



RÉPUBLIQUE FRANÇAISE

*Liberté
Égalité
Fraternité*

Institut national de l'information
géographique et forestière



CHANGER
D'ÉCHELLE



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

IGN
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE

CHANGER
D'ÉCHELLE

LIDAR HD ET IA

De l'IA 3D au sein de la classification du nuage de points

Webinaire MAGIS du 5/12/2024 – Floryne Roche

Floryne.Roche@ign.fr

Sommaire

1. Intro et contexte

- a. Le projet Lidar HD
- b. Les grandes phases du chantier
- c. Classer le nuage de points
- d. Vue d'ensemble

2. Caractéristiques du modèle IA

- a. Architecture
- b. Métriques

3. Rome ne s'est pas bâtie en un jour

- a. Le sujet bâti.
- b. Les ponts
- c. Le sol
- d. Et puis après...

4. Bilan & Perspectives

1. Contexte

Deux mots d'intro sur le projet Lidar HD

- Le projet lidar HD pourquoi?

Optimiser la gestion de crise (inondation par ex.), estimer le volume de bois, connaître la position des obstacles aériens, ...

- Le projet lidar HD c'est quoi?

Couverture de la France (hors Guyane) en lidar, à 10 impulsions par m², en 5 ans, à des saisons variées, 9 classes de points. Novateur sur le territoire français !

- Le lidar HD, comment?

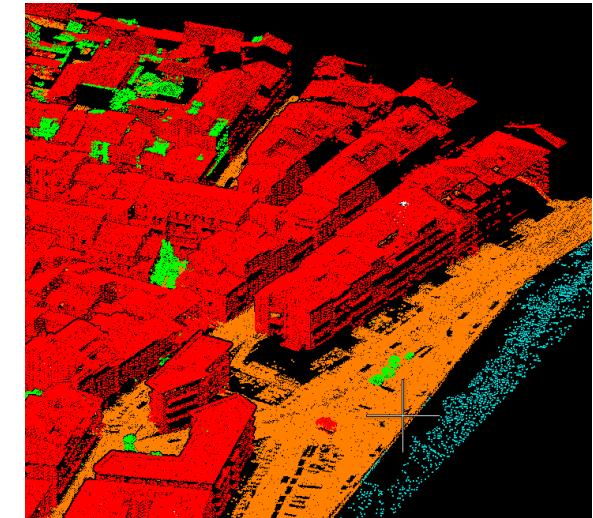
De la production en propre et de la sous-traitance, de la production et des contrôles pour partie automatique et pour partie manuels.

- Le projet lidar HD, ça embarque qui?

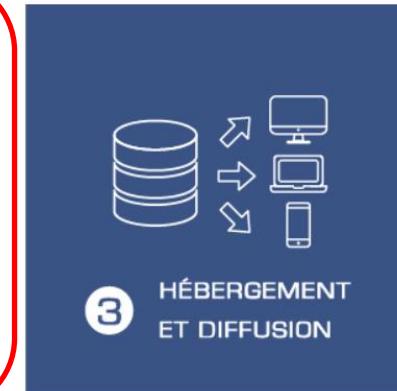
De nombreuses équipes à l'IGN, le MTECT, le MASA, la DGPR, des régions, des départements des collectivités, ...



Carte des données disponibles
[Téléchargement des produits dérivés | LiDAR HD - IGN](#)



Les grandes phases du chantier



Décrire en 3D l'occupation du territoire : classer le nuage

Classification du nuage de points en 9 classes:

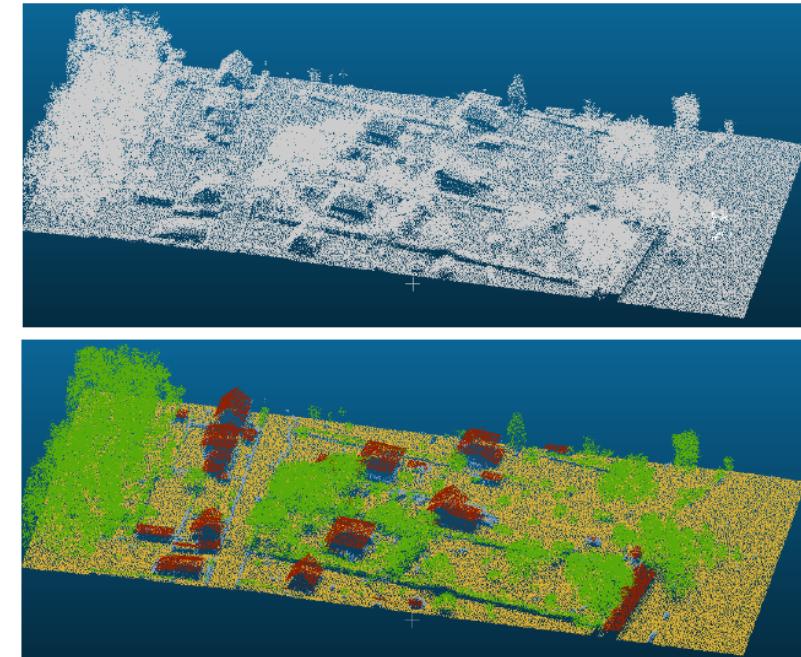
- Sol
- Végétation basse, moyenne et haute
- Eau
- Pont
- Bâtiments
- Sursol pérenne (Antennes, Eoliennes, Lignes électriques...)
- Artefacts

⇒ Définition du sol (MNT)

⇒ Identification des types d'objets présents sur le territoire
(visualisable en 2D)

⇒ ...

➤ **Comment classer ce nuage de points?**



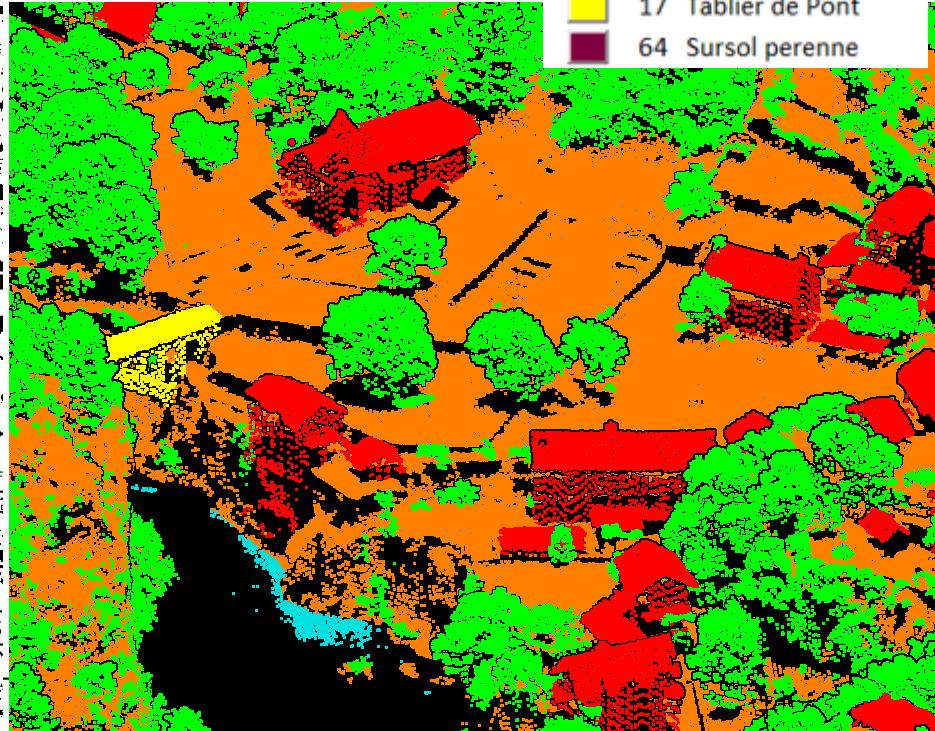
Nuage de point avant et après classification

