



UNIVERSITÉ  
LAVAL

Faculté de foresterie, de géographie  
et de géomatique

Fonds de recherche  
sur la nature  
et les technologies

Québec



Sécurité publique

Québec



CENTRE DE RECHERCHE

EN DONNÉES ET INTELLIGENCE  
GÉOSPATIALES

# Simulation d'acquisitions LiDAR aéroportées en milieu forestier pour l'entraînement de réseaux de Deep Learning

---

Olivier STOCKER<sup>(1)</sup>, Éric GUILBERT<sup>(1)</sup>, Jean-François CÔTÉ<sup>(2)</sup>, Antonio FERRAZ<sup>(3)</sup>, Thierry BADARD<sup>(1)</sup>

(1) Université Laval

(2) Ressources naturelles Canada

(3) Jet Propulsion Laboratory NASA

TRIDIFOR – 20 SEPTEMBRE 2022

# Sommaire

---



PROBLÉMATIQUE



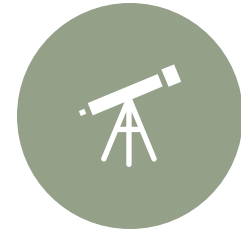
MÉTHODOLOGIE



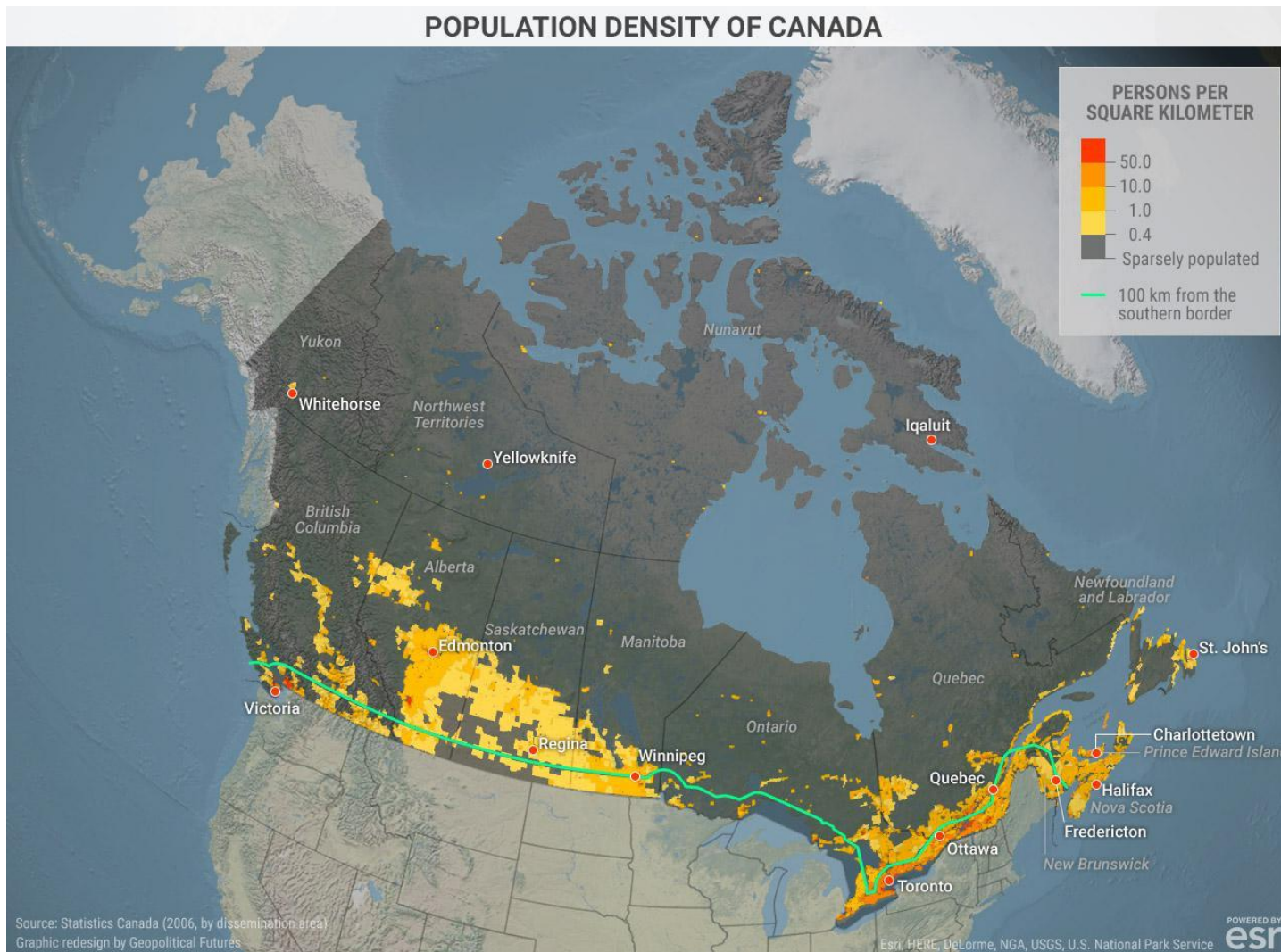
MODÉLISATION



SIMULATION



PERSPECTIVES



Source : [reddit](#)

## Suivi du territoire

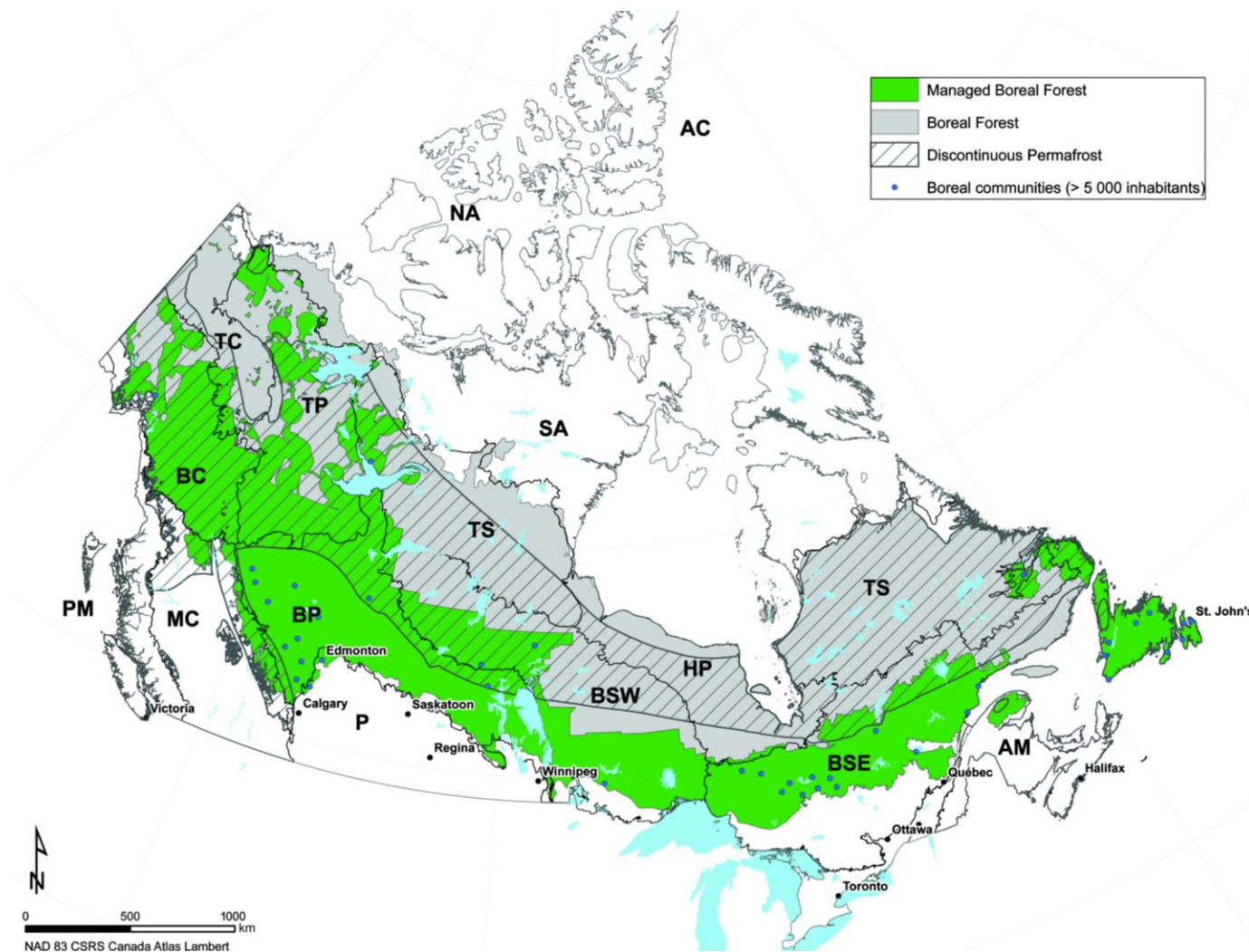
- Population canadienne inégalement répartie sur le territoire



## Suivi du territoire

- Population canadienne inégalement répartie sur le territoire
- Grande quantité de **territoires non aménagés**

Source : [1]



