

# Traitements automatisés des données issues d'un sondeur acoustique multifaisceaux pour la cartographie des fonds marins

Soutenance de Thèse de Marie LAMOURET en présence de :

M Pierre BOISSERY

M Guillaume MATTE

Pr Caroline FOSSATI

Pr Nadège THIRION-MOREAU

Pr Ali MANSOUR

M Christophe VIALA

Pr Jérôme MARS

# Plan

## Contexte des travaux

- 1 – Données issues du sondeur multifaisceaux
- 2 – Estimation de valeurs bathymétriques par réseau de neurones convolutifs
- 3 – Classification des données pour l'identification des habitats marins

Conclusion générale

Activités de Seaviews

Cartographies des fonds marins

Cartographies bathymétriques

Cartographies des habitats marins

Méthodes cartographiques

Analyse et traitements des données

# Activités de Seaviews



**SEAVIEWS**

: Bureau d'études spécialisé en cartographie pour le milieu marin

## Prestation de services :

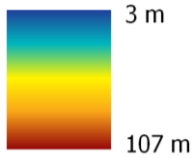
- Production de cartes bathymétriques, de cartes des habitats marins
- Réalisation de suivis écologiques, d'orthophotographies sous-marines, etc.

## Projets de R&D :

- Mise au point de matériel → Plateforme de surface pour réaliser des photogrammétries
- Développement de méthodes → Classification des habitats marins au sondeur multifaisceaux (SMF)
- Exploitation de données → Imagerie de la colonne d'eau pour l'étude des bancs de poissons
- Nouvelles applications → Photogrammétrie sous-marine géoréférencée à l'aide du SMF

# Cartographies des fonds marins

## LEGENDE



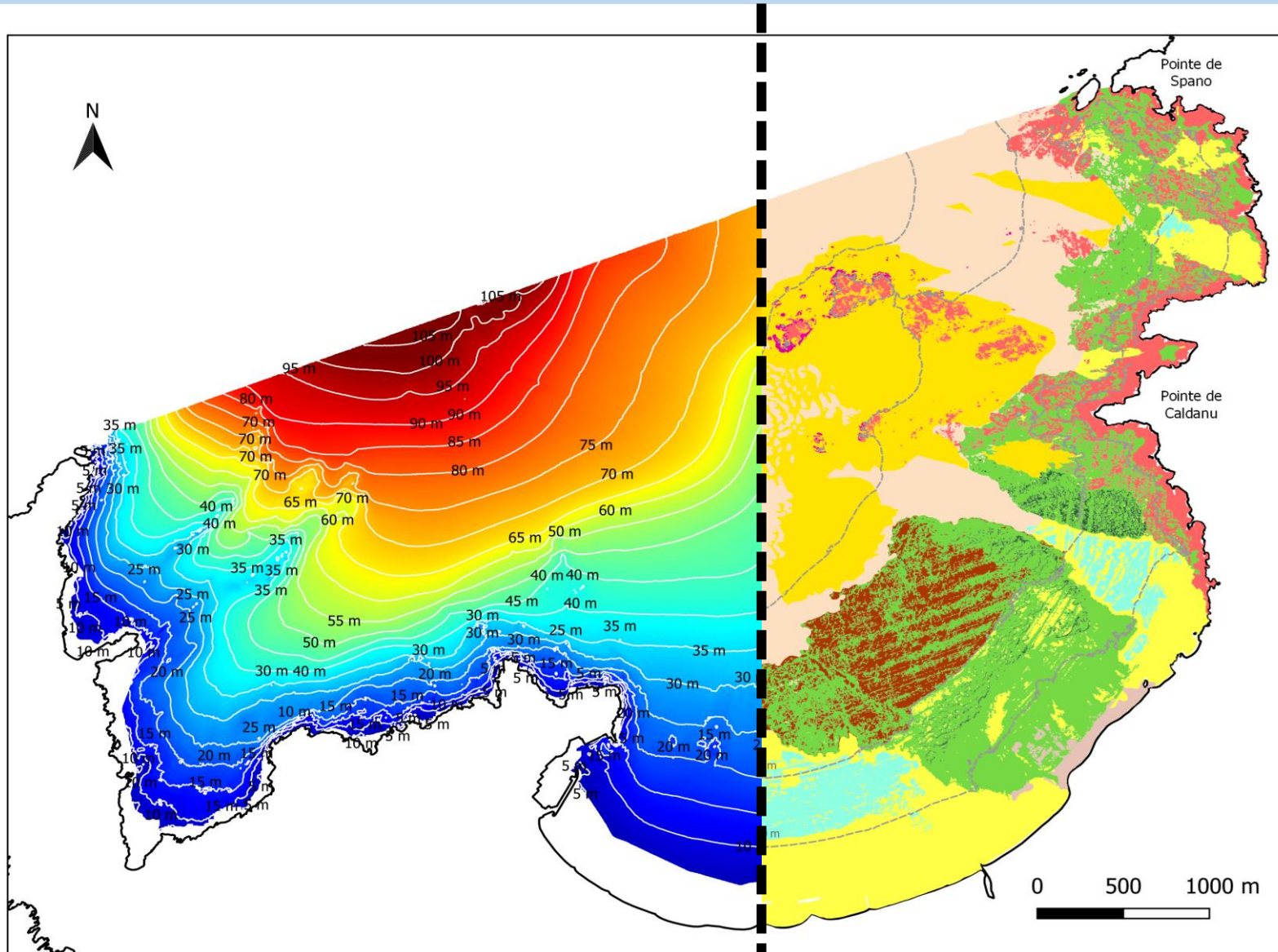
Maître d'ouvrage : STARESO  
Maître d'oeuvre : Seaviews

Acquisition des données : juillet 2018

Matériel d'acquisition : sondeur multifaisceaux R2Sonic 2022  
Traitement des données : View SMF, ViewMap

Edition des cartes : QGIS

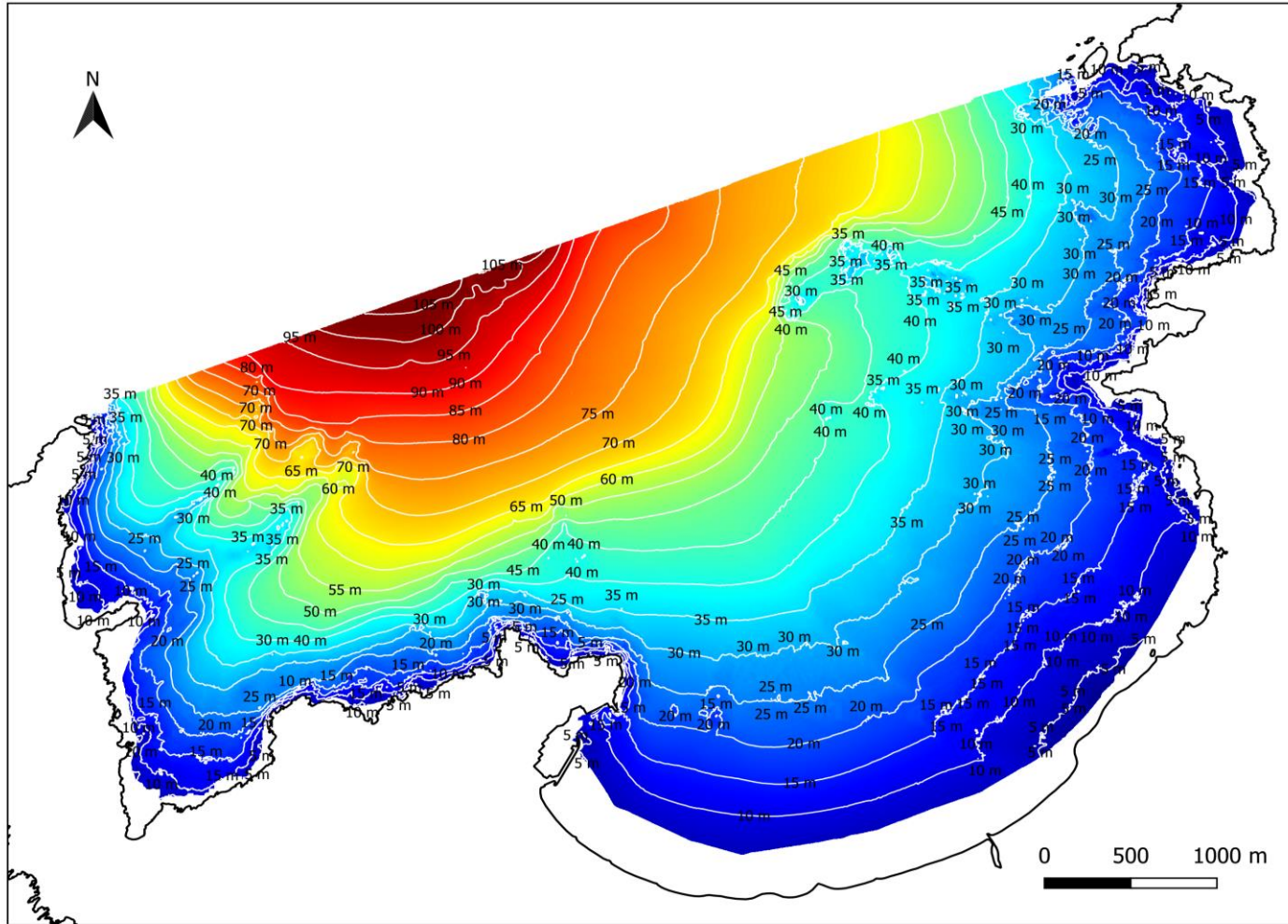
Projection : RGF93



## LEGENDE

- Sables fins bien calibrés
- Sédiment grossier
- Sédiments meubles
- Galets infralittoraux
- Roche infralittorale à algues photophiles
- Biocénose coralligène
- Herbier à *Posidonia oceanica*
- Matte morte de *Posidonia oceanica*
- Intermattes de *Posidonia oceanica*
- Association à *Cymodocea nodosa*
- Substrat artificiel

# Cartographie Bathymétrique



Carte bathymétrique :

- Profondeurs de la mer
- Description du paysage sous-marin

Besoins :

- Carte de sécurité de navigation
- Suivi de travaux sous-marins
- Suivi temporel
- Études océanographiques

# Cartographie des habitats marins

« Habitat écologique » : espace homogène en

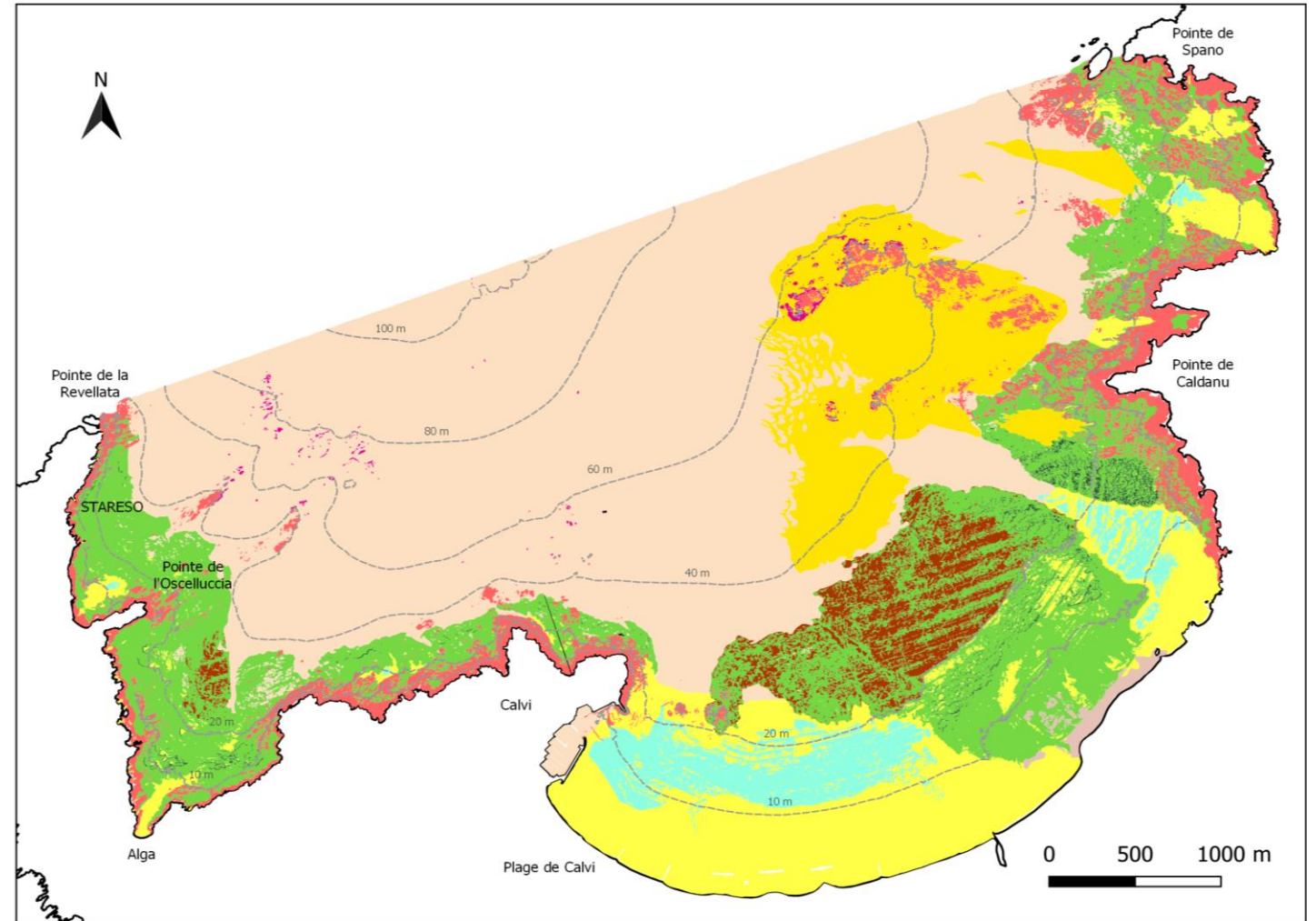
- Propriétés physiques
- Faune et flore
- Rôles

Carte des habitats :

- Inventaire des habitats écologiques
- Répartition dans une aire d'étude

Besoins :

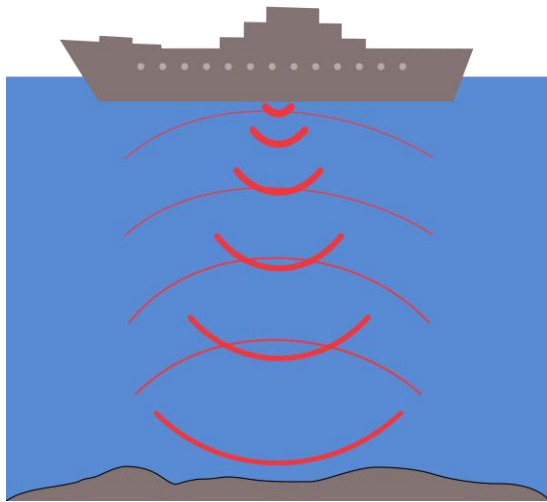
- Définir la gestion d'une zone
- Préparer des aménagements



# Méthode cartographique

1

Acquisition de données  
par écho-sondage



2

Analyses et traitements  
des données

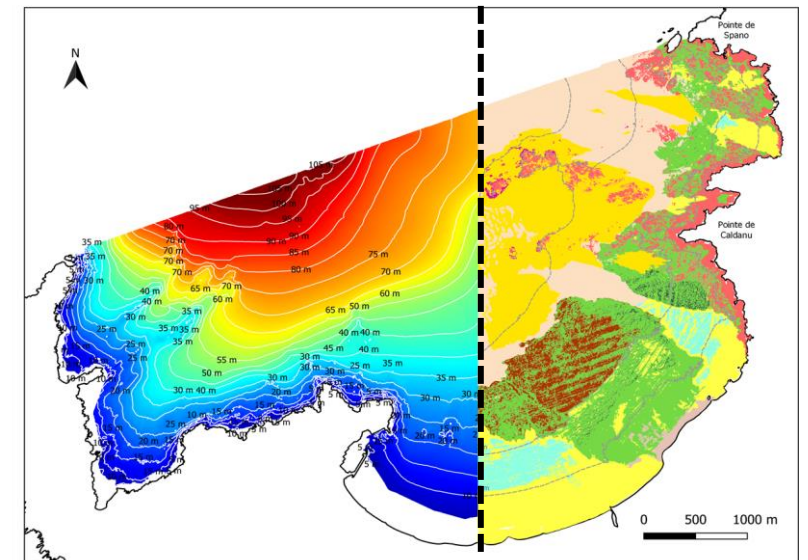


- Filtrage de données
- Transformation
- Extraction d'information
- Classification
- Détection
- Mise en forme

...

