French Research Institute for Exploitation of the Sea

Ifremer

5th Zooplankton symposium, Pucón Chile 14-18 March 2011, Workshop 5

Comparison of traditional microscopy and digitalized images analysis to identify and delineate pelagic fish egg spatial distribution

Elvire Antajan, Stéphanie Lelièvre, Sandrine Vaz,

IFREMER, BP 699, 62200 Boulogne sur Mer, France

This work was funded by the EU under the INTERREG 4a scheme and ERDF

CHARM
CHANNEL INTEGRATED APPROACH FOR MARINE RESOURCE MANAGEMENT

Goal and objectives

Study the winter distribution of spawning areas in the Channel and southern North Sea

Problems:

– High number of samples is needed

– Identification criteria are limited (size, pigmentation, oil globules, need to match with the species fished)

– Lack of experts, time consuming

– Sample archiving (formalin toxicity issue, need storage space)

Objectives:

– Digitalization of samples with the ZooScan integrated system

– Building a classification model for automatic identification of fish egg species

– Compare maps of spawning areas obtained with ZooScan identifications and microscope identifications by taxonomist

Study area – Annual IBTS survey

International Bottom Trawl Survey

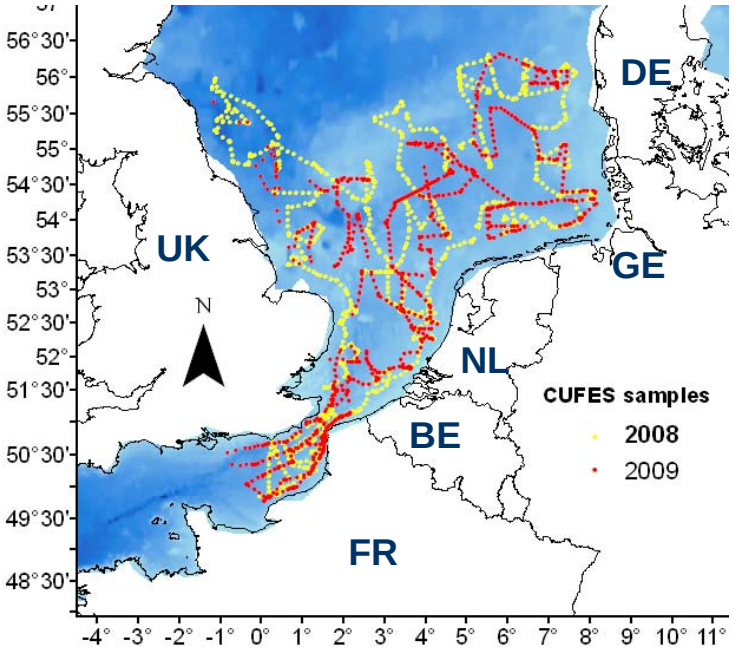
– fish abundance & distribution

– recruitment indices

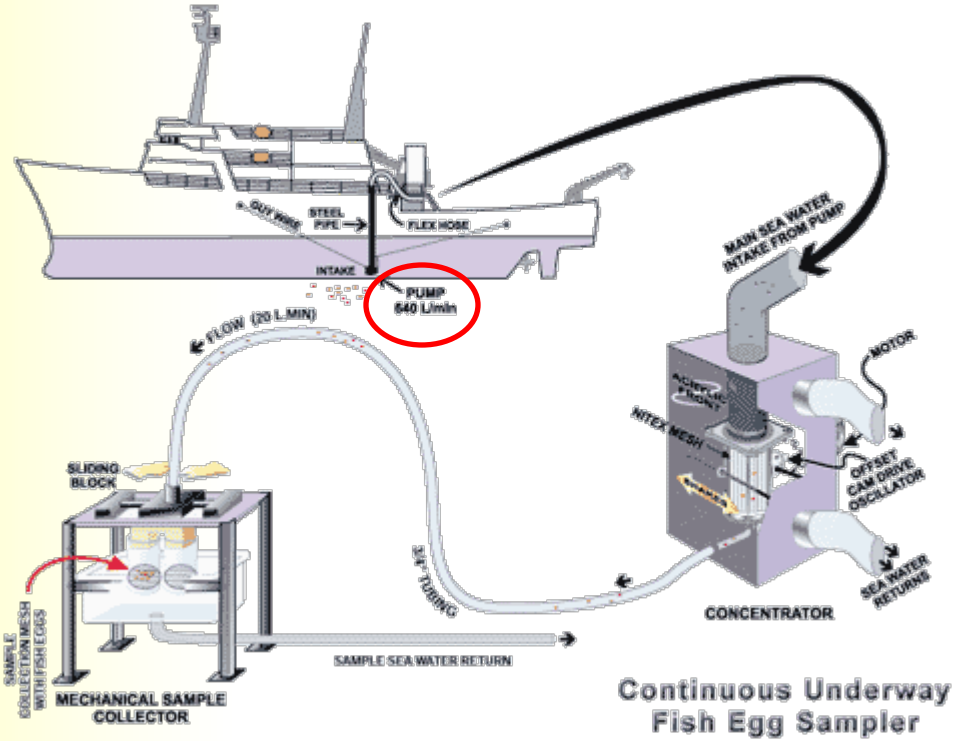
IBTS 2008 & 2009

Eastern Channel & Southern North Sea

15 January – 15 February

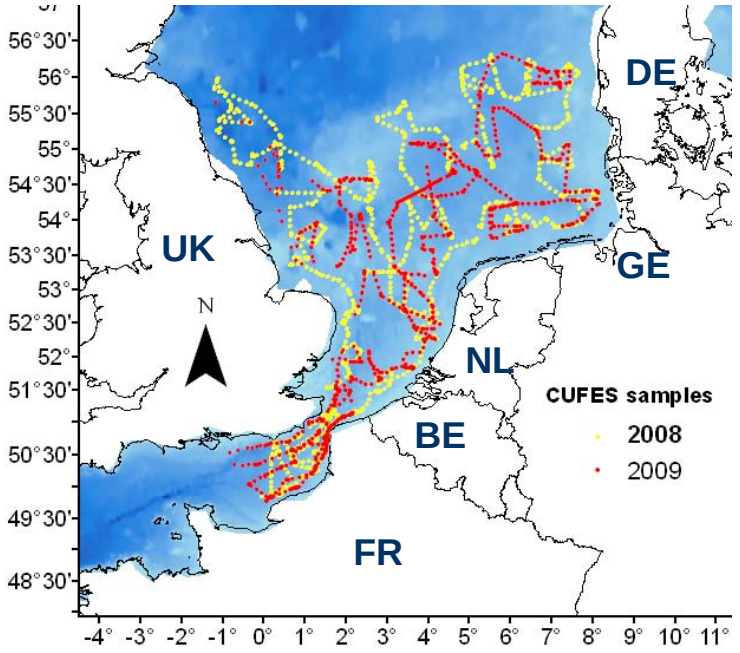


Continuous Underway Fish Egg Sampler (CUFES)



