Earth Observation Data Centre (EODC) Collaborative Exploitation of Sentinel data

What is needed to provide flexible service to users?

Norbert Pfeifer, Christian Briese, Wolfgang Wagner

norbert.pfeifer@geo.tuwien.ac.at

Department of Geodesy and Geoinformation
Vienna University of Technology
www.geo.tuwien.ac.at

Outline

- Earth Observation Data Centre (EODC),
  a solution for collaborative exploitation of Sentinel data.

- What is needed to provide flexibility of services to the users?

- Motivation / is there a problem ?
  - Context NMCA
  - Context Traditional Academic Earth Observation
  - Context MW RS / Sentinel 1
  - Context Big Data

- EODC
  - Mission
  - Services
 - Communities
  - Infrastructure

National Mapping and Cadastral Agencies

- EuroSDR – pan-European organization for Spatial Data Research

- EuroSDR – linking NMCA (research) and academia

- EuroSDR workshop (26-27 November 2014, Ordnance Survey UK)
  **Efficient capturing of 3D objects at a national level
  with a focus on buildings and infrastructure**

  - *Processing 3D point clouds* (from airborne laser scanning) with ~10 pts/m2 and cadastral building footprints to obtain building models for the entire Netherlands 40.000km2 = $4*10^{11}$ points, one 3D point ~ 30 Byte, *12 TeraByte processing job*
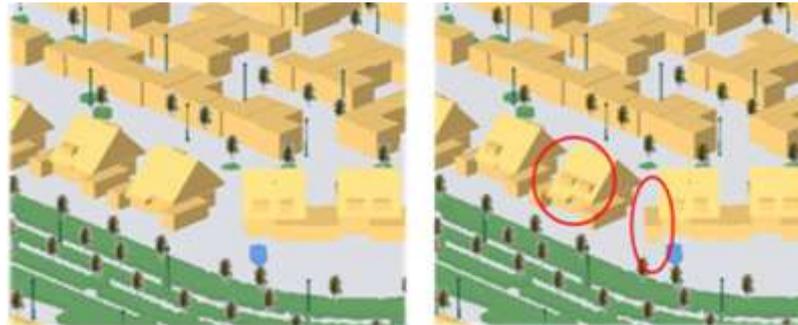

    Stoter et al., 2011. PFG

  - One key element of "**efficiency**" and of "**national level**" is standardized regular Earth Observation

Earth Observation Today

- Scientists download EO data (internet), store on own file system (server?), process with own software
  - Software: mix of commercial, open source and in-house developed
- Advantage
  - Each organisation has full control over data and software
  - Protection of IPRs
  - Impression / Illusion of independence
- Limitations & disadvantages
  - Limited IT capabilities
    - Data storage, processing power
  - Everybody solves the same basic processing steps
    - Data input & management, quality control, georeferencing, radiometric calibration, …
  - Cooperation with external scientists/organisations is not encouraged

Traditional Approach has worked, e.g., for ENVISAT ASAR

- TU Wien has build up an ENVISAT ASAR data processing system for
  - Scientific exploitation
  - Demonstration of operational capabilities
- Continental to global scale products
  - Soil moisture
  - Water bodies
  - Freeze/thawing
- Level 1 data volume
  - GM mode: 202.492 images $\rightarrow$ 1,74 TB
  - WS mode: 116.483 images[*] $\rightarrow$ 23,49 TB
- IT Infrastructure
  - 600 TB file server
  - 1 PB robotic tape library
  - 12 multi-core processing machines

