

Algorithme Lvox

Problème de l'estimation des données manquantes

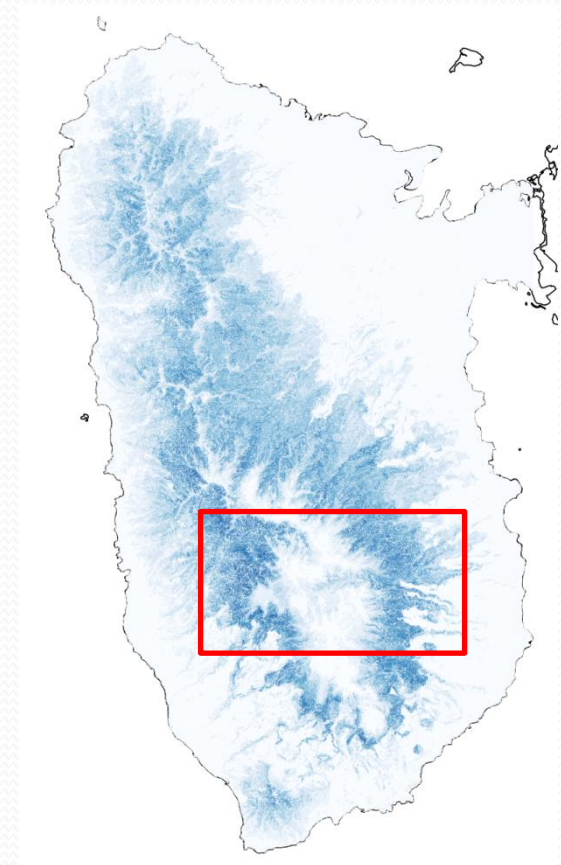
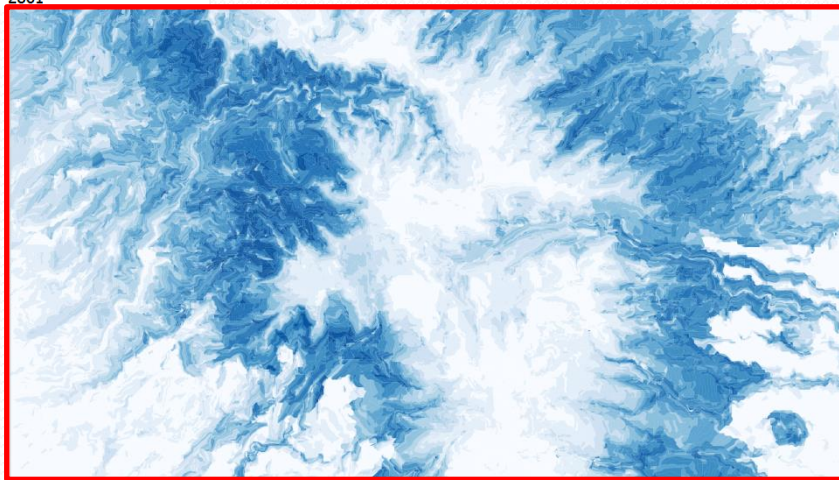
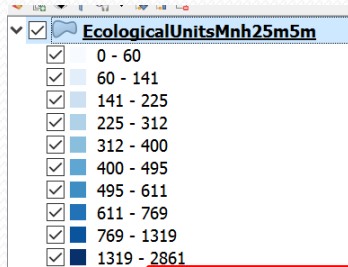
Enguerran Grandchamp, LAMIA, UA

Plan

- Objectifs opérationnels
- Principe LiDAR
 - Acquisition des données
 - LiDAR terrestre
 - MultiScan
- Algorithme LVox
 - Fusion des scans
 - Calcul des densités
 - Profils
 - PAD
- Intégration LiDAR aérien
- Outils
- Problèmes rencontrés
 - Estimation des données manquantes
 - Principe / validation

Objectifs opérationnels

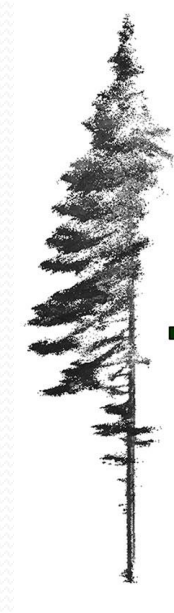
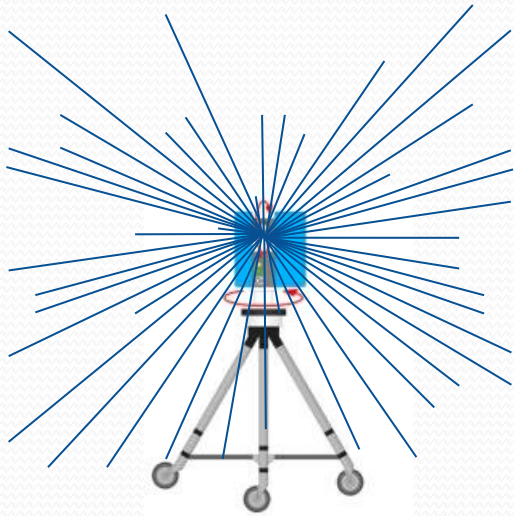
- Meilleure caractérisation des forêts
- Estimation de la biomasse



