

Big data challenge in Earth Observation

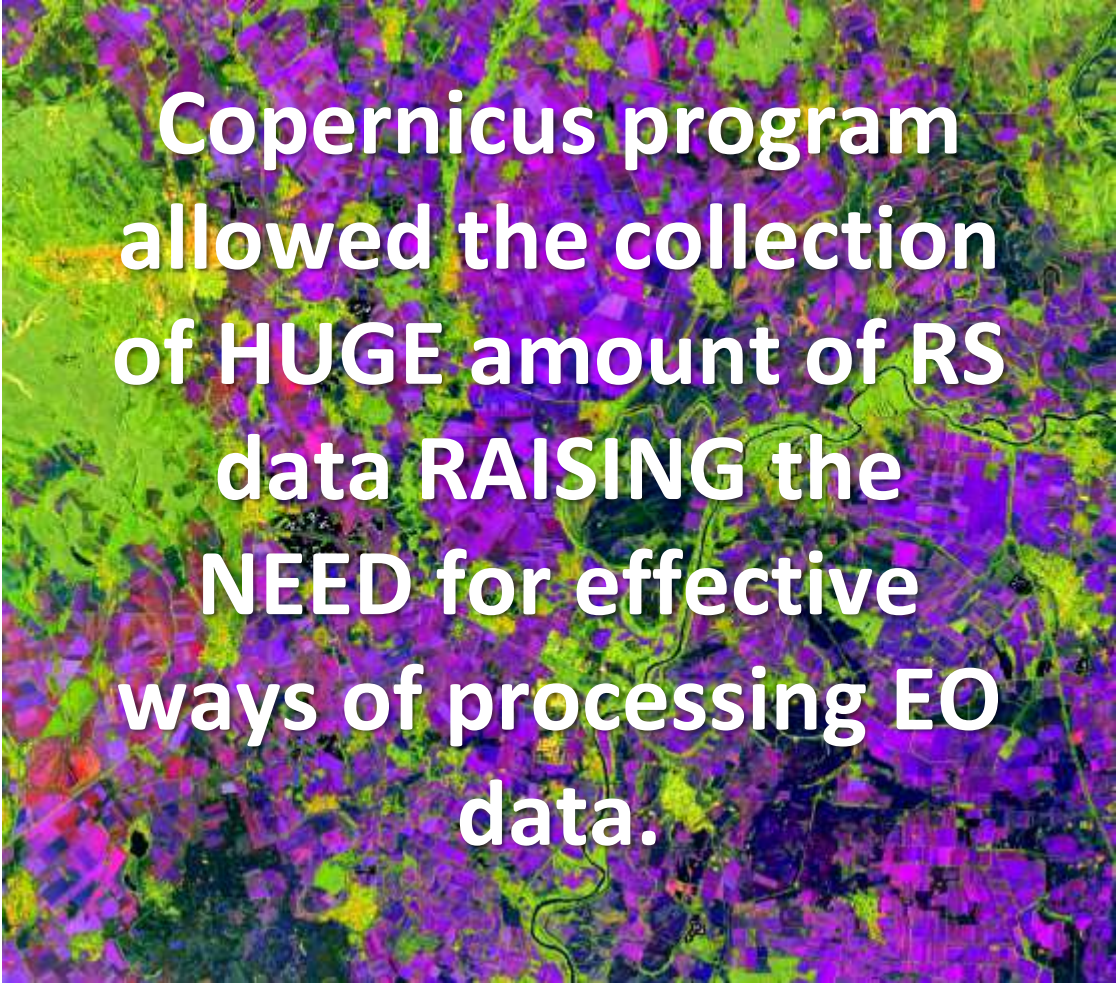
Dedicated supercomputer based Sentinel-1 processing chain

F. Collivignarelli¹ – Gy. Harka² – F. Holecz¹ – A. Peternier¹ - L. Ronczyk² – B. Zaválnij³

¹sarmap SA

²University of Pécs

³Coker College

A satellite image showing a dense urban or industrial area with a complex pattern of purple, green, and yellow, representing different land use or intensity levels. The text is overlaid on this image.

**Copernicus program
allowed the collection
of HUGE amount of RS
data RAISING the
NEED for effective
ways of processing EO
data.**

Sentinel-1 Intensity temporal composite

Outline

- **Big data issue,**
- **New challenges of EO,**
- **Dedicated supercomputer based Sentinel-1 processing chain**
 - MAPscape upscale,
 - Container virtualization,
 - Job Automated System,
 - User interface.
- **Future direction.**

Big Data

High management capabilities.

