Zoo/PhytoImage

A software for automatic analysis of plankton samples based on R and ImageJ

K. Denis, X. Irigoien, R. François, V. Rousseau, J.-Y. Parent, Ch. Lancelot & Ph. Grosjean

(affiliations, see abstract)

Context (1) We need a *lot* of data!

- Oceans are large, tri-dimensional (four dimensions with time) and distribution of life beings is not homogeneous, especially plankton.

- Thus, any *reasonable* study must gather an *unreasonable* number of samples to draw a good picture of plankton distribution.

SCOR (Scientific committee for Oceanographic Research) currently funds an international group of experts to solve the problem (SCOR WG 130).

Context (2) very large plankton diversity

- Plankton diversity is huge (hundreds of species in a single sample is not an uncommon situation).

- Manual processing of samples is the bottleneck currently.

  We need to automatize processing of plankton samples.
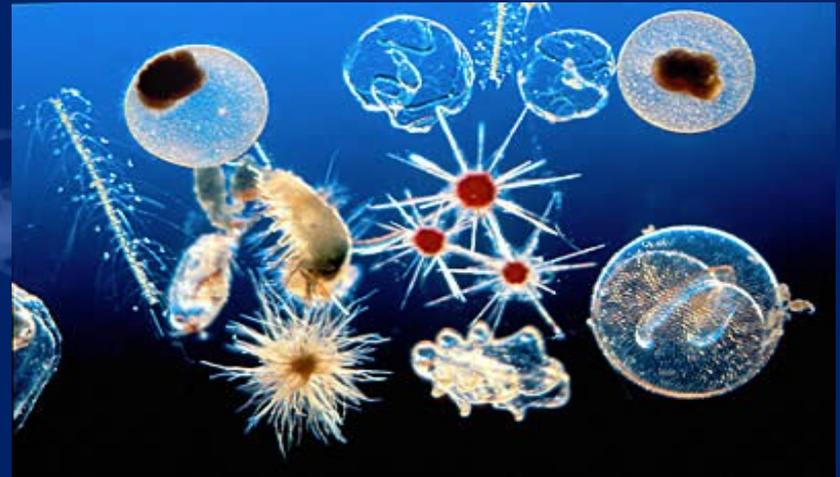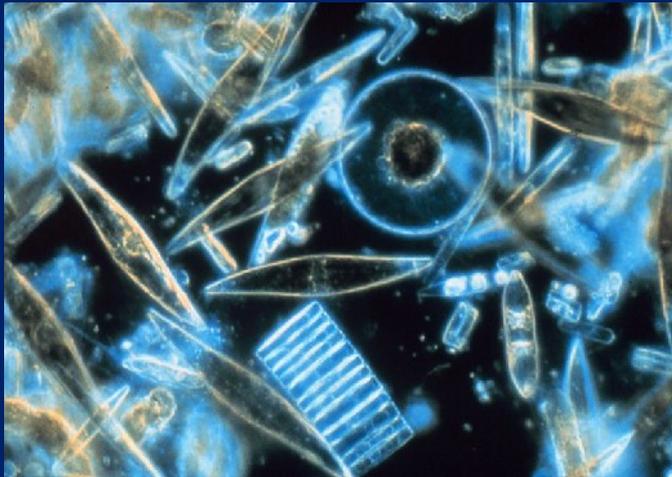
Context (3) current technology

- Genetic analyses give only (but detailed!) qualitative data.

- Particle counters give only quantitative data.