CUFES: a pumping device to collect pelagic eggs of fish from a moving vessel

- water pumped at 5 m depth
- collector mesh size: 500 µm

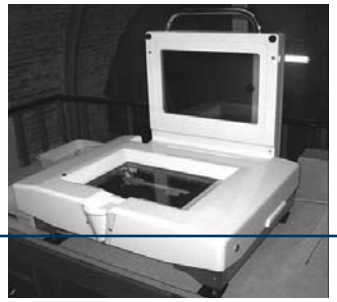


The CUFES operated continuously during the survey :

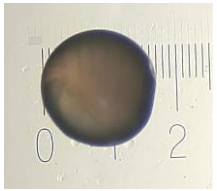
- sequential sampling interval: 30 min
- >1000 samples collected

Sample identification

	CUFES 2008	CUFES 2009
Microscope	1 048 samples	1 103 samples
Image analysis (ZooScan)	1/3 analyzed	All analyzed
DNA analysis	1/3 analyzed	

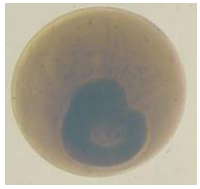


Morphological criteria for microscope identification

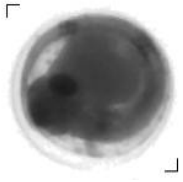


Egg size

Oil globule



Stage



Pigmentation

Sample processing with ZooScan

http://www.zooscan.com

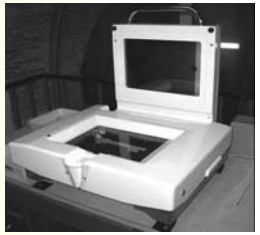


Image acquisition

↓

Image process & Feature extraction

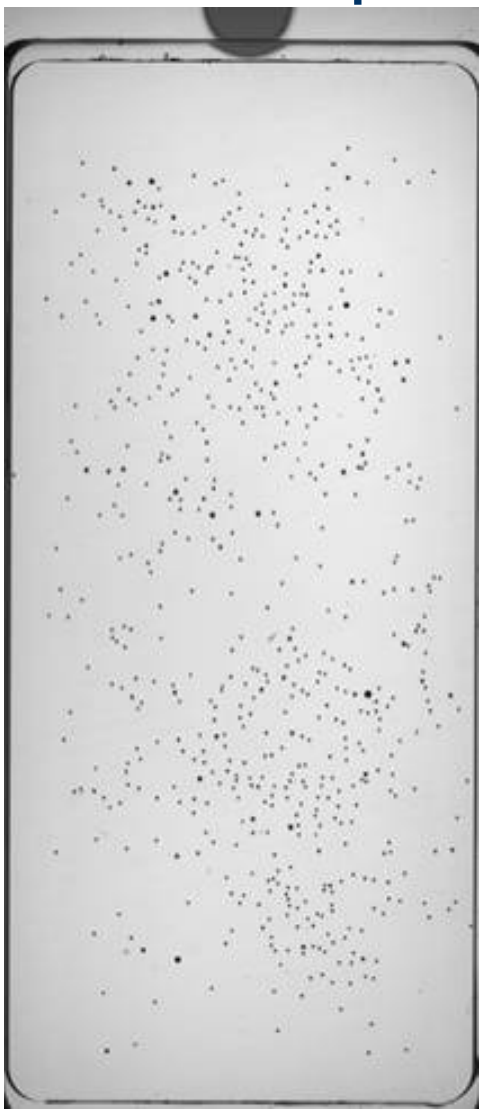
↓

Learning set building

↓

Validation of classifier performances on an **independent Test set**

CUFES sample



Sample processing with ZooScan

http://www.zooscan.com

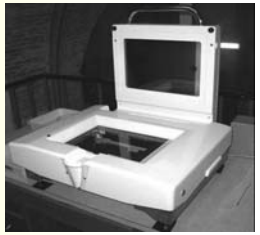


Image acquisition

↓

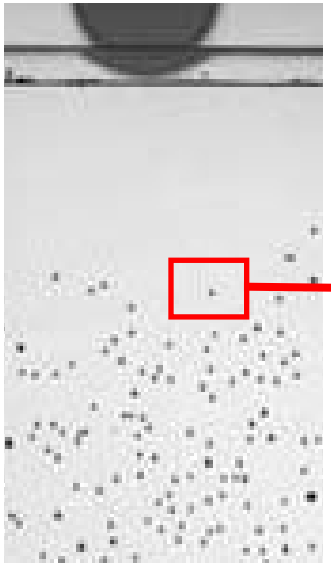
Image process & Feature extraction

↓

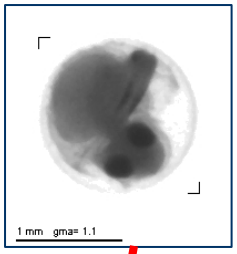
Learning set building

↓

Validation of classifier performances on an **independent Test set**



vignette



Text file

!Item	Label	Area	StdDev	Perim.	Major	Minor	...	Ident
59	cufes100_1	11073	62.72	398.37	120.50	117.00	...	11_Sole_stA
61	cufes100_1	14940	50.00	474.01	142.27	133.70	...	10_Cod_stA
221	cufes1000_1	26692	42.48	638.62	186.02	182.70	...	13_Plaice_stA
227	cufes1000_1	25269	58.75	622.52	181.32	177.44	...	13_Plaice_stB
172	cufes1000_1	12542	57.18	554.42	127.54	125.21	...	11_Sole_stA
163	cufes1000_1	12345	61.71	655.23	129.13	121.72	...	10_Cod_stA
182	cufes1000_1	12671	63.48	629.09	127.80	126.24	...	10_Cod_stB
192	cufes1000_1	11281	62.57	611.54	121.67	118.05	...	10_Cod_stB
167	cufes1000_1	10879	55.38	559.21	119.25	116.16	...	09_Whiting_stA
106	cufes1000_1	5979	47.02	314.33	87.68	86.82	...	01_Dab_stB
116	cufes1000_1	5112	42.56	281.46	82.21	79.18	...	06_Rockling_stA

