

# Visualiser et analyser les données du drone marin PAMELi

*Utiliser R/Python pour l'analyse de données*



**TP Coursus Master Ingénieur – Mars 2021**

Christine Plumejeaud,  
Clémence Chupin,  
U.M.R. 7266 LIENSS

# *PARTIE I – Introduction*

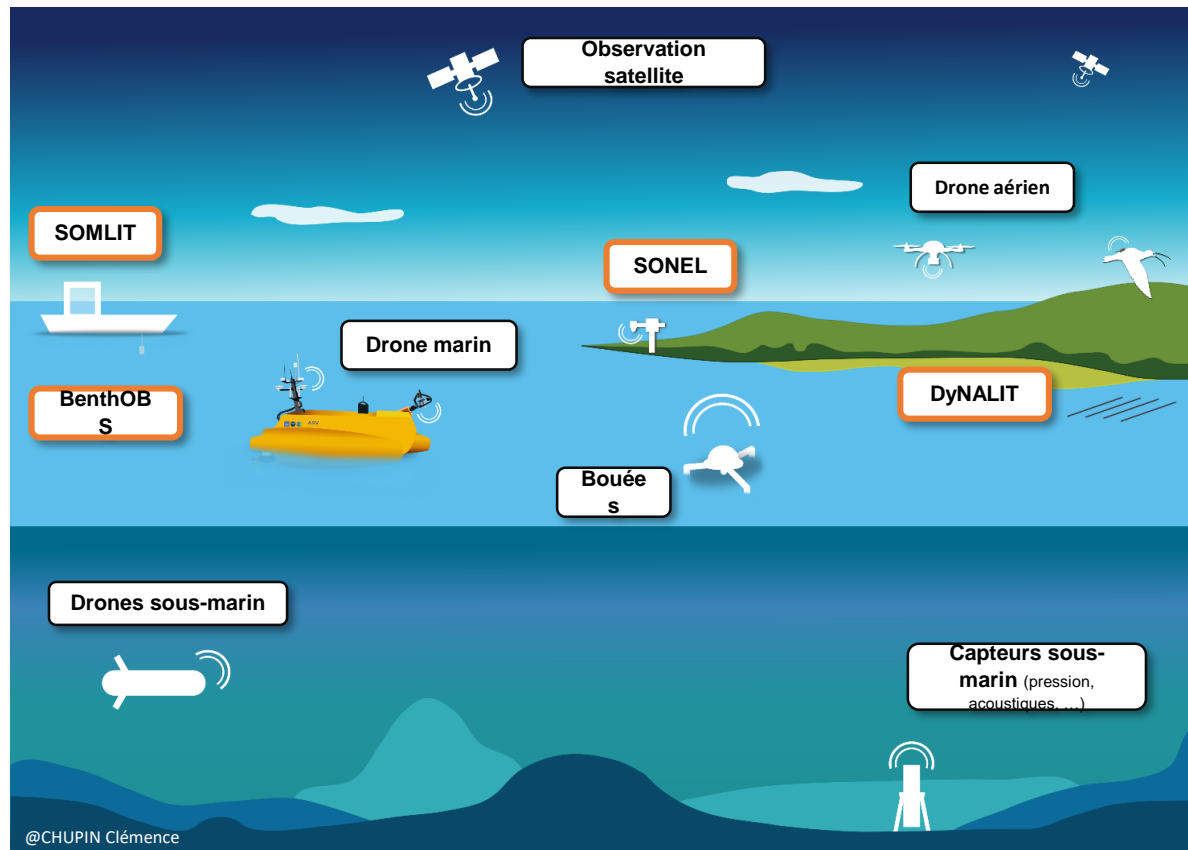


**TP Coursus Master Ingénieur – Mars 2021**

Christine Plumejeaud,  
Clémence Chupin,  
U.M.R. 7266 LIENSS

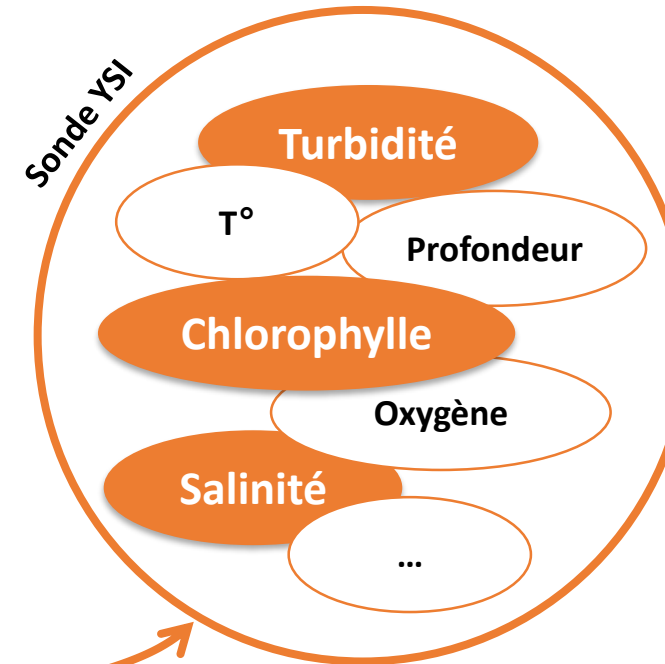
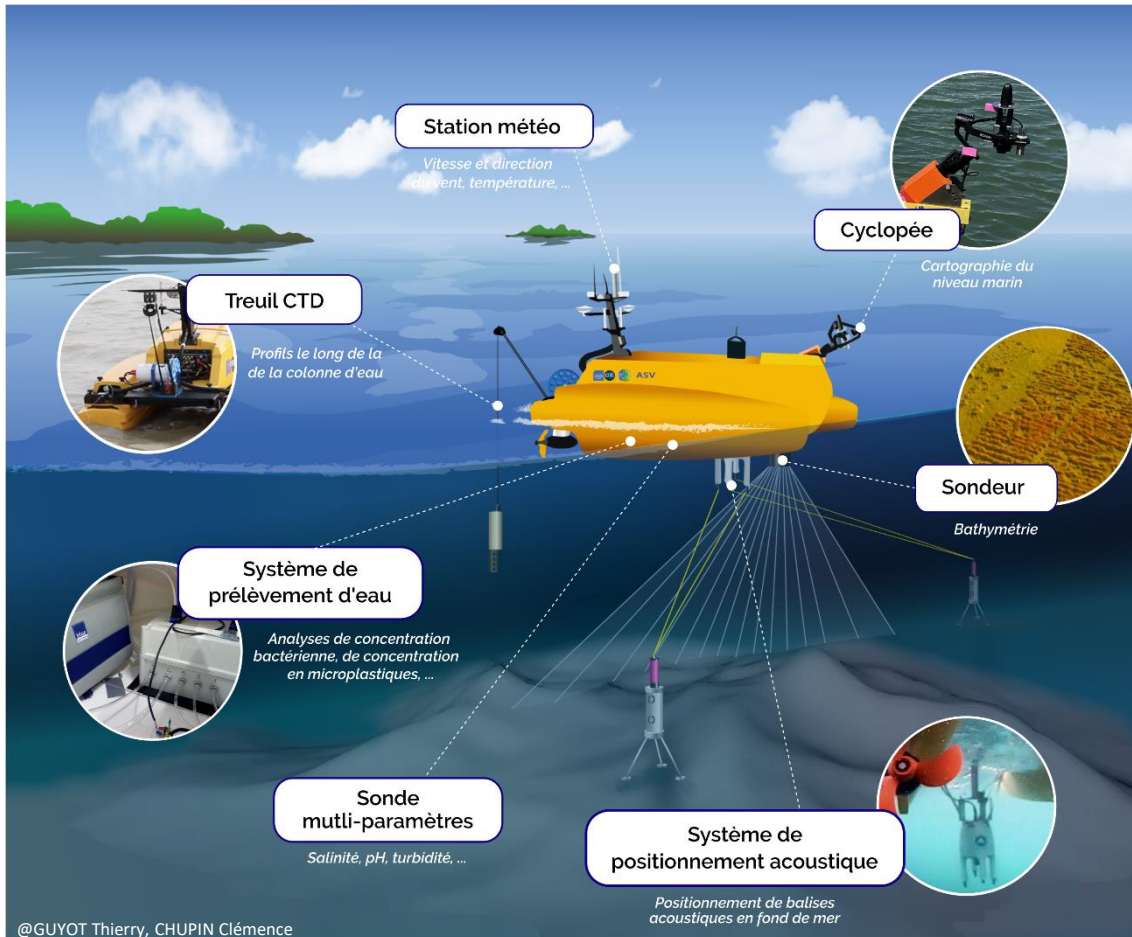
# PAMELi – Une plateforme autonome pour l'exploration du littoral

Le projet **PAMELi** (Plateforme Autonome Multi-capteur pour l'Exploration du Littoral) est né du besoin d'acquisition simultanée de nombreux paramètres afin de **rendre compte de la dynamique complexe du littoral**. Le projet s'appuie sur des équipes et des observatoires déjà en place (SONEL, SOMLIT,...) et complète le panel des techniques d'observation du littoral en place ou à venir.

