

# Pytroll



ECMWF, November 29th 2017

Martin Raspaud, SMHI



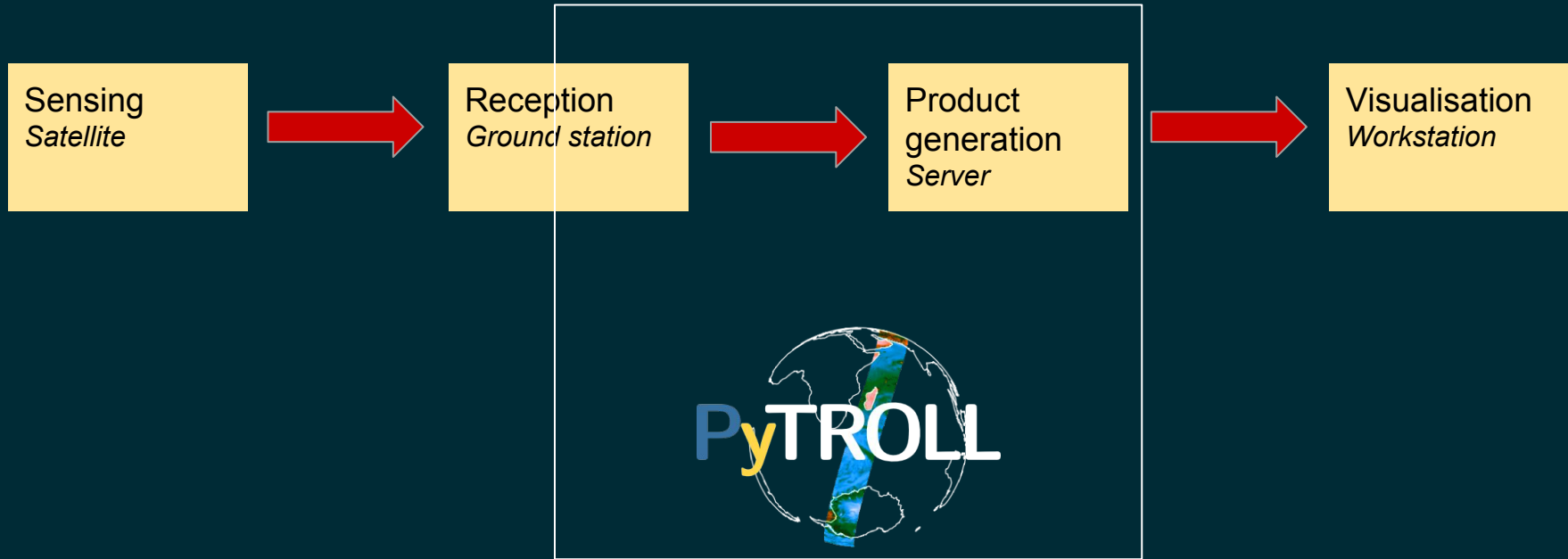
# What is Pytroll ?

- Collection of FOSS python packages (GPLv3)
- For EO satellite data reading/processing/writing



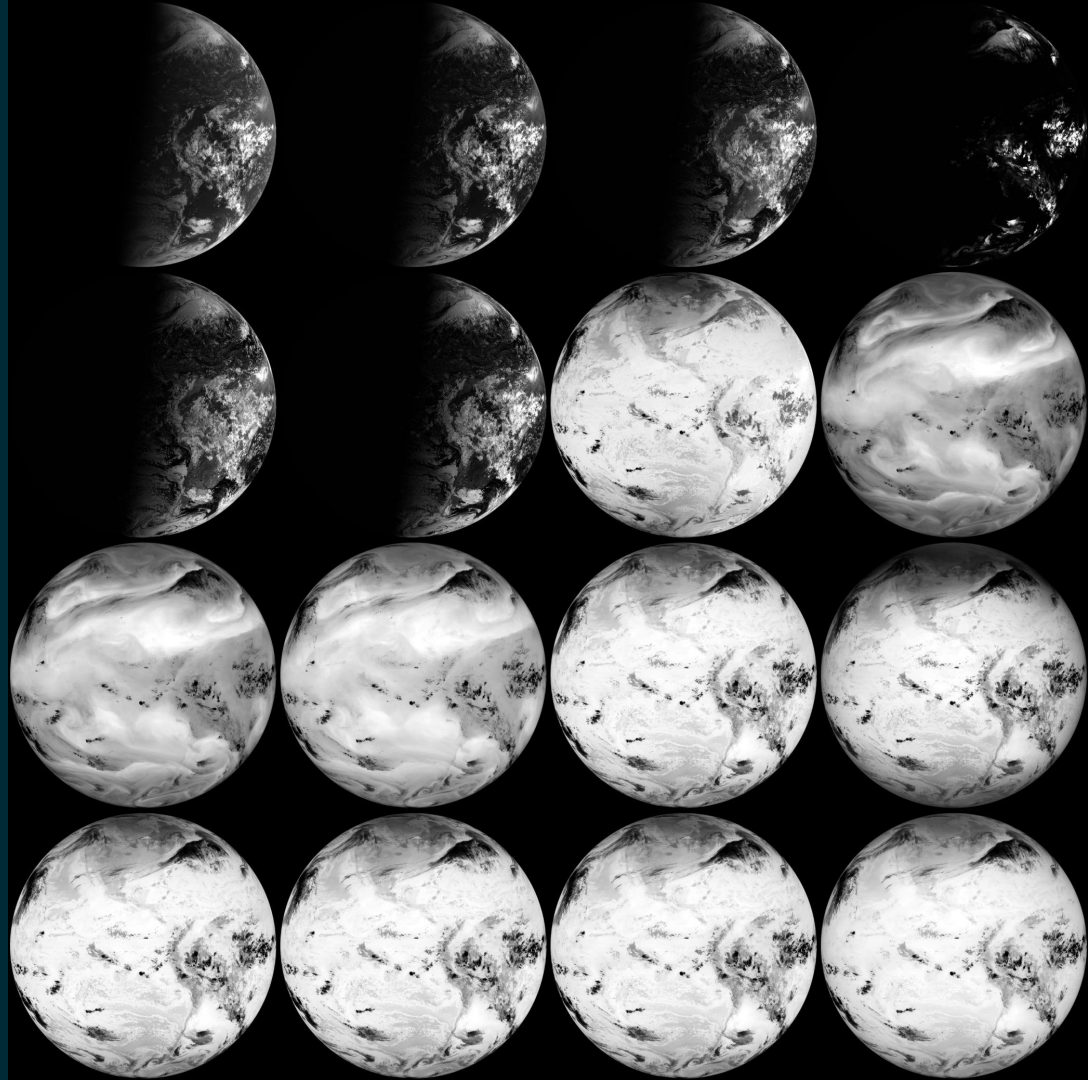
- SatPy
- Pyresample
- Trollflow
- Trollmoves
- Pyorbital
- Pyspectral
- TrollImage
- Pytroll-Schedule
- ...

