IMDIS 2016
International Conference on Marine Data and Information Systems

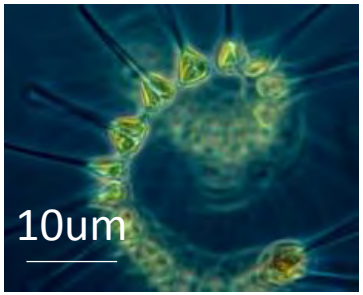
High resolution and automated flow cytometry data management

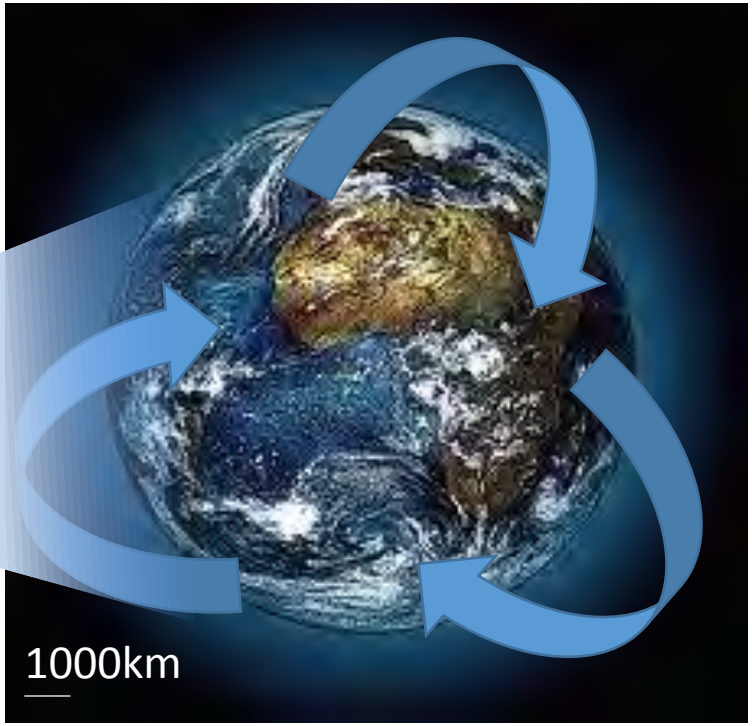
Gdansk (Poland) - October 11-13, 2016

Soumaya Lahbib (Data manager MIO), M. Dugenne (PhD candidate MIO), M. Libes (Computer Engineer OSU), C. Sammari (Professor INSTM), M. Belhassen (Scientist INSTM), G. Grégori (Research scientist MIO), Melilotus Thyssen (Research scientist MIO)

Phytoplankton

Thousands of species < 1000 µm catalyze the most important geochemical processes for sustaining life on earth AND at a minute scale.

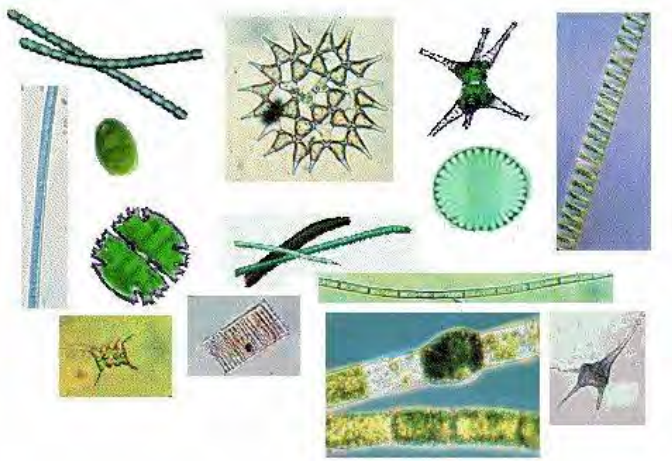




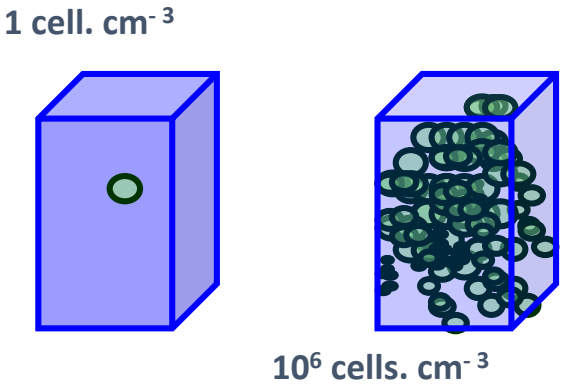
Phytoplankton produces between 45 and 57 Pg C Yr^{-1} of the NPP on earth (~45%) but represents <2% of its biomass.
Very high turn-over rate !

