Towards the use of CNNs to constrain the phytoplankton community response to environmental changes

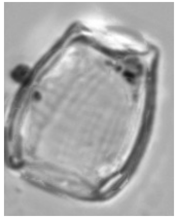
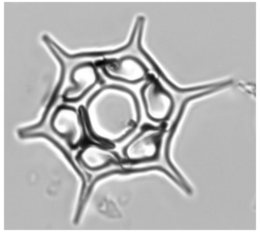
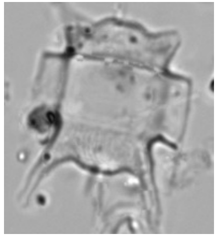
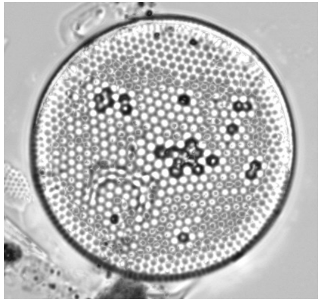
– An example from the Mediterranean Sea –

Camille Godbillot

CEREGE: Luc Beaufort, Yves Gally, Ross Marchant, Baptiste Pesenti, Thibault de Garidel-Thoron

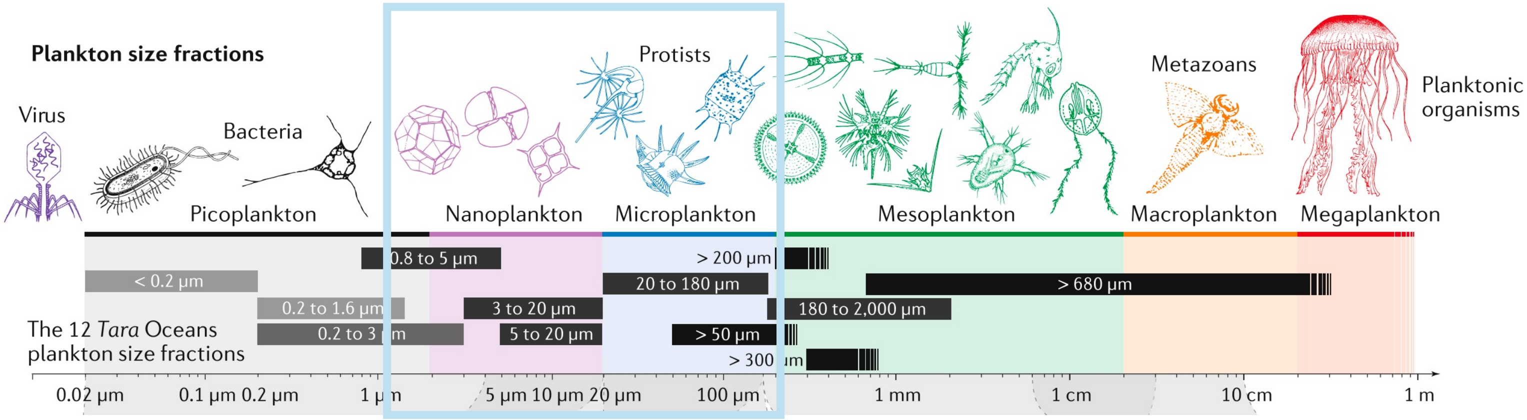
MIO: Karine Leblanc, Julien di Pane, Thang D.Q. Le, Cristele Chevalier

CEFREM: Xavier Durrieu de Madron

RESEARCH INTERESTS

Geochemical and morphological response of the phytoplankton community to environmental changes



Sunagawa *et al.*, 2020

Changes in assemblage + morphology

Species migrations
Morphological changes

ENVIRONMENT ⇄ **PHYTOPLANKTON** ⇄ **HIGHER TROPHIC LEVELS**

Biological carbon pump

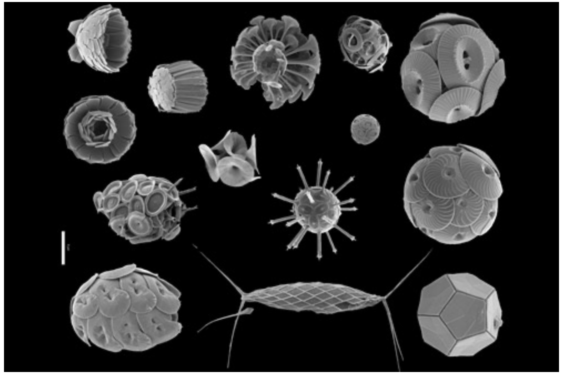
Predation

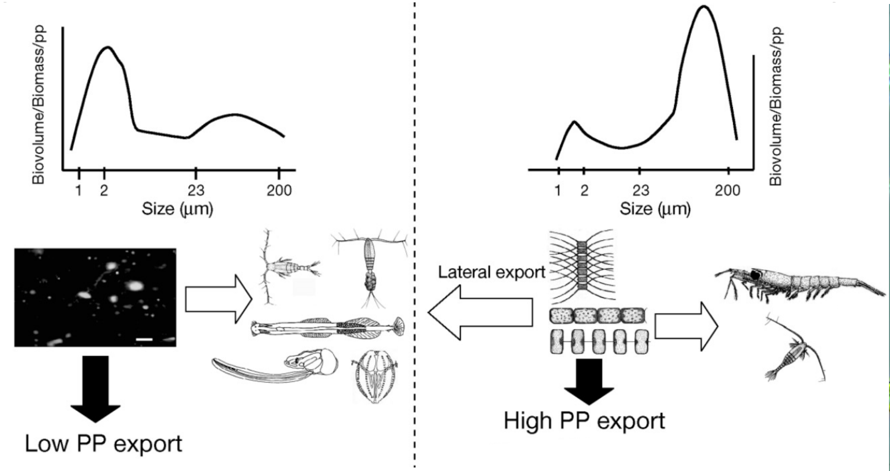
Research interests

Geochemical and morphological response of the phytoplankton community to environmental changes

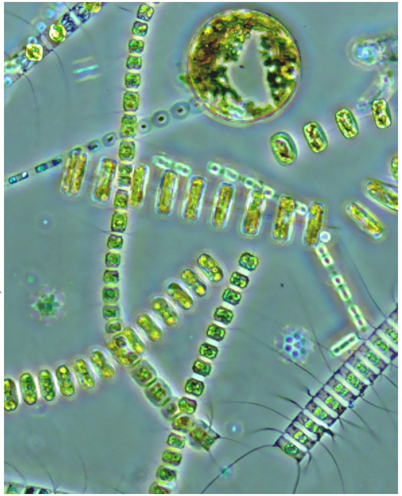
Coccolithophore-dominated community

Diatom-dominated community

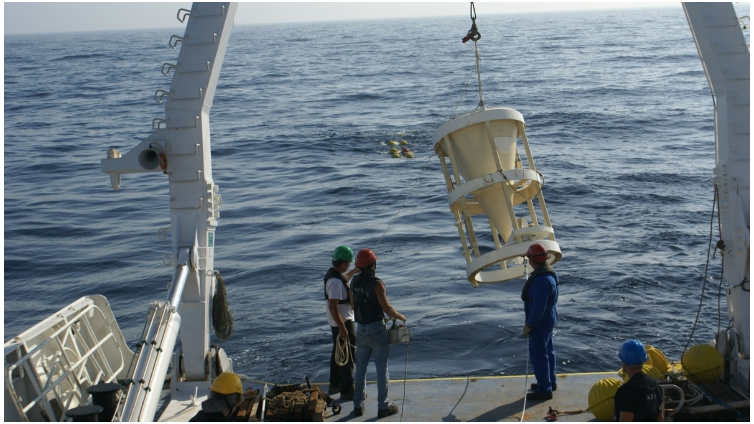




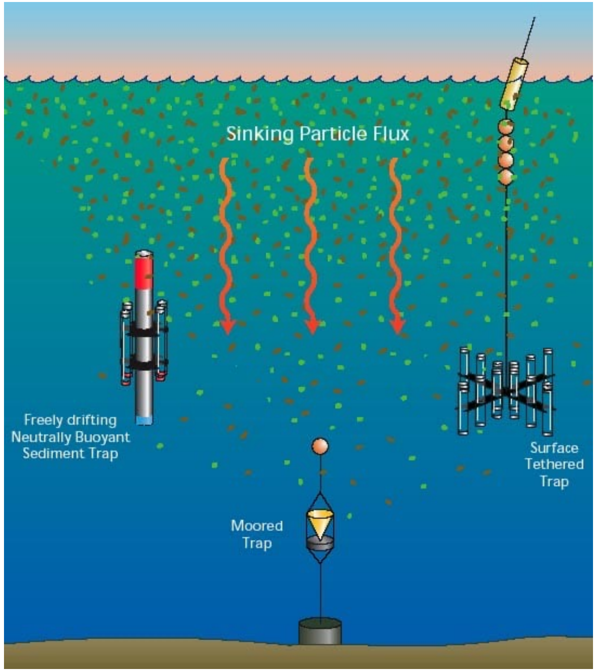
Iriarte & González, 2004



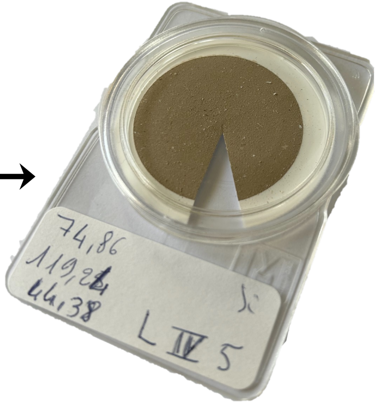
Research interests – Study of sediment traps



→



→

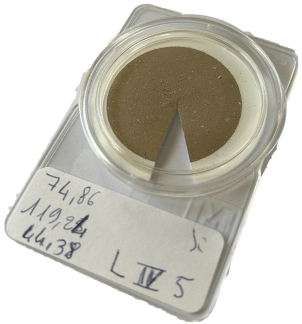


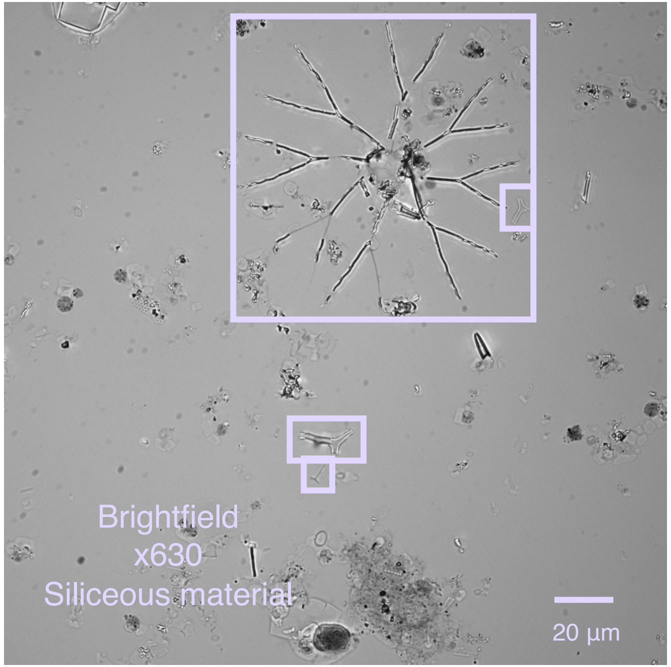
Sediment traps yield insight into the:

