



16/12/2015

# An improved infrastructure for the exploitation of EO-data sources at VITO

*Erwin Goor - VITO nv*

# Company Profile

VITO - Flemish Institute for Technological Research, Belgium

750 employees, 26 nationalities, 140 mio€ turnover in 2014

Remote Sensing unit: 80+ employees

Energy



Materials



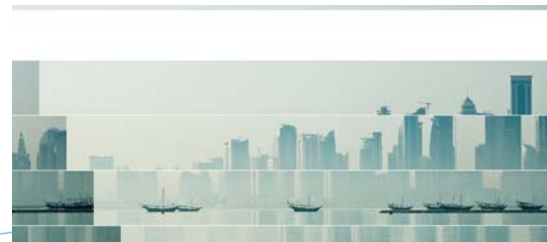
Chemistry



Health



Land use



# VITO Remote Sensing Products and Solutions

## Platforms



UAV



AIRBORN

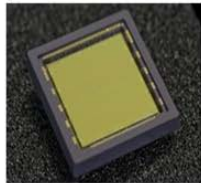


HALE UAV



SATELLITE

## Sensors



## Value Added Services & Information Products



## Markets



Vegetation



Agriculture



Water



Forest



Environment & Security

# PROBA-V User Segment (incl. SPOT-VEGETATION, METOP-AVHRR S10, ...)

<http://proba-v.vgt.vito.be/>



<http://www.vito-eodata.be/>





# Global Land Service

## Global Monitoring of Continental Surfaces

Describe the state, and the dynamics of vegetation, energy budget and water cycle



LAI	VPI
FAPAR	VCI
FCOVER	DMP
NDVI	Burnt Area

- ✓ User-driven service
- ✓ Sustainable system
- ✓ Reliable information
  
- ✓ Basic variables
- ✓ Validated products
- ✓ Near Real Time



**Free and open access to products:**  
<http://land.copernicus.eu/global/access>

Supported by several R&D – operational - capacity building projects

Food security – Agriculture – Re-insurance – Biodiversity – Coastal zone management - ...



Stimulating Innovation for  
Global Monitoring of Agriculture



# The role of VITO in scientific exploitation of EO data

Grow towards

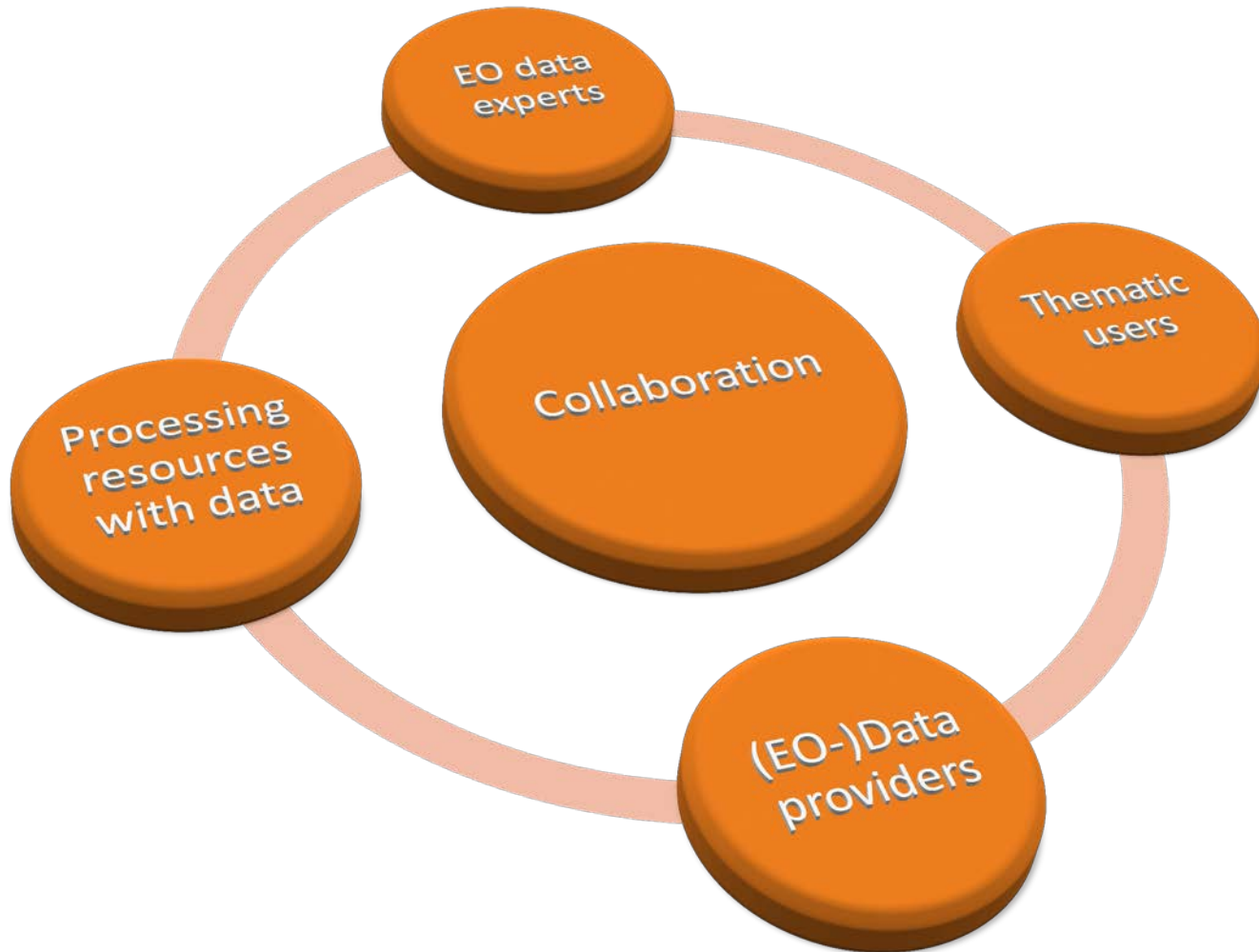
“Scalable R&D/Operations platform”

