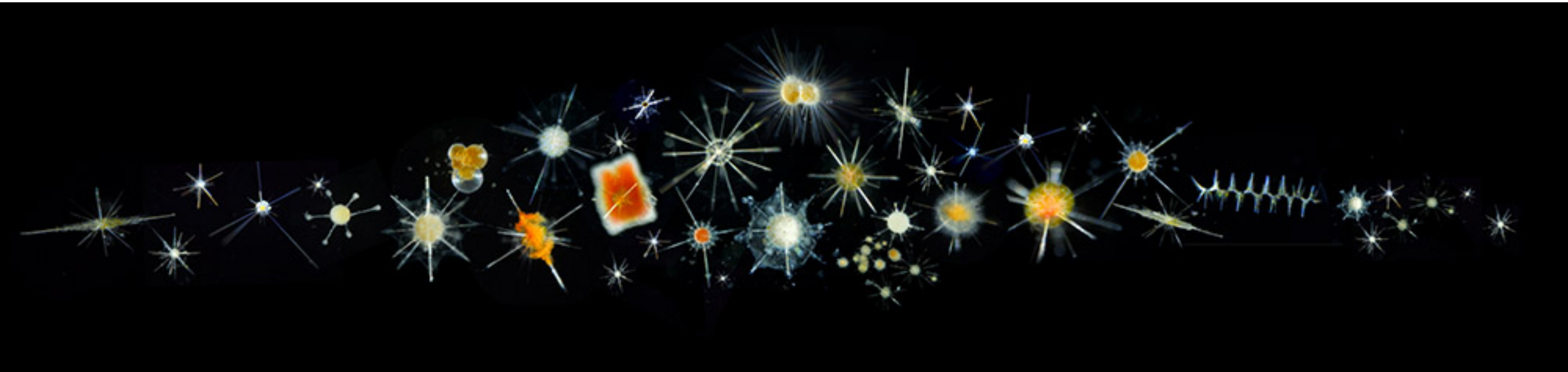
European Marine Board, 2022-12-15
Jean-Olivier Irisson (with input from many colleagues!)

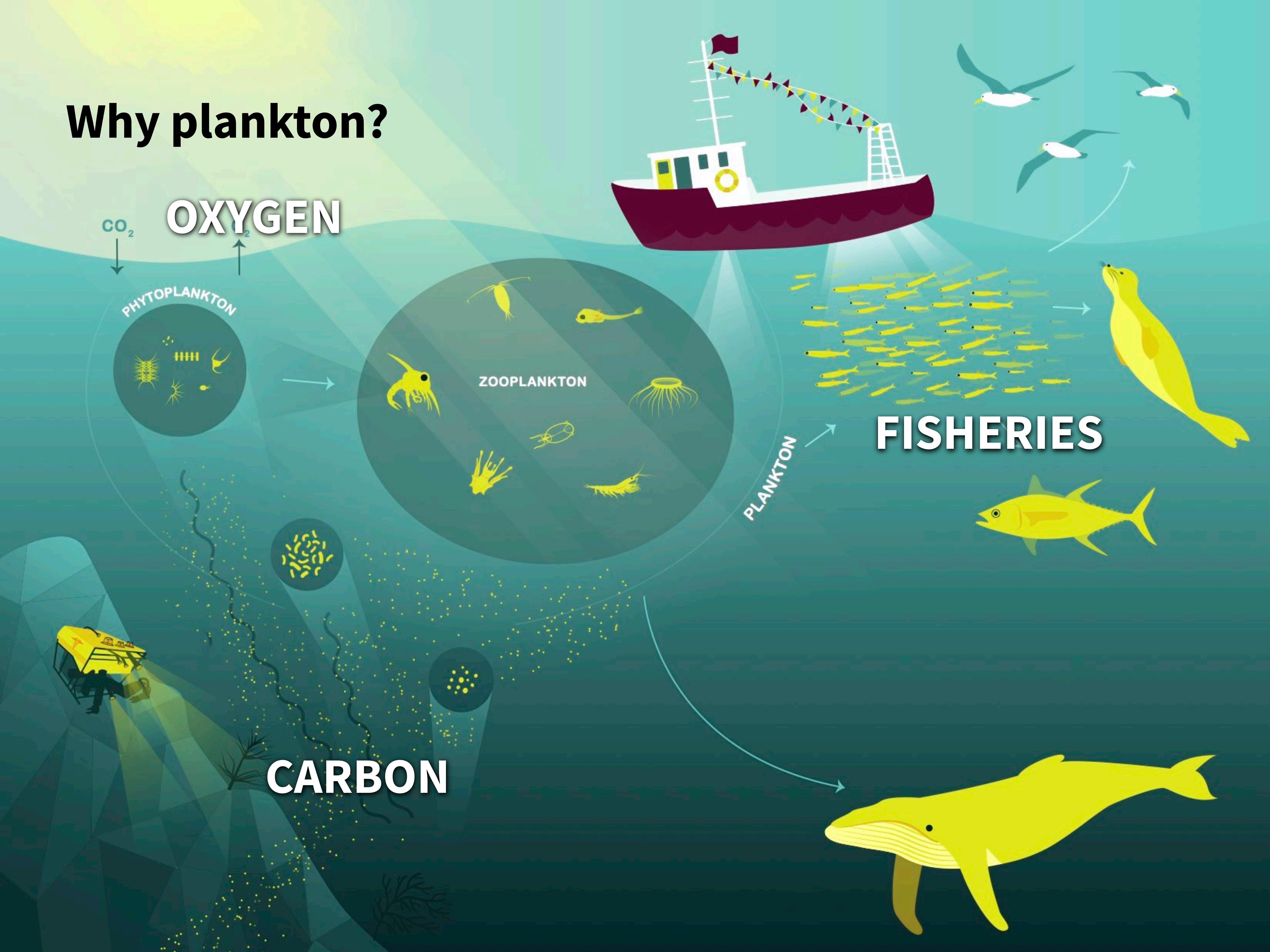
SCIENCES SORBONNE UNIVERSITÉ

Machine learning to unlock the potential of plankton imaging big data

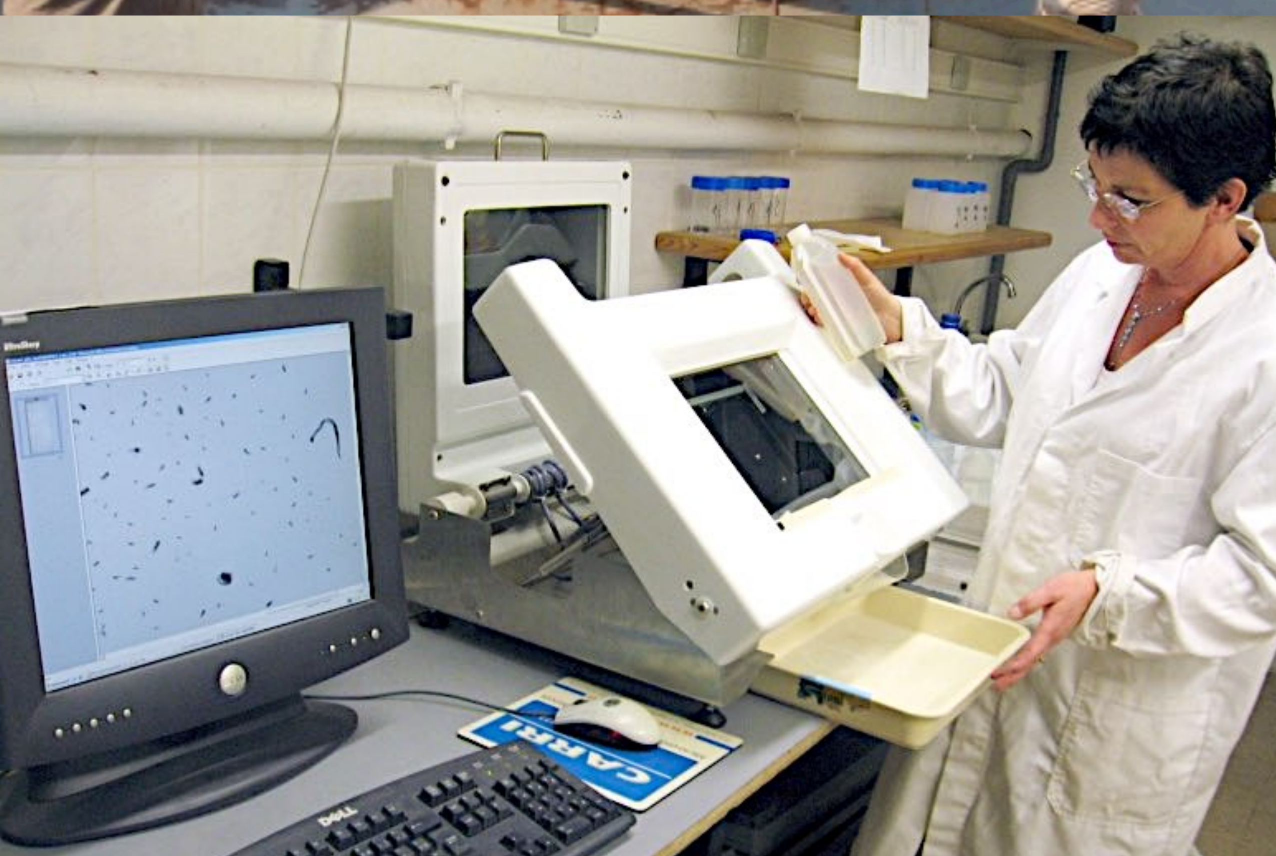
From pictures to knowledge

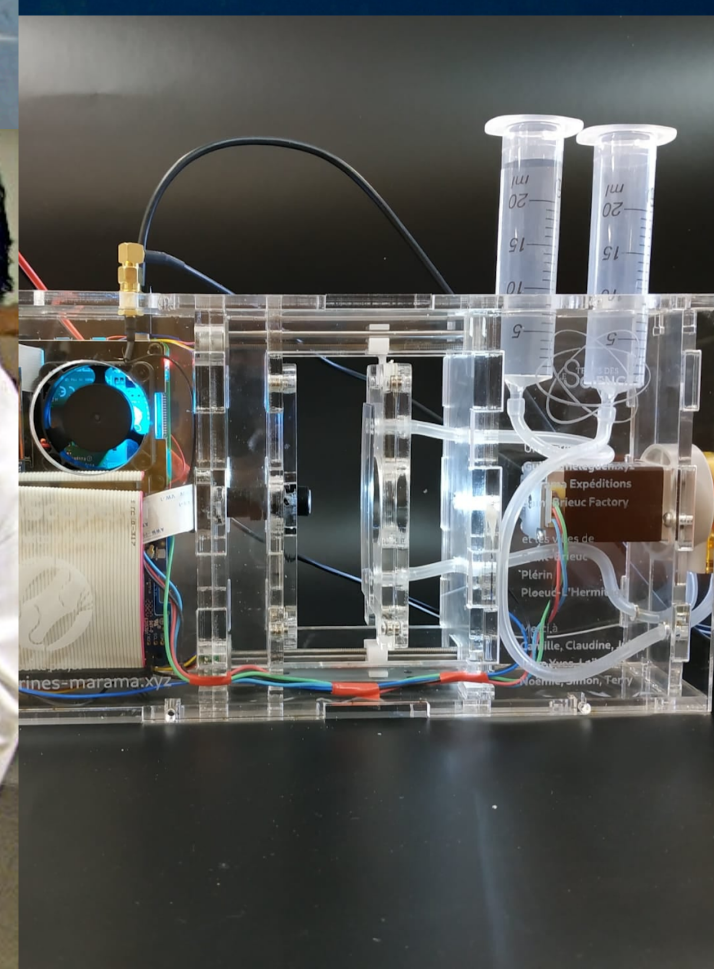


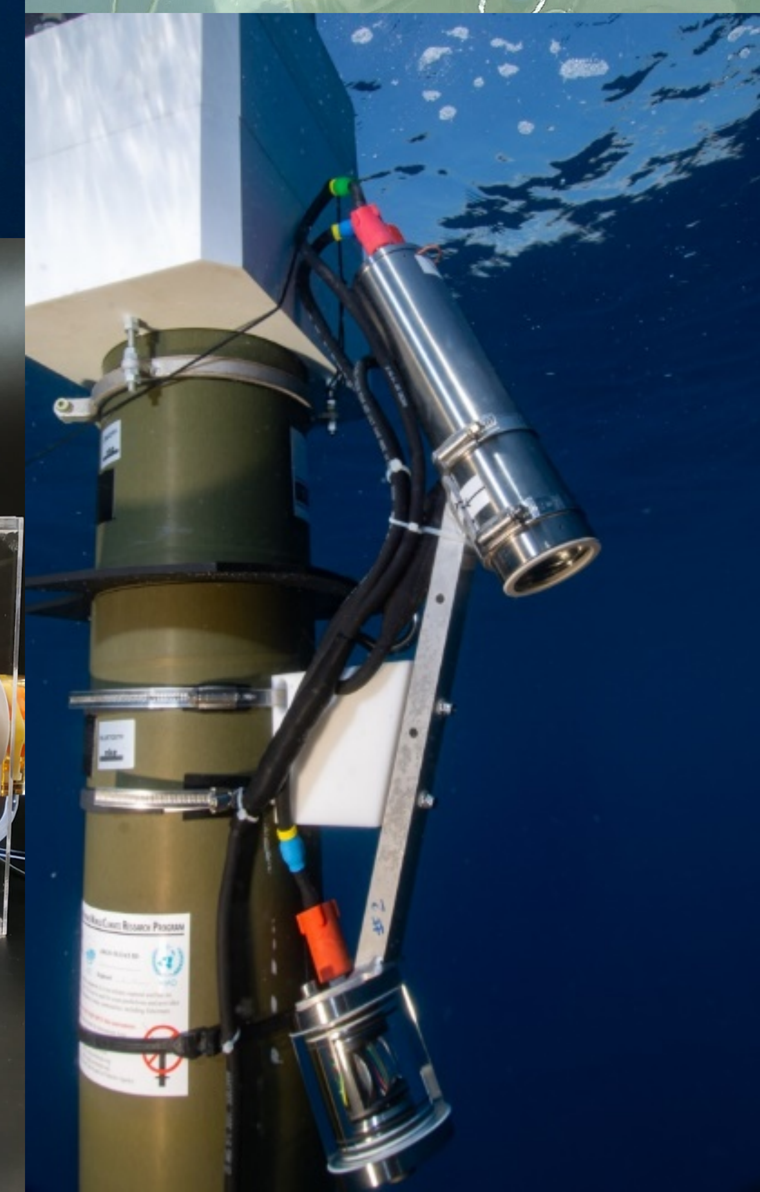
Why plankton?

