The use of biological sensors on observatories, mooring lines or vehicles for long time series records

The use of biological sensors on observatories, mooring lines or vehicles for long time series records

The use of biological sensors on observatories, mooring lines or vehicles for long time series records

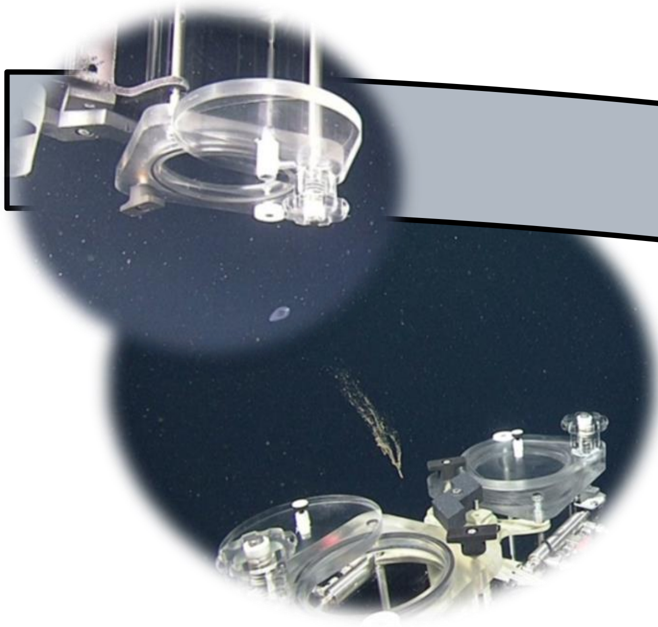
Can bioluminescence be a biological proxy *in situ* ?

The use of biological sensors on observatories, mooring lines or vehicles for long time series records

Can bioluminescence be a biological proxy *in situ* ?

- **Bioluminescence kinetics can discriminate groups of organisms**
  - Classification(K-means), pattern-recognition of bioluminescence kinetics (with Monique) from laboratory experiments with results applied to DORADO AUV dataset
- **Bioluminescence can be understood and predicted using models based on environmental variables**
  - Tree-based model (decision tree) on AUV DORADO dataset

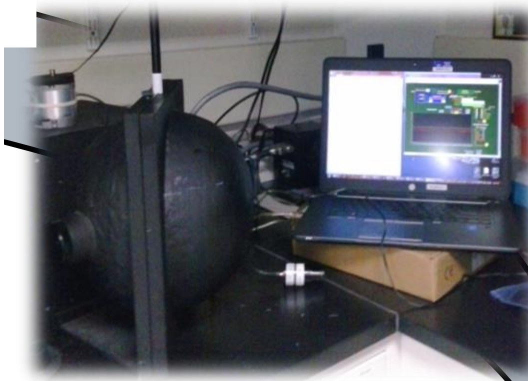
1- Bioluminescence to discriminate groups of organisms



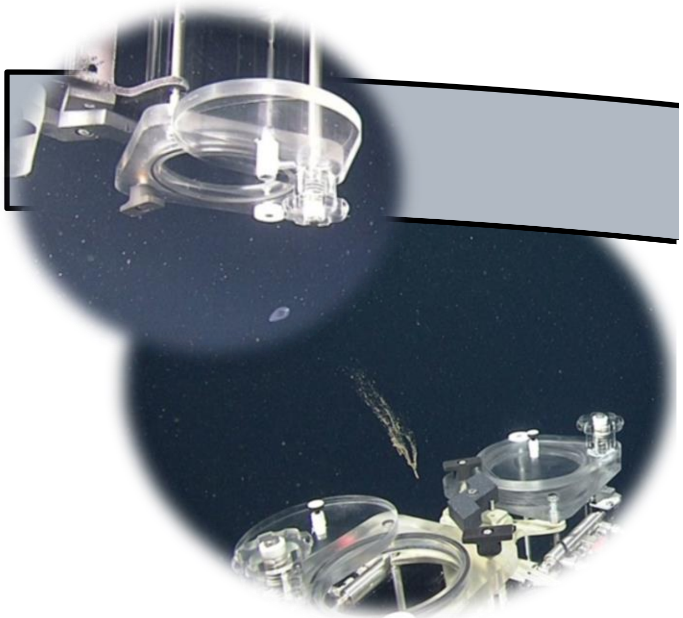
Bioluminescent organisms:

- Copepods (metridinidae)
- Beroe sp.
- Euphausiid
- Dinoflagellates species
- Larvaceans (+/- house)
- Marine snow
- Bacteria

Haddock lab



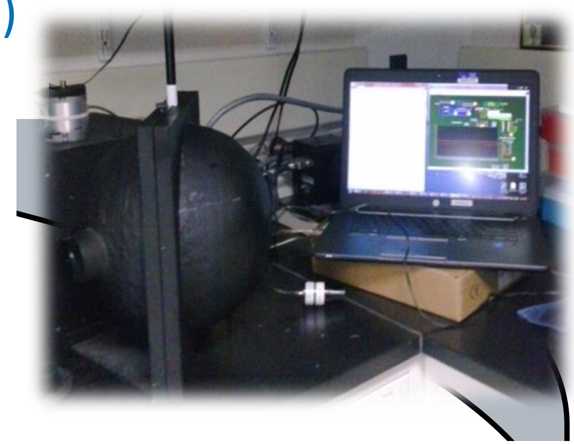
1- Bioluminescence to discriminate groups of organisms

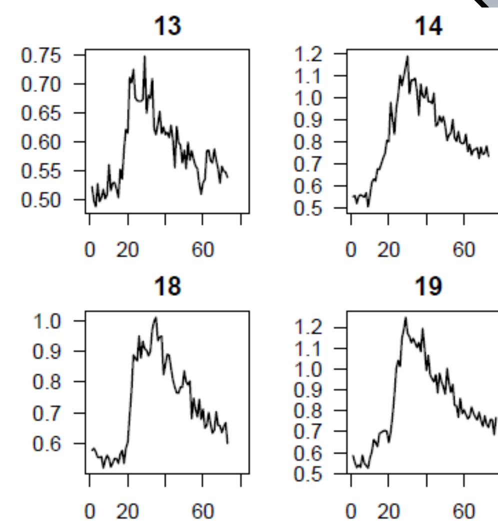


Bioluminescent organisms:

- Copepods (metridinidae)
- Beroe sp.
- Euphausiid
- Dinoflagellates species
- Larvaceans (+/- house)
- Marine snow
- bacteria

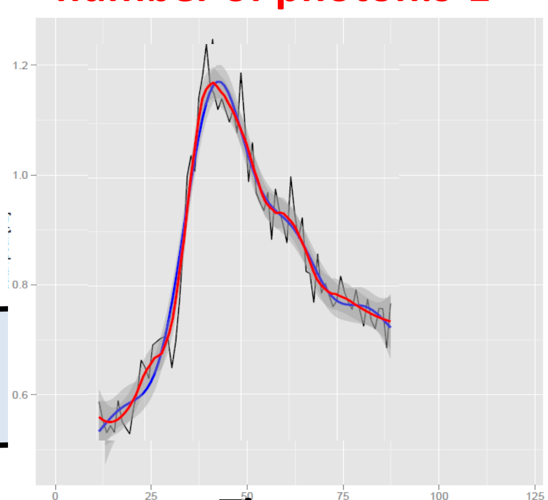
Haddock lab



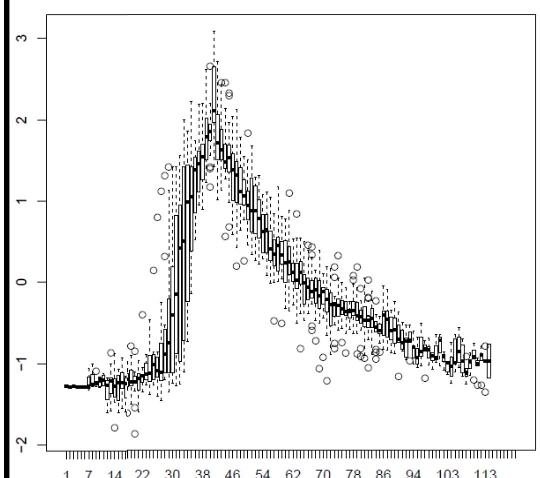


Several peaks for each individual

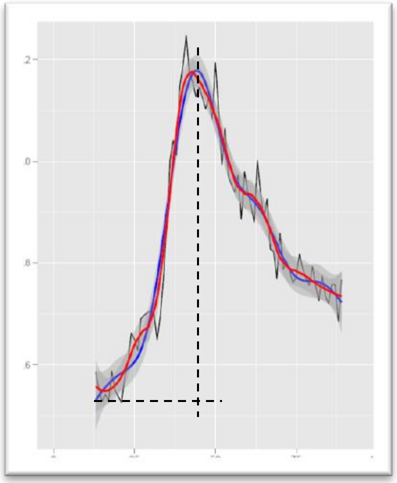
Bioluminescence = number of photon.s-1



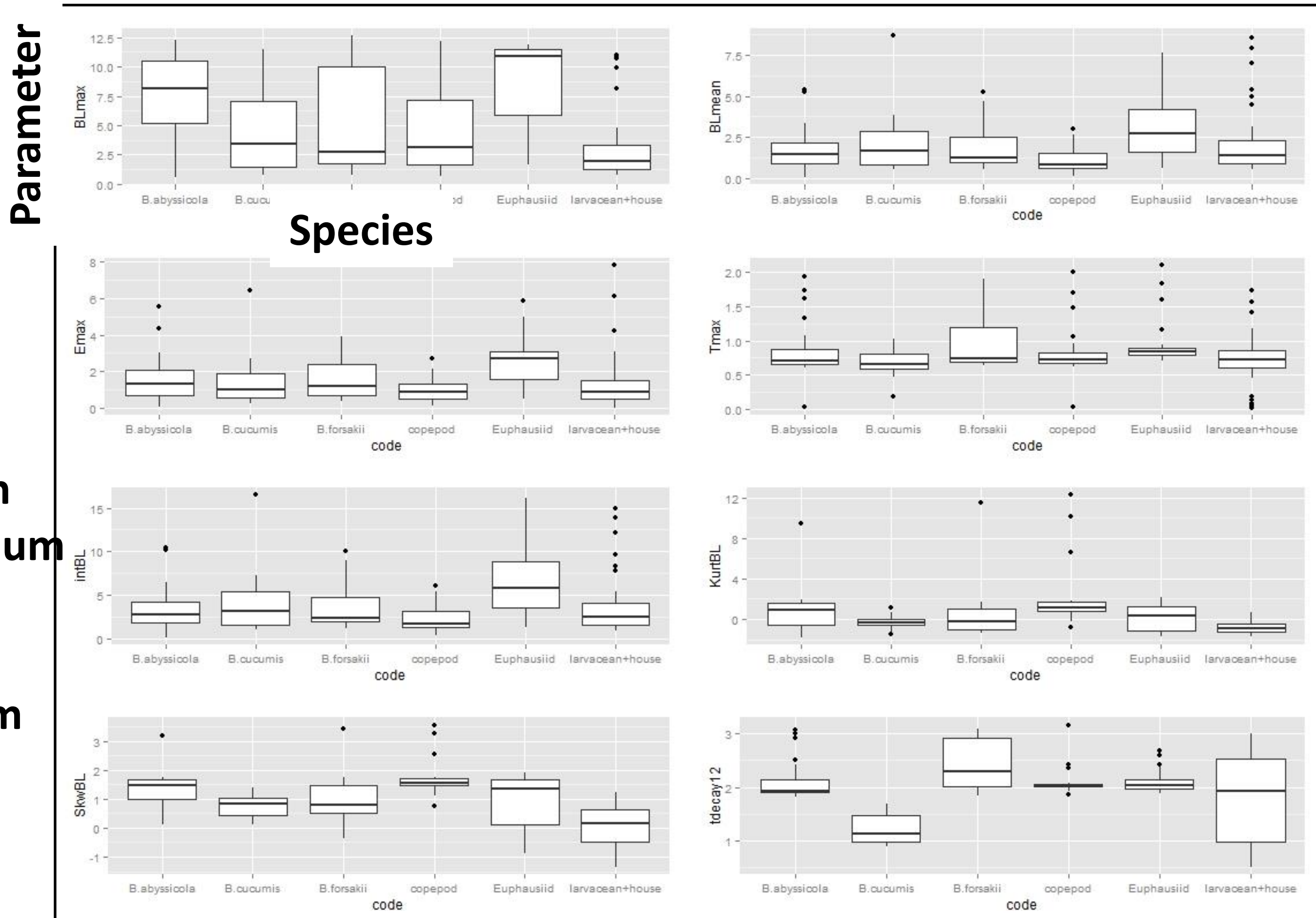
intra-species variability