&

“Information dashboards/tools”

on non-co-located & heterogeneous

BIG DATA – Time series



# SERVICES

analysis services

download services

Processing services

On-demand raster data query

- ✓ On the fly statistical analysis
- ✓ Queries on user defined regions



On-demand processing

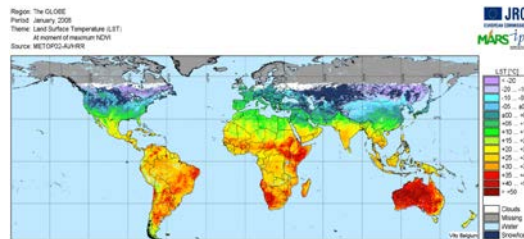
- ✓ Scalable distributed processing
- ✓ Deployed on cloud infrastructures with data locality



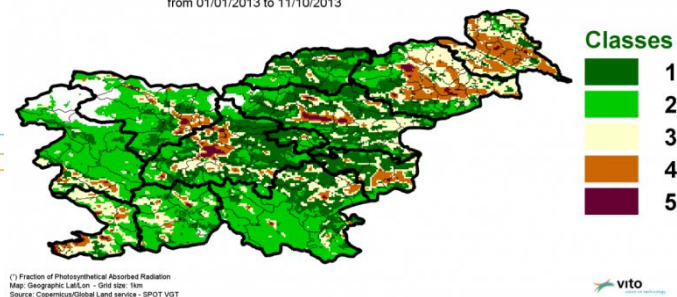
Systematic production & delivery of Earth Observation products

[WWW.VITO-EODATA.BE](http://WWW.VITO-EODATA.BE)

[LAND.COPERNICUS.VGT.VITO.BE](http://LAND.COPERNICUS.VGT.VITO.BE)



SLOVENIA - Remote sensing information support  
 fAPAR(\*) cluster classification  
 from 01/01/2013 to 11/10/2013



16/12/2015

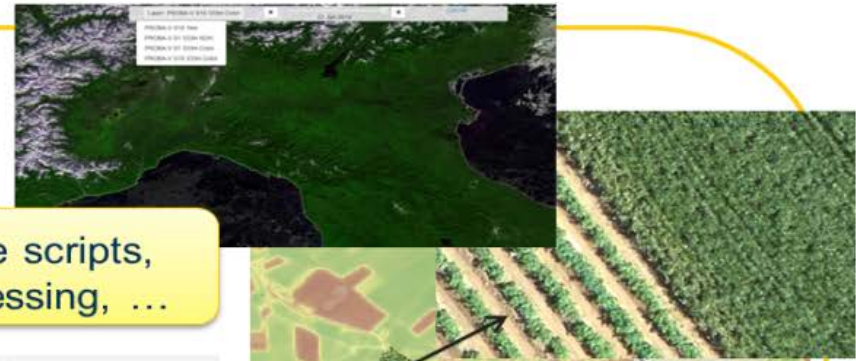
© 2015, VITO NV



# Web-based Applications

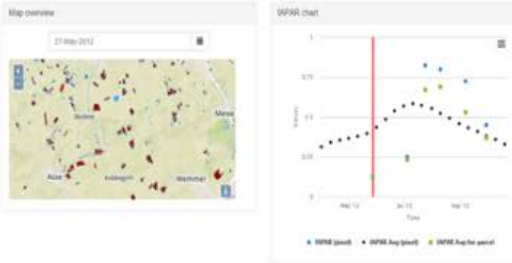


IDE's, Interactive scripts,  
on-demand processing, ...



