Using SNAP in Cloud processing services for GEP

Φ-week event

10 September 2019

ESA-ESRIN

Terradue

- Terradue is an European Space Agency spin-off started in 2006
  - Based in Rome, staff of 16 from 6 nationalities
- Providing support to application builders in Earth sciences
  - To use satellite EO data as information source
  - Cloud PaaS, complemented with APIs for Cloud bursting
- Business model: Platform-centered, a collaborative workplace “**Ellip**” for value adders to interact & co-create

Thematic Exploitation Platforms

- R&D activities to create an ecosystem of interconnected Thematic Exploitation Platforms

- Users access a work environment containing the data and resources required, as opposed to downloading and replicating the data “at home”

The fundamental principle is to move the User to the data and tools

Geohazards Exploitation Platform | GEP

GEP designed in the context of:

- Geohazards Supersite initiative (GSNL)
- CEOS Disasters Working Group

User-driven model for partnership and community building

Started from Int. Forum on Satellite EO and Geohazards organised by ESA and GEO in Santorini in 2012 (140+ participants)

Geohazards Exploitation Platform | GEP

Platform based on virtualization & federation of EO data

- Provide services & support to the geohazards community

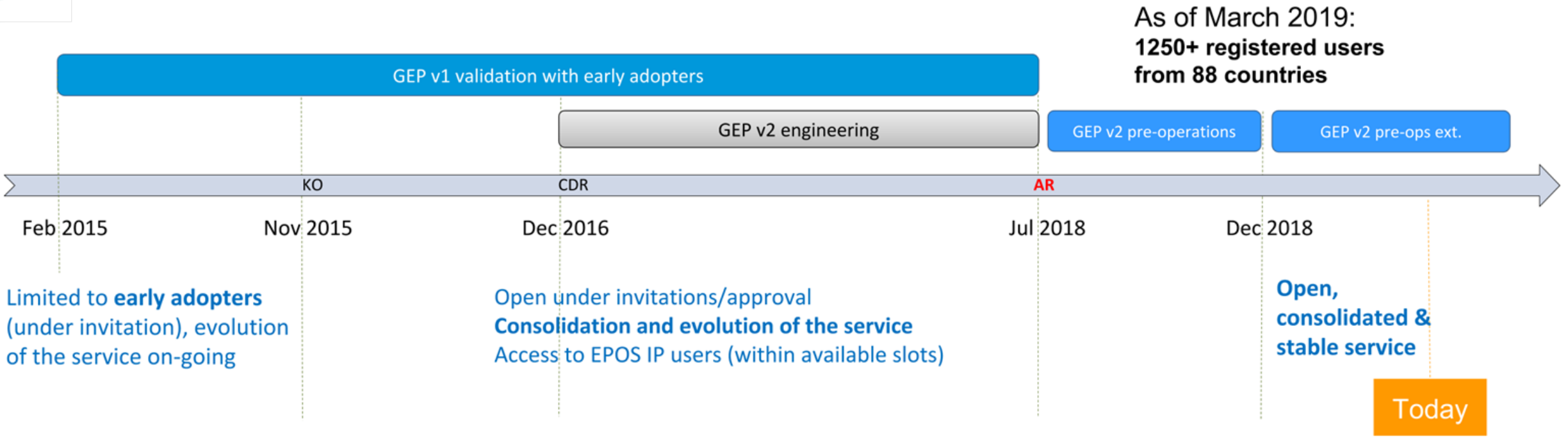
On-demand & systematic processing services

- Cloud Compute power, managing multi-tenant resources

Access to Copernicus Sentinels repositories

- Plus access to hundred TBs of EO data archives (ERS and ENVISAT), and other EO missions (ALOS-2, Cosmo-Skymed and TerraSAR-X) under CEOS WG Disaster and the GSNL agreements

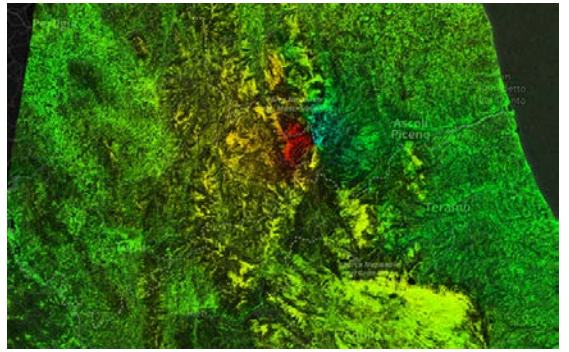
Geohazards Exploitation Platform | GEP

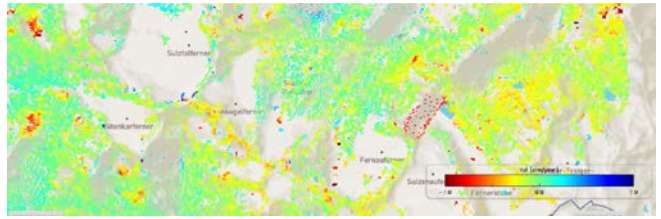


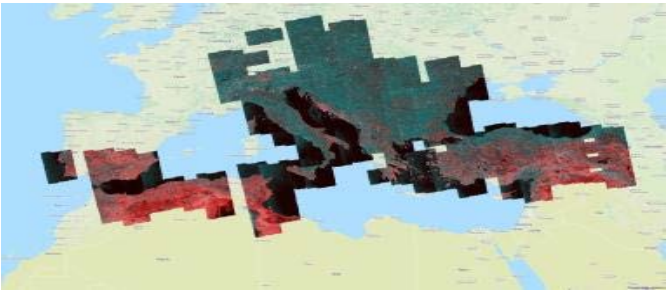
Want to apply as early adopter of the GEP Early Adopters Programme ?
contact@geohazards-tep.eu

GEP | Enriched Services Portfolio

- 25+ on-demand services using Optical & SAR data grouped in Thematic Apps, according to the defined goals of Community Managers
- New basic services providing full resolution and change detection imagery for rapid online visualization
- 8 systematic services delivering continuously updated information layers on GEP, including the large scale production of Sentinel-1 InSAR browse images at both 100m and 50m resolution over tectonic regions and volcanoes



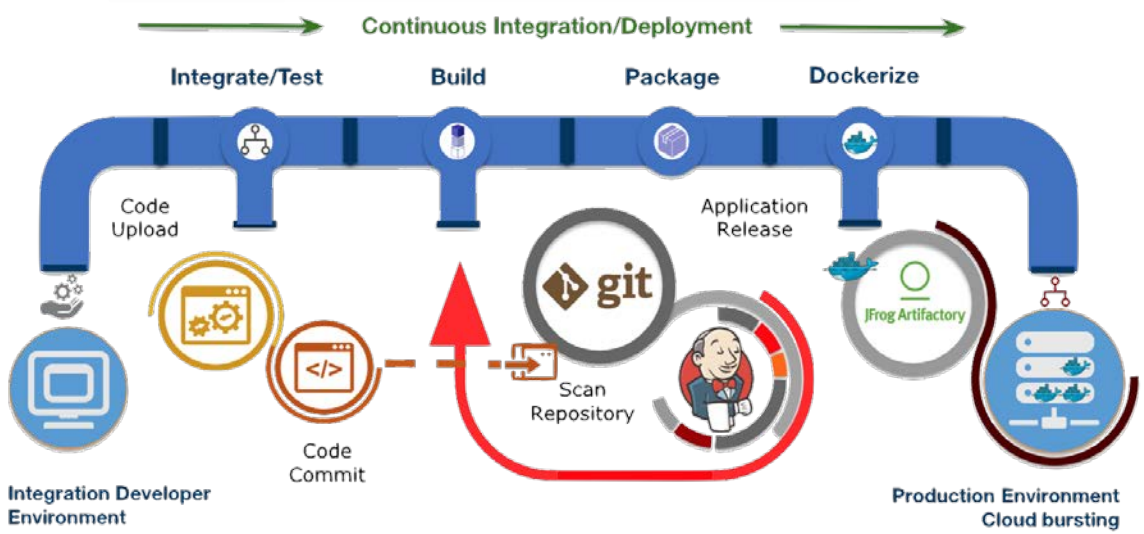




GEP | Application Integration

- Applications developed in any programming language supported
  - C/C++, Java, Python, Matlab and IDL

