

LA CLASSIFICATION DE PARCELLES AGRICOLES BRETONNES

La prise en compte de données sur plusieurs années permet-elle d'améliorer la classification des parcelles agricoles sous-représentées ?

INTRODUCTION

CONTEXTE

Ce travail s'appuie sur le jeu de données Breizhcrops¹ créé pour entraîner des algorithmes à la classification de parcelles agricoles. Il présente de nombreux défis : données manquantes, déséquilibre de classes, autocorrélation spatiale, etc.

QU'EST-CE QUE LA CLASSIFICATION ?

La classification a pour objectif de regrouper des objets spatiaux (souvent des pixels) selon leur appartenance à des classes d'occupation du territoire.

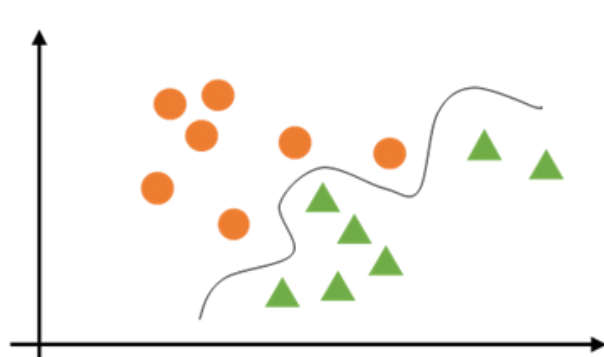
Classification supervisée (*classification*)

- Nombre de classes connu
- Utilisée pour classer des futures données



Classification non supervisée (*clustering*)

- Nombre de classes inconnu
- Utilisée pour comprendre et explorer des données



OBJECTIF

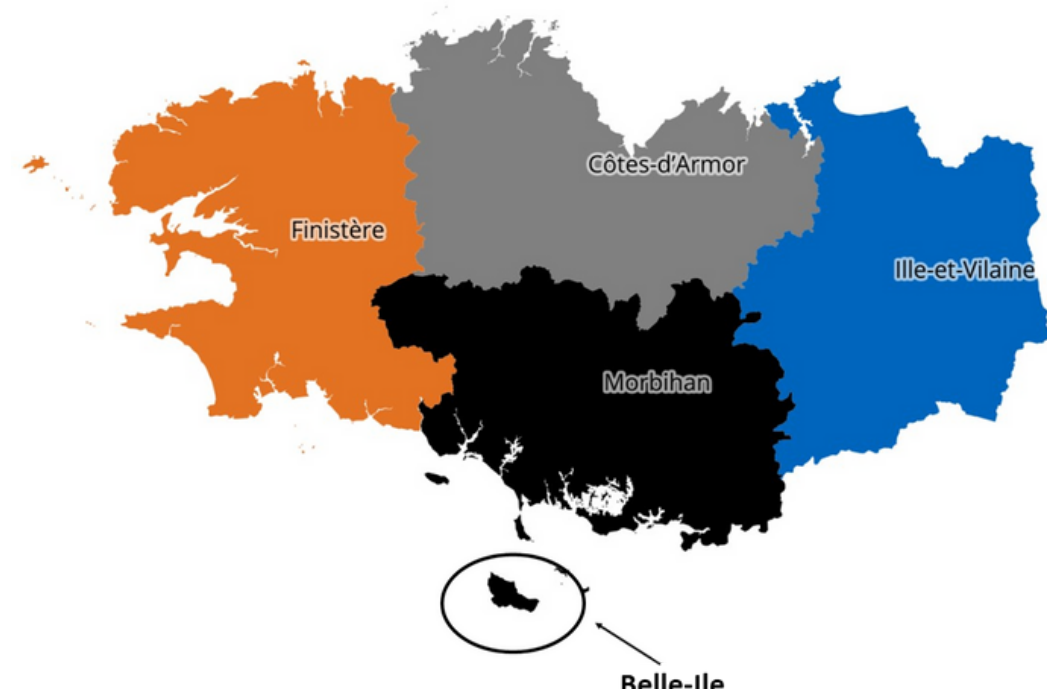
Cette étude se concentre sur la sous-représentation, appelée aussi déséquilibre de classes. Dans le cas de la classification d'occupation des sols agricoles, il pourrait être intéressant d'étudier plusieurs années. En effet, la rotation des cultures est fréquente pour ne pas épuiser le sol. L'information temporelle redondante et pluriannuelle permet donc ainsi de multiplier les sources d'observations des échantillons sous-représentés.