## Dashboards - Info

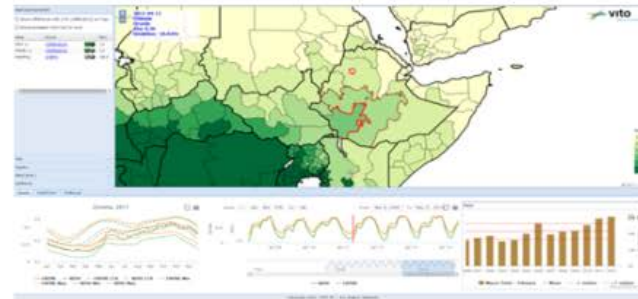
- ✓ Self service
- ✓ highly customisable by users
- ✓ E-collaboration



Full-resolution viewing

## Time Series viewer

- ✓ Near realtime updates
- ✓ Integrate data from various sources
- ✓ Powerful analysis



Data feed  
towards  
3<sup>rd</sup>-party  
applications

<http://tsviewer.vito-eodata.be>

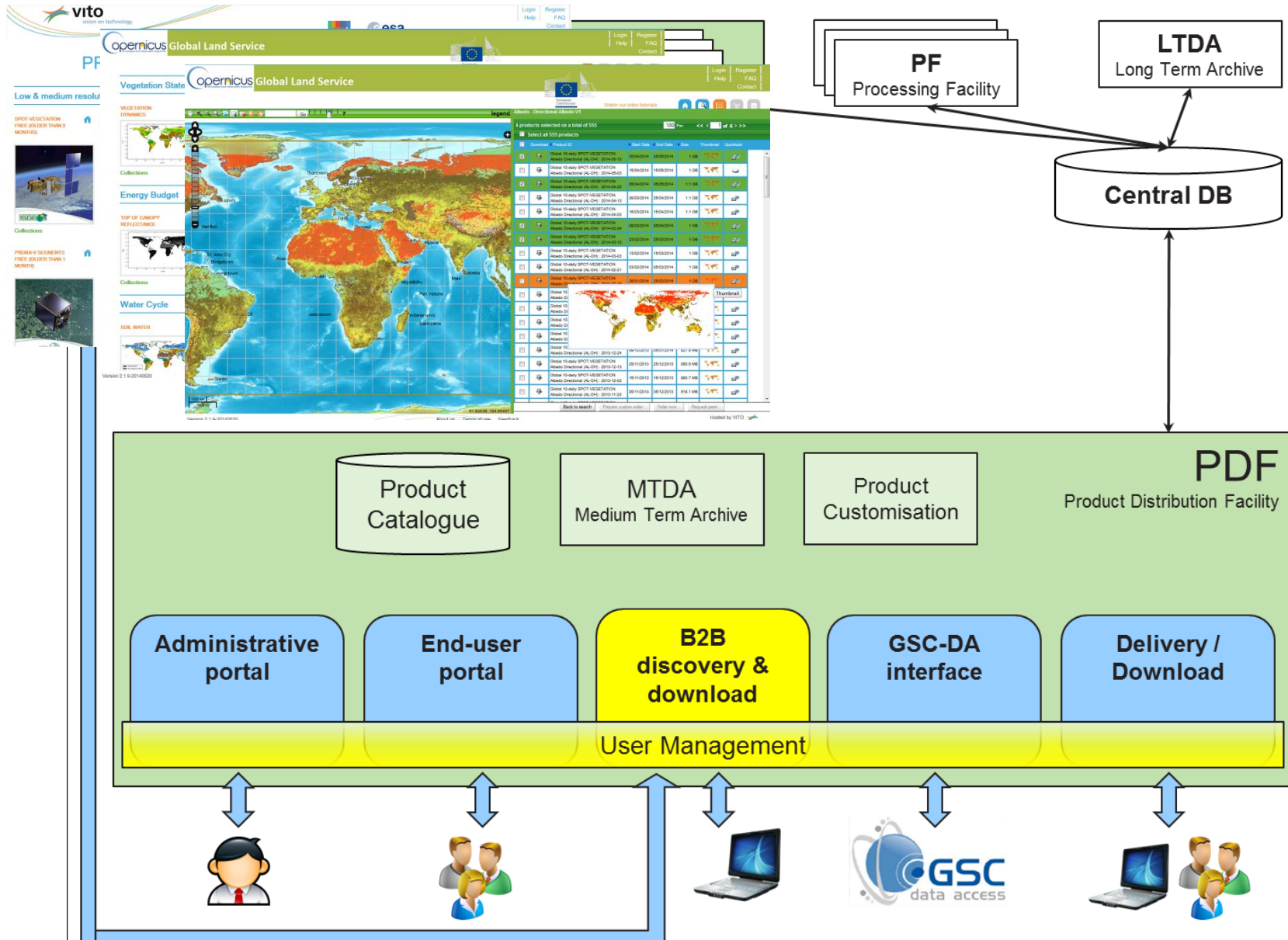
analysis  
service

download  
services

processing  
services

# Product Distribution Facility: data distribution

» Web based portal: <http://www.vito-eodata.be> & [land.copernicus.eu/global/access](http://land.copernicus.eu/global/access)





# Full resolution viewing



# Hadoop based Processing Engine



- » Use of Cloudera/Hortonworks Hadoop distribution for distributed EO-data processing: data locality
- » WPS processes encapsulated in Oozie workflows
- » Use of Spark for fast data extraction

