

WP5 – Training session: RGeostats

Ellip Solutions for IMR dataset

INTAROS – General Assembly

Haus der Wissenschaft,
Sandstrasse 4/5
28195 Bremen, Germany
14.00-17.30
January 10th 2019

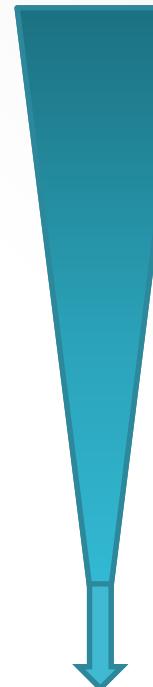
 ORS Fabien, ARMINES (lead)
RENARD Didier, ARMINES (co-lead)

Outline

Thursday 10th

1. Creating iAOS Processing Services
2. Geostatistics and RGeostats
3. Ellip Notebooks using RGeostats
4. Ellip Worflow using RGeostats
5. IMR Case Study - RGeostats in Action!

Details level



Friday 11th

Geostatistics Course & Exercises



Part #1

Creating iAOS Processing Services



3



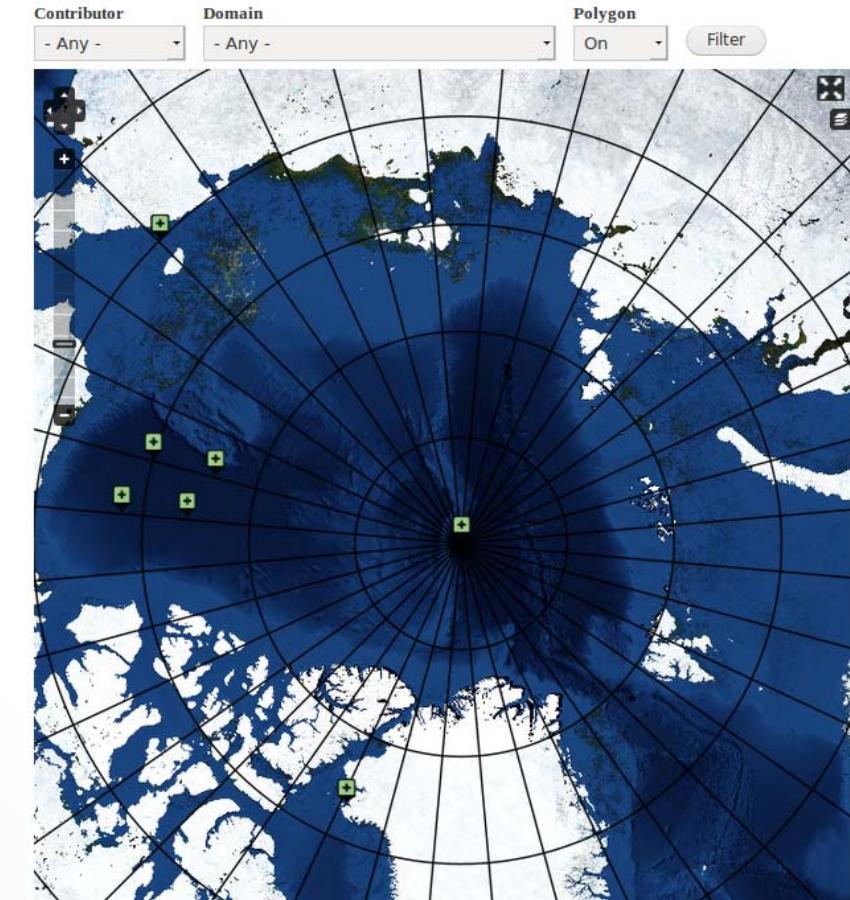
INTAROS

Creating iAOOS Processing Services

Why?

- Visualizing and analyzing Pan-Arctic data on the iAOOS
- Studying spatial and temporal correlations
- Mapping multiple variables
- Combining different data sources
- Learning more from available data with a deeper analysis

Map of observing sites



Creating iAOS Processing Services

For who?

- Developers:
 - INTAROS partners (data providers, iAOS developers, WP6 projects...)
 - INTAROS arctic scientific community (researchers, scientists...)
- iAOS users (*visitors*):
 - Scientists (climate, meteorology, biology, oceanography,...)
 - Companies (tourism, earth resources, environments...)
 - Non Governmental Organization (climate change, pollution, life preservation...)
 - European commission (politics)
 - World citizens



Creating iAOS Processing Services

For who?

- Developers:
 - INTAROS partners (data providers, stakeholders...)
 - Arctic scientific community (researchers, scientists)
- iAOS users (*visitors*):
 - Scientists (climatologists, meteorologists, biologists, oceanologists,...)
 - Companies (tourism, earth resources, environments)
 - Non Governmental Organization (climate change, pollution, life preservation)
 - European commission (politics)
 - Curious world citizens



The rest of this talk is mainly targeted to **developers**



Creating iAOs Processing Services

New Ellip Solution provided by TERRA)UE (sept. 2018)



<https://notebooks-qa.terradue.com>

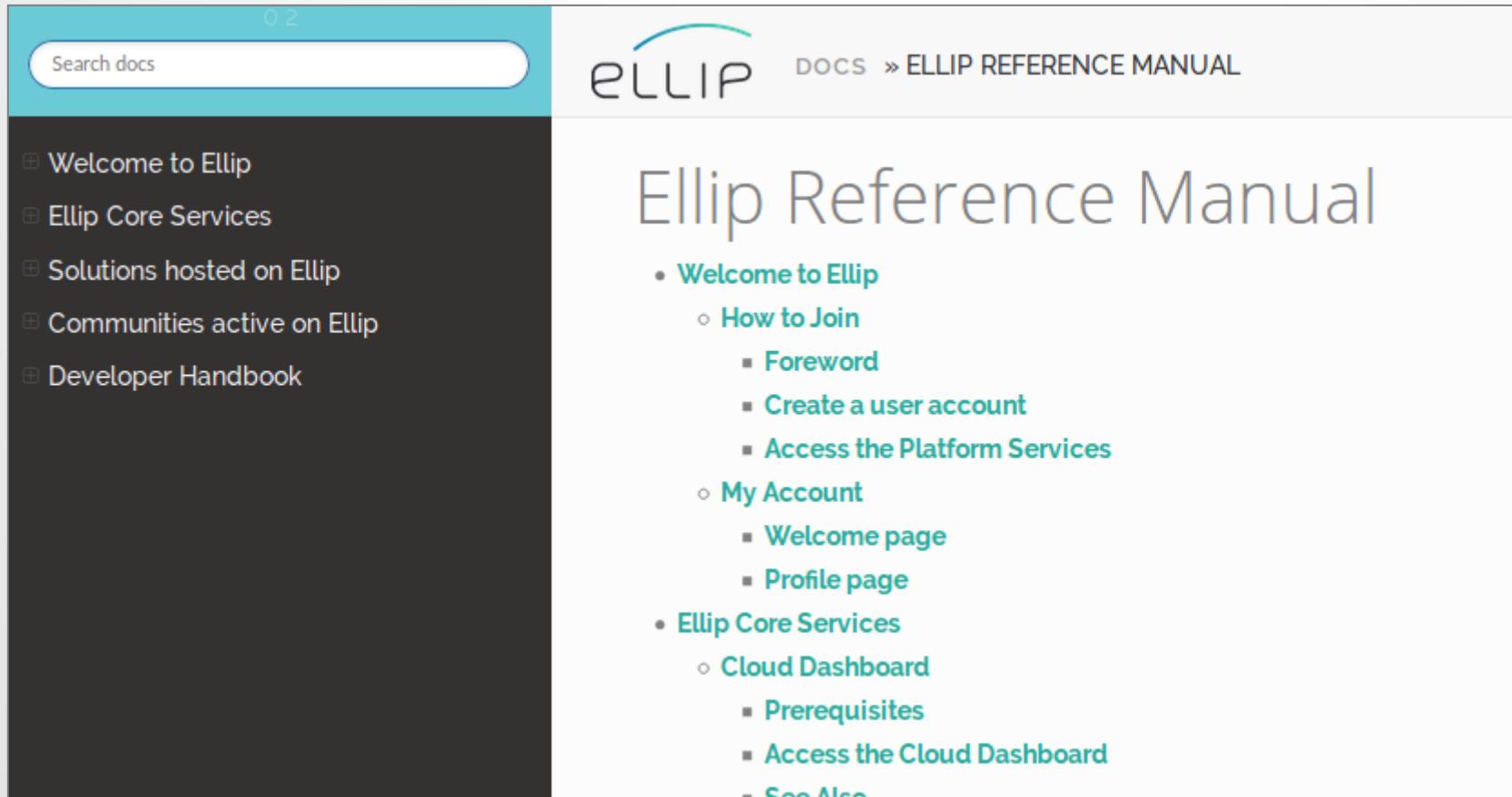
Objectives:

- Make data processing service **design easier**
- **Build, deploy and monitor** data processing services



Creating iAOOS Processing Services

Full set of Ellip documentation



The screenshot shows a documentation page for 'ELLIP'. The top navigation bar includes a search bar labeled 'Search docs' and a breadcrumb trail 'DOCS » ELLIP REFERENCE MANUAL'. The main content is titled 'Ellip Reference Manual' and lists several sections and sub-sections. The sections are: 'Welcome to Ellip' (with sub-sections 'How to Join', 'Foreword', 'Create a user account', 'Access the Platform Services'), 'My Account' (with sub-sections 'Welcome page', 'Profile page'), 'Ellip Core Services' (with sub-sections 'Cloud Dashboard', 'Prerequisites', 'Access the Cloud Dashboard', 'See Also'). The sidebar on the left contains a table of contents with links to 'Welcome to Ellip', 'Ellip Core Services', 'Solutions hosted on Ellip', 'Communities active on Ellip', and 'Developer Handbook'.

- [Welcome to Ellip](#)
 - [How to Join](#)
 - [Foreword](#)
 - [Create a user account](#)
 - [Access the Platform Services](#)
 - [My Account](#)
 - [Welcome page](#)
 - [Profile page](#)
- [Ellip Core Services](#)
 - [Cloud Dashboard](#)
 - [Prerequisites](#)
 - [Access the Cloud Dashboard](#)
 - [See Also](#)

<https://docs.terradue.com/ellip>



Creating iAOs Processing Services

Become an Ellip user

1. Request iAOs access to Terradue's Ellip Notebooks solution
2. Install Google Chrome browser
3. Sign-in to the Terradue Portal

SIGN-IN TO ACCESS TERRADUE CLOUD PLATFORM SERVICES

Use your Terradue account, or use an Identity Provider trusted by Terradue

Use your Terradue account credentials

Username or Email

Password (forgot your password)

Sign in

New here? [Create a free account](#)

Or, use your credentials from a trusted Identity Provider

 [Sign in with ESA EO](#)

 [Sign in with Ever-EST](#)

 [Sign in with Co-ReSyF](#)

 [Sign in with NextGEOSS](#)

<https://www.terradue.com>

Creating iAOs Processing Services

Configure the workshop

1. Open Google Chrome and sign-in to Terradue:

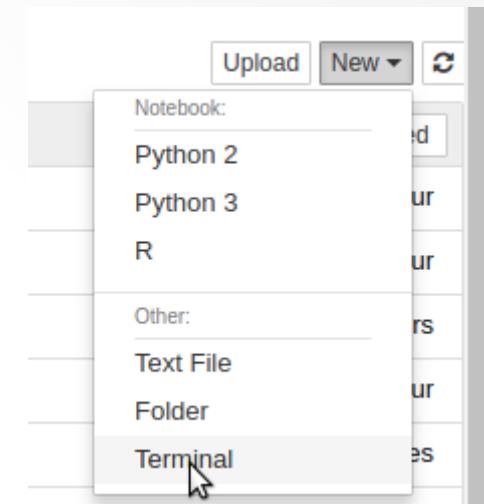
<https://www.terradue.com>

2. Browse to Jupyter Notebook and open a Terminal:

<https://notebooks-qa.terradue.com>

2. Execute the 3 following commands:

```
$ rm -f configure_workshop.sh
$ wget http://rgeostats.free.fr/doc/Files/configure\_workshop.sh
$ chmod +x configure_workshop.sh
$ ./configure_workshop.sh
```

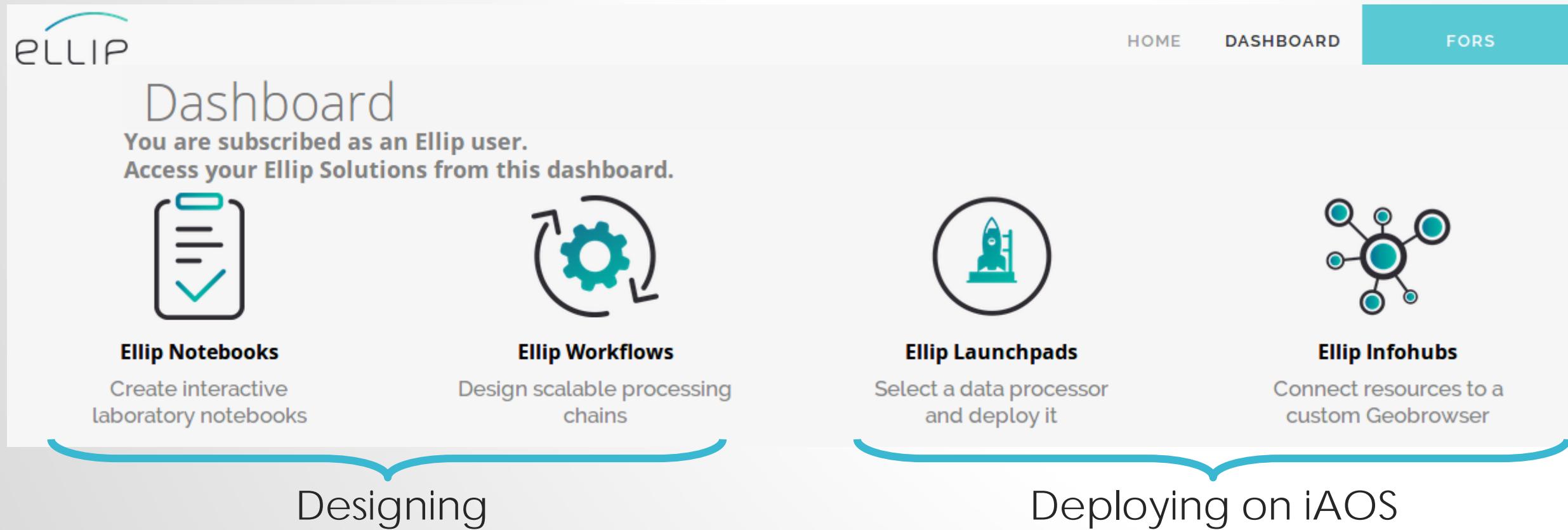


This may take several minutes...



Creating iAOs Processing Services

4 steps to create and use data processors

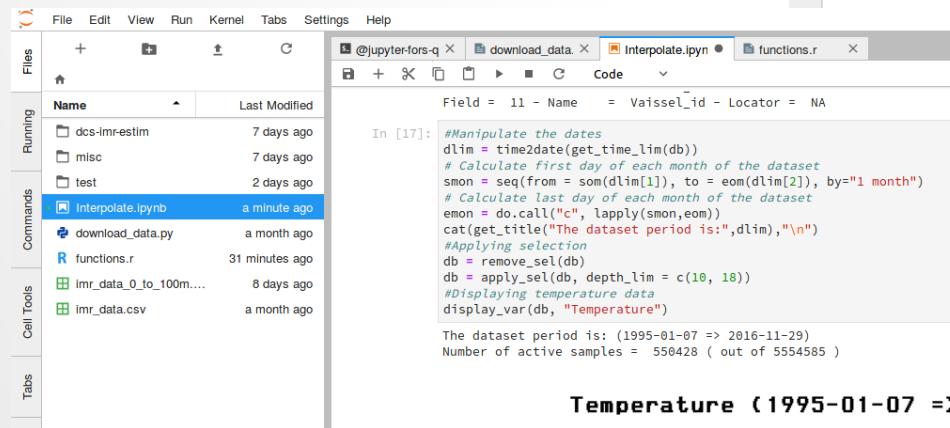


Creating iAOs Processing Services

1. Designing data processors

1. Access **Jupyter Notebooks** and your lab:
<https://notebooks-qa.terradue.com>
2. Configure your Jupyter environment
3. Write your Notebooks (Python or R)
4. Test your Notebooks and analyze results

- Python Kernel
- R Kernel
- Terminal



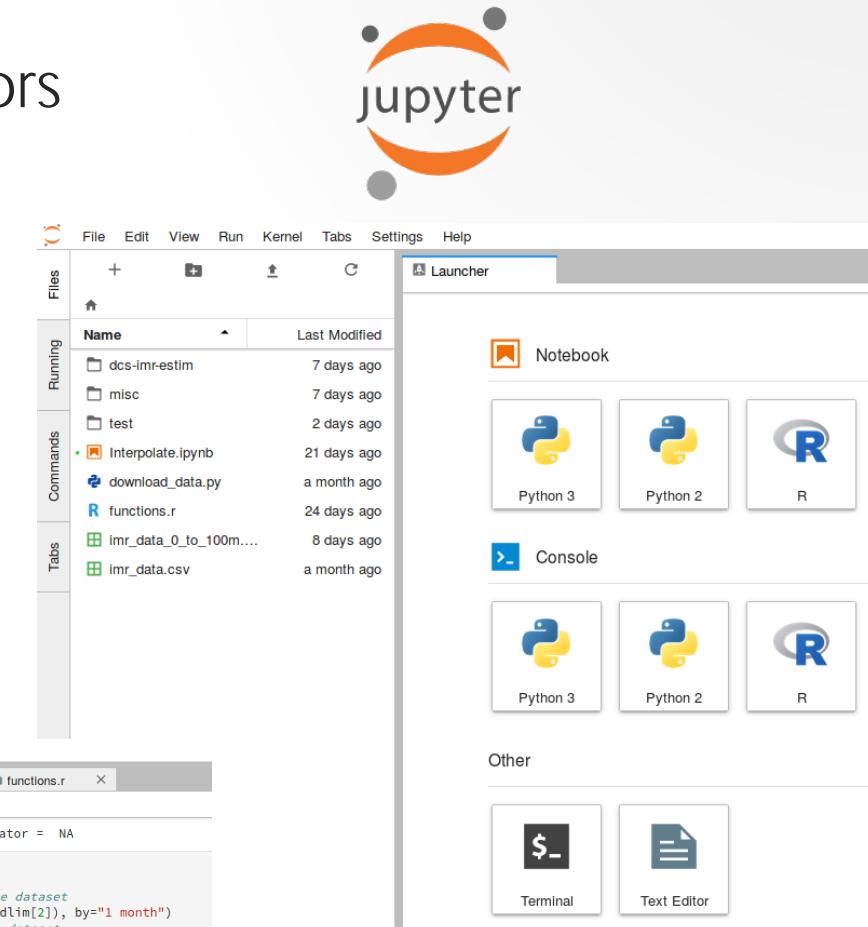
```

Field = 11 - Name = Vaissel_id - Locator = NA
In [17]: #Manipulate the dates
dlim = time2date(get_time_lim(db))
# Calculate first day of each month of the dataset
smon = seq(from = som(dlim[1]), to = eom(dlim[2]), by="1 month")
# Calculate last day of each month of the dataset
emon = do.call("c", lapply(smon, eom))
cat(get_title("The dataset period is:", dlim), "\n")
#applying selection
db = remove_sel(db)
db = apply_sel(db, depth_lim = c(10, 18))
#displaying temperature data
display_var(db, "Temperature")

The dataset period is: (1995-01-07 => 2016-11-29)
Number of active samples = 550428 (out of 5554585 )

```

Temperature (1995-01-07 =>



Creating iAOS Processing Services

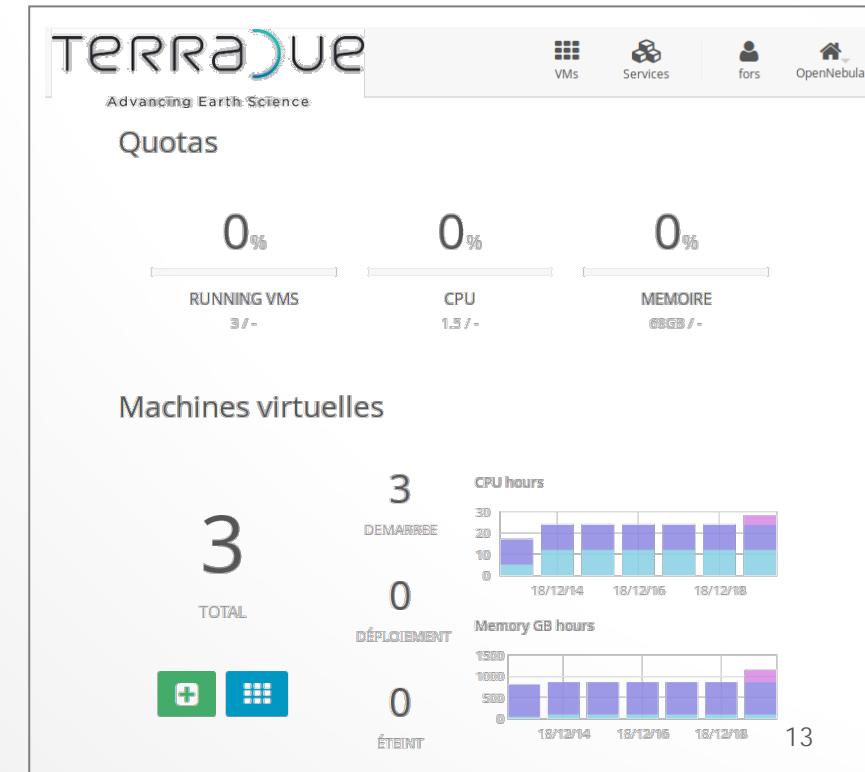
2. Building your workflow (1/2)

1. Access to your Cloud Dashboard <https://cloud.terradue.com/>
2. Create an **Ellip Workflows** Virtual Machine (VM)



While the VM is deploying...