4Vs: Volume, Velocity, Variety, Veracity

Baumann et al 2016.

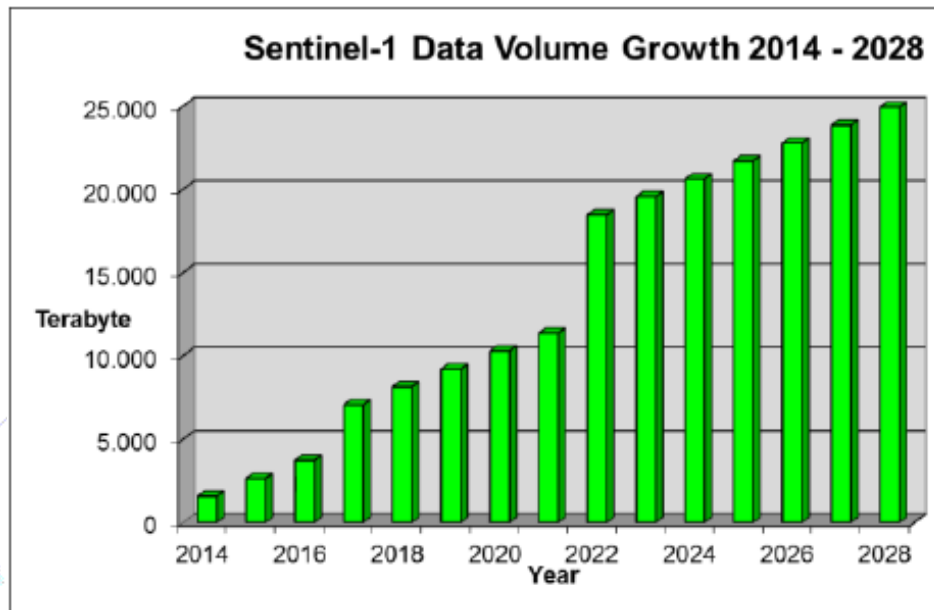


Figure 5-Expected-Sentinel-1-Data-Volume-Growth-at-DLR-PAC-2014-2028.png



4 Components

Effective utilization of EO data

Data

ESA,
Copernicus
Programme,
Sentinel mission,
Sentinel-1 all-
weather, day and
night radar
imaging



Software



MAPscape

Successfully
implemented:
www.rice.org



Interface

- Container virtualization
 - Job automatization system



HPC

Widely
available
service



Challenges :

1. Desktop driven software development implementation on HPC environment (CPU challenge),
2. Very diverse architectures and environment of supercomputers
3. Limited user knowledge

Supercomputer architecture and performance

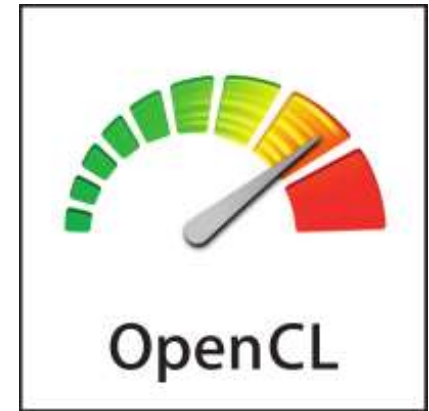
Számítógép	Budapest klaszter	Budapest 2 - Phi klaszter	Szeged klaszter GPU-val	Debrecen klaszter	Debrecen 2 (Leo) - GPU klaszter	Debrecen 3 (Apollo) - Phi klaszter	Pécs - UV számítógép	Miskolc - UV számítógép
Típus	HP CP4000SL	HP SL250s	HP CP4000BL	SGI ICE8400EX	HP SL250s	HP Apollo 8000	SGI UV 1000	SGI UV 2000
CPU magok száma	768	280	2400	1536	1344	1080	1152	352
Memória / node	66 GB	63 GB	132 GB	47 GB	125 GB	125 GB	6 TB	1.4 TB
GPU / Phi gyorsítókártyák, koprocesszorok	-	28 db Intel Xeon Phi	12 db Nvidia M2070	-	204 db Nvidia K20x + 48 db Nvidia K40x	88 db Intel Xeon Phi	-	-
Linpack teljesítmény (Rmax)	5 Tflops	27 Tflops	20 Tflops	18 Tflops	254 Tflops	106 Tflops	10 Tflops	8 Tflops
Compute node-ok száma	32	14	50	128	84	45	1	1

1. Challenge > MAPscape upscale

It enables the generation of backscattering coefficient (σ^0) from Single Look Complex (SLC) or Ground Range Detected data.