WPS client (VITO or 3<sup>rd</sup> party)

Federated workflows

WPS server

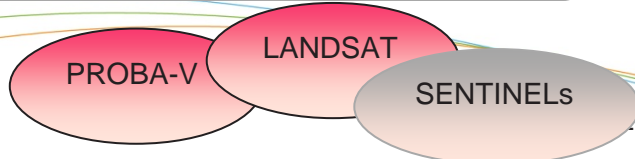
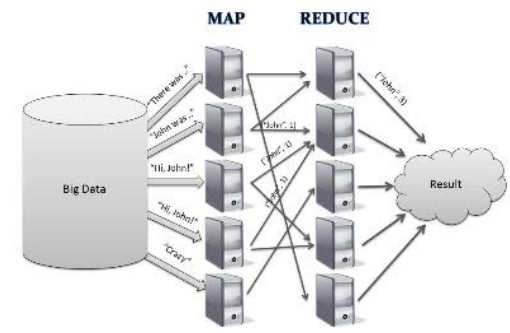
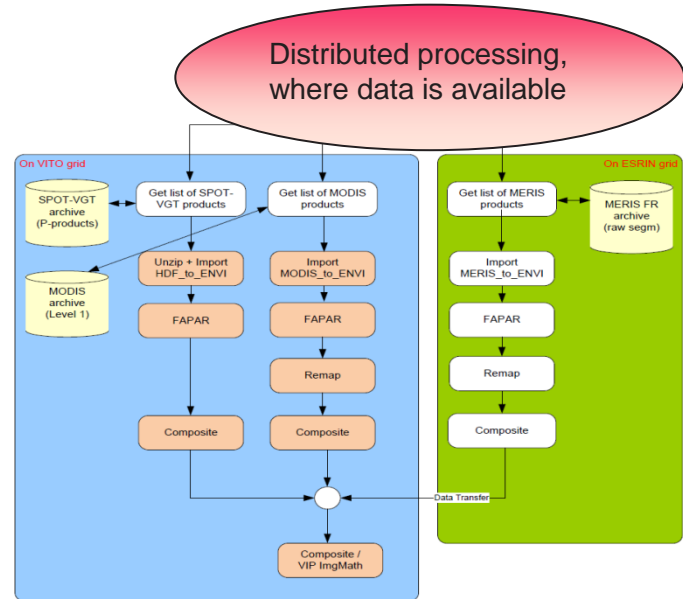
Workflows  
Oozie

Distributed processing  
mapReduce

Distributed storage  
HDFS

Data  
Manager

Data Distribution Facilities





# First 'operational' PoD service

Composite On Demand

Launch job

Your jobs

About

Composite On Demand

Launch job

Your jobs

About

Completed jobs

11/05/2015 13h51 - 'SUCCEEDED'

<b>Bands</b>	ndvi
<b>Composite algorithm</b>	maximumNdvi
<b>Composite frequency</b>	7
<b>Composite unit</b>	3
<b>End date</b>	2014-02-01T00:00:00
<b>File format</b>	geotiff
<b>ROI bottom right X</b>	31.17920
<b>ROI bottom right Y</b>	7.03125
<b>ROI top left X</b>	14.30420
<b>ROI top left Y</b>	19.68750
<b>Spatial resolution</b>	1km
<b>Start date</b>	2014-01-01T00:00:00
<b>Description</b>	<a href="#">empty</a>

