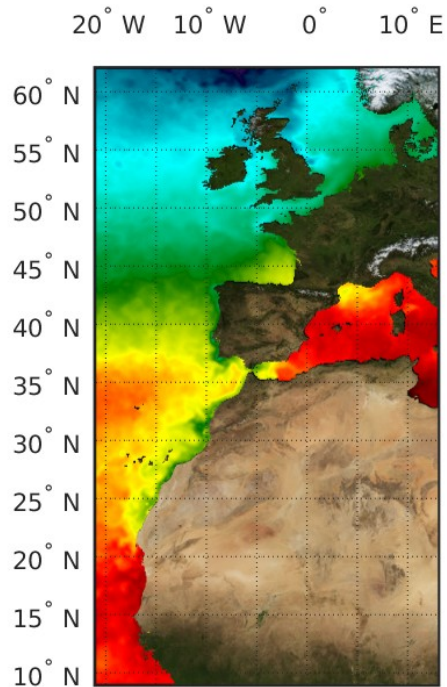




Ifremer/RDAC and EU-GDAC

Jean-François Piollé, Emmanuelle Autret,
Cédric Prevost, Dominique Briand (Ifremer)

- Ifremer is National French Marine Institute / CERSAT is its Satellite Data Center
- operates a Producer / RDAC & GDAC center since Medspiration project (2005)
- As a **producer or regional and global (L3/L4) – Odyssea products**
 - Regional products inherited from Medspiration
 - Mediterranean Sea
 - South Africa
 - Brazil/Tropical Atlantic
 - CMEMS products : **only distributed by CMEMS Dissemination unit** (*cf. CMEMS report*)
 - L4 Europe North Western Shelves/Iberian sea/Canary islands
 - Global multi-sensor L3S
- as a **R/G DAC**
 - Distribution of O&SI SAF, push to PODAAC mirror (*cf. OSI SAF report*)
 - Mirror some datasets from GDAC (used in Odyssea multisensor merging)
 - support projects requiring combination of multiple source of data, microwave data, etc...
 - Central Repository of In situ Radiometer Network Data (*cf: W. Wimmer*)



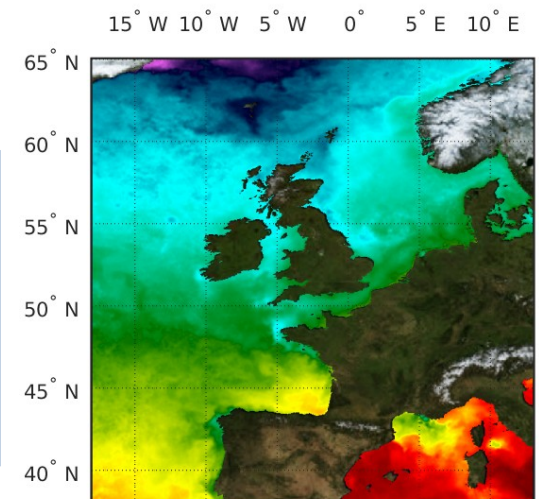
Extended West Europe Area (ATL)

- “ATL” : NWS + Iberian-Biscay-Irish (IBI) areas
- NRT Product
 - Update planned this year with addition of VIIRS, S3A&B, MSG Indian, GOES-16, Himawari
- Reprocessing for Copernicus Marine environment Monitoring Service (CMEMS) - 1982-2018
 - **Input observations** : AVHRR Pathfinder Version 5.3 (PFV53) L3C (1982-2014), extended to 2018 by including the real time AVHRR18-19G data
 - To be replaced with CCI SST data
 - **Method** : Kalman smoother (Tandéo et al., 2011)

European North West Shelf (NWS) 1982-2017

Reprocessing for Copernicus Marine environment Monitoring Service (CMEMS)

NWS products will be retired this year (superceded by ATL series)



- Service migrated to central Ifremer Petascale infrastructure
 - More robust and sustained infrastructure
 - Full history of data available
- Access
 - FTP
 - OSI SAF (+HTTP) : request to <http://osi-saf.org>
 - CMEMS products : <http://marine.copernicus.eu/>
 - Others: <http://cersat.ifremer.fr/data/collections/ghrsst>
 - Thredds, OpeNDAP and WMS protocols for OSI SAF and Medspiration products : <http://tdso.ifremer.fr/thredds/CERSAT.html>
 - Remote processing with Jupyterlab : request to cersat@ifremer.fr



Jupyterhub on HPC

Experimental service, on demand
(cersat@ifremer.fr)