# Usual workflow



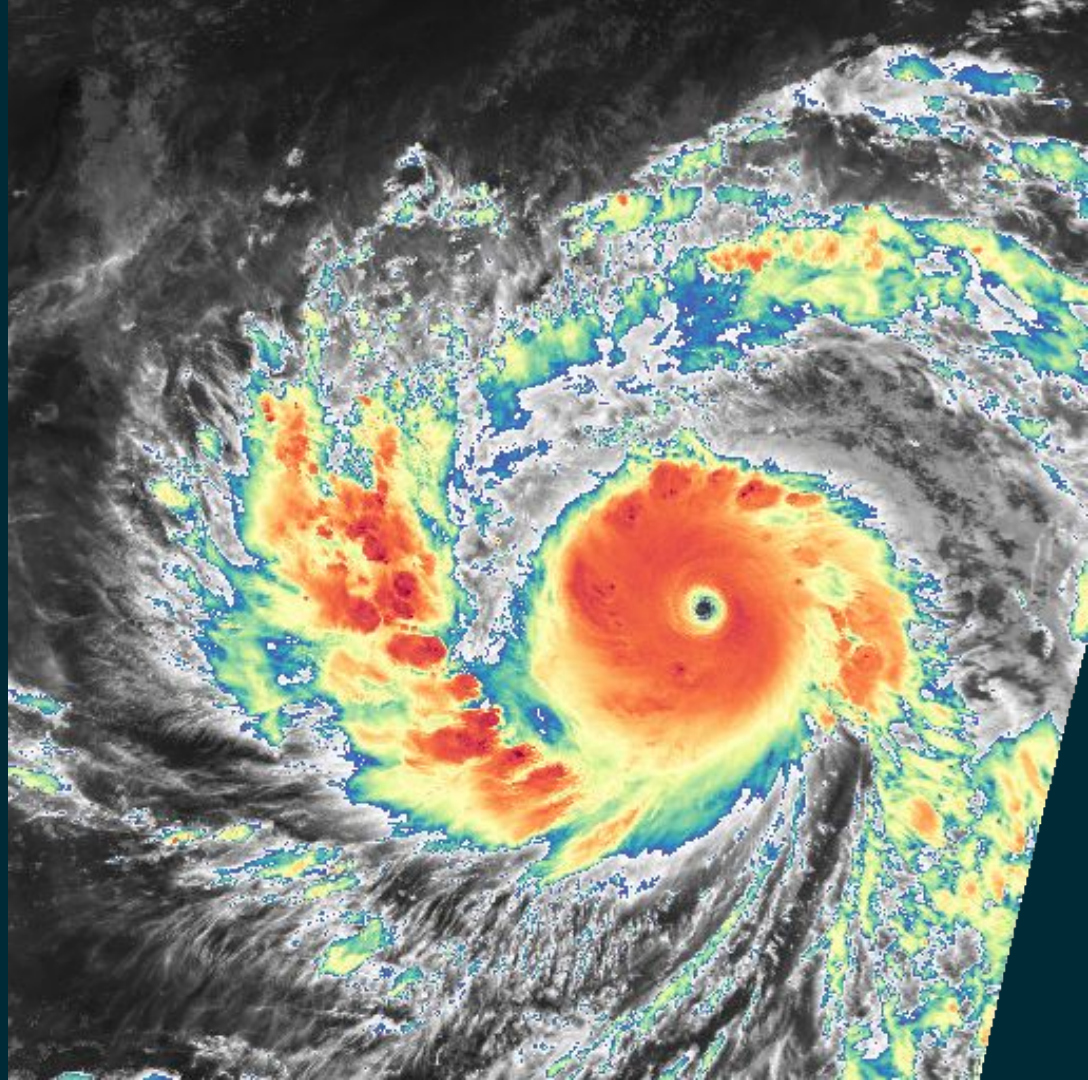
# Weather Satellite Data

- Each channel covers a narrow frequency band
- Different frequencies expose different features