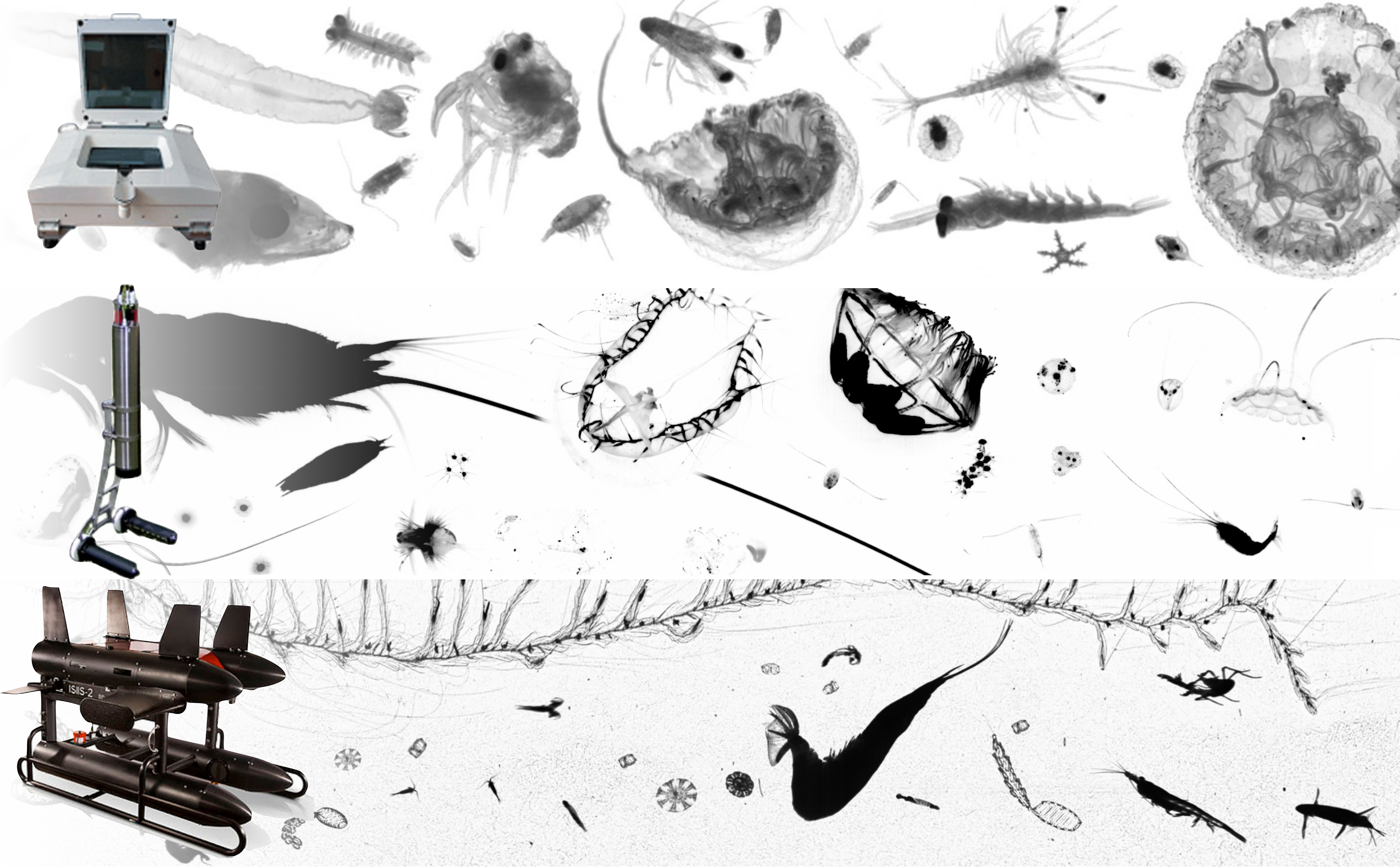


Many instruments





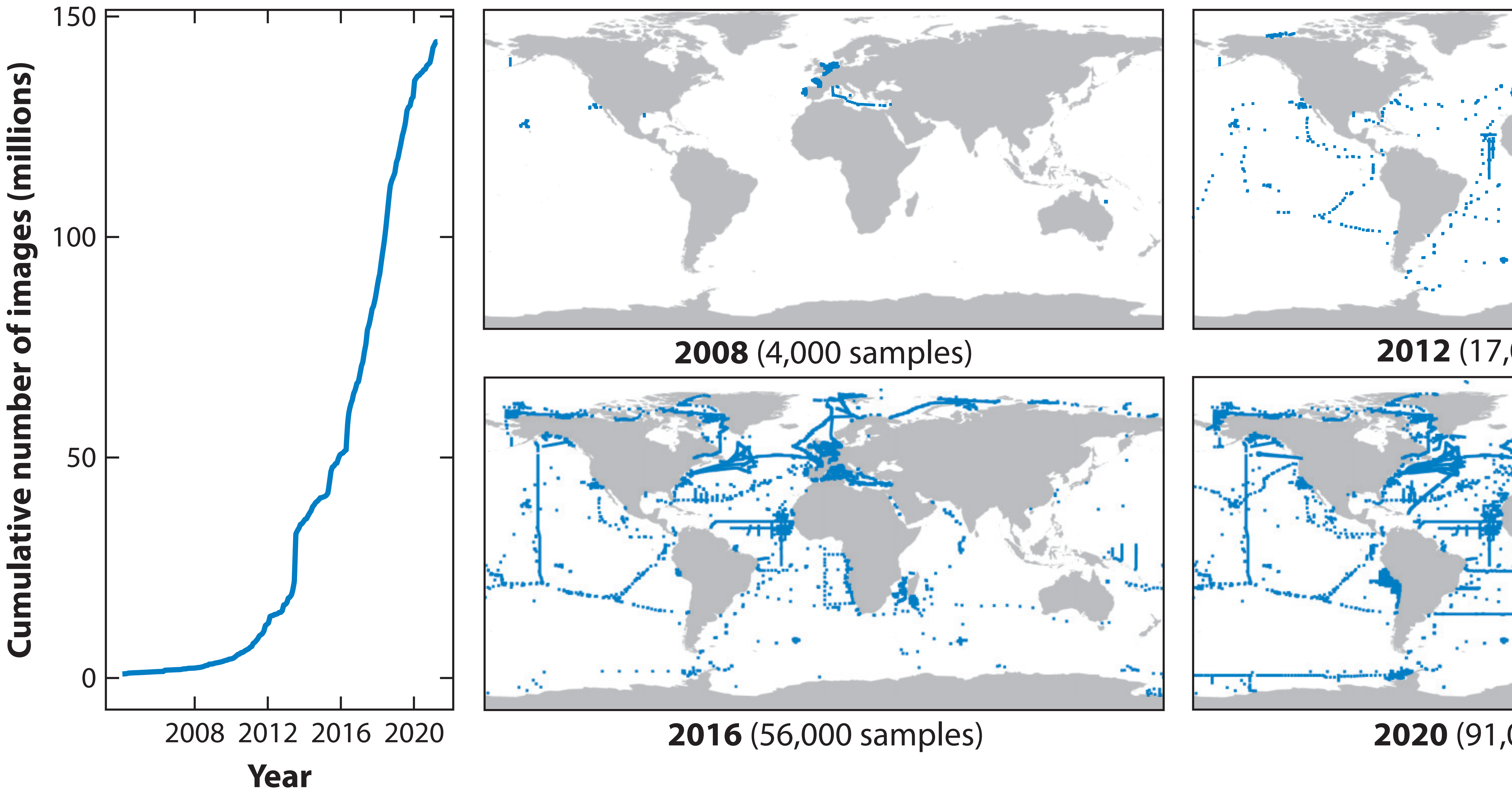




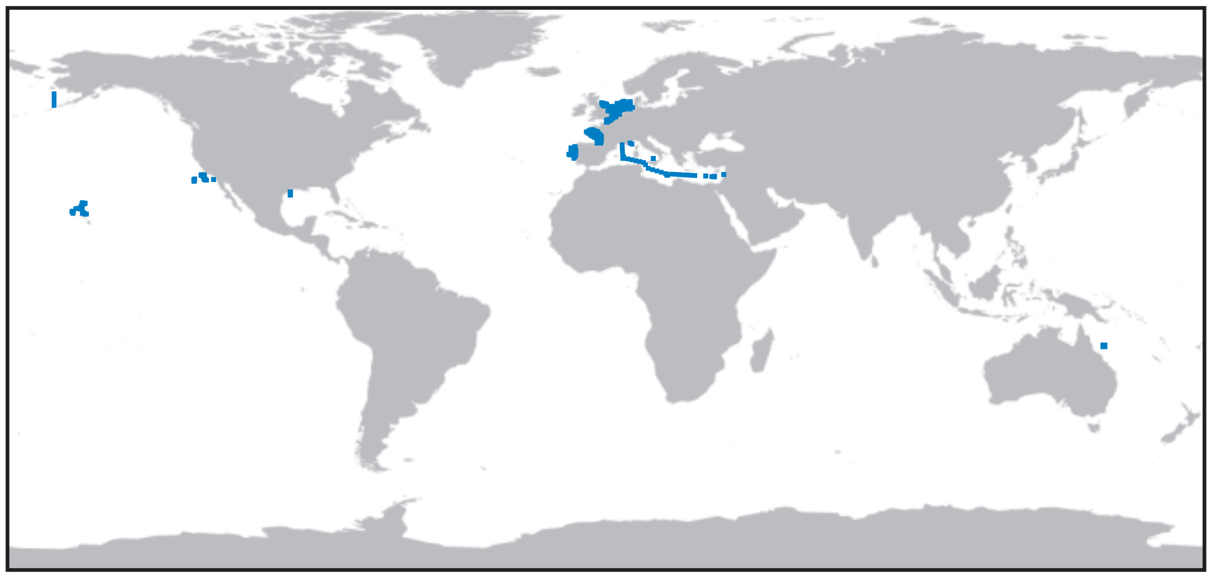
Loads of data

ZooScan = 1 Bpx/y, 1.5M objects/y
UVP = 8.6Bpx/y, ~10M objects/y
ISIIS = 25Tpx/y, 100M objects/y

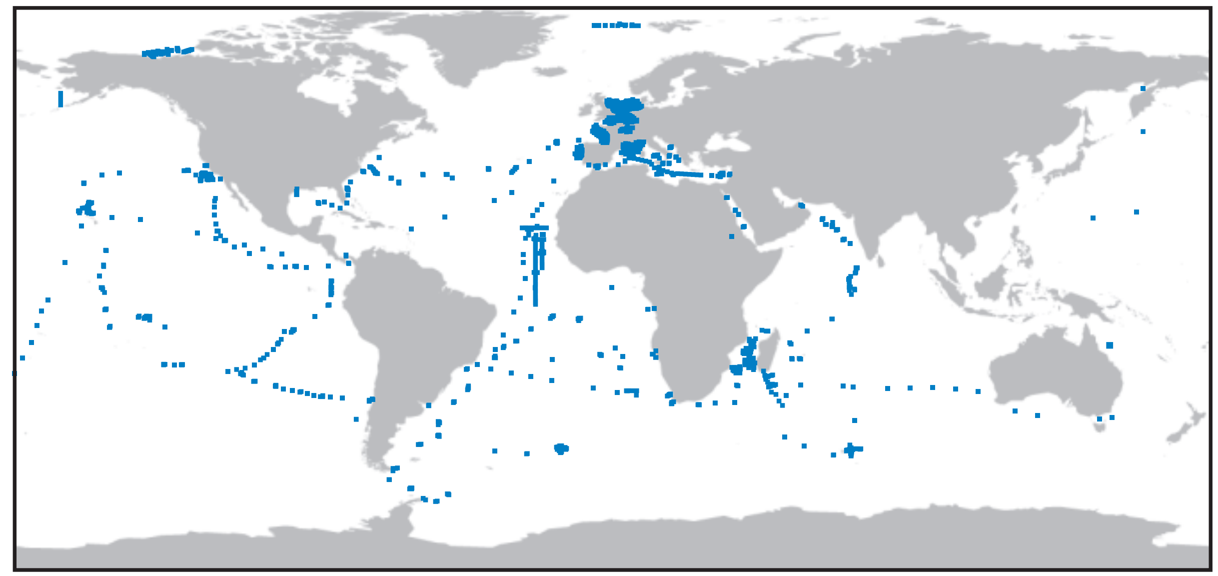
Steep growth in data acquisition



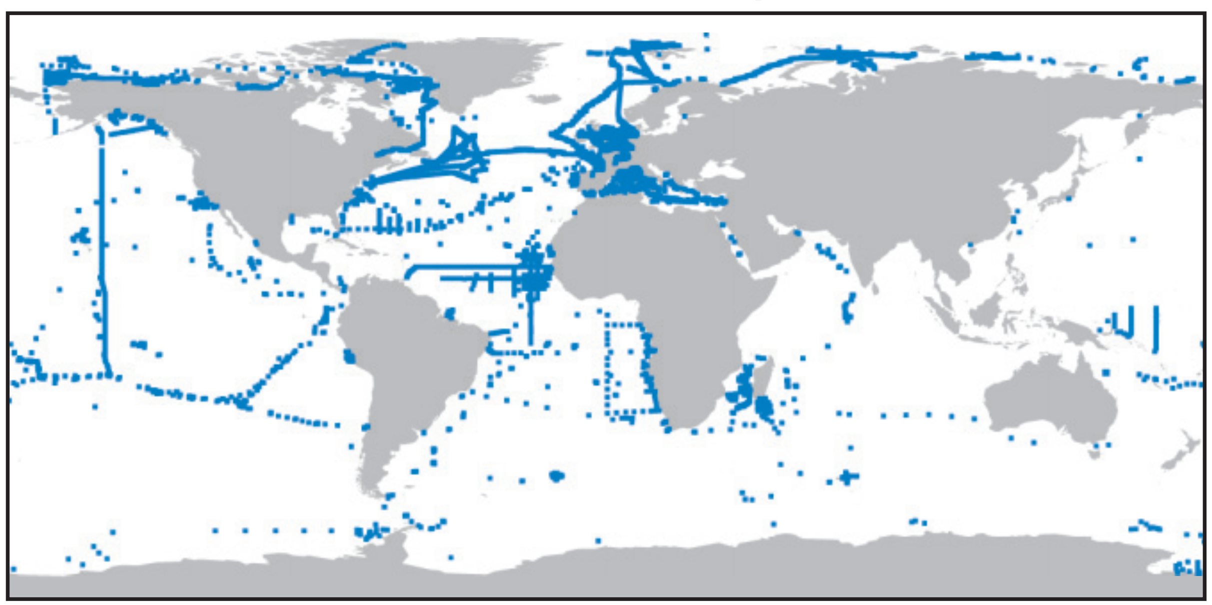
Steep growth in data acquisition



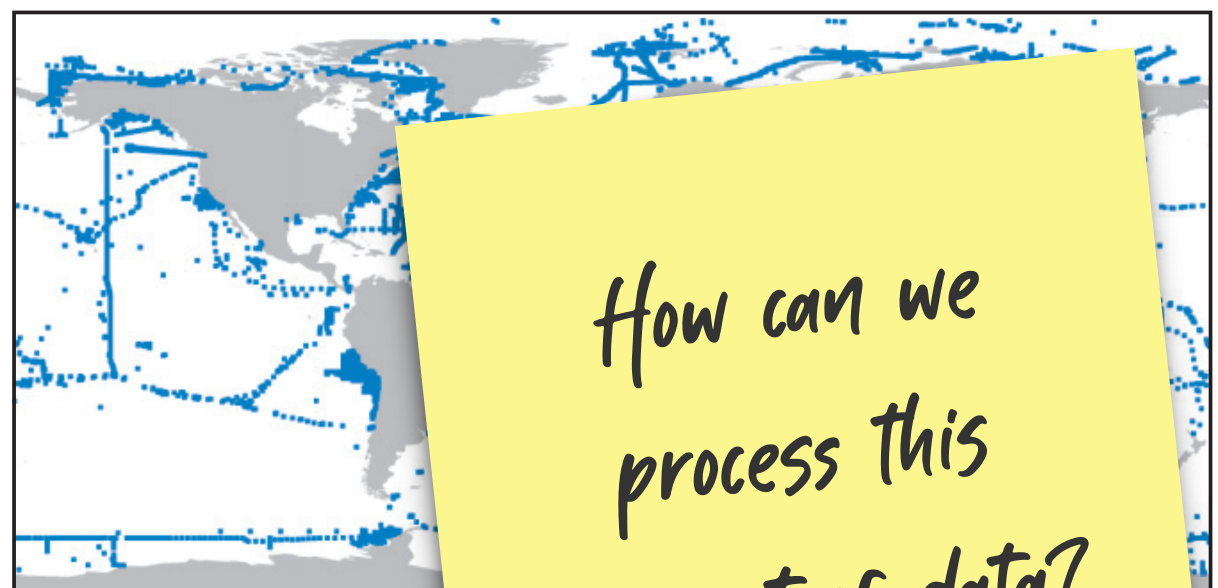
2008 (4,000 samples)



2012 (17,000 samples)



2016 (56,000 samples)

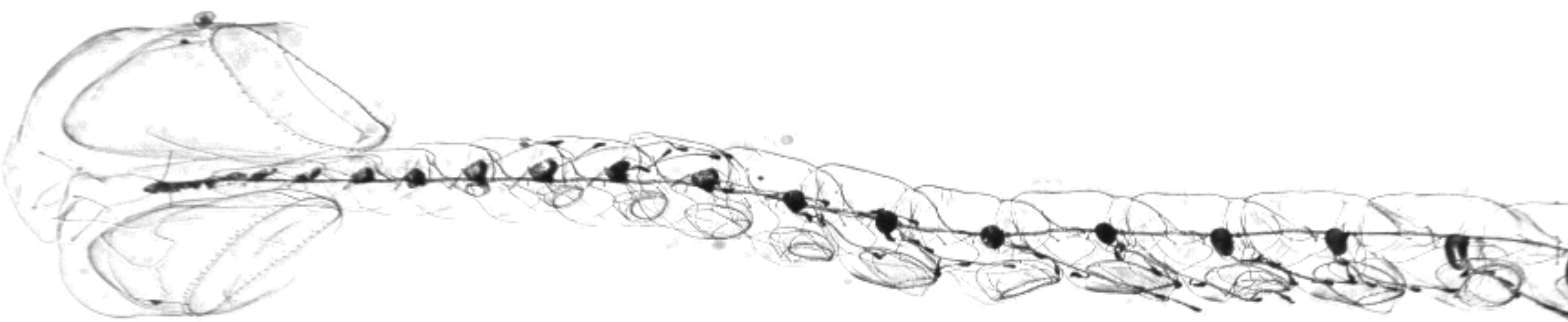


2020

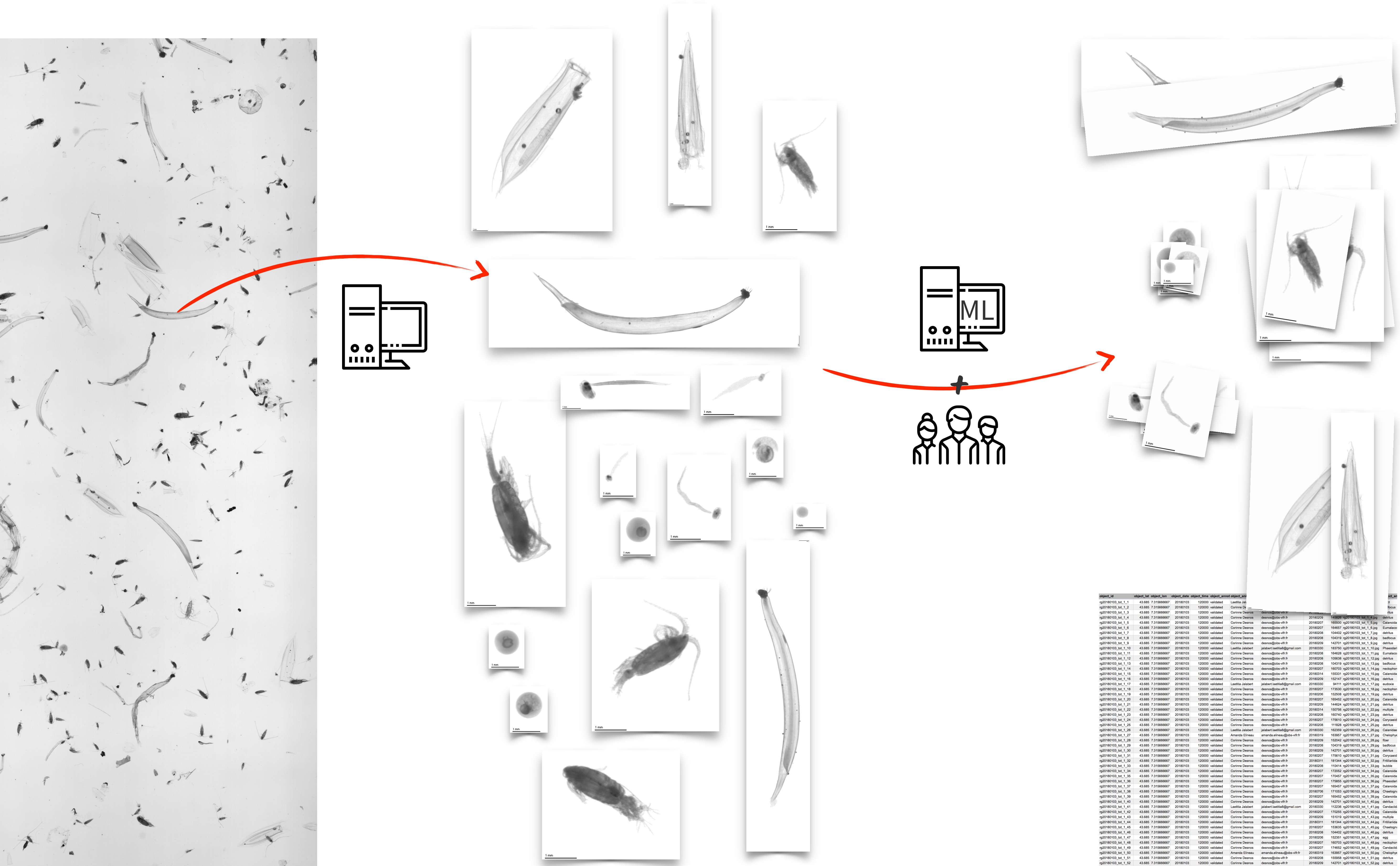
How can we process this amount of data?

