

Extraction automatique de structures anthropiques

Adrien Acchiardi, Gaël Chauchadis, Zoé Laurent--Iranmher, Elsa Supiot

Encadré par Nathalie Debese et Rodéric Moitié

7 mars 2025



Objectifs

- Extraction d'objets anthropiques: **écluses, déversoirs, maisons éclusières**
- On effectue un **levé hydrographique** sur le lac de Guerlédan
- On s'appuie sur des attributs **géomorphométriques**

↳ Comment **détecter** des structures anthropiques **automatiquement** ?



Figure- Vue aérienne du lac de Guerlédan

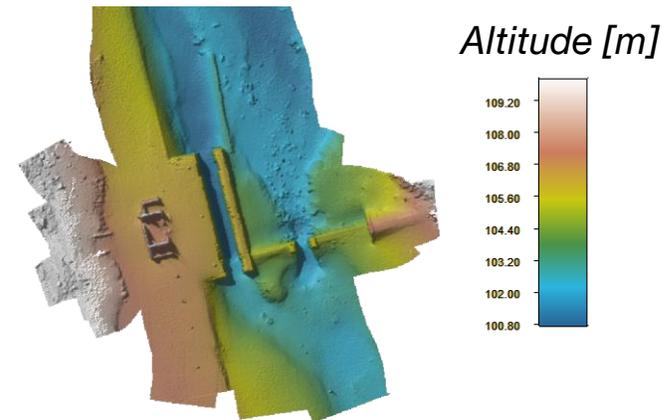


Figure- MNT de l'écluse de Cuilleret



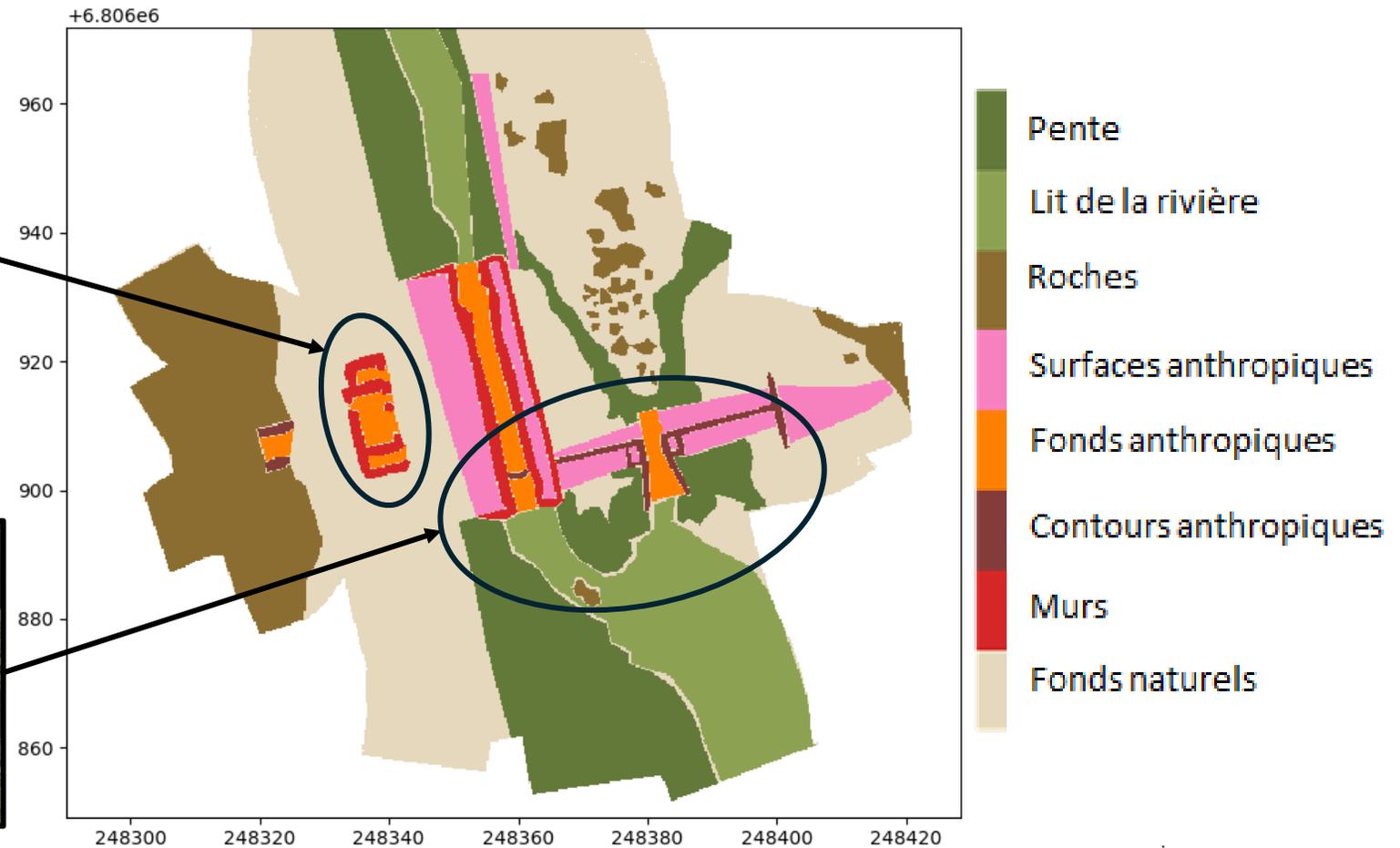
Plan

1. Description de la zone
2. Choix des attributs
3. Optimisation et choix de l'algorithme