Sample processing with ZooScan

http://www.zooscan.com

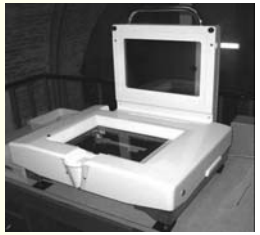


Image acquisition

↓

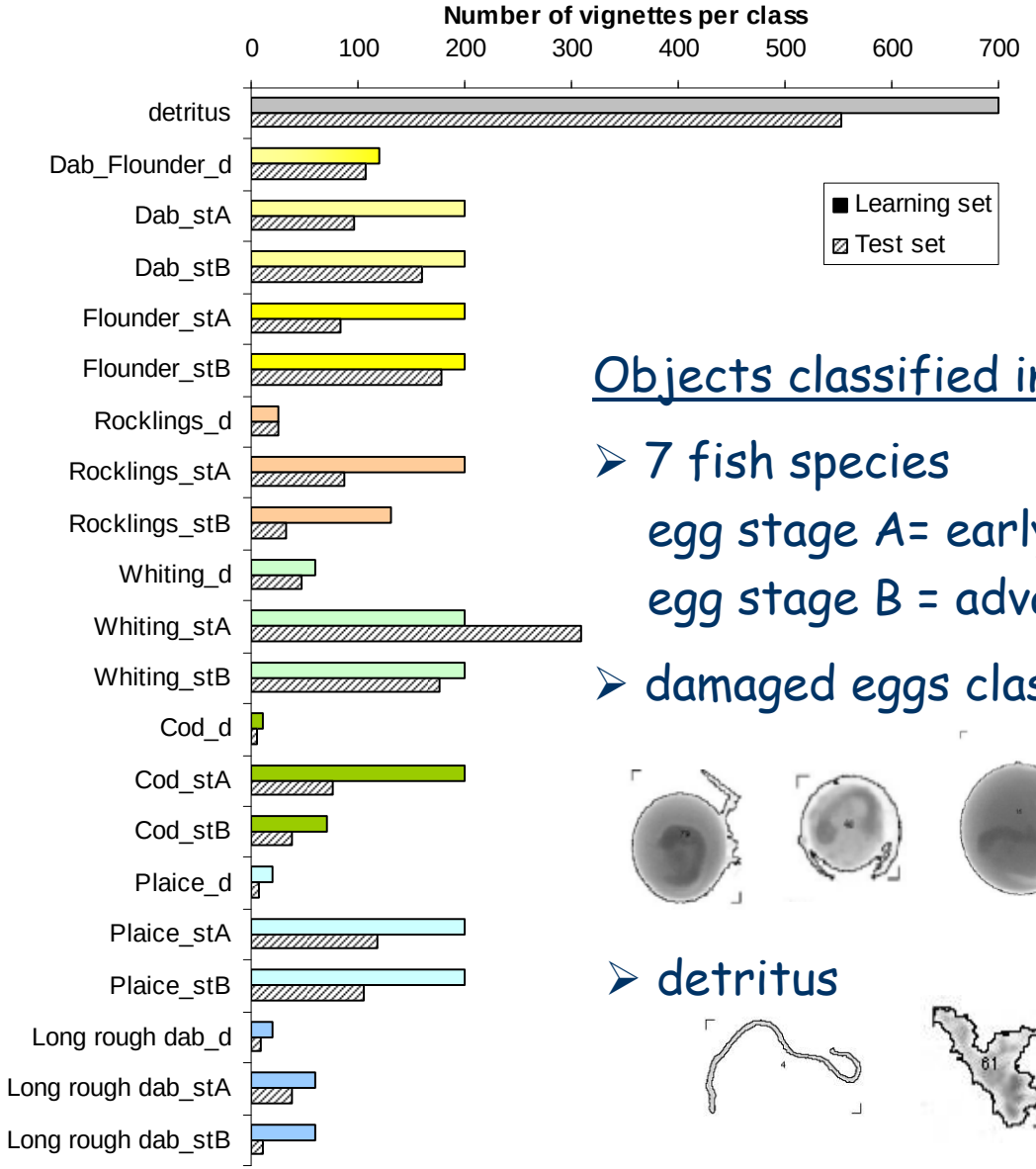
Image process & Feature extraction

↓

Learning set building

↓

Validation of classifier performances on an **independent Test set**

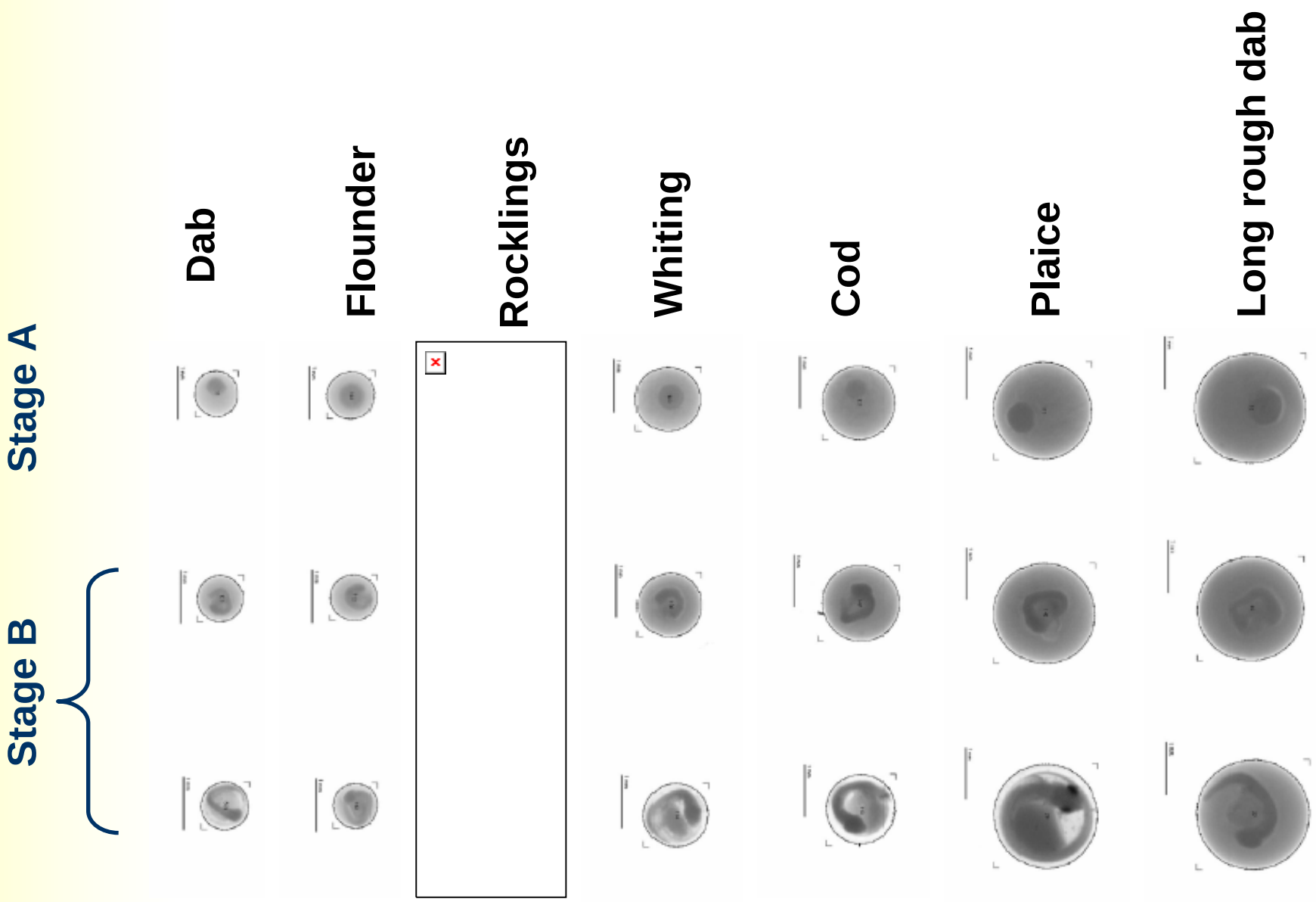


Objects classified into groups:

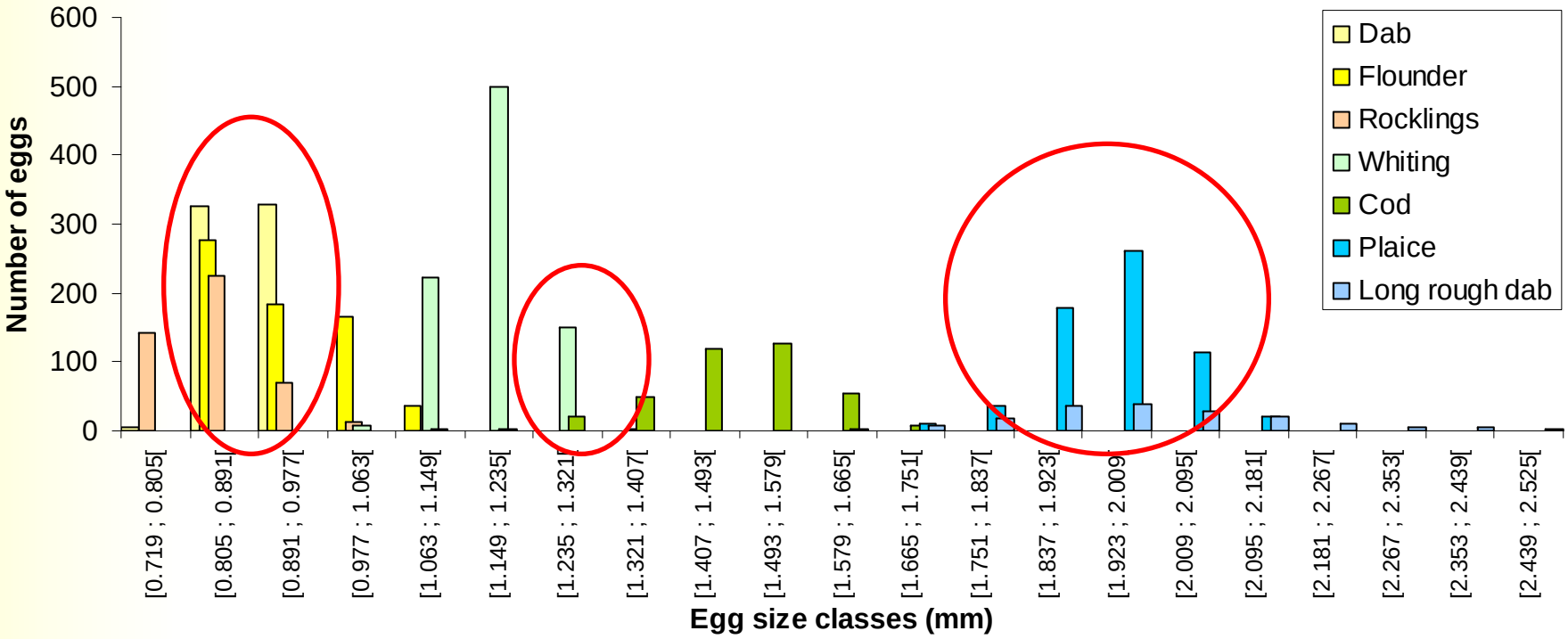
- 7 fish species
  - egg stage A= early stages
  - egg stage B = advanced stages
- damaged eggs classes

- detritus

The 7 species selected



Size classes analyses of fish egg species



Egg size => confusion risk for 3 species groups

- dab, flounder and rocklings
- cod and whiting
- plaice and long rough dab

Sample processing with ZooScan

http://www.zooscan.com

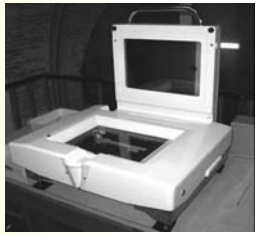


Image acquisition

↓

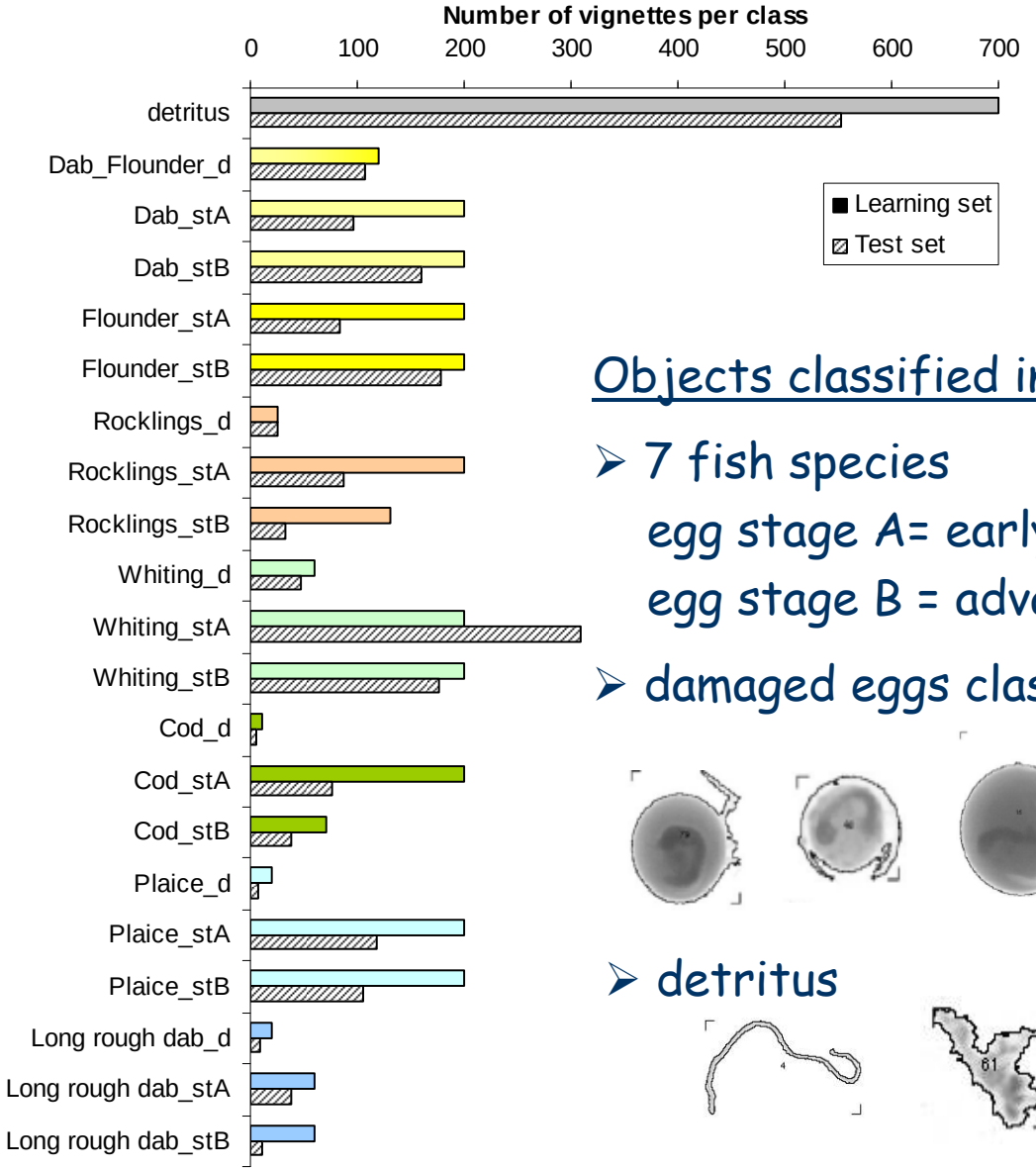
Image process & Feature extraction

↓

Learning set building

↓

Validation of classifier performances on an **independent Test set**



Objects classified into groups:

- 7 fish species
  - egg stage A= early stages
  - egg stage B = advanced stages
- damaged eggs classes

- detritus

Sample processing with ZooScan

http://www.zooscan.com

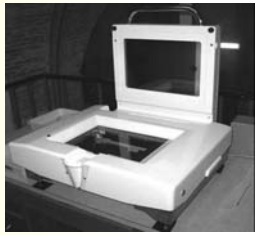


Image acquisition

↓

Image process & Feature extraction

↓

Learning set building

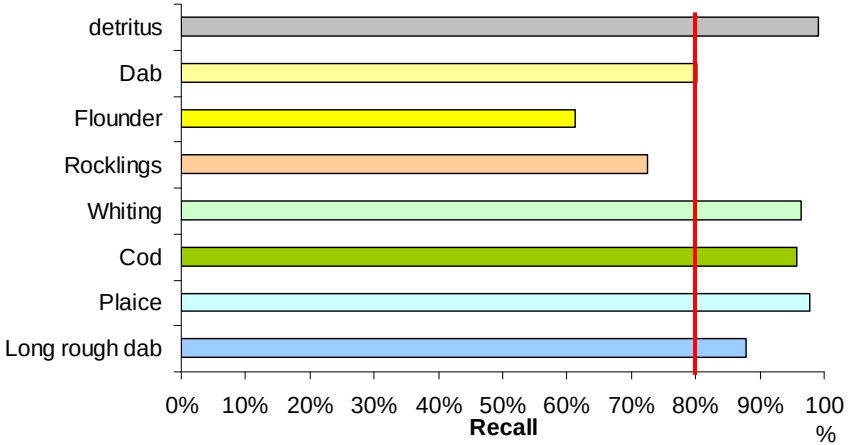
↓

Validation of classifier performances on an **independent Test set**

Error rate = 0.13

Confusion matrix

	Detritus	Dab & Flounder_d	Dab	Flounder	Rockling_d	Rockling	Whiting_d	Whiting	Cod_d	Cod	Plaice_d	Plaice	Long rough dab_d	Long rough dab	Total in test set	Recall	Contamination
Detritus	547	2	0	0	3	0	0	0	0	0	0	0	0	0	552	0.99	0.01
Dab & Flounder_d	6	82	4	11	1	3	1	0	0	0	0	0	0	0	108	0.76	0.25
Dab	0	0	206	35	0	16	0	0	0	0	0	0	0	0	257	0.80	0.32
Flounder	0	4	80	161	0	16	0	2	0	0	0	0	0	0	263	0.61	0.30
Rockling_d	0	15	0	0	4	6	0	0	0	0	0	0	0	0	25	0.16	0.56
Rockling	0	1	15	16	1	87	0	0	0	0	0	0	0	0	120	0.73	0.32
Whiting_d	0	5	0	0	0	0	38	3	0	1	0	0	0	0	47	0.81	0.12
Whiting	0	0	0	8	0	0	4	468	0	5	0	0	0	0	485	0.96	0.02
Cod_d	0	0	0	0	0	0	0	0	2	0	2	0	1	0	5	0.40	0.00
Cod	0	0	0	0	0	0	0	3	0	109	0	2	0	0	114	0.96	0.08
Plaice_d	0	0	0	0	0	0	0	0	0	0	5	2	0	0	7	0.71	0.44
Plaice	0	0	0	0	0	0	0	0	0	2	2	219	0	1	224	0.98	0.04
Long rough dab_d	0	0	0	0	0	0	0	0	0	0	0	0	7	2	9	0.78	0.13
Long rough dab	0	0	0	0	0	0	0	0	0	1	0	5	0	43	49	0.88	0.07
Total in prediction set	553	109	305	231	9	128	43	476	2	118	9	228	8	46	2265		