La classification Lidar, exemple

Avant classification



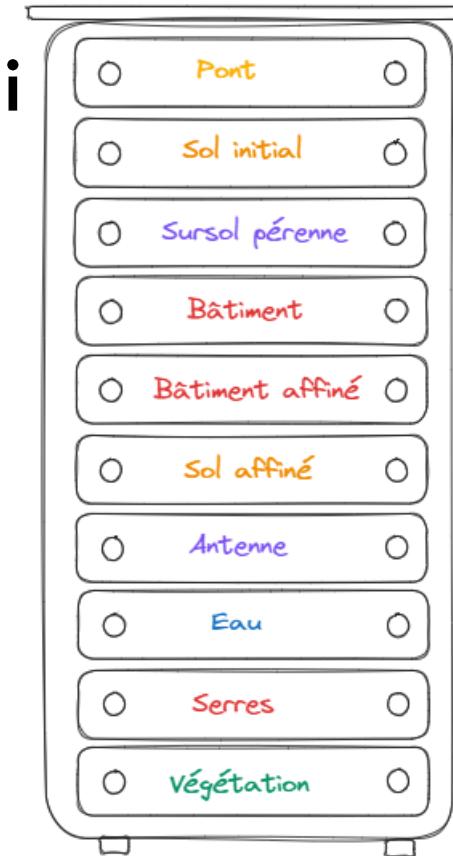
Après classification



1	Non classes
2	Sol
3	Vegetation basse
4	Vegetation interme
5	Vegetation haute
6	Batiment
9	Eau
17	Tablier de Pont
64	Sursol perenne

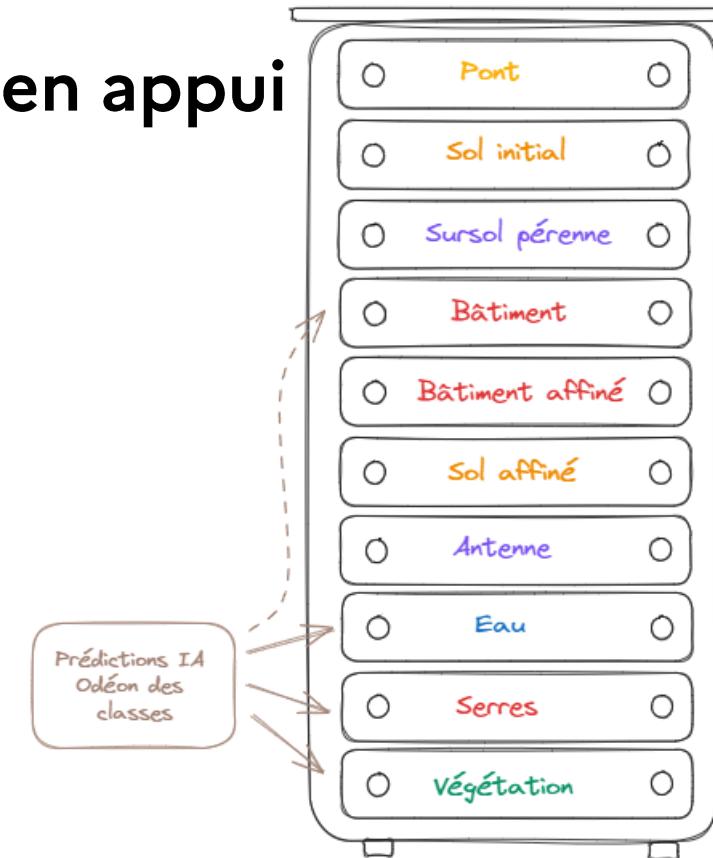
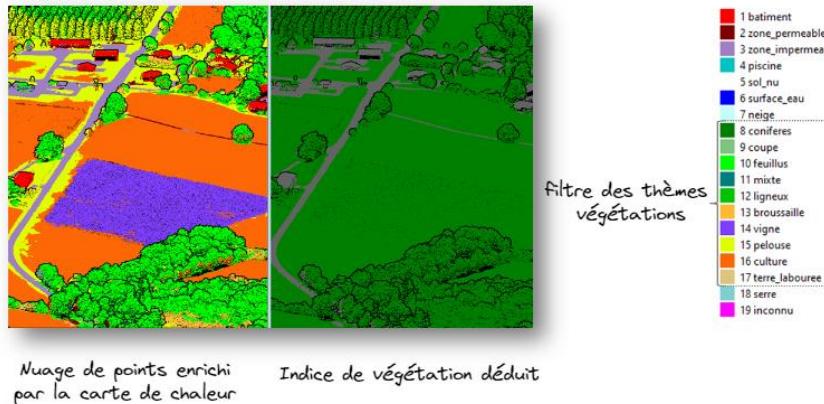
Un processus hybride : l'IA en appui

- Des **macro Terasolid**, avec des méthodes explicites de détection



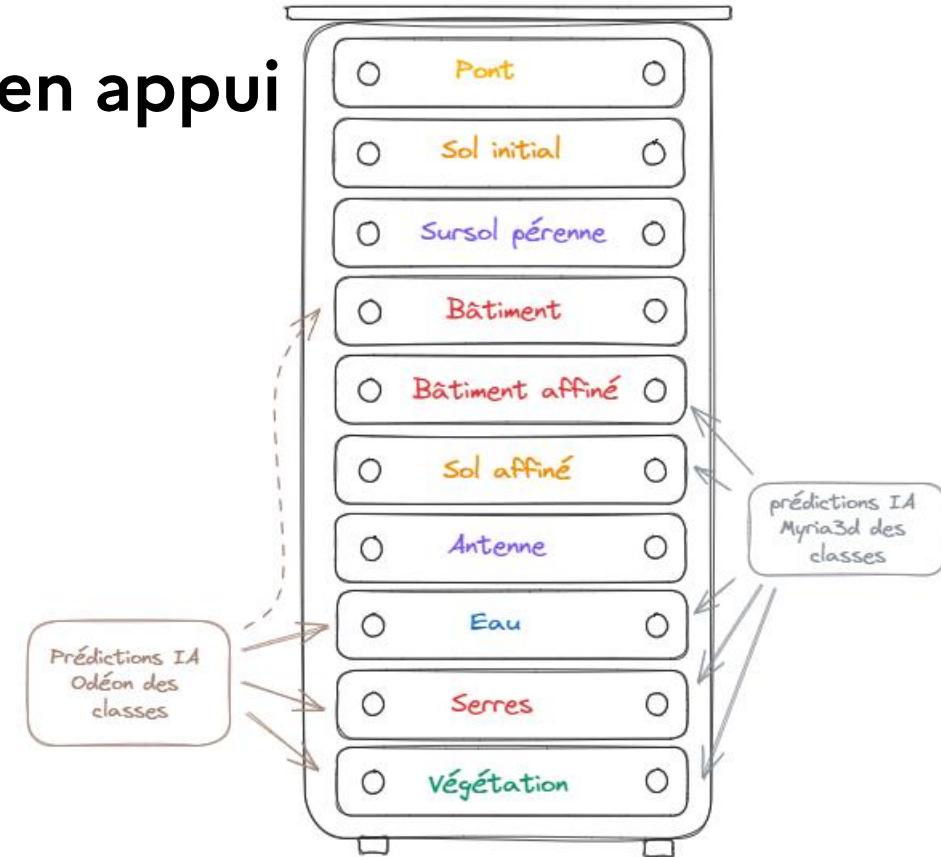
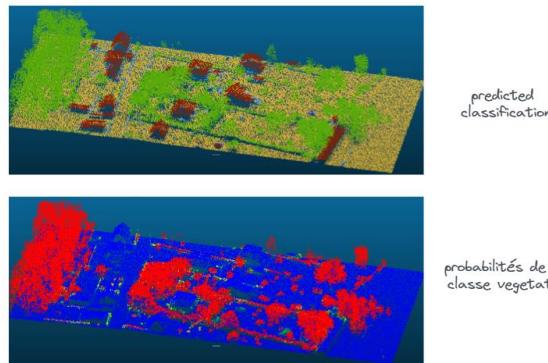
Un processus hybride : l'IA en appui

