



XI Jornadas Ibéricas de
Infraestructuras de Datos Espaciales
26 | 30 octubre 2020



Instituto Geográfico Nacional
CENTRO NACIONAL DE INFORMACIÓN GEOGRÁFICA



Hacia la implantación de una infraestructura de datos de imágenes de satélite en el Instituto Geográfico Nacional con Open Data Cube y QGIS



Bruno Pérez Martín

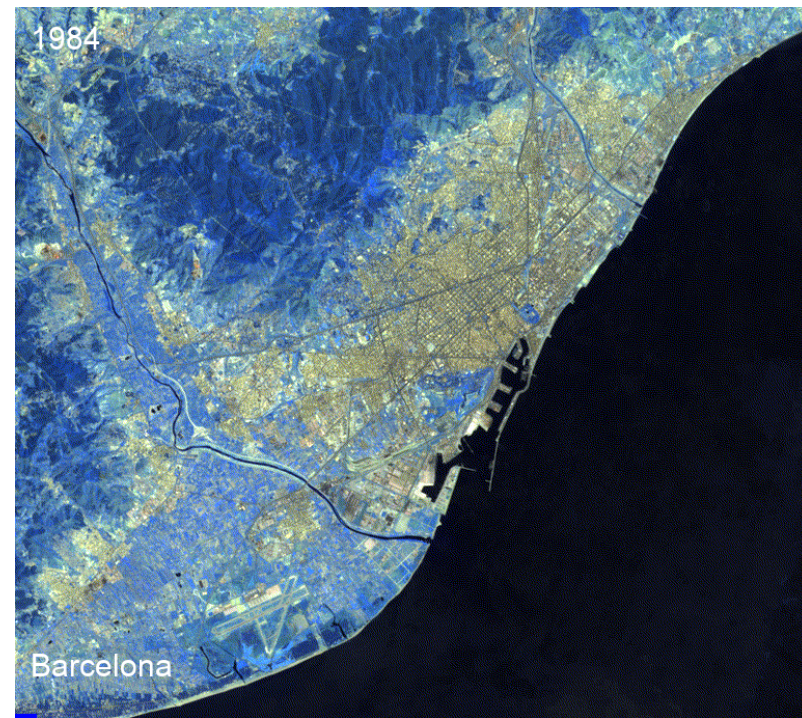
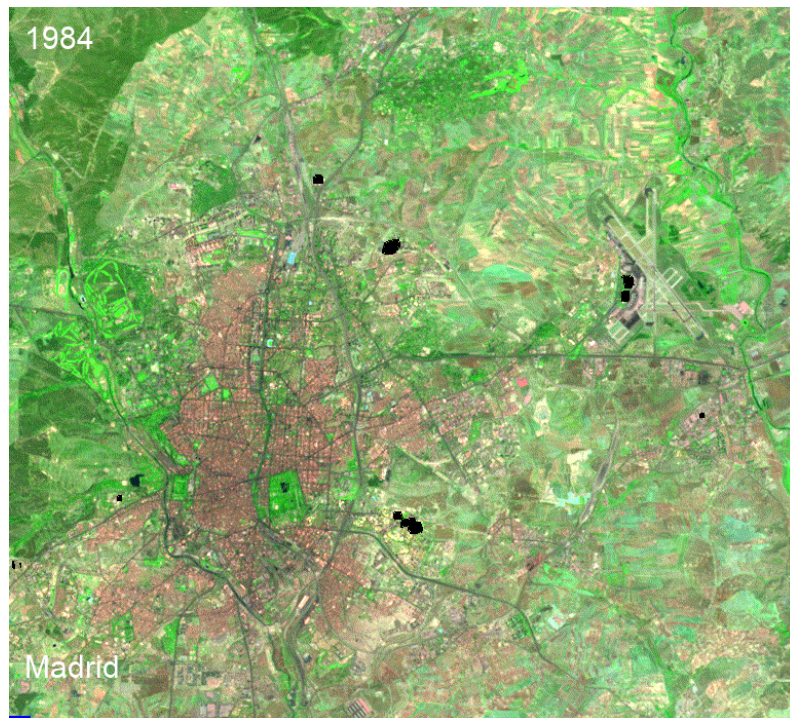
Damián Ortega Terol

www.ign.es





Wu, Q., (2020). geemap: A Python package for interactive mapping with Google Earth Engine. The Journal of Open Source Software, 5(51), 2305.
<https://doi.org/10.21105/joss.02305>



Wu, Q., (2020). geemap: A Python package for interactive mapping with Google Earth Engine. The Journal of Open Source Software, 5(51), 2305.

<https://doi.org/10.21105/joss.02305>

Por qué un datacube

Teledetección v1.0

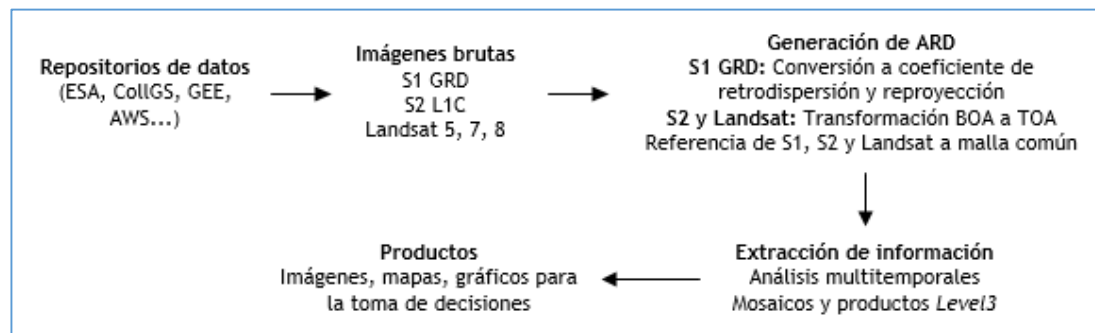


Teledetección v2.0



Copyright: Geoscience Australia

Para poder explotar la multitemporalidad es necesario la generación de Analysis Ready Data (ARD) para su ingestión en los *datacubes*