4. Résultats



Description de la zone

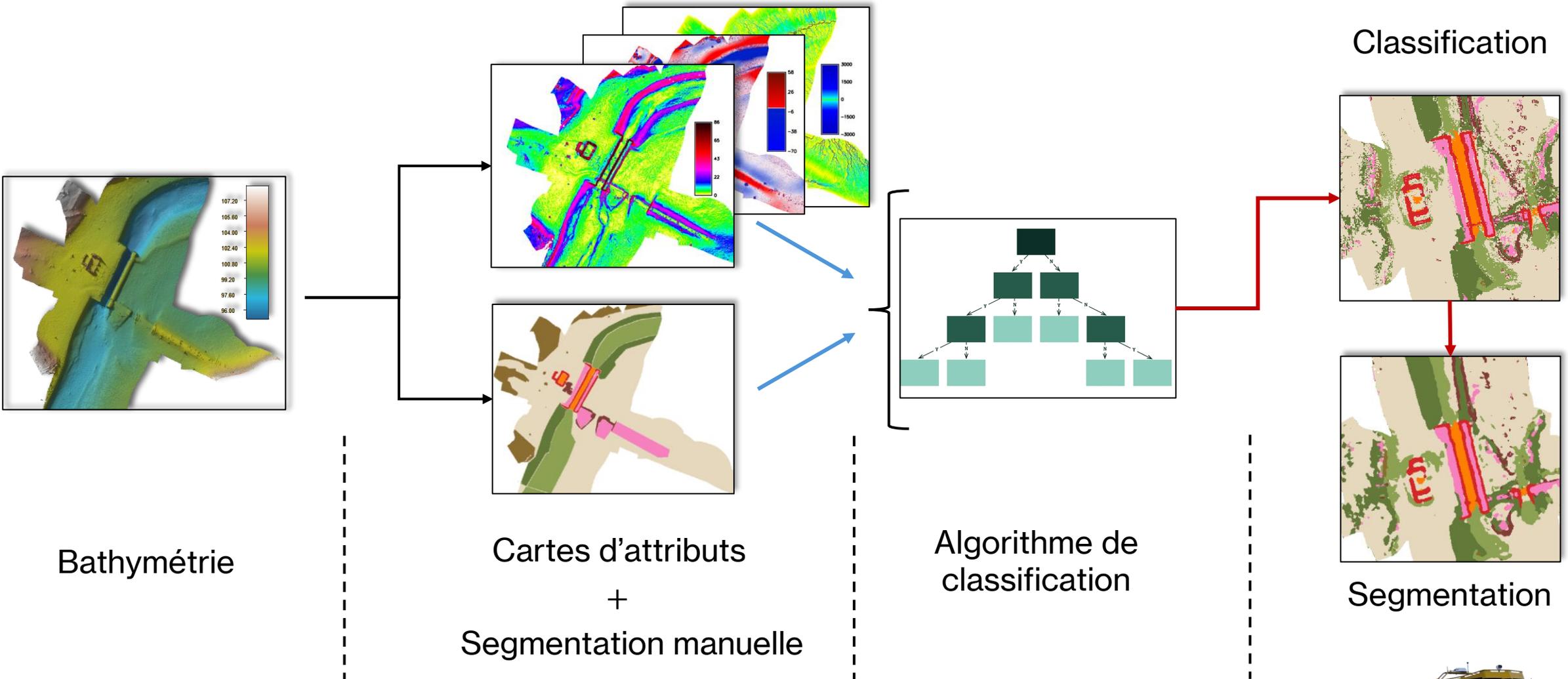
Éléments à identifier / Segmentation manuelle



Classification



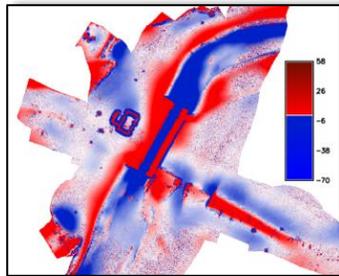
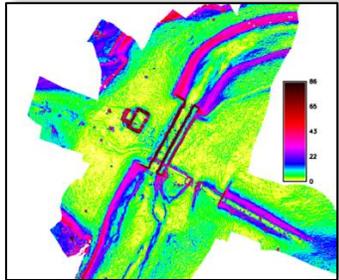
Segmentation



Choix des attributs

Une approche empirique

Pente

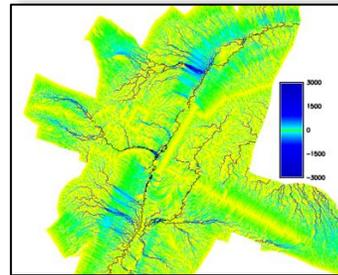


TPI

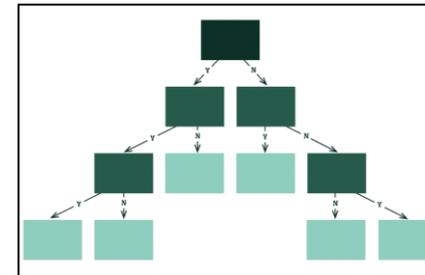
(Topographic Position Index)

2 attributs **fondamentaux**

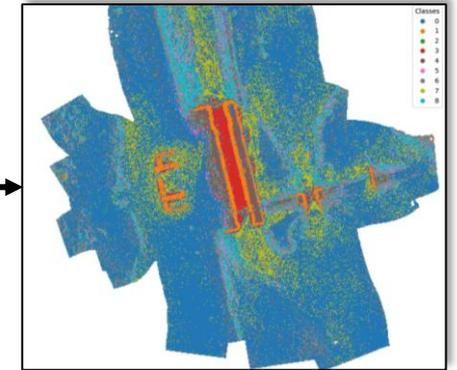
Attribut évalué



*Algorithme de classification
(arbre de décision)*



Classification



Attribut écarté

Attribut retenu



Choix des attributs

Attributs retenus

- Le TPI
- La pente
- L'accumulation
- L'accumulation inverse
- Les classes géomorphon