Phytoplankton observation is complex

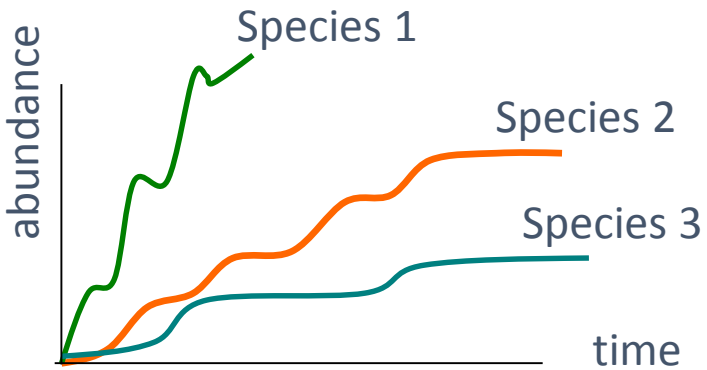
Morphology and size



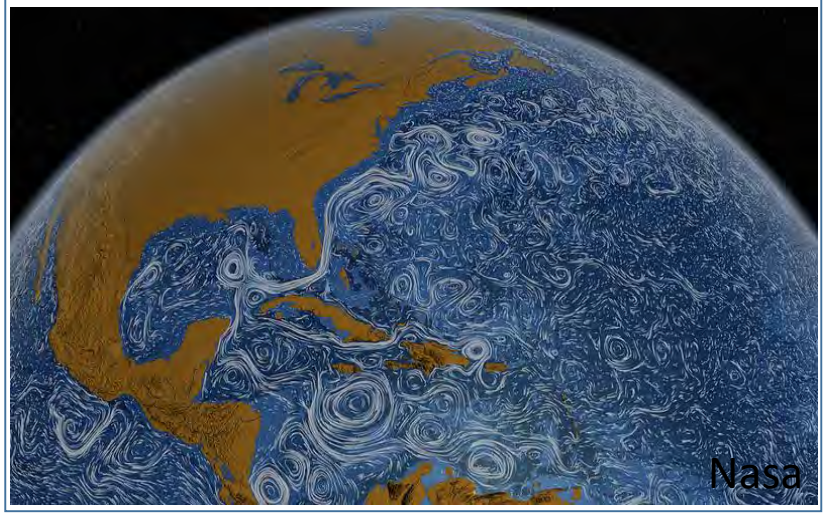
Abundances



Growth rates

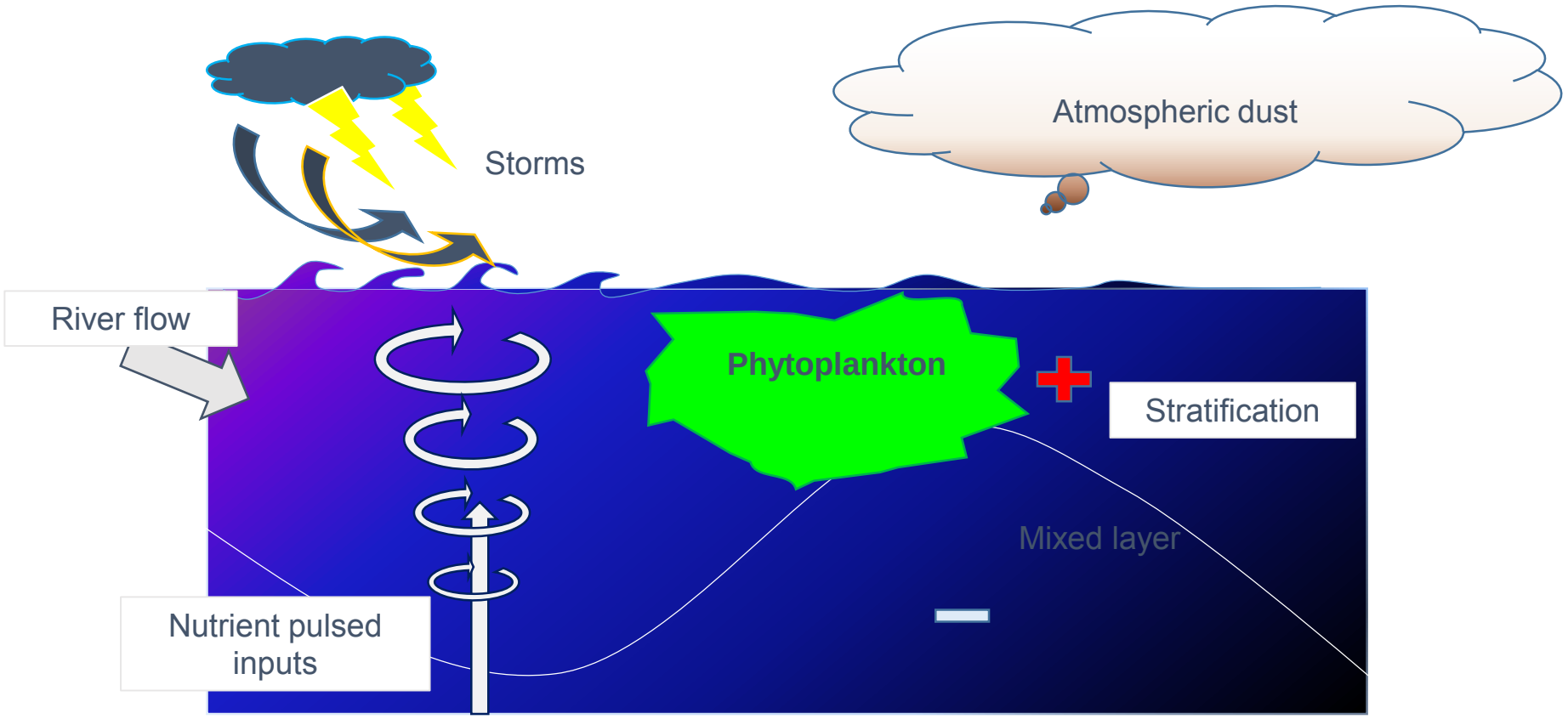


Turbulence

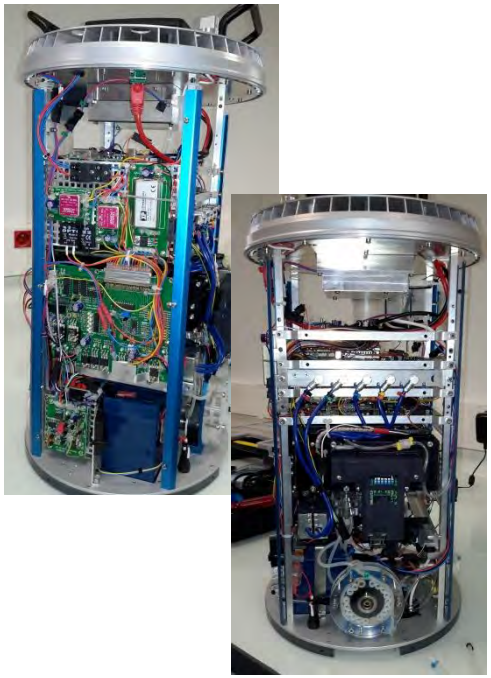


Serious lack in understanding and quantifying the role of phytoplankton in the biogeochemical processes

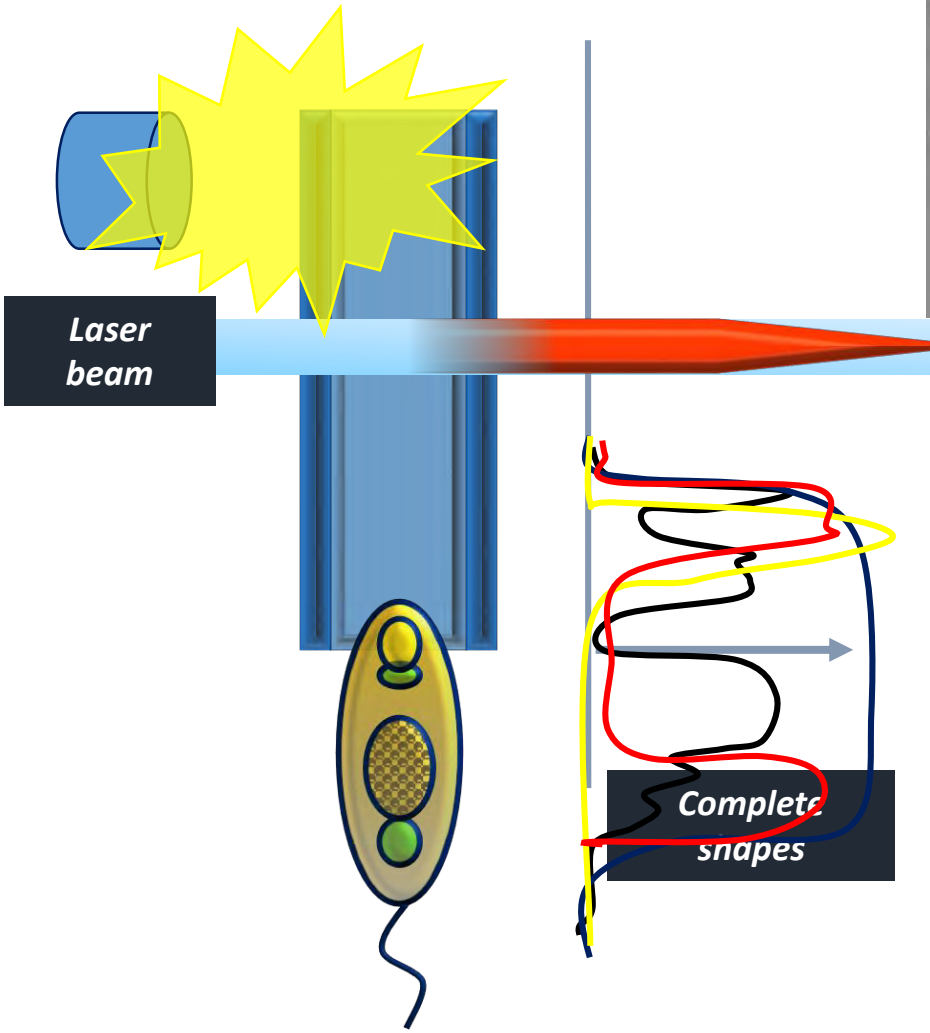
Short term variation and sporadic events impacts are nearly unknown

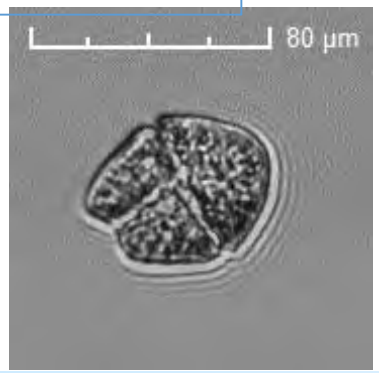


New technology for the resolution of phytoplankton functional diversity at hourly and regional scales



