

# EVALUATION PAR VALIDATION INDEPENDANTE DES PREDICTIONS DES PARAMETRES FORESTIERS REALISEES A PARTIR DE DONNEES LIDAR AEROPORTE

Alain Munoz<sup>1</sup>, Jérôme Bock<sup>1</sup>, Jean-Matthieu Monnet<sup>2</sup>, Jean Pierre Renaud<sup>3</sup>, Anne Jolly<sup>3</sup>, Catherine Riond<sup>1</sup>

1 : ONF, Département Recherche Développement et Innovation, Chambéry (alain.munoz@onf.fr)

2 : Irstea, Centre de Grenoble, UR EMGR, Saint-Martin-d'Hères (jean-matthieu.monnet@irstea.fr)

3 : ONF, Département Recherche Développement et Innovation, Nancy (anne.jolly@onf.fr)

## Résumé

L'utilisation des données LiDAR aéroporté pour modéliser, caractériser et cartographier la ressource forestière est effective depuis une dizaine d'années dans les pays d'Amérique du Nord ou d'Europe du Nord. Néanmoins, en France, il existe peu d'exemples d'applications forestières du LiDAR sur de grandes surfaces, dans des contextes forestiers souvent plus complexes et hétérogènes que les forêts boréales et les grandes pinèdes du nord de l'Europe. Dans cette étude, la méthode dite « surfacique », qui consiste à calibrer des modèles empiriques de régression mettant en relation des paramètres dendrométriques mesurés sur le terrain avec des métriques dérivées des données LiDAR, est évaluée sur trois sites forestiers français. Les modèles de prédiction se basent sur les métriques usuelles dérivées du nuage de points LiDAR, ainsi que sur des métriques calculées à partir d'une individualisation préalable des houppiers des arbres dominants. Une validation indépendante, à l'échelle du peuplement, a permis d'évaluer la fiabilité des modèles de prédiction sur des surfaces de l'ordre d'un hectare, pertinentes pour le gestionnaire. Les résultats montrent qu'il est possible de caractériser la ressource forestière dans différents peuplements, allant des hêtraies ou pinèdes régulières de plaine aux sapinières-pessières hétérogènes de montagne. L'erreur des modèles (RMSE) estimée par validation croisée à l'échelle de la placette est de 14 à 22 % ( $R^2$  de 0,75 à 0,90) pour la surface terrière, de 7 à 8 % ( $R^2$  de 0,76 à 0,96) pour le diamètre dominant et de 17 à 32 % ( $R^2$  de 0,50 à 0,88) pour la densité de tiges. La validation indépendante à l'échelle du peuplement montre une bonne fiabilité des modèles quel que soit le paramètre forestier étudié, et tant qu'ils sont utilisés dans leur domaine de validité. Les écarts entre les moyennes des prédictions et les mesures de référence terrain sont de l'ordre de 10 à 15 % pour la surface terrière, de 6 à 8% pour le diamètre dominant et de 10 à 33% pour la densité de tiges. Ces résultats démontrent, sur des surfaces unitaires d'environ un hectare, que l'inventaire par LiDAR aérien peut atteindre une précision équivalente à celle des méthodes traditionnelles de terrain, y compris pour des forêts complexes.

**Mots-clés :** LiDAR, forêt, dendrométrie, surface terrière, diamètre dominant, densité de tiges, validation indépendante

## Abstract

*Airborne LiDAR data has been used operationally for more than ten years to model, characterize and map forest resource has been known in Northern countries of America and Europe. However, in France, where topography is more complex and forests more heterogeneous compared to the boreal and large pine forests of Northern Europe, only few examples of LiDAR applications over large areas can be found. In this study, an "area based" approach, that establishes empirical models relating forest measurements collected in the field to plot level LiDAR metrics, was used in three contrasted forest sites. Models were established using the usual metrics derived from the LiDAR point cloud, as well as metrics calculated from dominant tree crowns individualized by applying a CHM segmentation algorithm. At stand level, an independent model validation was used to evaluate predictions accuracy on areas of around one hectare, which are relevant to forest managers. Results showed that it is possible to estimate forest resource in contrasted stands, ranging from beech or pine even-aged plain stands, to mixed fir-spruce mountain stands. The cross-validation residual error (RMSE) at the plot level was 14 to 22% ( $R^2$  from 0.75 to 0.90) for basal area, 7 to 8% ( $R^2$  from 0.76 to 0.96) for dominant diameter, and 17 to 32% ( $R^2$  from 0.50 to 0.88) for stem density. Independent validation at the stand level showed a good reliability of the models for all studied parameters, as far as they were used within their validity domain. The differences between the averaged predictions and the field reference measurements were about 10 to 15% for basal area, 6 to 8% for dominant diameter and 10 to 33% for stem density. These results showed that, on area of around one hectare, airborne LiDAR inventories can reach an accuracy equivalent to that of traditional field methods, even in complex forests.*

**Keywords:** LiDAR, forest, forest mensuration, basal area, dominant diameter, stem density, independent validation

## 1. Introduction

Une gestion durable des forêts implique de savoir caractériser et spatialiser la ressource au sein d'un massif forestier. Ces informations permettent de définir des règles de gestion qui sont, en France, synthétisées dans des documents de programmation : les

aménagement pour les forêts publiques et les plans simples de gestion pour les forêts privées.

Les informations principales relevées sur le terrain pour rédiger ces documents cadres sont en général : la surface terrière, qui est un indicateur de la quantité de bois sur pied, la densité de tiges, et des indicateurs renseignant sur le stade de développement des peuplements comme par exemple le diamètre des

arbres dominants ou la surface terrière des gros bois (Dubourdieu, 1997).

Historiquement, les méthodes utilisées pour recueillir ces informations consistaient à réaliser des inventaires pied à pied où le diamètre à 1,30 m de hauteur est mesuré pour chaque arbre. Compte tenu du temps important nécessaire pour de telles opérations, les méthodes ont progressivement évolué vers la mesure de quelques parcelles « caractéristiques », puis vers des inventaires statistiques par échantillonnage de placettes (Duplat et Perrotte, 1981). Le temps d'inventaire est considérablement réduit avec de tels sondages statistiques. En revanche la précision dépend du nombre de placettes inventoriées, du coefficient de variation du paramètre étudié et de l'erreur de mesure elle-même. Pour des placettes de surface fixe, l'erreur de mesure de la surface terrière est de l'ordre de quelques pourcents. Dans le cas d'inventaire par placettes à angle fixe, l'erreur dépendra du type d'appareil utilisé et de l'observateur. A l'extrême, dans le cas d'inventaire par jauge d'angle, outil généralement utilisé par les gestionnaires, les erreurs ponctuelles d'estimation de la surface terrière peuvent atteindre 20 à 50% (Bock et al., 2014). Pour des inventaires pied à pied sur des surfaces de quelques hectares, Duplat et Perrotte (1981) estiment que 95% des résultats en surface terrière sont situés dans un intervalle de -15 à +10% autour de la valeur réelle. Quel que soit le type d'inventaire, il est très difficile de disposer de cartographie décrivant la variabilité de la ressource à l'intérieur d'une parcelle sauf à réaliser des sondages statistiques très denses (d'une densité supérieure à 1 placette/ha) qui sont très coûteux.

Depuis une vingtaine d'années, la recherche forestière s'est intéressée aux possibilités offertes par la technologie LiDAR (Light Detection And Ranging), pour apprécier de nombreux paramètres utiles pour le gestionnaire forestier (Nelson et al., 1988 ; Næsset, 1997, 2002, 2004 ; Magnussen et Boudewyn, 1998 ; Woods et al., 2008 ; Sherrill et al. 2008). Ces études portent le plus souvent sur des forêts d'Amérique du Nord ou d'Europe du Nord et les cas d'application sur des peuplements français, généralement plus complexes et hétérogènes, sont peu fréquents (Bock et al., 2009 ; Monnet et al., 2010).

Le principe de la modélisation de paramètres forestiers à partir de données LiDAR consiste généralement à calibrer des modèles à l'aide de placettes d'inventaire terrain, puis à prédire et cartographier ces paramètres sur l'ensemble de la zone d'intérêt. Il s'agit le plus souvent de modèles empiriques qui utilisent une approche surfacique (Naæsset, 2004). Les erreurs de prédiction pour la surface terrière varient de 10 à 30% (Naæsset, 2004 ; Heurich et Thoma, 2008 ; Woods et al., 2008 ; Magnussen et al., 2012 ; Treitz et al., 2012 ; Watt et al., 2013). Pour la densité de tiges, l'erreur de prédiction varie de 15 à 50% (Næsset, 2004 ; Heurich et Thoma, 2008 ; Woods et al., 2008 ; Treitz et al., 2012 ; Watt et al., 2013 ; Ediriweera et al., 2014). Quant au diamètre dominant, paramètre qui a été moins traité dans la littérature, l'erreur de prédiction varie de 10 à 17% (Heurich et Thoma, 2008 ; Ediriweera et al., 2014). Pour la surface terrière des gros bois, paramètre principalement utilisé par les gestionnaires français, il n'existe pas de référence. Les résultats cités ont été obtenus lors d'études portant sur des types forestiers comme les peuplements d'épicéas et de pins du Sud-

Est de la Norvège (Naæsset, 2004 ; Magnussen et al., 2012), les forêts des Grands Lacs en Ontario (Woods et al., 2008 ; Treitz et al., 2012), les peuplements naturels mélangés d'épicéas et de hêtres en Bavière (Heurich et Thoma, 2008), les jeunes plantations de douglas en Nouvelle Zélande (Watt et al., 2013) ou les forêts pluviales australiennes (Ediriweera et al., 2014).