L'objectif est donc de classer les parcelles agricoles selon leur culture principale par des méthodes de classification, en essayant de palier à la sous-représentation.

DONNÉES

Jeu de données Breizhcrops

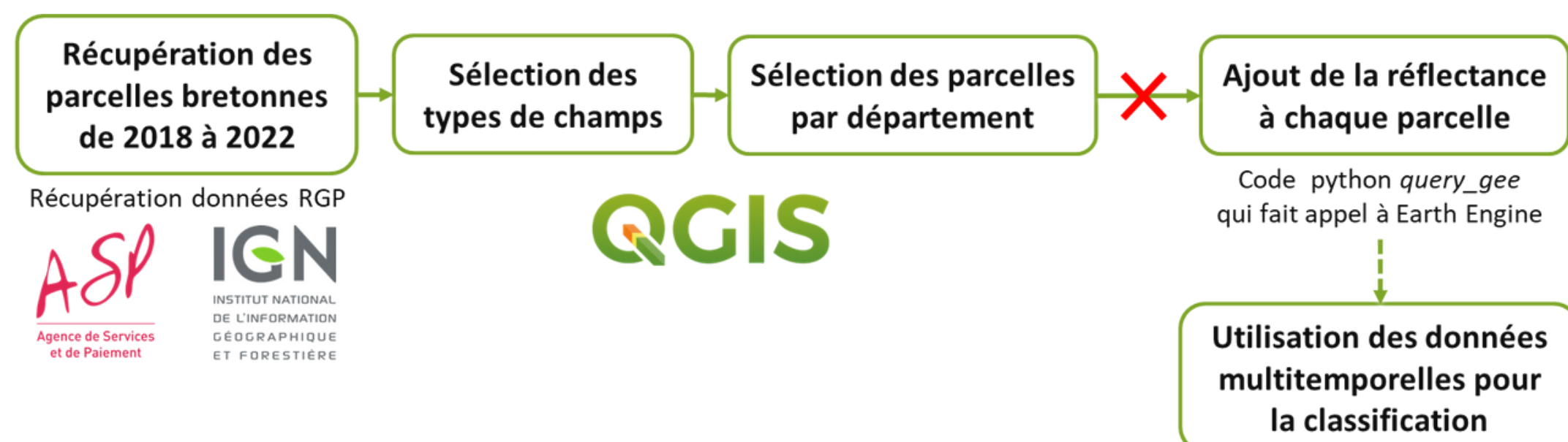
5 shapefiles des parcelles agricoles bretonnes : 1 par département + Belle-Île-en-Mer



Attributs des shapefiles :

- identifiant de chaque parcelle
- label (= groupe de culture)
- réflectances associées

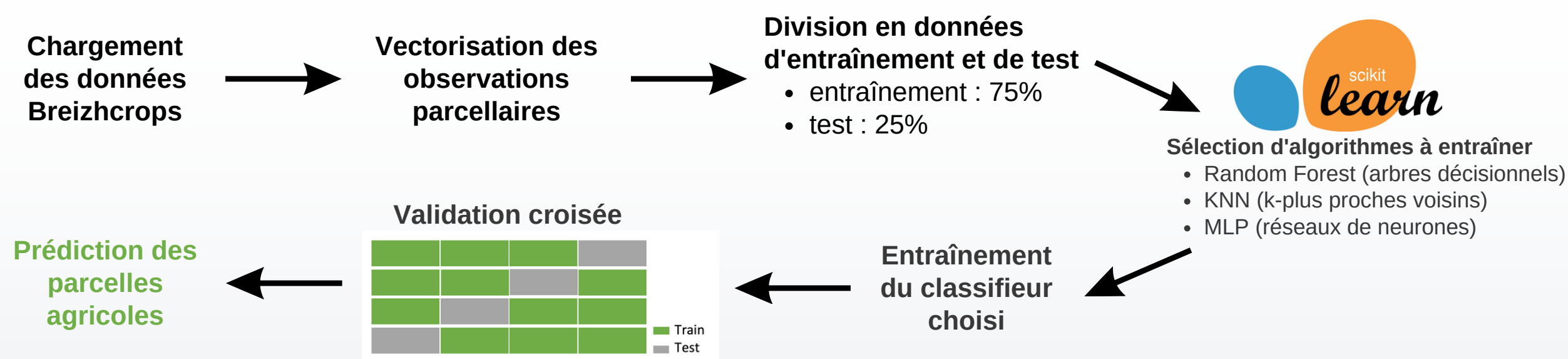
Construction de données sur le modèle de *Breizhcrops*