[*] About 80 % of all WS mode images registered in ESA archive
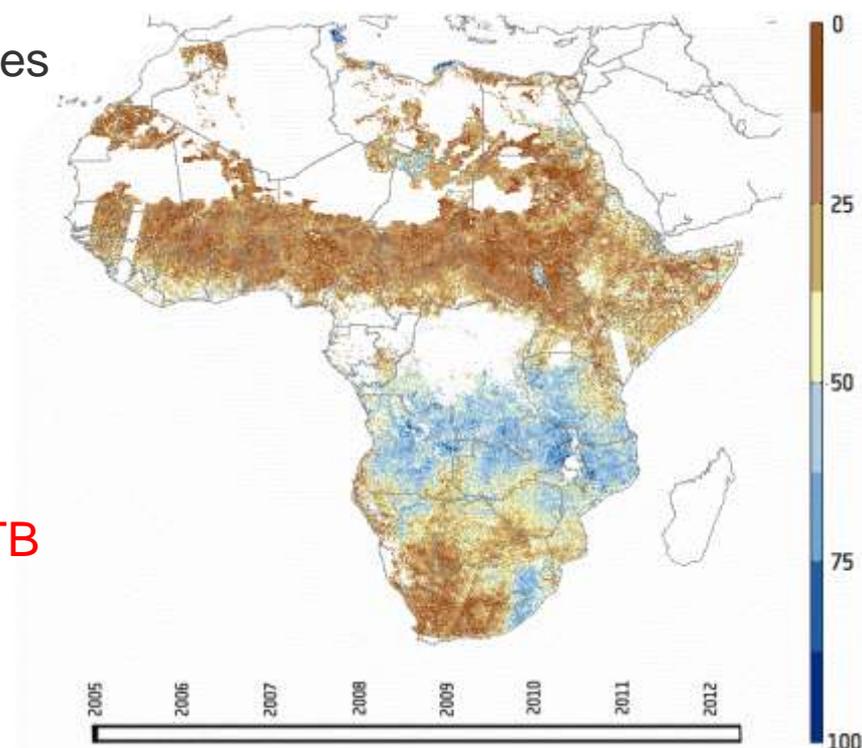
1km ASCAT soil moisture over Africa produced by TU Wien in the ESA funded SHARE & Tigernet projects

Traditional approach does not work for Sentinel-1

- Sentinel-1 data volume
  - One satellite, over landmasses (IW, 15 minutes duty cycle per orbit)
  - SLC data acquisitions only 3,75 minutes during each duty cycle

Data-volume estimates (Single Polarization, raw format, excl. annotation)

Product Type	Data rate [MB/s]	Data acq. per orbit	Data volume per orbit	Data volume per day	Data volume per year	Data volume 7.5 years	Data volume 20 years
IW L1 SLC	127.554	3.75 min	28.7 GB	419.0 GB	152.9 TB	1.1 PB	3.1 PB
IW L1 GRD-HR	32.418	**15 min**	**29.2 GB**	**426.0 GB**	**156.0 TB**	**1.2 PB**	**3.1 PB**
IW L1 GRD-MR	5.190	**15 min**	**4.7 GB**	**68.2 GB**	**25.0 TB**	**186.7 TB**	**497.8 TB**
IW L1 BRW	0.007	**15 min**	**6.3 MB**	**92.0 MB**	**33.6 GB**	**251.8 GB**	**671.5 GB**
Total	-	-	**62.6 GB**	**913.3 GB**	**333.9 TB**	**2.5 PB**	**6.7 PB**

Average Revisit Time with S-1A + S-1B Satellites

~ 3 days

~ 6 days

Average Revisit time for two Sentinel-1 satellites © ESA
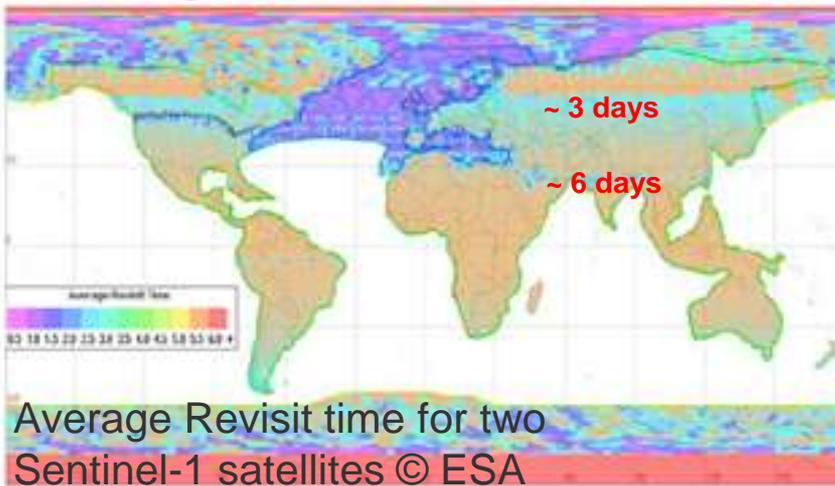
Comparable in size to the complete ASAR data volume

