



Validation de l'estimation de la hauteur des jeunes peuplements à partir de données LiDAR aériennes

Nuria Sánchez López

► To cite this version:

Nuria Sánchez López. Validation de l'estimation de la hauteur des jeunes peuplements à partir de données LiDAR aériennes. Sciences de l'environnement. 2013. hal-01862582

HAL Id: hal-01862582

<https://hal.univ-lorraine.fr/hal-01862582>

Submitted on 27 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-memoires-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



UNIVERSIDAD
POLITECNICA
DE VALENCIA

VALIDATION DE L'ESTIMATION DE LA HAUTEUR DES JEUNES PEUPELEMENTS FORESTIERS A PARTIR DE DONNEES LIDAR AERIEN

Mémoire de stage, soutenu à Nancy le 3 septembre 2013

Nuria Sánchez López

**Master 2 STVE – Sciences et Technologies du vivant et de
l'environnement. Spécialité Fonctionnement et Gestion des
écosystèmes (FGE)**

Maître de stage : Nicolas GOMEZ, technicien R&D à Nancy.

Co-encadrants : Jean- Pierre RENAUD, chargé de R&D à Nancy
Jerôme BOCK, chargé de R&D à Chambéry.

Tuteur universitaire : Mathieu FORTIN, Ingénieur de recherche au
Laboratoire d'Etude des Ressources Forêt-Bois (LERFoB UMR 1092)

Structure d'accueil:

ONF, service R & D pôle de Nancy
11, rue de l'Île de Corse
54000 Nancy

AGRADECIMIENTOS

Es complicado dar las gracias en pocas líneas cuando hay tantas personas que han ayudado a que este trabajo sea posible. Es por ello que, de entrada, quisiera manifestar mis gratitud mas sincera a todos aquellos que de muy diversas formas me han apoyado y han estado dispuestos a ayudarme a lo largo de estos seis meses: compañeros y amigos de promoción, compañeros y amigos de Nancy así como todos los profesionales y profesores que han estado ahí para orientarme.

Agradecer especialmente a mi supervisor, Nicolas Gomez, toda su dedicación, ayuda, disponibilidad y paciencia en estos seis meses de trabajo. A Jérôme Bock por sus consejos, orientación y valiosas críticas constructivas en todo momento. A Anne Jolly, por sus recomendaciones en la redacción final de esta memoria. Y a Jean Pierre Renaud por su ayuda incondicional, sus recomendaciones y sugerencias, así como el buen ambiente y compañía que han hecho muy llevaderos los días de trabajo. A todos ellos gracias enormemente además por su paciencia en la relectura de esta memoria y la corrección de todos mis errores en francés.

En general, agradecer a todos los miembros del Departamento R&D del ONF en Nancy su buena acogida, compañía y ayuda.

Y finalmente, a mi familia y amigos, por estar siempre ahí y apoyarme desde la distancia.

TABLE DE MATIERES

TABLE DE FIGURES	iv
TABLEAUX	iv
1. INTRODUCTION.....	1
1.1 Contexte de l'étude.....	1
1.1.1 L'Office National des forêts (ONF) : L'Organisme d'accueil.....	1
1.1.2 La BDR et la hauteur dominante.....	2
1.1.3 Light Detection And Ranging (Lidar): Fonctionnement.....	3
1.1.4 Light Detection And Ranging(Lidar): Etat de l'art sur son utilisation en foresterie.....	4
1.2 Objectif.....	6
2. MATERIEL ET METHODE	6
2.1. Méthode d'analyse.....	6
2.2. La zone d'étude	7
2.3. Données	7
2.3.1. Données de terrain.....	7
2.3.2. Données LiDAR.....	8
2.4. Modélisation	10
2.4.1. Régression linéaire simple et multi variable	10
2.4.2. Validation des modèles	11
2.5. Etude de la robustesse.....	11
2.5.1. Effet de l'incertitude de la localisation des placettes	11
2.5.2. Effet de la variation de la densité de points LiDAR	12
2.6. Choix de la variable indépendante.....	12
2.7. Cartographie	13
3. RESULTATS	14
3.1. Modélisation	15
3.1.1. Régression linéaire simple vs régression linéaire multiple	15
3.2. Etude de la robustesse.....	16
3.2.1. Effet de l'incertitude dans la localisation des placettes.....	16
3.2.2. Effet de la densité de points LiDAR	16
3.3. Sélection de la variable prédictive.....	16
3.4. Cartographie	18
4. DISCUSSION	19
4.1. Modélisation	19
4.2. Etude de la robustesse.....	20
4.2.1. Effet de l'incertitude dans la localisation des placettes	20
4.2.2. Effets de la densité de points	21
4.3. Cartographie	22
5. CONCLUSION	23
6. BIBLIOGRAPHIE	24
ANNEXES	27
RESUME	

TABLE DE FIGURES

Figure n° 1 : la technique du LIDAR multi-échos ; A) Schéma d'un vol LIDAR ; B) Impact du laser sur la végétation ; C) Retour du laser et représentation des plusieurs échos (Bock et al., 2011).....	3
Figure n° 2 : A) capture de nuage des points Lidar dans le logiciel MARS7. Les différents couleurs renseignent sur la hauteur de chaque point (rouge plus hauts, bleu plus bas).....	4
Figure n° 3 : Schéma de l'approche « are based » (traduit de White et al., 2013).....	7
Figure n° 4: Nombre des placettes par intervalle de hauteur spécifié.....	8
Figure n° 5 : Processus de normalisation du nuage de points Lidar.....	8
Figure n° 6 : présentation des différents types d'indicateurs LiDAR (métriques) calculés à partir d'un nuage de points LIDAR (emprise de 15 m de rayon colorisée en fonction de la hauteur des points). Les indicateurs de densité (en bleu) sont calculés par tranches fixe (1, 8, 12 m ...) ou par tranche variable (d1, d2 ... d9 correspondant à la hauteur du percentile 10, 20 ... 90 %).	9
Figure n° 7 : Placettes d'essai découpées à partir de nuage de points LiDAR. Extraite et adapté de White et al. (2013)	10
Figure n° 8 : A) Taille des placettes pour lesquelles les métriques LiDAR ont été extrait afin de étudier la robustesse des modèles ; B) Schéma illustratif de la procédure développée pour étudier la robustesse des modèles en variant le centre du rayon original (centre signalé avec une X). On n'a pas désigné toutes les placettes extraites.	11
Figure n°9 : Maillage systématique des placettes carrées permettant de cartographier l'estimateur LiDAR de hauteur dominante ; le point central signale le centre de parcelle ; en orange, périmètre qui définit la surface (200 m ²) des cellules à partir de laquelle la valeur de la métrique a été estimée.	13
Figure n° 10 : Maillage systématique des placettes circulaires disposées en quinconce (méthode 2). En rouge, emprise circulaire (200 m ²) des cellules à partir de laquelle la valeur de la métrique a été estimée.....	14
Figure n° 11 : relation entre la hauteur de prédiction pour l'ensemble de la zone survolé (rouge) et les placettes d'essai (vert). Exemple avec le percentile 95.	14
Figure n° 12 : Effet de la densité de points (par m ²) sur le RMSE (en mètres) selon les bandes du vol utilisé.....	17
Figure n°13 : Variation de RMSE selon le rayon d'extraction utilisé pour le calcul des métriques.....	17
Figure n° 14 : variation du RMSE par métrique quand le centre de la placette est changé d'une façon aléatoire autour du centre original (± 2 mètres) obtenu du GPS différentiel.	17
Figure n° 15 : Effet de la densité de points (par m ²) sur le RMSE (en mètres) des modèles conformément à une réduction progressifs du pourcentage des points des fichiers .las complets.....	17
Figure n° 16 : Détail de la carte de hauteur en appliquant le modèle du H95 dans l'ensemble de la zone d'étude A) avec la méthodologie 2 (placettes circulaires) ; Taille de pixel 8 mètres B) avec la méthodologie 1 (placettes carrées). Taille de pixel 14.14 mètres.....	19
Figure n° 17 Diagramme illustratif du manque de co-régistration parmi les données de terrain et les données du LIDAR. S'il y a un décalage des coordonnées parmi celles du terrain et du Lidar on peut avoir des prédictions erronées (Mauro et al., 2010)	21

TABLEAUX

Tableau n° 1 : Définition des classes BDR selon les critères de l'Agence Bourgogne-Est	2
Tableau n° 2 : densité de points Lidar par rapport à la variation du pourcentage de points des fichiers originaux	12
Tableau n° 3: densité de points Lidar par rapport aux nombre de vols utilisés	12
Tableau n° 4 : Critères de classification pour la stabilité	13
Tableau n° 5 : Critères de classification pour RMSEcv	13
Tableau n° 6 : Reclassification pour le critère stabilité	13
Tableau n° 7 : Reclassification pour le critère RMSEcv	13
Tableau n° 8 : Résumé de paramètres de la régression linéaire multiple et après cross –validation pour trois exemples de modèles évaluées avec le mieux RMSE . Hop =Hauteur dominante de prédiction ; RMSE : erreur quadratique moyenne ; R^2_{ad} : R carré ajusté ; pval_int : p-value ; RMSEcv : erreur quadratique moyenne après cross-validation.	15
Tableau n° 9 : Résumée de paramètres de la régression linéaire simple et après cross –validation pour chaque variable explicative évalué (métrique) ; Hop =Hauteur dominante predite; RMSE : erreur quadratique moyenne ; R^2_{ad} : R carré ajusté ; pval_int: p-value ; RMSEcv : erreur quadratique moyenne après cross-validation ; Effcv : efficience.	15

Tableau n° 10 : résultat de la classification des métriques selon la stabilité et la valeur du RMSEcv; en rose, les métriques sans stabilité et/ou avec une RMSEcv mauvaise; en vert les métriques avec une bonne ou très bonne stabilité et avec un bon ou très bon RMSEcv. _____ 18

Tableau n° 11: moyenne et écart-type des résidus (différence entre la hauteur dominante mesurée sur le terrain et la hauteur dominante prédite selon la carte correspondante) ; M1 et M2 : méthode de cartographie 1 et 2 respectivement. _____ 18

1. INTRODUCTION

Le suivi et la description des **jeunes peuplements forestiers** est important et essentiel pour la qualification et planification de travaux sylvicoles. Afin d'effectuer ce suivi, ONF a créé en 2005 la "**Base de Données Régénérations**" (désignée dans ce rapport par «**BDR**»), qui permet de suivre les flux de régénération et de quantifier les surfaces en travaux dans les jeunes peuplements. Le critère d'évaluation principal est la hauteur des peuplements. Une telle base permet de planifier les travaux, d'évaluer leur coût, ou de calculer des indicateurs du renouvellement des forêts: critère essentiel justifiant la gestion durable des forêts.

Mais, la mesure de la hauteur des arbres sur le terrain est souvent fastidieuse, coûteuse et peu précise. Des erreurs de mesures de 1.5 à 2 m sont fréquemment constatées en forêt (IFN Suisse, 2009). De plus, il existe souvent, au sein de parcelles issues de régénération naturelle, une forte hétérogénéité des hauteurs parfois difficile à appréhender depuis l'intérieur de ces peuplements denses. Actuellement, et dans le cadre de la BDR, ce diagnostic se fait à dire d'expert forestier: c'est-à-dire le plus souvent « à l'œil » (plus ou moins bien calibré ...) ou au mieux à l'aide de placettes statistiques de mesure de hauteur dominante (diagnostic SYLVIE) (Tran and Vallance, 1997). D'autre part, même avec un appareil de mesure bien calibré, il existe un effet observateur non négligeable sur l'appréciation de la hauteur des peuplements. Dans ce contexte, l'utilisation des **LiDAR** (Light Detection and Ranging) pourrait être un plus pour une estimation économique, fiable et sans biais observateur.

En effet le LiDAR est un système actif de télédétection basé sur l'émission et la réception d'un faisceau laser. Beaucoup d'études ont été développées dans les dernières 20 années pour évaluer la ressource forestière à partir du LiDAR (Hudak et al., 2009). Les résultats sont abondants et assez prometteurs, notamment pour l'estimation rapide, précise et complète de la hauteur des peuplements forestiers dont la précision est souvent inframétrique (Holmgren and Honsson, 2004 ; Hall et al., 2005; Bock et al., 2011). Le LiDAR permet d'avoir une bonne approche du couvert de la végétation en offrant une information sur la structure 3D des forêts. Il est ainsi possible de disposer de cartes exhaustives de hauteur des peuplements sur de grande surface, chose qui serait impossible ou très difficile et coûteux à réaliser avec des diagnostics de terrain classiques.

Cependant, la gamme de calibration des modèles existants ne couvrent pas l'ensemble des jeunes peuplements concernées par la BDR. Par exemple, le modèle de Bock et al. (2009) a été calibré dans des hêtraies du Nord Est de la France entre le stade perchis (10-12 m de haut) et le stade de haute futaie (plus de 40 m de haut). Or les peuplements concernés par la BDR concernent tous les peuplements qui n'ont jamais été éclaircies, c'est-à-dire des peuplements entre 0 à 16 m de haut.

Il apparaît donc important d'évaluer, les possibilités de diagnostiquer le stade de développement des jeunes peuplements à l'aide du LiDAR.

L'objectif de ce stage est donc vérifier que le LiDAR permet d'estimer la hauteur de jeunes peuplements.

1.1 Contexte de l'étude

1.1.1 L'Office National des forêts (ONF) : L'Organisme d'accueil

L'ONF, créée en 1964, est un établissement public à caractère industriel et commerciale. Il gère les forêts publiques, domaniales et communales en France et assure donc la gestion durable de 10 millions d'ha de forêts et espaces boisés en métropole et dans les DOM (départements d'outre-mer).

L'ONF dispose d'un service R&D (Recherche et Développement) dont l'objectif est de mettre au point et de développer des outils afin de répondre aux nouveaux enjeux de la gestion

durable des écosystèmes forestiers. C'est dans ce cadre-là que j'ai effectué mon stage dans le pôle de Nancy dont une thématique principale, outre celles liées aux changements climatiques, est la télédétection, notamment LiDAR aussi bien terrestre qu'aérien (ONF, 2013).

1.1.2 La BDR et la hauteur dominante

Dans le cadre d'une gestion en futaie régulière¹ ou d'un suivi de la régénération par surface, la BDR (Base de Données de Régénération) est un outil qui permet le suivi des régénérations, naturelles ou artificielles.

Elle vise à faciliter la programmation des interventions dans les jeunes peuplements ($H < 16$ mètres) et permet d'avoir des éléments d'estimation sur les moyens à allouer en travaux. La BDR permet aussi le suivi et le contrôle du renouvellement des forêts gérées par l'ONF, c'est donc un outil essentiel à la gestion durable. La BDR est fondée sur la description des surfaces à régénérer ou en régénération, inscrites dans différentes classes en fonction de leur hauteur (voir [Tableau n° 1](#)). Les amplitudes des classes sont de plus variables: très rapprochée dans les 1^{ers} stades (30 cm à 50 cm), l'amplitude atteint 3 à 6 m dans les stades les plus avancés.

Dans la BDR, les unités basiques de gestion sont les UES (unité élémentaire de suivi) qui peuvent découper les parcelles de gestion en plusieurs unités homogènes. Le classement des UES s'effectue selon trois principaux critères: la hauteur des semis/plants, l'essence objective du peuplement et le type d'itinéraire sylvicole. En fonction de ces trois critères, les UES en régénération s'inscrivent dans une des cinq classes nationales (Duez, 2011). Mais, chaque agence territoriale de l'ONF peut modifier les classes et les adapter à ses besoins. Pour notre étude,

Classe	Description
0	Inscrite à l'état d'assiette de l'année pour la coupe d'ouverture en régénération (avec ou sans semis)
1	Intervention en travaux / coupe et semis dont la hauteur est inférieure à 30 cm ou des plants ayant moins d'une saison de végétation.
2	$0.3 < H \leq 1.5$ m
3	$1.5 < H \leq 3$ m
4	$3 < H \leq 6$ m
5	$6 < H \leq 12$ m et n'ayant pas bénéficié de la première éclaircie
6	$12 \text{ m} < H$ et n'ayant pas bénéficié de la première éclaircie
7	Dès que la première éclaircie a été réalisée (jusqu'à la première coupe de régénération et son passage en classe 0)

Tableau n° 1 : Définition des classes BDR selon les critères de l'Agence Bourgogne-Est

nous allons utiliser le classement de l'Agence Bourgogne-Est puisque notre site d'étude se situe dans cette région. (Voir Carte annexe 1 et [Tableau n° 1](#)).

La classification des UES se fait en croisant les trois critères mentionnés. Par exemple, une UES classée comme 5HET1 est un peuplement de Hêtre, inscrit dans la classe 5 de la BDR et dont l'itinéraire sylvicole est un itinéraire peuplement (chiffre 1).

La classification des UES se fait en croisant les trois critères mentionnés. Par exemple, une UES classée comme 5HET1 est un peuplement de Hêtre, inscrit dans la classe 5 de la BDR et dont l'itinéraire sylvicole est un itinéraire peuplement.

Le classement BDR se fait actuellement « à l'œil » par le forestier de terrain qui estime des zones avec une hauteur homogène, après avoir, en principe, parcouru la parcelle forestière. Cette méthode est en partie satisfaisante car l'œil du forestier permet d'intégrer plusieurs paramètres : la hauteur (plus ou moins bien évaluée), l'essence objectif (qui peut se trouver en dessous d'un étage dominant concurrent) et le type d'intervention (qui dépend de la station, du niveau de concurrence constaté et de l'autécologie des essences présentes). En revanche, ce diagnostic peut conduire à des biais d'appréciation de stade de développement quand il est fait un peu trop rapidement, parfois de manière simplifiée « depuis le bureau », ou dans le cas

¹ Futaie régulière : groupes des arbres issues des graines ou plants et dont la majorité ont la même âge.

de peuplements très hétérogènes. D'autre part, il y a un « effet observateur » non négligeable pour apprécier la notion d'homogénéité d'un peuplement, pour cartographier ces peuplements et les caractériser en hauteur.

Ce critère de classement BDR à partir de la hauteur des peuplements est une notion qui demande à être précisée. La hauteur qui permet le plus de répondre aux objectifs de la BDR est la notion de hauteur dominante, c'est-à-dire, la moyenne des hauteurs de la strate dominante et co-dominante compris comme la moyenne de la hauteur des 100 plus gros arbres à l'hectare (Pardé and Bouchon, 1988). C'est cette hauteur qui guide les interventions en travaux et les éclaircies dans les guides sylvicoles ce sera donc cette hauteur qui sera prise en compte pour le classement BDR de notre étude.

Cependant, il y a des limitations par rapport à l'utilisation de la définition de hauteur dominante dans le cadre de jeunes peuplements. Par exemple, ils ont une disposition favorable à changer de statut rapidement du fait de leur grande variabilité d'accroissement en hauteur entre les premières et les dernières années. L'existence des préexistants dans les parcelles en régénération est habituel, surtout dans le cas du hêtre. Les préexistants sont des arbres qui appartiennent aux plus gros mais qui ne seront pas parti de la strate dominante au final de la rotation. Ou l'existence des espèces avec une dynamique différente à celle de l'essence objective qui peuvent être part aussi dans les premiers stades de la strate dominante et co-dominante.

1.1.3 Light Detection And Ranging (Lidar): Fonctionnement

Le LiDAR est une technologie qui permet de déterminer la distance depuis un émetteur de faisceau alumineux (Laser) jusqu'à un objet en utilisant le temps de retour des échos². Le principe de fonctionnement est exactement le même que celui du RADAR, mais les ondes de radio sont remplacées pour des ondes lumineuses (Young, 2011).

Dans cette étude les données proviennent d'un LiDAR aéroporté (ou ALS pour ses sigles en anglais « Airborne laser scanning »), c'est-à-dire avec un émetteur laser embarqué dans un avion. Un système ALS est constitué d'un vecteur aérien (un avion), d'un laser transmettant les ondes, d'un récepteur (assurant la réception de la lumière rétrodiffusée), d'un système de géoréférencement avec un GPS différentiel (global position system) et une centrale inertielle de navigation INS (inertial navigation system) ou IMU³ (inertial measurement unit).

Les coordonnées et la direction exacte de l'émission de chaque impulsion sont connues, en tenant en compte l'information fournie par le système de géoréférencement. Le capteur calcule l'espace parcouru par chaque impulsion et la triplette de coordonnées (X, Y et Z) de chaque retour. Au final, les données se présentent sous la forme d'un nuage de points en 3D (longitude, latitude, altitude) (Voir Figure n° 2). La quantité d'énergie du signal reçu ou

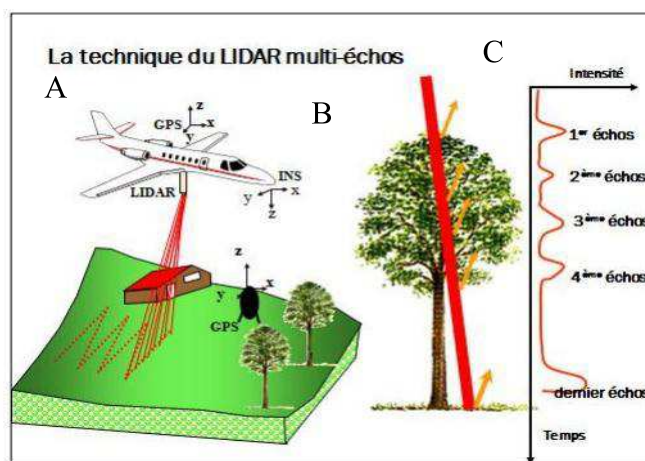


Figure n° 1 : la technique du LIDAR multi-échos ; A) Schéma d'un vol LIDAR ; B) Impact du laser sur la végétation ; C) Retour du laser et représentation des plusieurs échos (Bock et al., 2011).

² Retours ou pics d'énergie pour une même impulsion laser

³ Une IMU est un système permettant de mesurer en temps réel l'inclinaison de l'avion par rapport au plan horizontal (roulis, tangage)

intensité, peut également être enregistrée. Elle renseigne sur la nature du matériau renvoyant l'impulsion.

Le LiDAR est basé sur l'émission d'impulsions. Il est capable d'enregistrer soit plusieurs retours (échos) pour chaque impulsion. C'est ce type de LiDAR à enregistrement d'échos discrets (LiDAR multi-écho) qui a été utilisé dans cette étude. Il existe des systèmes LiDAR qui enregistrent le retour d'onde complet: on parle alors de Full Wave Form (Lim et al., 2003). Le Full Wave Form ne sera pas abordé dans le cadre de ce stage car les études sur ce type de capteur sont encore très récentes et sont du domaine de la recherche plus fondamentale.

Dans le cas du LiDAR multiéchos, une impulsion laser est émise à très haute fréquence (100-150 KHz) depuis l'avion (voir [Figure n° 1](#)). Le signal intercepte une première strate de végétation (une feuille ou une branche, par exemple). Une partie du signal est reflétée en direction du récepteur LiDAR et enregistrée comme le premier retour. La partie d'énergie qui n'a pas été reflétée continue jusqu'à une autre strate qui peut à son tour renvoyer tout ou partie de l'énergie incidente. Ainsi, la structure de la végétation est décrite de façon discrète en captant l'information tridimensionnelle de différentes couches de végétation et du sol. Le nombre des retours dépend de la surface. Sur les surfaces compactes, sans végétation, on obtient un seul écho. Mais sur la végétation le capteur est capable d'enregistrer plusieurs échos (entre 3 et 5 selon les derniers modèles les plus utilisées) pour une impulsion laser émise.

On considère que les premiers retours proviennent de la cime de la couverture végétale, alors que les derniers sont classifiés le plus souvent comme points sol (White et al., 2013).

