

# COMPARAISON DE MÉTHODES DE SPATIALISATION POUR L'AGRÉGATION PAR PARCELLE DES ESTIMATIONS DE PARAMÈTRES FORESTIERS PAR LIDAR AÉROPORTÉ

Jean-Matthieu Monnet<sup>1,2</sup>, Alain Munoz<sup>3</sup>

1: Irstea, UR EMGR, 2 rue de la Papeterie - BP 76, F-38402 St-Martin-d'Hères, France

2: Univ. Grenoble Alpes, F-38402 Grenoble, France

3: ONF, Département Recherche Développement et Innovation, 42 quai Charles Roissard, 73000 Chambéry, France

## Résumé

Alors que les méthodes de modélisation des paramètres forestiers à partir de placettes de terrain et de données LiDAR aéroporté ont fait l'objet de nombreuses publications dans la dernière décennie, la question de leur utilisation et de leur évaluation à l'échelle de la parcelle forestière (surface de l'ordre de quelques hectares) reste peu documentée. La présente étude s'appuie sur un jeu de donnée d'inventaire en plein de 35 parcelles forestières sur 380 ha pour comparer différentes stratégies de spatialisation lors de l'application des modèles et d'agrégation des estimations de paramètres forestiers par parcelle. Les résultats montrent une diminution des erreurs entre les estimations à l'échelle de la placette et celles à l'échelle de la parcelle de 15 à 6.4% pour la surface terrière, de 26 à 7.7% pour la densité de tiges et de 6.5 à 3.4% pour le diamètre. À l'échelle de la parcelle, la précision de l'inventaire basé sur les données LiDAR se révèle similaire à celle d'un inventaire en plein, pour la surface terrière. Pour la spatialisation lors de l'application des modèles, le plus important est de respecter la taille des placettes de calibration terrain, alors que pour l'agrégation par parcelle le traitement des bordures reste délicat quel que soit le paramètre forestier.

**Mots clés :** Inventaire forestier, LiDAR aéroporté, Spatialisation, Parcelle forestière

## Abstract

*While the use of field plots and airborne LiDAR data for the estimation of forest parameters has been intensively investigated in the past ten years, the issue of the evaluation of their accuracy at the compartment level (surface of a few hectares) remains poorly documented. Based on a full-calliper inventory of 35 compartments representing 380 ha, the present study compares different strategies for the mapping of LiDAR predictions and their aggregation by compartment. Results show that the prediction error decreases between estimations at the plot and those at the compartment levels : from 15 to 6.4% for basal area, 26 to 7.7% for stem density and 6.5 to 3.4% for mean diameter. At the compartment level, a LiDAR-based inventory thus displays an accuracy similar to a full-calliper inventory, for basal area. For the mapping step, it is crucial to respect the size of field plots used for calibration, whereas for the aggregation step the handling of compartment borders remains tricky for all forest parameters.*

**Keywords :** Forest inventory, Airborne laser scanning, Upscaling, Small-area estimation

## 1. Introduction

Actuellement, l'inventaire forestier lié aux besoins de planification à moyen terme repose principalement sur une description statistique obtenue à partir d'un échantillonnage de placettes sur le terrain (Duplat et Perrotte, 1981). Cette méthode produit des estimations fiables à l'échelle du massif forestier tout en limitant l'effort de terrain requis. Cependant l'information obtenue n'est pas exploitable à l'échelle de l'unité de gestion forestière (dénommée dans la suite de l'article par le terme général « parcelle »), dont la taille est de l'ordre de quelques hectares (McRoberts, 2010). Depuis une dizaine d'années, la télédétection par scanner laser (ou LiDAR) aéroporté a démontré son potentiel pour la caractérisation de la structure tridimensionnelle des forêts, jusqu'au niveau de l'arbre. Plus particulièrement, l'approche dite surfacique

(Næsset, 2002; Hollaus et al., 2009) consiste à combiner la description du couvert végétal par le nuage de points laser et un réseau de placettes terrain de référence pour produire des cartes continues de paramètres forestiers. Les relations entre les descripteurs du nuage de points, appelés « métriques », et les paramètres forestiers comme la surface terrière, le volume sur pied, la hauteur dominante, le diamètre moyen et la densité de tiges sont calibrées statistiquement à partir des placettes de terrain.