Access to complete EU-GDAC data
archive

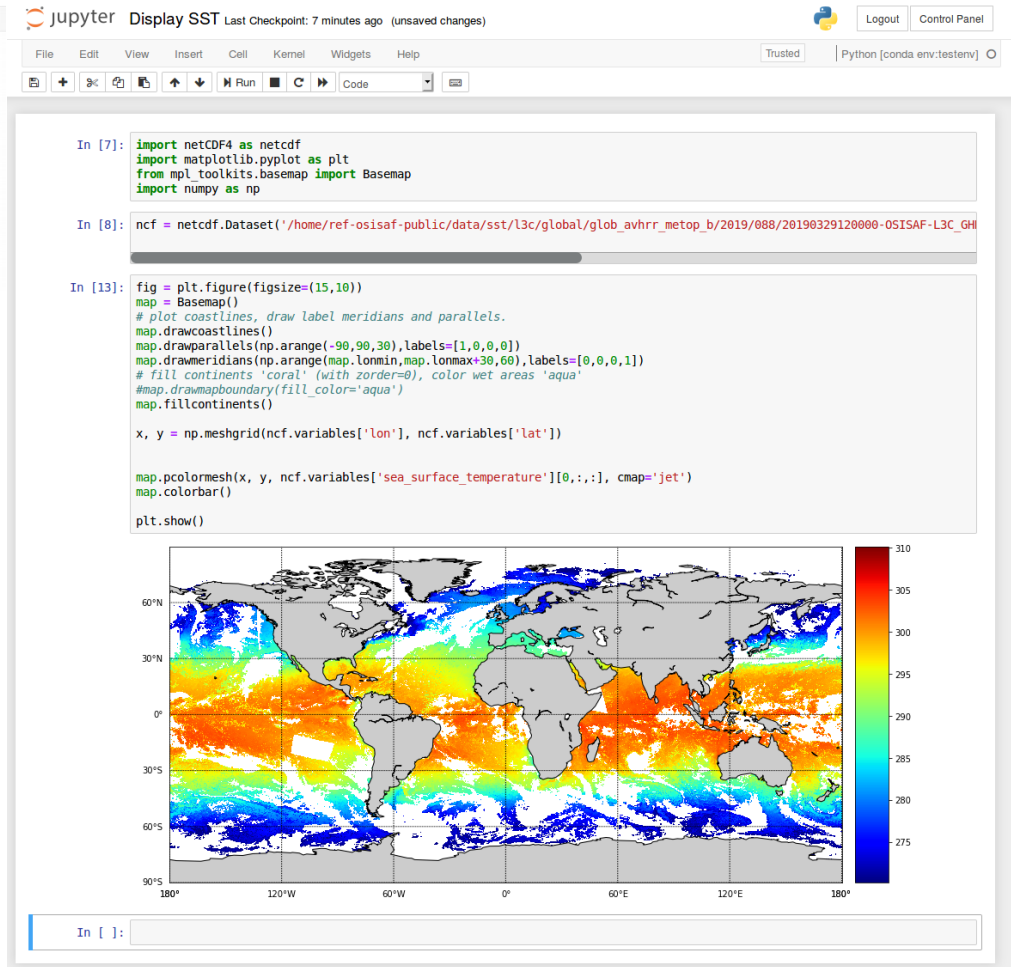
Different resource usage
configuration

Predefined or customizable conda
environments

Spawner options

Select a job profile:

- Datarmor - 1 core, 2GB RAM, 2 hours
- Datarmor - 1 core, 2GB RAM, 2 hours
- Datarmor - 1 core, 60GB RAM, 4 hours
- Datarmor - 4 cores, 8GB RAM, 2 hours
- Datarmor - 8 cores, 16GB RAM, 2 hours
- Datarmor GNU Octave 4.2.2 - 1 core, 10GB RAM, 2 hours
- Spark 2.2.0 for Python 2.7 - 28 cores * 100GB, 2 hours
- Spark 2.2.0 for Python 3.6 - 28 cores * 100GB, 2 hours
- Spark 2.3.1 for Python 2.7 - 28 cores * 100GB, 2 hours
- Spark 2.3.1 for Python 3.6 - 28 cores * 100GB, 2 hours
- SDAP Nexus with Spark 2.2.0 - 28 cores * 100GB, 2 hours