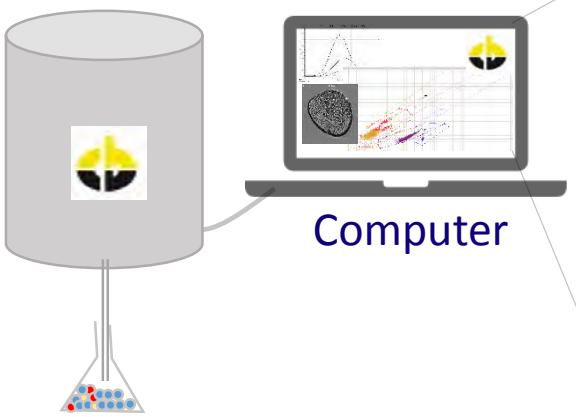
CytoBuoy
flow cytometry solutions





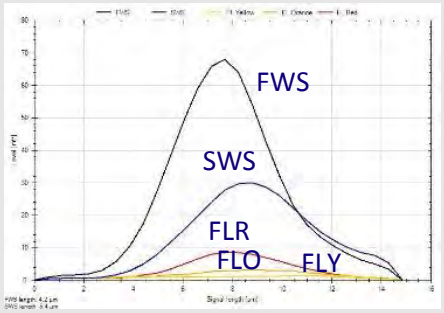
Fluorescences
Scatters
Abundances

Phytoplankton functional groups resolution

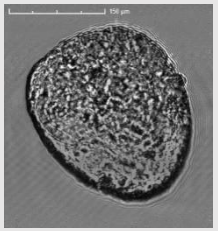


Computer

1 cell

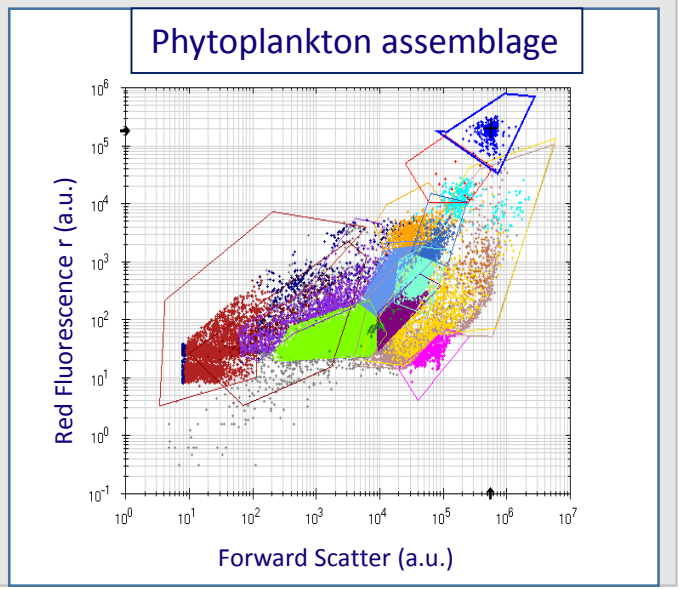


Pulse shape



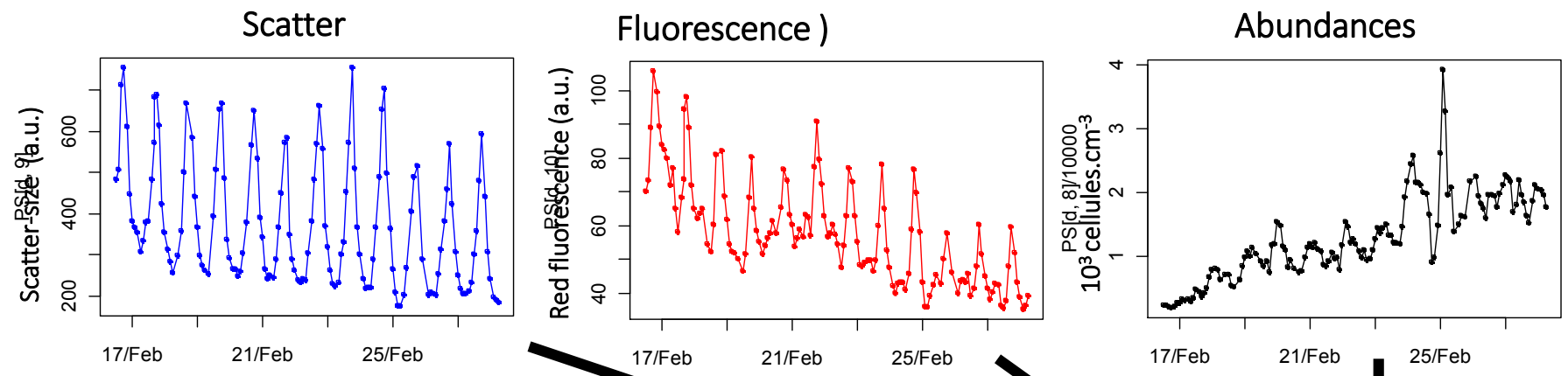
Picture

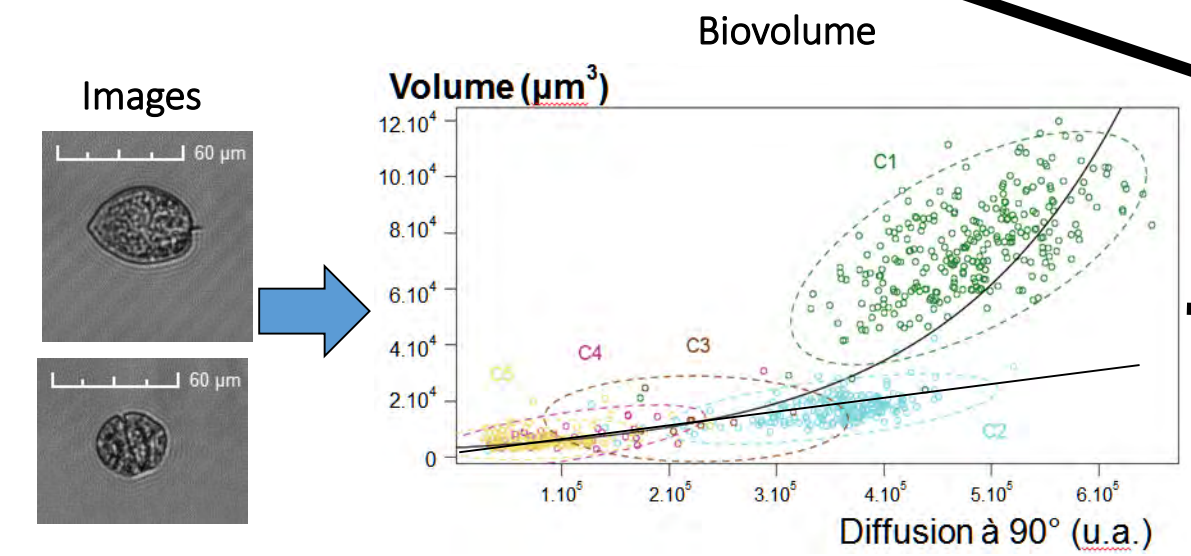
$1x10^6$ cells



- Phytoplankton functional groups/Phytoplankton abundance per group
- Fluorescences/scatter per cell/Size estimation after calibration of scatter
- Phytoplankton images (taxonomical identification >20 µm)

Additional information extracted from the single cell approach:





size-structured matrix population model:
In situ growth rate
per phytoplankton cluster

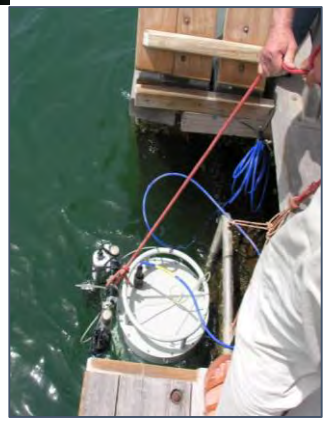
André *et al.,* 1999; Sosik *et al.*, 2003;
Thyssen *et al.*, 2009 ,2010,
Dugenne et al., 2014, 2015.

Several scientific experiences were conducted with a relative autonomie of 6 months

Scientific vessels Coastal platforms Ships of opportunity Buoys








Malkassian et al. 2011
Dugenne et al. 2014
Thyssen et al. 2008, 2009a.b, 2012, 2014, 2015

cnrs dépasser les frontières

mio Mediterranean Institute of Oceanography

CHROME

Continuous High Resolution Observation of the Mediterranean Sea:

https://chrome.mio.univ-amu.fr/

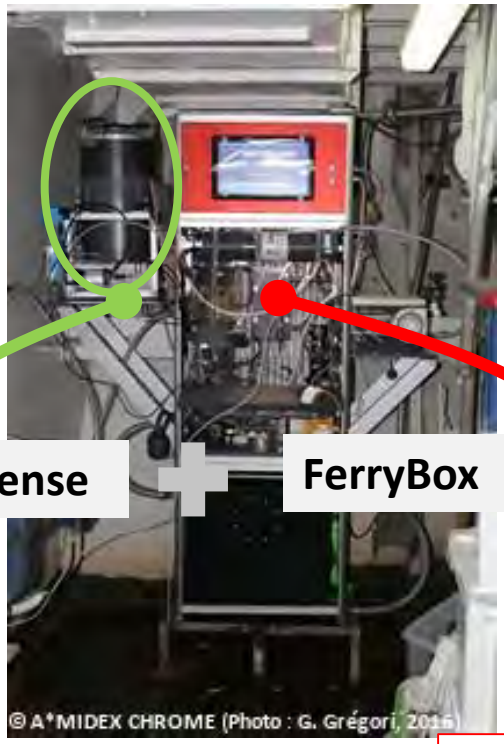
Understanding of the ecological and biogeochemical functioning in relation to meso-scale dynamics at the Mediterranean sub-basin scale and weekly scale.

A*MIDEX CHROME Project

Data Acquisition = one analysis every 30 min.

— C/F CARTHAGE trajectories

Pont 1 – C/F CARTHAGE



CytoSense + **FerryBox**

mio Institut Méditerranéen d'Océanologie

INSTM

./30min

./1min

- **Phytoplankton functional groups**
- **Phytoplankton abundance per group**
- **Fluorescences/scatter per cell**
- **Size estimation after calibration of scatter**
- **Phytoplankton images (taxonomical identification >20 µm)**

– **Temperature**
– **Salinity**
– **Fluoresence**
– **Turbidity**
– **pH**
– **pCO_2**
– **Oxygen**

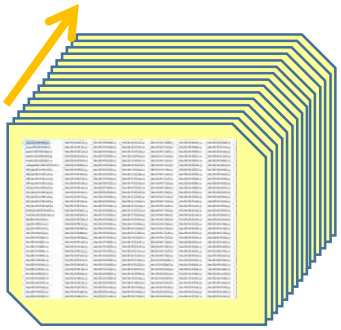
Data acquisition & analysis

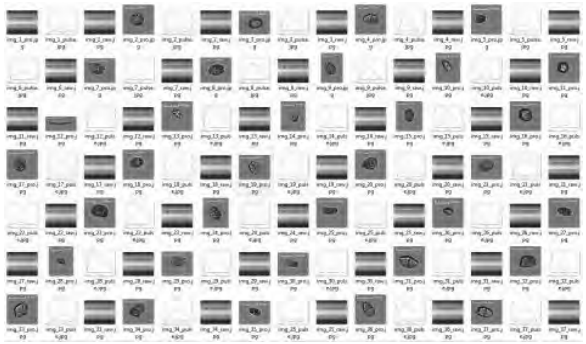
Measurements by CytoSense

CytoClus

Output

Manual clustering





Backup

<u>Seperate statistical CSV files:</u>
<u>Average</u> values of optical properties and <u>Counts</u>
+ Pictures

Issues

1. **Huge quantity of data**
2. **Data memory size consumming**
3. **Not a dedicated FCM database**

→ THE CYTOBASE DATABASE

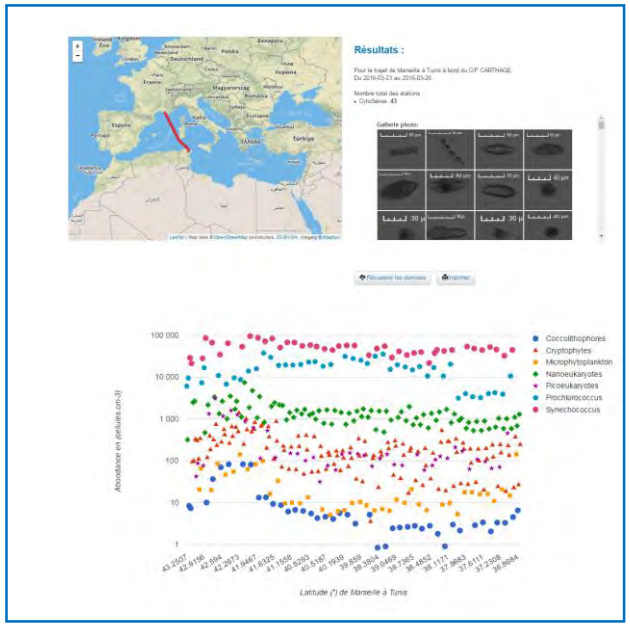
FCM Data management Workflow

CytoClus

R

Table Données
Table photo
Collection photos

Acquisition | **Analysis** | **Consolidation** | **Expert QC**



Accessibility

CYTOBASE

Integration

Data consolidation

Cytobase Input Processor (M. Dugenne, 2015)

Cytobase data

Parcourir... 11 fichiers sélectionnés.