Dans la plupart des cas, les erreurs de prédiction sont évaluées par validation croisée. L'évaluation à l'aide d'échantillon indépendant n'est que rarement mise en œuvre et en général reste à l'échelle de la placette (Næsset, 2004 ; Jensen et al., 2006 ; Gautam et al., 2013).

L'originalité de cette étude réside dans la mise en place d'une procédure de validation indépendante des modèles de prédiction de ressource forestière, à l'aide de grappes de placettes ou d'inventaires pied à pied représentant des surfaces élémentaires de 1 ha, dont la taille est pertinente du point de vue du gestionnaire forestier. D'autre part, les modèles de prédiction proposés combinent les métriques standards qui décrivent la distribution verticale du nuage de points LiDAR avec d'autres métriques issues d'une approche reposant sur l'individualisation de couronnes d'arbres (Hyyppä et al., 2012 ; André et al., 2015). Alors que la contribution des métriques de distribution est souvent difficile à interpréter, on peut s'attendre à ce que les métriques « arbre » apportent une information semi-empirique directement liée aux caractéristiques physiques des arbres dominants du peuplement.

Cette étude qui porte sur trois sites forestiers, permet de plus d'évaluer la robustesse de notre approche dans des peuplements forestiers variés, allant des hêtraies de plaine jusqu'aux sapinières-pessières de montagne.

## 2. Matériels et méthodes

### 2.1. Sites d'étude et données LiDAR disponibles

Trois forêts publiques, ont été choisies en raison de leur diversité de types de peuplement et de paramètres d'acquisition LiDAR (figure 1 et tableau 1).

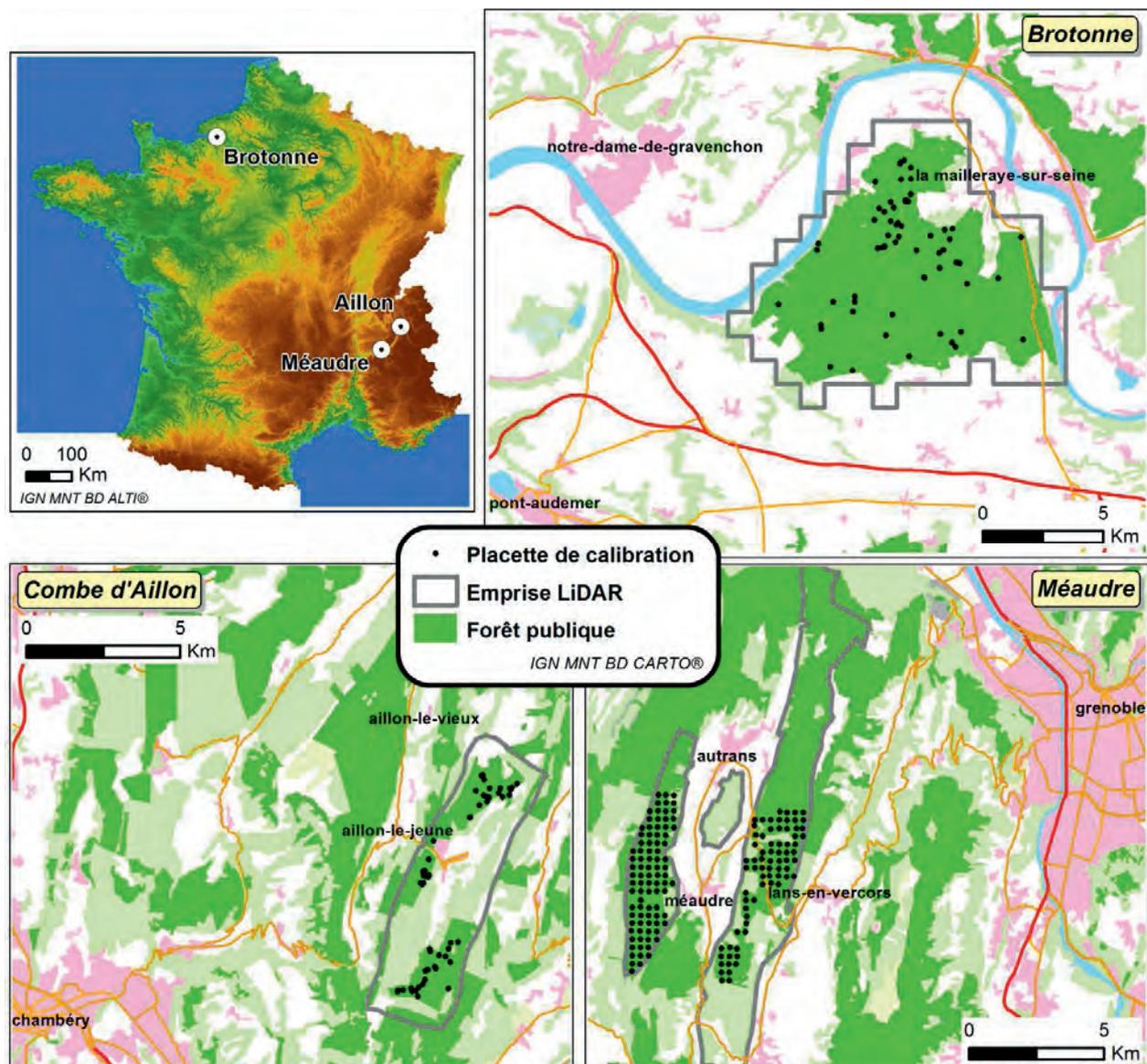
La forêt départementale de la Combe d'Aillon (620 ha) et la forêt communale de Méaudre (1800 ha) sont situées respectivement dans les Parcs Naturels Régionaux des Bauges, en Savoie (73) et du Vercors, en Isère (38). Il s'agit de deux forêts de montagne des Alpes du nord couvertes de peuplements irréguliers majoritairement composés de sapins (*Abies alba*), d'épicéas (*Picea abies*) et de hêtres (*Fagus sylvatica*). Les données LiDAR ont été acquises en juillet 2011 pour la Combe d'Aillon et en septembre 2010 et août 2011 pour Méaudre, avec une densité d'émission moyenne de 12 impulsions/m<sup>2</sup>.

La forêt domaniale de Brotonne (10600 ha), située dans le Parc Naturel Régional des Boucles de la Seine Normande (76), est une forêt régulière de plaine majoritairement composée de hêtres (*Fagus sylvatica*) et de pins sylvestres (*Pinus sylvestris*). Les données LiDAR ont été acquises en deux fois en avril 2010 et mars 2011 avec une densité d'émission moyenne de 4 impulsions/m<sup>2</sup>.

	Aillon (73)	Méaudre (38)	Brottonne (76)
<b>Caractéristiques forestières :</b>			
Surface	620 ha	1 800 ha	10 600 ha
Région	Montagne	Montagne	Plaine
Structure de peuplement	Irrégulière	Irrégulière	Régulière
Essences	Mélange sapins (48%), épicéas (15%) et hêtres (25%)	Mélange sapins (40%), épicéas (29%) et hêtres (22%)	Mono-essence hêtres (67%) ou pins sylvestres (14%)
<b>Caractéristiques LiDAR :</b>			
Origine des données ©	Conseil Général de Savoie	Irstea (projet Interreg Manfred)	GIP Seine Aval
Date(s) de(s) vol(s)	juillet 2011	sept. 2010 et août 2011	avril 2010 et mars 2011
Densité moyenne d'émission (cv *)	12 impulsions/m <sup>2</sup> (40%)	12 impulsions/m <sup>2</sup> (45%)	4 impulsions/m <sup>2</sup> (87%)

\* coefficient de variation (cv) = (Ecart-type / Moyenne) x 100

**Tableau 1 : Caractéristiques principales des sites d'étude.**



**Figure 1 : Localisation des sites d'étude. Situation générale (en haut à gauche) et zoom sur les sites de Brottonne (en haut à droite), de la Combe d'Aillon (en bas à gauche) et de Méaudre (en bas à droite).**

## 2.2. Données de référence pour la calibration des modèles

Des modèles de prédiction ont été calibrés à l'aide de placettes à surface fixe (tableau 2) où tous les arbres de plus de 17,5 cm de diamètre ont été inventoriés (diamètre, essence, état sanitaire) et positionnés (azimut et distance par rapport au centre de la placette). La position du centre des placettes a été relevée à l'aide d'un récepteur GNSS dont la précision, après correction différentielle, est estimée entre 1 et 5 m. Puis un recalage manuel a été effectué afin de faire correspondre le mieux possible la position des arbres relevés sur le terrain avec les couronnes des arbres visibles sur le modèle numérique de hauteur (MNH) LiDAR. Le recalage évite que le mauvais géoréférencement n'induisse des erreurs dans la modélisation (Monnet et Mermin, 2014).

Au total, on dispose de 54 placettes de 15 m de rayon sur la forêt de la Combe d'Aillon, mesurées entre avril 2011 et novembre 2012, soit un écart par rapport au vol LiDAR de moins d'une saison de végétation. Sur la forêt de Méaudre, 113 placettes de 17 m de rayon ont

été mesurées entre mai et octobre 2013, soit deux à trois saisons de végétation après la mission LiDAR. Enfin sur la forêt de Brotonne, 60 placettes de 15 m de rayon ont été mesurées durant l'été 2013, soit deux à trois saisons de végétation d'écart avec l'acquisition LiDAR. L'ensemble de ces placettes de calibration a été implanté dans l'emprise des forêts publiques en s'assurant qu'aucune exploitation forestière n'avait eu lieu entre la mission LiDAR et les mesures de terrain.