**Start of processing** 11/05/2015 13h51

**End of processing** 11/05/2015 13h42

**Status** **SUCCEEDED**

Hide job runs

Relaunch job

« < 1 2 3 4 5 > »

## Job runs

Jobrun 11/05/2015 13h39 - 'SUCCEEDED'

**Start date** 01/01/2014 - **End date** 31/01/2014 - **Status**  
SUCCEEDED

**PrepareComposite:** OK

Start time: 11/05/2015 13h39

End time: 11/05/2015 13h39

**CalculateComposite:** OK

Start time: 11/05/2015 13h39

End time: 11/05/2015 13h42

**UploadResults:** OK

Start time: 11/05/2015 13h42

End time: 11/05/2015 13h42

**DistributeViaPDF:** OK

Start time: 11/05/2015 13h42

End time: 11/05/2015 13h42

**Cleanup:** OK

Start time: 11/05/2015 13h42

End time: 11/05/2015 13h42

## Results

No results available

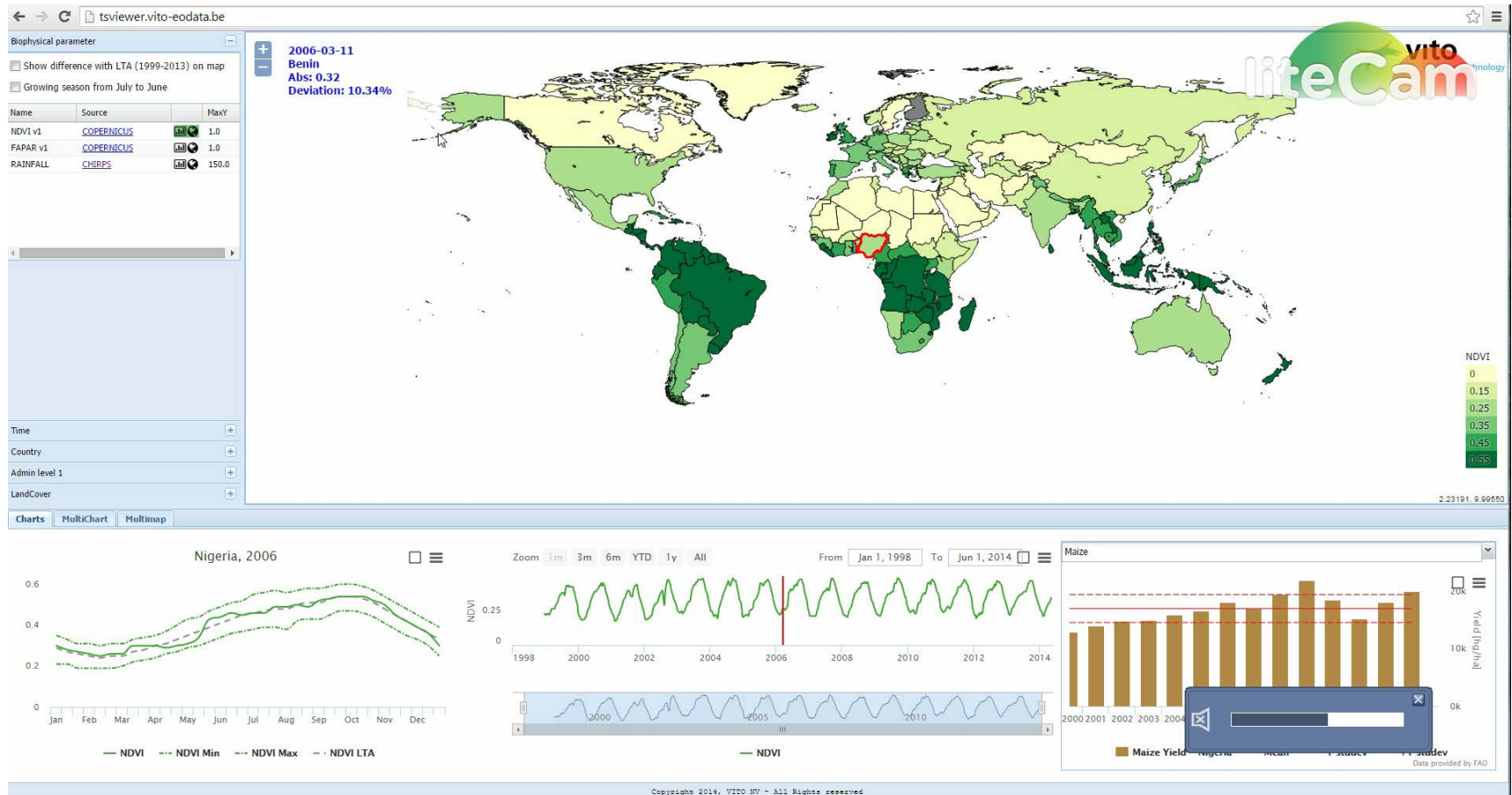
# Develop – debug – test – operate ‘your’ applications