3

Élaboration de produits  
cartographiques



# Analyse et traitements des données



Seaviews développe ses logiciels pour :

- ① L'acquisition des données
- ② Les traitements des données
- ③ La production de cartes

| Actuellement  | En cours / A venir  |
|---|---|
| Méthodes +/- manuelles  | Méthodes + automatisées   |
| <ul style="list-style-type: none"><li>+ L'opérateur a le contrôle sur le traitement</li><li>- Traitements longs</li><li>- Tâches complexes ou répétitives</li><li>- Subjectivité de l'opérateur</li></ul> | <ul style="list-style-type: none"><li>+ Gain temps</li><li>+ Traitements plus reproductibles</li><li>+ Aide à la décision de l'opérateur</li><li>+ Analyses plus poussées</li><li>+ Développement de nouvelles applications</li></ul> |



# Plan

Contexte des travaux

**1 – Données issues du sondeur multifaisceaux  
(SMF)**



**Principe d'acquisition au SMF**

**Chaîne d'acquisition complète**

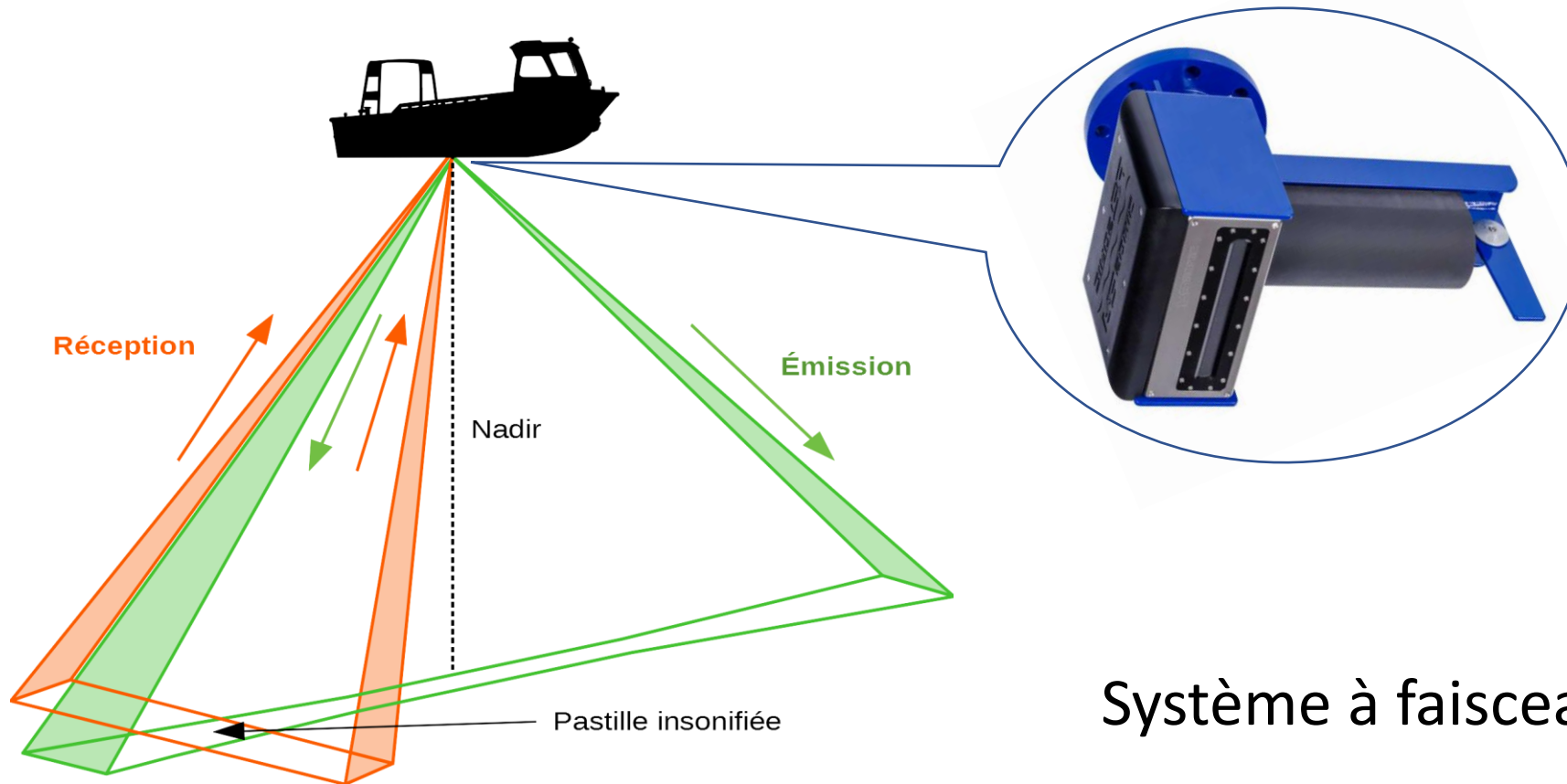
**Les trois types de données acquises au SMF**

2 – Estimation de valeurs bathymétriques  
par réseau de neurones convolutifs

3 – Classification des données pour  
l'identification des habitats marins

Conclusion générale

# Principe d'acquisition au sondeur multifaisceaux

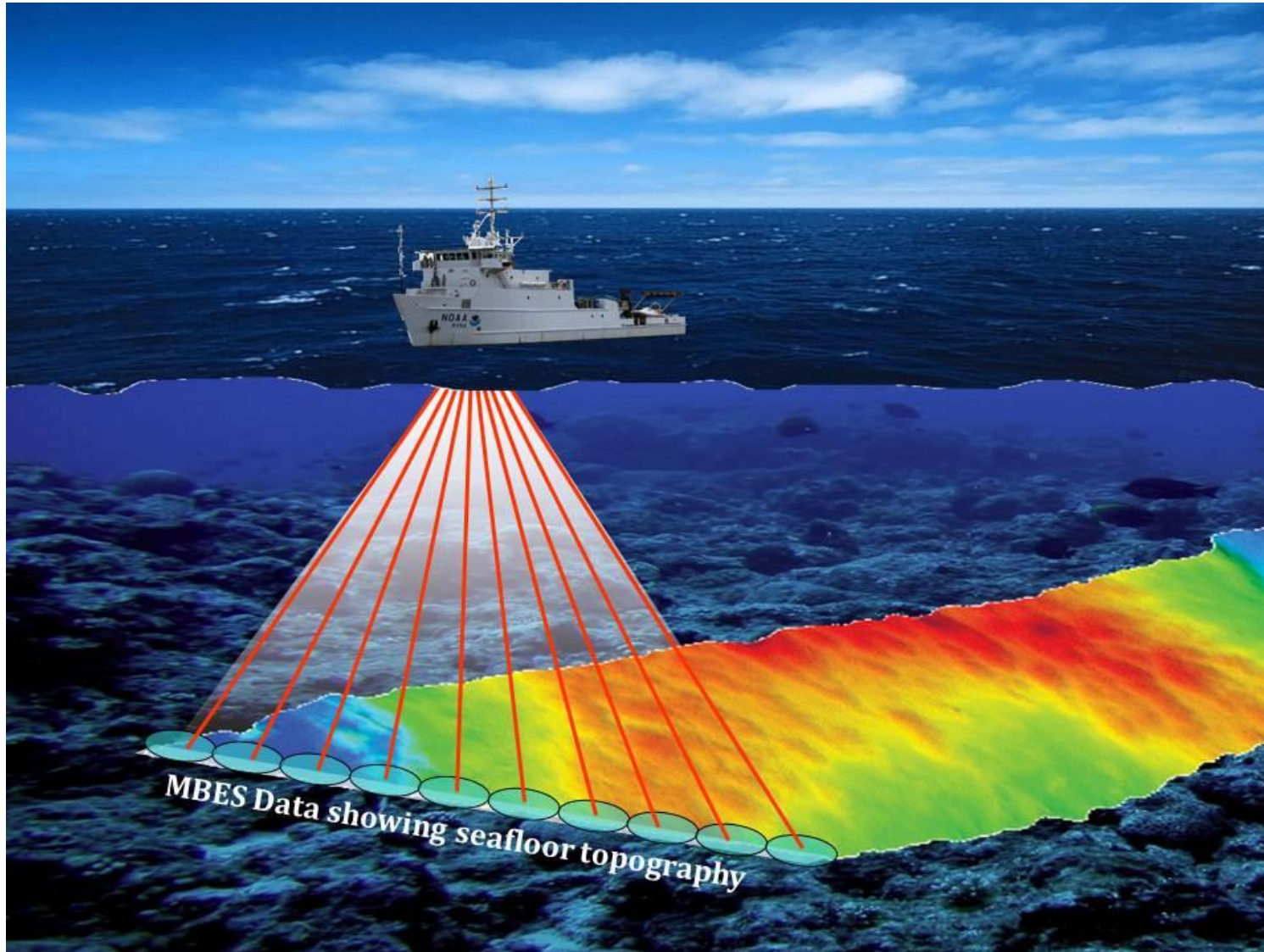


Systeme à faisceaux croisés :

Une émission large transversalement au navire & fine longitudinalement

Plusieurs réceptions larges longitudinalement & fines transversalement

# Principe d'acquisition au sondeur multifaisceaux



→ Émission & Réception de signaux  
= **un ping**

→ Acquisition de N signaux le long  
d'une ligne perpendiculaire au navire  
= **une fauchée**

# Chaîne d'acquisition complète

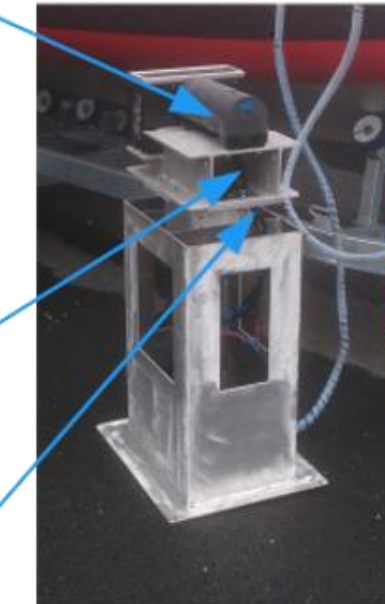
SMF R2SONIC 2022



Centrale d'attitude I2NS d'Applanix



Célérimètre



Regroupement sur un support unique

Interfaçage informatique



Récepteurs GPS

Base GPS



Logiciel de pilotage automatique du navire

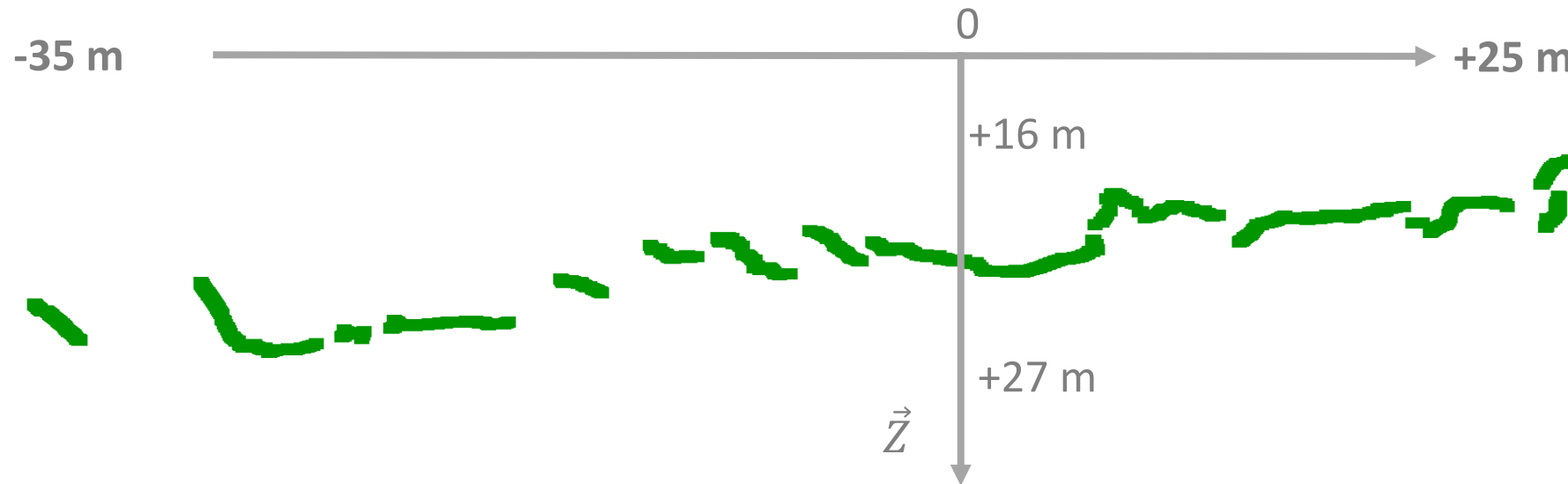
Intégration au navire Seaviews One



# Les trois types de données acquises au SMF

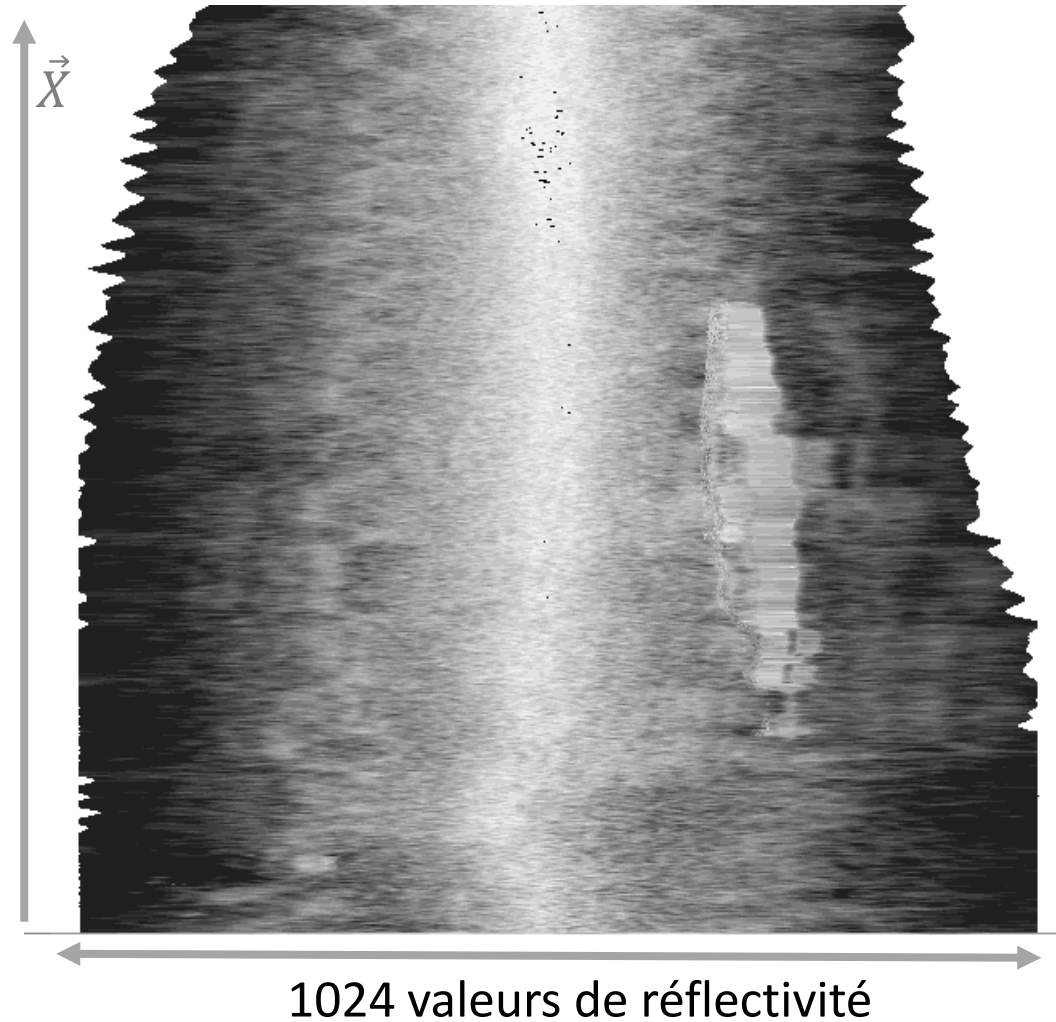
## ① Sondes bathymétriques

→ Temps aller-retour mis par un signal à rencontrer un obstacle



Un exemple des 1024 sondes bathymétriques d'un ping sur un fond rocheux

# Les trois types de données acquises au SMF

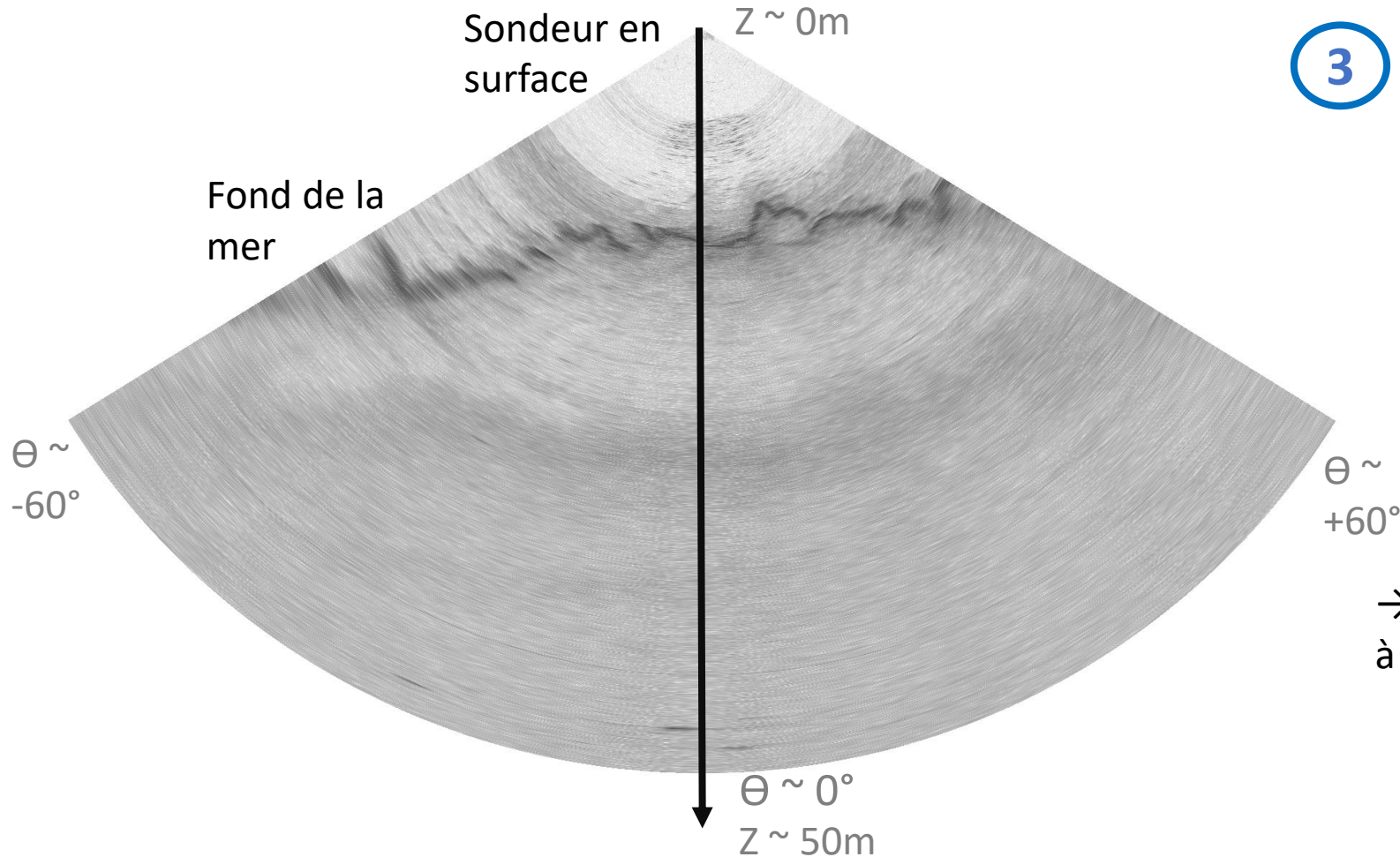


## ② Imagerie de réflectivité

- Modification de l'énergie du signal lors de la rencontre avec les fonds marins
- Signature de la nature des fonds marins
- L'accumulation de pings successifs forme une image

Imagerie de réflectivité (ici, une épave de sous-marin)