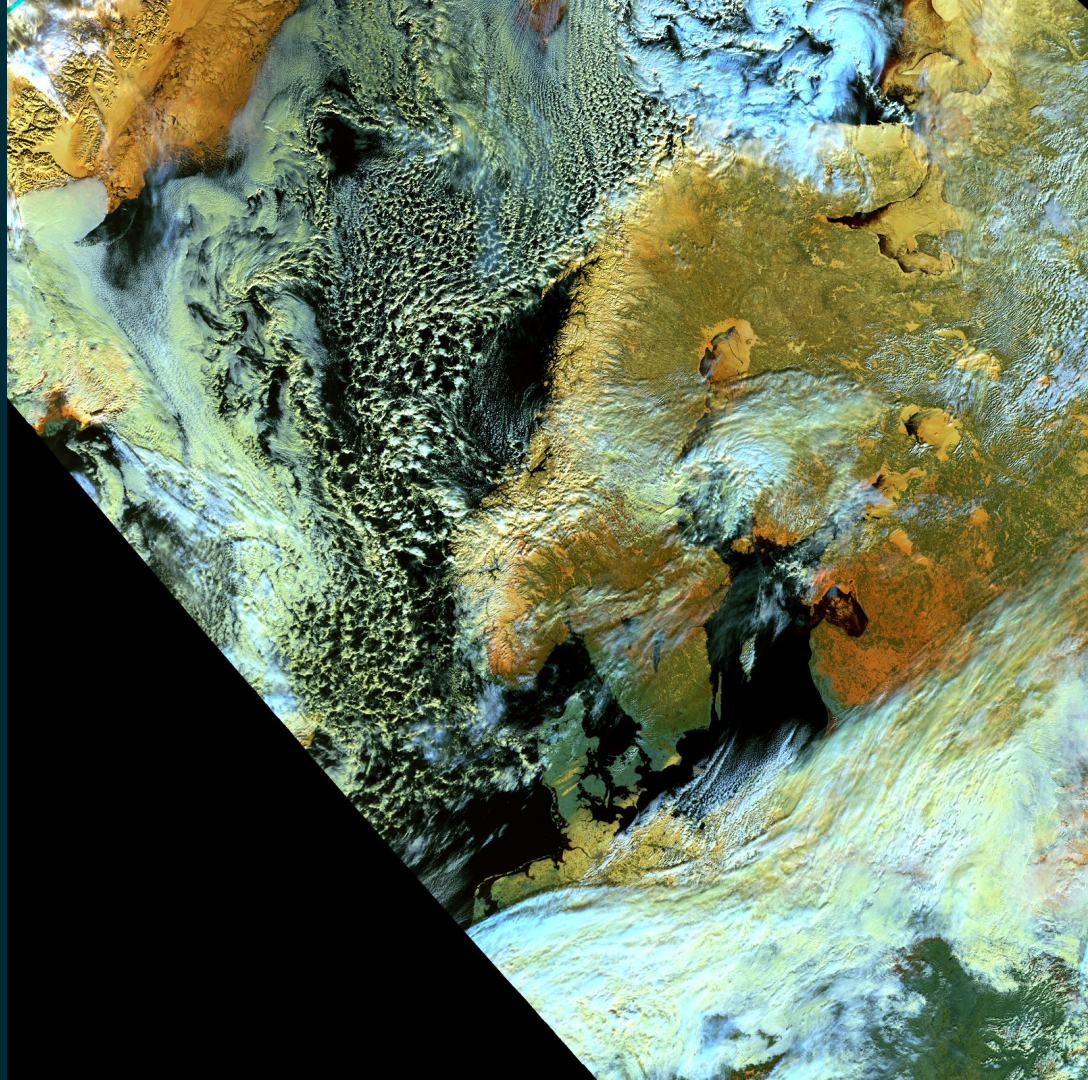
# Satellite Data

Look at a single channel for some feature



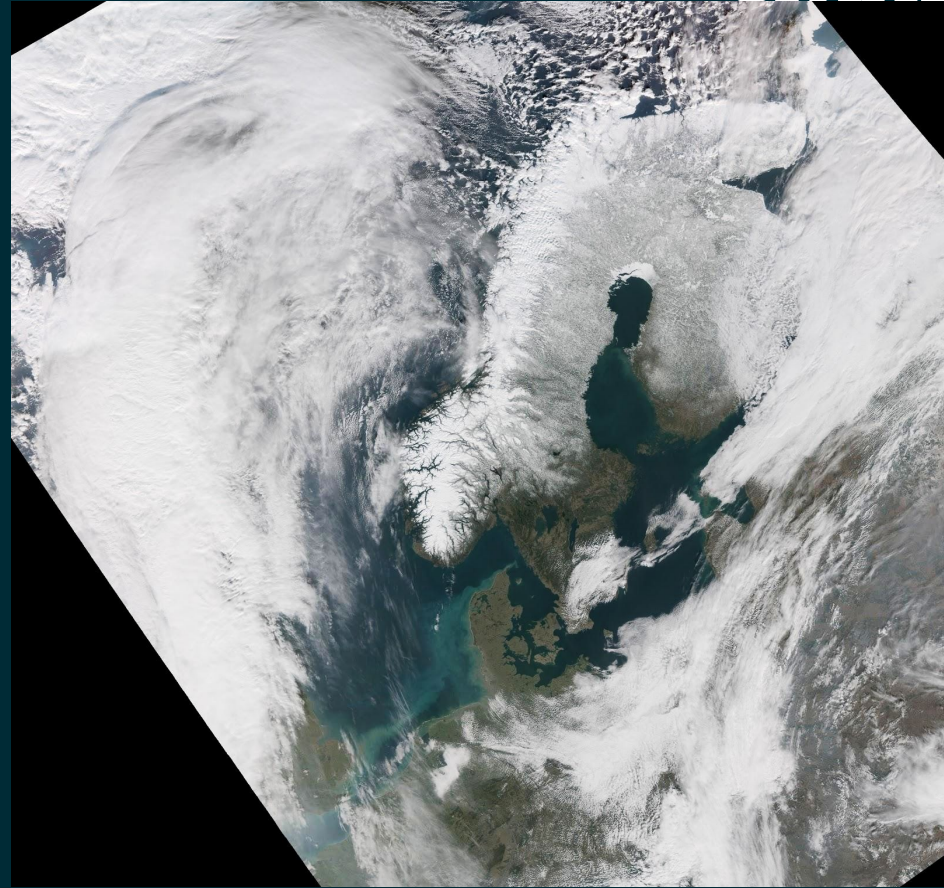
# Satellite Data

Combine channels for more features



# SatPy

- High level processing for satellite data
- Both GEO and LEO
- Indexing by name or wavelength
- Many built-in composites
- Imagers, SAR, sounders



# SatPy example

```
from glob import glob
from satpy.scene import Scene

# Load data by filenames
files = glob("/data/viirs_data/*.h5")
scn = Scene(reader="viirs_sdr", filenames=files)
```



# SatPy example

```
from glob import glob
from satpy.scene import Scene

# Load data by sensor
scn = Scene(sensor="modis", base_dir="/data/terra_data/")
```

# SatPy example

```
# Automatically load composites and their dependencies
scn.load(["true_color"])

# Resample multi-band data to a uniform grid
rs_scn = scn.resample("euron1")

# Save RGB geotiff
rs_scn.save_dataset("true_color")
```

# SatPy example

```
# Load single channels
scn.load(["M15", 0.6])

# Show a channel
scn.show("M15")

# Channel arithmetics
array = scn["M15"] + scn[0.6]
```

# How does it work ?

- Unified reader interface across sensors
- Default (usable) configuration files included
- Default RGB combinations included
- Integrated resampling features
- Replaced the base data structure with xarray and dask support in the github repo (feature-xarray branch)

# dask.distributed

Terminal

Bokeh Application: workers

Dask -- Documentation -- Scheduler Docs -- Tasks -- Workers

Processing and Pending

Memory Usage (%)

#	host	cores	processes	memory	cpu	memory %
0	127.0.0.1	4	1	16 GiB	39.1 %	29.2 %
1	127.0.0.2	4	1	16 GiB	37.8 %	29.2 %
2	127.0.0.3	4	1	16 GiB	35.7 %	29.2 %
3	127.0.0.4	4	1	16 GiB	38.2 %	29.2 %

Bokeh Application: status

Dask -- Documentation -- Scheduler Docs -- Tasks -- Workers

Memory Use: 0.00 MB

Task Stream

Progress -- total: 0, in-memory: 0, processing: 0, ready: 0, waiting: 0, failed: 0