Quatre paramètres dendrométriques ont été étudiés. La surface terrière (G) correspond à la somme des surfaces des sections des arbres (mesurées à 1.30 m de hauteur) rapportée à la surface d'inventaire et exprimée en m<sup>2</sup>/ha. GGB est la surface terrière des gros bois, arbres de plus de 42,5 cm de diamètre (Courbaud et Gauquelin, 2006). Le diamètre dominant (D0) correspond à la moyenne des diamètres des 100 plus gros arbres dominants par hectare (Pardé et Bouchon, 1988). Enfin, la densité de tiges des arbres de plus de 17,5 cm de diamètre (N) correspond au nombre d'arbres rapporté à l'hectare (t/ha).

Site	Paramètre dendrométrique	DONNEES DE CALIBRATION				DONNEES DE VALIDATION (évaluation)			
		Echantillon	Min	Moy	Max	Surfaces	Min	Moy	Max
Aillon (forêt publique)	G (m <sup>2</sup> /ha)	54 placettes d'un rayon de 15 m	17,5	43,3	74,7	8 quadrats de 1 ha (100x100 m) inventoriés en plein	20,9	35,5	51,3
	GGB (m <sup>2</sup> /ha)		0,0	27,6	60,3		4,9	22,8	35,8
	D0 (cm)		34,9	57,3	78,2		41,4	53,0	59,7
	N (t/ha)		99	349	1118		244	356	640
Méaudre (forêt publique)	G (m <sup>2</sup> /ha)	113 placettes d'un rayon de 17 m	5,7	30,1	64,8	18 grappes (≈ 1 ha) de 4 placettes statistiques	17,1	30,1	41,9
	GGB (m <sup>2</sup> /ha)		0,0	13,5	39,1		5,0	14,6	25,9
	D0 (cm)		26,8	46,9	68,5		38,8	47,4	60,8
	N (t/ha)		99	334	584		165	319	506
Brotonne (forêt publique)	G (m <sup>2</sup> /ha)	60 placettes d'un rayon de 15 m	2,7	20,33	43,4	14 grappes (≈ 1 ha) de 4 à 5 placettes statistiques	17,6	26,3	34,8
	GGB (m <sup>2</sup> /ha)		0,0	9,6	38,2		15,2	20,7	27,8
	D0 (cm)		18,1	39,8	68,8		41,9	55,2	71,7
	N (t/ha)		85	226	523		92	157	258
Méaudre (forêt privée)	G (m <sup>2</sup> /ha)	/	/	/	/	14 grappes (≈ 1 ha) de 4 placettes statistiques	24,4	42,2	62,2
	GGB (m <sup>2</sup> /ha)		/	/	/		0,0	13,8	39,8
	D0 (cm)		/	/	/		32,6	45,2	61,4
	N (t/ha)		/	/	/		265	521	785

**Tableau 2 :** Caractéristiques dendrométriques des données de calibration et de validation (G : surface terrière des arbres de plus de 17,5 cm de diamètre, GGB : surface terrière des gros bois de plus de 42,5 cm de diamètre, D0 : diamètre dominant, N : densité de tiges des arbres de plus de 17,5 cm de diamètre).

## 2.3. Calcul des métriques LiDAR

Deux types de métriques LiDAR ont été utilisés.

Le premier type de métriques est extrait du nuage de points LiDAR normalisé. À l'altitude de chaque point est retranchée l'altitude du sol modélisée par un réseau de triangles irréguliers (TIN) réalisé à partir des points LiDAR classés « sol ». De ce nuage de points caractérisant la distribution de la végétation en 3D, on extrait des placettes circulaires correspondant exactement à l'emprise géographique des placettes de calibration terrain. Différentes familles de métriques ont été calculées sur chaque placette : des statistiques de distribution des hauteurs des points LiDAR (quantiles de 5 à 95%, moyenne, écart-type), des densités relatives et absolues de points, des taux de pénétration et des indicateurs spatiaux (voir en annexe des exemples de métriques) (Bock et al., 2009). Chacune de ces métriques est calculée pour tous les points

LiDAR ou pour certains types de retours (les premiers, les derniers retours, les retours compris dans des tranches de hauteurs spécifiques). Ces calculs ont été réalisés à l'aide de l'outil Lasmanager (© ONF) qui permet de traiter de grands volumes de données.

Le second type de métriques est calculé à partir de segments représentant *a priori* les couronnes des arbres (emprise des houppiers) et de maxima locaux correspondant à la position supposée de la cime des arbres identifiés sur le MNH (figure 2). La méthode est basée sur la segmentation par « bassin versant » (Hyypä et al., 1999 ; Persson et al., 2002 ; Koch et al., 2006). Les MNH en entrée, représentant la hauteur normalisée au-dessus du sol de la végétation, ont été générés au pas de 50 cm en suivant la procédure mise au point par Khosravipour et al. (2013) (figure 3) à l'aide de l'outil Lastools (rapidlasso GmbH, 2013). La méthode de segmentation consiste à prendre l'opposé

du MNH afin que les couronnes des arbres représentent des bassins dont les points les plus bas correspondent aux apex. Les emprises de couronnes sont déterminées en utilisant les outils hydrologiques présents dans l'extension « Spatial analyst » d'ArcGis v10.0 (© ESRI). Les apex sont identifiés par la localisation de la hauteur maximum dans chaque couronne. Au final, les traitements réalisés ont permis d'obtenir une couche de polygones représentant les emprises des couronnes. Chaque couronne est décrite par des informations de hauteur du maximum local, de surface de section, de volume et d'accumulation de flux (« flowaccumulation » d'ArcGis). Enfin, sur chaque placette on calcule des métriques correspondant à des statistiques de distribution de hauteurs des maxima, de surfaces ou de volumes des couronnes, en prenant en compte toutes les couronnes dont les apex sont inclus dans la placette ou un sous-ensemble constitué seulement de celles au-dessus de certains seuils de hauteur ou de surface (voir en annexe des exemples de métriques).

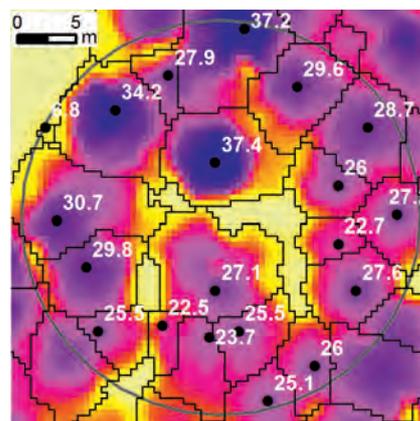


Figure 2 : Exemple de couronnes d'arbres (segments en noir) et de maxima locaux (points noirs) identifiés sur le MNH LiDAR sur une placette de 15 m de rayon (représenté par le cercle gris).

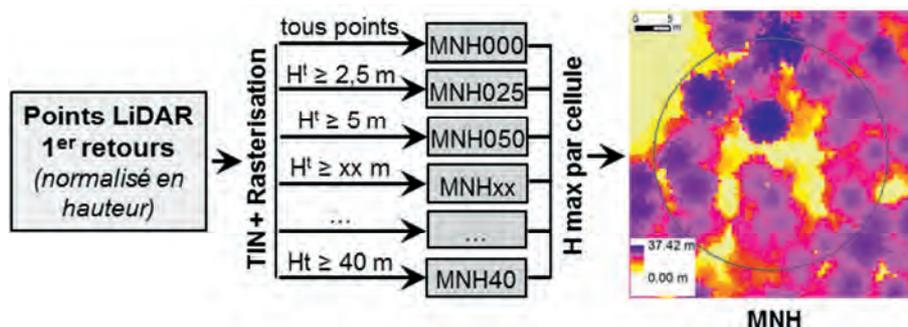


Figure 3 : Présentation de la méthode de création d'un MNH par la procédure mise au point par Khosravipour et al. (2013). Tout d'abord, un jeu de MNH est créé par triangulation et rasterisation de tous les premiers retours du nuage LiDAR puis en excluant les points au-dessous de certains seuils de hauteur (ex. 2,5 m, 5 m, 7,5 m ...). Le MNH final est obtenu par combinaison de ce jeu de données en conservant la hauteur maximale pour chaque cellule.

#### 2.4. Modélisation des paramètres dendrométriques

Les paramètres dendrométriques ont été modélisés pour chaque site à l'aide de régressions linéaires par moindres carrés car ces modèles sont simples à appliquer et leurs résultats facilement interprétables (White et al., 2013). Cette méthode, couramment utilisée pour prédire les paramètres forestiers (Næsset, 2004 ; Heurich et Thoma, 2008 ; Watt et al., 2013), met en relation un paramètre à expliquer représenté par une valeur mesurée sur le terrain, avec des variables explicatives correspondant aux métriques LiDAR calculées sur la même emprise que les placettes terrain (figure 4).

La sélection des modèles de prédiction s'inspire d'une méthode « pas-à-pas » en appliquant la procédure décrite ci-dessous, entièrement réalisée à l'aide du logiciel statistique R (R Core Team, 2014) :

1. Toutes les métriques LiDAR calculées sont testées dans un modèle à une variable afin de sélectionner celles qui ont un effet significatif ( $p$ -value  $\leq 0,05$ ).

2. Tous les modèles de deux à quatre variables sont construits par combinaison des métriques sélectionnées précédemment. Ceux non significatifs ou ne respectant pas les hypothèses de validité d'un modèle linéaire sont écartés. Ainsi la normalité des résidus a été contrôlée par un test de Shapiro, l'homoscédasticité des résidus a été analysée en

utilisant la librairie « gvlma » (Pena et Slate, 2014) et l'indépendance des variables explicatives a été acceptée lorsque le VIF (Variance Inflation Factor) calculé par la librairie « car » (Fox et Weisberg, 2011) était inférieur à 5.