Approche globale

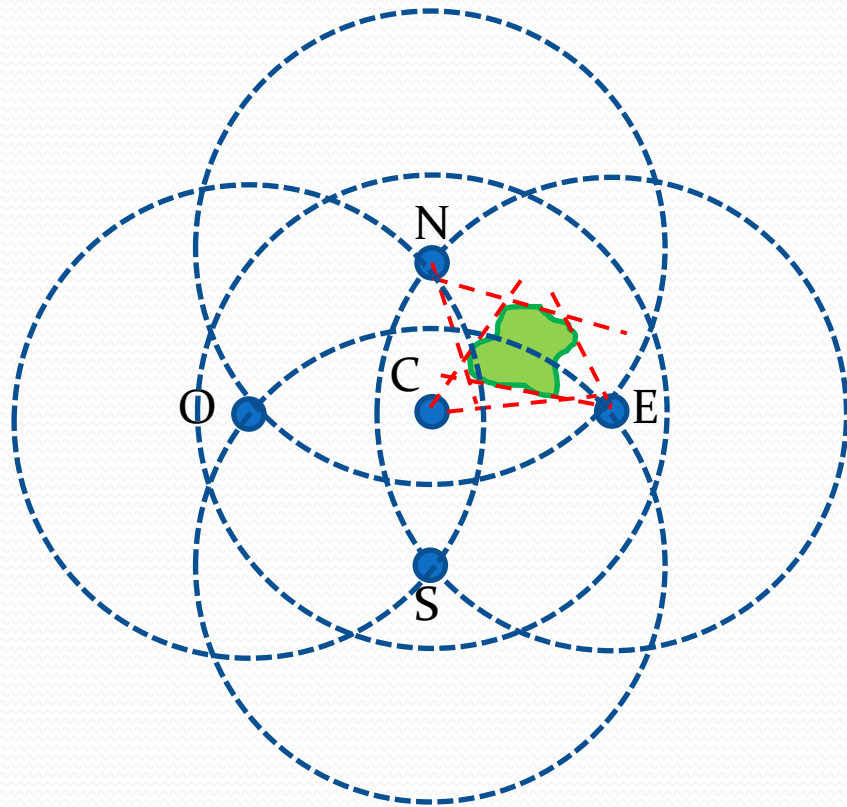
- Acquisition des données
 - LiDAR terrestre
- Traitement et analyse des données
 - Algorithme LVox
- Définition des modèles de biomasse à partir de relevés manuels et mise en correspondance avec le LiDAR terrestre.
- Mise en correspondance LiDAR terrestre et aéroporté
- Généralisation du modèle de biomasse

Principe d'acquisition : LiDAR terrestre



- Alternative relevé manuel
- A l'échelle d'une placette
- Calcul de distance par mesure Laser
- → Nuage de points en 3D, ou 3D+I

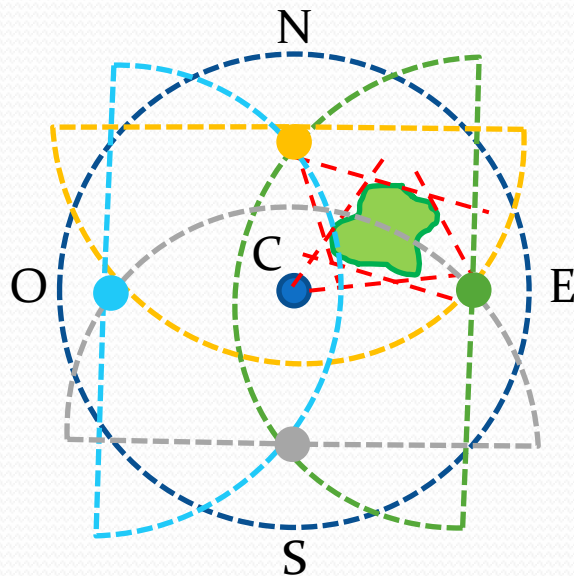
Acquisition des scènes



- Juillet 2015
- 19 scènes acquises
- 5 scans minimum par scène
- 20 Millions de points par Scan (140 Go)