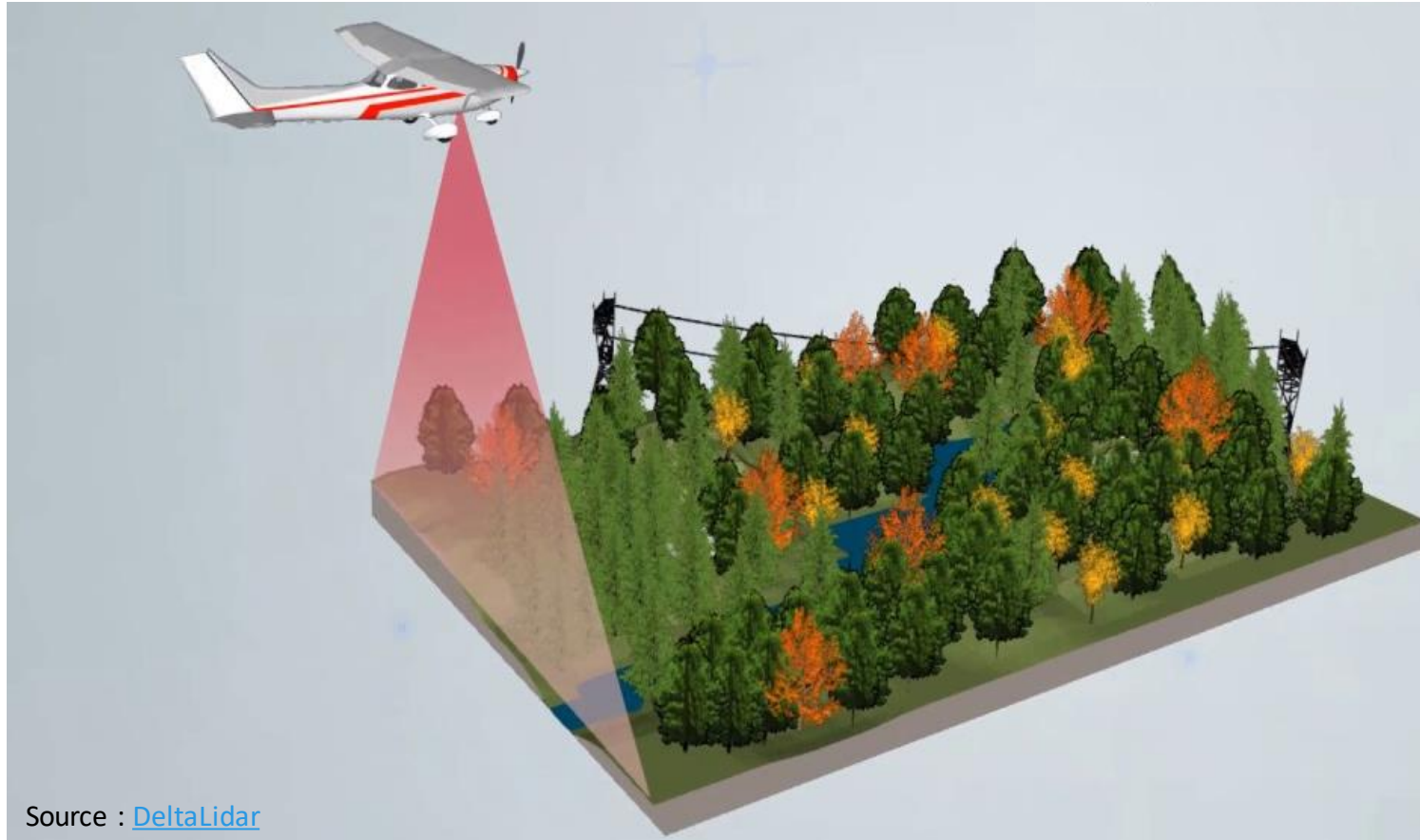
Source : [2]

## Suivi du territoire

- Population canadienne inégalement répartie sur le territoire
- Grande quantité de **territoires non aménagés**
- Grande quantité de **forêts non entretenues**

**Besoin de suivi  
de l'évolution  
du territoire**





Source : [DeltaLidar](#)

## Suivi du territoire

LiDAR aéroporté pour suivi des milieux forestiers :

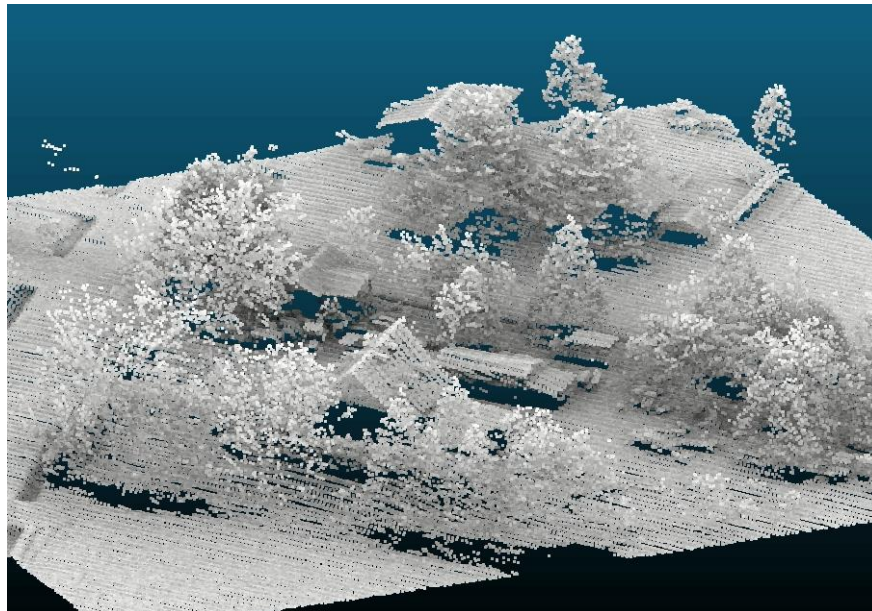
- Vecteur : ratio optimal entre couverture et résolution
- Capteur : pénètre la végétation et données informatives [Fassnacht et al., 2016]
- Multitude d'applications forestières



# Besoin d'automatisation

## Nuage de points LiDAR :

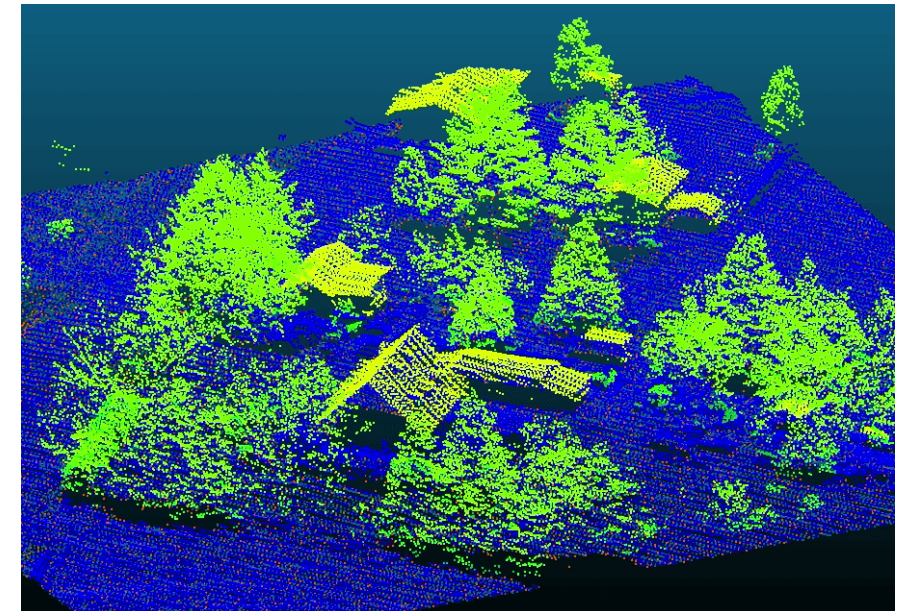
- Informations géométriques des *surfaces*
- Informatiquement *lourd*
- Non sémantisé



Sémantisation

## Sémantisation :

- Difficulté de compréhension et de sélection
- Grande quantité de données



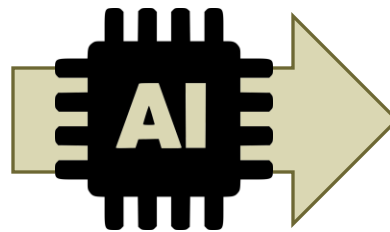
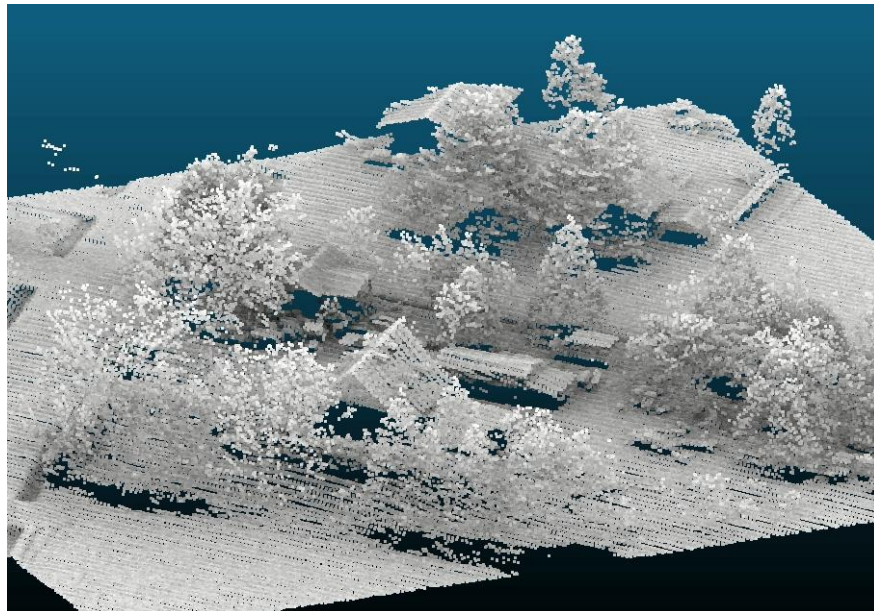




