Month	450
INCD Lisbon (NCG) (Compute)	Cloud Service	CPU Day	4500

Total capacity of resources offered through VA:

- Cloud: 1648 CPU core/year

- HTC/HPC: 370 CPU core/year

- Storage: 1104 TB/Year;

Conclusion

- C-SCALE puts together EO and the e-infrastructure partners to:
    - Facilitate the exploitation of Copernicus data leveraging on large resources and advanced technologies from e-infras and EOSC
    - Make Copernicus resources easily accessible to new research areas and EOSC in general
- C-SCALE will deliver a federation of EO and e-infrastructure services and resources
    - Create a very large distributed repository of EO data close to compute resources
    - EO data will be FAIR through the federation
    - Federation services accessible through the EOSC Portal
- C-SCALE federation will be co-designed with researchers
- Large amount of resources available through the Virtual Access mechanism

Thank you for your attention.