Avec les points sol, on peut construire un MNE (modèle numérique d'élévation), c'est-à-dire, une structure numérique de données qui représente la distribution spatiale de l'altitude de la surface du terrain. Avec les altitudes des points végétations un MNS (modèle numérique de surface) peut être construit aussi. Finalement, par soustraction de MNS-MNE on obtient un modèle numérique de hauteur de la végétation (MNH). On peut donc connaître la hauteur du couvert par rapport au sol (H).

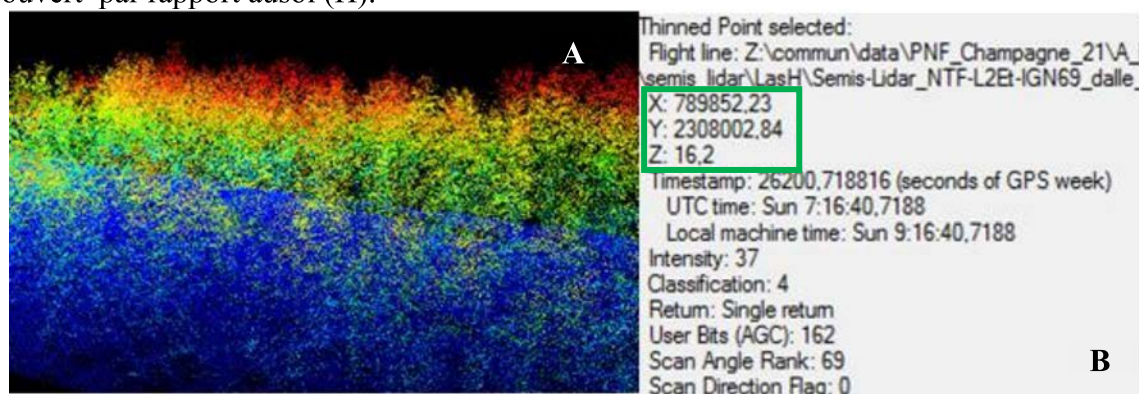


Figure n° 2 : A) capture de nuage des points Lidar dans le logiciel MARS7. Les différents couleurs renseignent sur la hauteur de chaque point (rouge plus hauts, bleu plus bas).

B) information basique associé à chaque point dans une couche qui a été déjà normalisée (coordonnées Z sous le sol).

1.1.4 Light Detection And Ranging(Lidar): Etat de l'art sur son utilisation en foresterie

Plusieurs variables dendrométriques ont déjà été estimés et modélisés et ceci à différentes échelles, avec différentes méthodologies et appliqués aux contextes locaux concrets avec une grande précision et exactitude (Treitz et al., 2012). Les variables les plus souvent modélisées sont : la biomasse et le volume (Nelson et al., 1998; Naesset, 2011), la surface terrière (Means et al., 2000 ; Hudak et al., 2006 ; Chen et al., 2007) la taux de couvert (Estornell Cremades, 2011), et bien sûr, la hauteur (Naesset and Bjercknes, 2001 ; Heurich and Thoma, 2008; Bock

et al., 2009) Les meilleurs résultats pour la hauteur avec coefficients de corrélation r^2 supérieures à 0.85 et RMSE (erreur quadratique moyenne) de moins de 1 mètre (Bock et al., 2009 ; Treitz et al., 2012) mais pour la biomasse ou le volume, par exemple, on a des bons résultats qui se trouvent parmi le 0.75 et 0.8 de coefficient de corrélation r^2 (Næsset, 2011; González-Ferreiro et al., 2012). D'une manière générale, ces études ont porté en majorité sur des peuplement de résineux (Holmgren and Jonsson, 2004; González-Ferreiro et al., 2012) mais il y a quelques études sur les feuillus aussi (Heurich and Thoma, 2008 ; Bock et al., 2009) (Voir annexe 6).

En 2001, (Næsset and Bjercknes, 2001) ont proposé une procédure en deux étapes afin d'obtenir l'estimation de la hauteur des peuplements à partir de données LiDAR en utilisant des placettes de jeunes peuplements. Ils obtiennent une modèle logarithmique avec un RMSE⁴ de 3.7 mètres et un r^2 de 0.83. Dans une étude suivante, Næsset (2002), a testé une procédure semblable pour plusieurs variables pour des forêts adultes et à grande échelle et avec des résultats améliorées, avec un RMSE⁵ de 1.07 m et un r^2 de 0.93 (également avec un modèle logarithmique).

Cette procédure est connue comme l'approche surfacique (area-based). Il s'agit d'installer un échantillon de placette de calibration sur l'ensemble de la zone d'étude de façon aléatoire ou de façon stratifiée (pour couvrir l'ensemble de la variabilité de la variable à prédire). Les mesures de référence (diamètre, hauteur, position des arbres ...) sont réalisées sur le terrain, en général sur des placettes circulaires. La position est enregistrée aussi précisément que possible au GPS. Le nuage de points LiDAR est alors extrait sur l'emprise exacte de la placette. Des métriques, c'est-à-dire généralement des statistiques de distribution des points LiDAR, sont alors calculées. Des modèles mathématiques sont ensuite établis pour prédire les variables dendrométriques à partir des métriques LiDAR. Dans la deuxième étape, ces modèles sont appliquées à l'ensemble de la zone d'étude afin d'estimer la variable et faire la cartographie correspondante (White et al., 2013) (Voir Figure n° 3).

Cette méthodologie a été utilisée depuis les dernières dix années pour estimer et cartographier différents variables. Dans le cadre de la première étape (estimation des variables), des bons résultats ont été obtenus pour la hauteur dominante avec des r^2 supérieurs à 0.8 et des RMSEs de moins de 2 mètres (Næsset, 2004 ; Heurich and Thoma, 2008 ; Næsset, 2009; Zaldo et al., 2010; González-Ferreiro et al., 2012). Parmi des métriques habituellement utilisées comme les variables explicatives des modèles on trouve : les percentiles de hauteur⁶(Næsset and Bjercknes, 2001 ; Holmgren and Jonsson, 2004 ;Heurich and Thoma, 2008), les hauteurs des points maximaux et moyens (Gobakken and Næsset, 2009; Treitz et al., 2012), les coefficients des variation ou les densités des points verticaux (Næsset, 2004), ou des indicateurs spatiaux de maximum local (Bock et al. 2009).

Certaines études montrent que les spécifications et la configuration du laser et du vol ainsi que la planification d'échantillonnage du terrain jouent un rôle très important dans les interactions entre le couvert forestier et le signal LiDAR (White et al., 2013), la précisions des estimations ou le coût du projet. Par exemple, une haute densité de points LiDAR nécessite des plans de vol plus longs donc, le coût d'acquisition augmente. La hauteur haute du vol influe sur la distribution du nuage des points sur la partie haut du couvert (Næsset, 2004). Quand l'angle du scan laser augmente le nombre de retours du sol diminue (Lovell et al., 2005). L'introduction de placettes d'essai larges fait diminuer le biais introduit par l'erreur de positionnement du GPS (Mauro et al., 2010). L'utilisation des systèmes GPS très précis réduit

⁴ Valeur du RMSE obtenu après transformation inverse pour revenir en mètres parce que le modèle original est sous la forme logarithmique. Il existe un biais associé à cette de-transformation (factor correction= $e^{\text{RMSE}^2/2}$).

⁵ De-transformation inverse du RMSE aussi.

⁶ Hauteur du point situé au quantile x. Ainsi le percentile x% est la limite de hauteur qui permet de séparer les x% de points les plus bas

le erreur de positionnement aussi mais il peut faire augmenter d'une façon exponentielle le coût des inventaires (White et al., 2013).

A l'Office National des Forêts (ONF) le LiDAR a été utilisé la première fois en 2007. Différentes études successives ont permis de mettre au point un modèle qui prédit la hauteur dominante avec une RMSE de prédiction de 75 cm (Dez, 2008 ; Martins, 2009 ; Bock et al., 2009), un modèle assez robuste qui n'était pas influencé par l'effet densité de points, la localisation des placettes ou le type de MNE utilisé en fonction de différents interpolations des points sol utilisés. Ce modèle utilise un indicateur spatial qui est la moyenne des 6 maximums locaux obtenus après application d'un rayon d'exclusion dépendant de la hauteur (Bock et al. 2009). Ce modèle est relativement générique puisqu'il a été testé sur d'autres sites feuillus et résineux avec des erreurs de l'ordre de 1 m à 1.5 m (Diop, 2011). Cependant, la gamme de calibration de ces modèles se situe entre 10 et 40 m de haut. Les jeunes peuplements sont donc très peu pris en compte et les jeunes régénérations aux stades semis à gaulis ne sont pas représentés. Or, l'estimation de hauteur dans les semis et gaulis est primordiale pour une utilisation dans le cadre de BDR.

1.2 Objectif

Le but principal de cette étude est de mettre au point une méthodologie pour estimer la hauteur dominante des jeunes peuplements forestiers, à partir d'un nuage de points LiDAR, notamment pour pouvoir les classer dans la BDR. Pour cela, un modèle mathématique de prédiction de la hauteur dominante des peuplements à partir du nuage sera calibré grâce à des placettes de mesure sur le terrain. Ce travail de modélisation sera discuté grâce à une analyse de sensibilité qui permettra de définir le domaine de validité de ce modèle. Il s'agira également de mettre au point une méthode de cartographie pour fournir au gestionnaire des cartes de hauteur dominante à partir de données lidar acquises sur une grande surface.

Afin de mener à bien cet objectif général, le déroulement du stage a été le suivant :

- Prise de contact avec la définition des enjeux du stage (BDR, LiDAR, variabilité des jeunes peuplements....).
- Travail de terrain avec la mise au point d'un protocole et d'un plan d'échantillonnage.
- Analyse statistique afin de trouver le meilleur modèle de prédiction de la hauteur estimée sur le terrain à partir d'indicateurs sélectionnés provenant du nuage de point lidar.
- Analyse critique des limites de la méthode, définition de son domaine d'application.
- Cartographie selon les critères définis antérieurement.

2. MATERIEL ET METHODE

2.1. Méthode d'analyse

La méthode générale à suivre est basée sur la démarche based-area (surfactive) en deux étapes proposée par Naesset (2002) (Voir Figure n° 3). 101 placettes de calibration ont été implantées, sur la zone d'étude (voir carte annexe 2), selon un plan d'échantillonnage décrit dans le point 2.3.1 et l'Annexe 1 afin de développer une relation empirique entre le nuage de points LiDAR et la hauteur dominante mesurée sur le terrain. Le principe consiste alors à calibrer un modèle linéaire pour prédire la hauteur dominante à partir de métriques de points LiDAR acquis sur l'ensemble de la zone d'étude. Si l'analyse de sensibilité de ce modèle se révèle satisfaisante, il sera appliqué sur l'ensemble de la zone d'étude pour disposer d'une carte de hauteur sur les parcelles avec des jeunes peuplements.

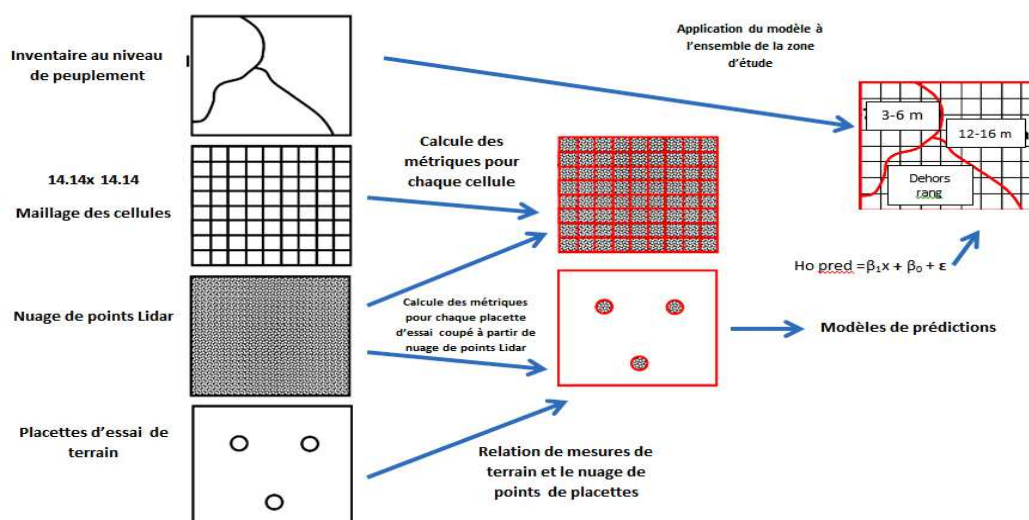


Figure n° 3 : Schéma de l'approche « are based » (traduit de White et al., 2013)

2.2. La zone d'étude

La zone d'étude se situe dans le Département de Cote d'Or (voir Carte annexe 1), dans la région naturelle de la Montagne Bourguignonne. La zone a été survolée en mars de 2012 à l'occasion du projet de création du projet de Parc National Forestier. C'est donc la disponibilité de données LIDAR qui nous a guidé dans le choix de ce site.

La majorité des placettes a été placées sur la Forêt Domaniale de Châtillon. La Forêt Domaniale a une surface de 8894,87 ha, dont l'altitude est comprise entre 258 et 457 m. Le hêtre est l'essence majoritaire ; il représente 54% du couvert, les chênes sessile et pédoncule 33%, les autres feuilles 6%, en enfin les résineux 3%. Majoritairement ces parcelles sont traitées en futaie régulière.

2.3. Données

2.3.1. Données de terrain

Les données du terrain ont été mesurées en essayant d'obtenir des mesures les plus précises possibles. Elles nous serviront à calibrer et à valider les modèles

Au total 101 placettes ont été installées, dont 22 en forêt communale afin d'avoir dans notre échantillon une bonne répartition de modes de gestion forestière qui peuvent être différents en fonction du propriétaire (commune ou état) (Voir Carte annexe 2).

Plan d'échantillonnage et phase de terrain

On a conçu un échantillonnage stratifié en fonction des groupes de hauteur dont la hauteur varie entre 3 et 16. La limite de 16 mètres répond à une raison sylvicole. Selon les critères de la BDR Bourguignonne, la « sortie » d'un peuplement de la classe 6 arrive au moment où la première éclaircie est réalisée. Cela se fait, normalement, à une hauteur dominante de 14 à 16 mètres (Office National des Forêts, 2005). D'un autre côté, la limite de 3 mètres est fixée par les traitements sylvicoles locaux. Les régénérations de hêtre se font en plusieurs coupes et sur plusieurs années, au final les derniers semenciers sont prélevés lorsque la hauteur dominante de la régénération a atteint 3m. Il y a donc très peu de peuplement complètement ouvert avec une hauteur dominante de la régénération inférieure à 3m.

La localisation des placettes d'essai dans les placettes en régénération a été choisie en utilisant les photos aériennes du terrain, une pre-carte de hauteur raster qui a été faite après pré-traitement de données LiDAR et la vérification visuelle sur le terrain. La carte de hauteur a été faite par pixels de 5 mètres, en calculant le percentile 95 de hauteur à partir des données

LiDAR disponibles (voir section 2.3.2). Le percentile a été cartographié par intervalles de hauteur.. Cependant, l'utilisation des photos aériennes et la carte du percentile 95 n'est pas très précise donc la vérification finale pour localiser les intervalles de hauteur qui nous intéressent a dû être faite sur le terrain. Etant donné que les unités de gestion en régénération ne sont pas très nombreuses, on a pu localiser jusqu'à 3 placettes d'échantillonnage sur les parcelles en régénération échantillonnées (Voir carte Annexe 2).

Les peuplements en cours de récolte, c'est-à-dire avec un mélange de semenciers et de régénération présents sur la même parcelle, ont été volontairement écartés de l'étude en raison des difficultés d'appréciation de la hauteur dominante dans ces peuplements multi-étagés.

Les variables dendrométriques mesurées sont: la hauteur, le diamètre, l'essence des arbres, la densité par placettes. Les coordonnées du centre de chaque placette ont été relevées au GPS différentiel Trimble GeoXT. Après correction différentielle, en post processus, l'erreur de positionnement attendue est d'entre 1 et 4 mètres (DTCB- R&D Pôle de Chambéry, 2012). C'est à cause de cette erreur qu'on a cherché des peuplements homogènes donc les placettes de terrain ont été sélectionnées dans un environnement estimé homogène sur environ 30 m de rayon.

Pour l'intérêt de l'étude, ci-dessous sont détaillés les critères de mesures référant à la hauteur dominante. Pour avoir plus de détails sur les critères de mesure des autres variables consulter l'annexe 1.

La hauteur dominante

La hauteur dominante, considérée comme la moyenne des arbres de la strate dominante, s'estime de différentes façons selon la définition exacte et la méthode de mesure utilisée. Ici la méthode utilisée sera celle qui est la plus couramment utilisée en France : ce sera donc la moyenne des 100 plus gros arbres par hectare (Pardé and Bouchon, 1988). L'ONF utilise une estimation traditionnelle non biaisée de la hauteur dominante qui consiste à mesurer sur une placette les $n-1$ plus gros arbres sur n ares (Pardé and Bouchon, 1988).

On va travailler pour l'ensemble de calculs, avec la hauteur de l'arbre le plus gros de la placette (de 200 m²). Dans le 90% des cas l'essence cet arbre est le hêtre. La Figure n° 4 présente, le nombre des placettes par groupes de hauteurs. On a un nombre des placettes relativement constant le long de l'intervalle défini entre 3 et 16 mètres.

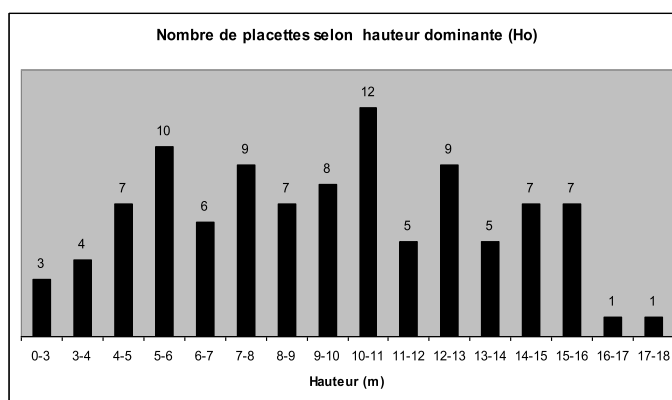


Figure n° 4: Nombre des placettes par intervalle de hauteur

2.3.2. Données LiDAR

Le vol LiDAR, permettant la prise de données, a eu lieu sur trois jours en mars 2012. La densité moyenne de points émis (« first return ») est de 14.5 par mètre carré.

Les points ont été classés par le prestataire (l'entreprise qui a effectué le vol LiDAR pour le parc naturel) comme points végétation ou points sol en

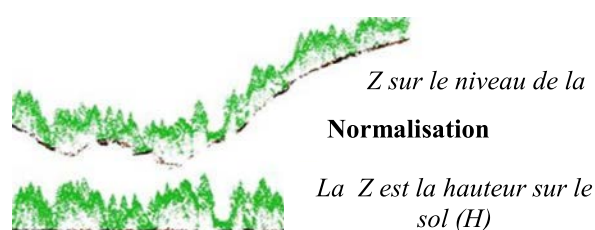


Figure n° 5 : Processus de normalisation du nuage de points Lidar

utilisant le logiciel TerraScan⁷. La surface de la zone couverte est de 381 km². (Voir carte annexe 1 et annexe 3)

Traitement de données LiDAR

Notre intérêt réside dans la mesure de la hauteur. Une étape importante est la normalisation du nuage de points afin de transformer la coordonnée Z (sur le niveau de la mer) dans une donnée H (hauteur sur le sol). Cette opération a été réalisée avec le logiciel Lastools⁸. En connaissant des points sol, logiciel réalise un TIN (Triangulated Irregular Network) avec les points sol. Il soustrait ensuite l'altitude de chaque point végétation à l'altitude du TIN à l'aplomb du point donc on connaît la hauteur sur le sol de chaque point végétation (Voir Figure n° 5).

Calcul des métriques du nuage de points LiDAR sur les placettes de mesure

A partir du nuage de point LiDAR normalisé, plus de 500 indicateurs (métriques) ont été calculés (voir Figure n° 1 et annexe 2) à l'aide de Lasmanager ONF⁹, pour les différents types d'échos (First, Last, Only, All, sol ou végétation uniquement ...).

- des indicateurs de densités calculées dans différentes tranches de hauteur variable ou fixe,
- des indicateurs de distribution de hauteur (Coefficients de variation, Skewness, ...)
- des statistiques classiques (moyenne, écart-type, ...).
- des indicateurs spatiaux (voir annexe 2)

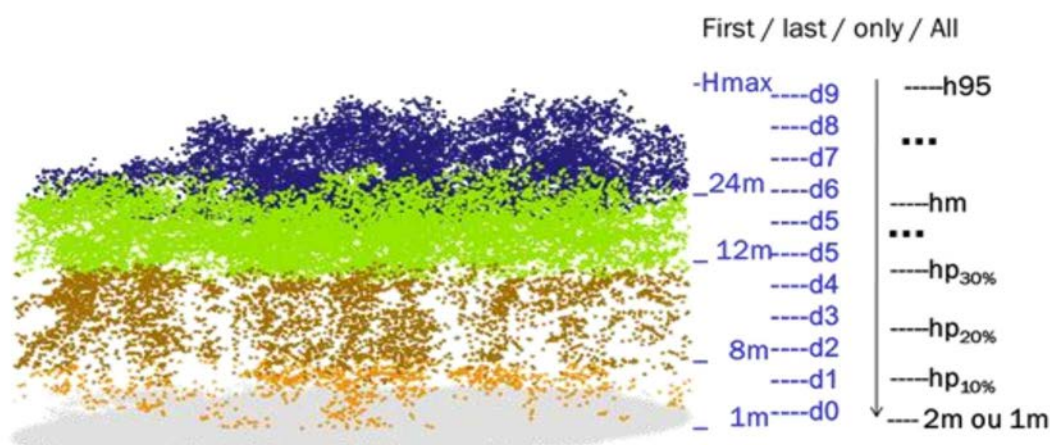


Figure n° 6 : présentation des différents types d'indicateurs LiDAR (métriques) calculés à partir d'un nuage de points LIDAR (emprise de 15 m de rayon colorisée en fonction de la hauteur des points). Les indicateurs de densité (en bleu) sont calculés par tranches fixe (1, 8, 12 m ...) ou par tranche variable (d1, d2 ... d9 correspondant à la hauteur du percentile 10, 20 ... 90 %).

Les indicateurs spatiaux testés sont dérivés de l'indicateur Hmv6_Bd mis au point en Forêt Domaniale (FD) de Haye et sur la FD de Languimberg lors de 2 précédents stages (Dez, 2008 ; Martins, 2009). Pour mémoire, le Hmv6 est la hauteur moyenne des 6 plus hauts points de la placette, sélectionnés en tenant compte d'un rayon d'exclusion simulant la taille des couronnes des arbres (Bock et al. 2009). Ce rayon (Re) est calculé d'après des relations allométriques reliant la hauteur de l'arbre (H) à son diamètre de houppier. Cette relation a été

⁷ Application de la famille des logiciels Terrasolid pour le traitement des nuages des points LiDAR.

⁸ Lastools est un ensemble des outils pour le traitement des données LiDAR. Il dispose d'une variété des extensions pour le maniement, l'analyse, le développement et la publication de fichiers .LAS.

⁹ Logiciel développé à l'ONF qui permet notamment le calcul des indicateurs utilisés dans la modélisation de H0 à partir de fichiers *.las ou Grid (Pierre-Yves, 2012).

établie sur plus de 800 hêtres et chênes mesurés par l'INRA¹⁰ ($r^2 = 0.54$) selon la formule : $Re = \exp[0.80 - 1.27 \cdot \log(H)]$.