# Besoin d'automatisation

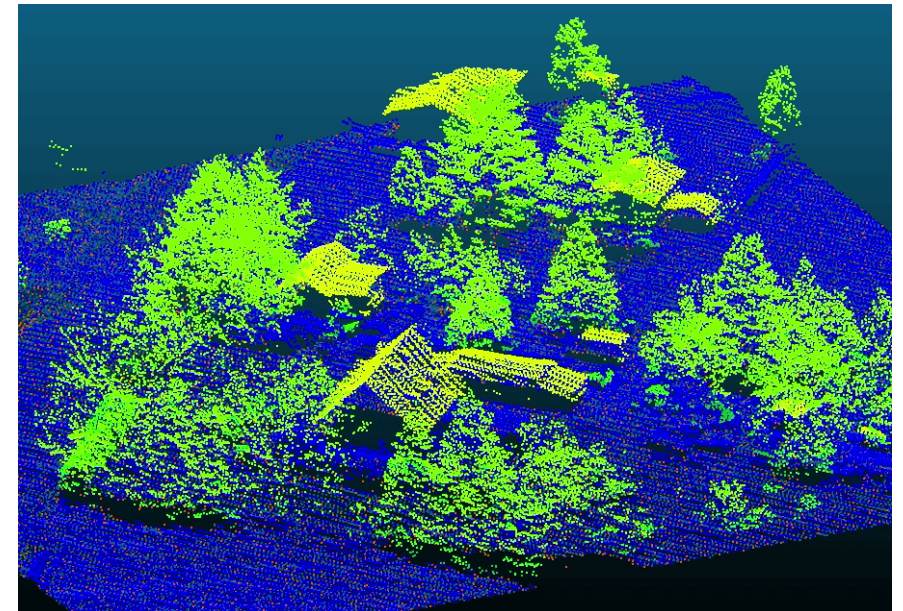
## Nuage de points LiDAR :

- Informations géométriques des *surfaces*
- Informatiquement *lourd*
- Non sémantisé



## Sémantisation :

- Difficulté de compréhension et de sélection
- Grande quantité de données
- Utilisation du **Machine Learning**







# Problématique : manque de vérité terrain

Deep Learning sur nuage de points :

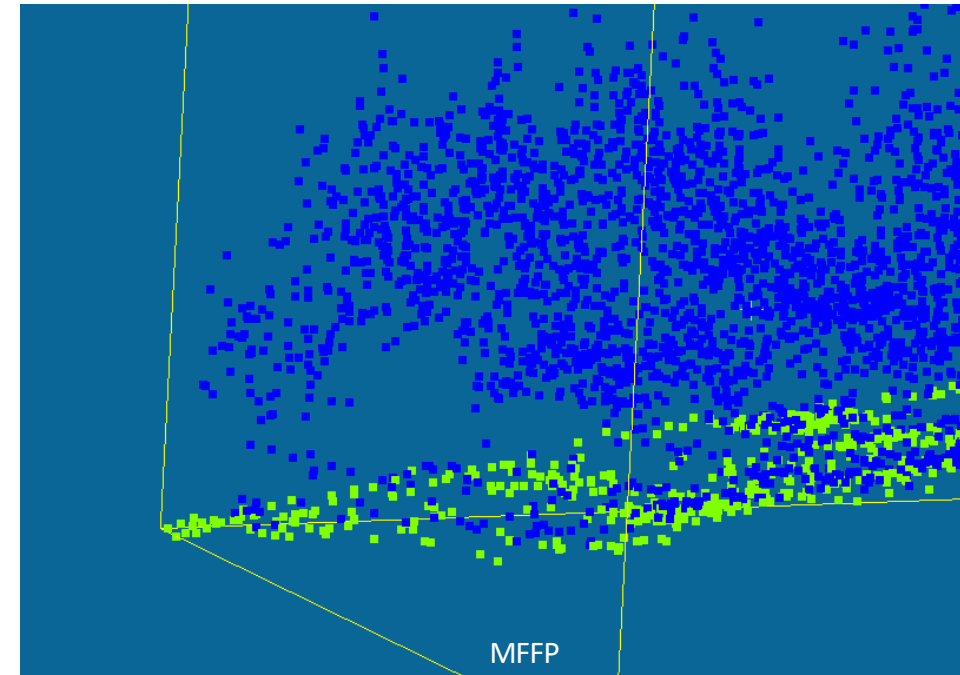
- Supérieur aux approches de ML classiques
- Nécessite grande quantité d'annotations



Paris-Lille 3D, Source : [4]

...en milieu forestier :

- Annotations **pauvres** et de **faible qualité**
- Environnement 3D complexe





# Solution envisageable

## Simulation de vérité terrain :

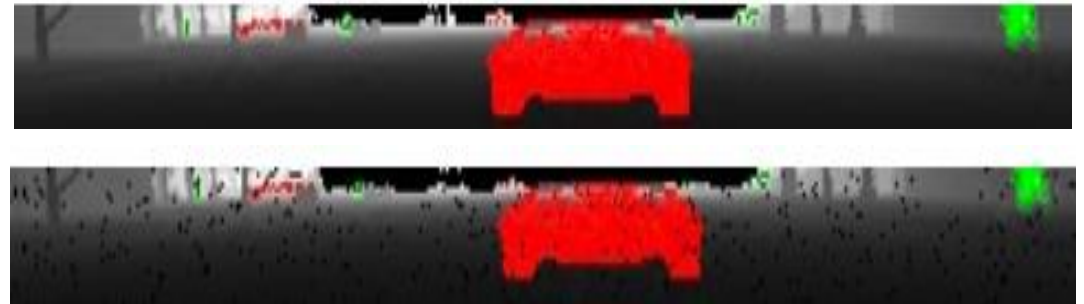
- [Wu et al., 2015] : Utilisé couramment
- [Dosovitskiy et al., 2015] : Transfert possible, même avec modélisation éloignée
- [Zhao et al., 2020b] : Simulation LiDAR



Source : [5]



Source : [6]



Source : [7]



# Simulation pour le DL

Critères de la simulation :

1. Réduction de l'écart entre distribution statistique **simulée** et **réelle**
2. Maximisation de l'informativité de la simulation

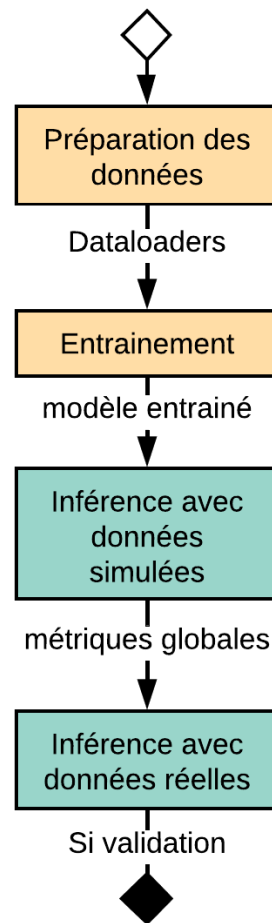
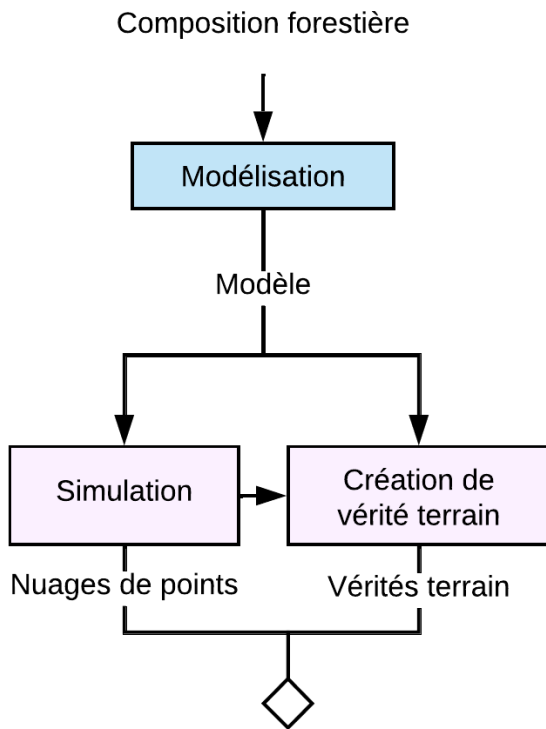


Source : Philippe Giguère



# Axe de recherche

**Correction du manque de données** d'apprentissage par **simulation** de données **LiDAR ALS** pour l'**entraînement** de réseau de **DL**



# Méthodologie

1. Modélisation
2. Simulation
3. Entrainement
4. Validation des modèles



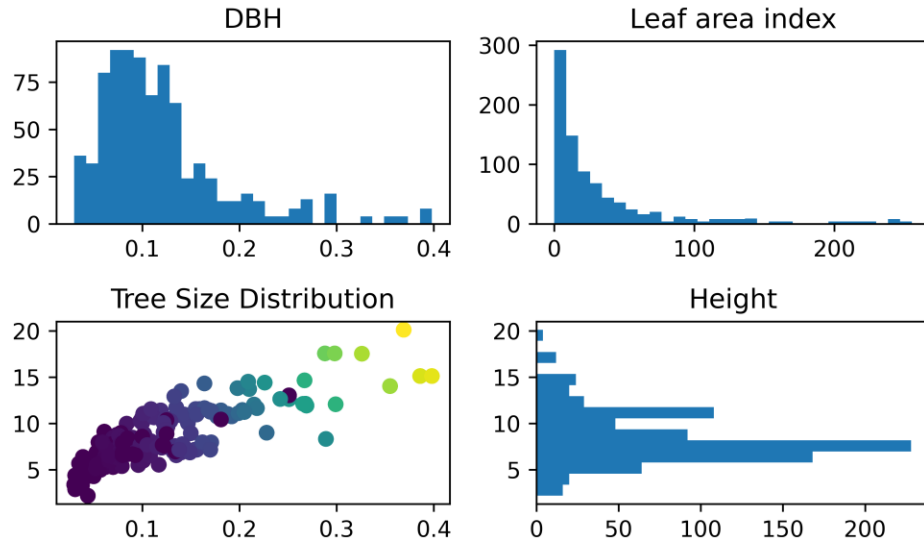
# Modélisation

---

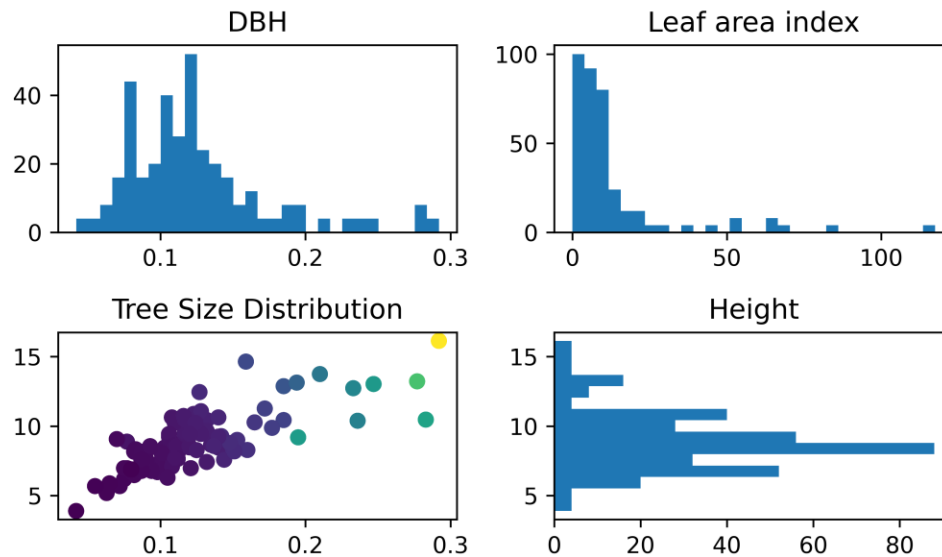
1. Choix de la composition forestière
2. Modèles d'arbres
3. Modélisation de la structure spatiale de la forêt
4. Modélisation du sol