jupyter Display SST Last Checkpoint: 7 minutes ago (unsaved changes)

```
In [7]: import netCDF4 as netcdf
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap
import numpy as np

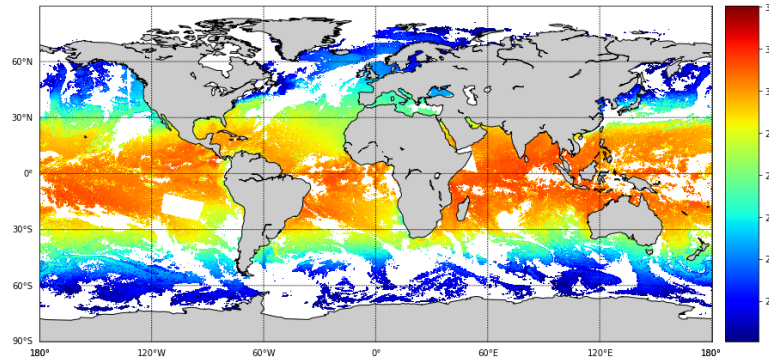
In [8]: ncf = netcdf.Dataset('/home/ref-osisaf-public/data/sst/l3c/global/glob_avhrr_metop_b/2019/088/20190329120000-OSISAF-L3C_GH

In [13]: fig = plt.figure(figsize=(15,10))
map = Basemap()
# plot coastlines, draw label meridians and parallels.
map.drawcoastlines()
map.drawparallels(np.arange(-90,90,30),labels=[1,0,0,0])
map.drawmeridians(np.arange(map.lonmin,map.lonmax+30,60),labels=[0,0,0,1])
# fill continents 'coral' (with zorder=0), color wet areas 'aqua'
map.drawmapboundary(fill_color='aqua')
map.fillcontinents()

x, y = np.meshgrid(ncf.variables['lon'], ncf.variables['lat'])

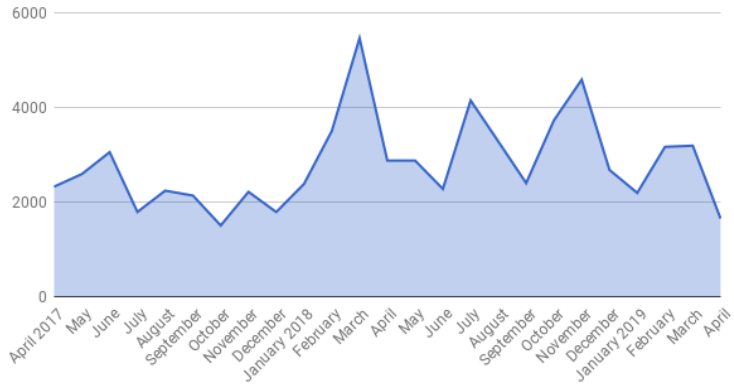
map.pcolormesh(x, y, ncf.variables['sea_surface_temperature'][0,:,:], cmap='jet')
map.colorbar()

plt.show()
```





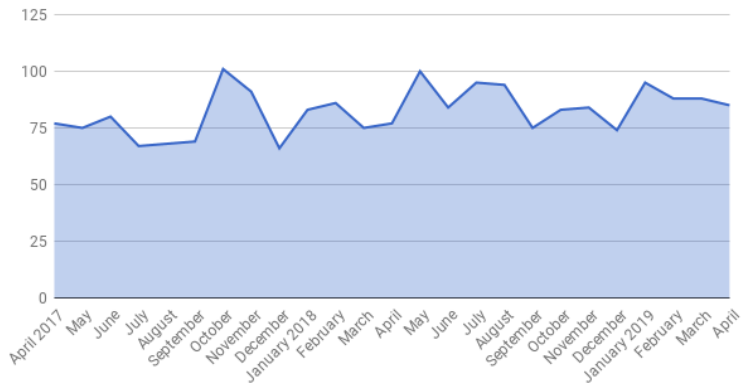
Volume distributed, in TB



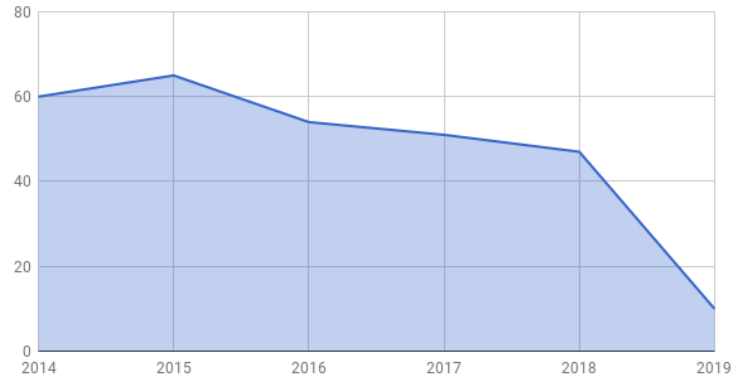
Non European data is less than 10 % of download

Mirroring of « light » products (« GDAC ») seems not so relevant anymore, unless added value or services

Number of unique visitors



New registered users every year (excl. CMEMS & OSI SAF)





- Build upon GDS success
 - Still quite unique consensus in a community
 - Used as an explicit model/reference in many other projects : CMEMS, Glob<*>, ESA CCI data standards
 - People are watching what we're doing !
- Growing number of sensors and products, multiple versions and reprocessings
 - Impossible to evaluate all alternatives
 - Need fair, objective, quantitative and independant evaluation criteria, intercomparison, tools, reference data resources
 - Traceability
 - log of changes and issues
 - Multiple copies
 - propagation of information (L2P > L3 > L4 > Applications)
 - Reproductability : references, metadata, practices, ...
- Data on the cloud is here now
 - Commercial services (Google, Amazon,...) are now far ahead what can be implemented or offered by academic and national organizations
 - How do ease this while keeping some control on what data are made available?
 - Extra care should be paid to what we produce and distribute as it gets quickly out of control
 - Many challenges ahead: new ways to format/store data – NetCDF not well suited for massive processing