UNION EUROPÉENNE
Région PAYS de la LOIRE
CE PROJET EST COFINANCÉ PAR LE FONDS EUROPÉEN DE DÉVELOPPEMENT RÉGIONAL

WEAMEC MARINE ENERGY

Ifremer

FIRST APPROACH TO UNSUPERVISED CLASSIFICATION OF BROADBAND FISHERIES ACOUSTIC DATA FOR ECOSYSTEM MONITORING

ALICIA MAURICE[1], LAURENT BERGER[2], MATHIEU DORAY[1]

[1]: IFREMER, UNITÉ ECOLOGIE ET MODÈLES POUR L'HALIEUTIQUE, NANTES

[2]: IFREMER, IMN/NSE/A, BREST

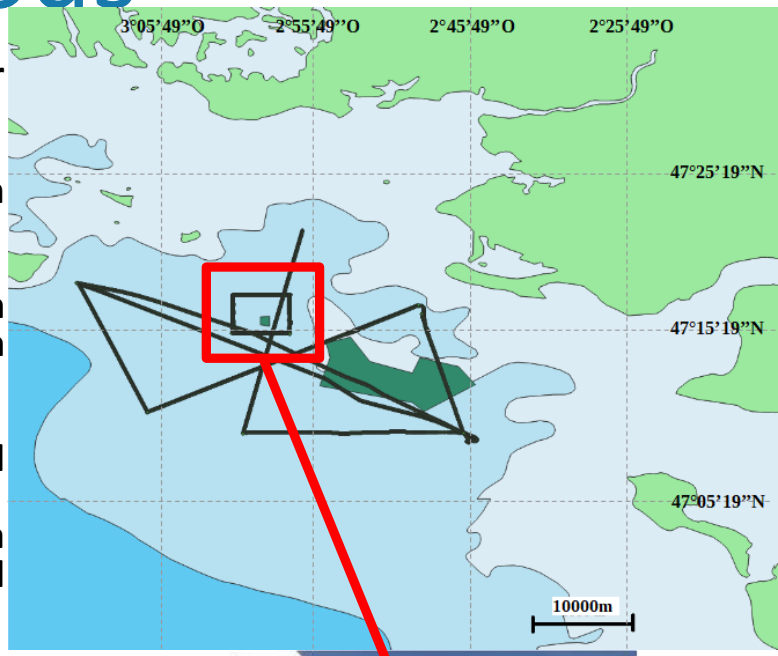
CE PROJET EST COFINANCÉ PAR LE FONDS EUROPÉEN DE DÉVELOPPEMENT RÉGIONAL

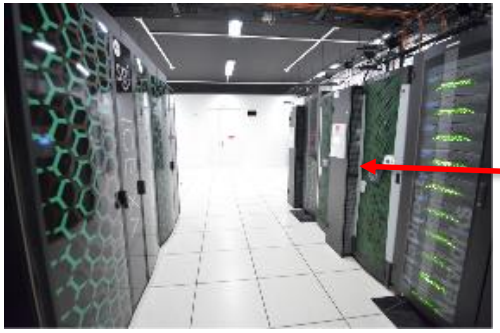
Introduction

- **Increase of broadband acoustic images (BAI) collection**
- **Detailed spectral information**
- **But curse of dimensionality, limited availability of labeled samples and taxa mixing**
- **Unsupervised classification approaches to BAI big data challenge**
- **New methodologies, (open) softwares, (super) computers and common reference datasets for benchmarking**
- **Hyperspectral image community legacy**
- **First results of unsupervised classifications of BAI near an offshore windmill in the Bay of Biscay (BoB, France, EchoSonde project)**

Material and methods

- Software / hardware
  - Python / py_movies3D package / datarmor HPC supercomputer
- Methods
  - Fine scale echo-integration (1m vert / 3m horiz) at low threshold (-90dB)
  - Spatial-spectral classification
  - Several clustering and dimension reduction techniques (sklearn library and Rasti et al[1]) on absolute MVBS
- Data
  - BoB « echoscapes » (validated typical frequency spectra)
  - PHOENIX2018 survey BAIs Day/Night data acquisition around offshore windmill installation
- Platform and echosounders
  - Thalassa vessel fitted with Simrad EK80 (6 main frequencies: 18kHz (CW), 38kHz, 70kHz, 120kHz, 200kHz and 333kHz (FM))
  - Pings emitted sequentially, 2 sec for a complete cycle