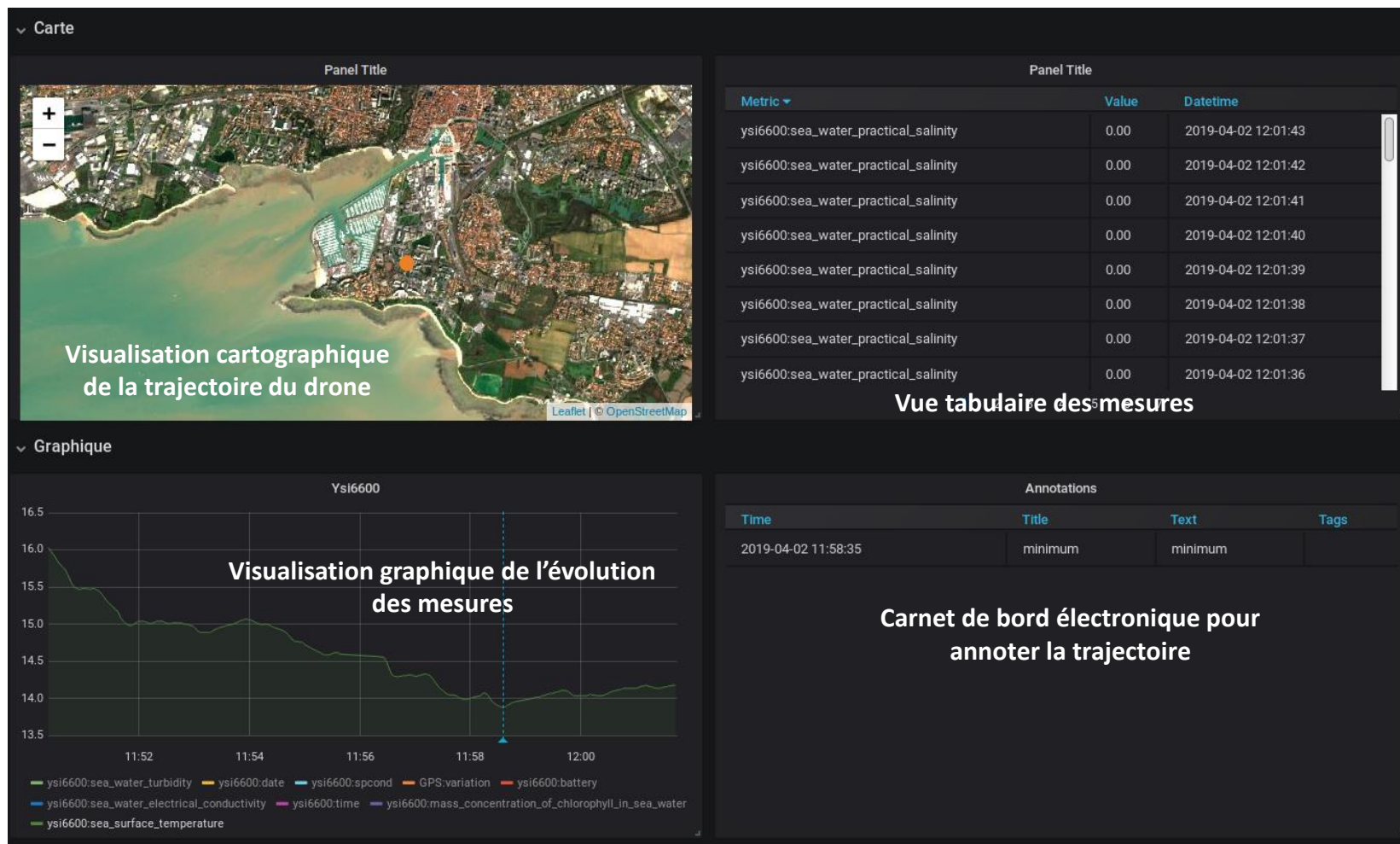


*Pour comprendre le littoral, de nombreux systèmes de mesure sont déployés.*

# PAMELi – Un drone marin multi-capteurs



# PAMELi – Visualiser les données en temps réel



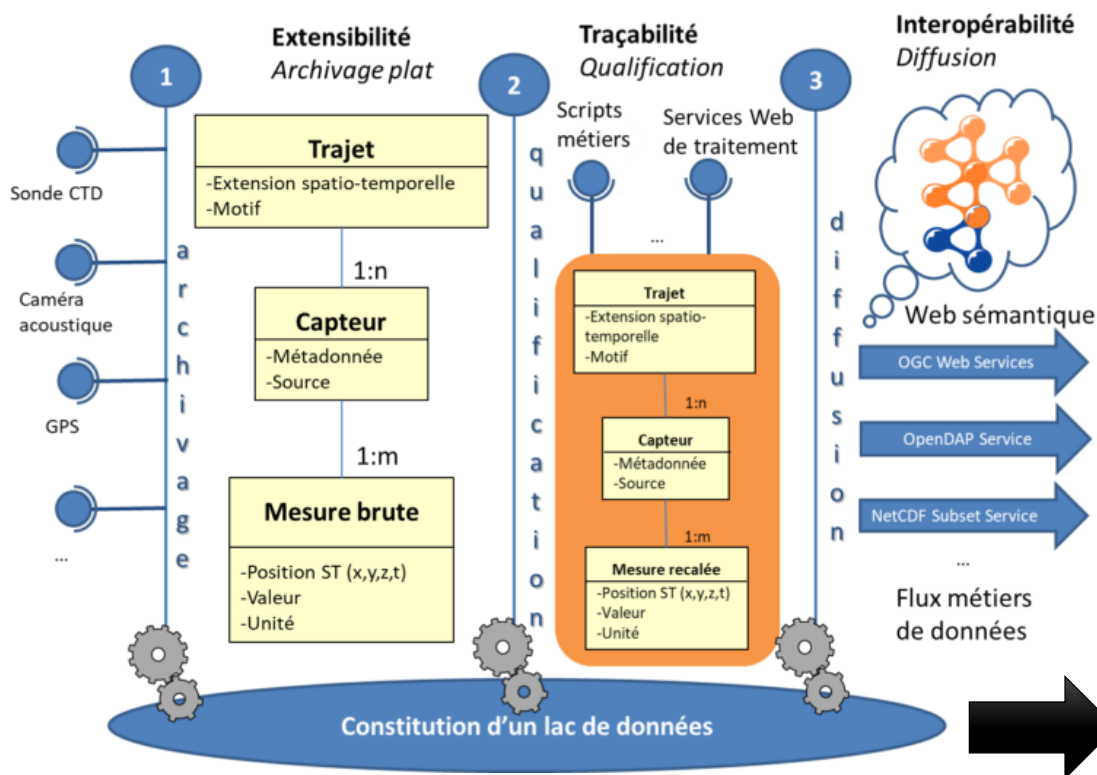
Une interface Graphana permet de visualiser en temps réel les données scientifiques acquises lors d'une sortie terrain.



# PAMELi – Une base de données multidisciplinaire

Les données acquises sont **enregistrées dans une base de données temporaire à bord du drone** avant **validation, transfert et archivage** dans une **base de données multi-disciplinaire à terre**. L'objectif est de construire une base de données spatio-temporelle interopérable avec d'autres systèmes d'observation, et de la **diffuser via des services Web**, accessibles aussi bien par les scientifiques que par le grand public.

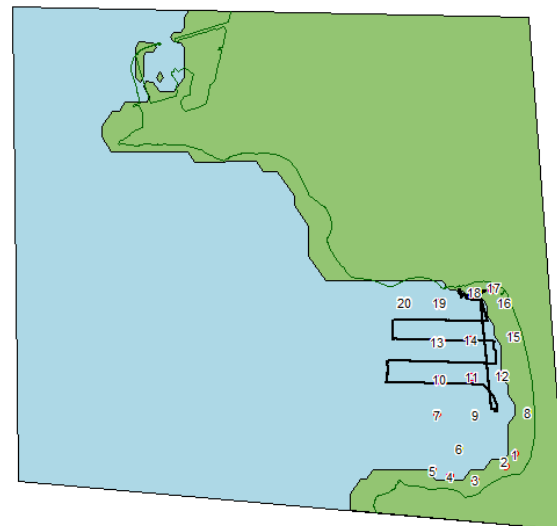
Le projet PAMELi



# Mesures terrain en baie d'Aytré

- **Contexte** : La baie d'Aytré connaît une pollution chronique, avec un taux de bactéries *Escherichia coli* important.
- Lors de la sortie terrain, le drone a **mesuré différents paramètres grâce à la sonde multi-paramètres YSI** et notamment :
  - La turbidité
  - La température
  - La salinité
  - La chlorophylle
- Des **prélèvements d'eau** ont également été réalisés pour des analyses microbiologiques et de micro-plastiques.

Espace d'étude sur la baie d'Aytré, 25-02-2021



Source : UMR 7266 LIENS

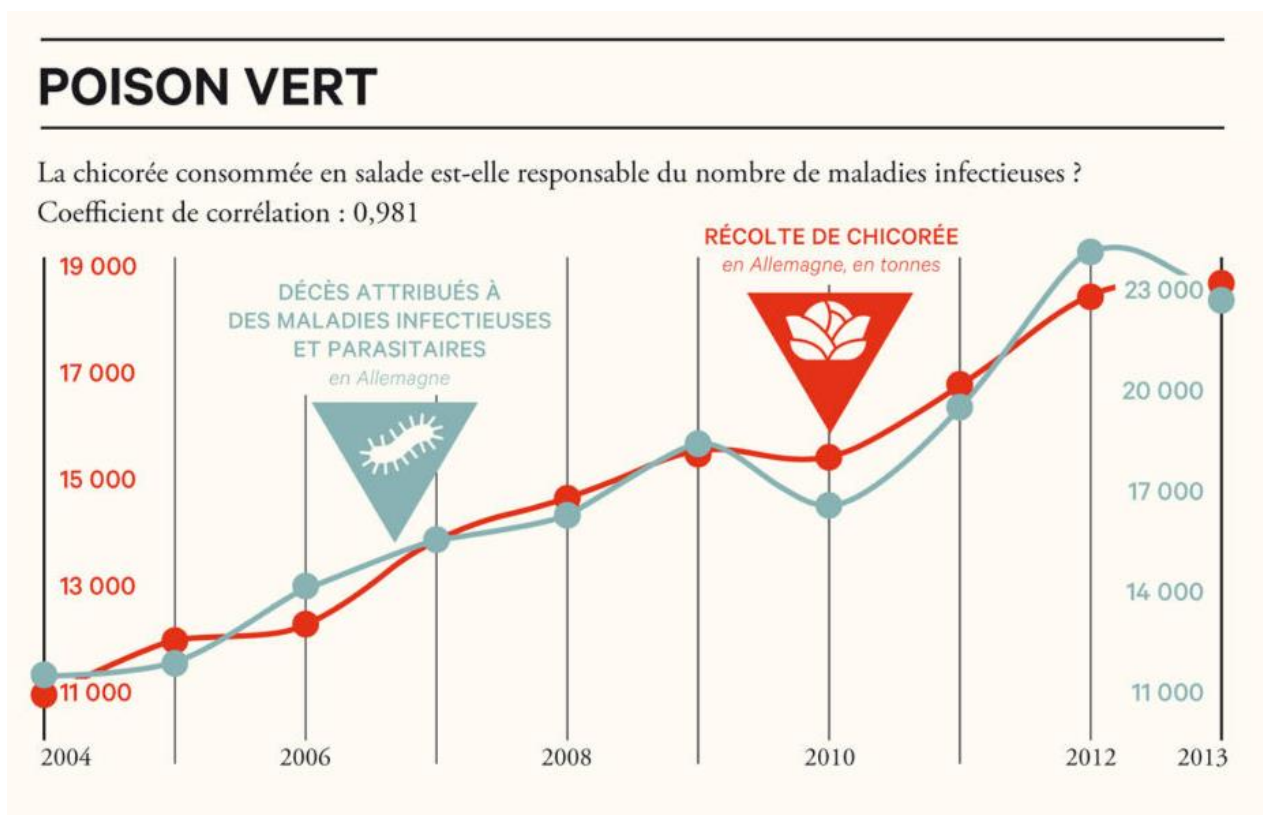


# Analyse de données

**Attention lors de la comparaison, l'analyse de paramètres : corrélation n'est pas causalité !**

<https://www.courrierinternational.com/grand-format/statistiques-les-correlations-de-labsurde>

Analyse de données



« Ce n'est pas parce que deux courbes se ressemblent qu'il y a un lien entre elles ! »



# Analyse de données

## R

- Utilisé par les chercheurs et enseignants **pour l'analyse de donnée**, mais est de plus en plus utilisé par les entreprises.
- **Accessible** même aux personnes n'ayant jamais codé.
- De **nombreuses librairie** sont disponibles, avec une documentation détaillée, mais il faut parfois du temps pour trouver la bonne librairie.
- Permet de **représenter**, analyser facilement et rapidement un jeu de données.
- **L'environnement de codage intégré** – R Studio – permet de prendre en main le code facilement.

## Python

- Utilisé par les **programmeurs** pour de l'analyse de données ou l'application de techniques statistiques.
- Accessible pour des personnes ayant des petites notions de programmation.
- **Syntaxe simple** et système d'indentation qui rendent le code clair et lisible.
- **Beaucoup de librairies** sont disponibles mais il y a parfois des questions de dépendances, et elles ne sont pas aussi développées que R pour l'analyse de données.
- Peut être **utilisé pour d'autres applications** que l'analyse de données (développement WEB, ...).

Source : <https://www.datacamp.com/> "Choosing Python or R for Data Analysis? An Infographic" (09/03/2021)

# Format des données de la BDD

date	sensor_name	property_name	property_value	unit_code	unit_name	point_id	pt_longitude	pt_latitude	geom_point	local_time	is_reference
2021-02-25	ysiexo2	ODOSat_%sat	93.750	%Sat	percent_saturation	36256	-1.132644	46.12575	c(-1.13264430888341, 46.1257502617766)	2021-02-25 13:58:01	FALSE
2021-02-25	ysiexo2	fDOM_rfu	5.890	RFU	RelativeFluorescenceUnit	36256	-1.132644	46.12575	c(-1.13264430888341, 46.1257502617766)	2021-02-25 13:58:01	FALSE
2021-02-25	ysiexo2	pH	8.480	pH	pH_units	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	salinity_ppt	22.680	ppt	parts_per_thousand	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	fDOM_rfu	5.890	RFU	RelativeFluorescenceUnit	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	temp_C	12.747	C	Celsius_degrees	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	conductivity_mS/cm	27.525	mS/cm	milliSiemens/centimeter	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	turbidity_NTU	31.530	NTU	NephelometricTurbidityUnit	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	ODOSat_%sat	93.780	%Sat	percent_saturation	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	depth_m	0.000	m	meters	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	chlorophyll_mg/L	6.170	ug/L	microgramms/liter	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	ysiexo2	turbidity_NTU	31.720	NTU	NephelometricTurbidityUnit	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE
2021-02-25	ysiexo2	conductivity_mS/cm	27.526	mS/cm	milliSiemens/centimeter	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE
2021-02-25	ysiexo2	depth_m	0.000	m	meters	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE
2021-02-25	ysiexo2	pH	8.480	pH	pH_units	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE

- Les données de la sortie PAMELi ont été **extraites sous forme de table (fichier .CSV)** avec en en-tête : la date, le nom du capteur, le nom du paramètre, sa valeur, son unité, l'identifiant du point, la longitude et la latitude, le point géométrique, l'heure, et un indice pour différencier les points de prélèvement des autres.

# Lire les données

R

```
data<-read.csv("aytre_25fev2021_ysiexo2_gps.csv",  
               header=TRUE,  
               sep=";",  
               dec="," ,  
               stringsAsFactors=FALSE,  
               encoding = "UTF-8")
```

Python

```
import pandas as pd  
  
dataframe = pd.read_csv("aytre_25fev2021_ysiexo2_gps.csv",  
                        sep=';' ,  
                        decimal=',' ,  
                        parse_dates=['date', 'local_time'])
```

# Accéder aux variables

R

# Pour rappel, la variable *data* contient l'ensemble du tableau

`data$property_name`

`data$property_value`

date	sensor_name	property_name	property_value	unit_code	unit_name	point_id	pt_longitude	pt_latitude	geom_point	local_time	is_reference
2021-02-25	yslexo2	ODOsat_%sat	93.750	%Sat	percent_saturation	36256	-1.132644	46.12575	c(-1.13264430888341, 46.1257502617766)	2021-02-25 13:58:01	FALSE
2021-02-25	yslexo2	fDOM_rfu	5.890	RFU	RelativeFluorescenceUnit	36256	-1.132644	46.12575	c(-1.13264430888341, 46.1257502617766)	2021-02-25 13:58:01	FALSE
2021-02-25	yslexo2	pH	8.480	pH	pH_units	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	salinity_ppt	22.680	ppt	parts_per_thousand	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	fDOM_rfu	5.890	RFU	RelativeFluorescenceUnit	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	temp_C	12.747	C	Celsius_degrees	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	conductivity_mS/cm	27.525	mS/cm	milliSiemens/centimeter	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	turbidity_NTU	31.530	NTU	NephelometricTurbidityUnit	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	ODOsat_%sat	93.780	%Sat	percent_saturation	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	depth_m	0.000	m	meters	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	chlorophyll_mg/L	6.170	ug/L	microgramms/liter	36424	-1.132641	46.12575	c(-1.13264091826368, 46.1257523333333)	2021-02-25 13:58:02	FALSE
2021-02-25	yslexo2	turbidity_NTU	31.720	NTU	NephelometricTurbidityUnit	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE
2021-02-25	yslexo2	conductivity_mS/cm	27.526	mS/cm	milliSiemens/centimeter	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE
2021-02-25	yslexo2	depth_m	0.000	m	meters	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE
2021-02-25	yslexo2	pH	8.480	pH	pH_units	36590	-1.132640	46.12575	c(-1.13264042218137, 46.1257527224264)	2021-02-25 13:58:03	FALSE