The screenshot shows the Hue Job Browser interface. The top navigation bar includes 'Query Editors', 'Data Browsers', 'Workflows', 'Search', and 'Security'. The 'Job Browser' section is active, displaying a search bar with 'Username: driesj' and a 'Text' search field. A legend at the top right indicates job statuses: Succeeded (green), Running (orange), Failed (red), and Killed (black). The main area contains a table of job logs.

Logs	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted
	1431941104609_4810	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	52s	09/03/15 00:59:20
	1431941104609_4703	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	54s	08/31/15 02:48:24
	1431941104609_4699	be.vito.eodata.ipot.ComputeStatistics	FAILED	driesj	%	%	root.driesj	N/A	46s	08/31/15 02:20:00
	1431941104609_4551	be.vito.eodata.ipot.ComputeStatistics	FAILED	driesj	%	%	root.driesj	N/A	1m:3s	08/18/15 05:28:09
	1431941104609_4550	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	1m:5s	08/18/15 05:22:26
	1431941104609_4225	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	1m:31s	07/14/15 07:34:29
	1431941104609_4224	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	1m:27s	07/14/15 07:00:00
	1431941104609_4166	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	1m:6s	07/14/15 03:11:42
	1431941104609_4164	be.vito.eodata.ipot.ComputeStatistics	SUCCEEDED	driesj	100%	100%	root.driesj	N/A	53s	07/14/15 02:44:41
	1431941104609_4162	be.vito.eodata.ipot.ComputeStatistics	FAILED	driesj	%	%	root.driesj	N/A	48s	07/14/15 02:35:28
	1431941104609_4158	be.vito.eodata.ipot.ComputeStatistics	FAILED	driesj	%	%	root.driesj	N/A	50s	07/14/15

# Global Satellite Time Series Viewer (prototype)

- <http://tsviewer.vito-eodata.be>
- Aim: quickly explore & review satellite time series and complementary indicators for agriculture and environmental monitoring



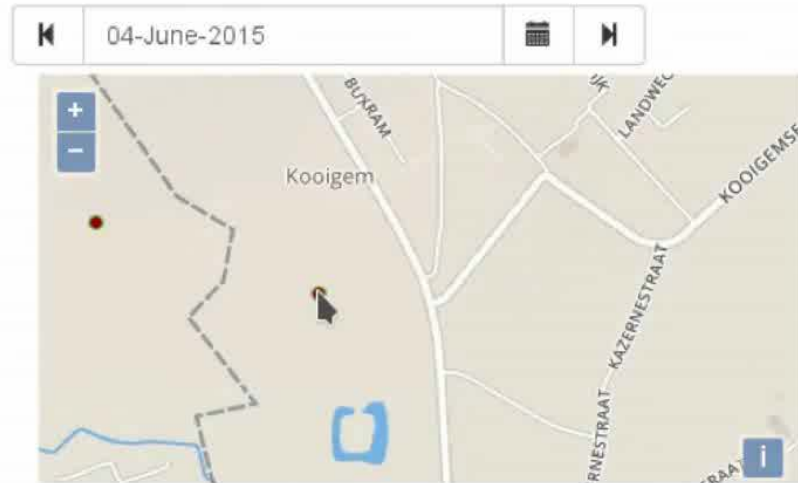
<http://land.copernicus.eu/global/>

<http://www.vito-eodata.be>

[Fa0.org](http://Fa0.org)

# User-tailored dashboards

## Map overview



## DataViewer

### Metadata

<b>Field name</b>	<b>Community</b>
AGT-03 Vremderhoeve	Broechem
<b>Latitude</b>	<b>Longitude</b>
51.1721	4.5659