L'évaluation de la précision des modèles d'estimation se fait généralement par validation croisée, ce qui permet d'optimiser le jeu de données disponible mais peut également se traduire par une sur-estimation de la précision (Strunk et al., 2012). Par ailleurs, la précision est ainsi estimée à l'échelle de la placette de calibration, alors que cette échelle spatiale ne correspond

pas au produit final intéressant les utilisateurs. En effet, une fois le modèle de prédiction obtenu, il doit être appliqué à l'ensemble de la couverture LiDAR pour produire une cartographie continue. Cette spatialisation suppose d'effectuer au préalable un découpage de l'espace en un ensemble de supports spatiaux pour lesquelles les modèles sont supposés valides. La solution la plus fréquemment retenue est de découper l'espace en carrés contigus, ou pixels, de surface similaire aux placettes de calibration (White et al., 2013). Ceci permet ensuite de présenter la cartographie au format raster. Cela pose cependant la question de la différence de forme entre les supports spatiaux qui servent à calibrer le modèle et ceux utilisés pour l'appliquer.

Par ailleurs certaines opérations de gestion forestière, comme la planification des coupes et l'estimation des volumes récoltables, se font à l'échelle de la parcelle. Il est donc nécessaire d'agrèger les estimations obtenues par parcelle forestière. Cette opération n'est cependant pas triviale. Étant donné la taille et la forme des parcelles, ainsi que la résolution de la carte raster, le nombre de pixels se trouvant sur des bordures représente une proportion significative du nombre total de pixels constituant la parcelle. Contrairement aux inventaires statistiques basés sur un échantillonnage, l'utilisation de la moyenne des pixels pour calculer la valeur agrégée peut donner des résultats biaisés dans le cas d'un inventaire basé sur un modèle (Breidenbach et Astrup, 2012). La majorité des travaux actuels sur ces questions d'agrégation se concentrent sur la comparaison d'estimateurs à partir du calcul théorique de leur incertitude (McRoberts, 2010; Breidenbach et Astrup, 2012; Magnussen et al., 2012) car il existe peu de données de référence terrain à cette échelle. Par ailleurs les surfaces d'agrégation considérées sont le plus souvent de l'ordre de plusieurs km<sup>2</sup>, situation dans laquelle l'influence des bordures est marginale, contrairement au cas des parcelles.

L'objectif de cet article est d'évaluer la précision d'un inventaire LiDAR à l'échelle de la parcelle, en considérant différentes options pour les étapes de spatialisation et d'agrégation. L'étude se base sur un jeu de données d'inventaire en plein unique en France, incluant 35 parcelles forestières pour 380 ha.

## 2. Données

### 2.1. Zone d'étude

La zone d'étude (583 ha) comprend les forêts publiques des communes de Prénovel et Les Piards (figure 1), situées dans le département du Jura (France). La topographie présente un fort micro-relief, avec des altitudes comprises entre 900 et 1060 m. Les forêts sont traitées en futaie jardinée, et principalement constituées de peuplements mélangés de sapin pectiné (*Abies alba*), épicéa commun (*Picea abies*) et hêtre commun (*Fagus sylvatica*). Elles sont gérées par l'Office National des Forêts (ONF).