Chaîne de traitement pour l'ajout des données des années 2018 à 2022

La création de ces nouveaux jeux de données était trop couteuse en temps et n'a pas pu être réalisée.

TRAITEMENT DES DONNÉES



Chaîne de traitement pour la classification



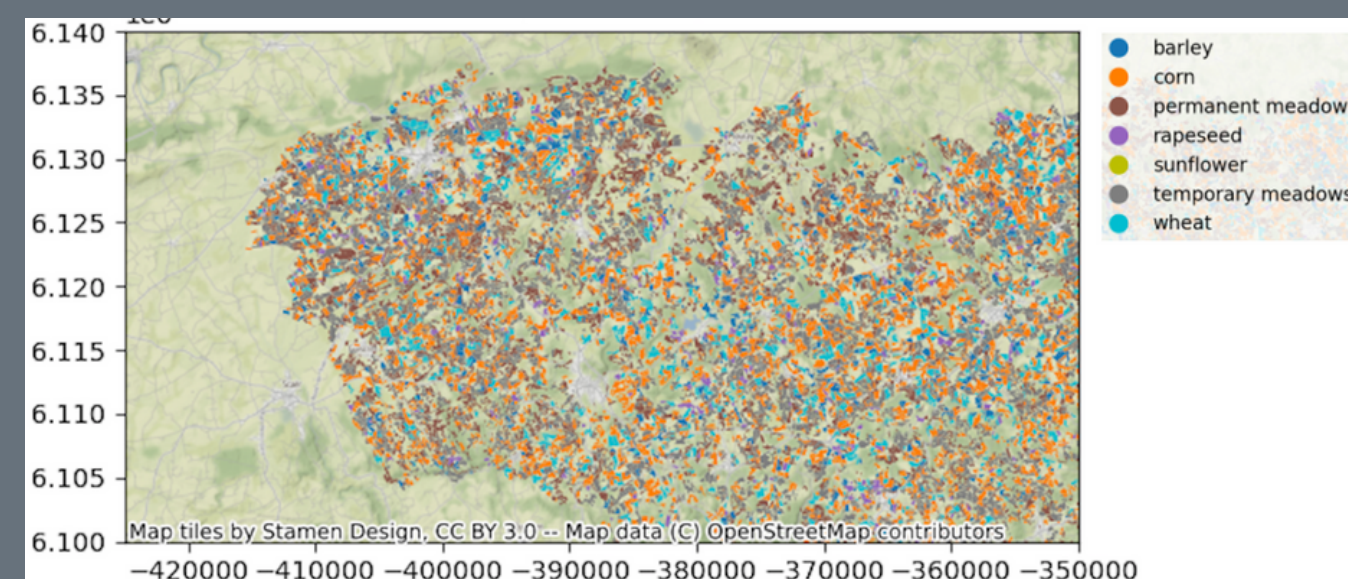
Seules les données de 2017 sont utilisées car la production des données des années suivantes était trop longue (données volumineuses).

La validation croisée a donc été mise à l'écart au profit de la comparaison de 3 algorithmes sur des métriques (matrice de confusion, accuracy, F-score, temps de calcul).