Les nuages de points LiDAR avec la hauteur de chaque point a été extrait sur l'emprise exacte de chaque placette à l'aide du logiciel Lasmanager. La taille initiale a été exactement la même que celles des placettes d'échantillonnage (8 mètres de rayon). Cependant pour l'évaluation de la robustesse du modèle (voir section 2.5), on a fait le même processus mais en variant la taille ou la localisation du centre des placettes.

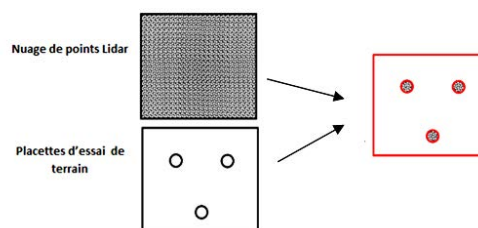


Figure n° 7 : Placettes d'essai découpées à partir de nuage de points LiDAR. Extraite et adapté de White et al. (2013)

C'est sur ces nuages de points découpés, à partir du nuage global, que le calcul des métriques pour chacune de ces placettes a été effectué. L'annexe 2 décrit l'ensemble des métriques calculées.

2.4. Modélisation

Au vu de la bibliographie et du comportement qu'on imagine de la hauteur vis-à-vis des métriques LiDAR la modélisation a été basé sur un modèle linéaire.

Nos critères généraux pour la sélection du modèle ont été : **la simplicité** (principe d'économie¹¹), **la robustesse** (afin de travailler avec des métriques stables) et l'obtention d'une **RMSE** (erreur quadratique moyenne) **petite**. Le RMSE est utile pour la comparaison de plusieurs estimateurs puisqu'il permet de montrer la différence entre les valeurs de prédiction du modèle et les valeurs observées (données de terrain) aux mêmes unités de mesures et/ou en pourcentage. Plus la RMSE est faible, plus la valeur prédite par le modèle est bonne.

Dans ce cadre-là, on a testé la régression linéaire simple et la régression linéaire multiple, et ceci avec nos différents métriques. Pour ces analyses on a utilisé le logiciel R (logiciel libre de traitement des données et d'analyse statistiques).

2.4.1. Régression linéaire simple et multi variable

On a fait des modèles linaires simples ($Y = \beta_1 X + \beta_0 + \epsilon$) où la variable explicative X chacune des métriques obtenues à partir du logiciel Lasmanager. Ensuite, on a trié en fonction des RMSE obtenus et on a fait une sélection de quelques métriques qui pouvaient devenir la variable explicative de notre modèle finale. Les critères pour cette sélection ont été:

1. Ne que sélectionner une métrique par famille. Par exemple, une seule parmi: H95, H95_fst, H95_fst_nosoil, H95_fstlst_nosoil, H95_fst_sup1m, H95_fst_sup2m, H95_lst, etc) (Voir Annexe 2 - description des familles de métriques).
2. Choisir des métriques donnant un modèle avec une RMSE petite (< 1 mètre).
3. Choisir des métriques qui ont été déjà utilisées pour l'estimation de la hauteur selon la bibliographie avec des résultats acceptables (en fonction du R^2 et RMSE).

Pour la régression multivariable ($Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_0 + \epsilon$), on a eu besoin de chercher des variables explicatives de différente nature, c'est-à-dire, variables de différent signification (métriques de hauteur, de densité, taux de pénétration, etc). Pour cela, une ACP (analyse de composante principaux) a aussi été utilisé afin de chercher les métriques qui ne sont pas corrélés entre eux. Ensuite on a calculé quelques modèles avec une procédure « stepwise » (ou régression pas à pas) sur R.

¹⁰ Institut national de Recherche agronomique française.

¹¹ Les hypothèses suffisantes les plus simples sont les plus vraisemblables

A chaque étape de la procédure, on examine à la fois si une nouvelle variable doit être ajoutée selon un seuil d'entrée fixé, et si une des variables déjà incluses doit être éliminée selon un seuil de sortie fixé. Cette méthode permet de retirer du modèle d'éventuelles variables qui seraient devenues moins indispensables du fait de la présence de celles nouvellement introduites. La procédure s'arrête lorsqu'aucune variable ne peut être rajoutée ou retirée du modèle selon les critères choisis.

2.4.2. Validation des modèles

Pour la validation des modèles une validation croisée (cross validation) a été réalisée. Il s'agit d'une méthode d'estimation de fiabilité d'un modèle fondé sur une technique de tirage et de remise aléatoires dans l'échantillon de départ. En effet comme le nombre de placettes n'était pas suffisamment important, toutes les placettes ont été utilisées pour la calibration et la validation des modèles.

Il y a différents types mais on a utilisé la 10 fold cross-validation sur R. On a divisé 10 fois l'échantillon, puis on a sélectionné un des 10 échantillons comme ensemble de validation et les (10-1) autres échantillons constitueront l'ensemble d'apprentissage. On calcule l'erreur quadratique moyenne. Puis on répète l'opération en sélectionnant un autre échantillon de validation parmi les (10-1) échantillons qui n'ont pas encore été utilisés pour la validation du modèle. L'opération se répète ainsi 10 fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. La moyenne des 10 erreurs quadratiques moyennes est enfin calculée pour estimer l'erreur de prédiction (RMSEcv).

2.5. Etude de la robustesse

La robustesse d'un modèle est sa capacité à ne pas être altéré/modifié par une petite variation dans les données ou dans les paramètres du modèle choisi pour l'estimation. On recherche donc les métriques permettant de construire les modèles les plus stables.

2.5.1. Effet de l'incertitude de la localisation des placettes

On a utilisé un GPS qui a une erreur de positionnement compris entre 1 et 4 mètres sous couvert forestier (DTCB- R&D Pôle de Chambéry, 2012). Donc la correspondance complète entre les placettes de mesures et les nuages des points découpés, ne peut pas être assurée. C'est pour cela qu'il est important d'étudier les métriques les plus stables dans l'environnement de nos placettes (il est rappelé que les placettes de terrain ont été sélectionnées dans un environnement estimé homogène sur environ 30 m de rayon).

Les métriques les plus stables auront la priorité lors de la sélection des variables explicatives du modèle. On a testé deux méthodes :

- On a extrait les métriques pour différents rayons (Voir Figure n° 8 A) en maintenant le même centre de la placette qu'on a obtenu par le positionnement GPS. On a exporté les métriques et on a calculé les modèles linéaires simples en obtenant une nouvelle RMSEcv pour chacun.

- On a fait varier (dans une limite de 2 m) les coordonnées X et Y du centre des placettes d'une façon aléatoire (15 fois). La taille du rayon est de 8 mètres (Voir

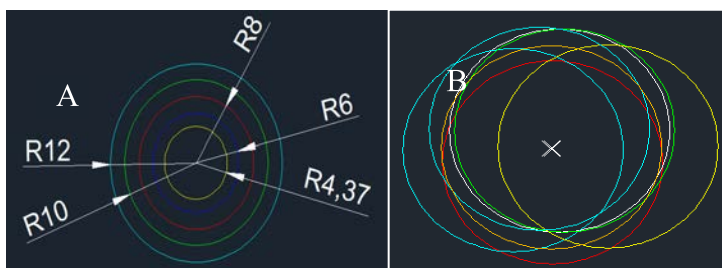


Figure n° 8 : A) Taille des placettes pour lesquelles les métriques LiDAR ont été extrait afin de étudier la robustesse des modèles ; B) Schéma illustratif de la procédure développée pour étudier la robustesse des modèles en variant le centre du rayon original (centre signalé avec une X). On n'a pas désigné toutes les placettes extraites.

Figure n° 8 B). Pour chaque nouvelle placette on a fait la même procédure, c'est-à-dire, on a extrait les métriques et les caractéristiques des modèles linéaires ont été de nouveau calculées.

La variable dépendante est la hauteur dominante (Ho) obtenue de l'échantillonnage du terrain à partir des placettes originales de 8 mètres. Si les peuplements sont homogènes, on considère que la variation des métriques ne doit pas être très significative et rester constante autour des zones adjacentes. Le paramètre évalué pour la comparaison est le RMSEcv (erreur quadratique moyenne après la validation croisée).

2.5.2. Effet de la variation de la densité de points LiDAR

La densité de points conditionne grandement le coût de la campagne LiDAR, une densité de points élevés nécessitant un temps de vol important (Bock et al., 2011). C'est pour cela qu'il est intéressant de tester sa variation pour voir l'influence dans les résultats de la modélisation. La quantité de points existants (14.5 pts/m²) permet de faire une comparaison en faisant varier cette densité.

Deux approches ont été utilisées pour faire varier la densité de points. La première approche consiste à éclaircir le nuage de point LiDAR en enlevant aléatoirement 1 point sur X. Cependant, cette méthode n'est pas complètement réaliste puisque la zone a été couverte par 3 croisements de plan de vol successif. Cette méthode présente l'inconvénient de sélectionner des points dans toutes les bandes de vol. Pour être plus réaliste, nous avons sélectionné sur les placettes, les points issus du premier vol, puis nous avons rajouté les points issus du deuxième vol croisé et du troisième vol croisé. Ainsi, nous disposons de 3 densités de points correspondant réellement à des plans de vols différents (Voir [Tableau n° 2](#) et [Tableau n° 3](#)). Puis la qualité prédictive des modèles a été comparée sur la base de la RMSEcv.

Densité (%)	10	20	30	40	50	60	70	80	90	100
Points/m2	1,5	2,9	4,31	5,8	7,2	8,7	10,2	11,75	13,3	14,65

Tableau n° 2 : densité de points Lidar par rapport à la variation du pourcentage de points des fichiers originaux

Bande Vol	1	2	3
Points/m2	4,8	9,5	14,7

Tableau n° 3 : densité de points Lidar par rapport aux nombre de vols utilisés

2.6. Choix de la variable indépendante

Afin de sélectionner la meilleure métrique et d'avoir un modèle robuste et stable, **quatre situations différentes** avec à chaque fois **deux critères** de jugement ont été utilisés :

-**Les quatre situations** sont issus de la variation du rayon des placettes, de la variation de la position du centre des placettes, de la réduction du nombre de point dans le nuage LIDAR et de la variation des bandes de vols utilisées (résultats dans les [Figures n°12, 13 14 et 15](#)).

-**Les deux critères** sont la valeur moyenne de la RMSEcv (si elle est petite ou grande) ainsi que sa stabilité associée à ces situations. Concrètement la stabilité ont été classée en différents groupes de couleurs pour chaque métrique selon si le comportement de son RMSEcv observé est très stable (les RMSEcv obtenus dans la situation ne sont pas statistiquement différents), stable, moyennement stable ou s'il n'y a pas de stabilité (Voir [Tableaux 4 et 5](#)). De même la valeur moyenne, dans la situation testée, de la RMSEcv a été classée (Tableau n°4)

Puis des points ont été attribués à chaque couleur. Enfin une somme (ce qui fait une note sur 20) a été faite permettant de résumer le comportement de la métrique pour les quatre situations étudiées (Voir [Tableau n° 6 et 7](#)).

Stabilité	ID	Critère de classification
Très stable		Tous les RMSE
Stable		Il y a un RMSE différent significativement
Stabilité moyenne		Il y a deux RMSE différents ou un très différent
Pas de stabilité		3 ou plus RMSE différents

Tableau n° 4 : Critères de classification pour la stabilité

RMSEcv	ID	(m)
Petite	P	<0.8
Moyenne	M	0.8-1
Grand	G	>1
Variable	V	

Tableau n° 5 : Critères de classification pour RMSEcv

Deuxièmement, chaque couleur a été reclassifiée. C'est-à-dire, à chaque couleur on a associé une valeur et on a sommé l'ensemble des valeurs pour la métrique. On a classifié selon différents intervalles en fonction de la somme obtenue (haute ou basse) pour le critère étudié (Voir Tableau 6 et 7).

Id	PTs	Classif	Somme
Vert	5	Très bonne	>18
Jaune	3	Bonne	15-18
Orange	2	Moyenne	12-15
Rouge	0	Pas de stabilité (PS)	<12

Tableau n° 6 : Reclassification pour le critère stabilité

Id	PTs	Classif	Somme
P	5	Très bonne	>18
P M	4	Bonne	14-18
M	3	Moyenne	9-13
G	2	Nulle	<9
V	0		

Tableau n° 7 : Reclassification pour le critère RMSEcv

Une fois qu'on a choisi la variable indépendante et avant d'appliquer les modèles à la cartographie, un diagnostic du modèle est réalisé afin de vérifier les hypothèses de la régression linéaire : la linéarité (test Reset), l'homoscédasticité (test Breush-Pagan), la normalité de résidus (test Shapiro-Wilks) et l'existence des points extrêmes (test de Bonferroni).

2.7. Cartographie

Afin de cartographier la hauteur dominante dans la zone survolée (381 km²), le modèle sélectionné sera utilisé pour calculer la hauteur de prédiction.

On va tester deux méthodologies différentes.

La **méthode 1** : à partir des fichiers .las du nuage de points LiDAR normalisés, le logiciel Lasmanager, permet d'obtenir un maillage systématique de **placettes carrées** auxquelles on associe la valeur de la métrique et les coordonnées X et Y du centre de chaque placette (voir Figure n°9). La dimension de ces carrées est de 14.14x14.14 mètres (≈ 200 m²). Ces dimensions sont utilisées pour les pixels (ou cellules) des cartes une fois que le maillage est intégré au SIG.

La **méthode 2** sera celle utilisé à l'ONF. On utilise Lasmanager pour calculer les métriques sur des **placettes circulaires** disposées en quinconce (voir Figure n° 10). On obtient un fichier avec le centre de chaque placettes circulaire (coordonnée X et Y) et la valeur de la métrique associé. Dans notre cas, la séparation entre les placettes est de 8 mètres. On intègre ensuite le maillage au SIG où une interpolation IDW (interpolation de pondération par l'inverse de la distance)¹². La taille de la grille de sortie choisie est également de 8 mètres.

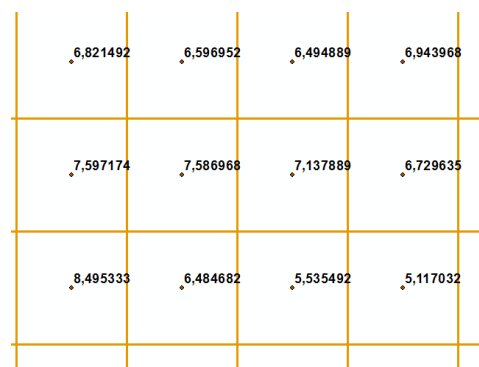


Figure n°9 : Maillage systématique des placettes carrées permettant de cartographier l'estimateur LiDAR de hauteur dominante ; le point central signale le centre de parcelle ; en orange, périmètre qui définit la surface (200 m²) des cellules à partir de laquelle la valeur de la métrique a été estimée.

¹² Détermine les valeurs de cellule via la combinaison pondérée de manière linéaire d'un ensemble de points d'échantillonnage. La pondération est une fonction inverse de la distance. La surface qui est interpolée doit être

Avant intégrer les maillages au GIS (pour chacune des méthodologies), le modèle choisi est appliqué à l'ensemble des placettes en utilisant le logiciel R (on dispose d'un fichier .txt avec les coordonnées de centre de chaque placettes et la métrique calculée pour chacun).

On obtient une carte de hauteur de l'ensemble de la zone survolée. Pour des raisons de disponibilité d'information sur les limites des parcelles de gestion, nous n'avons toutefois réalisé cette carte que sur les forêts publiques (voir carte annexe 4, 7).

On obtient une carte de hauteur de l'ensemble de la zone survolée réduite à la surface correspondant aux forêts publiques (voir carte annexe 4, 7).

Nous avons comparé les deux méthodologies de cartographie sur les placettes d'échantillonnage. En utilisant les coordonnées du centre de nos placettes de terrain, nous avons exporté la hauteur dominante prédite selon chaque méthodologie pour ensuite les comparer avec la hauteur dominante « réelle » des placettes d'échantillonnage. On a calculé les résidus (la différence entre la valeur exportée de la carte et la valeur que nous avons mesuré sur le terrain), sa moyenne et son écart-type. On considère que quand les résidus sont petits, l'ajustement de la méthode de la cartographie est meilleur.

Notre modèle a une limite d'application en fonction des groupes de hauteurs échantillonnés donc la prédiction de la hauteur en dehors de cette limite ne peut pas être assurée. Si on regarde la Figure n° 11, on pourrait garantir la prédiction de la hauteur pour la zone où les couleurs rouge et vert coïncident (c'est-à-dire, pour l'intervalle d'entre 3 et 16 mètres).

Il est nécessaire de définir les unités de gestion qui sont en régénération. Donc, nous avons calculé le pourcentage de la surface de chaque parcelle qui est dans l'intervalle. On sélectionne ensuite les parcelles pour lesquelles les peuplements de hauteur comprise entre 3 et 16 mètres représentent plus de 30 % de sa surface (voir carte annexe 5).

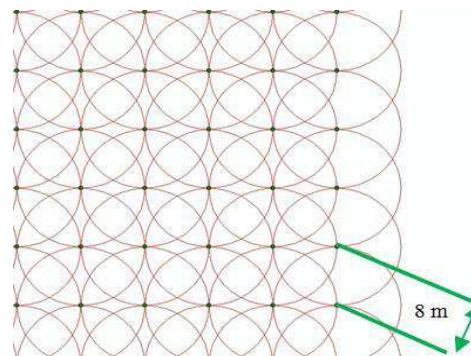


Figure n° 10 : Maillage systématique des placettes circulaires disposées en quinconce (méthode 2). En rouge, emprise circulaire (200 m²) des cellules à partir de laquelle la valeur de la métrique a été estimée.

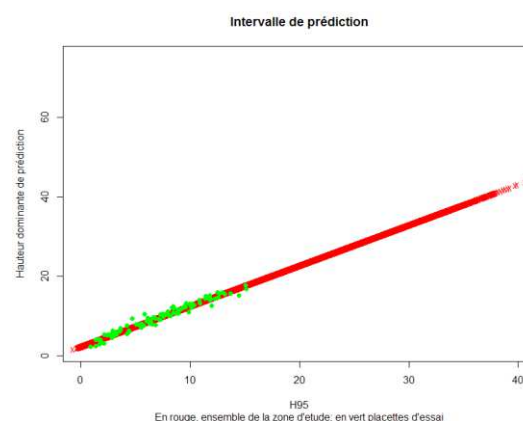


Figure n° 11 : relation entre la hauteur de prédiction pour l'ensemble de la zone survolée (rouge) et les placettes d'essai (vert). Exemple avec le percentile 95.

3. RESULTATS

Dans la suite de l'étude les résultats des différents types de modèles de prédiction de hauteur dominante des jeunes peuplements à partir de données LiDAR sont présentés. Il s'agit d'analyser la robustesse, les performances de prédiction de modèles linéaires simples ou multiples. L'analyse de sensibilité des modèles à l'incertitude de positionnement des placettes de calibration ou à la densité de points est également présentée. Le meilleur modèle sera alors retenu pour réaliser une cartographie sur l'ensemble de la zone d'étude selon les deux démarches testées.

celle d'une variable dépendante de l'emplacement. Dans ce calcul sur SIG, on a utilisé un seul point sur une distance variable.

3.1. Modélisation

3.1.1. Régression linéaire simple vs régression linéaire multiple

Les Tableau n° 9 et 8 résument les paramètres obtenus des modèles testés pour les placettes de 8 mètres de rayon et les coordonnées X et Y du centre de la placette.

Dans le Tableau n° 9, le RMSE de cross validation (RMSEcv) des métriques sélectionnées varie de 0.68 (pour la Hmv3¹³) à 0.97 mètres (pour le percentile 75), le R² varie entre 0.93 et 0.97 et le biais est pratiquement inexistant. On peut donc considérer ces résultats satisfaisants. Pour l'analyse multivariable, les trois meilleurs modèles sont montrés dans le

Modèle		RMSE	R ² ad	pval_int	RMSEcv
Y = β ₁ X ₁ + β ₂ X ₂ + β ₃ X ₃ + ... + β ₀ + ε					
m1	Hop = 1,03589*Hmv3 + 1,5592*N_m + + dhm_lst*(-5,51195) + 0,57527	0,6721	0,9696	2,20E-16	0,682
m2	Hop = 0,910037*H99_fs1st + 0,002438*Hcv_lst + + N_m*0,819173 + 0,884151	0,697	0,967	2,00E-16	0,714
m3	Hop = 0,913567*Hmv135 + 0,841704*Hcv_lst + + N_m*0,841704 + 0,039385	0,647	0,972	2,00E-16	0,663

Tableau 8. On a limité le nombre de variables explicatives à 3 (afin de ne pas compliquer le calcul de la hauteur et de ne pas augmenter le temps calcul de la cartographie). Dans ce cadre, la RMSE de cross validation a diminué très légèrement par rapport à la modélisation linéaire monovariable (de 0.687 à 0.663 mètres). Il s'agit quand même d'une amélioration négligeable (2.4 cm) qui ne compensera pas l'augmentation du temps de calcul due à l'intégration de deux variables supplémentaires dans les calculs.

Tableau n° 8 : Résumé de paramètres de la régression linéaire multiple et après cross –validation pour trois exemples de modèles évaluées avec le mieux RMSE . Hop =Hauteur dominante de prédiction ; RMSE : erreur quadratique moyenne ; R²ad : R carré ajusté ; pval_int : p-value ; RMSEcv : erreur quadratique moyenne après cross-validation.

En conclusion, les résultats entre la régression linéaire simple et multivariable ne sont pas trop différents (en termes du RMSEcv). La hauteur est une mesure directe provenant du laser donc on pense qu'un modèle avec une métrique explicative du type de hauteur pourrait bien expliquer son comportement. **En conséquence, en appliquant le critère de simplicité du modèle, l'étude va se concentrer sur l'analyse linéaire simple. On remarquera la faible variation des RMSEcv des 11 modèles sélectionnés (30 cm).**

Rappel : l'annexe 2 donne la liste et la définition des métriques utilisées dans la modélisation.

Métrique	Modèle	RMSE	R ² ad	pval_int	RMSEcv	Biais	Effcv
Hmv3	Hop = 0.9987*Hmv3+0.6857	0.6808	0.9689	1.23E-04	0.6871	-1.00E-04	0.9955
Hmv5	Hop = 1.0074*Hmv5+0.7978	0.6901	0.968	9.99E-06	0.6969	-3.00E-04	0.9953
Hmv135	Hop= 1.0056*Hmv135+0.7825	0.6921	0.9678	1.54E-05	0.6998	-3.00E-04	0.9953
Hmv6	Hop = 1.0105*Hmv6+0.8478	0.6964	0.9674	3.37E-06	0.7033	-4.00E-04	0.9953
H99	Hop = 1.0059*H99+1.4628	0.7116	0.966	3.98E-14	0.7209	-0.0011	0.995
Hmax	Hop = 0.9831*Hmax+0.4944	0.7272	0.9645	9.07E-03	0.7352	-3.00E-04	0.9948
H95	Hop = 1.0206*H95+2.2490	0.7831	0.9588	7.39E-24	0.7922	-0.0013	0.994
H90	Hop = 1.0381*H90+2.6653	0.8249	0.9543	1.32E-28	0.8349	-0.0017	0.9933
H80	Hop = 1.072*H80+3.235	0.9088	0.9445	1.07E-33	0.918	-0.0019	0.9919
dhm	Hop = 5.244*dhm+2.742	0.9472	0.9397	2.09E-25	0.9555	-5.00E-04	0.9913
H75	Hop = 1.087*H75+3.509	0.963	0.9377	1.70E-35	0.9716	-0.0018	0.991

Tableau n° 9 : Résumée de paramètres de la régression linéaire simple et après cross –validation pour chaque variable explicative évalué (métrique) ; Hop =Hauteur dominante prédite; RMSE : erreur quadratique moyenne ; R²ad : R carré ajusté ; pval_int : p-value ; RMSEcv : erreur quadratique moyenne après cross-validation ; Effcv : efficence.