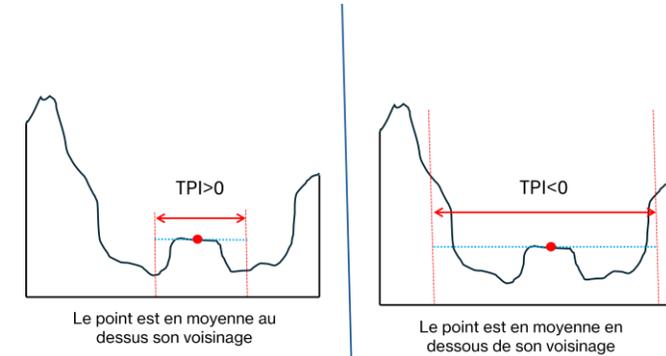
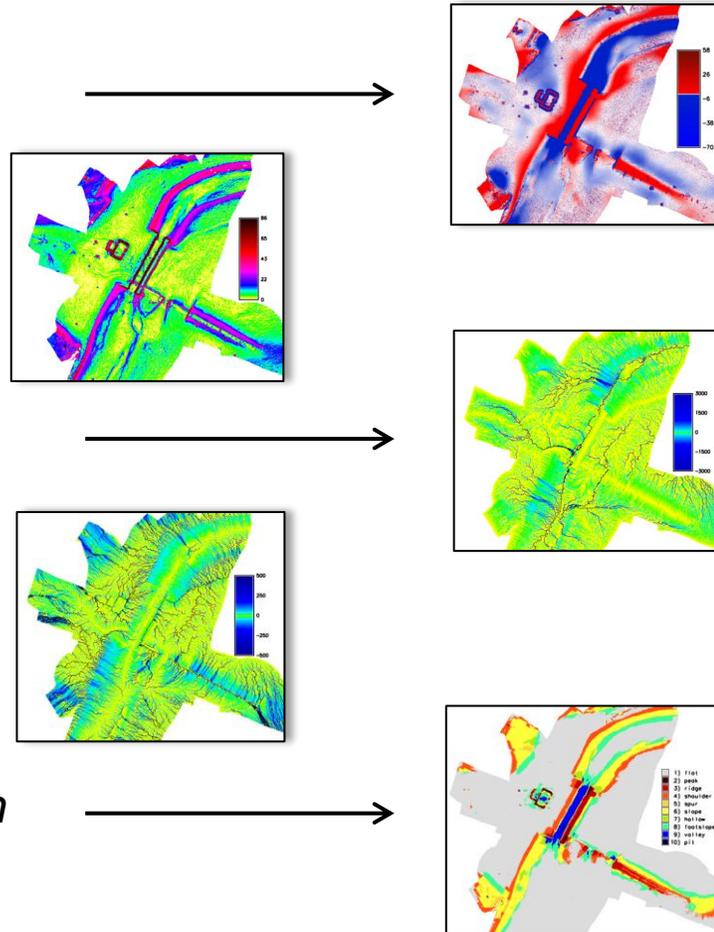


Figure- Principe du TPI

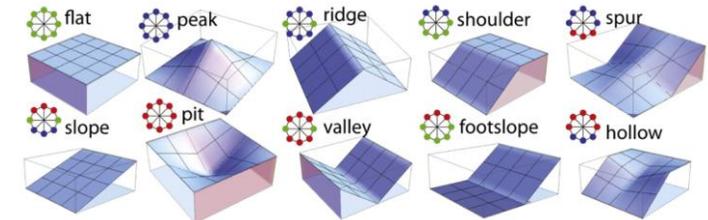


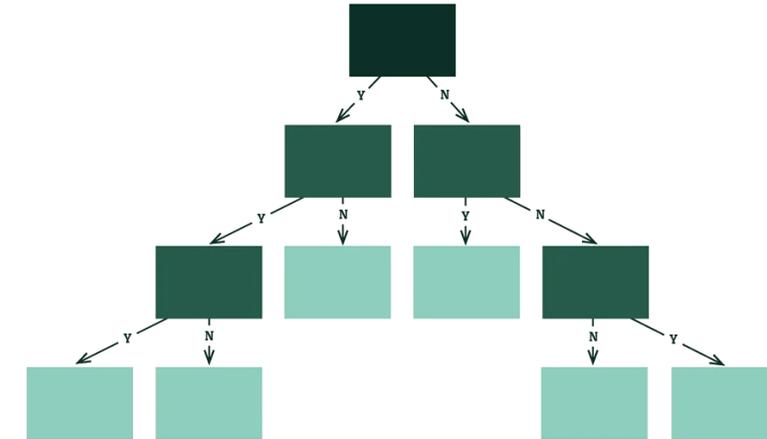
Figure- Les 10 classes de géomorphon [1]



Présentation des principaux algorithmes

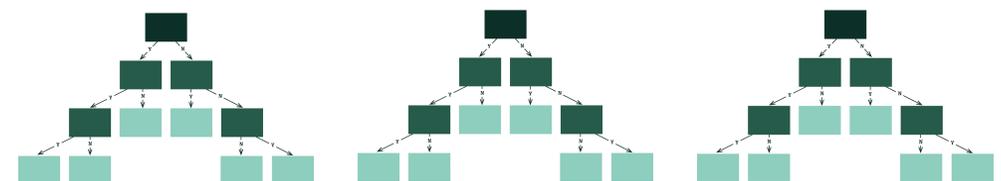
- **Arbre de décision**

- Algorithme **supervisé** : Nécessite un échantillon d'**entraînement**
- Détermine les critères de séparation pour **homogénéiser** les nœuds
→ Homogénéité mesurée par une fonction d'impureté
- Possibilité de limiter la **profondeur** pour éviter le **surapprentissage**



- **Random Forest**

- Construction de plusieurs arbres en **parallèles**
- Chaque arbre apprend sur un **sous-ensemble aléatoire** de l'échantillon d'entraînement
- Le choix final est un **compromis** des différents arbres