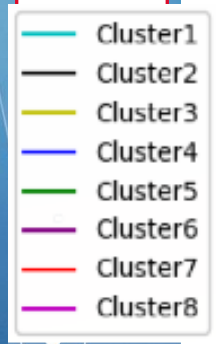
Datarmor HPC supercomputer

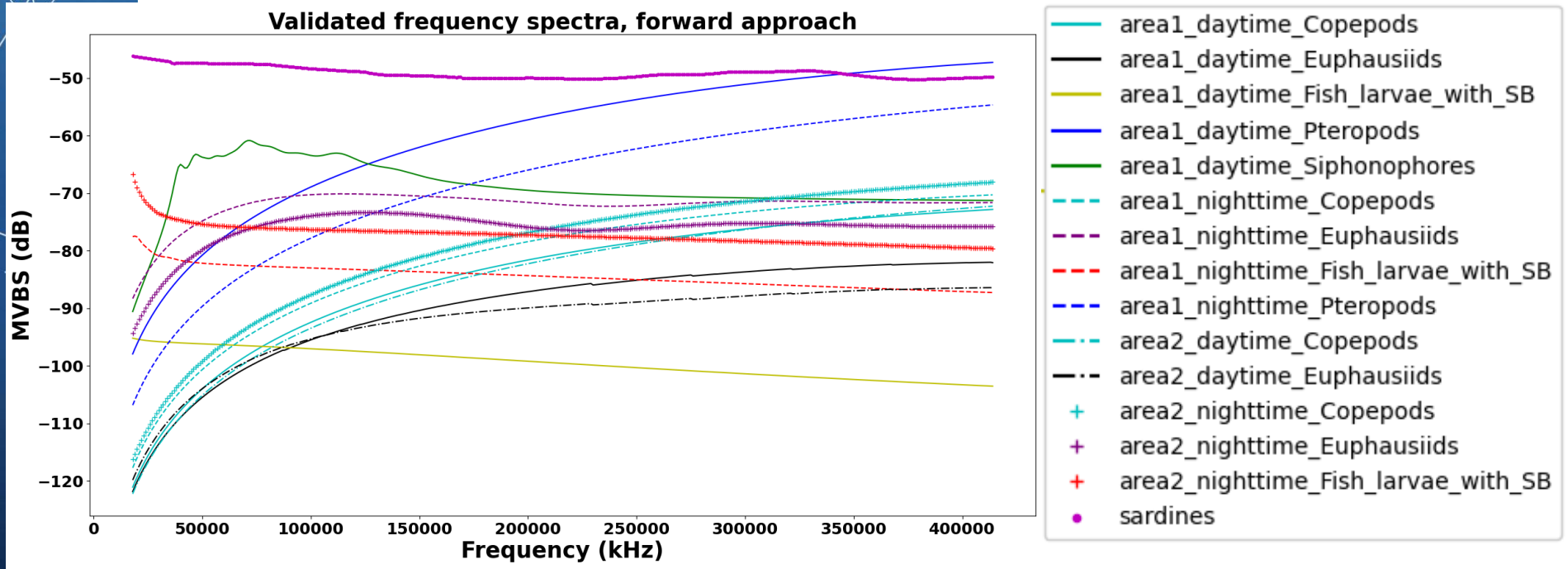
[1] Feature Extraction for Hyperspectral Imagery by Rasti et al, IEEE Geoscience and Remote Sensing Magazine, 2020

Results (1): « echoscape » clustering

Obj: separate BoB typical « echoscapes »

- **Few reference labeled broadband acoustic datasets available=>use of unsupervised learning methods**
- **K-Means method fast and widely used but major drawbacks=>look for other methods to cluster the data**
- **Firstly, test of the efficiency of those methods on a set of validated in situ data (modeled by forward approach)=>8 clusters are to be identified [1]**