It includes the following automated steps:

1. Co-registration including Digital Elevation Model (DEM)
2. Time series speckle filtering
3. Terrain geocoding and radiometric calibration
4. Radiometric normalisation
5. Anisotropic Non-Linear Diffusion filtering



It enables the generation of coherence (interferometric correlation, γ) from Single Look Complex (SLC) data.

It includes the following automated steps:

1. Co-registration including DEM
2. Generation of coherence including DEM
3. Terrain geocoding
4. Anisotropic Non-Linear Diffusion filtering

2. Challenge > platform/distribution independent environment for MAPscape on supercomputers

Container virtualization:

technology, where the MAPscape running environment can be adjusted for different supercomputers

The library and kernel architectures of different supercomputers are different a lot. Our aim is to deliver a software which can be used on various supercomputers.

Own virtualization system

3.1 Challenge > Limited user knowledge (supercomputer usage)

Job Automated System (JAS)

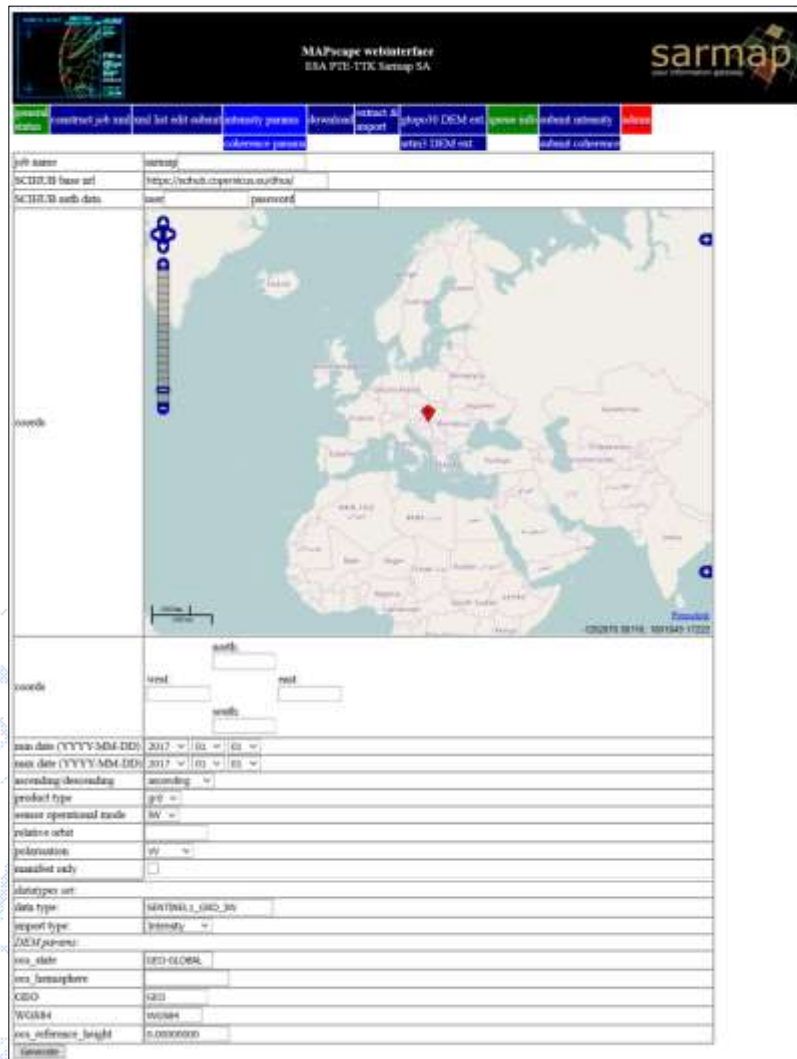
Totally automated JOB management in HPC environment,

