EUMETSAT's Drought & Vegetation Data Cube

A prototype

Christine Träger-Chatterjee[1], Daniel Lee[1], Mike Grant[1],

Maurizio Bottaccio[2], Carla Sofia Barroso[3], David Fairbairn[4],

Joerg Trentmann[5], Markus Ziese[5], Mark Higgins[1]

[1] EUMETSAT; [2] B-Open Solutions;

[3] Portuguese Institute for Sea and Atmosphere; [4] ECMWF;

[5] Deutscher Wetterdienst

Why a Data Cube?
What is a Data Cube?
Lessons learnt from building the D&V Cube
User Feedback on the D&V Cube

EUMETSAT

Why a data cube?

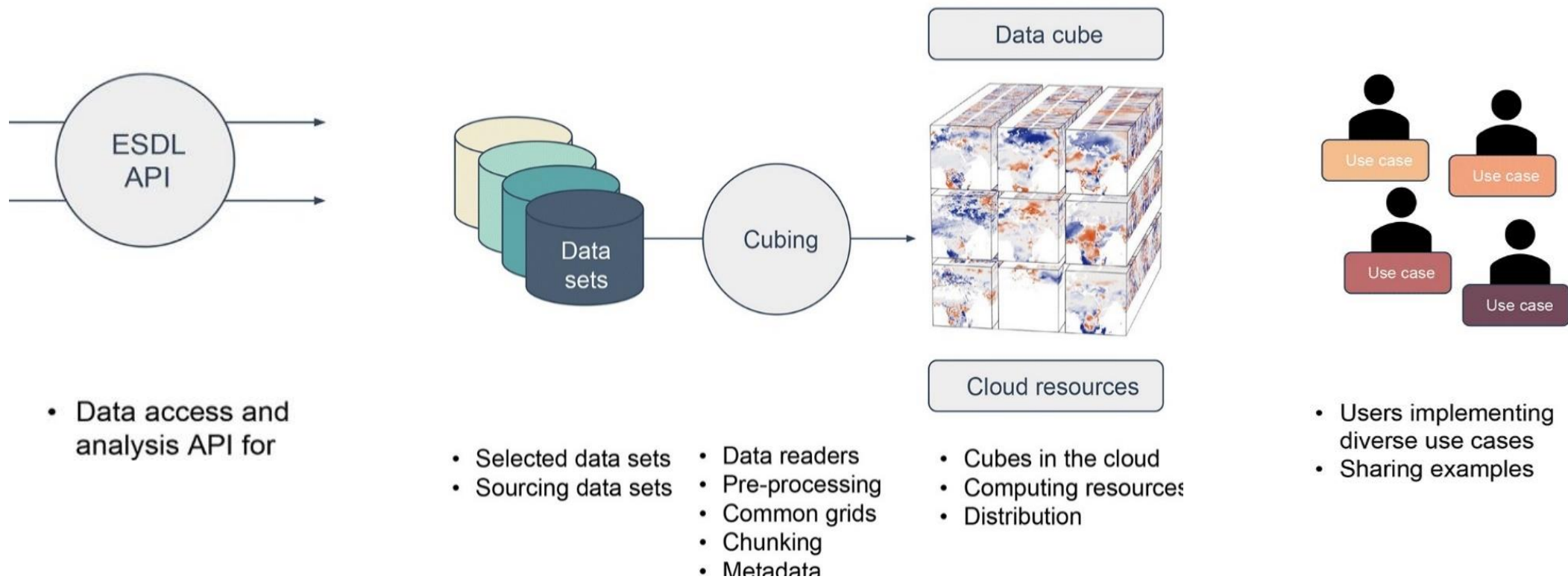## Satellite Application Facilities

- EUMETSAT has a network of different Satellite Application Facilities (SAFs)

- Each SAF is a consortium of entities from EUMETSAT member states

- SAFs are dedicated centers of excellence for processing satellite data

  - research, development and operational activities

  - each SAF is specialized in a different applications

- SAFs provide 80% of EUMETSAT's product portfolio

- Many (climate) data records ➔ real treasure

# Status

The idea of Data Cubes

ESDL API

- Data access and analysis API for

Data sets → Cubing → Data cube / Cloud resources → Use cases

- Selected data sets
- Sourcing data sets

- Data readers
- Pre-processing
- Common grids
- Chunking
- Metadata

- Cubes in the cloud
- Computing resources
- Distribution

- Users implementing diverse use cases
- Sharing examples

Earth system data cubes unravel global multivariate dynamics Miguel D. Mahecha et al.; https://doi.org/10.5194/esd-11-201-2020

Facilitating data access and data handling saves time, which can be used to answer scientific questions.