Upload complete

Warning:

Upload successful

Create Inputs folder

Project and samples context | Raw data | Size conversion

Metadata

Project	Project date	PI	Cytometer ID
Enter project name	2015-08-18	Enter PI name	Enter cytometer ID
Station	**Depth**	**Latitude**	**Longitude**
Filename	**Filename model** BERRE_082013_3F_FLR9 2013-12-17 13(		
Samples operator	**Standards reference**	**Clustering method**	**Observation type**
Enter name of operator	Enter standards beads ref	Automated	In situ

Data

2013-12-17T14:17:00Z	2013-12-17T14:17:00Z	16,056.60	Synechococcus	BERRE_082013_12S_FLR9 2013-12-17 14u17.cyz	1.99	FL Red	10	
2013-12-17T14:17:00Z	2013-12-17T14:17:00Z	16,056.60	Cryptophytes	BERRE_082013_12S_FLR9 2013-12-17 14u17.cyz	1.99	FL Red	10	
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61	Beads 2 mu	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10	
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61	Microphytoplankton	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10	
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61	Picooeukaryotes 2	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10	
2013-12-17T14:43:00Z	2013-12-17T14:43:00Z	16,056.61	Picoeukaryotes 1	BERRE_082013_16F_FLR9 2013-12-17 14u43.cyz	2.04	FL Red	10	

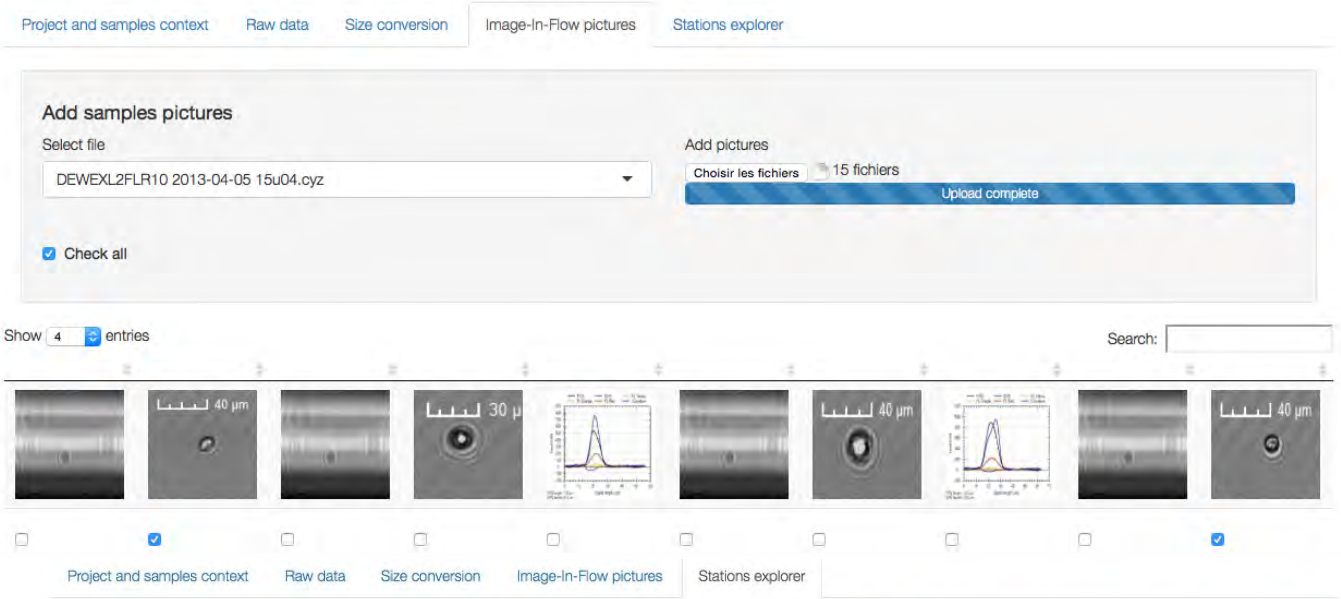
Please associate each selection set to trigger, PMt's amplification and standardized phytoplankton category

NB: All incompatible entries will be removed

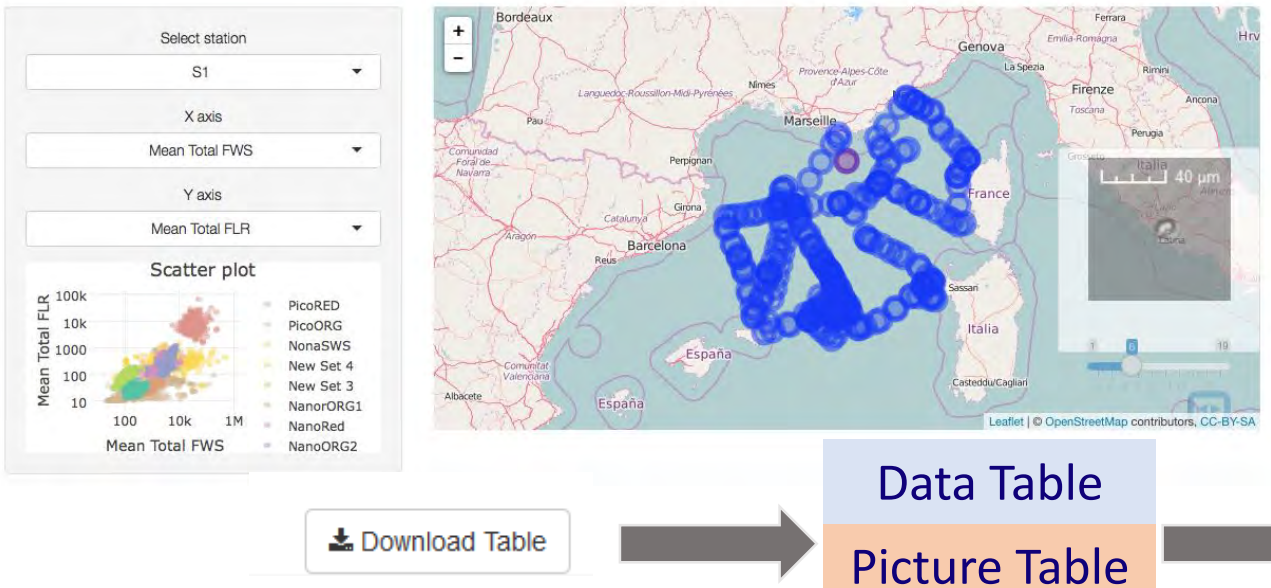
Expert name	Trigger	PMT's amplification	Standardized name	
Cluster	Channel/Level	SWS	Cluster	
Beads 2 mu	FL Red 10	0 50 100	Standard beads	Associate
		FLO 0 50 100		