Acquisition des scènes

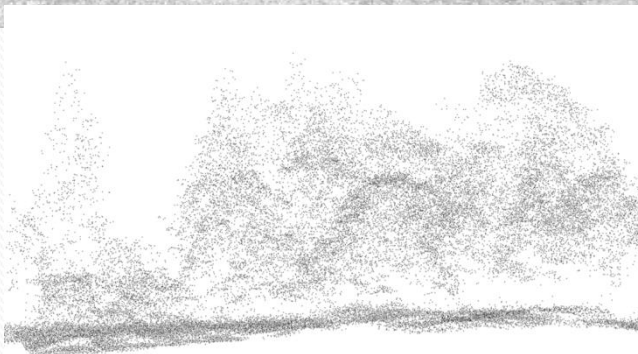
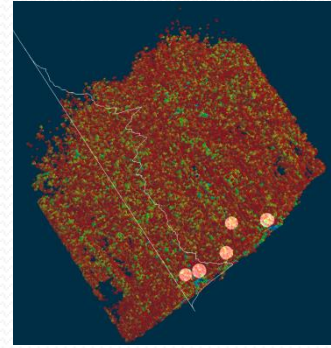
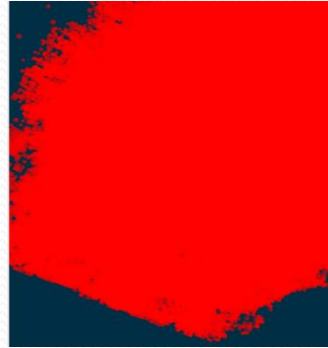


- Juillet 2015
- 19 scènes acquises
- 5 scans minimum par scène
- 20 Millions de points par Scan (140 Go)