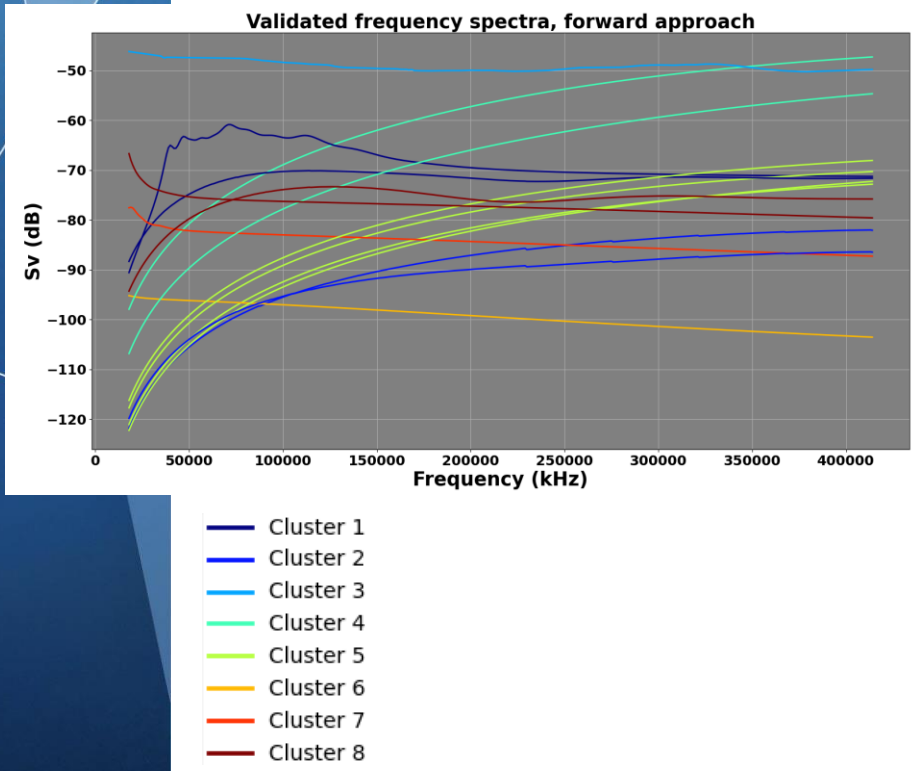


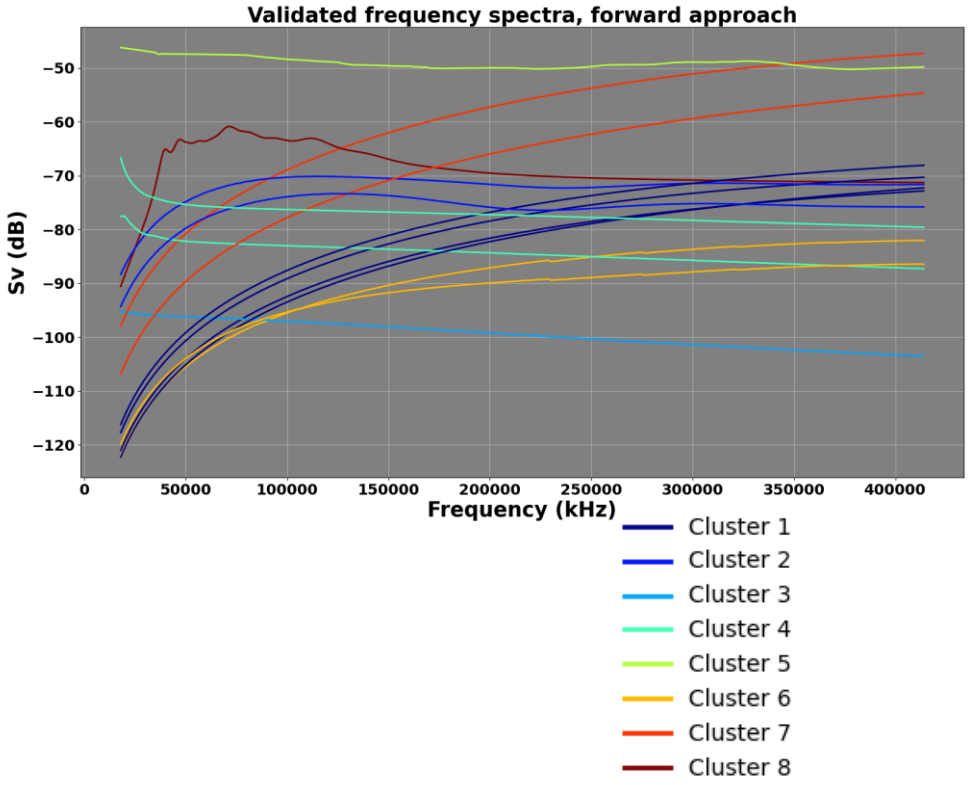
[1] Broadband acoustic evidences of the mesoscale distribution of gas-bearing siphonophores in the eutrophic Bay of Biscay, A. BLANLUET, IFREMER, 2019