EUMETSAT

Idea / Vision
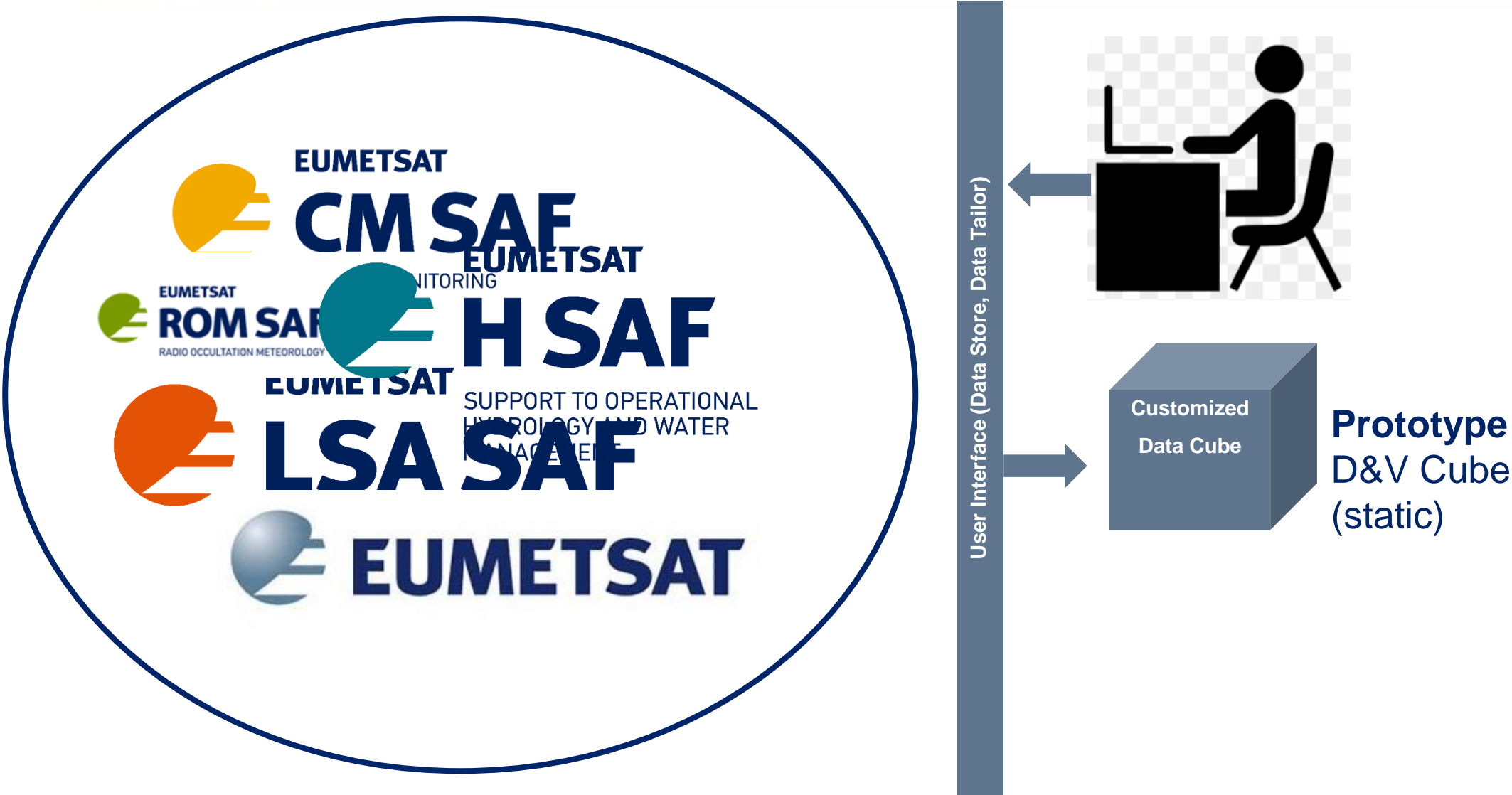
Prototype D&V Cube at glance