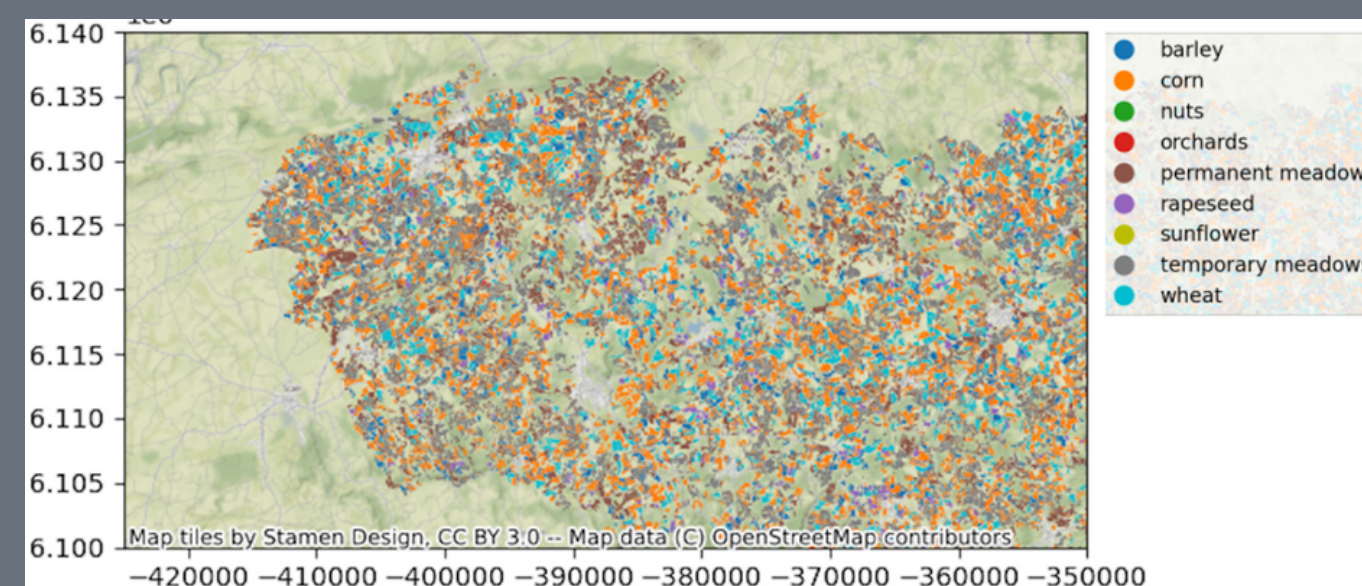
RÉSULTATS

	Random Forest	KNN (3 voisins)	MLP
Accuracy (%)	44	25.8	74.1
F-score (%)	32.4	25.9	73.7
Temps de calcul (s)	5	2692 (≈ 45 min)	854 (≈ 15 min)

Performances moyennes sur 3 itérations par algorithme (entraînements sur le Finistère, Ille-et-Vilaine, Morbihan et tests sur les Côtes-d'Armor)



Classification des parcelles agricoles dans le Morbihan — Vérité terrain



Classification des parcelles agricoles dans le Morbihan — Prédiction MLP

CONCLUSION ET PERSPECTIVES

Apports :

- Mise en évidence des méthodes de récupération des données des années ultérieures à 2017
- Établissement d'une chaîne de traitement pour la classification des parcelles agricoles

Poursuites de recherche :

- Entraîner les algorithmes sur des données pluriannuelles de la région Bretagne
- Étudier la classification des parcelles agricoles sur d'autres régions françaises
- Prendre en compte des attributs spatiaux comme des informations sur la géométrie et la topologie des parcelles, le contexte tel que le climat, la distance à l'eau ou la culture de l'année précédente

RÉFÉRENCES BIBLIOGRAPHIQUES

(1) Marc Rußwurm, Charlotte Pelletier, Maximilian Zollner, Sébastien Lefèvre, and Marco Körner. *Breizhcrops : A time series dataset for crop type mapping*, 2019.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. *Scikit-learn : Machine learning in python*. *the Journal of machine Learning research*. 12 :2825–2830, 2011.

Félix Quinton. *Suivi de la rotation des cultures à partir de séries temporelles d'images satellite*, 2021.