- Continuous Integration and Deployment Environment with automatic packaging and deployment in production environments



sobloo

CREODIAS

ONDA

GEP | Cloud Platform - Computing

- Continuous Integration and Deployment Environment with automatic packaging & deployment in production environments

- Improved Production Center, with (auto)scalability allowing cost-effective data processing on Cloud Computing

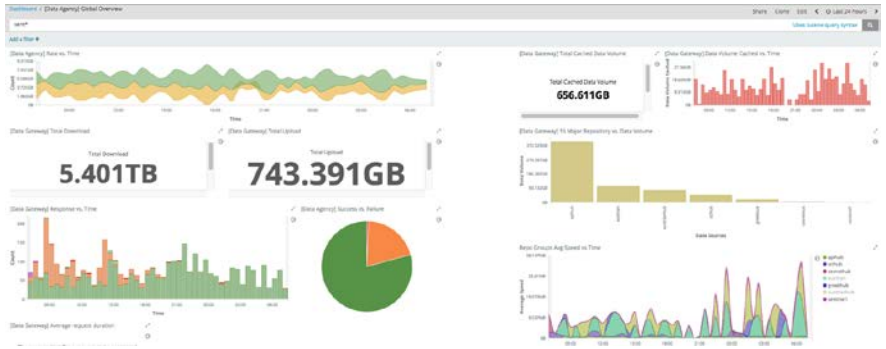
- Deployment in multiple Cloud-based processing environments with no lock-in on a Cloud provider

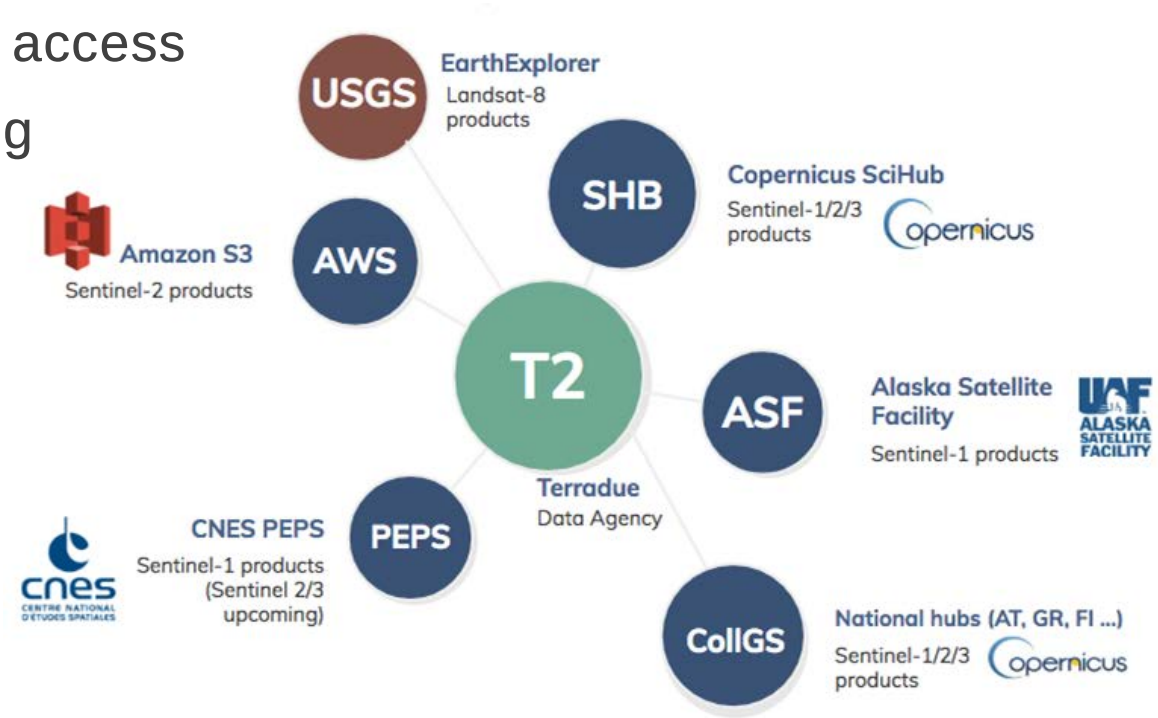
GEP | Cloud Platform - Data

Enhanced Data Gateway using OpenSearch

- Automatic multi-sourcing to optimise data access
- Programmable and systematic data caching
- Data usage accounting
- Personal cloud storage (repository)

Daily figures





Supported Data Providers

TERRADUE
Advancing Earth Science

Using SNAP in GEP

ESA asked us to provide a feedback:

1. In what context are you using SNAP, which parts and to do what?
2. What other software do you use along with SNAP? Do they have to interact with each other?
3. What works and what doesn't?
4. What performances do you get (time, quality)?
5. What extra SNAP feature would help your work?
6. Did SNAP help you achieve everything that you expect it to?