- Download data from SciHub,
- Construct input parameters of MAPscape software,
- Call the processing steps on supercomputer,

```
select HPC
login daniel@2 /usr/bin/ls | select
submit job on the HPC
057254 post-gis license- sar/2014 PD 0-00 1 24 2017-10-17T19:40:41 (Priority)
submit job on HPC queue
Job name:
JAS version 2.1.4

sar-jas-run
sar-jas-grp-bin
sar-jas-ctrl
sar-jas-verifyname
sar-jas-runlog
sar-jas-defin
sar-jas-queue
sar-jas-down
sar-jas-extract-and-import
sar-jas-import
sar-jas-dm-ctrl
sar-jas-dm-grp10
sar-jas-ctrl-ctrl-grp10
sar-jas-run-queue
sar-jas-attenu-local
sar-jas-attenu-queue
sar-jas-coherence-local
sar-jas-coherence-queue
sar-jas-cancel-queue
sar-jas-queue1st
sar-jas-local1st
sar-jas-queue1st
sar-jas-queue1st
```

Rapid Sentinel-1 processing chain on supercomputer for testing (already exists)



SARscape webinterface
ESA PTG-TTK SarMap SA

job name: sarmap

SARscape base url: <https://scratch.sarmap.com/ptg/>

SARscape auth data: user: password:

search: [Map of Europe with a red pin over France]

coordinates: [lat, lon, east, west, south, north]

main date (YYYY-MM-DD): 2017 | 01 | 01

aux date (YYYY-MM-DD): 2017 | 01 | 01

ascending/descending: ascending

product type: p0

sensor operational mode: SV

relative orbit: []

polarization: []

swath/shot width: []

datatypes set: []

data type: SENTINEL_1_GRID_30

report type: intensity

DSM parameters:

sea state: IFT-Global

sea hemisphere: []