¹³ Hmv3_Bd: moyenne des 3 plus haut de la placette avec application d'un rayon d'exclusion (semblable au Hmv5_Bd, dont calcule est expliqué dans la section 2.3.2 Données LiDAR) .

3.2. *Etude de la robustesse*

3.2.1. Effet de l'incertitude dans la localisation des placettes

Il n'y a pas d'évolution majeure de la RMSEcv lorsque l'on réduit la taille des placettes (voir Figure n°13). En revanche, la RMSEcv accroît lorsqu'on augmente la taille des placettes jusqu'à 12 mètres de rayon ($\approx 450 \text{ m}^2$). Cette surface est plus du double de la taille des placettes originales du terrain (200 m^2). On constate donc que l'homogénéité de notre placette de mesure de 8 m permet la stabilité du RMSE mais que rapidement (12m de rayon) le RMSE peut varier fortement. Deux variables semblent particulièrement sensibles : dhm et Hmax.

Lorsque l'on applique une erreur aléatoire de 2 m autour du centre des placettes, il y a des variables qui apparaissent insensibles comme H75, H80, H90 et H95 (voir Figure n° 14). En revanche, il existe des variables plus sensibles aux changements aléatoire de position du centre des placettes comme : hmax, dhm, H99 ou Hmv3, Hmv135. Les autres variables présentent des comportements intermédiaires.

3.2.2. Effet de la densité de points LiDAR

Généralement, l'erreur de prédiction (RMSEcv) a tendance à augmenter lorsque la densité de points est réduite artificiellement en enlevant un point sur X (voir Figure n° 15). Cependant, il faut réduire fortement la densité de point (jusqu'à moins de 2 pts/m²) pour constater une évolution forte de la RMSE.

Comme il s'agit d'une réduction artificielle de la densité, nous avons choisi de tester une réduction réelle de la densité de points en ne sélectionnant qu'un vol, 2 vol ou tous les vols croisés sur la zone d'étude. La Figure n° 12 montre que la réduction d'un tiers la densité d'émission initiales n'a pas d'influence sur précision des variables prédites. En revanche, pour certains variables comme Hmax, hdm, etc ... il y a une augmentation de la RMSEcv. Pour les autres variables, on pourrait se contenter d'un seul vol sur la zone à 4.8pts/m² en moyenne.

Les métriques correspondant aux moyennes de maxima locaux ont les variations les plus fortes quand il y a une diminution de la densité de points. Mais par contre, les percentiles ont une RMSE invariable donc, pour travailler avec des densités des points faibles, ils doivent être pris en compte.

Si on examine la Figure n° 12, il y a une observation intéressante à signaler. En fait, il y a une erreur plus petite avec deux bandes de vol (9.5 points/m²) qu'avec toutes les bandes de vols combinées (14.65 points/m²). Il est certain que la différence est faible. Une explication possible serait l'imprécision introduite en post-processus par le prestataire, pour l'ajustement spatial tous les bandes dans les zones de superposition.

3.3. *Sélection de la variable prédictive*

Le Tableau 10 présente une synthèse sur le choix des variables. Selon les critères de sélection décrits dans la section 2.6 et l'annexe 4, on a obtenu une table classifiée par couleur et avec une valeur associée pour chaque métrique en fonction du RMSEcv et le critère de stabilité évalué.

On remarque qu'il y a trois métriques (Hmv6, Hmv5 et H95) qui se détachent pour être choisies comme variable prédictive de la hauteur dominante des jeunes peuplements. Les métriques de moyenne de maxima locaux (Hmv6, Hmv5) ont une RMSE plus petite que le percentile 95.

En revanche, le percentile 95 semble plus stable vis-à-vis des erreurs de positionnement de la placette ou de la densité de points émise. Au final, le choix de la variable explicative dépend du critère qu'on veut prioriser.

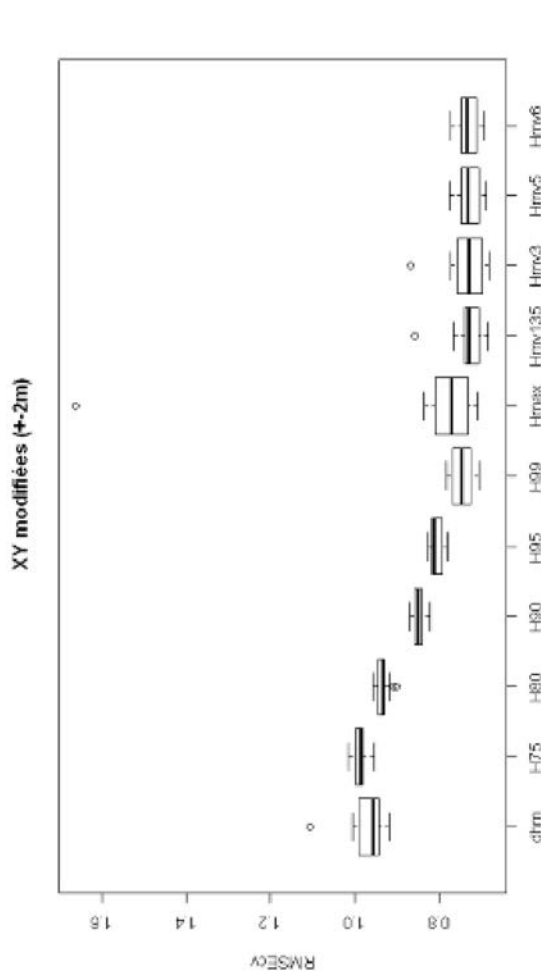


Figure n° 14 : variation du RMSE par métrique quand le centre de la placette est changé d'une façon aléatoire autour du centre original (± 2 mètres) obtenu du GPS différentiel.

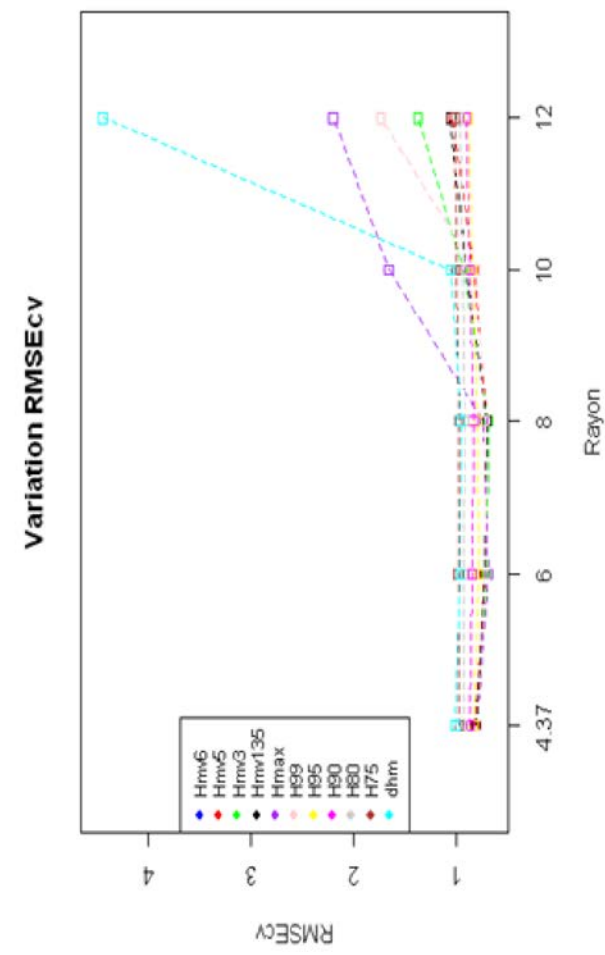


Figure n°13 : Variation de RMSE selon le rayon d'extraction utilisé pour le calcul des métriques

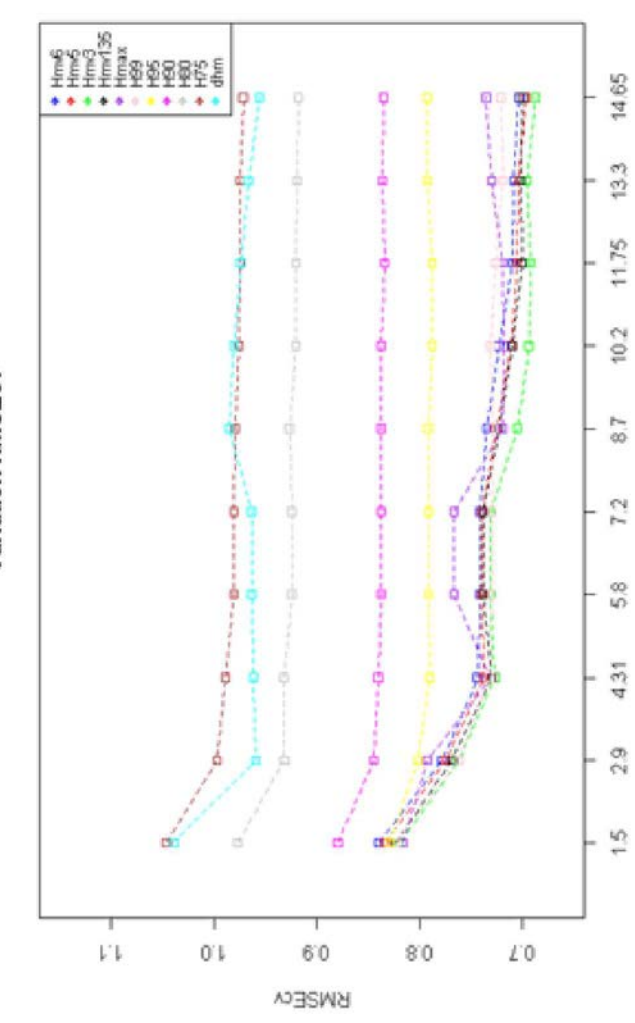


Figure n° 15 : Effet de la densité de points (par m²) sur le RMSE (en mètres) des modèles conformément à une réduction progressive du pourcentage des points des fichiers .las complets.

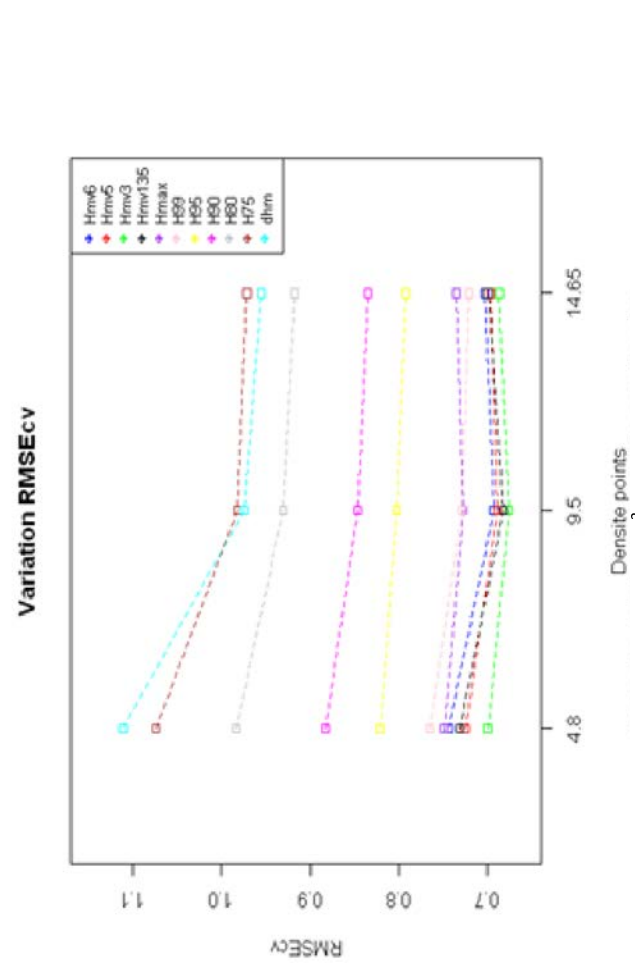


Figure n° 12 : Effet de la densité de points (par m²) sur le RMSE (en mètres) selon les bandes du vol utilisé

On a vu que le RMSEcv est assez bon pour les onze métriques testées (et inférieur à l'erreur de mesure de terrain) et, par contre, la stabilité et la robustesse du modèle sont des aspects à prendre en compte (par exemple, à cause de l'incertitude dans la localisation des placettes pour l'erreur du GPS). En plus, si on doit choisir entre la Hmv5 et le percentile 95, on voit que dans la littérature (annexe 6), les percentiles de hauteur sont largement utilisés pour prédire la hauteur dominante pendant que la définition de Hmv5 est moins intuitive est difficile à calculer.

En conséquence, on va choisir le percentile 95 comme la variable explicative de notre modèle linéaire simple à appliquer sur l'ensemble de la zone d'étude.

Avant d'appliquer le modèle pour cartographier la zone d'étude, on a fait le diagnostic pour vérifier les hypothèses de la régression linéaire (Voir annexe 5). Le diagnostic démontre que la variance est égale dans tout le graphique (hypothèse d'hétéroscédasticité), que les résidus ont un comportement normal, et que le modèle n'a pas de valeurs extrêmes. L'hypothèse de la linéarité n'ajuste pas parfaitement (mais si on observe le graphique Ho vs Hopred de l'annexe pour le H95, on peut assurer que les données suivent un comportement linéaire). Globalement on peut considérer que le diagnostic vérifie les hypothèses de la régression linéaire donc on peut appliquer le modèle.

3.4. Cartographie

On a appliqué le modèle avec le percentile 95 comme variable explicative à l'ensemble de la forêt publique couverte par le vol LiDAR en appliquant les deux méthodologies définies dans la section 2.7. On obtient 2 cartes différentes (carte annexe 4 et 7). La hauteur dominante mesurée sur le terrain (placettes d'échantillonnage) est alors comparée avec la hauteur dominante prédite une fois que les modèles ont été appliqués à l'ensemble de la zone d'étude pour les deux cartes. Les résidus, leur moyenne et leur écart-type sont calculés (Tableau n° 11). La différence entre les deux n'est pas trop significative. Il est certain que les résidus par la deuxième méthodologie sont légèrement plus faibles. Cependant, en observant la Figure n° 16, on peut dire que, même s'il y a quelques variations compréhensibles dues à la différence de méthode et aussi à la taille pixels, globalement les deux cartes montrent une hauteur de prédiction semblable si on classe par intervalles de hauteur.

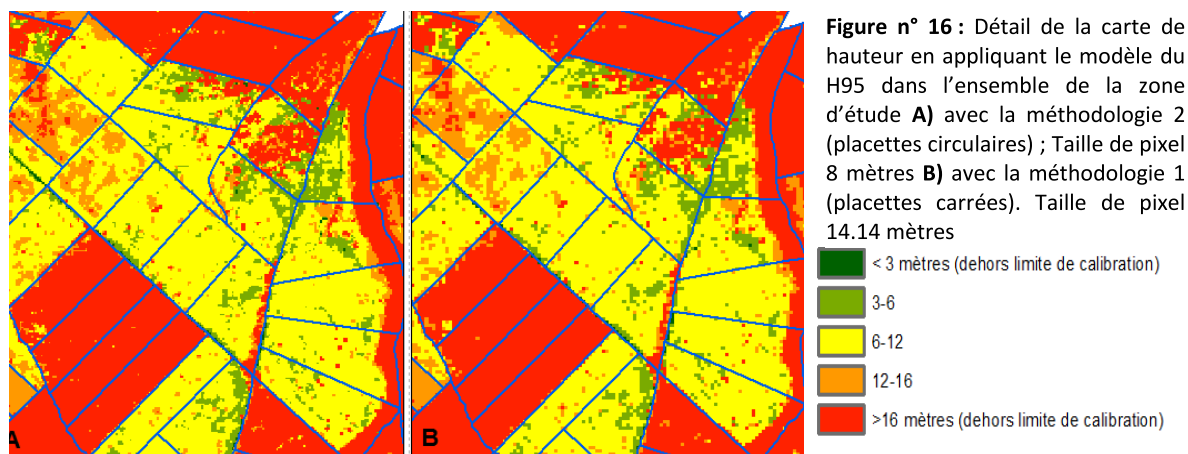
Il existe des pixels classés dans l'intervalle de hauteur estimé (3-16 mètres) dans plusieurs parcelles. Ils ne sont pas assez nombreux et peuvent représenter une surface de la placette assez faible (voir Figure n° 16). Ces endroits peuvent être par exemple des zones lisières ou des cloisonnements. Afin de délimiter les placettes en régénération (ou des unités élémentaires de suivi en régénération) on a seulement sélectionné les parcelles avec au moins 30 % de la surface avec une hauteur de prédiction entre 3 et 16 mètres. Comme exemple, cette démarche est présentée dans les cartes en annexes 6 et 7 pour la méthodologie 2. La classification est faite en fonction des classes BDR en fonction des critères de l'Agence Bourgogne-Est (voir Tableau n° 1)

Metrique	STABILITE	RMSEcv
Hmv6_Bd	Bonne	Très bon
Hmv5_Bd	Bonne	Très bon
Hmv3_Bd	Moyenne	Très bon
Hmv135	Bonne	Très bon
Hmax	Pas de stabilité	Bon
H99_fst.nosoil	Bonne	Bon
H95_fst.nosoil	Très Bonne	Bon
H90_fst.nosoil	Très Bonne	Moyenne
H80_fst.nosoil	Très Bonne	Moyenne
H75_fst.nosoil	Bonne	Null
dHmean_sup2m	Pas de stabilité	Null

Tableau n° 10 : résultat de la classification des métriques selon la stabilité et la valeur du RMSEcv; en rose, les métriques sans stabilité et/ou avec une RMSEcv mauvaise; en vert les métriques avec une bonne ou très bonne stabilité et avec un bon ou très bon RMSEcv.

Hoterrain-Hopred	M1_H95	M2_H95
Moyenne	0,746	0,666
Ecartype	0,541	0,506

Tableau n° 11: moyenne et écart-type des résidus (différence entre la hauteur dominante mesurée sur le terrain et la hauteur dominante prédite selon la carte correspondante) ; M1 et M2 : méthode de cartographie 1 et 2 respectivement.



4. DISCUSSION

4.1. Modélisation

Parmi près de 500 modèles monovariante testés, les 11 meilleurs ont une erreur de prédiction (RMSEcv) toujours inférieur à 1 m, avec des r^2 de l'ordre de 0.94 à 0.97 (voir Tableau n° 9).

Parmi ces modèles parcimonieux (une seule variable), certains sont de plus indépendant de la densité de points lidar émise ou de l'erreur de positionnement des placettes. On peut retenir en particulier, deux bons modèles de prédiction de la hauteur dominante des peuplements :

- Celui qui utilise la hauteur du percentile 95 des points first de la végétation (H95_fst_nosoil) avec une RMSEcv de 0.792 et un r^2 de 0.958
- Celui qui utilise la moyenne de 5 maxima locaux (Hmv5_Bd) avec une RMSEcv de 0.6969 et un r^2 de 0.968

Cette précision et la robustesse des modèles permettent de conclure qu'on peut donc bien prédire la hauteur de jeunes peuplements à partir d'un nuage de point LiDAR.

Si on compare aux données de la bibliographie (voir annexe 5), on constate que les performances de nos modèles sont très satisfaisant, puisque notre RMSE est équivalente (0.69 et 0.79) et que nos modèles sont en général plus parcimonieux (une seule variable explicative).

Un des premiers auteurs qui a commencé à modéliser les variables dendrométriques à partir de données LiDAR aériennes a été Erik Naesset. Donc, Næsset (2004), calcule la hauteur dominant par un modèle logarithmique. Après une de-transformation des paramètres du modèle au forme linéaire, on a une RMSE de 1.07^{14} mètres et un R^2 de 0.934.

Holmgren and Jonsson, (2004), dans leur étude sur l'estimation de plusieurs variables dendrométriques sur la forêt en Suède, obtient une RMSE de 0.8 m mais il s'agit de la hauteur moyenne pas très bien définissable. Heurich et Thoma (2008) obtient une RMSE de 0.647 et un R^2 de 0.99 mais le modèle est construit avec sept variables explicatives, ce qui démontre la difficulté de travailler avec beaucoup de variables. Zaldo et al., (2010), en utilisant le percentile 80, obtiennent une RMSE de 1.710 mètres pour la estimation de la hauteur maximale d'un peuplement de hêtre.

González-Ferreiro et al. (2012), estiment la hauteur dominante des peuplements de *Pinus radiata* avec le percentile 95 comme variable explicative. La RMSE obtenue est beaucoup plus élevée (1.988 mètres) et le R^2 plus faible (0.846) que nos modèles mais la densité de points est plus faible aussi (0.5 point/m²).

¹⁴ Valeur du RMSE obtenu après transformation inverse pour revenir en mètres parce que le modèle original est sous la forme logarithmique. Il existe un biais associé à cette de-transformation (factor correction = $e^{RMSE^2/2}$).

A l'ONF, la hauteur dominante des peuplements à partir des données LiDAR aériennes a commencé à être estimée en 2008. Selon Martins, (2009), en utilisant la Hmv5 et d9¹⁵ dans un modèle linéaire multi variable, la RMSE obtenue est 0.761 mètres et une R^2 de 0.99. Ce modèle apparaissait comme indépendant de la densité de point, de la position de la placette, de la façon de modéliser le sol pour obtenir la hauteur des points LiDAR. De plus, l'effet type de peuplement (régulier ou irrégulier) n'était pas significatif grâce à l'introduction dans le modèle de la variable d9. Cependant la gamme de validité de ce modèle se situe entre 8 et 40 m de haut.

Une étude précédente (Diop, 2011), avait conclu qu'il était difficile d'utiliser tels quels les modèle de Naesset ou Heurich sans recalibration sur un jeu de placettes. Diop a montré que l'indicateur Hmv5 présentait l'avantage d'avoir des performances équivalente ou meilleures comparés aux modèles de Naesset et Heurich, avec l'avantage de ne pas nécessiter de recalibration. Une des perspectives intéressantes de ce travail consisterait à évaluer la généralité de nos modèles en les testant sur d'autres jeux de données équivalents obtenus dans sur d'autres sites couvert au LiDAR et d'autres contextes forestiers (peuplement résineux ou mixtes).

Tous ces résultats différents mettent en évidence l'importance des caractéristiques locales existantes dans chaque zone d'étude, la topographie, le type de peuplement (âge et essence), le type de modèle utilisé (linéaire, potentiel, exponentiel), le protocole développé, les variables explicatives choisies et, bien sûr, les paramètres du vol LiDAR. En plus, il y a des différences entre les définitions de la hauteur dominante qui rendent difficile à avoir un même critère pour comparer les métriques Lidar d'une études à l'autres (Lovell et al., 2005).

4.2. Etude de la robustesse

4.2.1. Effet de l'incertitude dans la localisation des placettes

La co-registation géographique exacte des données ALS et des données du terrain est essentielle pour avoir une prédiction précise des variables des peuplements forestiers (Gobakken and Næsset, 2009). Mais, la relation spatiale des placettes du terrain ne peut pas coller exactement avec les placettes LiDAR en raison des erreurs de positionnement propres des navigateurs GPS sous la couvert forestier (Mauro et al., 2011) Dans notre cas, le GPS utilisé a une erreur de positionnement estimée entre 1 et 4 mètres en forêt (DTCB- R&D Pôle de Chambéry, 2012).

Si on lance un regard à la bibliographique, on confirme que la sensibilité attendue pour les erreurs des positionnements doit décroître quand les rayons des placettes augmentent (Gobakken and Næsset, 2009). Cependant, la taille de nos placettes (2 ares, donc assez petite) est justifié par le type de peuplement à étudier (jeunes et en régénération). En conséquence, la mesure sur le terrain des placettes plus larges peut devenir dans une opération assez lourde. En fait, White et al., (2013) recommandent une taille des placettes comprise entre 8 et 14 mètres pour l'approche surfacique (area-based). On se situe donc dans cette limite.