Python

`dataframe['property_name']`  
OU `dataframe.property_name`

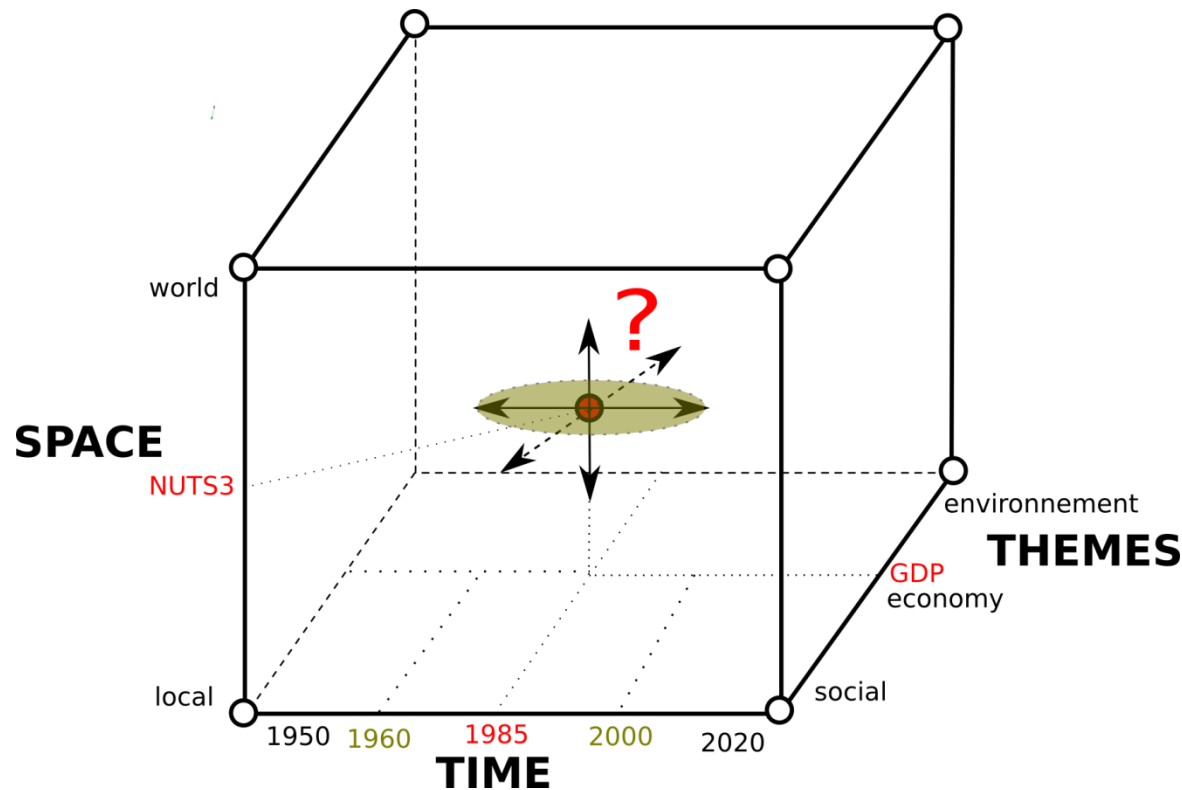
`dataframe['property_value']`

# Pour rappel, la variable *dataframe* contient l'ensemble du tableau

# Analyses temporelles

- Est-ce que des (ir)régularités saisonnières ou annuelles sont observables ?

Analyse temporelle





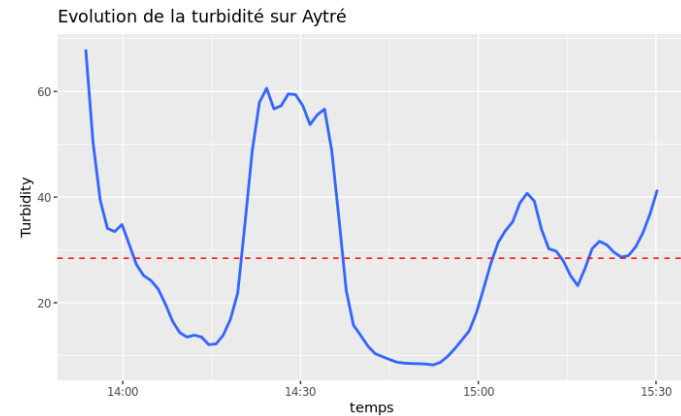
# Représentation temporelle

**Question :** Comment évolue la turbidité au cours de la mission ?

R

# On utilise ggplot

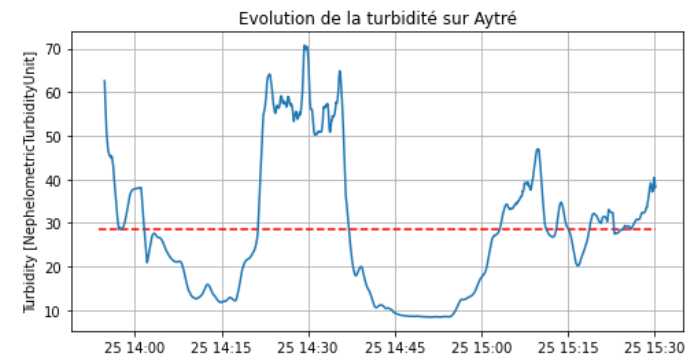
```
g <- ggplot(oneparam, aes(x=local_time, y=property_value))
g <- g + geom_smooth(method = 'loess', formula = 'y ~ x',
  span = 0.1, se=FALSE)
g <- g + geom_hline(yintercept=mean(oneparam$property_value),
  linetype="dashed", color = "red")
g <- g + ggtitle("Evolution de la turbidité sur Aytré")
g <- g + xlab("temps")
g <- g + ylab("Turbidity")
```



Python

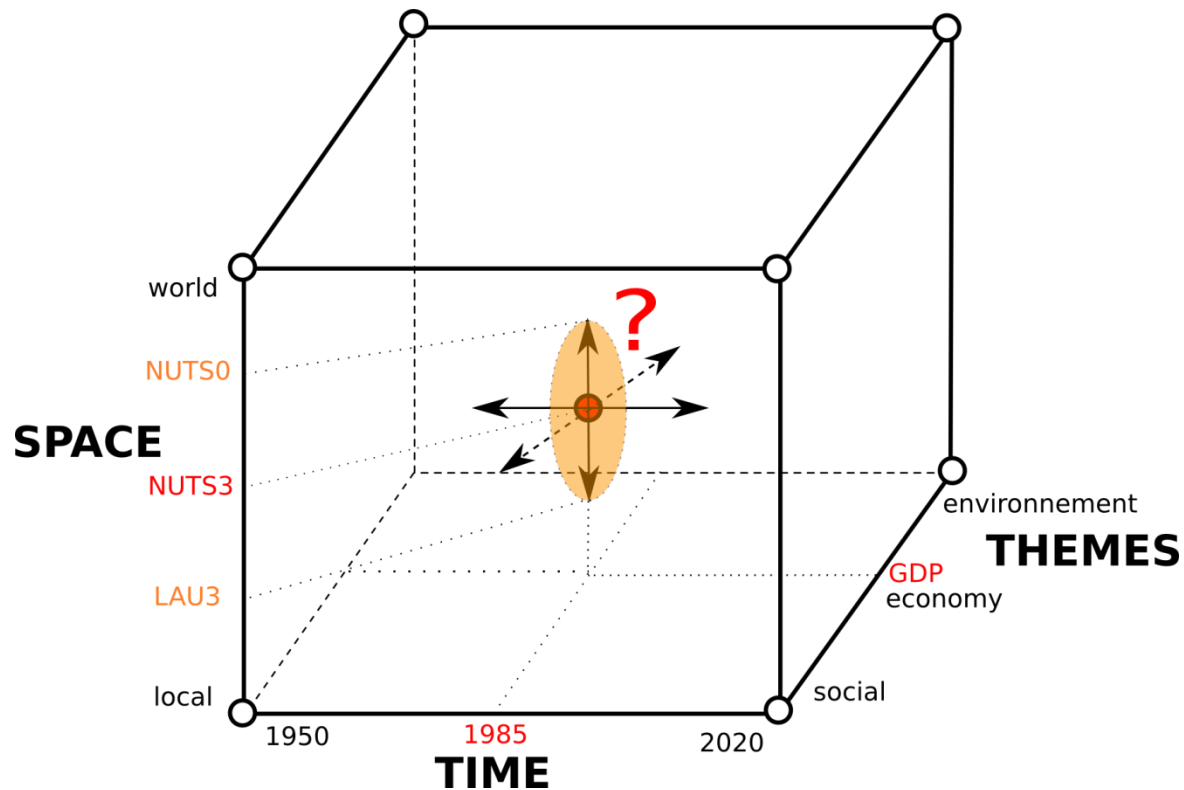
# On utilise matplotlib

```
plt.figure(figsize=(8, 4))
plt.plot(turbidite_reduit.index, turbidite_reduit.
  property_value.rolling(60).mean(), ms=1)
plt.hlines(turbidite_reduit.property_value.mean(), turbidite_reduit.index[0],
  turbidite_reduit.index[-1], linestyle='--', color='red')
plt.title('Evolution de la turbidité sur Aytré', fontsize=12)
plt.grid()
plt.ylabel(f'Turbidity [{turbidite_reduit.unit_name[0]}]', fontsize=10)
```



# Analyses spatiales

- Auto-corrélation spatiale ?



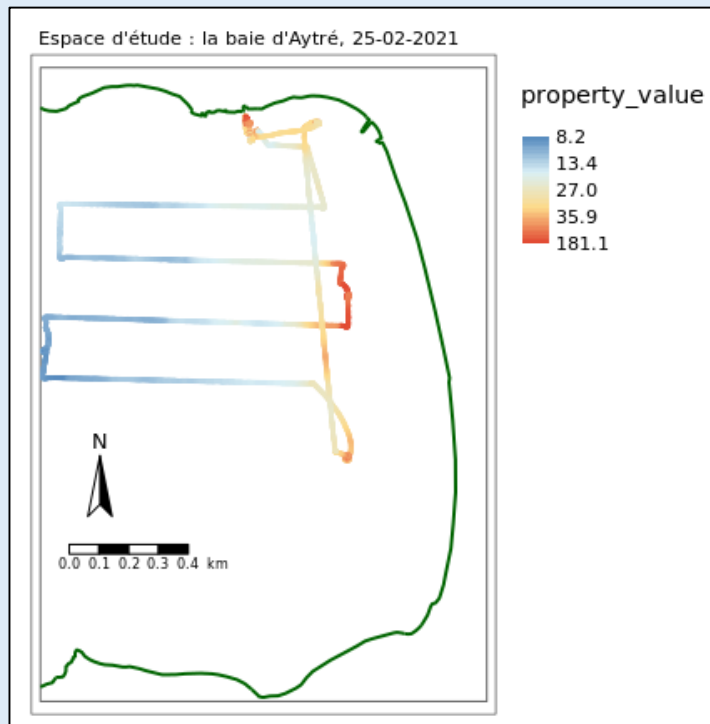
# Représentation spatiale (Trait de cote et turbidité)