CEO: CEO

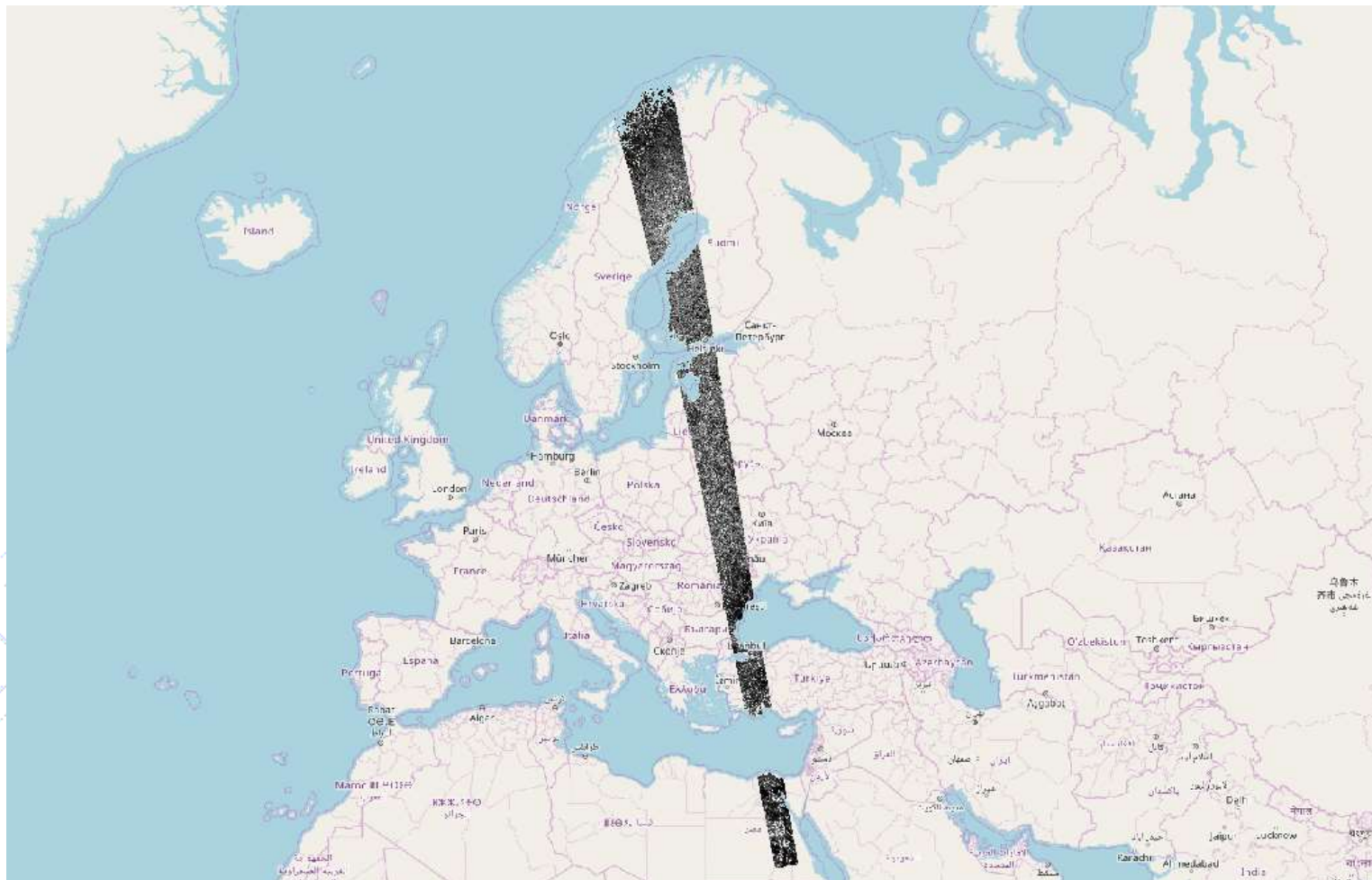
WGS84: WGS84

sea reference height: 0.000000

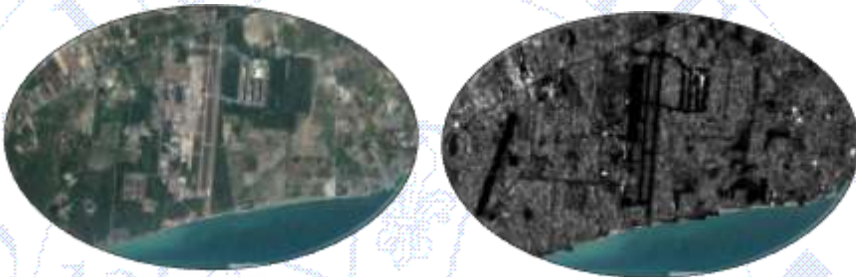
Generate

- State-of-the-art algorithms
- Exploit supercomputer capabilities,
- Performance of JAS script

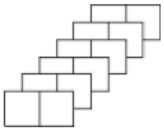
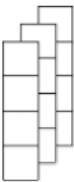
EXTREM STRESS TEST in HPC (32 adjacent frames (137 GB) GTOPO DEM)

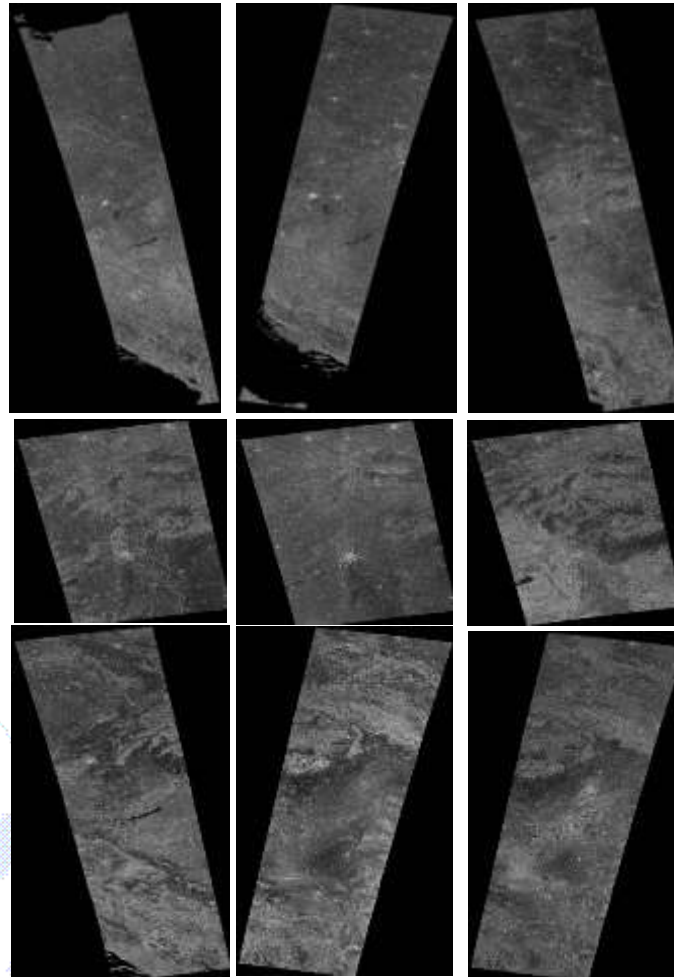


EXTREM STRESS TEST in HPC (24 adjacent frames SRTM DEM)