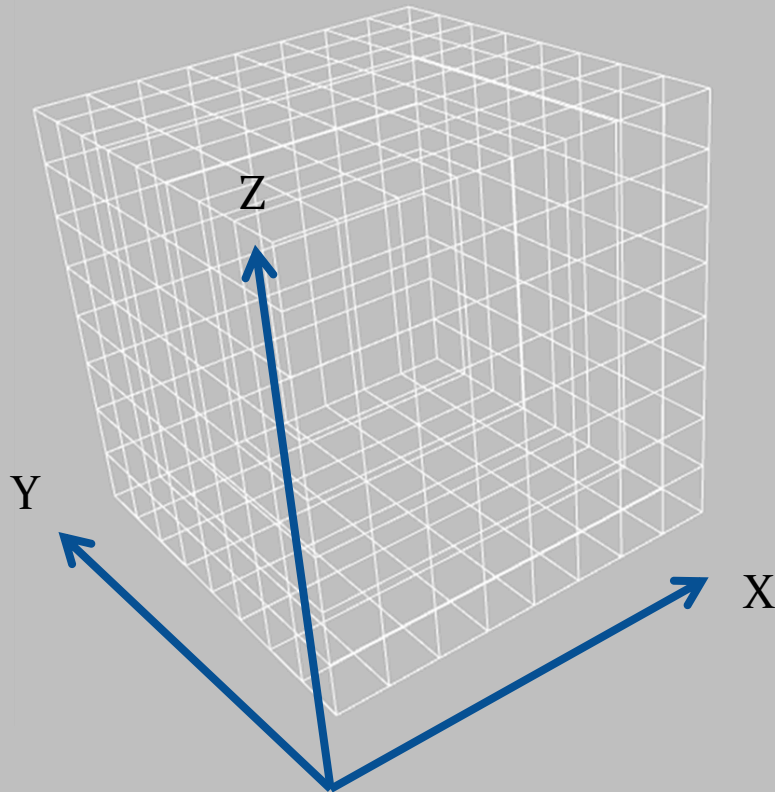


Données

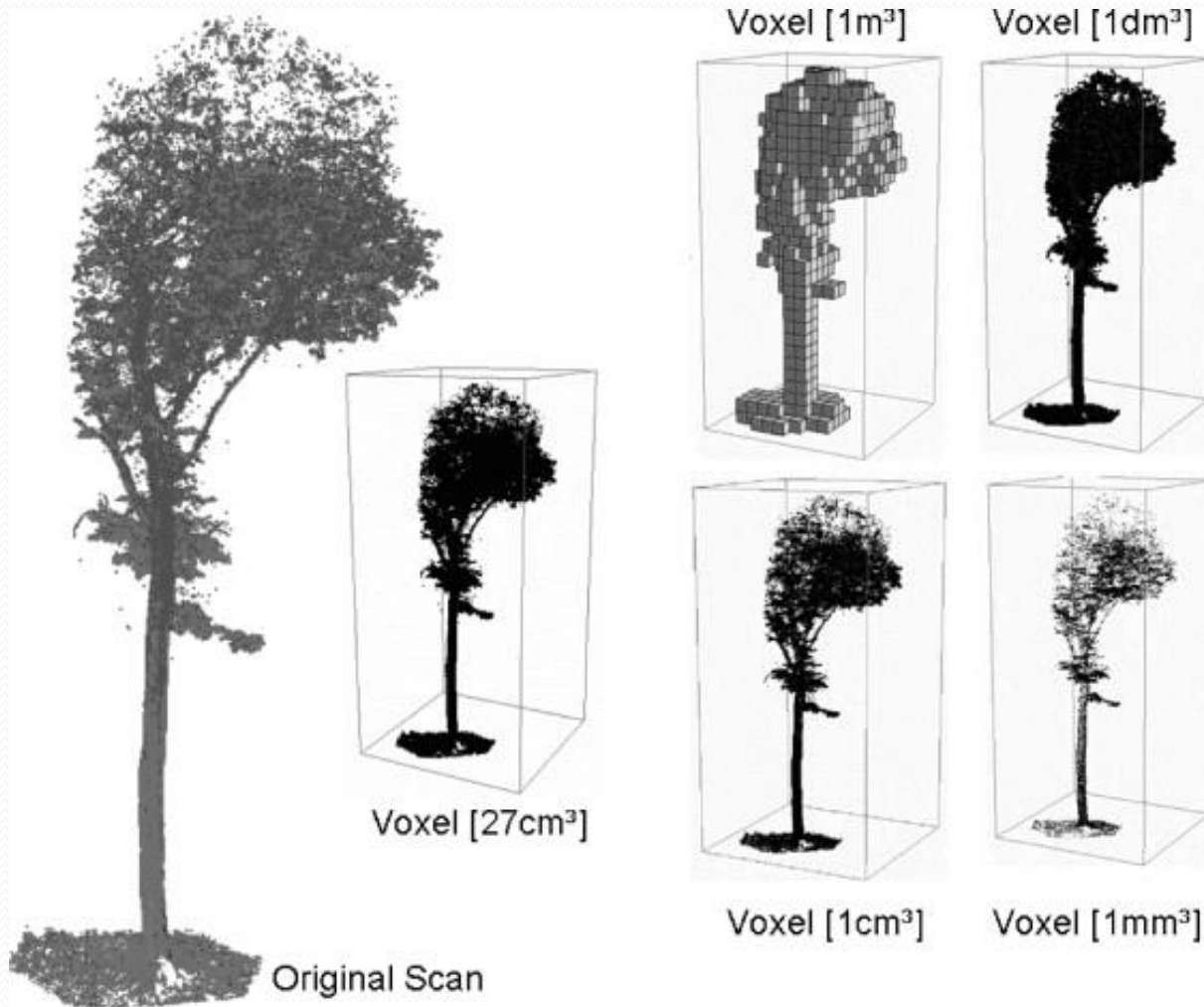
- Données brutes 3D
- Projection sur une section



Lvox : décomposition en voxels



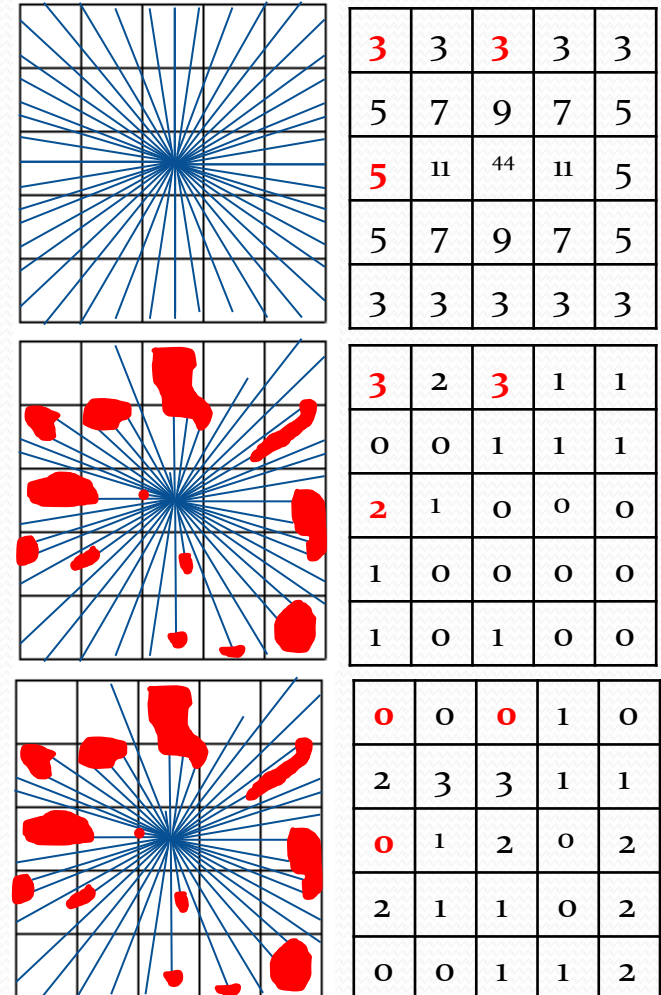
Représentation par Voxel



Siedel et al. (2012)
Agric. For. Meteorol.