D'autre part, l'utilisation du GPS est justifiée pour des raisons économiques et techniques parce que des autres types de systèmes plus précis (comme la cinématique temps réel ou RTK¹⁶ par exemple) peuvent être très coûteux et la réception des GPS RTK est tout aussi aléatoire sous couvert forestier. Il faudrait coupler la position de points de références hors forêt au GPS RTK, puis « lancer » des placettes de calibration en forêt à l'aide d'un

¹⁵ d9 densité relative entre dans la 10° tranche (10 tranche de 2 m à Hq95)(tous les points)

¹⁶ Technique de positionnement par satellite basée sur l'utilisation de mesures de phase des ondes porteuses des signaux émis par le système GPS ou GLONASS. Une station de référence fournit des corrections en temps réel permettant d'atteindre une précision de l'ordre du centimètre.

théodolite. Cette opération est réalisable mais avec un coût très élevé et limite également les zones échantillonnables.

Pour essayer de résoudre cette incertitude de localisation et éviter les problèmes du décalage des coordonnées du terrain et LiDAR (voir Figure n° 17), il est possible parfois de faire une recalibration des placettes moyennant la prise sur le terrain des coordonnées des arbres remarquables ou de repères microtopographique comme des structures archéologiques, des « galettes » de chablis, des limites de parcelle, des cibles installées spécifiquement ou des autres éléments qui peuvent être reconnus facilement sur les données LiDAR. Cependant, dans le cadre des jeunes peuplements, avec des masses si homogènes et avec une densité de tiges/ha abondante cette opération n'offre pas toujours des bons résultats (erreurs résiduelles de l'ordre de 1 à 2 m).

Il est néanmoins rassurant de constater que nos modèles sont relativement indépendants de la position des placettes (dans une limite de +/- 2m).

D'autre part, les statistiques de hauteur LiDAR sont assez insensibles à la taille des placettes d'échantillonnage et à l'erreur du GPS (Gobakken and Næsset, 2009 ; Frazer et al., 2011), ce qui est en accord avec nos résultats de stabilité des métriques quand on fait varier la taille des placettes pour l'extraction des métriques. On a constaté d'un autre côté, qu'à chaque fois que la taille de la placette est agrandie, la probabilité d'introduire des arbres dominants qui ne sont pas ceux mesurés sur le terrain augmente donc il faut faire attention lorsqu'on étudie la robustesse avec des tailles de placettes très différentes par rapport à celles mesurées sur le terrain. En plus, des métriques comme la Hmax ou le percentile 99 sont plus exposées à la variabilité du peuplement en hauteur et aussi, à l'existence des obstacles ou des animaux, par exemple, qui sont captés pour le laser (surtout la Hmax).

4.2.2. Effets de la densité de points

Il existe quelques études qui montrent que l'estimation de la hauteur dominante n'est pas affectée par la réduction de la densité de points. Par exemple, Treitz et al., (2012), étudient l'influence que une diminution de la densité de points a sur différents métriques. Ils arrivent à la conclusion qu'une densité de 0.5 pt/m² (laquelle peut être considéré faible) est suffisante pour l'estimation de variables dendrométriques au niveau des peuplements. Cette conclusion est réaffirmée par González-Ferreiro et al. (2012) qui signalent que pour l'estimation de la hauteur moyenne et dominante, la surface terrière, le volume ou la biomasse, la densité de points peut être assez faible (≈ 0.5 points/m²) sans avoir une augmentation de l'erreur significative. Ils ont développé des modèles mathématiques, avec une approche surfacique et

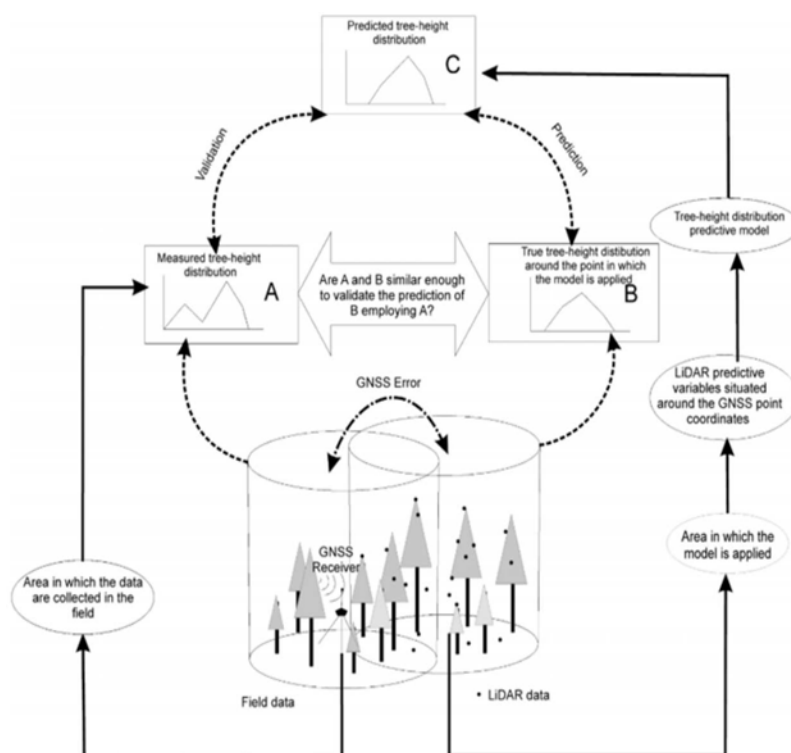


Figure n° 17 Diagramme illustratif du manque de co-régistration parmi les données de terrain et les données du LiDAR. S'il y a un décalage des coordonnées parmi celles du terrain et du Lidar on peut avoir des prédictions erronées (Mauro et al., 2010)

pour deux densités de points différents (8 et 0.5 point/m²). Pour réduire la densité de points ils ont fait une sélection aléatoire de retours LiDAR dedans les cellules d'une grille de 1 m².

En appliquant notre méthode de réduction artificielle de la densité de points, nous aboutissons aux mêmes résultats. La réduction de la densité de points a tendance à faire augmenter l'erreur d'estimation de la hauteur dominante. Mais l'augmentation de l'erreur atteint au maxima (20%) lorsqu'on réduit de 90 % la densité de point initiale et la RMSE de prédiction reste le plus souvent inférieure à 1 m quel que soit l'estimateur LiDAR utilisé. Il est certain qu'il y a des métriques plus instables que d'autres comme la Hmax ou le H99 (voir Figure n° 15). Cependant, cette observation peut être mise en relation aussi avec les résultats de Coops et al. (2007), qui souligne une forte corrélation entre les hauteurs dominantes LiDAR et mesurées, malgré une densité de points inférieure à 1 points/m².

L'intérêt de notre étude réside dans un vol LiDAR à très haute densité de points qui a été obtenue par 3 vols croisés successifs. Il a donc été possible de tester réellement l'effet de la densité de points émis en sélectionnant les points issus du 1^{er} survol, de 2 ou les 3 vols successifs. Il est donc possible de comparer en condition réelle, et non plus en simulant artificiellement un nuage de points moins dense, l'effet d'une réduction de 1/3 ou 2/3 de la densité de points. Les tests que nous avons réalisés ne montrent aucun effet significatif de la densité de points sur les performances de nos modèles.

La connaissance de cet aspect est intéressant parce que les densités des points LiDAR faibles se traduisent dans une réduction des coûts d'acquisition des données (Treitz et al., 2012; Bock et al., 2011).

4.3. Cartographie

On a mis en œuvre deux méthodologies différentes pour cartographier l'ensemble de la zone d'étude. Pour vérifier la qualité de l'estimation sur l'ensemble de la zone, il serait recommandable d'échantillonner de nouvelles zones pour avoir des placettes d'essai indépendantes. Cependant, pour des raisons de délai de l'étude, nous avons utilisé nos placettes d'échantillonnage initiales.

Nos placettes d'essai sont circulaires alors que la cartographie de la première méthode a été faite à partir d'un calcul des métriques pour un maillage des placettes carrées. Donc, il existe la possibilité d'introduire une erreur d'estimation par rapport à la forme des placettes. Avec la méthode 2, qui calcule les métriques pour des placettes circulaires de la méthode 2, cette problématique peut être minimisée. Mais il est certain que quand on exporte les résultats au SIG, on va « rasteriser » le résultat en pixels carrés. On est toujours obligés de passer d'une unité d'évaluation circulaire à un autre carré. Cependant, et par rapport aux résultats obtenus de moyenne et écart type de résidus (voir [Tableau n° 11](#)), on peut dire que la méthode 2 est légèrement plus précise si on veut connaître la hauteur dominante concrète. Cependant, la représentation de hauteur prédite pour les deux méthodologies ([Figure n° 16](#)) montre des différences peu évidentes à apprécier, surtout quand on fait des classifications de hauteur par intervalles.

Un aspect très important dans la méthode 1 est la dimension des pixels (ou cellules) de la cartographie. Il n'est pas recommandable de choisir une taille de pixels petite. Les maillages des cellules larges peuvent contenir un nombre plus important d'impulsions laser et avoir aussi une distribution plus uniforme (Næsset, 2002). En Finlande, des pixels de 16 mètres sont utilisés pour des placettes du terrain avec un rayon de 9 mètres (White et al., 2013), ce qui n'est pas loin de notre choix (placettes de terrain de 8 mètres et taille des pixels de 14.14 m). Il est recommandé d'utiliser au moins une taille équivalente de pixels à celle des placettes d'échantillonnage.

Dans la méthode 2, la taille des pixels choisie (8 mètres) est justifiée par rapport à la séparation des placettes du maillage qui a été choisie (8 mètres aussi). Cependant, après avoir

fait la cartographie, on croit important de laisser une séparation plus grande entre les placettes pour assurer le minimum de superposition (mais en assurant toujours un recouvrement complet de la surface).

D'un autre côté, en faisant une carte de classes BDR (carte annexe 6) à partir des hauteurs dominantes prédites, on est limité par des intervalles de prédictions mais aussi par des critères différents à la hauteur, lesquels sont utilisés pour la classification BDR (comme les interventions sylvicoles faites). Par exemple, l'appartenance aux classes 6 et 7 est soumise à la réalisation de la première éclaircie donc il est nécessaire d'avoir ce type d'information des gestionnaires pour pouvoir la combiner avec la couche de hauteur. Une autre possibilité serait d'essayer de combiner l'information issue du LiDAR avec des autres systèmes de télédétection optiques à très fine résolution spatiale (décimétrique) afin de pouvoir reconnaître automatiquement les trouées dans le couvert révélant ce type d'interventions. Mais le problème de l'identification des classes 6 et 7 ne doit pas faire oublier quand même que pour les autres classes, on peut faire un seuillage simple sur la hauteur.

Jusqu'à la classe 3 de la BDR (voir Tableau n° 1), les peuplements à classer se trouvent en dessous de 3 mètres de hauteur. Notre limite d'échantillonnage a écarté une hauteur inférieure à 3 mètres pour des raisons sylvicoles locales par rapport aux coupes de régénération sur le hêtre (il y a très peu des peuplements complètement ouverts à cette hauteur) (voir section 2.3.1. Données de terrain). Donc si on veut estimer la hauteur de ces très jeunes peuplements une démarche plus précise devrait être mise en place surtout par rapport au protocole et plan d'échantillonnage. Par exemple, Estornell Cremades, (2011) estime la hauteur de la couche arbustive d'une forêt méditerranéenne à partir de données LiDAR aériennes dont la hauteur est centrée juste entre 0.8 et 2.5 mètres.

Cependant, les classes BDR 4, 5 et 6 (les classes que l'on peut cartographier avec la démarche suivie) correspondent aux peuplements dont la hauteur détermine en grande partie les interventions sylvicoles à faire.

5. CONCLUSION

Le but de cette étude était de valider une méthodologie pour pouvoir estimer la hauteur de jeunes peuplements forestiers à partir des données LiDAR aériennes, en tenant en compte de toutes les différences de comportements (en croissance, en gestion sylvicole) qu'il existe par rapport aux peuplements matures. Nos résultats nous indiquent que notre méthode est très efficace dans l'estimation de la hauteur dominante des jeunes peuplements pour des hauteurs comprises entre 3 et 16 mètres, avec une RMSE de précision inframétrique.

Parmi les différentes variables explicatives (métriques) étudiées, les percentiles de hauteurs sont très stables (par rapport à une variation de densité de points et une variation de la taille et localisation des placettes) et les moyennes de maxima locaux offrent une RMSE plus faible.

Une diminution de la densité de points n'introduit pas une erreur significative dans l'estimation de la hauteur dominante. Ainsi, une bonne estimation de la hauteur dominante de ces jeunes peuplements peut être obtenue avec une densité de points faible. D'autre part, les deux méthodologies de cartographie ne donnent pas de résultat différent. Pour une application opérationnelle, il faudrait pouvoir pré-cartographier les jeunes peuplements correspondant à toutes les classes BDR. Cependant, certaines classes sont difficiles à cartographier avec un critère d'échantillonnage uniquement basé sur la hauteur comme variable de stratification. Pour pouvoir mettre en place une application plus opérationnelle, il serait utile d'avoir d'autres données. Il serait donc intéressant de tester de nouveaux outils (combinaison avec des orthophotos, données de terrain, développement de méthodes de traitement des données Lidar complémentaires) pour pouvoir parvenir à cartographier l'ensemble de la surface BDR.

Malgré ces limites, l'information que l'on peut fournir dans la gamme de calibration de 3 à 16 mètres est déjà très intéressante. En effet, des applications concrètes peuvent alors se

développer, notamment pour les gestionnaires forestiers pour pouvoir prévoir les dates des travaux sylvicoles de façon précise. De plus, en croisant ces informations avec les résultats déjà obtenus à l'ONF dans de précédentes études (Martins, 2009 ; Diop, 2011 ; Bock et al., 2011) cette gamme de calibration augmente jusqu'à 40 mètres.

6. BIBLIOGRAPHIE

- Andersen, H.-E., McGaughey, R.J., Reutebuch, S.E., 2005. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sens. Environ.* 94, 441–449.
- Bock, J., Dez, G., Dupouey, J.-L., Georges-Leroy, M., Jolly, A., Martins, F.S.R.V., Renaud, J.-P., 2009. Towards site index mapping in deciduous stands using multi-echo LIDAR data 5.
- Bock, J., Renaud, J.-P., Dez, G., Martins, F., Lenoy, M., Dambrine, E., Dupouey, J.-L., 2011. Le Lidar, outil performant pour cartographier la hauteur des peuplements forestiers. *RDV Tech.* 39–40.
- Chen, Q., Gong, P., Baldocchi, D., Tian, Y.Q., 2007. Estimating basal area and stem volume for individual trees from lidar data. *Photogramm. Eng. Remote Sens.* 73, 1355.
- Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J.T., 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. *Trees* 21, 295–310.
- Dez, G., 2008. Utilisation du Lidar aéroporté pour l'estimation de la hauteur dominante des peuplements forestiers - Application en forêt de Haye (Rapport fin d'études). ONF- Pôle Recherche & Développement Velaine en Haye; AgroParisTech - ENGREF; INRA.
- Diop, C.A.B., 2011. Modèle de prédiction de biomasse à partir de données de scan laser aéroporté (Lidar) (Rapport fin d'études). ONF- Pôle Recherche & Développement Velaine en Haye.
- DTCB- R&D Pôle de Chambéry, 2012. Analyse et optimisation du positionnement GNSS en forêt. Récepteurs Trimble gamme SIG/Cartographie. Office National des forêts.
- Duez, F., 2011. Adaptation de la Base de Données (BDR) sur l'île de la Réunion et proposition pour la gestion des Tamarinaies de production. (Rapport fin d'études). ONF - Pôle aménagement à la Réunion, La Réunion.
- Estornell Cremades, J., 2011. Análisis de los factores que influyen en la precisión de un MDE y estimación de parámetros forestales en zonas arbustivas de montaña mediante datos LiDAR. *Riunet*.
- Frazer, G.W., Magnussen, S., Wulder, M.A., Niemann, K.O., 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sens. Environ.* 115, 636–649.
- Gobakken, T., Næsset, E., 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* 39, 1036–1052.
- González-Ferreiro, E., Diéguez-Aranda, U., Miranda, D., 2012. Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry* 85, 281–292.
- Hall, S.A., Burke, I.C., Box, D.O., Kaufmann, M.R., Stoker, J.M., 2005. Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. *For. Ecol. Manag.* 208, 189–209.
- Heurich, M., Thoma, F., 2008. Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*) forests. *Forestry* 81, 645–661.
- Holmgren, J., Jonsson, T., 2004. Large scale airborne laser scanning of forest resources in Sweden. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 36, 157–160.
- Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M., Gessler, P.E., Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Can. J. Remote Sens.* 32, 126–138.
- Hudak, A.T., Evans, J.S., Stuart Smith, A.M., 2009. LiDAR Utility for Natural Resource Managers. *Remote Sens.* 1, 934–951.
- IFN Suisse, 2009. Inventaire Forestier national suisse. URL http://www.waldwissen.net/technik/inventur/wsl_landesforstinventar/index_FR

- Lim, K., Treitz, P., Wulder, M., St-Onge, B., Flood, M., 2003. LiDAR remote sensing of forest structure. *Prog. Phys. Geogr.* 27, 88–106.
- Lovell, J.L., Jupp, D.L.B., Newnham, G.J., Coops, N.C., Culvenor, D.S., 2005. Simulation study for finding optimal lidar acquisition parameters for forest height retrieval. *For. Ecol. Manag.* 214, 398–412.
- Mangeonjean, L., 2011. Etude de l'accroissement de la hauteur dominante en peuplements feuillus selon 2 méthodes (Forêt de Haye - Meurthe et Moselle) (Rapport fin d'études). ONF- Pôle Recherche & Développement Velaine en Haye.
- Martins, F., 2009. Estimativa da altura dominante a partir do sistema de varredura a laser aerotransportado (LIDAR) (Rapport fin d'études). ONF- Pôle Recherche & Développement Velaine en Haye, Nancy.
- Mauro, F., Valbuena, R., Manzanera, J.A., García-Abril, A., 2010. Influence of Global Navigation Satellite System errors in positioning inventory plots for tree-height distribution studies. This article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. *Can. J. For. Res.* 41, 11–23.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *PE RS- Photogramm. Eng. Remote Sens.* 66, 1367–1371.
- Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* 80, 88–99.
- Næsset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand. J. For. Res.* 19, 164–179.
- Næsset, E., 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sens. Environ.* 113, 148–159.
- Næsset, E., 2011. Estimating above-ground biomass in young forests with airborne laser scanning. *Int. J. Remote Sens.* 32, 473–501.
- Næsset, E., Bjerknes, K.-O., 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sens. Environ.* 78, 328–340.
- Nelson, R., Krabill, W., Tonelli, J., 1998. Estimating forest biomass and volume using airborne laser data. *Remote Sens. Environ.* 24, 247–267.
- Office National des Forêts, 2005. Le hêtre en Lorraine, Guide des sylvicultures.
- ONF, 2013. Office National des Forêts. URL <http://www.onf.fr/>
- Pardé, J., Bouchon, J., 1988. Dendrométrie. École Nationale du Génie Rural, des Eaux et des Forêts, École nationale du génie rural; des eaux et des forêts (France).
- Pierre-Yves, R., 2012. Adaptation d'une méthode de modélisation de la hauteur dominante de peuplements forestiers à partir des données LiDAR de type raster. (Rapport fin d'études). ONF- Pôle Recherche & Développement Velaine en Haye, Nancy (France).
- Tran, H., Vallance, M., 1997. Sylvie: Diagnostic sylvicole des peuplements réguliers. Protocole de terrain et logiciel. Office National des Forêts, Département des Recherches Techniques.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., Etheridge, D., 2012. LIDAR sampling Density for Forest Resource Inventories in Ontario, Canada - Google Académico. *Remote Sens.* 830–848.
- White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D., Woods, M., 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach.
- Young, J., 2011. Lidar for dummies. Autodesk and DTL solutions special editions.
- Zaldo, V., Moré, G., Pons, X., 2010. Estimación y cartografía de parámetros ecológicos en tres especies forestales (*Quercus ilex* L. subsp *ilex*, *Fagus sylvatica* L. y *Pinus halepensis* L.) con datos LiDAR. *Rev. Teledetección* 34 55–68.

ANNEXES

ANNEXE 1 – Protocole et plan d'échantillonnage

La localisation des placettes d'essai a été choisie en regardant les orthophotos du terrain, une pre-carte de hauteur qui a été fait après traitement de données LiDAR et la vérification visuelle sur le terrain pour pouvoir reconnaître les parcelles en régénération.

La hauteur dominante, entendue comme la moyenne des arbres de la strate dominante, s'estime à différentes façons selon la définition et la méthode qui est utilisée. Etant donné qu'une définition assez acceptée en France est la moyenne de 100 arbres les plus gros par hectare, l'ONF utilise l'estimation traditionnelle non biaisée de la hauteur dominante qui consiste à mesurer sur une placette les $n-1$ plus gros arbres sur n ares (Pardé and Bouchon, 1988).

Pour appliquer strictement la définition de la hauteur dominante, seulement l'arbre le plus gros sur 2 ares devrait être mesuré. Lorsque l'essence de l'arbre le plus gros ne correspondait pas à l'essence objectif de la parcelle le deuxième voir le troisième plus gros arbre de la parcelle a été mesuré.

On a fait ça parce qu'on a un intérêt pour connaître aussi la hauteur dominante de l'essence objectif (Ho_{ob}). La raison vient de la propre définition de la hauteur BDR : hauteur moyenne de l'essence objectif. Ça pourrait être intéressant pour des études futures, mais aussi pour une autre approche de ce même projet si la hauteur dominante ne colle pas bien avec la hauteur BDR qu'on cherche à estimer. Dans notre cas, la Ho_{ob} sera utilisée pour l'évaluation de la robustesse de métriques.

Le ruban métrique a été utilisé pour la mesure de la circonférence à 1,30 mètres et pour la hauteur, une perche télescopique.

Après le traitement des données de terrain on va travailler, pour l'ensemble de calculs, avec la hauteur de l'arbre le plus gros de la placette (de 200 m²), sans considérer l'essence de cet arbre là. On peut regarder sur la [Figure 1](#), le nombre des placettes par groupes des hauteurs.

Accroissements

Le vol Lidar a été effectué en mars de 2012 et les mesures de terrain entre mars et avril de 2013. Il existe donc une année de végétation de décalage qu'il faut évaluer pour pouvoir évaluer la hauteur lors du vol Lidar. Pour ça, la dernière saison de végétation des arbres dominants (les mêmes arbres mesurés pour la hauteur) a été estimé dans les groupes des peuplements jusqu'à 5 mètres

Mangeonjean, (2011), l'accroissement en hauteur du hêtre (l'essence majoritaire dans notre échantillonnage) a plus relation avec la séquence de végétation (espèce très sensible aux précipitations) qu'avec la fertilité du sol et, en plus, il n'y a pas des différences statistiquement significatives parmi les accroissements des peuplements jeunes et adultes pour

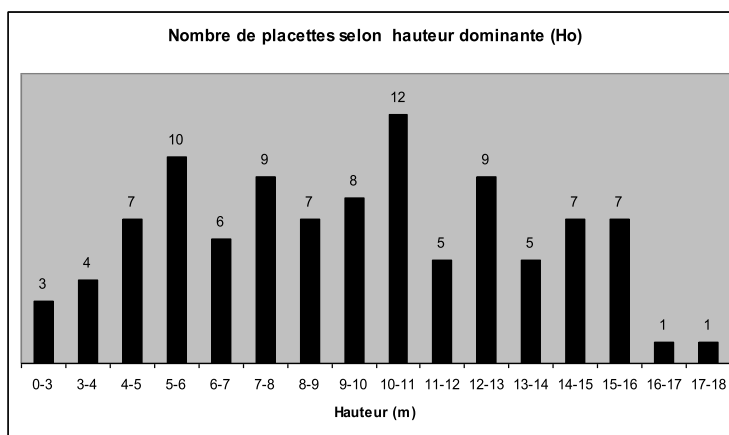


Figure1 : Nombre des placettes par intervalle de hauteur

une saison de végétation. En conséquence, et dû à la difficulté d'arriver à l'extrémité terminale des arbres plus hauts et à l'inexistence de cartes fiables de fertilité du sol, la moyenne d'accroissements (au peu près de 40 cm) a été faite et extrapolé à l'ensemble des peuplements même s'il y a des cas où l'espèce objective n'est pas le hêtre et/ou l'essence dominante. Mais, conformément aux peuplements observés sur le terrain, et le faible nombre de ce type des placettes, on considère que le biais introduit est minimum.