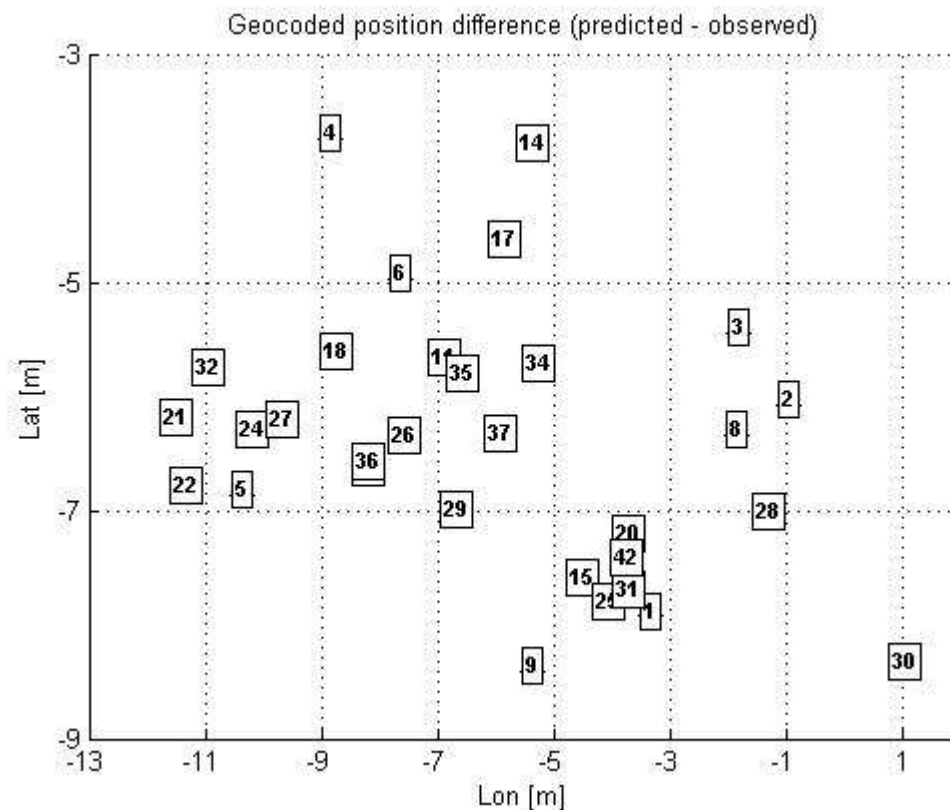


SYSTEMATIC STRESS TEST on supercomputer (different combination of 24 products)