Variable		Datasource	Temporal Coverage	Spatial Resolution	Temporal Resolution
Global Radiation		Satellite	Jan 1983 – Dec 2020	0.05°	Daily & monthly
Direct normal Solar Radiation		Satellite	Jan 1983 – Dec 2020	0.05°	Daily & monthly
Sunshine Duration		Satellite	Jan 1983 – Dec 2020	0.05°	Daily & monthly
Land Surface Temperature		Satellite	Jan 2004 – Dec 2020	0.05°	Hourly
Reference Evapotranspiration		Satellite	Jan 2004 – Dec 2020	0.05°	Daily
NDVI		Satellite	Mar 2007 – Dec 2020	0.01°	10-daily
Fractional Vegetation Cover		Satellite	Jan 2004 – Dec 2020	0.05°	Daily
Leaf Area Index		Satellite	Jan 2004 – Dec 2020	0.05°	Daily
Fraction of absorbed photosynthetically active radiation		Satellite	Jan 2004 – Dec 2020	0.05°	Daily
Soil Wetness Index (root zone)		Satellite	Jan 1992 – Dec 2020	0.1°	Daily
Precipitation		In situ	Jan 1982 – Dec 2020	1°	Monthly
T2m		Re-analysis	Jan 1979 – Oct 2020	0.1°	Monthly

Spatial Coverage:	Europe
Grid:	Regular Lat / Lon
Data Format:	CF compliant netCDF4

Users are able to

- access via THREDDS

- download the entire cube (3TB) or subsets / subsamples

- analyse the data on the VM using the CM SAF R Toolbox (and other Tools, e.g. Jupyter Notebooks)
→ work in progress

11 May 2021
Information Day

**May 2021 – August 2021**

Exploration Phase

SALGEE meeting November 2021
D&V session during
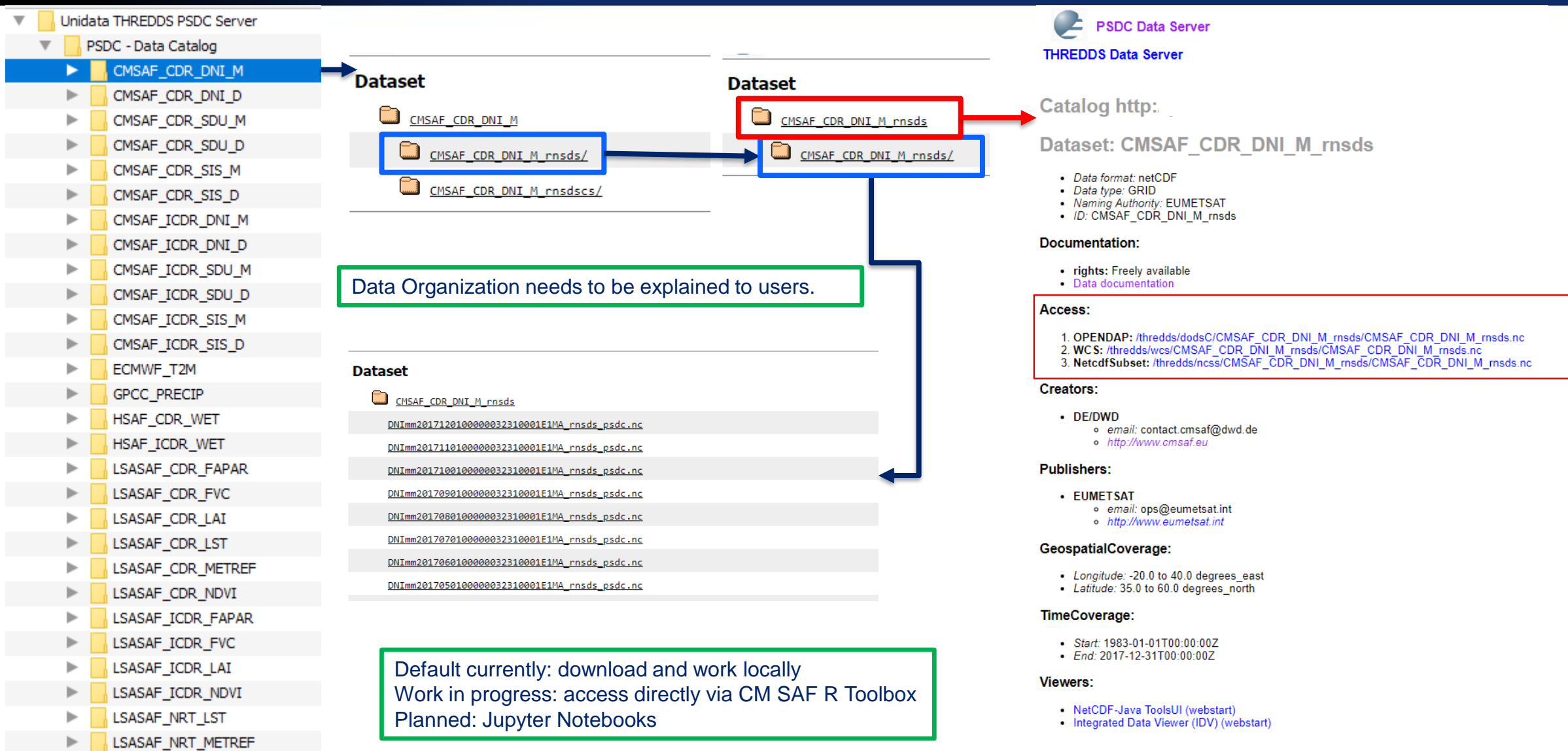
Creating the cube