Data consolidation

Picture selection



Stations Explorer



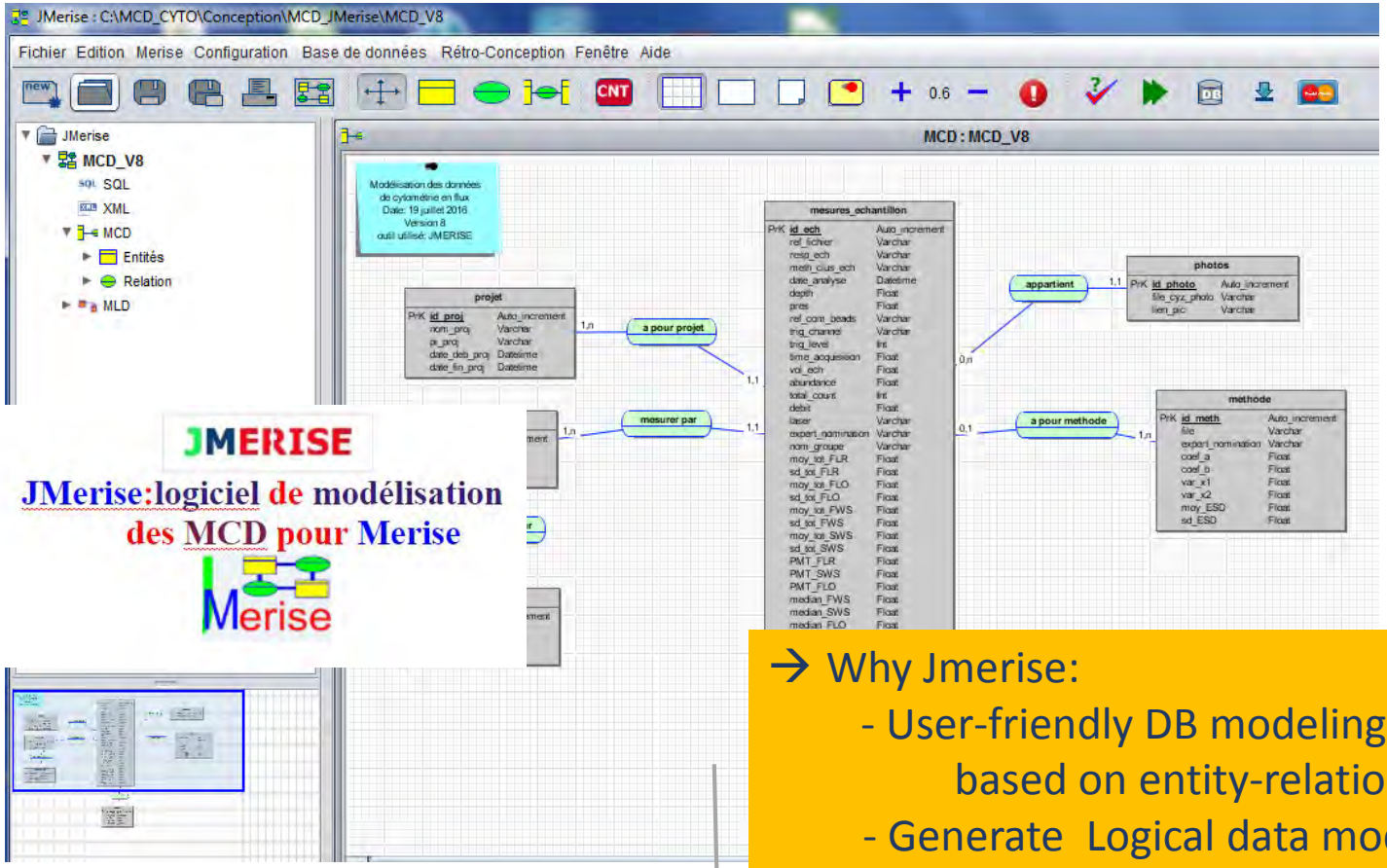
Download Table

Data Table

Picture Table

CYTOBASE modeling

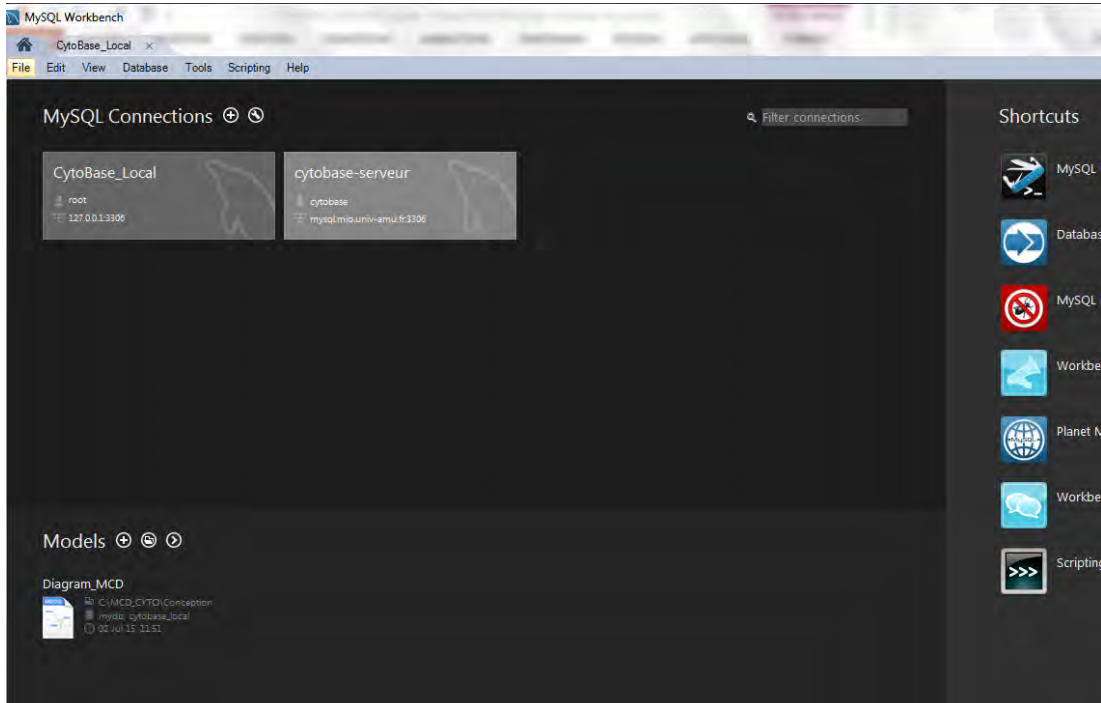
→ User requirements acquisition
→ Management rules clarification
→ First modeling versions
→ Validation and/or modifications



→ Why Jmerise:
- User-friendly DB modeling programme: based on entity-relation methods
- Generate Logical data model and Mysql script

CYTOBASE modeling

MySQL Workbench

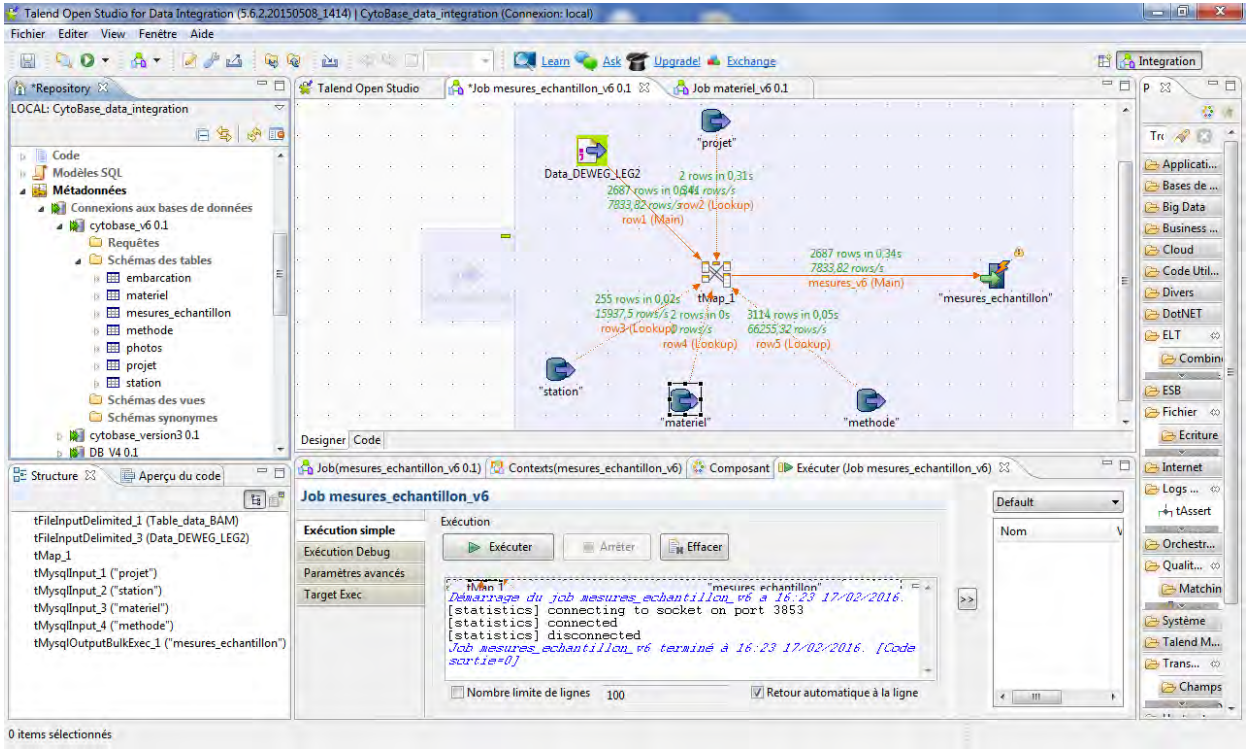


→ Easy communication with DB
→ Management and administration of DB
→ Editing and quick querying

CYTOBASE data integration

ETL

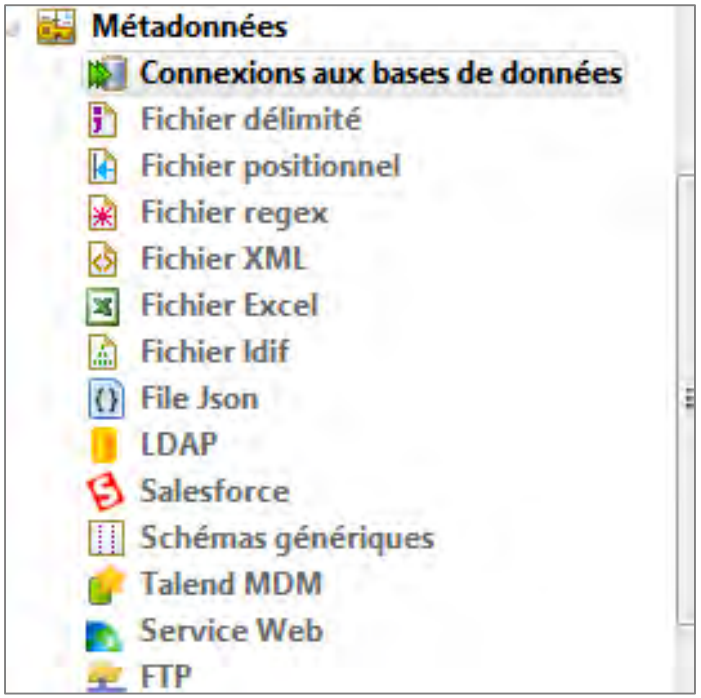
talend



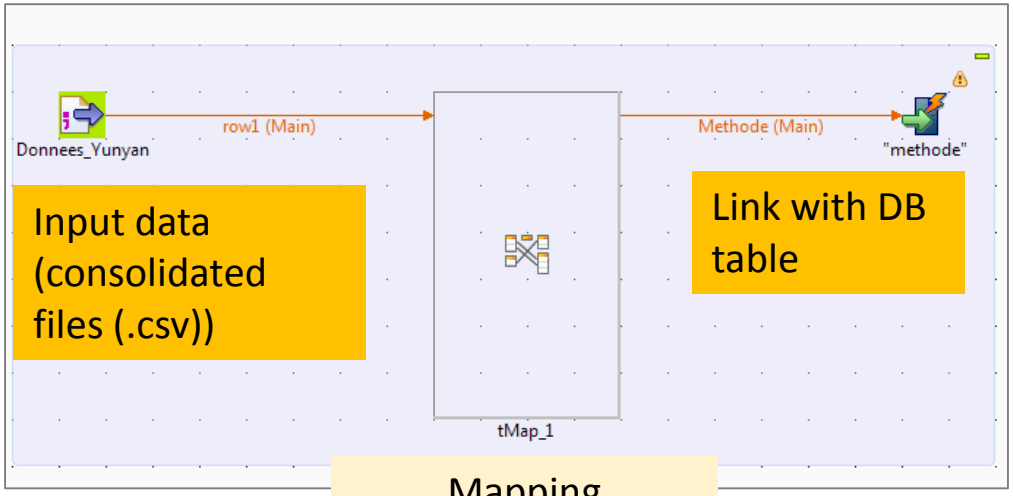
→ ETL (Extract, Transform, Load) processing development platform