### Abies Balsamea n=833



### Picea Mariana n=364



# Choix de composition

Composition forestière : Pessière noire à mousses

1. Epicéa Noir : *Picea Mariana*
2. Sapin Baumier : *Abies balsamea*

Bibliothèque d'arbre : Jean François Côté

- 1197 arbres
- 827 feuillages non nuls



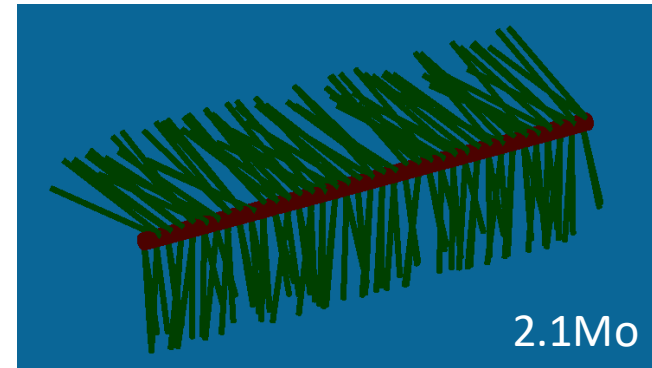
# Modèles d'arbres

## Origine :

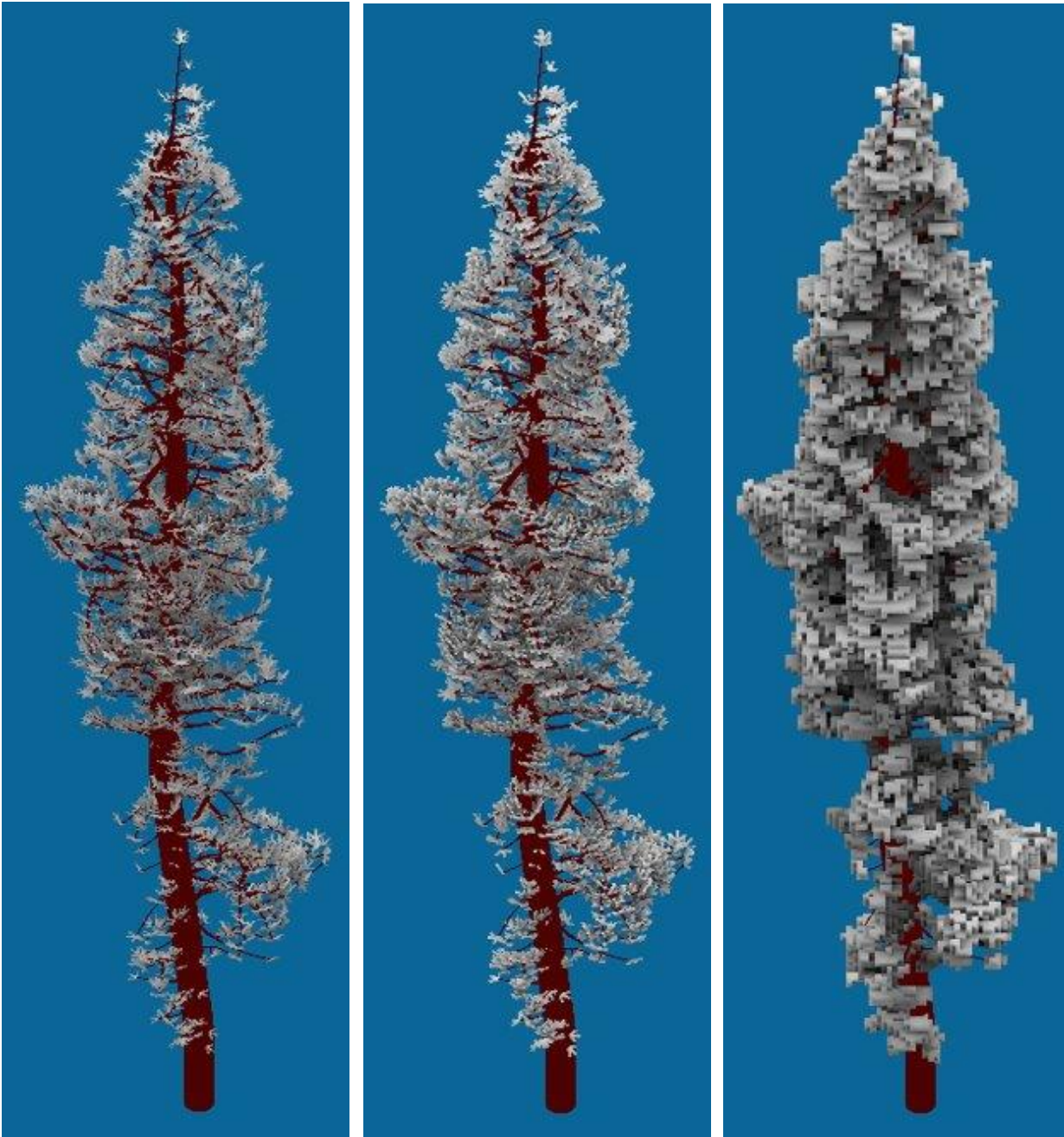
- Modélisation à partir de TLS
- Estimation du **feuillage** via **intensité**

## Stockage :

- Structure ligneuse [.obj]
- 5 types de **pousses** par **espèces** [.obj]
- Tableau de **position** des **pousses**



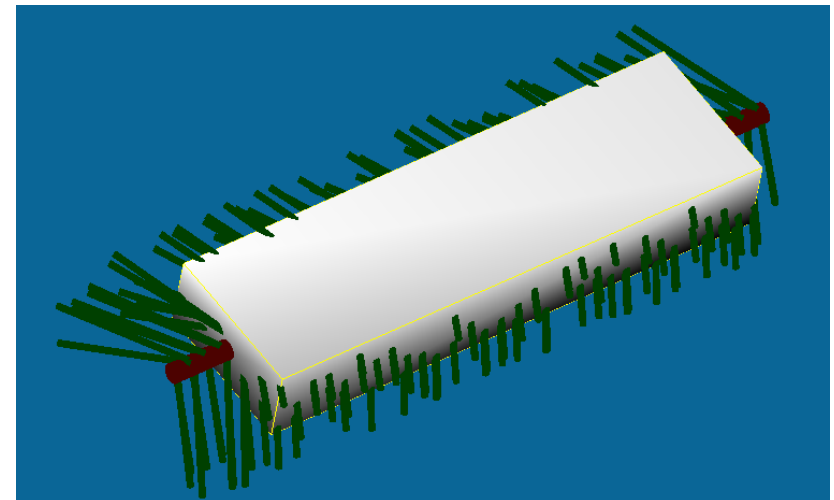
**Modèles volumineux**



# Modèles d'arbres

## Simplifications :

- Objet 2D : 4 points, 2 triangles
- Objet 3D : 8 points, 12 triangles
- Voxels à médium turbide