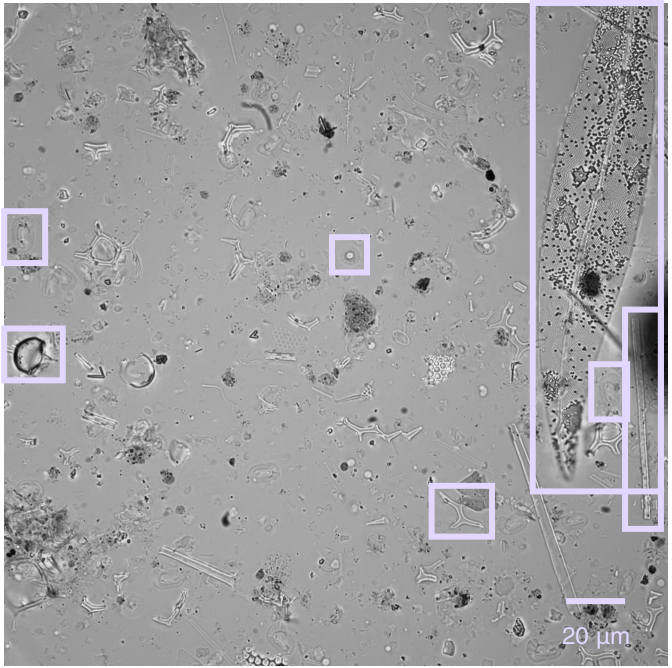
- **Biological production at the surface**
- **Carbon export** to the sea floor
- **Seasonal changes** in particle deposition

Context

Initial sediment →

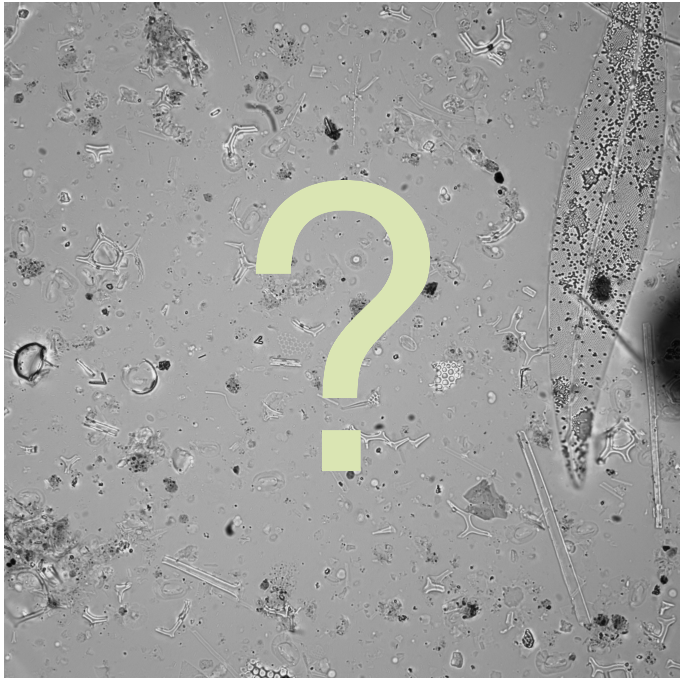






Obtain **count** data + object **morphology** *i.e.* length, [mass] (using birefringence)

Context



Can we make this data acquisition automatic?

- Overlapping particles
- Differences in size and shapes
- Look similar to the background

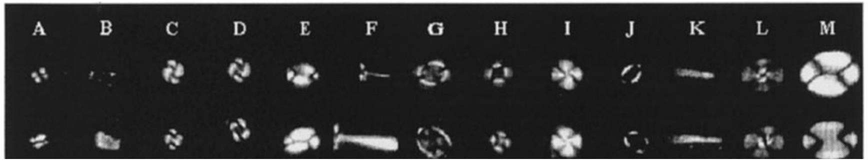
New object detection workflows can now be tested

Overview of the methods used at CEREGE

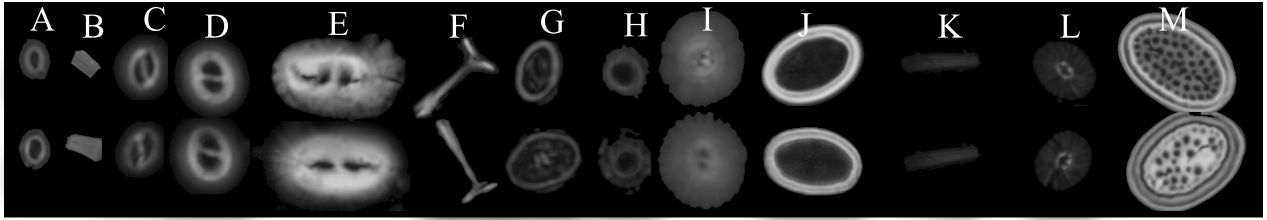
CONTEXT – OVERVIEW OF THE METHODS USED AT CEREGE

Image acquisition methods have improved, the classification methods have as well.

1996



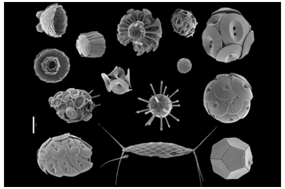
2020



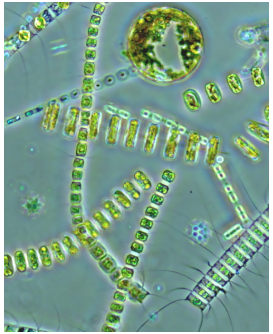
Year	Lens resolution	Camera (res. Spat.)	Pixel	Camera (niv. gris)	Polarisation	Focus	Slide preparation method	Patern recognition
1996	50X	0.5 Mpixel	0.3 µm	8 bits	Linear	Auto	Smear slides	CNN
2004	50X	0.5 Mpixel	0.3 µm	8 bits	Linear	Auto	Smear slides	Dyn CNN + hierarchy
2012	100X	4 Mpixel	**0.062 µm**	**14 bits**	Linear	Auto	Smear slides	Dyn CNN + hierarchy
2014	100X	4 Mpixel	0.062 µm	14 bits	**Rotative / circular**	Auto	Random settling	Dyn CNN + Random Forest
2020	100X	4 Mpixel	0.059 µm	16 bits	**Bidirectional Circular**	**Multi-focus**	Random settling 8 mini-lamelles	ResNet

CONTEXT – OVERVIEW OF THE METHODS USED AT CEREGE

Plateforme de micropaléontologie automatisée labellisée PRT (AMU, CNRS, INSERM) depuis 2021



Groupe étudié	depuis	Personnel contributeur AI	Materiel impliqué
Coccolithes fossiles	1995	LB, CB, BSM, 4 thèses	4 micro-auto
Coccolithophores : plancton	2005	LB, 1 thèse, 1 post-doc	*idem*
Coccolithophores cultures	2017	LB, 1 Post-Doc	1 microinversé auto
Foraminifères planctoniques	2015	TdGT, 1 thèse, 1post-doc	MISO + 2 binos auto
Foraminifères benthiques	2016	TdGT, LL, 1 thèse	*idem*
Radiolaires	2019	1 post doc, LB	1 micro-auto
Pollens	2018	DB, 2 thèses	*idem*
Diatomées	2022	1 post-doc	*idem*

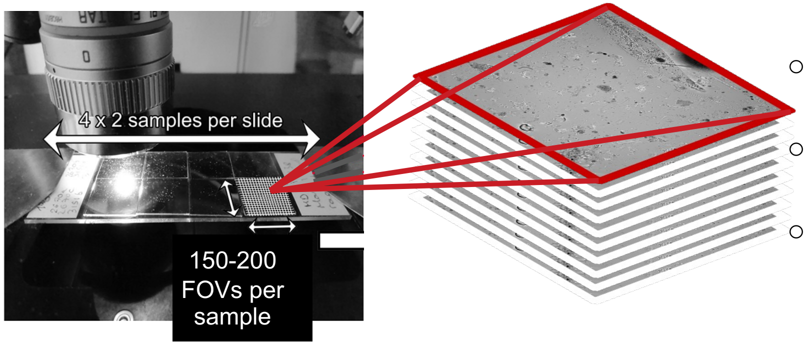


LB: Luc Beaufort (DR), CB: Clara Bolton (CR), BSM : Baptiste Sucheras-Marx (MdC)
TdGT: Thibault de Garidel-Thoron (CR), LL : Laetitia Licari (MdC)
DB: Doris Barboni (CR)

ITA: Yves Gally (IR) – Jusqu'en 2022

Context – Overview of the methods used at CEREGE

1) **Image acquistion**
2) Object detection
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images



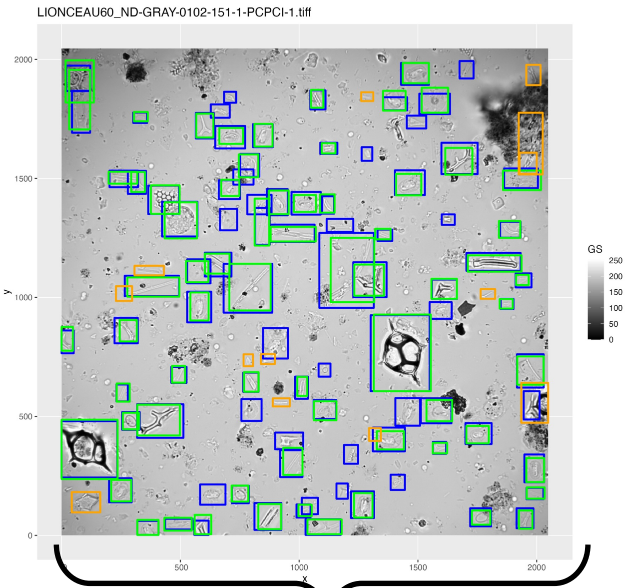
- Motorised optical microscope
- X630 – x1000 observations
- Polarized or brightfield lighting

Sediment samples → ***Image acquisition:***
- > 200 images per sample
- Z-stacking
- ~ 1h per sample
- Possible to preset 16 samples

→ **Images**

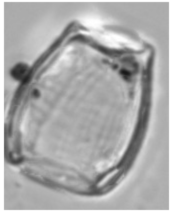
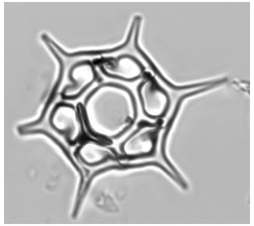
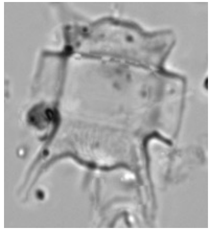
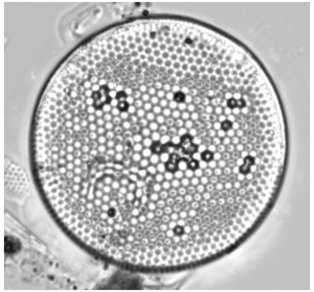
CONTEXT – OVERVIEW OF THE METHODS USED AT CEREGE