TERRADUE
Advancing Earth Science

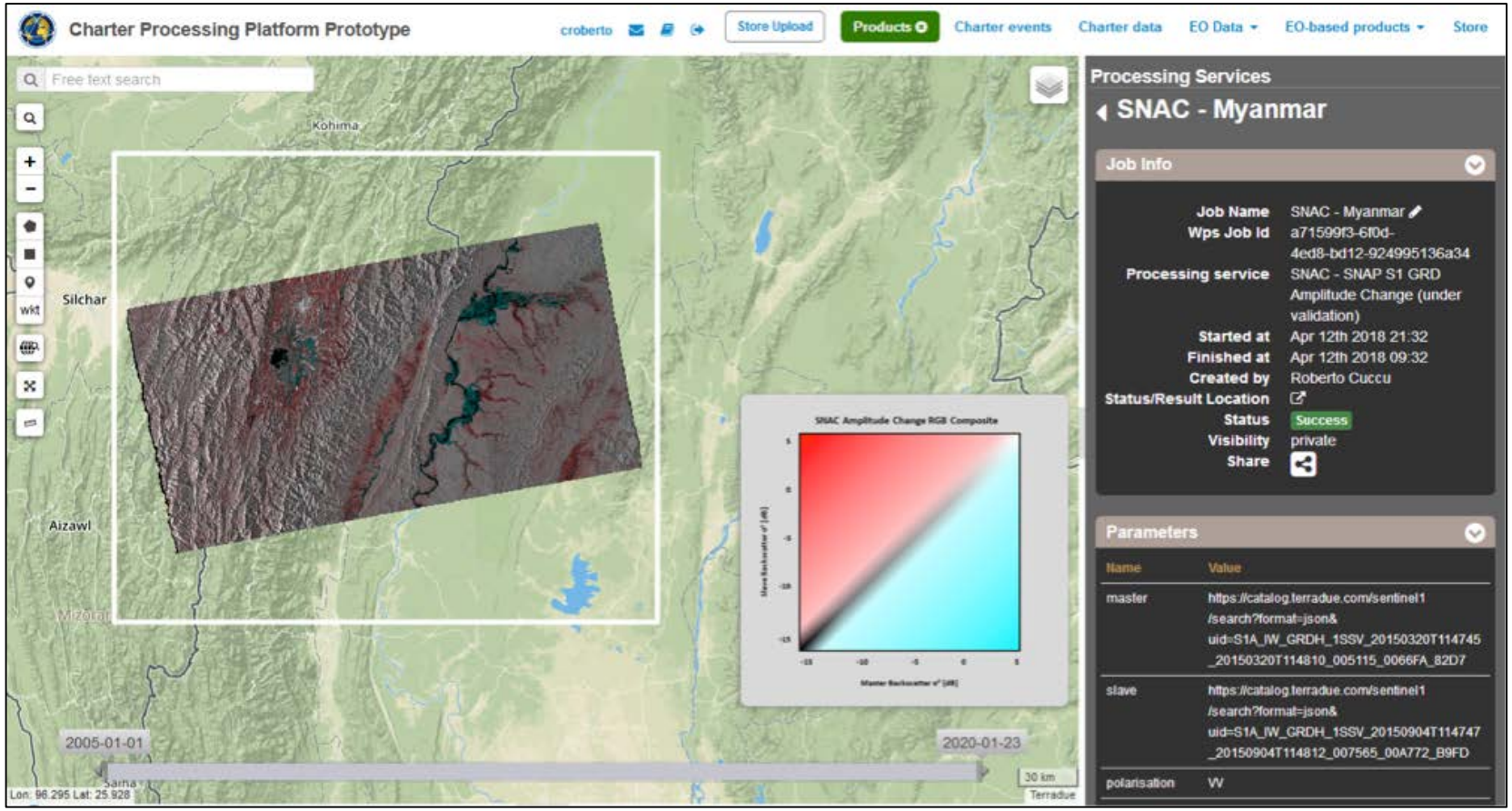
Some examples of SNAP services

- SNAC
- COIN
- COMBI
- Active fire detection with Sentinel-3
- Burned area assessment with Sentinel-2
- CSK interferogram generation

SNAC Sentinel-1 Amplitude Change

SNAC generates RGB composite of backscattering from a pair Sentinel-1 GRD IW and EW products (e.g. pre- and post-event)

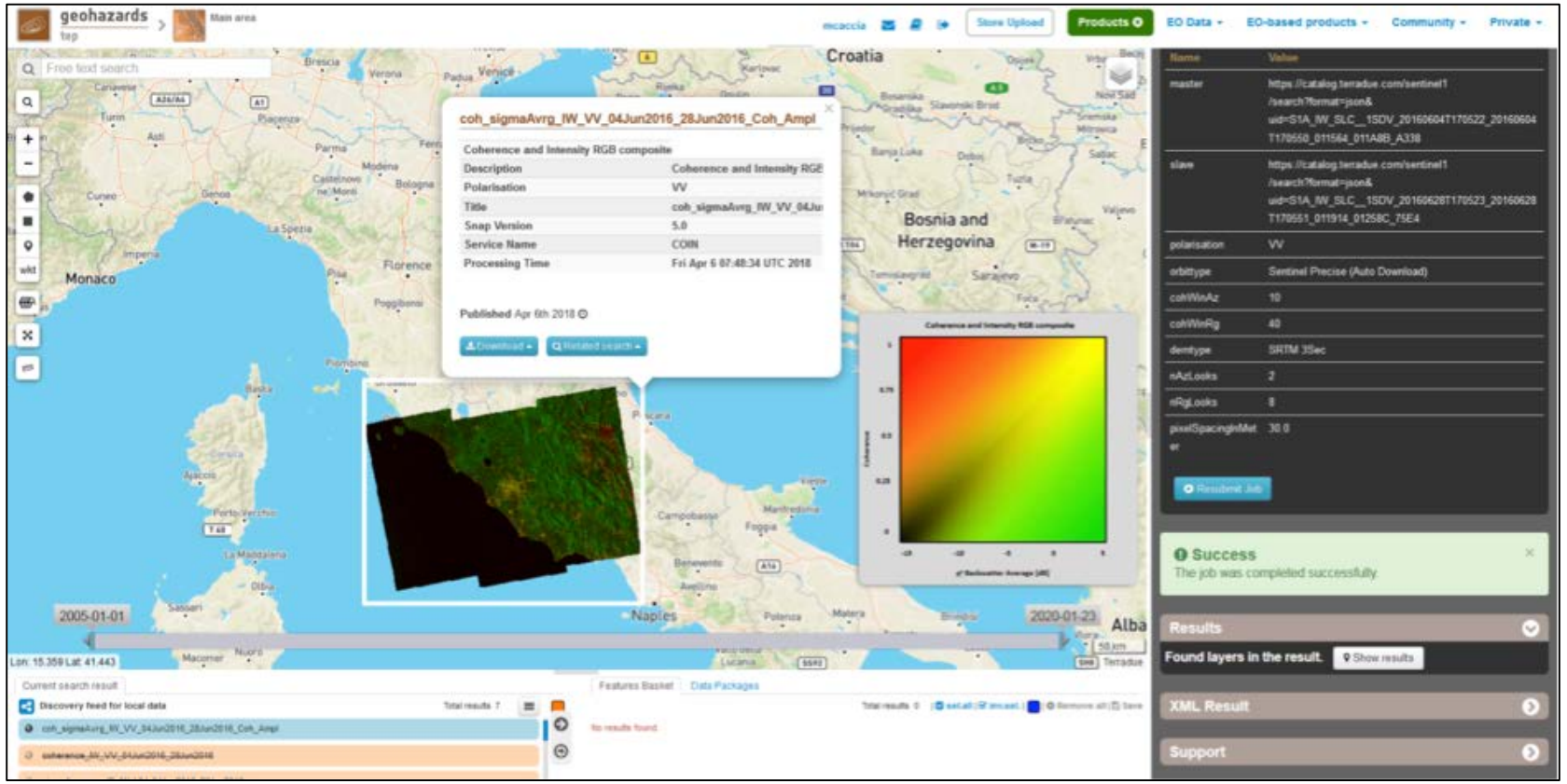
https://terradue.github.io/doc-tep-geohazards/tutorials/rss_snap_s1_snac.html



COIN Coherence and Intensity change for S1

COIN produces geocoded composites of coherence and amplitude images from a pair of Sentinel-1 TOPSAR IW data pairs.