02-custom-etl

jupyter 02-custom-etl (unsaved changes)

```

In [9]: from dask.distributed import Client, progress

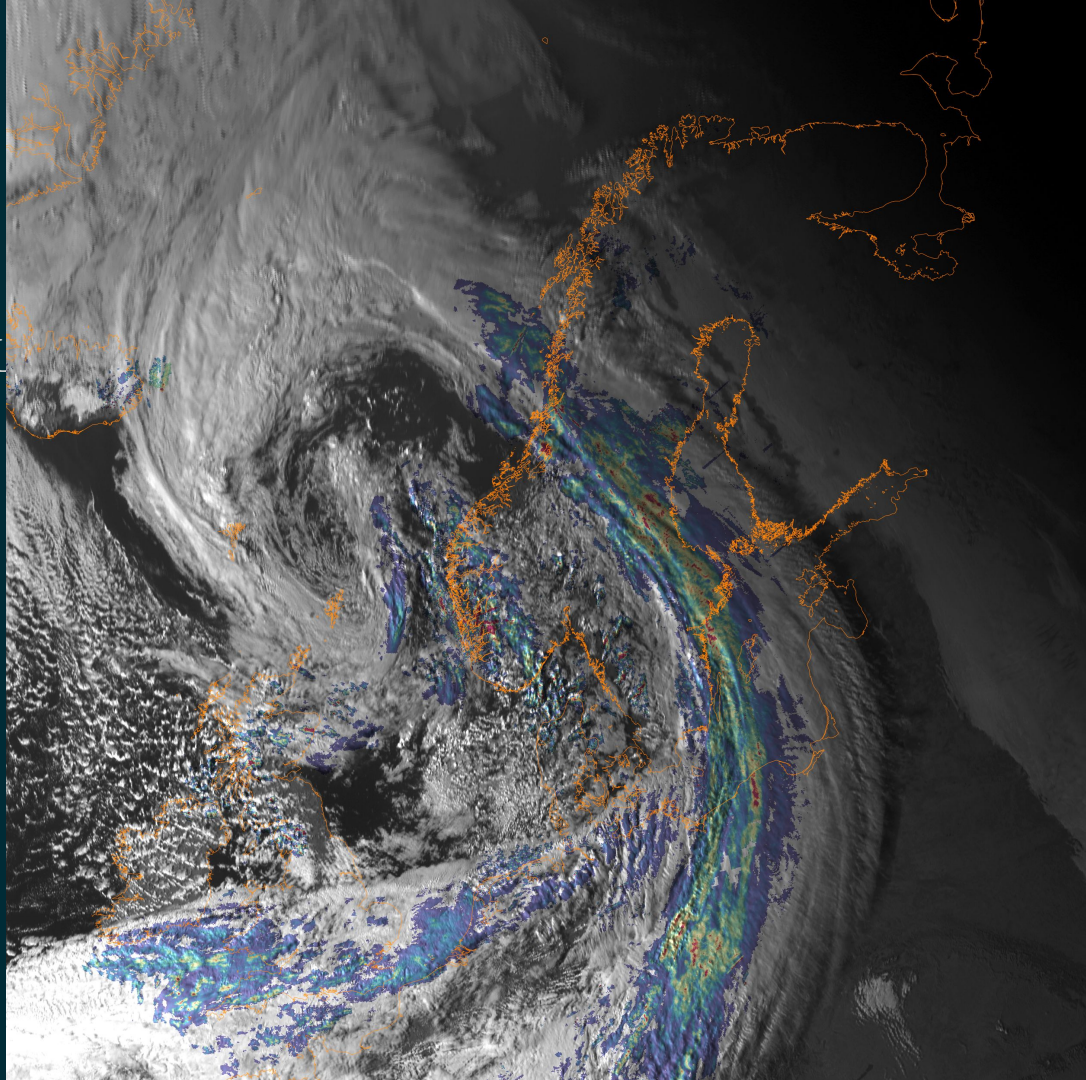
        client = Client('localhost:8786')