3. Install SSH key pairs
4. Configure the VPN client
(openvpn)



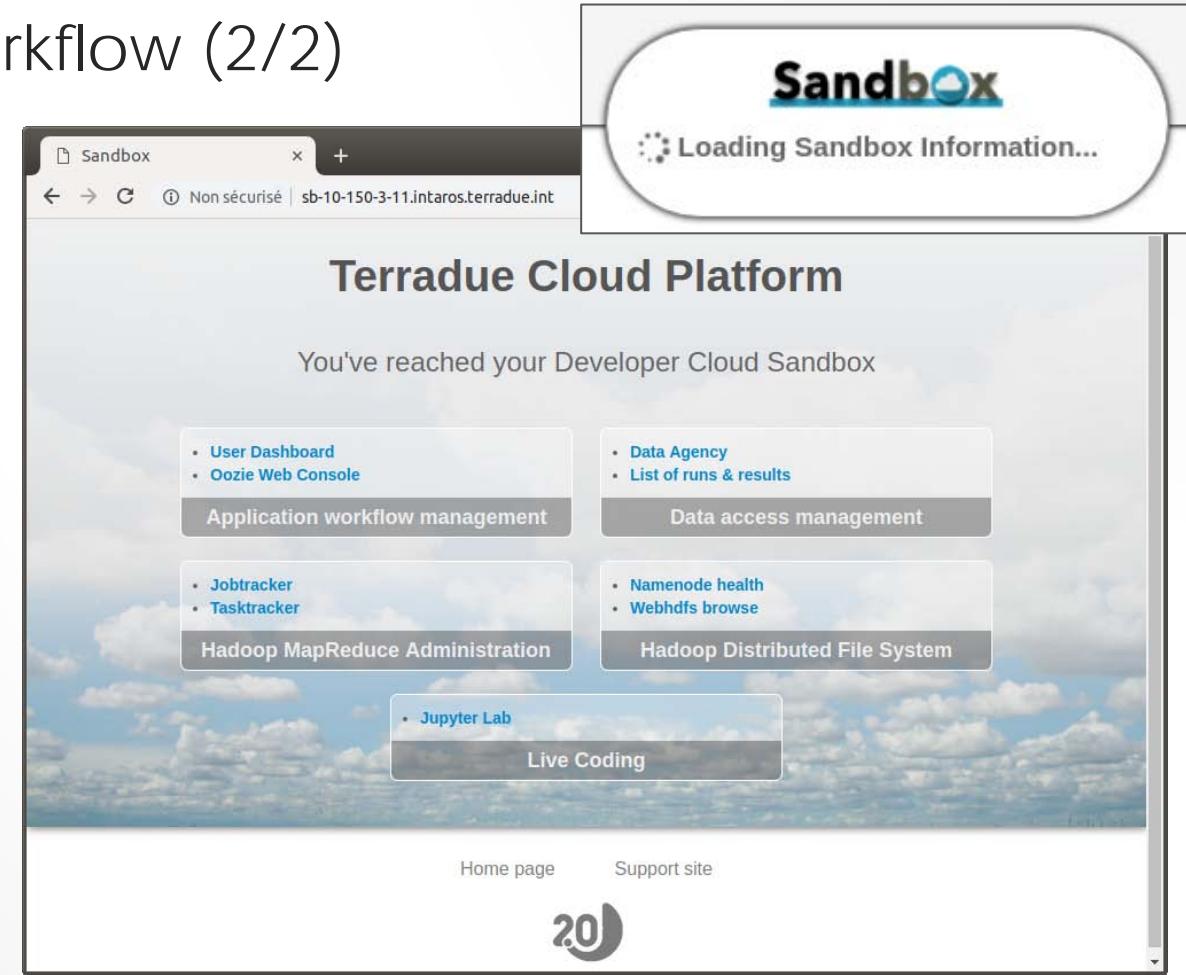
Creating iAOOS Processing Services

2. Building your workflow (2/2)



Workflows

1. Connect to Terradue's **VPN**
2. Connect via **ssh** to your VM
3. Configure your VM system environment
4. Create an Ellip workflow application using **ciop**, **git** and **mvn**
5. Test your application with the **Sandbox**



The screenshot shows the Terradue Cloud Platform interface. At the top, a window titled "Sandbox" is open, displaying the message "Loading Sandbox Information...". Below this, the main dashboard is titled "Terradue Cloud Platform" and displays the message "You've reached your Developer Cloud Sandbox". The dashboard is organized into several sections: "Application workflow management" (User Dashboard, Oozie Web Console), "Data access management" (Data Agency, List of runs & results), "Hadoop MapReduce Administration" (Jobtracker, Tasktracker), "Hadoop Distributed File System" (Namenode health, Webhdfs browse), and "Live Coding" (Jupyter Lab). At the bottom of the dashboard, there are links for "Home page" and "Support site", and the INTAROS logo.



Creating iAOs Processing Services



Launchpads

1. Manage successive workflow releases (**ciop-release**)
2. Package it in a Web Processing Service (**WPS**)
3. Make it available to geo-portals

3. Deploying your workflow

Select a data processor and deploy it

NextGEOSS

Data processors

<input type="checkbox"/> gridded- data.armines.nextgeoss.terradue.com	Cluster 2
<input type="checkbox"/> maxent.wur.nextgeoss.terradue.com	

Available clusters

Cluster 1	Deploy
	Undeploy
Cluster 2	Deploy
	Undeploy



Creating iAOs Processing Services

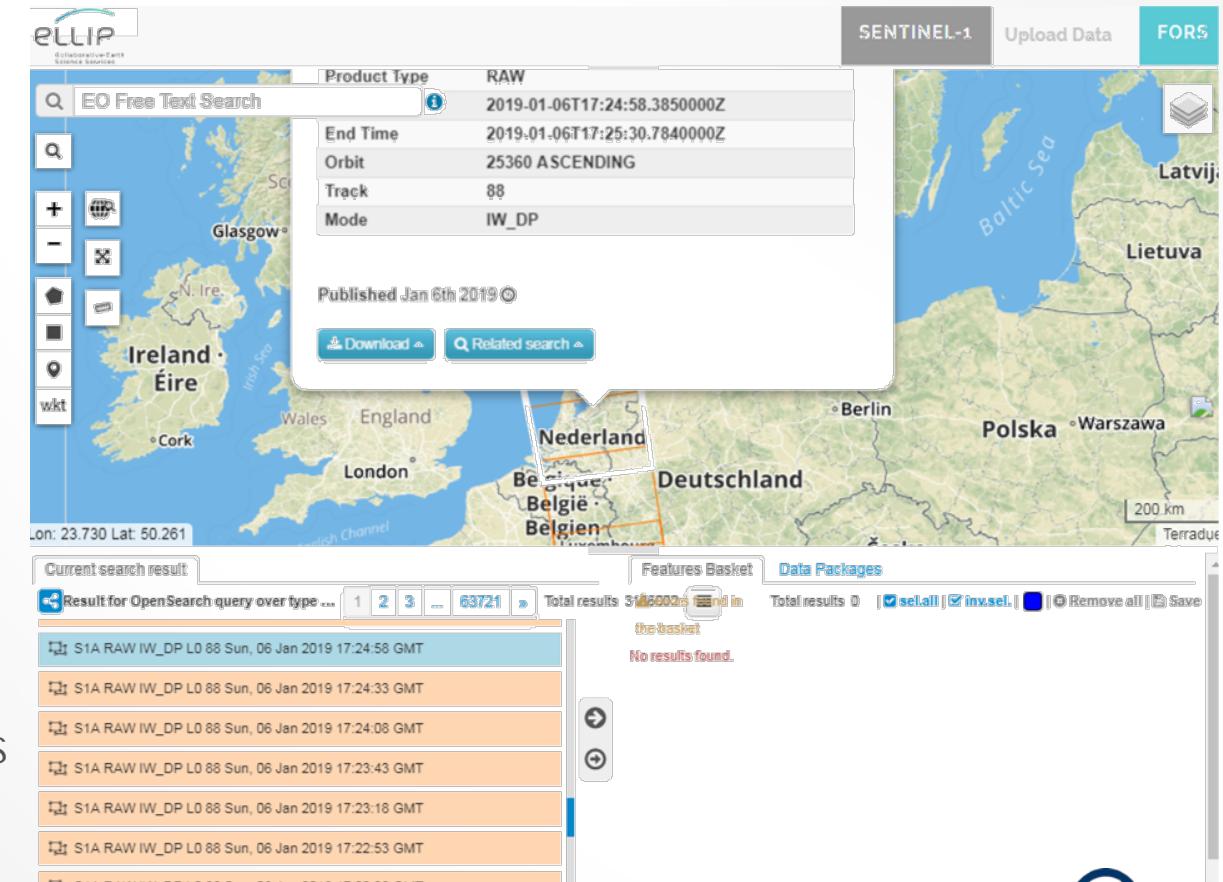
4. Running your workflow

1. Create a geo-browser application
2. Connect to your data
3. Launch your deployed Web Processing Services



Infohubs

This could be an example of interface for iAOs visitors



The screenshot shows the ELLIP web interface. At the top, there are buttons for 'SENTINEL-1', 'Upload Data', and 'FORIS'. The main area features a map of Europe with a bounding box drawn around the Netherlands and Belgium. A callout box displays search parameters: Product Type (RAW), End Time (2019-01-06T17:24:58.385000Z), Orbit (25360 ASCENDING), Track (88), and Mode (IW_DP). Below the map, a table lists search results for an 'EO Free Text Search' for 'S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:24:58 GMT'. The table shows 63721 results, with the first few entries listed as S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:24:58 GMT, S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:24:33 GMT, S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:24:08 GMT, S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:23:43 GMT, S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:23:18 GMT, and S1A RAW IW_DP L0 88 Sun, 06 Jan 2019 17:22:53 GMT. The interface includes buttons for 'Download' and 'Related search'.



INTAROS

Part #2

Geostatistics and RGeostats



17

Who are we?



60 Bd St Michel in Paris

MINES ParisTech



Training of engineers



Research centers

Fontainebleau
Earth sciences & Applied mathematics

Corbeil - Evry

**Sophia
Antipolis**

Centre de Géosciences



Who are we?

Centre de Géosciences in Fontainebleau: Géostatistique



Names in Geostatistics

- D.G. Krige (mining engineer) & H.S. Sichel (statistician):
Gold mining (1950)
- L. Gandin:
Objective Analysis of Meteorological Fields (1965)
- B. Matérn:
Importance of spatial dependence and variation in forestry (1960)
- G. Matheron:
 - Formalization of the theory of regionalized variables, Ecole des Mines de Paris (1962)
 - Inception of Mathematical Morphology discipline
- Many fellows / students:
 - M. David, A. Journel, J. Davis, D. Merriam, R. Webster, N. Cressie, ...



Geostatistics

- Describe the spatial characteristics of the variable (variogram): classification, spatial correlation
- Estimation, interpolation (kriging)
- Simulations: possible alternative scenarios
- Appraisal of uncertainty
- Risk assessment



Applications

- Soil Science, Topography, Geology
- Mining, Petroleum
- Hydrology
- Biology, Epidemiology, Health, Ecology
- Environmental monitoring and assessment
- Forestry, Fishing, Agronomy
- Atmospheric Sciences
- Any discipline with spatial data



RGeostats

- Created in Centre de Géosciences of MINES ParisTech
- Initiated for Fish Industry within a European project: GEFA (2001)
- Named RGeoS and expanded possibilities:
 - Based on a separate commercial library Geoslib (written in C)
 - Mapping its functions in R language (using Rcpp layer)
- Renamed in RGeostats in 2014
- Used for testing new methodologies and/or for teaching purpose
- (Free) Download from: <http://cg.ensmp.fr/rgeostats>
- Provides demonstration scripts, FAQ, forum for posting questions, asking for help and benefiting from users experience



Trainings

More extensive courses in Geostatistics are available:

- *Les méthodes de la Géostatistique* (in French): October 2019 in Paris & Fontainebleau
- C.F.S.G. (in English) : ten months from September to July
- *Mining professional training* (during academic year in English)
- *Introduction to Geostatistics* (in English): 18-22 Fev 2019 in Paris

More information:

<http://www.geosciences.mines-paristech.fr>

Contact:

nathalie.dietrich@mines-paristech.fr



Bibliography

G. Matheron:

- 1962-1963: *Treatise of applied geostatistics* (in French), Technip and BRGM editions, Paris
- 1965: *Regionalized variables and their estimation* (in French), Masson, Paris
- 1967: *Elements for a theory of porous media* (in French), Masson, Paris
- 1968: *Treatise of applied geostatistics* (in Russian), MIR, Moscow
- 1969: *Theory of random sets* (in French), Ecole des Mines de Paris
- 1969: *Geostatistics course* (in French), Mines Paris
- 1969: *Universal kriging* (in French), Mines Paris
- 1970: *Mathematical morphology* (in French), Mines Paris
- 1970: *The theory of regionalized variables and its applications*, Mines Paris
- 1972-1975: *Random sets and integral geometry*, Wiley, New York
- 1978-1989: *Estimating and choosing*, Springer, Berlin



Bibliography

From others:

- Armstrong M., Galli A., Beucher H., Le Loc'h G., Renard D., Doligez B., Eschard R., Geffroy F., *Plurigaussian Simulations in Geosciences*, 2011. 176p.
- Chilès J.P., Delfiner P. *Geostatistics : modeling spatial uncertainty*. N.Y. : Wiley, 1999. 695p. 2nd edition 2012.
- Journel A., Huijbregts C. *Mining geostatistics*. London : Academic Press, 1978. 600p.
- Lantuéjoul C. *Geostatistical simulation : models and algorithms*. Berlin : Springer, 2001. 256p.
- Rivoirard J. *Introduction to disjunctive kriging and non-linear geostatistics*. Oxford : Clarendon Press, 1994. 181p.
- Wackernagel H. *Multivariate geostatistics: an introduction with applications*. 3rd ed. Berlin : Springer, 2003. 387p.



Part #3

Ellip Notebooks using RGeostats

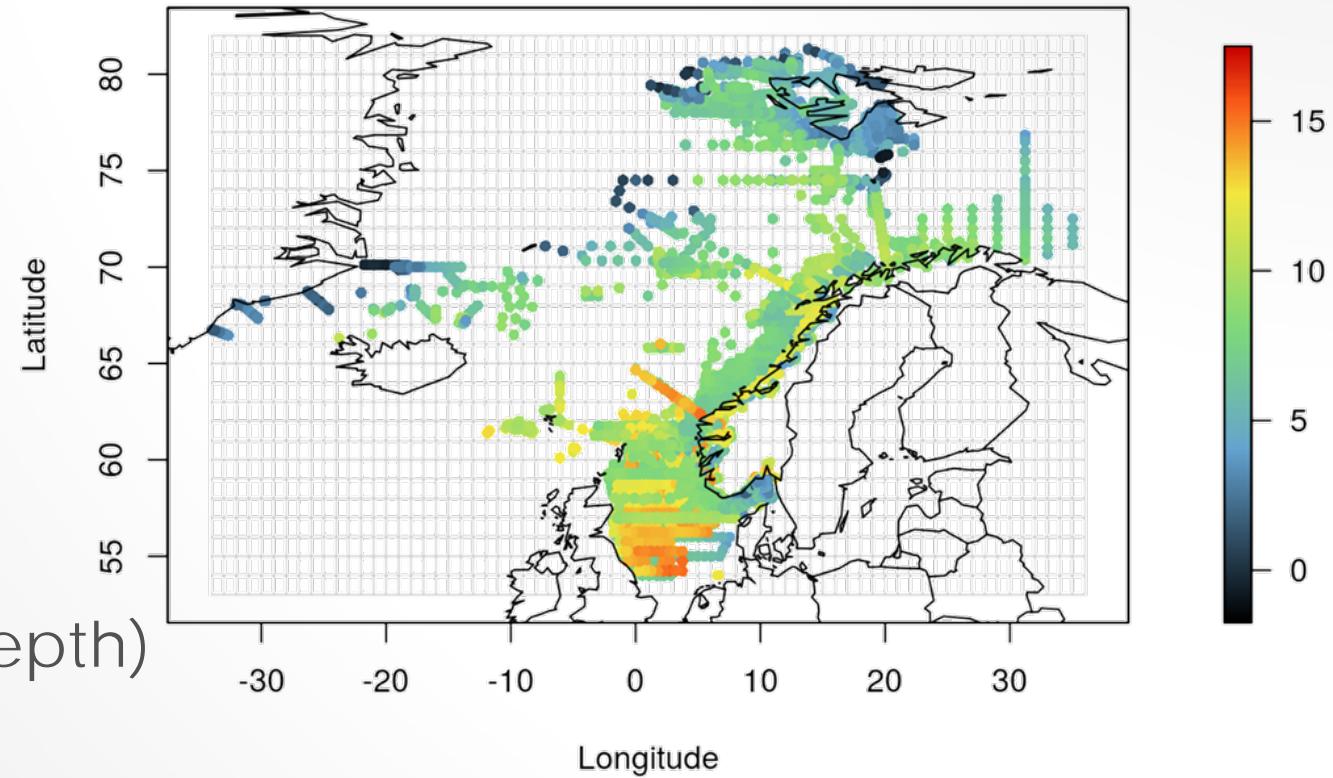


27

Ellip Notebooks using RGeostats

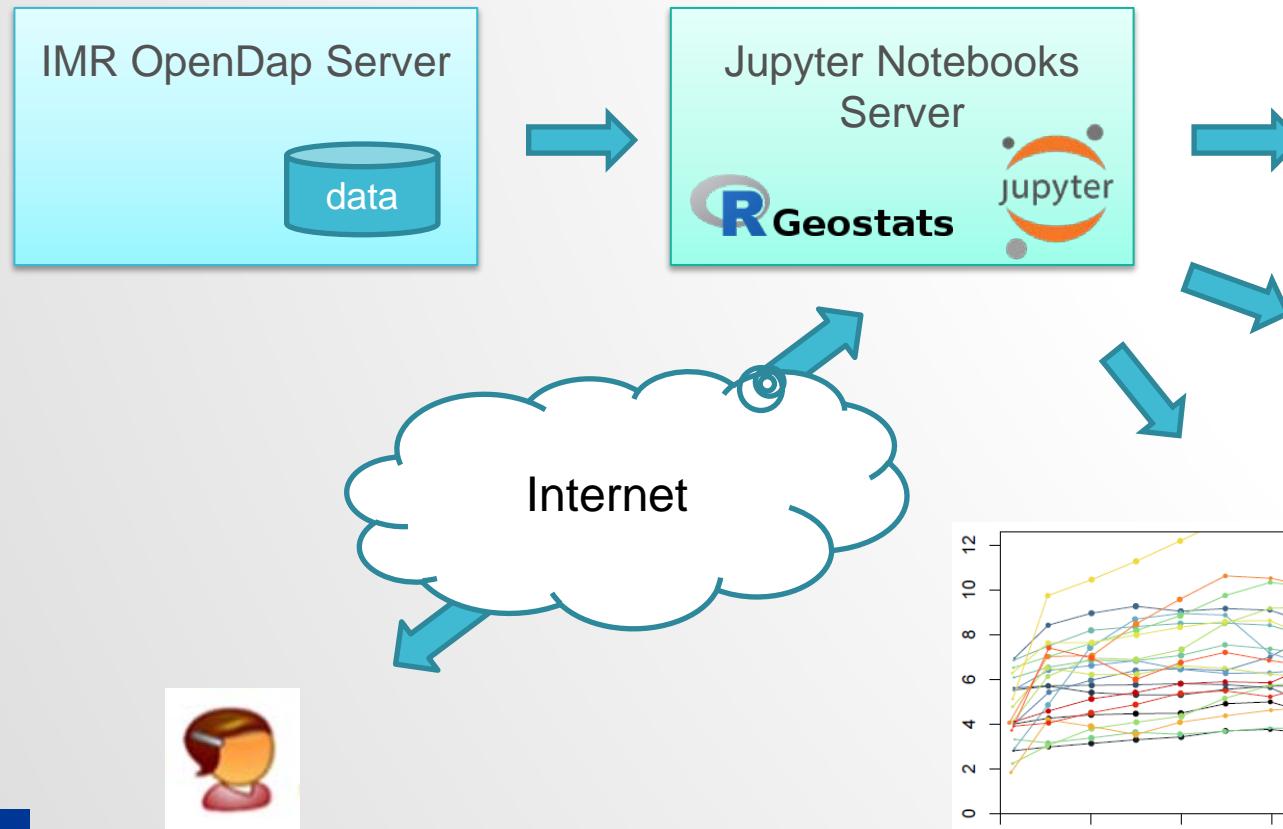
IMR Dataset global overview

- 7 vessels
- from 1995 to 2016
- 3 variables measured:
 - Temperature
 - Salinity
 - Conductivity
- 63 500 positions {long, lat}
- 63 500 vertical profiles (in depth)
- A few million samples
- 84 NetCDF files (~60 Mb each)

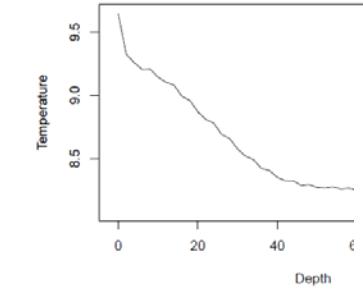


Ellip Notebooks using RGeostats

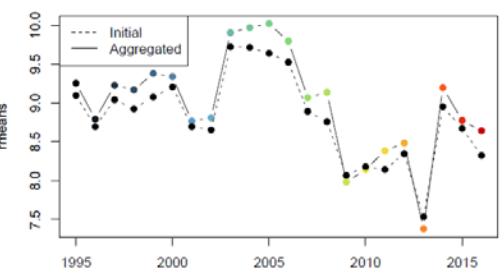
Work in the cloud!