# Les trois types de données acquises au SMF



## 3 Imagerie de la colonne d'eau

→ les N signaux d'un ping en amplitude

→ Image composée de 256 signaux colonnes à échantillonnage spatial de l'ordre de 1cm

Imagerie de la colonne d'eau (ici, un banc de poissons)

# Plan

Contexte des travaux

1 – Données issues du sondeur multifaisceaux

**2 – Estimation de valeurs bathymétriques  
par réseau de neurones convolutifs**

3 – Classification des données pour  
l'identification des habitats marins

Conclusion générale

**De la donnée bathymétrique à la carte**

**Estimer les sondes par des approches « vision »**

**Utilisation de réseau de neurones convolutifs (CNN)**

**Composition d'un CNN**

**Préparation des données et labellisation**

**Architecture et paramétrage du CNN**

**Phase d'apprentissage/validation**

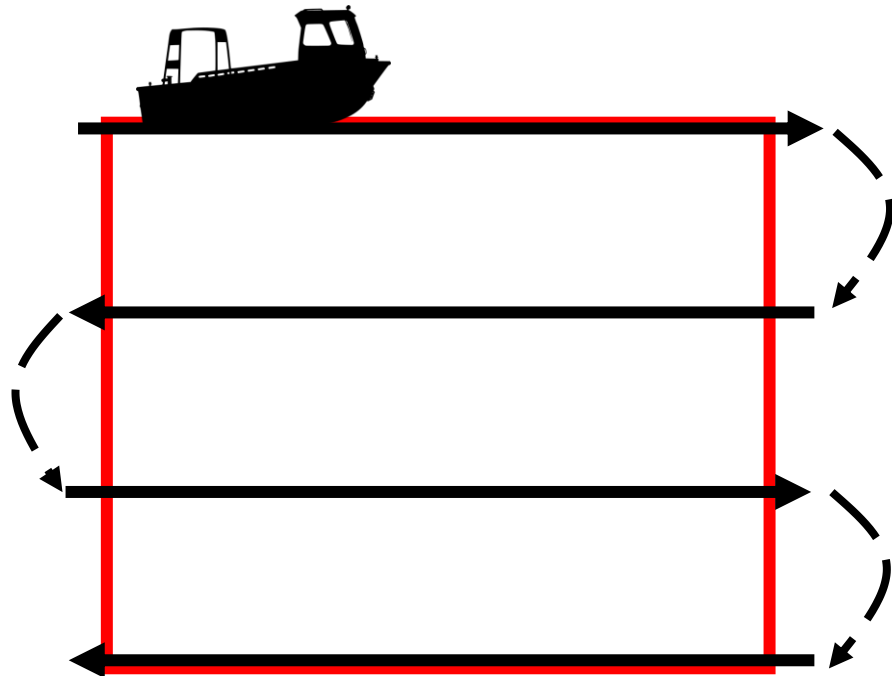
**Phase d'application à de nouvelles données**

**Conclusions & Perspectives**



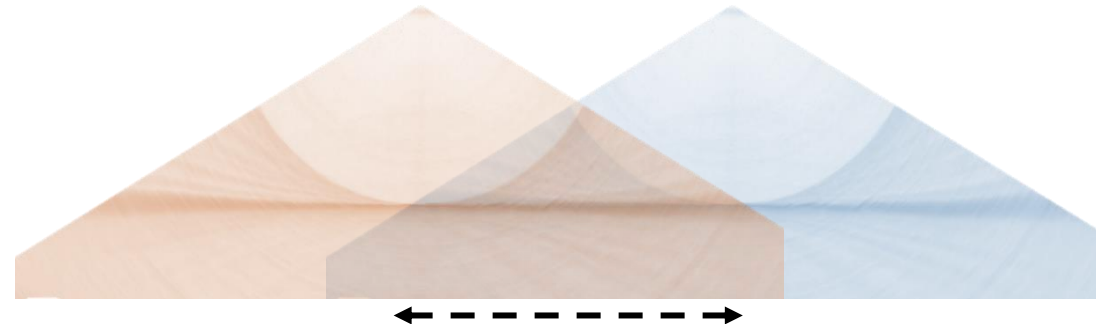
# De la donnée bathymétrique à la carte

## ① Acquisitions sur le terrain



Navigation sur la zone à lever

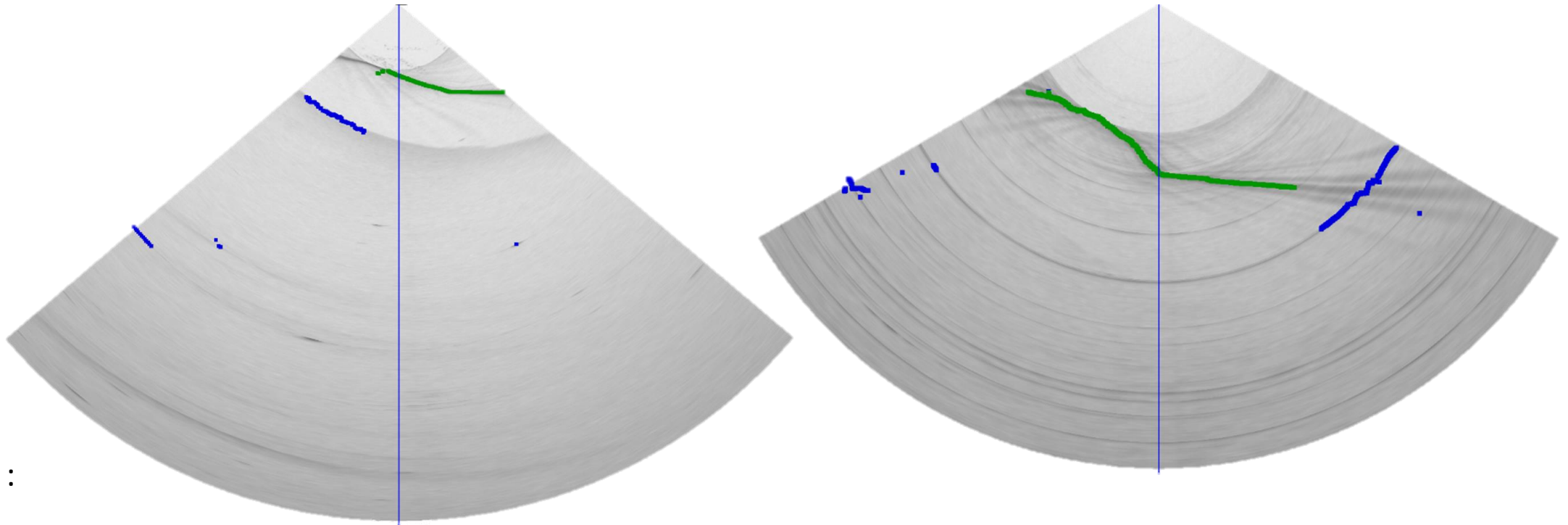
Espacement entre les transects  
avec 50% de recouvrement



# De la donnée bathymétrique à la carte

## 2 Passer en revue toutes les données

Neutraliser les sondes erronées en bleu



Légende :

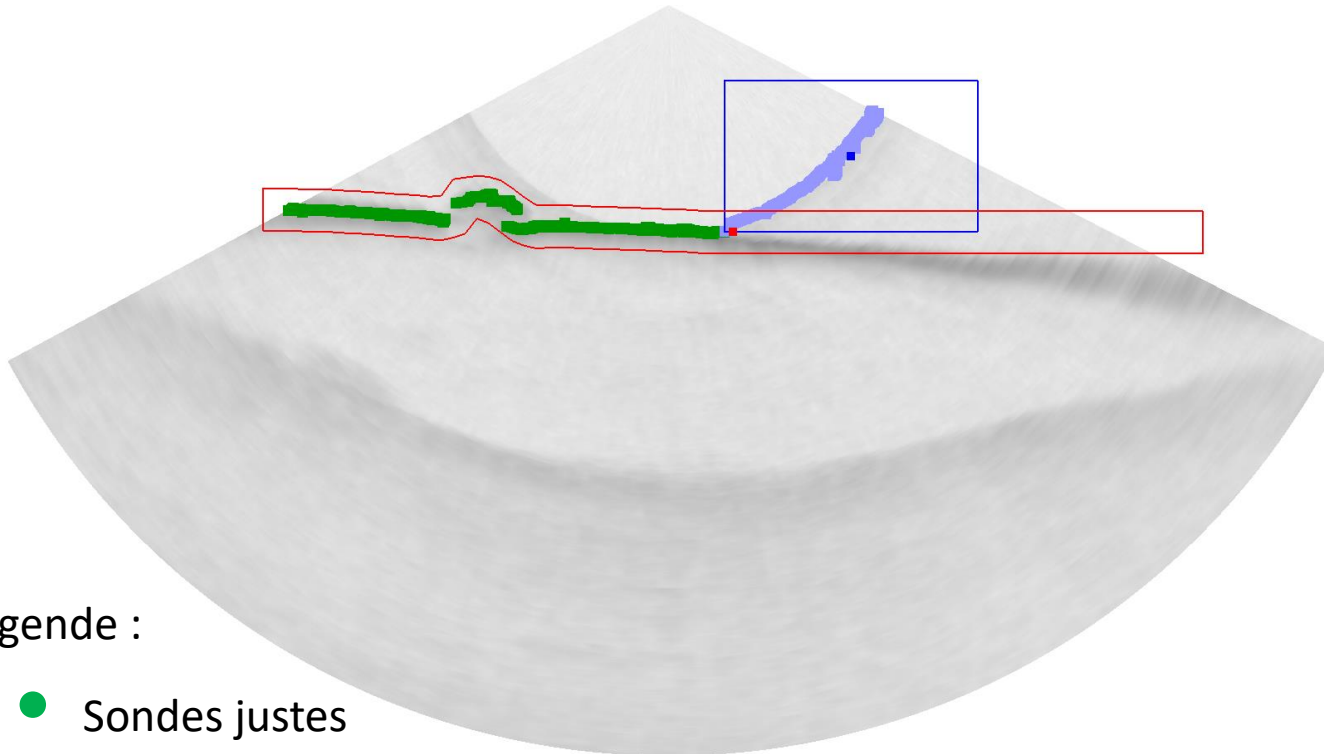
- Sondes justes
- Sondes erronées

Exemples d'erreurs de détection du fond commises par le sondeur

# De la donnée bathymétrique à la carte

## 2 Passer en revue toutes les données

Masquage des sondes erronées par méthode manuelle ou semi-manuelle :



Légende :

- Sondes justes
- Sondes erronées

Très long : 1/1

Temps de traitement / acquisition

Décimation du jeu de données

Risque de « trou » dans la carte

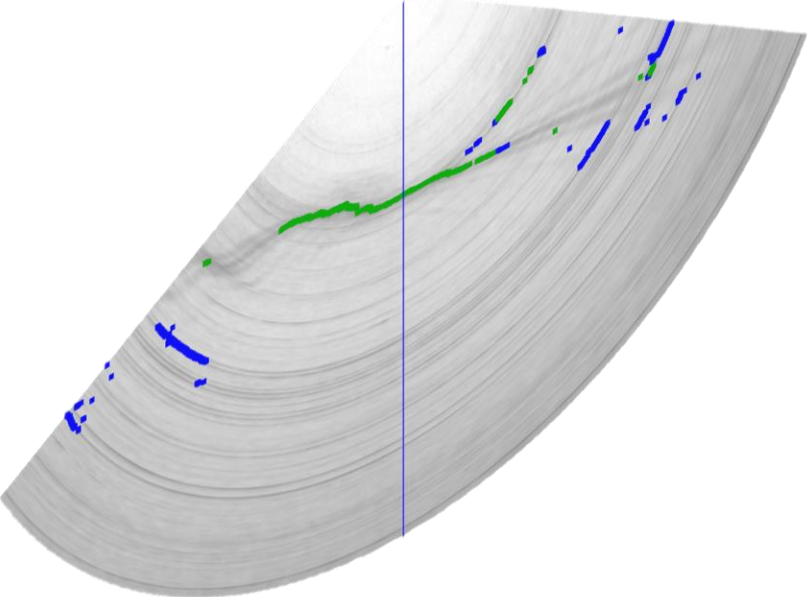
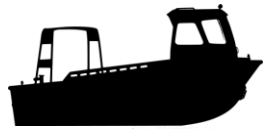
« Frustrant »

Les images colonne d'eau sont « bonnes »

# De la donnée bathymétrique à la carte

## 3 Réaliser un produit cartographique

Pings avec sondes filtrées

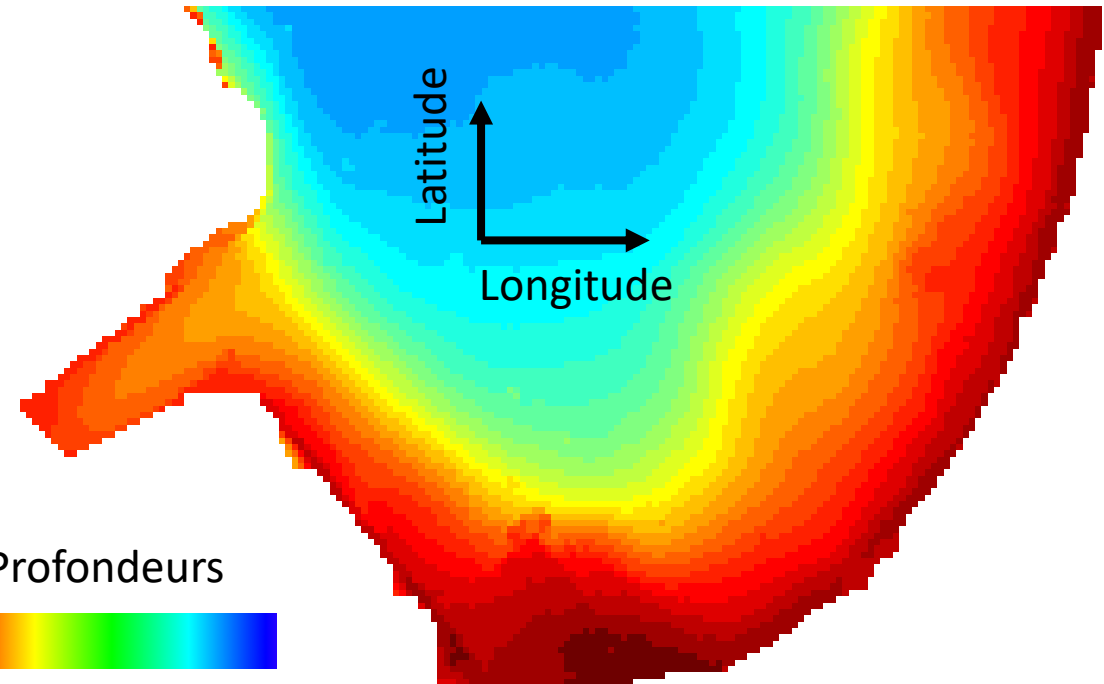


Changement de coordonnées



Projection dans un maillage

Modèle Numérique de Terrain (MNT) à mailles carrées,  
couleurs artificielles de bathymétrie



faibles

grandes

# Estimer les sondes par des approches « vision »

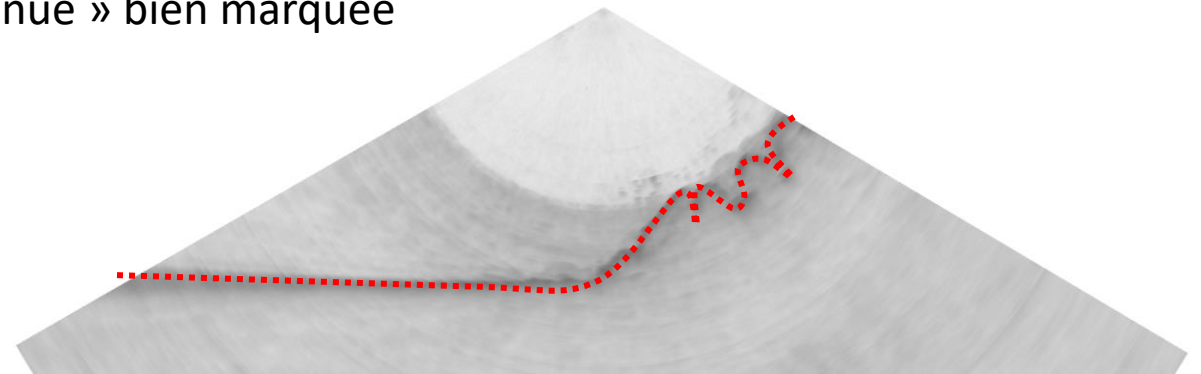
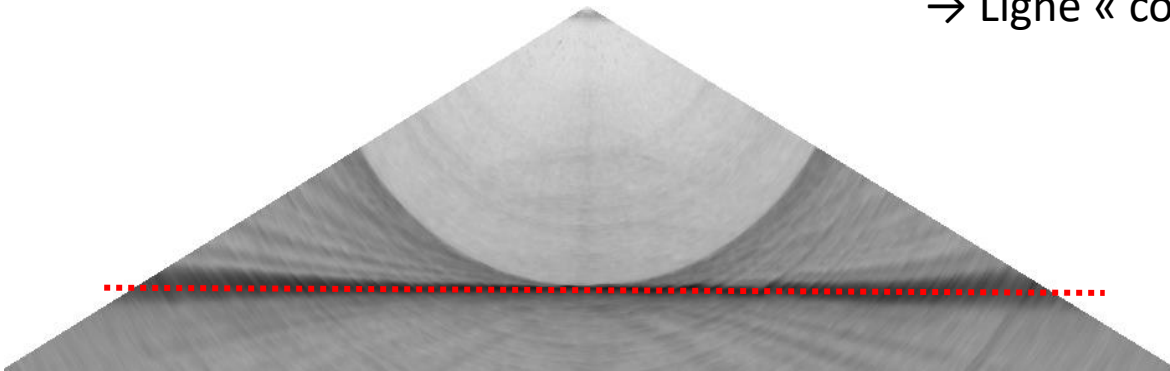
**L'imagerie de la colonne d'eau peut-elle permettre de déduire les sondes bathymétriques ?**

Avec les objectifs :

- Être exempt d'erreurs
- Utiliser toute l'ouverture de la fauchée
- Réduire les temps d'acquisition
- Être plus rapide que la revue systématique des données

Car à l'œil :

- Le fond est la caractéristique la plus fréquente
- Ligne « continue » bien marquée