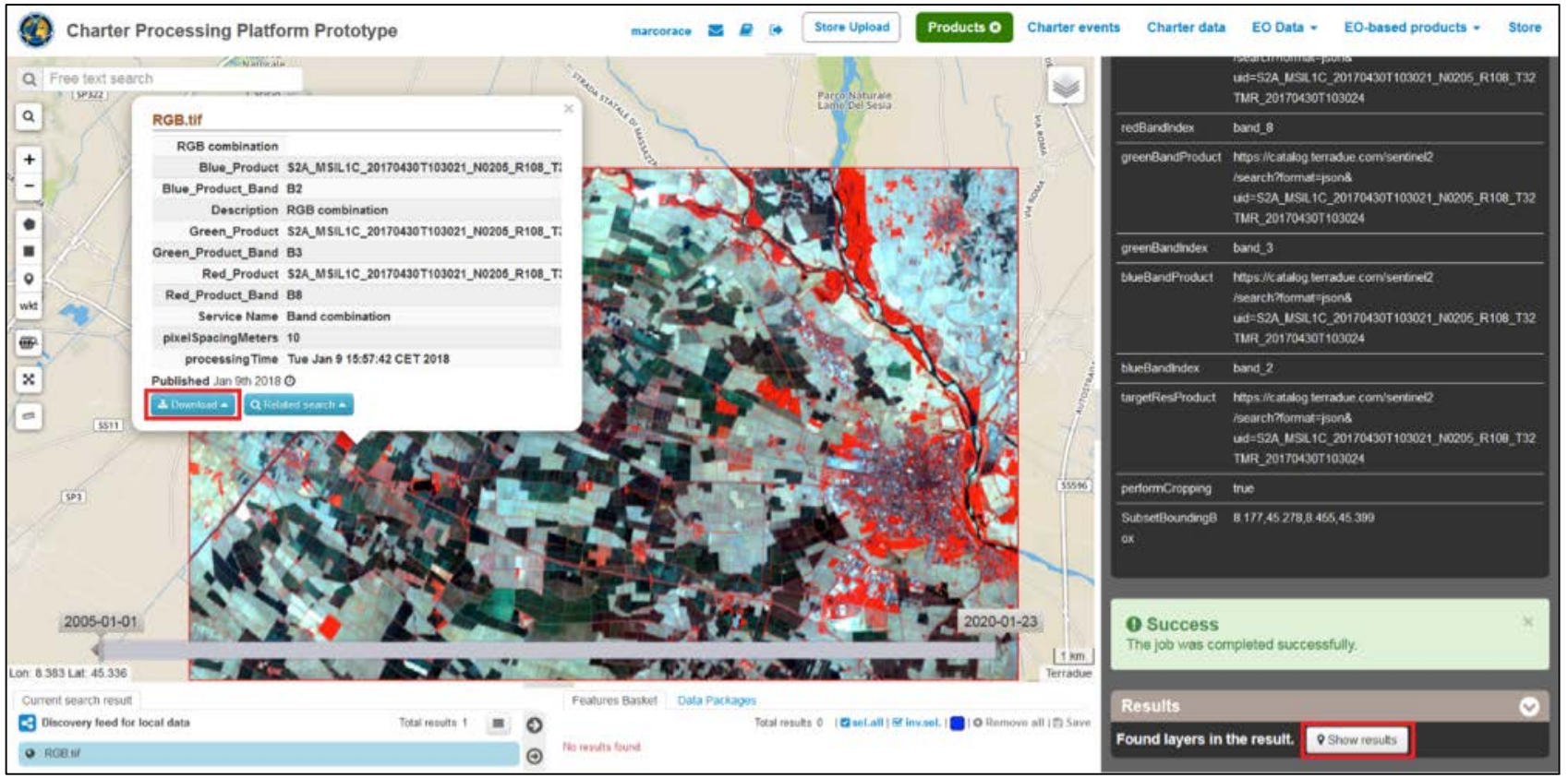
https://terradue.github.io/doc-tep-geohazards/tutorials/rss_snap_s1_coin.html



COMBI Band Combination

RGB band combination from single or multiple EO data product user-defined band combinations from multi-mission Optical and SAR data.

Missions: ALOS, ALOS-2, Kanopus-V, KOMPSAT-2, KOMPSAT-3, KOMPSAT-5, GF2, Landsat 8, Pleiades 1A/1B, RADARSAT-2, RapidEye, Resurs-P, Sentinel-1, Sentinel-2, SPOT 6, SPOT 7, TerraSAR-X, VRSS1 and UK-DMC 2.

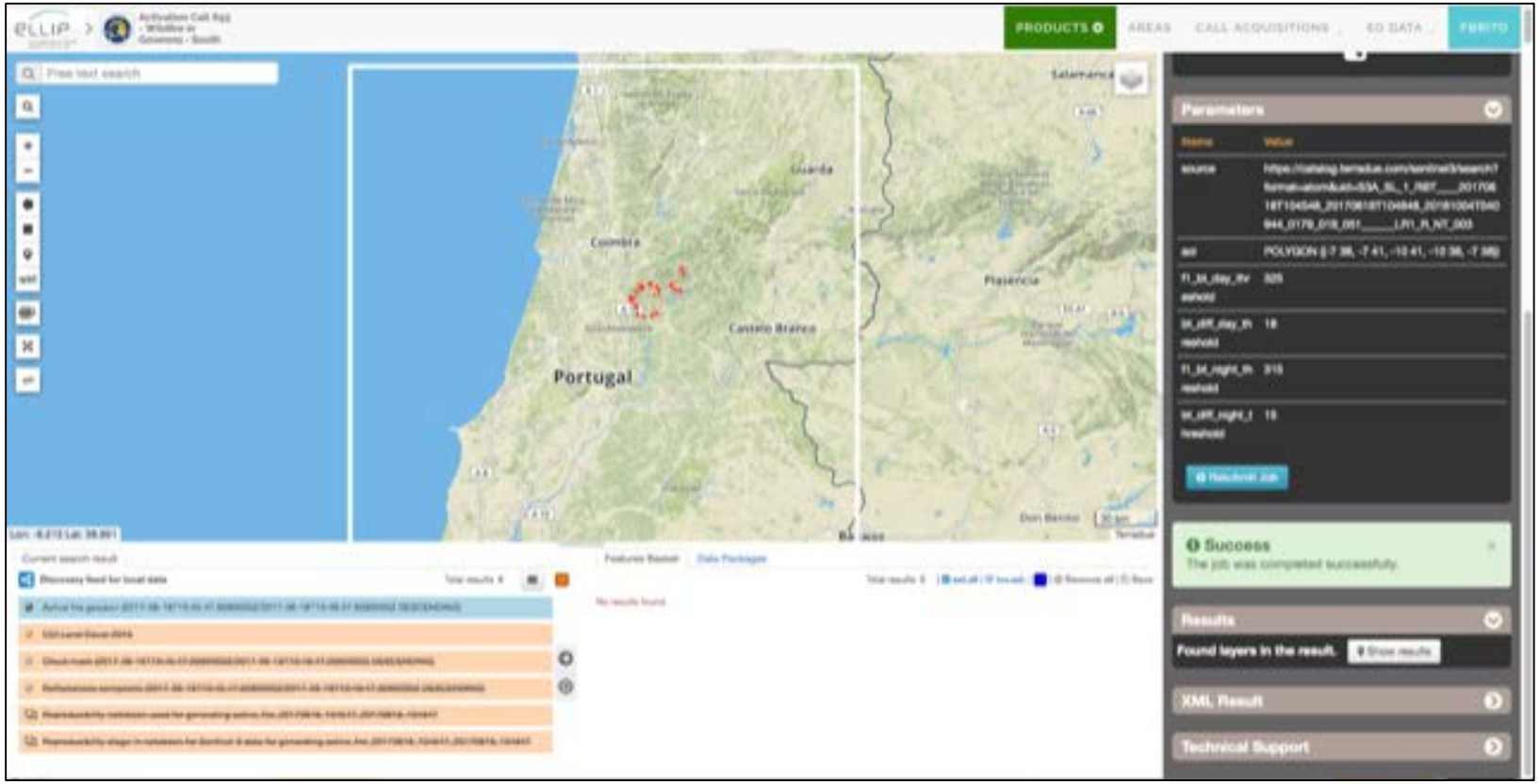


https://terradue.github.io/doc-tep-

Active fire detection with Sentinel-3

Uses Sentinel-3 SLSTR to detect hot spots to generate a product including:

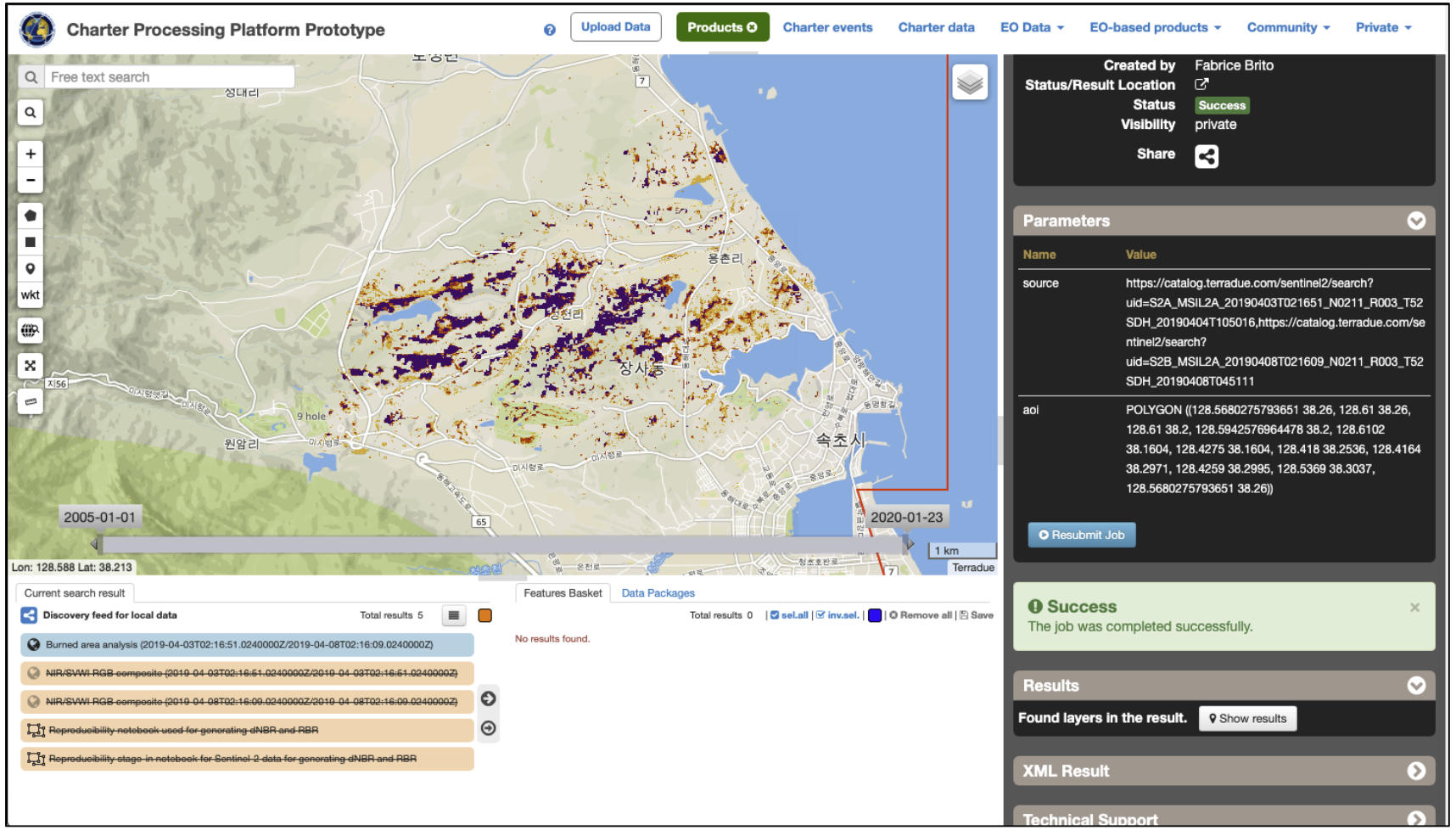
- A geojson with the hotspots
- A RGB composite for descending acquisitions
- A cloud mask
- CCI Land Cover



Burned area assessment with Sentinel-2

Sentinel-2 burned area assessment including:

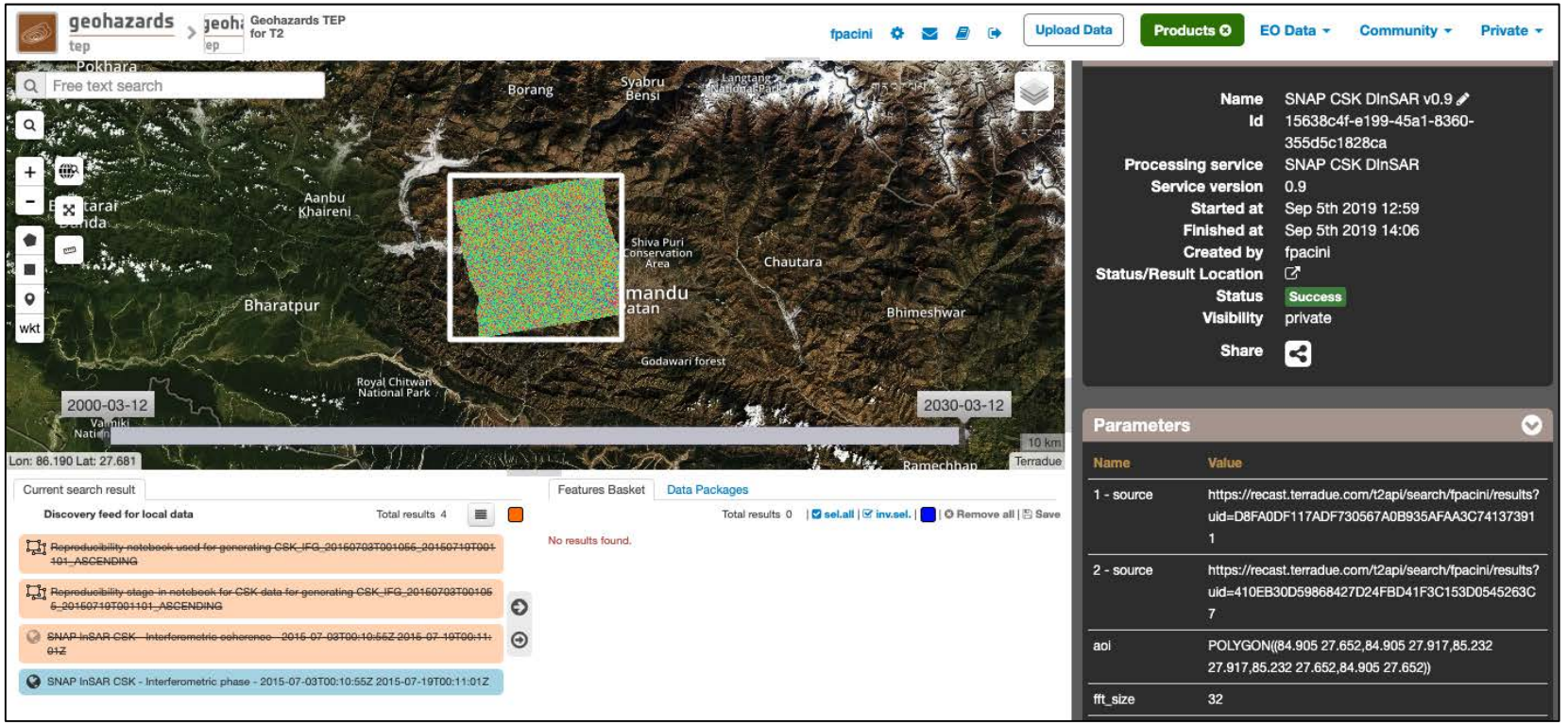
- RGB composite B12, B11, B8A
- dNBR delta normalized burn ratio
- RBR Relativized Burn Ratio



CSK interferogram generation

COSMO-SkyMed Differential SAR Interferometry using SNAP.

This service performs an InSAR workflow on a pair (master, slave) of COSMO-SkyMed single look complex (L1A SCS) acquisitions producing interferograms and coherence map



Other software

- **Libraries and toolboxes:**
  - Orfeo Toolbox
  - Gdal
  - pandas/geopandas
  - NumPy
- **Environments**
  - Jupyter
  - Xarray
  - Dask
- **Processing**
  - YARN, OOZIE
  - Kubernetes

TERRAQUE
Advancing Earth Science

The importance of Snappy

"I CERTAINLY didn't set out to create a language that was intended for mass consumption," says **Guido van Rossum**, a Dutch computer scientist who devised **Python**, a programming language, in 1989.