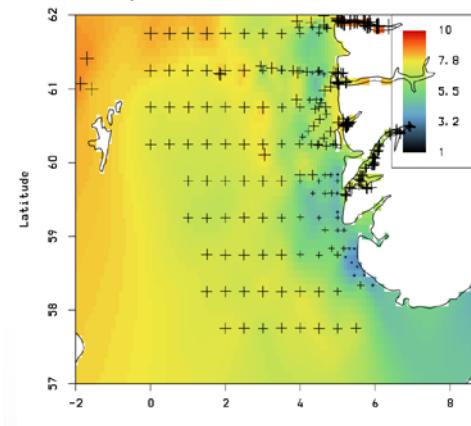
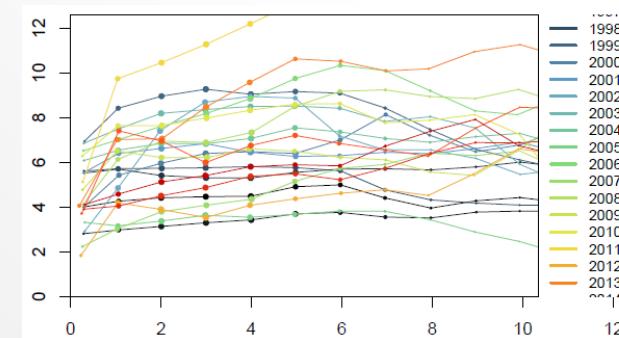
Temperature for Depth (0 => 98)



Mean of Temperature at 25 m



Estimation at 20m depth
Temperature (2001-04-01 => 2001-05-24)



Ellip Notebooks using RGeostats

Available unitary jobs for IMR dataset

One **Python** job for fetching data:

1. Subsetting by location, depth or time period from IMR OpenDAP server

Independent RGeostats **R scripts** jobs:

2. Display basemap of a variable
3. Subsetting / Filtering
4. Statistics by cells
5. Multi-variable correlation
6. 2D aggregation along time or depth
7. Variable plotting through time or depth
8. Interpolation map using kriging
9. ...

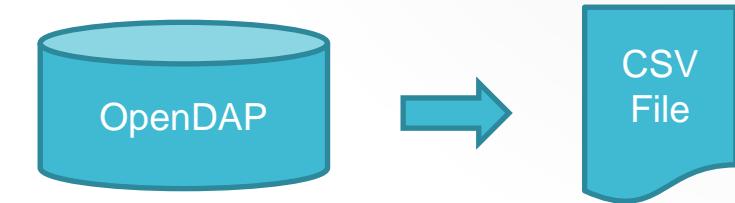


Ellip Notebooks using RGeostats

Python job for downloading data

Python script fetching data into a CSV file

IMR OpenDAP server:
Thredds server



Catalog URL:

http://opendap1-test.nodc.no/thredds/catalogs/physics/physics_point_yearly.xml

File URLs:

http://opendap1-test.nodc.no/thredds/dodsC/physics/point/yearly/58xx_CTD_yyyy.nc

Variables dumped in CSV:

Longitude, Latitude, Depth, Time, Vessel_id, Vessel_name,
Profile_id, **Temperature**, **Salinity**, **Conductivity**

Ellip Notebooks using RGeostats

Download known issues

IMR OpenDAP access directly through R:

- **Rdap** package: Not operational and not maintained
 - **Pydap** module through **reticulate** package: no translation of Pydap indexing within reticulate
- ⇒ Current workaround: **netCDF4** python module
- ⇒ Fetching / subsetting data with python script
 - ⇒ Creating CSV intermediate file
 - ⇒ Loading this CSV file for creating the RGeostats Db

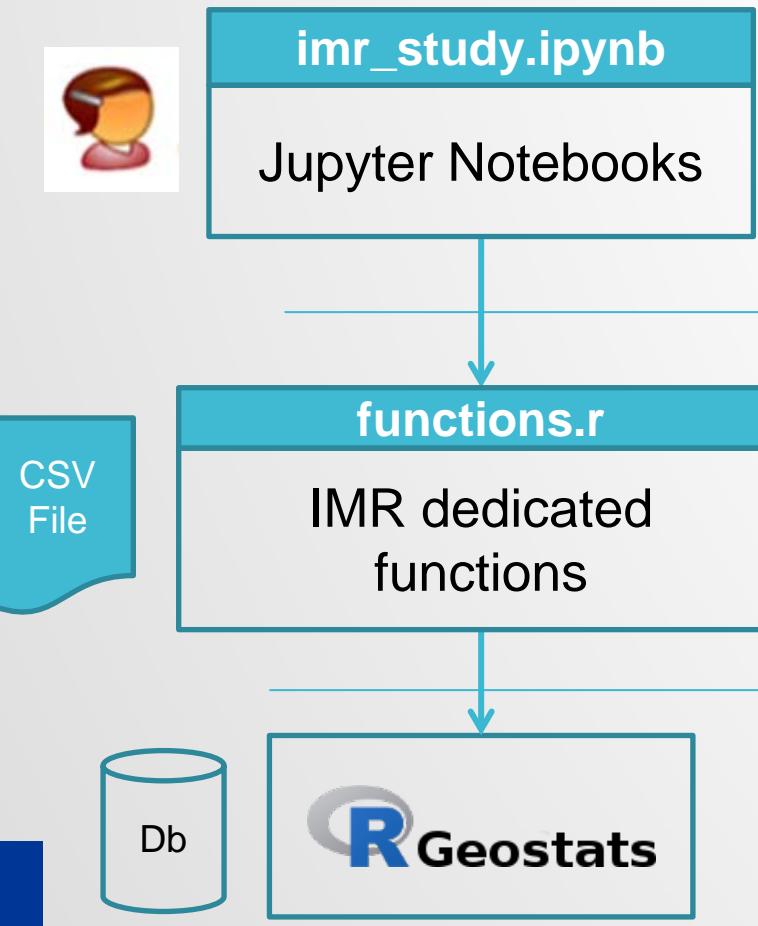
But:

- No numpy advanced indexing support
- No meta dataset covering all boats (yet)
- IMR Remote access to Hyrax OpenDAP server is slow (has to be solved)



Ellip Notebooks using RGeostats

Wrapping RGeostats functions



Unitary jobs for each notebook cell:
Statistics / Aggregation / **Interpolation**

read_csv / write_csv / load_data
date2time / time2date / aggregate_depth
apply_sel / remove_sel / **interpolate_2D**
display_var / display_stats / **display_result**

*db.create / db.locate / db.sel / db.plot
vario.calc / model.auto / kriging
blockstat / db.regularize*

Ellip Notebooks using RGeostats

Interpolation service example : **estimate.R**

```

# Load the CSV file
dbr = read_csv(csv_file)

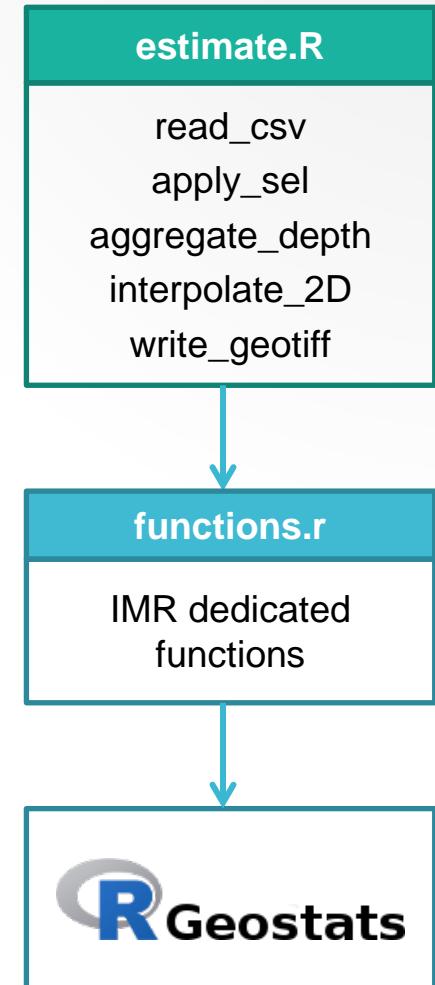
# Select measures in the required interval (time and aera)
dbr = apply_sel(dbr, ...)

# Aggregate all measures around 20m depth
dbr = aggregate_depth(dbr, 20, ...)

# Interpolate Temperature values
res = interpolate_2D(dbr, "Temperature", ...)

# Display results
display_result(dbr, res,...)

```



Part #4

Ellip Worflow using RGeostats



35

Ellip Worflow using RGeostats

Application creation

New workflow:

```
mvn archetype:generate
```

Type: 3 (R workflow with 2 nodes)

Group Id: **com.terradue**

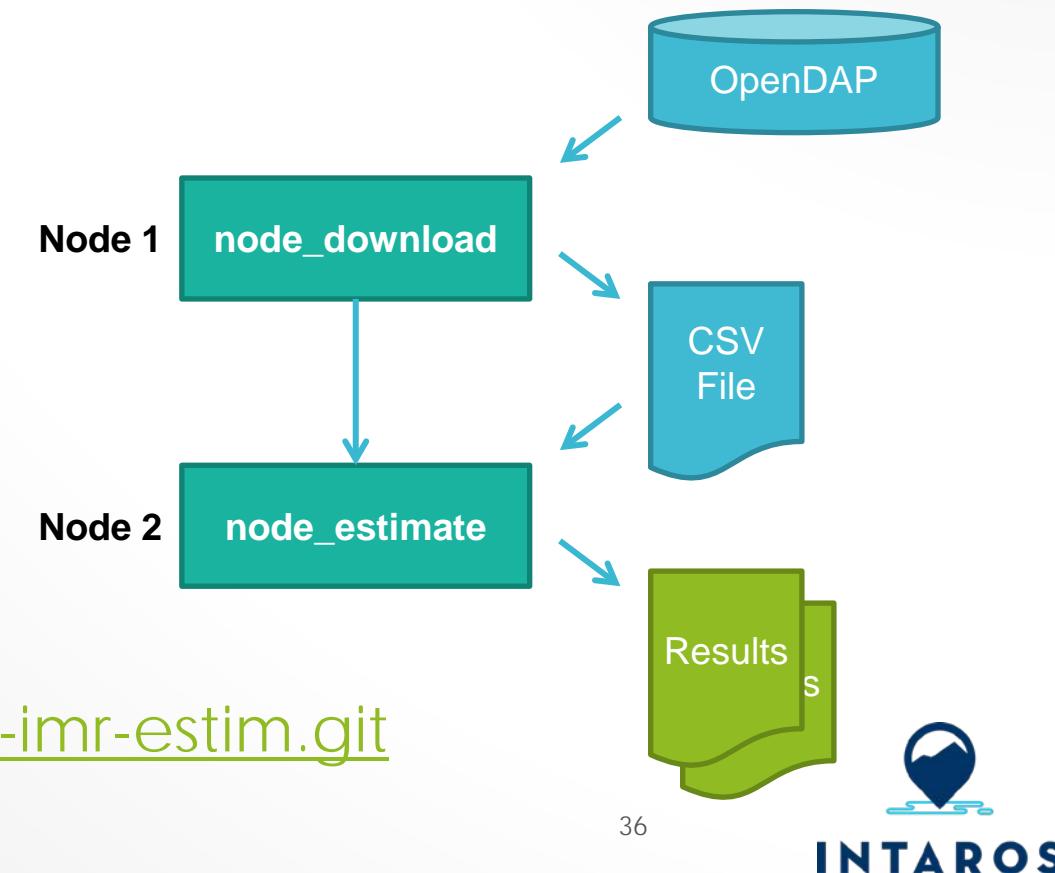
Community: **ec-intaros**

Name: **dcs-imr-estim**

Git repository:



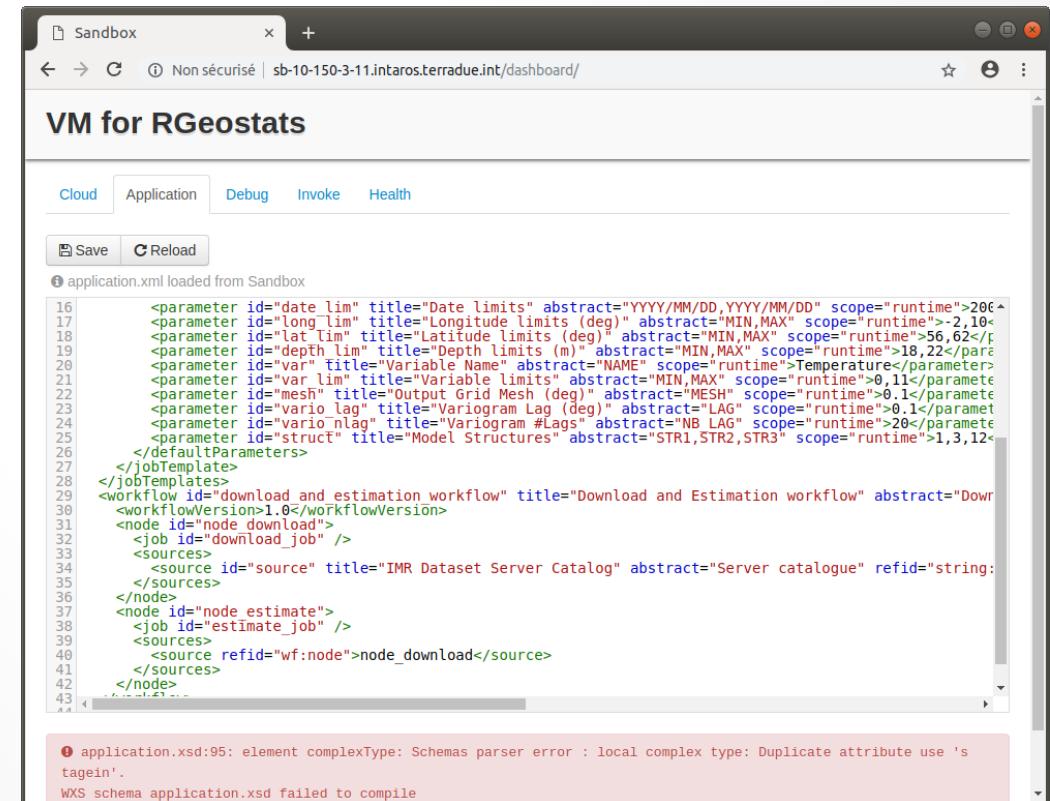
<https://gitlab.com/ec-intaros/dcs-imr-estim.git>



Ellip Worflow using RGeostats

Application.xml file

- Job Templates:
 - Streaming executable = R entry points:
node_download/run.R
node_estimate/run.R
 - Job parameters
(long_lim: Longitude limits
vario_nlag: Variogram lags)
- Workflow sequence:
 - Nodes identification
 - Data sources
 (catalog URL or results from previous node)



```

<parameter id="date_lim" title="Date limits" abstract="YYYY/MM/DD,YYYY/MM/DD" scope="runtime">2016-01-01,2016-12-31</parameter>
<parameter id="long_lim" title="Longitude limits (deg)" abstract="MIN,MAX" scope="runtime">-2,10</parameter>
<parameter id="lat_lim" title="Latitude limits (deg)" abstract="MIN,MAX" scope="runtime">56,62</parameter>
<parameter id="depth_lim" title="Depth limits (m)" abstract="MIN,MAX" scope="runtime">18,22</parameter>
<parameter id="var" title="Variable Name" abstract="NAME" scope="runtime">Temperature</parameter>
<parameter id="var_lim" title="Variable limits" abstract="MIN,MAX" scope="runtime">0,11</parameter>
<parameter id="mesh" title="Output Grid Mesh (deg)" abstract="MESH" scope="runtime">0.1</parameter>
<parameter id="vario_lag" title="Variogram Lag (deg)" abstract="LAG" scope="runtime">0.1</parameter>
<parameter id="vario_nlag" title="Variogram #lags" abstract="NB_LAG" scope="runtime">20</parameter>
<parameter id="struct" title="Model Structures" abstract="STR1,STR2,STR3" scope="runtime">1,3,12</parameter>
</defaultParameters>
</jobTemplate>
</jobTemplates>
<workflow id="download_and_estimation_workflow" title="Download and Estimation workflow" abstract="Download and estimation workflow">
<workflowVersion>1.0</workflowVersion>
<node id="node_download">
<job id="download_job" />
<sources>
<source id="source" title="IMR Dataset Server Catalog" abstract="Server catalogue" refid="string:IMR Dataset Server Catalog"></source>
</sources>
</node>
<node id="node_estimate">
<job id="estimate_job" />
<sources>
<source refid="wf:node">node_download</source>
</sources>
</node>
</workflow>

```

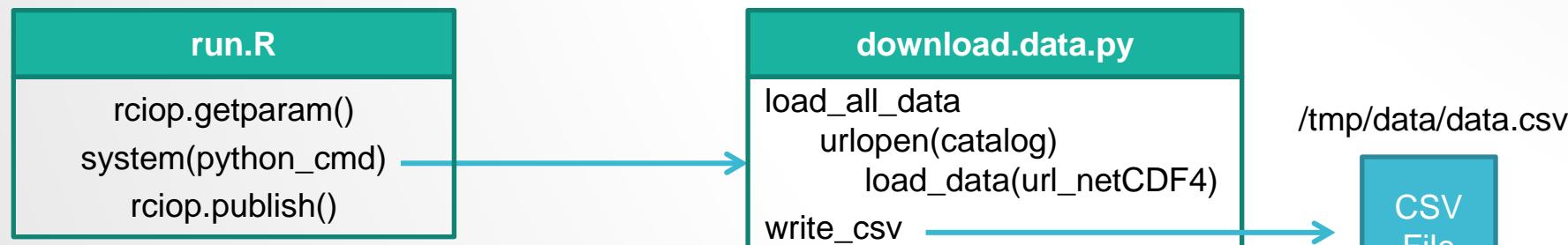
application.xsd:95: element complexType: Schemas parser error : local complex type: Duplicate attribute use 'stagein'.
 WXS schema application.xsd failed to compile



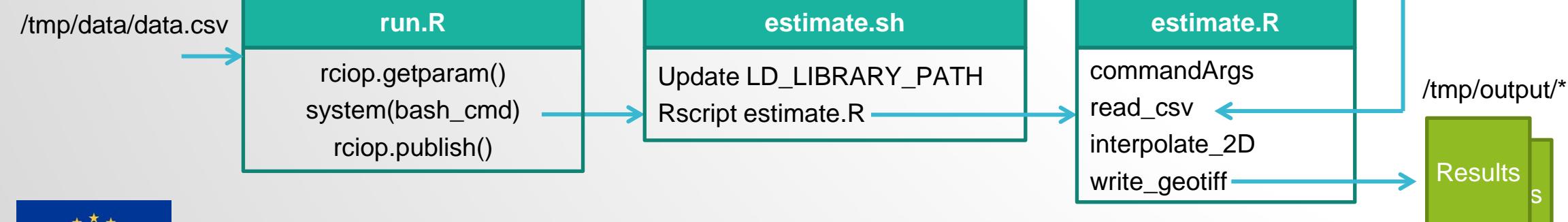
Ellip Worflow using RGeostats

Nodes description

- node_download:



- node_estimate:



Ellip Worflow using RGeostats

Run the workflow

1. Synchronise the application with your git repository (**git pull**)
2. Install the application (**mvn clean install**)
3. Launch the application (**ciop-run**)



Ellip Worflow using RGeostats

Run the workflow

```
$ ci op- run
```

```
2019-01-06 09:13:51 [INFO] - Workflow submitted
2019-01-06 09:13:51 [INFO] - Closing this program will not stop the job.
2019-01-06 09:13:51 [INFO] - To kill this job type:
2019-01-06 09:13:51 [INFO] - ci op- stop 0000065-190101000008245-oozi e-oozi -W
2019-01-06 09:13:51 [INFO] - Tracking URL: 2019-01-06 09:13:51 [INFO] - http://sb-10-150-3-11.intaros.terradue.int:11000/oozi/e/?job=0000065-190101000008245-oozi e-oozi -W
```

```
Node Name      : node_download
Status         : OK
```

```
Node Name      : node_estimate
Status         : OK
```