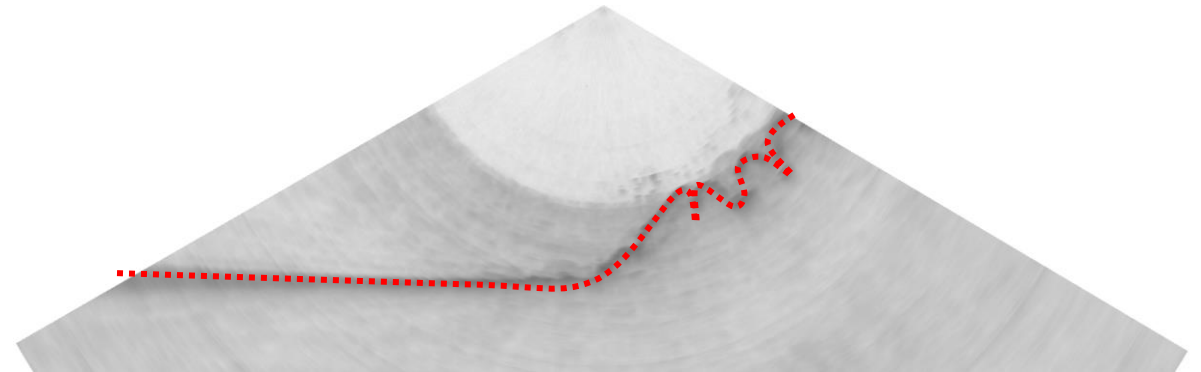
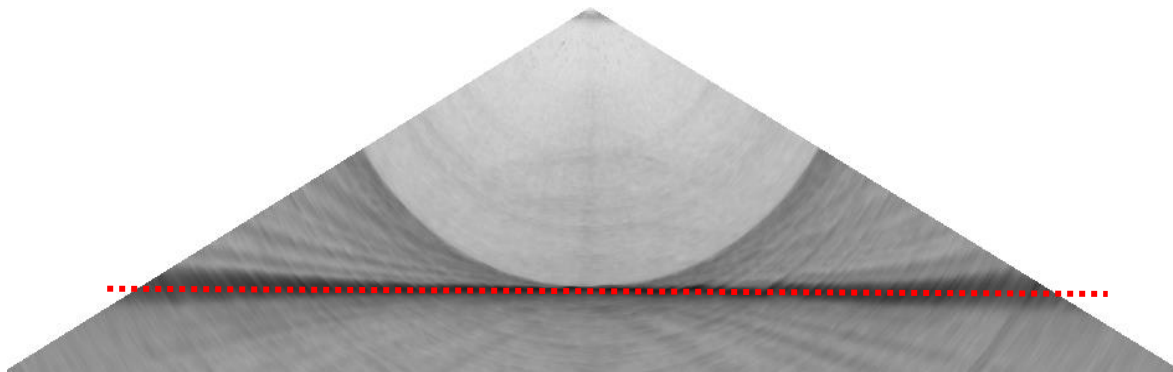
# Estimer les sondes par des approches « vision »

## Proposition :

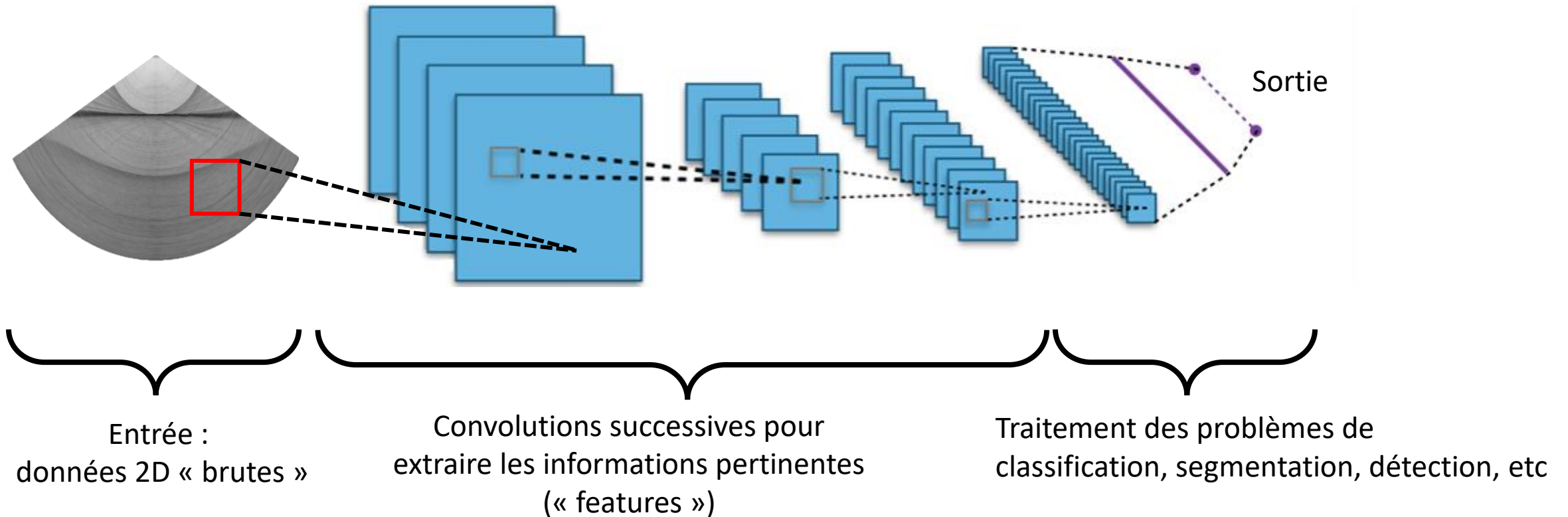
Méthode par apprentissage supervisé à base de **réseaux de neurones convolutifs (CNN)**

Problématique originale de par :

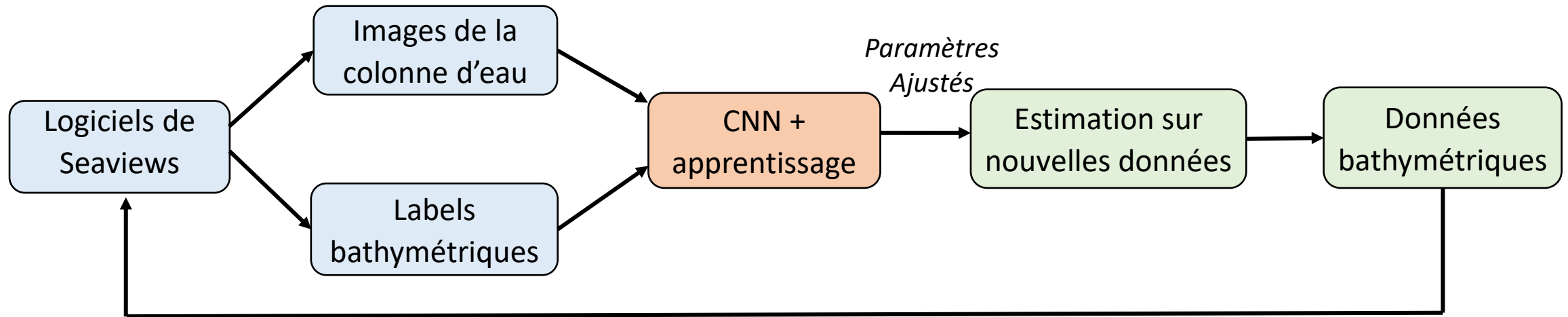
- L'exploitation des images de la colonne d'eau
- L'application de CNN sur des données SMF
- La rareté des problèmes de type régression multi-sorties



# Composition d'un réseau de neurones convolutifs



# Mise en place d'un apprentissage



Étapes à réaliser :

- Phase 1 : Construire un jeu d'apprentissage
- Phase 2 : Labelliser une partie des données
- Phase 3 : Choix d'une architecture CNN
- Phase 4 : Apprentissage
- Phase 5 : Application sur de nouvelles données

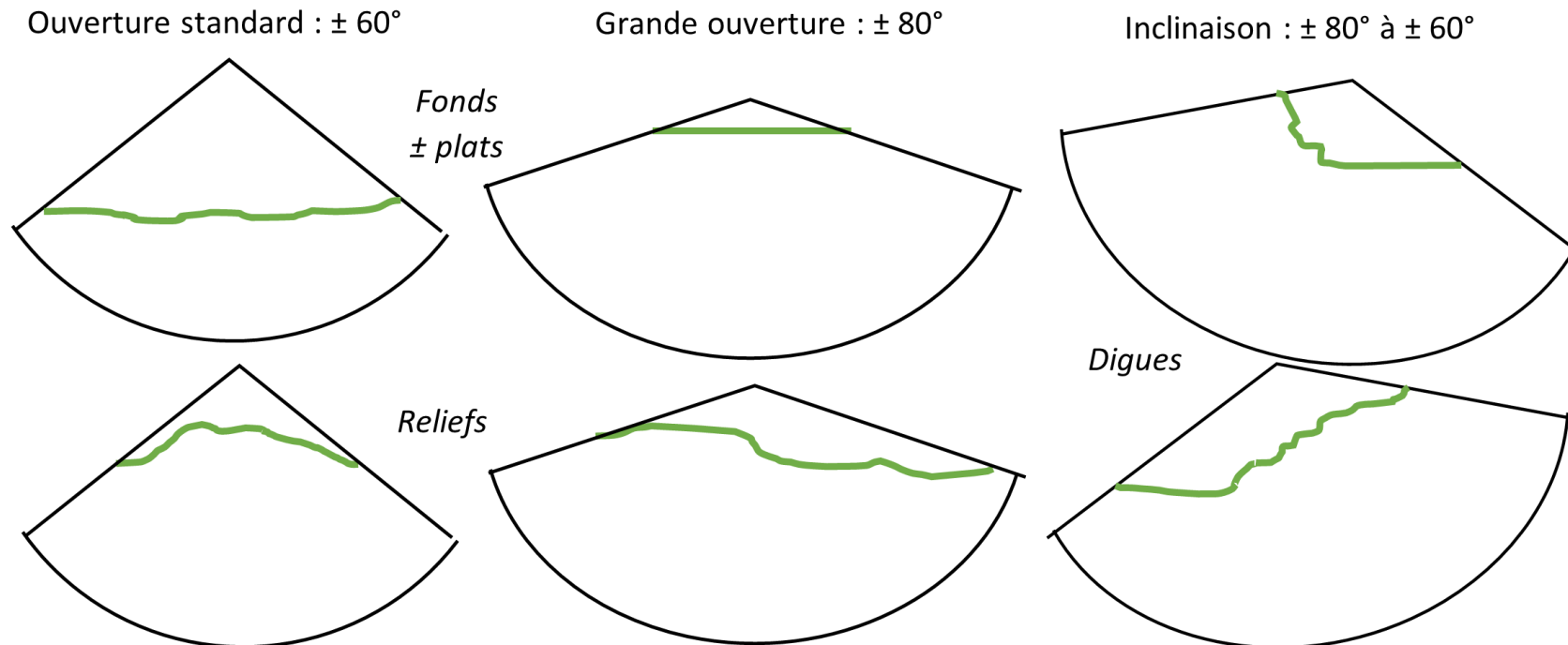


# Préparation des données et labellisation

## Phase 1 : Construire un jeu d'apprentissage pertinent

→ Sélectionner des données avec un maximum de variabilité

→ Re-dimensionner toutes les données à un format carré de 256 colonnes à 256 valeurs



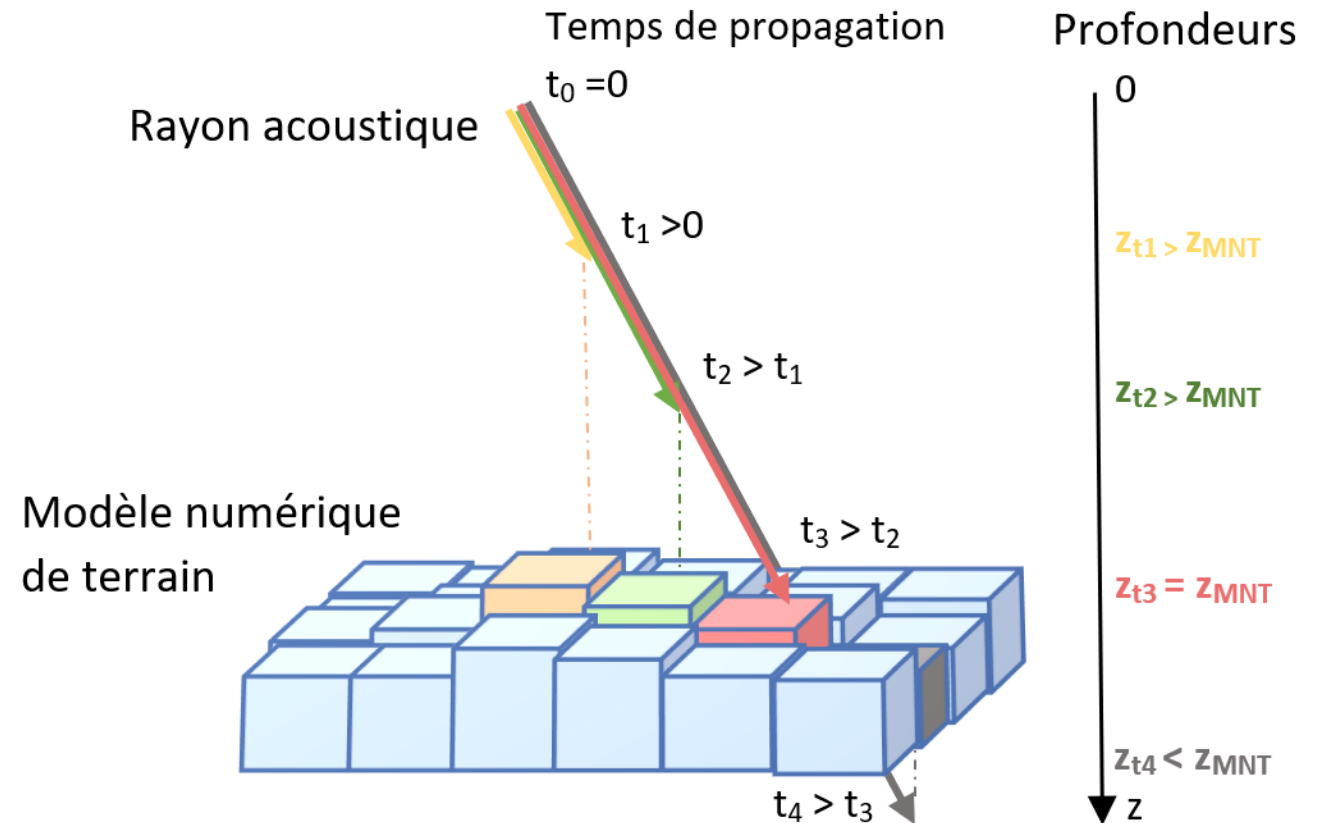
# Préparation des données et labellisation

## Phase 2 : Labelliser automatiquement et sans erreur les images de la colonne d'eau à partir d'un modèle numérique de terrain (MNT)

→ Construire un MNT bathymétrique concernant les données à labelliser

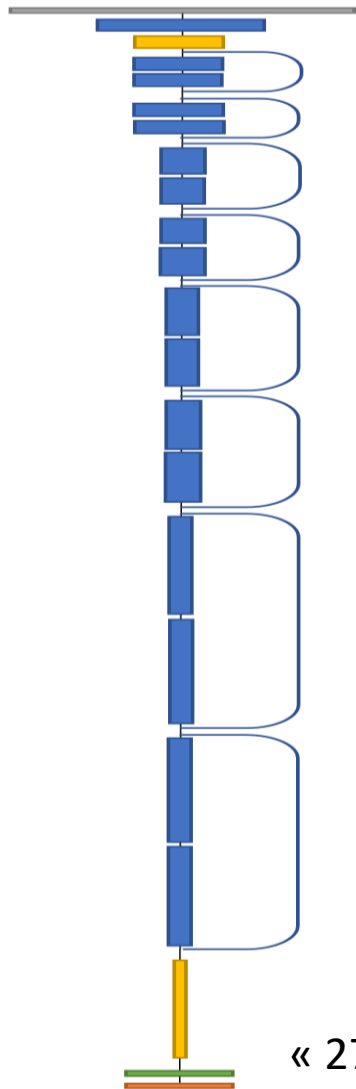
→ Calculer la trajectoire du rayon acoustique de chaque voie de chaque ping à labelliser

→ Chercher l'intersection entre la trajectoire du rayon et le MNT

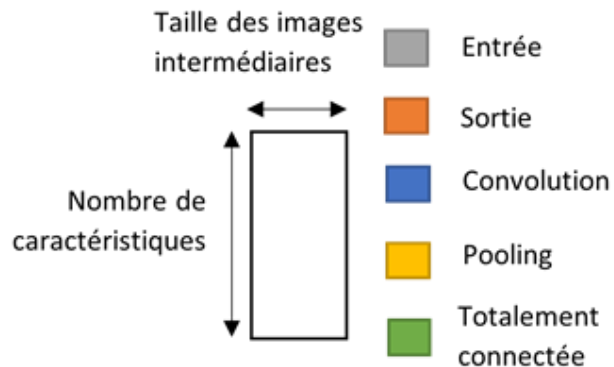


# Architecture et paramétrage du CNN

ResNet (2015) \*



Entrée : 256 x 256



Sortie : 256 valeurs numériques  
« 27, 26.8, ..., 19.1, 19.15, ..., 25.3, 26 »

## Phase 3 : Choix d'une architecture CNN

→ Architecture ResNet 50 :

- 50 convolutions successives
- 25,6 millions de paramètres à ajuster
- Connexions de raccourcis pour accélérer l'apprentissage

→ Apprentissage depuis zéro :

- Pas de transfert d'apprentissage possible
- Initialisation des paramètres avec une loi uniforme qui facilite leur ajustement

\* He, K., Zhang, X., Ren, S. and Sun, J. (2015). *Deep residual learning for image recognition.*