Results (1): « echoscape » clustering

Obj: separate BoB typical « echoscapes »

- **Best compromise between quality of the results and computation time=>Modified Locally Linear Embedding (LLE) dimension reduction technique [1] + Kmeans clustering method**
- **All clusters are correctly identified with this method**



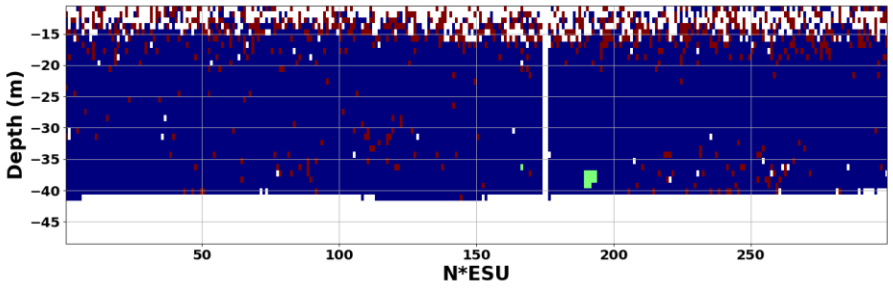


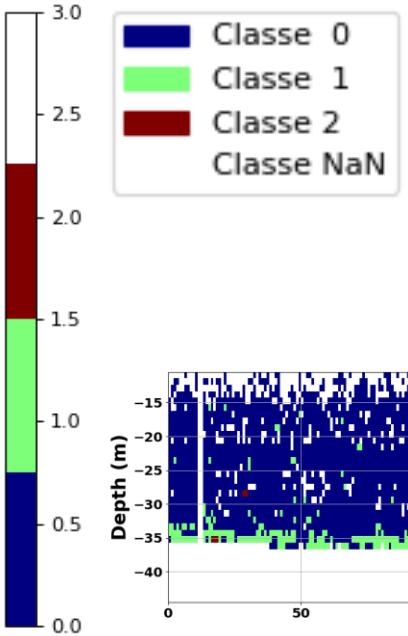
[1] LLE: Modified Locally Linear Embedding Using Multiple Weights, Zhang Z. & Wang L., MIT Press, 2007

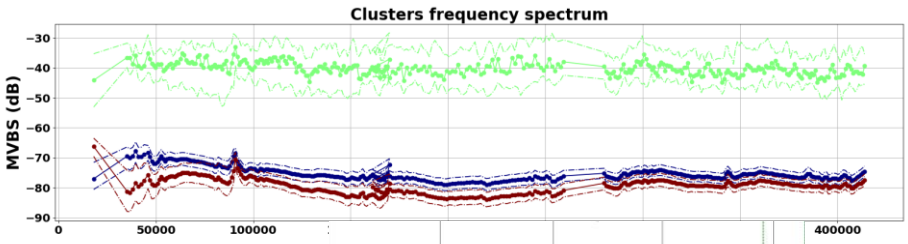
Results (2): BAI clustering

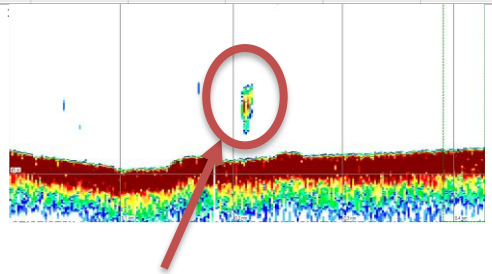
Obj: segment echogram to isolate backscatters

- **Small fish schools are completely isolated when running LLE+Kmeans**
- **A bias occurs for very small fish schools (because of two main factors: the position of the transducers on the vessel and the insonification sequence) => Confidence intervals are then essential**