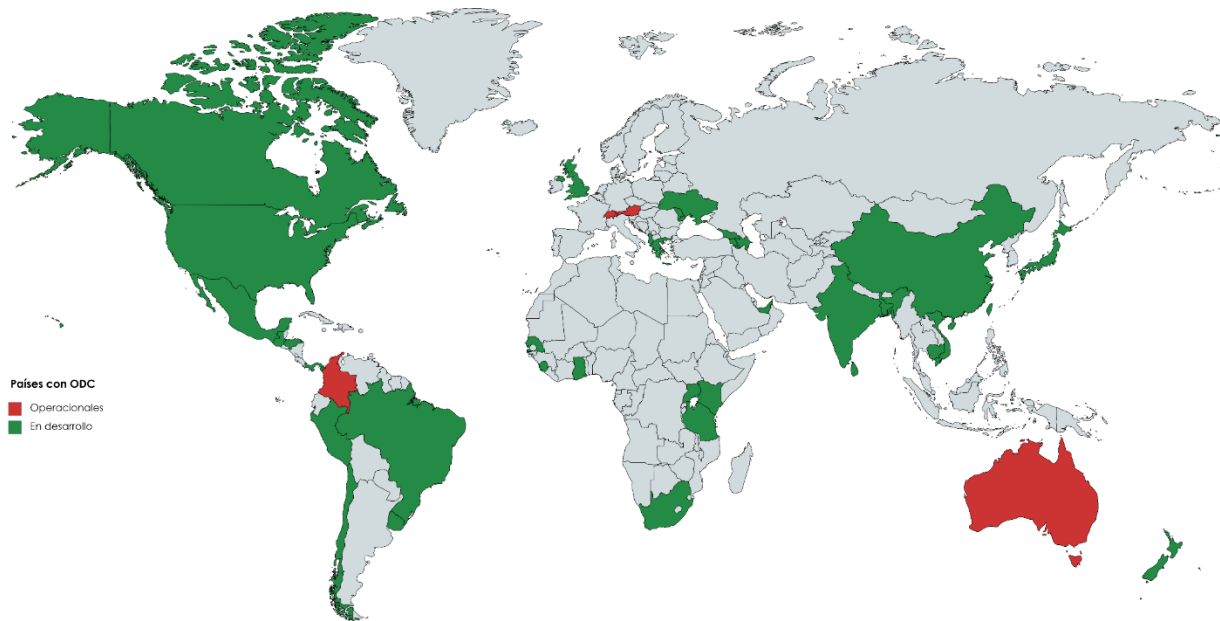


Por qué OpenDataCube

El núcleo del código del *Australian Geoscience DataCube* se liberó en 2017 como parte de una colaboración internacional junto con otros organismos entre los que se cuentan CEOS y el USGS y pasó a denominarse Open Data Cube.



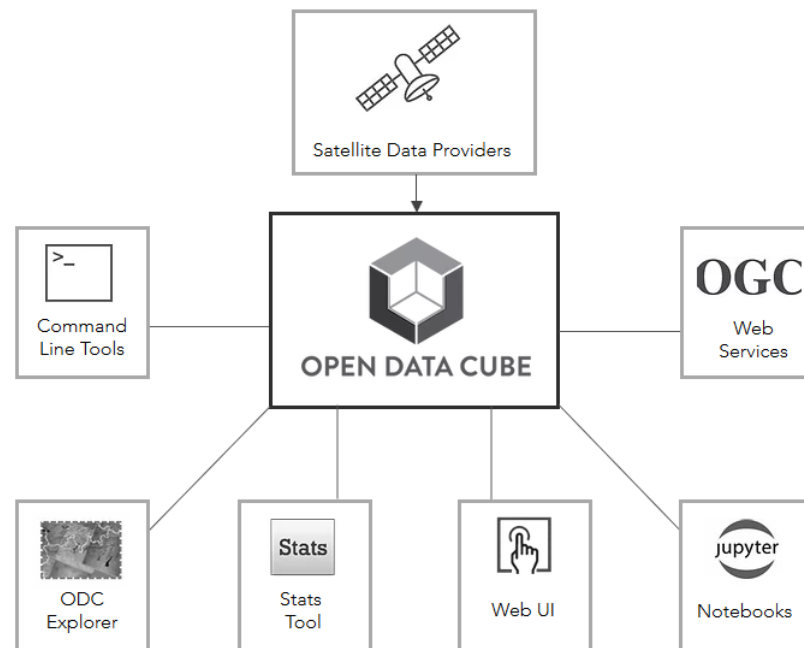
Desde entonces el interés en la tecnología ha ido aumentando



Created with mapchart.net

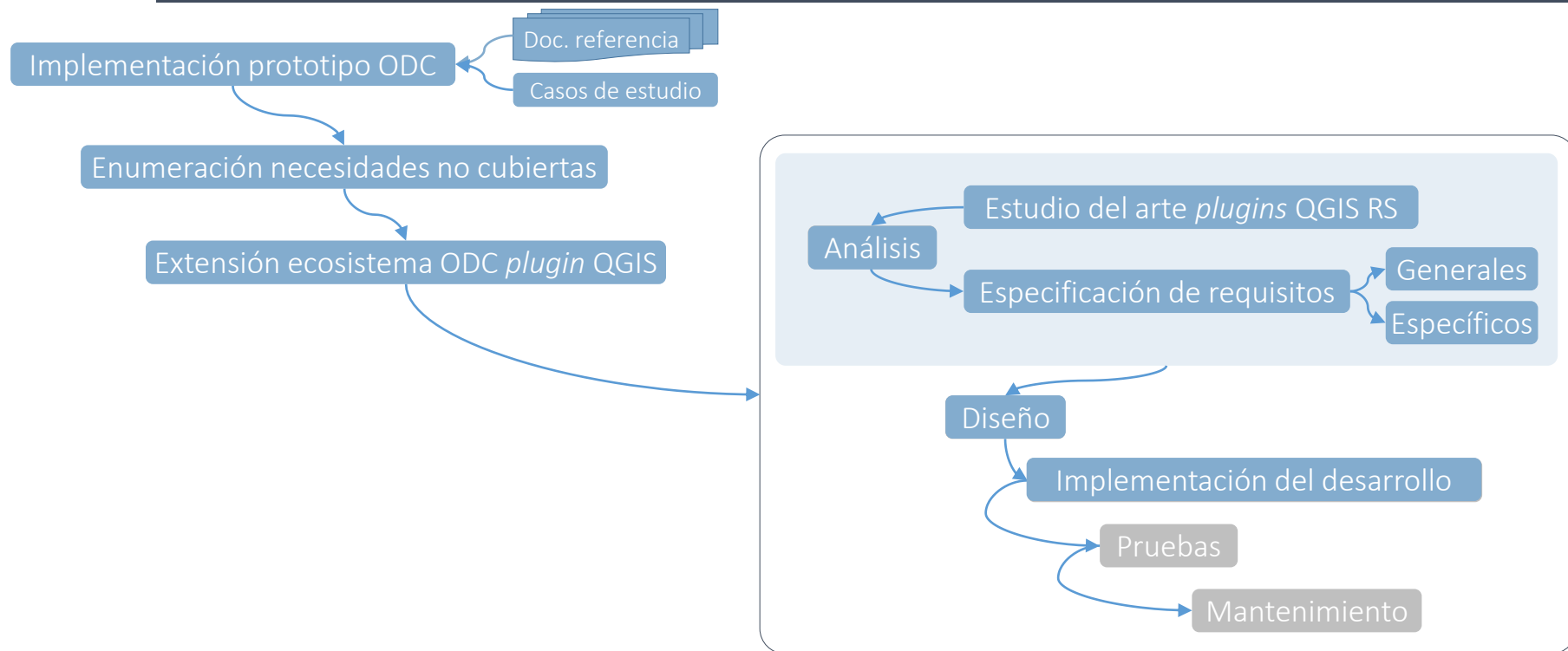
Funcionalidades de Open Data Cube

- **ODC Core:** ecosistema integrado de datos raster para el análisis multitemporal de imágenes de satélite.
- Web User Interface.
- Jupyter notebooks para prototipado de aplicaciones.
- Servicios OGC (WMS/WMTS), incluyendo soporte para formato COG.