        future = client.compute(best)
    
```

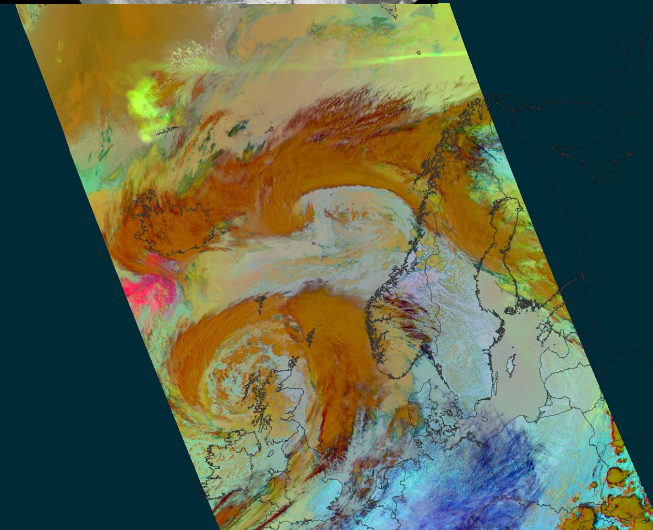
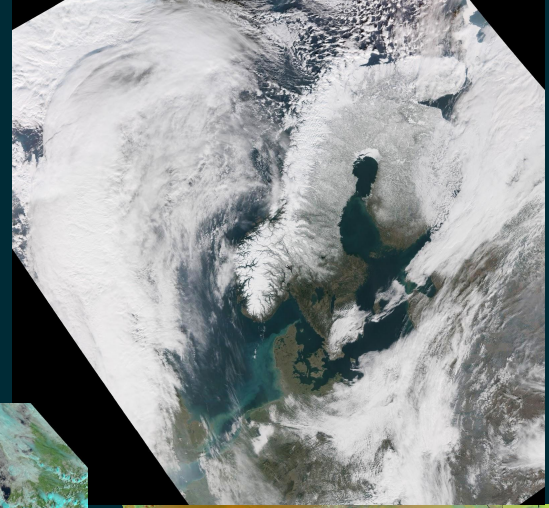
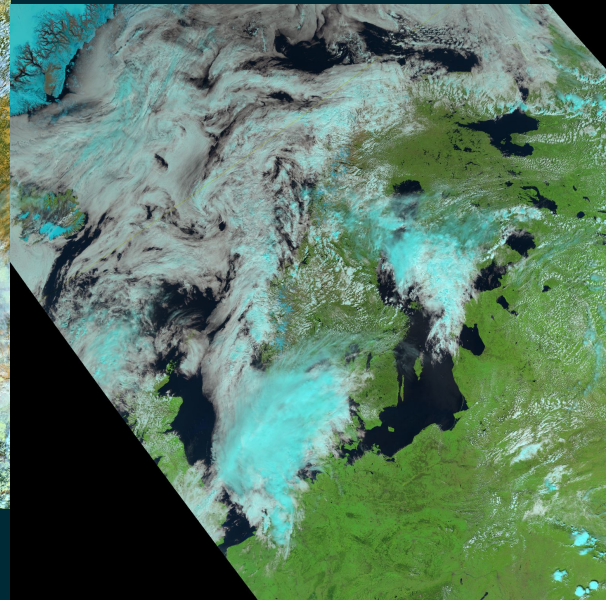
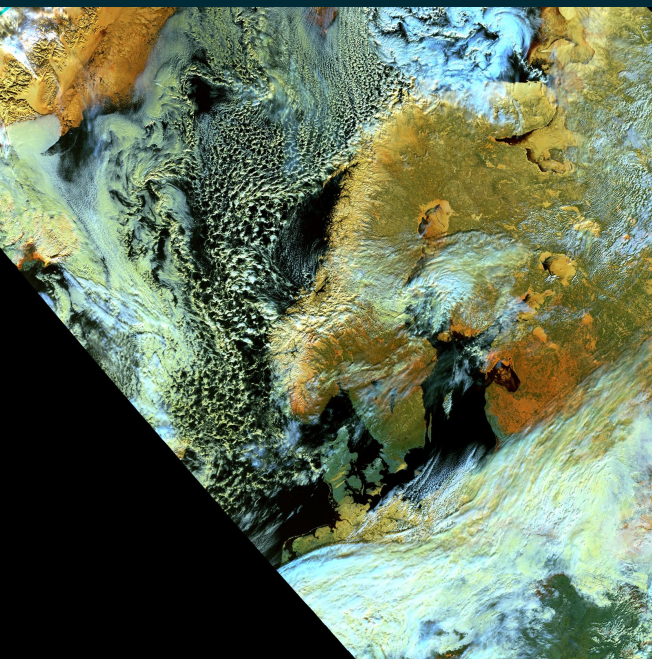
Larger Example