**Question :** Comment évolue la turbidité dans la Baie d'Aytré ?

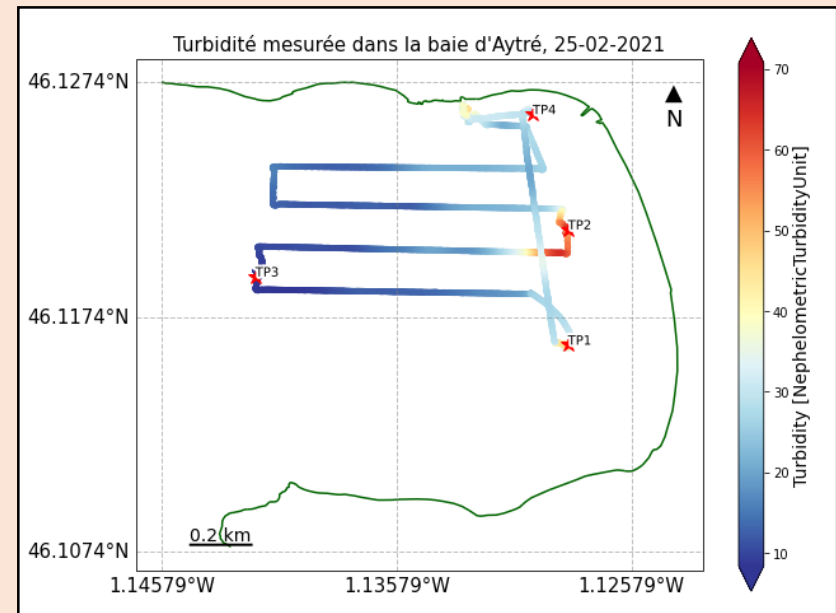
R

Python

# On utilise [tmap](#)

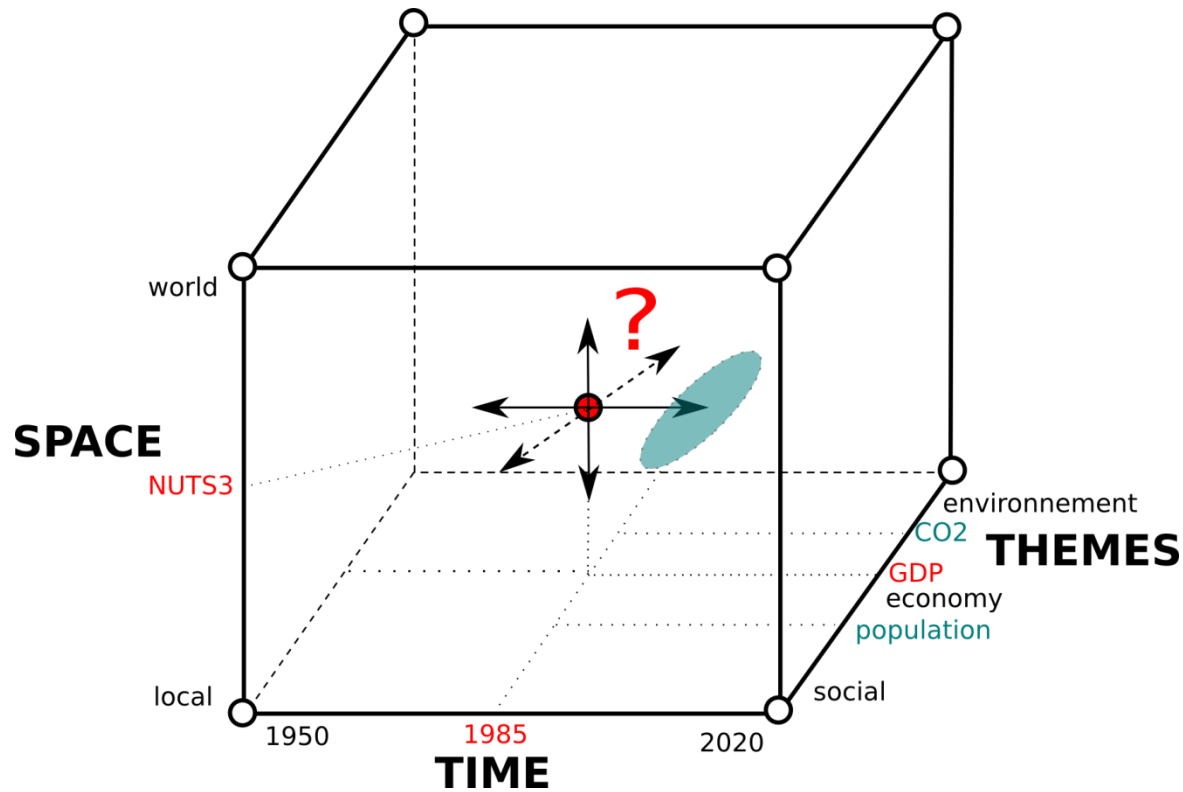


# On utilise [cartopy](#)



# Analyse thématique

- Est-ce qu'une relation connue entre variable existe ?



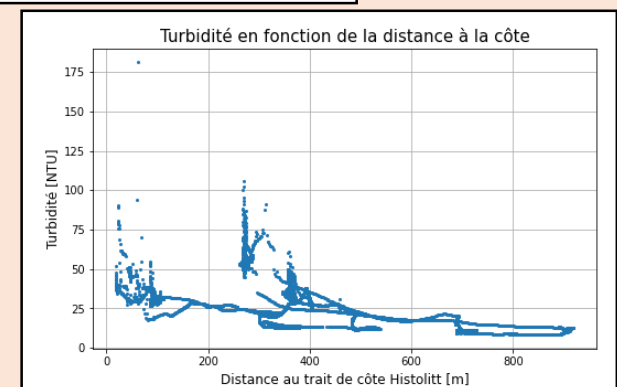
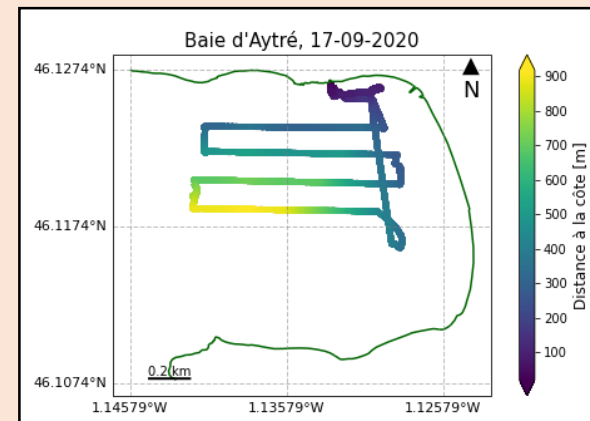
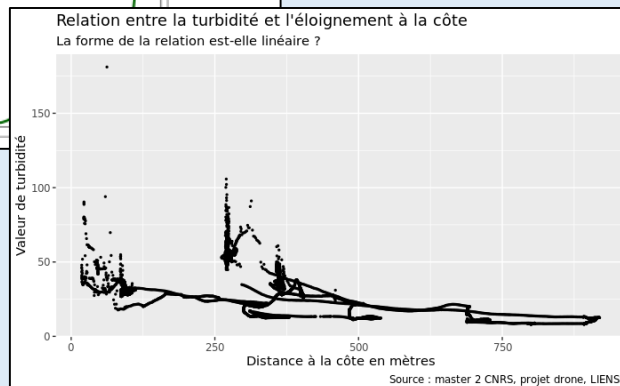
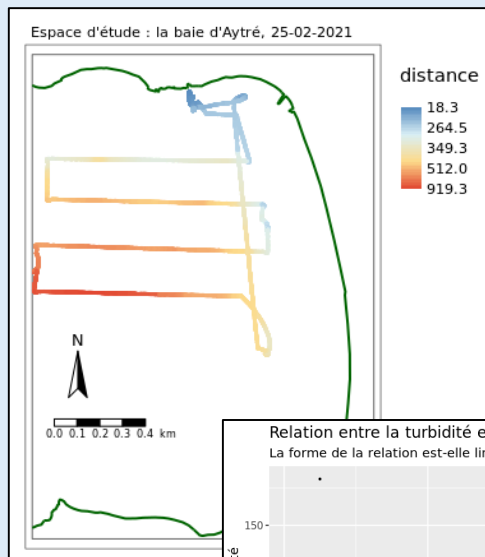
# Analyse thématique

**Question :** Y a-t-il un lien entre la distance à la côte et la turbidité ?

R

Python

Analyse thématique





# *PARTIE II – Manipulation des données*



**TP Coursus Master Ingénieur – Mars 2021**

Christine Plumejeaud,  
Clémence Chupin,  
U.M.R. 7266 LIENSS

# Accéder aux données et aux codes

- Le TP va se dérouler sur un « bureau virtuel »
  - Pour se connecter ...
- Les données sont stockées ...
- Possibilité de faire le TP avec
  - R avec R Studio (code « ... »)
  - Python avec Jupiter Notebook (code « ... »)

# Objectifs du TP

## 1. Lire les données de la base de données

- Lire les données de la sortie terrain de février 2021 (données spatialisées).
- Créer, puis lire, un fichier GeoJSON avec la localisation des points de prélèvements (avec éventuellement les résultats des analyses en labo).

## 2. Faire une carte du trajet suivi par PAMELi

- En utilisant les données SHAPEFILE (trait de côte, emprise de la zone, trajet de PAMELi), tracer le trajet de PAMELi durant la mission de février.
- Faire un zoom sur la zone à échantillonner.

## 3. Analyser l'évolution des paramètres mesurés

- Représenter l'évolution des paramètres dans le temps.
- Créer des statistiques pour comparer les différents paramètres (RATIO, STD).
- Représenter l'évolution spatiale des paramètres.
- Calculer la distance à la côte, et croiser avec les différents paramètres.

N'hésitez pas à tester d'autres méthodes de représentation/analyse des données pour répondre à vos propres questions !

# Librairies utiles

## R

- **Pour manipuler les données :** tidyverse
  - `library(tidyverse)`
  - `library(lubridate)`
- **Pour les données temporelles :** lubridate
- **Pour les données spatiales :** sf
  - <https://r-spatial.github.io/sf/articles/sf1.html>
  - `library(sp)` #pour bbox
- **Pour la cartographie**
  - `library(cartography)`
  - `library(tmap)`

## Python

- **Pour lire et manipuler les données :**
  - pandas
  - Geopandas (données spatiales)
  - Shapely (données spatiales)
- **Pour l'analyse de données :**
  - Numpy
  - Scipy
- **Pour l'affichage et la cartographie :**
  - matplotlib
  - Cartopy

# Sauver / reprendre son travail sous R

R

<https://mtes-mct.github.io/parcours-r/m1/sauvegarder-son-travail.html>

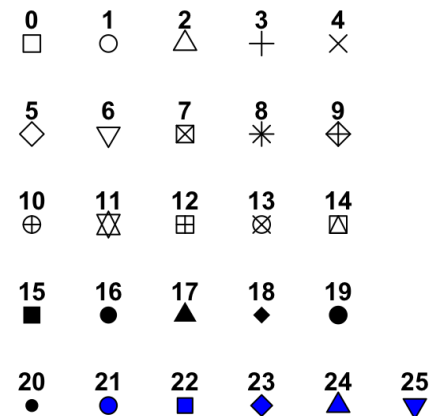
```
save(list = ls(), file = "outputs/env_entier.RData")  
# sauvegarde de tout l'environnement sur le répertoire choisi  
  
rm(list = ls()) # suppression de notre environnement dans R  
  
load("outputs/env_entier.RData")  
# chargement de l'environnement stocké sur l'ordinateur
```



# Les couleurs et les styles sous R

R

- <https://www.statmethods.net/advgraphs/parameters.html>
- <https://bookdown.org/hneth/ds4psy/D-2-apx-colors-essentials.html>
  - Nom de la couleur : « black »
  - Code RGB : (0,0,0)
  - Code hexadecimal : #000000
  - Code R : 0



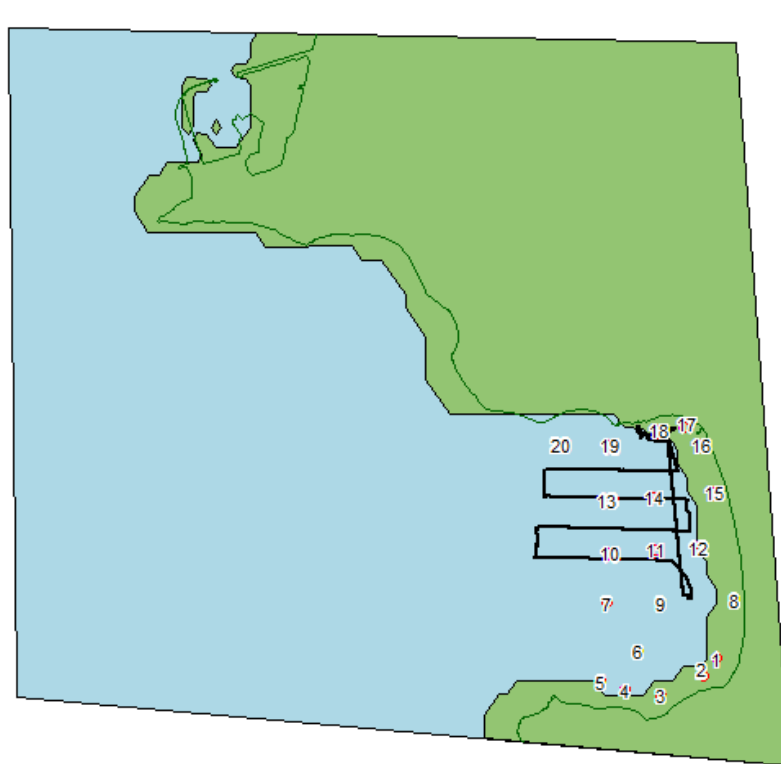
- <http://www.sthda.com/french/wiki/les-differents-types-de-points-dans-r-comment-utiliser-pch>
  - pch : forme du point
  - lty : forme de ligne
  - cex : text and symbol size

## 2. Cartographie de toute la mission

R

- Ajouter les points d'échantillonnage (sampling)

Espace d'étude sur la baie d'Aytré, 25-02-2021



Source : UMR 7266 LIENSS

# Fichier GeoJSON

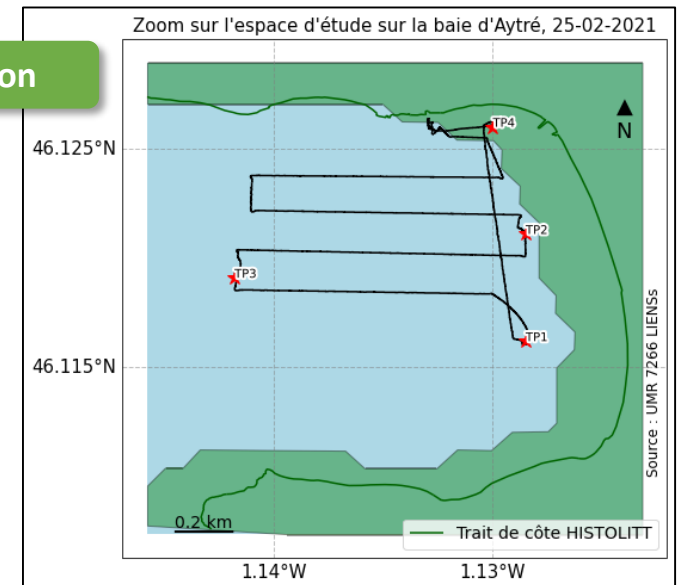
## Fichier GeoJSON

```
{
  "type": "FeatureCollection",
  "name": "vraiechantillons_4326",
  "crs": { "type": "name", "properties": { "name": "urn:ogc:def:crs:OGC:1.3:CRS84" } }, = Projection (WGS84 en °)
  "features": [
    { "type": "Feature", "properties": { "id": 1, "nom": "TP1", "%microplastique": 21.0, "geometry": { "type": "Point", "coordinates": [ -1.1285, 46.1162 ] } },
    { "type": "Feature", "properties": { "id": 1, "nom": "TP2", "%microplastique": 26.0, "geometry": { "type": "Point", "coordinates": [ -1.1285, 46.1211 ] } },
    { "type": "Feature", "properties": { "id": 1, "nom": "TP3", "%bacteries": 0.4, "geometry": { "type": "Point", "coordinates": [ -1.1418, 46.1191 ] } },
    { "type": "Feature", "properties": { "id": 1, "nom": "TP4", "%bacteries": 7.8, "geometry": { "type": "Point", "coordinates": [ -1.130, 46.126 ] } }
  ]
}
```