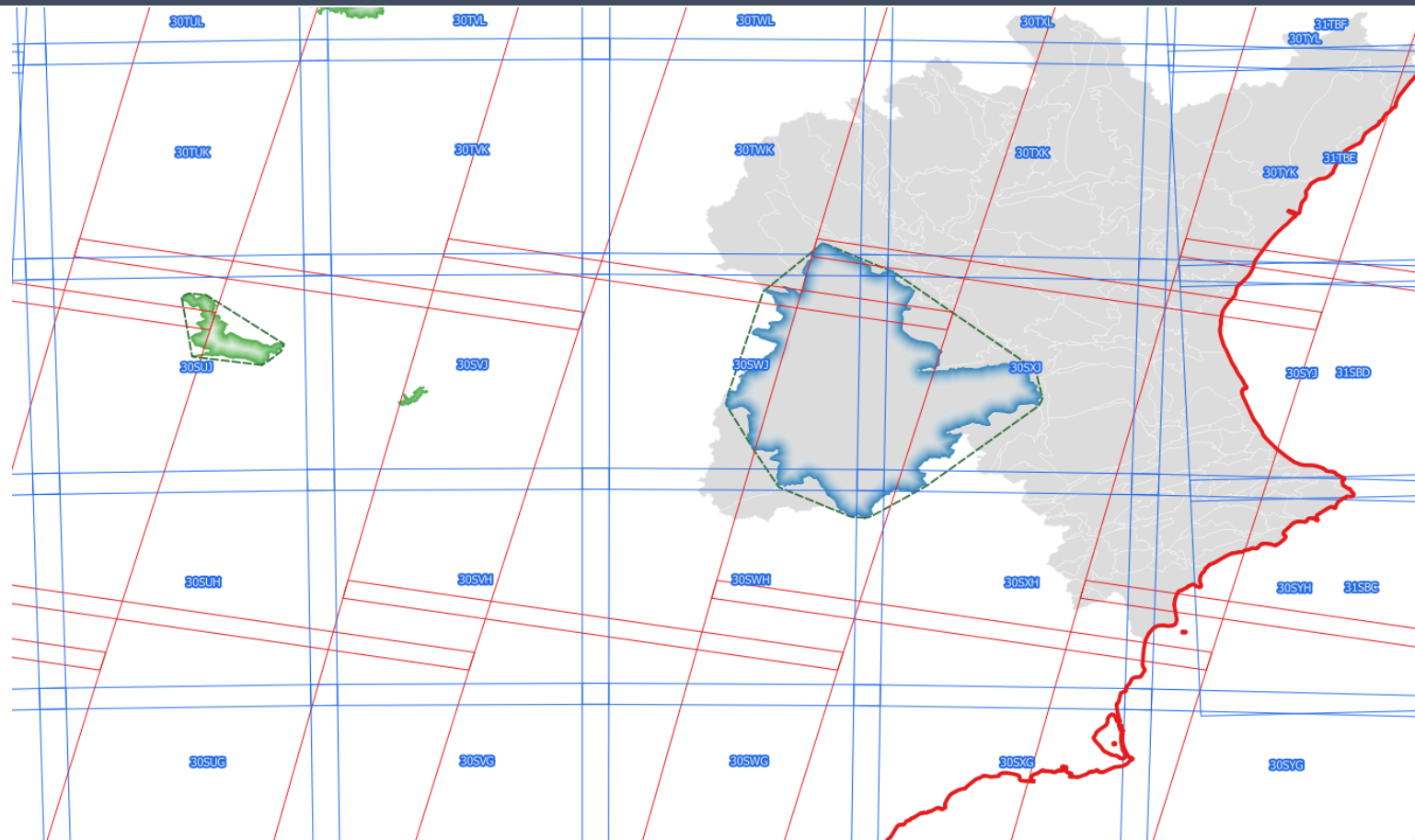


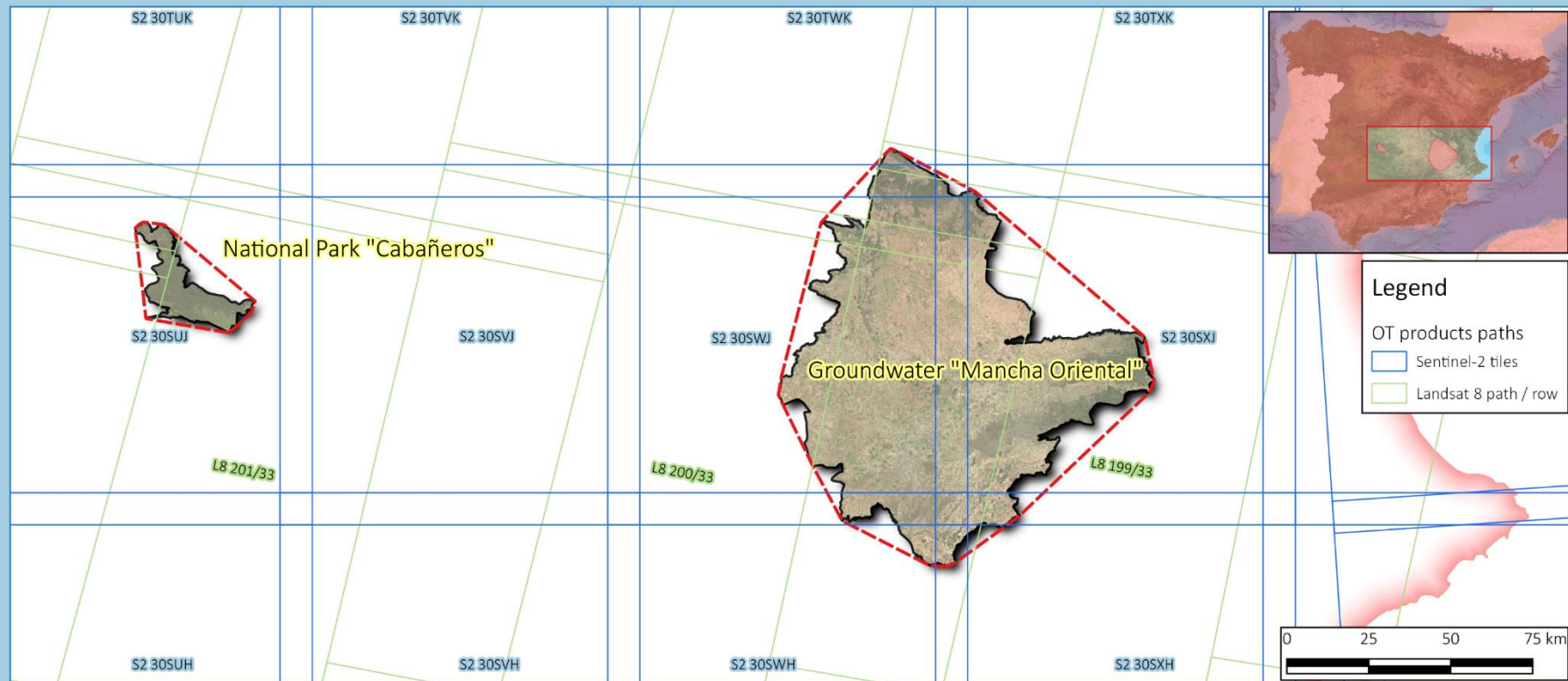
Copyright: OpenDatacube

Metodología desarrollo



Casos de estudio





Estudio de la documentación de referencia

Chatenoux B., Richard J.-P., Guigoz Y., Poussin C., Giuliani G. (2019) **Bringing Open Data Cube into Practice**. Workshop Material. GRID-Geneva & University of Geneva. 65 pages. DOI: [10.13140/RG.2.2.17703.91044](https://doi.org/10.13140/RG.2.2.17703.91044)