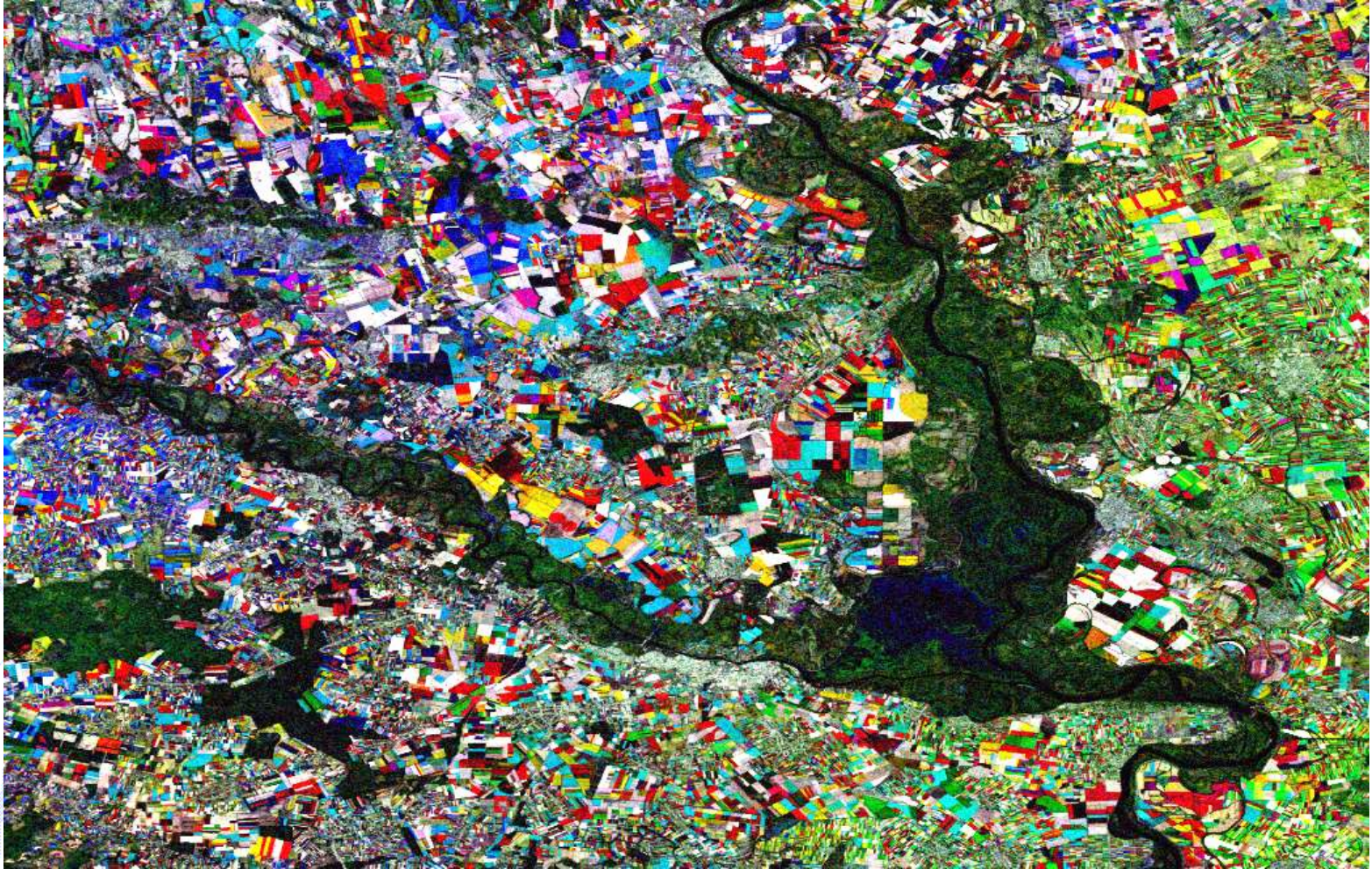
261	2	6	1	
621	6	2	1	
612	6	1	2	
134	1	3	4	



Geocoding accuracy of Sentinel-1 data processed by MAPscape on supercomputer



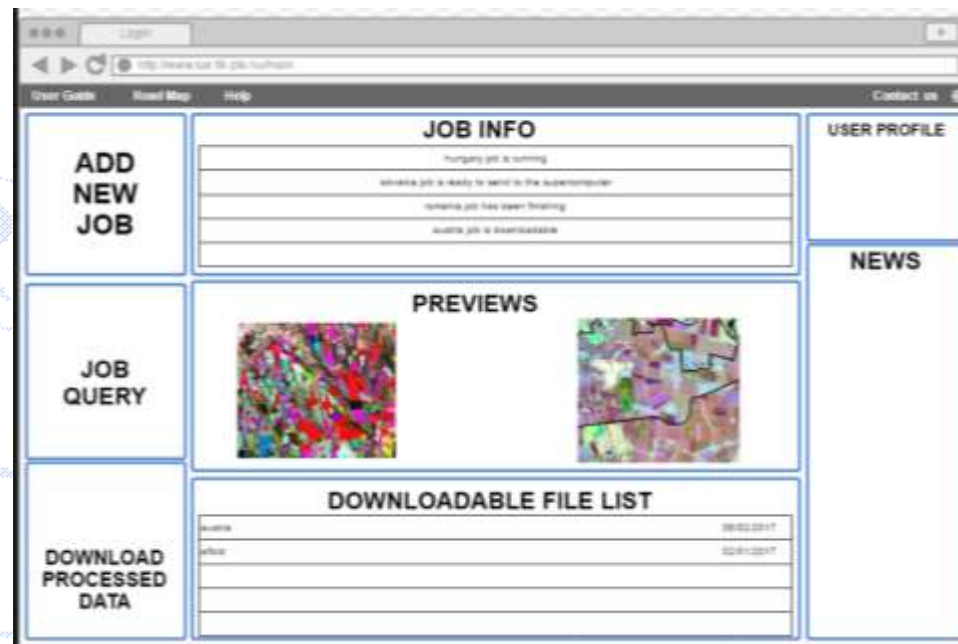
Coherence process land cover/land use test