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images



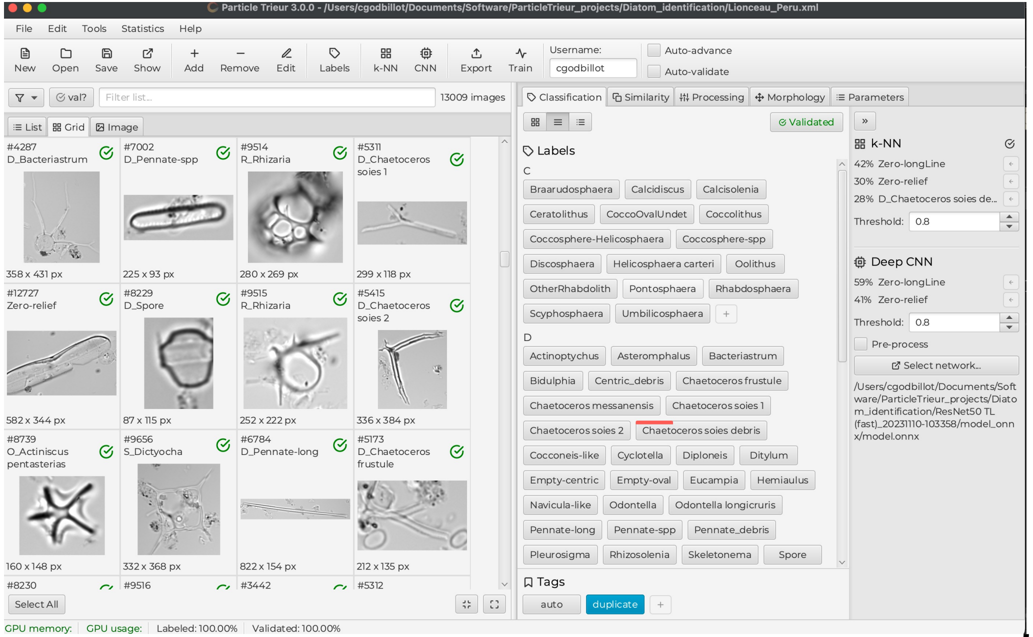
Images → ***Object detection:***
- Detection of regions of interest

→ **Crops**

CONTEXT – OVERVIEW OF THE METHODS USED AT CEREGE

1) Image acquistion
2) Object detection
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) **Object classification**
   - Image library construction
   - Model training
   - Inference on new images



ParticleTrieur

Crops → ***Object identification:***
- Use of CNNs to classify and extract morphological features

→ **Taxon counts + morphology**

Object detection protocol

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

Most recent workflow

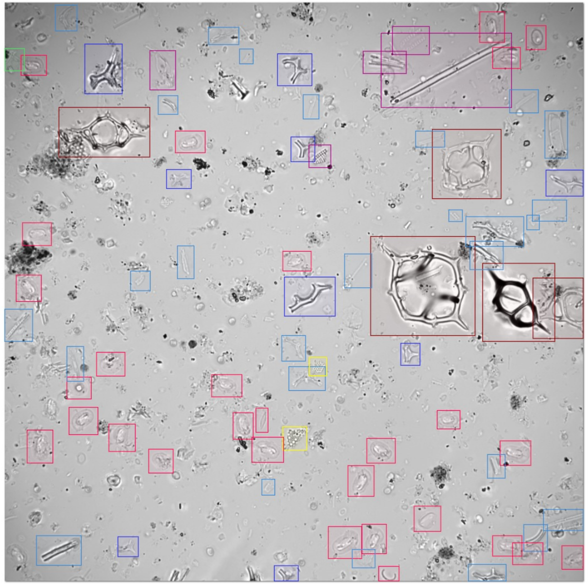
Preprint available on ESS Open Archive:

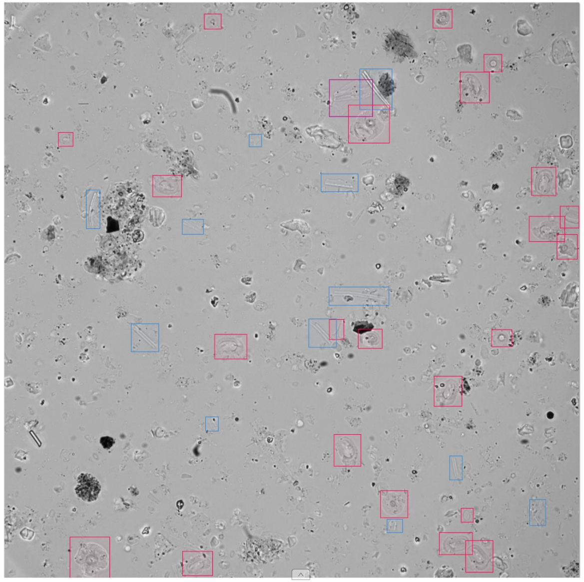
A new method for the detection of siliceous microfossils on sediment microscope slides using convolutional neural networks

OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
    - Image annotation using CVAT (Computer Vision Annotation Tool)
    - Model training
    - Inference on images
    - Cropping
3) Object classification
    - Image library construction
    - Model training
    - Inference on new images

1) Image annotation using CVAT (Computer Vision Annotation Tool):





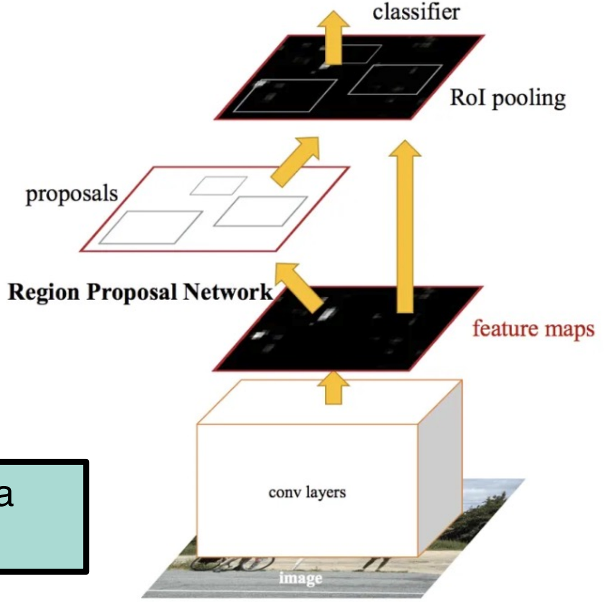
- 1 category "Microfossil" (merged all microfossils into a single category, no difference between silicoflagellates and diatoms for instance)
- 298 Images annotated (239 for training, 59 for testing): Mediterranean sediment + a sediment core from the coast of Peru
- 12 269 bounding boxes drawn

OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
    - Image annotation using CVAT (Computer Vision Annotation Tool)
    - Model training
    - Inference on images
    - Cropping
3) Object classification
    - Image library construction
    - Model training
    - Inference on new images

1) Image annotation using CVAT (Computer Vision Annotation Tool)
2) **Model training:**
    - Model:
        - ✓ Faster-RCNN with a ResNet50 backbone pre-trained on COCO
    - Augmentations:
        - ✓ Random horizontal flip
        - ✓ Random vertical flip
        - ✓ Brightness
        - ✓ Contrast
    - Implemented in PyTorch

Performed by a separate network, not a selective search algorithm



Faster R-CNN Model Architecture.Taken from: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2016.

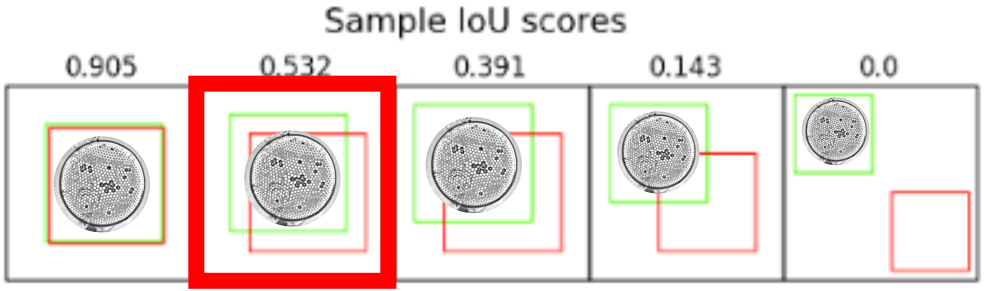
OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

1) Image annotation using CVAT (Computer Vision Annotation Tool)

2) **Model training:**

How to evaluate model performance ?



Ground-truth bounding box

Predicted bounding box

IoU = Area of Intersection of the two bboxes / Area of Union

→ In general, studies consider that the model has detetcted the ground-truth bbox when **the IoU > 0.5**

OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

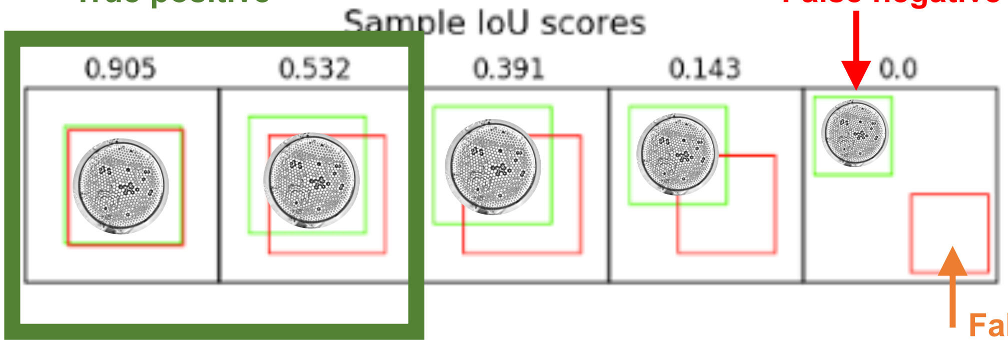
1) Image annotation using CVAT (Computer Vision Annotation Tool)

2) **Model training:**

How to evaluate model performance ?

True positive

False negative



False positive

Ground-truth bounding box

Predicted bounding box

IoU = Area of Intersection of the two bboxes / Area of Union

→ In general, studies consider that the model has detetcted the ground-truth bbox when **the IoU > 0.5**

→ Three cases can occur in **object detection:**

- True Positive : a **predicted** box has an IoU > 0.50 with a **ground-truth** box
- False Negative : a **ground-truth box** has **no corresponding predicted box** (IoU < 0.50)
- False Positive : a **predicted** box corresponds to **no ground-truth box** (IoU < 0.50)

OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

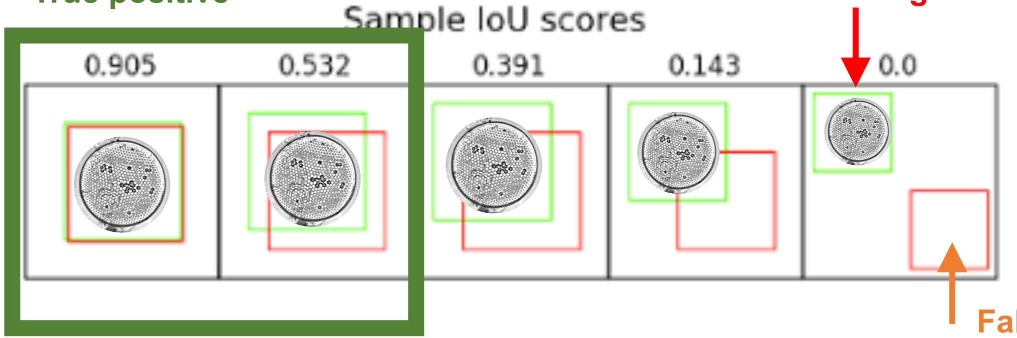
1) Image annotation using CVAT (Computer Vision Annotation Tool)

2) **Model training:**

How to evaluate model performance ?