# Satellite data formats

- HDF4 - MODIS L1B
- HDF5 - VIIRS (SDR, Compact, L1)
- HDF5 - FY3
- NetCDF - NWCSAF
- NetCDF - Sentinel & GOES
- AAPP
- NOAA GAC
- Metop PDF
- HRIT/LRIT
- SAR
- **Add your own reader!**

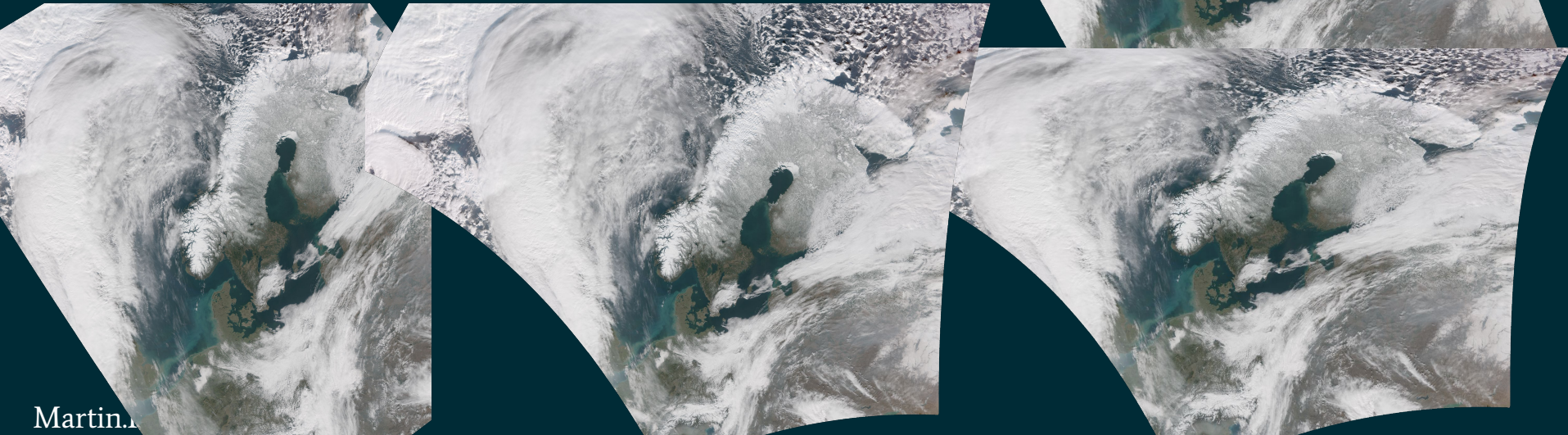


# SatPy built-in composites



# Pyresample for resampling

- Unprojected & Grid
- Proj.4, more than 130 different projections
- NN, bilinear, EWA, gaussian averaging methods





# Satpy and Metadata

```
<xarray.DataArray 'astype-e3d8a3ed181f81e4a0296bd6fe458299' (y: 3712, x: 3712)>  
dask.array<concatenate, shape=(3712, 3712), dtype=float64, chunksize=(53, 3712)>
```