Bruno Chatenoux, Jean-Philippe Richard, Yaniss Guigoz, Charlotte Poussin, Gregory Giuliani

Bringing Open Data Cube into Practice



<http://www.opendatacube.org>



Harnessing the information power of satellite data

<https://www.opendatacube.org/resources>

Resources

ODC documentation, videos, tutorials, and guides

SOCIAL MEDIA & FORUMS



Medium

Follow our blog on Medium



Videos

Open Data Cube videos



GIS Stack Exchange

Developers Forum



Slack

Find us on Slack



Twitter

Follow @opendatacube on Twitter

<https://github.com/opendatacube>



Open Data Cube

<https://www.opendatacube.org/> opendatacube@googlegroups.com

Repositories 39 Packages People 4 Projects

Pinned repositories

datacube-core

Open Data Cube analyses continental scale Earth Observation data through time

Python 259 120

datacube-dataset-config

Scripts for indexing data into ODC instances

Python 9 10

datacube-notebooks

Extra documentation about using ODC with Jupyter Notebooks

Jupyter Notebook 31 22

Implementación del prototipo ODC

Sistema

Procesador: Intel(R) Xeon(R) CPU E5-1620 v3 @ 3.50GHz 3.50 GHz

Memoria instalada (RAM): 64,0 GB

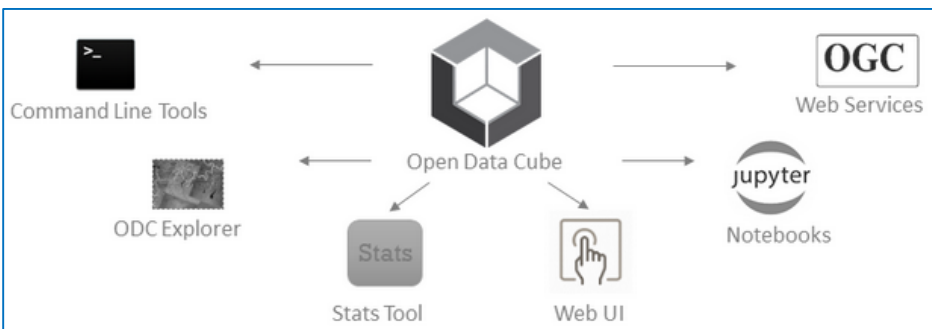


Welcome to **VMware Workstation 14 Player**

ubuntu

```
(datacube_env) localuser@ubuserver1804:~/Datacube/data_cube_notebooks$ lsb_release -a
No LSB modules are available.
Distributor ID: Ubuntu
Description:    Ubuntu 18.04.4 LTS
Release:        18.04
Codename:       bionic
```