Machine learning for data acquisition

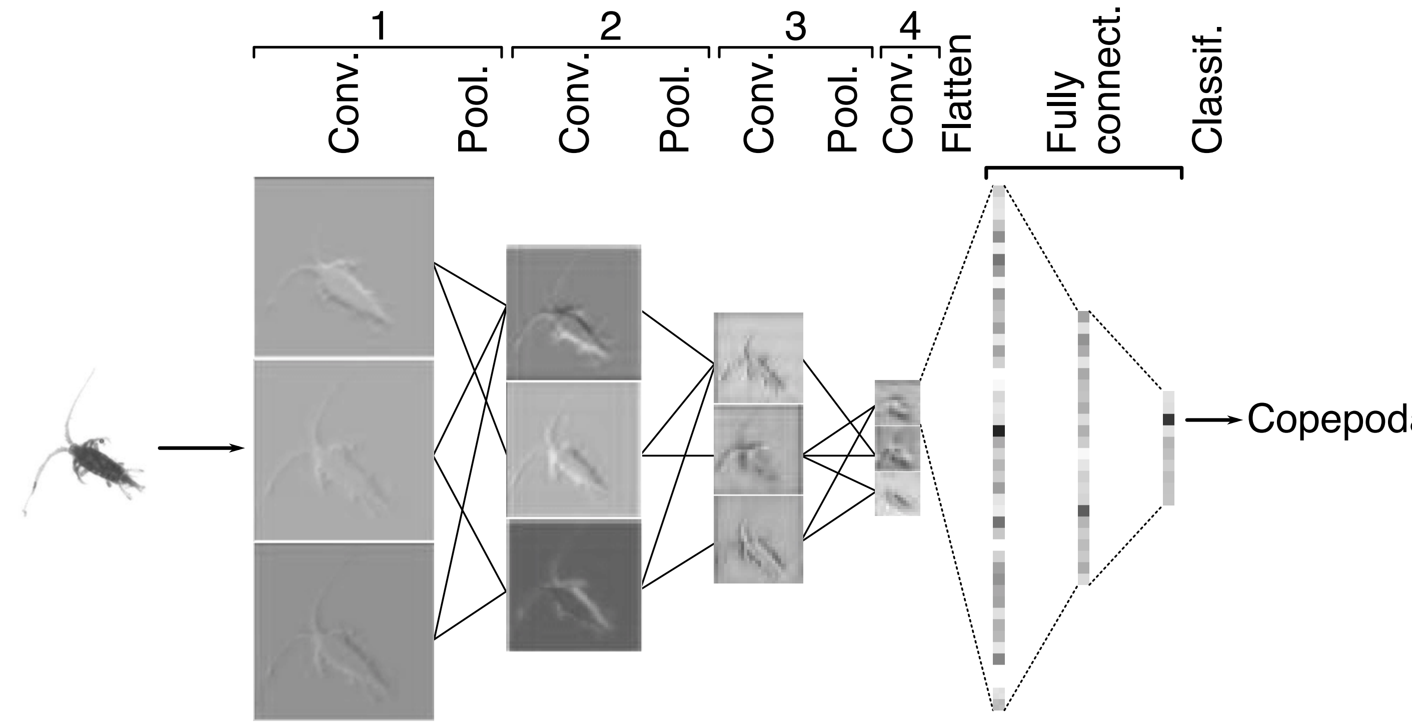
From pictures to numbers



Quantitative imaging and ML-assisted sorting



Measure + classify *vs.* Deep learning



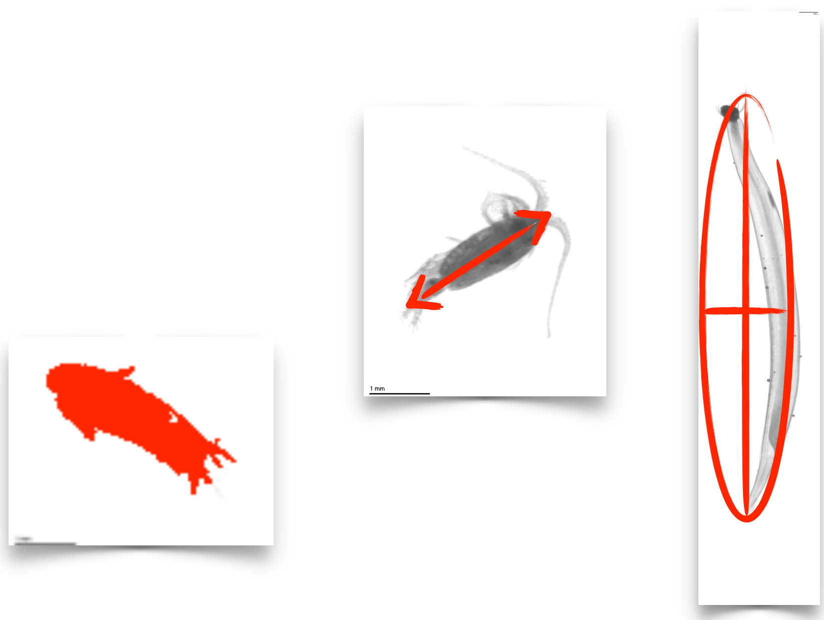
Software to **extract features**
Area (ESD)
Mean/SD of grey
Feret diameter
Major/minor, angle

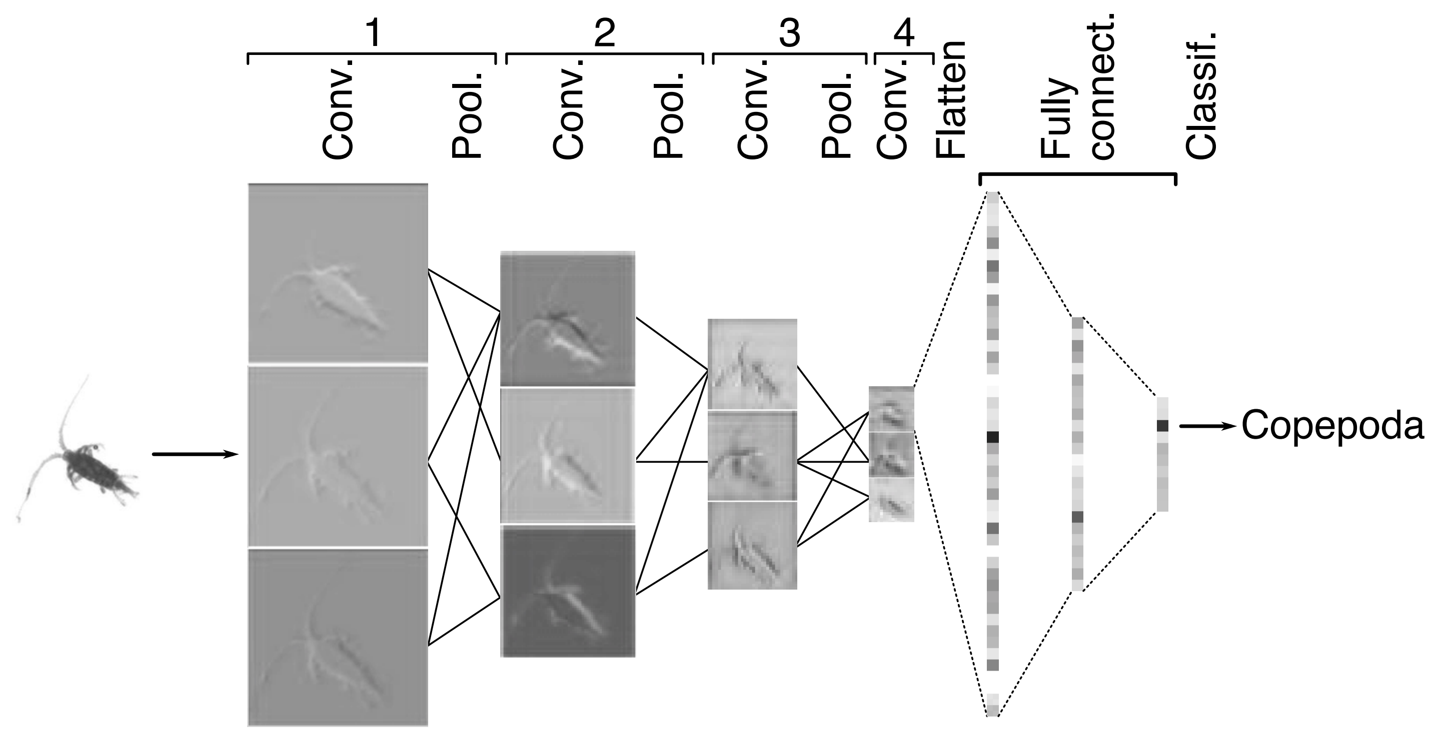
+ a **classifier**

A **feature extractor**
Convolutions
Pooling

+ a **classifier**
Flattening
Fully connected layers

Measure + classify *vs.* Deep learning





Software to **extract features**
Area (ESD)
Mean/SD of grey
Feret diameter
Major/minor, angle

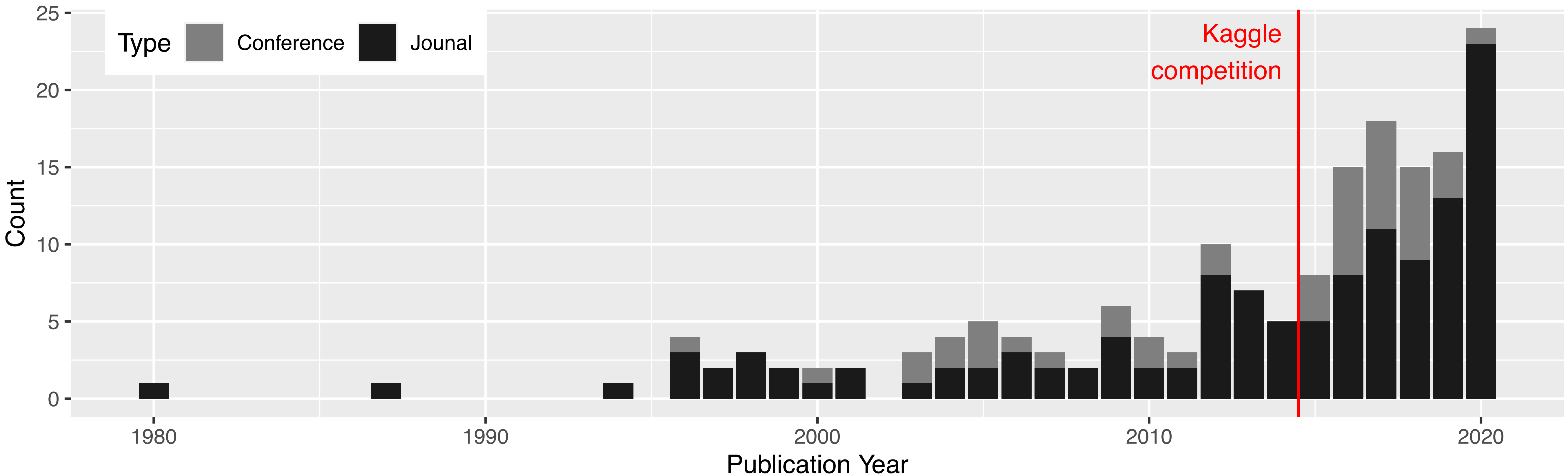
+ a **classifier**

A **feature extractor**
Convolutions
Pooling

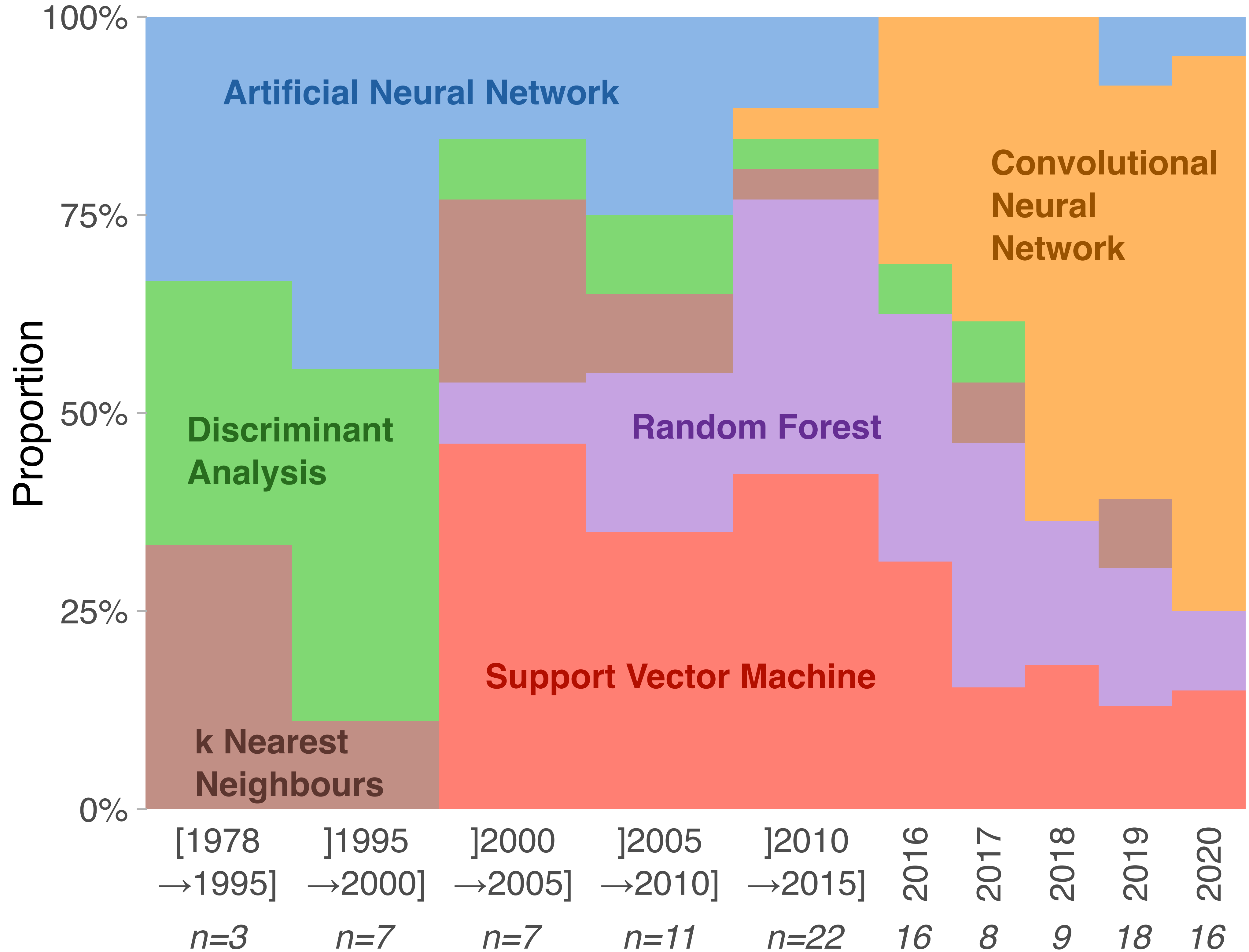
+ a **classifier**
Flattening
Fully connected layers