1- Bioluminescence to discriminate groups of organisms



Kinetic parameters have been described for each individual and the **inter-species** variability is observed

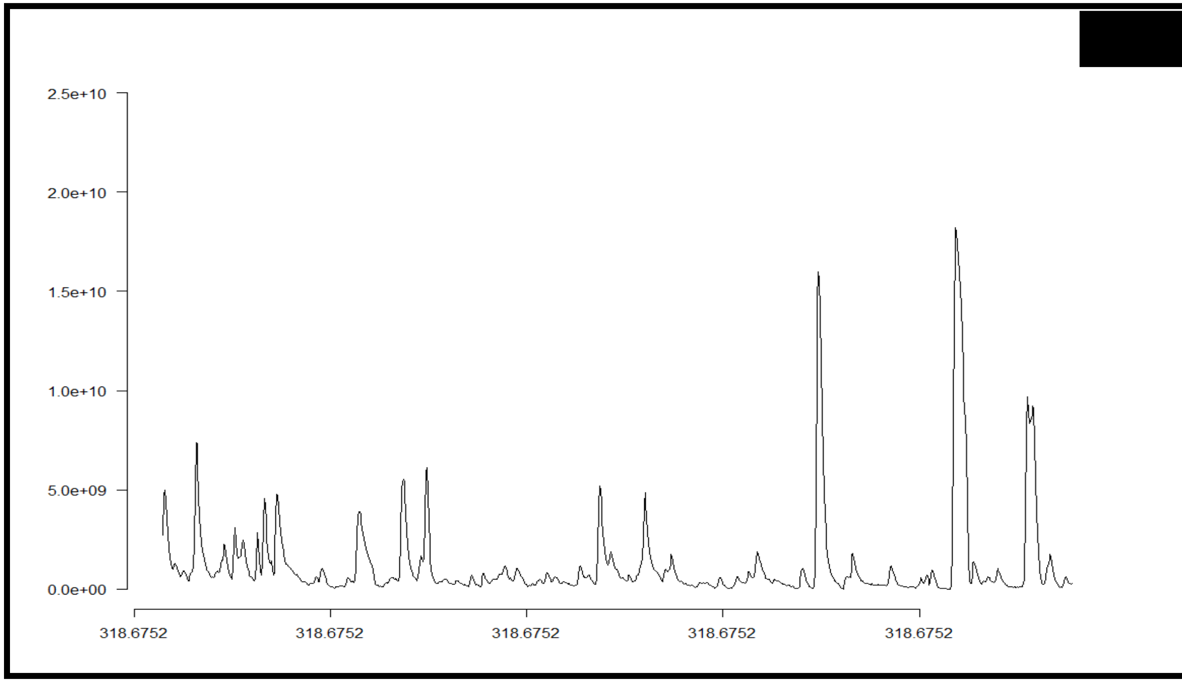


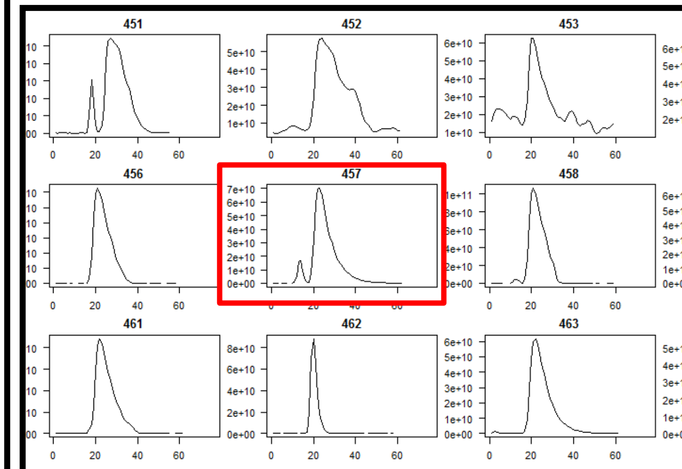
Parameters:

- Maximum intensity
- Integration of photon
- Integration to maximum
- Kurtosis
- Skewness
- Time to the maximum
- Decay time

What are the most relevant parameters to describe those species ?