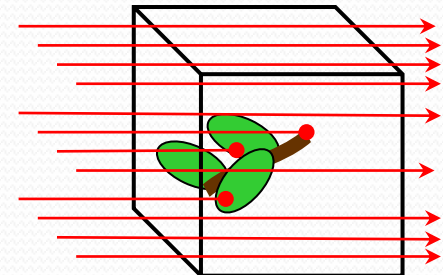
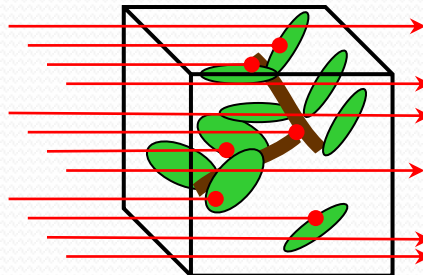
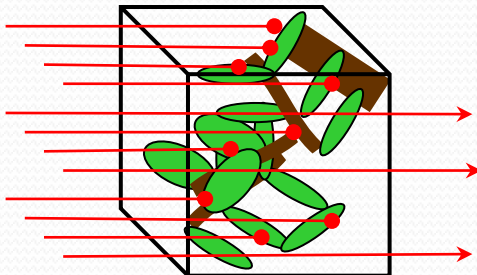
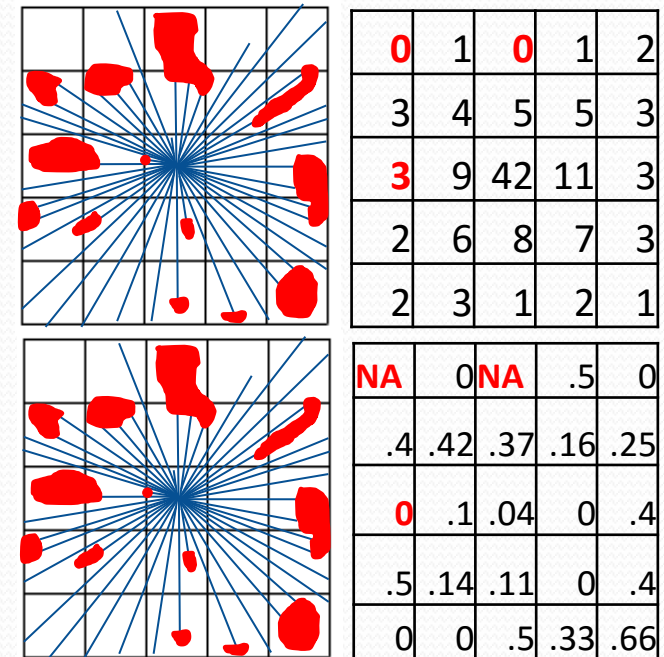
LVox : Nt, Nb, Ni

- Nt : nombre de rayons théoriques que le voxel doit recevoir sans occultation
- Nb : nombre de rayons bloqués avant le voxel
- Ni : nombre de rayons interceptés par le voxel



Lvox : densité

- $N_t - N_b - N_i$ = nombre de rayons traversant
- Densité $d = N_i / (N_t - N_b)$ (VDI : Vegetal Density Index)
- *Hypothèse* : L'interception est proportionnelle à la densité de végétation
- Si $N_t = N_b$: Voxel sans hit (NA)
- → 5 Matrices par Scan



LVox : fusion des n scans

- Critère choix unique

- $d_{xyz} = \max_{s=1,n}(d_{xyz}^s) = \max_{s=1,n}(Ni_{xyz}^s / (Nt_{xyz}^s - Nb_{xyz}^s))$

- $d_{xyz} = d_{xyz}^{s_0} | S_0 = \underset{s=1,n}{\text{ArgMax}}(Nt_{xyz}^s - Nb_{xyz}^s)$

- $d_{xyz} = d_{xyz}^{s_0} | S_0 = \underset{s=1,n}{\text{ArgMax}}((Nt_{xyz}^s - Nb_{xyz}^s) / Nt_{xyz}^s)$

- $d_{xyz} = d_{xyz}^{s_0} | S_0 = \underset{s=1,n}{\text{ArgMax}}(Ni_{xyz}^s)$

- Calcul d'une nouvelle densité

- $d_{xyz} = \sum_{s=1}^n Ni_{xyz}^s / \sum_{k=1}^n (Nt_{xyz}^s - Nb_{xyz}^s)$

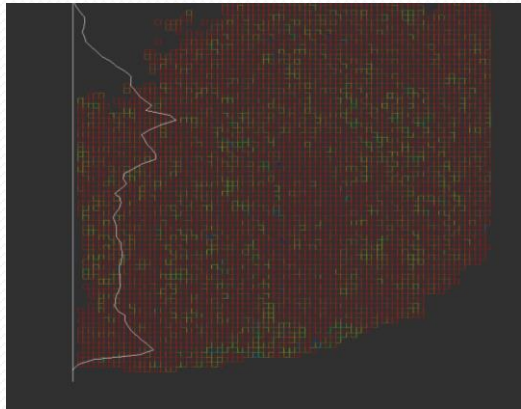
- Seuil sur Nt-Nb : 10 ?

- Prise en compte des scans en fonction de leur position par rapport au voxel ?