True positive

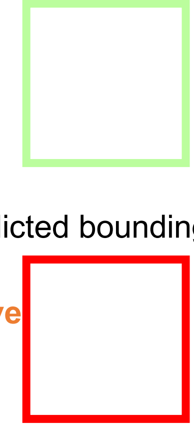
False negative



False positive

Ground-truth bounding box

Predicted bounding box



IoU = Area of Intersection of the two bboxes / Area of Union

To measure model perfomance, you can measure :

Mean Average Precision $= \frac{True\ Positive}{Trupe\ Positive + False\ Positive}$: is the model precise? Does it generate many false positives?

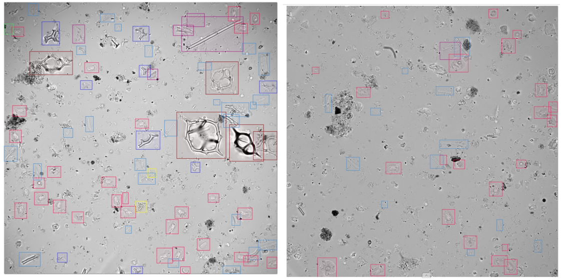
Mean Average Recall $= \frac{True\ Positive}{True\ Positive + False\ Negative}$: did the model find all the ground-truth boxes?

OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

1) Image annotation using CVAT (Computer Vision Annotation Tool)

2) Model training:



How to evaluate model performance ?

Precision

```
IoU metric: bbox
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=300 ] = 0.751
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=300 ] = 0.567
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=300 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=300 ] = 0.389
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=300 ] = 0.587
```

75% of the boxes that the model predicts actually correspond to ground-truth boxes

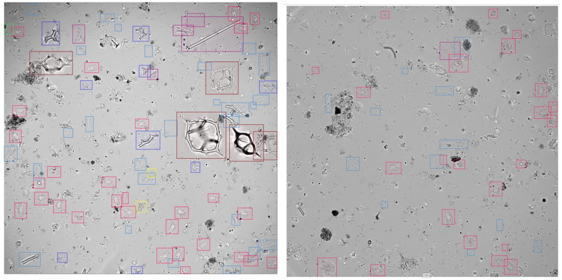
OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

1) Image annotation using CVAT (Computer Vision Annotation Tool)

2) **Model training:**

How to evaluate model performance ?



Recall

```
IoU metric: bbox
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.181
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 30 ] = 0.405
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=300 ] = 0.578
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=300 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=300 ] = 0.505
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=300 ] = 0.657
```

Over a range of IoUs, 58% of the ground-truth bounding boxes were found by the model

- the number increases if you only consider the large objects
- the value for recall is averaged over a range of different IoU thresholds

OBJECT DETECTION PROTOCOL

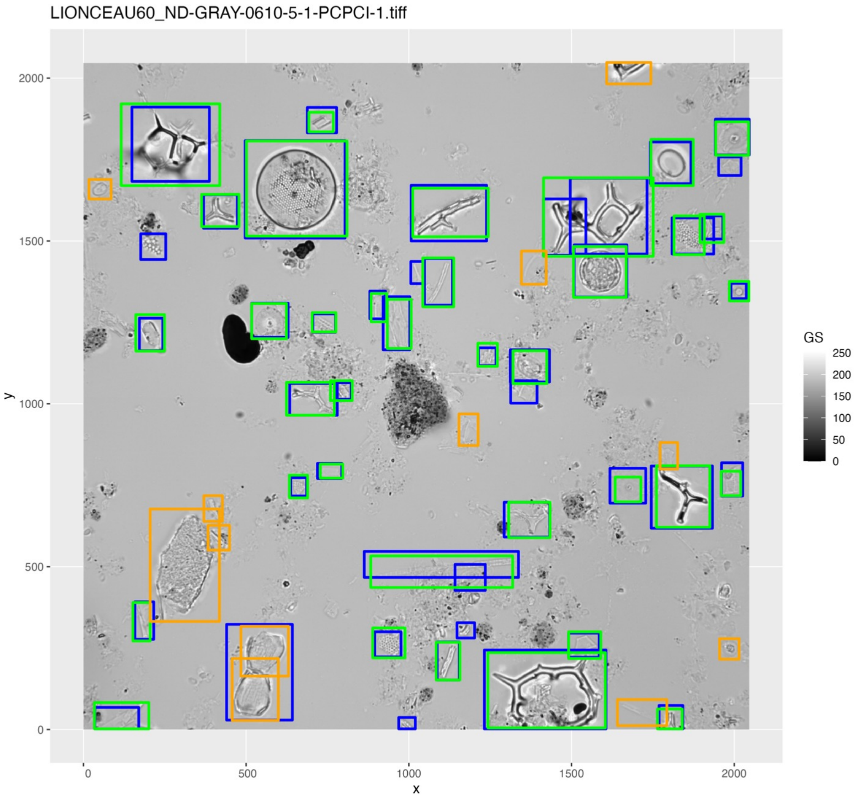
How does this translate on a test image set? **→ RECALL:**

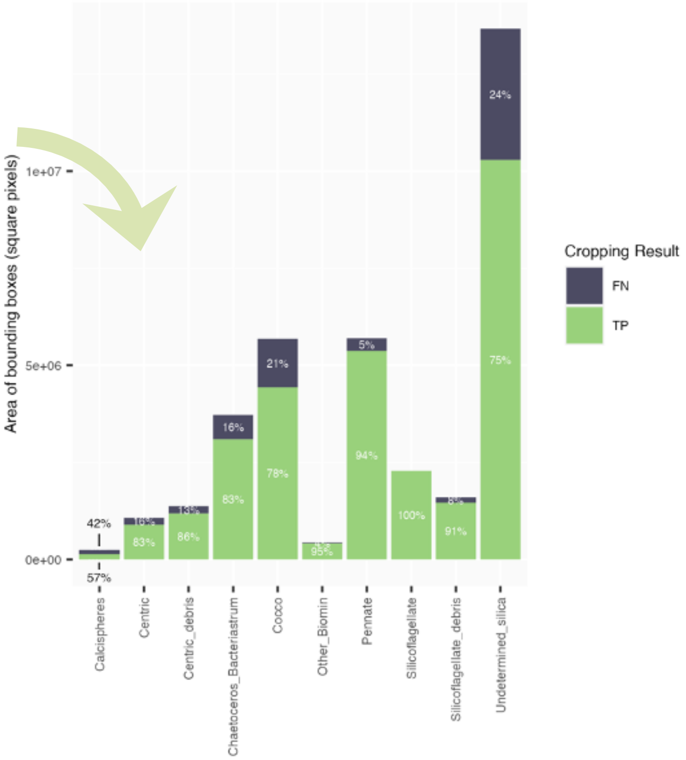
Ground-truth bbox

True positives

False positives

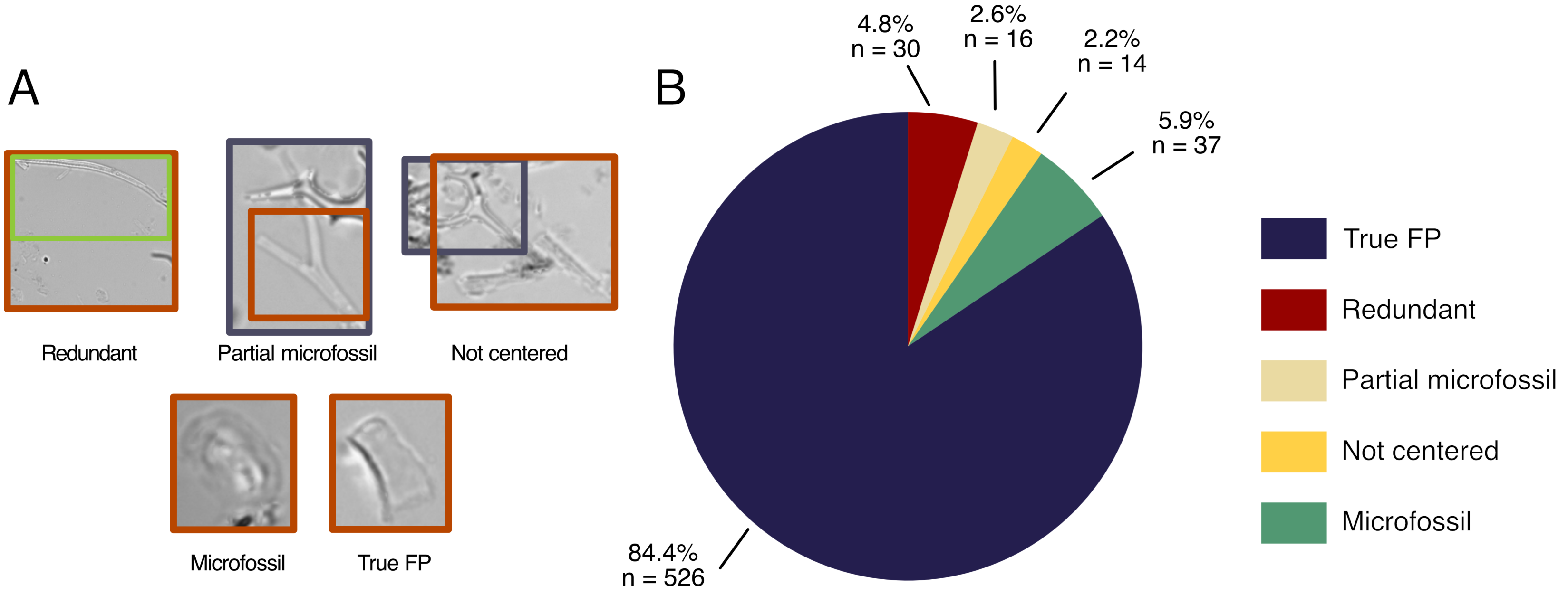
Godbillot et al. (submitted)





OBJECT DETECTION PROTOCOL

How does this translate on a test image set? **→ PRECISION:**



OBJECT DETECTION PROTOCOL

1) Image acquistion
2) **Object detection**
   - Image annotation using CVAT (Computer Vision Annotation Tool)
   - Model training
   - Inference on images
   - Cropping
3) Object classification
   - Image library construction
   - Model training
   - Inference on new images

ADVANTAGES

- **Perform inference and crop 200 images in 30s** → very efficient
- **Does not miss rare species**

→ **Now possible to study automatically** the production of different phytoplanction groups in the same sample

DRAWBACKS

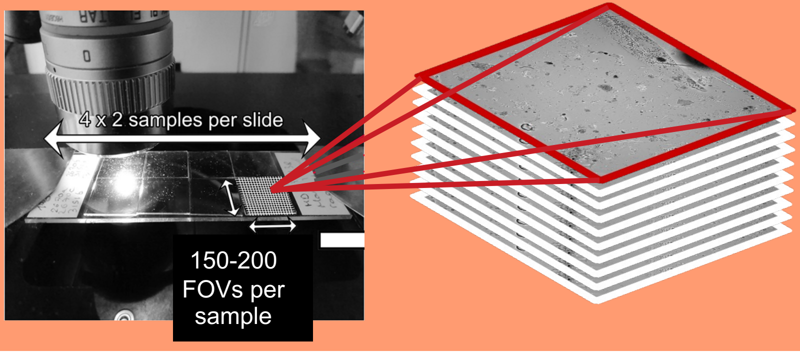
- **Classification step is independent of detection step** (because of rare species)
  - → doubles the efforts to build an image library and train models
- **No real-time cropping possible**
  - → requires memory to store the images
  - → can only work on the “hyperfocused” image → loss of information → cannot go down to species level.
- **Capacity for generalization to other types of plankton images (filters etc.) is limited**