Image analysis, combined with machine learning is considered as a potential solution to get both qualitative and quantitative data.
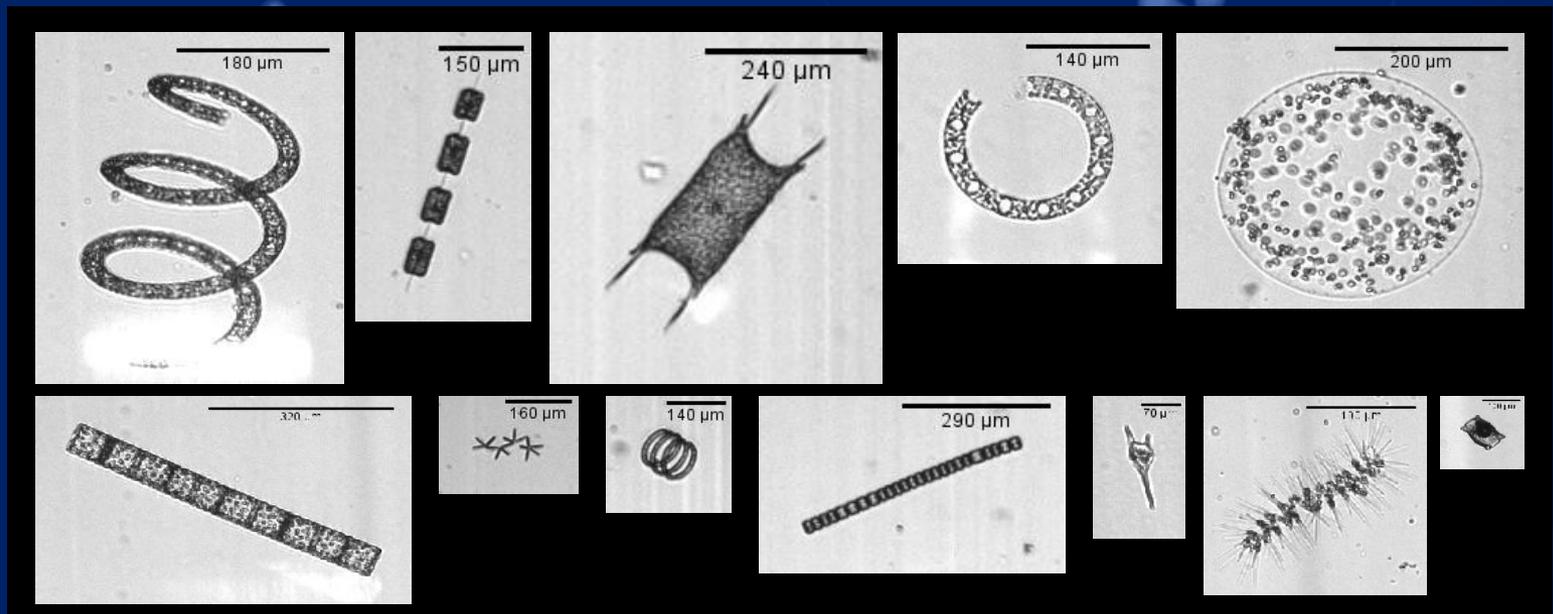
Context (4) studies

- Belgium, project AMORE III, development of a system to monitor phytoplankton continuously aboard the oceanographic ship "Belgica", funded by Belgian Science Policy.

- France, IFREMER, survey of phytoplankton in coastal waters (RePhy network).

- Spain, AZTI, spatial and temporal distribution of surface zooplankton in the Bay of Biscay.

- ...

What we need: these maps are much more informative and appealing...