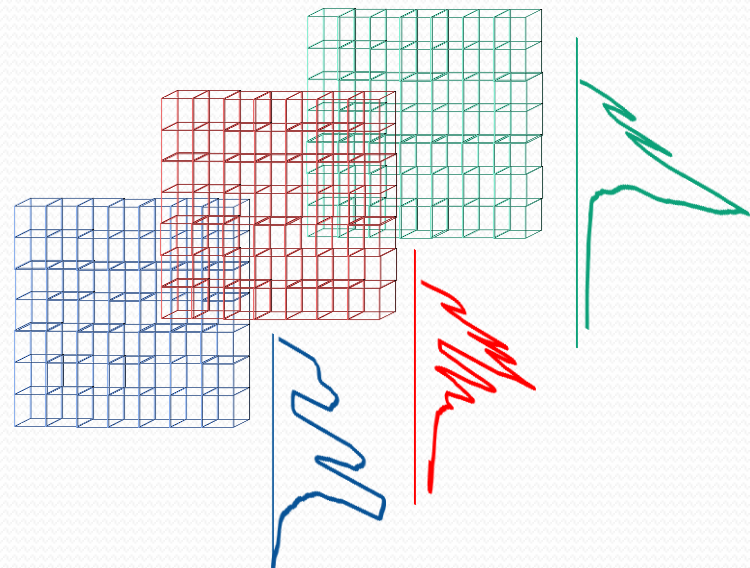
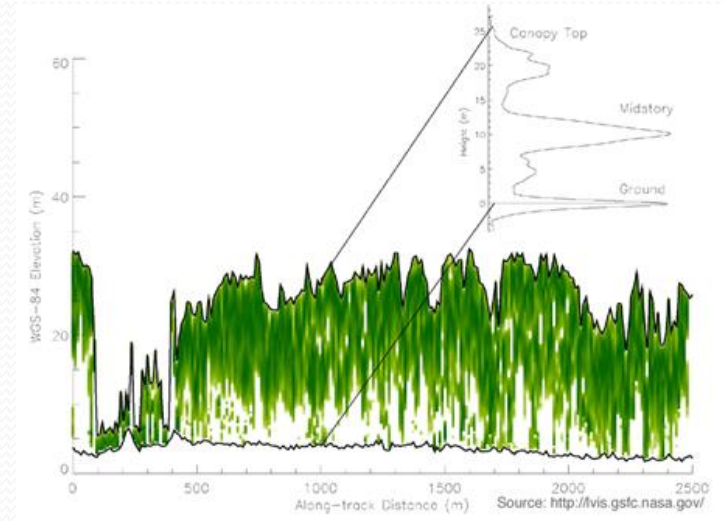
LVox : profil

- Vertical (z)

- Global

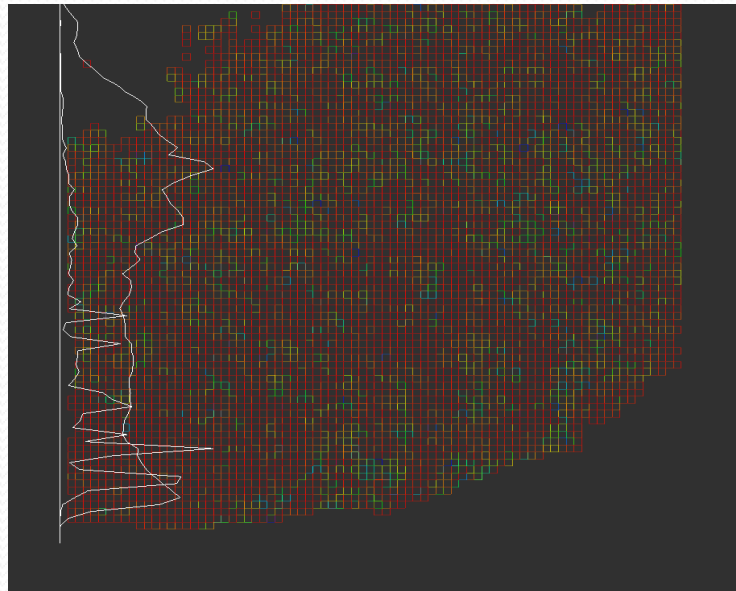


- Par tranche (selon x ou y)
 - → variabilité



Lvox : PAD

- Hypothèse :
- Plant Area Density (m^2/m^3)
- Densité surfacique dans un voxel
- Profil PAD