3. Les RMSE de prédiction et le  $R^2$  ajusté (en tenant compte du nombre de variables dans le modèle) sont calculés pour tous les modèles par validation croisée basée sur trois groupes, à l'aide de la fonction « cv.lm » de la librairie « DAAG » (Mairdonald et Braun, 2014).

4. Sur chaque modèle, des analyses de variance complémentaires permettent de tester l'effet de certains paramètres externes : paramètres d'environnement (pente, altitude), du vol (densité de points LiDAR) ou des paramètres dendrométriques (densité de tiges et pourcentage de feuillus).

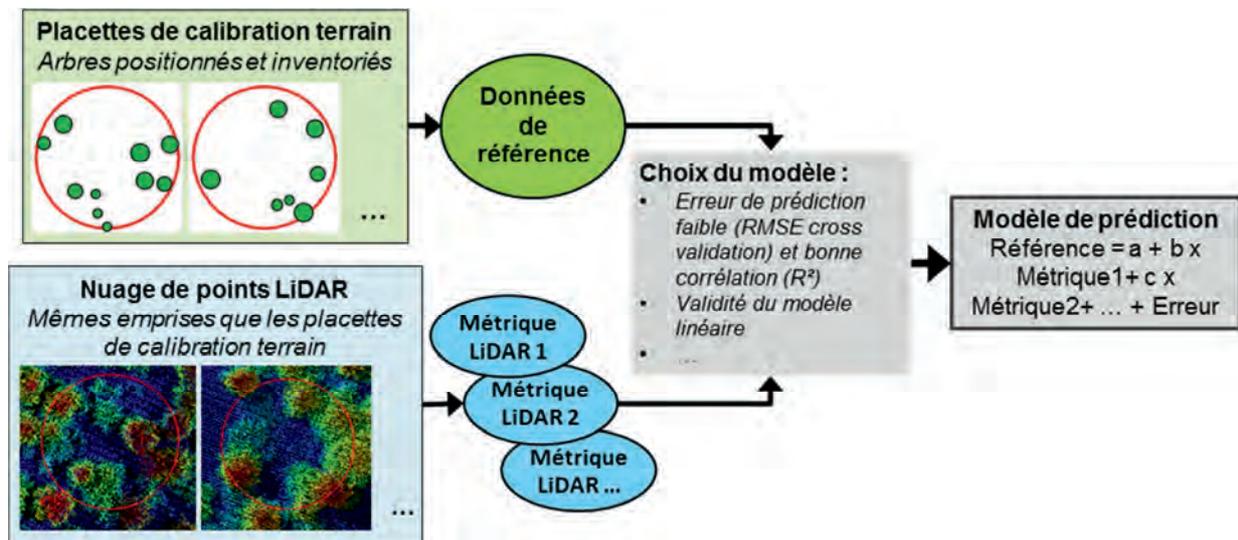
Au final, le modèle sélectionné est le résultat de l'optimisation de ces différents critères : RMSE de prédiction la plus faible,  $R^2$  ajusté le plus élevé, tests d'hypothèses des modèles linéaires satisfaisants, VIF inférieur à 5, effet limité de certains paramètres externes, diagnostic visuel permettant de s'assurer de la bonne corrélation linéaire entre chaque variable explicative et le paramètre à expliquer.

Contrairement à la plupart des études qui utilisent des procédures de sélection automatique de métriques (Næsset et al., 2004 ; Heurich et Thoma, 2008), notre démarche consiste à tester toutes les combinaisons de

variables (Monnet et al., 2010) et sélectionner des modèles où les variables sont non-corrélées (Allouis et al., 2013) et les résidus relativement indépendants des paramètres d'acquisition, forestiers ou topographiques. On espère ainsi sélectionner des modèles plus robustes, qui limitent les risques de comportement aberrant en prédiction.

Une fois les modèles sélectionnés, il est possible de les appliquer à l'ensemble de la zone d'intérêt. Les

métriques LiDAR sont calculées sur des emprises de même surface que les placettes de calibration terrain (par exemple un cercle de 15 m de rayon) et centrées sur les mailles d'une grille régulière (par exemple une grille de 20 x 20 m). Le modèle est appliqué afin d'obtenir une prédiction du paramètre forestier sur chaque point de la maille. Enfin, la rasterisation des prédictions ponctuelles permet d'obtenir une cartographie continue.



**Figure 4 :** Illustration de la modélisation des paramètres dendrométriques à l'aide d'un modèle linéaire mettant en relation la référence terrain mesurée sur des placettes de calibration en fonction de variables explicatives correspondant à des métriques LiDAR calculées sur la même emprise.

## 2.5. Données de validation et évaluation des modèles de prédiction

Deux types de données de référence, collectées sur le terrain indépendamment des placettes de calibration, ont été employés (tableau 2). Sur la forêt de la Combe d'Aillon, 8 quadrats de 1 ha (100x100 m) ont été inventoriés pied à pied. Toutes les tiges d'un diamètre supérieur ou égal à 17,5 cm ont été mesurées. Dans les forêts de Méaudre et Brotonne, un système d'inventaire statistique par grappes de 4 à 5 placettes circulaires non chevauchantes, et de 15 m de rayon, a été installé. Sur ces placettes, toutes les tiges d'un diamètre supérieur ou égal à 17,5 cm ont été mesurées. Puis l'estimation des paramètres dendrométriques (G, GGB, D0, N) a été effectuée par la moyenne des mesures de chaque grappe. Ces grappes de placettes sont équivalentes à un inventaire statistique sur peuplement d'environ 1 ha et permettent d'obtenir, avec une économie de moyens, des données de référence sur une surface équivalente aux inventaires en plein réalisés sur la forêt de la Combe d'Aillon.

L'implantation des surfaces de validation a été réalisée à partir des cartes de prédiction LiDAR afin de s'assurer d'un échantillonnage couvrant une large gamme de surfaces terrières. L'essentiel de ces surfaces de validation a été installé dans l'emprise des forêts publiques, c'est-à-dire dans le domaine de calibration des modèles de prédiction. De plus, sur la forêt de Méaudre, 14 grappes supplémentaires de placettes ont été implantées en forêt privée dans des peuplements plus capitalisés et plus régularisés. Ainsi il est possible d'évaluer, sur ce massif, la robustesse des modèles en dehors de leur domaine de calibration. Lors des

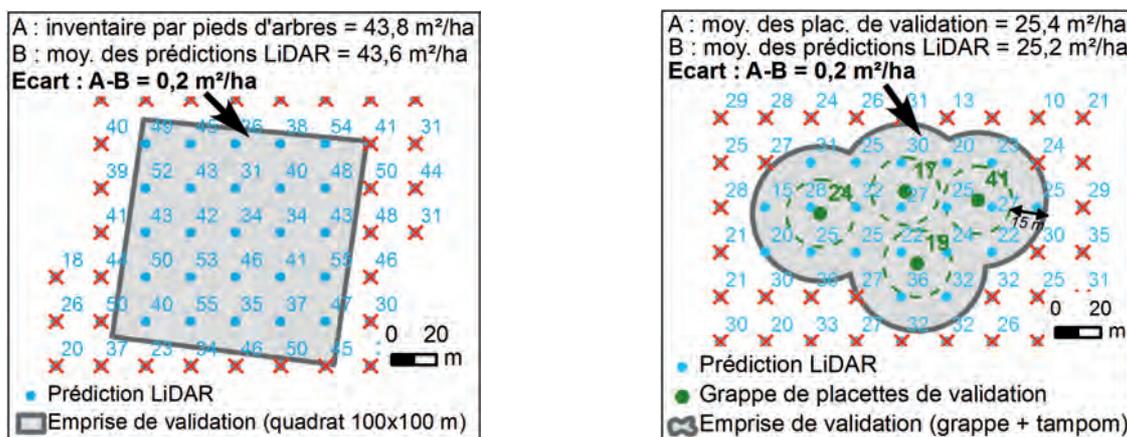
inventaires des surfaces de validation, il a été vérifié qu'aucune trace d'exploitation pouvant indiquer une coupe de bois entre la mission LiDAR et la mesure terrain n'était présente.

Sur la forêt de la Combe d'Aillon, les mesures ont été réalisées durant l'automne 2012, soit un écart de moins d'une saison par rapport aux placettes de calibration. Les 14 grappes de placettes de la forêt de Brotonne ont été inventoriées au cours du printemps 2014 représentant un écart d'une saison de végétation avec les mesures de calibration. Enfin, sur la forêt de Méaudre, l'effort d'inventaire des 32 grappes de placettes a été réparti sur les étés 2011, 2012 et 2013, correspondant à un décalage de zéro à deux saisons de croissance par rapport aux données de calibration.

L'évaluation de la qualité des prédictions a été réalisée par la comparaison entre les prédictions des modèles LiDAR et les données de référence. Pour cela, les modèles statistiques ont été appliqués sur des placettes d'un rayon identique à celui des placettes de calibration, soit de 15 m (forêts de Combe d'Aillon et de Brotonne) ou de 17 m (forêt de Méaudre), disposées sur un maillage systématique régulier de 20x20 m. La valeur estimée par LiDAR correspond à la moyenne des valeurs estimées sur les centres du maillage se trouvant à l'intérieur du polygone de validation d'environ 1 ha, constitué soit par les limites des quadrats, soit par l'enveloppe des grappes de validation incluant la zone tampon de 15 m (figure 5). Sur les surfaces de validation constituées par des grappes, l'application d'une zone tampon de 15 m a permis d'obtenir par une méthode simple une superficie proche de l'hectare mais qui n'est pas strictement identique d'une zone à l'autre.