Programming languages



Spatial databases



Cloud platforms



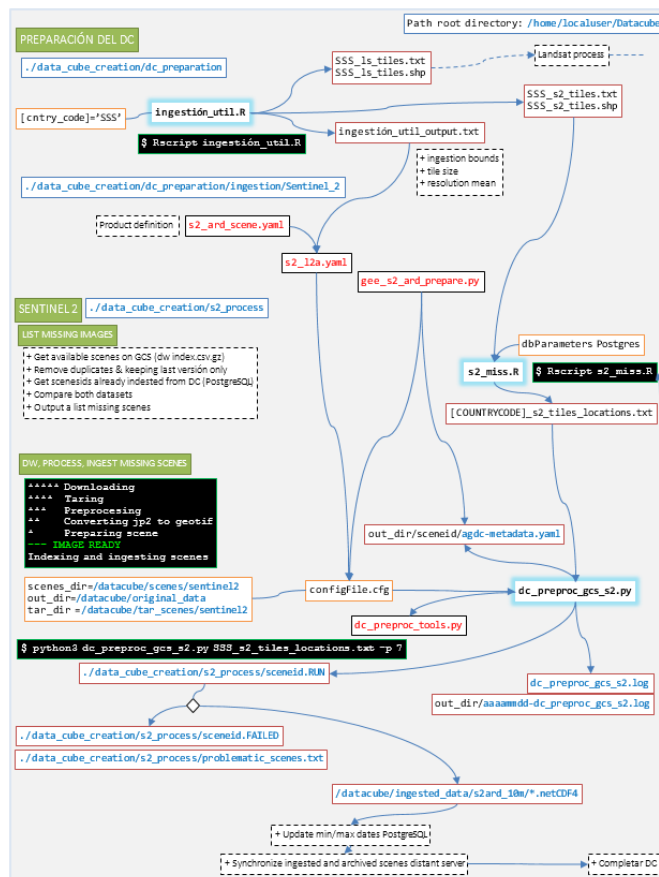
Geospatial / Scientific Libraries



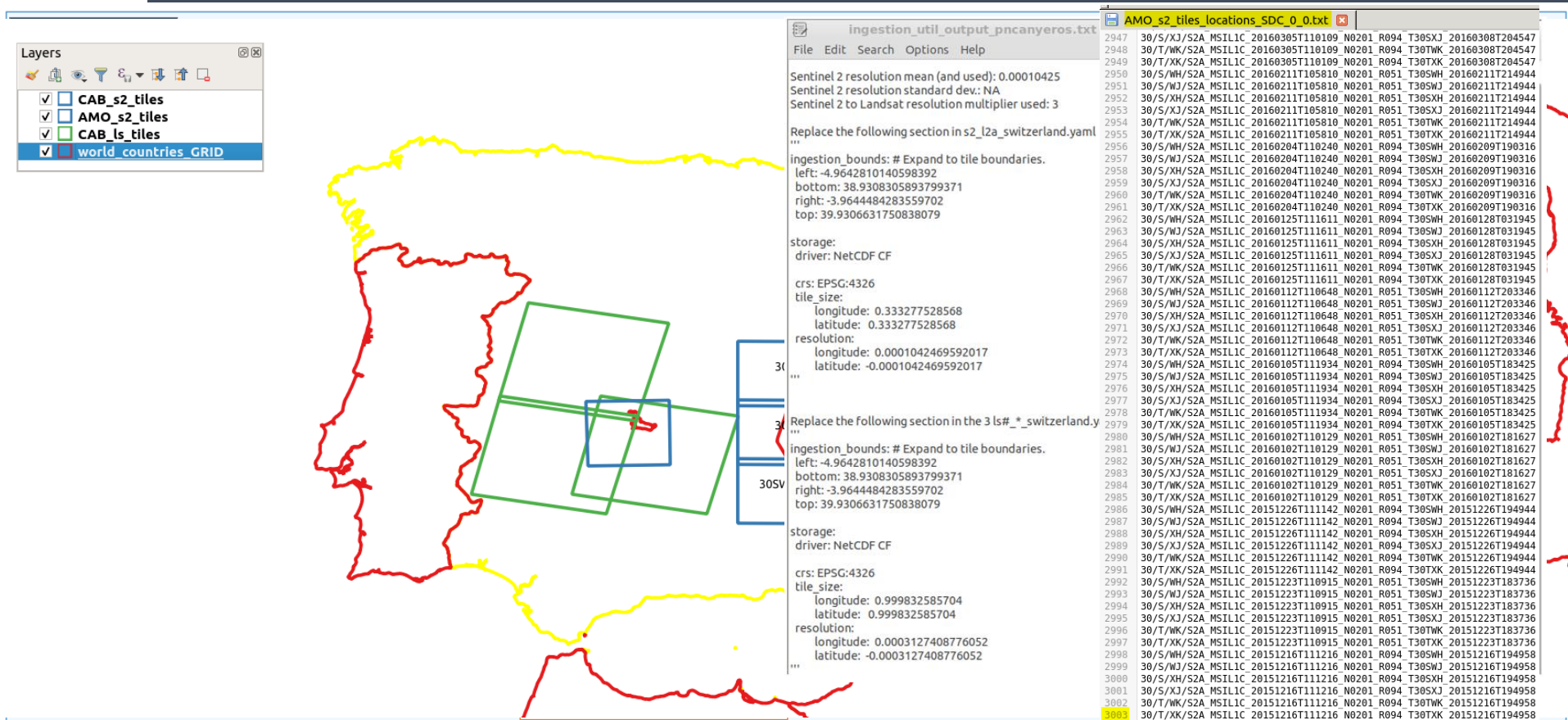
DC Graphical Web UI



Implementación del prototipo ODC. Ingestión de escenas. Flujoograma extendido.



Implementación del prototipo ODC. Ingestión de escenas paso a paso. Caso de uso escenas Sentinel 2 Level 1C





Welcome to the Open Data Cube

CEOS is using the power of the Open Data Cube to help address the needs of satellite data users, giving them a better picture of their land resources and land change.

- Ease of use and access to satellite-based data
- Multiple dataset interoperability and spatial consistency
- Use of "Analysis Ready" Data Products
- A Shift in Paradigm from Scenes to Pixels

Implementación del prototipo ODC. Aplicaciones.

In [3]: # Select the product

In [8]: # Select measurements

In [11]: # Select geographical extent

In [13]: # Select time period

In [16]: # Resume configuration parameters in a format ready to be copy/pasted to a new cell,
and in a txt file to be loaded with the '%load config_cell.txt' magic.

```
str = '''# Configuration
```

```
platform = '{}'  
product = '{}'  
measurements = {}
```

```
min_lon = {}  
max_lon = {}  
min_lat = {}  
max_lat = {}
```

```
start_date = datetime.strptime('{}', '%Y-%m-%d')  
end_date = datetime.strptime('{}', '%Y-%m-%d')'''.\  
format(get_platform(dc, [product_sel.value])[0],  
       product_sel.value,  
       measur_sel,  
       min_lon, max_lon, min_lat, max_lat,  
       start_date.value.strftime('%Y-%m-%d'), end_date.value.strftime('%Y-%m-%d'))  
print(str)  
with open('config_cell.txt', 'w') as text_file:  
    print(str, file=text_file)
```

```
# Configuration
```

```
platform = 'SENTINEL 2'  
product = 's2_l2a_sdc_mancha_oriental'  
measurements = ['blue', 'green', 'red', 'slc']
```

```
min_lon = -1.85071  
max_lon = -1.752895  
min_lat = 39.125786  
max_lat = 39.187549
```

```
start_date = datetime.strptime('2019-01-01', '%Y-%m-%d')  
end_date = datetime.strptime('2020-02-23', '%Y-%m-%d')
```

Time Series Water Detection Analysis

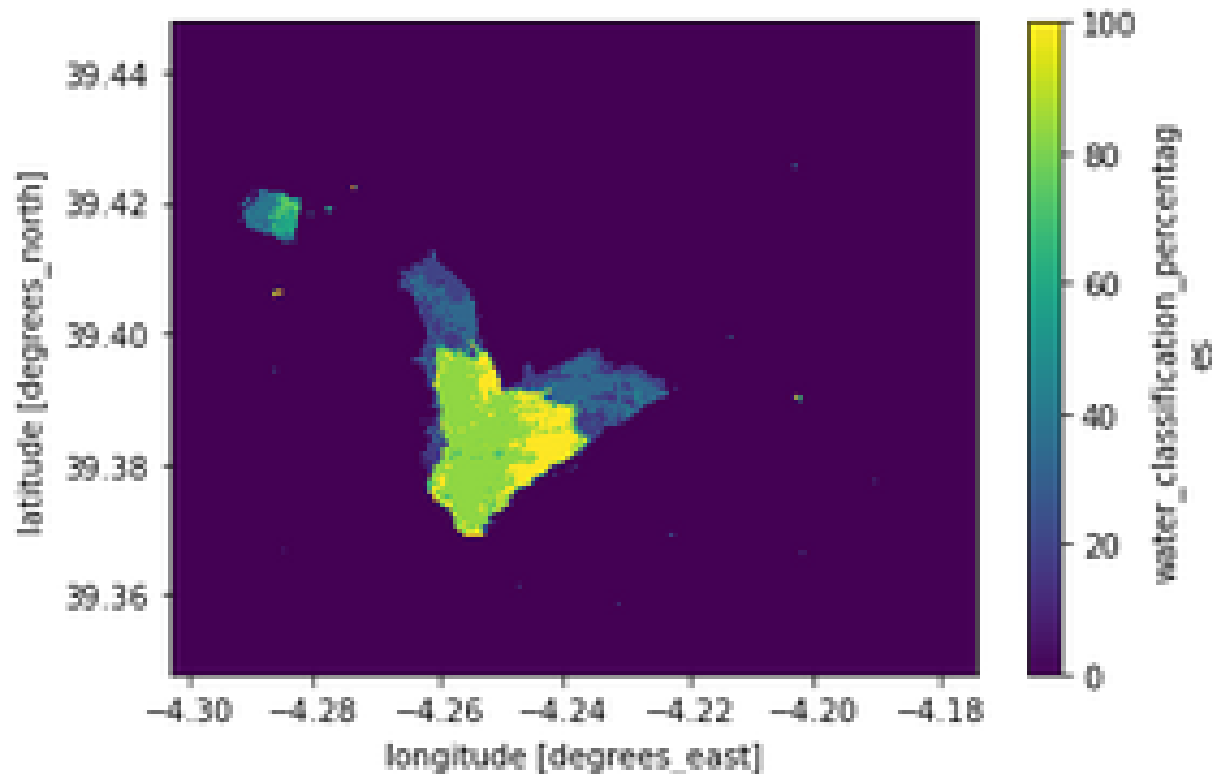
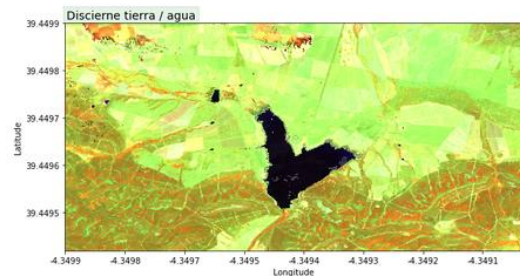
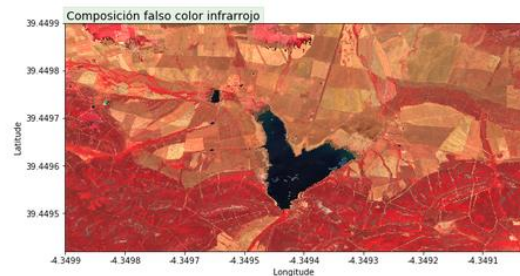
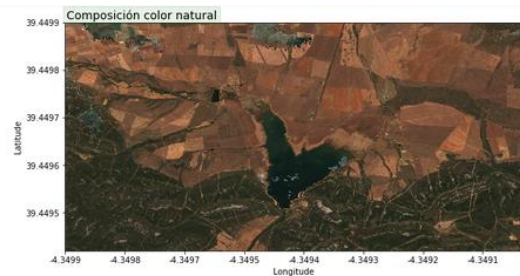