talend

→ Connectivity with DB and multiple file types

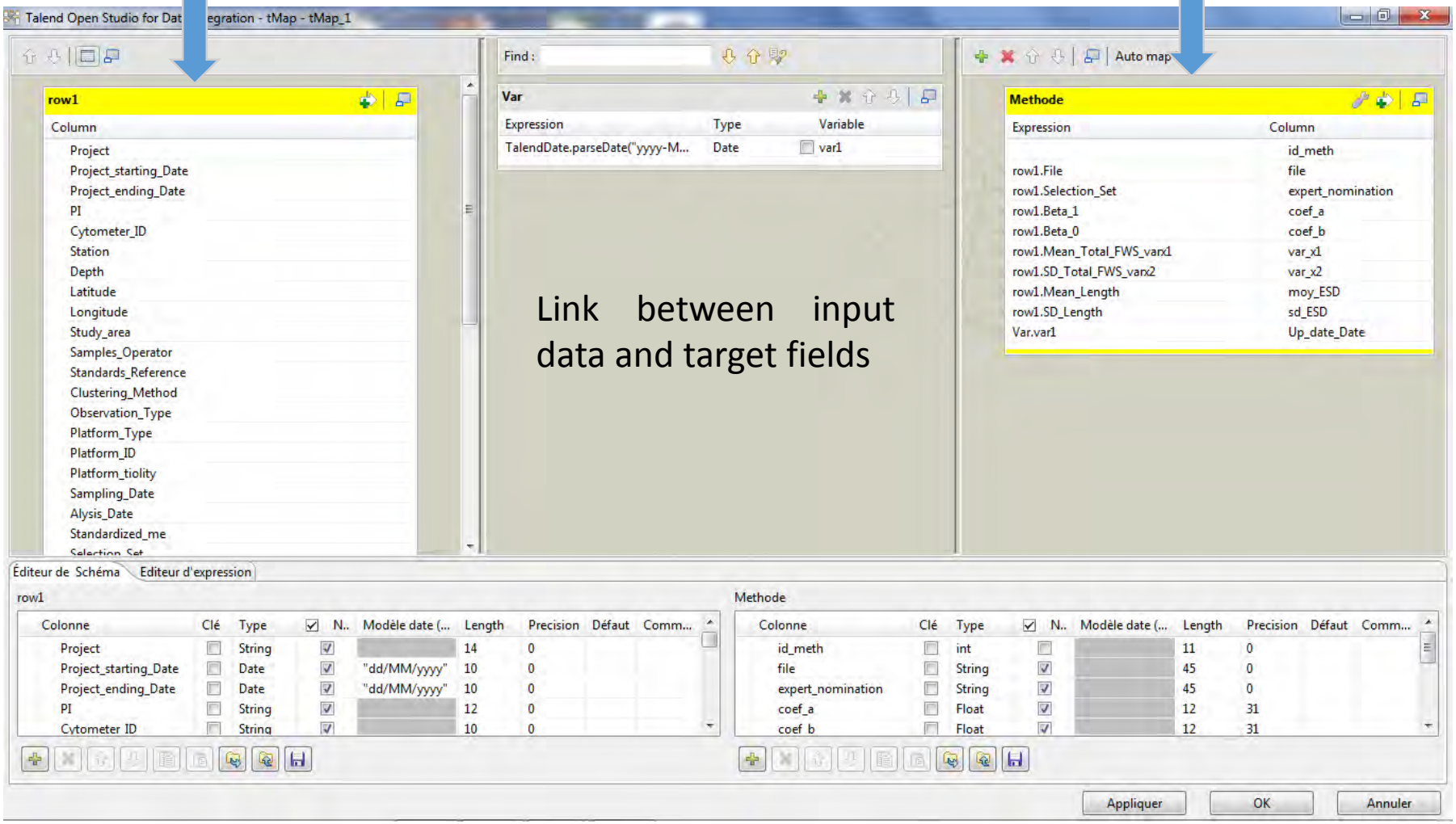


→ Data integration optimized processing time



Input data

DB table



Link between input data and target fields

CYTOBASE data accessibility

Javascript library

Google Developers dygraphs

Javascript charts

CHROME +

https://chrome.mio.univ-amu.fr/chrome-cytobase/

(Free access)

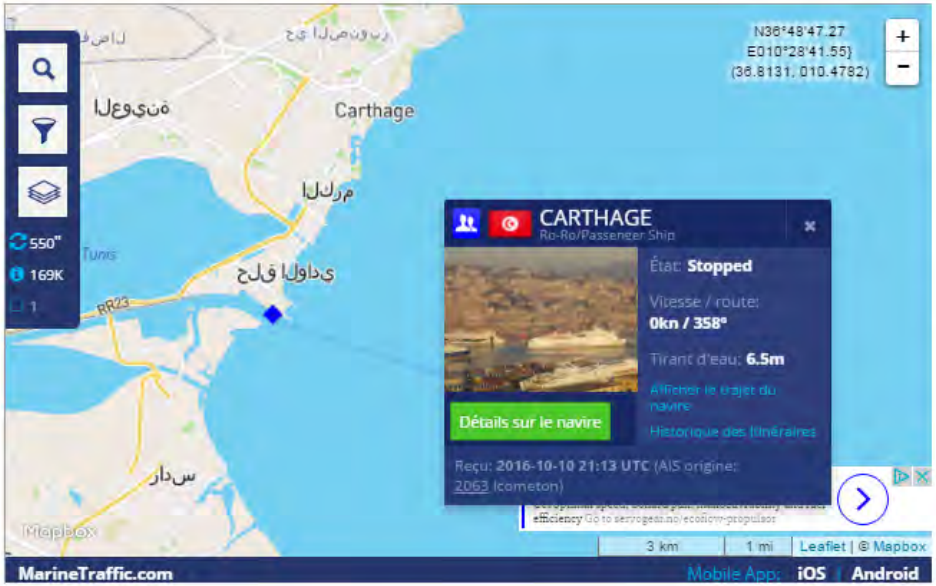
http://www.mio.univ-amu.fr/cytobase/

(Access within the MIO)

CYTOBASE data accessibility

CHROME

Continuous and High Resolution Observation of the MEditerranean Sea



Zone d'étude *(Area of Study)*:
Mediterranean Sea - Western basin

Trajet *(Transect)*:
MRS-TUN : du 24-03-2016 au 25-03-2016

Date de début *(Start Date)*
23/03/2016

Date de fin *(End Date)*
26/03/2016

Note: *Please refer to the date mentionned in "Transect"*

Instruments :

- [x] CytoSense
- [x] Photos cellules phytoplanctoniques
- [ ] FerryBox *(Données pas disponibles)*

Rechercher

https://chrome.mio.univ-amu.fr/chrome-cytobase/

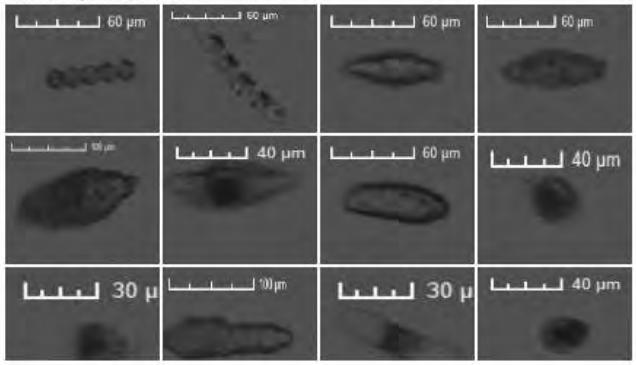
Résultats :

Pour le trajet de Marseille à Tunis à bord du C/F CARTHAGE
Du 2016-03-23 au 2016-03-26

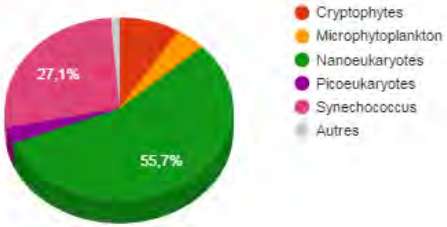
Nombre total des stations :

- CytoSense :**43**

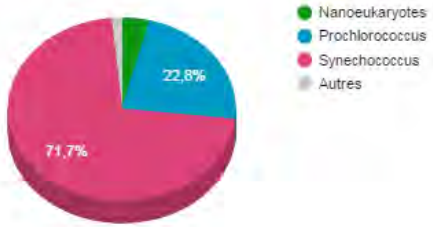
Gallerie photo:

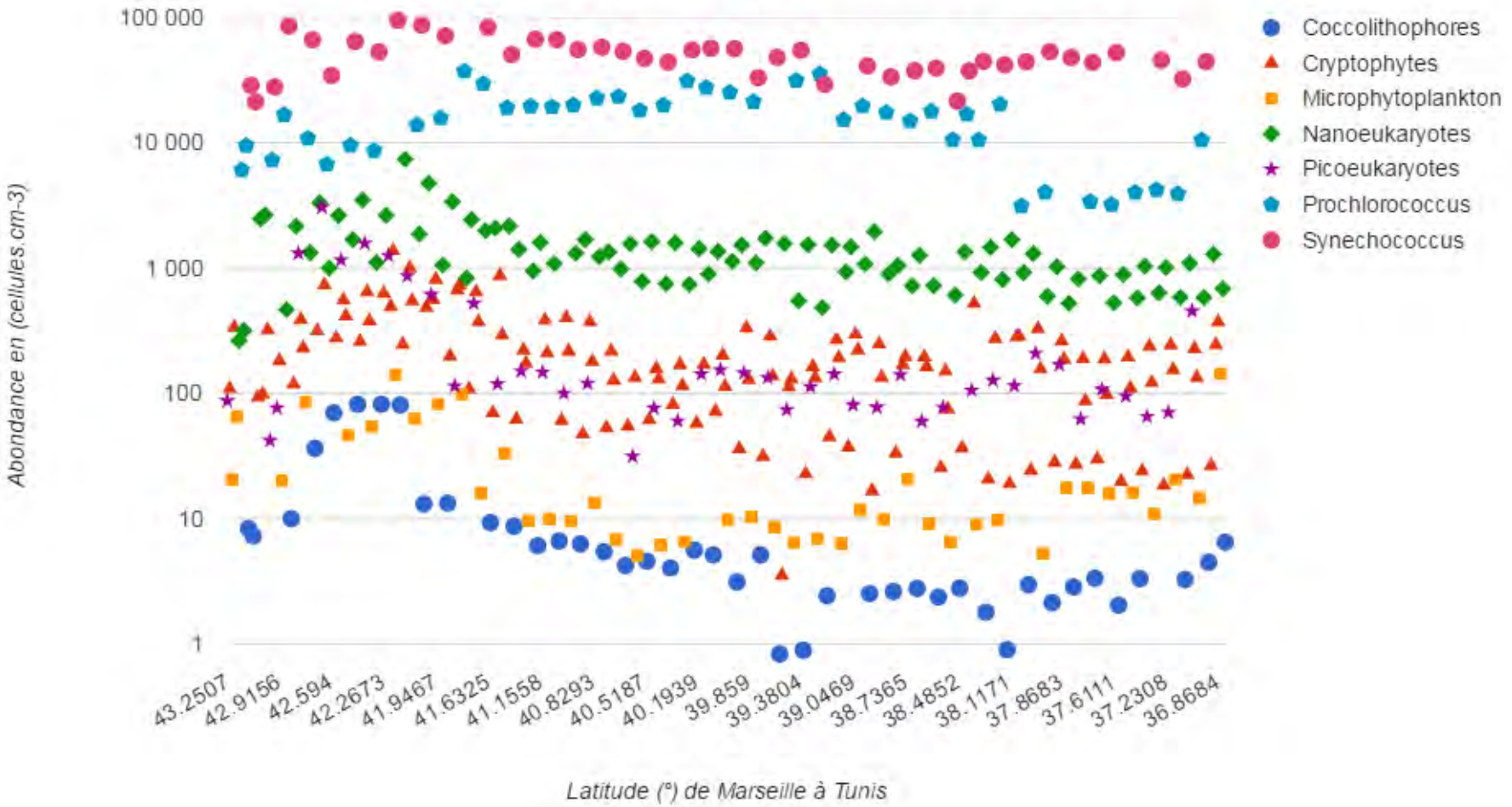


Total Fluorescence Rouge (u.a..cm^{-3})



Total Abondance (cell.cm^{-3})




100 000
10 000
1 000
100
10
1
Abondance en (cellules.cm-3)
43.2507
42.9156
42.594
42.2673
41.9467
41.6325
41.1558
40.8293
40.5187
40.1939
39.859
39.3804
39.0469
38.7365
38.4852
38.1171
37.8683
37.6111
37.2308
36.8684
Latitude (°) de Marseille à Tunis
Coccolithophores
Cryptophytes
Microphytoplankton
Nanoeukaryotes
Picoeukaryotes
Prochlorococcus
Synechococcus

Cytometric distribution parameters per group

CYTOBASE data accessibility (Access within the MIO)

FLOW CYTOMETRY DATABASE

Zone d'étude: Mediterranean Sea - Western basin

Nom et date du Projet: DEWEX LEG2 : du 01-04-2013 au 30-04-2013

Date de début: 31/03/2013

Date de fin: 30/04/2013

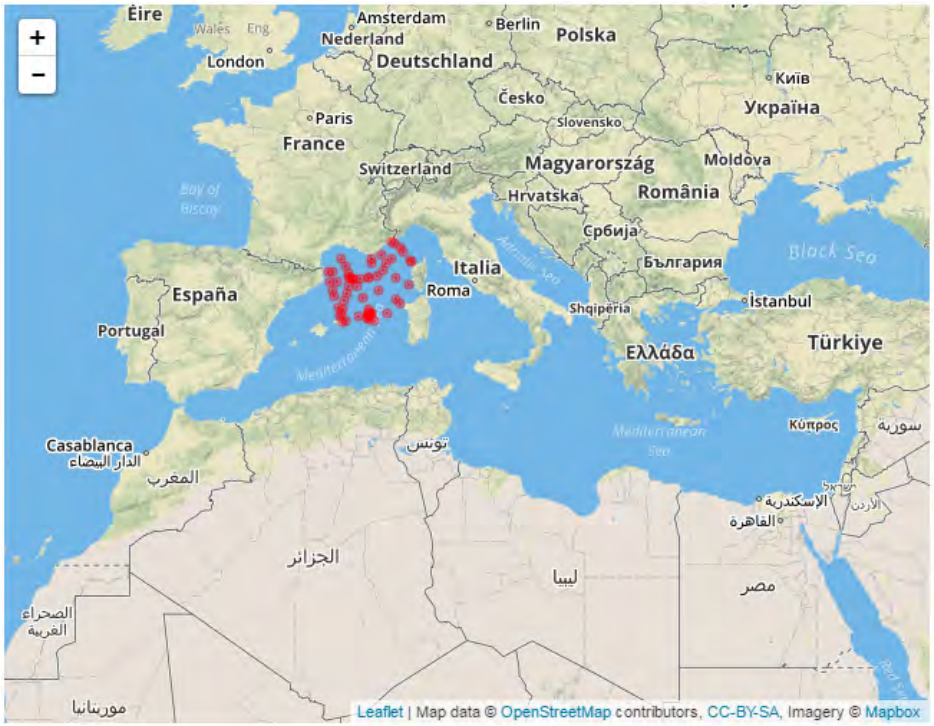
Flow cytometer :

- ○ CytoSense
- ◉ FACS Calibur
- ○ BD Influx
- ○ All instruments

- ○ Phytoplankton groups
- ◉ Heterotrophic Bacteria

Search

FLOW CYTOMETRY DATABASE



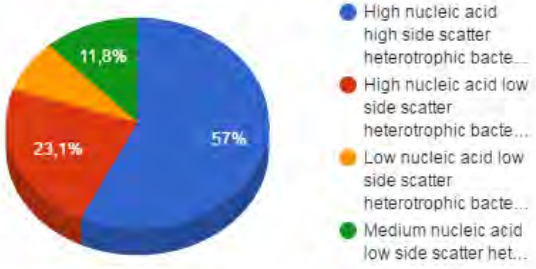
Results :

For the period from **2013-03-31 to 2013-04-30**
Project name **DEWEX LEG2**
Flow Cytometer : **FACS Calibur**
Station number : **59**

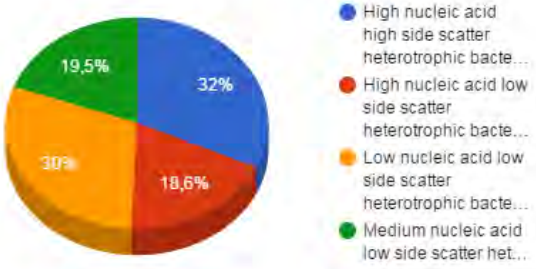
No pictures

Download Data | Print

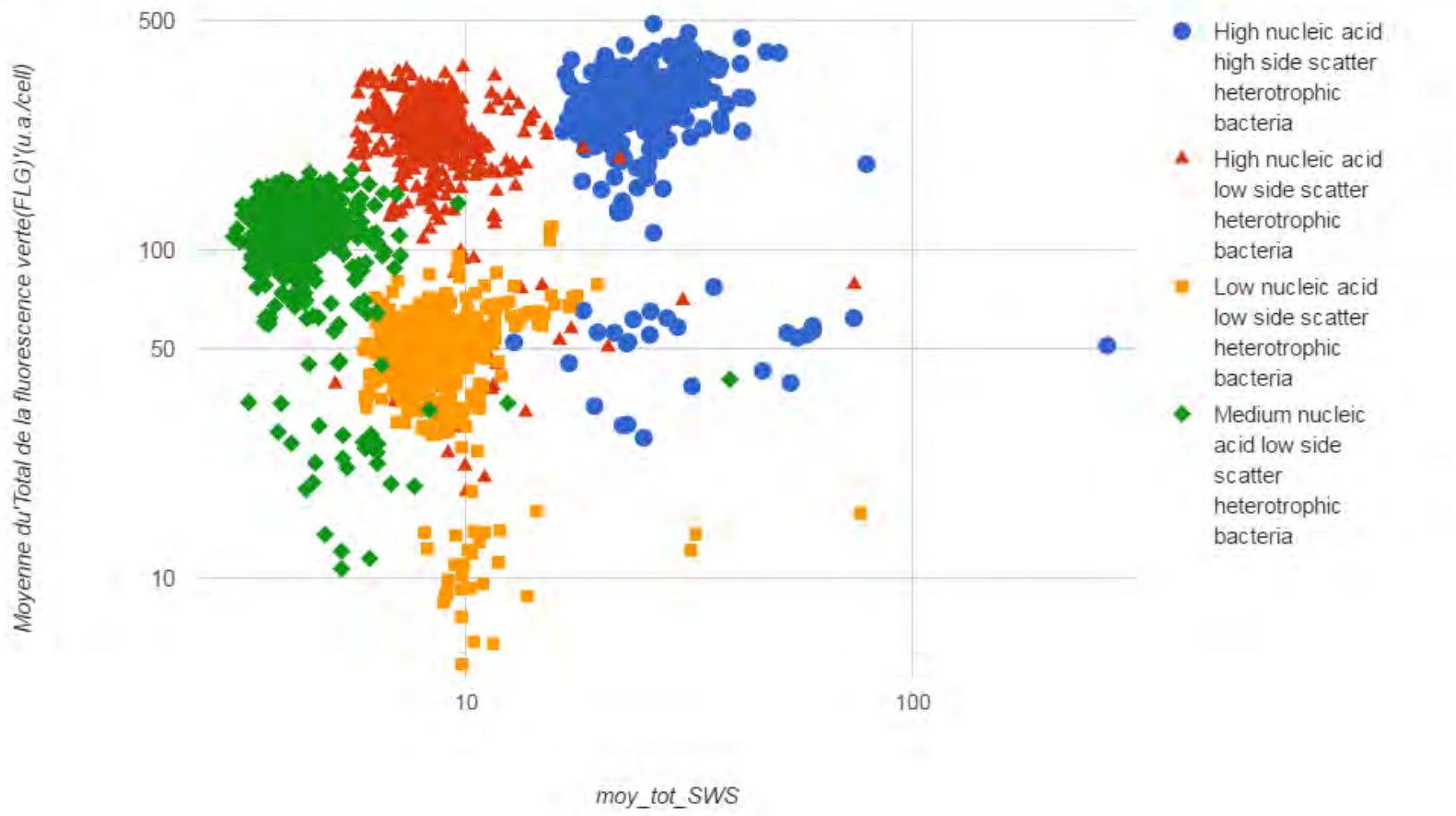
Contribution to Green Fluorescence (u.a..cm^{-3})