Plankton image classification is a challenging ML problem

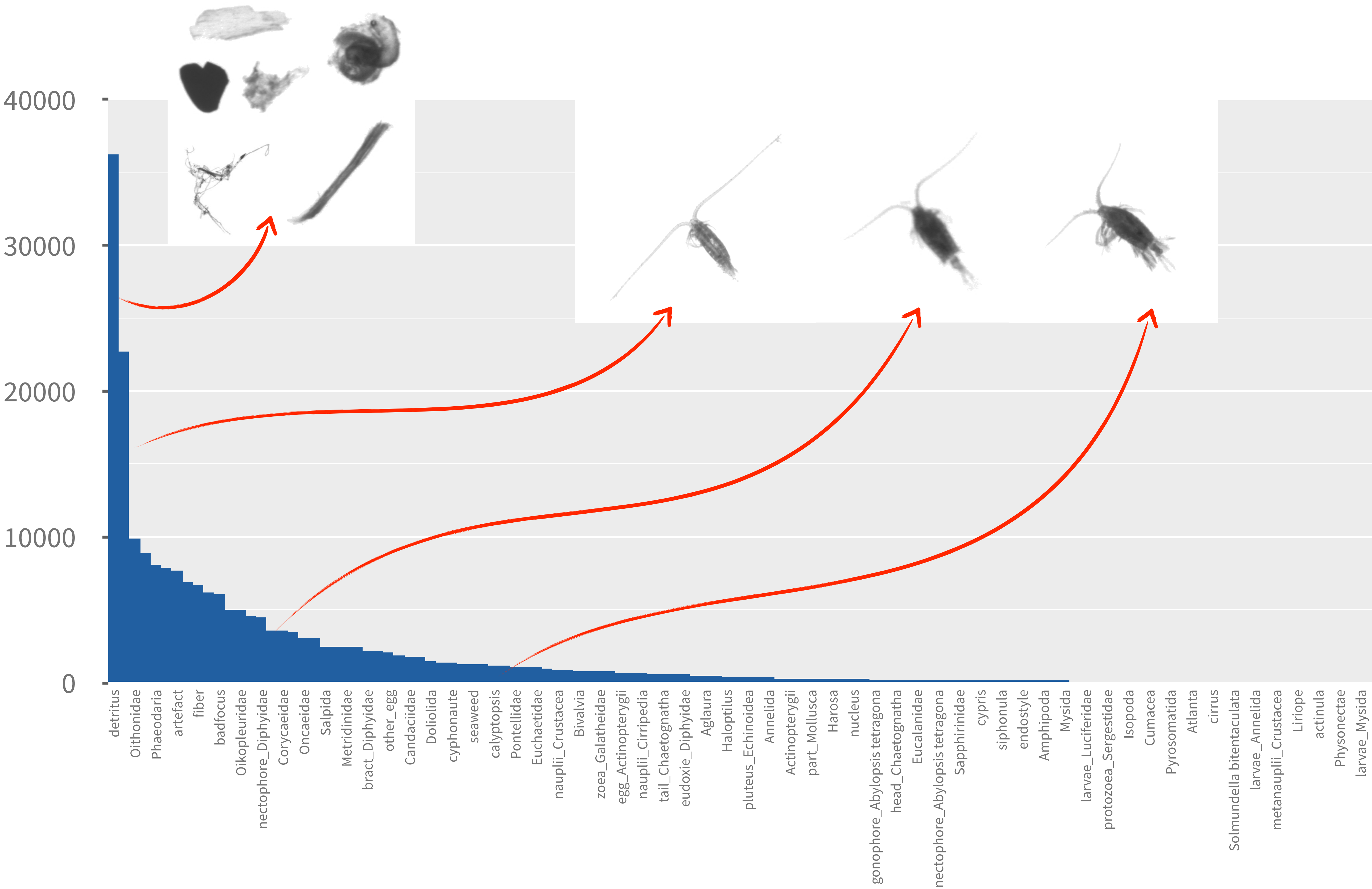
Total: 175 papers!

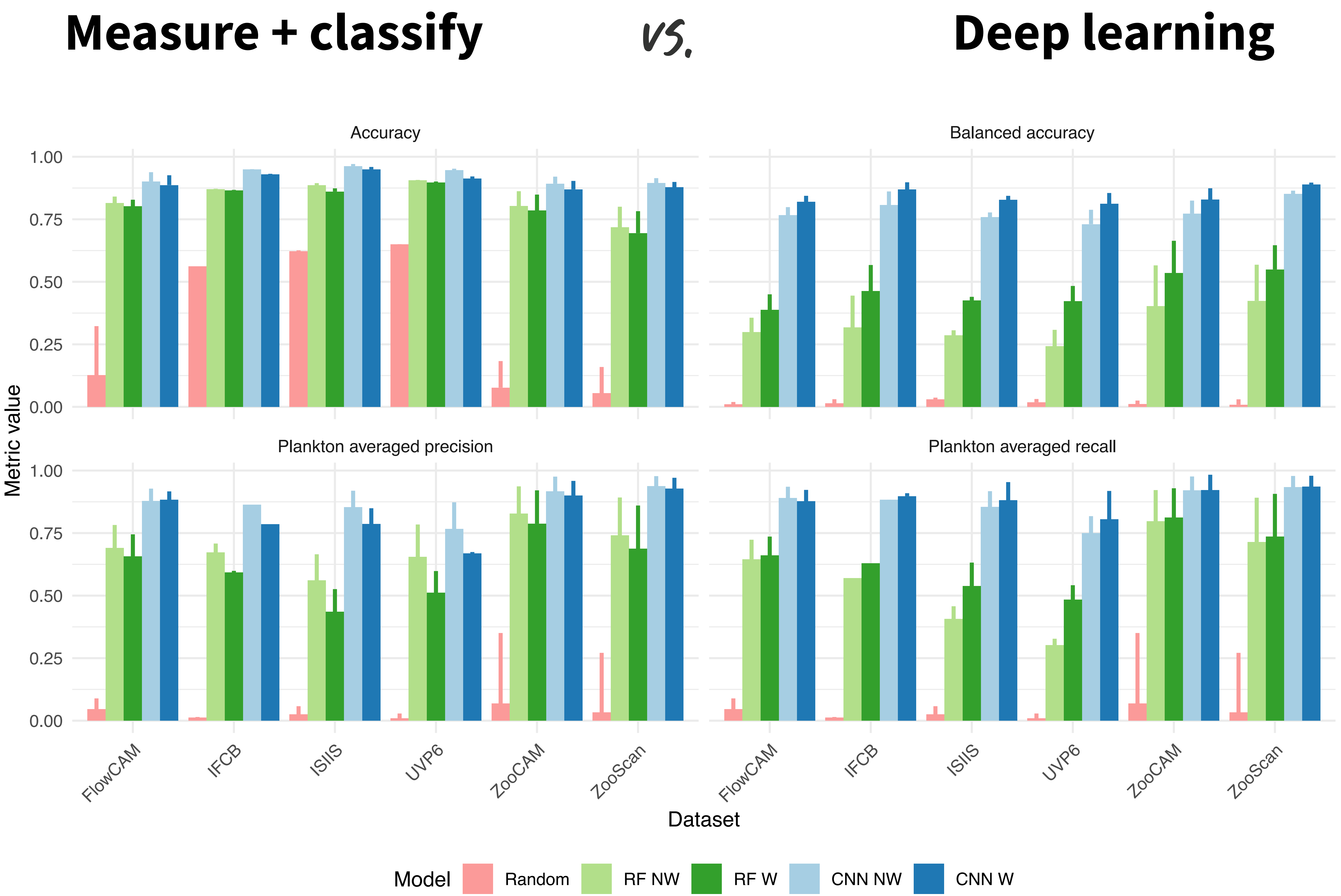


Evolution of machine learning techniques



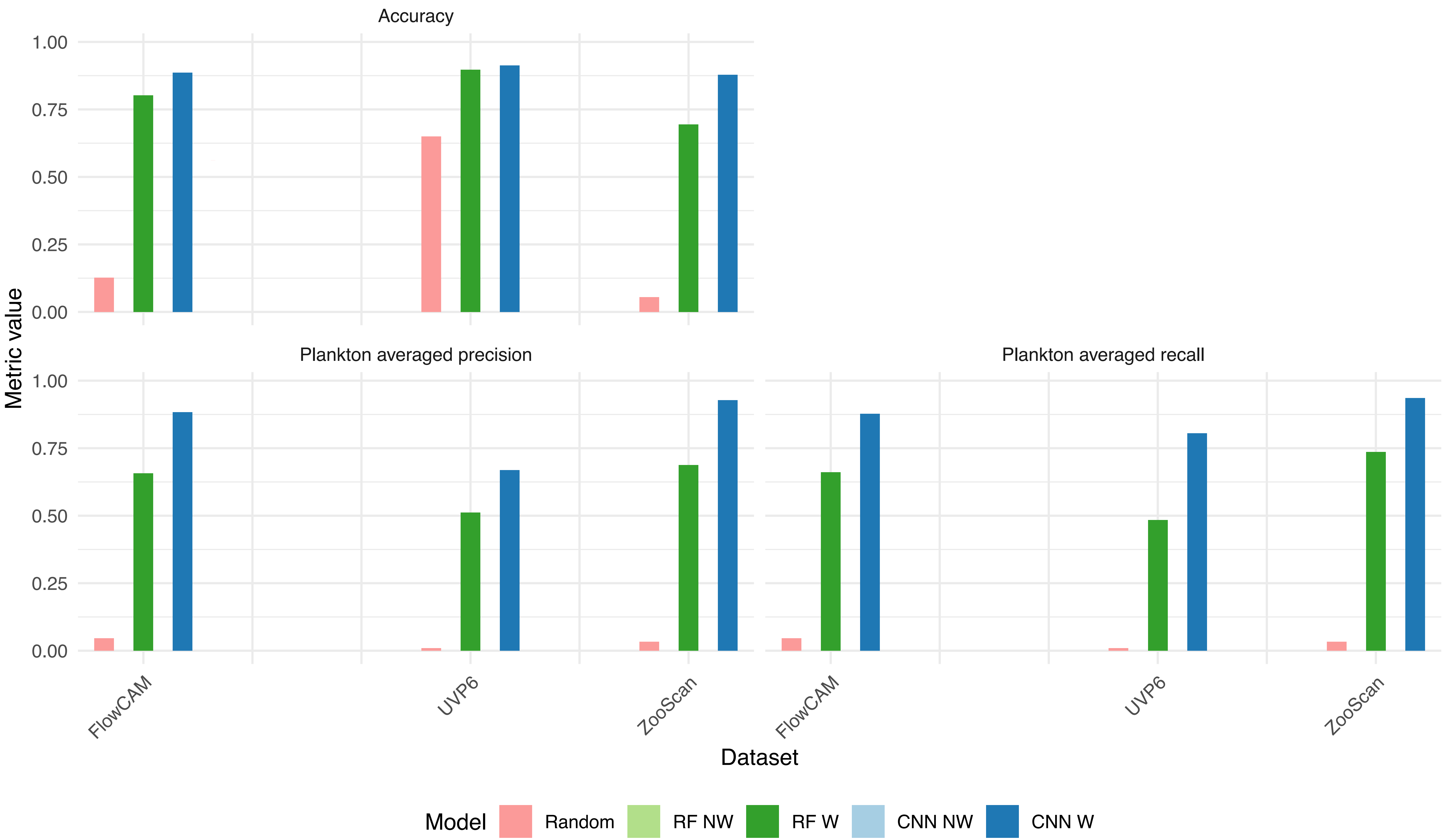
Why is it hard?



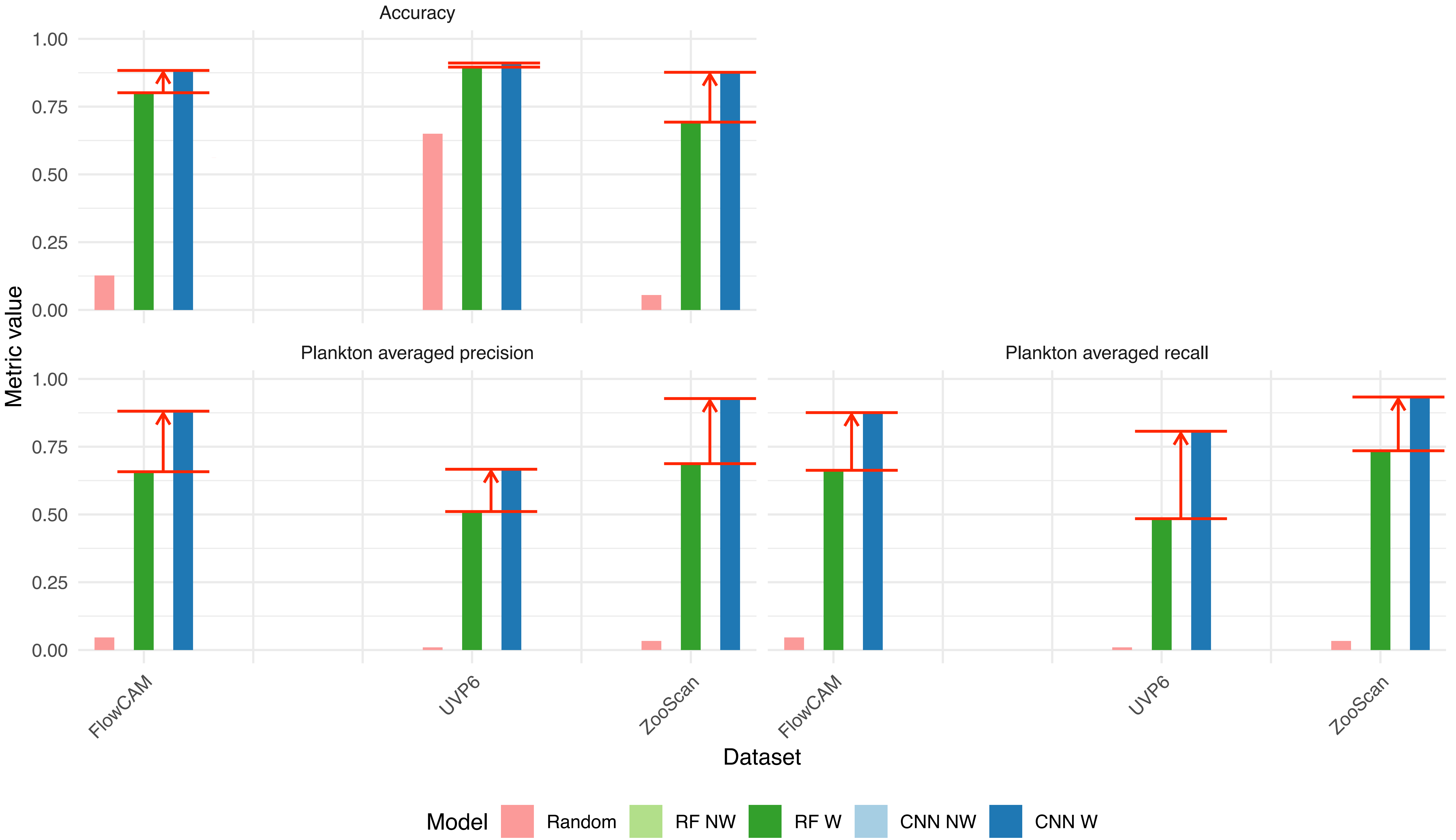


Measure + classify
vs.
Deep learning
Accuracy
Balanced accuracy
Plankton averaged precision
Plankton averaged recall
Metric value
1.00
0.75
0.50
0.25
0.00
FlowCAM
IFCB
ISIIS
UVP6
ZooCAM
ZooScan
Dataset
Model
Random
RF NW
RF W
CNN NW
CNN W