Ongoing projects and perspectives

Short-term:
Compare Faster R-CNN and YOLO results

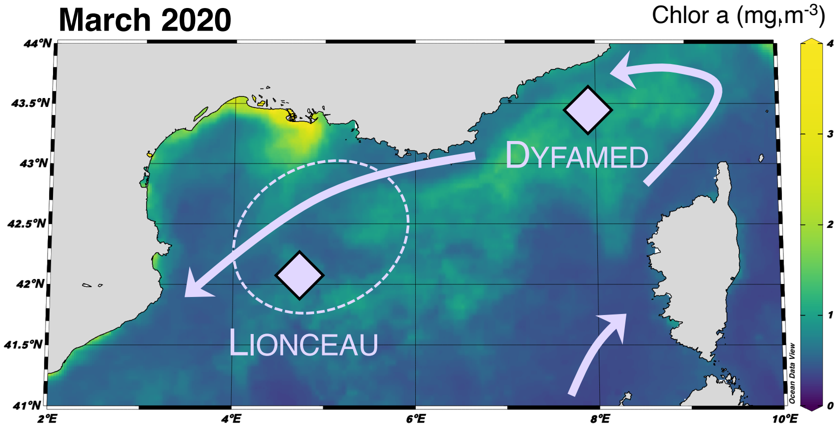
Medium-term:
Explore ways to build “artificial slides” to increase rare species → build new library to perform **detection + classification in a single step**

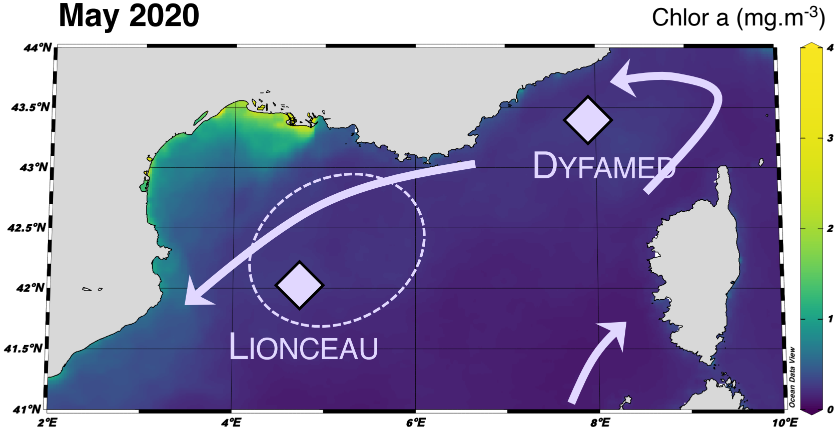
Long-term:
Work on the image stack in real time → use entire stack to extract more information → species (but loss of images if ever models are improved after acquisition)



Example results from the Mediterranean Sea

Applications



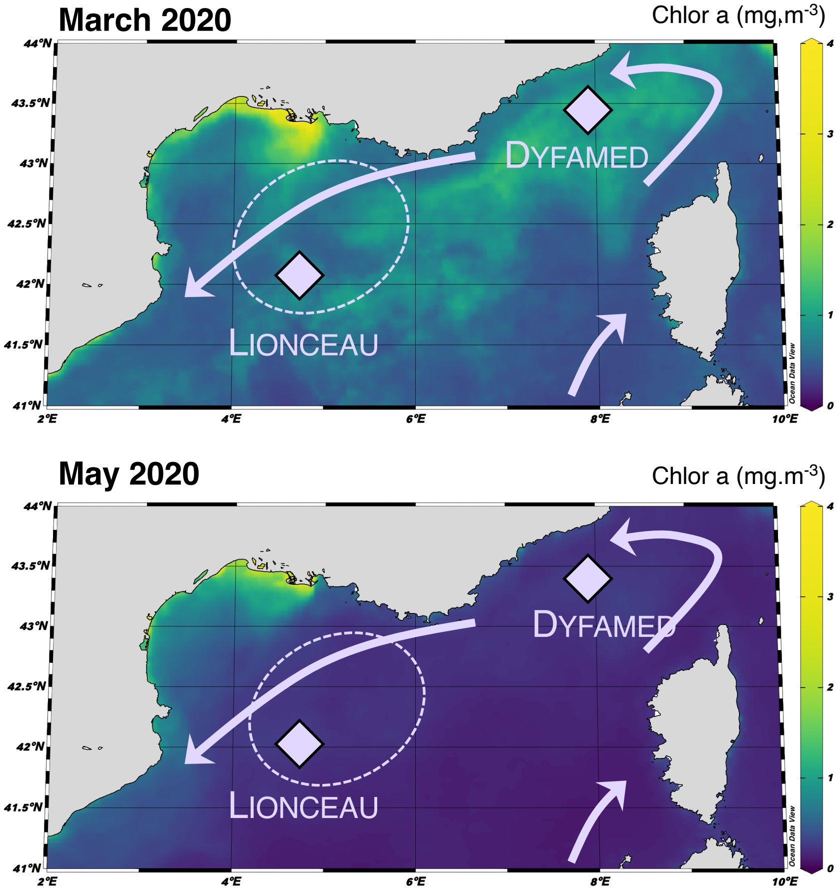


www.oceancolour.org

Two sediment trap series from the NW Mediterranean (2010-2020); approx. 2 weeks per point

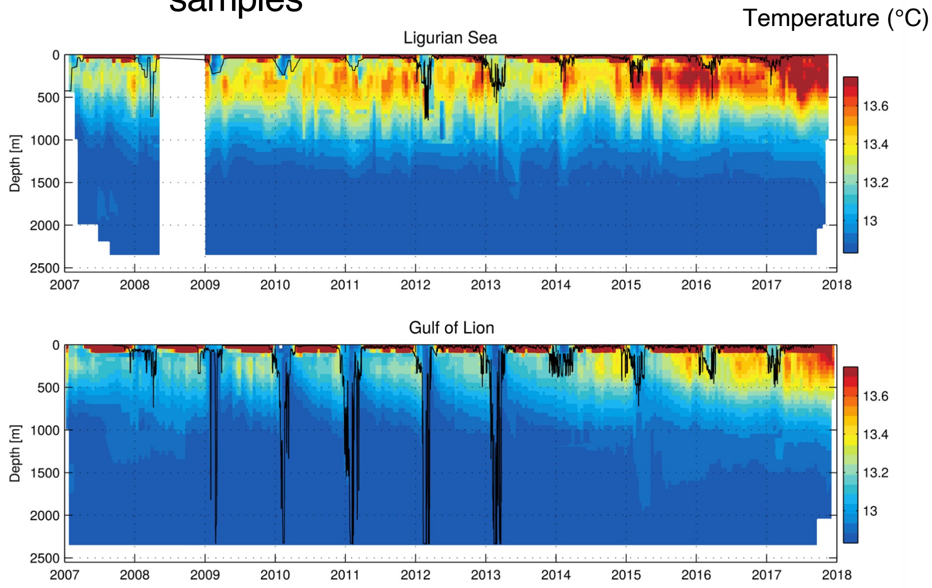
- Dyfamed (Ligurian Sea; 1000m depth), 246 samples
- Lionceau (Gulf of Lion; 2400m depth), 80 samples

APPLICATIONS



Two sediment trap series from the NW Mediterranean (2010-2020); approx. 2 weeks per point

- Dyfamed (Ligurian Sea; 1000m depth), 246 samples
- Lionceau (Gulf of Lion; 2400m depth), 80 samples



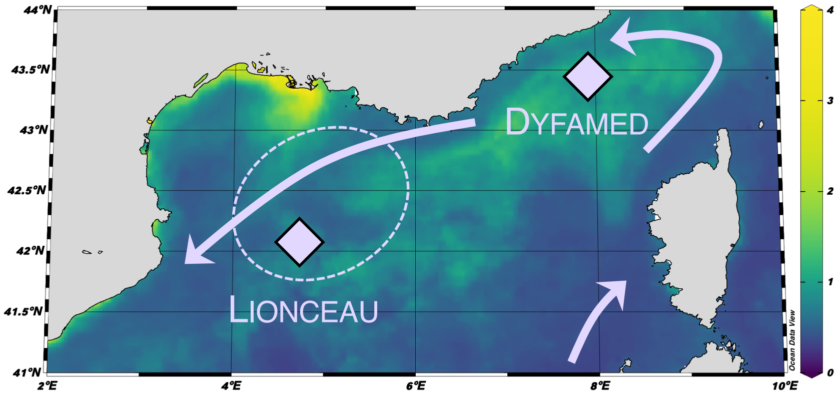
Margirier et al. 2020

www.oceancolour.org

Applications

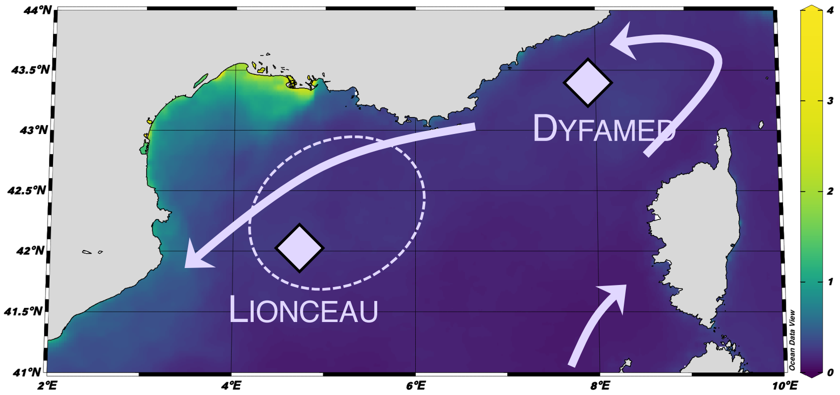
March 2020

Chlor a ($mg.m^{-3}$)



May 2020

Chlor a ($mg.m^{-3}$)

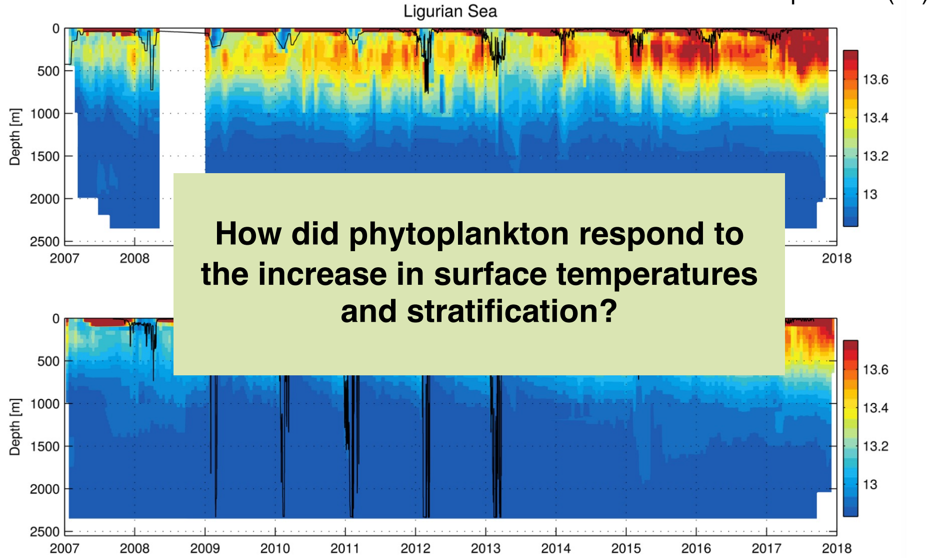


www.oceancolour.org

Two sediment trap series from the NW Mediterranean (2010-2020); approx. 2 weeks per point