IW: Interferometric Wide swath mode
SLC: slant range geometry, single look complex
GRD: ground range geometry
FR/HR/MR: full/high/medium resolution

New Paradigm in EO Data Processing is Needed

- **Main reasons**
  - Data volume and data transfer rates
  - Increasing complexity of algorithms with increasing resolution
  - Higher scientific standards
    - Algorithms must be validated with big data sets and competing algorithms
    - Algorithms ensembles needed

- **Solutions**
  - Bring software to data
  - Cooperation & specialisation

- **IT Solutions exists**
  - Tools for collaboration
  - Virtualisation
  - Parallelisation
  - Cloud Computing
  - Big Data

Challenge is to change the behaviour of people & organisations!

Open-mindedness, willingness to share, transparency, and participatory decision making are probably crucial if one wants to succeed

Earth Observation Data Centre (EODC)

- A bottom-up initiative with top-down guidance
  - First proposed by an Austrian EO expert group but with a strong international orientation from the beginning
  - Two feasibility studies supported by the Austrian Aeronautics and Space Agency (FFG-ALR) and ESA
- Cooperation model was developed starting from the example of an existing cooperation between public and private partners for open source developments
  - Adapted to our goals and needs
  - Strongly shaped by legal considerations
    - European and national regulations on state aid, public procurement law, Anti trust law, personnel law
- Key challenges
  - Build the trust that everybody will benefit from the cooperation
  - Obtain financial commitments from cooperation partners

EODC Mission Statement

- EODC will work together with its partners from science, the public- and the private sectors in order to foster the use of EO data for monitoring of water and land

- EODC shall act as a community facilitator

- Joint developments
  - Cloud infrastructur
  - Operational data services
  - Software
    - Open Source

Satellite	Temporal	Product
Sentinel-1	static	Water Bodies
		Forest / Non-Forest
		Rice
	dynamic	Water Bodies
		Rice and Growing Status
		Soil Moisture
		Soil Water Index
		Start of Soil Water Season
		Dry / Wet Indices
		Flooding
		Snow
		Freeze / Thaw
Sentinel-2	static	Forest / Non-Forest
		Biomass Indices
		Land Cover
		Glacier Area
	dynamic	Forest Functional Parameters
		Evapotranspiration
		Lakes Water Quality
Sentinel-3	static	Glacier Area
	dynamic	Snow
		Lake / River Levels

Primary EO data products considered by EODC

EODC Services

Community-Building Services

- Communication platforms
- Development of standards
- Project management
- Support of Communities

Software Services

- Licence management
- Software development
- Sale of commercial software
- User support

Data Services

- Acquisition
- Distribution
- Archiving
- Near-real-time services
  - Satellite raw data
  - Geodata
  - Model forecasts

Platform Services

- Scientific platforms
- Development platforms
- Processing of Big Data
- Cloud-based web applications
  - Visualisations
  - User-tailored Apps
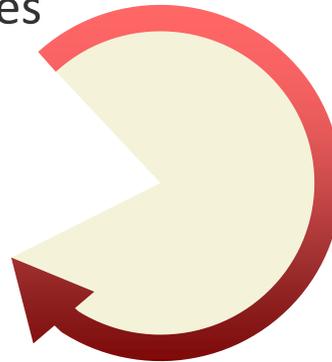
Cooperation Model

- Work is done within the "Communities"
- Strategic decisions taken by
  - EODC Executives
  - Advisory Board
  - Shareholder Assembly
- Democratic voting principles
  - No Veto