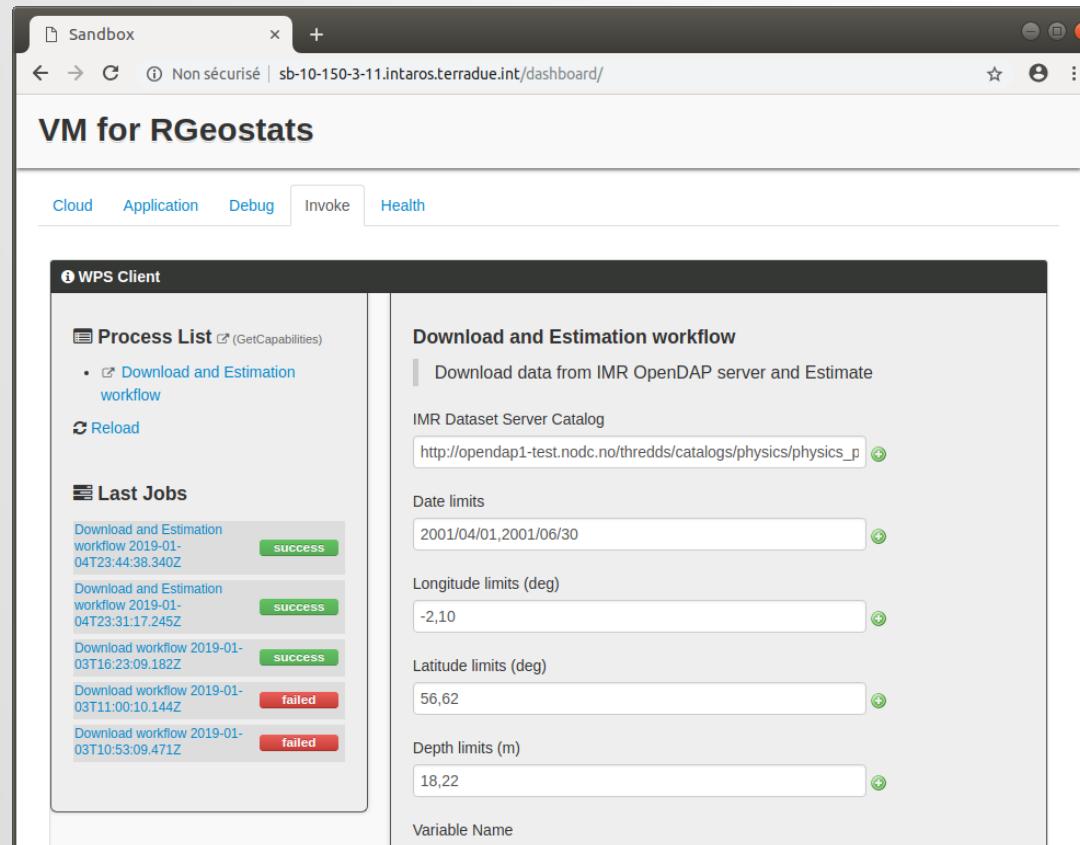
```
Publ i shing resul ts...
```

```
2019-01-06 09:19:52 [INFO] - Workflow completed.
2019-01-06 09:19:52 [INFO] - Output Metalink: http://sb-10-150-3-11.intaros.terradue.int:50070/webhdfs/v1/ci op/run/download\_and\_estimation\_workflow/000065-190101000008245-oozi e-oozi -W/results.metalink?op=OPEN
```



Ellip Worflow using RGeostats

Invoke the workflow (1/2)



VM for RGeostats

WPS Client

Process List (GetCapabilities)

- Download and Estimation workflow

Reload

Last Jobs

- Download and Estimation workflow 2019-01-04T23:44:38.340Z success
- Download and Estimation workflow 2019-01-04T23:31:17.245Z success
- Download workflow 2019-01-03T16:23:09.182Z success
- Download workflow 2019-01-03T11:00:10.144Z failed
- Download workflow 2019-01-03T10:53:09.472Z failed

Download and Estimation workflow

Download data from IMR OpenDAP server and Estimate

IMR Dataset Server Catalog

http://opendap1-test.nodc.no/thredds/catalogs/physics/physics_p

Date limits

2001/04/01,2001/06/30

Longitude limits (deg)

-2,10

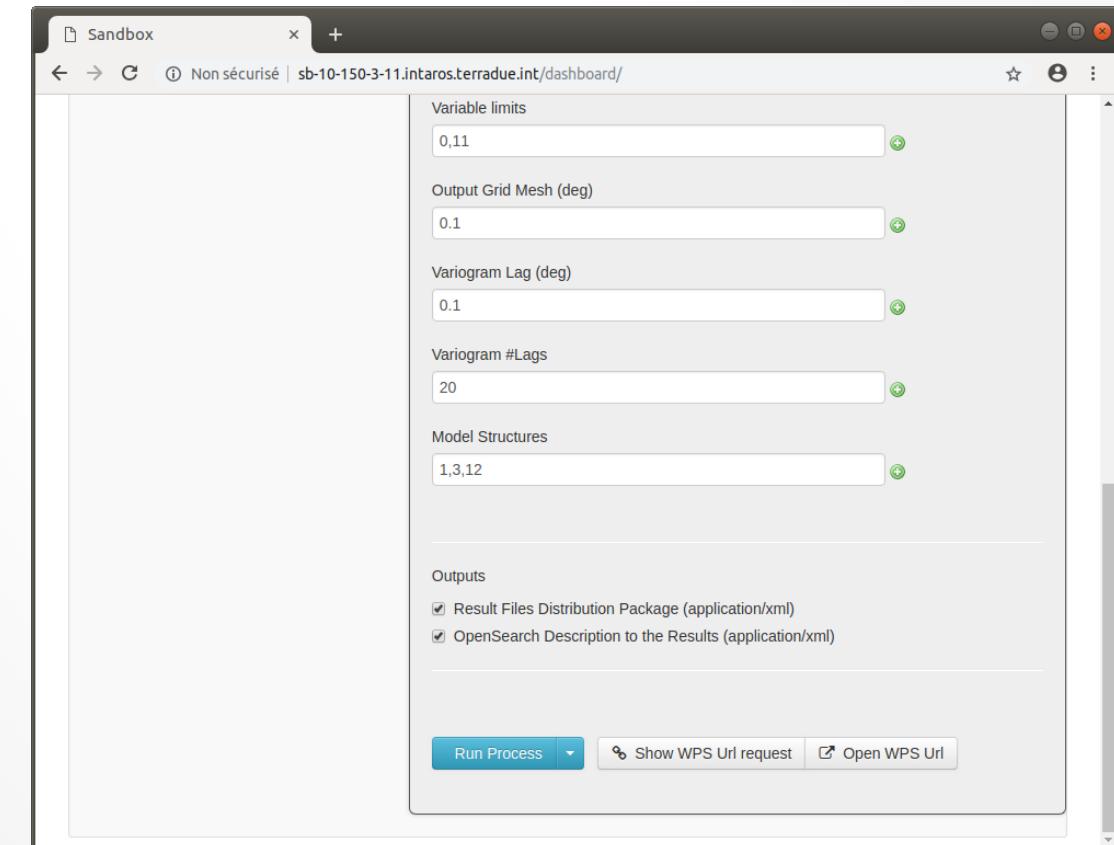
Latitude limits (deg)

56,62

Depth limits (m)

18,22

Variable Name



Variable limits
0,11

Output Grid Mesh (deg)
0.1

Variogram Lag (deg)
0.1

Variogram #Lags
20

Model Structures
1,3,12

Outputs

Result Files Distribution Package (application/xml)

OpenSearch Description to the Results (application/xml)

Run Process

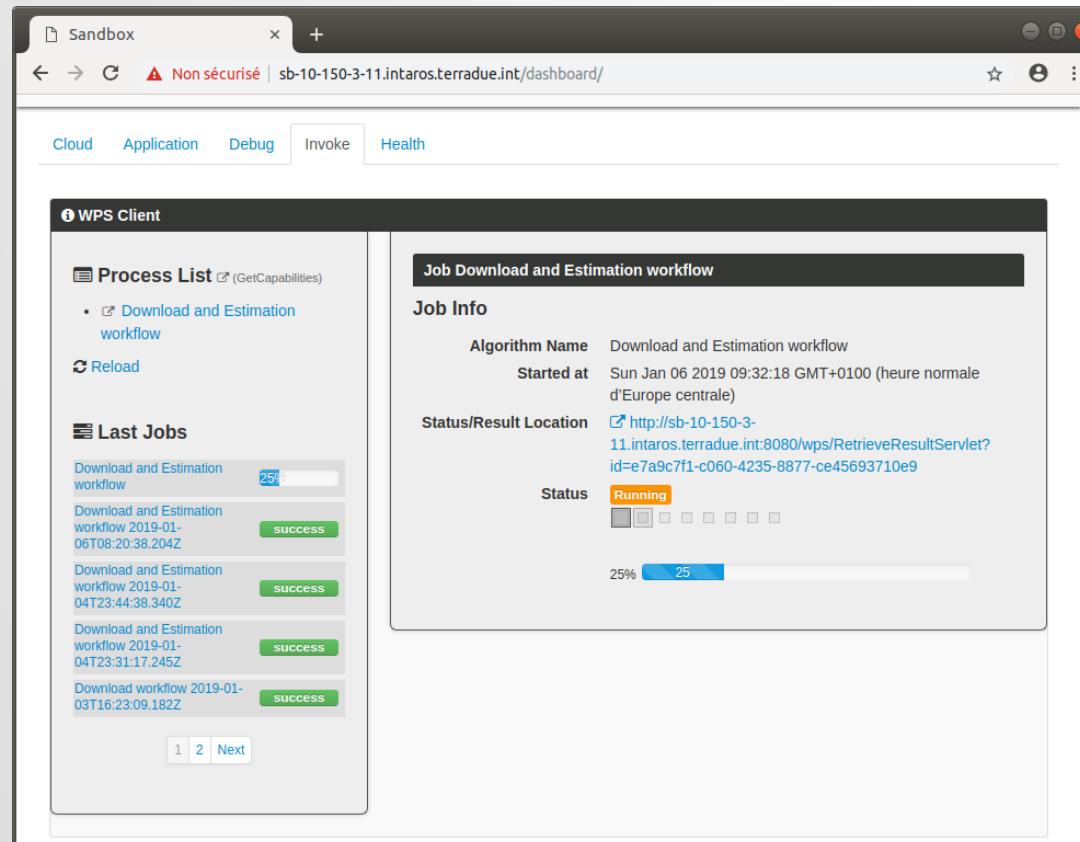
Show WPS Url request

Open WPS Url



Ellip Worflow using RGeostats

Invoke the workflow (2/2)



WPS Client

Process List (GetCapabilities)

- Download and Estimation workflow

Last Jobs

Workflow	Status
Download and Estimation workflow	25%
Download and Estimation workflow 2019-01-06T08:20:38.204Z	success
Download and Estimation workflow 2019-01-04T23:44:38.340Z	success
Download and Estimation workflow 2019-01-04T23:31:17.245Z	success
Download workflow 2019-01-03T16:23:09.182Z	success

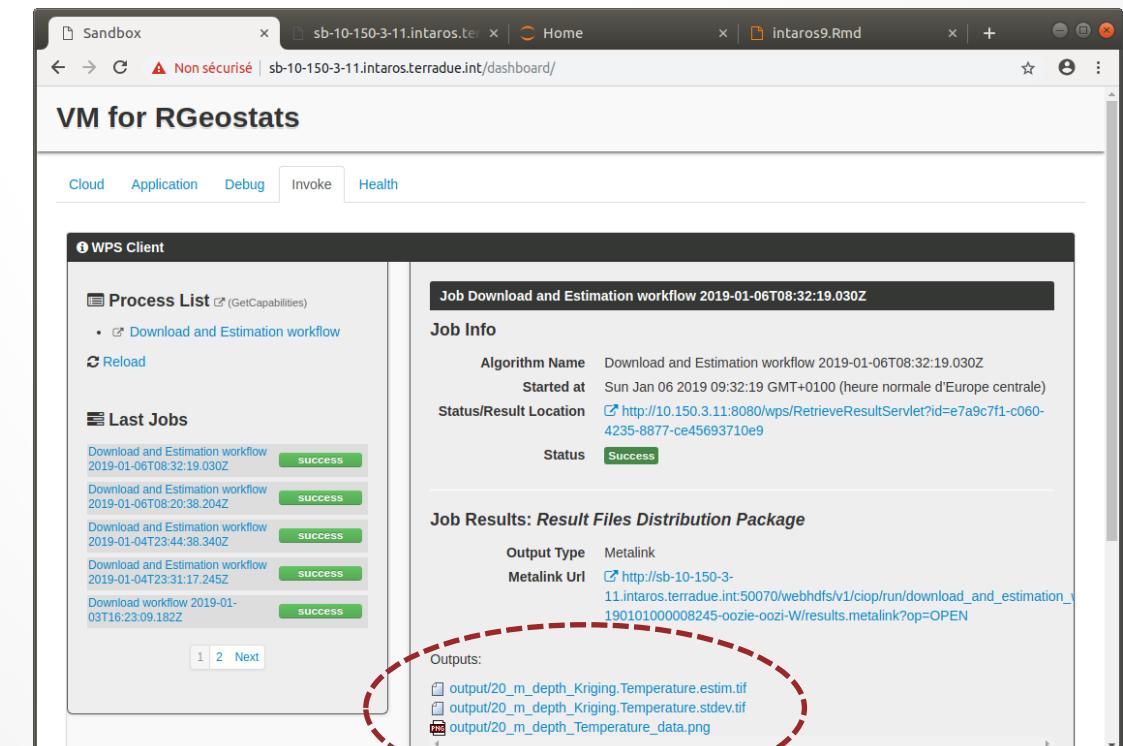
Job Info

Algorithm Name: Download and Estimation workflow
 Started at: Sun Jan 06 2019 09:32:18 GMT+0100 (heure normale d'Europe centrale)

Status/Result Location: <http://sb-10-150-3-11.intaros.terradue.int:8080/wps/RetrieveResultServlet?id=e7a9c7f1-c060-4235-8877-ce45693710e9>

Status: Running

Progress: 25%



VM for RGeostats

WPS Client

Process List (GetCapabilities)

- Download and Estimation workflow

Last Jobs

Workflow	Status
Download and Estimation workflow 2019-01-06T08:32:19.030Z	success
Download and Estimation workflow 2019-01-06T08:20:38.204Z	success
Download and Estimation workflow 2019-01-04T23:44:38.340Z	success
Download and Estimation workflow 2019-01-04T23:31:17.245Z	success
Download workflow 2019-01-03T16:23:09.182Z	success

Job Info

Algorithm Name: Download and Estimation workflow 2019-01-06T08:32:19.030Z
 Started at: Sun Jan 06 2019 09:32:19 GMT+0100 (heure normale d'Europe centrale)

Status/Result Location: <http://10.150.3.11:8080/wps/RetrieveResultServlet?id=e7a9c7f1-c060-4235-8877-ce45693710e9>

Status: Success

Job Results: Result Files Distribution Package

Output Type: Metalink
 Metalink Url: http://sb-10-150-3-11.intaros.terradue.int:50070/webhdfs/v1/ciop/run/download_and_estimation_19010100008245-oozie-oozi-W/results.metalink?op=OPEN

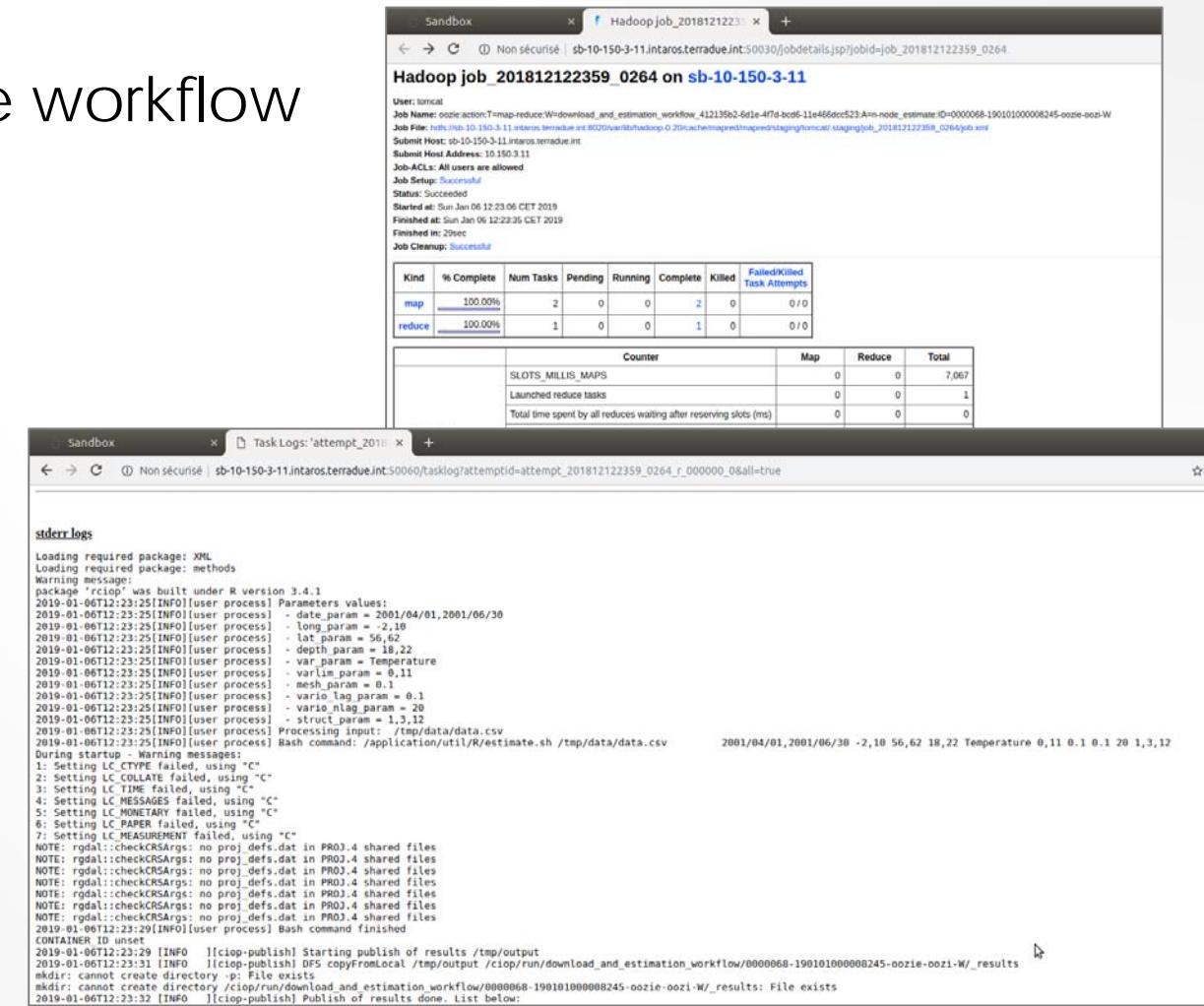
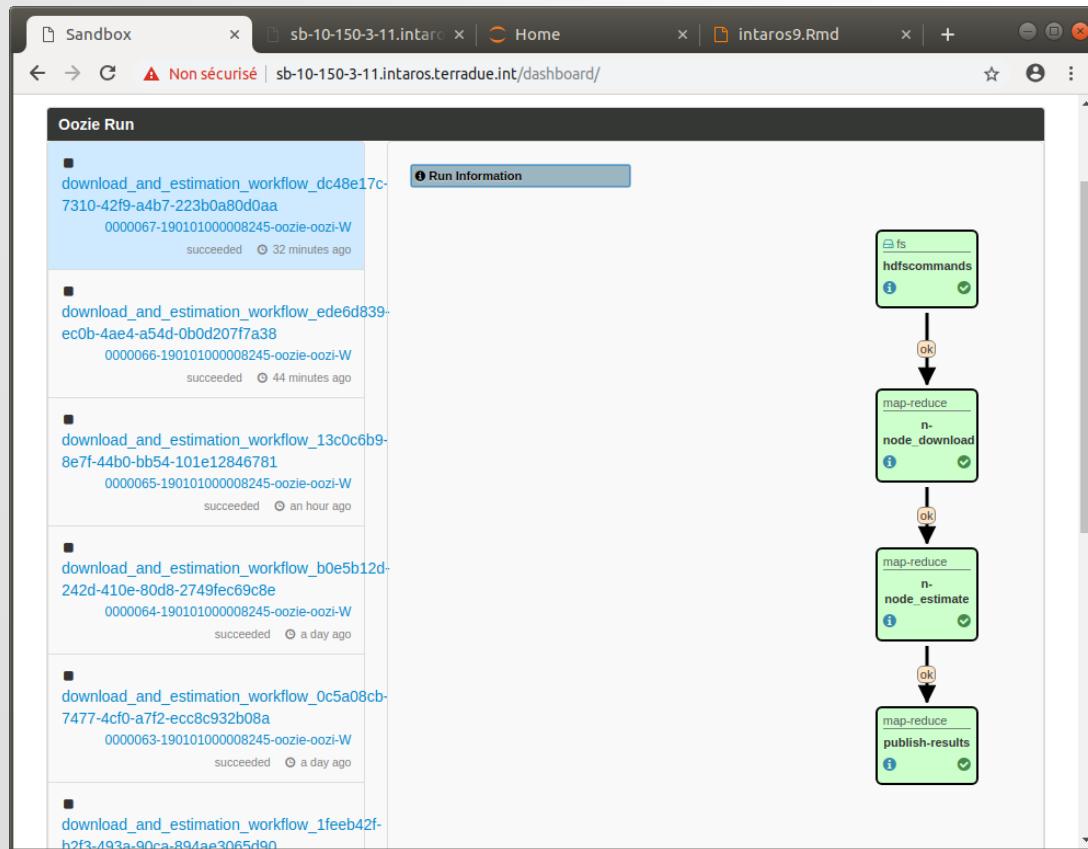
Outputs:

- output/20_m_depth_Kriging.Temperature.estim.tif
- output/20_m_depth_Kriging.Temperature.stdev.tif
- output/20_m_depth_Temperature_data.png



Ellip Worflow using RGeostats

Debug the workflow



Ellip Worflow using RGeostats

Deploy the workflow

```
$ ci op- release

***  
*** Application release for 'dcs- imr- estim'  
***  
...  
...  
Writing objects: 100% (1/1), 159 bytes, done.  
Total 1 (delta 0), reused 0 (delta 0)  
To https://gitlab.com/ec-intaros/dcs-imr-estim.git  
 * [new tag] 1.0 -> 1.0  
***  
-> Stage 1. Checking your repository 'dcs- imr- estim' DONE  
-> Stage 2. Setting the release and development versions DONE  
-> Stage 3. Releasing 'dcs- imr- estim' DONE  
-> Stage 4. Syncronising the remote repository DONE
```



Free text search 

Ellip Worflow using RGeostats

Useful commands (1/3)

Connect to your VM

```
sudo openvpn --config client.ovpn  
ssh 10.150.3.11
```

=> Connect to VPN
=> Connect to your VM with
no password (ssh key pairs)

Creating and installing application

```
mvn archetype:generate  
mvn clean install
```

=> Create a new workflow
=> Install the application



Ellip Worflow using RGeostats

Useful commands (2/3)

Local git repository

```
git init
git config credential.helper store
git add -A .
git commit -m <msg>
git branch <branch>
git checkout <branch>
```

=> Initialize a local repository
=> Store credentials (no more password)
=> Add content to a local repository
=> Commit changes to the local repo
=> Create a branch named <branch>
=> Work on the branch named <branch>

Remote git repository interaction

```
git remote add origin <url>
git push
git pull
git clone
```

=> Synchronize with remote repository
=> Push local changes to server
=> Update local repo from the server
=> Create a local repo from the server



Ellip Worflow using RGeostats

Useful commands (3/3)

Ciop tool

ciop-getparam
ciop-log
ciop-copy
ciop-publish
ciop-run
ciop-release

- => Retrieve job's parameter value (user input)
- => Log something to job log files
- => Copy online source file to local file system
- => Publish job productions as output results
- => Execute a job or a complete workflow
- => Create and deploy a new application release



Part #5

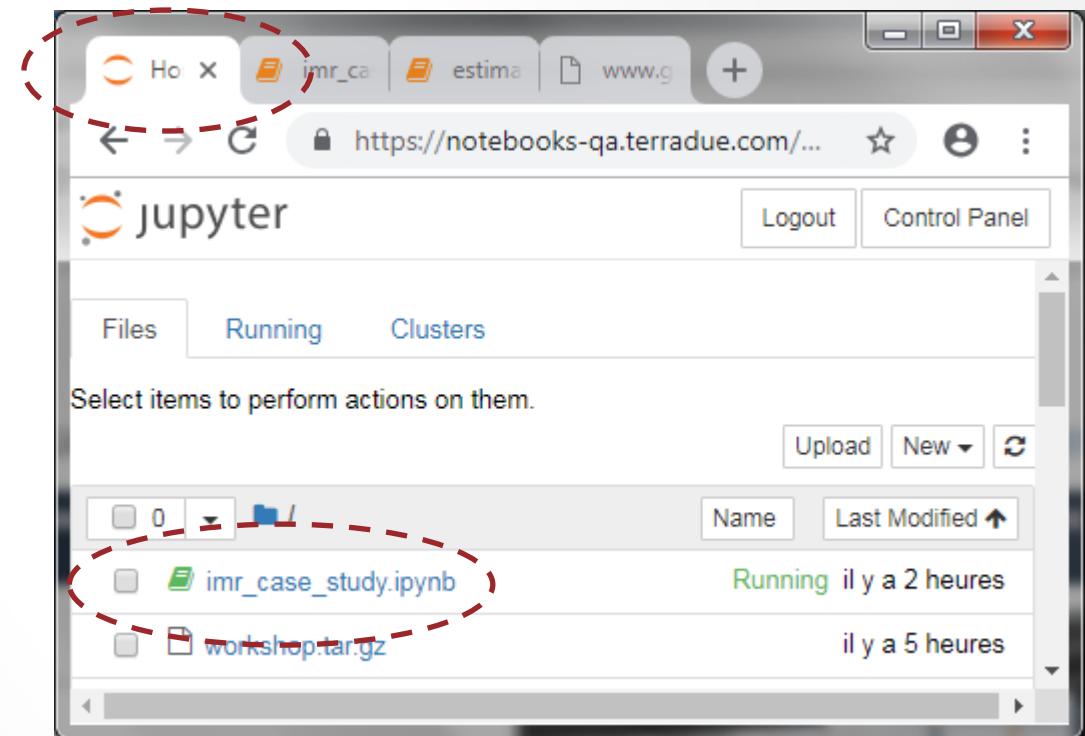
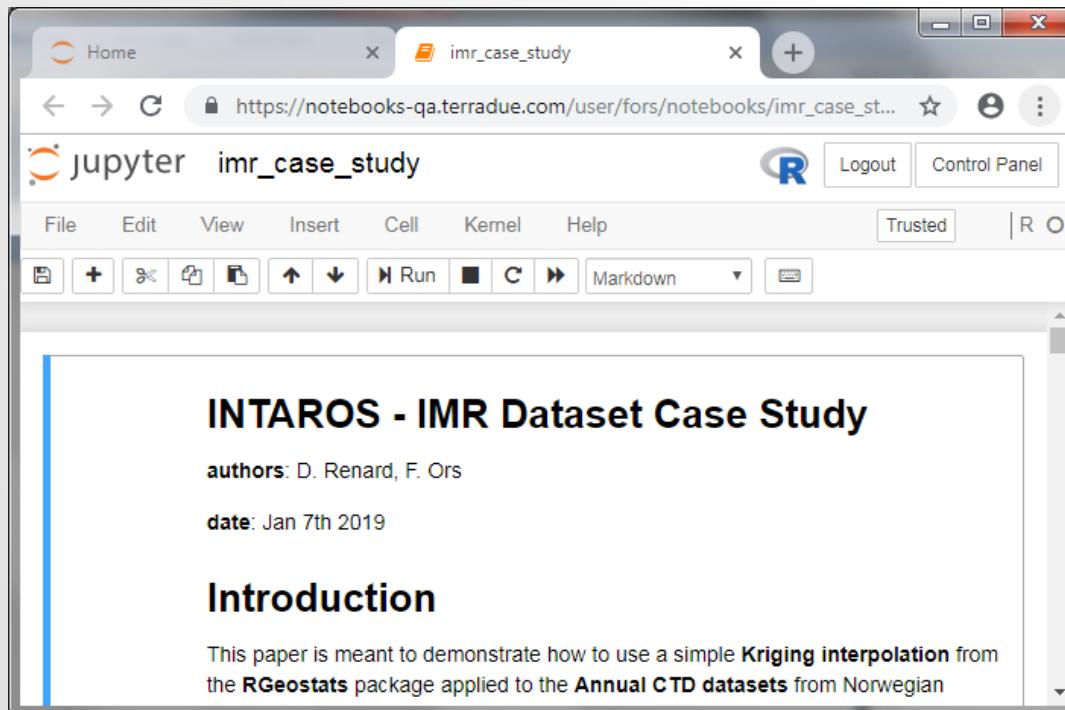
IMR Case Study – RGeostats in Action!



49

Getting Ready for Action!

- Click on the **Home** tab of your Jupyter Notebooks (in Chrome)
- Click on the **imr_case_study.ipynb**



Note: You may prefer the 'lab view': replace **tree** by **lab** in URL

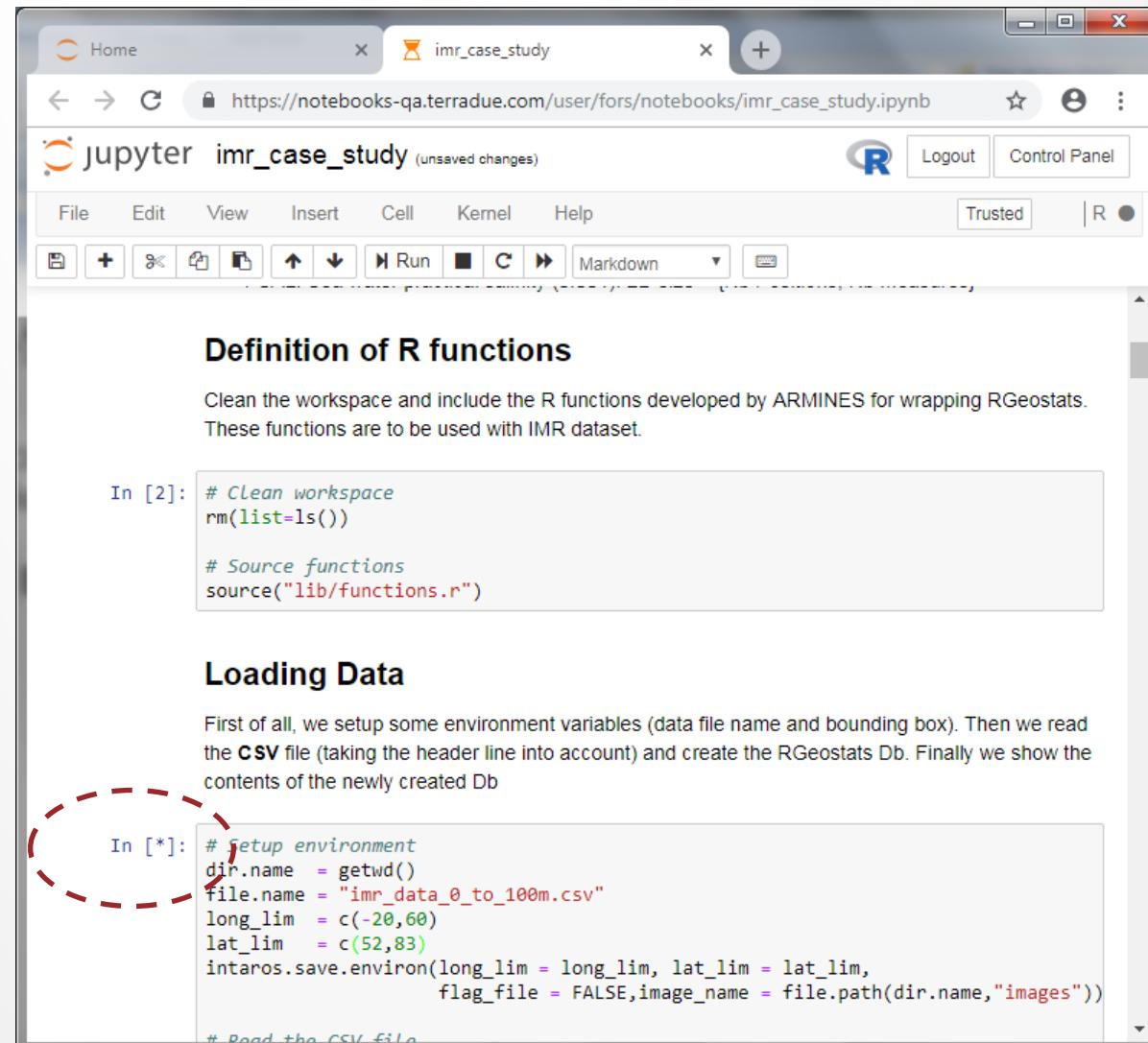
Getting Ready for Action!

- Execute all code cells in the **Introduction** section:

=> Hitting **Shift + Enter**

- Definition of R functions
- Loading Data (slow)
- Dataset global statistics (2 cells)

Note: The symbol **In [*]** indicates that a cell is running



In [2]:

```
# Clean workspace
rm(list=ls())

# Source functions
source("lib/functions.r")
```

In [*]:

```
# setup environment
dir.name  = getwd()
file.name = "imr_data_0_to_100m.csv"
long_lim  = c(-20,60)
lat_lim   = c(52,83)
intaros.save.environ(long_lim = long_lim, lat_lim = lat_lim,
                     flag_file = FALSE,image_name = file.path(dir.name,"images"))

# Read the CSV file
```

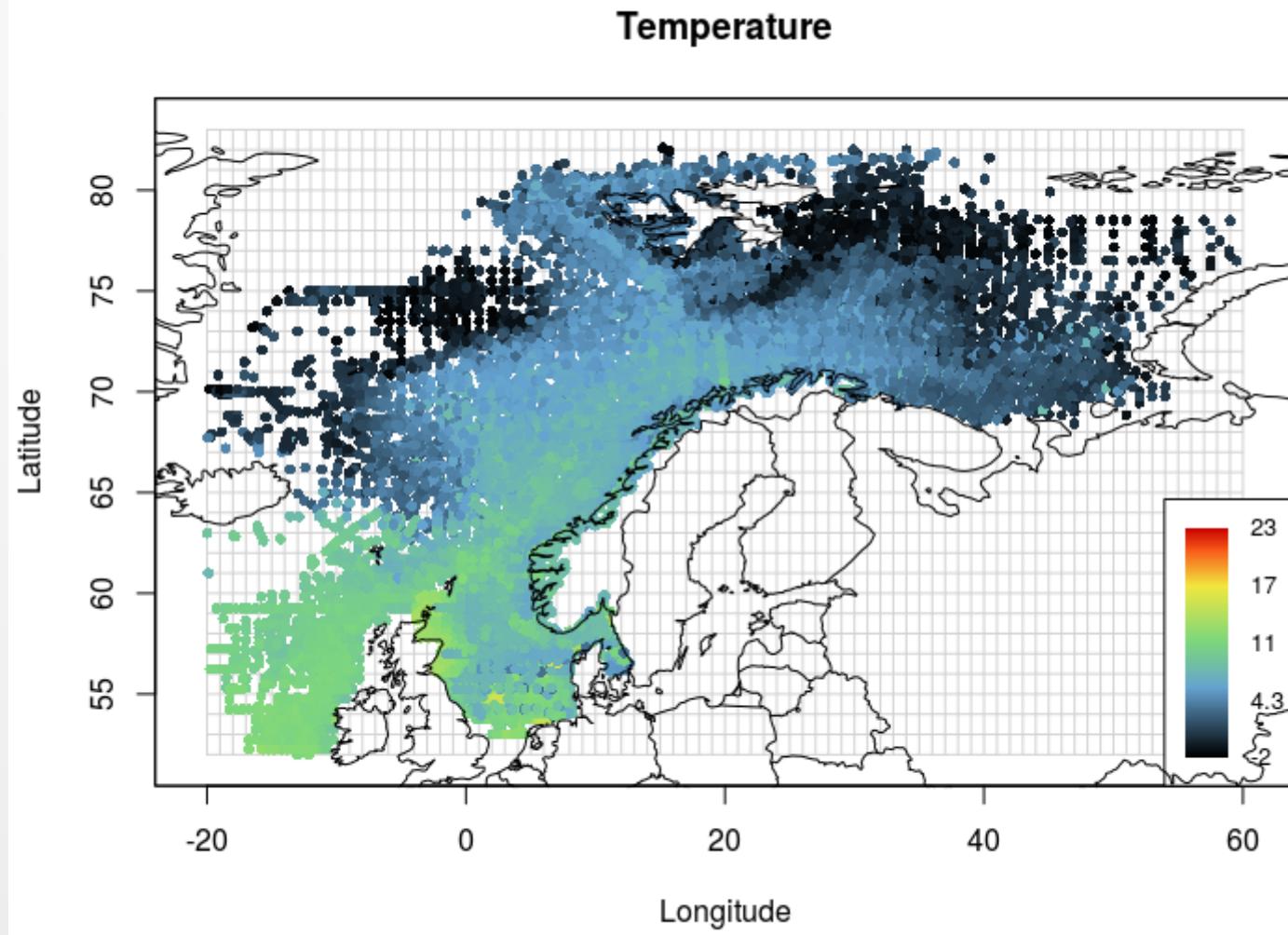


Studying Temperature variable



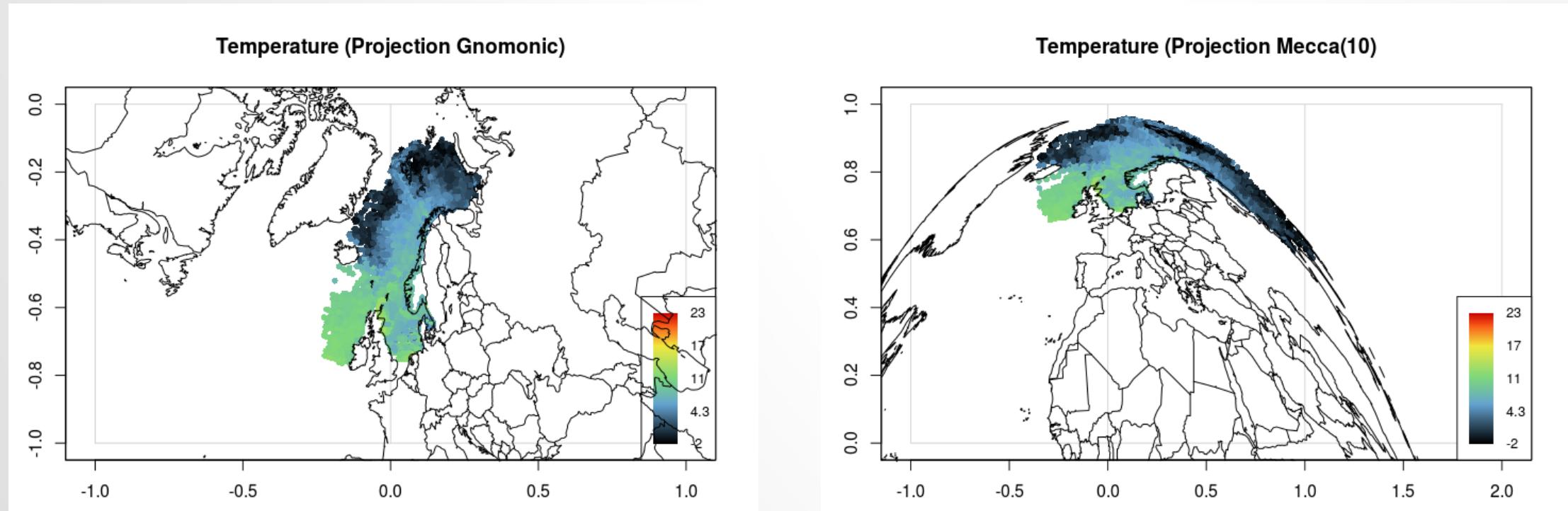
All Database

- Display Temperature Variable
- 5 million samples displayed



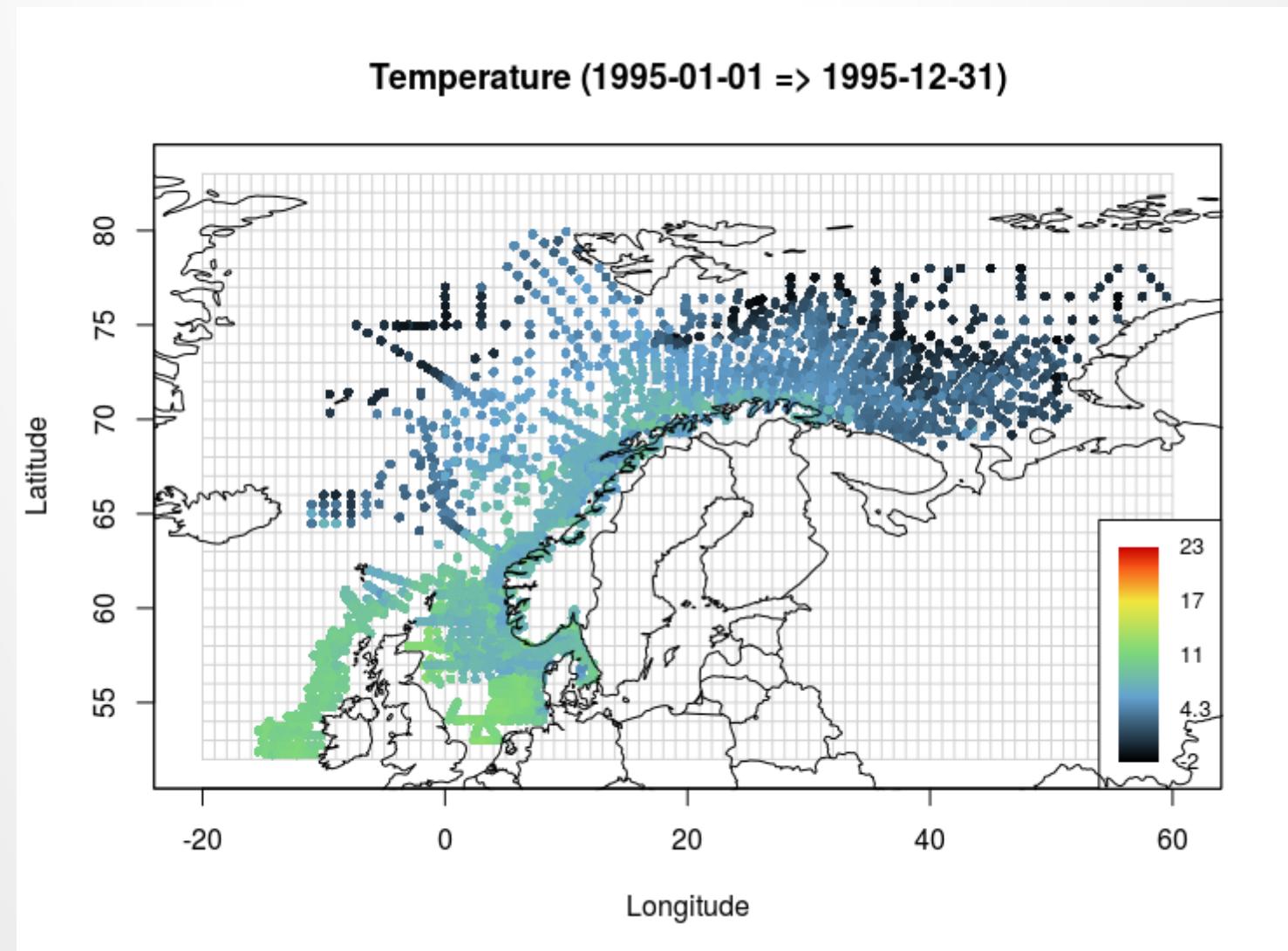
Different Projections

- Define New Projections (Gnomonic and Mecca)