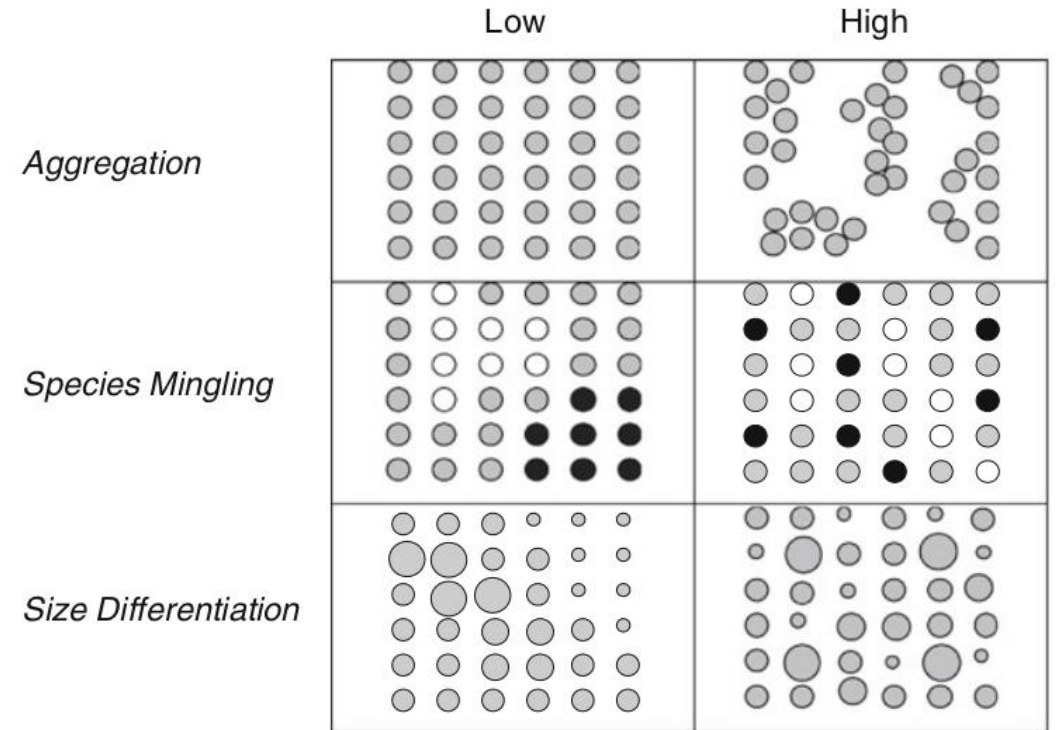




# Modélisation : Structure spatiale

## Descripteurs de structure spatiale :

- NNSS : Nearest Neighbors summary Statistics
- Processus ponctuels



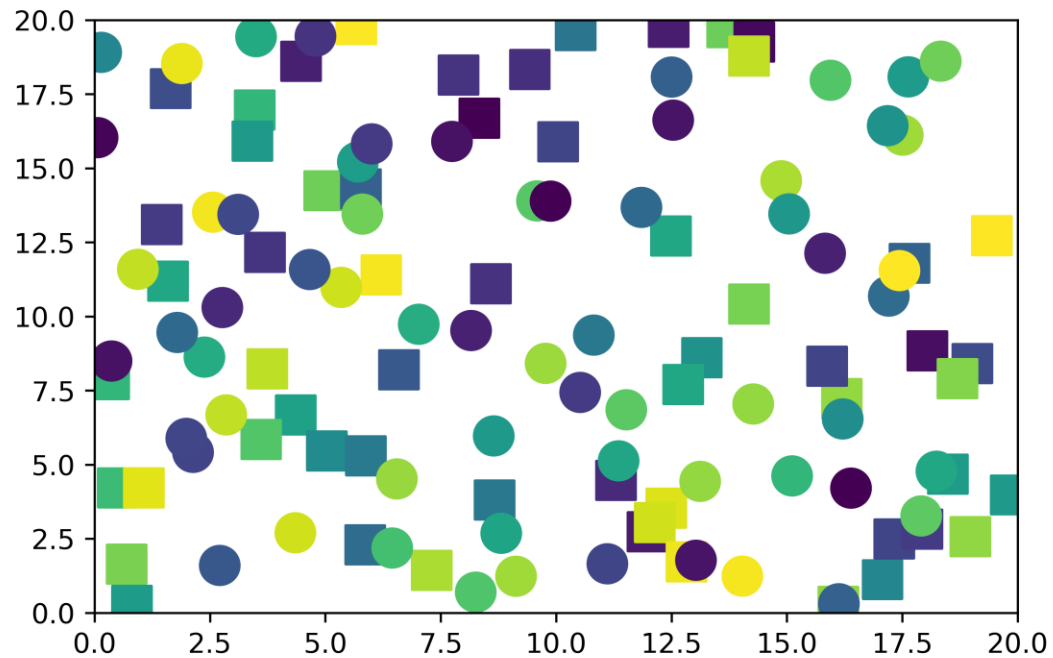
Source : [8]



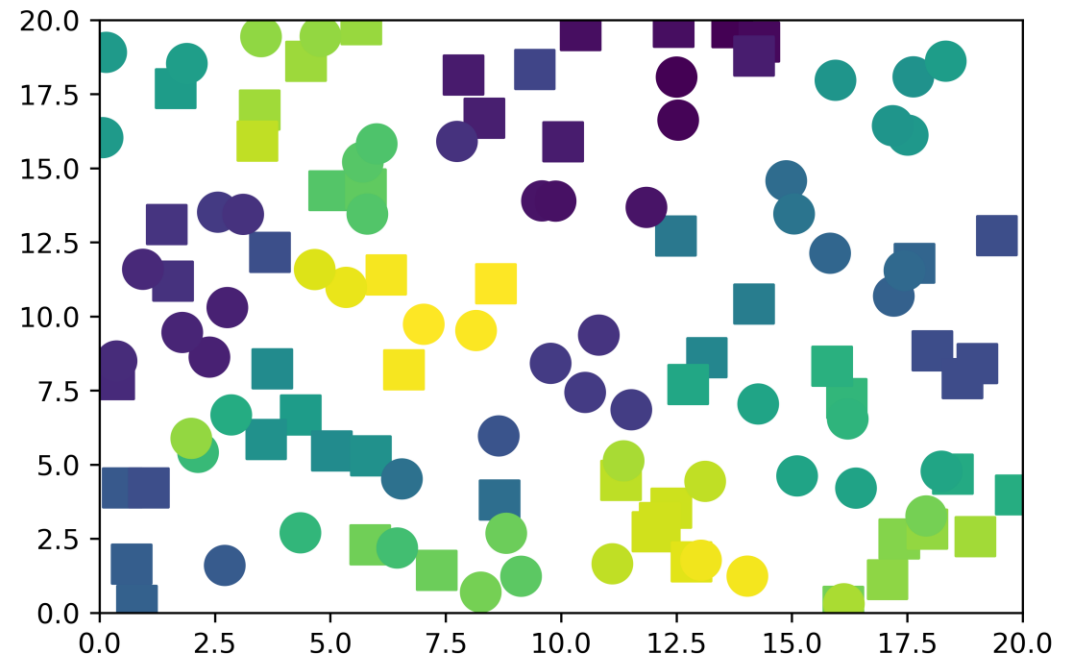
# Modélisation : Structure spatiale

Implémentation de la méthode de Pommerenning et Stoyan 2008 (NNSS):

- Mélange de hauteurs



Aléatoire



Optimisé

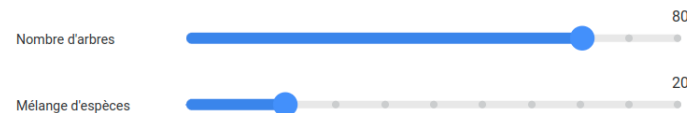


# Modélisation : Structure spatiale

## Baseline aléatoire :

- Paramètres variables (combinatoire pour génération de n placettes):

1. Nombres d'arbres : 1 à 121
2. Mélange d'espèces : 100% Ab et 0% Pm à 0% Ab et 100% Pm

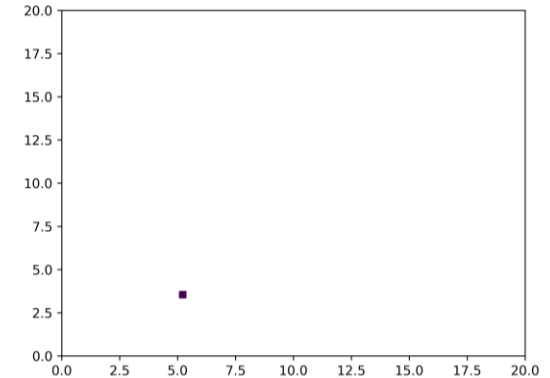


- Paramètre fixe :

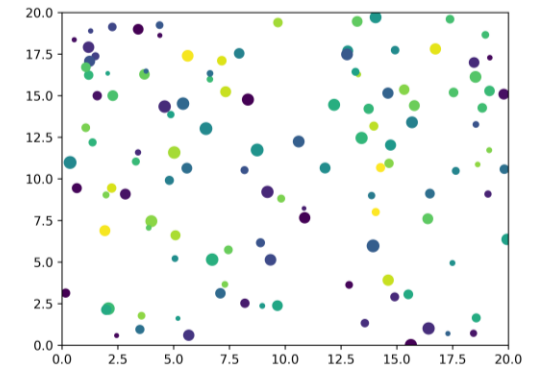
- Taille de placette : 20 m par 20 m

- Paramètres aléatoires :

- Positionnement en X : uniforme de 0 à 20 m
- Positionnement en Y : uniforme de 0 à 20 m
- Hauteur : uniforme de 2 à 20 m
- DBH : uniforme de 0.03 à 0.43 cm