Il a donc été nécessaire de s'assurer que ces variations n'avaient pas d'effet significatif sur les résultats. A l'échelle de la surface de 1 ha, la fiabilité des modèles est appréciée par l'analyse des écarts entre la référence terrain et la valeur estimée par LiDAR sur la même zone. La moyenne de ces écarts représente le biais et permet de dégager une tendance de

surestimation ou de sous-estimation. L'erreur quadratique moyenne (RMSE) et son coefficient de variation (CV) permettent d'apprécier l'erreur de prédiction. Le coefficient de détermination ( $R^2$ ) permet d'évaluer la part de variance expliquée par le modèle issu des métriques LiDAR.



**Figure 5 :** Exemples d'évaluation de la prédiction de G. A gauche, sur un quadrat de 1 ha (100x100 m) de la forêt de la Combe d'Aillon, la moyenne des estimations LiDAR (en bleu) se trouvant sur l'emprise de validation (en gris) est comparée aux résultats de l'inventaire pied à pied réalisé sur le quadrat. A droite, sur une grappe de placettes de la forêt de Brotonne, la moyenne des estimations LiDAR (en bleu) se trouvant sur l'emprise de validation (en gris) est comparée à la moyenne statistique de la grappe de placettes de validation (chiffres verts). Dans les deux cas, les estimations LiDAR en dehors de l'emprise de validation (croix rouges) sont exclues du calcul.

### 3. Résultats et discussion

#### 3.1. Modèles de prédiction

Les modèles retenus pour la prédiction de G ont des erreurs (RMSE de validation croisée) variant de 14% pour la forêt de Méaudre à 22% pour les résineux de la forêt de Brotonne. Ils permettent d'expliquer entre 75 et 90% de la variance (tableau 3). Pour GGB et N, les erreurs obtenues sont beaucoup plus importantes. Elles varient respectivement de 22% à 46% ( $R^2$  de 0,80 à 0,94) et de 17% à 32% ( $R^2$  de 0,50 à 0,88). Au contraire, l'estimation de D0 présente ici des erreurs faibles et constantes comprises entre 7 et 8% ( $R^2$  de 0,76 à 0,94) pour les trois sites étudiés.

Ces résultats sont en concordance avec ceux rencontrés dans la littérature. Pour la surface terrière, Magnussen et al. (2012) avaient obtenu une erreur d'environ 14%, mais en utilisant des modèles exponentiels, plus complexes que les régressions linéaires utilisées ici, et basés sur des métriques dérivés de MNH LiDAR. Naesset (2004) et Woods et al. (2008), avaient obtenu des erreurs variant de 16 à 23% à l'aide de modèles linéaires avec transformation de variable et calibrés par grands types de peuplements connus a priori. Pour la densité de tiges, Treitz et al. (2012) ont réussi à obtenir des erreurs plus faibles en pourcentage, d'environ 15%, mais sur des peuplements d'épicéas dont la densité moyenne est très élevée (1600 tiges/ha) soit une erreur d'environ 250 tiges/ha. Pour D0, les résultats de notre étude sont meilleurs que ceux rencontrés dans la littérature : Heurich et Thoma (2008) avaient obtenu des erreurs variant de 12 à 13% et Ediriweera et al. (2014) des erreurs de 9 à 17%.

Il est important de noter que sur la forêt de Brotonne, composée de peuplements mono essence de hêtre ou de pin sylvestre, la calibration de modèles spécifiques aux feuillus et aux résineux a permis d'améliorer la prédiction des paramètres G, GGB et N. Pour D0, cette stratification feuillus / résineux n'a pas apporté

d'amélioration de la précision. Heurich et Thoma (2008) avaient également constaté une réponse différente entre les feuillus et les résineux. La calibration de modèles spécifiques, sur des hêtraies et des pessières, leur avait permis d'améliorer les prédictions pour tous les paramètres dendrométriques, sauf D0. Cependant, la calibration de modèles spécifiques nécessite de connaître a priori la distribution des essences sur un massif afin d'appliquer ensuite le modèle adapté. Sur la forêt de Brotonne, une description d'aménagement fournit, pour chaque unité de gestion, l'essence dominante (hêtre ou pin). Il est donc possible d'y appliquer les modèles spécifiques pour cartographier l'ensemble des parcelles relativement pures. En revanche, sur les forêts de la Combe d'Aillon et de Méaudre, les peuplements sont intimement mélangés en feuillus et résineux. L'approche de modélisation stratifiée ne peut donc pas être mise en œuvre. Cependant, le taux de mélange n'a pas d'effet significatif sur les modèles retenus sur ces deux sites.

Les écarts temporels entre l'acquisition LiDAR et les mesures de calibration, et la croissance des peuplements qui en découle, sont pris en compte lors de la modélisation. En effet, le paramétrage des modèles est établi à partir des placettes de calibration, permettant ainsi de prédire les données forestières à la date de calibration. Malgré tout, ces écarts sont source d'erreur car l'accroissement des peuplements n'est pas homogène. Il dépend notamment de la fertilité de la station et de la maturité des peuplements qui peuvent varier d'une placette à l'autre. D'autre part, sur les sites de Méaudre et Brotonne, l'acquisition LiDAR a été réalisée en deux années successives entraînant une variation de cet écart temporel au sein du même site. Il aurait donc été préférable de réaliser une modélisation par année de vol, mais le type d'échantillonnage des placettes n'a pas permis de mettre en œuvre cette approche.

Paramètre dendro. étudié	Site	Erreurs des modèles de prédiction à la placette			Evaluation indépendante des modèles (≈ 1 ha)		
		Métriques LiDAR utilisés <sup>(1)</sup>	R <sup>2</sup> aj. <sup>(2)</sup>	RMSE prédiction (CV) <sup>(3)</sup>	R <sup>2</sup>	Biais <sup>(4)</sup>	RMSE (CV) <sup>(5)</sup>
G (m <sup>2</sup> /ha)	Aillon (publique)	Arbre.Som.H.Apex2_s15.fa40 + Arbre.SD.Vol.Emp.Cour_h35.fa40 + Arbre.Som.Surf.Cour_h20.fa40	0,78	7,2 (16%)	0,80	-0,2 ns	4 (10%)
	Méaudre (publique)	Arbre.Som.H.Apex_h12.5.fa40 + Arbre.Som.Surf.Cour/√N + Arbre.Quant.0.Vol.Emp.Cour._s2.5.fa15 + Moy.Altitude	0,76	4,3 (14%)	0,62	-1,8 ns	3,9 (13%)
	Brottonne (publique)	<b>Pour les feuillus :</b> Arbre.Som.Vol.Emp.Cour_h20 + Arbre.Som.H.Apex_h12.5	0,90	2,9 (15%)	0,50	0,6 ns	3,4 (14%)
		<b>Pour les résineux :</b> Arbre.Som.Surf.Cour_h13.5.fa20 + Arbre.Moy.Vol.Cour2_s20.s30	0,75	4,9 (22%)	0,79	3,1 ns	4,5 (15%)
Méaudre (privée)	Arbre.Som.H.Apex_h12.5.fa40 + Arbre.Som.Surf.Cour/√N + Arbre.Quant.0.Vol.Emp.Cour._s2.5.fa15 + Moy.Altitude	/	/	0,67	8,0 ***	10,3 (24%)	
GGB (m <sup>2</sup> /ha)	Aillon (publique)	Arbre.Som.H.Apex2_s40 + Arbre.Som.Vol.Emp.Cour_h30	0,83	6,1 (22%)	0,86	-0,6 ns	4,3 (19%)
	Méaudre (publique)	Arbre.Som.Vol.Cour_vol700 + Arbre.Quant.60.Surf.Cour_s2.5 + Arbre.Quant.95.Vol.Cour2_s30.s50 + Moy.Altitude	0,80	4,1 (30%)	0,68	-0,9 ns	3,8 (26%)
	Brottonne (publique)	<b>Pour les feuillus :</b> Arbre.Som.H.Apex/√N_h25 + Arbre.Som.H.Apex/√N_h20 + Arbre.Moy.Vol.Cour	0,94	2,8 (32%)	0,73	1,1 ns	3,5 (17%)
		<b>Pour les résineux :</b> Arbre.Quant.0.Vol.Cour2_s25 + Arbre.SD.H.Apex_h13.5 + Arbre.Som.H.Apex/√N_h25.fa40	0,82	4,7 (46%)	0,46	-0,8 ns	4,2 (20%)
Méaudre (privée)	Arbre.Som.Vol.Cour_vol700 + Arbre.Quant.60.Surf.Cour_s2.5 + Arbre.Quant.95.Vol.Cour2_s30.s50 + Moy.Altitude	/	/	0,86	1,5 ns	5,0 (36%)	
D0 (cm)	Aillon (publique)	Arbre.Quant.70.H.Apex_vol500.fa40+ Arbre.Som.H.Apex_s50 + Arbre.Moy.Surf.Cour_fa40	0,84	4,2 (7%)	0,76	-1,0 ns	3,7 (7%)
	Méaudre (publique)	Arbre.Moy.Quad.Vol.Cour_h15.fa15 + Nuage.Dens.Rel_6m.12m + Arbre.Quant.60.Surf.Cour_fa20	0,76	3,7 (8%)	0,60	-1,2 ns	3,7 (8%)
	Brottonne (publique)	Nuage.Quant.80.H + Nuage.Quant.05.Hfst.1m + Nuage.Quant.60.Hlst.2m	0,94	3,4 (8%)	0,83	-1,1 ns	3,4 (6%)
	Méaudre (privée)	Arbre.Moy.Quad.Vol.Cour_h15.fa15 + Nuage.Dens.Rel_6m.12m + Arbre.Quant.60.Surf.Cour_fa20	/	/	0,74	-1,3 ns	3,8 (8%)
N (t/ha)	Aillon (publique)	Arbre.Som.H.Apex_s5 + Nuage.Vol_12m.22m + Arbre.Som.Vol.Cour2_s20.s30	0,88	74 (21%)	0,96	6 ns	35 (10%)
	Méaudre (publique)	Arbre.Nb.Apex_s10.fa40 + Arbre.Quant.50.Vol.Cour_h13.5.fa20 + Nuage.Dens.Rel_t6.lst	0,68	58 (17%)	0,76	-14 ns	43 (13%)
	Brottonne (publique)	<b>Pour les feuillus :</b> Arbre.Som.H.Apex/√N_h25 + Arbre.Nb.Apex_h12.5	0,66	55 (26%)	0,01	15 ns	45 (33%)
		<b>Pour les résineux :</b> Arbre.Moy.H2.Apex_vol600 + Arbre.Nb.Apex_vol100.fa15	0,50	79 (32%)	0,63	29 ns	54 (26%)
Méaudre (privée)	Arbre.Nb.Apex_s10.fa40 + Arbre.Quant.50.Vol.Cour_h13.5.fa20 + Nuage.Dens.Rel_t6.lst	/	/	0,67	121 ***	145 (28%)	