- Des **macro Terasolid**, avec des méthodes explicites de détection
- Des données issues de la **carte de chaleur de l'OCSGE**,
 - dont l'indice de végétation – permettant une amélioration du passage à l'échelle



Un processus hybride : l'IA en appui

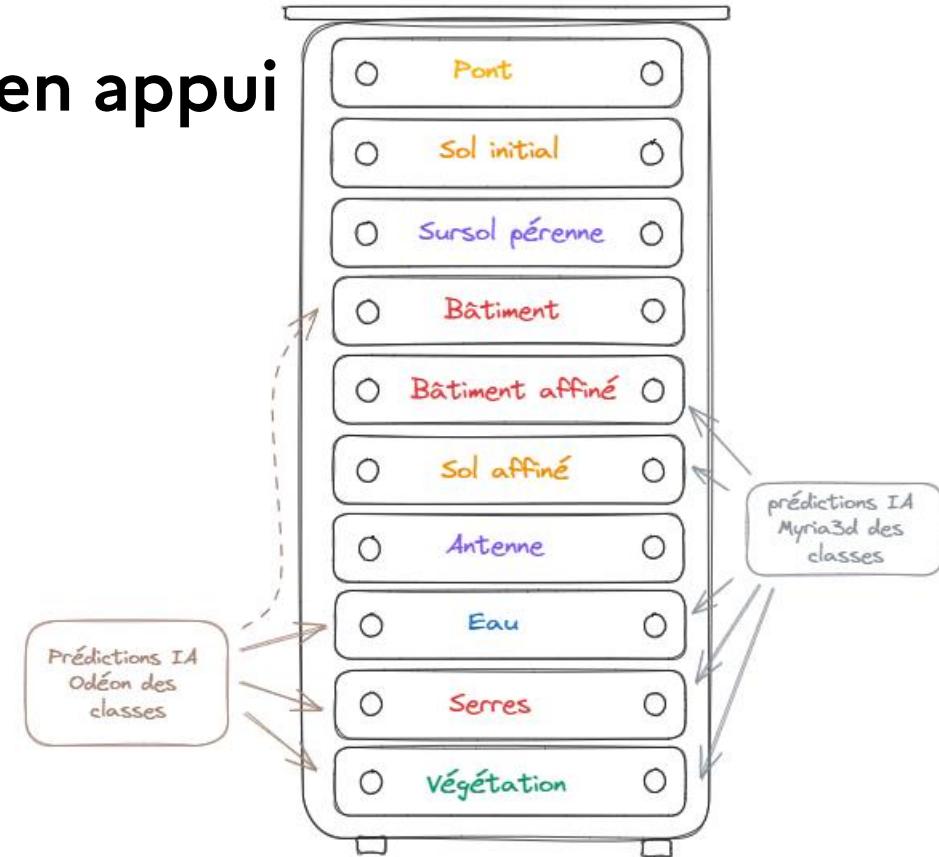
- Des **macro Terasolid**, avec des méthodes explicites de détection
- Des données issues de la **carte de chaleur de l'OCSGE**, - dont l'indice de végétation – permettant une amélioration du passage à l'échelle
- Des **prédictions IA 3D** sur le nuage de point pour améliorer ou conforter certaines classes



Un processus hybride : l'IA en appui

- Des **macro Terasolid**, avec des méthodes explicites de détection
- Des données issues de la **carte de chaleur de l'OCSGE**, - dont l'indice de végétation – permettant une amélioration du passage à l'échelle
- Des **prédictions IA** sur le nuage de point pour améliorer ou conforter certaines classes
- De la donnée **BD Topo**

Un **processus hybride** (algorithmes, donnée vecteur existantes, IA), donnant des meilleurs résultats qu'en s'appuyant sur une méthode unique.



2. Caractéristiques de l'IA

Myria3d: classer le nuage de points

Architecture

Données initiales

Jeu de données multiclasses Lidar HD

- 151 km², sur 3 départements
- Colorisation: RVB+IR
- 7 classes
- Processus automatique + Correction
- Bonne qualité (quelques erreurs + brouillard)
- Train/Val/Test: 130/10/10 km²

Note: échec du transfert de modèle depuis des données suisses

Juin 2022

Architecture du modèle

RandLa-Net

- Similaire à PointNet++
- Réécriture pour gérer des nuages de taille arbitraire
- Validé par une communauté open-source (Pytorch-Geometric)
- Accessible (Github)

Octobre 2022

Code

Myria3D

- Pytorch & Pytorch-Lightning
- Focus sur les spécificités du Lidar HD
- CI/CD (Tests+Docs+Docker)
- **Open Source:** [myria3d/README.md at main · IGNF/myria3d · GitHub](#)



En continu

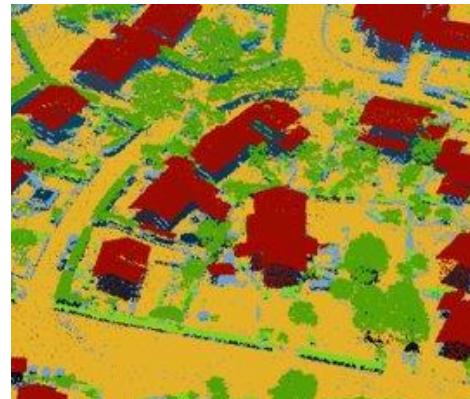
- Architecture du système conçue pour des mises à jour régulières
- Travail de choix des architectures de réseaux de neurones qui a bénéficié de la proximité **d'équipes de recherche très actives sur la classification 3D par deep learning**
- Davantage de données d'entraînement seront disponibles au fur et à mesure de l'avancement de la production du Lidar HD
- Publication anticipée du modèle pour les utilisateurs de Lidar HD (Github).

Métriques

									Classes rares
		Unclassified	Ground	Vegetation	Building	Water	Bridge	Antennas+Others	
TRAIN	Proportion (% of pts.)	7.7%	46.3%	43.6%	2.1%	0.2%	0.03%	0.07%	
VAL	Proportion (% of pts.)	8.4%	40.1%	48.8%	2.6%	0.00%	0.01%	0.04%	
TEST	Proportion (% of pts.)	9.9%	50.6%	38.9%	0.6%	0.03%	0.01%	0.01%	
	Class IoU	35%*	87%	98%	93%	18%	17%	58%	
	Mean IoU				57.8%				