- Collecting the data from different providers

- Harmonize grid → re-gridding
  - Different methods available
  - Changing the data!
  - Impact on accuracy
  - Error measures and flags!!

- Harmonize metadata
  - Not trivial
  - cf-conventions helpful, but additional standard names needed

- Harmonize file format
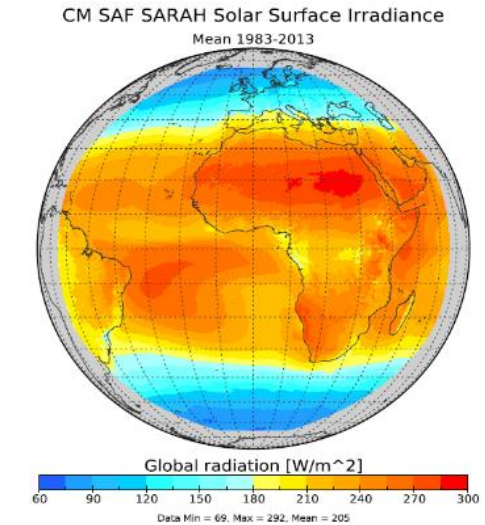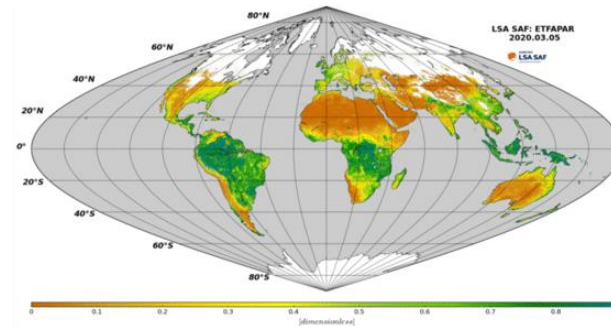
- Issues to consider:
  - Data are changed while re-gridding
    - Documentation not strictly applicable
    - Who is the provider of the data? EUMETSAT or the SAF?
   - ...

Accessing the D&V Cube

Data Organization needs to be explained to users.

Default currently: download and work locally
Work in progress: access directly via CM SAF R Toolbox
Planned: Jupyter Notebooks

The Exploration Phase

- [Cube Information Page](#)
- Information Day: well attended
- Live-Touch Points
- Exchange with other Cube Explorers
- Get support from data experts

EUMETSAT

Example Application: Downscaling Soil Moisture

Agricultural and hydrological applications would benefit from SM with a sub-kilometer spatial resolution and daily revisit time