## Table après lecture

	id	nom	%microplastique	%bacteries	geometry
0	1	TP1	21.0	NaN	POINT (-1.12850 46.11620)
1	1	TP2	26.0	NaN	POINT (-1.12850 46.12110)
2	1	TP3	NaN	0.4	POINT (-1.14180 46.11910)
3	1	TP4	NaN	7.8	POINT (-1.13000 46.12600)

## Représentation

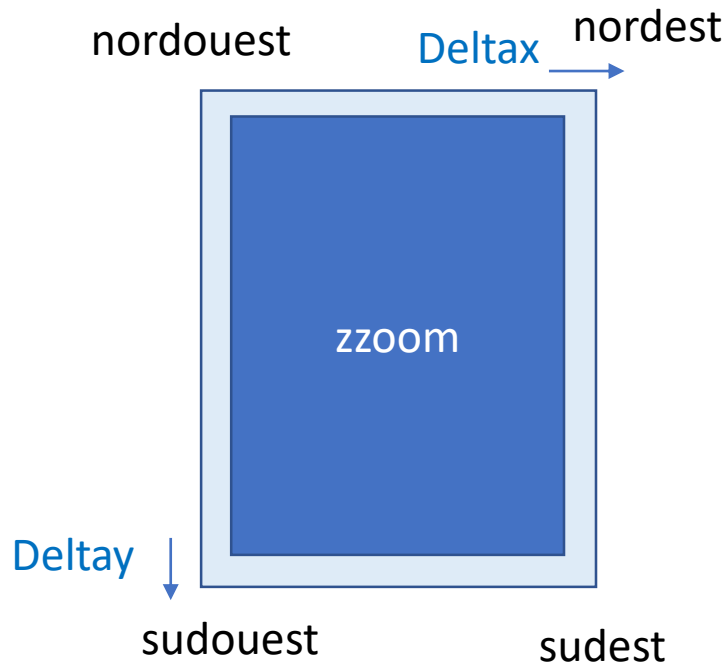


## 2. Cartographie de l'espace d'étude

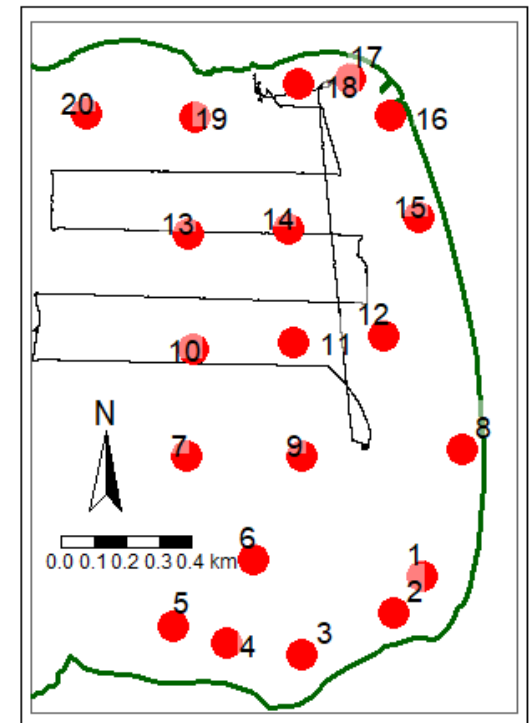
**R**

- Zoomer sur les points d'échantillonnage (sampling)

```
zzoom <- bbox(SpatialPoints(cbind(sampling$coord_x,
sampling$coord_y)))
```



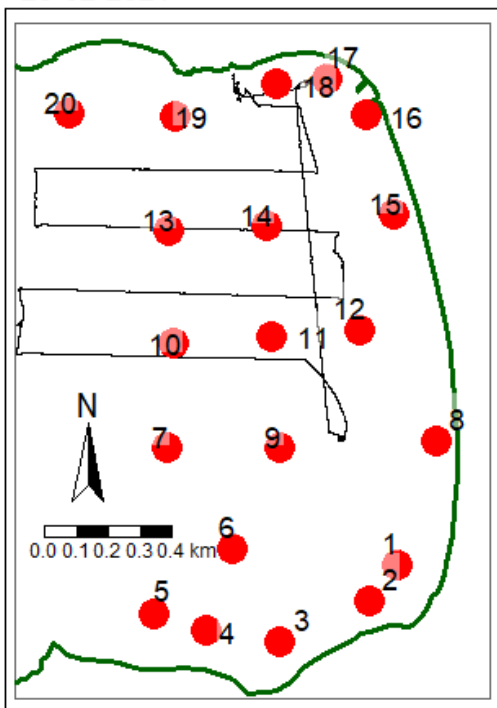
Zoom sur l'espace d'étude : la baie d'/  
25-02-2021



# 3. Evolution des paramètres

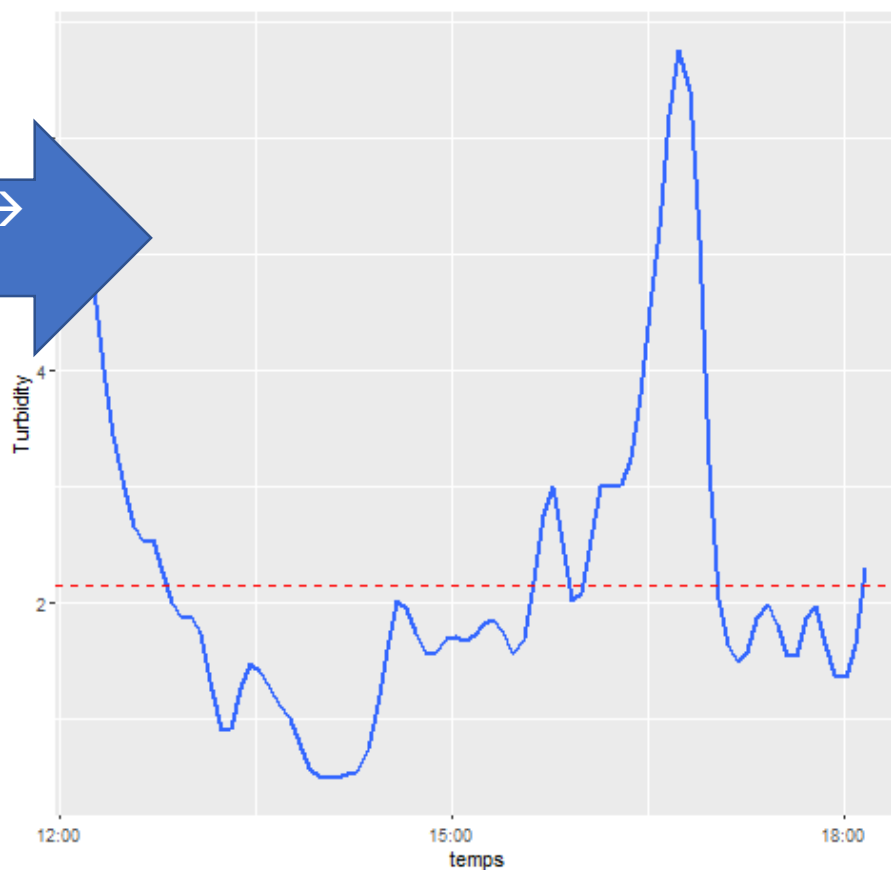
R

Zoom sur l'espace d'étude : la baie d'



Zoomsf →  
datasfz

Evolution de la turbidité au cours de la mission

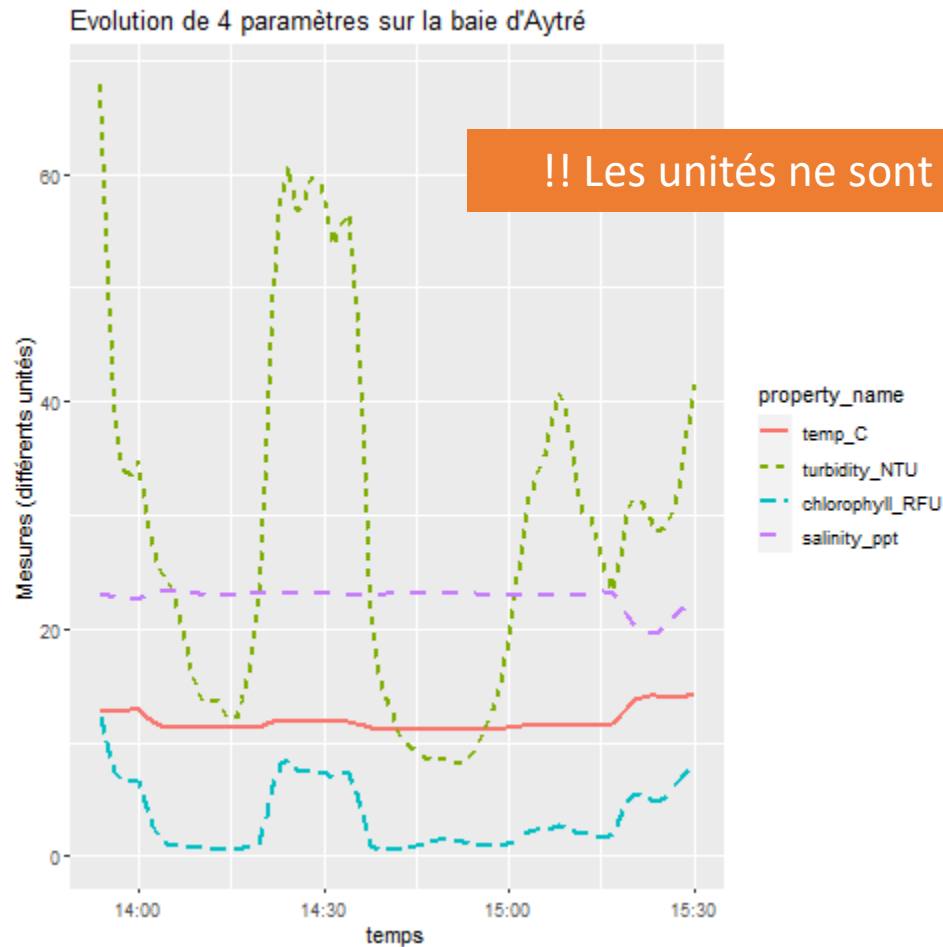




### 3. Evolution des paramètres

R

Mémo



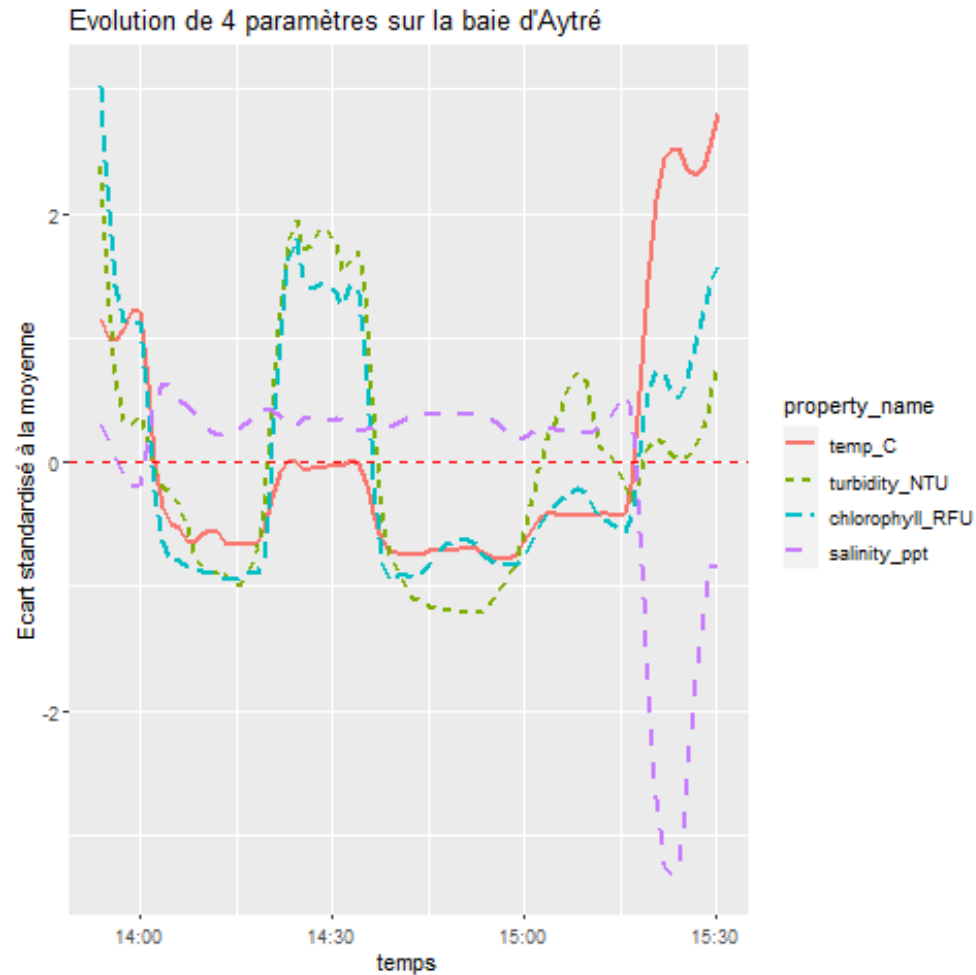
### 3. Evolution RELATIVE des paramètres