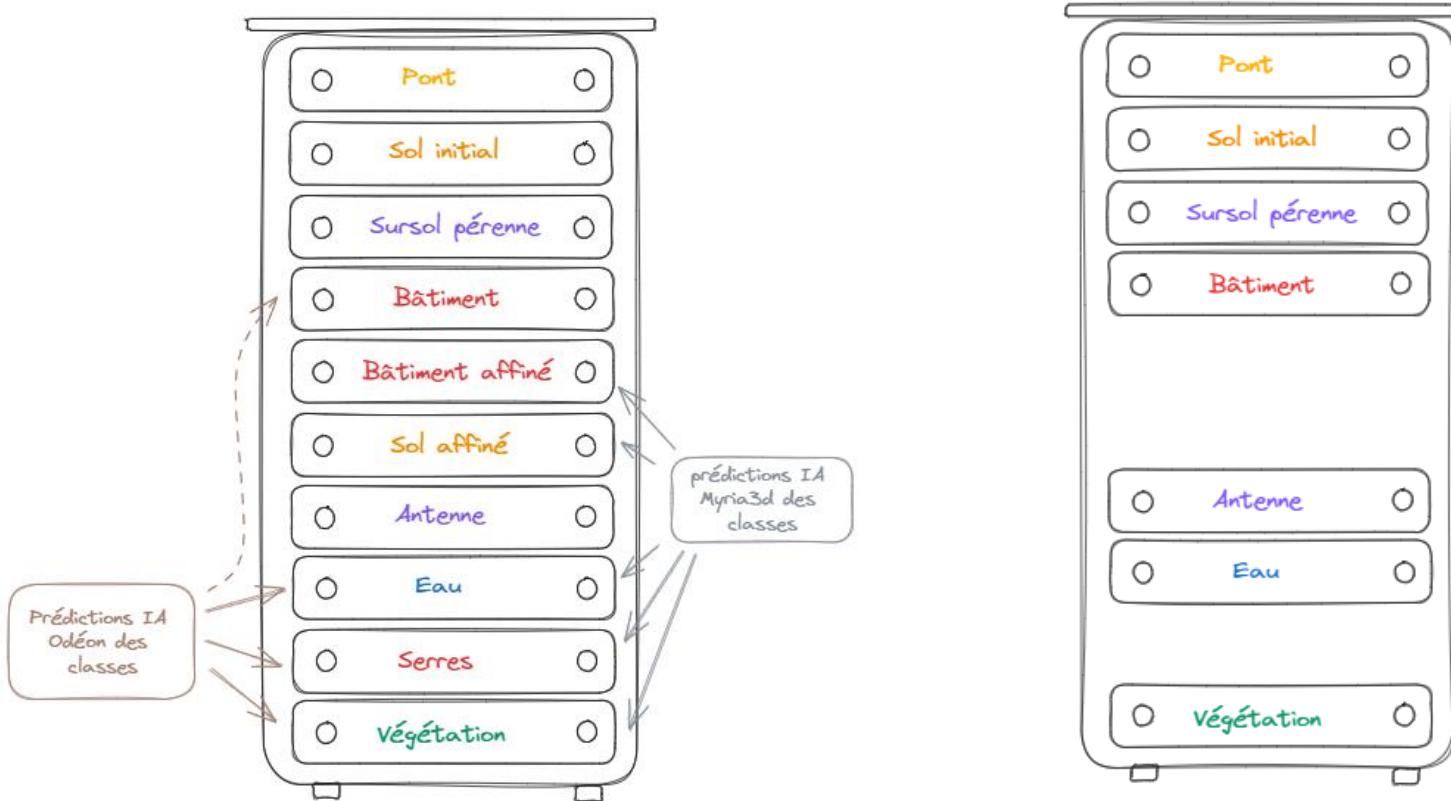
→ Bonnes performances sur les classes *Bâtiment*, *végétation*, *Sol* (mIoU= 93%)

True labels



Predicted labels







3. Et Rome ne s'est pas bâtie en 1 jour...

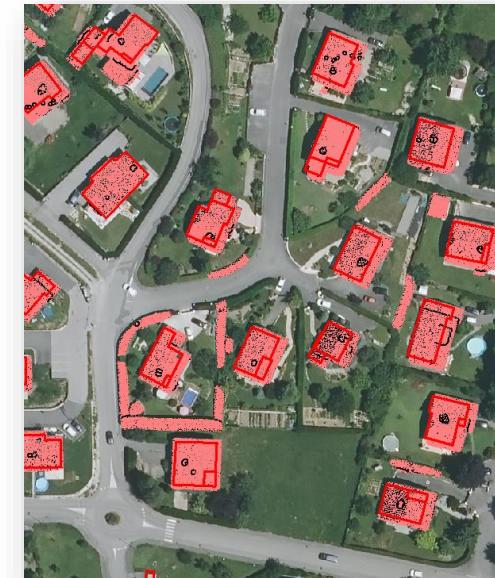
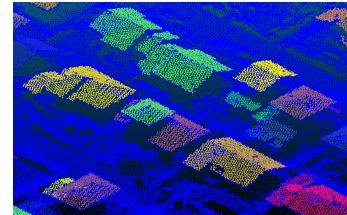
Au commencement, les bâti (sept 2021)

Constat initial

- Un enjeu fort
- Une inspection humaine chronophage et fastidieuse
- De la donnée d'apprentissage
- Périmètre délimité, gain quantifiable, possibilité d'extension du périmètre, positionnement « tout bénéf » dans la chaîne de production

Orientations de conception:

- Un module s'insérant entre la partie automatique et la partie manuelle, pour **cibler les reprises opérateurs.**
 - Reduction du **temps opérateur**
 - Limitation de la **partie triviale** de la tâche
 - **Prise en compte du processus** dans la conception de la solution.



Module bâti

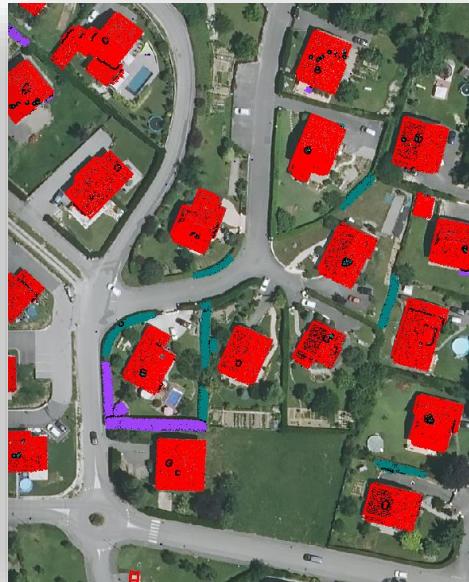


Faux positifs typiques : haies, camions, zones de fortes pentes.

En rose: bâtiment identifié dans TScan



Module bâti

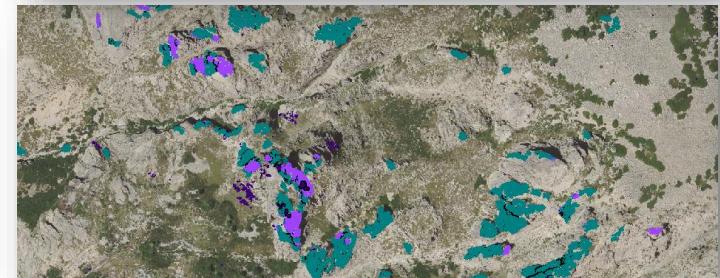


Faux positifs typiques : haies, camions, zones de fortes pentes.

Rouge = validé

Vert = réfuté

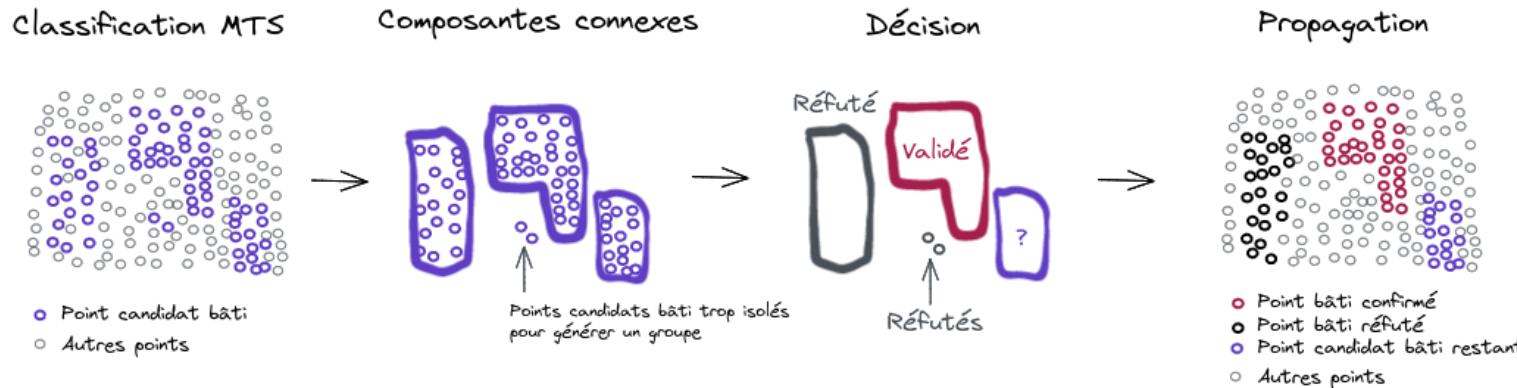
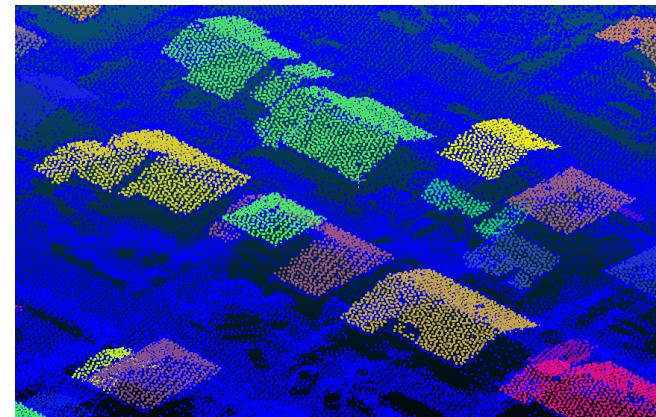
Violet = Autres