2- Reverse application on DORADO-AUV dataset





KINETIC PARAMETERS

- **Bioluminescence record by bathyphotometer during the survey**
- **Peaks extraction**
- **Computation of all parameters**

2- Reverse application on DORADO-AUV dataset

- **Unsupervised classification**

- **K-means with K=3 clusters**
  based on ecological knowledges and assumptions looking at the peaks shapes

- **Minimize the variability W_{CK} within clusters C_K**

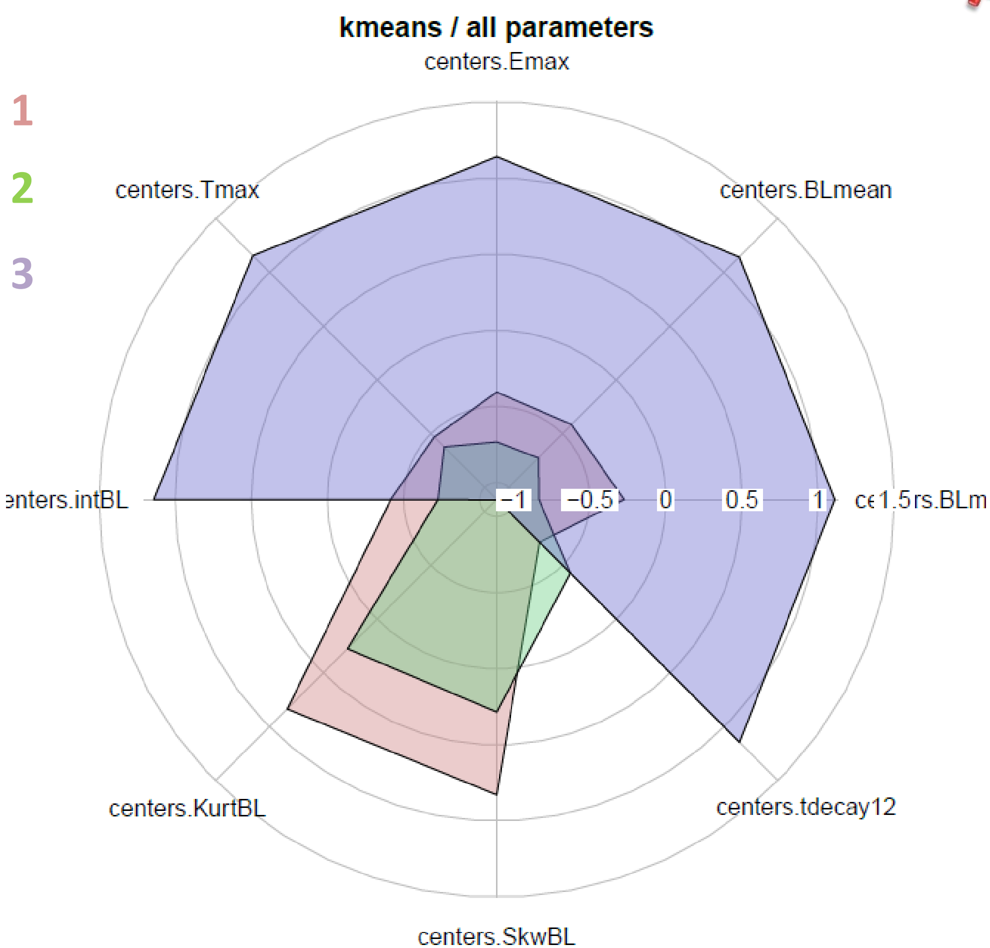
$$minimize_{C1,\dots,CK}\left\{\sum_{k=1}^{K} W(C_k)\right\}$$