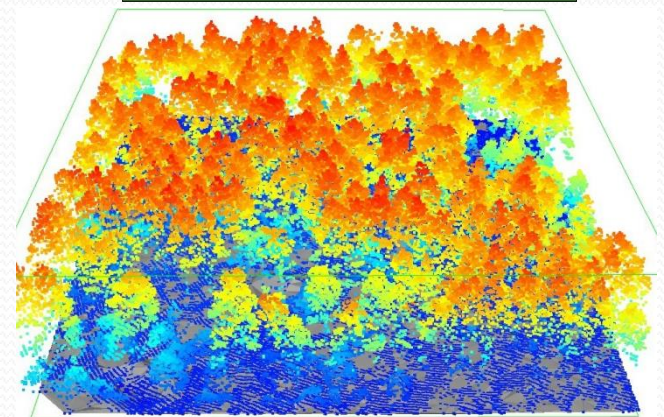
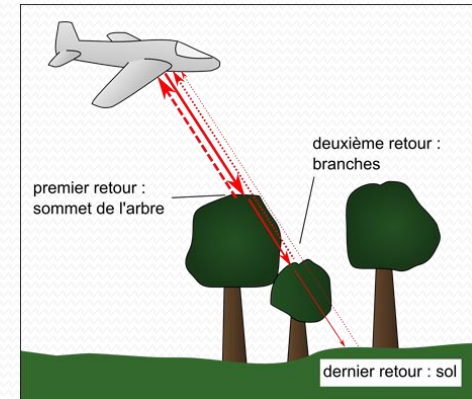
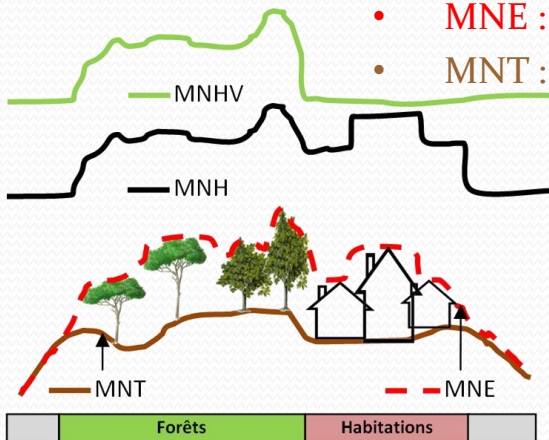


LVox : perspectives

intégration du LiDAR aéroporté

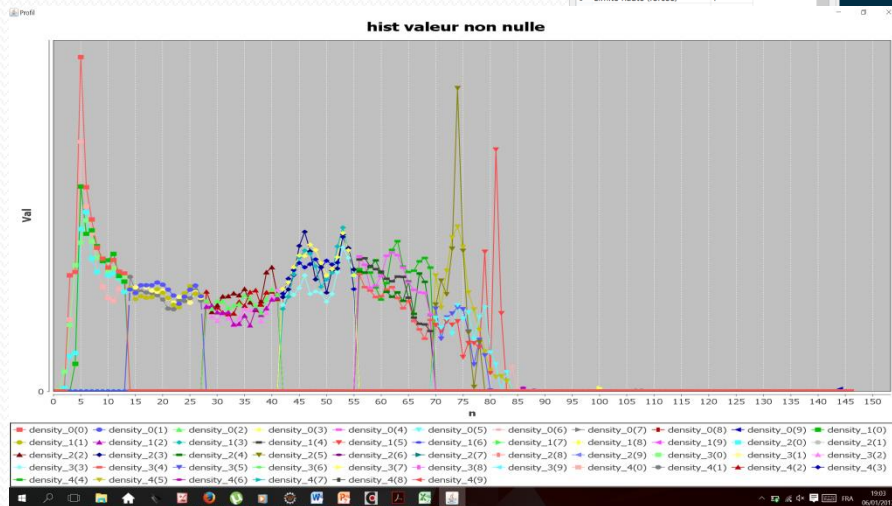
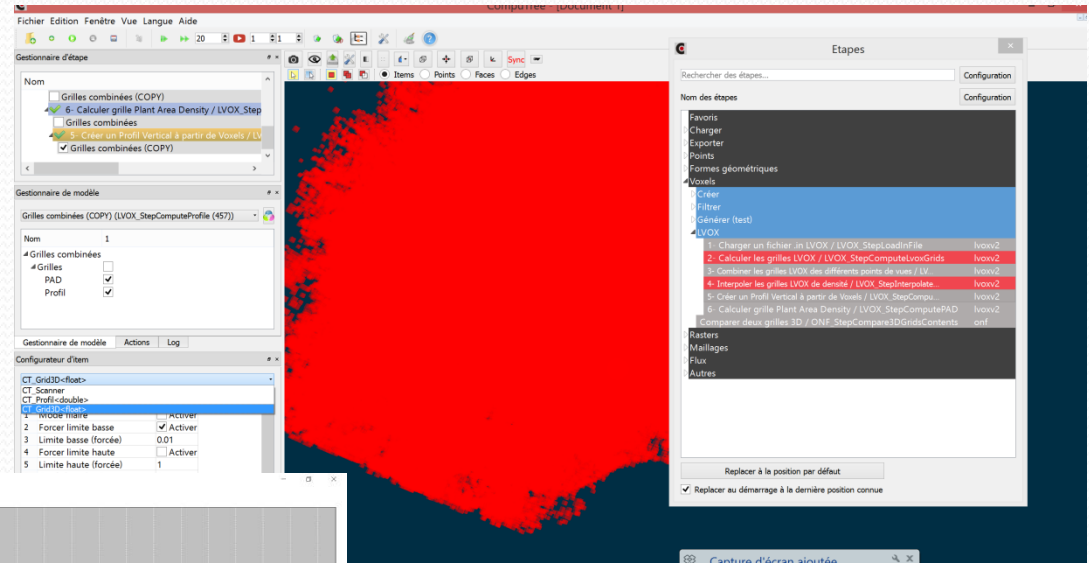
- A l'échelle d'une région
- Plusieurs retours possibles
- Adaptation de la méthode de calcul des Nt Nb Ni