# Phase d'apprentissage / validation

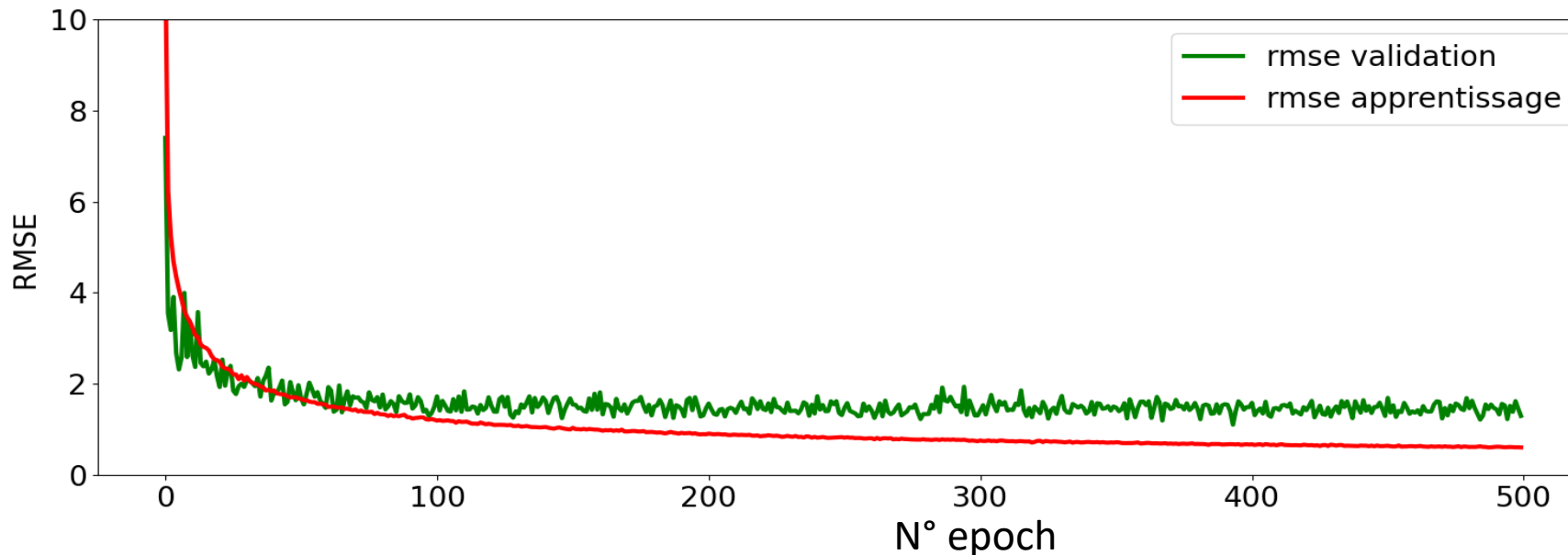
Quantité de données : 56 000 images de la colonne d'eau labellisées

Durée de l'entraînement : 1 mois, 500 passages complet du jeu de données

Suivi de l'apprentissage:

$$RMSE = \sqrt{\sum_{n=1}^{256} (Sorties_n - Labels_n)^2}$$

Division des données en :  
80 % pour l'apprentissage  
20 % pour la validation



RMSE : Racine de l'erreur quadratique moyenne

# Phase d'application à de nouvelles données

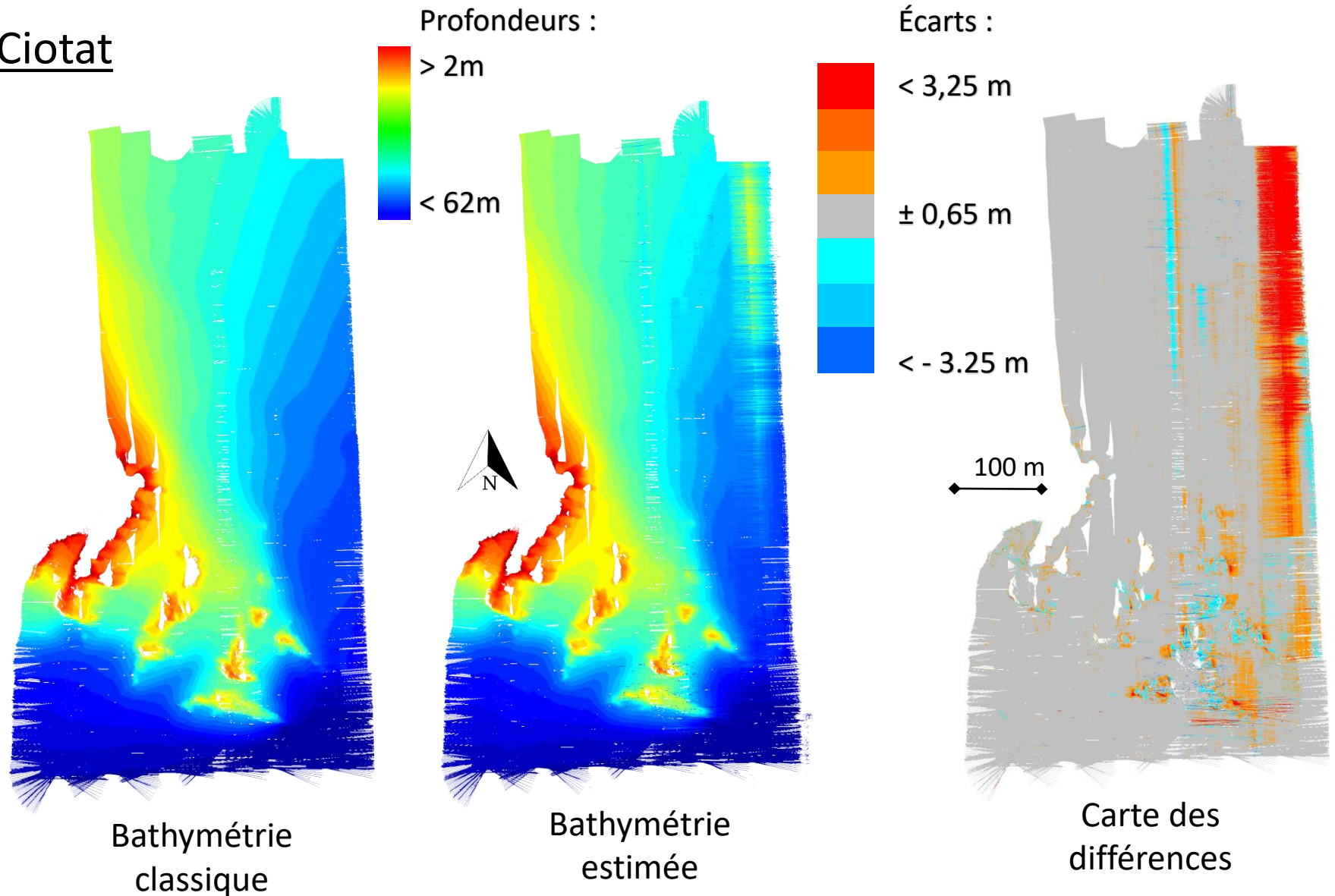
## Zone « Les Pierres », La Ciotat

→ 12 transects Nord-Sud  
sans recouvrement

→ Fonds plats en pente  
avec présence de reliefs rocheux

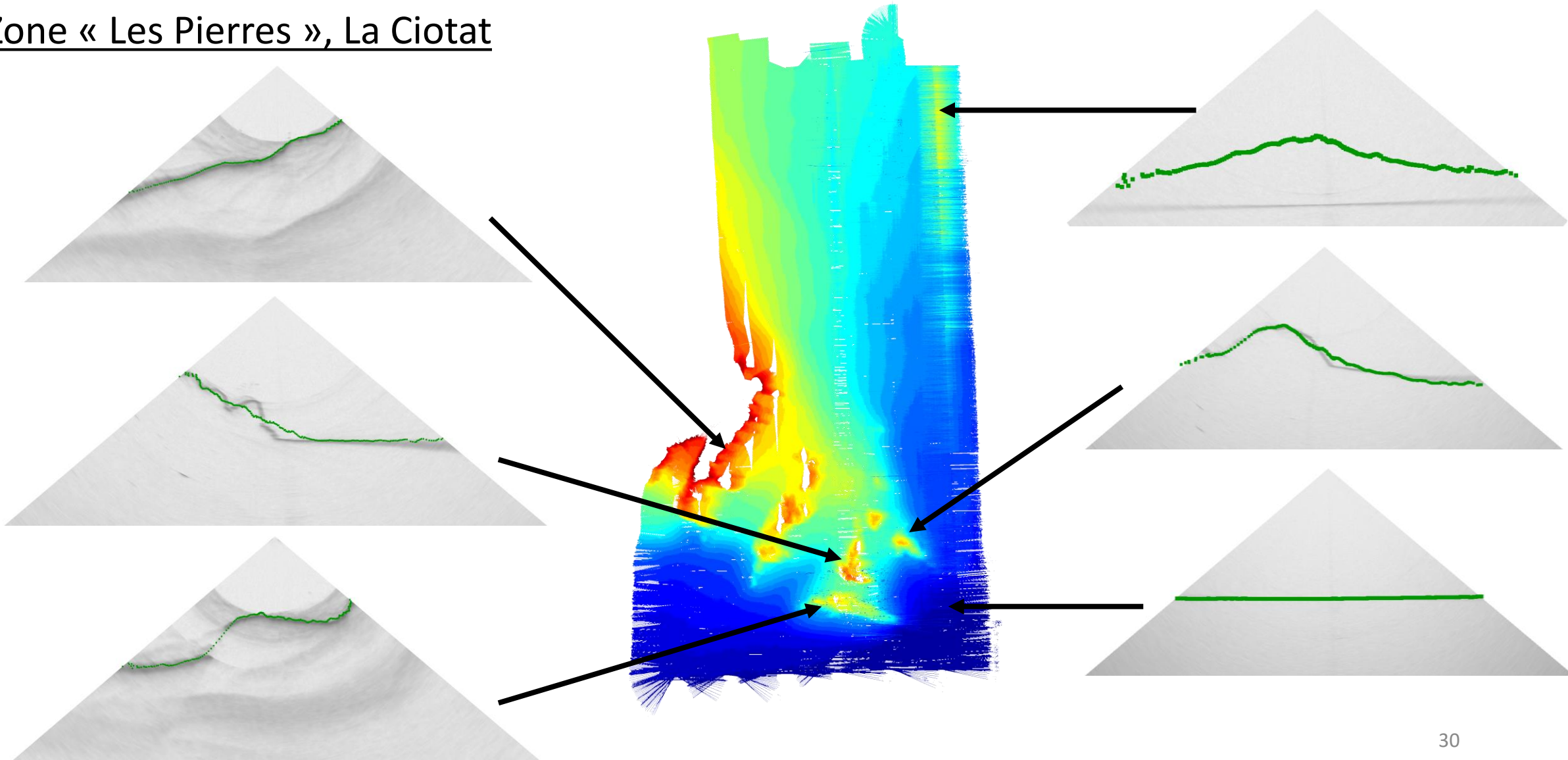
→ Taille originale des images :  
256 x 15 000  
après rééchantillonnage :  
1 ech / 64 cm

→ Bons résultats visuellement



# Phase d'application à de nouvelles données

Zone « Les Pierres », La Ciotat



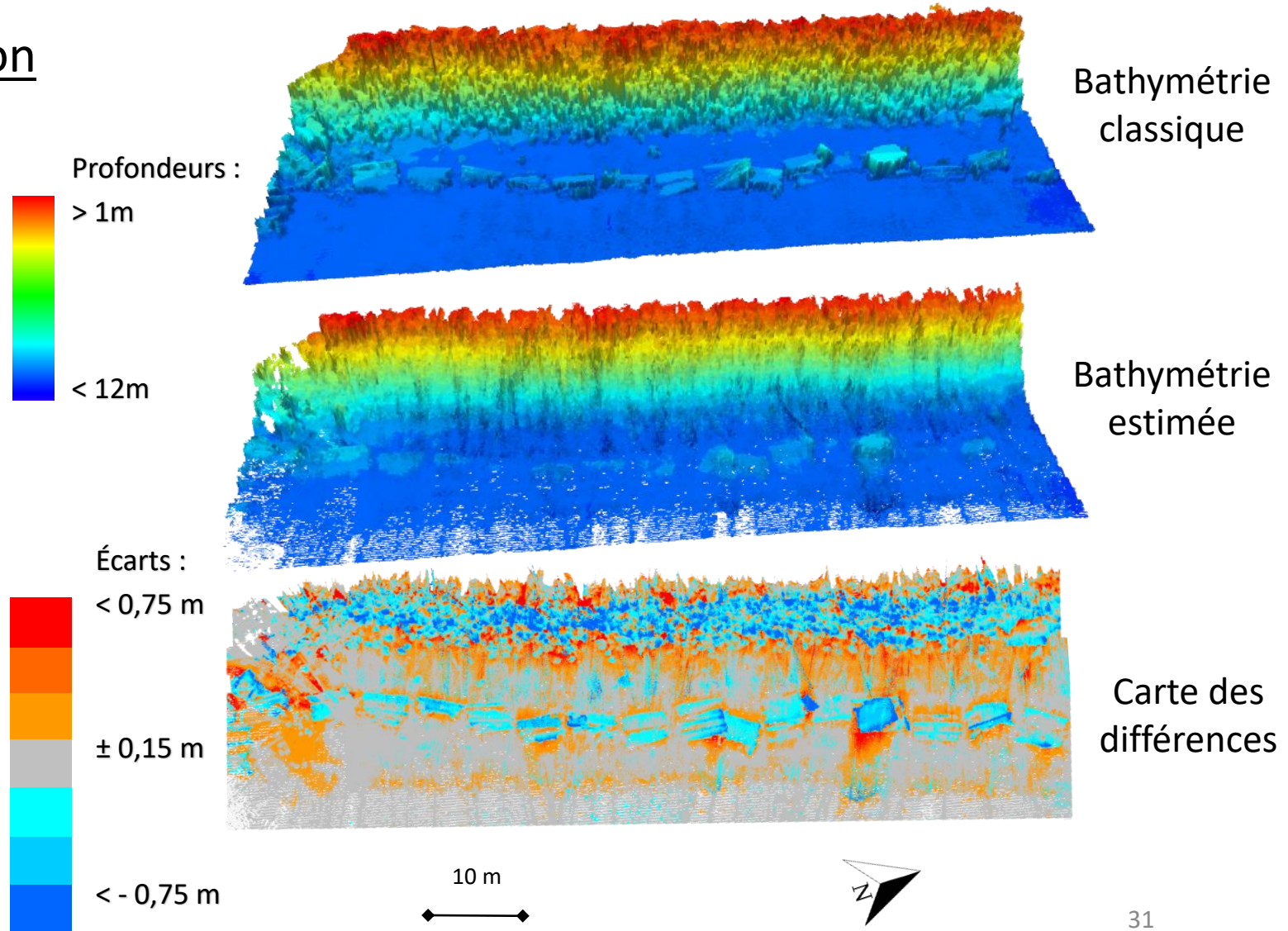
# Phase d'application à de nouvelles données

## Extrait de la grand digue de Toulon

→ 4 transects le long de la digue avec recouvrement

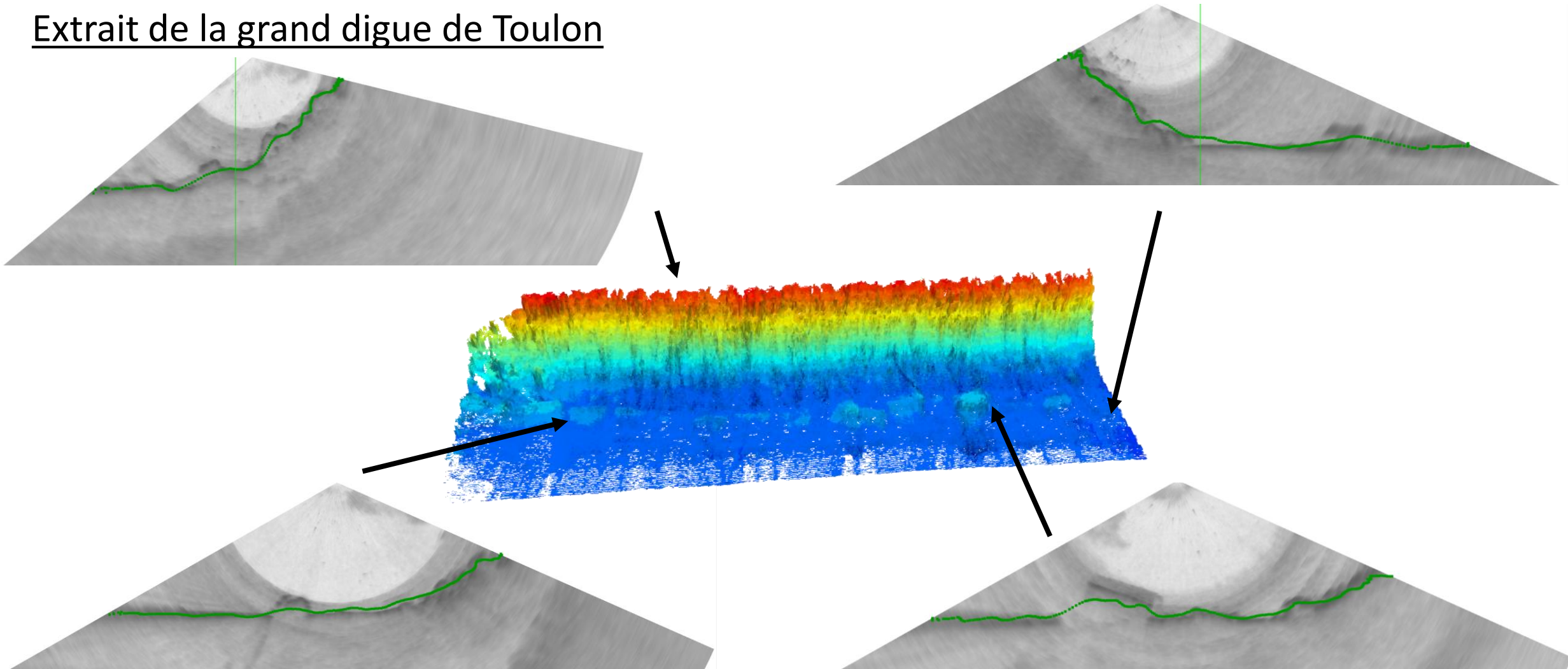
→ Digue rocheuse et blocs épars

→ Taille originale des images : 256 x 1550  
après rééchantillonnage : 1 ech / 13 cm



# Phase d'application à de nouvelles données

Extrait de la grand digue de Toulon





# Conclusions

→ Traitement inédit :

- Coupler une méthode récente + données peu utilisées + application complexe
- Elaborer toute la chaîne de calculs, dont la labellisation automatique
- Réaliser un apprentissage supervisé depuis zéro

→ Résultats bons mais pas totalement satisfaisants :

+ Valide les objectifs :

- + Utilisation de la pleine ouverture angulaire de l'image
- + Réduction du temps de traitement
- + Réduction du taux d'erreurs

- Contrepartie :

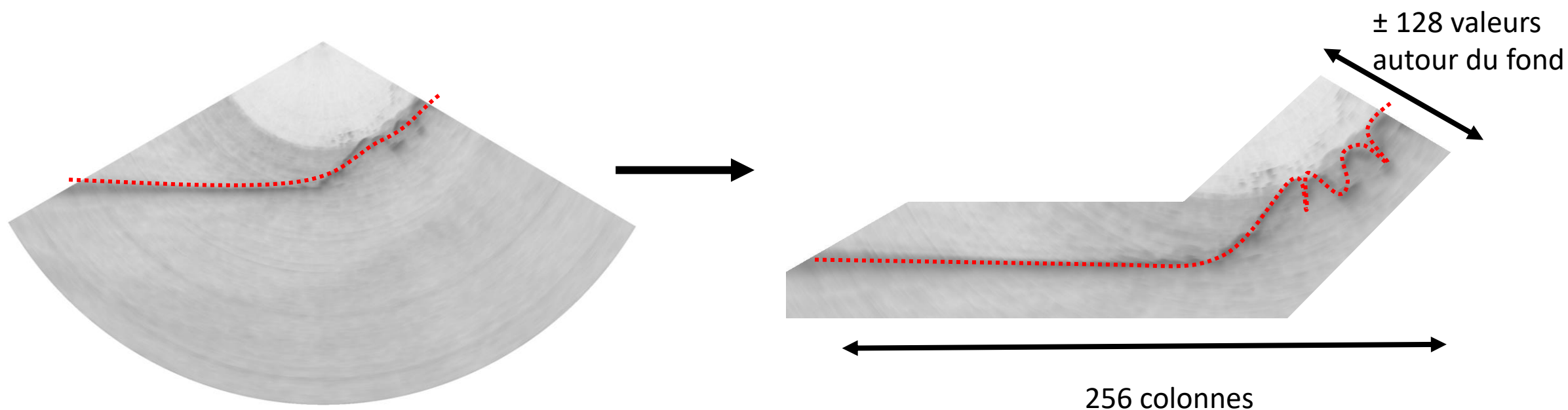
- Temps d'entraînement long
- Sous-échantillonnage de la donnée
- Dégradation de la précision dans les résultats

# Perspectives

## ① Améliorer la précision : travailler sur des images non ré-échantillonnées

→ « Progressive resizing » : 256 x 256 / 256 x 1000 / 256 x 2000 / ..