Logout

Trusted

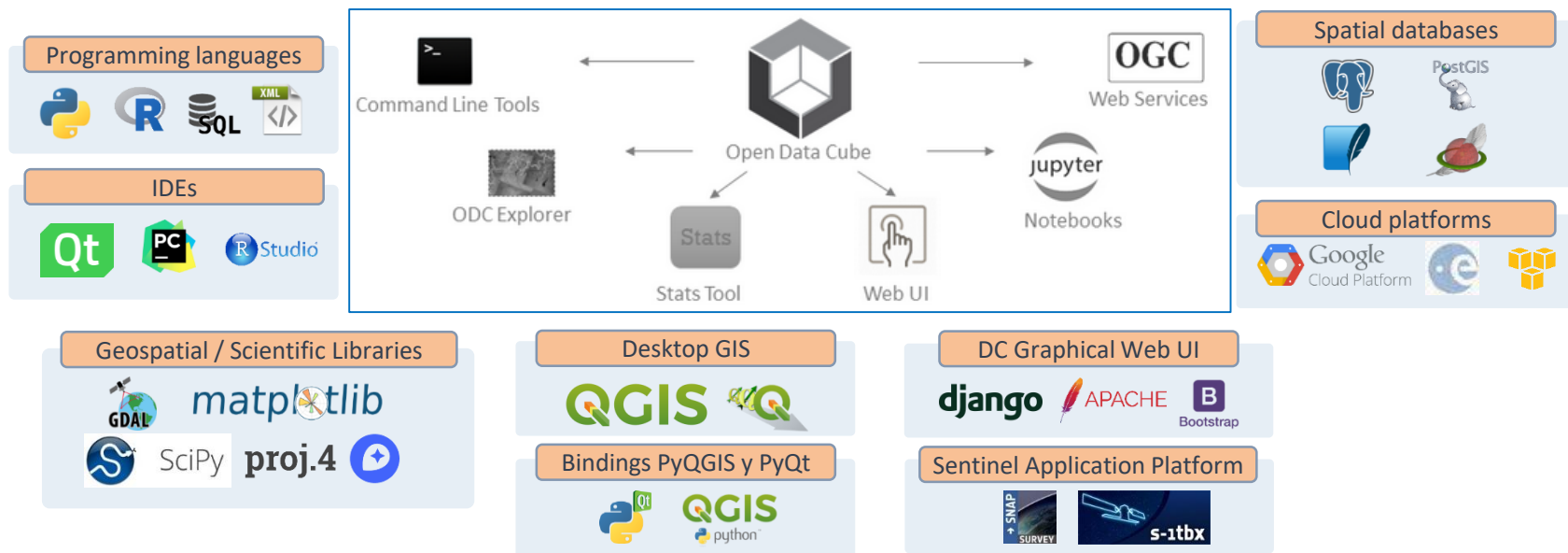
Python 3

Implementación del prototipo ODC. Aplicaciones.



- Profundizar en el conocimiento de experiencias de **generación de ARDs de la misión Sentinel-1 en el ecosistema ODC**. Estudio particularizado de productos **Sentinel-2 ARD** disponibles.
- **Acceso a distintas plataformas para consulta y descarga de datos OT**.
- Extensión de métodos para la **definición de ROIs**.
- Las labores de **mantenimiento del catálogo de escenas ingestadas** tras la indexación de productos en la base de datos PostgreSQL se verían facilitadas a partir de la implementación de una **interfaz de usuario**, permitiendo la consulta de información almacenada en el modelo de datos de forma sencilla.
- Existen ciertos parámetros de entrada necesarios para la ejecución de los scripts descritos. En unos casos se facilitan en ficheros de configuración independientes y en otros son sustituidos directamente en su código fuente Python. Disponer de **una interfaz de usuario para la recopilación de estos datos** facilitaría su ejecución repetida.
- Alguno de los procedimientos de preparación del DC y análisis y aplicaciones realizados a través de la interfaz Web de ODC o con Jupyter Notebook, podrían **extenderse a clientes SIG pesados**.
- **Integración de librerías geoespaciales de terceros programadas en Python**, ampliarían considerablemente las posibilidades de futuros desarrollos.

Extensión del ecosistema ODC a través del desarrollo de plugin para QGIS



Estudio del arte de herramientas QGIS en el ámbito de la teledetección



Especificación de requerimientos generales del desarrollo del complemento

Requerimientos generales:

- (i) El complemento se concibe como un **conjunto de herramientas escalable** cuyo objetivo principal es facilitar las operaciones contempladas en las tres fases del flujo ODC.
- (ii) Debe cumplir los [requisitos](#) y recomendaciones para su **publicación en el repositorio oficial de complementos de QGIS**.
- (iii) Se prestará especial atención a la **documentación de funcionalidades y código fuente del proyecto**; la documentación será generada con [Sphinx](#) utilizando el tema [Read the Docs](#) y será alojada en GitHub corporativo.