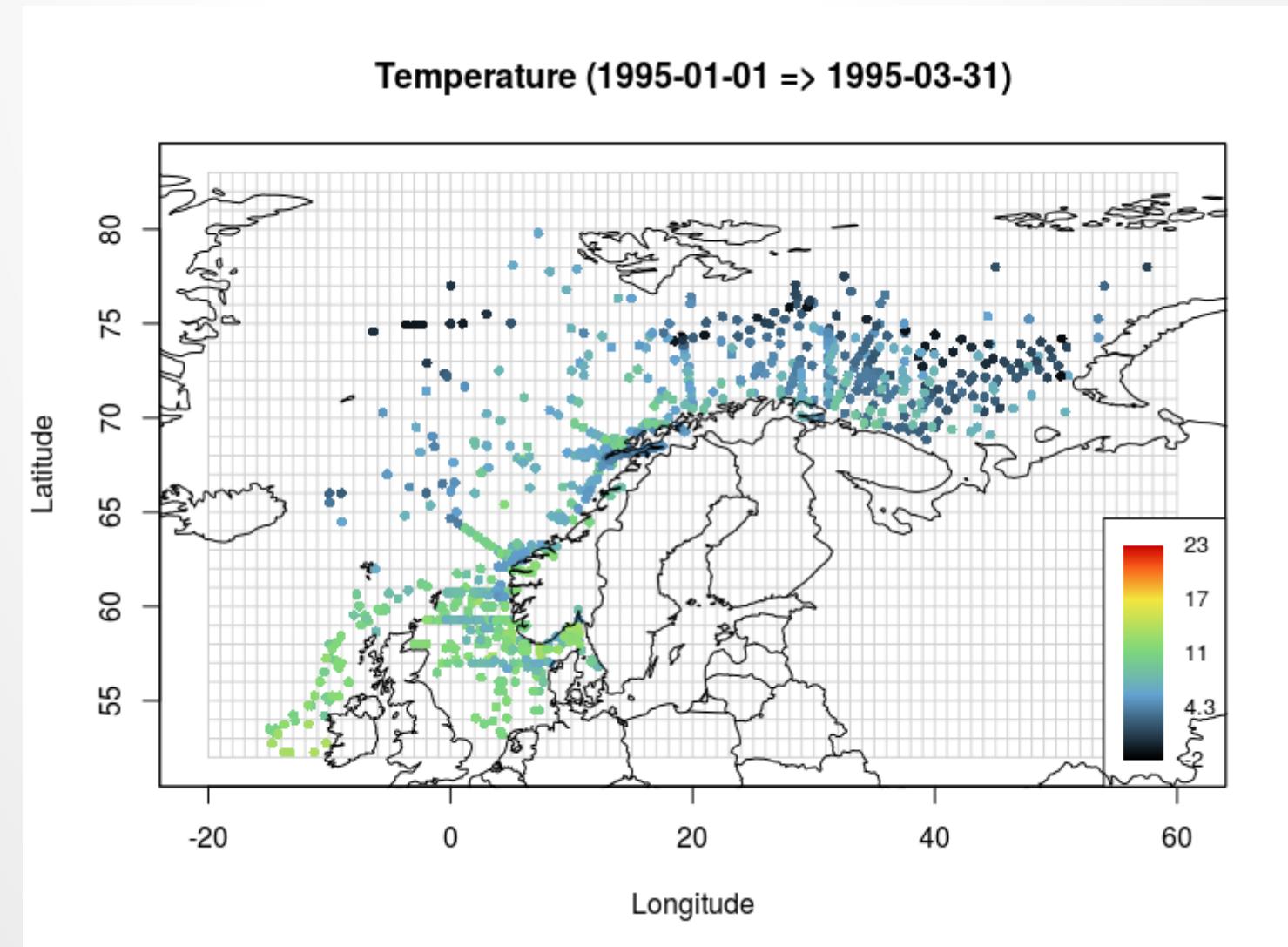
Year Campaign

- Define Target Year (1995)
- Display Variable



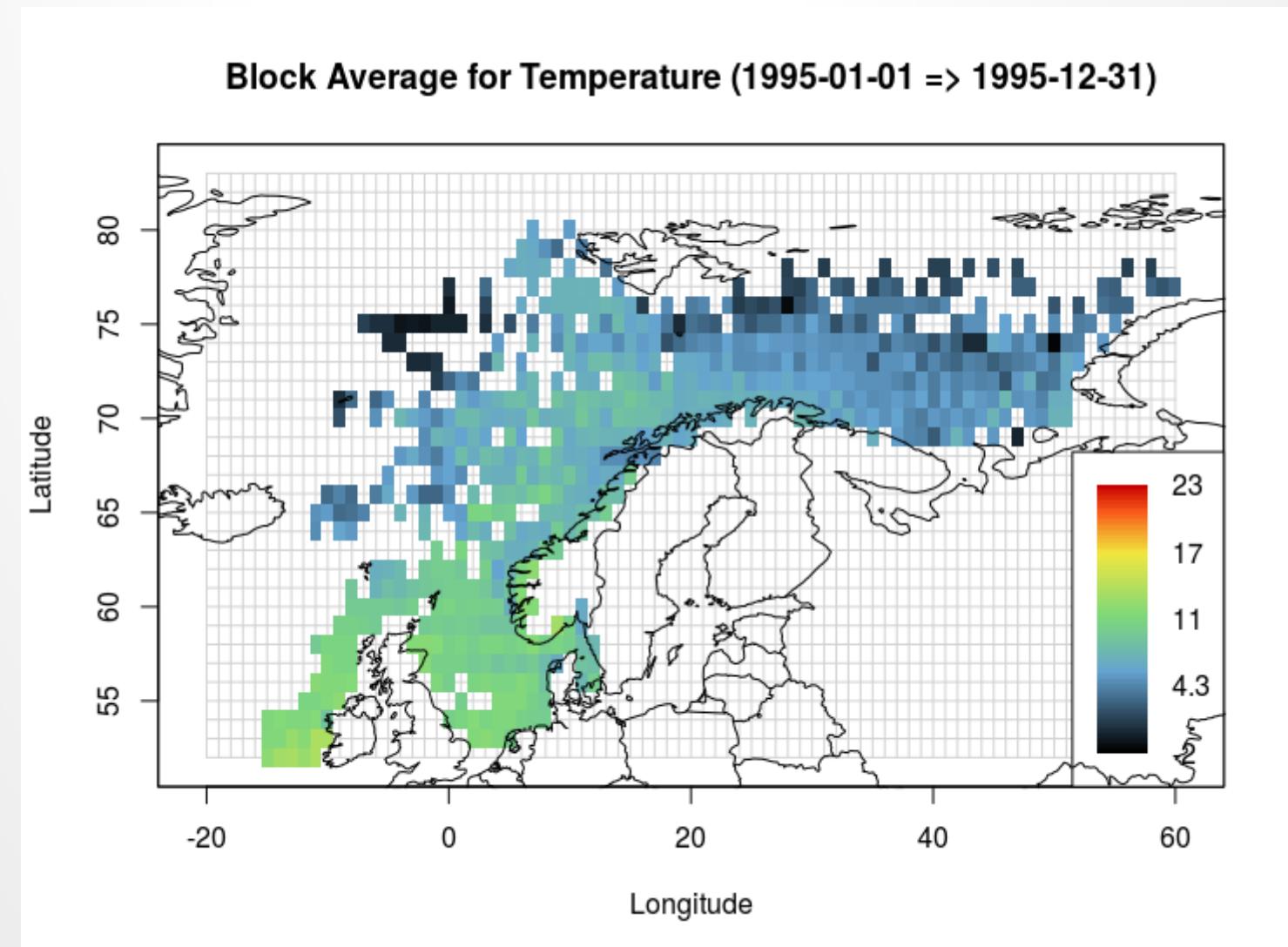
Year/Trimester at 20m Depth

- Select Depth Interval (20m)
- Define Target Year (1995)
- Define Target Trimester (1st)
- Display Variable



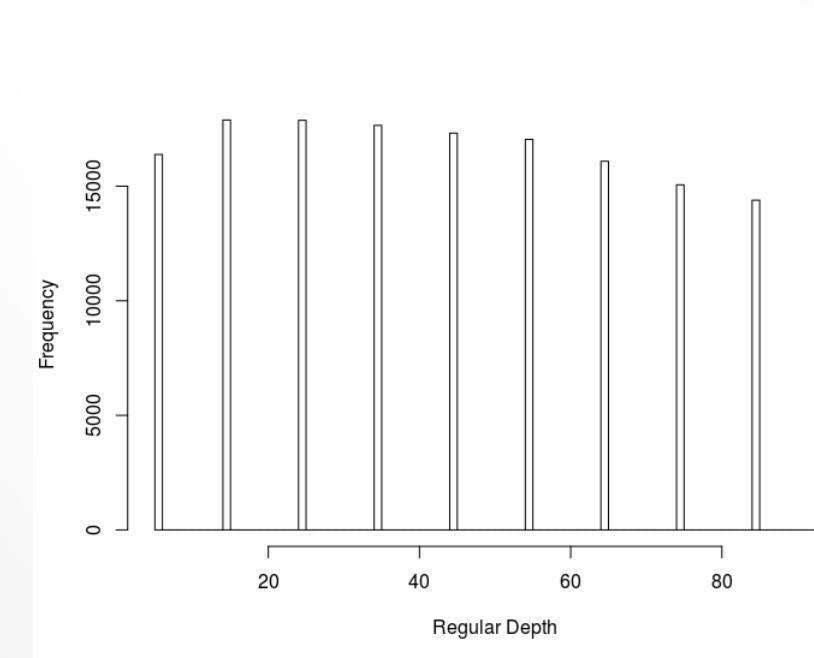
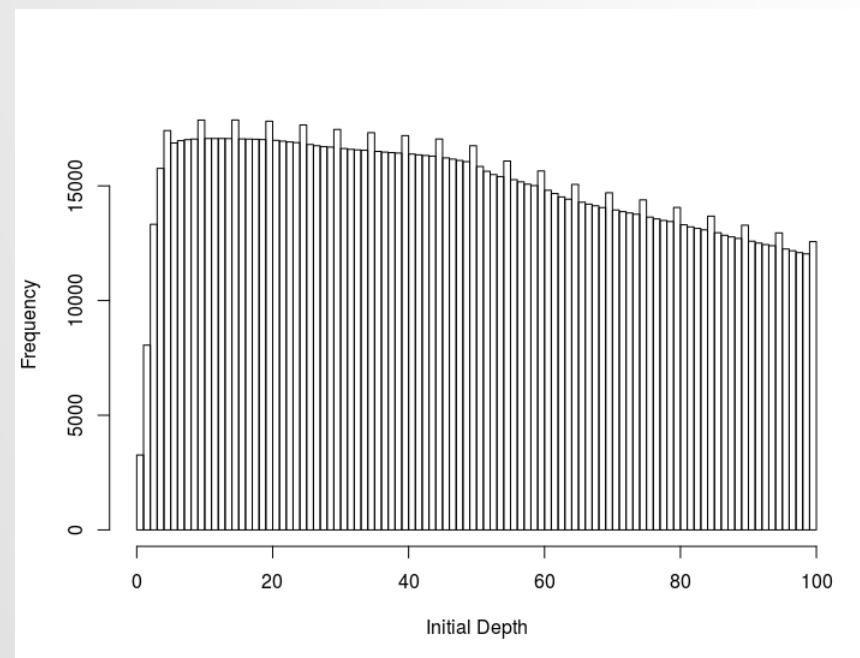
Block Average at 20m Depth

- Select Depth Interval (20m)
- Define Target Year (1995)
- Average per Cell (1 degree)
- Display Variable



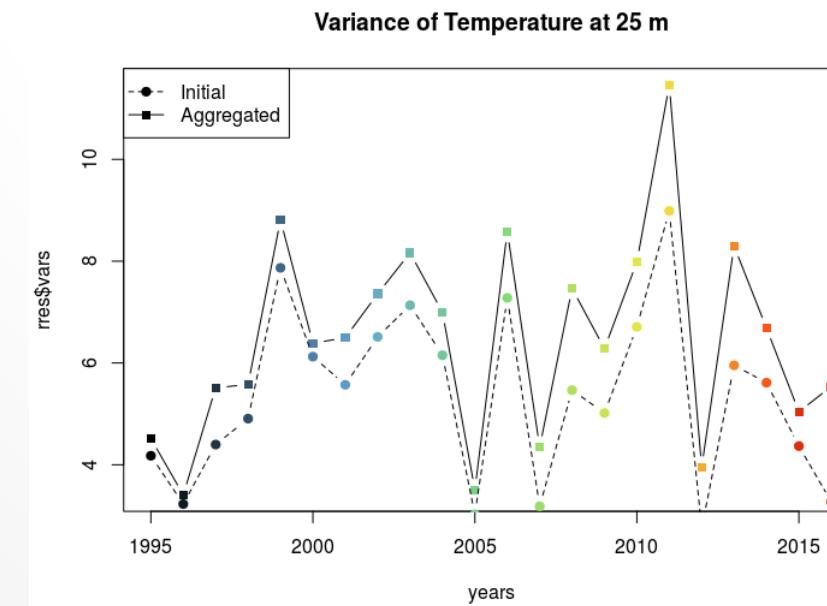
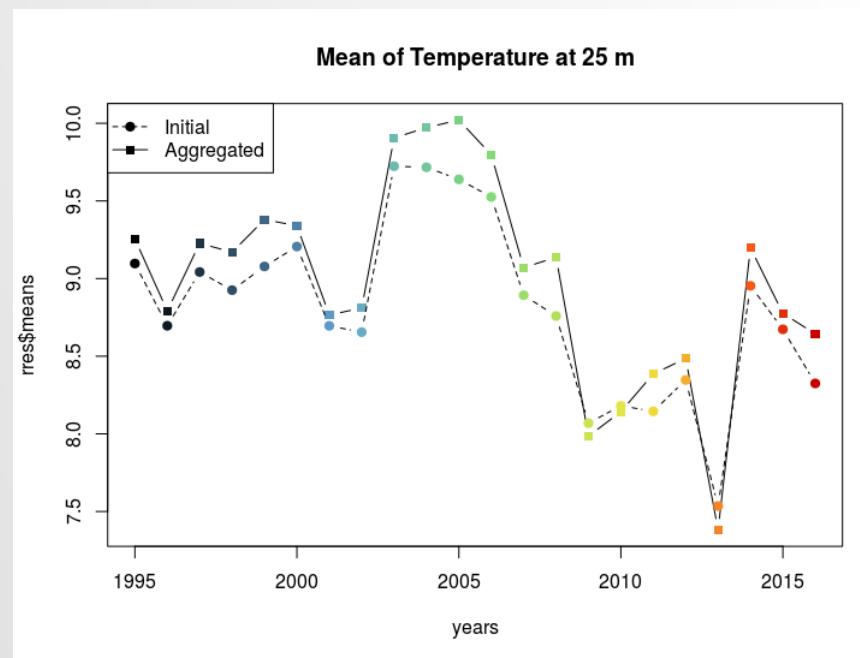
Histogram of Measurement Depths

- Focus on a Smaller Area (South West of Norway)
- Aggregate per Depth Intervals (10m)
- Histogram of Depths (before and after Aggregation)



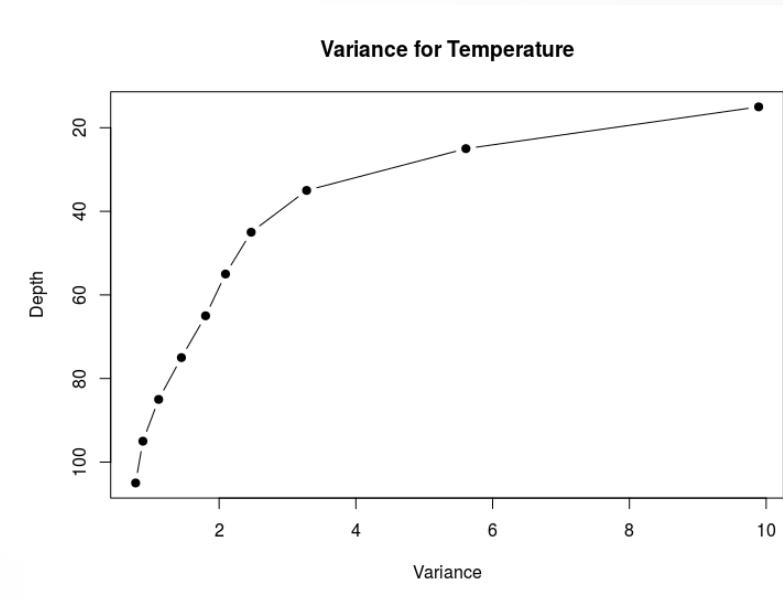
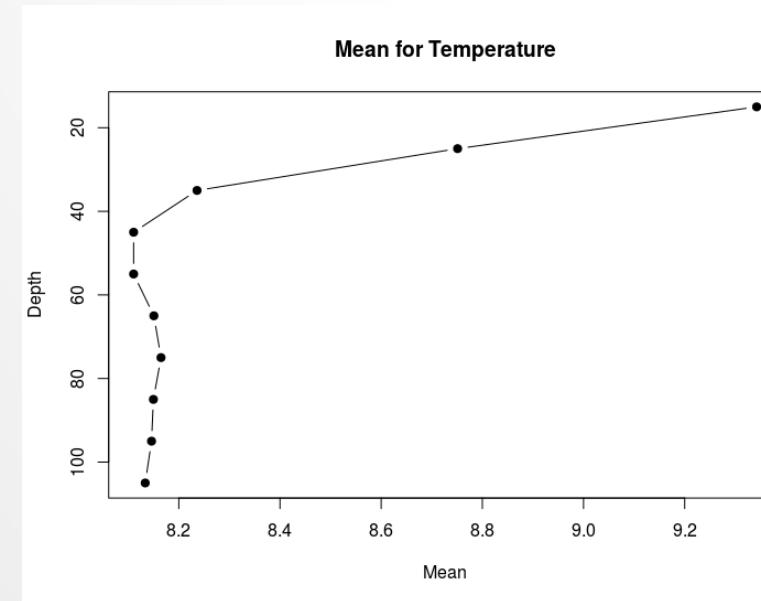
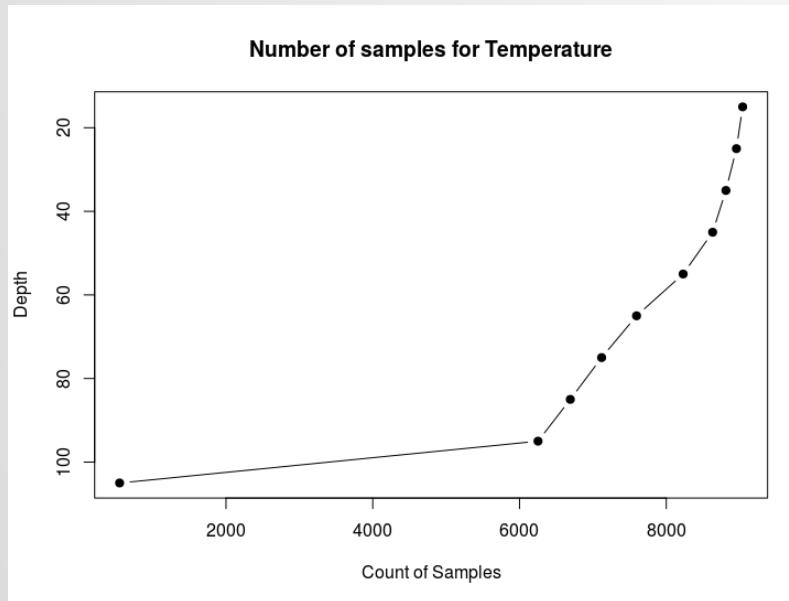
Statistics per Year

- Select Depth Interval (25m)
- Aggregate per Time Intervals (1 year)
- Mean and Variance of Temperature (before and after Aggregation)