Sample processing with ZooScan

http://www.zooscan.com

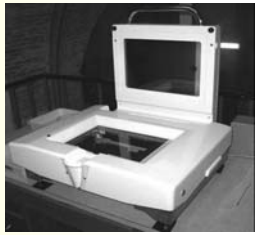


Image acquisition

↓

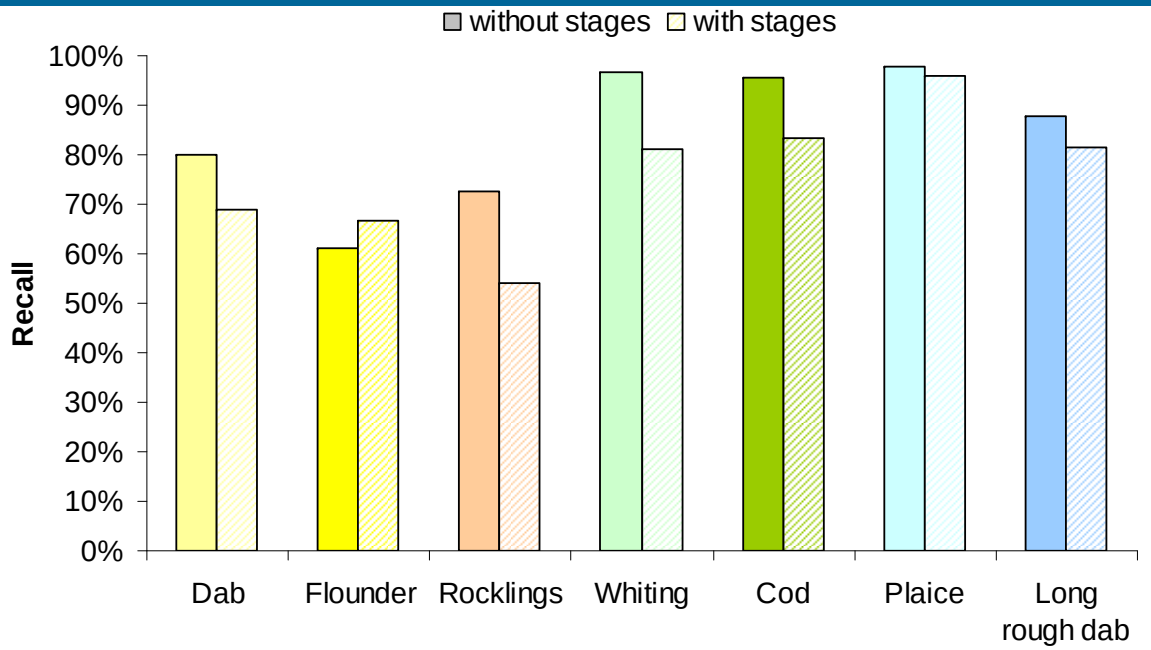
Image process & Feature extraction

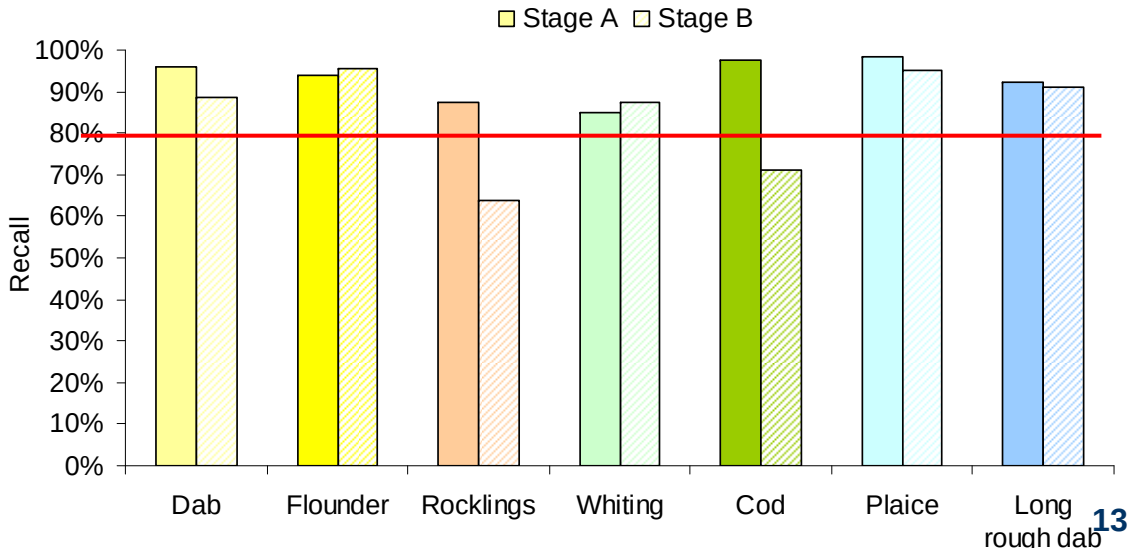
↓

Learning set building

↓

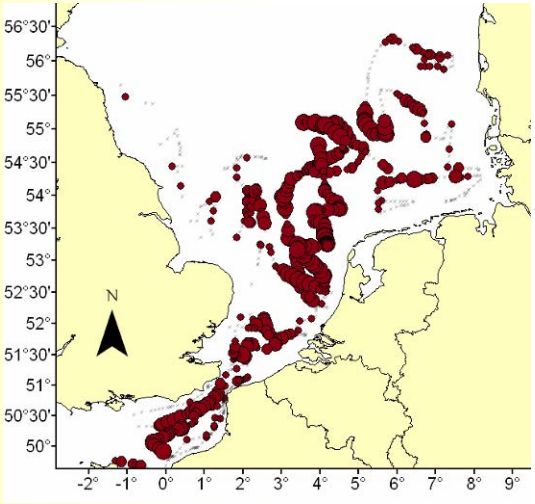
Validation of classifier performances on an **independent Test set**





Geostatistics: abundance data mapping

Principle: spatial auto-correlation described by the variogram

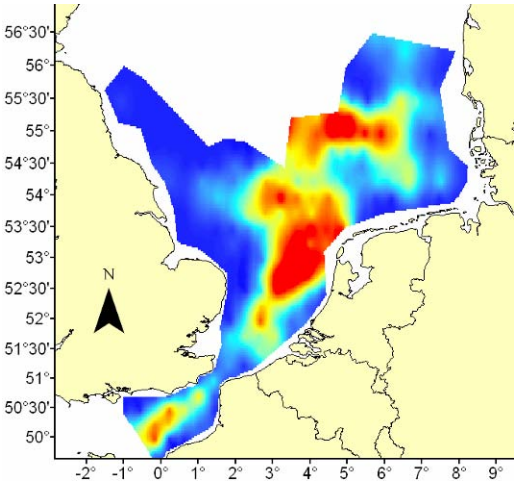


Station information

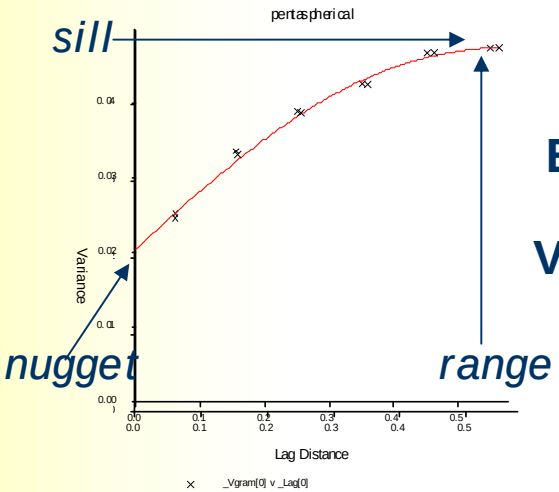
– Position

– Abundances

Interpolated spatial distribution



Kriging



Experimental variogram

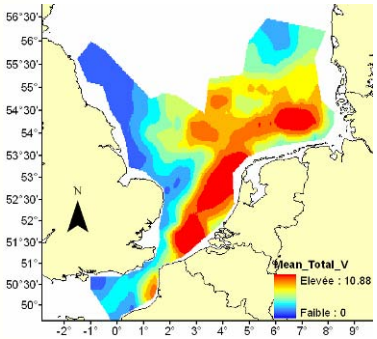
It describes how the abundance varies as the function of the distance

Variogram model Adjustment

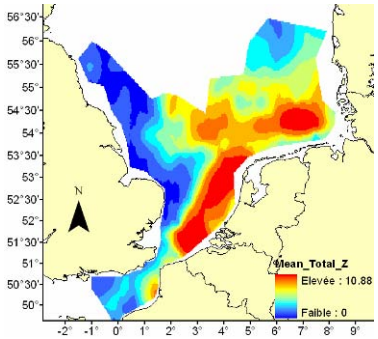
Eggs distribution mapping (2009)