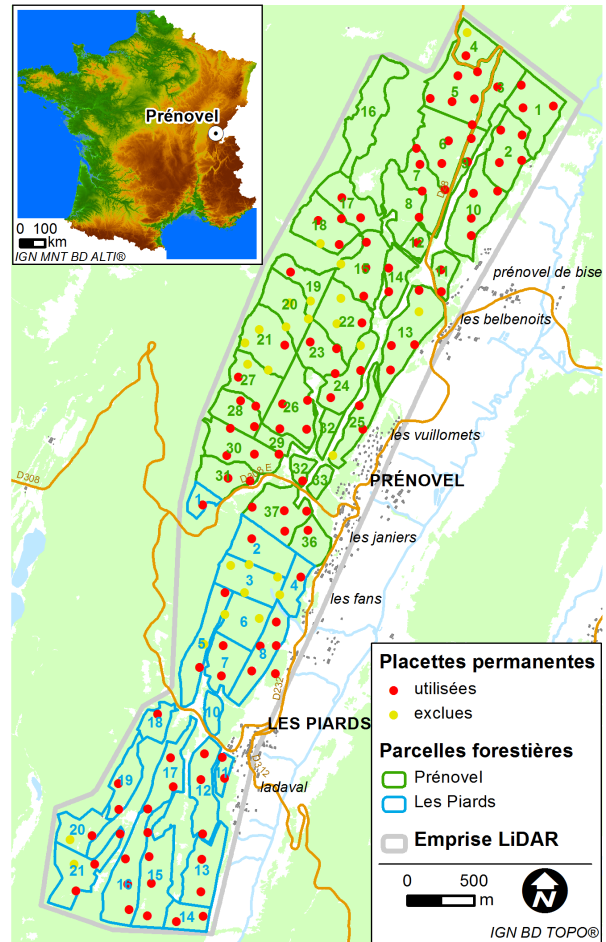


FIGURE 1: Présentation de la zone d'étude : parcellaire forestier et placettes permanentes.

### 2.2. Scanner laser aéroporté

L'acquisition LiDAR aérienne a eu lieu les 16 et 17 septembre 2012, avec un scanner à retour d'onde complète RIEGL LMS-Q560 embarqué dans un avion. La vitesse de vol était de 180 km.h<sup>-1</sup> à une hauteur de 500 m au-dessus du sol. Le recouvrement inter-bandes était de 60%, et la densité d'émission de 180 kHz avec un angle de scan de  $\pm 30$  degrés. La surface totale couverte est de 9.2 km<sup>2</sup>, avec une densité moyenne d'impulsions de  $9.3 \pm 2.3$  m<sup>-2</sup>, calculée à une résolution de 10 m. Le traitement des données brutes a été effectué par le prestataire : les échos ont été extraits et géoréférencés avec la suite logicielle RIEGL, puis le nuage de points a été classé selon les catégories « sol » et « végétation » avec le logiciel TerraScan.

### 2.3. Placettes forestières permanentes

Un réseau de 139 placettes permanentes a été installé en 2005, selon une grille régulière d'espacement 200 m. L'inventaire a été mis à jour entre le 9 septembre et le 6 octobre 2011. Le protocole se base sur des placettes circulaires emboîtées. Le centre des placettes est géoréférencé au moyen d'un récepteur GPS grand public. Les arbres dont le diamètre à hauteur de poitrine (DHP) est supérieur ou égal à 7.5 cm sont inventoriés

jusqu'à 10 m du centre, et jusqu'à 17 m si leur DHP est supérieur ou égal à 27.5 cm. Pour chaque arbre inventorié, le DHP, l'espèce ainsi que la distance et l'azimut par rapport au centre de la placette sont mesurés. Le géoréférencement des placettes a été amélioré en ajustant la localisation des arbres et le modèle numérique de hauteur LiDAR (Monnet et Mermin, 2014). Une placette n'a pas pu être localisée de manière satisfaisante, pour les autres la précision finale de la position est de l'ordre de un mètre. 26 placettes ont été exclues de l'analyse lors de cette étude. Elles sont situées dans des parcelles ayant été exploitées entre la date d'inventaire et la date de vol et leur modèle numérique de hauteur présente des différences notables avec la position des arbres inventoriés. Pour chacune des 113 placettes restantes (figure 1), trois paramètres forestiers sont calculés en ne prenant en compte que les arbres dont le DHP est supérieur ou égal à 17.5 cm (tableau 1), pour avoir des données en accord avec l'inventaire en plein (cf section 2.4).