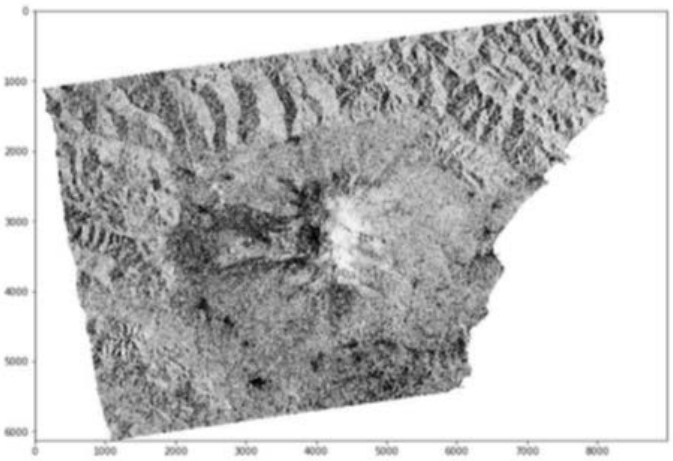
But nearly three decades on, his invention has overtaken almost all of its rivals and brought coding to the fingertips of people who were once baffled by it.

https://www.economist.com/graphic-detail/2018/07/26/python-is-becoming-the-worlds-most-popular-coding-language

Data Exploitation: Access to EO Products

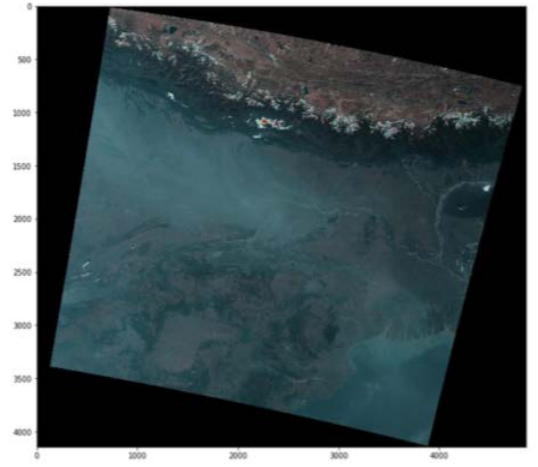
Using Snappy in notebooks to plot and analyse
Sentinel data

```
In [5]: import numpy as np
        import matplotlib
        import matplotlib.pyplot as plt
        import matplotlib.colors as colors

        %matplotlib inline

        def plotBand(product, band, vmin, vmax):

            band = product.getBand(band)

            w = band.getRasterWidth()
            h = band.getRasterHeight()

            band_data = np.zeros(w * h, np.float32)
            band.readPixels(0, 0, w, h, band_data)

            band_data.shape = h, w

            width = 12
            height = 12
            plt.figure(figsize=(width, height))
            imgplot = plt.imshow(band_data, cmap=plt.cm.binary, vmin=vmin, vmax=vmax)

            return imgplot

        plotBand(terrain, 'Sigma0_' + polarization, 0, 0.3)
Out[5]: <matplotlib.image.AxesImage at 0x7f6e2370e990>
```



- Step 4: Plot an RGB image

```
In [6]: red_radiance = reproject.getBand('Oa08_radiance')
        green_radiance = reproject.getBand('Oa05_radiance')
        blue_radiance = reproject.getBand('Oa04_radiance')

        w = red_radiance.getRasterWidth()
        h = red_radiance.getRasterHeight()

        red_radiance_data = np.zeros(w * h, np.float32)
        red_radiance.readPixels(0, 0, w, h, red_radiance_data)
        red_radiance_data.shape = h, w

        green_radiance_data = np.zeros(w * h, np.float32)
        green_radiance.readPixels(0, 0, w, h, green_radiance_data)
        green_radiance_data.shape = h, w

        blue_radiance_data = np.zeros(w * h, np.float32)
        blue_radiance.readPixels(0, 0, w, h, blue_radiance_data)
        blue_radiance_data.shape = h, w

        xmax=200
        red = (red_radiance_data*256/(xmax-np.amin(red_radiance_data)))
        green = (green_radiance_data*256/(xmax-np.amin(green_radiance_data)))
        blue = (blue_radiance_data*256/(xmax-np.amin(blue_radiance_data)))

        rgb_uint8 = np.dstack((red, green, blue)).astype(np.uint8)

        width = 12
        height = 12
        plt.figure(figsize=(width, height))
        img = Image.fromarray(rgb_uint8)
        imgplot = plt.imshow(img)
```

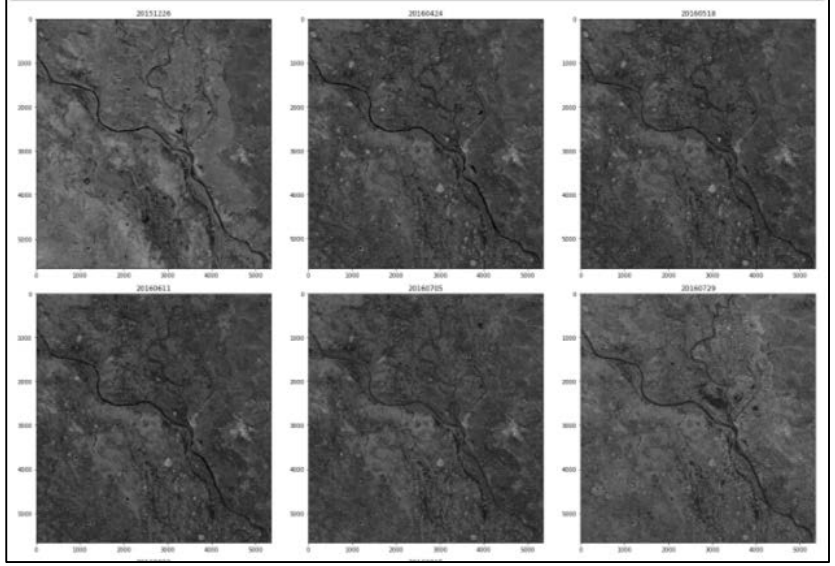


Data Exploitation: Time Series Analyses

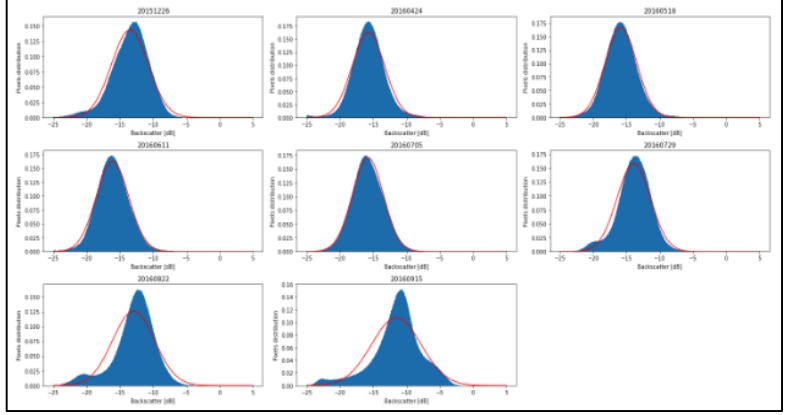
Using Snappy to analyse a stack of GRD Sentinel-1 data