Le module : grandes lignes

Les grandes lignes du fonctionnement du module:

- Regroupement des points candidats bâtis par composantes connexes.
- Décision pour chaque groupe de points.
- Propagation aux points.



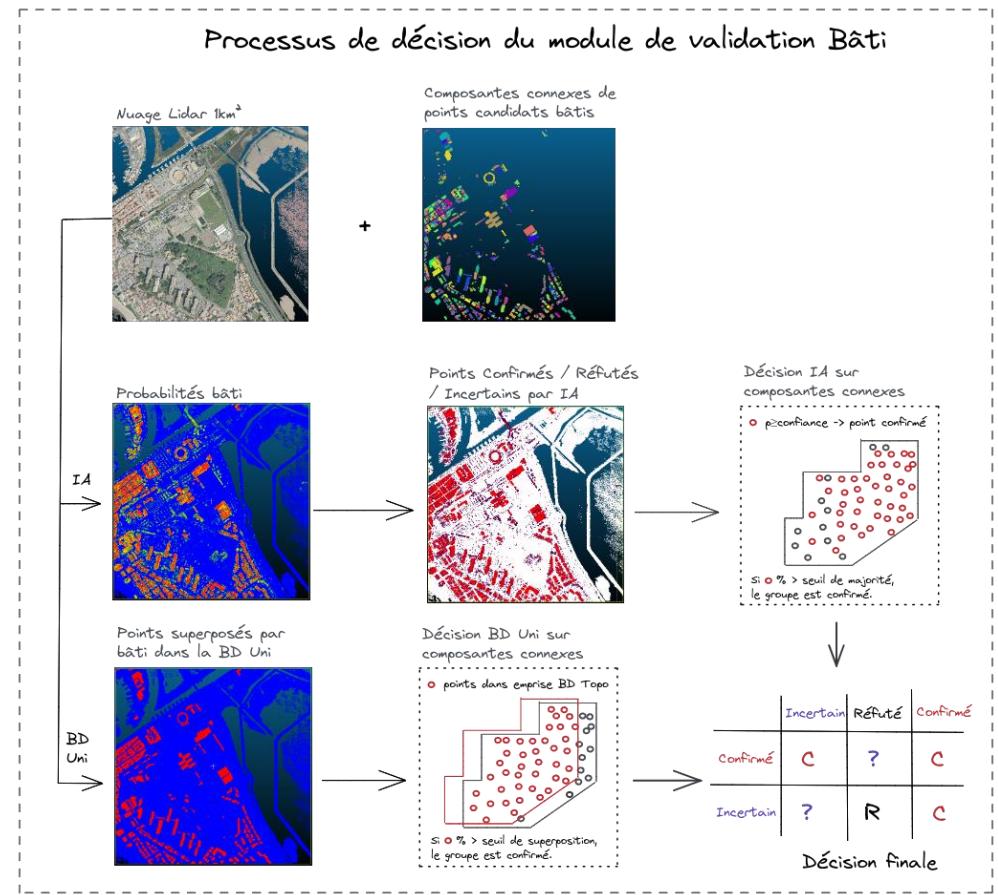
Le module: décisions

Décisions s'appuyant sur **deux sources de vérité**:

- Nuage de points traité par l'IA produisant des probabilités d'être du bâti
- Les emprises de la BD Uni
-> Prise de décision conjointe

Décisions s'appuyant sur des **seuils**:

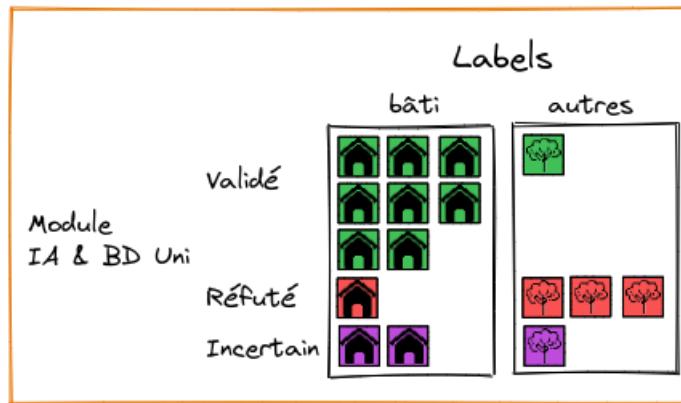
- Seuils de probabilité pour confirmation et réfutation des points
- Seuils de majorité proportionnelle pour confirmation ou réfutation d'un groupe de points via l'IA, confirmation d'un groupe de points via BD Uni.



Le module: optimisation

Les 5 seuils de décision sont optimisés pour **maximiser l'automatisation**, sous contrainte.

Décisions et erreurs



Contraintes et critère à maximiser

$$\text{Précision} = \frac{\text{green buildings} + \text{purple buildings}}{\text{green buildings} + \text{purple buildings} + \text{green trees}} \geq 98\%$$

$$\text{Rappel} = \frac{\text{green buildings} + \text{purple buildings}}{\text{green buildings}} \geq 98\%$$

$$\text{Automatisation} = \frac{\text{green buildings} + \text{red buildings}}{\text{green buildings} + \text{red buildings} + \text{purple buildings}}$$

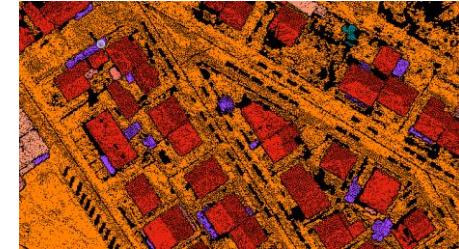
Synergie des 2 sources de données du fait de l'**optimisation conjointe**:

- L'IA compense les manquements de la BD Uni.
- La BD Uni rattrape les réfutations à tort du modèle IA. → Possibilité de moindre exigence IA en réfutation permettant d'augmenter le taux d'automatisation sans perte de qualité.

Les résultats (janvier 2022)

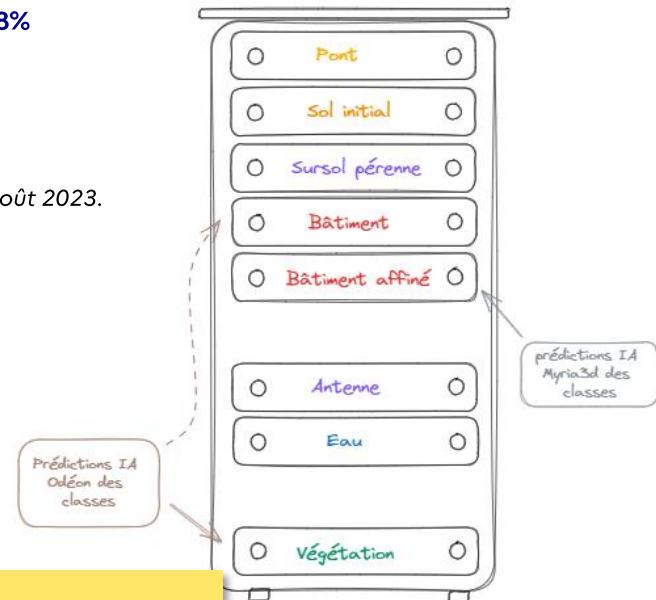
Résultats

- Automatisation à 85% des groupes de points, avec un **taux de rappel et de précision > 98%**
- Des améliorations possibles/anticipées/constatées sur:
 - Le nombre de groupe de points restant à inspecter (amélioration de la macro)
 - Les situations typiques d'erreurs (nouvelles données)
 - Le modèle lui-même (amélioration continue, modèle multi-classes) -> 96% d'automatisation en août 2023.