EODC cooperation partners

Infrastructure enabling EODC
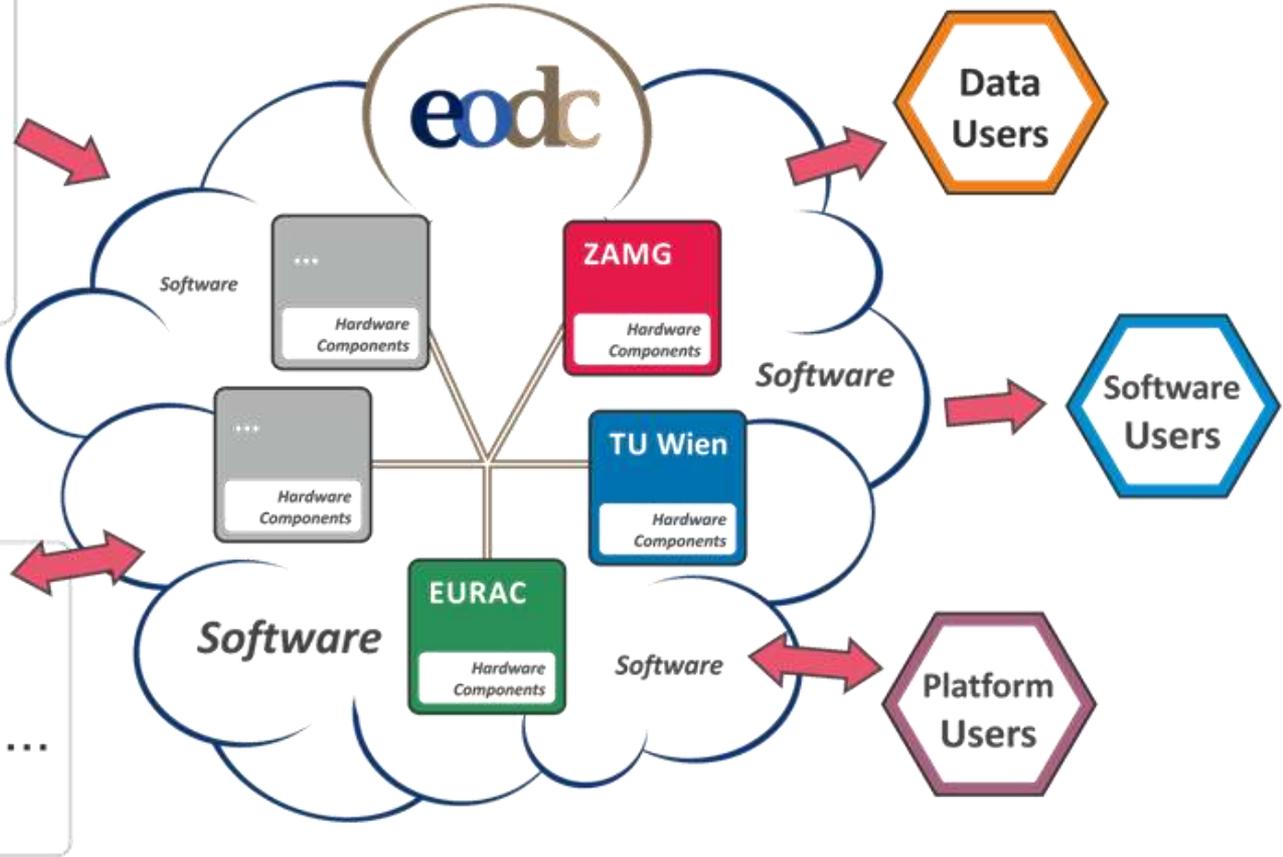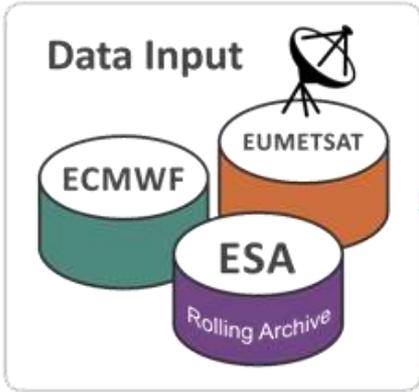
Connecting Science and Operations