- Dyfamed (Ligurian Sea; 1000m depth), 246 samples
- Lionceau (Gulf of Lion; 2400m depth), 80 samples

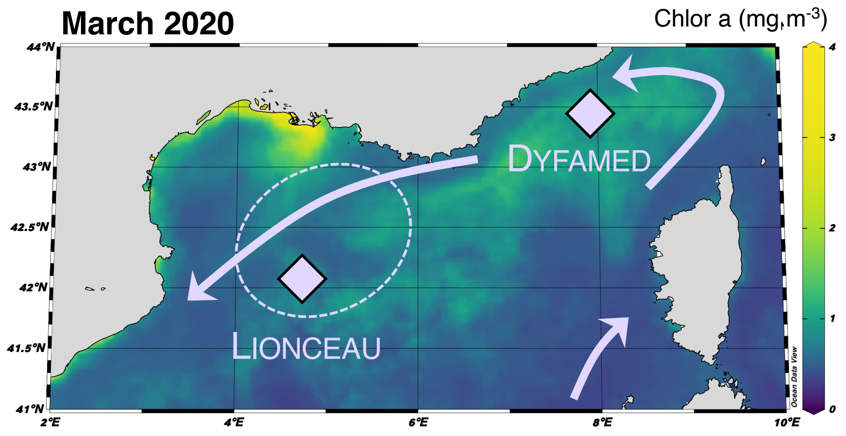
Temperature (°C)

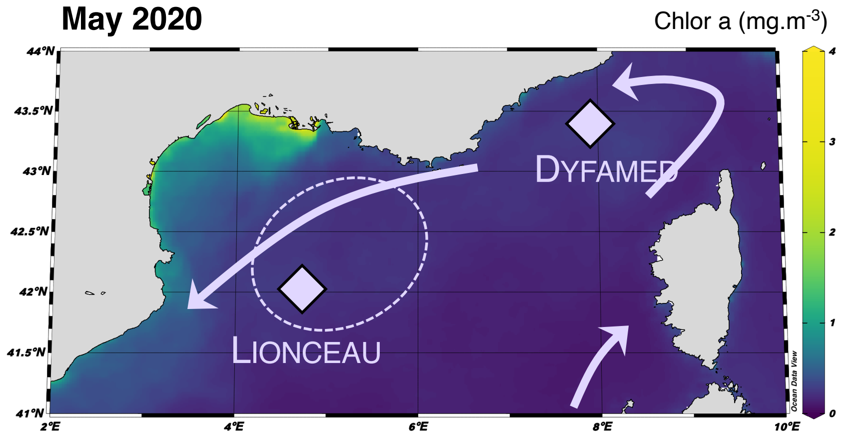


How did phytoplankton respond to the increase in surface temperatures and stratification?

Margirier et al. 2020

APPLICATIONS

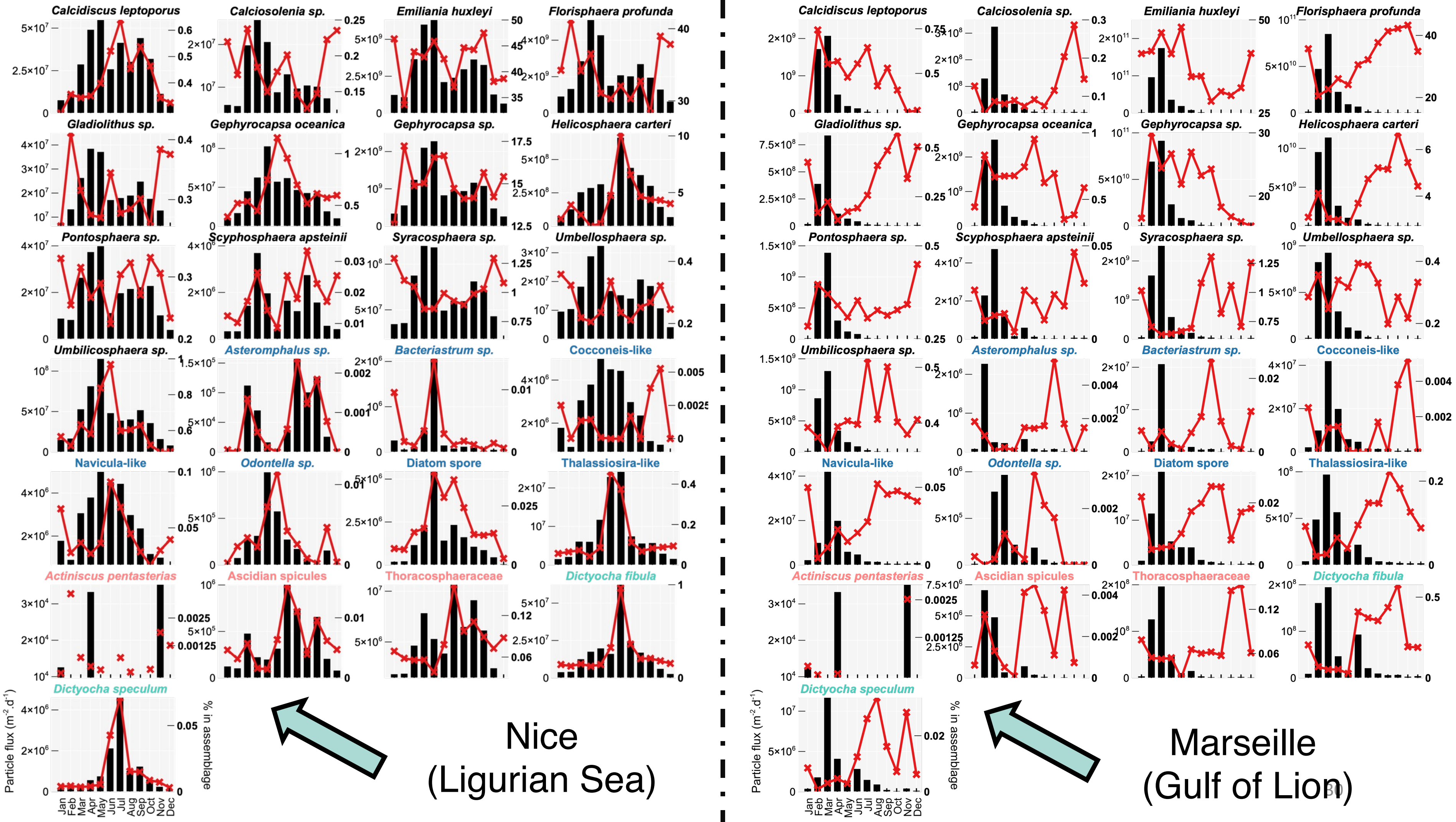


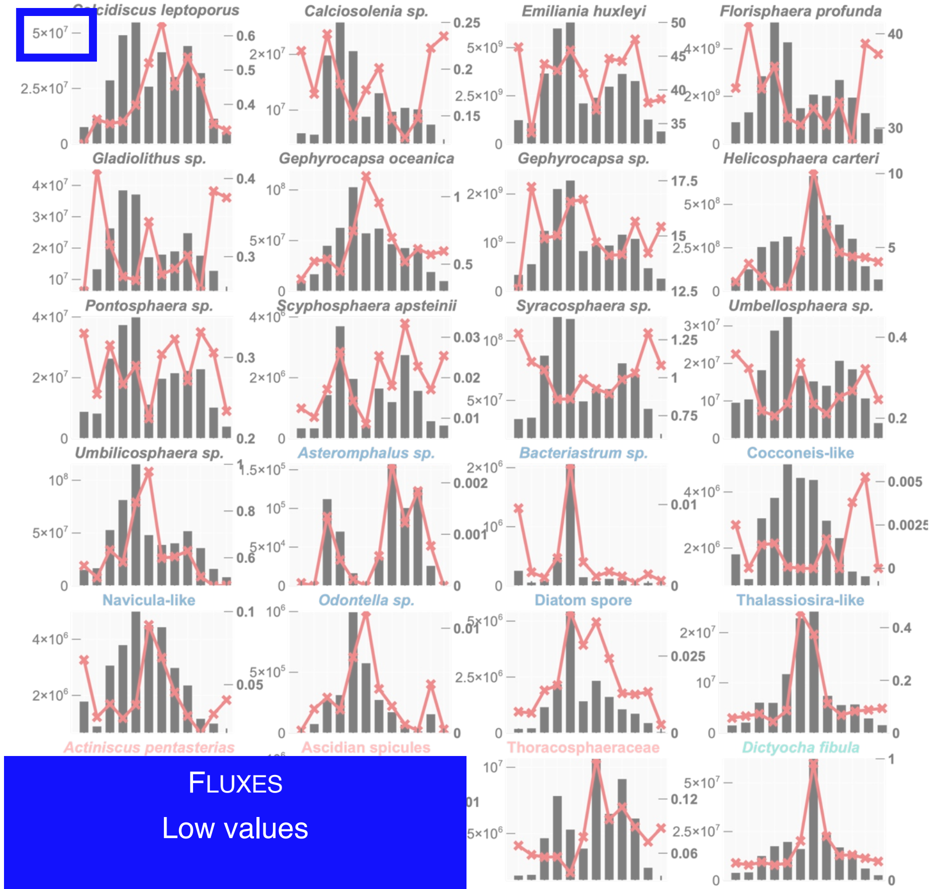


www.oceancolour.org

For the purpose of this study

- > 90 000 images
- A couple million particles sorted
- Several thousand plankton remains counted

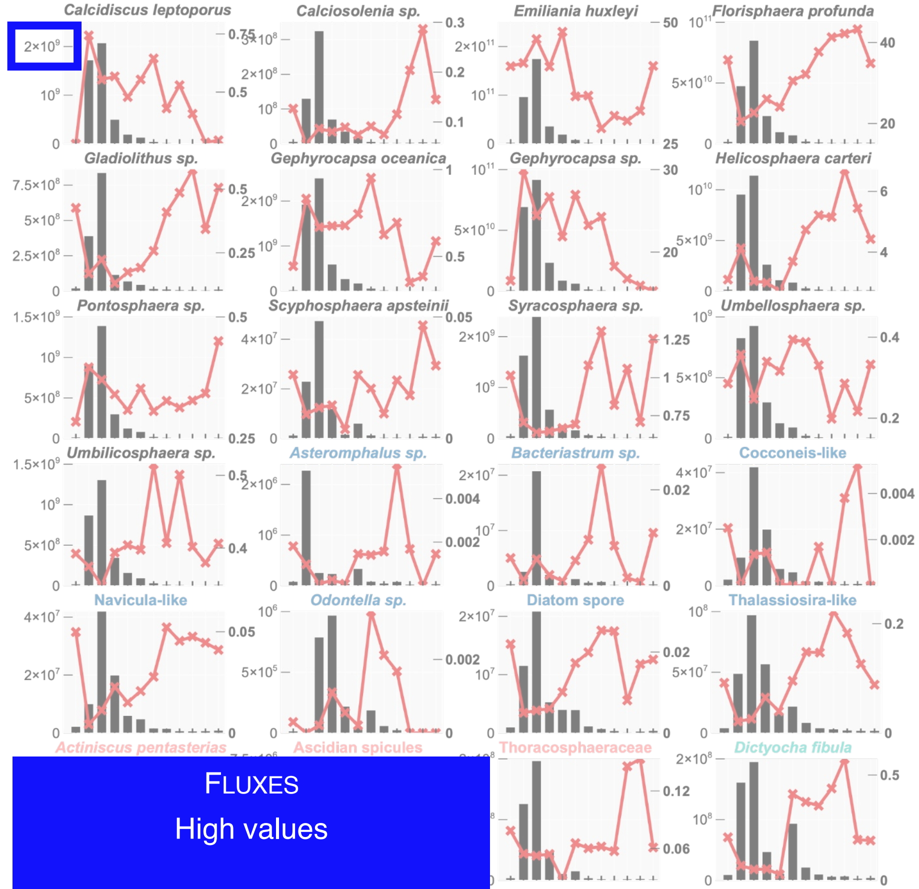

Calcidiscus leptoporus
Calciosolenia sp.
Emiliania huxleyi
Florisphaera profunda
Gladiolithus sp.
Gephyrocapsa oceanica
Gephyrocapsa sp.
Helicosphaera carteri
Pontosphaera sp.
Scyphosphaera apsteinii
Syracosphaera sp.
Umbellosphaera sp.
Umbilicosphaera sp.
Asteromphalus sp.
Bacteriastrum sp.
Cocconeis-like
Navicula-like
Odontella sp.
Diatom spore
Thalassiosira-like
Actiniscus pentasterias
Ascidian spicules
Thoracosphaeraceae
Dictyocha fibula
Dictyocha speculum
Particle flux (m-2.d-1)
% in assemblage
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
Nice
(Ligurian Sea)
Marseille
(Gulf of Lion)