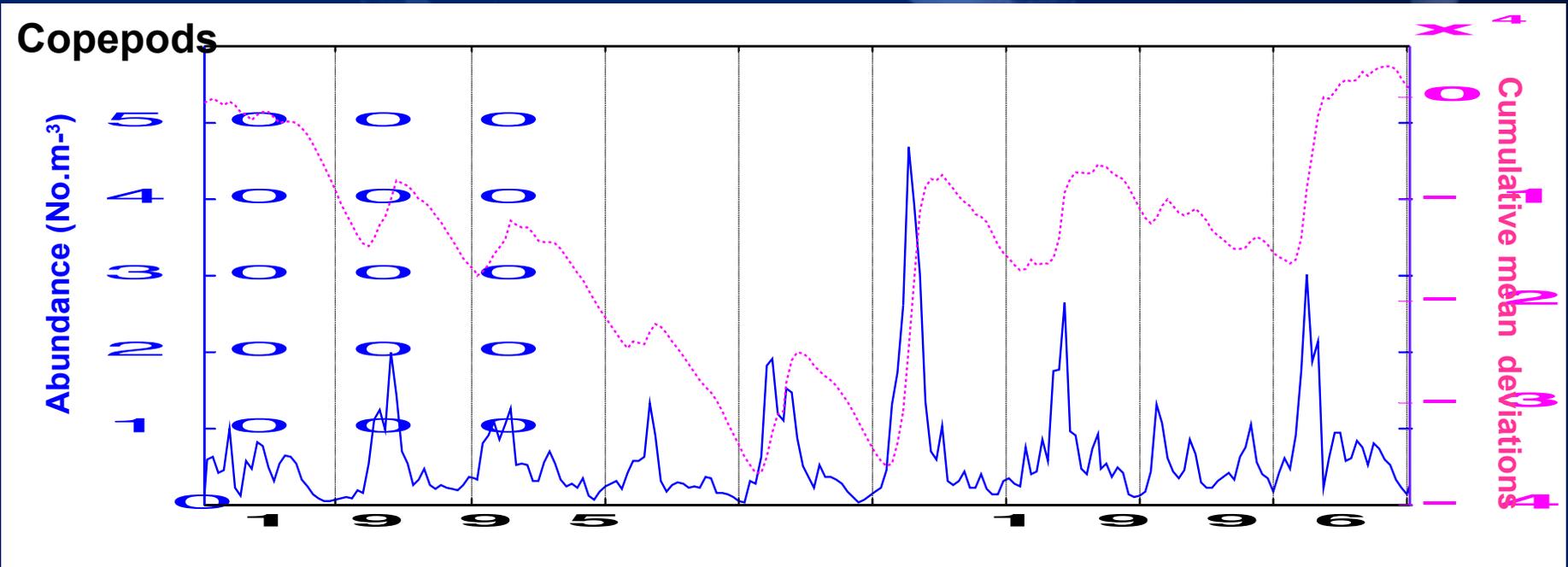Zooplankton abundance in the Bay of Biscay at two different dates.