- surface terrière : somme de la surface horizontale du tronc des arbres, mesurée à hauteur de poitrine (1.30 m), divisée par la surface de forêt occupée par les arbres, ici 907.9 m<sup>2</sup> ;
- diamètre dominant : moyenne des diamètres des 100 plus gros arbres par hectare, estimé pour une placette de 17 m de rayon par la moyenne des diamètres des huit plus gros arbres (Pardé et Bouchon, 1988) ;
- densité de tiges : nombre d'arbres divisé par la surface de forêt qu'ils occupent.

	Surface terrière (m <sup>2</sup> .ha <sup>-1</sup> )	Densité de tiges (ha <sup>-1</sup> )	Diamètre dominant (cm)
Minimum	8.2	55	25.0
Moyenne	29.5	278	49.6
Maximum	59.5	755	65.3

**Tableau 1 :** Statistiques forestières des placettes permanentes ( $n = 113$ ), pour les arbres de DHP supérieur ou égal à 17.5 cm.

#### 2.4. Inventaire en plein des parcelles

Les 35 parcelles forestières de la forêt communale de Prénovel (380 ha) ont fait l'objet d'un inventaire en plein de septembre à décembre 2011. Le protocole pour l'inventaire en plein est basé sur le parcours exhaustif de la parcelle avec pointage des arbres de plus de 17.5 cm de DHP par espèce et classe de diamètre de 5 cm d'intervalle. Cependant, les arbres déjà martelés, c'est-à-dire marqués car destinés à être exploités prochainement, ont été ignorés lors de l'inventaire. Il a donc été nécessaire de consolider les données. Pour cinq parcelles martelées avant l'inventaire mais n'ayant pas été exploitées à la date du vol LiDAR, les arbres martelés ont été rajoutés à l'inventaire en plein. Pour les 30 autres parcelles, des arbres ont été martelés et exploités entre l'inventaire et le vol LiDAR. Ceux-ci ont été déduits de l'inventaire. La surface terrière, la densité de tiges et le diamètre dominant

sont ensuite calculés pour chaque parcelle en prenant pour diamètre le centre de chaque classe et en utilisant la superficie mentionnée dans les documents d'aménagement (tableau 2).

	Superficie (ha)	Surface terrière (m <sup>2</sup> .ha <sup>-1</sup> )	Densité de tiges (ha <sup>-1</sup> )	Diamètre dominant (cm)
Min.	2.5	22.1	206	43.7
Moy.	10.8	29.6	258	49.9
Max.	18.0	40.1	376	58.0

**Tableau 2 :** Statistiques des parcelles forestières ( $N = 35$ ).

### 3. Méthode

#### 3.1. Chaîne de traitement

Les principales étapes de la chaîne de traitement (figure 2) sont détaillées dans les sections suivantes :

- calibration des modèles d'estimation à partir des placettes permanentes,
- spatialisations des estimations LiDAR en appliquant les modèles à des supports spatiaux qui couvrent l'ensemble de la zone d'étude,
- agrégation des estimations par parcelle.

L'évaluation des différentes options pour la spatialisations et pour l'agrégation se fait en comparant les valeurs estimées et les valeurs terrain à l'échelle de la parcelle, pour les trois paramètres forestiers considérés : surface terrière, densité de tiges et diamètre dominant.