FLUXES

Low values

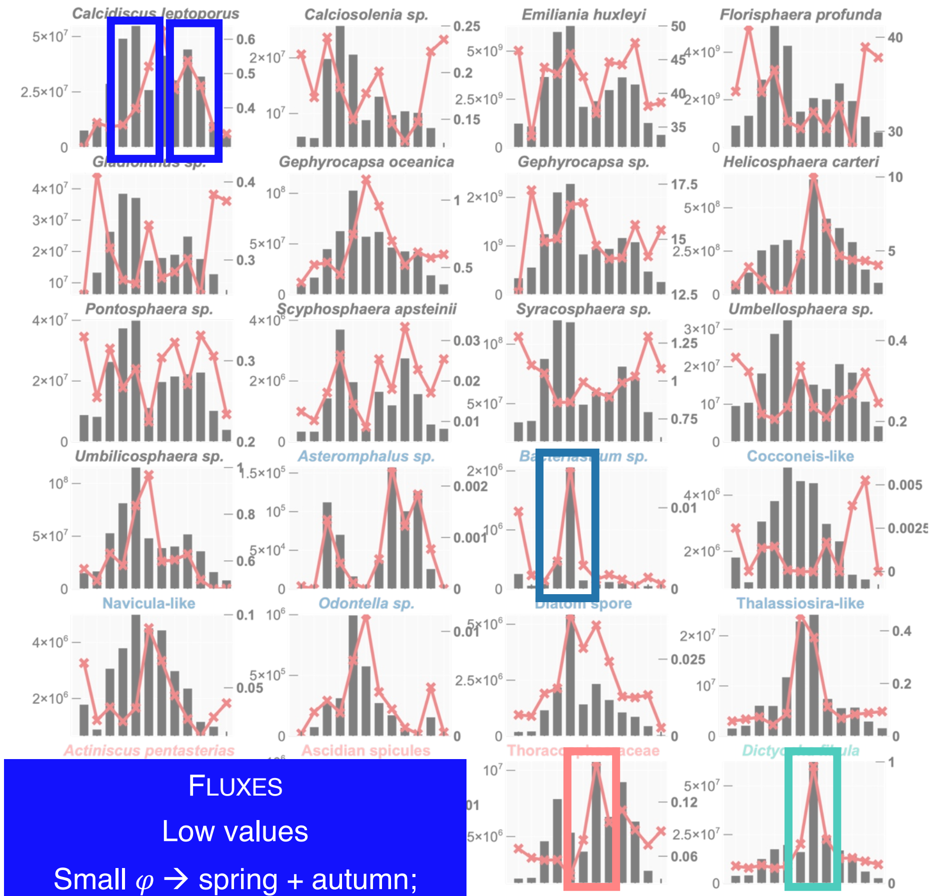
Nice
(Ligurian Sea)



FLUXES

High values

Marseille
(Gulf of Lion)

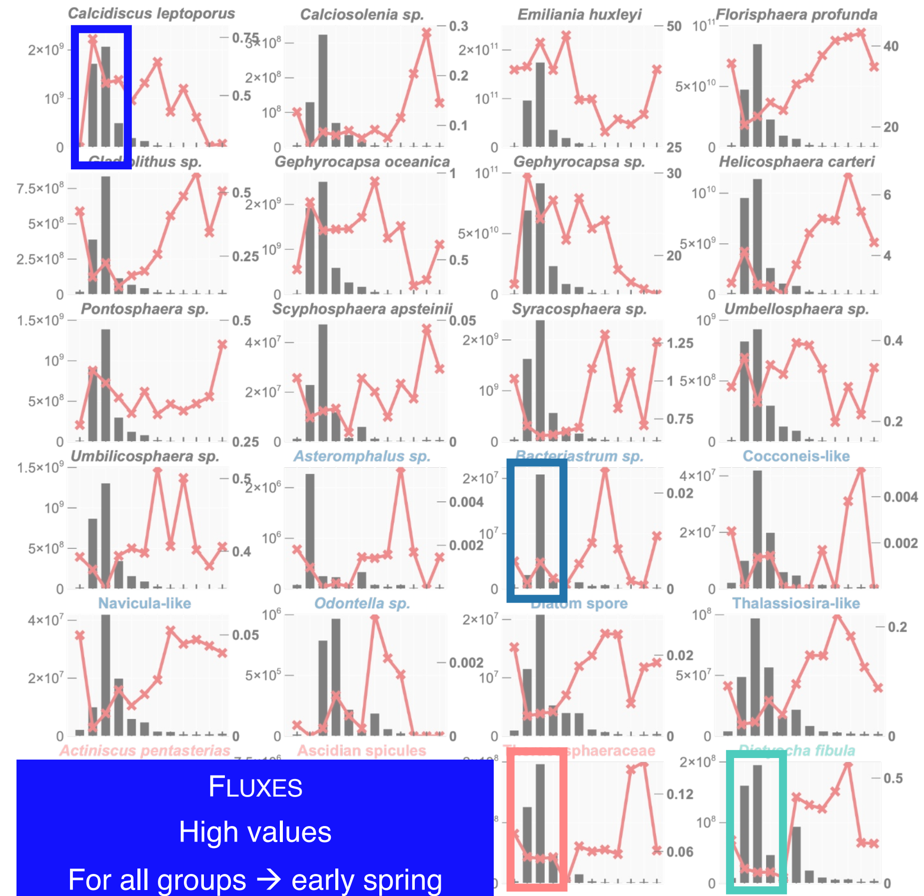


FLUXES

Low values

Small φ → spring + autumn;
large(r) φ → summer

Nice (Ligurian Sea)



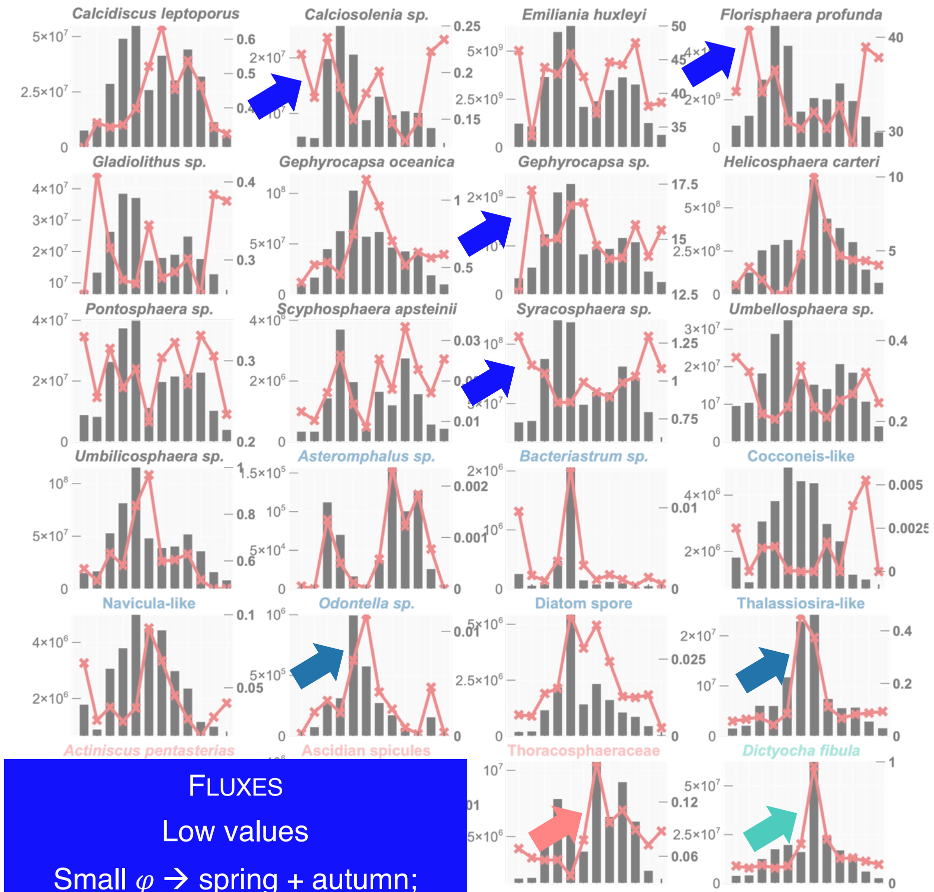
FLUXES

High values

For all groups → early spring

PERCENTAGES

Marseille (Gulf of Lion)



FLUXES

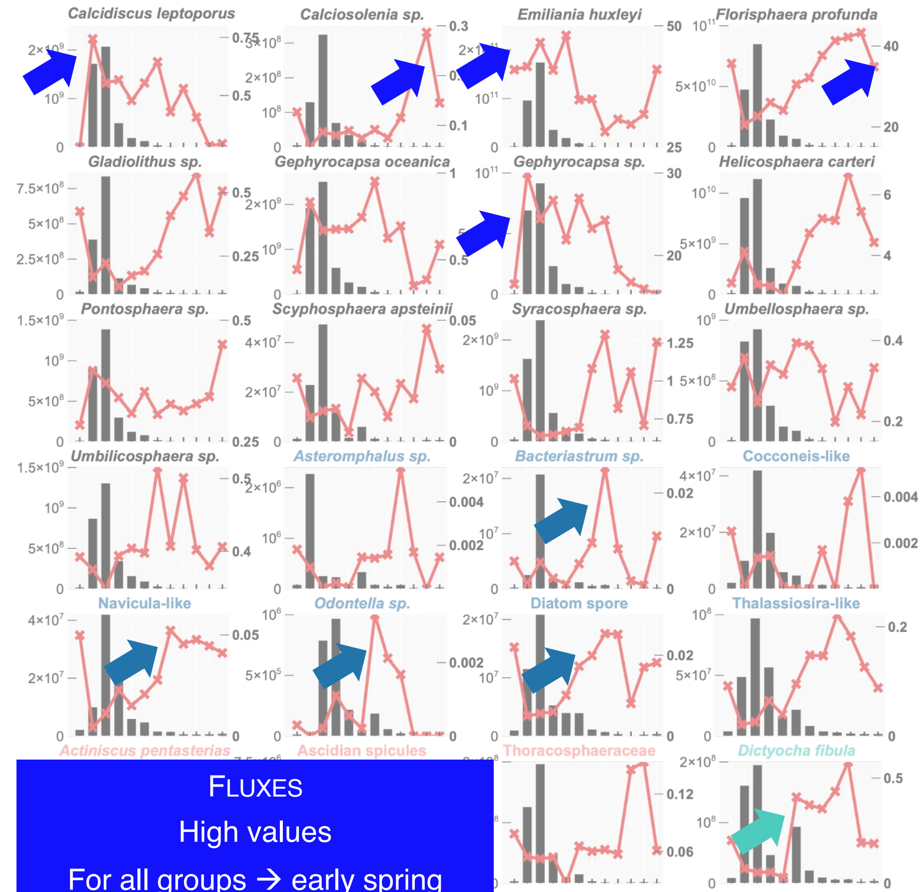
Low values

Small φ → spring + autumn;
large(r) φ → summer

PERCENTAGES

Small φ → winter + spring;
large(r) φ → summer + autumn

Nice
(Ligurian Sea)