1 : la définition des métriques LiDAR retenues pour les modèles est présentée en annexe ; 2 : R<sup>2</sup> ajusté du modèle de prédiction ; 3 : RMSE et coefficient de variation de la RMSE obtenus par validation croisée ; 4 : moyenne des écarts observés entre les prédictions LiDAR et les références de terrain (test t de Student : ns = non significatif (p.value > 0,05), \*\*\* = très fortement significatif (p.value ≤ 0,001)) ; 5 : RMSE calculée sur les écarts entre les prédictions LiDAR et les références de terrain.

**Tableau 3 : Synthèse des résultats des modèles de prédiction utilisés dans cette étude.**

Toutefois, sur ces deux forêts, la date de vol LiDAR n'a pas d'effet significatif sur la modélisation des paramètres dendrométriques.

Sur l'ensemble des modèles retenus, les métriques « arbres », issues de la segmentation de couronnes, sont prépondérantes (80% des métriques utilisées). Comme pour Hyypä et al. (2012) ou André et al. (2015), leur utilisation a permis d'améliorer la prédiction des paramètres forestiers par rapport à l'utilisation des seules métriques « nuage » calculées à partir des points LiDAR. Même sur la forêt de Brotonne, où la densité d'émission LiDAR est relativement faible, les métriques « arbres » se sont révélées plus performantes pour la prédiction de G, GGB et N.

### 3.2. Validation indépendante

Pour l'ensemble des paramètres dendrométriques étudiés, on s'intéresse à l'erreur relative, correspondant au coefficient de variation des écarts entre la valeur prédite et la référence de terrain.

A l'échelle du peuplement d'un hectare, dans l'emprise des forêts publiques, c'est-à-dire dans le domaine de calibration des modèles de prédiction, les erreurs relatives d'estimation de G sont de l'ordre de 10 à 15% (tableau 3). Elles sont comparables aux erreurs de mesures constatées lors d'inventaires pied à pied de parcelles forestières de plusieurs hectares (Duplat et Perrotte, 1981). L'estimation de GGB présente une erreur relative de 17 à 26%, ce qui peut paraître élevé mais reste néanmoins équivalent, voire meilleur, que l'erreur constatée sur la plupart des inventaires d'aménagement réalisés à l'aide du relascope à encoche (Bock et al., 2014). D0 est estimé avec une erreur relative faible, inférieure à 8%, qui correspond aux attentes globales des gestionnaires des sites étudiés. Dans les sites de montagne, N est estimé à l'échelle du peuplement avec une erreur relative assez faible, de 10 à 13%. En revanche, sur la forêt de Brotonne, les erreurs relatives sont fortes, de 26 à 33%, et sur les peuplements feuillus il n'y a pas de corrélation entre les prédictions et les mesures de référence ( $R^2 = 0.01$ ). Ces mauvais résultats, aussi bien à l'échelle de la placette que du peuplement, s'expliquent en partie par une acquisition LiDAR à plus faible densité sur Brotonne (4 pt/m<sup>2</sup>), en comparaison des sites de montagne (12 pt/m<sup>2</sup>). En effet, Treitz et al. (2012) avaient constaté une augmentation de l'erreur de prédiction de N, sur certains peuplements, avec la diminution de la densité de points LiDAR. De plus, il faut souligner la forte hétérogénéité de la couverture LiDAR (CV de 87% - tableau 1) faisant varier la densité d'émission de plus de 6 pt/m<sup>2</sup>, lorsque les bandes de vol se recouvrent, à moins de 2 pt/m<sup>2</sup> en dehors des recouvrements. Or, en dessous de 4 pt/m<sup>2</sup>, les approches « arbre » par détection d'apex ou de couronne et les calculs de métriques qui en découlent sont peu robustes (White et al., 2013).

La validation indépendante à l'échelle du peuplement d'environ 1 ha, dans le domaine de calibration, montre une bonne fiabilité des modèles de prédiction et une absence de biais significatif. Ces bons résultats sont tout de même à relativiser pour plusieurs raisons. Le faible échantillon, seulement 8 à 18 surfaces de validation par site, peut limiter la portée des résultats statistiques. D'autre part, les erreurs de mesure souvent constatées lors d'inventaires pied à pied ou statistiques (Duplat et Perrotte, 1981) rendent la référence de terrain discutable. Mais les oublis et les erreurs de mesures sur de petites surfaces sont

considérés dans cette étude comme négligeables. Enfin, les écarts temporels entre les mesures de calibration et de validation peuvent expliquer une partie de la variabilité des prédictions. Malgré tout, ces écarts sont faibles (de une à parfois deux saisons de végétation sur quelques grappes de validation de la forêt de Méaudre) et n'entraînent donc qu'un faible accroissement des peuplements. Par exemple, pour G, les accroissements annuels moyens sont estimés entre 0,6 et 0,9 m<sup>2</sup>/ha/an sur les sites d'étude d'après les données de l'Institut National de l'Information Géographique et Forestière (source Inventaire forestier - Tableaux personnalisés - Résultats sur la période 1988-2004).

Dans les forêts privées de Méaudre, constituées de peuplements régularisés plus capitalisés et plus denses que ceux de la forêt publique ayant servi à la calibration des modèles (tableau 2), les erreurs relatives d'estimation sont plus importantes pour G, GGB et N. Elles augmentent respectivement de 13 à 24%, 26 à 36% et 13 à 28% entre les prédictions en forêt publique et en forêt privée. En revanche, la prédiction de D0 reste fiable avec 8% d'erreur relative.

Le manque de généralité des modèles est la principale critique qui peut être faite à l'ensemble des études utilisant une approche de modélisation empirique. Même dans des contextes à première vue similaires, c'est-à-dire des peuplements mélangés de sapins, épicéas et hêtres, de la Combe d'Aillon et de Méaudre, avec des conditions d'acquisition LiDAR équivalentes (12 pt/m<sup>2</sup>), on constate que les métriques retenues et les coefficients associés sont différents. Au-delà de la composition en essences, la structure ou le type de peuplement sont également importants à prendre en compte dans la calibration des modèles (Naesset, 2004 ; Woods et al., 2008 ; Treitz et al., 2012). Dans notre étude, la stratification par type de peuplement n'a pas été utilisée car les aménagements indiquaient des structures majoritairement irrégulières sur les forêts de la Combe d'Aillon et de Méaudre. Sur les peuplements de la forêt de Brotonne, majoritairement réguliers, la prise en compte des stades de développement ou de l'âge aurait peut-être permis d'améliorer les modèles. Mais ces informations ne sont pas disponibles *a priori* pour une stratification. La modélisation empirique de paramètres dendrométriques à partir de données LiDAR implique donc une calibration pour chaque vol, chaque site, voire chaque grand type de peuplements caractérisable au préalable par des données externes (données issues de l'aménagement forestier, de photo-interprétation, voire du nuage de points LiDAR lui-même).

Ces résultats indiquent que l'inventaire par LiDAR aérien peut atteindre une précision équivalente à celle des inventaires classiques de terrain, à condition d'utiliser les modèles dans leur domaine de calibration.

## 4. Conclusion

La méthode mise au point dans le cadre de cet article a permis de construire des modèles « fiables », quand ils sont utilisés au sein du domaine de validité d'un site donné, pour caractériser la ressource forestière à l'échelle de la placette ou du peuplement. Cette approche consiste à tester toutes les combinaisons de métriques et sélectionner des modèles où les métriques sont non-corrélées et relativement indépendantes des paramètres externes du site tels que la topographie, la

densité d'émission LiDAR ou le mélange d'essences. L'originalité de la méthode réside également dans l'utilisation d'une approche mixte utilisant des métriques LiDAR, à l'échelle de la placette, calculées à partir du nuage de points ou dérivées d'un ensemble de couronnes d'arbres individualisées dans le MNH. Or, les métriques LiDAR issues des couronnes d'arbres améliorent les prédictions à l'échelle placette.

La « fiabilité » des modèles a également pu être évaluée à différentes échelles grâce au principe de calibration à partir de placettes circulaires et de validation indépendante sur des surfaces plus grandes d'environ un hectare. A l'échelle placette, les erreurs de prédiction de G et GGB, bien qu'assez élevées, restent néanmoins équivalentes, voire plus faibles, que celles constatées sur la plupart des inventaires d'aménagement réalisés à l'aide de relascope à encoche (Bock et al., 2014). A l'échelle du peuplement d'un hectare, les erreurs de prédiction de G, D0 et N, notamment sur les sites de montagne, sont comparables à celles des inventaires pied à pied (Duplat et Perrotte, 1981), c'est-à-dire autour de 10 à 20% d'erreurs. Une analyse complémentaire sur des surfaces de validation plus nombreuses et plus grandes, à l'échelle d'unités de gestion de plusieurs hectares, confirme cette tendance dans le cas particulier d'une forêt traitée en futaie jardinée (Monnet et Munoz, 2015).