Décision (janvier 2022)

- **GO** : Passage à l'échelle, utilisation du module dans la production prototype.
- **Amélioration continue.** (passage d'un modèle spécifique à un modèle multiclassé)
- *Investiguer l'utilisation du module pour le contrôle de sous-traitance.*
- *Investiguer l'utilisation du module pour la sous-détection.*



Qui?

- Données de la production
- Conception : équipe IA

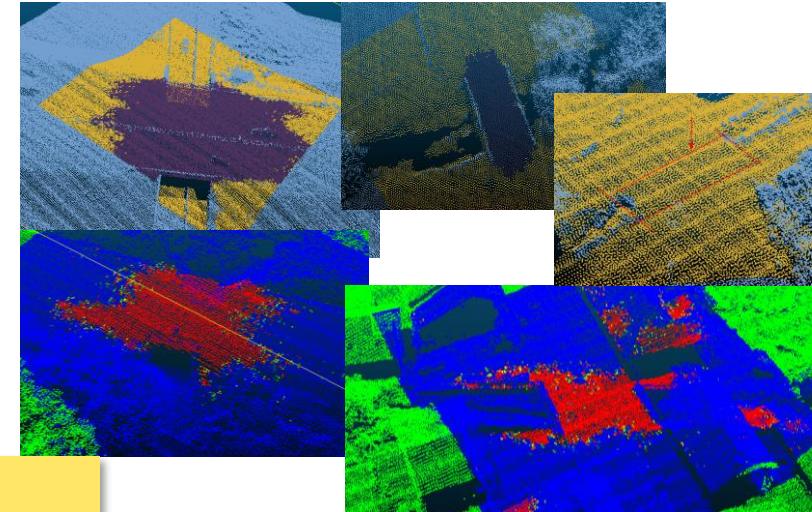
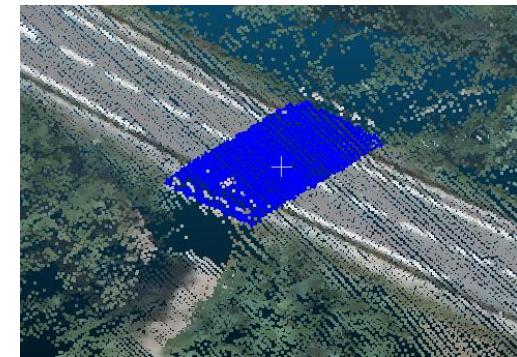
Le problème des ponts (été 2022)

- **Le besoin:** délimiter automatiquement les ponts
- **La données :** ~90 ponts exploitables classés par la production
- **La proposition:** différentes formulation de tâches (nuage classé, point de rupture, boîte englobante orientée)
- **Le résultat :** erreurs persistantes, incompatibles avec la qualité de délimitation de pont souhaitée.

⇒ Parfois, l'IA, ça ne fonctionne pas.

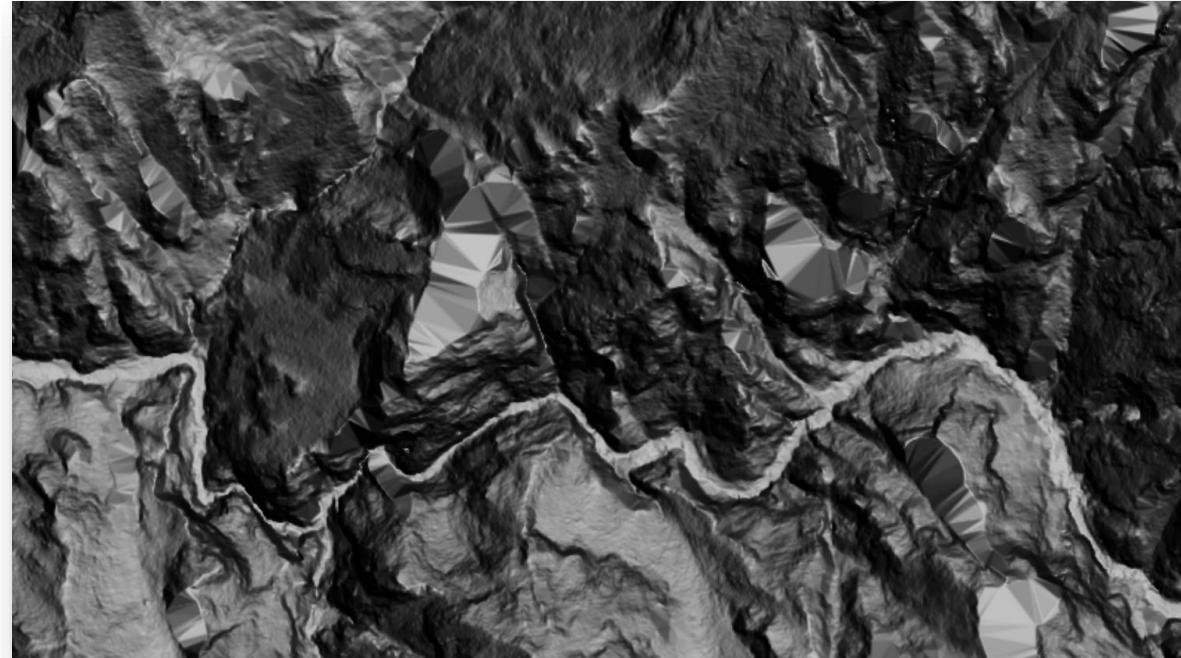
Qui?

- Annotations spécifiques de la production
- Conception : équipe IA



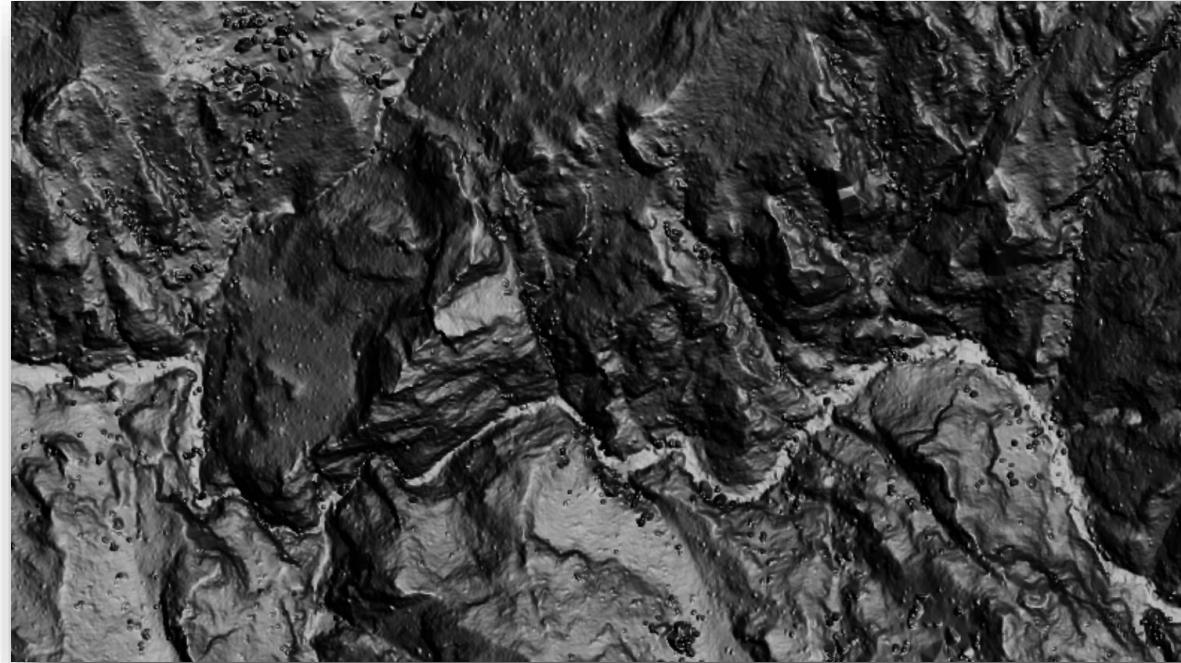
L'étonnante complémentarité de l'IA et des macros sur le sol