Measure + classify *vs.* Deep learning



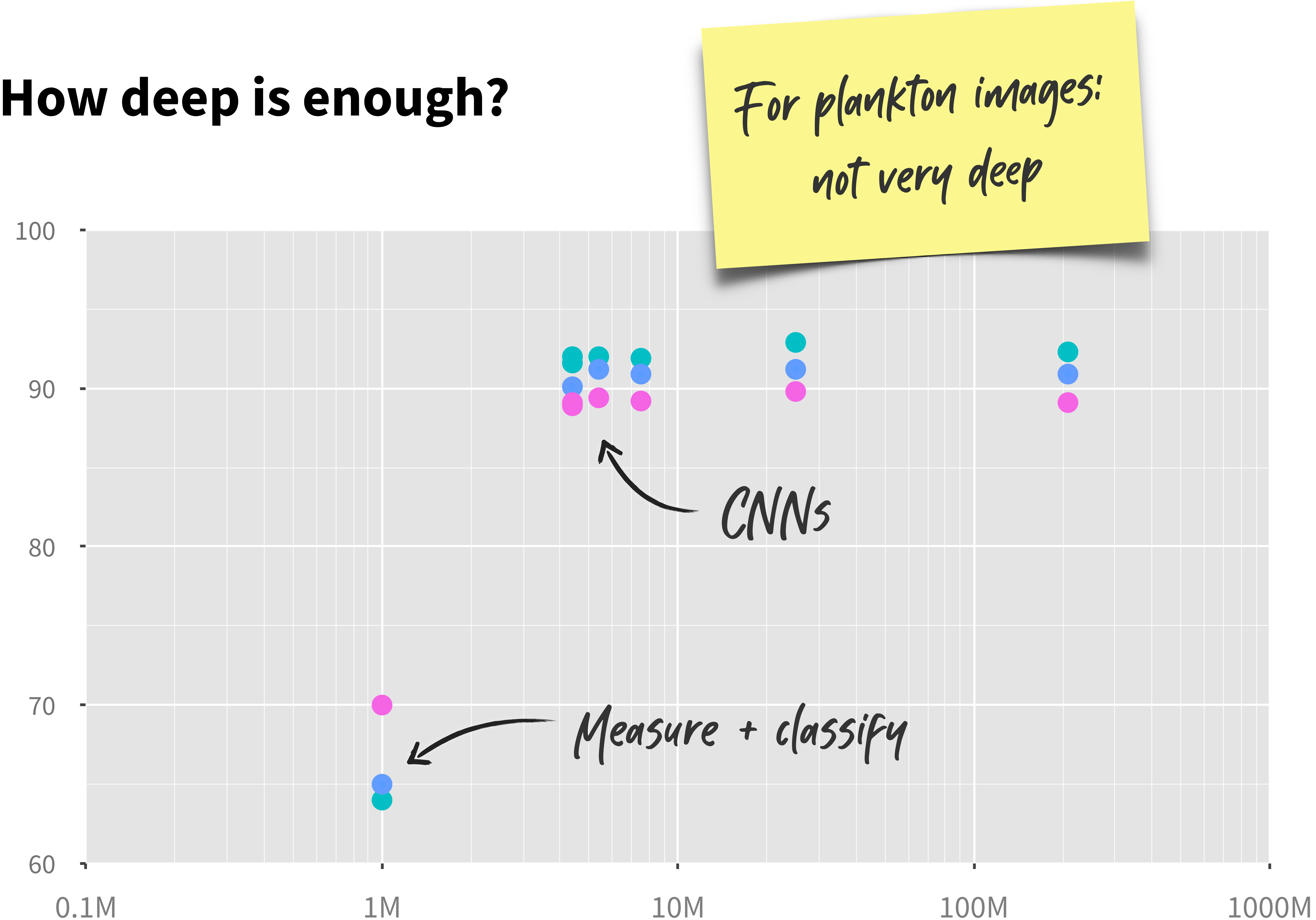
Measure + classify *vs.* Deep learning



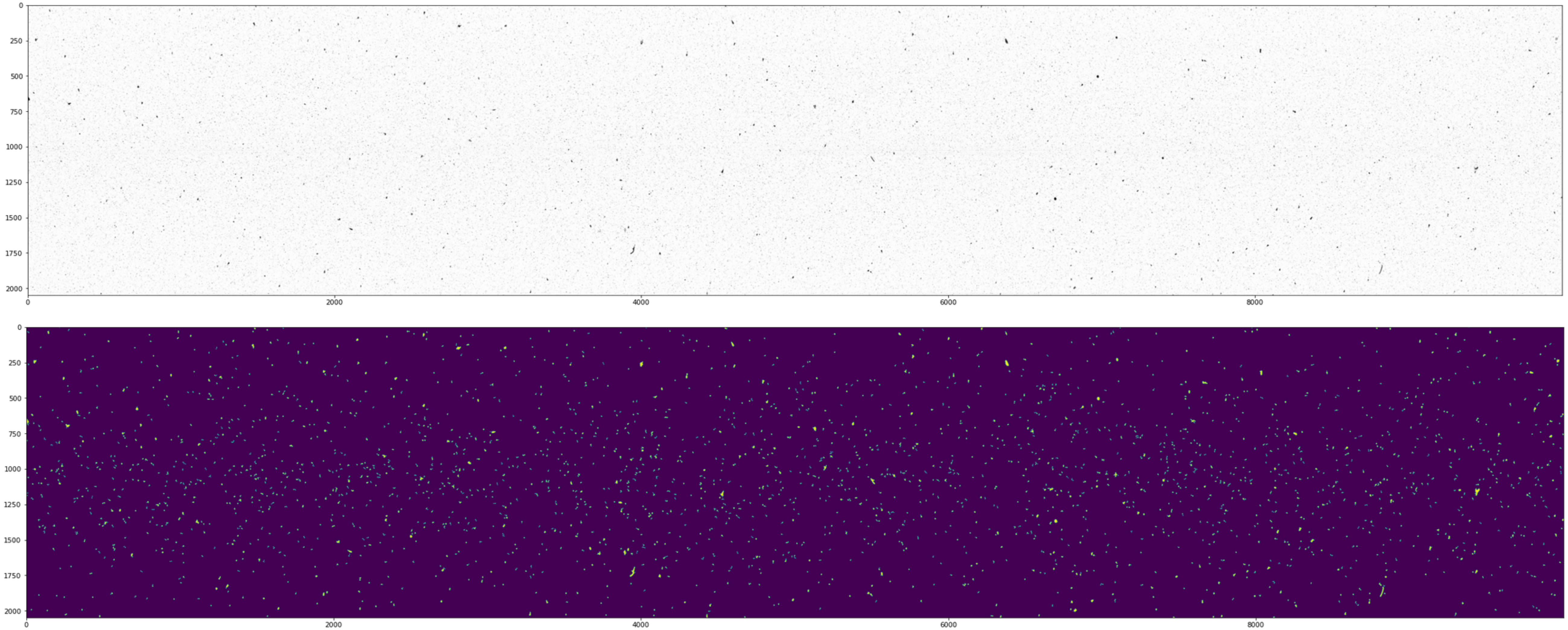
How deep is enough?

Model	Size	Accuracy	Avg. precision	Avg. recall
MobileNet v4 + 600	5.4M	89.4	91.2	92.0
MobileNet v4 + 1792	7.5M	89.2	90.9	91.9
EfficientNet v2 S + 600	25M	89.8	91.2	92.9
EfficientNet v2 XL + 600	208M	89.1	90.9	92.3
MobileNet v4 + 50	4.4M	88.9	90.1	901.6
MobileNet v4 features + PCA + RF	~4.4M	89.1	90.1	92.0

How deep is enough?



Sometimes, classification is not enough



0.35s ~ 2500 objects

6h ~ 150 millions objects

on cruise ~ 1.5 billion objects, among which ~1% are plankton (~20-30 millions)

Semantic segmentation

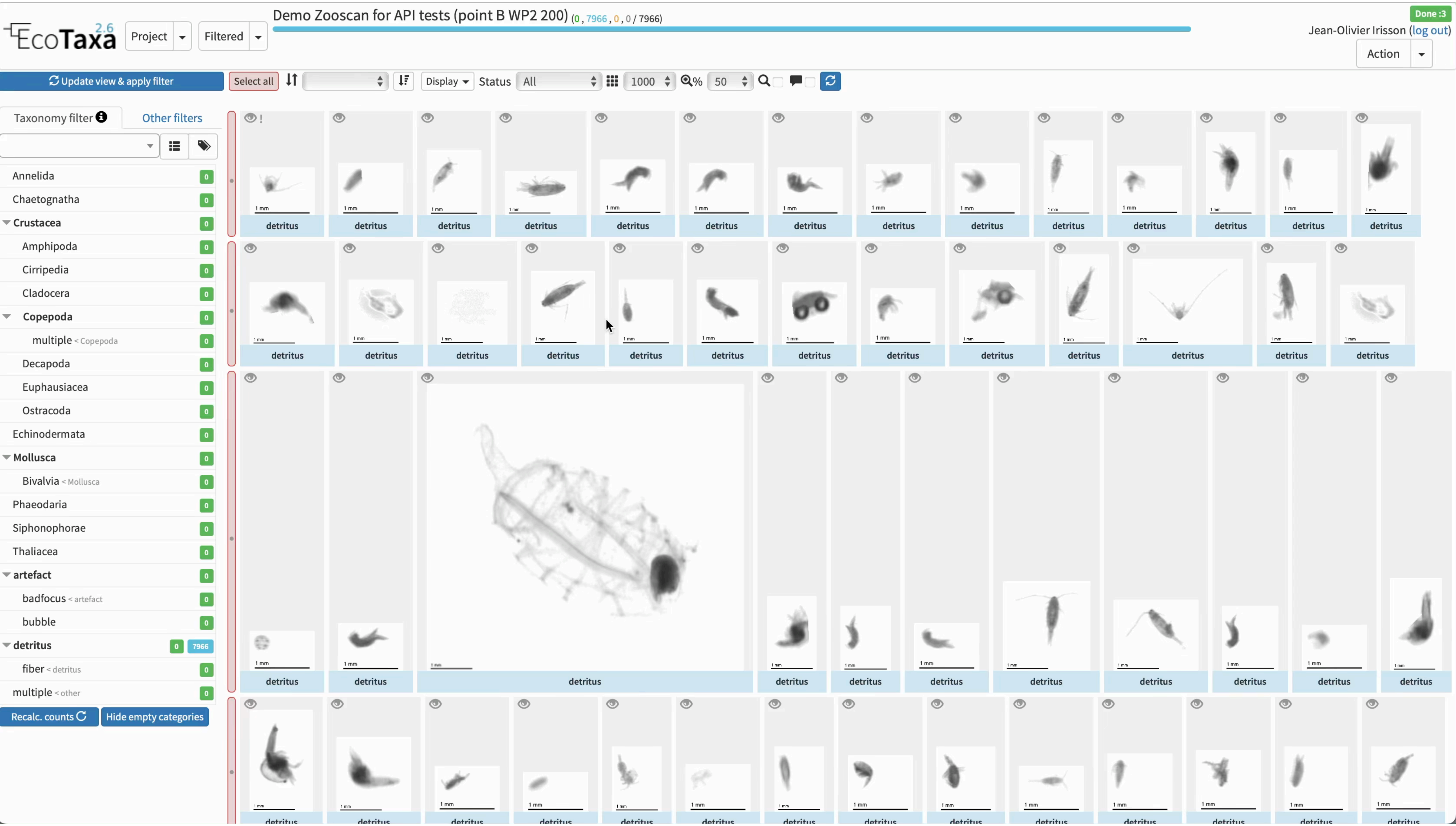
Extract only certain objects from a scene