L'intérêt de prendre en compte les accroissements de cette façon viendra par la suite du postérieur ajustement des résultats aux cartes de hauteur selon l'année (dans notre cas bien 2012 ou 2013). C'est-à-dire, il n'aura pas une influence sur l'erreur issue de la modélisation.

Densité

Dans chaque placette, la densité (nombre tiges/ha) a été prise et utilisé comme une variable de contrôle. Ça a été fait à partir des sous-placettes, c'est-à-dire, à partir de placettes plus petites que celles de 2 ares. Ce changement de taille des placettes a sa justification afin de réaliser une démarche plus rapide et pratique. Pour faire une bonne estimation, on a cherché un minimum de 10 tiges par sous-placette. En suivant le protocole Regelor de l'ONF, la surface des sous-placettes était de 6 m² (1,38m de rayon). Néanmoins, dans le cas où la surface n'était pas suffisante pour compter plus de 10 arbres, le rayon a été augmenté jusqu'à 4,37m (60m²) ou 5,64 m (100 m²) si besoin.

Coordonnées GPS

On a utilisé un GPS différentiel Trimble GeoXT qui a été localisé au centre des placettes afin d'obtenir les coordonnées X e Y. En disposition de ces coordonnées on pourra extraire les données Lidar de chaque placette. L'erreur du GPS a une écart d'entre 2-4 mètres dont il a fallu tester l'effet sur la fiabilité des résultats (évaluation de la robustesse d'écrit plus loin).

ANNEXE 2– METRIQUES

1.1 Exemple de famille de métriques.

Il s'agit du même statistique mais calculés en utilisant le nuage de point Lidar dans différents conditions, c'est-à-dire, en considérant différents nombre de points du nuage (soit tous les points, par différents type de impulsion, ou à partir d'une hauteur concret).

Ci-dessus nous avons une exemple pour le percentile 95 :

Métriques	
Code Lasmanager	Description
H95	Hauteur du point situé au quantile 0.95 (tous les points)
H95_fst	Hauteur du point situé au quantile 0.95 (first returns)
H95_fst-nosoil	Hauteur du point situé au quantile 0.95 (Points first classés autres que sol uniquement)
H95_fst-sup1m	Hauteur du point situé au quantile 0.95 (Points first situées strictement au-dessus de 1 mètre)
H95_fst-sup2m	Hauteur du point situé au quantile 0.95 (Points first situées strictement au-dessus de 2 mètres)
H95_fstlst	Hauteur du point situé au quantile 0.95 (first and last returns)
H95_lst	Hauteur du point situé au quantile 0.95 (last returns)
H95_lst-nosoil	Hauteur du point situé au quantile 0.95 (Points last classés autres que sol uniquement)
H95_lst-sup1m	Hauteur du point situé au quantile 0.95 (Points last situées strictement au-dessus de 1 mètre)
H95_lst-sup2m	Hauteur du point situé au quantile 0.95 (Points last situées strictement au-dessus de 2 mètres)
H95_nosoil	Hauteur du point situé au quantile 0.95 (Points classés autres que sol uniquement)
H95_sup1m	Hauteur du point situé au quantile 0.95 (Points situées strictement au-dessus de 2 mètres)
H95_sup2m	Hauteur du point situé au quantile 0.95 (Points situées strictement au-dessus de 1 mètre)

Table 1 – Métriques appartenant à la famille du percentile 95 ; « Code Lasmanager » : nomenclature utilisé sur le logiciel ; « Description » : petite définition de la métrique évaluée.

1.2 Description des métriques et signification

Selon (Bock et al., 2011) les métriques peuvent être classifiés en trois grandes catégories: les indicateurs de distribution de hauteur (percentiles, coefficients de variation, moyenne, l'écart type des hauteurs ...), les indicateurs de densité (nombre de points par tranche de hauteur, pourcentage de points, nombre total de points); et les indicateurs spatiaux (base sur des maxima locaux).

Ci-dessus (tableau 2) une petite explication des métriques utilisés ou nommées dans quelque section du rapport.

Métriques		
Code Lasmanager	ID	Description
d9	d9	Densité relative entre dans la tranche (10 tranches de mètres à H95) (tous les points)
dHmean_lst	dHm_lst	Moyenne des (H-Hmoy) des pints de la placettes (last returns)
dHmean_sup2m	dHm	Moyenne des (H-Hmoy) des points de la placettes (points situées strictement au dessus de 2 mètres)

H75_fst.nosoil	H75	Hauteur du point situé au quantile 0.75 (Points first classés autres que sol uniquement)
H80_fst.nosoil	H80	Hauteur du point situé au quantile 0.80 (Points first classés autres que sol uniquement)
H90_fst.nosoil	H90	Hauteur du point situé au quantile 0.90 (Points first classés autres que sol uniquement)
H95_fst.nosoil	H95	Hauteur du point situé au quantile 0.95 (Points first classés autres que sol uniquement)
H99_fst.nosoil	H99	Hauteur du point situé au quantile 0.99 (Points first classés autres que sol uniquement)
H99_fstlst	H99_fstlst	Hauteur du point situé au quantile 0.99 (first and last returns)
Hcv_lst	Hcv_lst	Coefficient de variation des hauteurs de points de la placette (last returns)
Hmax	Hmax	Hauteur du point le plus haut de la placette (tous les points)
Hmv135	Hmv135	Hauteur moyenne de maxima locaux (1,3,5) de la placette
Hmv3_Bd	Hmv3	Moyenne des 3 plus haut avec application de 1 x le rayon d'exclusion de Badeau
Hmv5_Bd	Hmv5	Moyenne des 5 plus haut avec application de 1 x le rayon d'exclusion de Badeau
Hmv6_Bd	Hmv6	Moyenne des 6 plus haut avec application de 1 x le rayon d'exclusion de Badeau
N.m2_lst.sup1m	N_m	Nombre de points de la placette (situées strictement au dessus de 1 mètre)

Tableau 2 - Description des métriques nommés dans le rapport issues du logiciel Lasmanager de l'ONF. En noir, les métriques qui ont été sélectionnées pour l'analyse de la robustesse.

Code Lasmanager : nomenclature utilisée sur le logiciel ; « ID » : abréviation utilisée dans le rapport pour des métriques ;
« Description » : petite définition de la métrique évaluée.

1.3 Calcul de HmvX (statistiques de maxima locaux)

Le Hmv5 est la hauteur moyenne des 5 plus hauts points de la placette. On identifie le point le plus haut de la placette, puis on ignore les points dans un rayon d'exclusion dépendant de la hauteur de ce point. Ce rayon est calculé d'après des relations allométriques hauteur de l'arbre-diàmetre du houppier établies sur plus de 800 hêtres et chênes mesurés par l'INRA ($r^2=0.54$). On cherche ainsi de suite les 2^e, 3^e, 4^e et 5^e points les plus hauts de la placette. Ça permet d'éviter de sélectionner plusieurs points dans un même arbre (Bock et al., 2011).

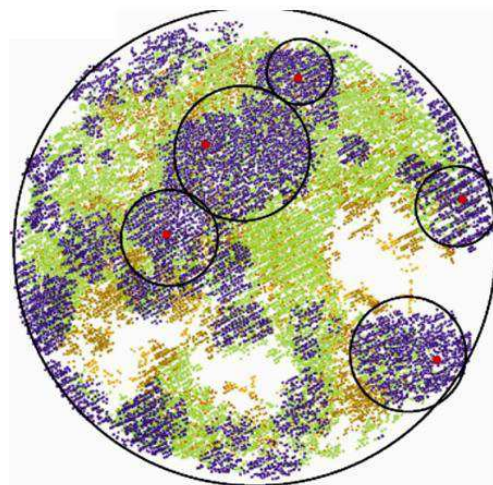
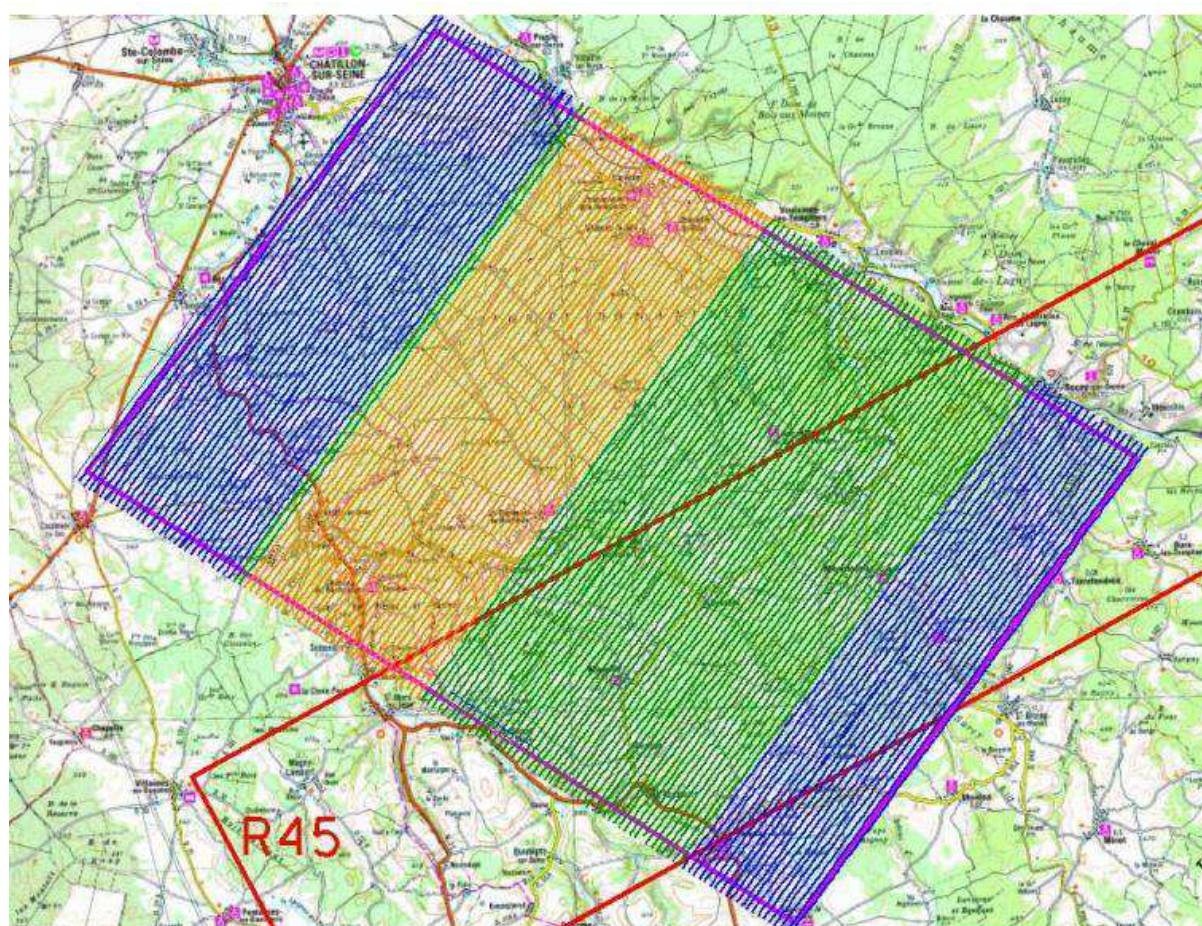


Figure- Schéma illustratif de la méthodologie utilisée pour le calcul de la métrique Hmv5_Bd (moyenne des 5 points signalés)

ANNEXE 3 – PARAMETRES DE VOL

Hauteur de vol:	550	m
Vitesse de vol:	100	kts
Fauchée d'une bande:	635	m
Recouvrement latéral:	71	%
Angle de scan:	60	°
Précision altimétrique (écart-type):	10	cm
Précision planimétrique (écart-type):	15-20	cm
Densité sur une bande:	4.1	impulsions/m ²

Le prestataire a fourni l'information dans les systèmes géodésiques RGF 93 Lambert 93, et NGF IGN69 NTF Lambert II étendu, NGF IGN69 et en format Las 1.2. Ce format contient les éléments suivant: X, Y, Z, classe, intensité, n° de ligne de vol, temps, n° de l'écho, l'angle du miroir.



Plan de vol réalisé le 13 (bleu), 14 (vert) et 15 (orange) mars 2012

ANNEXE 4– SELECTION DE LA VARIABLE EXPLICATIVE

La stabilité a été classée en différents groupes de couleurs pour chaque métrique et condition selon si le comportement observé est très stable, stable, moyenne ou il n’y a pas de stabilité (Voir Tableau a). Pour le RMSEcv une classification a été fait aussi mais en considérant la dimension du RMSEcv, c’est-à-dire, s’il est petite, moyenne, grand ou très variable (donc il n’y a pas un comportement uniforme) (Voir Tableau b).

Stabilité	ID	Critère de classification
Très stable		Tous les RMSE
Stable		Il y a un RMSE différent
Stabilité moyenne		Il y a deux RMSEs différents ou un très différent
Pas de stabilité		3 o plus RMSEs différents

RMSE	ID	(m)
Petite	P	<0.8
Moyenne	M	0.8-1
Grand	G	>1
Variable	V	

Tableau b: Critères de classification pour le RMSEcv

Tableau a : Critères de classification pour la stabilité

Métrique	Variation rayon		RMSEcv		Variation XY		RMSEcv		Variation Densité points en %		Variation bandes vol	
Hmv3_Bd			P	M			P				P	P
Hmv5_Bd			P	M			P				P	P
Hmv135			P	M			P				P	P
Hmv6_Bd			P	M			P				P	P
H99_fst.nosoil				M			M				P	P
Hmax			V				P	M				P
H95_fst.nosoil				M			P	M			M	M
H90_fst.nosoil				M			M				M	M
H80_fst.nosoil				M			M				M	M
dHmean_sup2m				G			G				G	G
H75_fst.nosoil				G			G				G	G

Deuxièmement, chaque couleur a été reclassifié pour chaque critère d’élection d’une façon séparé. C'est-à-dire, à chaque couleur on a associé une valeur et on a sommé l’ensemble des valeurs pour la métrique. On a classifié par intervalles de bonté en fonction de si la somme obtenue est haute ou basse pour le critère étudié (Voir Tableau c et d). La reclassification par intervalle a été soumis aux résultats obtenus.

Re-classification des résultats :

Id		PTs	Classif	Somme
Vert		5	Très bonne	>18
		3	Bonne	15-18
Orange		2	Moyenne	12-15
Rouge		0	Pas de stabilité (PS)	<12

Tableau c : Reclassification pour le critère stabilité

Id		PTs	Classif	Somme
P		5	Très bonne	>18
	M	4	Bonne	14-18
M		3	Moyenne	9-13
G		2	Null	<9
V		0		

Tableau d : Reclassification pour le critère RMSEcv

Métrique		Somme		Somme		STABILITE		RMSE	
		stabilité		RMSEcv					
Hmv6_Bd		16		19		Bonne		Très bon	
Hmv5_Bd		16		19		Bonne		Très bon	
Hmv3_Bd		14		19		Moyenne		Très bon	
Hmv135		16		19		Bonne		Très bon	
Hmax		8		14		Pas de stabilité		Bon	
H99_fst.nosoil		16		17		Bonne		Bon	
H95_fst.nosoil		20		14		Très Bonne		Bon	
H90_fst.nosoil		20		12		Très Bonne		Moyenne	
H80_fst.nosoil		20		12		Très Bonne		Moyenne	
H75_fst.nosoil		18		8		Bonne		Mauvais (MM)	
dHmean_sup2m		11		8		Pas de stabilité		Mauvais (MM)	

ANNEXE 5 – DIAGNOSE DE MODELE

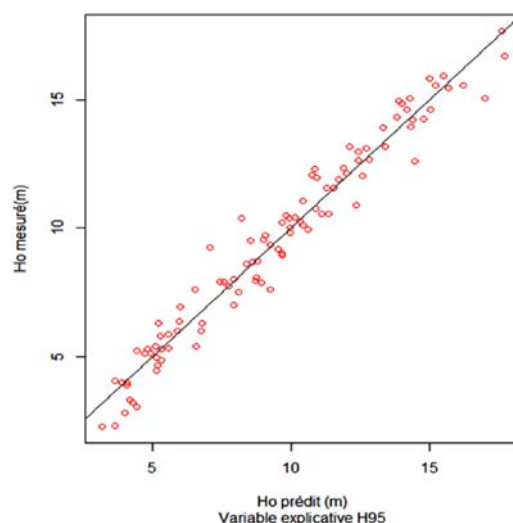
Modèle avec H95_fst.nosoil comme variable explicative

```
Call:
lm(formula = Ho ~ H95_fst.nosoil, data = modeles)

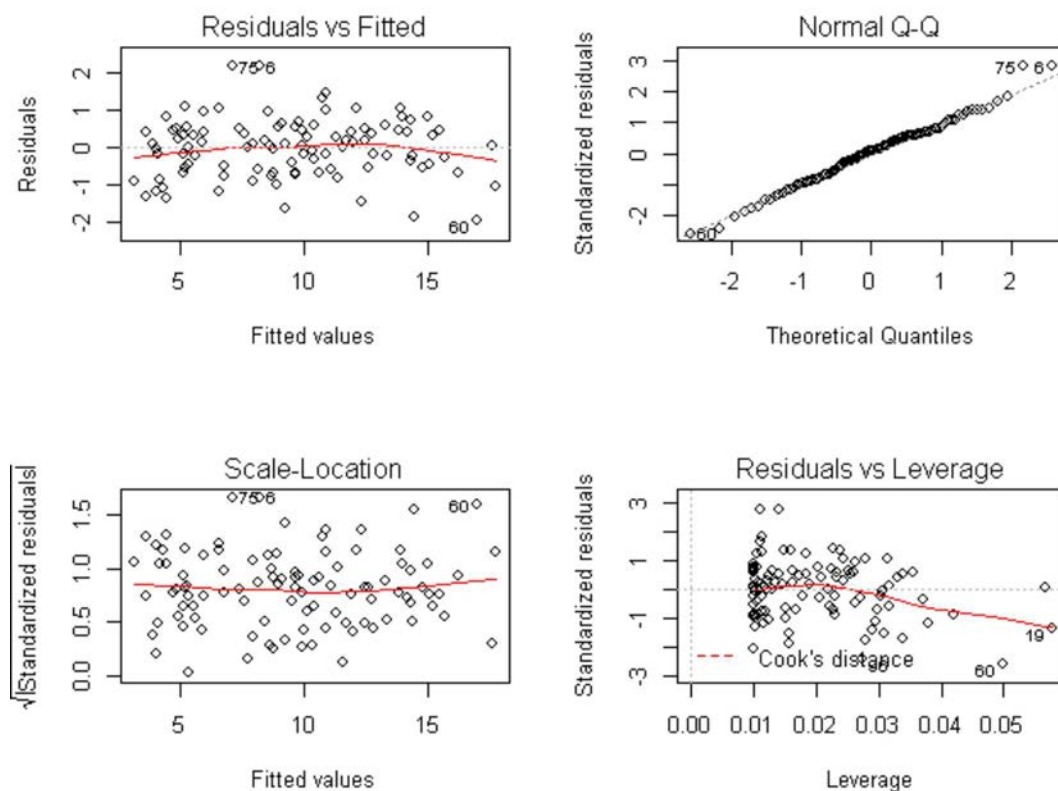
Residuals:
    Min       1Q   Median       3Q      Max
-1.9476 -0.5324  0.0626  0.4765  2.1802

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.24905     0.16874   13.33  <2e-16 ***
H95_fst.nosoil  1.02063     0.02116   48.24  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7831 on 99 degrees of freedom
Multiple R-squared:  0.9592, Adjusted R-squared:  0.9588
F-statistic: 2327 on 1 and 99 DF, p-value: < 2.2e-16
```



Métrique	β_0	β_1	Modèle	RMSE	R^2_{ad}	pval_int	RMSEcv	Biais	Effcv
H95	1.0206	2.249	Hop = 1.0206*H95+2.2490	0.7831	0.9588	7.39E-24	0.7922	-0.0013	0.994



Test de Breusch-Pagan de hétéroscédasticité

Breusch-Pagan test

```
data: Ho ~ H95_fst.nosoil
BP = 0.0501, df = 1, p-value = 0.8229
```

$\alpha = 0,05 < 0,8229 \rightarrow$ la variance est égale dans tout le graphe

Test Shapiro-Wilks (normalité)

Shapiro-Wilk normality test

```
data: modeles$residuals.H95_model
W = 0.9897, p-value = 0.63
```

Comme $p\text{-value} = 0,63 > 0,05 = \alpha \rightarrow$ les résidus ont un comportement normal

Test RESET de non linéarité


```

RESET test

data: Ho ~ H95fst.nosoil
RESET = 6.2825, df1 = 1, df2 = 98, p-value = 0.01384

```

Como $p\text{-value} = 0,01384 < 0,05 = \alpha \rightarrow$ le modèle n'ajuste pas parfaitement

Test de Bonferroni (valeurs extrêmes)

```

No Studentized residuals with Bonferonni p < 0.05
Largest |rstudent|:
  rstudent unadjusted p-value Bonferonni p
6 2.902519      0.0045726      0.46184

```

Comme $p\text{-value} = 0,461 > 0,05 = \alpha \rightarrow$ le modèle n'a pas des valeurs extrêmes

Modèles avec Hmv5_Bd comme variable explicative

```

Call:
lm(formula = Ho ~ Hmv5_Bd, data = modeles)

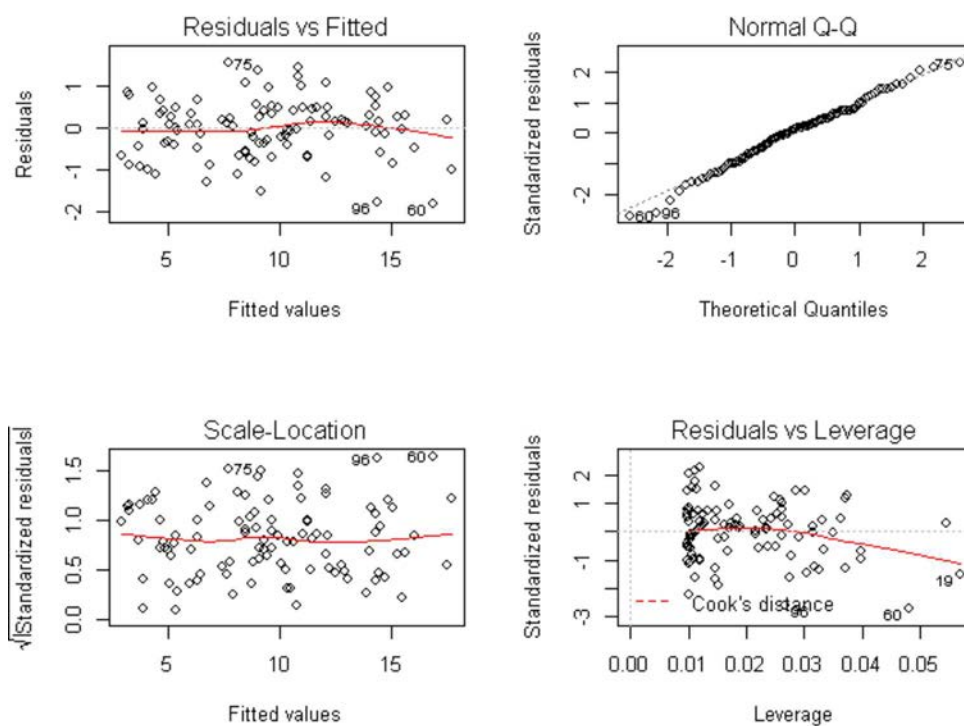
Residuals:
    Min       1Q   Median       3Q      Max
-1.81140 -0.43260  0.08665  0.43630  1.55578

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.79779     0.17195   4.64 1.07e-05 ***
Hmv5_Bd       1.00737     0.01831  55.00 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6901 on 99 degrees of freedom
Multiple R-squared:  0.9683, Adjusted R-squared:  0.968
F-statistic: 3025 on 1 and 99 DF, p-value: < 2.2e-16