→ Zoom : 2<sup>e</sup> étape sur une portion de l'image complète



# Perspectives

## ② Aller plus loin

→ CNN à données d'entrée 3D :

Séries d'images successives pour la continuité du fond le long de la fauchée et de la route du navire

→ Super résolution :

Estimer plus de sondes avec le même nombre de signaux

→ Traitement embarqué / temps réel :

Éviter l'acquisition et le stockage des images de la colonne d'eau

# Plan

Contexte des travaux

1 – Données issues du sondeur multifaisceaux

2 – Estimation de valeurs bathymétriques  
par réseau de neurones convolutifs

**3 – Classification des données pour  
l'identification des habitats marins**

Conclusion générale

**Réaliser une carte des habitats marins**

**Inventaire des classes à obtenir**

**Inventaire des données disponibles**

**Classification par la méthode de Seaviews**

**Évaluation du résultat**

**Classification automatique par apprentissage supervisé**

**Constituer un jeu d'apprentissage**

**Méthode des forêts aléatoires**

**Méthode des Séparateurs à Vastes Marges**

**Résultats par méthode**

**Conclusions**

# Réaliser une carte des habitats marins

**1**

## **Inventaire des classes à représenter**

→ Choisir un système de classification (local, national, international), le degré de précision

**2**

## **Inventaire des données disponibles**

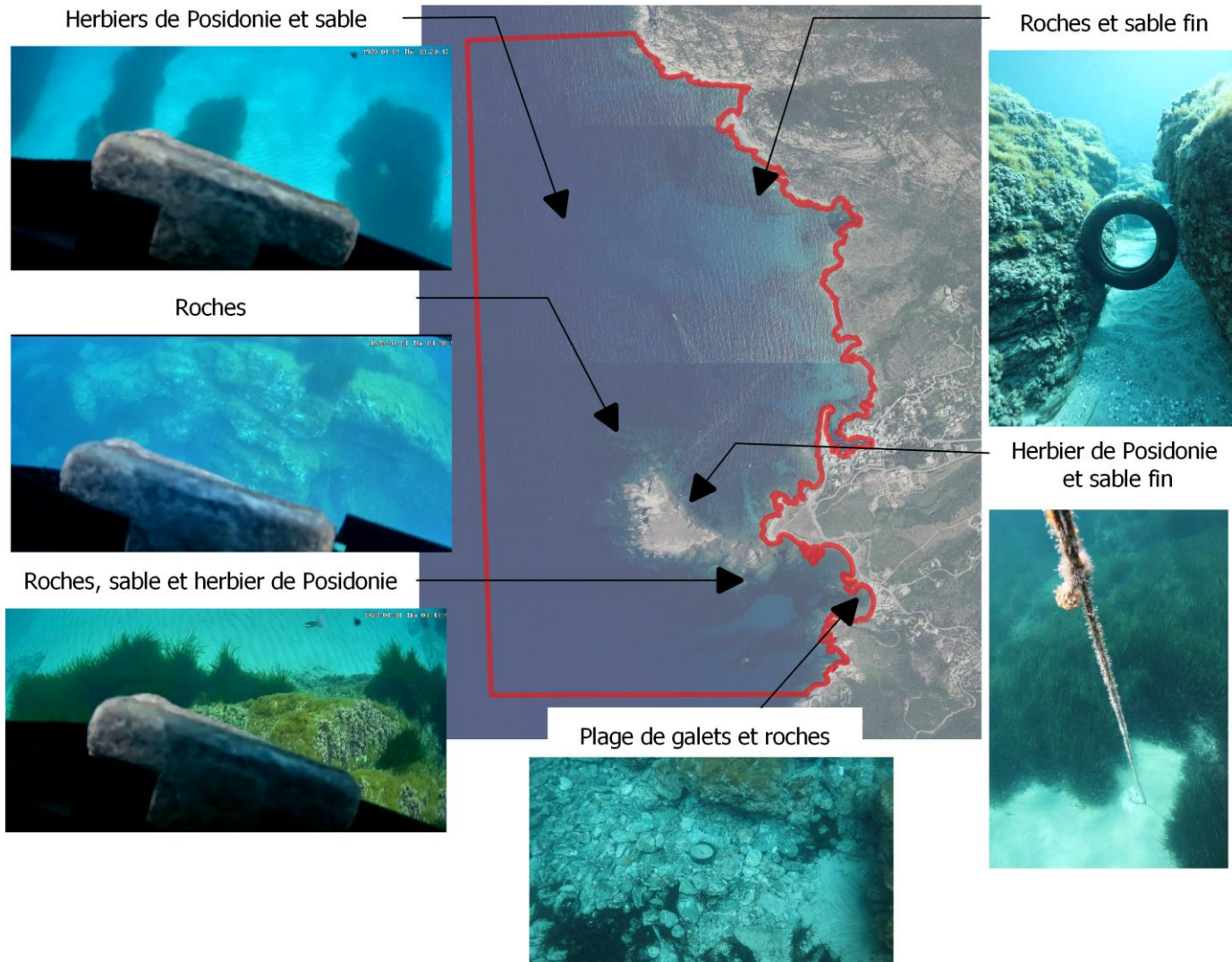
→ Données acquises, données du domaine public, données bibliographiques, etc.

**3**

## **Appliquer une méthodologie de classification**

→ Manuelle, assistée par un ordinateur, semi ou complètement automatisée

# Inventaire des classes à obtenir



→ **Système Européen de classification des habitats (EUNIS\*) :**

- A3.33 Roche à algues photophiles
- A5.46 Sédiments du détritique côtier
- A5.235 Sable fin infralittoral
- A5.535 Herbiers de Posidonie
- A2.111 Plage de galets
- J2.53 Digue ou enrochement artificiel

*\*European Nature Information System*

# Inventaire des données disponibles

Image aérienne\*



Bathymétrie SMF et Lidar

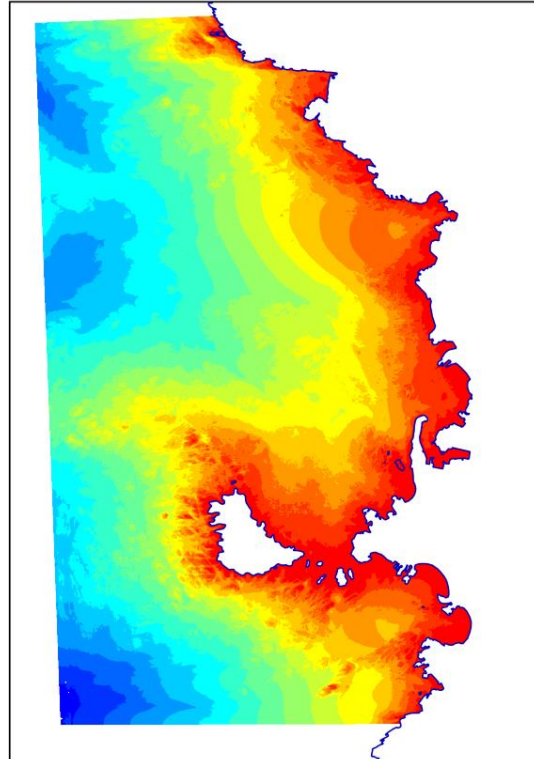
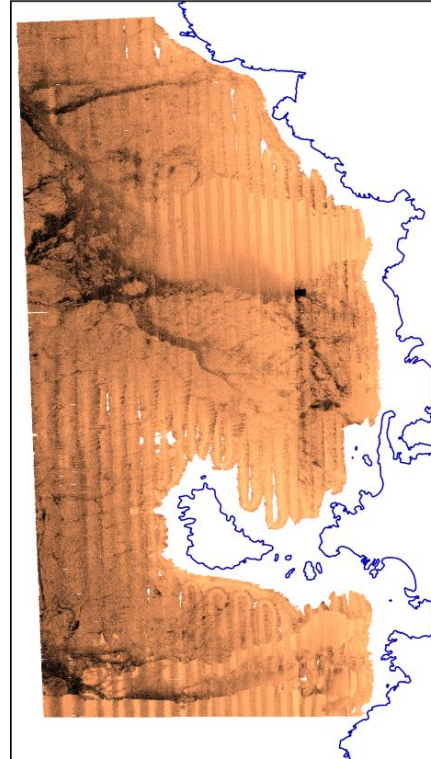
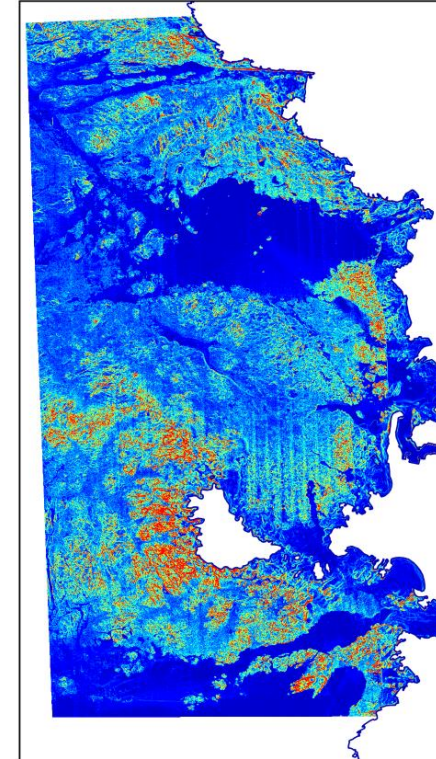


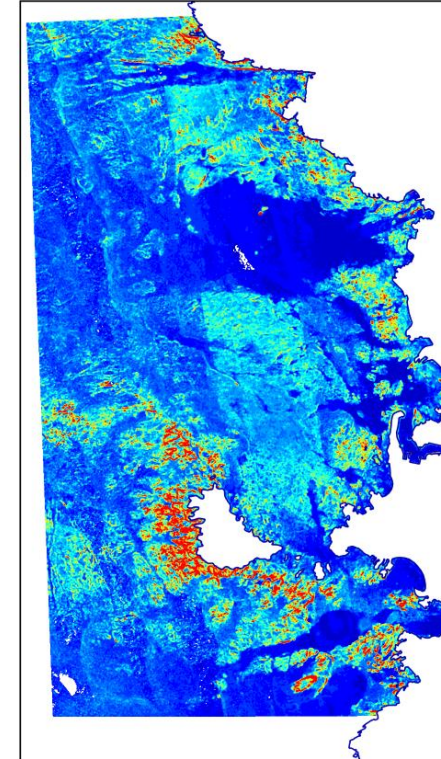
Image de réflectivité



Pente des fonds



Rugosité des fonds



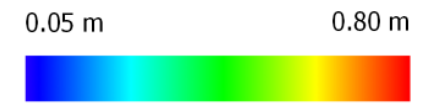
0 500 1000 m



Profondeurs



Réflectivité









Pente & Rugosité

\* Image aérienne BD ORTHO® 2019 de l'IGN

# Classification par la méthode de Seaviews

## Attribution des classes

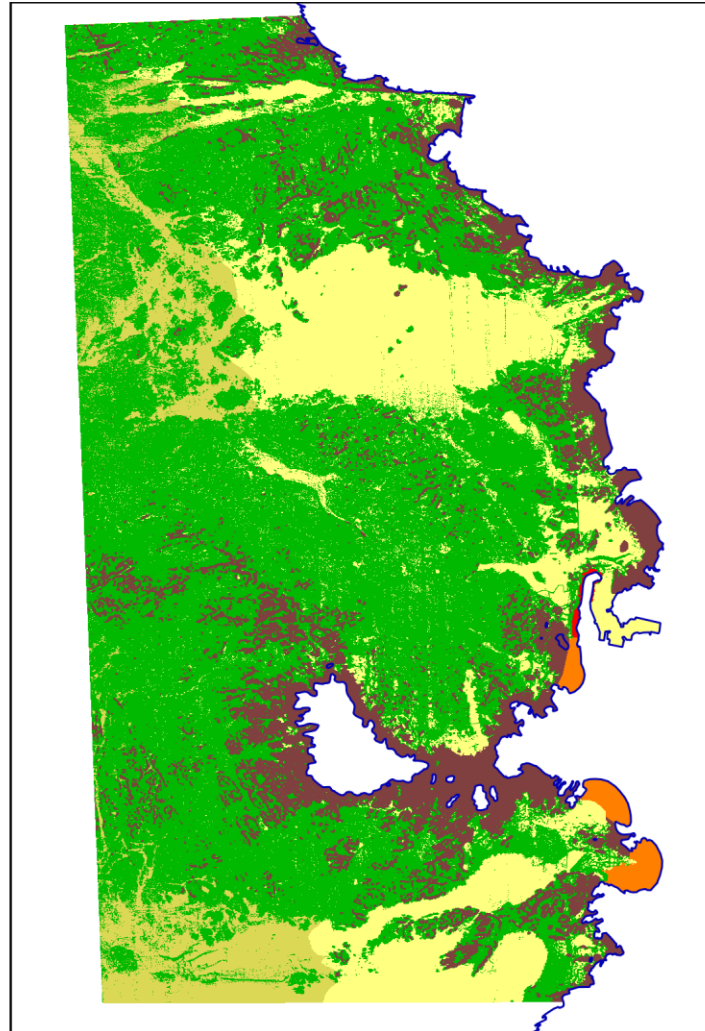
Description et ordre de priorité entre les classes :

-  **Digue**  
Définition manuelle d'une zone d'après l'image aérienne
-  **Plage de galets**  
Définition manuelle d'une zone d'après les vérités terrain
-  **Rochers à algues photophiles**  
Rugosité > 0.4 m
-  **Herbiers de Posidonies**  
Bathymétrie > 2.5 m  
Pente > 0.09 m
-  **Sables fins infralittoraux**  
Bathymétrie > 2.5 m & < 25 m
-  **Détritique côtier**  
Bathymétrie < 25 m
-  **Rochers à algues photophiles**  
Bathymétrie > 2.5 m

0 500 1000 m



## Carte de Seaviews



→ Méthode semi-manuelle

→ L'opérateur décrit les classes par des valeurs issues des données

→ Le logiciel construit un « arbre de classification » avec les descriptions

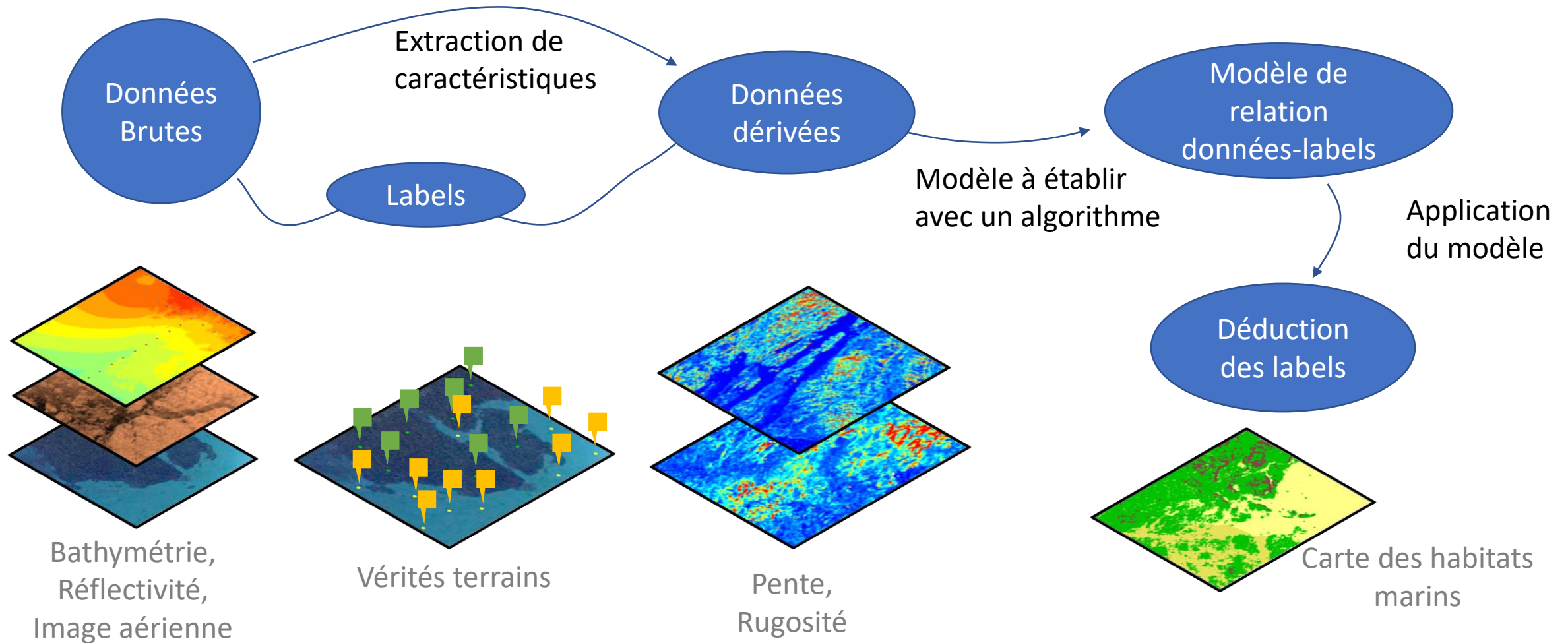
→ A chaque maille est attribuée une classe