Note the large number of stations sampled (black dots) and patchiness in the distribution (blue = low, yellow = high)
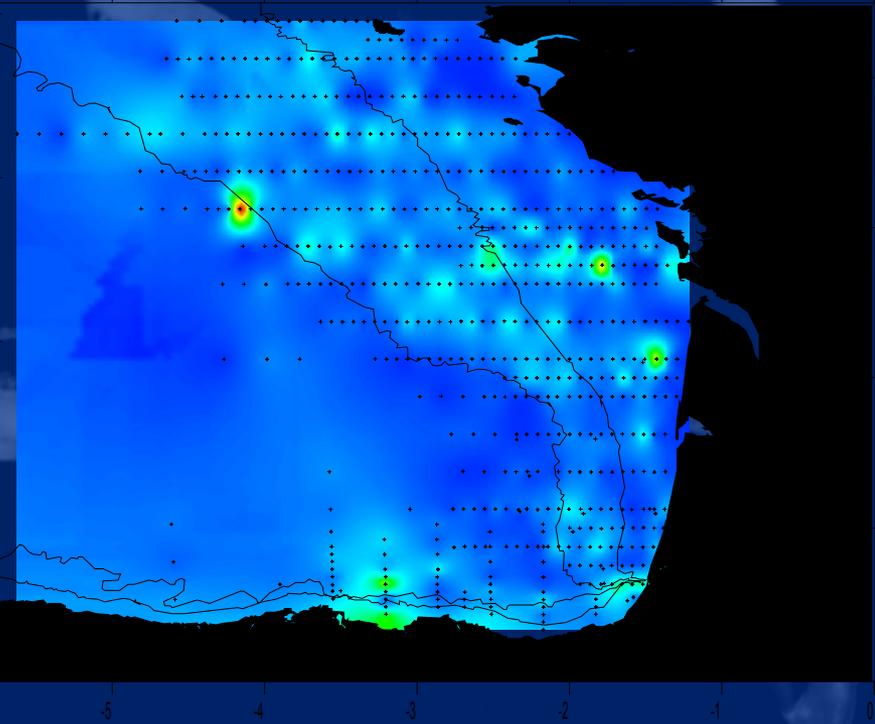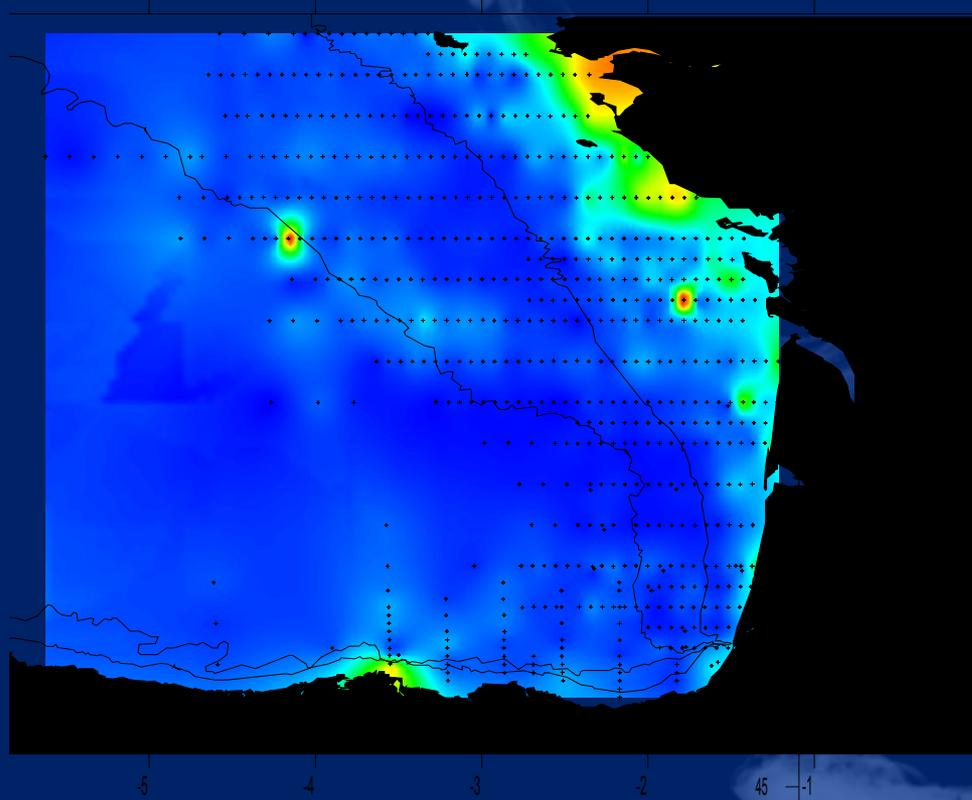
Data and graphs: X. Irigoien et al, AZTI. Automatic analysis of plankton digital images

... than this plot

Copepods

- Abundance of copepods (major zooplankton critters) at one station during 8 years

- Manual processing of the samples by a PhD student in roughly the same time as required for the previous plots!

The Zoo/PhytoImage software

http://www.sciviews.org/zooimage

Zoo/PhytoImage - What is it?

- Specialized software to analyse digital images of plankton, but can be used in other contexts (insects, bacteria, etc.)