Méthode classique seule



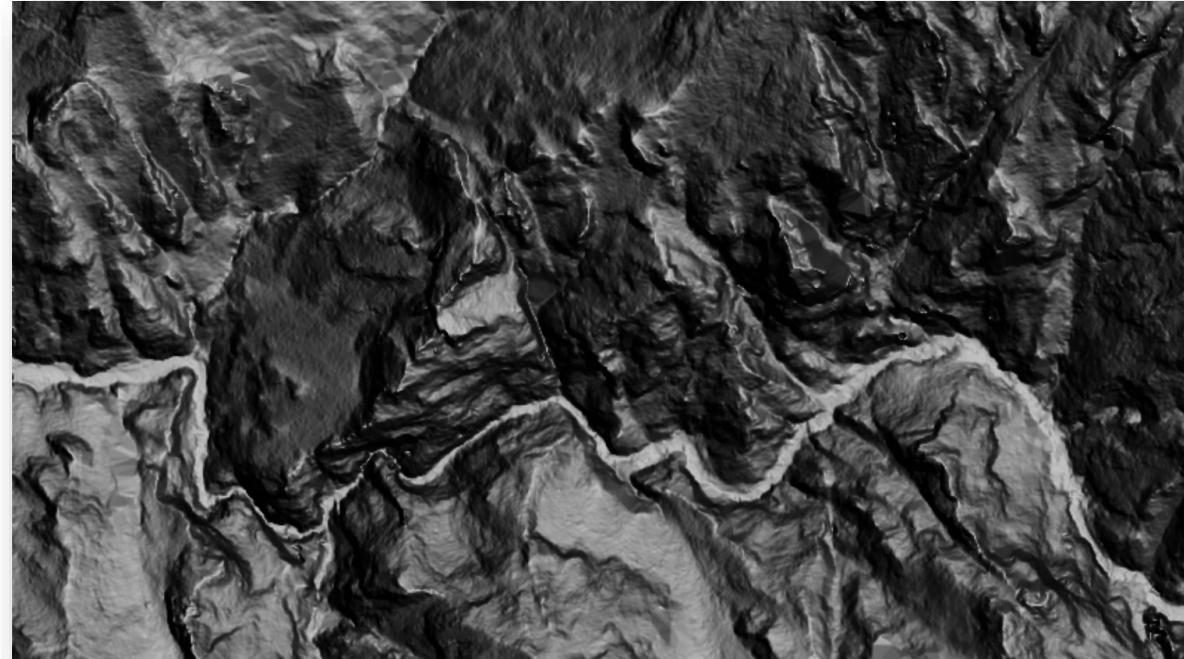
L'étonnante complémentarité de l'IA et des macros sur le sol

IA seule



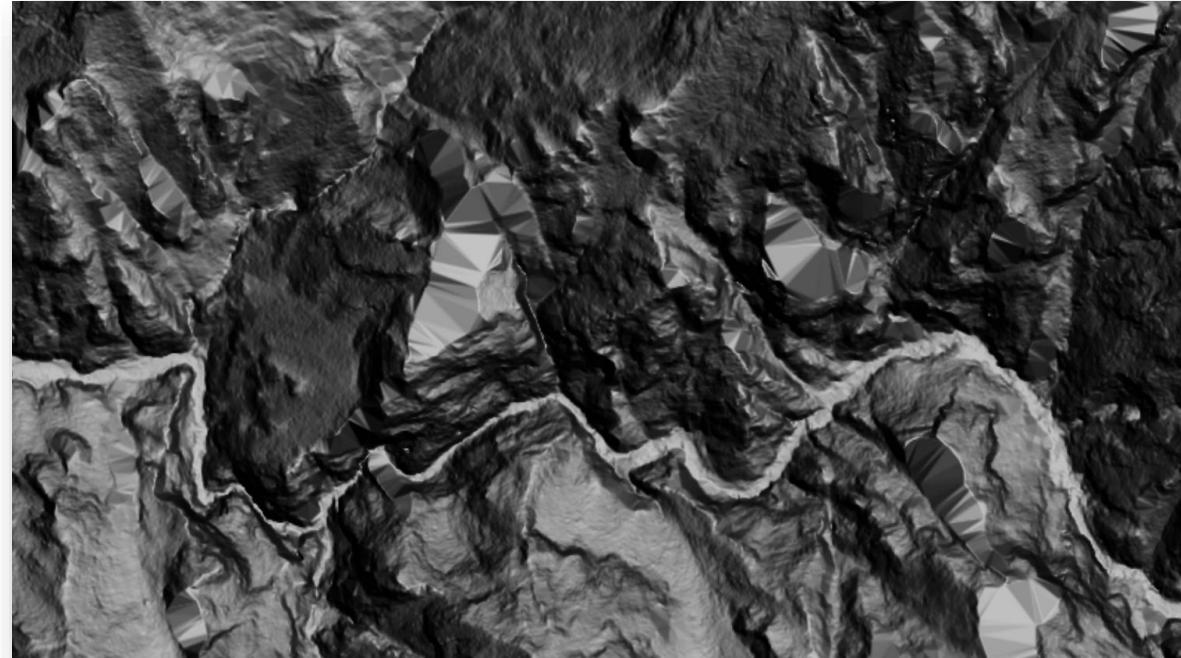
L'étonnante complémentarité de l'IA et des macros sur le sol

Hybridation des traitements:
Classique + IA



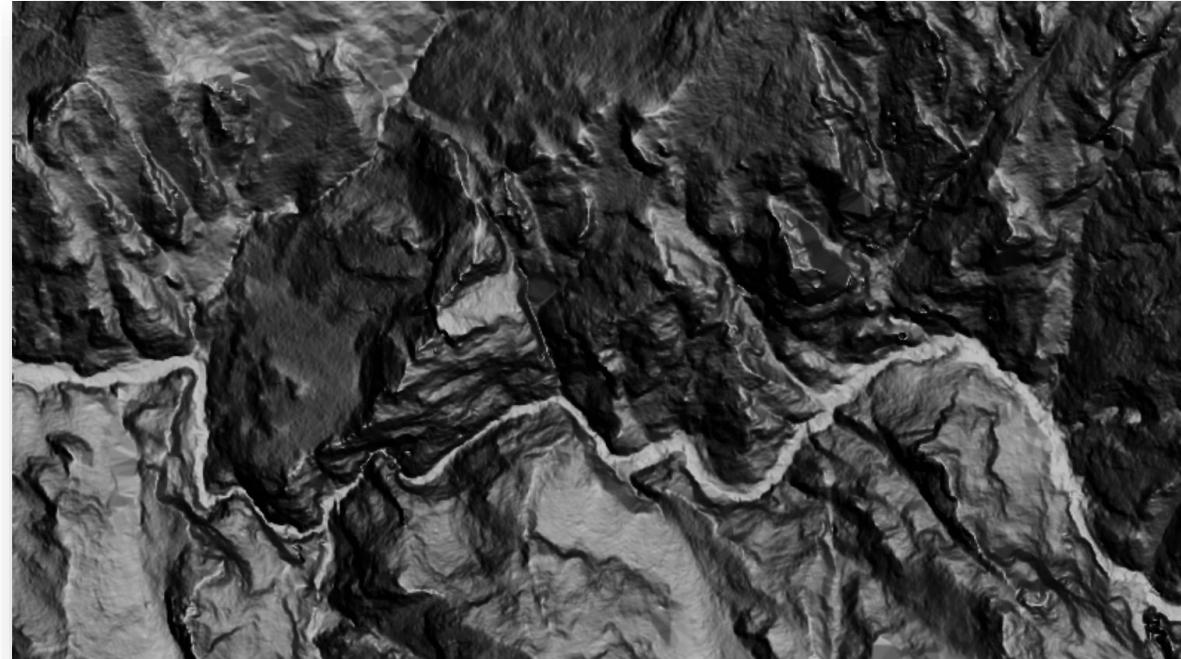
L'étonnante complémentarité de l'IA et des macros sur le sol