Detect or segment objects

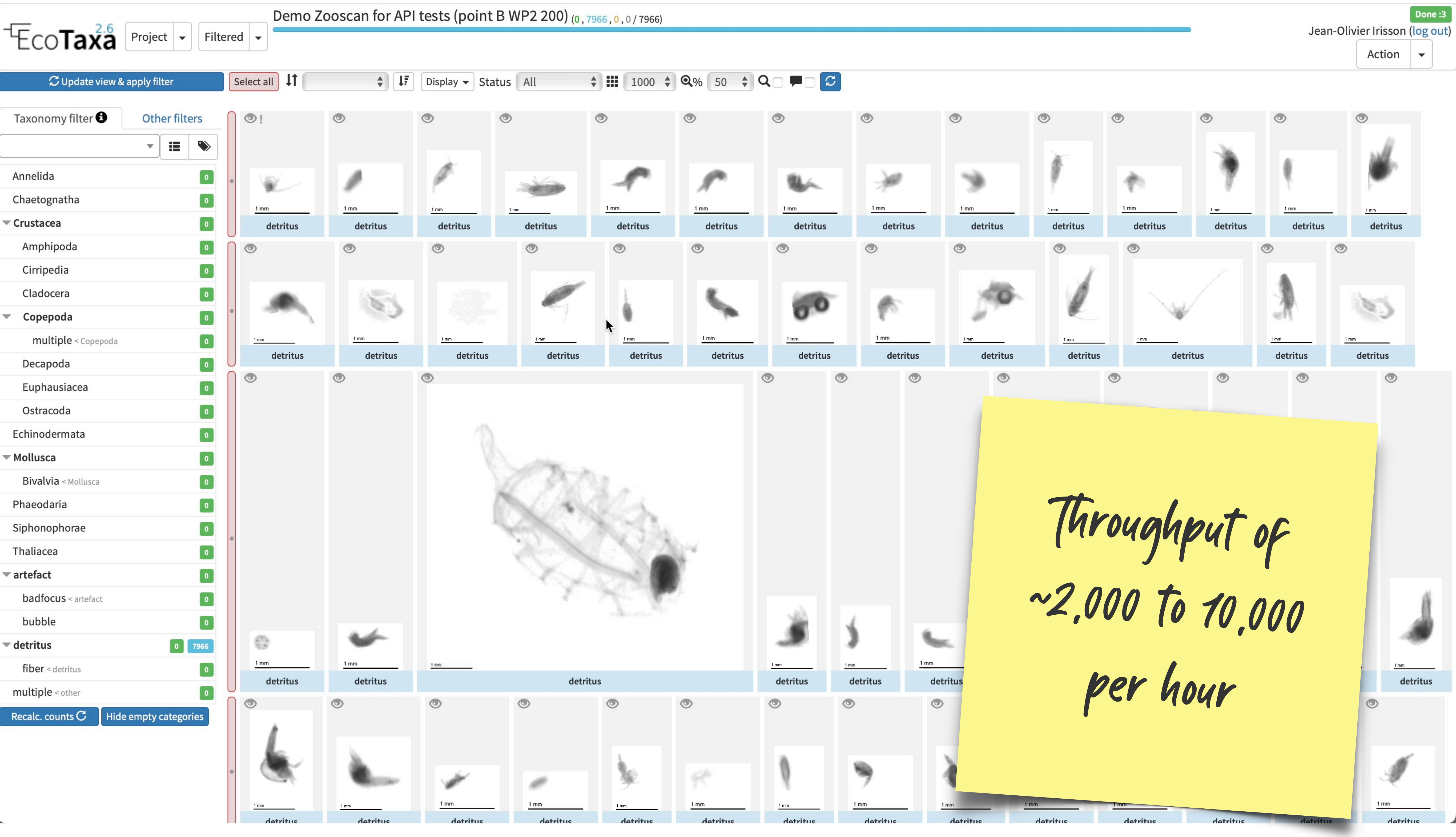
Classify them a the same time

Detectron2

EcoTaxa: ML-assisted image classification

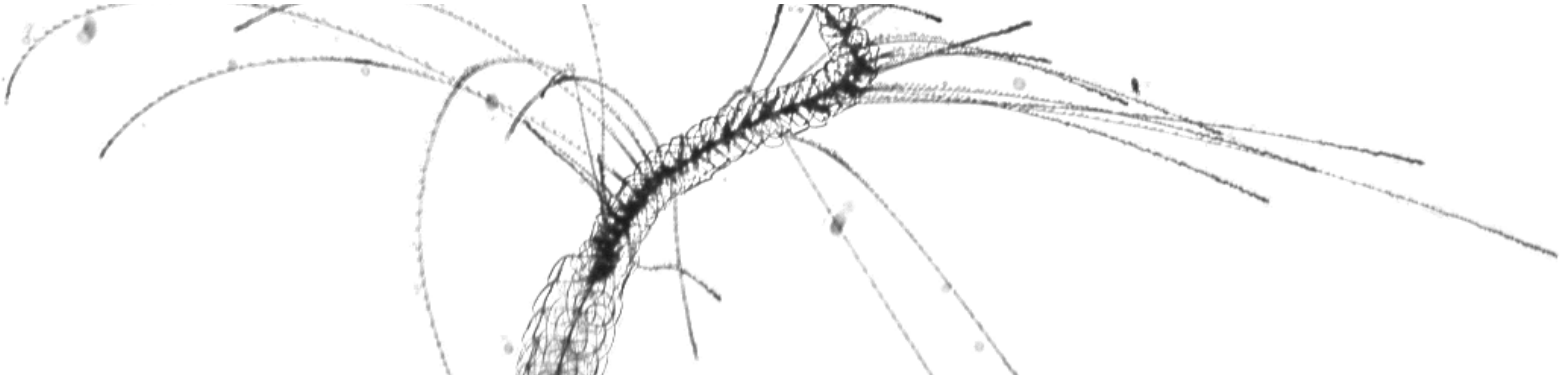


EcoTaxa: ML-assisted image classification



Machine learning for data acquisition

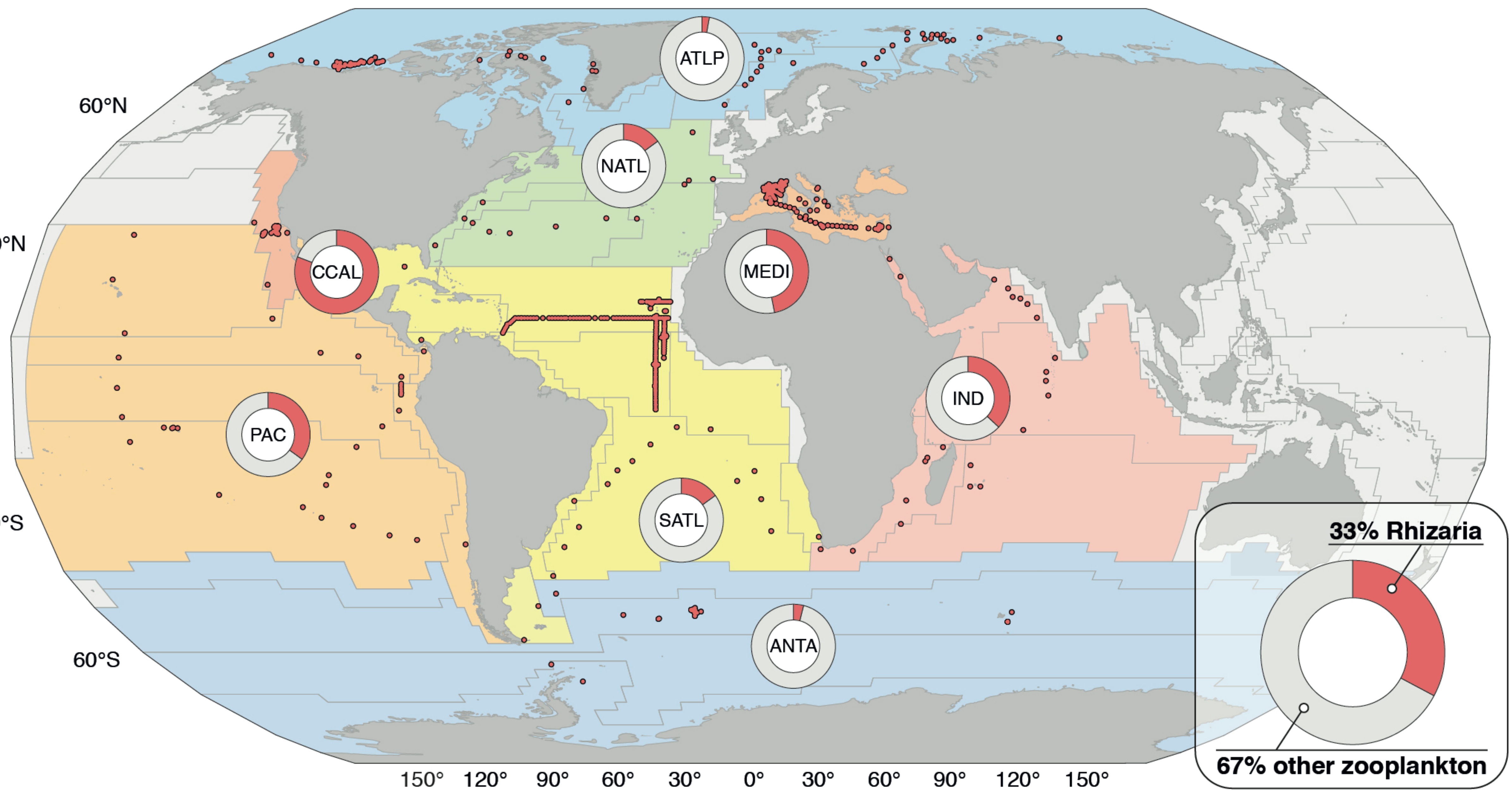
From numbers to knowledge



Temporal dynamics

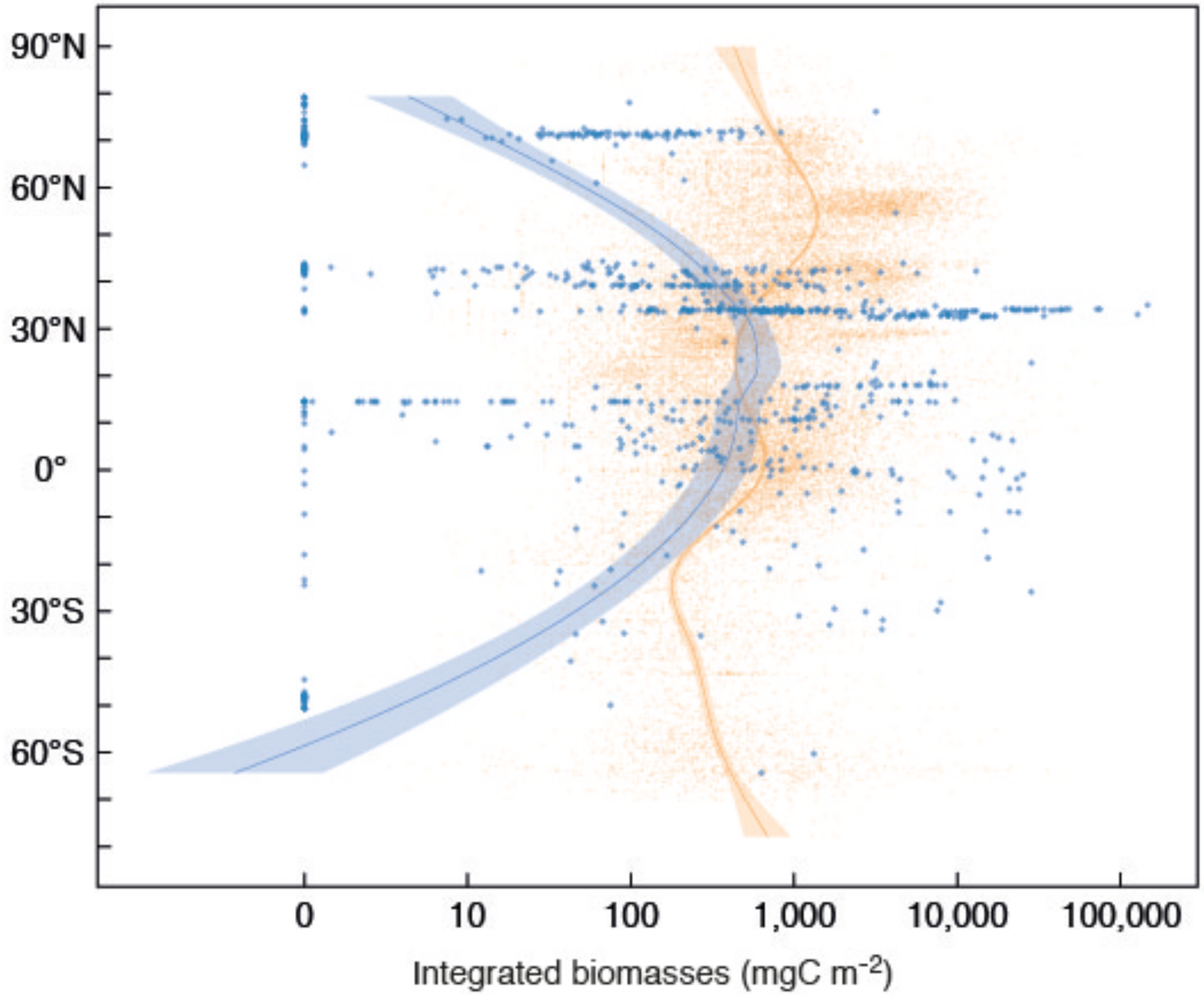
NB: performance metrics are hard to interpret!

Global biomass of fragile plankton



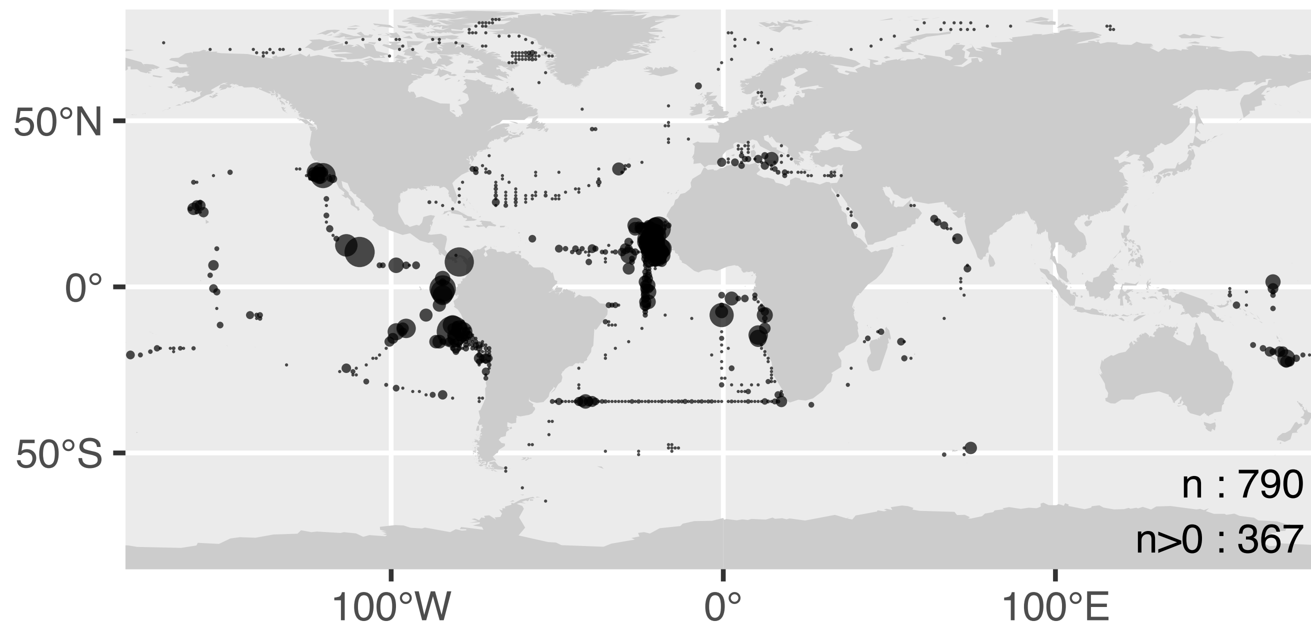
Biard T, Stemmann L, Picheral M, Mayot N, Vandromme P, Hauss H, Gorsky G, Guidi L, Kiko R, Not F (2016) In situ imaging reveals the biomass of giant protists in the global ocean. Nature 532:504.

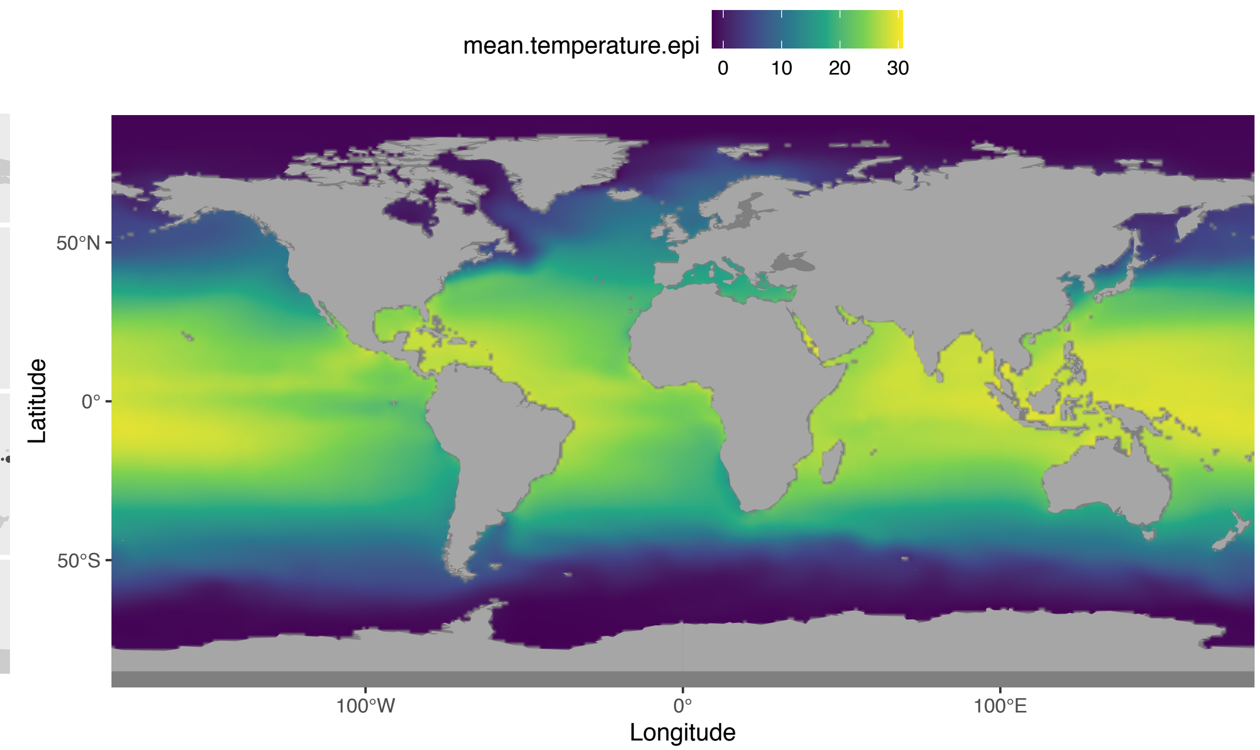
Global biomass of fragile plankton

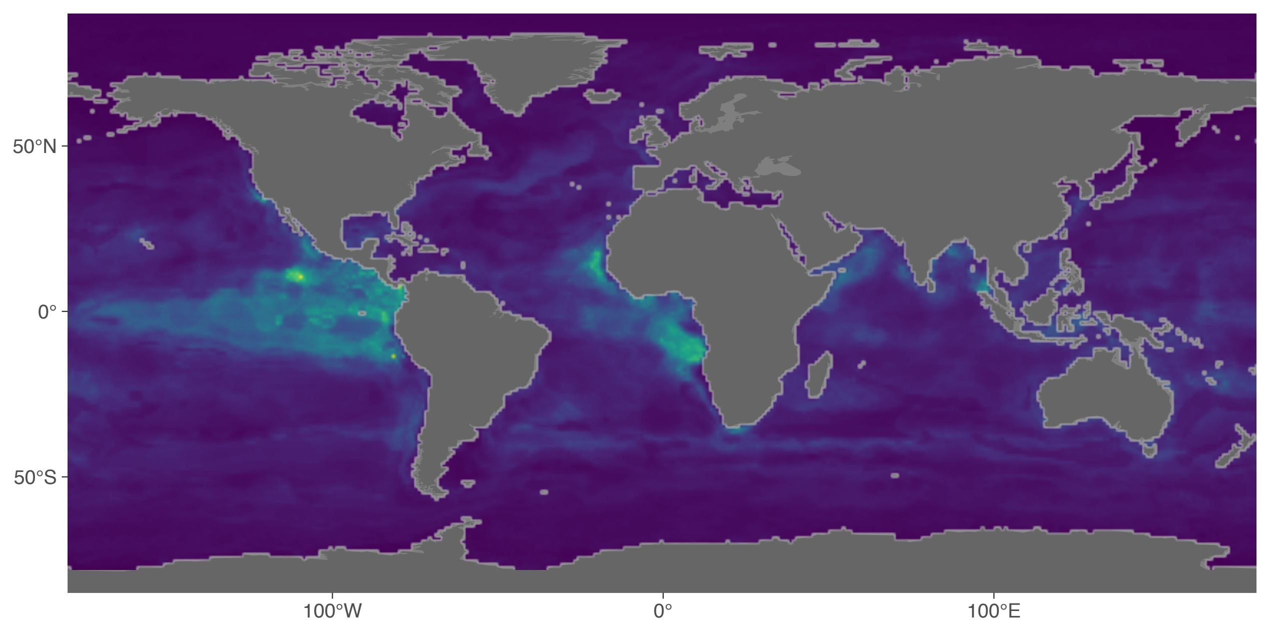


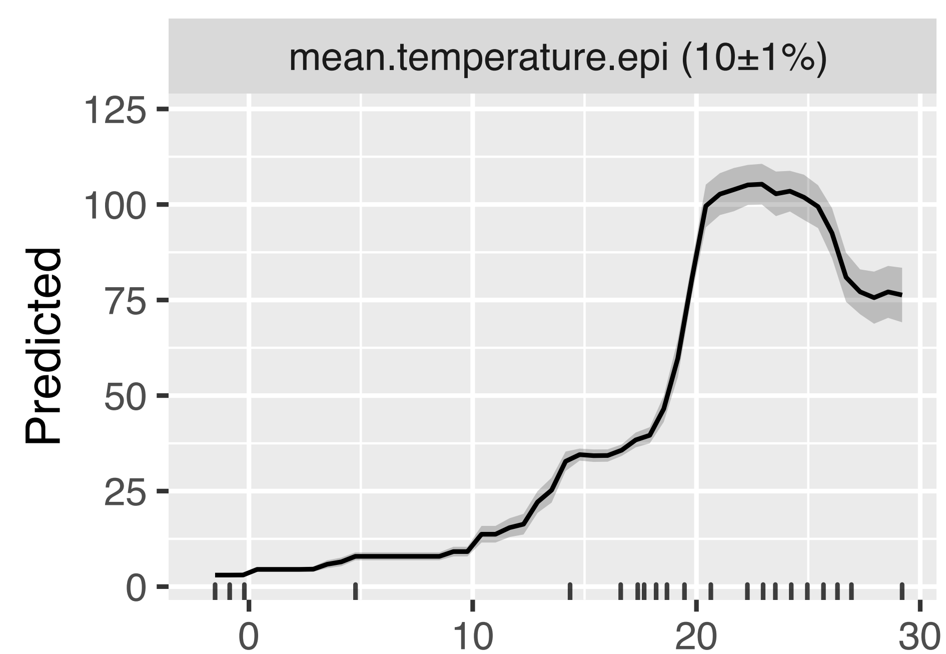
Biard T, Stemmann L, Picheral M, Mayot N, Vandromme P, Hauss H, Gorsky G, Guidi L, Kiko R, Not F (2016) In situ imaging reveals the biomass of giant protists in the global ocean. Nature 532:504.