Backscatter profiles for reference image used in flood change detection analysis

```
    band_data.shape = h, w

    imgplot = plt.imshow(band_data, cmap=plt.cm.binary_r, vmin=vmin, vmax=vmax)

    return imgplot


fig = plt.figure(figsize=(20,20))
i = 1

for lineartodb in lineartodbs :

        a=fig.add_subplot(330+i)
        imgplot = plotBand(lineartodb, 'Sigma0_VV_db', -25, 5)
        name = lineartodb.getName()
        timestamp = name.split("_")[5]
        date = timestamp[:8]
        a.set_title(date)
        i = i+1

plt.tight_layout()
fig = plt.gcf()
plt.show()

fig.clf()
plt.close()
gc.collect()
```




```
    band.ReadPixels(0, 0, w, h, band_data)
    band_data.shape = h * w

    ser = np.asarray(band_data, dtype='float')

    m, s = stats.norm.fit(ser) # get mean and standard deviation
    pdf_params.append([m,s])
    p5 = np.percentile(ser, 5) # return 5th percentile.
    p95 = np.percentile(ser, 95) # return 95th percentile.
    index1 = p95 - p5 # anomaly index 1
    index1s.append(index1)

    hist = plt.hist(ser, bins=2048, range=[-25, 5], normed=True)
    pdf_g = stats.norm.pdf(lnspc, m, s) # now get theoretical values in our interval
    pdf = plt.plot(lnspc, pdf_g, label="Norm", c='r') # plot it
    plt.xlabel('Backscatter [dB]')
    plt.ylabel('Pixels distribution')

    return hist

index1s = []
fig = plt.figure(figsize=(20,10))
i = 1

for lineartodb in lineartodbs :

        a=fig.add_subplot(330+i)
        a.patch.set_alpha(0.7)
        plotHistdB(lineartodb, 'Sigma0_VV_db')
        a.set_title(dates[i-1])
        i = i+1


plt.tight_layout()
fig = plt.gcf()
plt.show()
fig.clf()
plt.close()
gc.collect()
```



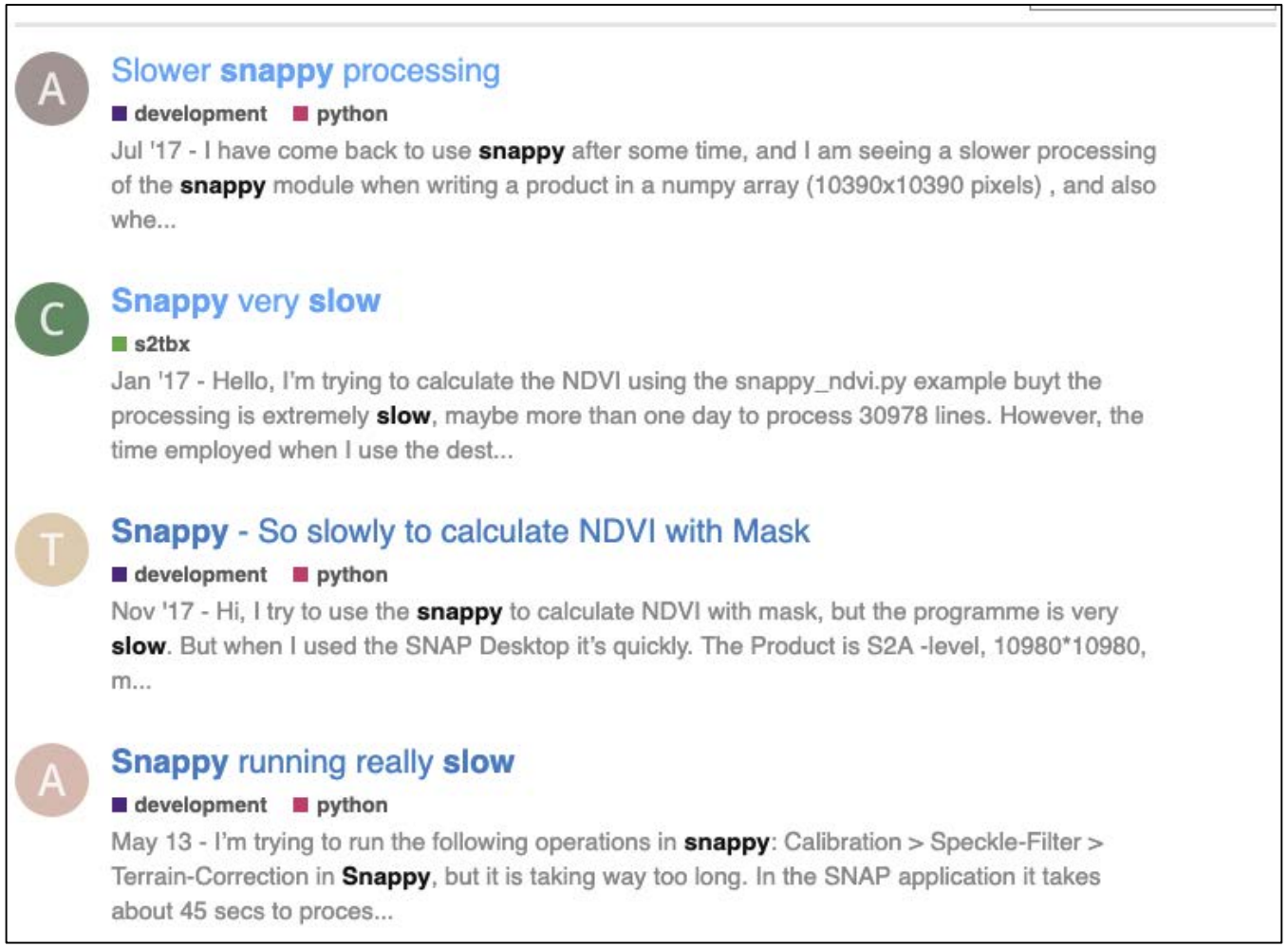
What doesn’t work: Snappy

Snappy is slow compared to GPT

Snappy has memory issues, results with Snappy are not the same as with GPT

Conclusion:

Snappy is barely used in our services, maybe just for inspecting band names for example. Sad.



A

Slower snappy processing

development python

Jul '17 - I have come back to use **snappy** after some time, and I am seeing a slower processing of the **snappy** module when writing a product in a numpy array (10390x10390 pixels) , and also whe...

C

Snappy very slow

s2tbx

Jan '17 - Hello, I'm trying to calculate the NDVI using the snappy_ndvi.py example buyt the processing is extremely **slow**, maybe more than one day to process 30978 lines. However, the time employed when I use the dest...

T

Snappy - So slowly to calculate NDVI with Mask

development python

Nov '17 - Hi, I try to use the **snappy** to calculate NDVI with mask, but the programme is very **slow**. But when I used the SNAP Desktop it's quickly. The Product is S2A -level, 10980*10980, m...

A

Snappy running really slow

development python

May 13 - I'm trying to run the following operations in **snappy**: Calibration > Speckle-Filter > Terrain-Correction in **Snappy**, but it is taking way too long. In the SNAP application it takes about 45 secs to proces...