#### 3.2. Métriques utilisées pour la modélisation

La modélisation et la spatialisations nécessitent de calculer des métriques décrivant les caractéristiques du nuage de points LiDAR sur l'emprise géographique de chaque support spatial, quelle que soit sa forme. Pour cela, le nuage de points est préalablement normalisé en retranchant à l'altitude de chaque point l'altitude du sol obtenue par projection verticale du point sur une surface TIN (triangulated irregular network) constituée par les points classés « sol ». Pour chaque support spatial, le nuage de points est extrait puis utilisé pour calculer différents indicateurs statistiques, appelés « métriques ». Celles-ci sont de deux sortes. Les métriques usuelles décrivant la distribution des hauteurs du nuage de points sont calculées au moyen du logiciel LASManager développé par l'ONF. À celles-ci s'ajoutent des métriques calculées à partir d'une segmentation préalable du modèle numérique de hauteur (MNH) sur chaque placette. Le MNH est calculé selon la méthode de Khosravipour et al. (2013), à résolution 0.5 m. Ensuite une segmentation par bassin versant (Gonzalez et Woods, 2008) est réalisée pour individualiser les houppiers des principaux arbres. Cette étape est réalisée au moyen des outils hydrologiques de l'extension *Spatial analyst* du logiciel ArcGIS 10.0. Les métriques « arbre » décrivent la distribution des surfaces, hauteurs et volumes sous le MNH des segments, après

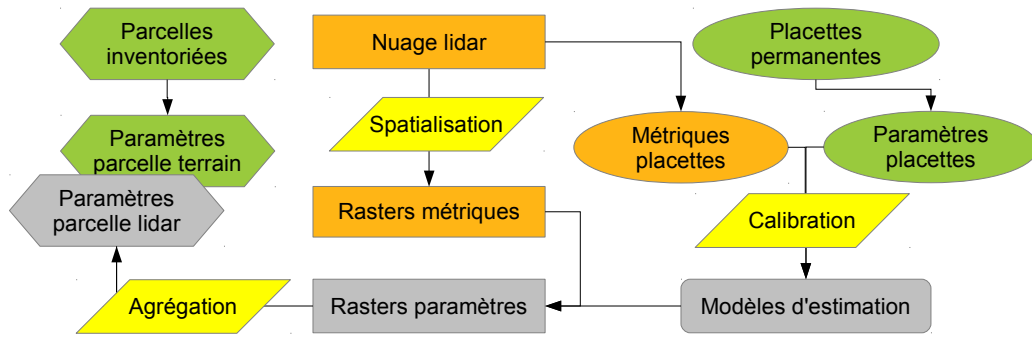


FIGURE 2: Schéma de la chaîne de traitement.

seuillage selon différents critères de hauteur ou de surface. Elles sont détaillées en annexe d'un autre article de ce numéro (Munoz et al., 2015).

### 3.3. Calibration des modèles d'estimation

Les métriques sont calculées sur l'emprise de chacune des placettes permanentes. Une procédure de sélection pas à pas est ensuite utilisée pour calibrer par la méthode des moindres carrés des modèles de régression linéaire (équation 1) :

$$y = b + \sum_{k=1}^m a_k \times x_k + \epsilon \quad (1)$$

avec  $y$  le paramètre forestier,  $(x_k)$  des métriques,  $(a_k)$  et  $b$  les coefficients et  $\epsilon$  l'erreur du modèle. Cette étape est effectuée au moyen du logiciel *R*. Tout d'abord les métriques qui minimisent l'erreur dans l'équation 1 avec  $m = 1$  sont sélectionnées. Les modèles de régression linéaire combinant deux ou trois de ces métriques sont alors tous testés (équation 1 avec  $m \in \{2, 3\}$ ). Ceux qui ne satisfont pas aux hypothèses du modèle linéaire (package *gvmlma*), ou comportant une variable dont le VIF (Variance Inflation Factor) est supérieur à 5 sont exclus. Pour les modèles restants, la précision en prédiction est estimée par validation croisée basée sur trois groupes (3-fold cross-validation). Pour chaque paramètre forestier, on retient finalement le modèle à deux ou trois variables avec le meilleur  $\text{adj-R}^2$ , après avoir vérifié qu'aucune variable externe telle que la densité de tiges, le pourcentage de feuillus, la densité d'impulsions LiDAR ou la topographie n'a d'effet significatif lorsqu'elle est ajoutée au modèle. Cette dernière vérification évite de sélectionner des modèles avec des résidus dépendants de ces facteurs externes. De tels modèles pourraient engendrer des artefacts spatiaux lors de la spatialisation, du fait de la variabilité des facteurs externes sur l'ensemble de la zone d'étude. La précision des modèles est évaluée en calculant la racine carrée de l'erreur quadratique moyenne  $eqm$  (équation 2) :