```

Métrique	β_0	β_1	Modèle	RMSE	R^2_{ad}	pval_int	RMSEcv	Biais	Effcv
Hmv5	1.0074	0.7978	Hop = 1.0074*Hmv5+0.7978	0.6901	0.968	9.99E-06	0.6969	-3.00E-04	0.9953



Test de Breush-Pagan de hétéroscédasticité

Breusch-Pagan test

```
data: Ho ~ Hmv5_Bd
BP = 0.7962, df = 1, p-value = 0.3722
```

$\alpha = 0,05 < 0,3722 \rightarrow$ la variance est égale dans tout le graphe

Test Shapiro-Wilks (normalité)

Shapiro-Wilk normality test

```
data: modeles$residuals.Hmv5_model
W = 0.9907, p-value = 0.7118
```

Comme $p\text{-value} = 0,7118 > 0,05 = \alpha \rightarrow$ les résidus ont un comportement normal

Test RESET de non linéarité

```

RESET test

data:  Ho ~ Hmv5_Bd
RESET = 3.2145, df1 = 1, df2 = 98, p-value = 0.07607

```

Como $p\text{-value} = 0,07607 > 0,05 = \alpha \rightarrow$ le modèle ajuste parfaitement

Test de Bonferroni (valeurs extrêmes)

```

No Studentized residuals with Bonferonni p < 0.05
Largest |rstudent|:
      rstudent unadjusted p-value Bonferonni p
60 -2.780326      0.0065121      0.65772

```

Comme $p\text{-value} = 0,65772 > 0,05 = \alpha \rightarrow$ le modèle n'a pas des valeurs extrêmes

ANNEXE 6 – COMPARAISON DES MODELES

Auteur		Métriques	Modèle	Indicateur	Variable	Comentaires
Næsset et Bjerknæs (2001)		H90, d1	log-log	R2 0,83 // RMSE* 0,131 (RMSEdest -3,706)	Hauteur dominante	Placettes de 200 m2
Næsset (2002)		H90, Hmax	log-log	R2 0,93 // RMSE* 0,07 (RMSEdes - 1,075)	Hauteur dominante	Jeunes peuplements; parcelles de 200m2
Næsset, (2004)		H90, d0f, d6f, d9f	log-log	R2 0,92 // RMSE 0,07 (RMSEdes -1,075)	Hauteur dominante	Placettes de 232 m2; Jeunes peuplements
Holmgren and Jonsson, (2004)		Percentiles de hauteur*	lineaire	RMSE 0,8	Hauteur moyenne de Lorey	placettes de 900 m2
Hall et al. (2005)		Hmax25	lineaire	RE 0,69	Hauteur moyenne de Lorey	Peuplements de Pinus ponderosa
Hall et al. (2005)		Hmax50, SDh3	lineaire	RE 1,18	Hauteur moyenne	Peuplements de Pinus ponderosa
Andersen et al., (2005)		Hmax, H25, H50, H75, D	linéaire	R2 0,98 // RMSEcv 1,5	Hauteur de la canopée (canopy height*)	Placettes entre 0,02 et 0,2 ha;
Heurich et Thoma, (2008)		H90, H70, Pril	lineaire	R2 0,95 // RMSEcv 4,3%	Hauteur dominante	Fagus sylvatica
Martins, (2009)		d9 / Hmv5	lineaire	R2 0,99 // RMSE 0,761	Hauteur dominante	Peuplements feuillus
Næsset, (2009)		H90	log-log	R2 0,93 // RMSE 0,0026 (RMSEdes -1,01)	Hauteur moyenne de Lorey	Peuplement d'épicea
Gobakken and Næsset, (2009)		Hmean; d1	log-log	R2 0,94// RMSE* 0,07 (RMSEdes - 1,075)	Hauteur moyenne de Lorey	placettes de 200 m2; Jeunes peuplements
Zaldo et al., (2010)		H80	lineaire	R2 0,823 //RMSE 1,710	Hauteur max. (top height)	Fagus sylvatica
Estornell Cremades, (2011)		P95	lineaire	R2 0,71// RMSE 0,13	Hauteur de la vegetation	Peuplements arbusives méditerranéan avec une hauteur petite
Treitz et al.(2012)		Hmax, H90	linéaire	R2 0,939// RMSE 0,87	Hauteur max. (top height)	Peuplements de feuillus; densité de points LiDAR 0,5 pts/m2
González-Ferreiro et al.(2012)		H95	lineaire	R2 0,846// RMSE 1,988	Hauteur dominante	Placettes 225 m2; Peuplements de Pinus radiata

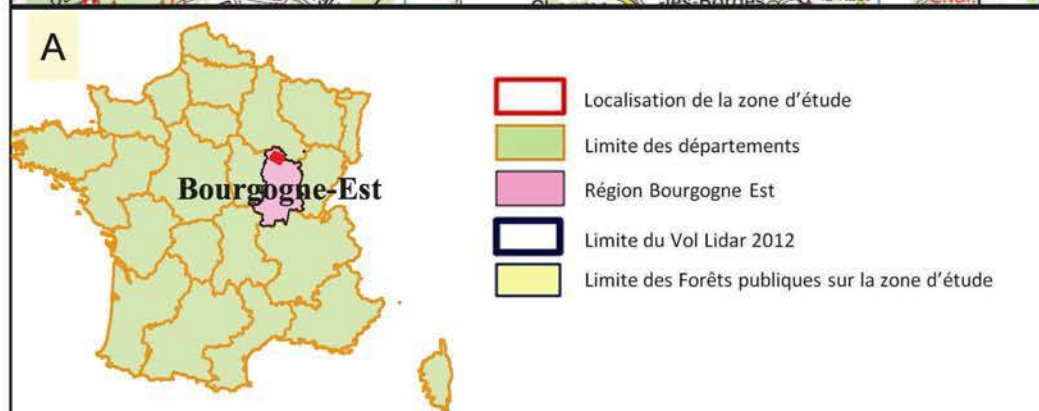
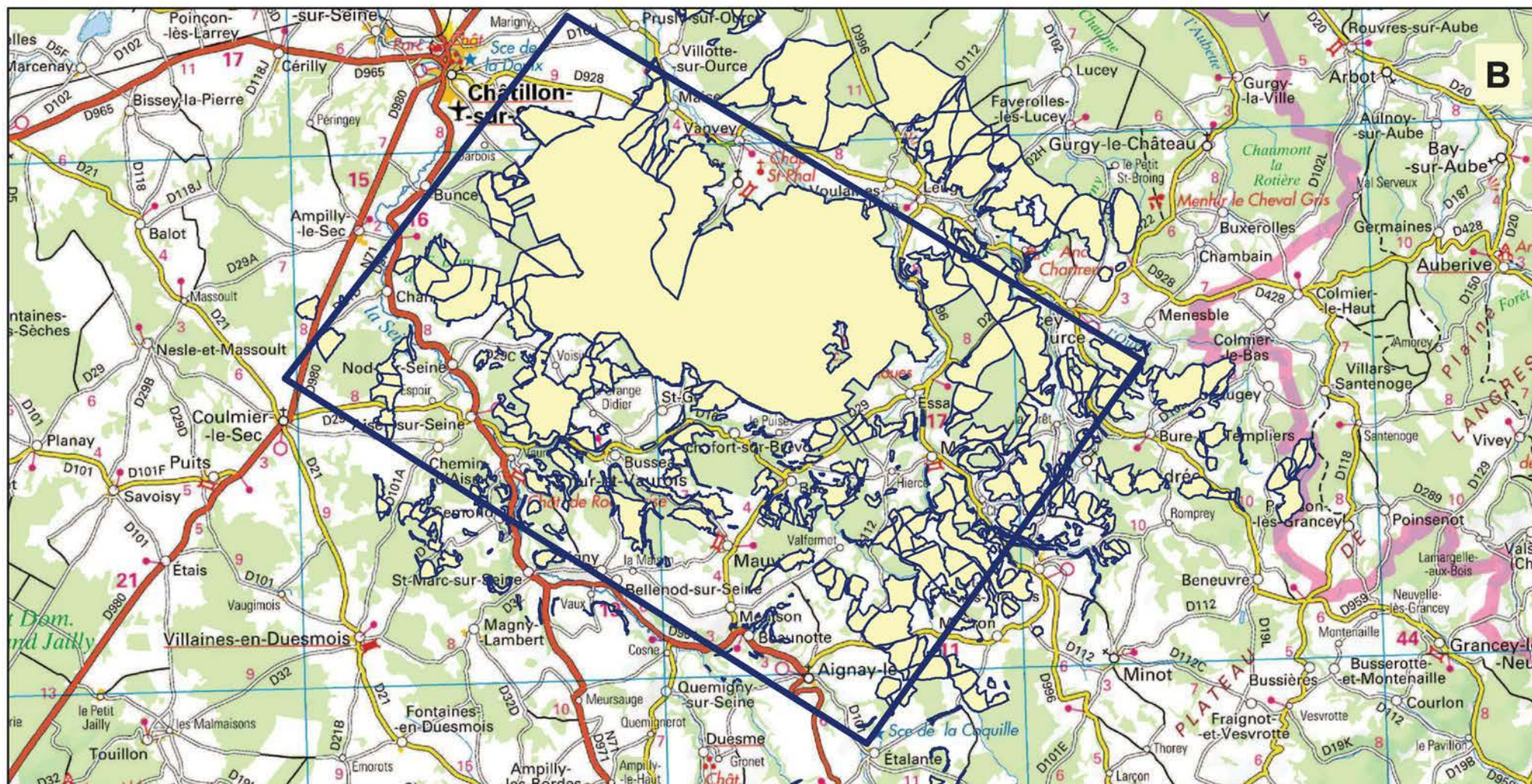
Tableau - Résumé bibliographique des indicateurs des quelque modèles mathématiques développées pour estimer la hauteur (soit hauteur dominante, moyenne de Lorey ou maximale).
RMSE : erreur quadratique moyenne ; RE : erreur standar ; R2 : coefficient de détermination; RERMSE* indique que le valeur de l'erreur quadratique moyenne est calculé pour le modèle logarithmique; RMSEdes : erreur des prédiction après de-transformation inverse du RMSE du modèle logarithmique (on introduit un biais au modèle quand on fait cette de-transformation donc il faudra corriger le modèle en multipliant pour un coefficient de correction ($e^{RMSE/2}$)) ; Percentiles de hauteur * : Holmgren and Jonsson, (2004) testent les percentiles 10, 20,..., 90, 95 et 100 et finalement ils sélectionnent le percentile avec le RMSE le plus petite (dans le rapport il n'est pas spécifié lequel) ; Canopy height* : hauteur plus haut quand la densité de la canopé est plus grand de 0,011kg/m³.

Metricque		Signification
Hmean		Moyenne arithmetique de la hauteur de la couvert
d1,		Indicateur 1 de densité (il y a 10 tranches verticaux construits; la hauteur de chaque tranche est définé comme le dixième de la distance entre le H95 et la valeur plus bas de hauteur de la couvert
Hmax		Hauteur maximale
d0f, d6f, d9f		Densité de la couvert qui correspond aux proportions de premiers retours au-dessus de la fransion 1, 6, et 9 (il y a dix fransions)
D		Pourcentage des premiers retours dedans de la canopée
Pril		Taux de pénétration à travers de la couche intermédiaire; à partir de toutes les premiers et deniers retours LiDAR qui sont sous la couche relative défini par 0,5*hauteur totale
H95, H90, H80		Percentiles de hauteur 95, 90, 80 ...
Hmv5		Moyenne des 5 plus haut avec application de 1 x le rayon d'exclusion de Badeau
d9		Densité relative entre dans la tranche (10 tranches de mètres à H95)
SDh3		Ecart type de la hauteur calculé à partir des premiers retours situés au-dessus de 3 mètres
Hmax50, Hmax 25		Hauteur moyenne du 50% (ou 25%) des premiers retours vegetation plus hauts

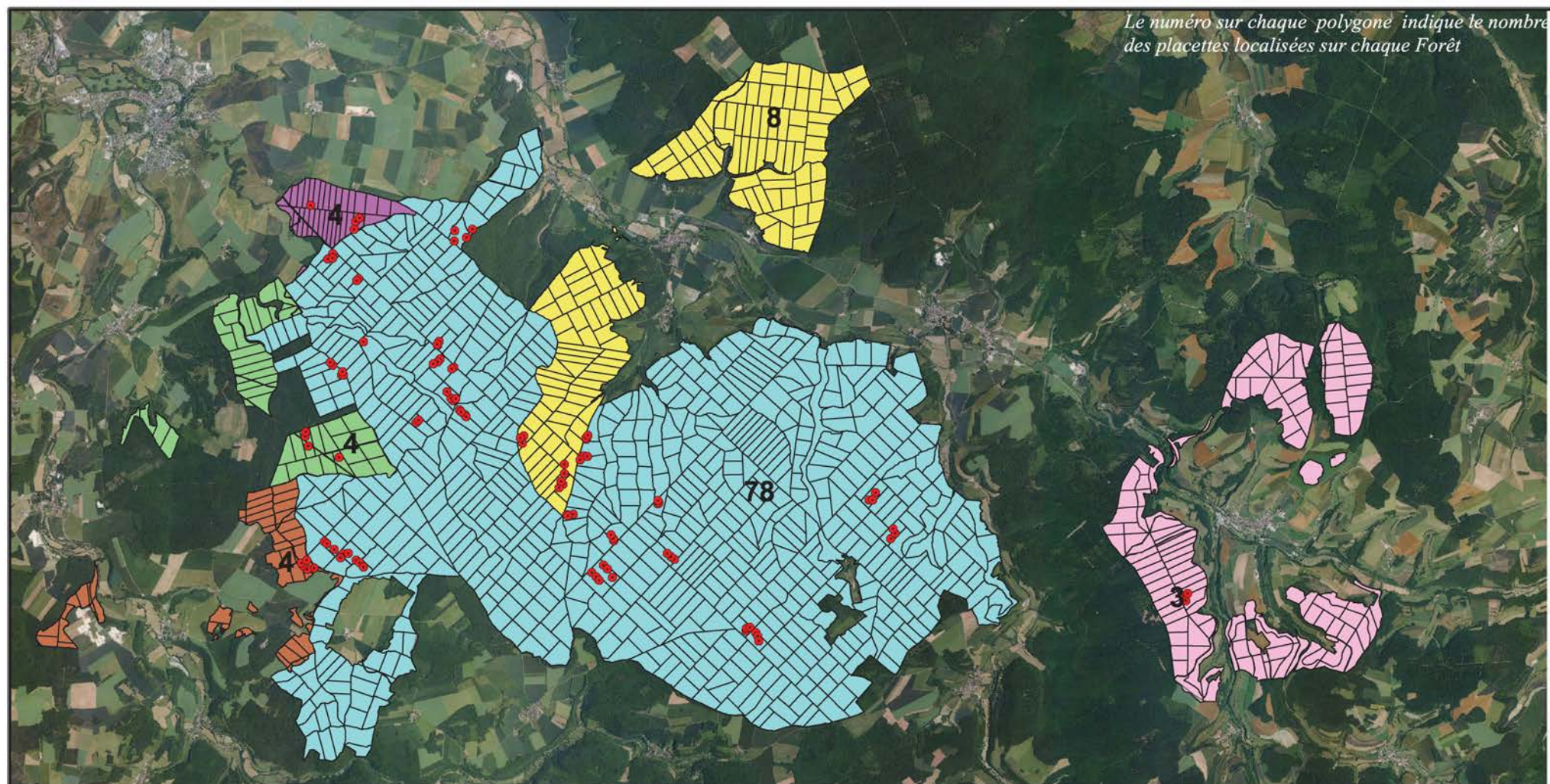
Dans les différents articles cités, il est habituel trouver des autres modèles en fonction de l'étude faite. On a sélectionné les modèles qui ont quelque similitude avec notre cas d'étude afin de pouvoir faire une comparaison. L'explication concrète de chaque métrique peut être consultée dans chaque article pour plus de détails.

ANNEXE

CARTOGRAFIQUE



		Office National des Forêts Pole R&D de Nancy	
Titre de la carte: Situation de la zone d'étude et limites du vol Lidar 2012			
Projet	Estimation de la hauteur de jeunes peuplements à partir de données Lidar aériennes		Carte Annexe 1
Auteur	Nuria Sánchez López		Signature:
Date	Août 2013		
		Echelle: A - 1: 2000000 B- 1: 200000	



- Placettes d'échantillonnage
- Parcelles
- Forêt communale de Buncey
- Forêt communale de Châtillon-Sur-Seine
- Forêt communale de Nod-Sur-Seine
- Forêt communale de Recey-Sur-Ource
- Forêt communale indivise de Vanvey-Villiers
- Forêt domaniale de Chatillon

0 1 300 2 600 5 200 7 800 10 400 Mètres



Office National des Forêts Pole R&D de Nancy

Titre de la carte: Localisation des placettes sur l'ensemble des Forêts de la zone d'étude

Projet Estimation de la hauteur de jeunes peuplements à partir de données Lidar aériennes

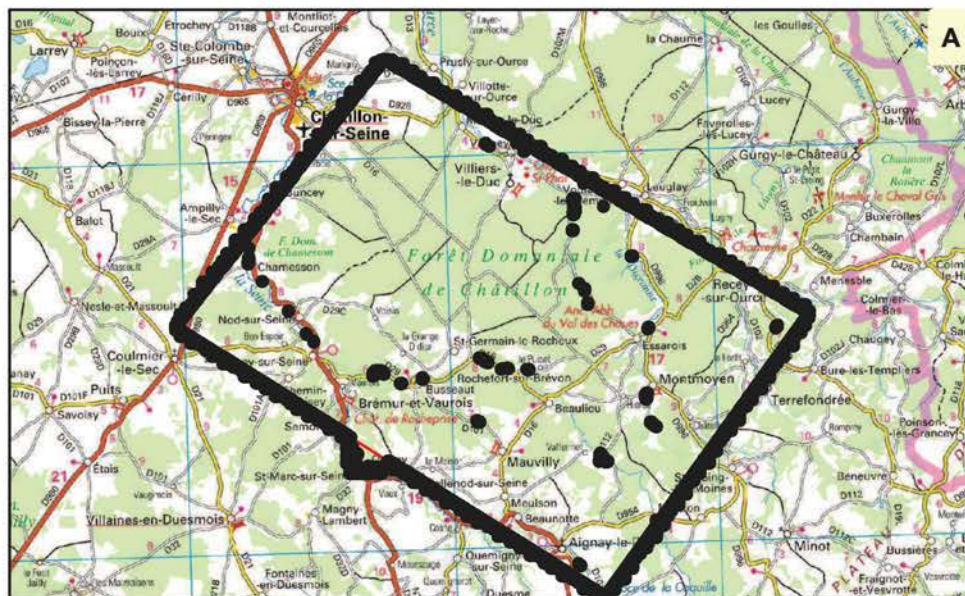
Carte Annexe **2**

Auteur Nuria Sánchez López

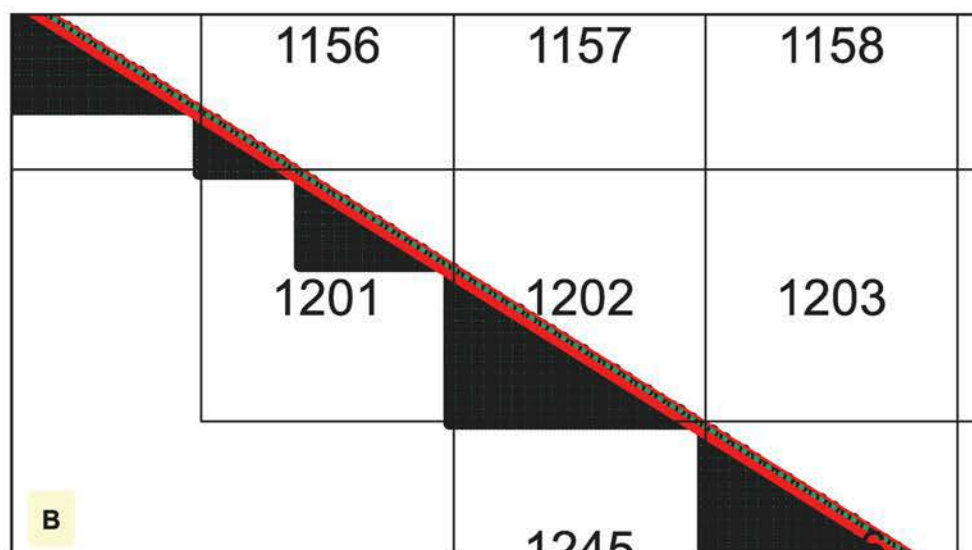
Signature

Date Août 2013

Echelle: 1:115 000



Limite du vol Lidar sur la zone d'étude et (en noir) localisation de points des placettes « NODATA »



Détail de points placettes « NODATA » correspondant aux limites du vol Lidar ; en rouge limites du vol Lidar ; en noir points placettes « NODATA » et couche de dalles Lidar fourni par le prestataire