EODC Infrastructure of TU Wien

Virtual Machines (VMs) for R&D
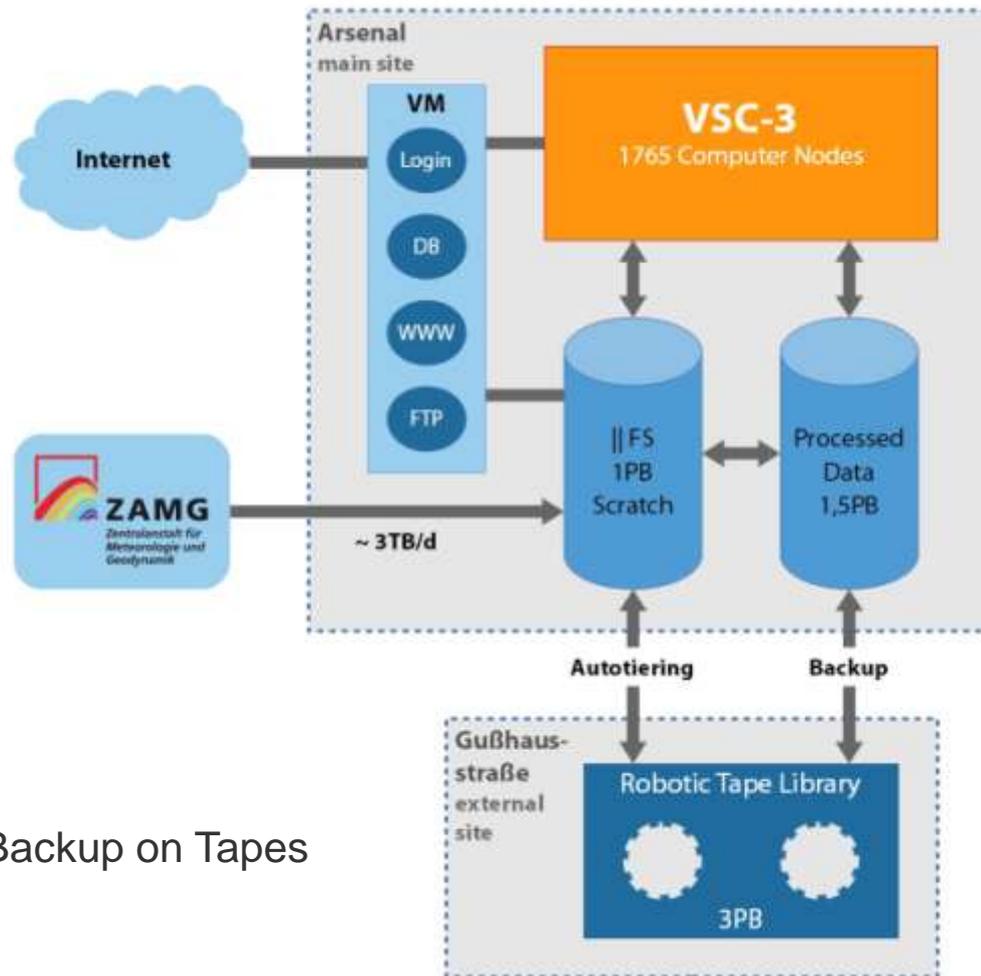
Supercomputing for NRT, Global Testing & Reprocessing

24/7 Operations & Rolling Archive

Data Storage

Backup on Tapes

20

Vienna Scientific Cluster 3

- Hosted by Science Center of TU Wien

- Cooperation of 8 Austrian Universities

- >32.000 cores with >600 Teraflops

- Cooling with mineral-oil in 35 tanks

- Power 540 Kilowatt

 - Matches the needs of ~1400 Austrian households

VSC-2

VSC-3

Current EODC work: Re-engineering of SAR Processing Line for S-1

- **Open source** code
- Fast parallel processing
- Optimized data formatting
- **Fast access in time** and spatial domain
- Efficient data archiving