$$eqm = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - \hat{y}_k)^2} \quad (2)$$

avec  $y_k$  la valeur du paramètre forestier mesurée sur la placette  $k$ ,  $\hat{y}_k$  celle estimée en validation croisée par le modèle et  $n = 113$  le nombre de placettes de calibration.

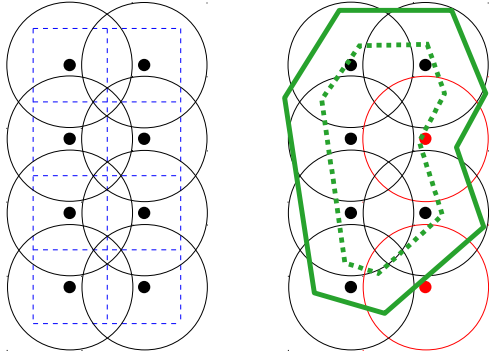
### 3.4. Spatialisation

Pour spatialiser les estimations des différents paramètres forestiers, c'est à dire réaliser la cartographie à partir du nuage de points LiDAR complet et des modèles de prédiction, les modèles sont appliqués à des supports spatiaux qui couvrent l'ensemble du domaine d'application. Les coefficients retenus dans les modèles sont ceux obtenus lors de la calibration avec l'ensemble des placettes de référence. Les métriques présentées à la section 3.2 sont calculées pour chacun de ces supports. Lors de cette étape, trois critères utilisés lors de la définition des supports spatiaux, et susceptibles d'influencer la précision des estimations à l'échelle de la parcelle, sont testés. Le premier est l'espacement entre le centre des différents supports. Si les métriques sont par la suite stockées au format raster, cet espacement correspond à la taille du pixel. Le deuxième critère concerne la forme du support. L'utilisation du cercle permet de respecter la forme des placettes qui ont servi à la calibration (White et al., 2013). Cependant dans le cas d'un stockage au format raster, il y aura risque de confusion entre l'emprise du pixel et l'emprise réelle de calcul de la valeur associée. Le troisième critère est la taille du support. Idéalement il faudrait que sa surface soit similaire à celle des placettes de calibration. Comme celles-ci sont le plus souvent circulaires avec pour rayon une valeur entière en mètres, il est peu probable que le carré de même surface ait également un nombre entier pour taille de côté. L'utilisation d'une taille de pixel non entière pour les rasters peut se révéler problématique lorsque l'on est amené à croiser ces données avec d'autres sources, par exemple des données topographiques ou de télédétection optique (Monnet et al., 2015). Dans cette étude, les combinaisons suivantes ont donc été considérées pour la forme et la taille des supports, en privilégiant des valeurs entières :

- cercle de rayon 17 m : forme et surface identiques aux placettes de calibration ;
- carré de côté 20 m : forme et surface différentes ;
- carré de côté 30 m : forme différente mais surface similaire (900 m<sup>2</sup> pour le carré, 907.9 m<sup>2</sup> pour le cercle).

Dans le cas des supports de forme carrée, l'espacement est pris égal à la taille de leur côté. Ils sont donc équivalents aux pixels. Dans le cas des supports circulaires, les

espacements suivants sont testés : 10, 20, 30 et 40 m. Dans le cas de 10 et 20 m, les supports se chevauchent et l'emprise des pixels qui servent à stocker l'information est plus petite que celle des supports spatiaux (figure 3, gauche). Avec un espacement de 40 m, les supports sont nettement disjoints et la surface de leur emprise est plus petite que celle des pixels associés. De plus certaines parties du nuage de points ne contribuent pas au calcul.