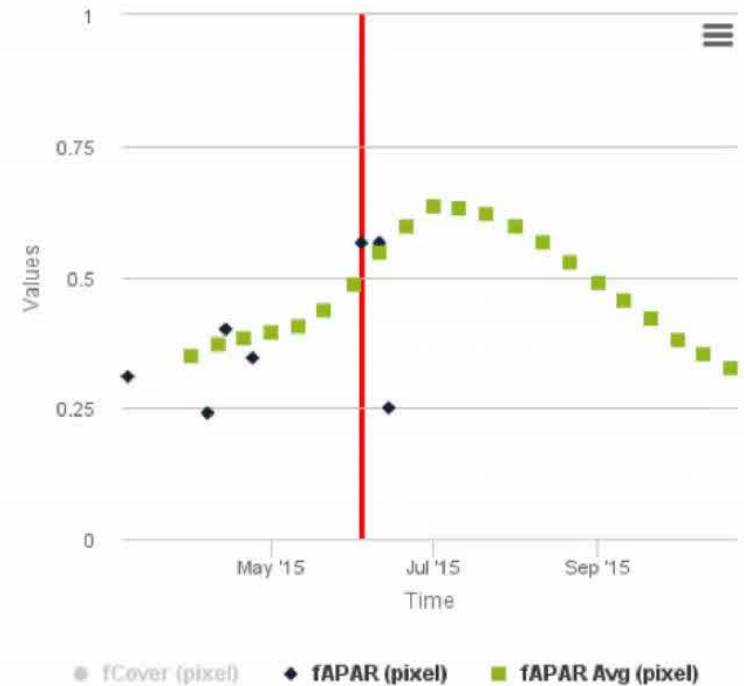
### General parameters

<b>Planting date</b>	<b>Harvest date</b>
2015-04-19	
<b>Plant separation</b>	<b>Row distance</b>
35	75

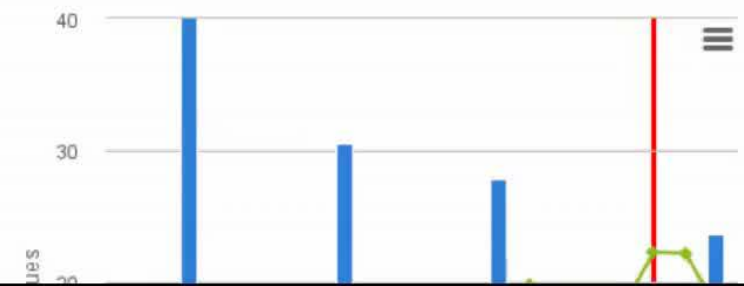
### Fenologies

Time	Date	Remarks
------	------	---------

## fCover/fAPAR



## Highchart





# Support of Python and R on the data/timeseries

```
In [1]: import matplotlib.pyplot as plt
import matplotlib
%matplotlib inline

from elasticsearchapi import ElasticsearchAPI

matplotlib.style.use('ggplot')

username = "ddaems"
password =

es = ElasticsearchAPI(username, password)
```