Current EODC work: Optimising S-1 Data Representation & Storage

- Equidistant projections are most suitable, with a global mean oversampling of 2% when using a system of seven continental grids
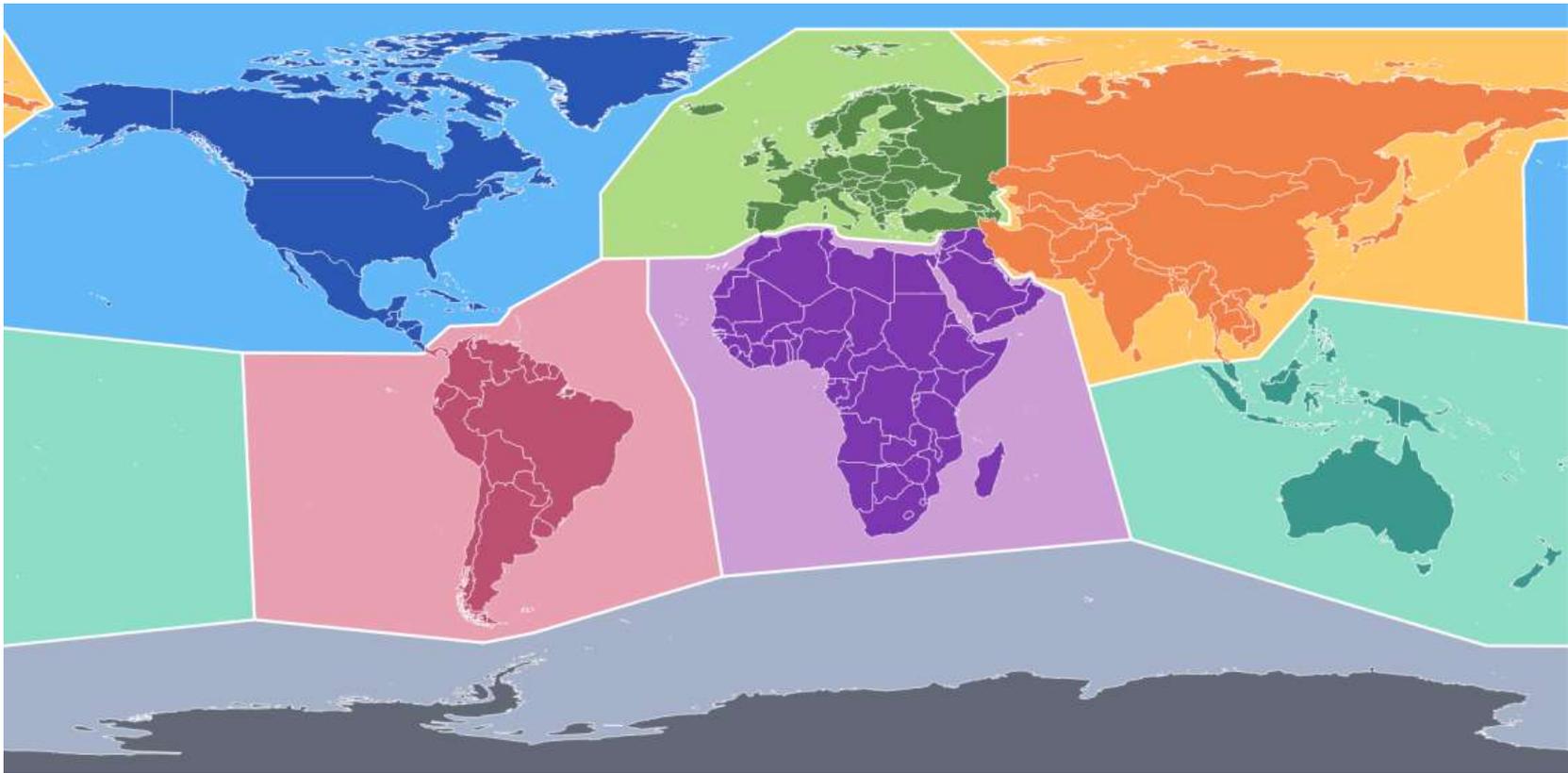
Bauer-Marschallinger, B., D. Sabel, W. Wagner (2014) Optimisation of Global Grids for High-Resolution Remote Sensing Data, Computers & Geosciences, 72, 84–93

EODC Summary

- **EODC** represents a **new model of cooperation** in earth observation, it strives to develop shared resources to the benefit of its cooperation partners

- EODC is still in development

- Big EO data can be an integral part of general Big Data activities and there is room for cooperation beyond different fields

Interested in the EODC Cooperation Network?

norbert.pfeifer@geo.tuwien.ac.at
wolfgang.wagner@geo.tuwien.ac.at
christian.briese@eodc.eu