## Coordinates:

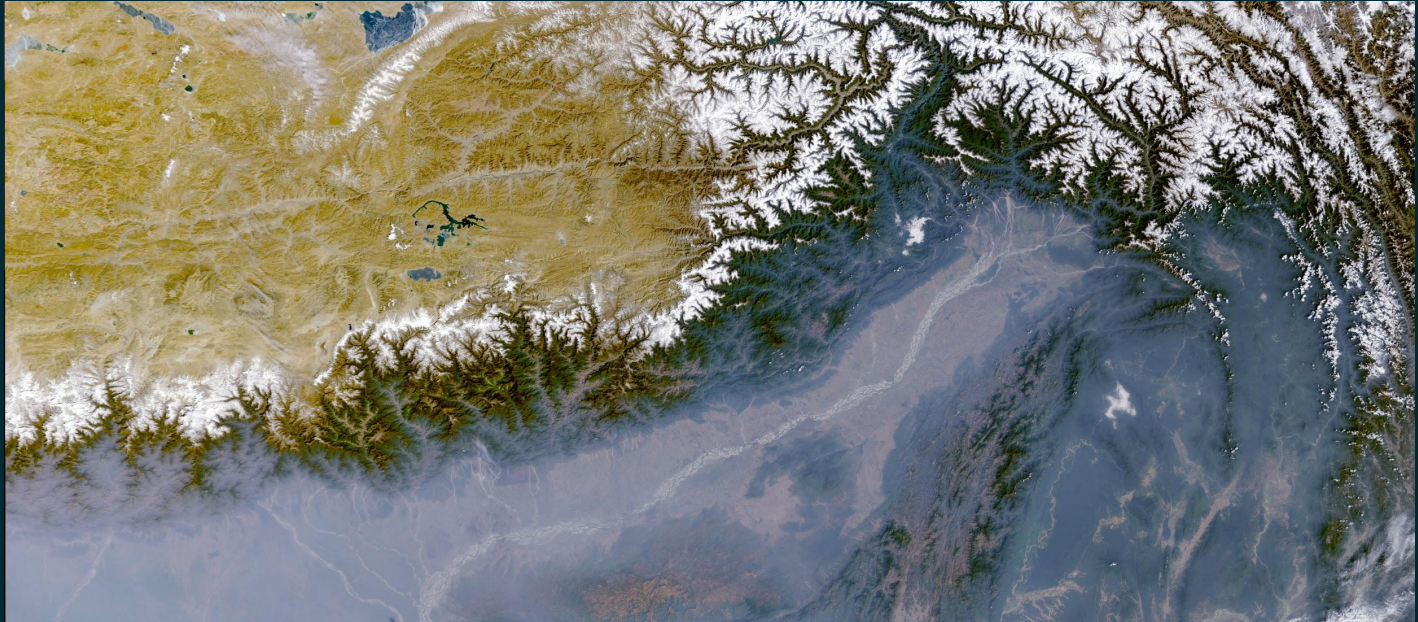
```
* x      (x) float64 -5.569e+06 -5.569e+06 -5.569e+06 -5.569e+06 ...  
* y      (y) float64  5.566e+06  5.566e+06  5.566e+06  5.566e+06  5.566e+06 ...
```

## Attributes:

```
units: %  
wavelength: (0.56, 0.635, 0.71)  
standard_name: toa_bidirectional_reflectance  
platform_name: Meteosat-10  
sensor: seviri  
satellite_longitude: 0.0  
satellite_latitude: 0.0  
satellite_altitude: 35785831.0  
start_time: 2017-01-19 09:30:10.553000  
end_time: 2017-01-19 09:42:41.403000  
area: Area ID: some_area_name\nDescription: On-the-fly ar...  
name: VIS006  
resolution: 3000.40316582  
calibration: reflectance  
polarization: None  
modifiers: ()  
ancillary_variables: []
```

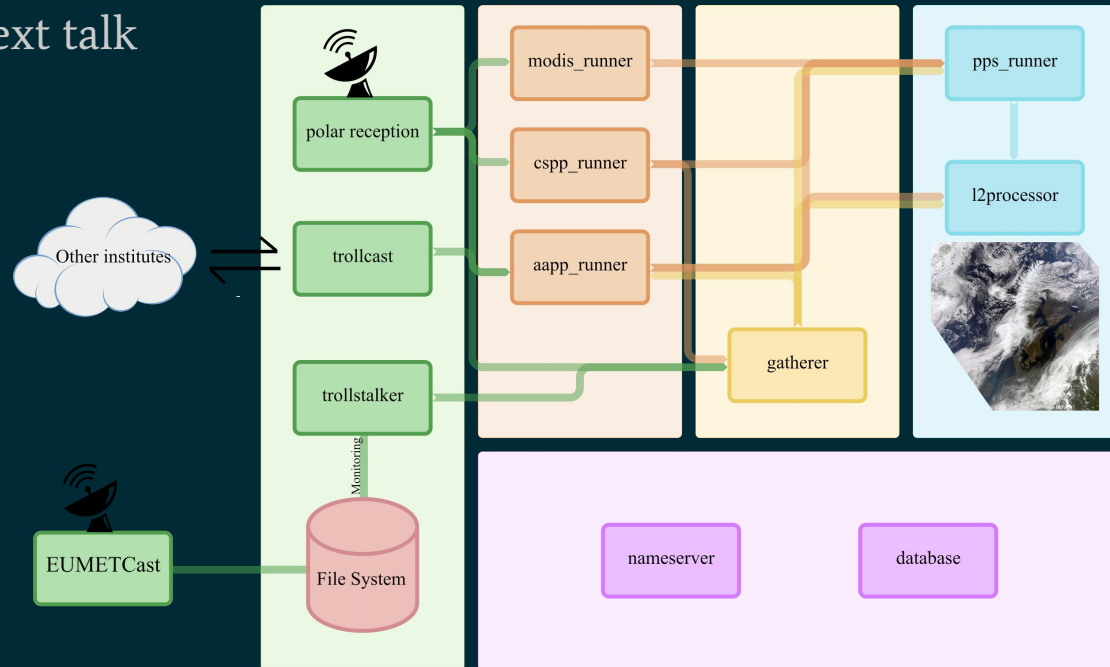
# Results

- Common image formats
- Geotiff, Ninjo
- NetCDF/CF



# Deployment and operations

- Python in operations at SMHI since 1997
- Pip with virtualenv, RPM packages on virtual server
- Running in Trollflow -> see next talk