2- Reverse application on DORADO-AUV dataset

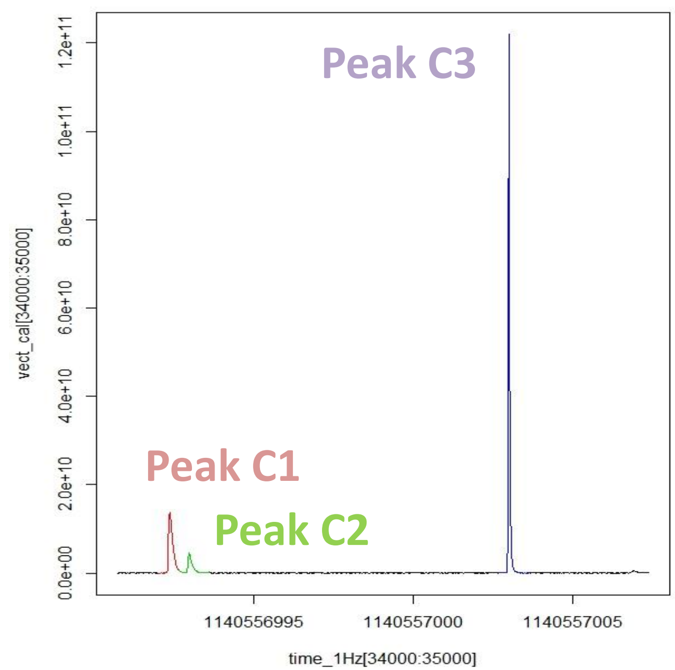
Cluster 1

Cluster 2

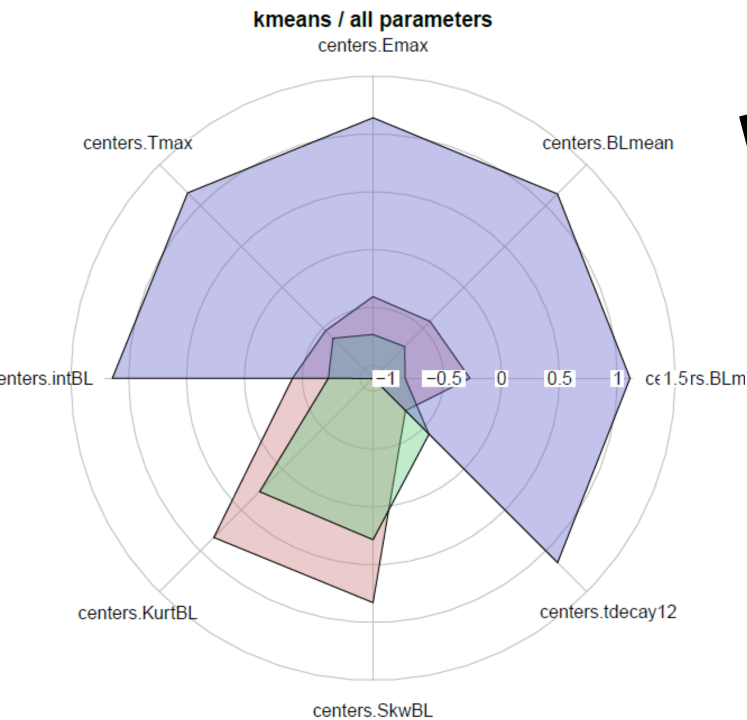
Cluster 3



PRELIMINARY RESULTS:
To be confirmed…

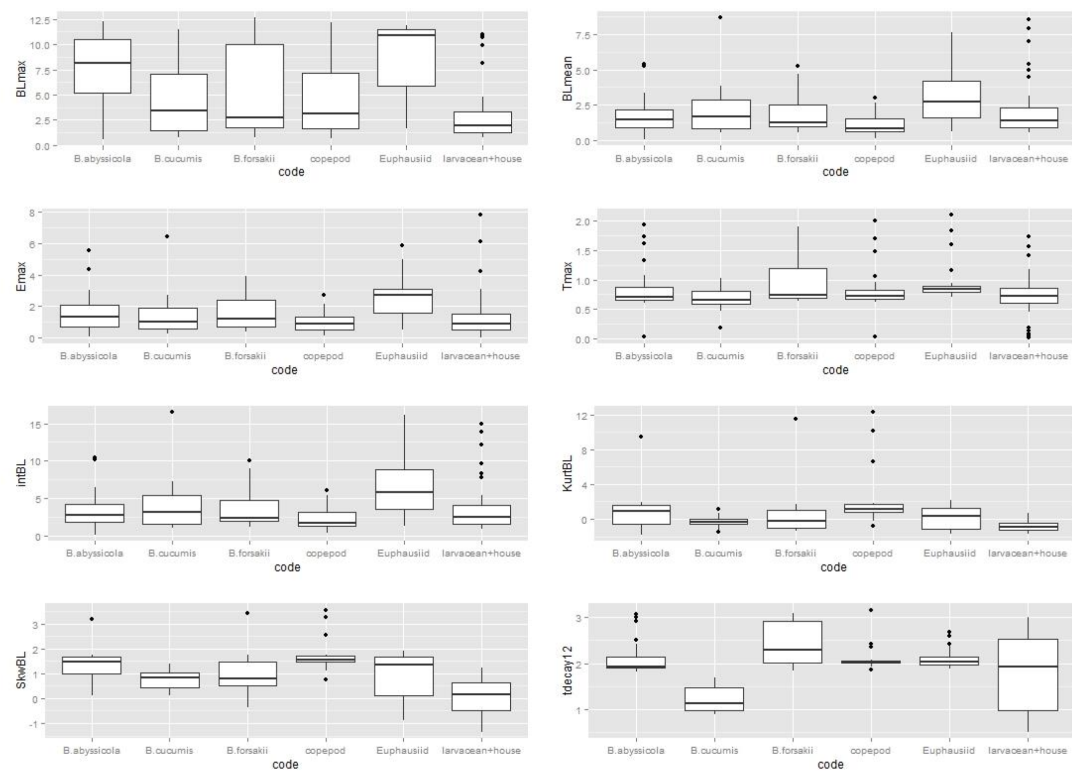


2- Reverse application on DORADO-AUV dataset



?

Which species belong to each cluster based on their kinetic parameters ?



3- Tree-based model to predict bioluminescence from environmental variables

Y(bioluminescence) = f(environmental variables)

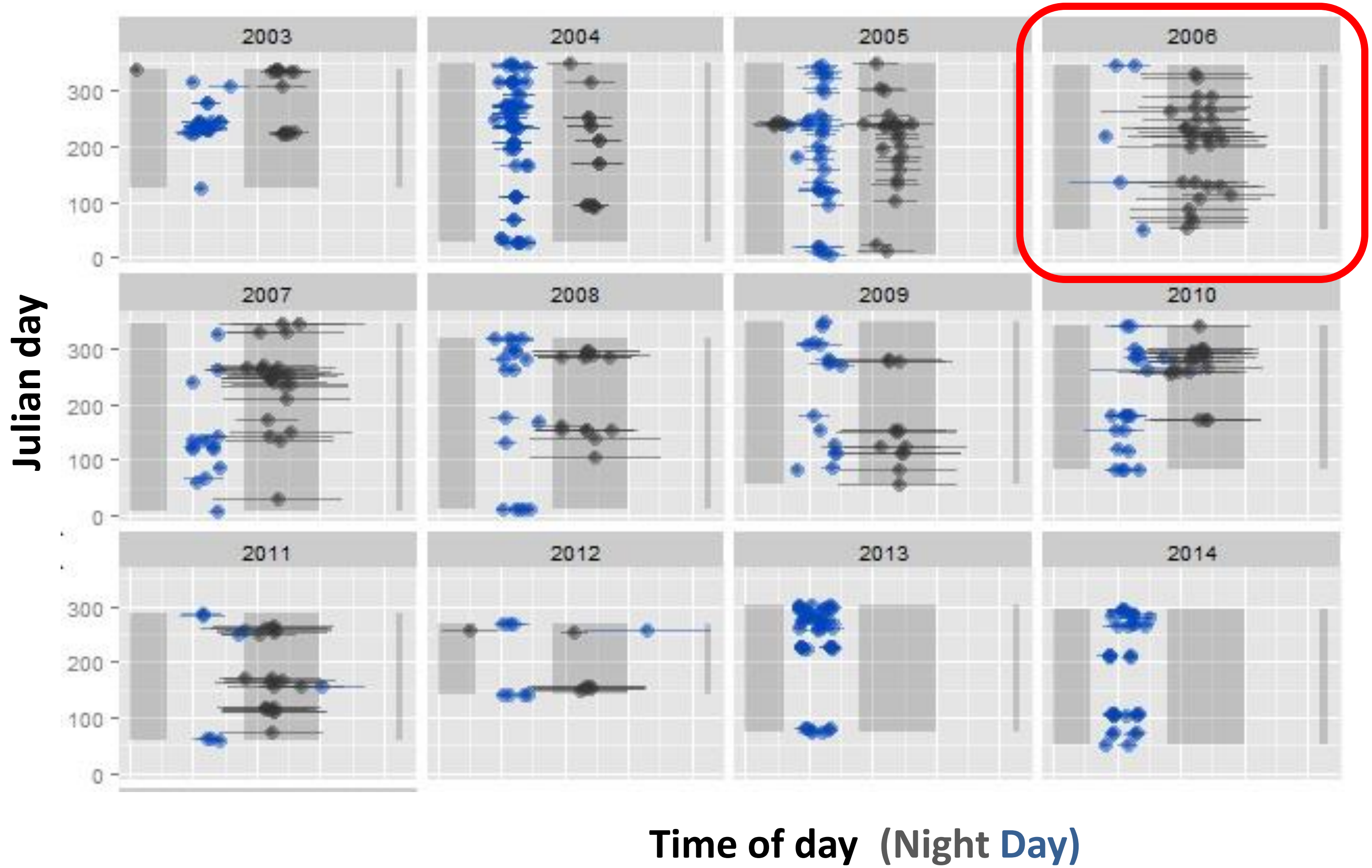
<u>Can AUV surveys be considered as time-series ?</u>