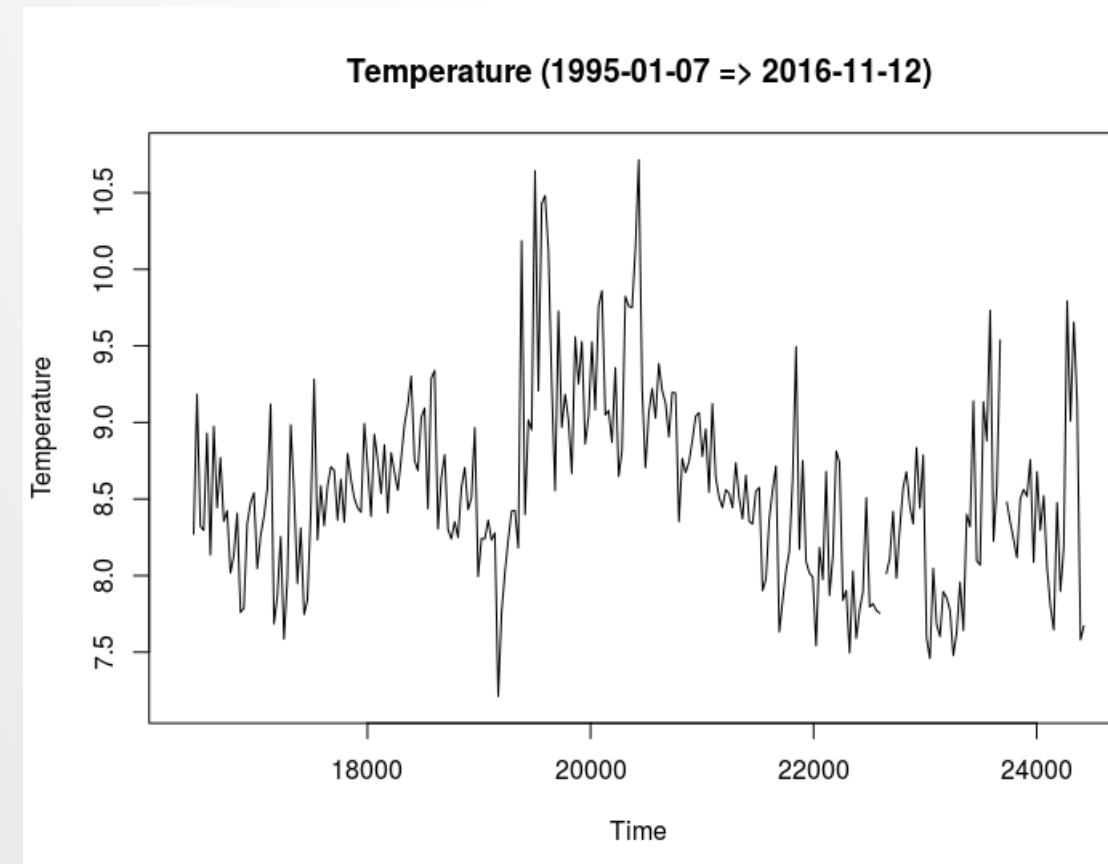
Statistics per Depth

- Select Year (2008)
- Average per Depth Interval (10m)
- Count, Mean and Variance of Temperature per Depth



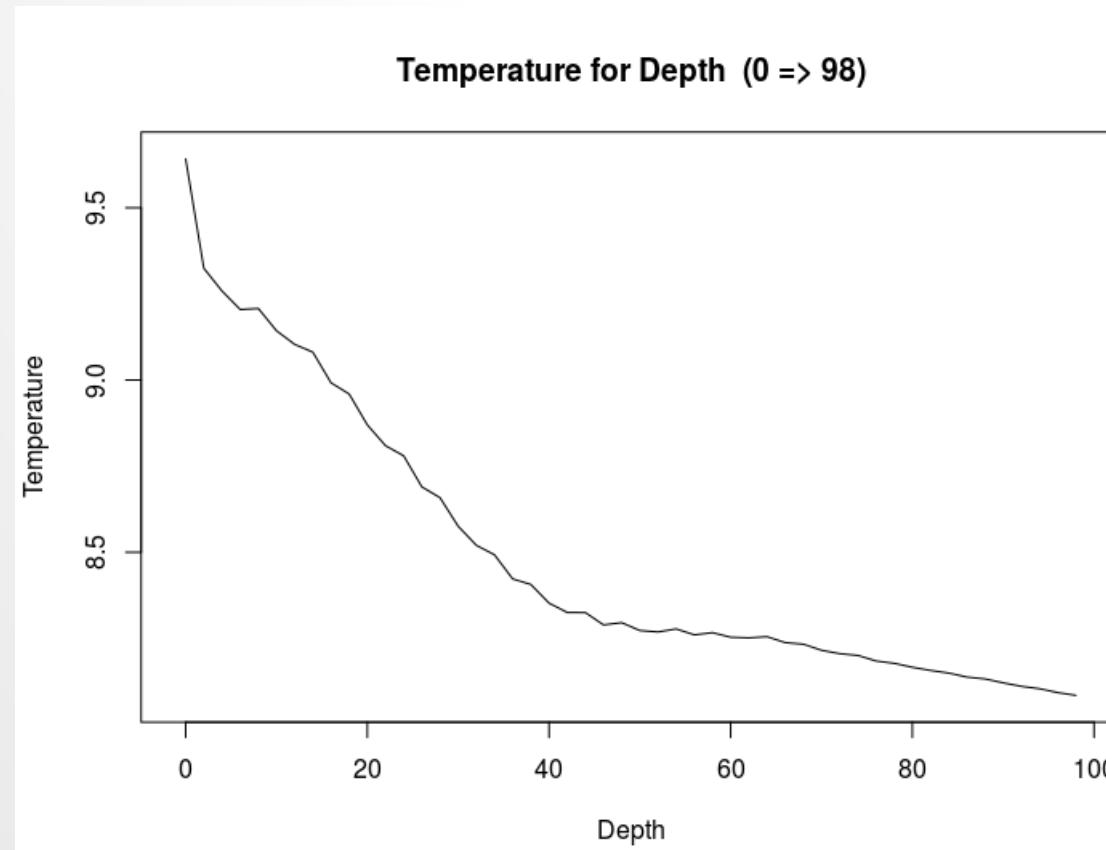
Variation along Time

- Aggregate by Time Intervals (30 days)



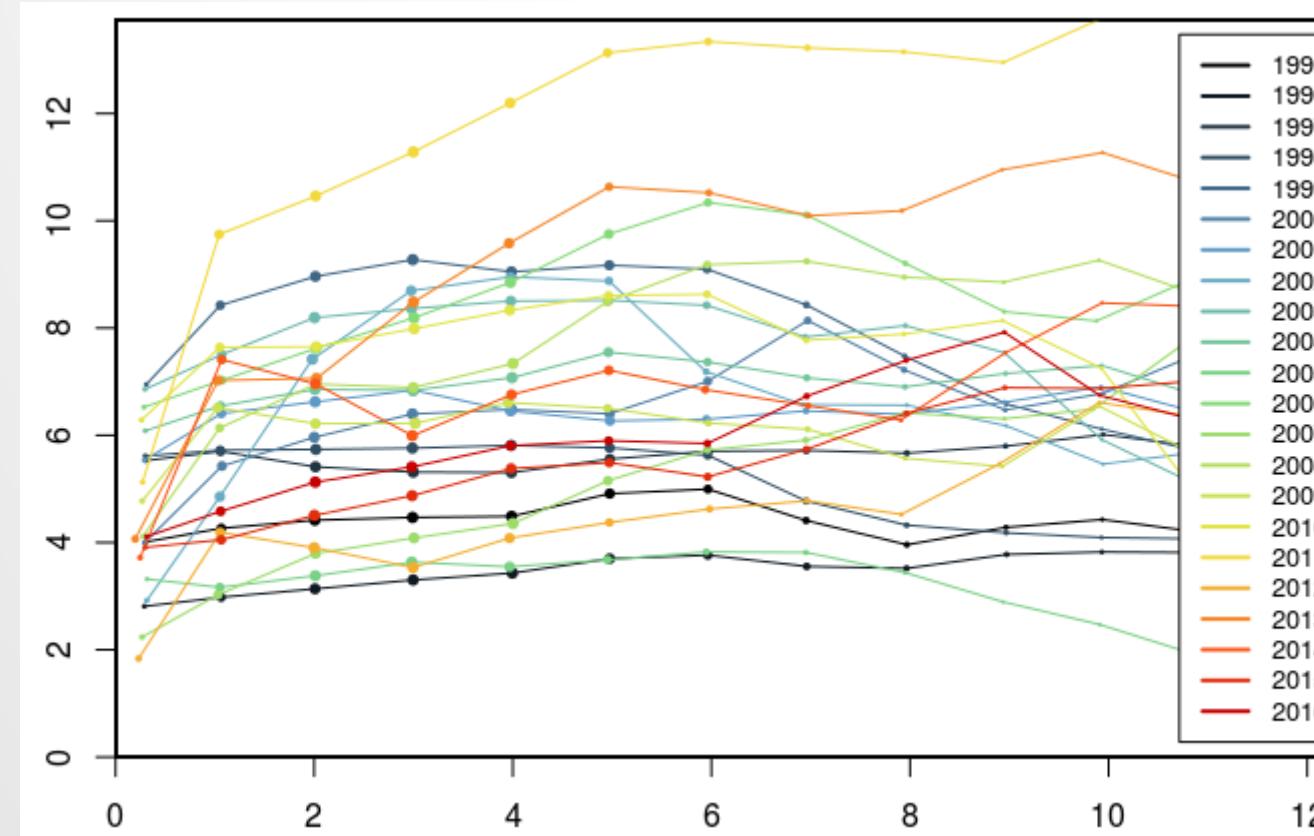
Variation along Depth

- Aggregate by Depth Intervals (2m)



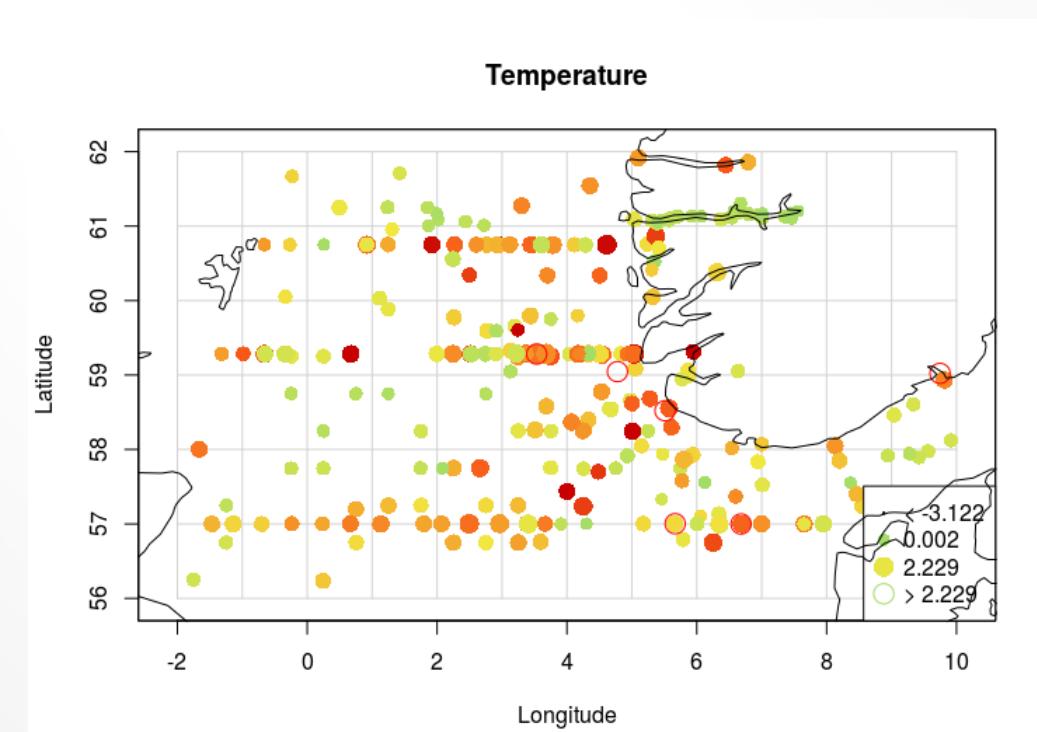
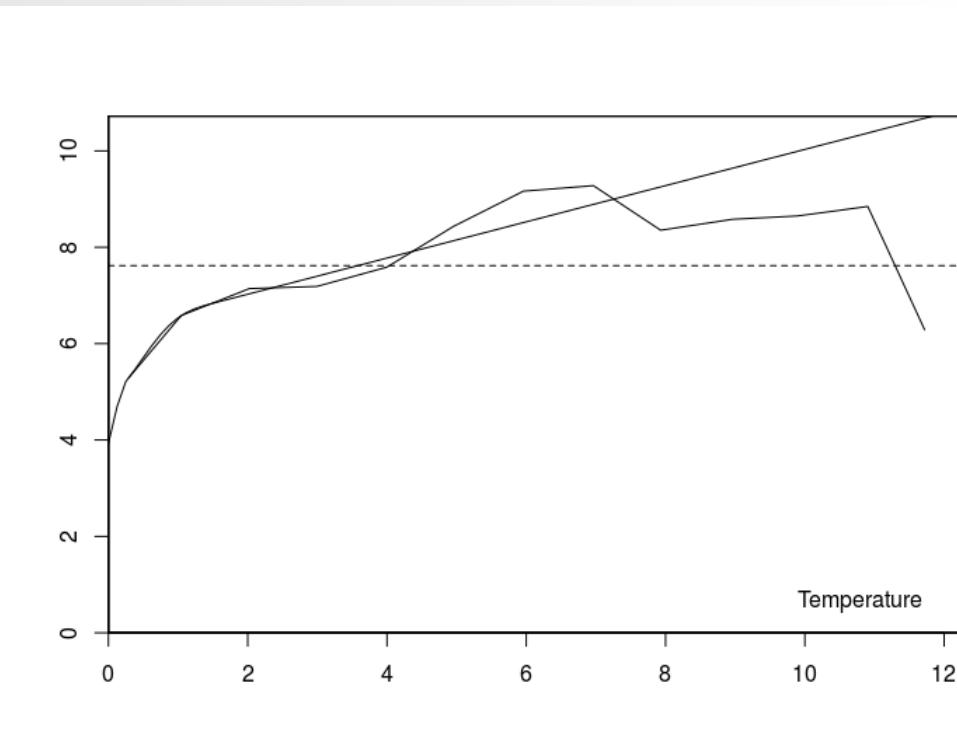
Horizontal Variogram per Year

- Select Depth Interval (25m)
- Omni-directional Variogram per Year for Temperature



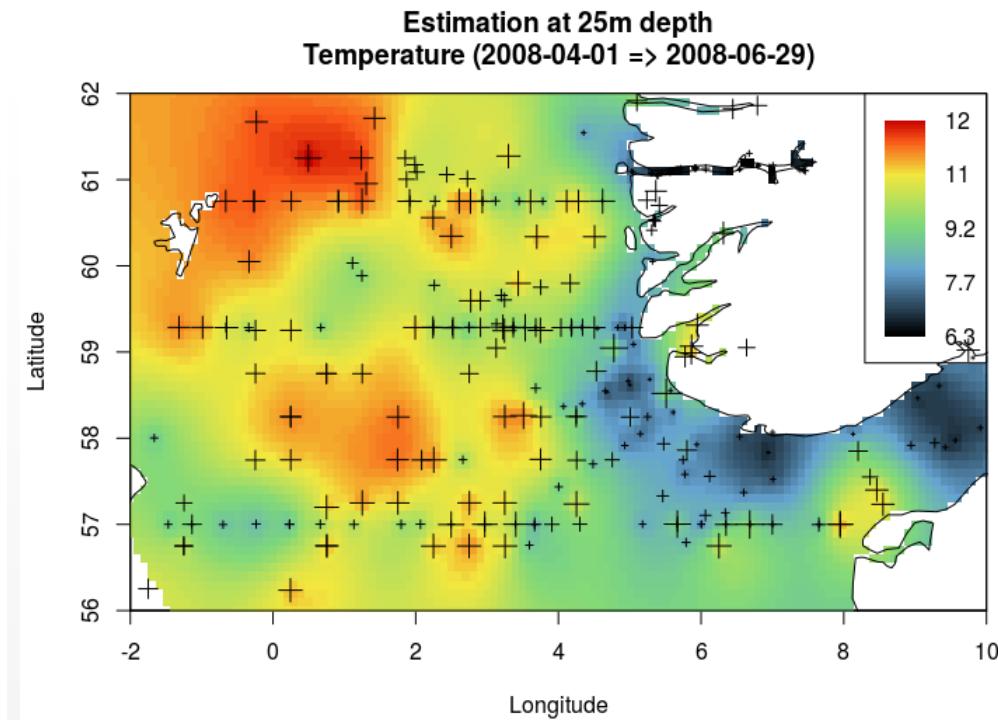
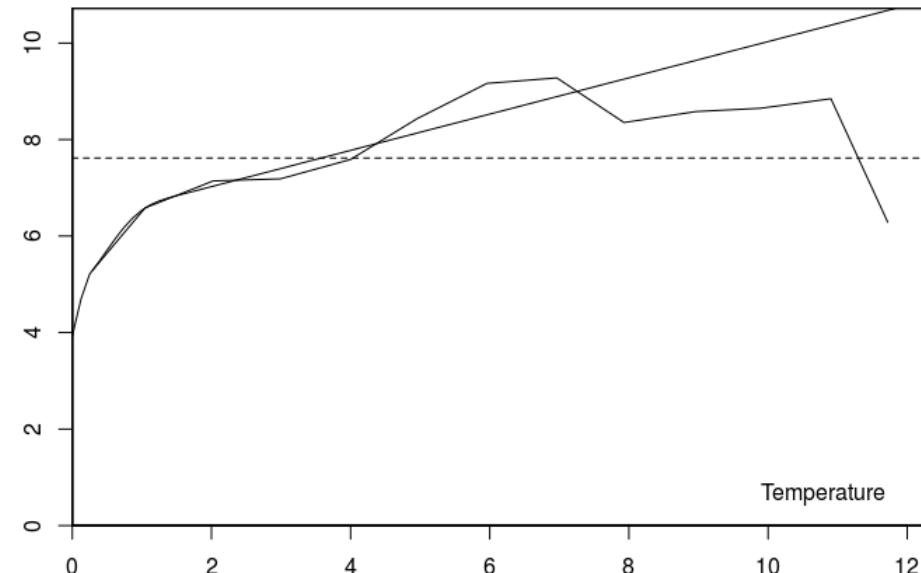
Cross-Validation

- From now: use Aggregated database per Depth Intervals (10m)
- Calculate Average Horizontal Variogram and Fit the Model of Temperature
- Perform the Cross-Validation at Data Samples



2-D Estimation of Temperature

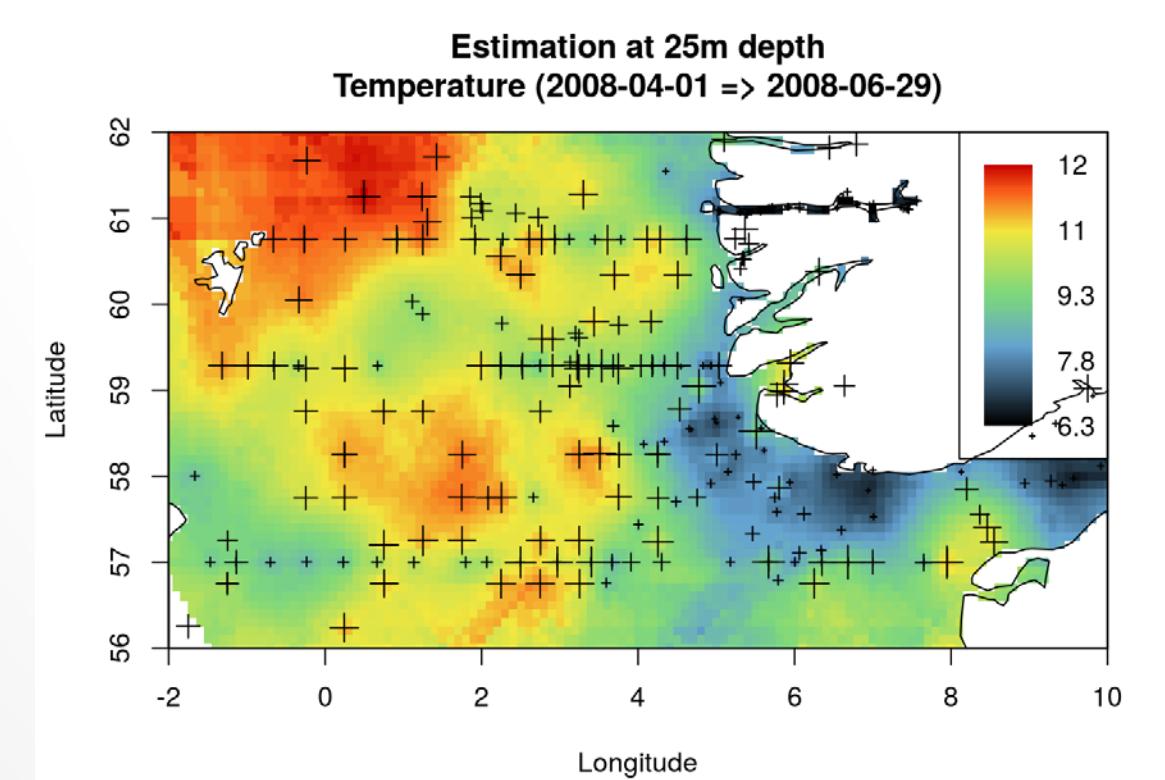
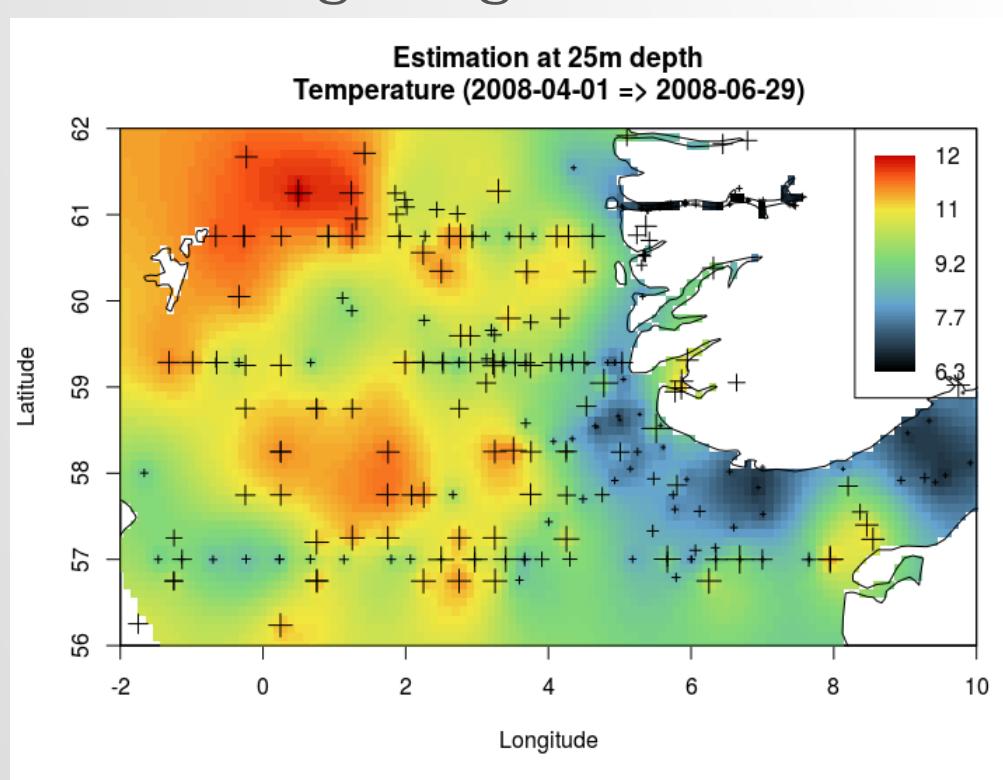
- Select Depth Interval (25m) and Focus on 2nd Trimester of Year 2008
- Calculate Average Horizontal Variogram and Fit the Model for Temperature
- Estimate the Temperature in 2-D in South-West of Norway