TERRAQUE
Advancing Earth Science

What doesn't work: Snappy

```python
class GraphProcessor():

    def __init__(self, wdir='.'):
        self.root = etree.Element('graph')

        version = etree.SubElement(self.root, 'version')
        version.text = '1.0'
        self.pid = None
        self.p = None
        self.wdir = wdir

    def view_graph(self):

        print(etree.tostring(self.root , pretty_print=True))

    def add_node(self, node_id, operator, parameters, source):

        xpath_expr = '/graph/node[@id="%s"]' % node_id

        if len(self.root.xpath(xpath_expr)) != 0:

            node_elem = self.root.xpath(xpath_expr)[0]
            operator_elem = self.root.xpath(xpath_expr + '/operator')[0]
            sources_elem = self.root.xpath(xpath_expr + '/sources')[0]
            parameters_elem = self.root.xpath(xpath_expr + '/parameters')

            for key, value in parameters.iteritems():

                if key == 'targetBandDescriptors':

                    parameters_elem.append(etree.fromstring(value))

                else:
                    p_elem = self.root.xpath(xpath_expr + '/parameters/%s' % key)[0]

                    if value is not None:
                        if value[0] != '<':
                            p_elem.text = value
                        else:
                            p_elem.text.append(etree.fromstring(value))
```

```python
    def run(self):

        os.environ['LD_LIBRARY_PATH'] = '.'

        print('Processing the graph')

        fd, path = tempfile.mkstemp()

        try:

            self.save_graph(filename=path)
            options = ['/opt/snap/bin/gpt',
               '-x',
               '-c',
               '2048M',
               path]

            p = subprocess.Popen(options,
                stdout=subprocess.PIPE, stdin=subprocess.PIPE, stderr=subprocess.PIPE)

            print('Process PID: %s' % p.pid)
            res, err = p.communicate()
            print (res, err)
        finally:
            os.remove(path)

            print('Done.')
```

We all had to write code to generate the graph XML and do a system call to /opt/snap/bin/gpt

TERRAQUE
Advancing Earth Science

What doesn't work: Snappy

```
def burned_area(**kwargs):

    options = dict()

    operators = ['Read',
                 'BandMaths',
                 'Write']

    for operator in operators:

        print 'Getting default values for Operator {}'.format(operator)
        parameters = get_operator_default_parameters(operator)

        options[operator] = parameters

    for key, value in kwargs.items():

        print 'Updating Operator {}'.format(key)
        options[key.replace('_', '-')].update(value)

    mygraph = GraphProcessor()

    for index, operator in enumerate(operators):

        print 'Adding Operator {} to graph'.format(operator)
        if index == 0:
            source_node_id = ''

        else:
            source_node_id = operators[index - 1]

        mygraph.add_node(operator,
                         operator,
                         options[operator], source_node_id)

    mygraph.view_graph()

    mygraph.run()
```

```
for index_swath, swath in enumerate(swaths):
    print 'process swath {}'.format(swath)

    # Read
    operator = 'Read'

    source_node_id = ''

    node_id = 'Read'

    parameters = get_operator_default_parameters(operator)
    print products[products.date == date].local_path.values[0]
    parameters['file'] = products[products.date == date].local_path.values[0]

    mygraph.add_node(node_id, operator, parameters, source_node_id)

    # TOPSAR-Split
    operator = 'TOPSAR-Split'

    source_node_id = node_id

    node_id = 'TOPSAR-Split'

    parameters = get_operator_default_parameters(operator)
    parameters['subswath'] = swath
    parameters['selectedPolarisations'] = 'VV'

    mygraph.add_node(node_id, operator, parameters, source_node_id)

    # Apply-Orbit-File
    operator = 'Apply-Orbit-File'

    source_node_id = node_id

    node_id = 'Apply-Orbit-File'

    parameters = get_operator_default_parameters(operator)
    parameters['orbitType'] = 'Sentinel Precise (Auto Download)'

    mygraph.add_node(node_id, operator, parameters, source_node_id)
```

Examples of such "wrapping code"

What is great: gpt processor

SNAP provides an excellent framework to develop additional processors

How to integrate a new processor

Created by Nicolas Ducoin
Last updated 2016-05-27 by Luis Veci

Processors in SNAP are implemented via Graph Processing Framework (GPF) Operators. An Operator consists of the block of code that will manipulate a data product and create a new data product as a result. The operator will be used within the graph processing framework.

A new Operator can be created by extending the base Operator class and implementing the Operator interface. An Operator basically takes a source product as input and creates a new target product within initialize().

The algorithm implementation for what your operator does will go inside **computTile()** or **computeTiles()**.

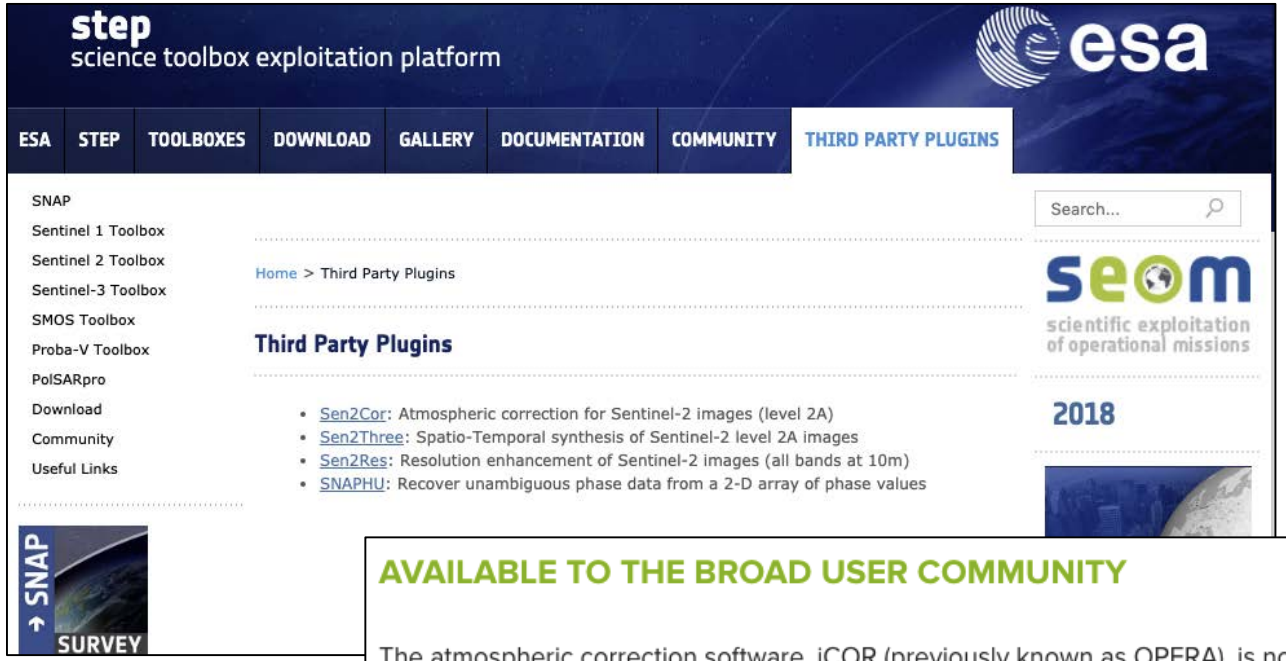
```
public interface Operator {
        OperatorSpi getSpi();
        Product initialize(OperatorContext context);
        void computeTile(Tile targetTile, ProgressMonitor pm);
        void computeTileStack(Rectangle targetTileRectangle, ProgressMonitor pm);
        void dispose();
}
```

What doesn't work: plugins

On the other hand, we see no value in providing plugins.

In this case, SNAP acts as a wrapper on top of CLI applications.

We go straight to the CLI.



AVAILABLE TO THE BROAD USER COMMUNITY

The atmospheric correction software, iCOR (previously known as OPERA), is now available to the broad user community through the ESA Sentinel Application Platform (SNAP) for the **atmospheric correction of Sentinel-2 and Landsat-8 data**.

By implementing iCOR in SNAP, **researchers can test iCOR for their own study areas and experiment** with the different functionalities the new tool has to offer. The iCOR SNAP plug-in can be **freely downloaded**.

iCOR
atmospheric image correction for land and water

Conclusion

Performances

- Processing performance are OK for us
- SNAP Graph developer learning curve could be better: several GOTCHAS here and there (e.g. Thermal Noise Removal)
- Split Graphs is a mandatory strategy

Extra SNAP feature would help

- Cloud Optimized Geotiff as an output format
- BEAM-DIMAP as an internal format
- ‘no data’ support

Does SNAP help you achieve everything that you expect it to?

- A single toolbox will never be able to gather all features and functions we need to build services
- We are used to cherry picking what works from toolboxes and libraries and avoid what doesn’t


TERRAQUE
Advancing Earth Science
Thank you !