L'avantage de l'inventaire par la méthode « LiDAR », par rapport aux méthodes classiques d'inventaire pied à pied ou de sondages statistiques, réside dans la possibilité de cartographier les paramètres forestiers. Cependant, compte tenu des erreurs parfois importantes constatées ponctuellement, notamment pour la prédiction de GGB ou N, il est recommandé d'utiliser les modèles pour des estimations dendrométriques sur des surfaces unitaires plus grandes, à l'échelle du peuplement ou de zones homogènes, et situées dans le domaine de calibration. Enfin, on constate, comme dans la plupart des études utilisant des modèles empiriques, que le point faible de cette méthode réside dans l'absence de généralité des modèles de prédiction. Cela implique la nécessité de calibrer les modèles pour chaque vol LiDAR et, en théorie, pour chaque grand type de peuplement. Cette stratégie est compatible avec les systèmes d'inventaire et de suivi d'aménagement par placettes permanentes préconisés en France dans les forêts de référence pour le domaine public (ONF, 2013). Il est donc possible de valoriser ces placettes permanentes, mesurées dans le cadre de la révision d'aménagement, pour la calibration de modèles LiDAR et obtenir ainsi, sans surcoût d'inventaire, une cartographie en continu et des estimations à l'échelle de la parcelle ou du peuplement. Les recherches se poursuivent également pour établir des modèles moins empiriques, plus génériques, comme cela a été fait pour la hauteur dominante (Bock et al., 2009) ou le volume (Hollaus et al., 2007 ; Bouvier et al., 2014).

### Remerciements

Ce travail a bénéficié d'un soutien financier de l'Agence Nationale de la Recherche (ANR) dans le cadre du projet FORESEE « Caractérisation de la ressource forestière pour les bioénergies » (ANR-10-BIOE-08-07). Les auteurs remercient aussi les partenaires ayant mis à disposition des données : le Conseil Général de

Savoie pour les données LiDAR de la Combe d'Aillon, le Groupement d'Intérêt Public Seine-Aval (GIPSA) pour les données LiDAR de la forêt de Brotonne ainsi que l'Irstea de Grenoble pour les données LiDAR de Méaudre (acquises dans le cadre du projet Interreg Manfred) et certaines placettes terrain utilisées lors des validations indépendantes.

Le Département recherche, développement et innovation de l'ONF bénéficie du soutien de l'Agence Nationale de la Recherche au titre du programme Investissements d'avenir portant la référence n° ANR-11-LABX-0002-01 (Laboratoire d'Excellence ARBRE).

### Références

- Allouis T., Durrieu S., Véga C., Couteron P., 2013. *Stem volume and above-ground biomass estimation of individual pine trees from LiDAR data: Contribution of full-waveform signals*. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 6(2), 924-934.
- André A.C., Renaud J.-P., Véga C., Munoz A., Bock J., St-André L., 2015. *Apport de variables issues de la segmentation d'arbres sur données LiDAR aéroportée pour l'estimation des variables dendrométriques de placettes*. Revue Française de Photogrammétrie et de Télédétection, ce numéro.
- Bock J., Dambrine E., Dez G., Dupouey J. L., Georges-Leroy M., Jolly A., Renaud J. P., 2009. *Towards site index mapping in deciduous stands using multi-echo LiDAR data*. IUFRO Division 4 Meeting A "Extending Forest Inventory and Monitoring over Space and Time", Quebec, Canada.
- Bock J., Munoz A., Fay J., 2014. *Précision des mesures en forêt*. Atelier Z-Forest, Paris, France - <http://foresee.fcba.fr/restitutions/> (dernière consultation le 16/01/2015).
- Bouvier M., Durrieu S., Fournier R. A., Renaud J. P., 2015. *Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data*. Remote Sensing of Environment, 156, 322-334.
- Courbaud B., Gauquelin X., 2006. *Guide des sylvicultures de montagne : Alpes du Nord Françaises*. Cemagref de Grenoble, 289p.
- Dubourdieu J., 1997. *Manuel d'aménagement forestier (gestion durable et intégrée des écosystèmes forestiers)*. Office National des Forêts, 244 p.
- Duplat P., Perrotte G., 1981. *Inventaire et estimation de l'accroissement des peuplements forestiers*. Section technique de l'Office national des Forêts - 432 p.
- Ediriweera S., Pathirana S., Danaher T., Nichols D., 2014. *Lidar remote sensing of structural properties of subtropical rainforest and eucalypt forest in complex terrain in north-eastern Australia*. Journal of Tropical Forest Science, 26(3):397 – 408.
- Fox J., Weisberg S., 2011. *An {R} Companion to Applied Regression, Second Edition*. Thousand Oaks CA: Sage. <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion> (dernière consultation le 05/05/2015).
- Gautam B., Peuhkurinen J., Kauranne T., Gunia K., Tegel K., Latva-Käyrä P., Rana P., Eivazi A., Kolesnikov A., Hämäläinen J., ... Joshi C., 2013.

- Estimation of Forest Carbon Using LiDAR-Assisted Multi-Source Programme (LAMP) in Nepal.* Proceedings of the International Conference on Advanced Geospatial Technologies for Sustainable Environment and Culture, Pokhara, Nepal, 12-13.
- 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.
- Hollaus M., Wagner W., Schadauer K., Maier B., Gabler K., 2009. *Growing stock estimation for alpine forests in Austria: a robust lidar-based approach.* Canadian Journal of Forest Research, 39(7): 1387-1400.
- Hyypää J.; Inkinen M., 1999. *Detecting and estimating attributes for single trees using laser scanner.* Photogramm. J. Finl., 16: 27–42.
- Hyypää J., Yu X., Hyypää H., Vastaranta M., Holopainen M., Kukko A., Kaartinen H., Jaakkola A., Vaaja M., Koskinen J., Alho P., 2012. *Advances in forest inventory using airborne laser scanning.* Remote Sensing, 4(5): 1190-1207.
- Jensen J. L. R., Humes K. S. Conner, T., Williams C. J., & DeGroot, J., 2006. *Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar.* Canadian Journal of Forest Research, 36(5): 1129-1138.
- Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., Hussin, Y.A., 2013. *Development of an algorithm to generate a Lidar pit-free canopy height model.* SilviLaser, 30:125 – 128.
- Koch B., Heyder U., Weinacker H., 2006. *Detection of individual tree crowns in airborne lidar data.* Photogrammetric Engineering & Remote Sensing, 72 : 357–363.
- Kodani E., Awaya Y., 2008. *Estimating mean height and stand volume in broad leaved forest stands using LiDAR.* Journal of Forest Planning, 13:239-244.
- Magnussen S., Boudewyn P., 2008. *Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators.* Canadian Journal of Forest Research, 28(7):1016-1031.
- Magnussen S., Næsset E., Gobakken T., Frazer G., 2012. *A fine-scale model for area-based predictions of tree-size-related attributes derived from LiDAR canopy heights.* Scandinavia. Journal of Forest Research, 27(3):312-322.
- Maindonald J. H., Braun W. J., 2014. *DAAG: Data Analysis And Graphics data and functions.* R package version 1.20. <http://CRAN.R-project.org/package=DAAG> (last data accessed: 05 May 2015).
- Monnet J. M., Mermin E., Chanussot J., Berger F., 2010. *Estimation of forestry parameters in mountainous coppice stands using airborne laser scanning.* Proceedings of Silvilaser 2010.
- Monnet J. M., Mermin E., 2014. *Cross-Correlation of Diameter Measures for the Co-Registration of Forest Inventory Plots with Airborne Laser Scanning Data.* Forests, 5(9): 2307-2326.
- Monnet J.M., Munoz A., 2015. *Comparaison de méthodes de spatialisations pour l'agrégation par parcelle des estimations de paramètres forestiers par LiDAR aéroporté.* Revue Française de Photogrammétrie et de Télédétection, ce numéro.
- Naesset E., 1997. *Estimating timber volume of forest stands using airborne laser scanner data.* Remote Sensing of Environment, 61(2) : 246-253.
- Næsset E., 2002. *Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data.* Remote Sensing of Environment, 80(1) : 88-99.
- Næsset E., 2004. *Practical large-scale forest stand inventory using a small-footprint airborne scanning laser.* Scandinavian Journal of Forest Research, 19:164–179.
- Nelson R., Krabill W., Tonelli J., 1988. *Estimating forest biomass and volume using airborne laser data.* Remote sensing of environment, 24(2) : 247-267.
- ONF, 2013. Note de service NDS-13-D-365 du 10 avril 2013 (*Elaboration des aménagements forestiers - Inventaire dendrométriques*), 5p.
- Pardé J., Bouchon J., 1988. *Dendrométrie.* 2ème édition. ENGREF Nancy – 328p.
- Pena E. A., Slate E. H., 2014. *gvlma : Global Validation of Linear Models Assumptions.* R package version 1.0.0.2. <http://CRAN.R-project.org/package=gvlma> (dernière consultation le 05/05/2015).
- Persson A., Holmgren J., Söderman, U., 2002. *Detecting and measuring individual trees using an airborne laser scanner.* Photogrammetric Engineering and Remote Sensing, 68(9) : 925-932.
- R Core Team, 2014. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria - <http://www.R-project.org> (dernière consultation le 09/02/2015).
- Rapidlasso GmbH, 2013. *LASTools Rapid LiDAR Processing.* URL: <http://www.rapidlasso.com>, Gilching, Germany (dernière consultation le 09/02/2015).
- Sherrill K. R., Lefsky M. A., Bradford J. B., Ryan M. G., 2008. *Forest structure estimation and pattern exploration from discrete return lidar in subalpine forests of the Central Rockies.* Canadian Journal of Forest Research, 38(8):2081-2096.
- Treitz P., Lim K., Woods M., Pitt D., Nesbitt D., Etheridge D., 2012. *LiDAR sampling density for forest resource inventories in Ontario, Canada.* Remote Sensing 4(4):830–848.
- Watt M. S., Meredith A., Watt P., Gunn A., 2013. *Use of LiDAR to estimate stand characteristics for thinning operations in young Douglas-fir plantations.* New Zealand Journal of Forestry Science, 43:18.
- White J. C., Wulder M. A., Varhola A., Vastaranta, M. Coops, N. C. Coops, B. D. Cook, D. Pitt, M. Woods, 2013. *A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach.* Natural Resources Canada – 39p
- Woods M., Lim K., Treitz P., 2008. *Predicting forest stand variables from LiDAR data in the Great Lakes – St. Lawrence forest of Ontario.* The Forestry Chronicle, 84(6):827-839.