→ Temps requis : ½ journée

*Viala C, Lamouret M, Abadie A (2021) Seafloor classification using a multibeam echo sounder: A new rugosity index coupled with a pixel-based process to map Mediterranean marine habitats. Applied Acoustics 179:108067*

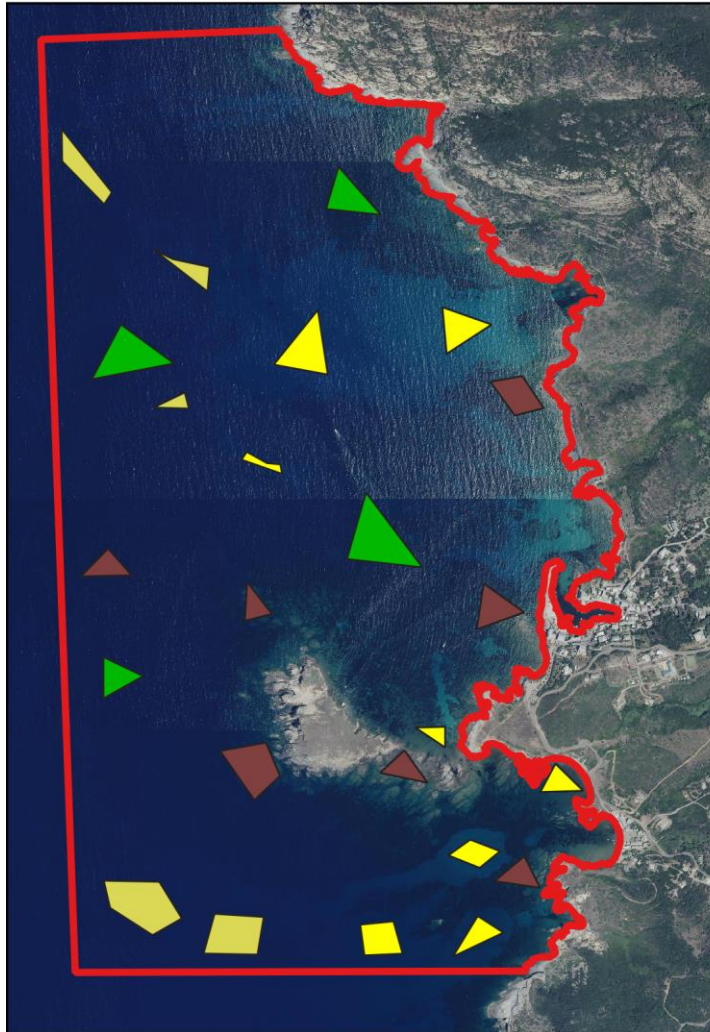


# Classification automatisée par apprentissage supervisé

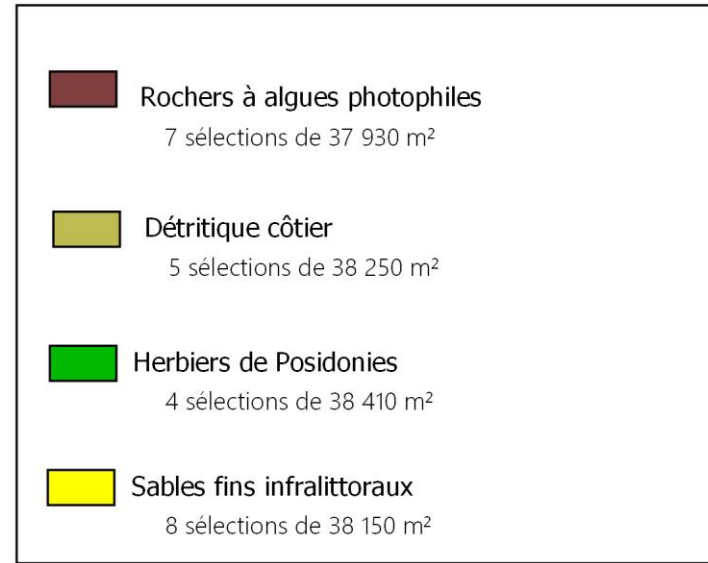


# Constituer un jeu d'apprentissage

Sélections des régions d'intérêt



Répartition des sélections



0 500 1000 m



→ Sélections manuelles de région d'intérêt pour chaque classe

→ Les sélections sont faites d'après :

- Les vérités terrains
- Les différents MNT
- Les connaissances

→ Quantité de sélections équivalentes par habitat

# Choix de méthodes à apprentissage

→ Intégration de 7 méthodes automatiques dans les logiciels :

## Apprentissage supervisé

Forêts aléatoires d'arbres décisionnels

Séparateurs à vastes marges

Classifieur bayésien normal

Réseau de neurones

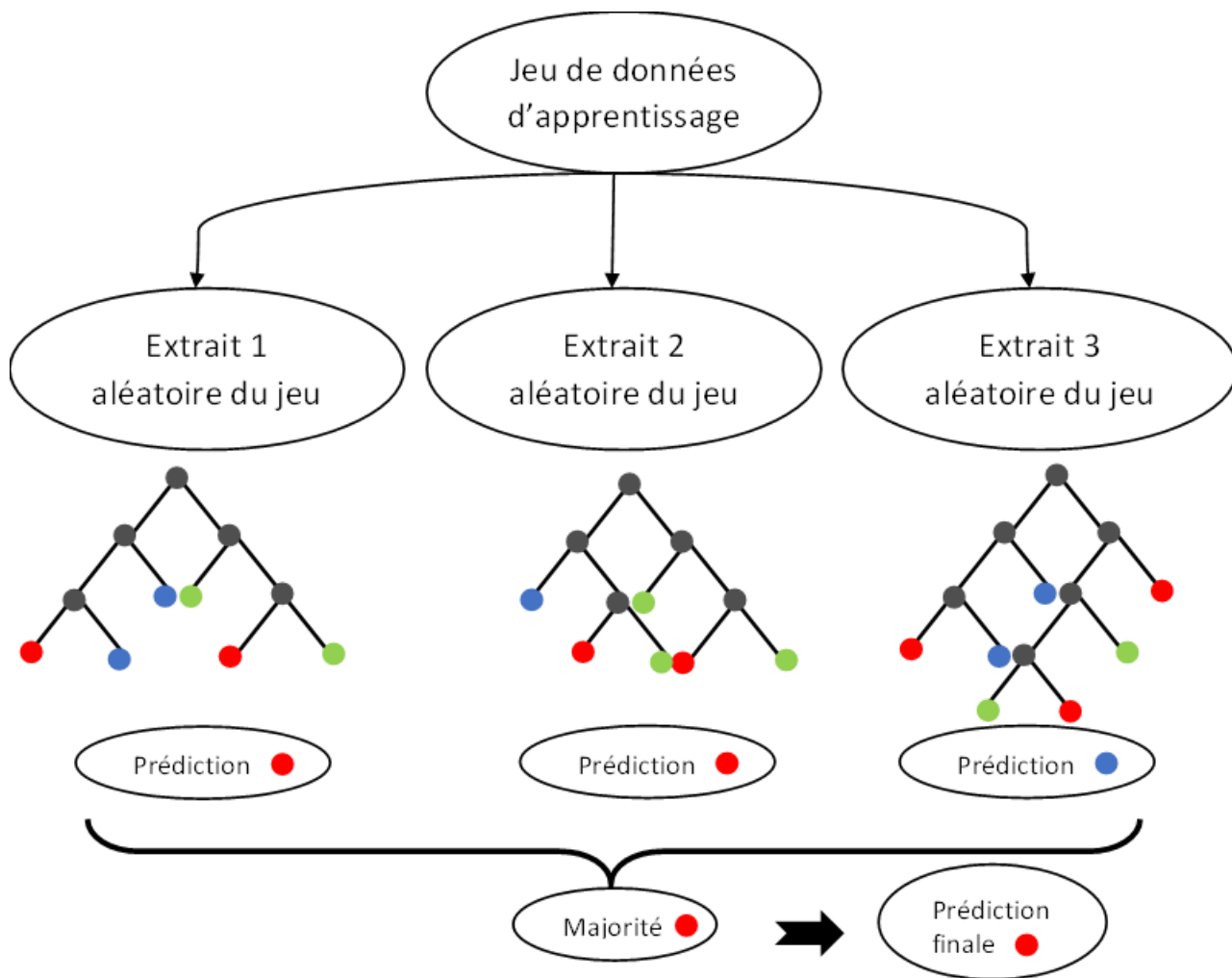
K-plus proches voisins

Spectral Angle Mapper

## Apprentissage non-supervisé

K-moyennes

# Méthodes des Forêts aléatoires d'arbres décisionnels (FAAD)



→ Le modèle crée un grand nombre d'arbres décisionnels à partir de sous jeux d'apprentissage

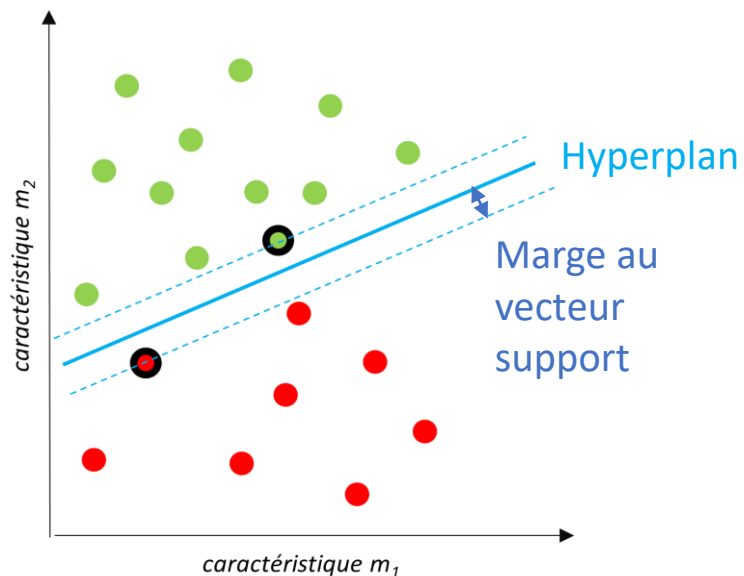
→ Les sous jeux sont obtenus par tirages aléatoires sur les données et les caractéristiques

→ Construction d'un arbre :  
Séries de questions sur les caractéristiques des données pour discriminer les classes

→ La prédiction finale dépend des prédictions de chaque arbre

# Méthode des Séparateurs à Vastes Marges (SVM)

Problème  
linéairement  
séparable

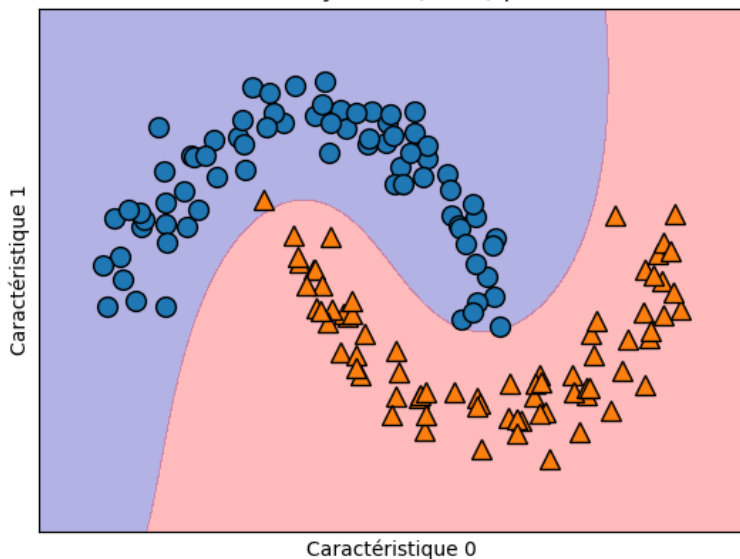


→ Le modèle recherche d'un hyperplan séparateur qui maximise les marges entre les deux classes.

Recherche des poids  $w$  et du biais  $b$  tels que  $\forall d$ :

$$w^T \cdot d + b = 0 \rightarrow d \in \text{hyperplan}$$

Marge souple  
et séparateur  
non linéaire



→ Marge souple : autoriser un taux d'erreur de classification pour simplifier le séparateur

→ Utiliser un noyau pour trouver un séparateur non linéaire

$$\text{Noyau RBF : } K(d^{(n)}, d^{(n')}) = \exp\left(-\gamma (d^{(n)} - d^{(n')})^2\right)$$

# Comparaisons des deux méthodes

→ Paramètres des méthodes :

|          | SVM          | FAAD   |
|----------|--------------|--------|
| Quantité | +            | ++     |
| Réglage  | systematique | unique |

→ Mise en pratique:

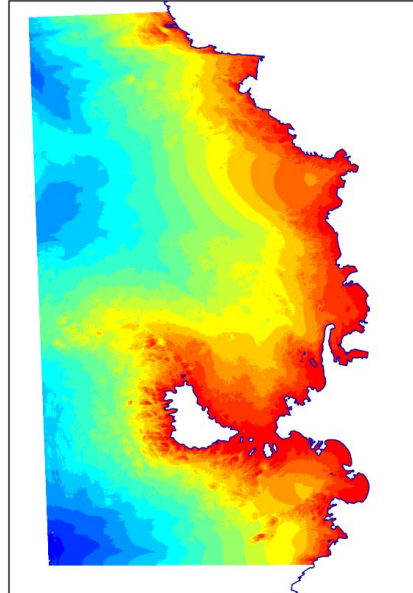
|            | SVM                                 | FAAD                                   |
|------------|-------------------------------------|--|
| Difficulté | réglage systématique des paramètres | non-répétitivité (caractère aléatoire) |
| Solution   | grille de recherche des valeurs     | augmenter le nombre d'arbres           |

# Lot de caractéristiques

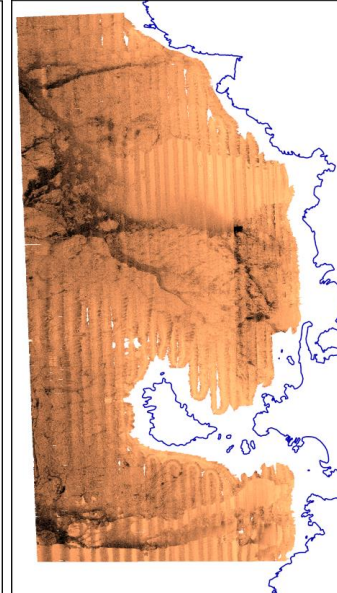
Image aérienne  
Rouge-Vert-Bleu



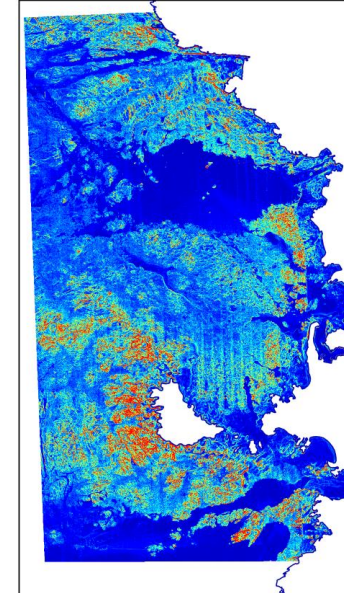
Bathymétrie



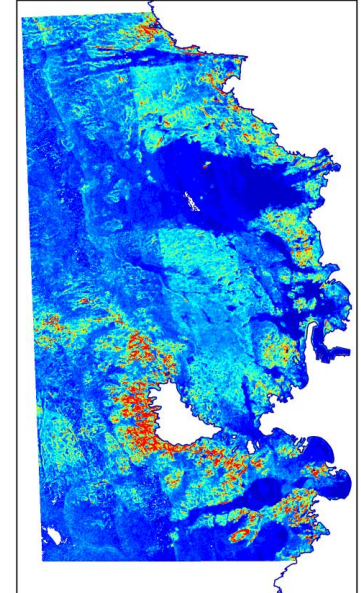
Réflectivité



Pente



Rugosité



|       |         |   |  |   |   |
|-------|---------|---|--|---|---|
| Lot A | X (RVB) | X |  | X | X |
| Lot B |         | X |  | X | X |
| Lot C | Bleu    | X |  | X | X |

# Résultats avec les forêts aléatoires d'arbres décisionnels

→ Variations des paramètres :

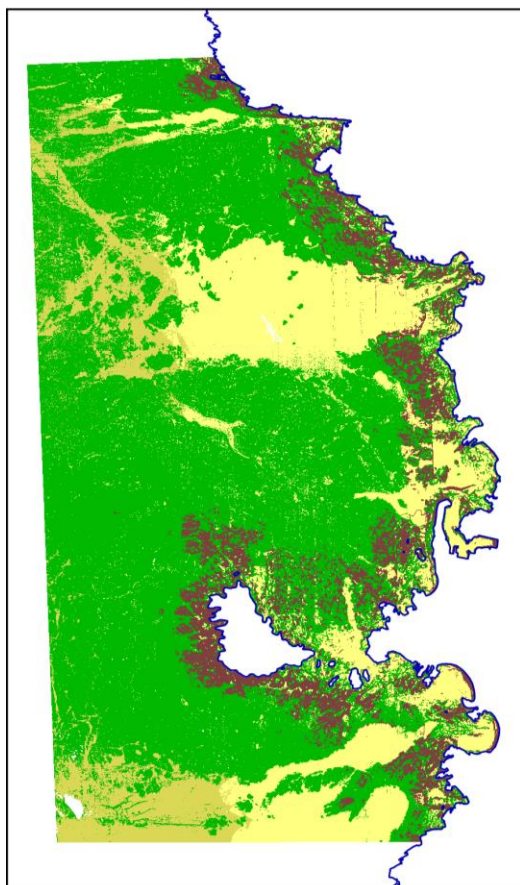
- Nombre d'arbres
- Profondeur des arbres
- Lot de caractéristiques