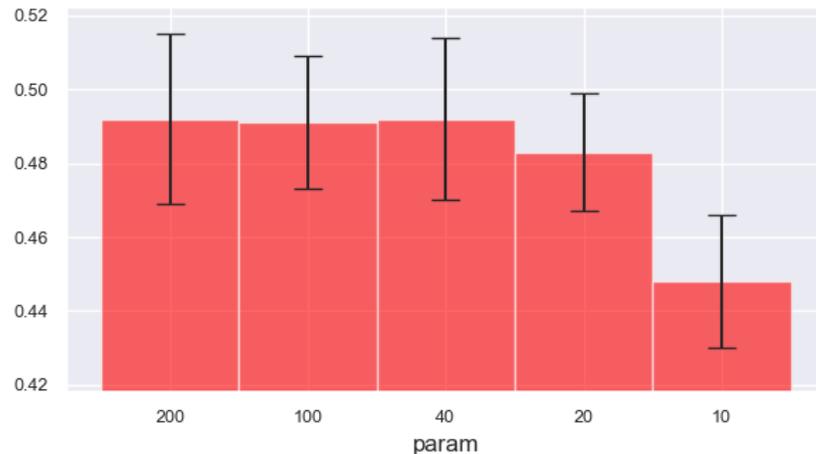


Méthodologie

- **Hyperparamètre** : paramètre interne de l'algorithme.
- **Optimisation** : Choix du meilleur **hyperparamètre** de l'algorithme.
→ Tests et évaluations pour différentes valeurs de paramètre.
- Regroupement des 8 classes en 2 classes : **Naturel / Anthropique**.
→ Permet d'évaluer les classifications à l'aide du **f1-score**.

| | Anthropique | Naturel |
|--------------------|-------------|---------|
| Déecté Anthropique | Va | Fa |
| Déecté Naturel | Fn | Vn |

Figure - Matrice de contingence



$$\text{F1-score} = \frac{2.Va}{2.Va+Fa+Fn}$$

- Pour les algorithmes **aléatoires** :
→ Méthode de **Monte Carlo** sur 20 itérations.
→ On prend en compte l'écart-type du F1-score.

Figure - F1-score pour l'algorithme de Random Forest pour différentes profondeurs d'arbres



Choix de l'algorithme

Résultat d'optimisation

| Algorithme | Arbre de decision (Tree) | Random Forest (RF) | Réseau de neurones (MLP) |
|------------------|--------------------------|---------------------|--------------------------|
| Hyperparamètre | Profondeur maximale | Profondeur maximale | Nombre de neurones |
| Valeur optimisée | 10 | 40 | 150 |

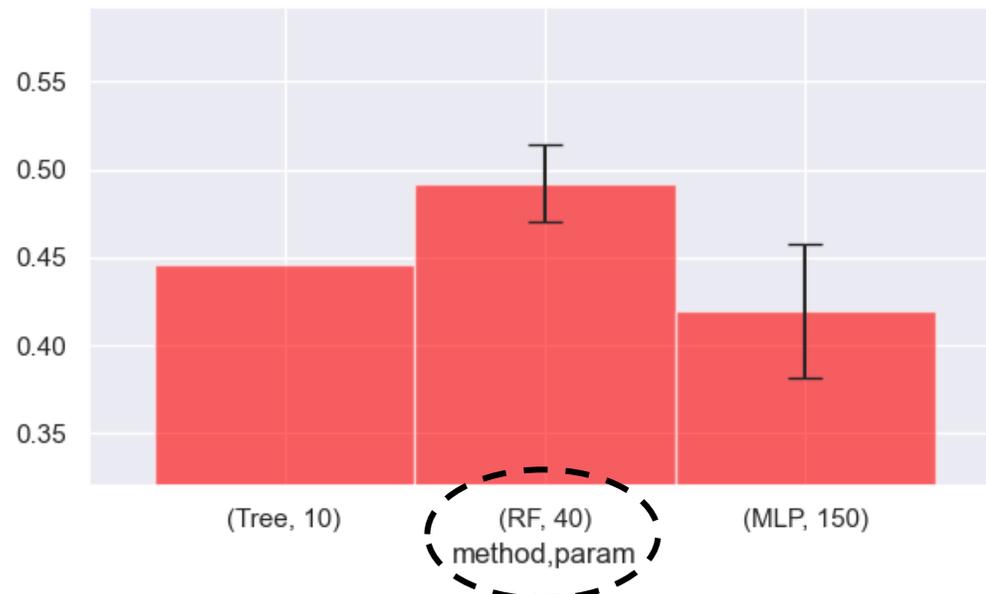
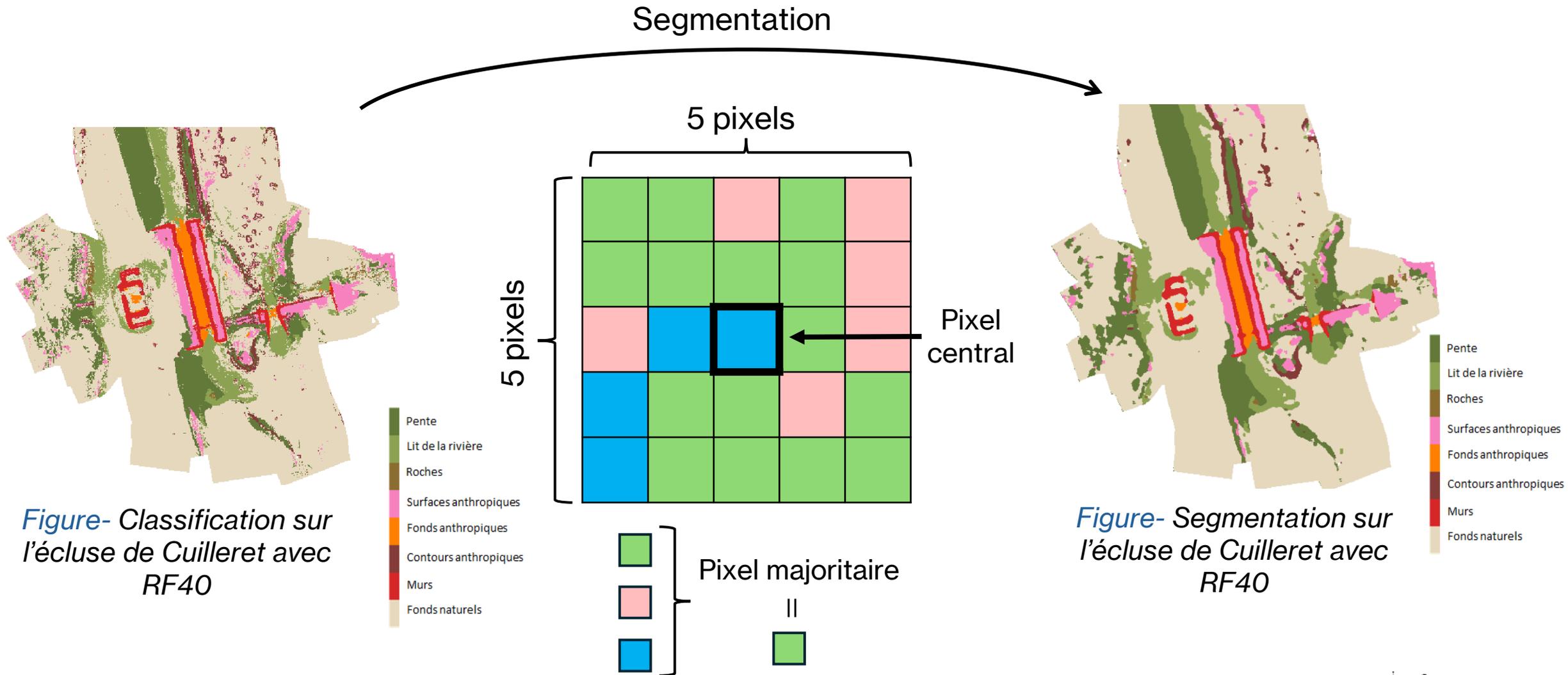