All species

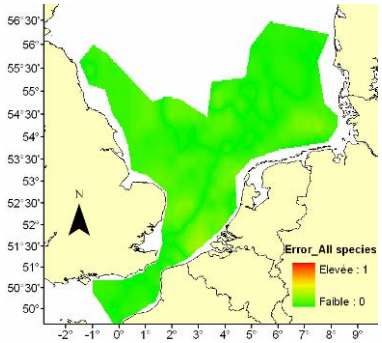
Microscope



Zooscan

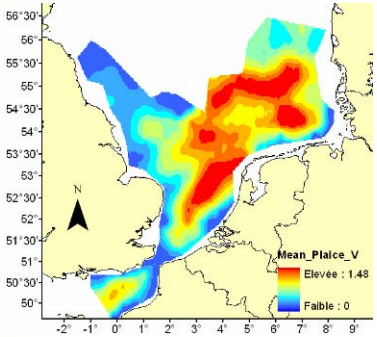
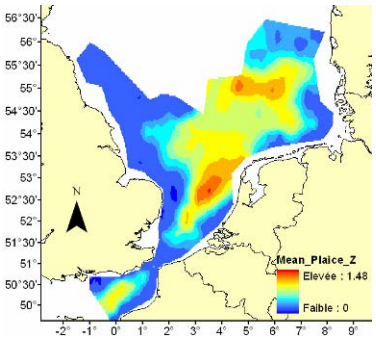
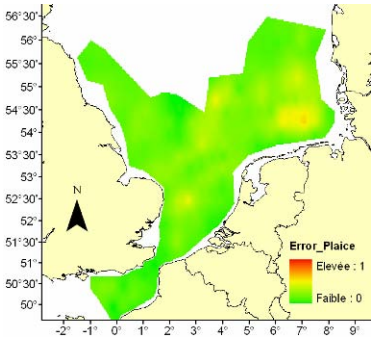
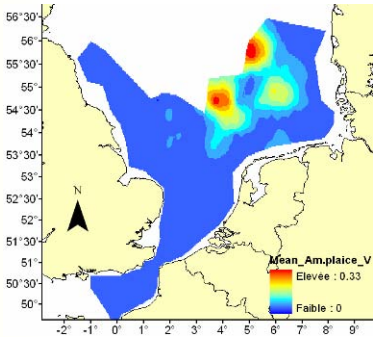
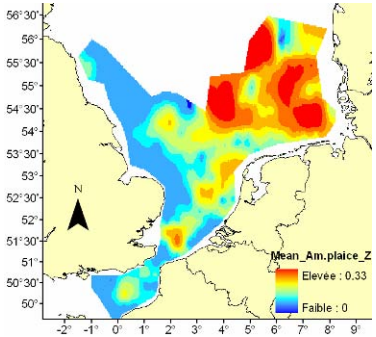
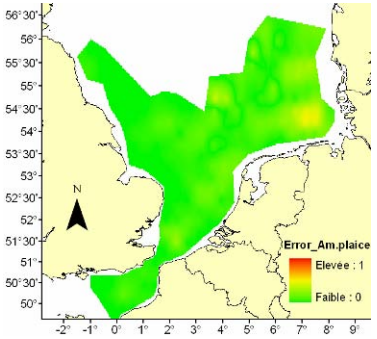


Zonal relative error

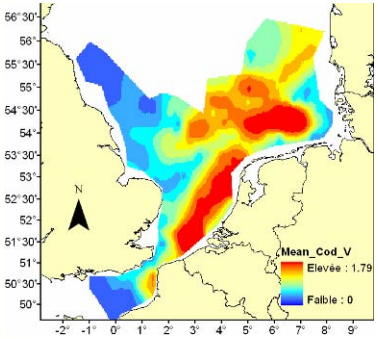
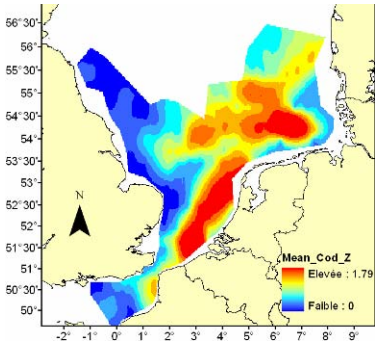
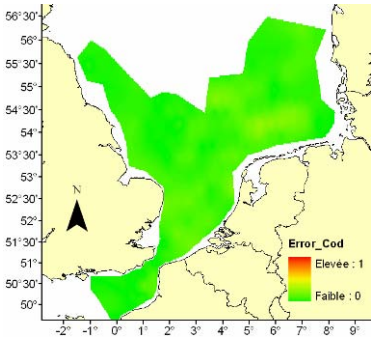
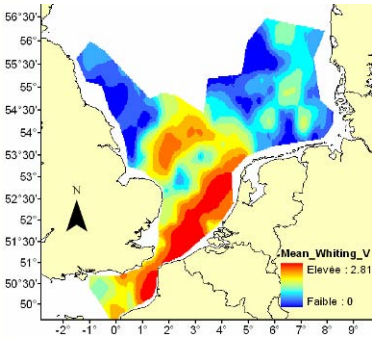
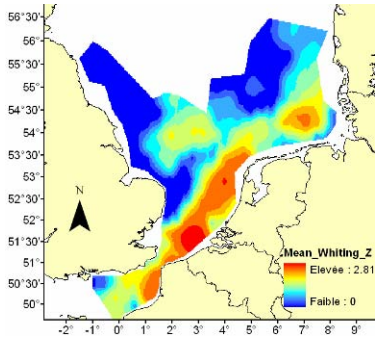
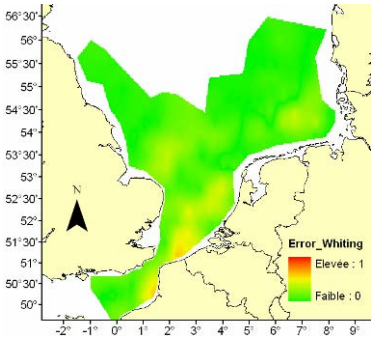


$$\text{Zonal relative error} = \frac{|N\mu - Nz|}{N\,\mu max}$$

Eggs distribution mapping (2009)