Herramientas previstas inicialmente:

- (i) Herramienta para la búsqueda y gestión de descargas de datos satelitales de distintas plataformas adaptada a las necesidades del proyecto.
- (ii) Herramienta para el procesamiento de datos Sentinel-1 y generación de los ARD necesarios para la ingestión.
- (iii) Herramientas para el mantenimiento del catálogo de escenas ingestadas.
- (iv) Herramientas que facilitan la preparación del DC y la generación de ficheros de configuración.

data_cube_creation [~/Datacube/data_cube_creation] - .../s2_process/configFile_cabanyeros.cfg - PyCharm

File Edit View Navigate Code Refactor Run Tools VCS Window Help

data_cube_creation s2_process configFile_cabanyeros.cfg

Project 1: Project

data_cube_creation ~D Plugins supporting *.cfg files found.

```

1 [mutt] # TO BE #####
2 smtp_url = "#####"
3 smtp_pass = "#####"
4 ml_from = "#####"
5 ml_to = "#####"
6 #####
7
8 [perf]
9 max_scenes = 150
10
11 [gsutil]
12 path = /home/localuser/gutil/gutil.py
13
14 [GCS]
15 uri_s2 = gs://gcp-public-data-sentinel-2
16
17 [S2]
18 jp2_postfx = _B02_10m.jp2, _B03_10m.jp2, _B04_10m.jp2, _B08_10m.jp2, _B05_20m.jp2, _B06_20m.jp2, _B07_20m.jp2, _B8A_20m.jp2,
19
20 [datacube]
21 venv = /home/localuser/Datacube/datacube_env/bin/python
22 datacube_bin = /home/localuser/Datacube/datacube_env/bin/datacube
23
24 prep_script_s2 = /home/localuser/Datacube/data_cube_creation/dc_preparation/ingestion/Sentinel_2/gee_s2_ard_prepare.py
25 s2_yaml = /home/localuser/Datacube/data_cube_creation/dc_preparation/ingestion/Sentinel_2/s2_l2a_pncabanyeros.yaml
26 scenes_dir = /mnt/hqfs/ODC/datacube_win/scenes/sentinel2

```

20:11 CRLF UTF-8 4 spaces ython 3.6 virtualenv at ~/Datacube/datacube...

Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help

ingestion_util_cabanyeros - QGIS

chuleta [localuser@ubuserver1... (2) Firefox data_cube_creation[-/... RStudio QGIS3

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

s2_miss_ESP_GCS.R x s2_miss_ESP_SDC.R x

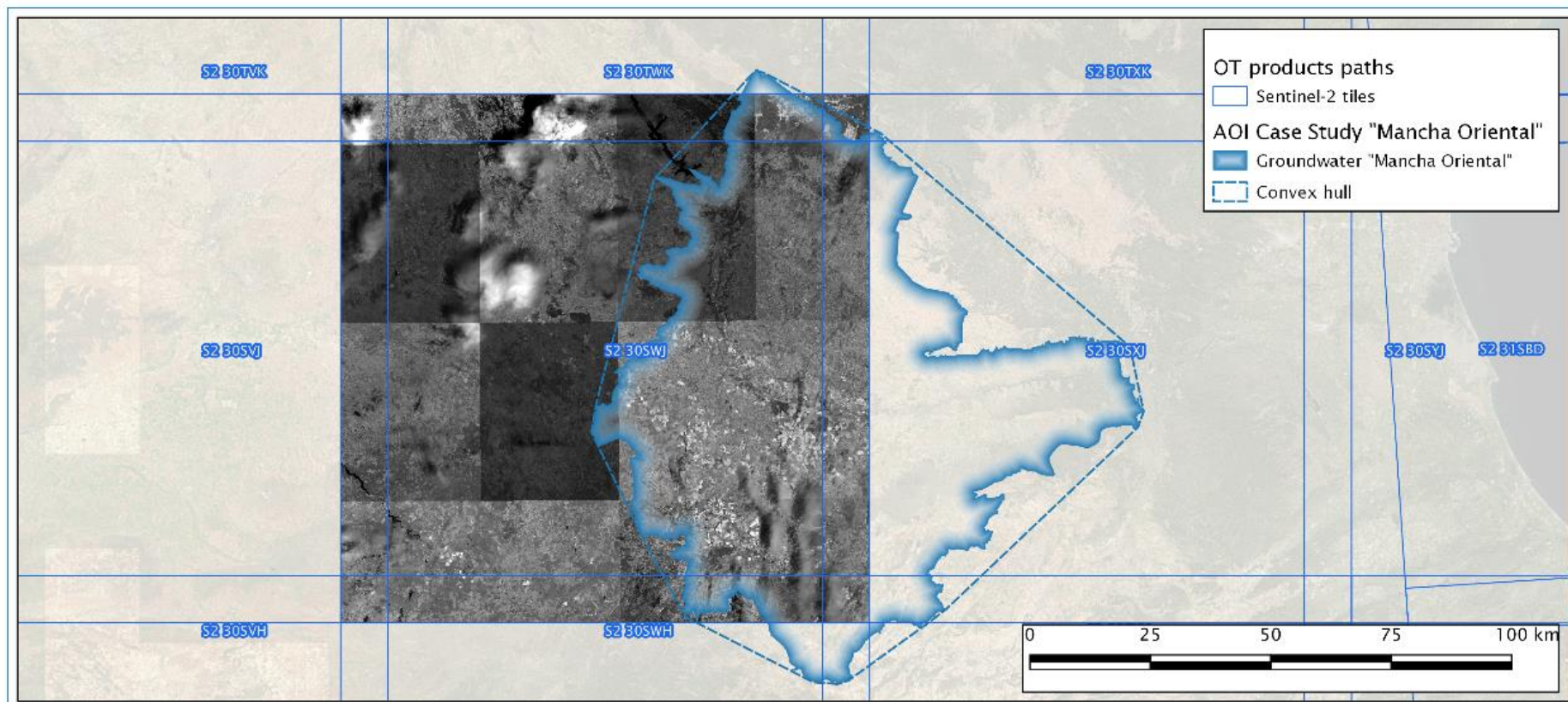
```

16 },
17 error=function(cond){
18   setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
19 }
20
21 library(data.table)
22 library(RPostgreSQL)
23 library(dplyr)
24 library(ggplot2)
25 library(forcats)
26
27 #####
28 # TO BE ADAPTED
29 tiles_file <- '../dc_preparation/ESP_s2_tiles.txt'
30 pbln_file <- 'problematic_scenes_ESP_SDC.txt'
31 gsutil_path <- '/gsutil/gutil.py'
32
33 db_name <- '#####' # PostgreSQL DB name (TO BE #####)
34 db_user <- '#####' # PostgreSQL user (TO BE #####)
35 db_pw <- '#####' # PostgreSQL password (TO BE #####)
36 #####
37
38 start.time <- Sys.time()
39
40 # GET AVAILABLE SCENES on GCS
41 cat('\n*** Downloading index.csv.gz\n')
42 system(paste('python2 ', gsutil_path, ' cp gs://gcp-public-data-sentinel-2/ ', tiles_file))
43 cat('\n*** Importing index.csv.gz\n')
44 gcs_df <- fread('gunzip -cq index.csv.gz', header = T, sep = ',')
45 #DOT: comentado para ver que lleva ese fichero
46 file.remove("index.csv.gz")
47
48 tiles_df <- fread(tiles_file, header = F, sep = ',')
49 colnames(tiles_df) <- 'tilename'
50
51 col_df <- gcs_df[, c("tilename", "index.csv.gz")]

```

35:23 (Untitled) R Script

Console



Remote Sensing Downloader

Search parameters Products Found Configuration

▼ AOI Definition

Area relation: Intersects

▼ ☒ Polygonal Area

Layer extent: MasaAguaSubterranea_Jucar [EPSG:25830]

Map tool draw polygon:

▼ ☐ Coordinates Bounding Box (EPSG: 4326)

North: 40,808228°

West: -2,525517° East: 0,531397°

South: 38,102480°

Map view extent:

▼ ☐ Tiles Sentinel-2 Spain Zone

☒ Tiles huso 30

Peninsula: 30STE||30STF||30STG||30STH||30STJ||30SUF||30S...