FLUXES

High values

For all groups → early spring

PERCENTAGES

Small φ → winter + spring;
large(r) φ → summer + autumn

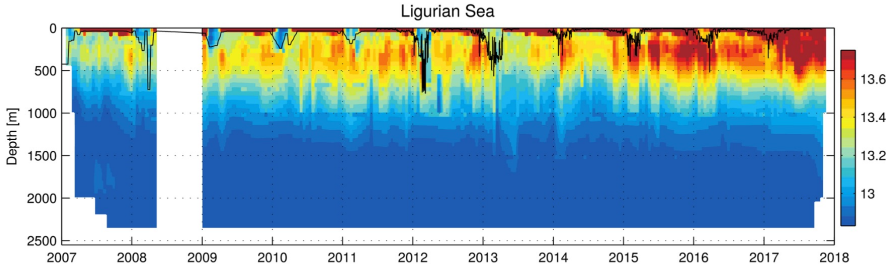
Marseille
(Gulf of Lion)

APPLICATIONS

FLUXES

Low values

Nice
(Ligurian Sea)

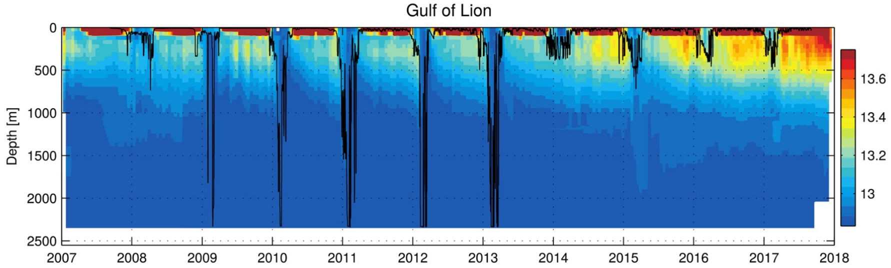


Margirier et al. 2020

FLUXES

High values

Marseille
(Gulf of Lion)



Margirier et al. 2020

Episodes of **deep convection** are important for the transport of material to the sea floor, and, by extension, for carbon storage at depth

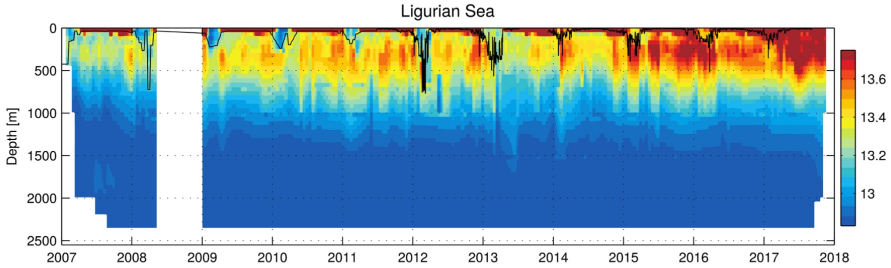
APPLICATIONS

FLUXES

Low values

Small φ → spring + autumn;
large(r) φ → summer

Nice (Ligurian Sea)



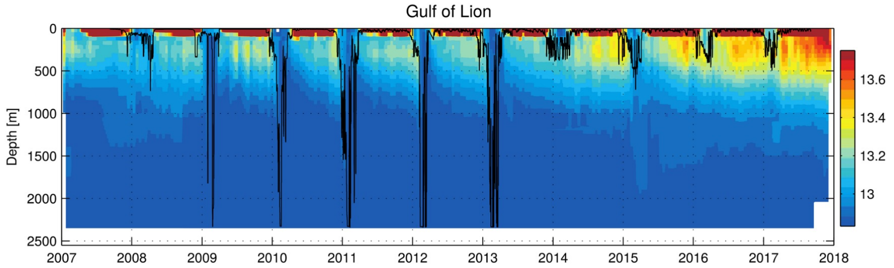
Margirier et al. 2020

FLUXES

High values

For all groups → early spring

Marseille (Gulf of Lion)



Margirier et al. 2020

Episodes of **deep convection** are important for the transport of material to the sea floor, and, by extension, for carbon storage at depth

In the Gulf of Lion, phytoplankton fluxes reflect deep water convection episodes. In the Ligurian Sea, they could also reflect, in part, the biological activity in the surface ocean

Applications

Fluxes

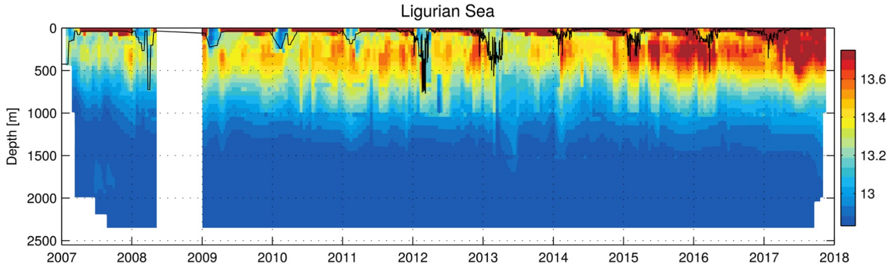
Low values

Small φ → spring + autumn;
large(r) φ → summer

Percentages

Small φ → winter + spring;
large(r) φ → summer + autumn

Nice (Ligurian Sea)



Margirier et al. 2020

Fluxes

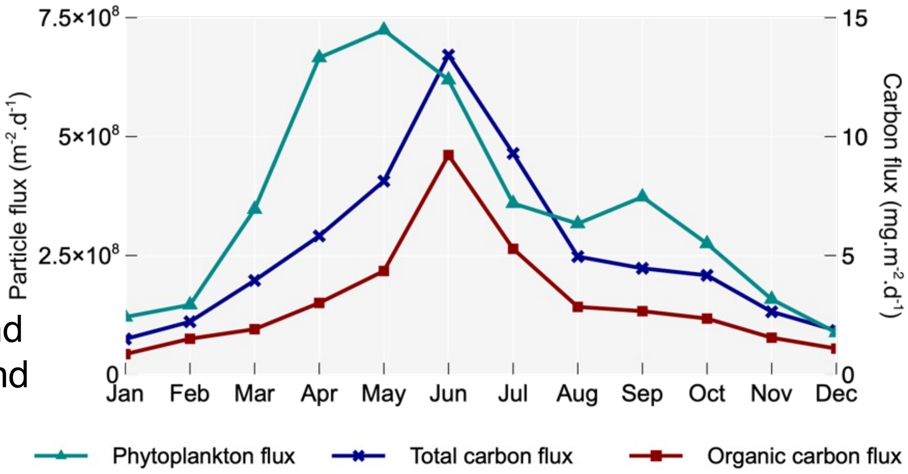
High values

For all groups → early spring

Percentages

Small φ → winter + spring;
large(r) φ → summer + autumn

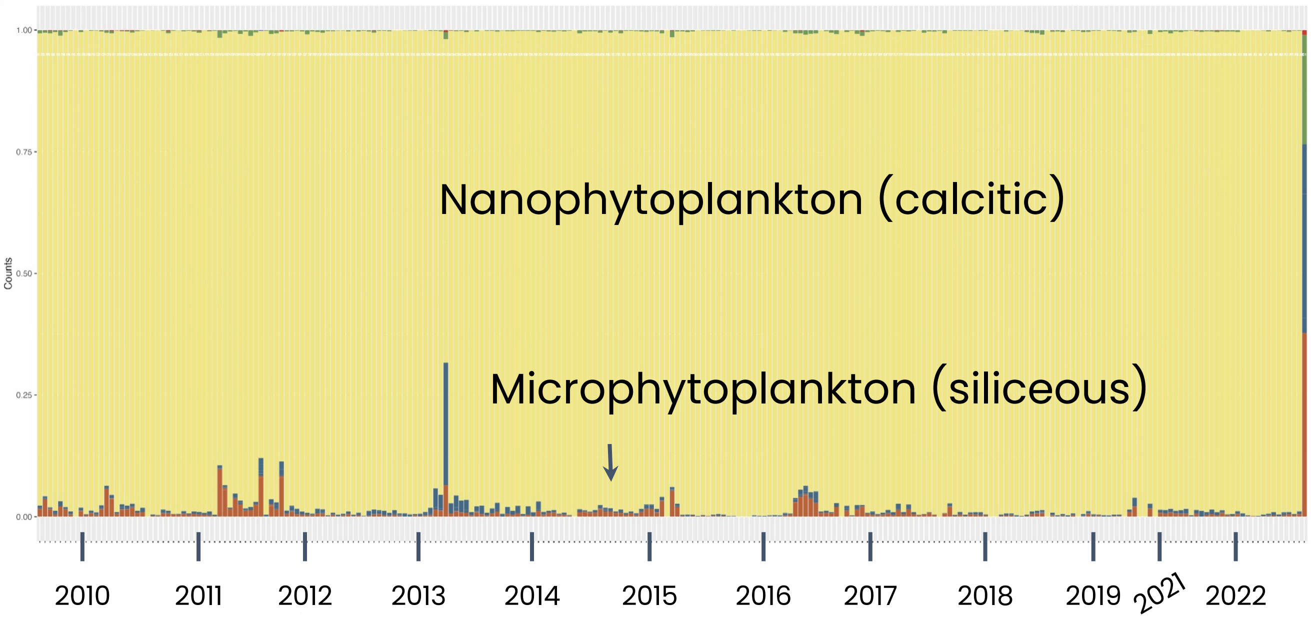
Marseille (Gulf of Lion)



Overall, while small phytoplankton dominate in winter and spring, larger phytoplankton dominates in the summer and fall:

→ Can this explain why carbon burial in the Ligurian Sea is highest in the summer?

Ligurian Sea example:



What is the impact of environmental change over the last decade;
If microphytoplankton decreases in the assemblage, will this impact carbon storage in the Mediterranean?

CONCLUSION

Use of AI in plankton studies:

Pros:

- Once trained, detection and classification workflows can **process data much more efficiently** than an expert
- Are a useful means of obtaining **standardized time-series** (*i.e.* results will not depend on who the observer was)
- Make it possible for **untrained researchers** to obtain high resolution time-series → increase number of observations

Cons:

- Training libraries are tedious to obtain
- Problems with precision when it comes to rare species
- Capacity for generalization to other types of plankton images is limited → image libraries are instrument-specific