-

- Open source (GPL 2 for most part of it), runs on Windows, Mac OS X and Linux

- R (statistics) and Java (image analysis, ImageJ)

- Image analysis, features extraction, vignettes extraction (all data stored in a single file per sample)

- Training set building with multiple levels for grouping and automatic classifier generation/analysis

- Calculation of ecologically meaningful parameters out of the samples after automatic classification in batch mode (numbers, size spectra, biovolumes/biomasses per group)

Zoo/PhytoImage – General overview

- Image analysis (of plankton) is a complex task

- It requires a complex and specialized software

- Data process flowchart in Zoo/PhytoImage
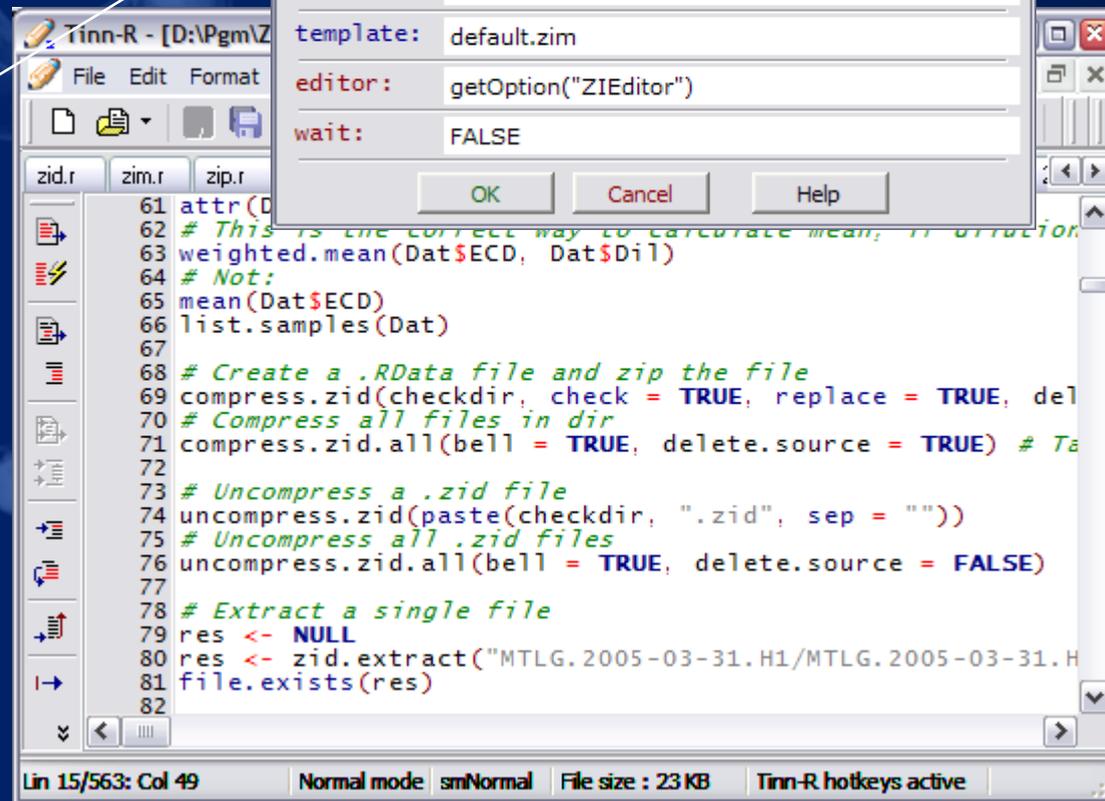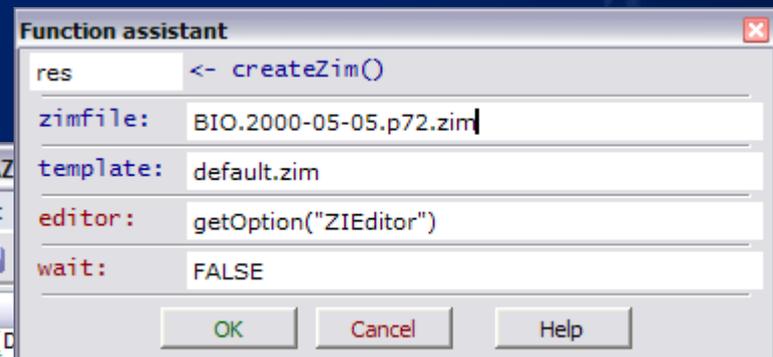