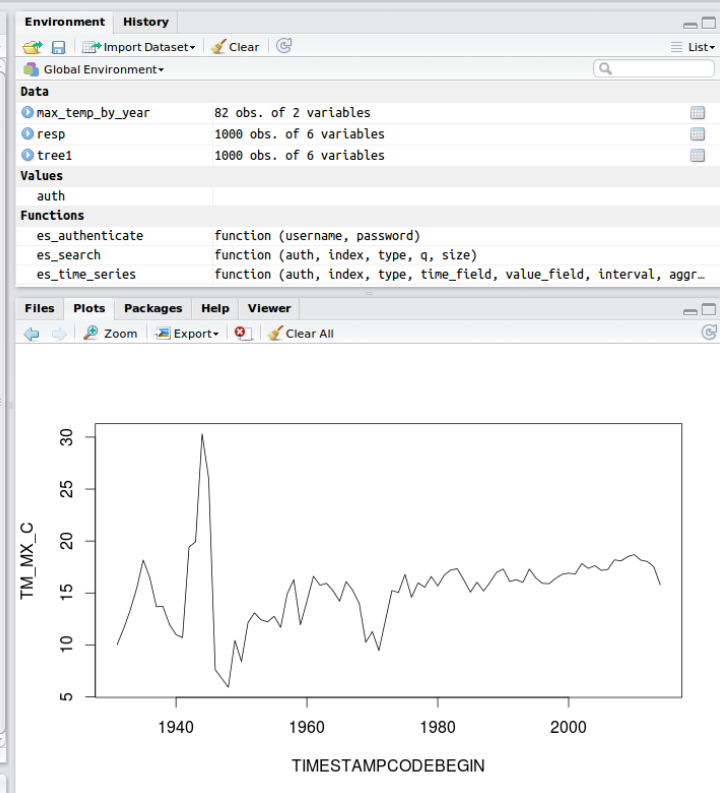
```
In [3]: df = es.search("pcfruit_nutrieten", "measure", "spectrum:0", 1000)
df.plot(x='wavelength', y='measure', style='o', ylim=(0,1))
```

```
Out[3]: <matplotlib.axes._subplots.AxesSubplot at 0x7f5c8ed4c490>
```



```
In [2]: # The parameters are: index, type, time, value, value, variable, aggregation
```

```
Out: RStudio
File Edit Code View Plots Session Build Debug Tools Help
Go to file/function
Project: (None)
Environment History
Global Environment
Data
max_temp_by_year 82 obs. of 2 variables
resp 1000 obs. of 6 variables
tree1 1000 obs. of 6 variables
Values
auth
Functions
es_authenticate function (username, password)
es_search function (auth, index, type, q, size)
es_time_series function (auth, index, type, time_field, value_field, interval, agrg_
Files Plots Packages Help Viewer
Zoom Export Clear All
sample.R* x es_api.R x
1 library(httr)
2 library(jsonlite)
3
4 library(elasticsearchapi)
5
6 auth <- es_authenticate("ddaems", "ddaems")
7 max_temp_by_year <- es_time_series(auth, "meteo", "measure", "TIMESTAMPCODEBEGIN", "TM_MX_C", "1y", "avg")
8 plot[max_temp_by_year, type='l']
9
10 resp <- es_search(auth, "pcfruit_nutrieten", "measure", "tree:1", 1000)
11 tree1 <- resp[ resp$tree == "1", ]
12
13 plot(tree1$wavelength, tree1$measure, ylim=c(0,1))
14
```



# Roadmap

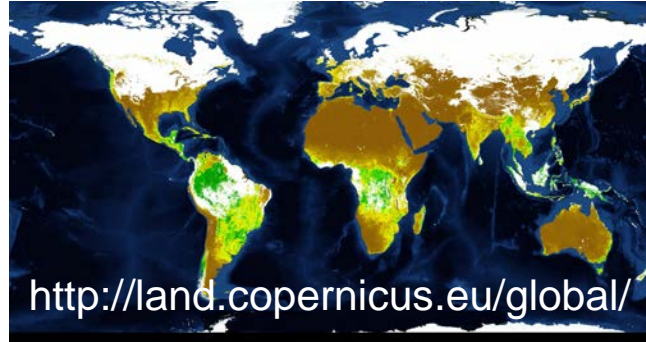
<http://proba-v.vgt.vito.be/>



Time series SPOT-VGT → PROBA-V

*Daily global coverage of land since 1998*

*1 km → 333 m → 100 m resolution*

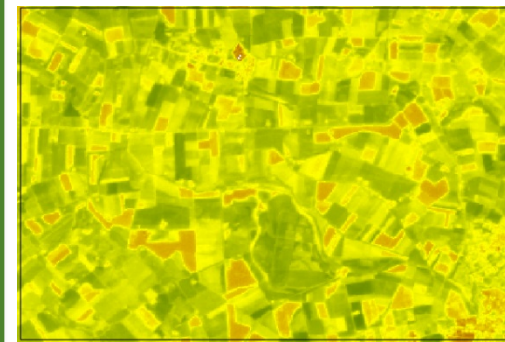


<http://land.copernicus.eu/global/>

Time series Copernicus Global Land

*Global coverage since 1998*

*1 km → 333 m resolution*



Bio-geophysical parameters

at high resolution

*On-demand service*

- R&D ESA project 'ESE' → Acceptance July 2015 → Support of pilots
- Operational '**PROBA-V Mission Exploitation Platform**' → First release Jan 2016
  - Time series viewer - full-res. Viewer - pre-defined PoD services – ESA Cloud Toolbox
  - Virtual Research Environment on scalable platform with access full archive
  - E-collaboration – Cal/Val - access to similar EO missions – Software - Documentation



European Space Agency

→ You can be involved

as external beta user!

as partner infrastructure in a federation



Belgian Science Policy Office



belspo

16/12/2015

© 2015, VITO NV

19



proba  
**VEGETATION**



# Thank you ...

Erwin Goor  
VITO, Belgium  
[erwin.goor@vito.be](mailto:erwin.goor@vito.be)  
+32.493.514274

**> PROBA-V Symposium 2016**

**26 - 28 January 2016  
Ghent, Belgium**

<http://congrexprojects.com/2016-events/16c01>