Moving Neighborhood

Comparing Unique (left) and Moving (right) Neighborhoods

- 2-D Moving Neighborhood: Maximum Number of Samples = 40

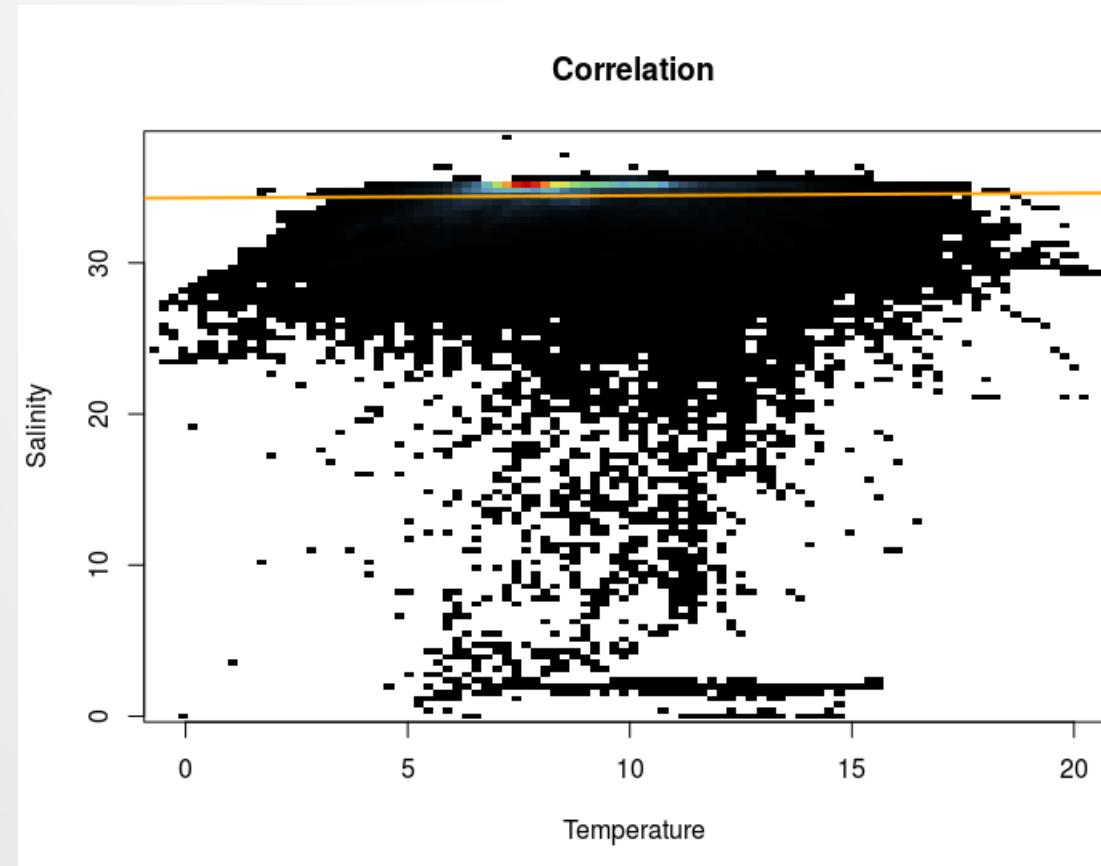


Bivariate Approach: Temperature & Salinity



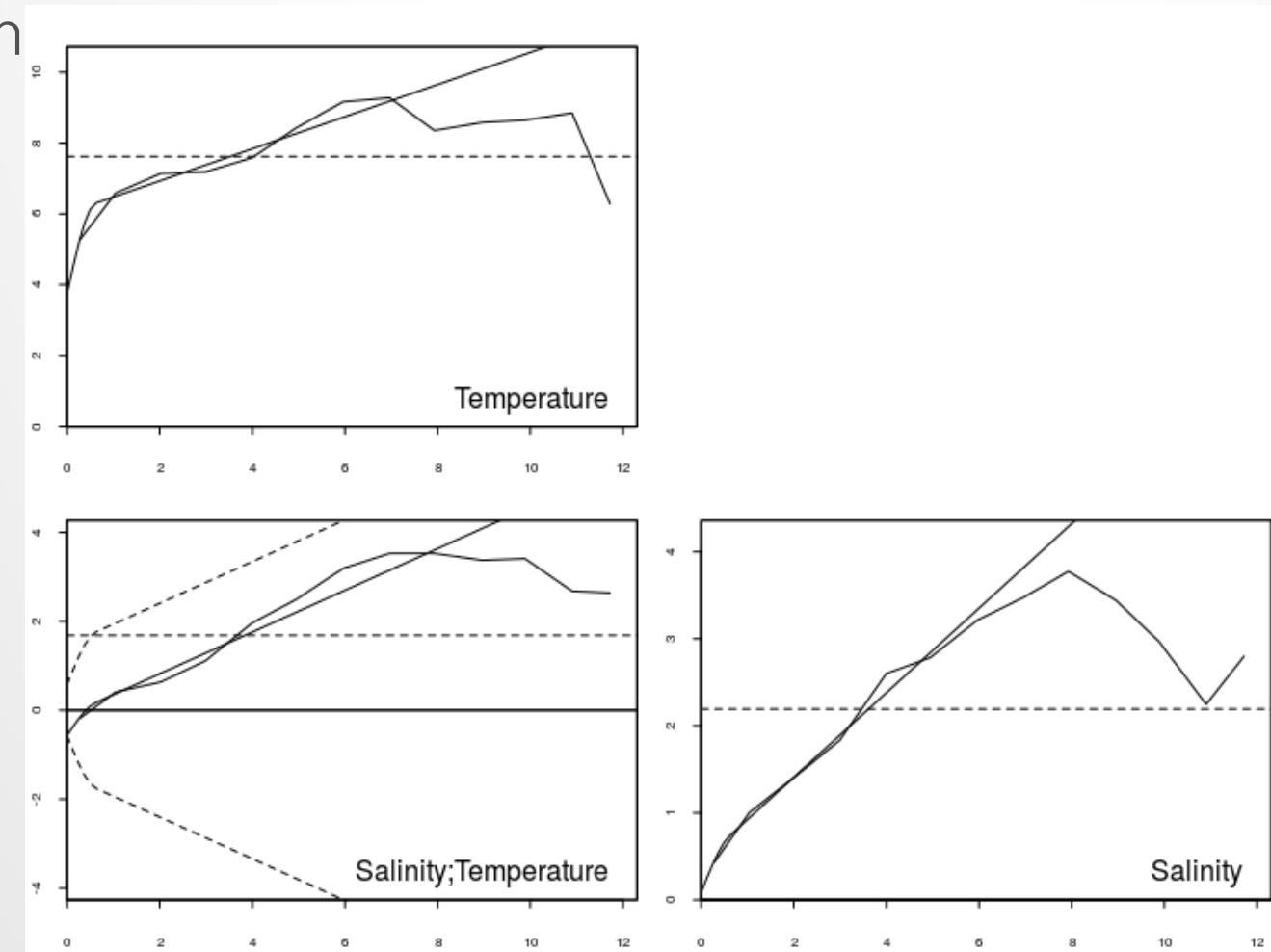
Temperature vs. Salinity

- Define New Variables of Interest
- Correlation and Linear Regression



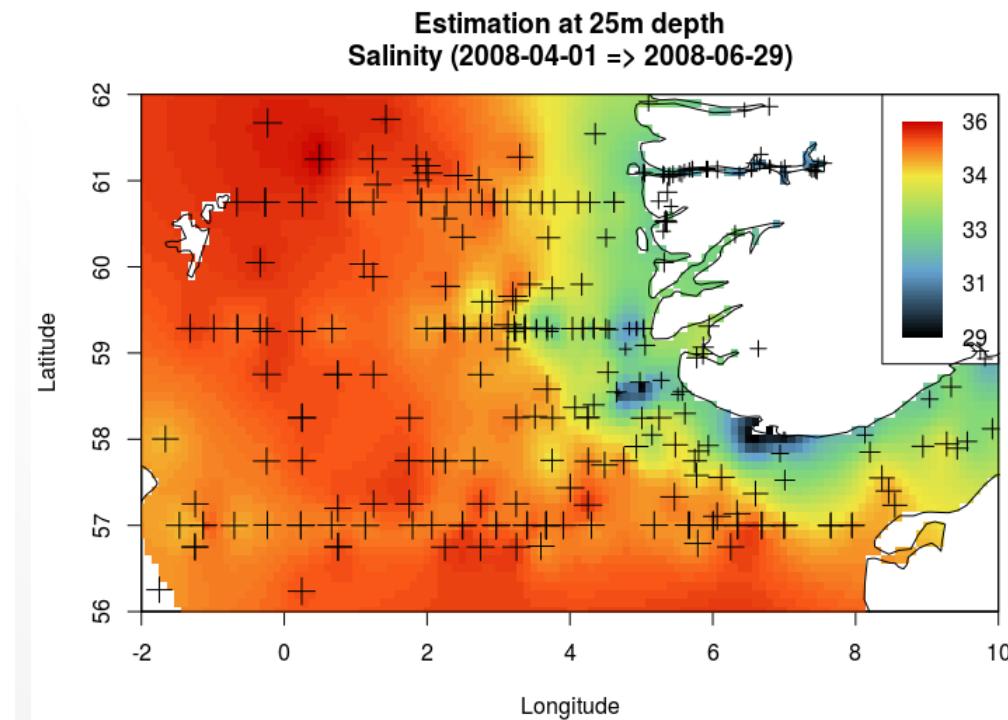
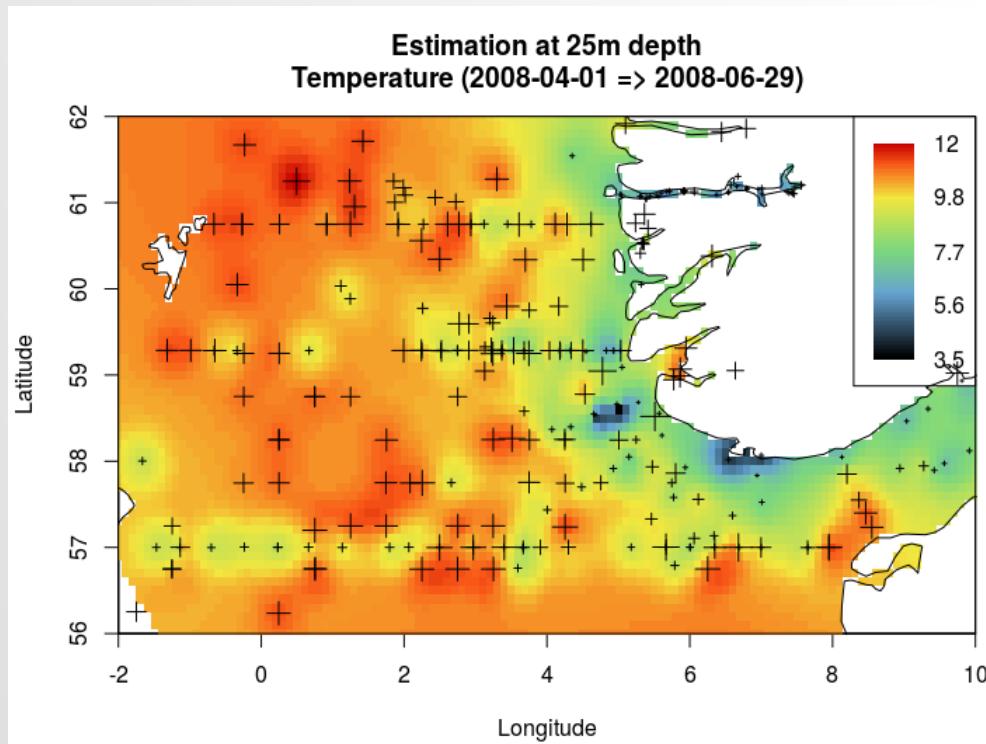
2-D Bivariate Variogram

- Select Depth Interval (25m) and Focus on 2nd Trimester of Year 2008
- Bivariate Variogram



2-D Estimation (Bivariate)

- Select Depth Interval (25m) and Focus on 2nd Trimester of Year 2008
- Calculate Average Horizontal Bivariate Variogram and Fit the Model
- Estimate the Temperature and the Salinity at 25m Depth by Cokriging

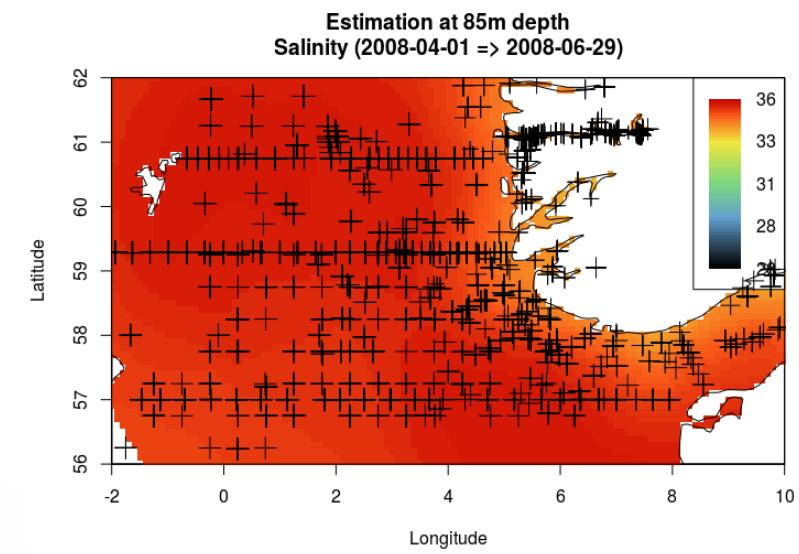
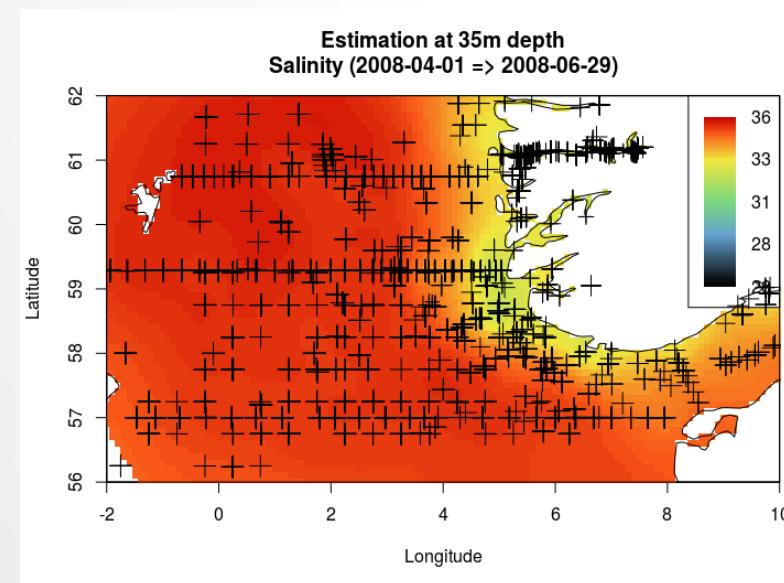
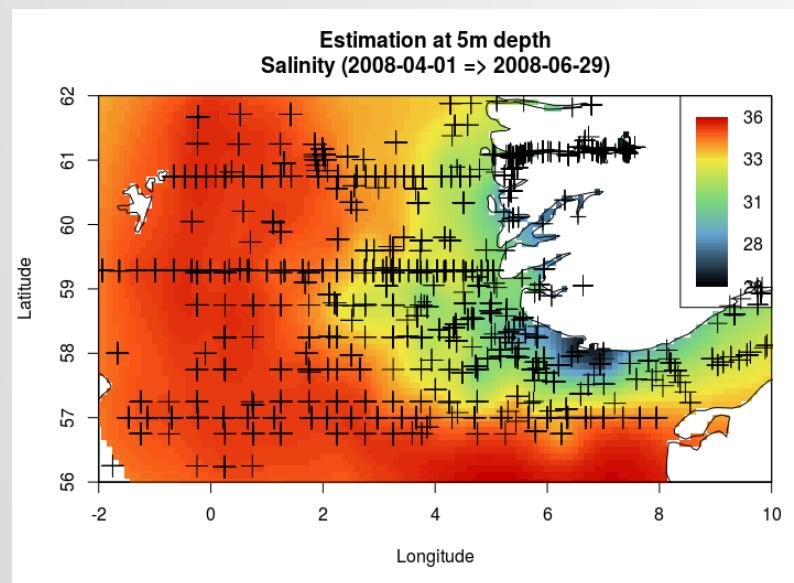


3-D Estimation of Salinity



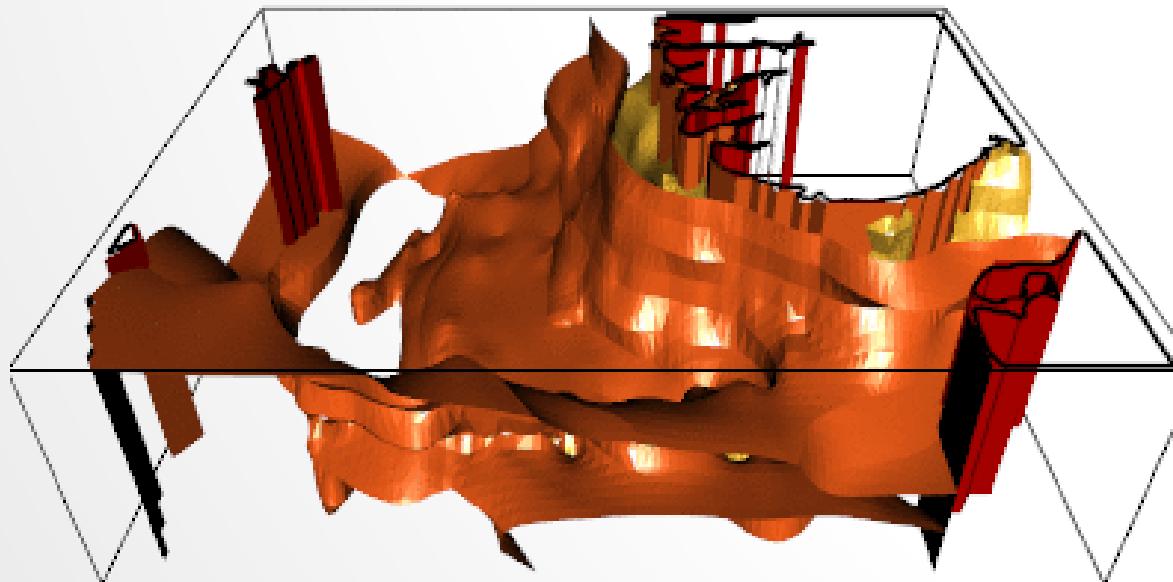
3-D Estimation of Salinity

- Define New Variable of Interest
- Focus on 2nd Trimester of Year 2008
- Calculate 3-D Variogram (by Bench) and Fit the Model for Salinity
- Estimate Salinity by Benches of 10m Thick from 5m to 95m Depth



3-D Estimation by Iso-Surfaces

- Show 3-D Estimation of Salinity with Iso-Surface



X



End of presentation

Thank you for attention!