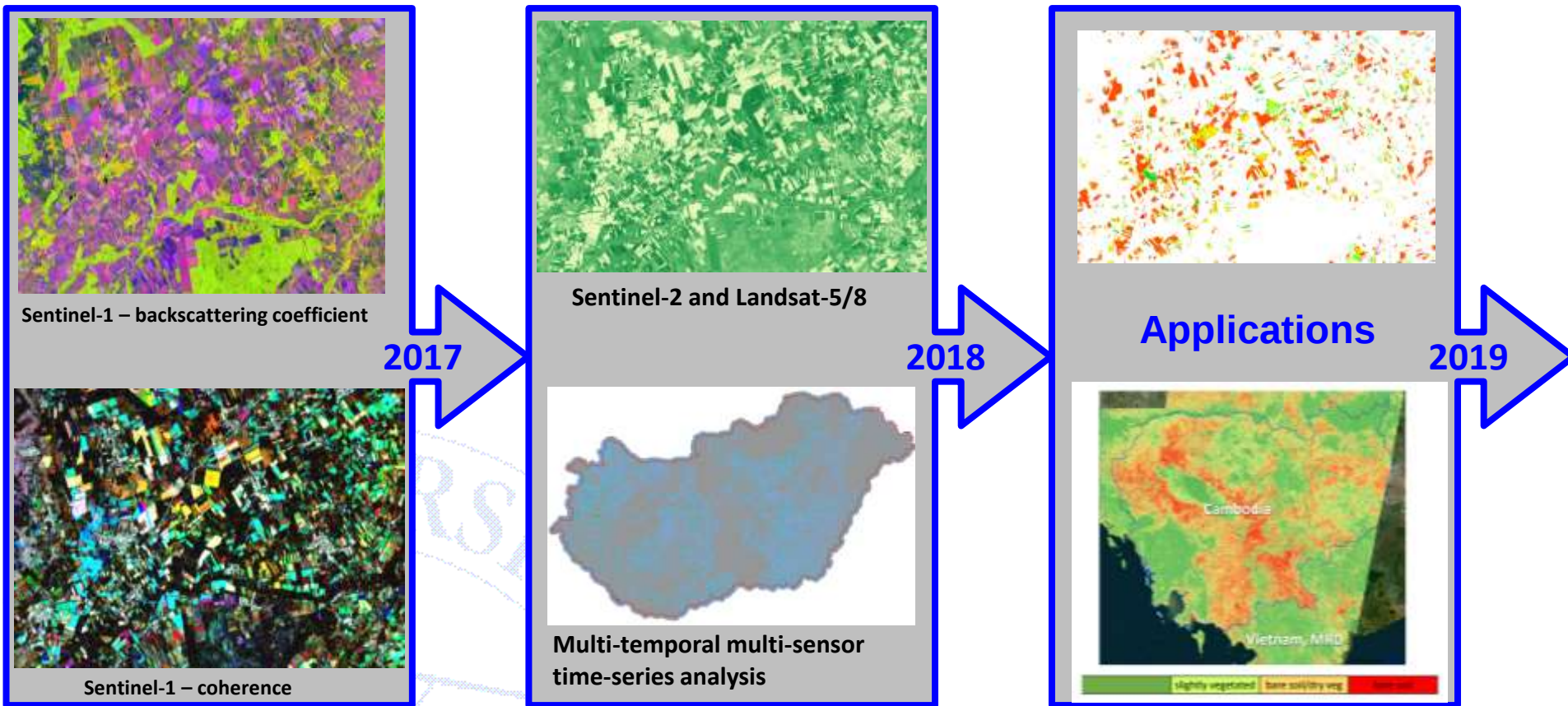
3.2 Challenge > Limited user knowledge (SAR data processing)

Web user-interface:

- User management,
- Service management,
- Notification sending,
- Gather Statistical Data



on Supercomputer



Financed by ESA

highly automated way

Köszönjük a figyelmet!

h2o@gamma.ttk.pte.hu