Relative Abundance (cell.cm^{-3})

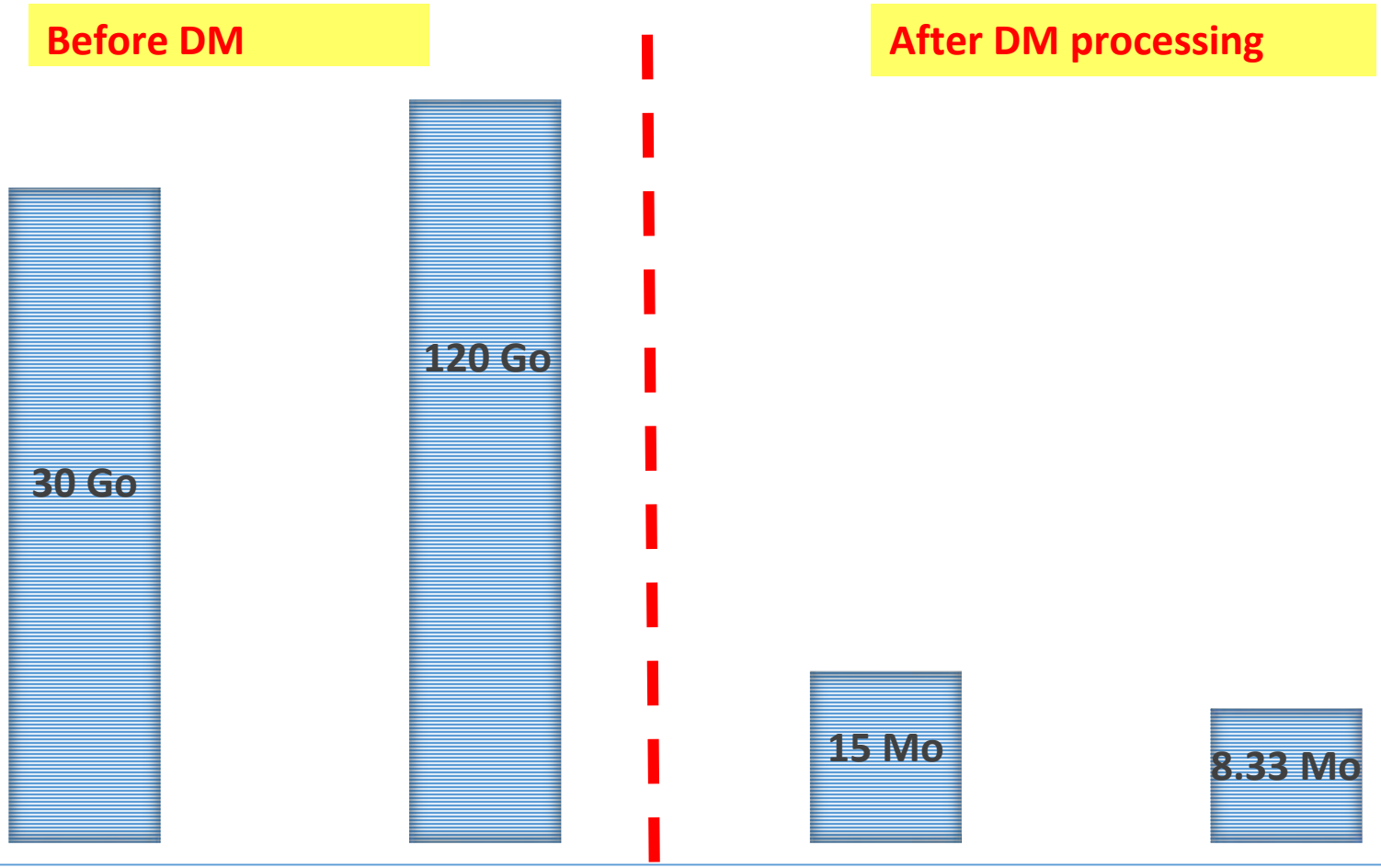


Cytometric distribution parameters per group

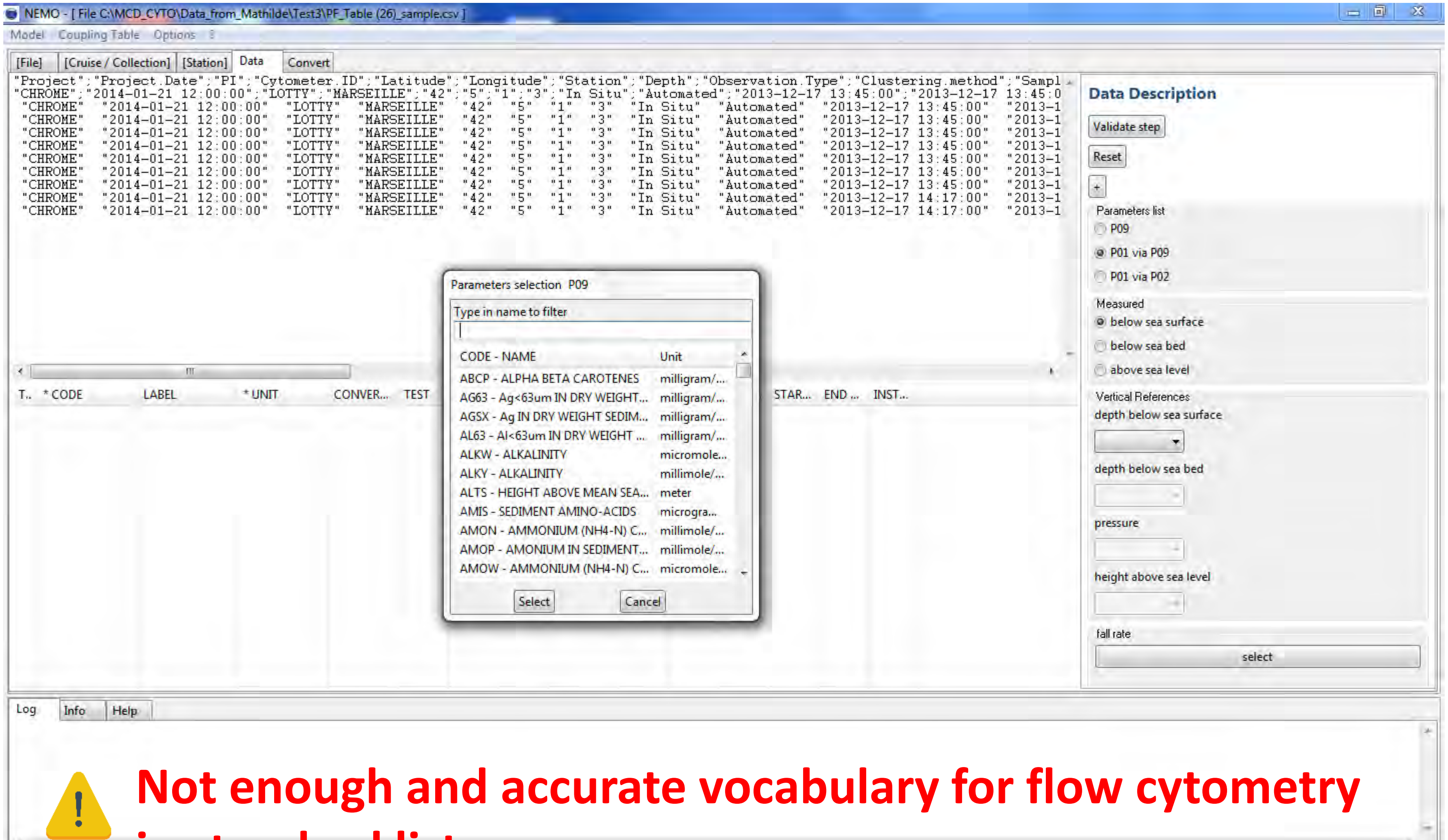


CYTOBASE data management method of the MIO

- Case of a 9 days Cruise : 1 sample / 20 min



Next Step!



Not enough and accurate vocabulary for flow cytometry in standard lists

Next Step!

➔WP9.5.2 of SEADATACLOUD (VLIZ, MIO, NERC-BODC and ICES):

Ingesting, validating, long-term storage and access of Flow Cytometry data

☛**We need to work with the automated and *in situ* FCM Community: Task 3.1 of JERICO Next**

Conclusion

- **Smart storage and sustainability of FCM data with net decrease on file size**

- **CYTOBASE is a dynamic and user friendly web-based interface**

- **The next step, we are working on FCM data interoperability with French national DB (SOMLIT) and international DB (SeaDataNet) (in compliance with INSPIRE Directive)**

- **It will be possible to link with task 5.4 (JERICO) and WP9 of SeaDataCloud so as to work with large FCM community.**

- **By working on FCM common vocabulary, we can adopt the best practice tools and data management method used in MIO (CYTOBASE) and in SeaDataNet (SeaDataCloud).**


Thank you for your attention
mio
Institut Méditerranéen d'Océanologie
Aix*Marseille université