→ Très rapide : 2 à 8 minutes de calcul des modèles

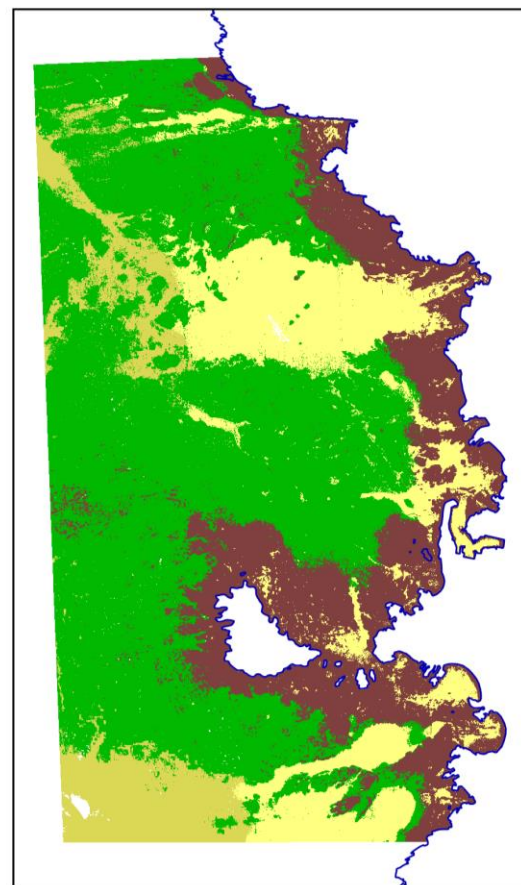
→ Taux d'erreur < 10 % entre les labels et l'apprentissage

→ Bons résultats : < 17 % d'erreurs avec la carte de Seaviews

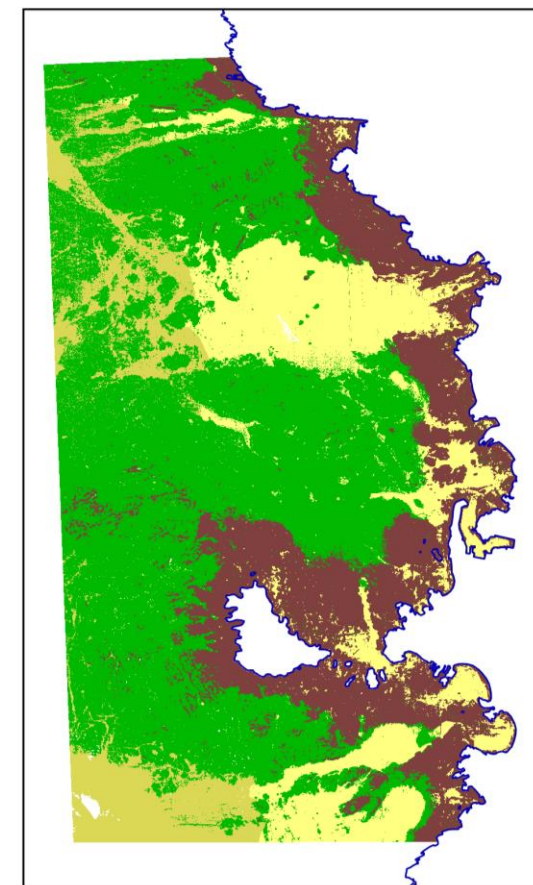
Lot B, n\_arbre=200, prof\_max=2



Lot A, n\_arbre=200, prof\_max=8



Lot B, n\_arbre=500, prof\_max=4



0 500 1000 m



Herbiers de Posidonies

Rochers à algues photophiles

Sables fins infralittoraux

Détritique côtier



# Résultat par les Séparateurs à Vastes Marges

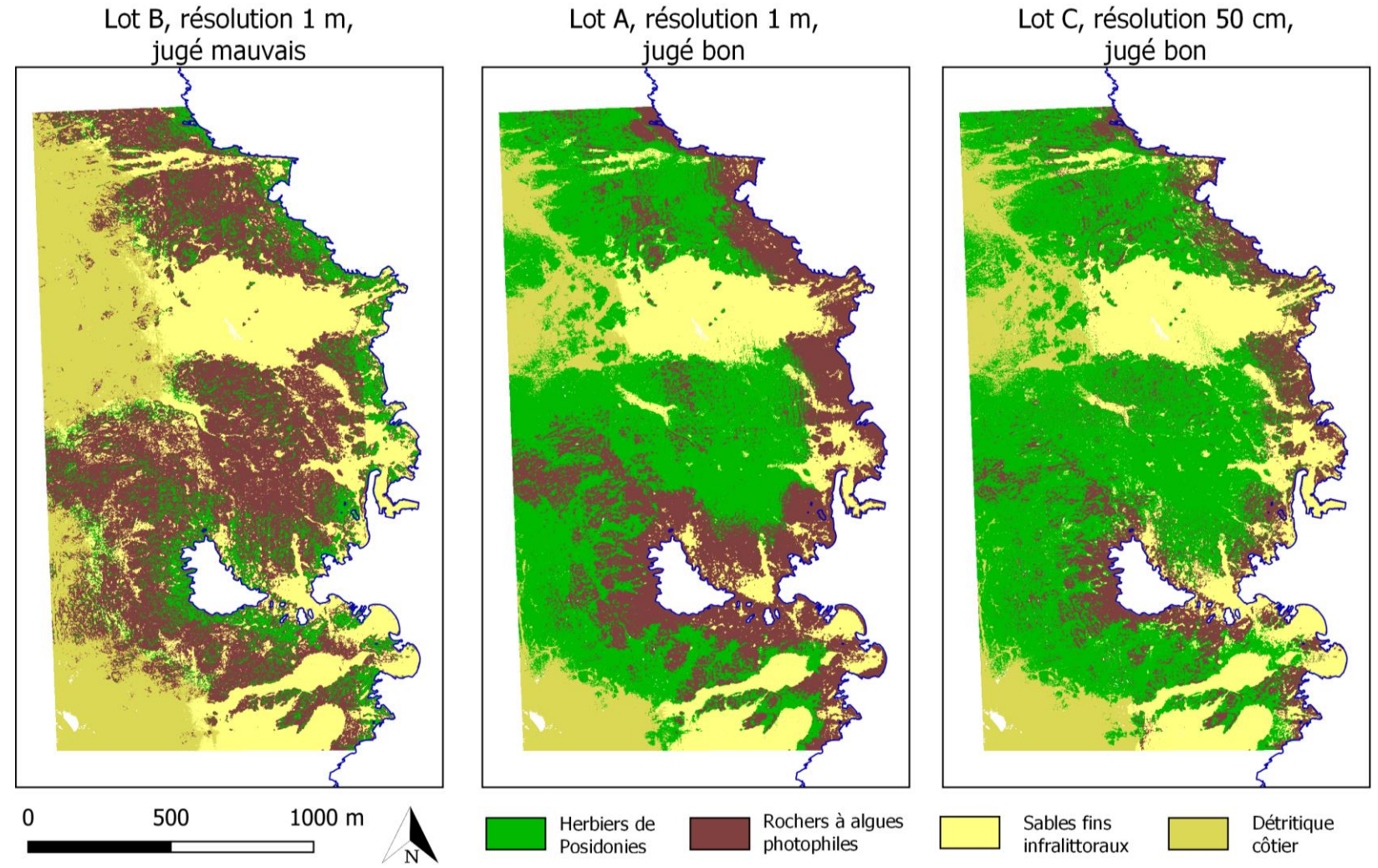
→ Variations des paramètres :

- Lot de caractéristiques
- Résolution de maille du MNT

→ Rapide : 4 à 15 minutes de calcul des modèles

→ Taux d'erreur < 15 % entre les labels et l'apprentissage

→ Qualité des résultats très variables



# Conclusions

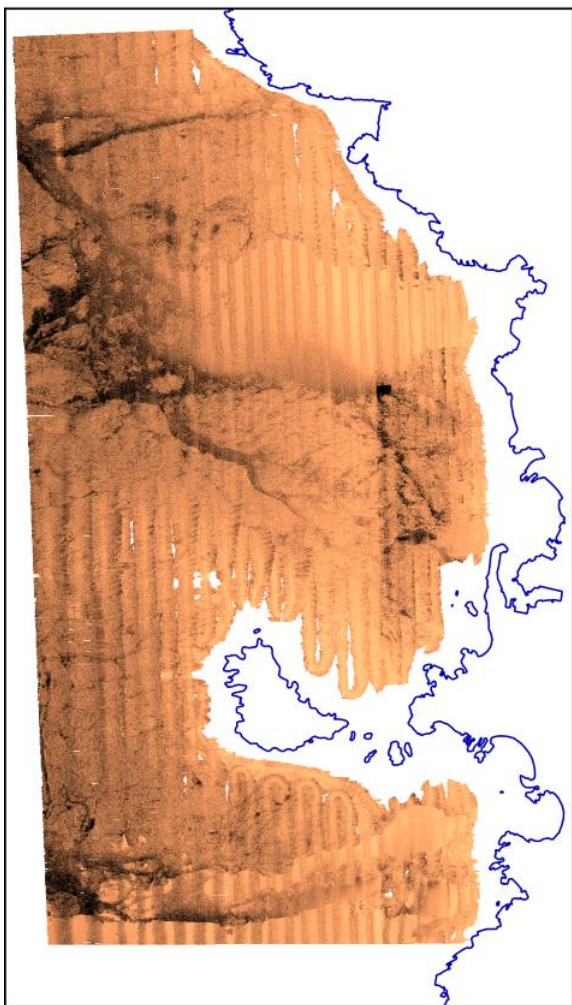
→ Comparaison de la méthode de Seaviews et les méthodes automatiques :

|                         | Méthode originelle de Seaviews   | Méthodes automatiques nouvellement intégrées |
|-------------------------|--|--|
| Temps                   | ½ journée au mieux   | 1-2 heures                                   |
| Données utilisables     | MNT bathymétrique, réflectivité, pente et rugosité   | MNTs + images aériennes                      |
| Données complémentaires | Imagerie aérienne / photogrammétrie, données de vérité terrain, données bibliographiques, anciennes cartes | Non  |
| Retouches du résultat   | Oui  | Non  |

→ Méthodes complémentaires l'une à l'autre : permet d'avoir un support de comparaison des résultats

# Perspectives

Imagerie de réflectivité



1

Pré-traitement des données

→ Nettoyer les défauts des données :  
Trace de navigation,  
Reflets du soleil sur les images aériennes

2

Gestion des données manquantes

→ Réaliser plusieurs modèles et fusion des résultats ?  
→ Compléter les données avec des valeurs non classifiantes ?  
→ Les forêts aléatoires et le classifieur bayésien peuvent s'y adapter

# Plan

## Contexte des travaux

1 – Données issues du sondeur multifaisceaux

2 – Estimation de valeurs bathymétriques  
par réseau de neurones convolutifs

3 – Classification des données pour  
l'identification des habitats marins

**Conclusion générale**



**Résumé des travaux**

**Perspectives**

# Résumé des travaux

→ Intégration de méthodes à **apprentissage supervisé** dans les logiciels de Seaviews

- Une approche par « **Deep Learning** » pour l'estimation bathymétrique
- Des algorithmes de « **Machine Learning** » pour la classification des fonds

→ Ces méthodes ont en commun de :

- Réaliser **automatiquement** une tâche après entraînement
- Nécessiter **une grande quantité de données** pour l'entraînement
- Devoir préparer **la labellisation** des données d'entraînement
- Permettre d'obtenir de **bons résultats** sur des données expérimentales bruitées

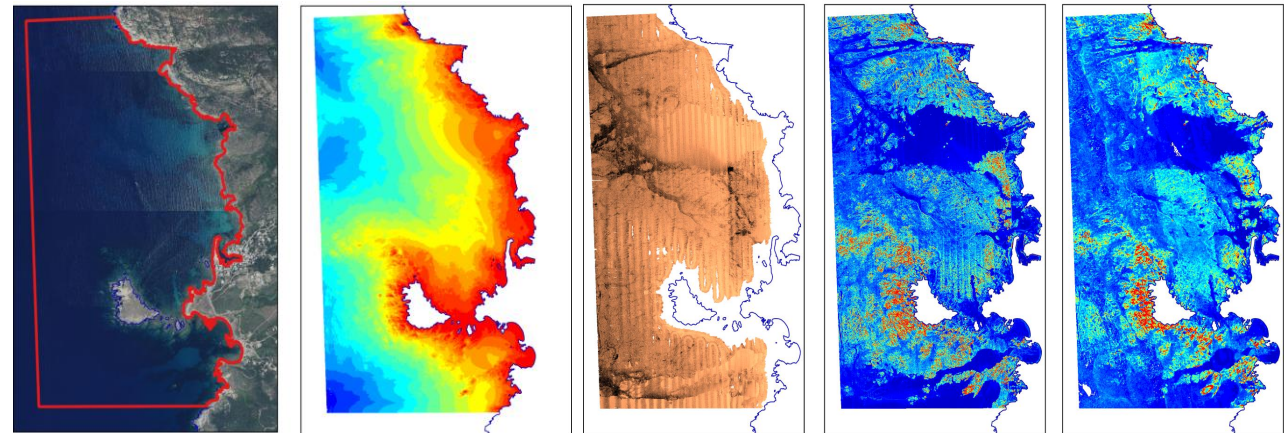
# Résumé des travaux

→ Points de divergence :

|                         | <b>Deep Learning / bathymétrie</b>  | <b>Machine Learning / classification</b>  |
|-------------------------|-------------------------------------|---|
| <b>Données d'entrée</b> | Image « brute » de la colonne d'eau | Caractéristiques déduites des données SMF |
| <b>Apprentissage</b>    | Unique et très long                 | Spécifique à une zone et très rapide      |
| <b>Labellisation</b>    | Automatique                         | Réalisée par un expert                    |



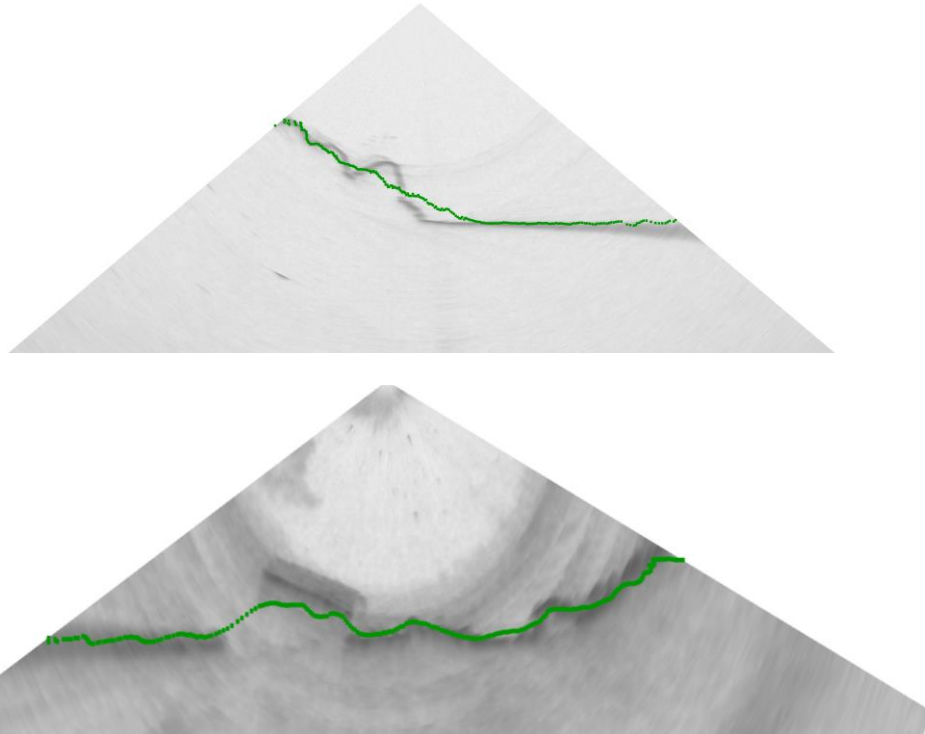
Donnée brute



Données brutes et caractéristiques extraites

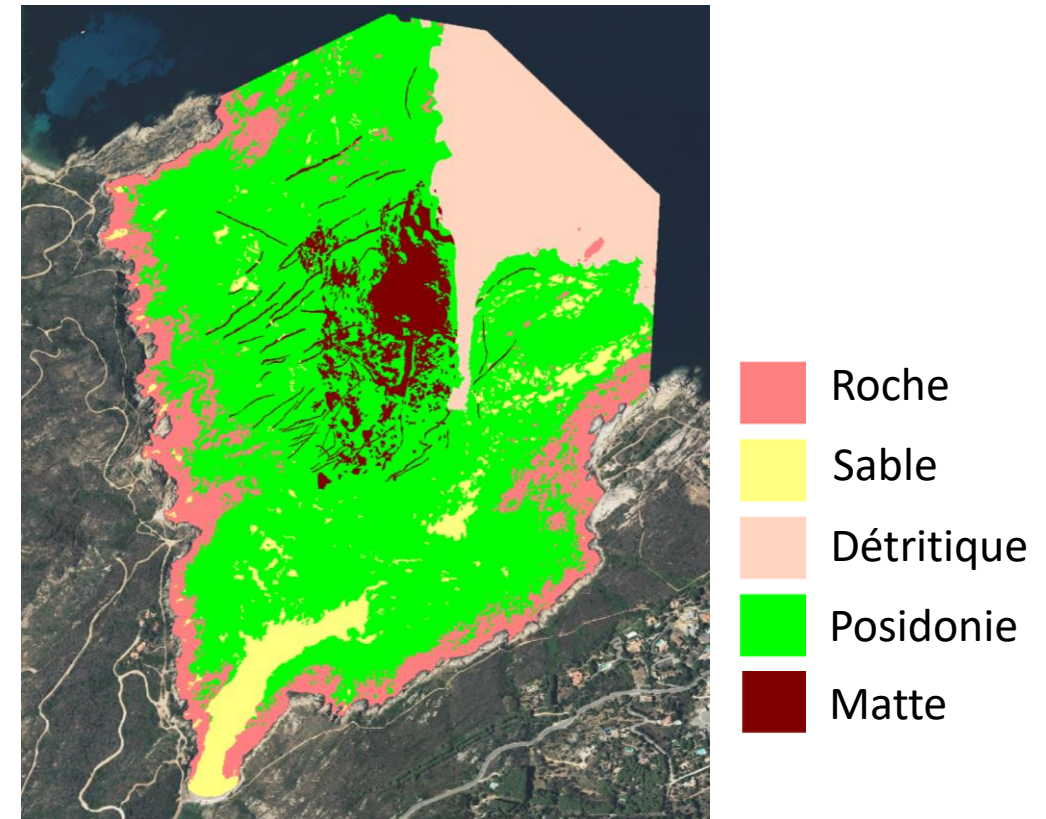
# Perspectives

→ Valider le prototype avec la 2<sup>e</sup> phase d'apprentissage en mode « zoom »



Estimation de la bathymétrie par réseau de neurones

→ Tester sur un plus grand nombre de jeux de données  
→ Améliorer le pré-traitement des données



Évolution de la matte morte de Posidonie dans la baie de l'Alga, Calvi

**Merci de votre attention**

Des questions ?