Figure - F1-score pour chaque algorithme optimisé

Algorithme retenu :
→ **Random Forest** avec
une profondeur de **40**





Pertinence des classes

| Classes | Proportion surfacique (%) |
|-----------------------|---------------------------|
| Fonds naturels | 55 |
| Murs | 1 |
| Contours anthropiques | 0.2 |
| Fonds anthropiques | 0.9 |
| Surfaces anthropiques | 3 |
| Roches | 14 |
| Lit de rivière | 16 |
| Pente | 10 |

Tableau- Répartition surfacique des classes de la segmentation manuelle

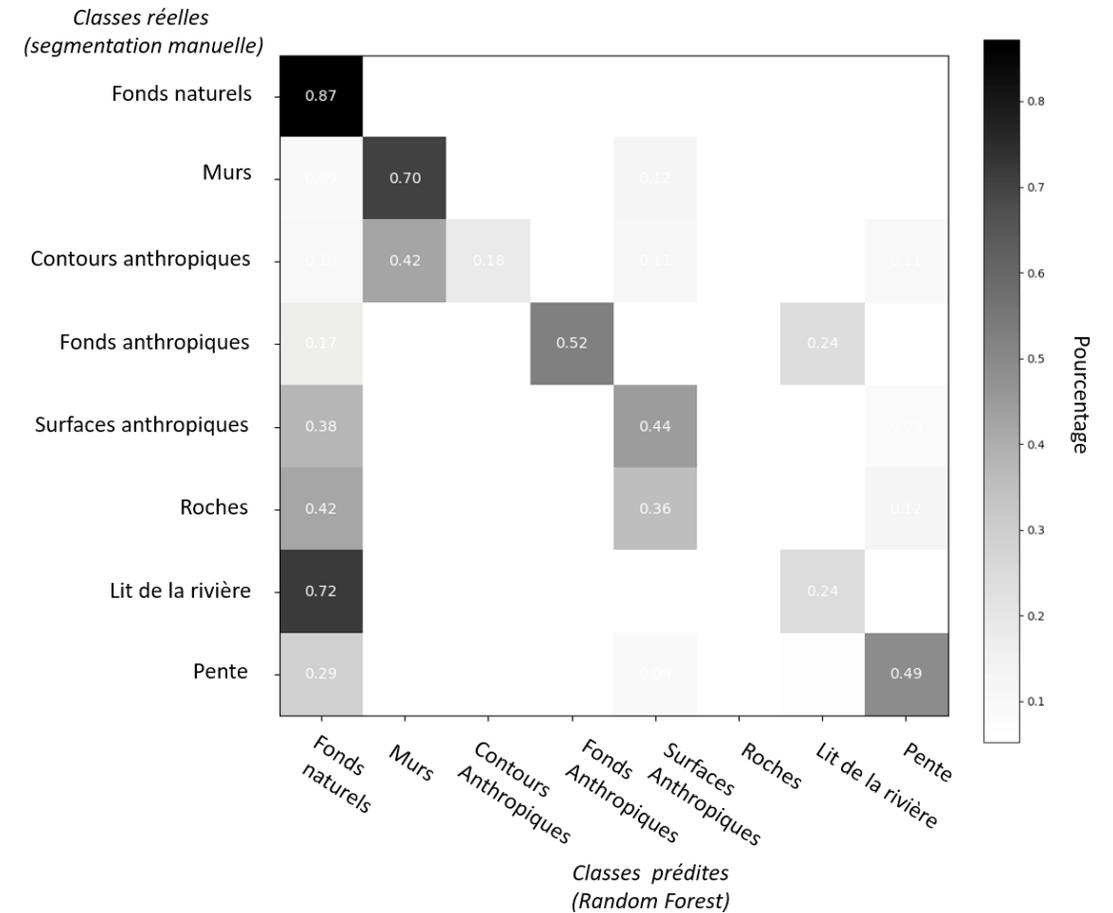


Figure- Matrice de contingence normalisée par les classes réelles



Application à un autre levé