<u>For mathematical analyses:</u>

– **strong sampling variability in time, depth AND space**

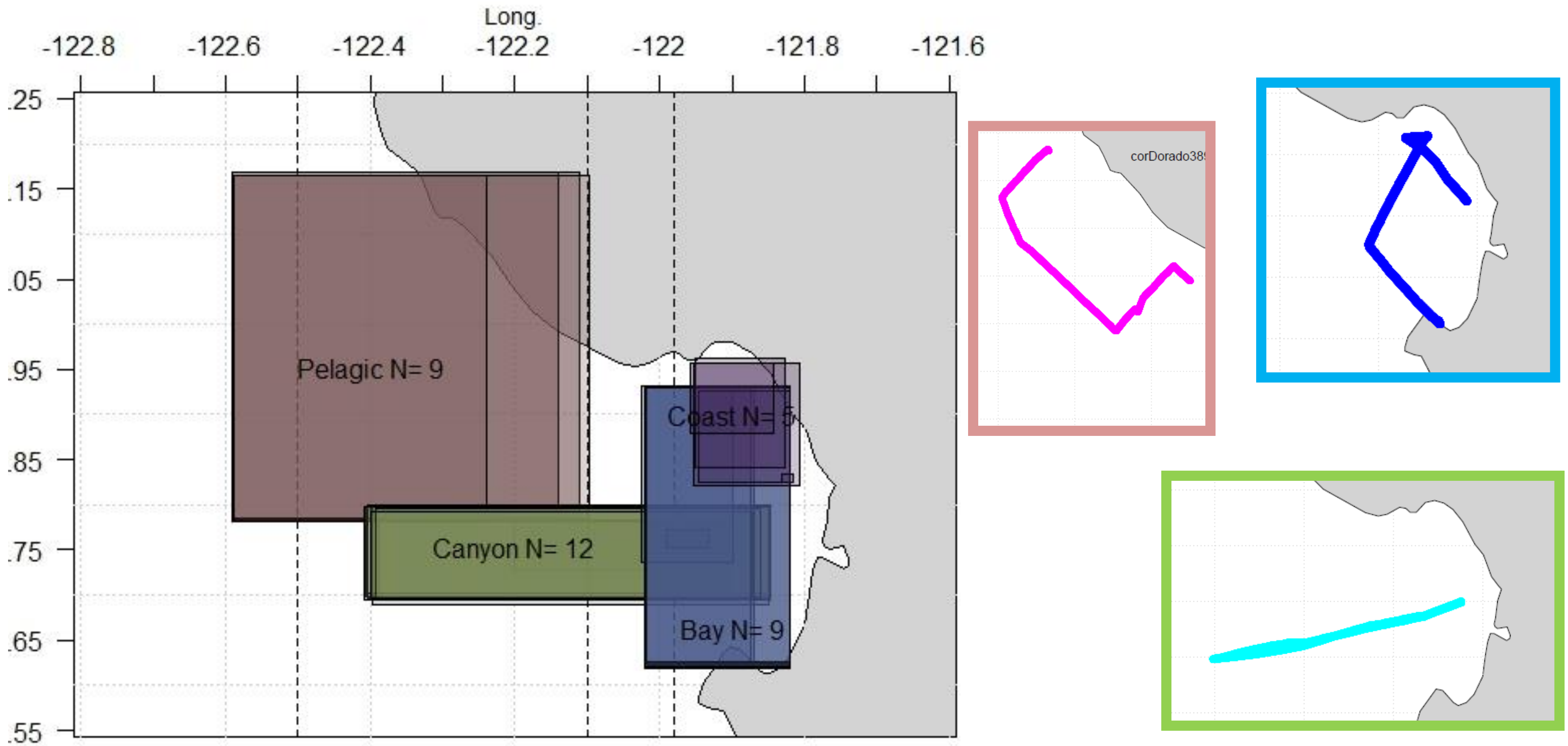
+ **strong sampling variability in time, depth AND space**