	Microscope	Zooscan	Zonal relative error
Plaice			
Long rough dab			

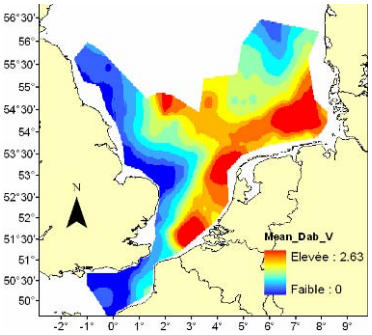
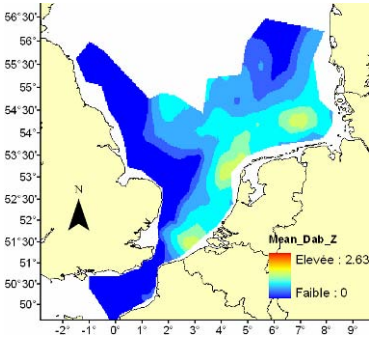
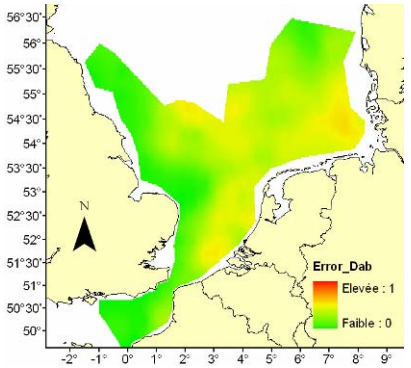
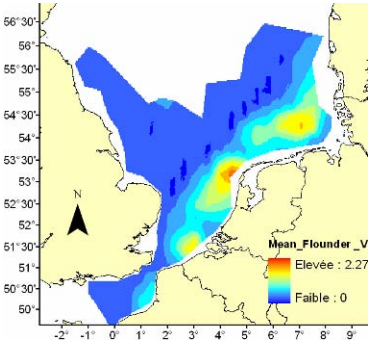
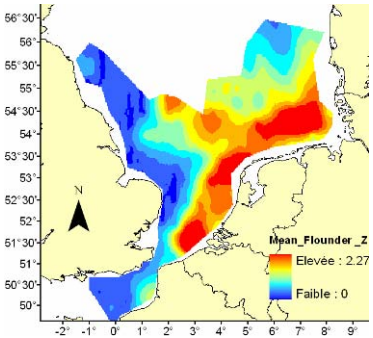
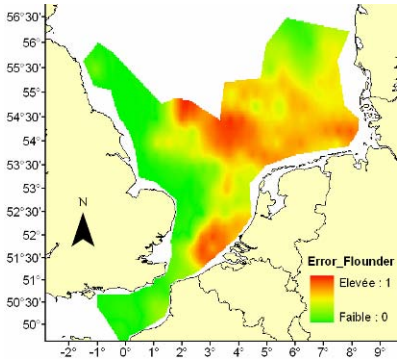
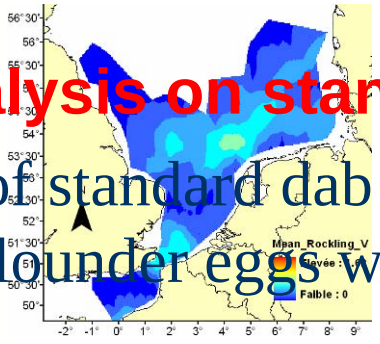
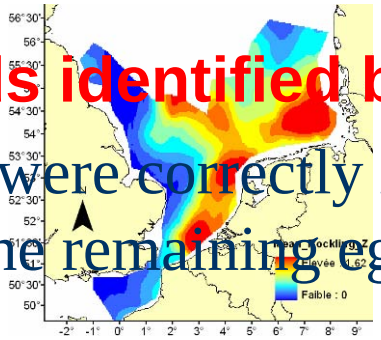
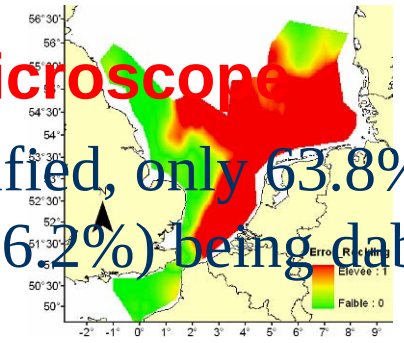
Eggs distribution mapping (2009)

	Microscope	Zooscan	Zonal relative error
Cod			
Whiting			

DNA analysis on standards identified by microscope

98% of standard cod eggs were correctly identified but only 71% of standard whiting eggs were, the remaining (29%) being of cod eggs.

Eggs distribution mapping (2009)

	Microscope	Zooscan	Zonal relative error
Dab			
Flounder			
Rocklings			

DNA analysis on standards identified by microscope

If 49.5% of standard dab eggs were correctly identified, only 63.8% of s flounder eggs were, the remaining eggs (36.2%) being dab eggs.

Conclusions

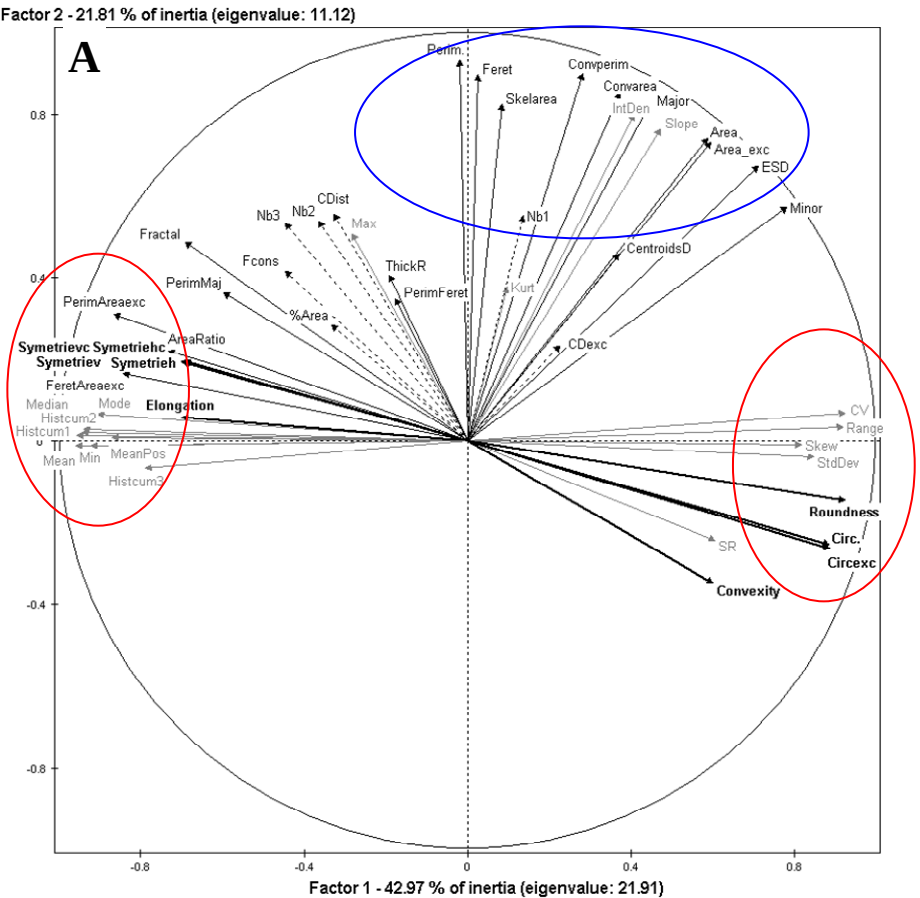
- With the ZooScan useful size spectra and biomass estimate of eggs may be rapidly obtained for ecological oriented studies.
- Patterns of distribution were similar for Plaice, Long rough dab, Cod and Whiting with microscope and ZooScan identifications
- A confusion exist between some species belonging to the same size range as for Dab and Flounder
- The presence of oil globules (Rocklings) does not help the ZooScan identifications
- ZooScan allows archiving of digital images of samples, to facilitate permanent records when conservation of the physical samples is not possible.

Perspectives to improve ZooScan identification performances

- **Improve our learning set by**
  - – Increasing number of items in classes that are under-represented (rocklings, cod and long rough dab)
  - – obtaining eggs directly from fish spawns under experimental and aquaculture conditions to avoid misidentification

Perspectives to improve ZooScan identification performances

➢**PCA analysis showed that among the 51 parameters some were highly correlated and redundant, resulting in an over fitted classifier, more sensitive to noise in the data**



Applying a method of selection of variables, on the learning set would probably increase the performance of the classifier and thus leading to a better recognition of fish egg species.


Thank you for your attention