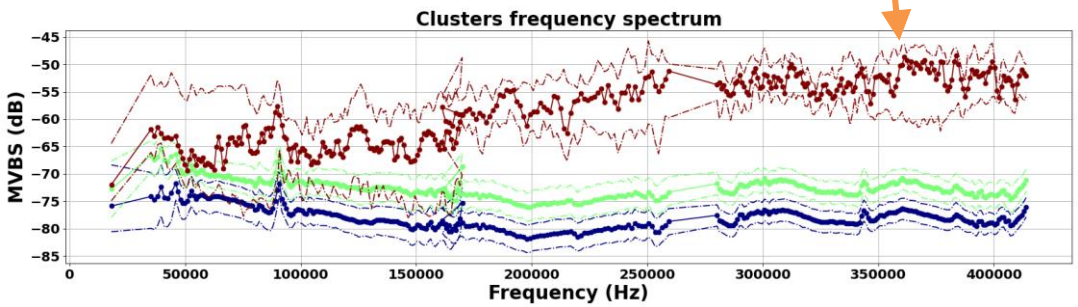








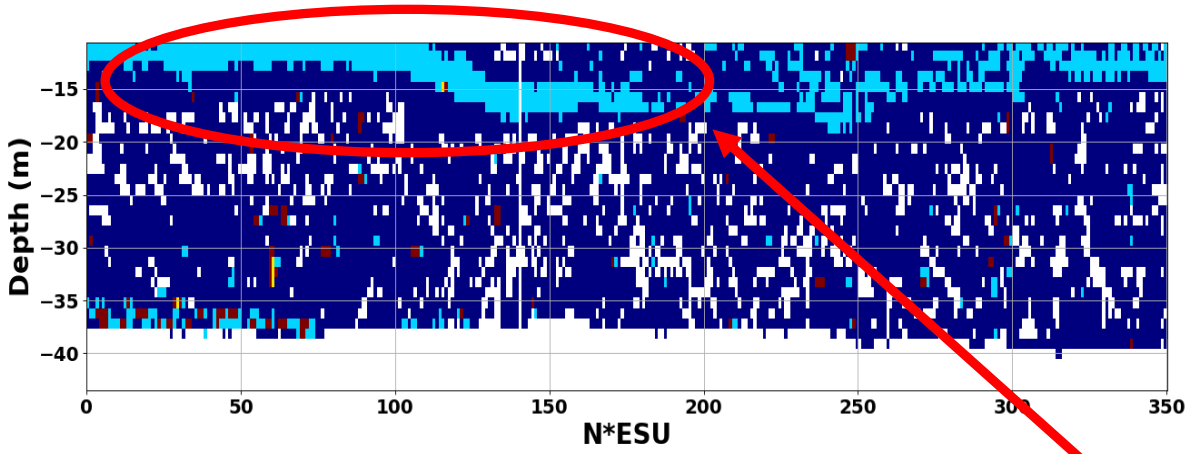
A 10 metre-long fish school

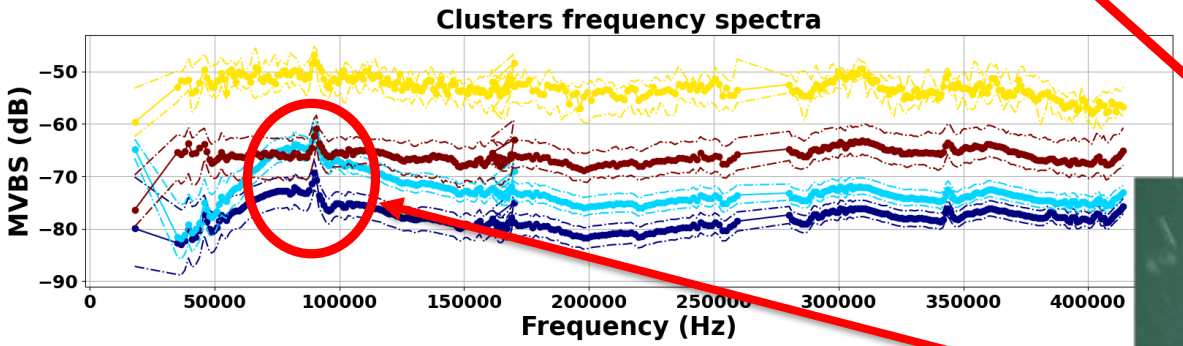


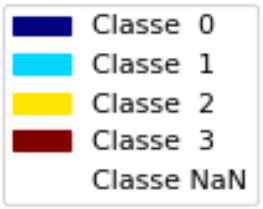
Results (3)

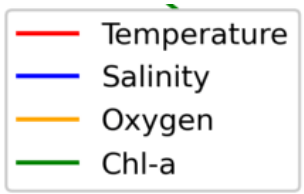
Obj: segment echogram to isolate backscatters