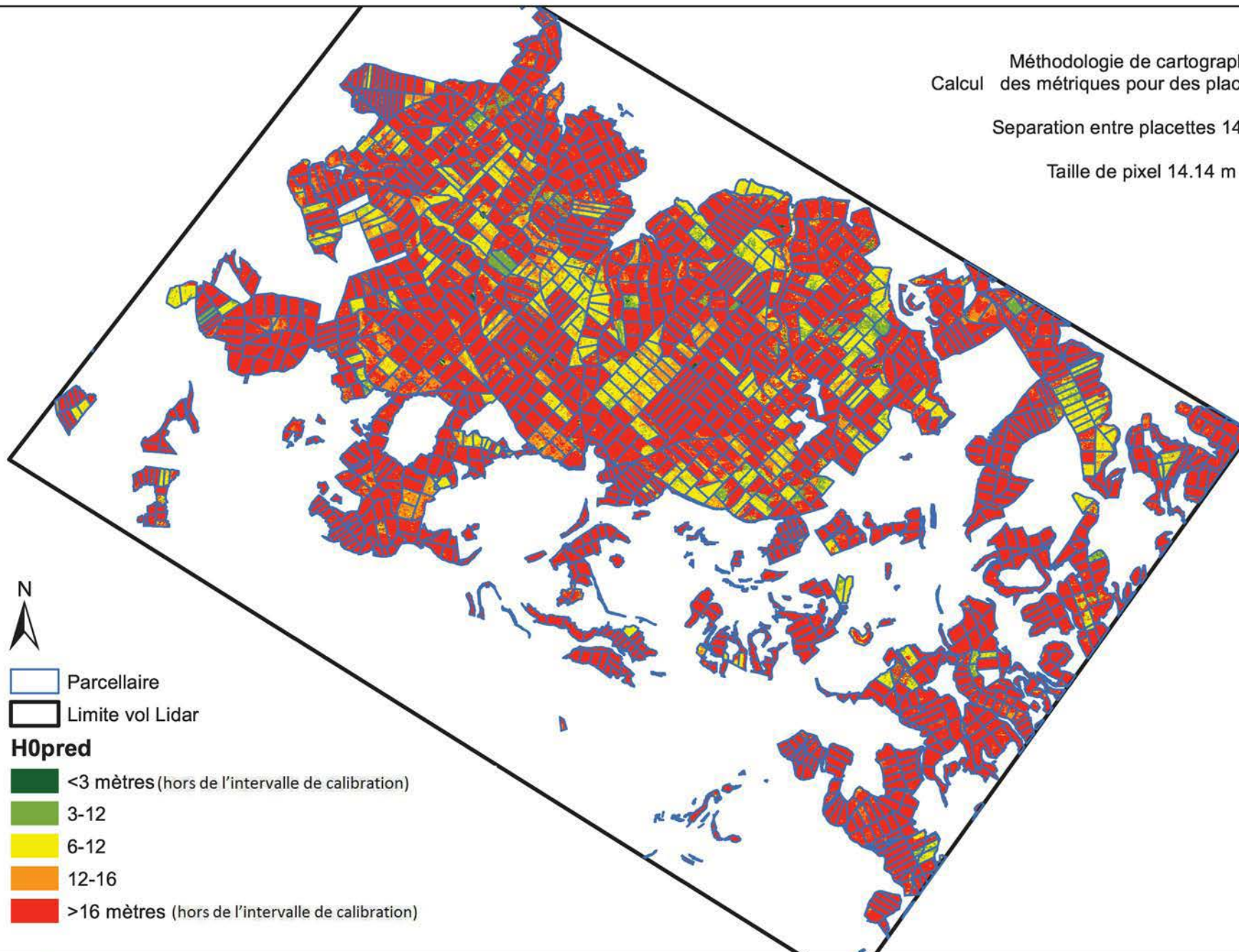
Détail de points placettes « NODATA » correspondant aux zones d'eau

		Office National des Forêts Pole R&D de Nancy	
Titre de la carte: Placettes « NODATA » obtenues du maillage systématique issu du Logiciel ONF « Lasmanager »			
Projet		Estimation de la hauteur de jeunes peuplements à partir de données Lidar aériennes	Carte Annexe 3
Auteur		Nuria Sánchez López	Signature:
Date		Août 2013	
		Echelle: A - 1:10000 B - 1:15000 C - 1:6000	

Méthodologie de cartographie:
Calcul des métriques pour des placettes carrées

Séparation entre placettes 14.14 m

Taille de pixel 14.14 m



Parcellaire



Limite vol Lidar

H0pred



<3 mètres (hors de l'intervalle de calibration)



3-12



6-12



12-16



>16 mètres (hors de l'intervalle de calibration)



Titre de la carte:

Hauteur dominante prédite par H95
Méthodologie 1

Projet

Estimation de la hauteur de
jeunes peuplements à partir
de données Lidar aériennes

Auteur Nuria Sánchez López

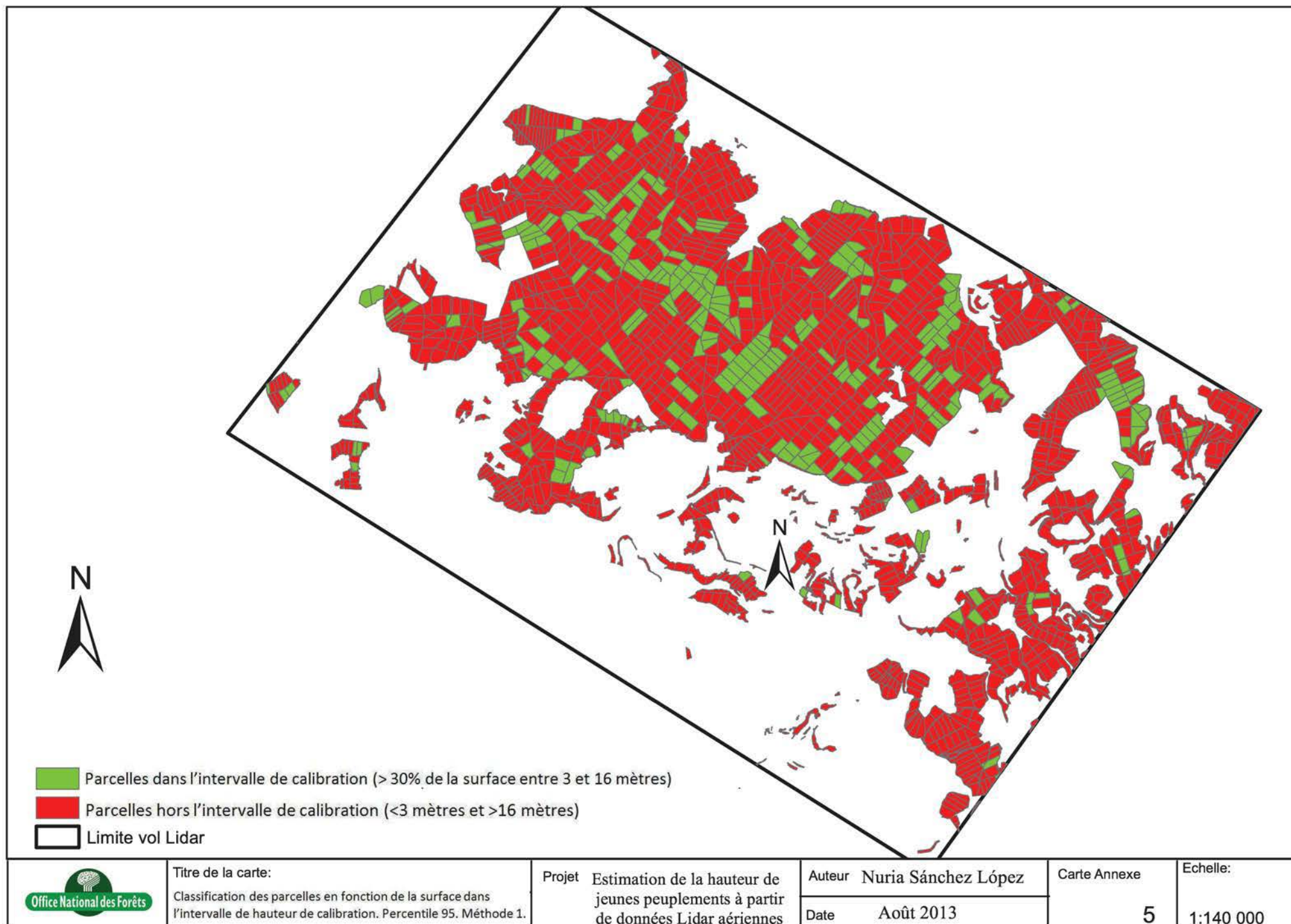
Date Août 2013

Carte Annexe

4

Echelle:

1:130 000

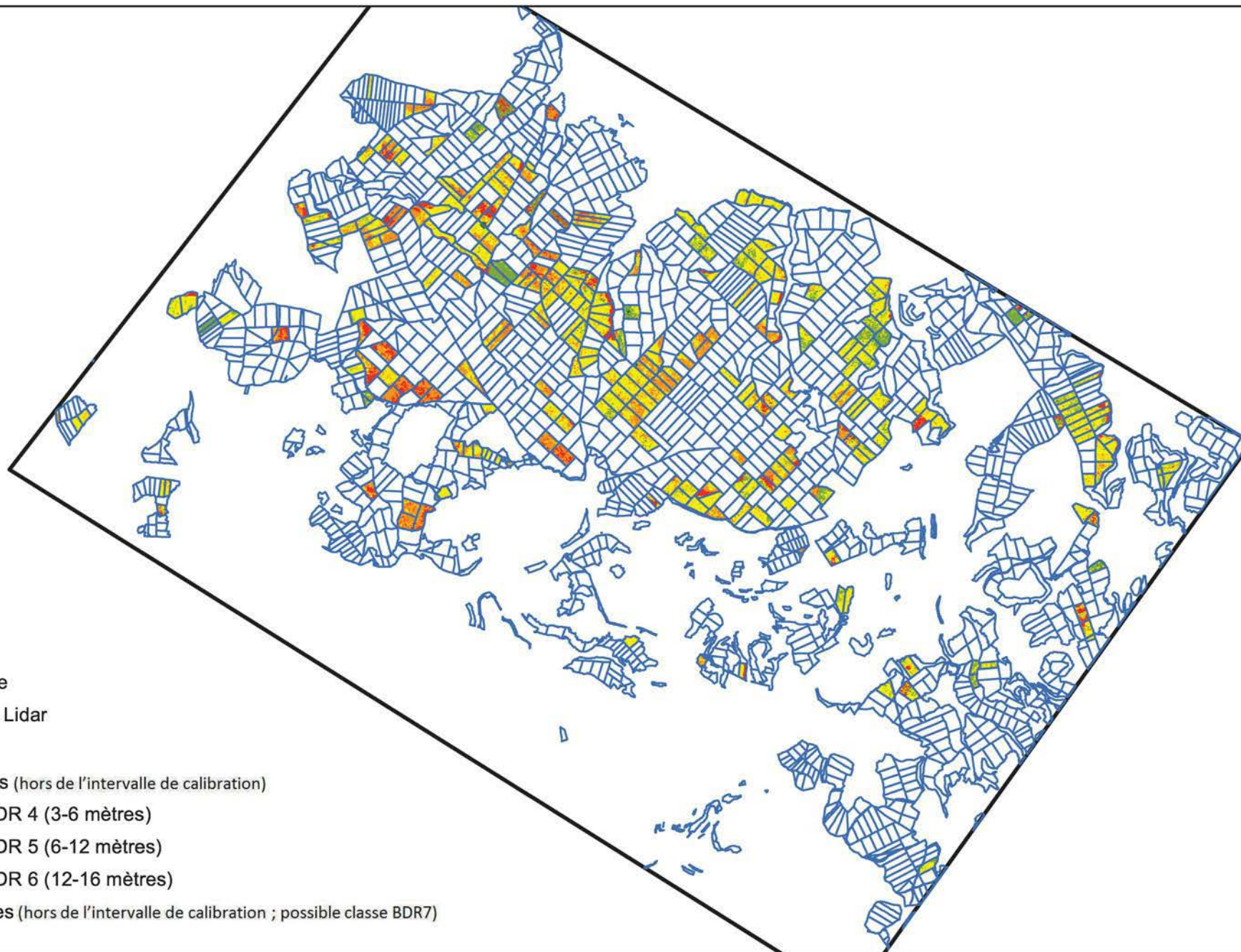




-  Parcelaire
 Limite vol Lidar

H0pred

-  < 3 mètres (hors de l'intervalle de calibration)
 Classe BDR 4 (3-6 mètres)
 Classe BDR 5 (6-12 mètres)
 Classe BDR 6 (12-16 mètres)
 >16 mètres (hors de l'intervalle de calibration ; possible classe BDR7)



Titre de la carte: Carte de hauteur dominante
prédite par H95 (méthode 1)
selon classes de hauteur BDR

Projet Estimation de la hauteur de
jeunes peuplements à partir
de données Lidar aériennes

Auteur Nuria Sánchez López

Date Août 2013

Carte Annexe

6

Echelle:

1:130 000

Methodologie de cartographie:
Calcul de métriques pour des placettes circulaires

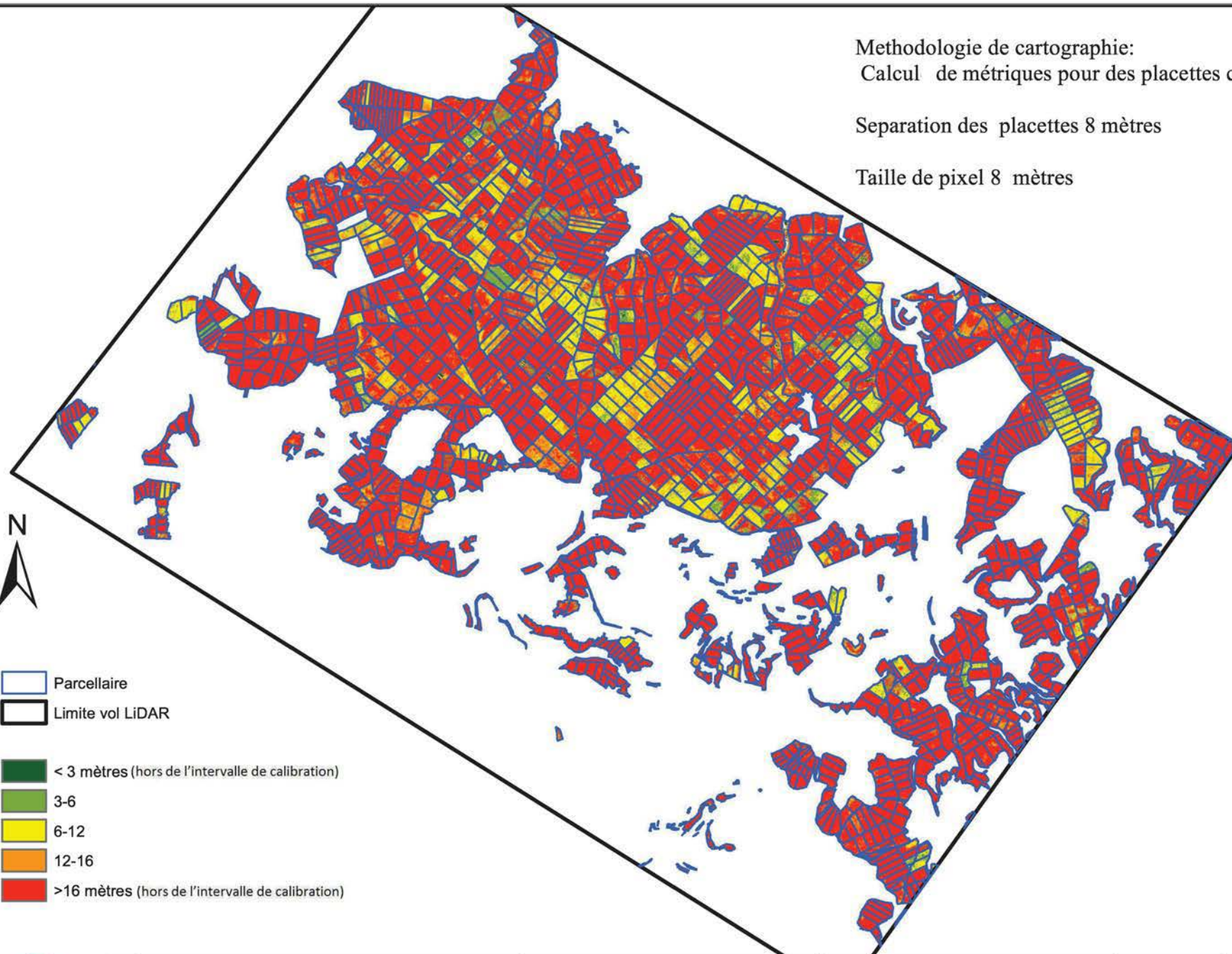
Separation des placettes 8 mètres

Taille de pixel 8 mètres



 Parcellaire
 Limite vol LiDAR

 < 3 mètres (hors de l'intervalle de calibration)
 3-6
 6-12
 12-16
 >16 mètres (hors de l'intervalle de calibration)



Titre de la carte: Hauteur dominante prédite par H95
Méthodologie 2

Projet Estimation de la hauteur de
jeunes peuplements à partir
de données Lidar aériennes

Auteur Nuria Sánchez López

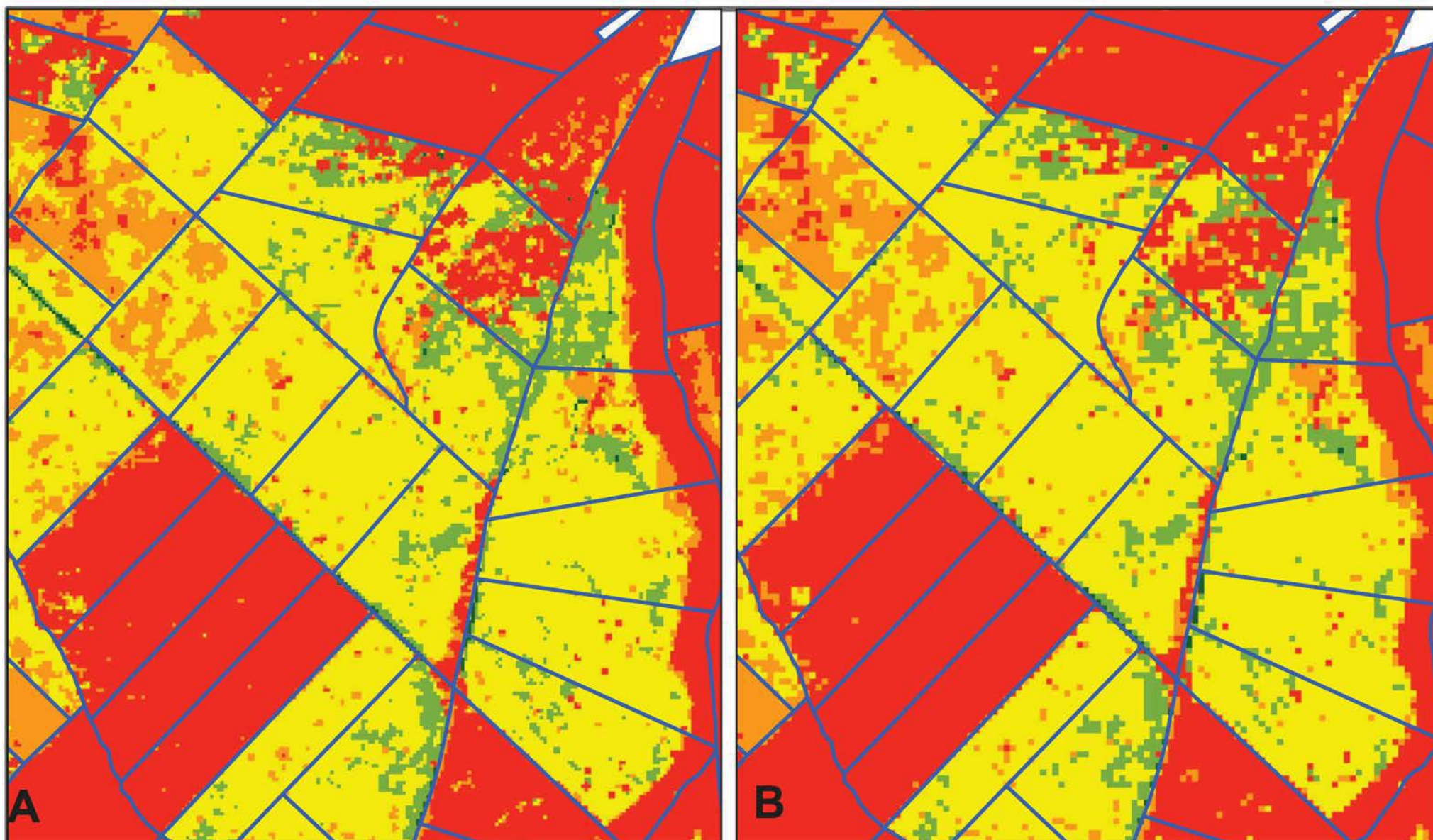
Date Août 2013

Carte Annexe

7

Echelle:

1:130 000



N

< 3 mètres (hors de l'intervalle de calibration)

3-6

6-12

12-16

>16 mètres (hors de l'intervalle de calibration)

Détail de la carte de hauteur en appliquant le modèle du H95 dans l'ensemble de la zone d'étude

A – Méthode 2 (placettes circulaires). Taille pixel 8 mètres

B- Méthode 1 (placettes carrées). Taille de pixel 14.14 mètres

	Titre de la carte:	Projet	Estimation de la hauteur de jeunes peuplements à partir de données Lidar aériennes	Auteur	Nuria Sánchez López	Carte Annexe	Echelle:
	Comparaison des deux méthodologies de cartographie			Date	Août 2013		

RESUME

VALIDATION DE L'ESTIMATION DE LA HAUTEUR DES JEUNES PEUPELEMENTS FORESTIERS A PARTIR DE DONNEES LIDAR AERIENNES

On a estimé la hauteur dominante de jeunes peuplements de hêtre en croisant l'information LiDAR aérienne et les données de terrain. A partir de 101 placettes de terrain, une série de métrique ont été extraites du nuage de points Lidar (percentiles, valeurs maximales et minimales, hauteurs de maxima locales, etc). Elles ont été utilisées comme les variables explicatives des modèles linéaires testés. L'incertitude du positionnement du GPS dans les placettes de terrain et l'effet de la variation de la densité de points Lidar ont été évalués afin de choisir le modèle le plus robuste et stable possible. Le modèle qui utilise le percentile 95 comme variable indépendante a été choisi ($RMSE_{cv} = 0.79$) et appliqué à l'ensemble de la zone d'étude obtenant ainsi un maillage de placettes qui ont la hauteur dominante de prédiction comme donnée associée. Deux méthodologies de calcul de carte (maillage de placettes circulaires et carrées) ont été appliquées sous un système d'information géographique afin de cartographier l'ensemble de la zone d'étude. Les deux méthodologies de cartographie testées semblent optimales pour l'échelle de notre travail.

Mots clés : hauteur dominante, jeunes peuplements, LiDAR, cartographie.

ABSTRACT

VALIDATION IN THE ESTIMATION OF THE HEIGHT OF YOUNG FOREST STANDS FROM AERIAL LIDAR DATA.

We estimated the dominant height of beech regeneration stand relating LiDAR aerial data and field data. The number of sample plots used was 101. LiDAR pulses were modeled using different metrics (percentiles, maximum and minimum, local maximum heights, etc) as explanatory variables. The uncertainty in the GPS location of the sample plots and the effect of the variation in the LiDAR returns density were evaluated in order to choose the most robust and stable model. The model using the 95th percentile as the independent variable was selected ($RMSE_{cv} = 0.79$) and applied to the whole study area getting a mesh of plots with dominant height as the associated predictor variable. Two different mapping methods (since a circular plot mesh and since square plot mesh) are applied into a geographic information system to map the dominant height in the study area. The two methods evaluated are optimal for our work scale.

Key words: regeneration stand, dominant height, LiDAR, cartography

RESUMEN

VALIDACION EN LA ESTIMACION DE LA ALTURA DE POBLACIONES FORESTALES EN REGENERACION A PARTIR DE DATOS LDAR AEREOS.

Se ha estimado la altura dominante de una población de haya en regeneración relacionando la datos LiDAR aéreos y datos de campo. A partir de 101 parcelas de muestreo, una serie de métricas han sido extraídas de la nube de puntos LiDAR (percentiles de altura, valores máximos y mínimos, alturas de máxima local, etc). Estas han sido utilizadas como las variables explicativas de los modelos evaluados. La incertidumbre en la localización del GPS de las parcelas de muestreo y el efecto de la variación de la densidad de retornos LiDAR ha sido evaluada con la finalidad de elegir el modelo más robusto y estable. El modelo que utiliza el percentil 95 como variable independiente ha sido elegido ($RMSE_{cv} = 0.79$) y aplicado al área de estudio obteniendo de este modo una malla de parcelas con la altura dominante como variable de predicción asociada. Dos métodos de cartografía (a partir de una malla de parcelas circulares y otra malla cuadradas) han sido aplicados en un sistema de información geográfica para poder cartografiar la zona de estudio en su conjunto. Los dos métodos evaluados han resultado ser óptimos para la escala de trabajo de este estudio.

Palabras clave: altura dominante, masas en regeneración, LiDAR, cartografía