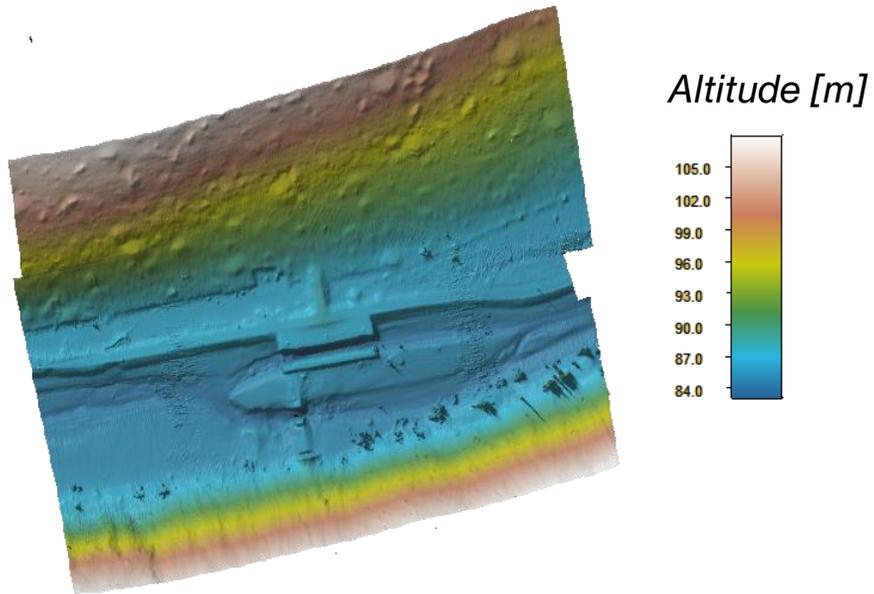


Figure- MNT de l'écluse de Caurel

F1-score = 0,16

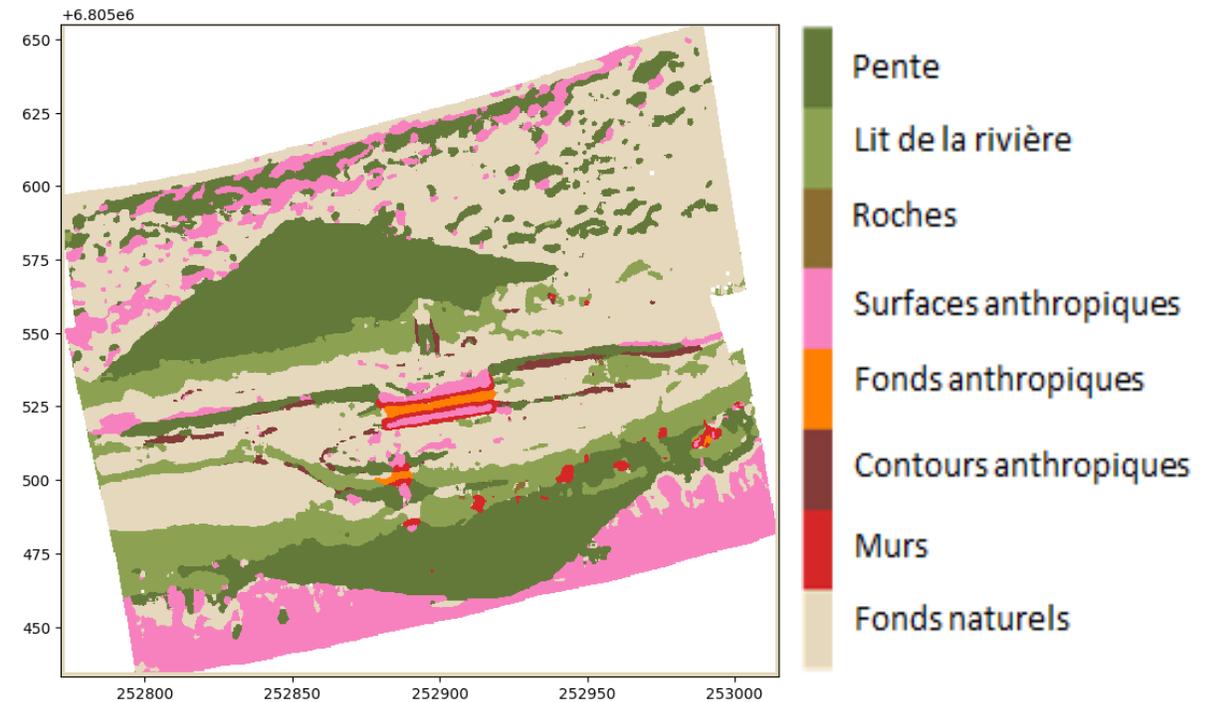


Figure- Segmentation sur l'écluse de Caurel avec RF40



- Segmentation finale **satisfaisante** et méthode **reproductible**
- **Biais** liés à l'ordre des choix (segmentation manuelle, attributs, algorithmes)
- Beaucoup de possibilités, **pas d'exhaustivité** des méthodes
- Piste non explorée prometteuse : ***Deep Learning***



Merci pour votre attention



Explication du F1-score

- **Sensibilité** : $\frac{Va}{Va+Fn}$
→ Part des pixels anthropiques détectés comme tels