Méthode classique seule



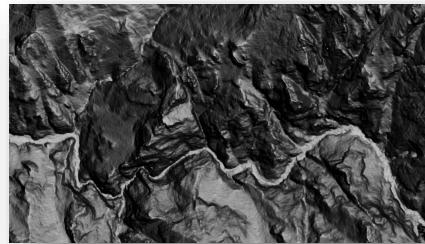
L'étonnante complémentarité de l'IA et des macros sur le sol

Hybridation des traitements:
Classique + IA



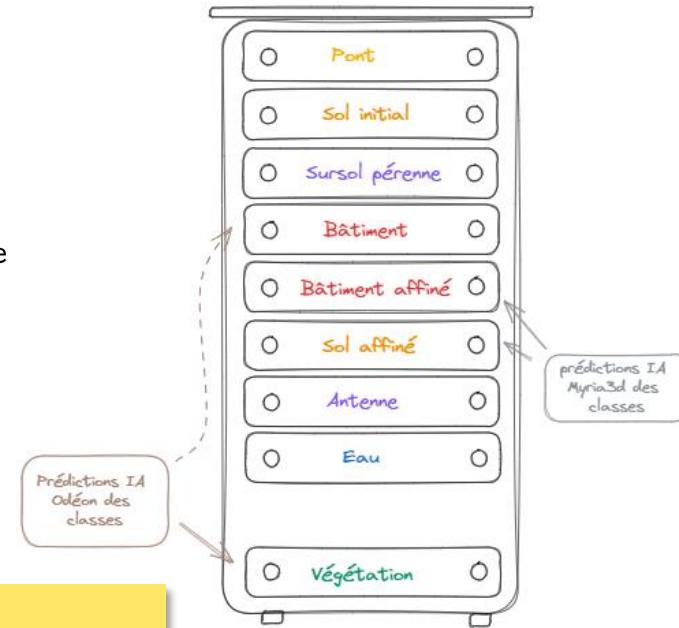
L'étonnante complémentarité de l'IA et des macros sur le sol

- **Le besoin:** améliorer la détection du sol
- **La données :** 150 km² répartis sur 3 départements, dont les gorges du Tarn
- **La proposition:** utiliser les données IA là où la densité de points sol est insuffisante
- **Le résultat :** Complémentarité des 2 sources de points sol.
- **Gain :** infini



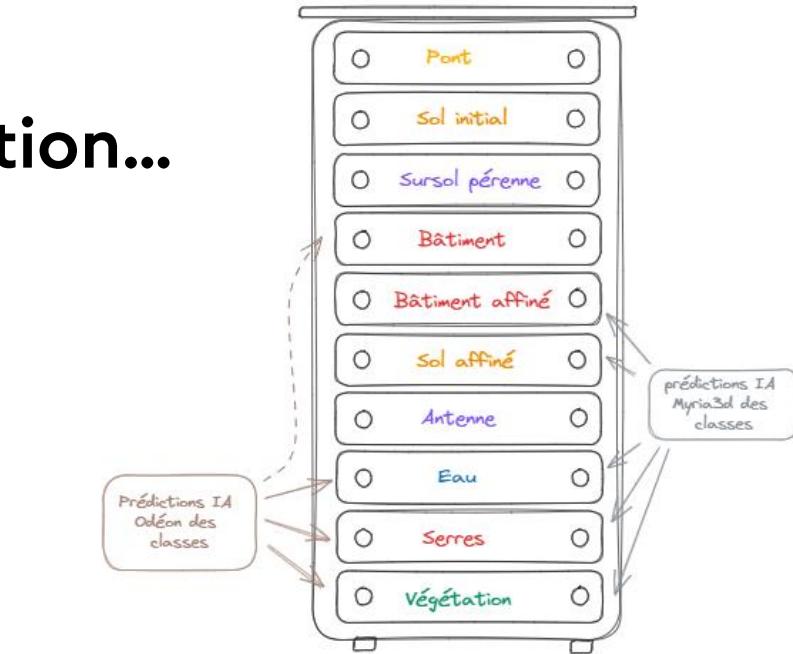
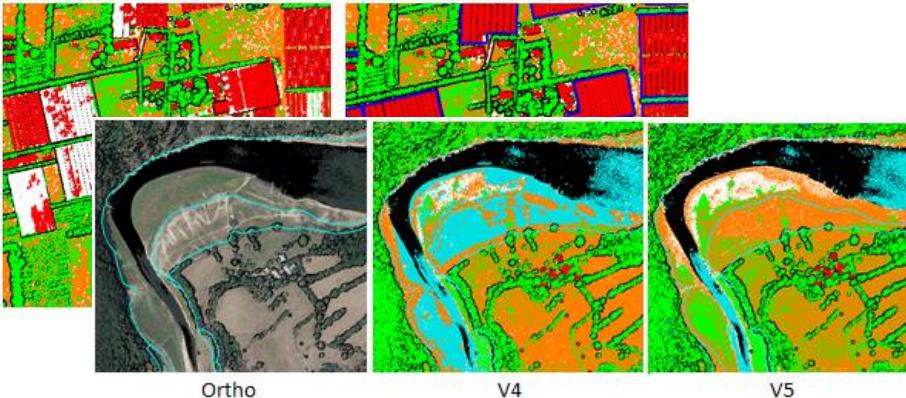
Qui?

- Données de la production + équipe IA
- Conception : équipe processus



Et encore: serres, hydro, végétation...

- **Le besoin:** améliorer la classification des serres, de l'hydrographie, de la végétation
- **La données :** Ø (modèles déjà entraînés)
- **La proposition:** mix de méthodes explicites, s'appuyant sur la sortie d'Odéon et celle de Myria3d
- **Le résultat :** des améliorations flagrantes



Qui?

- Données automatiques, modèles entraînés
- Conception : équipe processus

4. Bilan & Perspectives

Bilan & perspectives

Bilan

- L'IA améliore les résultats ! Et ça fonctionne, c'est en production! *Mais pas sur tout (ponts) ou pas partout (sol)*
- Un positionnement de l'IA Myria3d « **en complémentarité** »
 - des autres méthodes (explicites),
 - des autres données (BD Topo, carte de chaleur de l'OCSGE),
 - en appui sur/du savoir-faire (Lidar, macro TerraSolid)
- Une insertion de l'IA dans le processus **par petit pas** : en fonction de la donnée disponible, des problématiques soulevées, des résultats, avec une évolution des équipes qui l'opèrent.
- Des codes disponibles en open source.

Perspectives

- De la donnée d'apprentissage disponible. Réduire, concentrer, mettre un focus (cf. Partie II)
- Segmenter les essences forestières (Travail achevé)
- *Faire évoluer l'architecture : Super Point Transformer (non traité)*



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*

IGN
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE

CHANGER
D'ÉCHELLE

DES QUESTIONS SUR CETTE 1^{ÈRE} PARTIE?