*Pour comparer les différents paramètres, on crée de nouvelles statistiques*

- La variable **RATIO** sans unité de mesure représente l'écart à sa moyenne pour un indicateur
  - $\text{RATIO} = -(1 - \text{value} / \text{moyenne}(\text{valeurs}))$
  - RATIO vaut 0 quand le paramètre est moyen
  - **RATIO >0 quand les valeurs mesurées sont au dessus de la moyenne**
  - **RATIO <0 quand les valeurs mesurées sont au dessous de la moyenne**
- La variable **STD** sans unité de mesure représente l'écart standardisé à sa moyenne pour un indicateur
  - $\text{STD} = (\text{value} - \text{moyenne}(\text{valeurs})) \div \text{écart-type}(\text{valeurs})$
  - $\text{Mean}(\text{STD}) = 0$
  - $\text{Ecart-type}(\text{STD}) = 1$

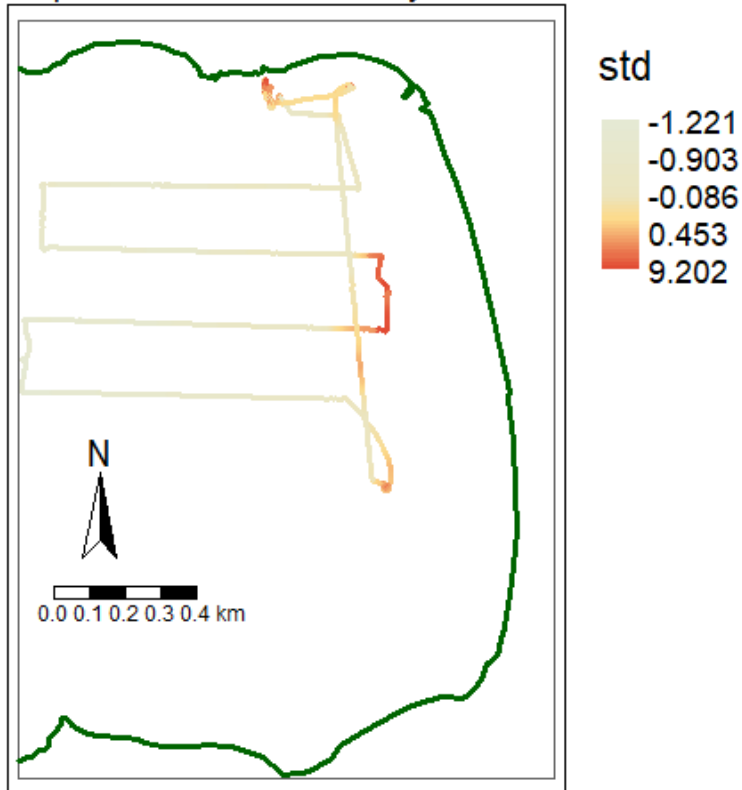
Possible que  
seulement parce que  
tous les paramètres  
mesurés restent  
positifs sur l'étude

### 3. Evolution conjointe des paramètres standardisés

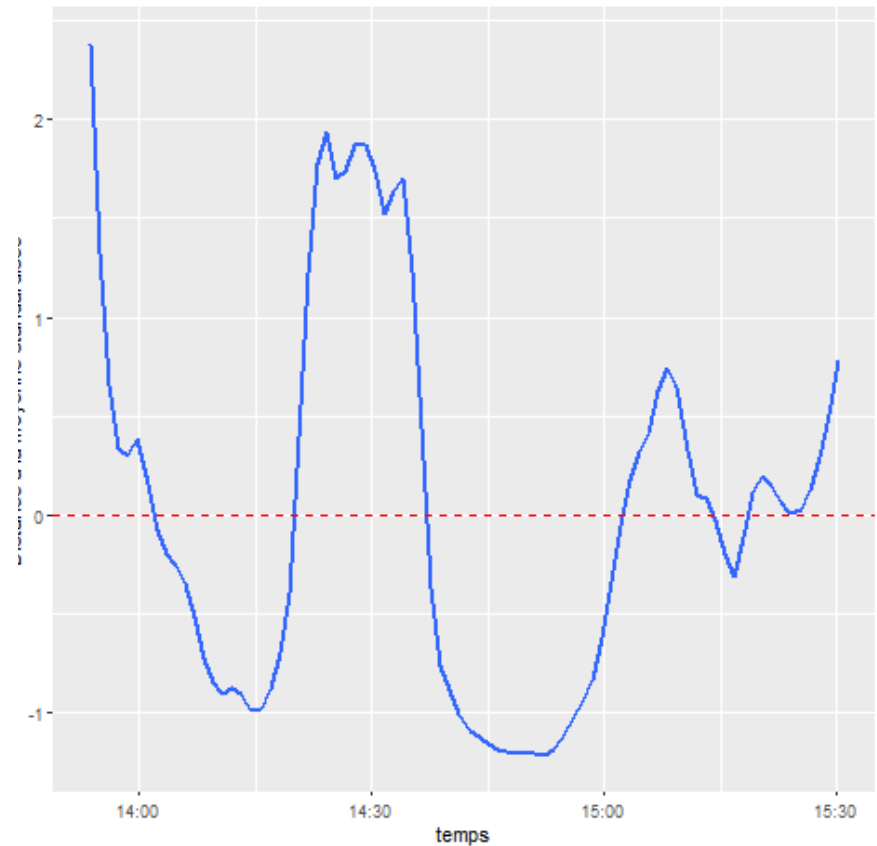


### 3. Evolution de la turbidité STD dans la baie

Espace d'étude : la baie d'Aytré, 25-02-2021



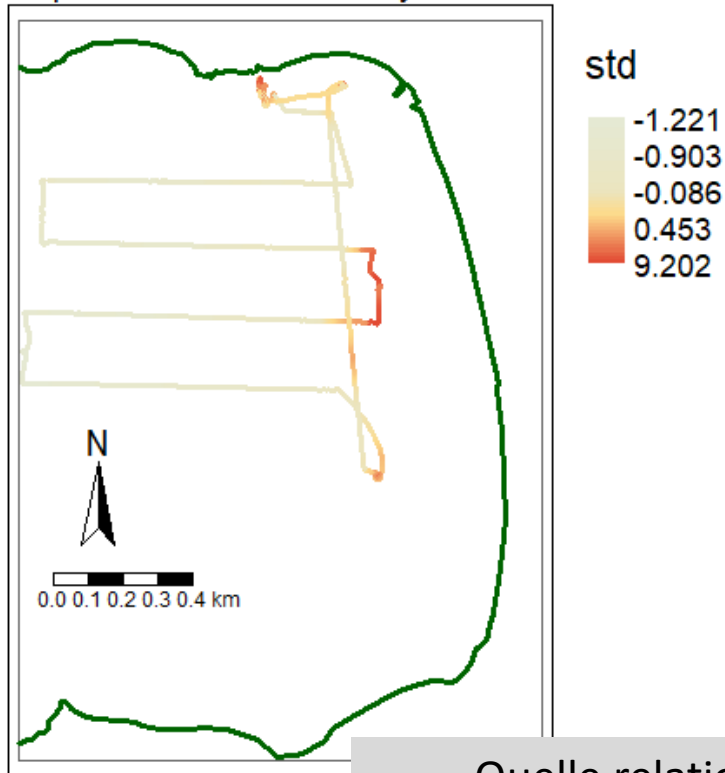
Evolution de la turbidité sur Aytré



### 3. Carte de température et turbidité STD

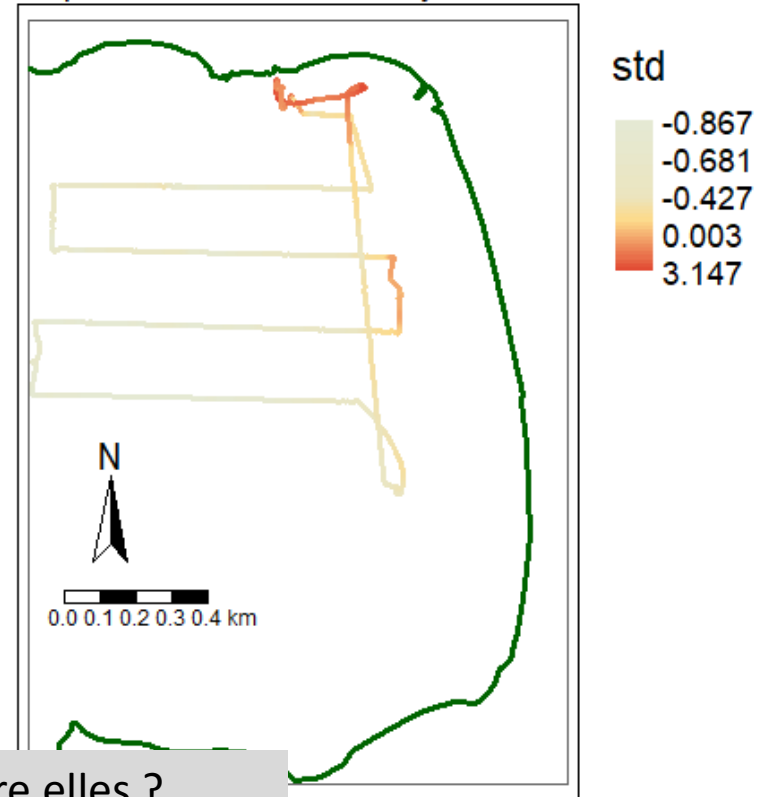
#### Turbidité

Espace d'étude : la baie d'Aytré, 25-02-2021



#### Température

Espace d'étude : la baie d'Aytré, 25-02-2021

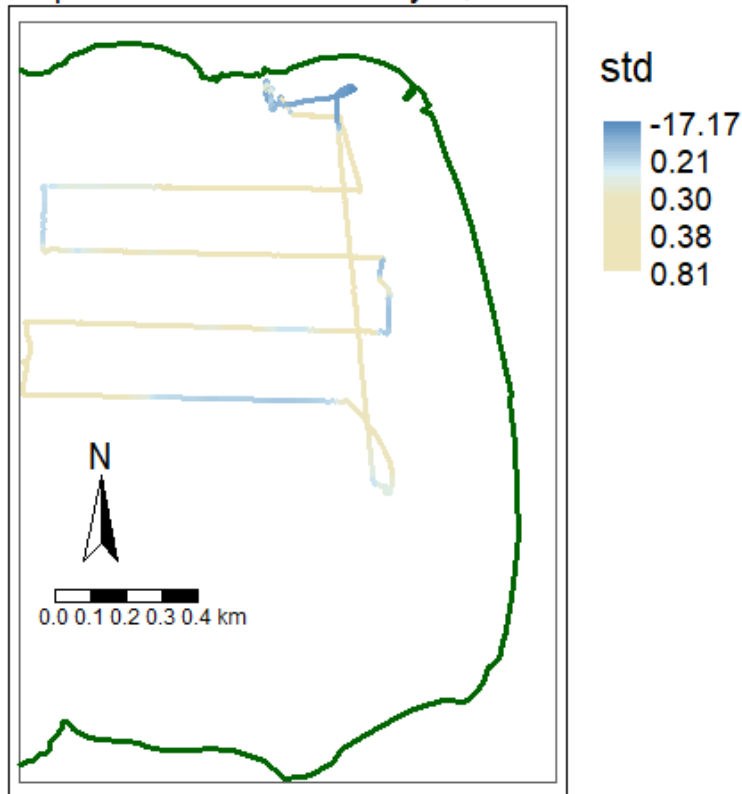


Quelle relation entre elles ?  
La faible bathymétrie par exemple ...

### 3. Carte de salinité et chlorophylle STD

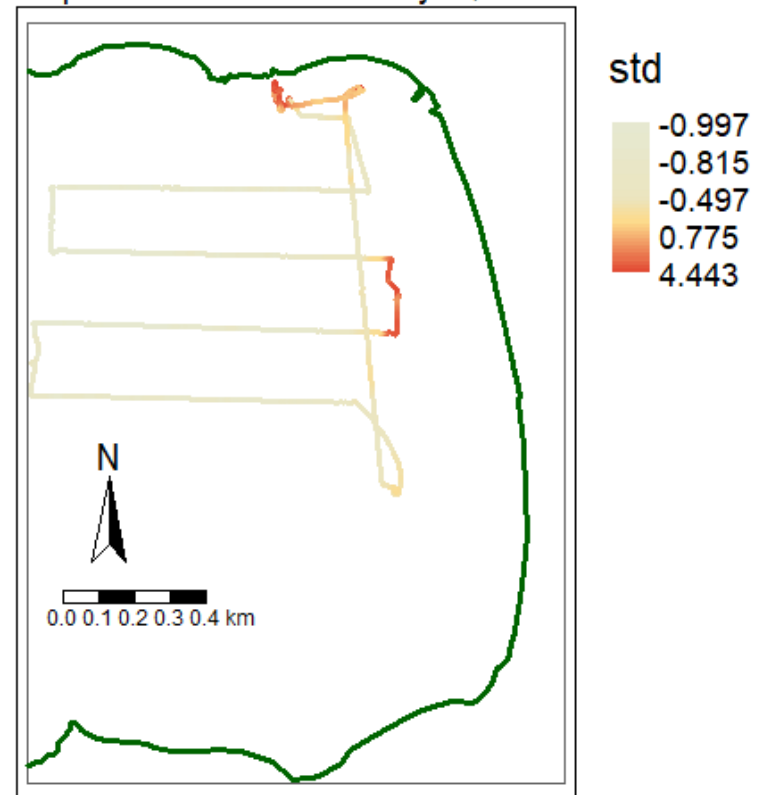
#### Salinité

Espace d'étude : la baie d'Aytré, 25-02-2021



#### Chlorophylle

Espace d'étude : la baie d'Aytré, 25-02-2021





# *PARTIE III – Pour aller plus loin dans l'analyse ...*



**TP Coursus Master Ingénieur – Mars 2021**

Christine Plumejeaud,  
Clémence Chupin,  
U.M.R. 7266 LIENSS

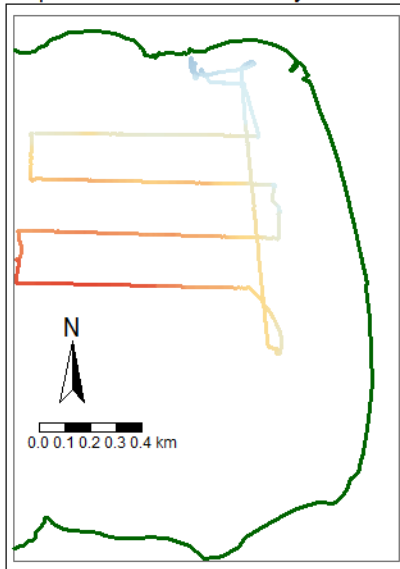
# Relations à explorer

1. **La turbidité** est assez intéressante car on peut y voir une information sur la remise en suspension des sédiments de la baie.
  - On pourra regarder par rapport à la pente de la baie, qui a un impact direct sur le déferlement lors de la marée montante?
  - On peut regarder par rapport à la distance à la côte.
2. On a souvent une corrélation entre la **turbidité et la chlorophylle** avant calibration car il s'agit dans les deux cas de capteurs optiques.
3. Calcul de l'**oxygène dissous attendu par effet thermique et salin**  
(Aminot & Kérouel, 2004)
  - [https://envlit.ifremer.fr/infos/actualite/2004/hydrologie\\_des\\_ecosystemes\\_marins\\_parametres\\_et\\_analyses](https://envlit.ifremer.fr/infos/actualite/2004/hydrologie_des_ecosystemes_marins_parametres_et_analyses)
  - L'oxygène dissous en surface est très rarement limitant ou éloigné de 100%.