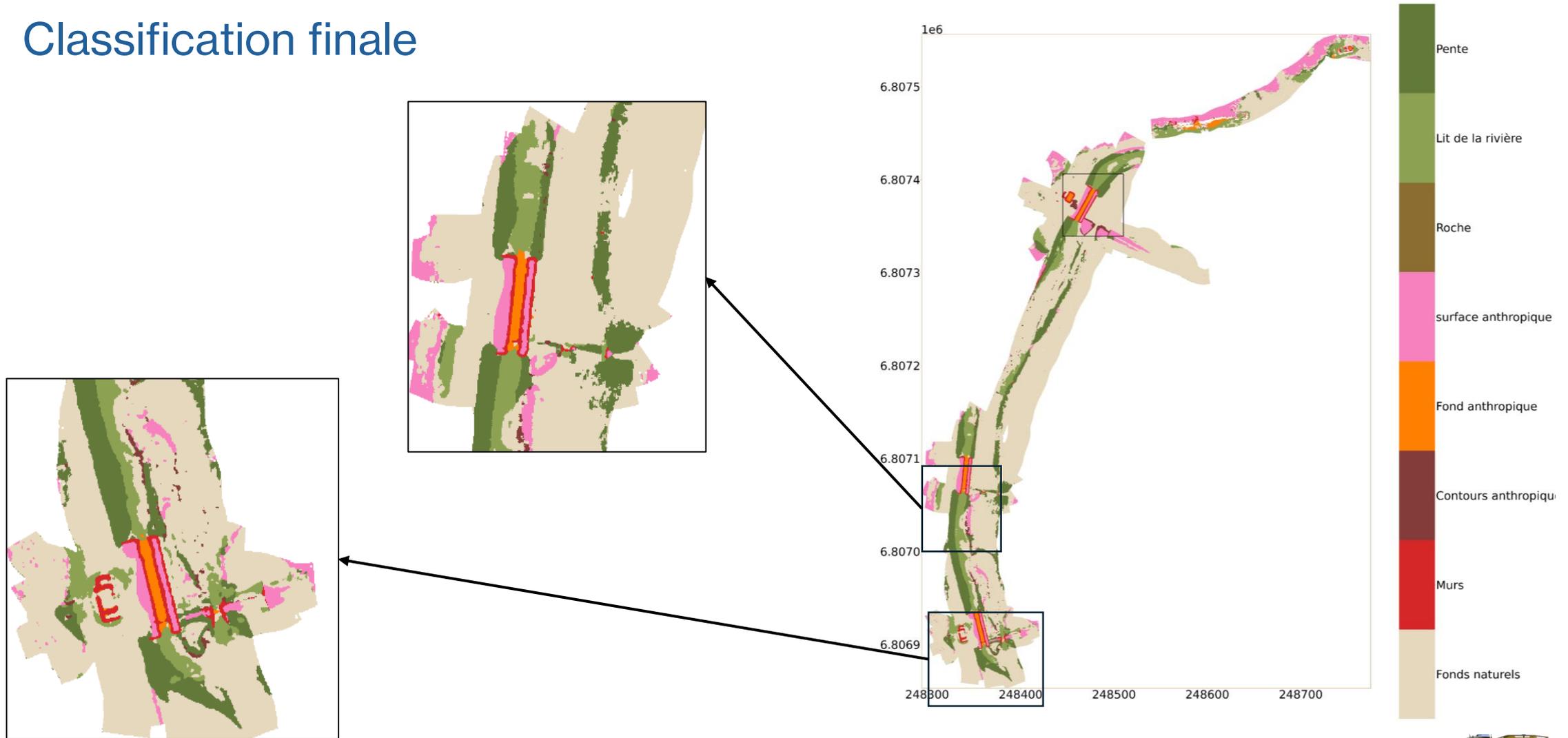
- **Précision** : $\frac{Va}{Va+Fa}$
→ Part des pixels détectés comme anthropique étant réellement anthropiques

- **F1-score** : $\frac{2.Va}{2.Va+Fa+Fn}$
→ Compromis entre la **précision** et la **sensibilité**

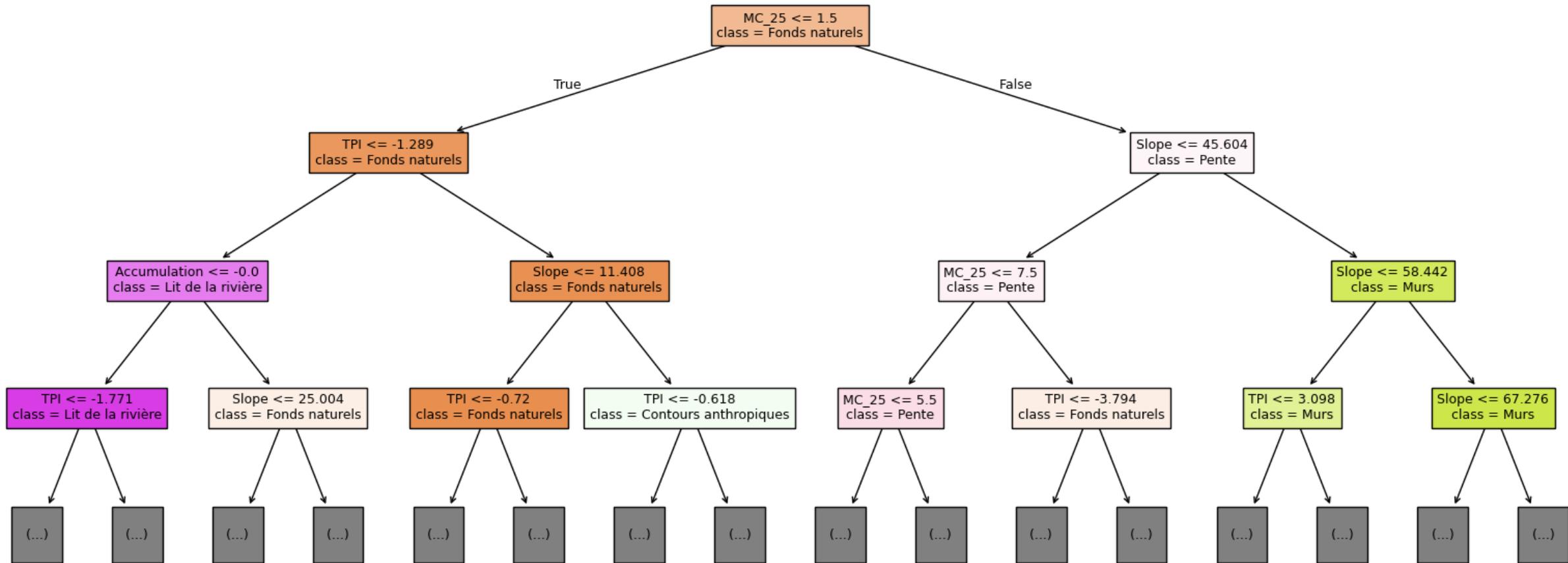
| | Anthropique | Naturel |
|--------------------|-------------|---------|
| Déecté Anthropique | Va | Fa |
| Déecté Naturel | Fn | Vn |