- MNHV : Modèle Numérique de Hauteur de la Végétation
- MNH : Modèle Numérique de Hauteur
- MNE : Modèle Numérique d'Élévation
- MNT : Modèle Numérique de Terrain



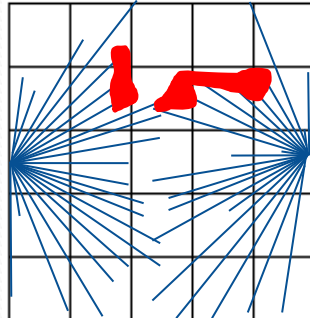
LVox : outils

- Computree
 - UdS, ONF
 - C++
- Code JAVA



Problème rencontré : estimation des données manquantes

- Valeur manquante : pas d'observation pour une variable (retours) donnée pour un individu (voxel) donné
- Origine
 - Occultation



Problème rencontré : estimation des données manquantes

- Problème
 - On ne peut pas faire de nouvelles mesures
 - Ne peuvent pas être ignorées (données pertinentes et informatives) au delà d'un certain seuil (à définir)
- Solutions
 - Calculer une valeur pour les données manquantes : imputation
 - Développer un modèle robuste ou un algorithme prenant en compte les données manquantes

Caractéristiques des données manquantes

- La probabilité qu'une observation soit manquante dépend-elle des mesures observées ou non observées ?
 - Si elle ne dépend d'aucune des deux : les données manquantes sont dites **complètement aléatoires (MCAR)**
 - $P(r \mid X_{\text{obs}}, X_{\text{nonObs}}) = P(r)$
 - Si elle ne dépend pas des données non observées : les données manquantes sont dites **aléatoires (MAR)**
 - $P(r \mid X_{\text{obs}}, X_{\text{nonObs}}) = P(r \mid X_{\text{obs}})$
- Approche statistique
 - Cas des données manquantes aléatoirement : les méthodes basées sur la vraisemblance restent valides si on ignore les données manquantes. D'autres méthodes comme les moments seront biaisés.
 - Cas des données manquantes non aléatoirement (MNAR) : on ne peut pas ignorer ces données
- Intuitivement il y a dépendance (au moins dans un voisinage à définir) : organisation spatiale du matériel, diminution du nombre de rayons avec la distance. **A Démontrer**
 - Modéliser le processus d'observation

Analyse des données complètes

- Avant ou après fusion des scans
 - Avant : n observations d'un même voxel
 - Après : 1 observation d'un même voxel
- Données complètes
 - Peu de voxel complets par construction (observation depuis des angles différents)

Imputation

- Par la moyenne ou la médiane
 - De la couche (z constant) : méthode actuelle
 - Inconvénients
 - même valeur pour toutes les données manquantes d'une couche
 - sous estime la variance
- Par tirage conditionnel
 - Moyenne du voisinage : lequel ? K-ppv ? LOCF (Last Observation Carried Forward) : propagation d'une observation. CMCF (Concept Most Common Attribute Value Fitting)
 - Construire un modèle de régression par rapport aux données renseignées et prédire les données manquantes.
- Imputation multiple : n tirages (5 en pratique) et répéter les analyses → unique estimateur
 - Algorithme EM (Espectation-Maximisation)
 - Méthode MCMC (Monte Carlo Markov Chain)
- Imputation par analyse factorielle
 - Exemple : estimation de la matrice de covariance à partir des observations, estimation paramètres ACP,

Autres méthodes possibles

- Prise en compte des scans en fonction de leur position par rapport au voxel ?
- Prise en compte de la signification physique des données
 - $N_t \geq N_b + N_i$
- Introduction de la position du voxel dans les variables d'observation ?
- Validation des approches
 - Validation croisée

Quelques références

- LiDAR et Forêt
 - Silvilaser : <https://silvilaser2015.teledetection.fr/>
 - <http://forestsat2014.com/>
 - Computree
 - <http://computree.onf.fr/?lang=fr>
 - <http://rdinnovation.onf.fr>
 - Logiciel LIDAR pour la foret
 - <http://www.liforest.com/#/>
 - <http://gisiana.info/gis-software/gis-for-forestry/>
 - <http://www.inigis.net/gis-for-forestry/>
 - **Rasterizing Perfect Canopy Height Models from LiDAR** [November 4, 2014](#), [Martin Isenburg](#)
- Estimation de données manquantes
 - Imputation des données manquantes : Comparaison de différentes approches. Mélanie Glasson-Cicognani & André Berchtold Université de Lausanne
 - Les données manquantes en statistique. N Meyer. Laboratoire de Biostatistique. Strasbourg