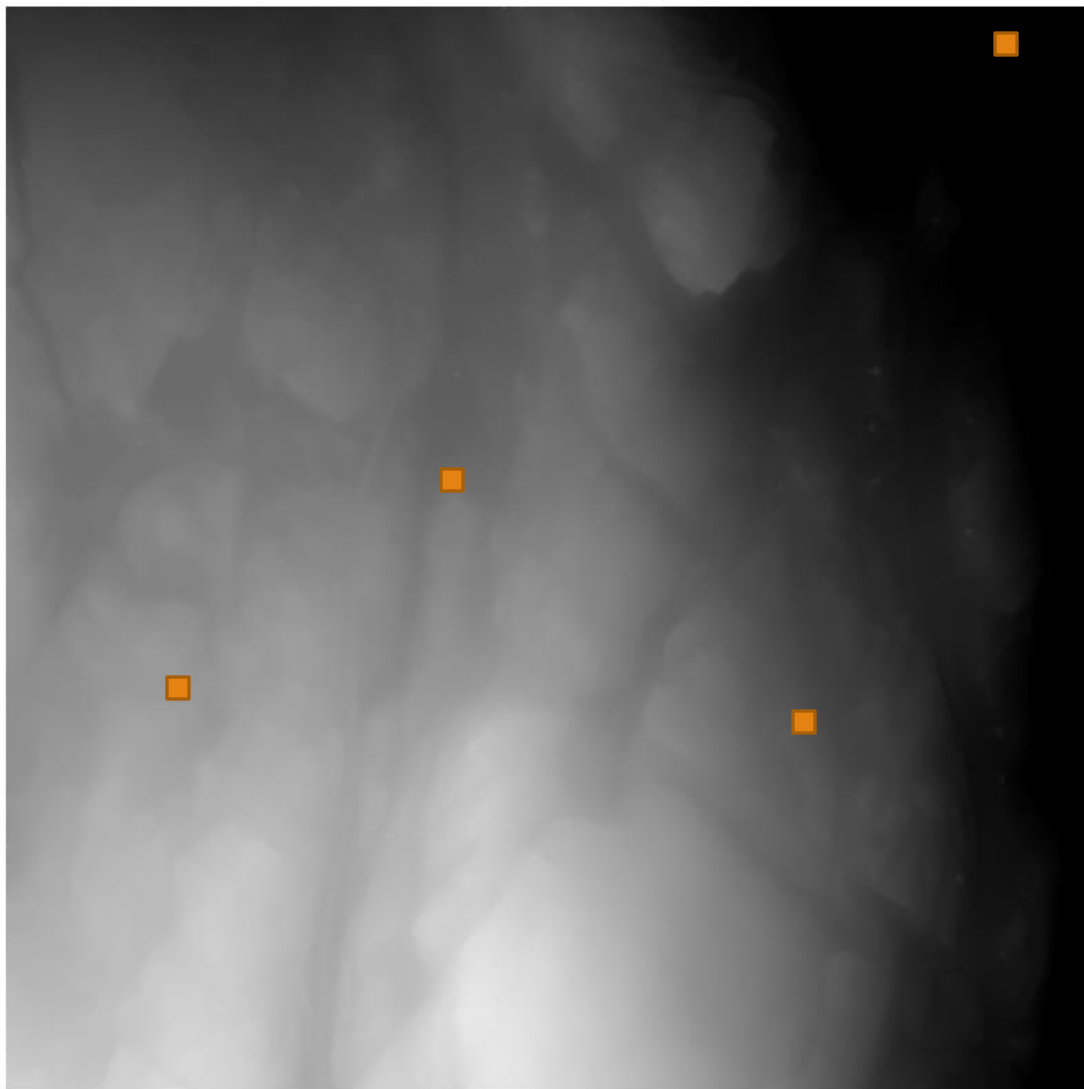


Placette 1 :  $N_{\text{arbres}}=1$ , Ab=100%, Pm=0%



Placette n :  $N_{\text{arbres}}=121$ , Ab=0%, Pm=100%





# Modélisation : Sol

À partir d'un MNT réel:

- Issu d'ALS
- Valeurs significatives :
  - Hauteur min : 150.319 m
  - Hauteur max : 309.480 m
  - Surface : 1 km<sup>2</sup>
  - Résolution : 10 cm
- Zones planes et zones de relief
- Transitions douces et abruptes



# Simulation

---

1. Outils de simulation
2. Paramètres d'acquisitions
3. Simulation puissance de calcul
4. Jeu de données actuel



## Choix de l'outil de simulation

---



Heidelberg LiDAR Operation Simulator :

- Rapidité de simulation
- Annotation des points avec ID de l'objet

Source : [10]



# Paramètres d'acquisition

**Objectif : 10 pts/m<sup>2</sup> minimum au sol**

## Paramètres de scan :

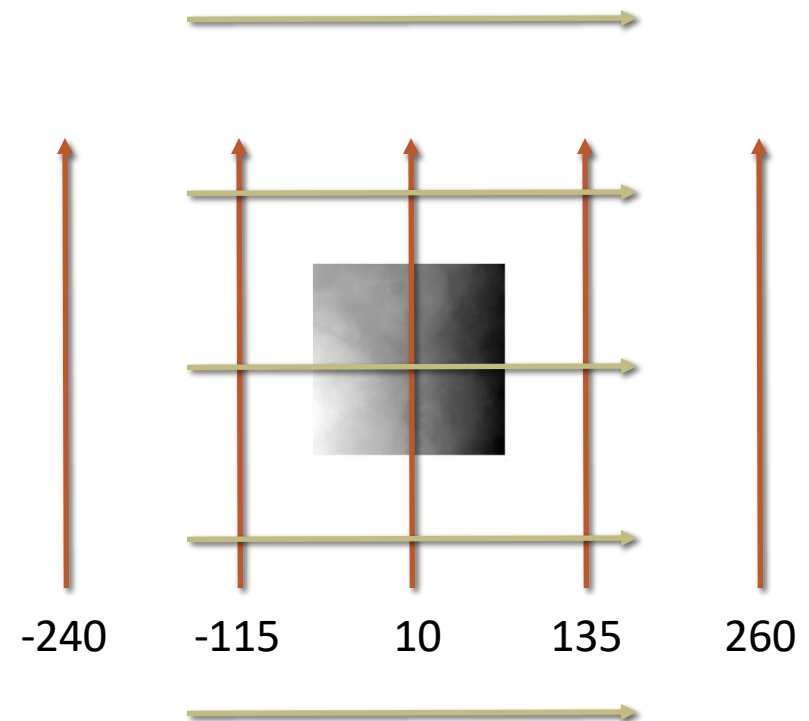
- Fréquence de pulsation : 500 000 Hz
- Angle de scan max 15 deg
- Fréquence de balayage : 100 Hz

## Paramètres du plan de vol :

- Hauteur de vol : 1200 m
- Vitesse 74.59 m/s (268.5 km/h)
- Trajet : 10 passes

## Paramètres de précision :

- Précision globale : 0 m







# Traitement de la simulation

---

## Objectif de la simulation :

- Env. 225 millions de points par classe
- 3 classes : Ab, Pm et sol

**675 millions de points**

## Nombre de placettes requises:

- Surface placette : 400 m<sup>2</sup>
- Densité d'acquisition moyenne : 18 pts/m<sup>2</sup>
- 10 passes

**10250 placettes  
(700+ millions points)**



# Traitement de la simulation

## Ressources informatiques nécessaires :

- Volume en RAM maximal estimé : 110 Go
- Principale perte de temps : Chargement en RAM
- Temps moyen par placette : 1h04

## Compute Canada :

- Répartition de la charge sur 1000 jobs



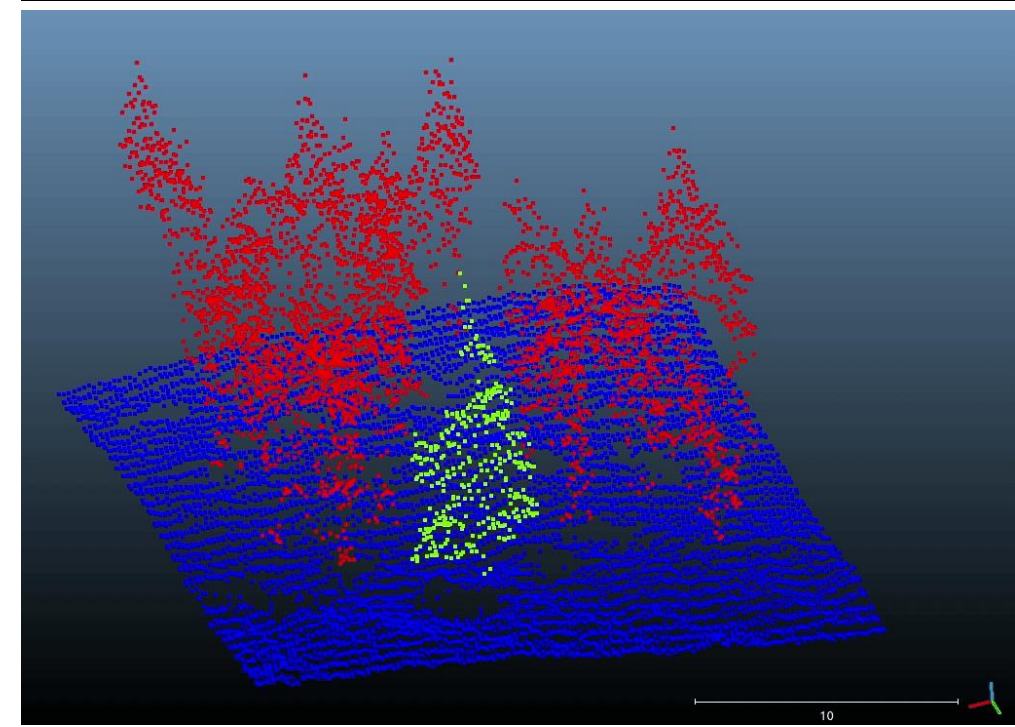
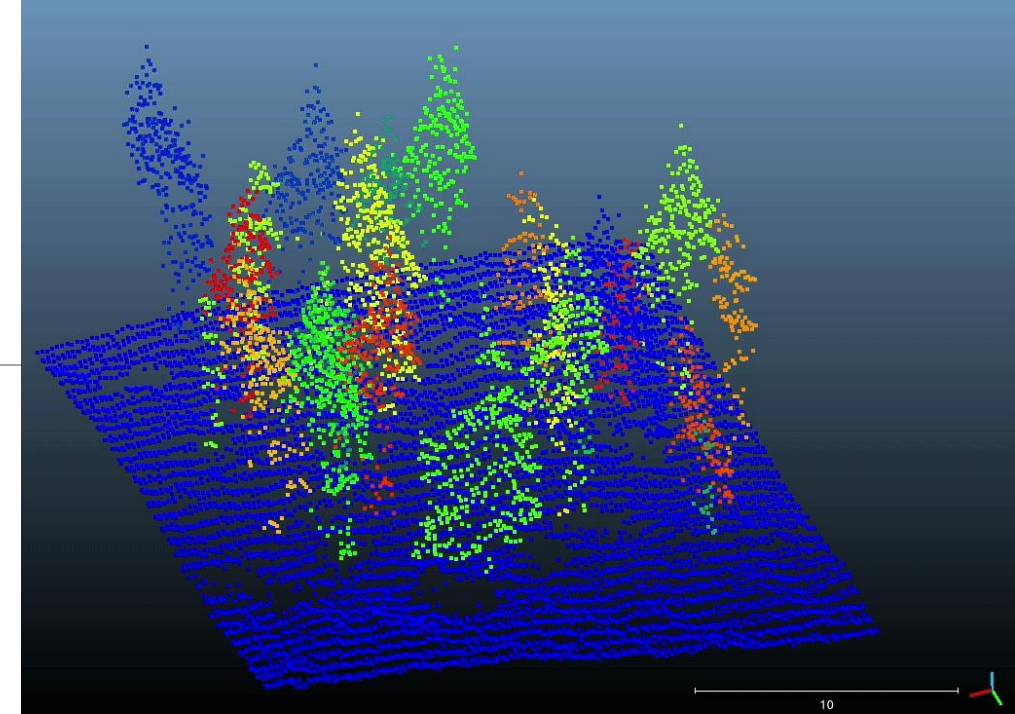
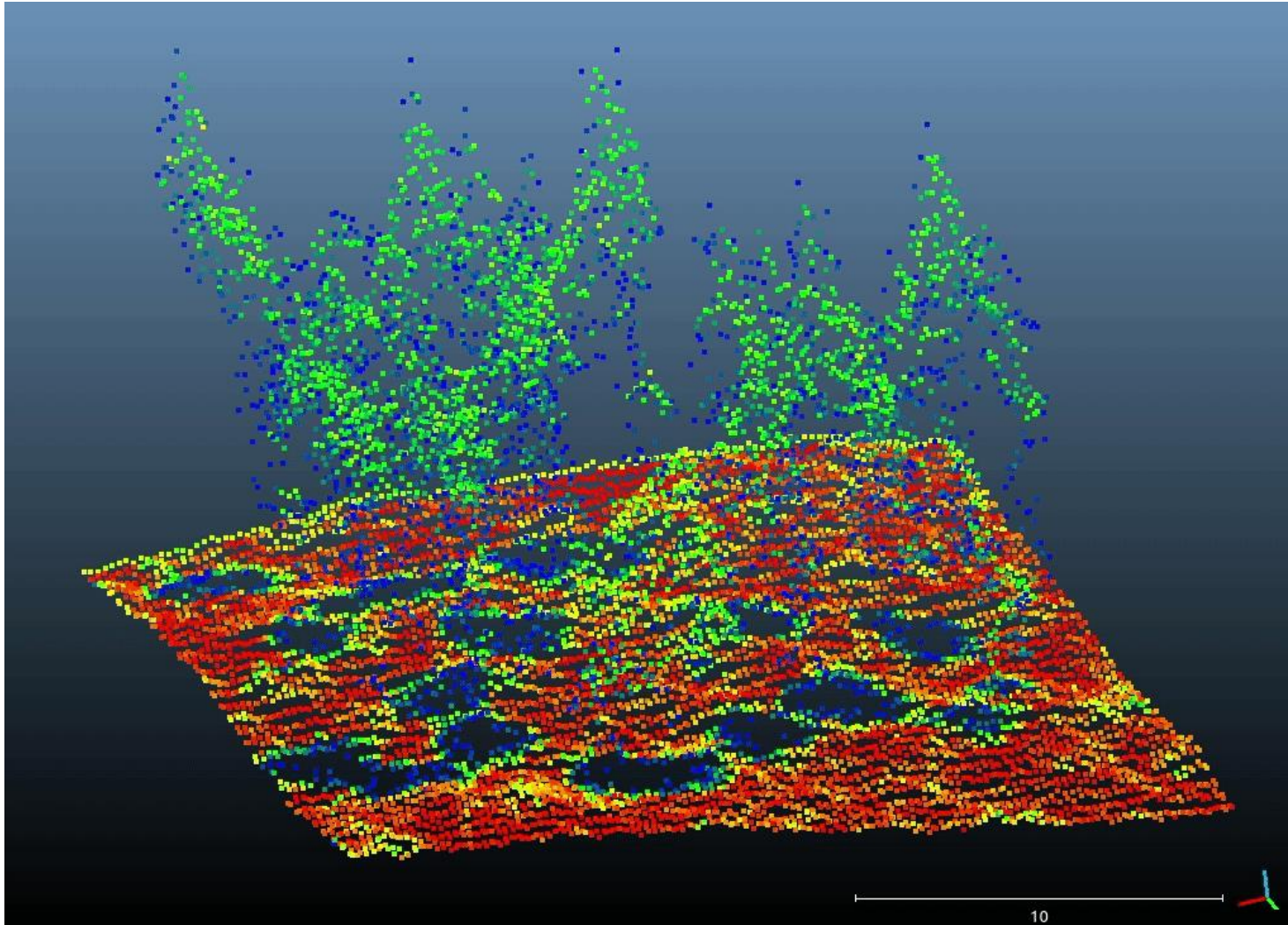
**compute** | **calcul**  
canada | canada