Total biomass macrozooplancton (>1mm)



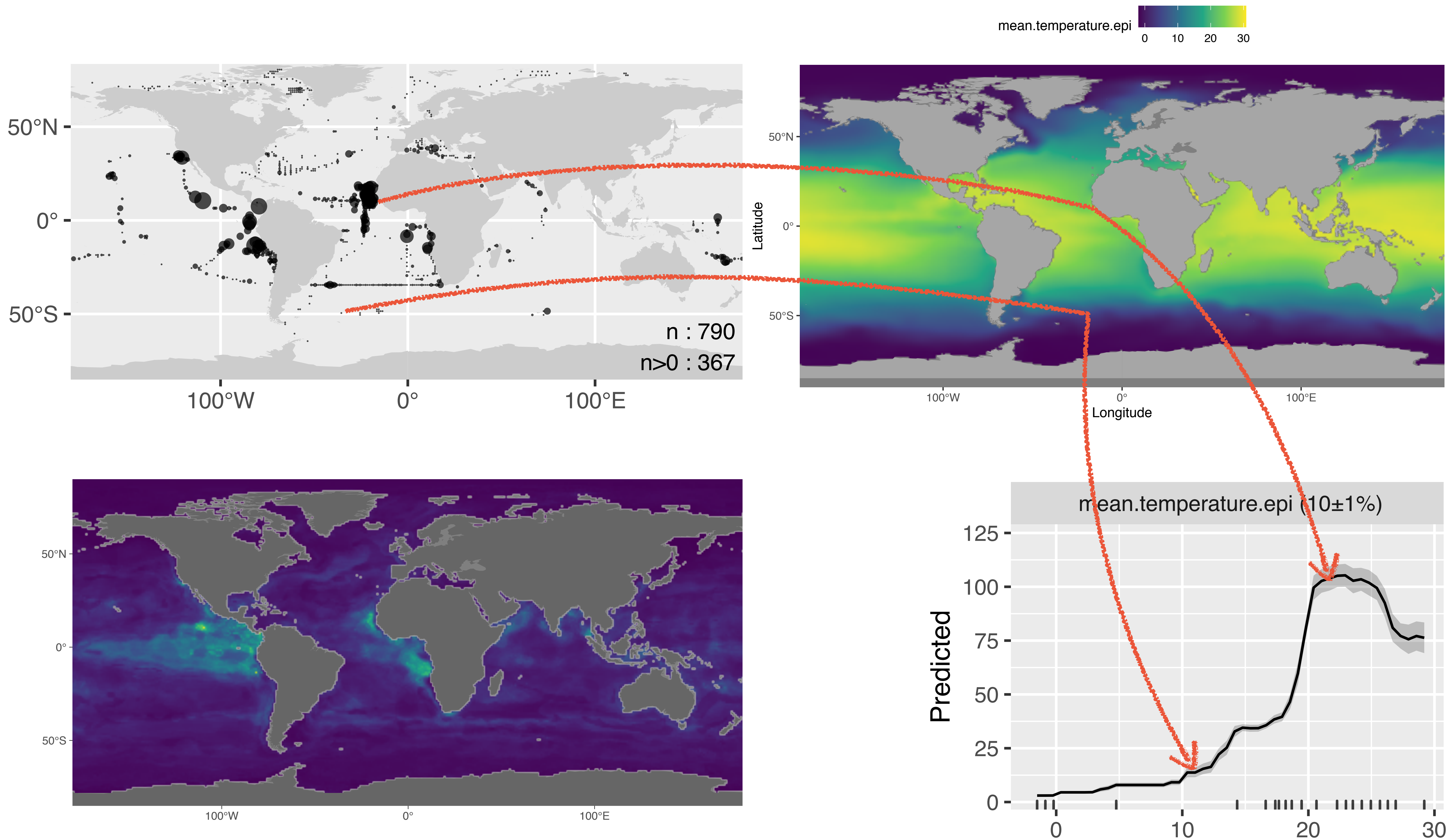






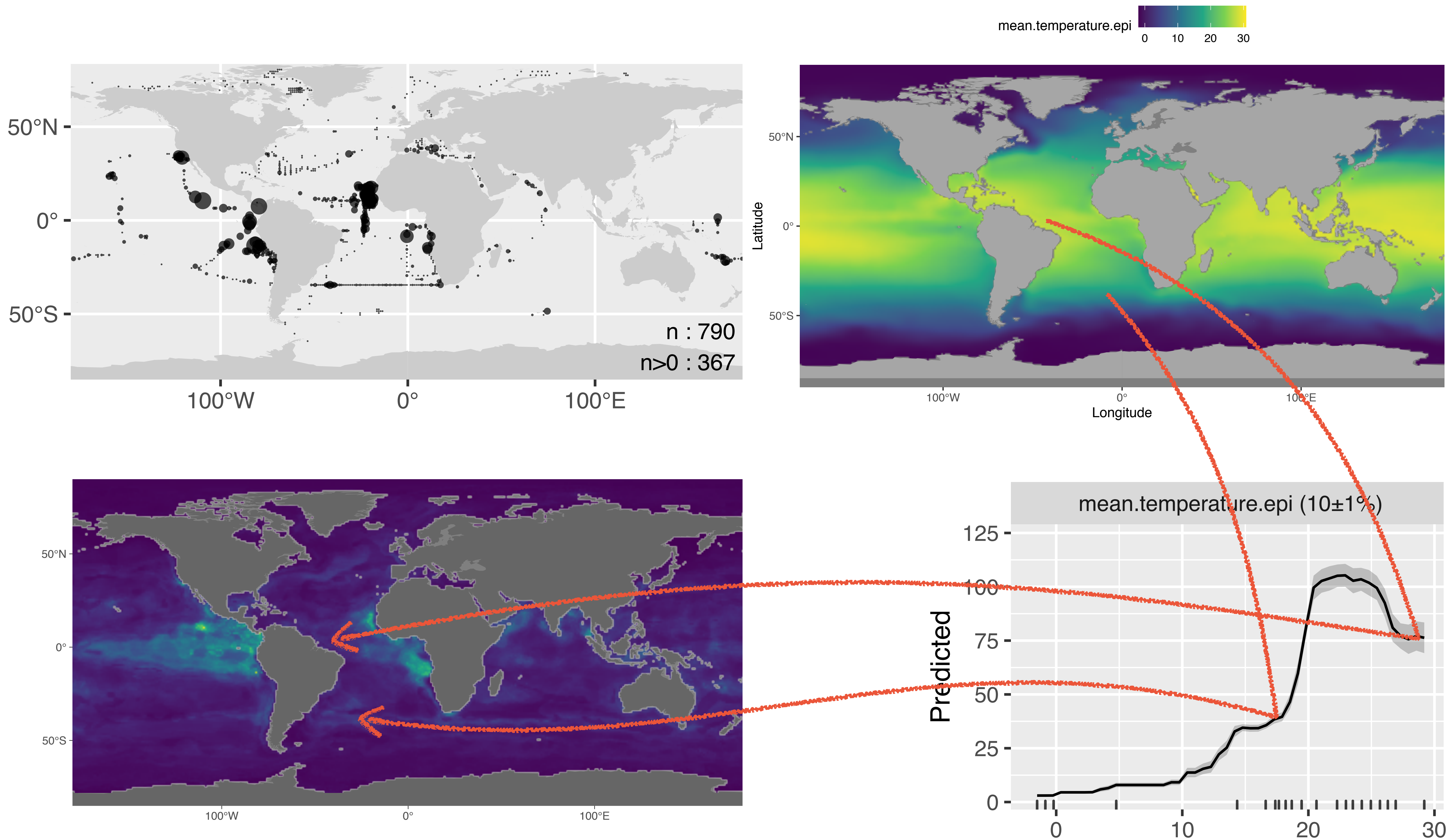
Drago L, Panaïotis T, Irisson J-O, […], Stemmann L, Kiko R (2022) Global Distribution of Zooplankton Biomass Estimated by In Situ Imaging and Machine Learning. Frontiers in Marine Science 9.

Total biomass macrozooplancton (>1mm)



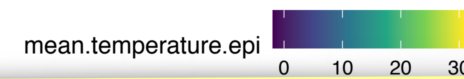
Drago L, Panaïotis T, Irisson J-O, [...], Stemmann L, Kiko R (2022) Global Distribution of Zooplankton Biomass Estimated by In Situ Imaging and Machine Learning. Frontiers in Marine Science 9.

Total biomass macrozooplancton (>1mm)

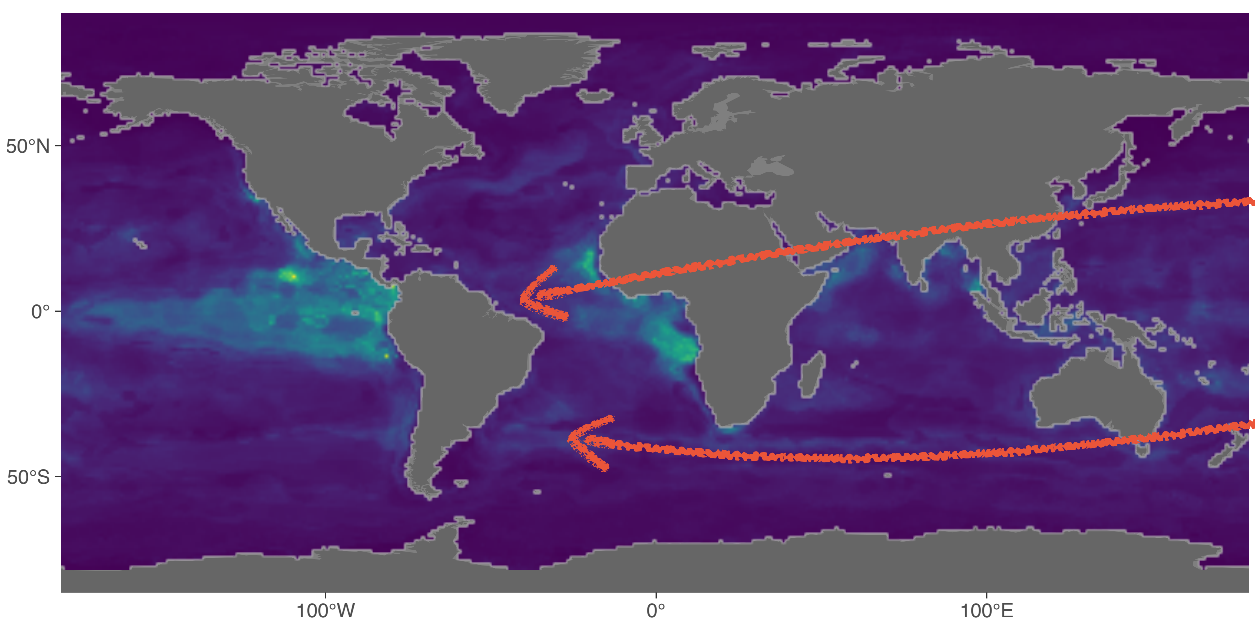


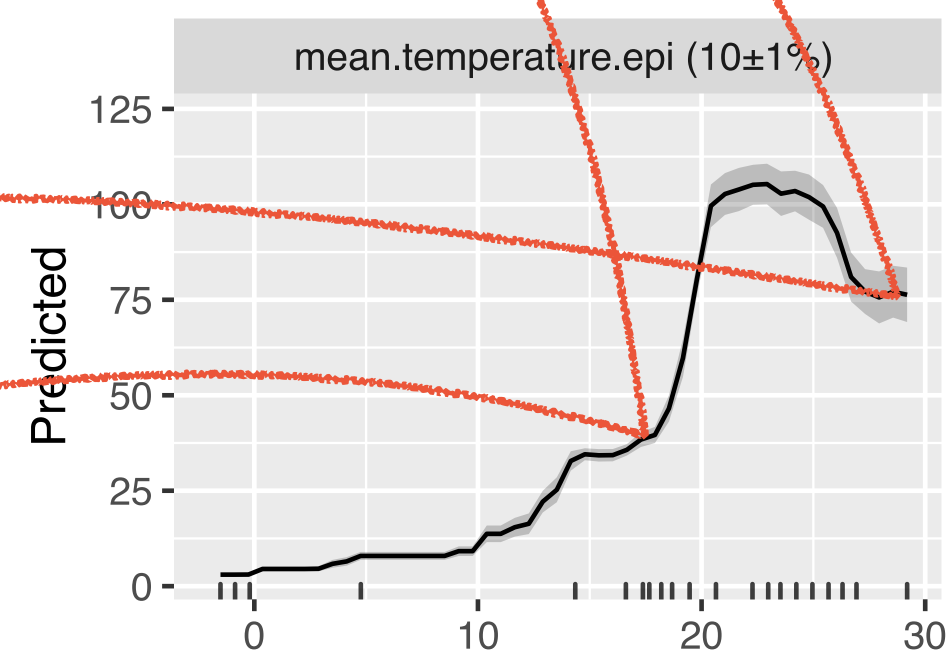
Drago L, Panaïotis T, Irisson J-O, [...], Stemmann L, Kiko R (2022) Global Distribution of Zooplankton Biomass Estimated by In Situ Imaging and Machine Learning. Frontiers in Marine Science 9.

Total biomass macrozooplancton (>1mm)



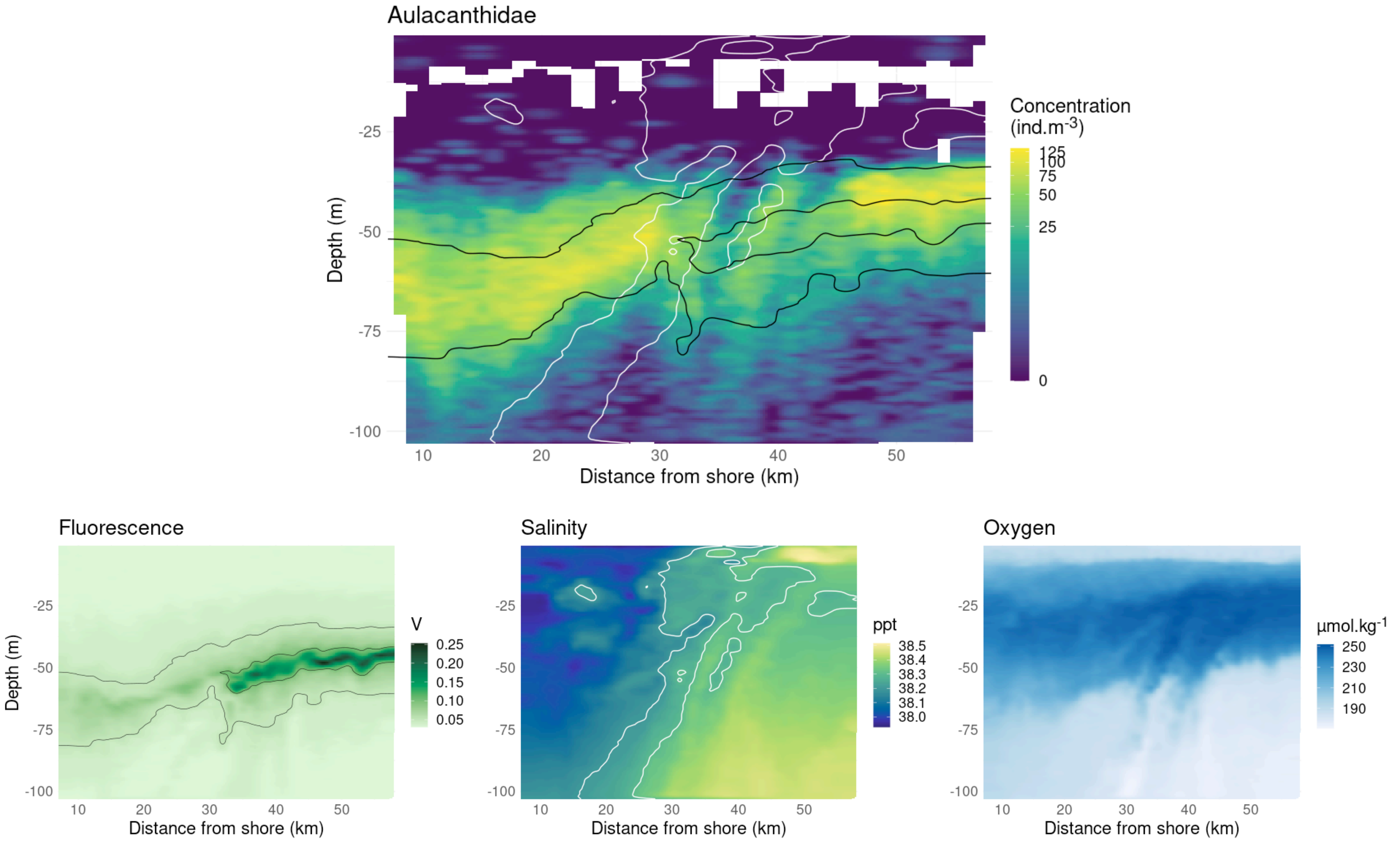
Study	Size range (mesh size)	Depth	Global estimates (PgC)
Moriarty et al., 2012	≥2 mm	0-350 m	0.02
Moriarty and O'Brien, 2013	≥200 µm	0-200	0.19
Buitenhuis et al., 2013	≥200 µm	Integrated	0.33-0.59
Buitenhuis et al., 2013	≥2 mm	0-500 m	0.22-1.52
Hatton et al., 2021	≥200 µm	0-200 m	0.53-31.57
Hatton et al., 2021	≥2 mm	0-200 m	0.02-2.64
This study	≥765 µm - 37.5 mm	0-200 m	0.229