# 1. Evolution de la turbidité/distance à la côte

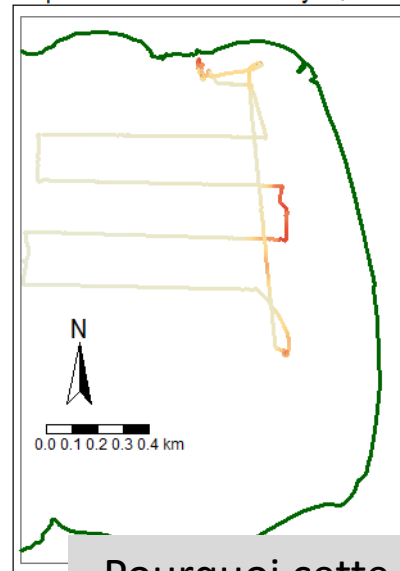
## Distance à la côte

Espace d'étude : la baie d'Aytré, 25-02-2021

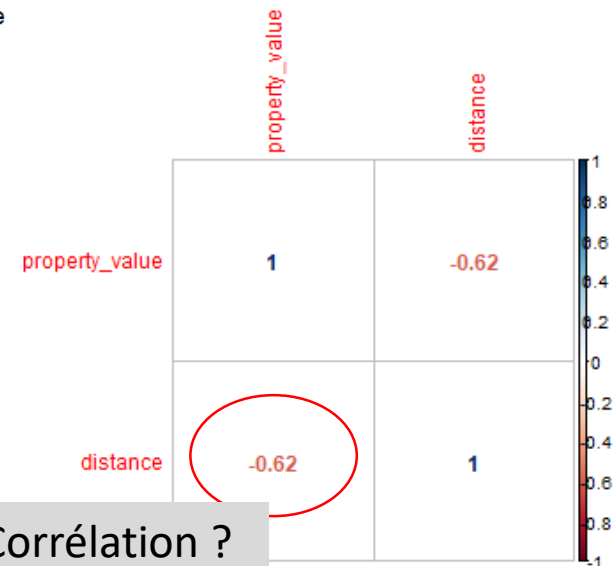
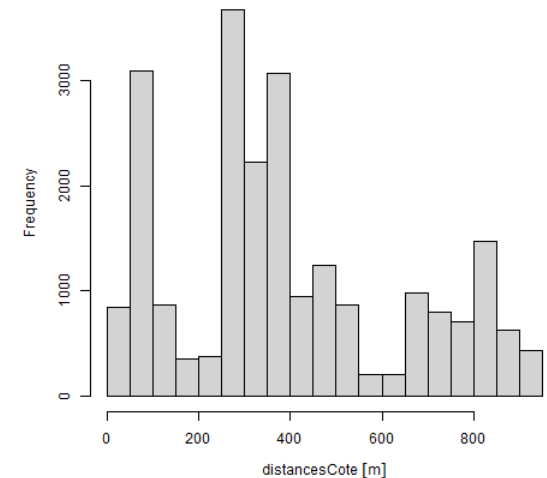


## Valeurs de turbidité NTU

Espace d'étude : la baie d'Aytré, 25-02-2021



Histogram of distancesCote



Pourquoi cette valeur ? Corrélation ?

# 1. Relation (linéaire?) turbidité/distance à la côte

R

- On peut tester l'application d'un modèle linéaire

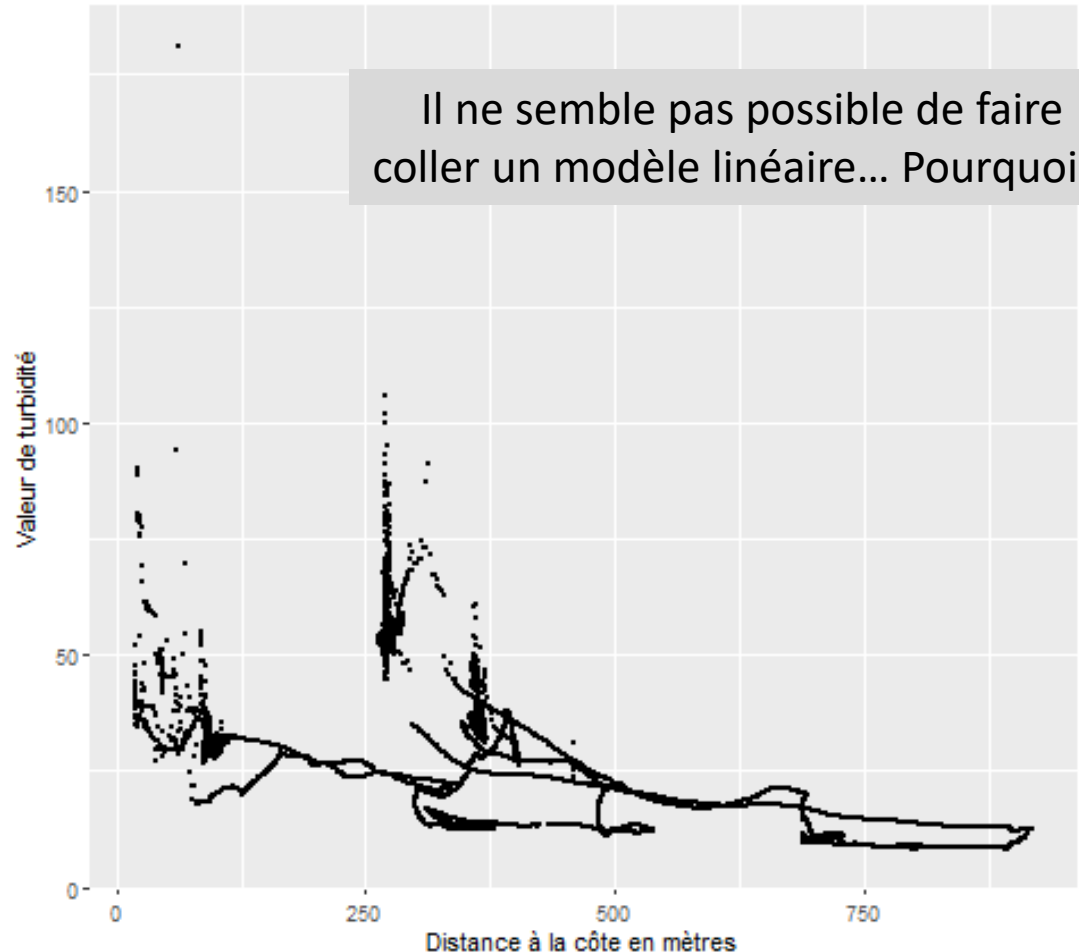
Model 1

```
(Intercept) 44.74 *
             [44.11; 45.37]
distance    -0.04 *
             [-0.04; -0.04]
```

```
R^2          0.38
Adj. R^2     0.38
Num. obs.    5752
```

\* 0 outside the confidence interval.

Relation entre la turbidité et l'éloignement à la côte  
La forme de la relation est-elle linéaire ?

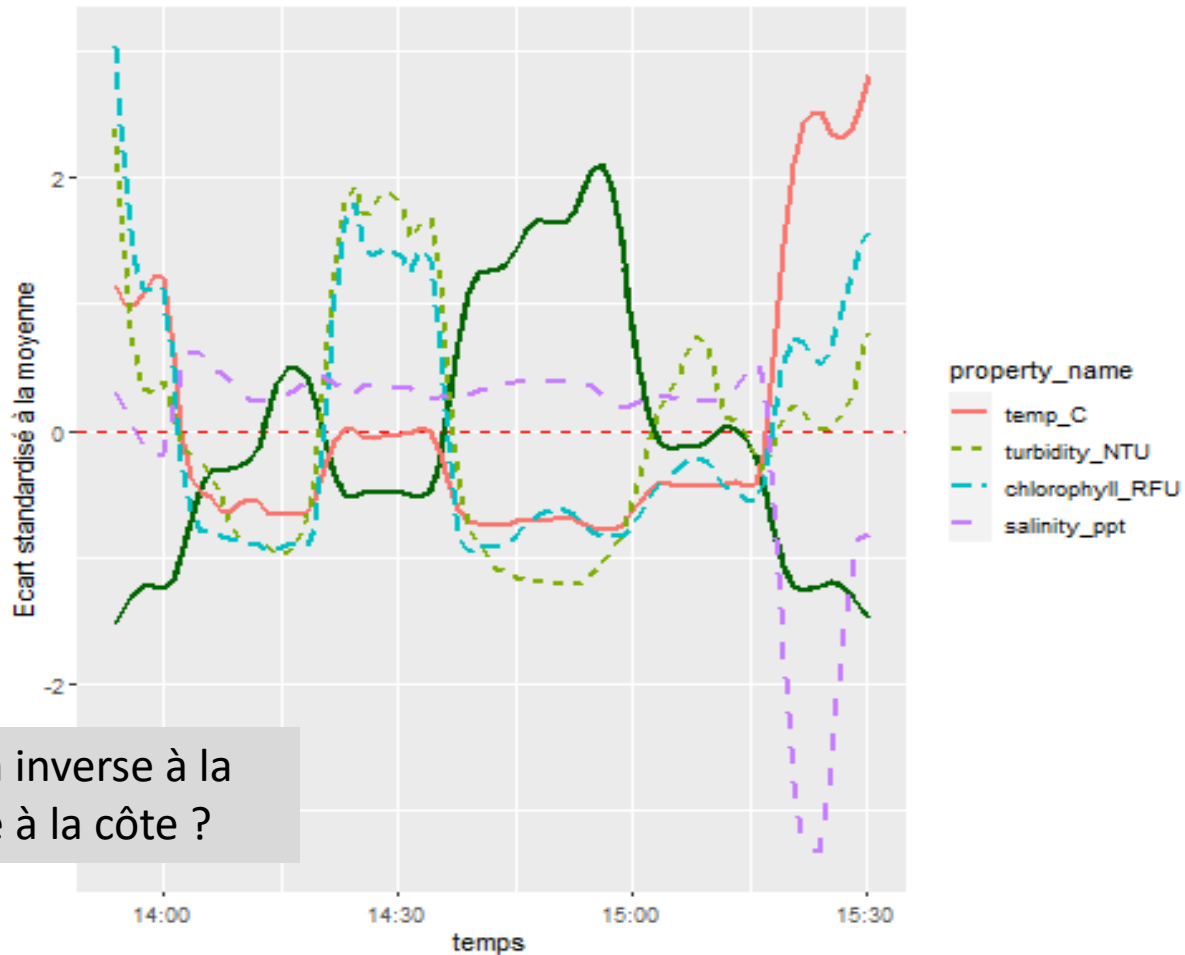


Source : master 2 CNRS, projet drone, LIENS5

Pour aller plus loin...

# 1. Evolution paramètres/distance à la côte

Evolution de 4 paramètres sur la baie d'Aytré, avec la distance standardisée à la côte

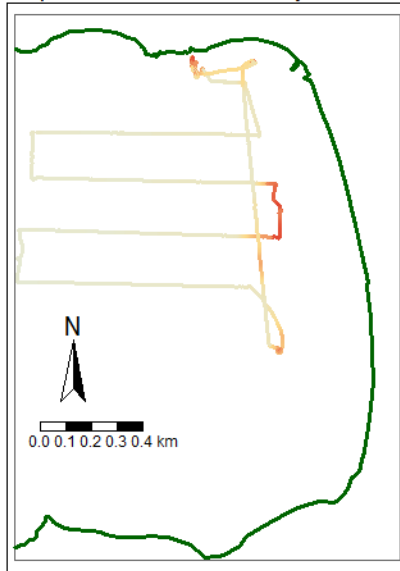


Evolution inverse à la distance à la côte ?