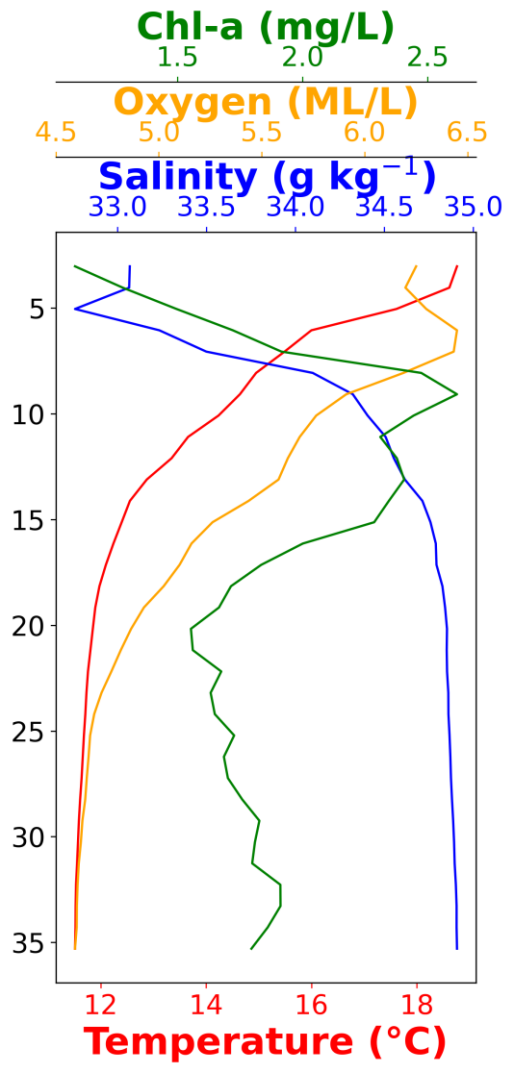
- **Identification of the thermocline when running LLE+Kmeans**

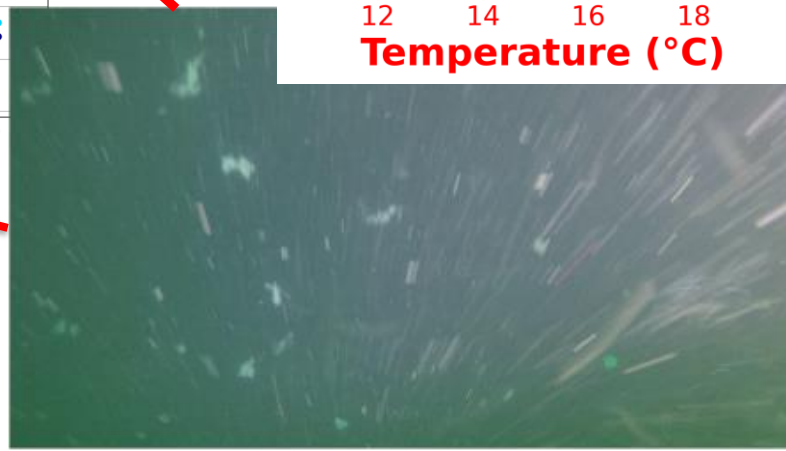












CE PROJET EST COFINANCÉ PAR LE FONDS EUROPÉEN DE DÉVELOPPEMENT RÉGIONAL

Discussion

- **Similarities with hyperspectral images processing method review / libraries=>need for shared reference labeled datasets and (open) frameworks to benchmark classification methods: « echoscapes » + BAIs**
- **Clustering on absolute MVBS as backscatter level discriminant**
- **Kmeans computationaly efficient, but equally-sized clusters**
- **LLE+Kmeans performed better than Kmeans:**
  - **spatial information accounted for in LLE: fish school separation, thermocline mapping ...**
  - **dimension reduction to alleviate the curse of dimensionality: improved clustering and computation time**

Conclusions

- **Python + MOVIES3D + Datarmor + spatial-spectral ordination/clustering to tackle BAI big data challenge**
- **EchoSonde dataset:**
  - **vessel-borne BAI to be combined with BAIs from acoustic-optic profiler (WBTTUBE) and physical, optical and biological data**
- **LLE+Kmeans method:**
  - **Test on larger dataset (CW dataset)**
- **Shallow approach promising, unsupervised deep learning will be tested for potential improvements**

This work was carried out within the framework of the WEAMEC (West Atlantic Marine Energy Community) EchoSonde project, with funding from the Pays de la Loire Region.