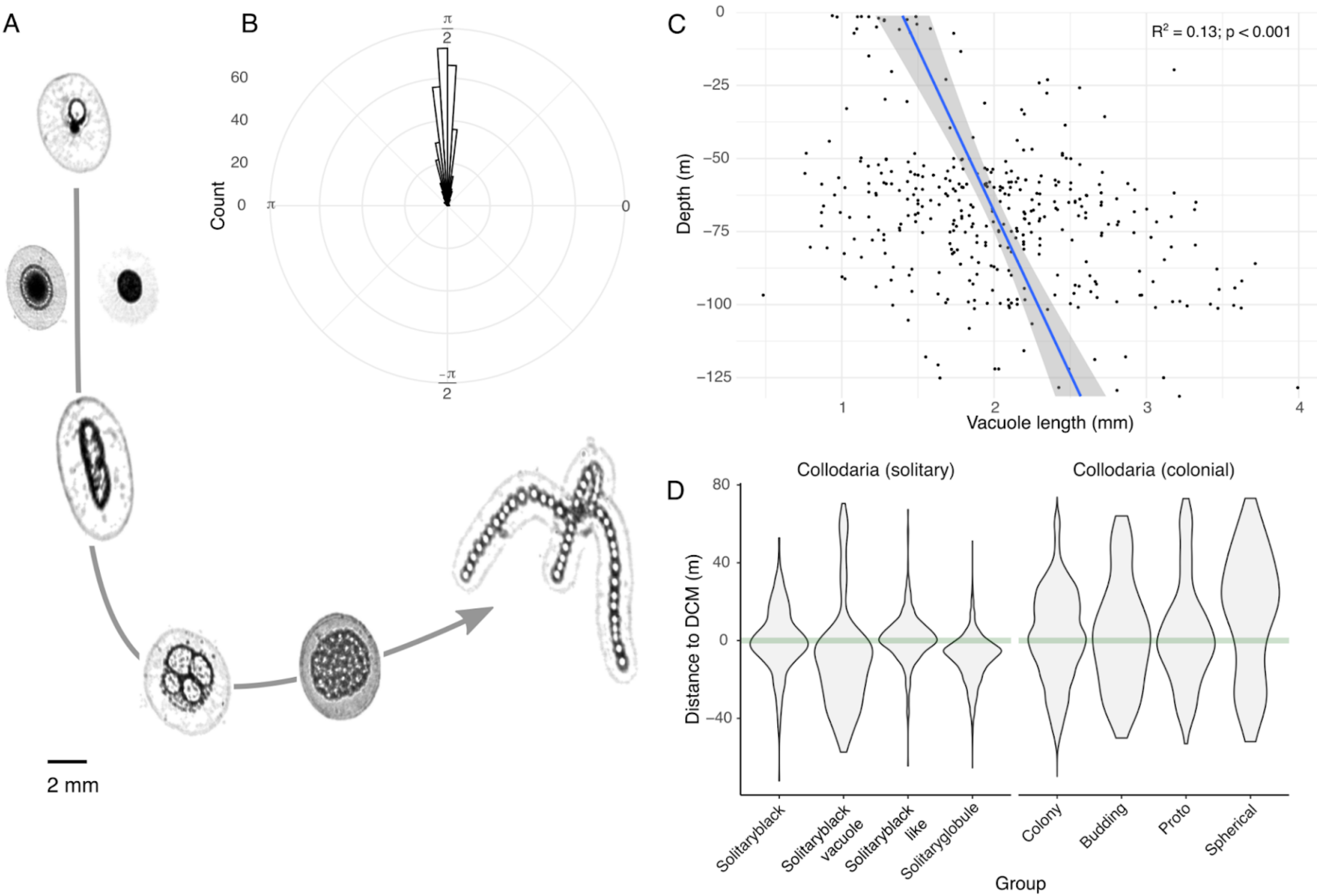


Drago L, Panaïotis T, Irisson J-O, […], Stemmann L, Kiko R (2022) Global Distribution of Zooplankton Biomass Estimated by In Situ Imaging and Machine Learning. Frontiers in Marine Science 9.

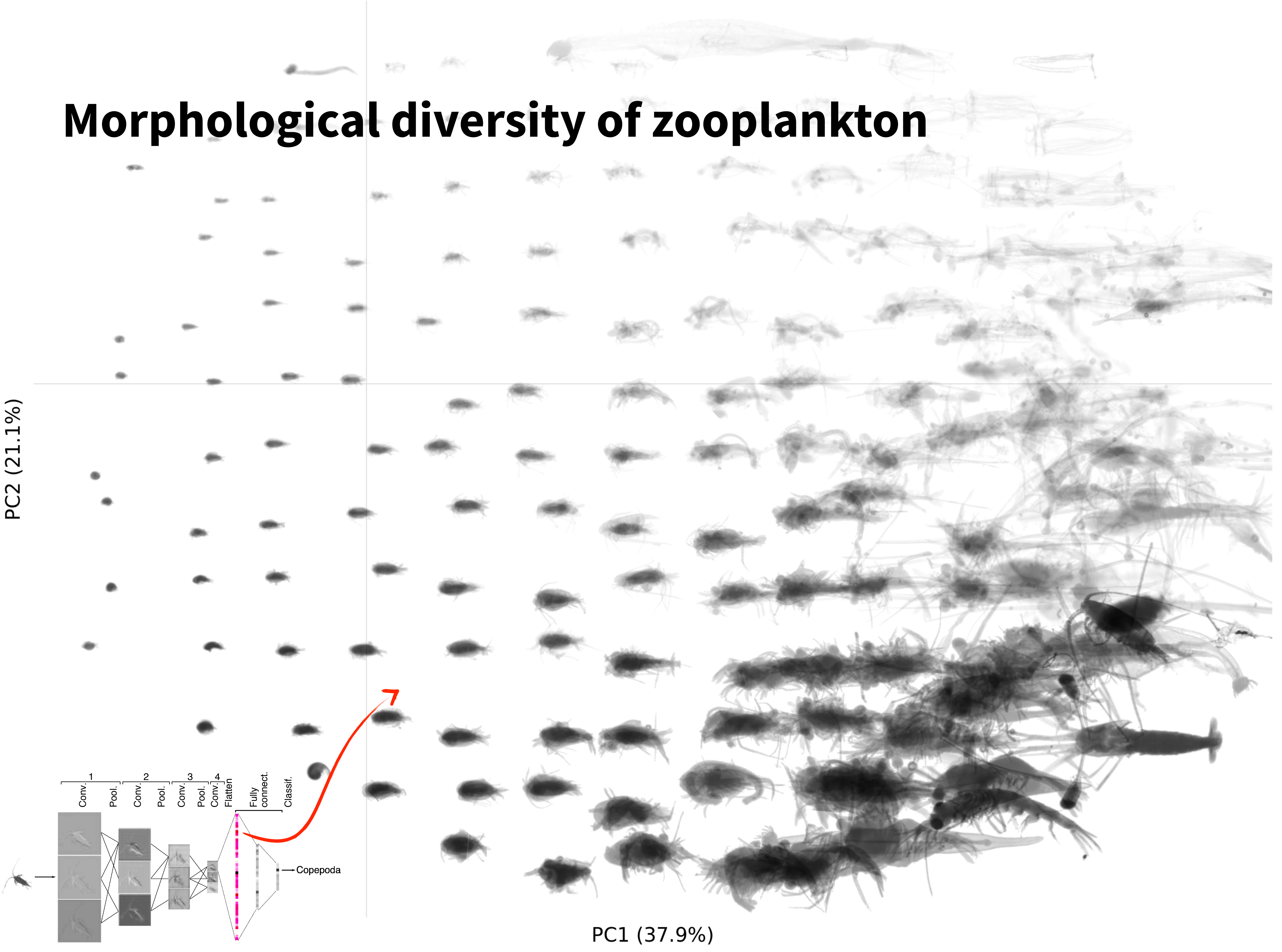
Physical-biological interactions at high resolution



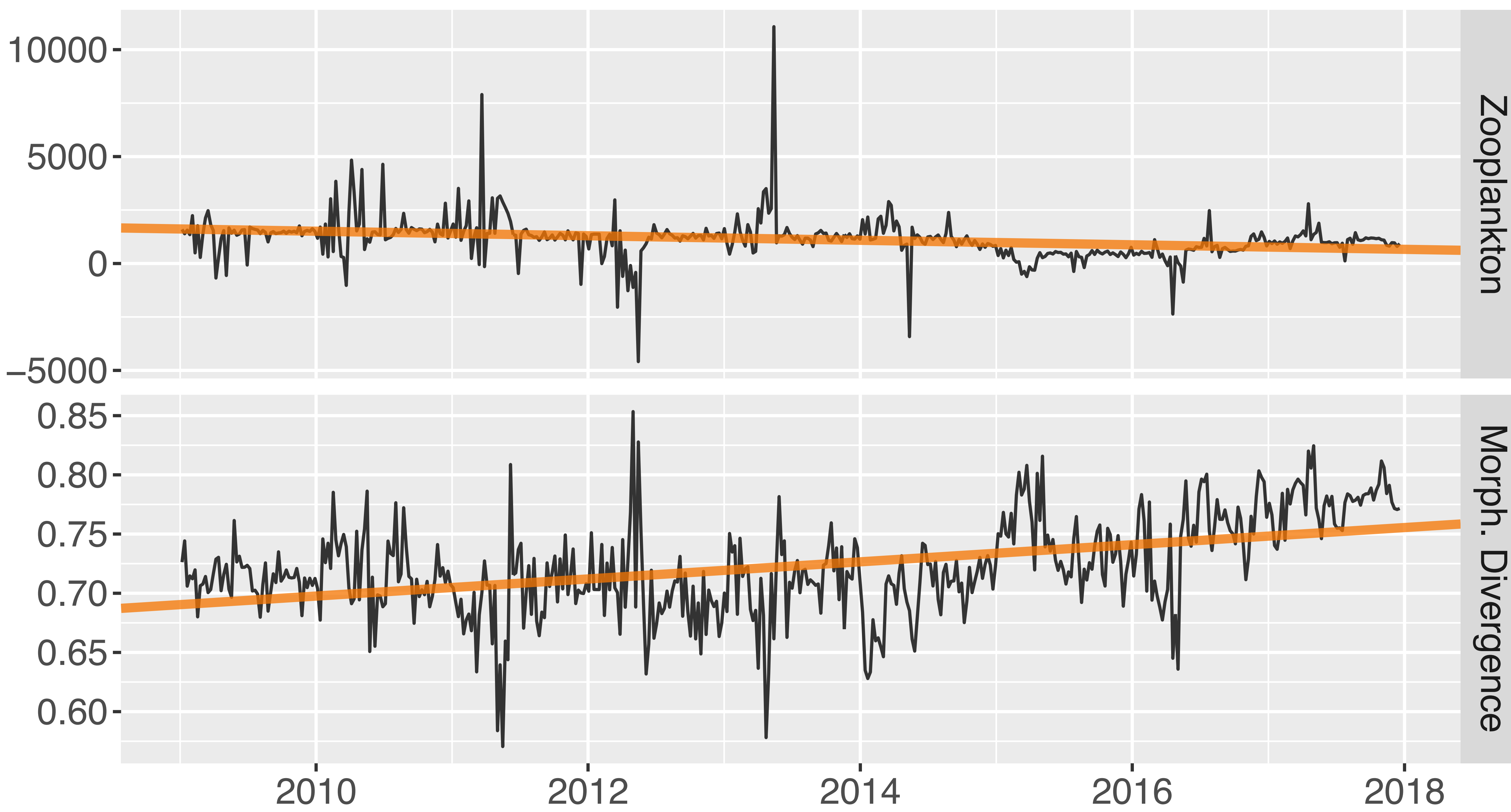
Life cycle and “behaviour” of single cell plankton



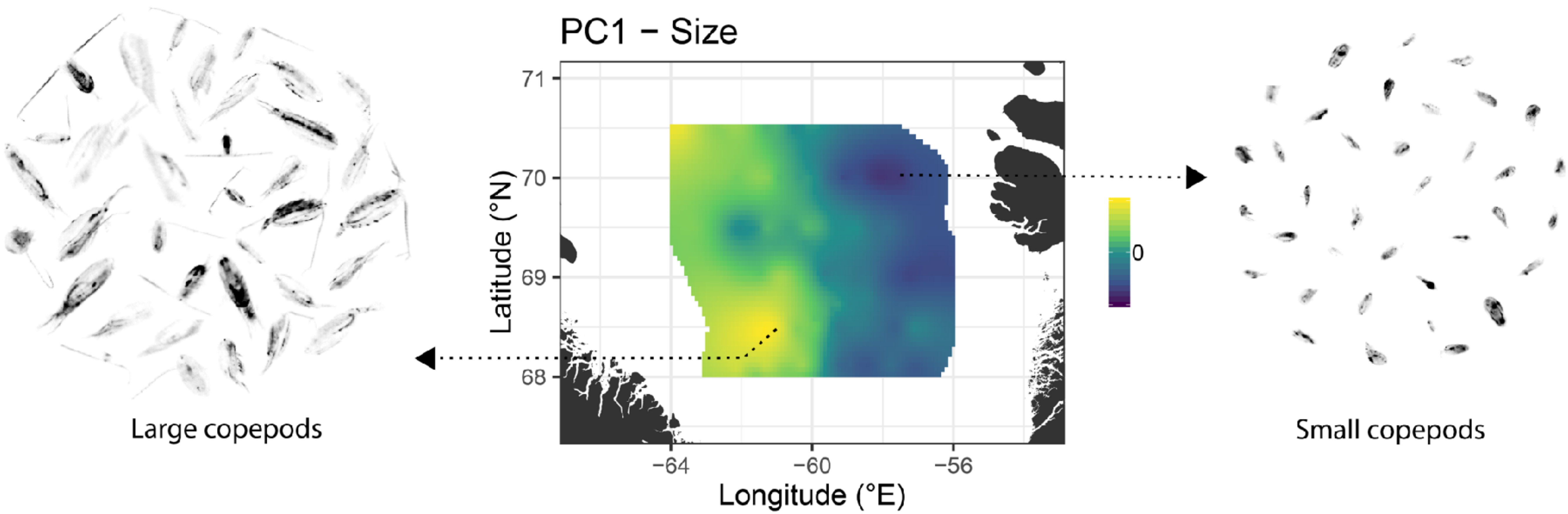
Morphological diversity of zooplankton



Morphological diversity of zooplankton

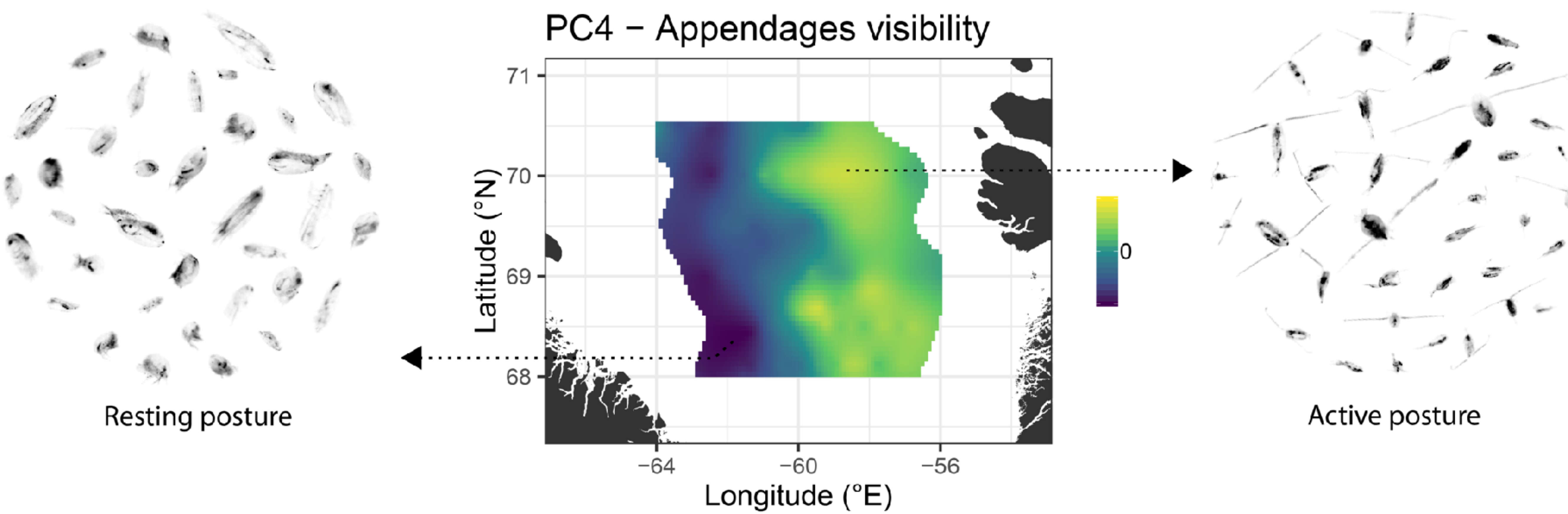


Community-level behaviour from individual-level morphology



Vilgrain L, Maps F, Picheral M, Babin M, Aubry C, Irisson J-O, Ayata S-D (2021) Trait-based approach using in situ copepod images reveals contrasting ecological patterns across an Arctic ice melt zone. Limnology and Oceanography 66:1155–1167.

Community-level behaviour from individual-level morphology



Vilgrain L, Maps F, Picheral M, Babin M, Aubry C, Irisson J-O, Ayata S-D (2021) Trait-based approach using in situ copepod images reveals contrasting ecological patterns across an Arctic ice melt zone. Limnology and Oceanography 66:1155–1167.


Merci