## 2. Lien chlorophylle / turbidité

### Turbidité

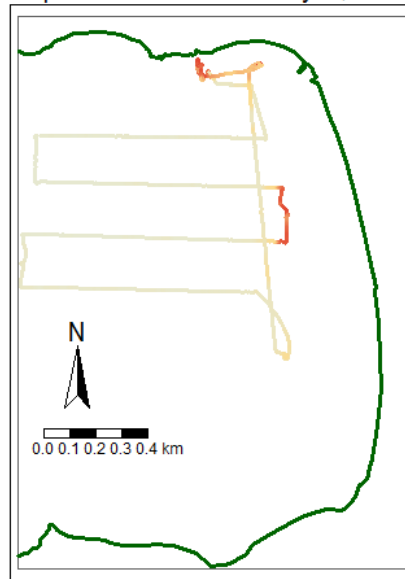
Espace d'étude : la baie d'Aytré, 25-02-2021



std  
-1.221  
-0.903  
-0.086  
0.453  
9.202

### Chlorophylle

Espace d'étude : la baie d'Aytré, 25-02-2021



std  
-0.997  
-0.815  
-0.497  
0.775  
4.443

### Corrélations

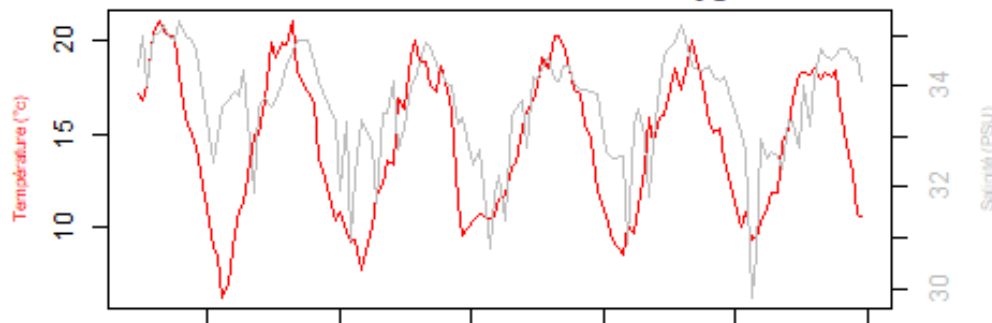
	turbidity	chloro	temp	salinity
turbidity	1	0.85	0.35	-0.07
chloro	0.85	1	0.64	-0.27
temp	0.35	0.64	1	-0.76
salinity	-0.07	-0.27	-0.76	1

Pourquoi cette corrélation forte ?

Pour aller plus loin...

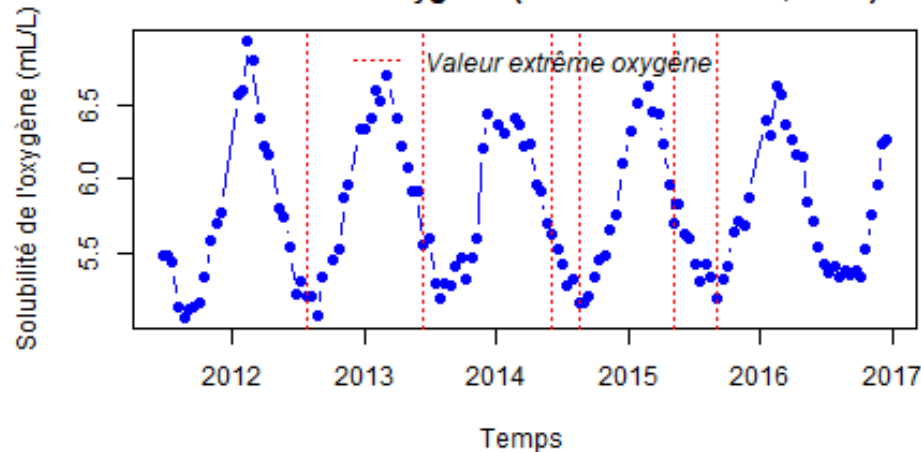
### 3. La relation entre O<sub>2</sub> dissous et (salinité, température)

**Température et Salinité :  
Influence sur la solubilité de l'oxygène**



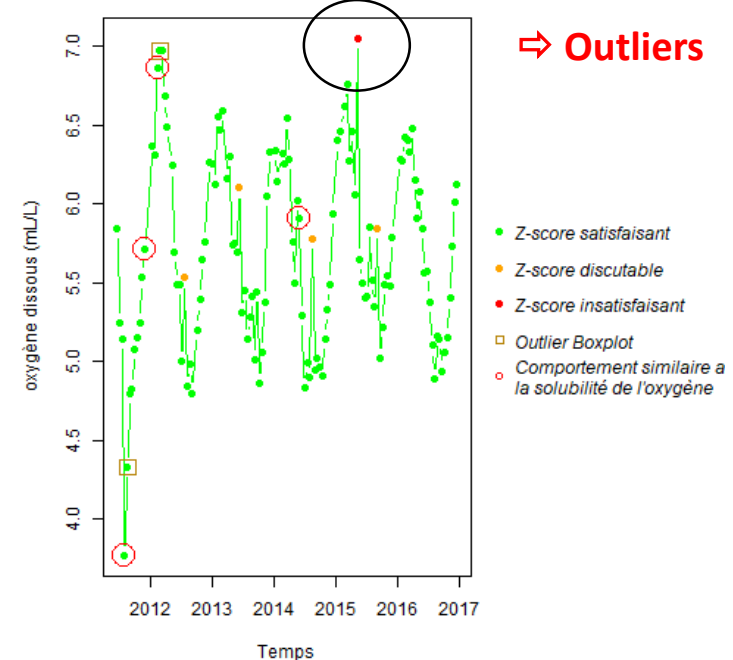
Calcul de l'oxygène attendu par effet thermique et salin (Aminot & Kérouel, 2004)  
Détection en ligne rouge des dates avec valeurs atypiques "outliers"

**Solubilité de l'oxygène (Aminot & Kérouel, 2004)**



Un **modèle physique** explique bien la **concentration d'oxygène dissous** en fonction des paramètres **Température et Salinité**.

**Similitude avec la solubilité de l'oxygène**



⇒ **Outliers**

### 3. La relation avec un paramètre biologique (Chlorophylle)

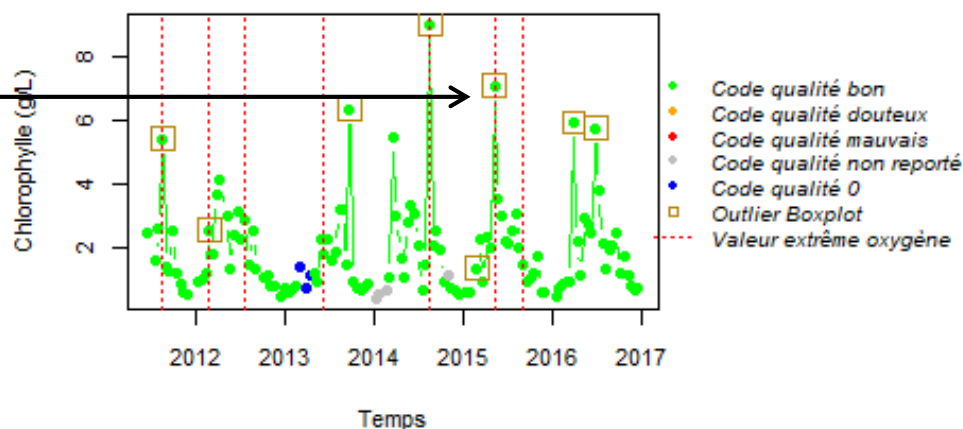
Le seul pic non expliqué par l'équation correspond à un pic de chlorophylle

Les pics de chlorophylle correspondent aux pics d'oxygène dissous :

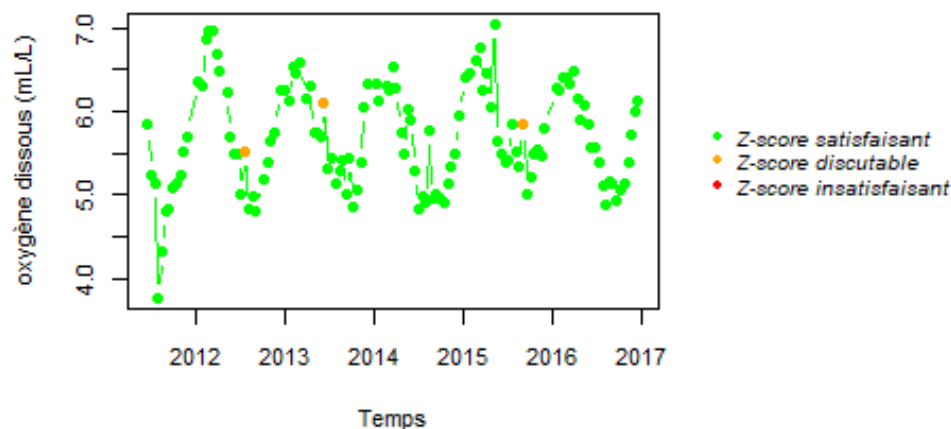
→ Le Zscore insatisfaisant peut être considéré comme concordant avec l'évolution de la chlorophylle.

Dans ce cas il est plausible de ne pas qualifier une valeur atypique comme réellement atypique. Ne reste que 3 valeurs qualifiées d'atypiques d'après le Zscore modifié.

Chlorophylle a / Codes qualite / Outliers du boxplot



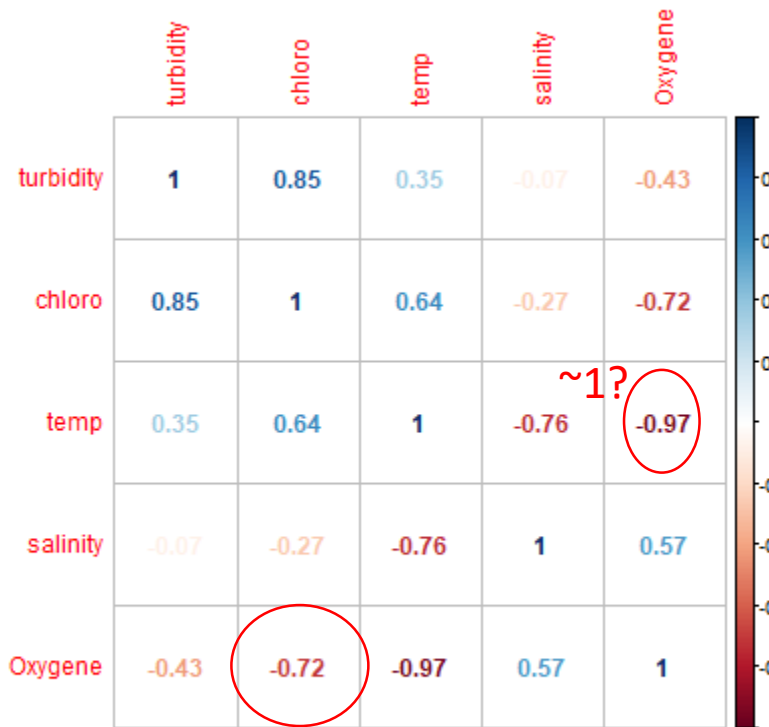
Analyse après comparaison



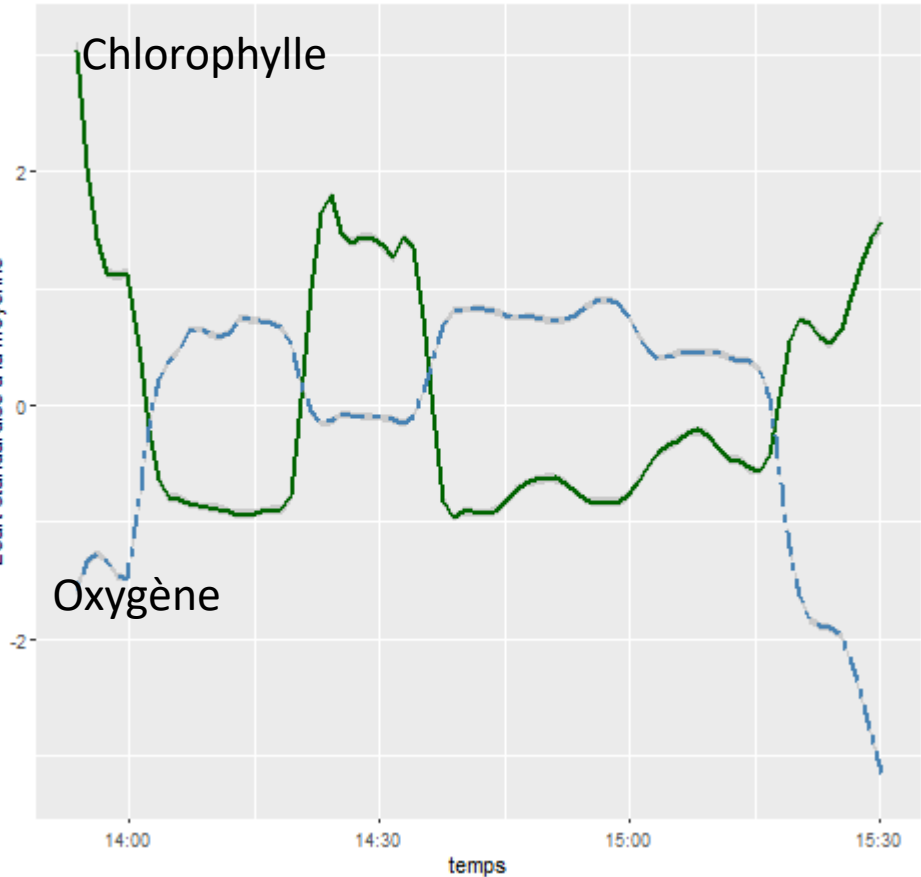


### 3. O<sub>2</sub> dissout *estimé* / chlorophylle

Evolution conjointe de la chlorophylle et l'oxygène dissous sur la baie d'Aytré



-0.72 ???



Pour aller plus loin...