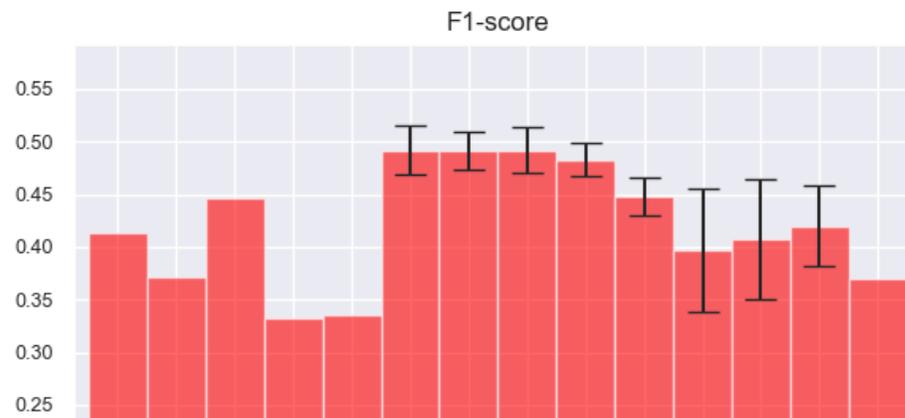
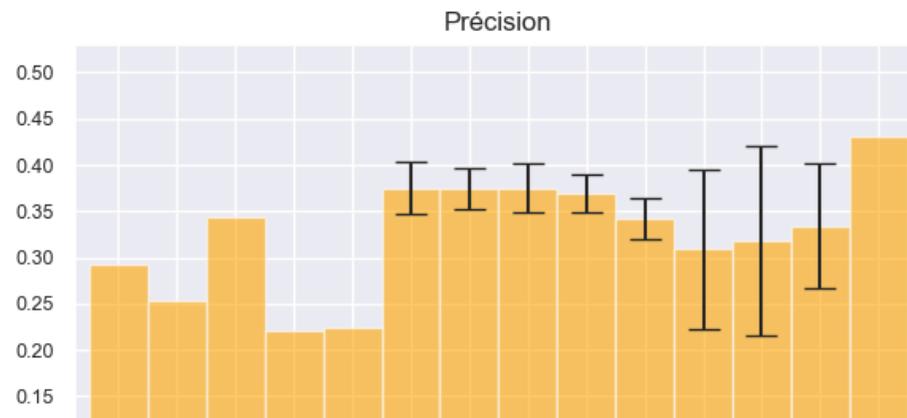
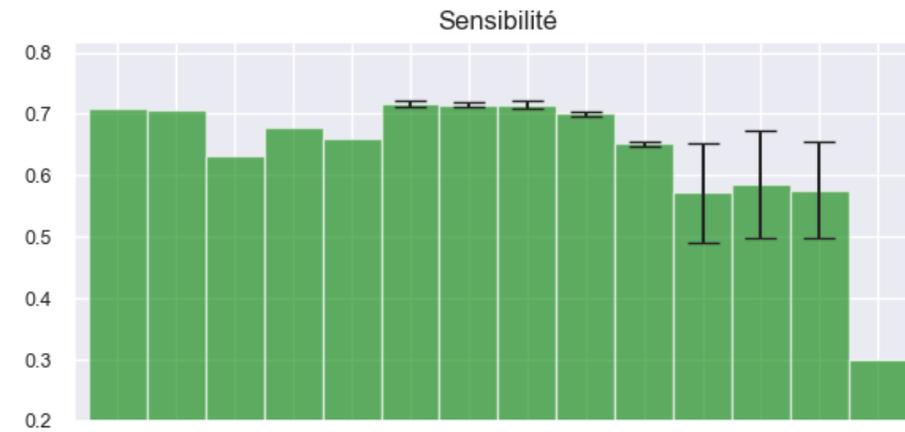
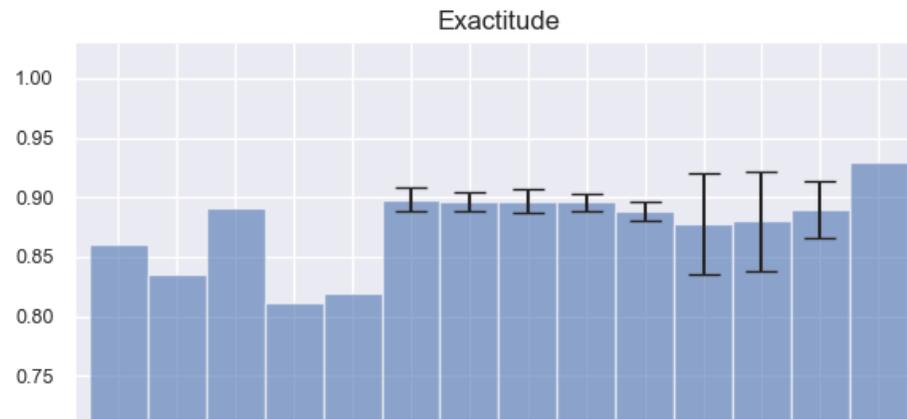
Classification finale



Arbre de décision



Scores



(Tree, 40)
(Tree, 20)
(Tree, 10)
(Tree, 8)
(Tree, 5)
(RF, 200)
(RF, 100)
(RF, 40)
(RF, 20)
(RF, 10)
(MLP, 50)
(MLP, 100)
(MLP, 150)
(SVC, Linéaire)

method,param

(Tree, 40)
(Tree, 20)
(Tree, 10)
(Tree, 8)
(Tree, 5)
(RF, 200)
(RF, 100)
(RF, 40)
(RF, 20)
(RF, 10)
(MLP, 50)
(MLP, 100)
(MLP, 150)
(SVC, Linéaire)

method,param



Indice de *Gini*

Probabilité qu'un point soit **mal classifié** s'il est attribué **aléatoirement** à une classe en ne tenant compte que des **proportions** de chaque classe dans l'échantillon.

$$Gini = 1 - \sum_{k=1}^n p_k^2$$

Avec :

- n le nombre de classes
- p_k la proportion d'éléments de la classe k dans l'ensemble

Pour une **feuille pure**, on a $\forall k \neq k_i, p_k = 0$ et $p_{k_i} = 1$

On a alors $Gini = 1 - 1^2 = 0$

→ **L'impureté est bien nulle**