**109 Jours de calculs**



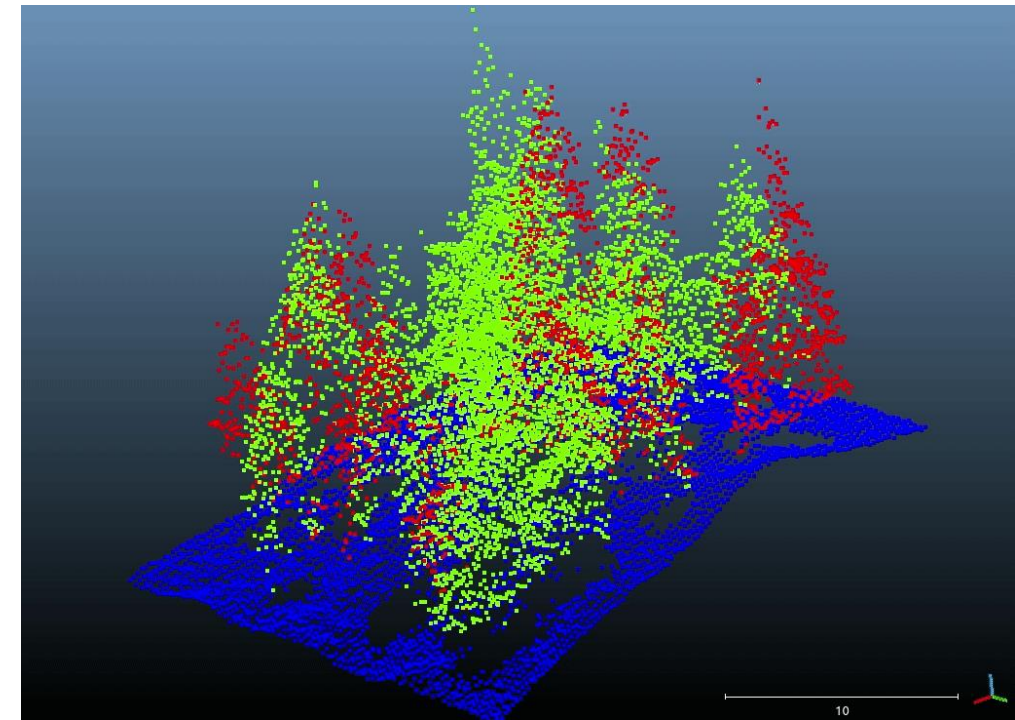
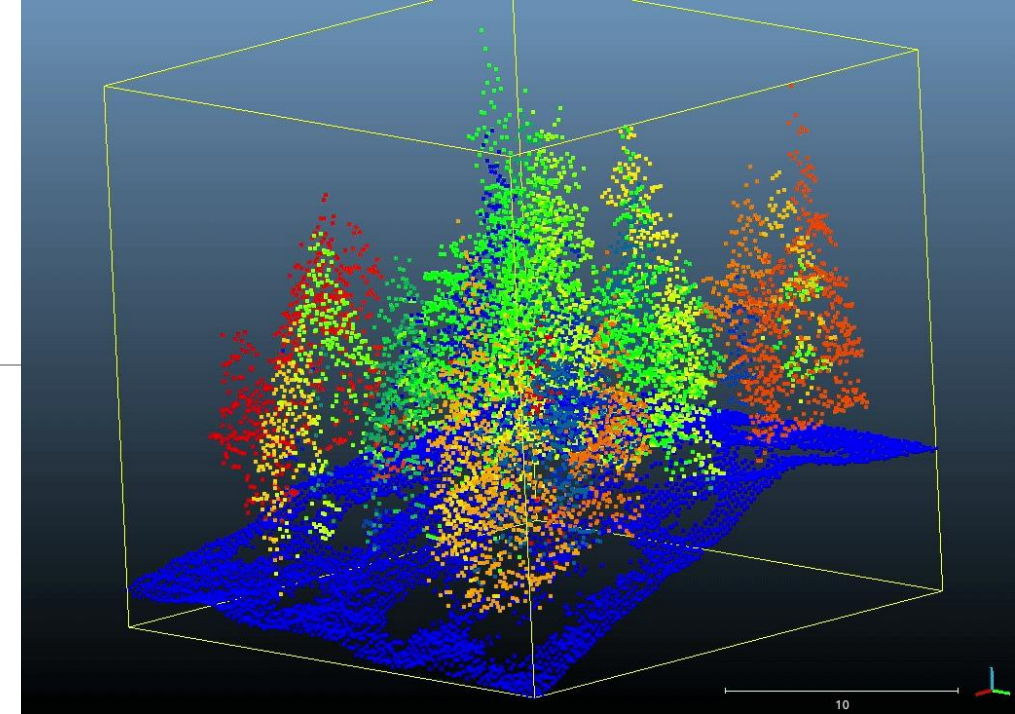
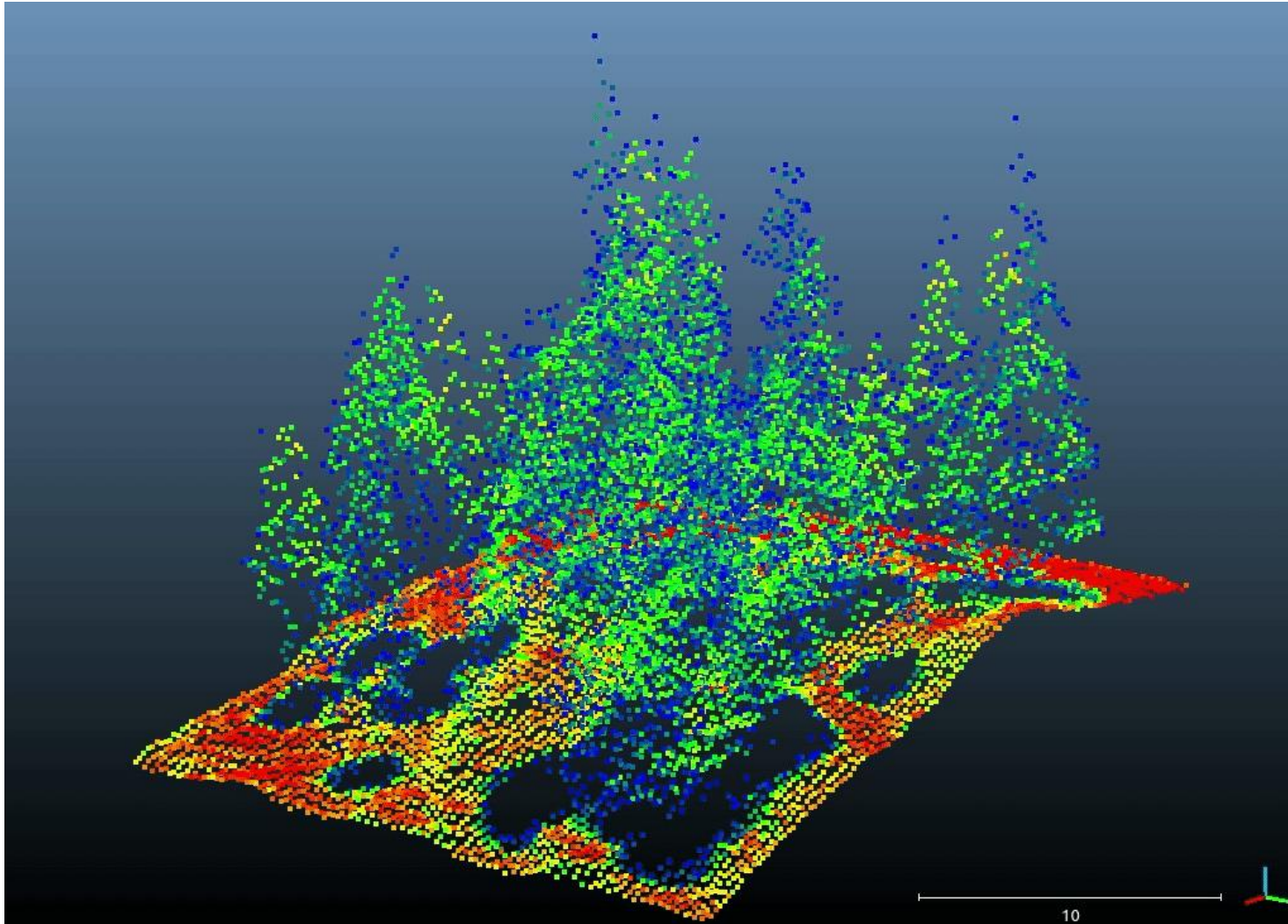
**3h de calculs**