## Annexes

### Définitions des métriques LiDAR retenues dans les modèles de prédiction :

Moy.Altitude : Altitude moyenne de la placette.

Arbre.Moy.H2.Apex\_vol600 : Moyenne des hauteurs d'apex au carré. Seul les apex dont le volume de couronne  $\geq 600 \text{ m}^3$  sont pris en compte.

Arbre.Moy.Quad.Vol.Cour\_h15.fa15 : Moyenne quadratique des volumes de couronne (volume du MNH calculé sur l'emprise de couronne). Seuls les apex dont la hauteur  $\geq 15 \text{ m}$  et le flowaccumulation  $\geq 15$  sont pris en compte.

Arbre.Moy.Surf.Cour\_fa40 : Moyenne des surfaces de couronne. Seul les apex dont le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.Moy.Vol.Cour : Moyenne des volumes de couronne (volume du MNH calculé sur l'emprise de couronne).

Arbre.Moy.Vol.Cour2\_s20.s30 : Moyenne des volumes de couronne (volume du MNH calculé sur l'emprise de couronne) au carré. Seuls les apex dont la surface de couronne est comprise entre 20 et 30  $\text{m}^2$  sont pris en compte.

Arbre.Nb.Apex\_h12.5 : Nombre d'apex. Seuls les apex dont la hauteur  $\geq 12,5 \text{ m}$  sont pris en compte.

Arbre.Nb.Apex\_s10.fa40 : Nombre d'apex. Seuls les apex dont la surface de couronne  $\geq 10 \text{ m}^2$  sont pris en compte.

Arbre.Nb.Apex\_vol100.fa15 : Nombre d'apex. Seuls les apex dont le volume de couronne  $\geq 100 \text{ m}^3$  sont pris en compte.

Arbre.Quant.0.Vol.Emp.Cour\_s2.5.fa15 : Quantile 0% des volumes d'emprise de couronne (surface de couronne x hauteur d'apex). Seuls les apex dont la surface de couronne  $\geq 2,5 \text{ m}^2$  et le flowaccumulation  $\geq 15$  sont pris en compte.

Arbre.Quant.0.Vol.Cour2\_s25 : Quantile 0% des volumes de couronne (volume du MNH calculé sur l'emprise de couronne) au carré. Seuls les apex dont la surface de couronne  $\geq 25 \text{ m}^2$  sont pris en compte.

Arbre.Quant.50.Vol.Cour\_h13.5.fa20 : Quantile 50% des volumes de couronne (volume du MNH calculé sur l'emprise de couronne). Seuls les apex dont la hauteur  $\geq 13,5 \text{ m}$  sont pris en compte.

Arbre.Quant.60.Surf.Cour\_fa20 : Quantile 60% des surfaces de couronne. Seuls les apex dont le flowaccumulation  $\geq 20$  sont pris en compte.

Arbre.Quant.60.Surf.Cour\_s2.5 : Quantile 60% des surfaces de couronne. Seuls les apex dont la surface de couronne  $\geq 2,5 \text{ m}^2$  sont pris en compte.

Arbre.Quant.70.H.Apex\_vol500.fa40 : Quantile 70% des hauteurs d'apex. Seuls les apex dont le volume de couronne  $\geq 500 \text{ m}^3$  et le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.Quant.95.Vol.Cour2\_s30.s50 : Quantile 95% des volumes de couronne (volume du MNH calculé sur l'emprise de couronne) au carré. Seuls les apex dont la surface de couronne est comprise entre 30 et 50  $\text{m}^2$  sont pris en compte.

Arbre.SD.Vol.Emp.Cour\_h35.fa40 : Ecart-type des volumes d'emprise de couronne (surface de couronne x hauteur d'apex). Seuls les apex dont la hauteur  $\geq 35 \text{ m}$  et le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.SD.H.Apex\_h13.5 : Ecart-type des hauteurs d'apex. Seuls les apex dont la hauteur  $\geq 13,5 \text{ m}$  sont pris en compte.

Arbre.Som.H.Apex/ $\sqrt{N}$ \_h20 : Somme des hauteurs d'apex divisée par la racine du nombre d'apex. Seuls les apex dont la hauteur  $\geq 20 \text{ m}$  sont pris en compte.

Arbre.Som.H.Apex/ $\sqrt{N}$ \_h25 : Somme des hauteurs d'apex divisée par la racine du nombre d'apex. Seuls les apex dont la hauteur  $\geq 25 \text{ m}$  sont pris en compte.

Arbre.Som.H.Apex/ $\sqrt{N}$ \_h25.fa40 : Somme des hauteurs d'apex divisée par la racine du nombre d'apex. Seuls les apex dont la hauteur  $\geq 25 \text{ m}$  et le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.Som.H.Apex\_h12.5 : Somme des hauteurs d'apex. Seuls les apex dont la hauteur  $\geq 12,5 \text{ m}$  sont pris en compte.

Arbre.Som.H.Apex\_h12.5.fa40 : Somme des hauteurs d'apex. Seuls les apex dont la hauteur  $\geq 12,5 \text{ m}$  et le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.Som.H.Apex\_s5 : Somme des hauteurs d'apex. Seuls les apex dont la surface de couronne  $\geq 5 \text{ m}^2$  sont pris en compte.

Arbre.Som.H.Apex\_s50 : Somme des hauteurs d'apex. Seuls les apex dont la surface de couronne  $\geq 50 \text{ m}^2$  sont pris en compte.

Arbre.Som.H.Apex2\_s15.fa40 : Somme des hauteurs d'apex au carré. Seuls les apex dont la surface de couronne  $\geq 15 \text{ m}^2$  et le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.Som.H.Apex2\_s40 : Somme des hauteurs d'apex au carré. Seuls les apex dont la surface de couronne  $\geq 40 \text{ m}^2$  sont pris en compte.

Arbre.Som.Surf.Cour/ $\sqrt{N}$  : Somme des surfaces de couronne divisée par la racine du nombre d'apex.

Arbre.Som.Surf.Cour\_h13.5.fa20 : Somme des surfaces de couronne. Seuls les apex dont la hauteur  $\geq 13,5 \text{ m}$  et le flowaccumulation  $\geq 20$  sont pris en compte.

Arbre.Som.Surf.Cour\_h20.fa40 : Somme des surfaces de couronne. Seuls les apex dont la hauteur  $\geq 20 \text{ m}$  et le flowaccumulation  $\geq 40$  sont pris en compte.

Arbre.Som.Vol.Cour\_vol700 : Somme des volumes de couronne (volume du MNH calculé sur l'emprise de couronne). Seuls les apex dont le volume de couronne  $\geq 700 \text{ m}^3$  sont pris en compte.

Arbre.Som.Vol.Cour2\_s20.s30 : Somme des volumes de couronne (volume du MNH calculé sur l'emprise de couronne) au carré. Seuls les apex dont la surface de couronne est comprise entre 20 et 30  $\text{m}^2$  sont pris en compte.

Arbre.Som.Vol.Emp.Cour\_h20 : Somme des volumes d'emprise de couronne (surface de couronne x hauteur d'apex). Seuls les apex dont la hauteur  $\geq 20 \text{ m}$  sont pris en compte.

Arbre.Som.Vol.Emp.Cour\_h30 : Somme des volumes d'emprise de couronne (surface de couronne x hauteur d'apex). Seuls les apex dont la hauteur  $\geq 30 \text{ m}$  sont pris en compte.

Nuage.Dens.Rel\_6m.12m : Densité relative des points LiDAR situés dans la tranche 6 à 12 m (nombre de points de la tranche / nombre de points total).

Nuage.Dens.Rel\_t6.lst : Densité relative des points LiDAR situés dans la tranche 7. Le nuage est découpé en 10 tranches de 2 m à la hauteur du quantile 95%, puis le nombre de points de la tranche 7 est divisé par le nombre total de points situé entre 2 m et la hauteur du quantile 95%. Seuls les derniers retours sont pris en compte dans le calcul.

Nuage.Quant.05.H\_fst.1m : Quantile 5% des hauteurs du nuage de points LiDAR. Seuls les premiers retours d'une hauteur  $> 1 \text{ m}$  sont pris en compte.

Nuage.Quant.60.H\_lst.2m : Quantile 60% des hauteurs du nuage de points LiDAR. Seuls les derniers retours d'une hauteur  $> 2 \text{ m}$  sont pris en compte.

Nuage.Quant.80.H : Quantile 80% des hauteurs du nuage de points LiDAR.

Nuage.Vol\_12m.22m : Volume de canopée entre 12 et 22 m de haut calculé selon la formule de Hollaus, 2009.