(diagram by Ben Tupper)
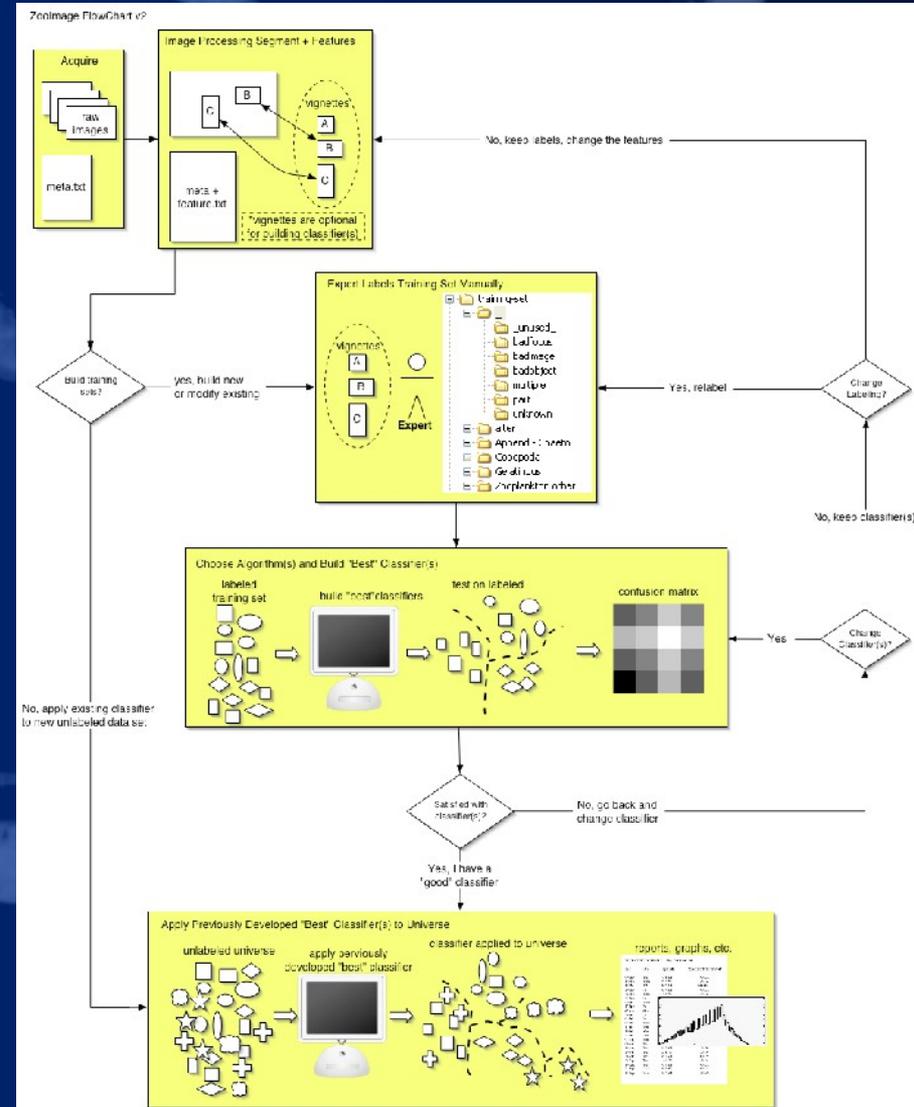
Accessibility

- Three levels, three user interfaces:
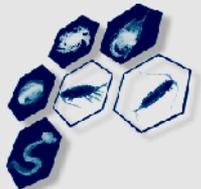
1. Full point&click assistant (complete beginner),

2. Dialog boxes prompting for command options (intermediate user),

3. Toolbox of functions for pure programmation (developers)

Current performances & future

- 20-30 taxa with 75-80% global accuracy on an average training set

- ZooImage analyses almost any kind of image, except Zooscan and VPR (for the moment)

- Developement highly dependent on contributions and actual use, but better algorithms, error correction, more graphs, more ecologically meaninful derived variables.

Thank you for your attention