Sabaghy et al. 2018 (*Remote Sens. Environ*)
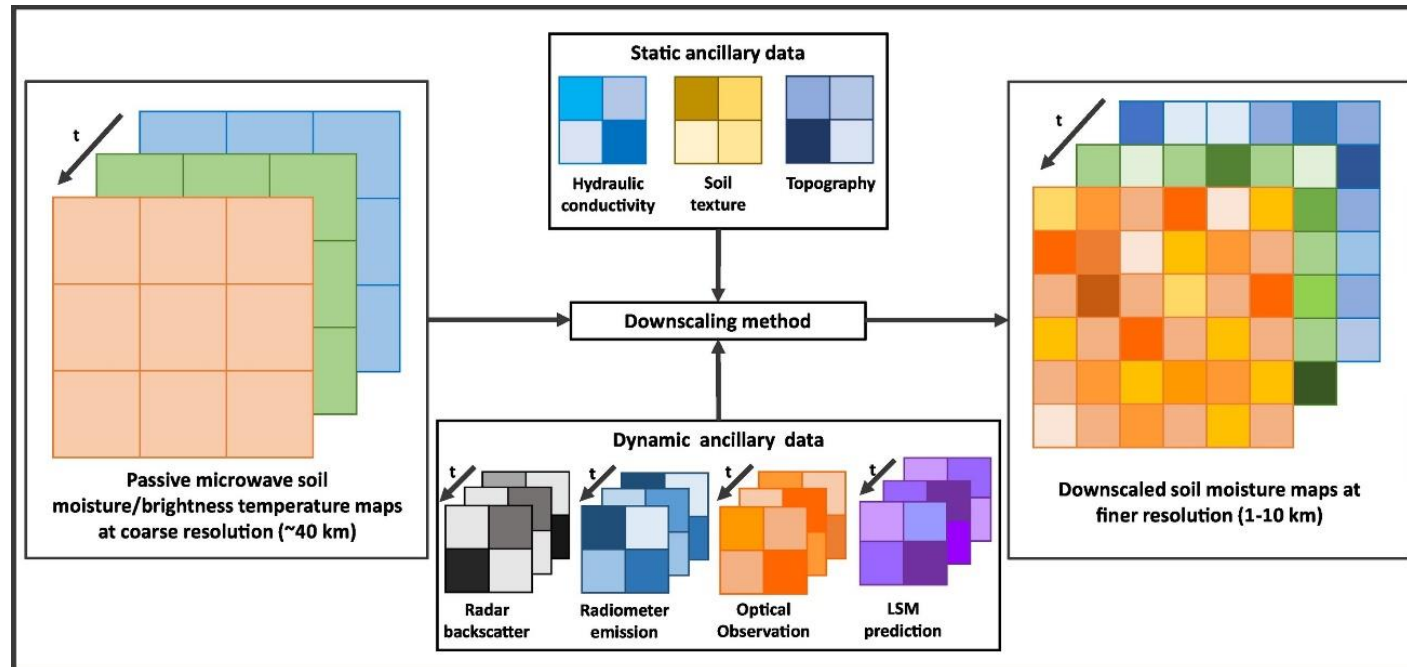
Idea:
Combine coarse resolution SM products and high resolution ancillary data, which drive/influence SM to get a high resolution SM.

Goal:
Downscale SSI SM from 25km to 5km spatial resolution

Dorigo et al. 2017 (*Remote Sens. Environ*)

Courtesy: Luca Zappa, TU Vienna

Example Application: Downscaling Soil Moisture

DATASET	Spatial res	Temporal res	Start date	End date	Source
Global radiation	0.05°	Daily	2018-01-01	2020-12-01	**EUMETSAT DC**
Direct normal solar radiation	0.05°	Daily	2018-01-01	2020-12-01	**EUMETSAT DC**
LAI	0.05°	Daily	2016-01-01	2020-12-31	**EUMETSAT DC**
Land cover (% of level 1 CLC)	100 m → 0.05°	Daily	2018		CLC
SM (combined, v05.2)	0.25°	~Daily	1978	2019-12-31	CCI

Datasets (predictors) used to feed and train the Random Forest regression

Courtesy: Luca Zappa, TU Vienna

TU WIEN
DEPARTMENT OF
GEODESY AND
GEOINFORMATION

EUMETSAT

Status & Outlook

- Still issues around data access via THREDDS
- Tutorials on accessing the D&V Cube using Python, R-based tools are being made available
- Collect user feedback
- Discuss applications with subject matter experts
- Continue the Exploration Phase?
- Further cubes for other application areas
  - Atmospheric Composition
  - Marine Applications
  - ….

EUMETSAT