3- Tree-based model to predict bioluminescence from environmental variables



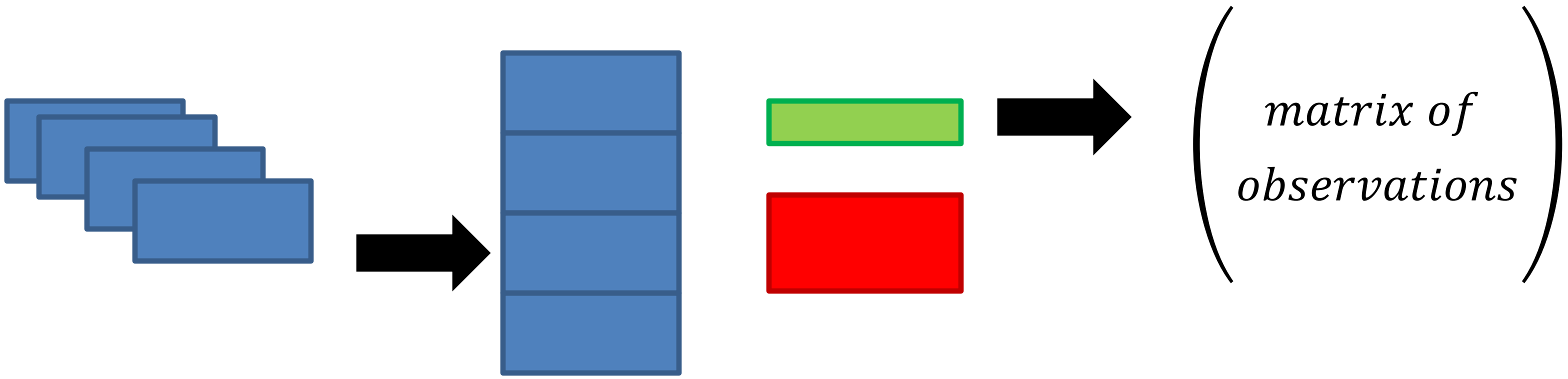
- **2006 "well" sampled in time**

3- Tree-based model to predict bioluminescence from environmental variables

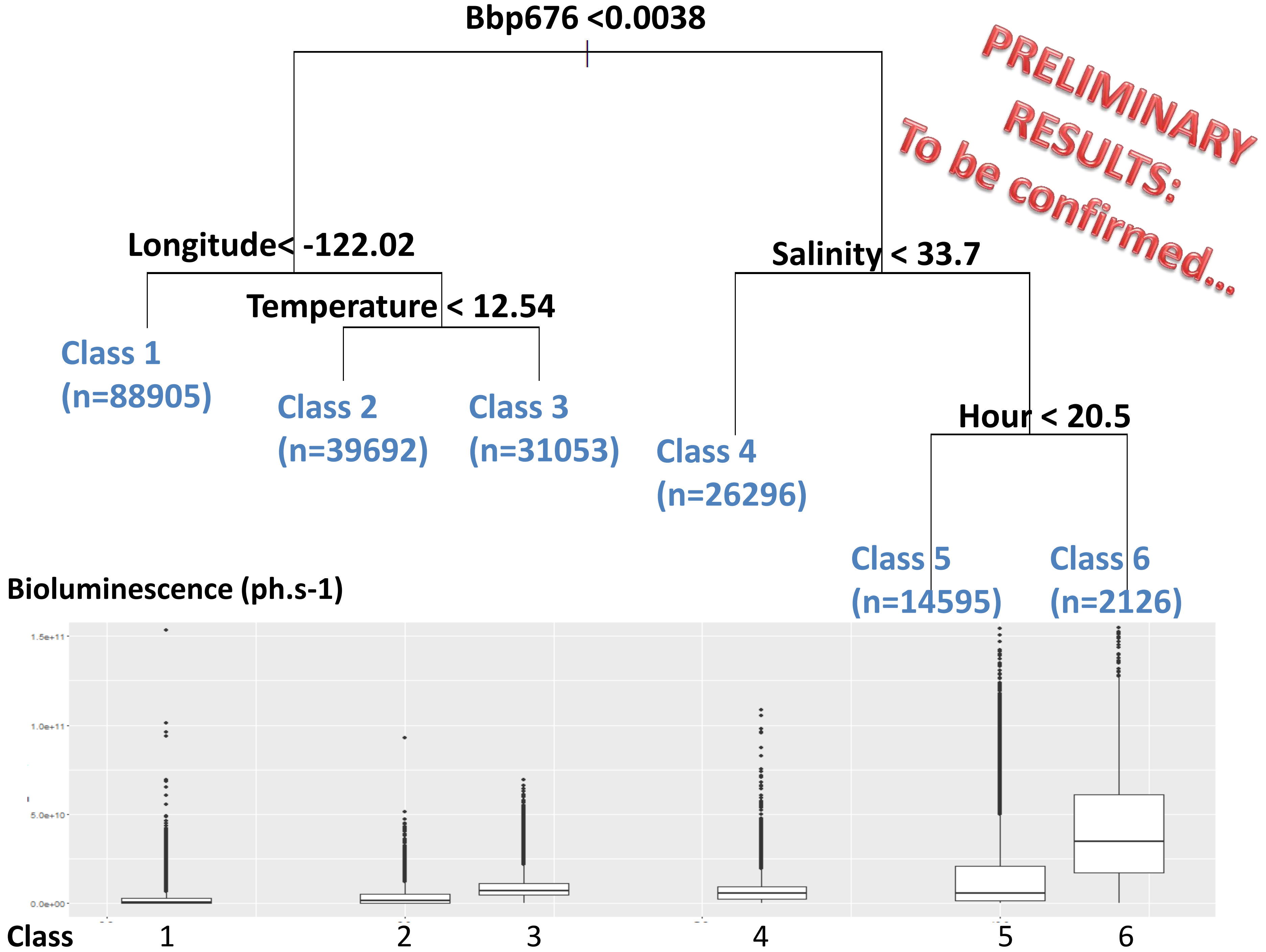


- **2006 “well” sampled in space**
- **Repetitive surveys**

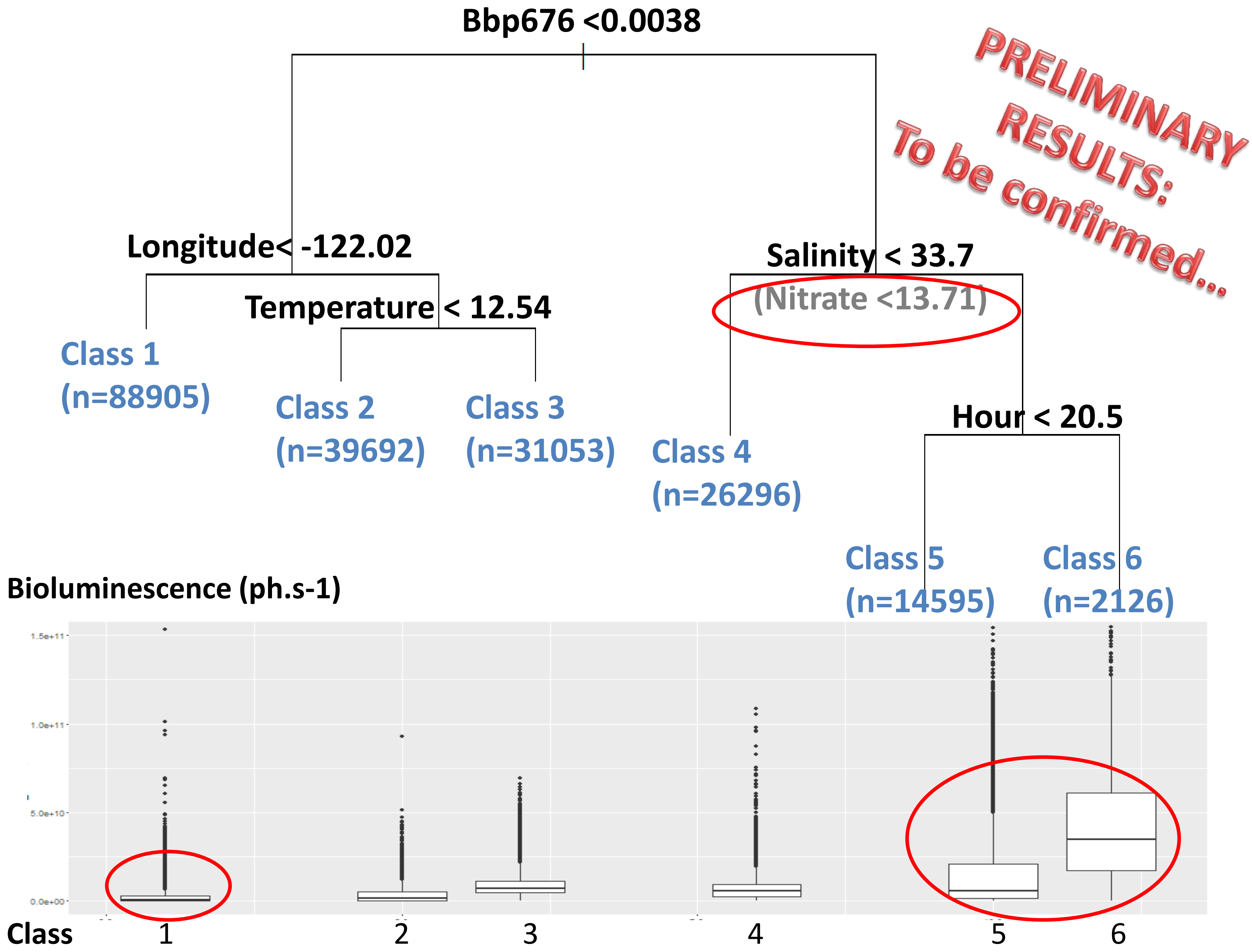
- **2006 surveys after correction of artefacts**
- **Y(bioluminescence) = f(Temperature + salinity + depth + oxygen + nitrate + bbp470 + bbp676 + latitude + longitude + hour of day)**
- **30% of data from each survey randomly = training dataset**
- **70% of data = efficiency of the model**



- **Binary recursive partitioning, maximize the variability within class, recursive process**
- **“Decision tree”**



Bbp676 <0.0038
PRELIMINARY RESULTS:
To be confirmed...
Longitude< -122.02
Temperature < 12.54
Salinity < 33.7
Hour < 20.5
Class 1
(n=88905)
Class 2
(n=39692)
Class 3
(n=31053)
Class 4
(n=26296)
Class 5
(n=14595)
Class 6
(n=2126)
Bioluminescence (ph.s-1)
1.5e+11
1.0e+11
5.0e+10
0.0e+00
Class
1
2
3
4
5
6


Bbp676 <0.0038
PRELIMINARY RESULTS:
To be confirmed...
Longitude< -122.02
Temperature < 12.54
Salinity < 33.7
(Nitrate <13.71)
Class 1
(n=88905)
Class 2
(n=39692)
Class 3
(n=31053)
Class 4
(n=26296)
Hour < 20.5
Class 5
(n=14595)
Class 6
(n=2126)
Bioluminescence (ph.s-1)
1.5e+11
1.0e+11
5.0e+10
0.0e+00
Class
1
2
3
4
5
6

Verifications :

- **Nodes are stable with recursive dataset (other 10% dataset within surveys)** **YES**

  avoid overfit as not reliable nodes

- **Percentage of good classification using the 70% of data for efficiency of the model** **~ 84%**

  classes have been defined by 1st to 3rd quartile using training distribution (boxplot without extreme values)

TO DO:

- **Improve the reliability of the model**
- **Try other mathematical methods (random forest)**
- **Ecological meaning of classes ? Relation with the knowledge of processes in the bay (front, thin layers …)**
- **Does the model work under highly variable environmental conditions ? Are we able to predict bioluminescence for different environmental conditions ? Upwelling strong/weak ?**
- **Application to detect/sample higher bioluminescence using AUV (cf thin layers)**