# Sharing your work

- Talk to us (Slack)
- Github is great

The screenshot shows the GitHub interface for the repository 'pytrol / satpy'. At the top, there are navigation links for 'Pull requests', 'Issues', and 'Gist'. Below this, the repository name is displayed along with statistics: 'Unwatch' (4), 'Star' (1), and 'Fork' (1). The 'Code' tab is selected, showing the repository's description: 'Python for weather satellite data http://satpy.readthedocs.org/en/latest/ — Edit'. A summary bar indicates '1,490 commits', '2 branches', '0 releases', and '9 contributors'. Below this, a section for 'Your recently pushed branches' highlights the 'feature-yaml' branch, pushed 'about 1 hour ago', with a 'Compare & pull request' button. The 'Branch: develop' dropdown is set to 'develop', and there are buttons for 'New pull request', 'Create new file', 'Upload files', 'Find file', and 'Clone or download'. The main content area lists recent commits by 'davidh-ssec', including changes to 'doc', 'etc', 'satpy', 'utils', and various configuration files like '.bumpversion.cfg', '.gitignore', and 'LICENSE.txt'.

Commit	Description	Time
davidh-ssec	Add OMPS so2_trm dataset	Latest commit 811f023 an hour ago
doc	Document the __init__.py files also	4 months ago
etc	Add OMPS so2_trm dataset	an hour ago
satpy	Rename "scaling_factors" to "factor" in reader configuration	18 hours ago
utils	Script to generate external calibration files for AVHRR instruments	a year ago
.bumpversion.cfg	Rename to satpy	4 months ago
.gitignore	Add config files for release utilities	6 months ago
.gitignore	Add .gitignore with python and C patterns	2 months ago
.travis.yml	Post travis notifications to #satpy	4 months ago
LICENSE.txt	Fix rpm building	a year ago
MANIFEST.in	Fix the manifest file to include the config files.	7 months ago
README	Add symlink from README.rst to README.	4 years ago
README.rst	Rename to satpy	4 months ago
changelog.rst	Rename to satpy	4 months ago
setup.cfg	Fix rpm building	a year ago

# Background

- Small teams
- Same needs
- Same tools

# Principles

- Modularity
- Reusability
- Maintainability
- KISS

Ease of use

# Pytroll is FOSS

Release early, release often

Packages available on

- Pypi
- Conda-Forge
- Github

> `pip install satpy`

The screenshot shows the GitHub repository page for `pytroll / satpy`. At the top, there are navigation links for Pull requests, Issues, and Gist. The repository name is `pytroll / satpy`, with 4 Unwatch, 1 Star, and 1 Fork. Below this, there are tabs for Code, Issues (0), Pull requests (0), Wiki, Pulse, Graphs, and Settings. The repository description is "Python for weather satellite data <http://satpy.readthedocs.org/en/latest/> — Edit". The repository statistics show 1,490 commits, 2 branches, 0 releases, and 9 contributors. A section for "Your recently pushed branches" shows a branch named `feature-yaml` pushed about 1 hour ago, with a "Compare & pull request" button. Below this, there are buttons for "Branch: develop", "New pull request", "Create new file", "Upload files", "Find file", and "Clone or download". The main content is a list of files and folders with their commit messages and timestamps:

File/Folder	Commit Message	Timestamp
<code>doc</code>	Document the <code>__init__.py</code> files also	4 months ago
<code>etc</code>	Add OMPS <code>so2_trm</code> dataset	an hour ago
<code>satpy</code>	Rename "scaling_factors" to "factor" in reader configuration	18 hours ago
<code>utils</code>	Script to generate external calibration files for AVHRR instruments	a year ago
<code>.bumpversion.cfg</code>	Rename to satpy	4 months ago
<code>.gitchangelog.rc</code>	Add config files for release utilities	6 months ago
<code>.gitignore</code>	Add <code>.gitignore</code> with python and C patterns	2 months ago
<code>.travis.yml</code>	Post travis notifications to #satpy	4 months ago
<code>LICENSE.txt</code>	Fix rpm building	a year ago
<code>MANIFEST.in</code>	Fix the manifest file to include the config files.	7 months ago
<code>README</code>	Add symlink from <code>README.rst</code> to <code>README</code> .	4 years ago
<code>README.rst</code>	Rename to satpy	4 months ago
<code>changelog.rst</code>	Rename to satpy	4 months ago
<code>setup.cfg</code>	Fix rpm building	a year ago

# Pytroll is in the cloud

- Github
- Readthedocs
- Travis-CI
- Appveyor
- Coveralls
- Landscape.io

...



# International collaboration

## Pytroll Workshops

- SMHI
- DMI
- FMI
- SSEC
- DWD
- MeteoSwiss
- KNMI
- Met.No
- ...



# Thank you!



[pytroll.org](https://pytroll.org)

[pytroll.slack.com](https://pytroll.slack.com)

[pytroll@googlegroups.com](mailto:pytroll@googlegroups.com)

[nbviewer.jupyter.org/github/pytroll/pytroll-examples](https://nbviewer.jupyter.org/github/pytroll/pytroll-examples)