# Jeu de données simulées



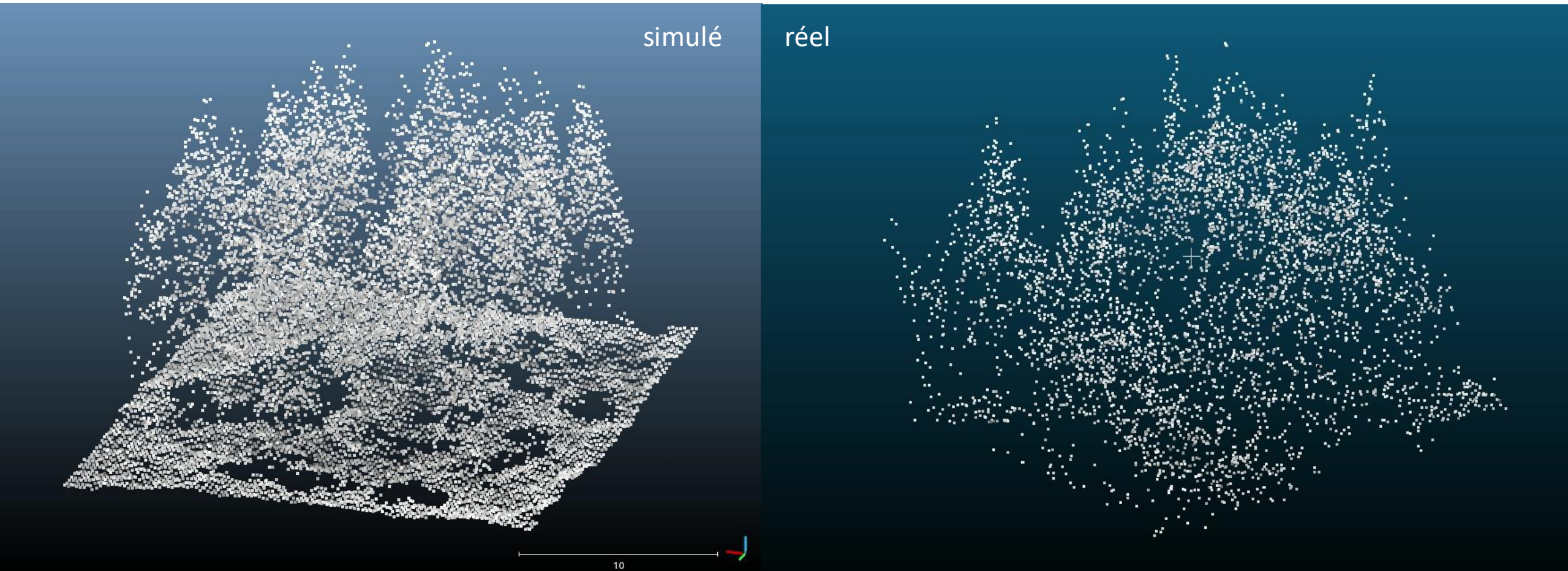


# Jeu de données simulées



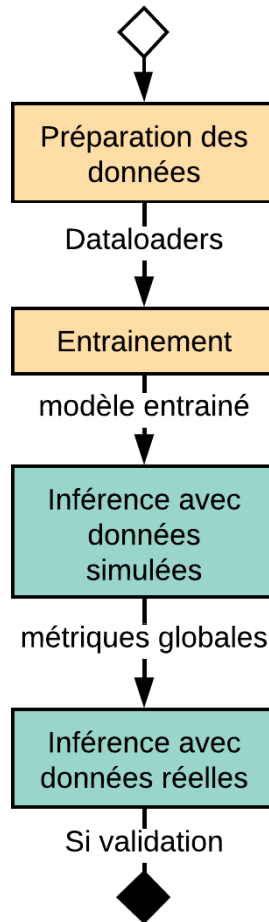


# Jeu de données simulées





# Axe 1 : entraînement et validation des modèles



## Entraînement :

- PointNet++ [11]
- KPConv [12]
- Point Transformer [13] : *Etat de l'art*

## Validation :

- Sur données **simulées** : Métriques globales et analyses visuelles
- Sur données **réelles** : Analyses visuelles



# Perspectives : Vue générale

## Axe 1 : Validation

Modélisation forêt simple  
Création de vérité terrain par simulation  
Entraînement réseau de segmentation sémantique  
Validation sur données réelles

## Axe 2 : Compréhension

Étude de sensibilité des paramètres :

- Modélisation
- Simulation

Entraînement de réseaux :

- Segmentation d'instances
- Segmentation panoptique

## Axe 3 : Généralisation

Enrichissement nomenclature  
Création benchmark  
Cas forêt tempérée  
Cas forêt tropicale



# Retombées du projet

---

**Premiers travaux de simulation LiDAR ALS pour la création de vérité terrain en milieu forestier**

Mise à disposition d'un **benchmark** => **Stimulation recherche** en **DL** pour le milieu forestier

**Recommandations de méthodologies d'acquisitions pour différents organismes et entreprises**



# Merci de votre attention



- [1] Ogle, S. M., Domke, G., Kurz, W. A., Rocha, M. T., Huffman, T., Swan, A., ... & Krug, T. (2018). Delineating managed land for reporting national greenhouse gas emissions and removals to the United Nations framework convention on climate change. *Carbon balance and management*, 13(1), 1-13.
- [2] Gauthier, S., Bernier, P., Burton, P. J., Edwards, J., Isaac, K., Isabel, N., ... & Nelson, E. A. (2014). Climate change vulnerability and adaptation in the managed Canadian boreal forest. *Environmental Reviews*, 22(3), 256-285.
- [3] Fassnacht, F., Latifi, H., Stereńczak, K., Lefsky, M., Straub, C., Waser, L., Ghosh, A., and Modzelewska, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 184. Review Lidar for forestry and species classification.
- [4] Roynard, X., Deschaud, J. E., & Goulette, F. (2018). Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *The International Journal of Robotics Research*, 37(6), 545-557.
- [5] Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., & Xiao, J. (2015). 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1912-1920).
- [6] Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., van der Smagt, P., Cremers, D., and Brox, T. (2015). FlowNet : Learning optical flow with convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- [7] Zhao, S., Wang, Y., Li, B., Wu, B., Gao, Y., Xu, P., Darrell, T., and Keutzer, K. (2020b). epointda : An end-to-end simulation-to-real domain adaptation framework for lidar point cloud segmentation. *CoRR*, abs/2009.03456
- [8] Gadow, K., González, J., Zhang, C., Pukkala, T., and Zhao, X. (2021). *Analyzing Forest Ecosystems*, pages 81–158. Springer Netherlands.
- [9] Pommerening, A., & Stoyan, D. (2008). Reconstructing spatial tree point patterns from nearest neighbour summary statistics measured in small subwindows. *Canadian journal of forest research*, 38(5), 1110-1122.
- [10] Winiwarter, L., Esmorís Pena, A. M., Weiser, H., Anders, K., Martínez Sánchez, J., Searle, M., and Höfle, B. (2021). Virtual laser scanning with helios++ : A novel take on ray tracing-based simulation of topographic full-waveform 3d laser scanning. *Remote Sensing of Environment*, page 112772.
- [11] Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in neural information processing systems*, 30.
- [12] Thomas, H., Qi, C. R., Deschaud, J., Marcotegui, B., Goulette, F., and Guibas, L. J. (2019). Kpconv : Flexible and deformable convolution for point clouds. *CoRR*, abs/1904.08889.
- [13] Zhao, H., Jiang, L., Jia, J., Torr, P. H. S., and Koltun, V. (2020a). Point transformer. *CoRR*, abs/2012.09164.

<https://polarsensing.net/>