Baleares:

Canarias:

Ceuta y Melilla:

▼ ☐ Scene id

Product id: 36cbea70-226a-4f0f-b4ef-f422f1d117c2

▼ Access Methods

Access method: Copernicus Open Access Hub (SentinelSat API)

▼ Search, sort and filter parameters

Platform name: Sentinel-2

Product types: S2MSI1C

Sensing period (start / end): 01/01/2019 / 31/12/2019

Cloud cover (min / max): 0 % / 100 %

Limit number results: 10

Remote Sensing

Proyecto Edición Ver Capa Configuración Complementos Vectrial Bácter Base de datos Web Malla Procesos Ayuda

user default

Search parameters

AOI Definition

Area relation:

Polygon

Layer

Map tool draw

Coordinate

West: -3

Map view extent

Access Method

Access method:

Search, sort and filter

Platform

Product

Sensing period (s)

Cloud cover

Limit number

ODC Proyecto piloto - QGIS

2019-02-08 11:02:21

2019-02-08 11:02:21

2019-02-15 19:36:31

18967

94

0,0751

INS-HOBS

30SWH

SH30W

SAFE

02.07

Sentinel-2

S2A_MSI_IC_20190208T110221_N0207_R094_T30SWH_20190214T155730.SAFE

Multi-Spectral Instrument

MSI

838.01 MB

GS2A_20190208T110221_018967_N02.07

S2MSI1C

2015-028A

DESCENDING

Sentinel-2A

Level-1C

S2A_MSI_IC_20190208T110221_N0207_R094_T30SWH_20190214T155730

S2A_OPER_MSI_1C_TL_EPA_20190214T155730_A018967_T30SWH_N02.07

ae12b879-d83c-4793-bc7a-8d9a22971225

S2A_OPER_MSI_1C_TL_EPA_20190214T155730_A018967_T30SWH_N02.07

S2A_OPER_MSI_1C_DS_EPA_20190214T155730_S20190208T110853_N02.07

ae12b879-d83c-4793-bc7a-8d9a22971225

20201022_T131045_dataset_Sentinel-2_S2MSI1C_20190101_20191231_cc_0_100 :: Objetos totales: 899, Filtrados: 899, Seleccio...

title

link

link_alternative

link_icon

summary

dataakesensingstart

beginposition

endposition

ingestondate

orbitnumber

relativeorbitnumber

doudcoverpercentage

sensoroperationalmode

tleid

hv_order_tleid

format

processingbaseline

platformname

filename

instrumentname

instrumentshorthandname

size

s2dataakeid

producttype

platformidentifier

orbitdirection

platformserialidentifier

processinglevel

identifier

level1cididentifier

uuid

granuleididentifier

datastrididentifier

id

Drive unit: C:

ask free: 250,00 GB

ing start

Title

10:58:29.024000 30SWJ

10:47:59.024000 30SWJ

10:58:51.024000 30SWJ

10:50:31.024000 30SWJ

10:57:19.024000 30SWJ

10:46:49.025000 30SWJ

10:57:41.024000 30SWJ

10:50:31.024000 30SWJ

10:56:39.024000 30SWJ

10:46:29.024000 30SWJ

10:56:31.025000 30SWJ

10:50:31.024000 30SWJ

10:56:19.024000 30SWJ

10:46:29.024000 30SWJ

100.geojson

100.geojson

Escriba para localizar (Ctrl+K)

vordena: -2,26,32,19

ce: 1:3617853

rplicaf: 100%

itack: 0,0 °

Representar

EPSG:4326

Mostrar todos los objetos espaciales

XI Jornadas Ibéricas de Infraestructuras de Datos Espaciales 26 | 30 octubre 2020

[Remote Sensing Downloader tool](#)[ARD Sentinel-1 preprocess](#)[IGN Data Cube utilities](#)[PostgreSQL Data Cube management](#)

Welcome to IGN-DC's plugin QGIS documentation

Contents:

- [Remote Sensing Downloader tool](#)
 - [Area of interest definition](#)
 - [Search parameters](#)
 - [Process dialog](#)
 - [This file](#)
- [ARD Sentinel-1 preprocess](#)
- [IGN Data Cube utilities](#)
- [PostgreSQL Data Cube management](#)

Indices and tables

- [Index](#)
- [Module Index](#)
- [Search Page](#)

[Next](#) ➞

© Copyright 2020, Damián Ortega Terol / Bruno Pérez Martín (IGN)

Built with [Sphinx](#) using a [theme](#) provided by [Read the Docs](#).

- Tamaño y número de escenas necesarias para cubrir la Península Ibérica con escenas de la misión Sentinel 2 producto S2MSIL1C. A partir de los datos ingestados en el prototipo desarrollado se estima que el volumen de almacenamiento aumenta aproximadamente un 36 %, tras aplicar sen2cor para obtener el producto S2MSIL2A y el paso a formato NetCDF.
- Datos desde 2015 hasta octubre 2020.

Cobertura nubosa (%)	Nº escenas	Tamaño (TB)	Cobertura nubosa (%)	Nº escenas	Tamaño (TB)
[0 – 100]	12431	6.25	[0 – 50]	8928	4.74
[0 – 90]	11319	5.95	[0 – 40]	8301	4.40
[0 – 80]	10613	5.61	[0 – 30]	7623	4.04
[0 – 70]	10050	5.33	[0 – 20]	6862	3.62
[0 – 60]	9426	5.05	[0 – 10]	5798	3.02

- El IGN está analizando diferentes alternativas para canalizar datos satelitales de diferentes misiones a los usuarios de las Administraciones Públicas españolas. Entre las alternativas, el IGN ha acometido los trabajos de implantación de un prototipo de ODC en dos casos de estudio preseleccionados y se ha desarrollado un manual de instalación traducido al español y una máquina virtual configurada.
- Identificación de necesidades no cubiertas, para lo cual se pretende dar salida con un complemento para QGIS que extienda las funcionalidades del ecosistema ODC, utilizando distintas APIs y librerías geoespaciales de referencia en este ámbito.
- El desarrollo de plugin de QGIS se concibe como un conjunto de herramientas escalable, con la posibilidad de incluir una herramienta para la consulta y gestión de descarga de productos, generación de ARDs y tareas de preparación del DC y para la gestión del catálogo de escenas ingestadas. Todo ello basado en soluciones de software libre.
- Escalar estas soluciones a la totalidad del territorio español constituye el reto fundamental de las líneas futuras de actuación, cuyo objetivo final podría ser la implantación de una infraestructura de datos de imágenes de satélite en el Instituto Geográfico Nacional. Un diseño compatible a nivel peninsular haría más útil esta posibilidad.

www.ign.es

GRACIAS
POR SU ATENCIÓN

Bruno Pérez Martín

bpmartin@mitma.es

Damián Ortega Terol

damian.ortega@correo.gob.es