**FIGURE 3:** Gauche : schéma de la spatialisation dans le cas d'un support spatial de rayon 17 m, espacement de 20 m. Les métriques sont calculées sur l'emprise des cercles en trait plein, mais sont ensuite stockées dans des pixels en trait pointillé, dont l'emprise est, dans ce cas, plus réduite. Droite : schéma de l'agrégation pour une parcelle forestière (polygone en trait plein) contenant les centres de sept supports spatiaux. Lorsqu'un tampon de -10 m est appliqué au polygone (trait pointillé), les supports dont le centre est en rouge sont exclus du calcul.

### 3.5. Agrégation

#### 3.5.1. Prise en compte des bordures

Pour le calcul des paramètres forestiers à l'échelle de la parcelle, l'approche la plus simple consiste à faire la moyenne des valeurs des paramètres estimés pour les supports dont le centre est inclus dans la parcelle. Ceci correspond au « synthetic regression estimator » (SRE, noté  $\hat{Y}_S$ ) utilisé par Breidenbach et Astrup (2012). Afin d'évaluer l'influence des bordures, seuls les supports se trouvant à une certaine distance de la bordure de la parcelle sont considérés (équation 3) :

$$\hat{Y}_{S,j} = \frac{1}{\text{card}(V(j))} \sum_{i \in V(j)} \hat{y}_i \quad (3)$$

avec  $\hat{Y}_{S,j}$  la valeur estimée par LiDAR pour la parcelle  $j$ ,  $V(j)$  l'emprise de la parcelle éventuellement diminuée par le tampon négatif,  $\text{card}(V(j))$  le nombre de supports dont le centre est inclus dans  $V(j)$ , et  $\hat{y}_i$  l'estimation LiDAR sur l'entité  $i$ . Les limites de distance testées sont 0, 10 et 20 m. Cela revient à appliquer aux polygones d'emprise des parcelles une zone tampon négative avant la sélection des supports contribuant au calcul (figure 3, droite). L'emprise initiale des parcelles est donnée par le fichier numérique du parcellaire de l'ONF.

#### 3.5.2. Correction du biais

L'utilisation de la moyenne lors de l'agrégation (SRE) est potentiellement source de biais pour les inventaires basés sur des modèles (McRoberts, 2010; Breidenbach et Astrup, 2012). Celui-ci peut être corrigé en retranchant pour chaque parcelle le résidu moyen constaté sur les placettes de calibration incluses dans la parcelle (generalized regression estimator, GREG, noté  $\hat{Y}_G$ , équation 4) :

$$\hat{Y}_{G,j} = \hat{Y}_{S,j} + \frac{1}{n_k} \sum_{k \in V(j)} (y_k - \hat{y}_k) \quad (4)$$

avec  $n_k$  le nombre de placettes de terrain contenues dans la parcelle  $j$ . Le risque est d'augmenter la variance de l'erreur dans le cas où le nombre de placettes par parcelle est faible. Cet estimateur est calculé uniquement dans le cas suivant : cercle de 17 m de rayon, espacement de 20 m et agrégation avec un tampon négatif de 10 m.

#### 3.5.3. Évaluation

La précision des estimations des paramètres forestiers à l'échelle de la parcelle est évaluée selon le biais et la racine carrée de l'erreur quadratique moyenne  $EQM$  (équations 5).

$$\begin{aligned} \text{Biais} &= \frac{1}{N} \sum_{j=1}^N (Y_j - \hat{Y}_j) \\ EQM &= \sqrt{\frac{1}{N} \sum_{j=1}^N (Y_j - \hat{Y}_j)^2} \end{aligned} \quad (5)$$

avec  $Y_j$  la valeur du paramètre forestier calculée à partir de l'inventaire en plein sur la parcelle  $j$ ,  $\hat{Y}_j$  celle estimée par agrégation des valeurs LiDAR et  $N = 35$  le nombre de parcelles.

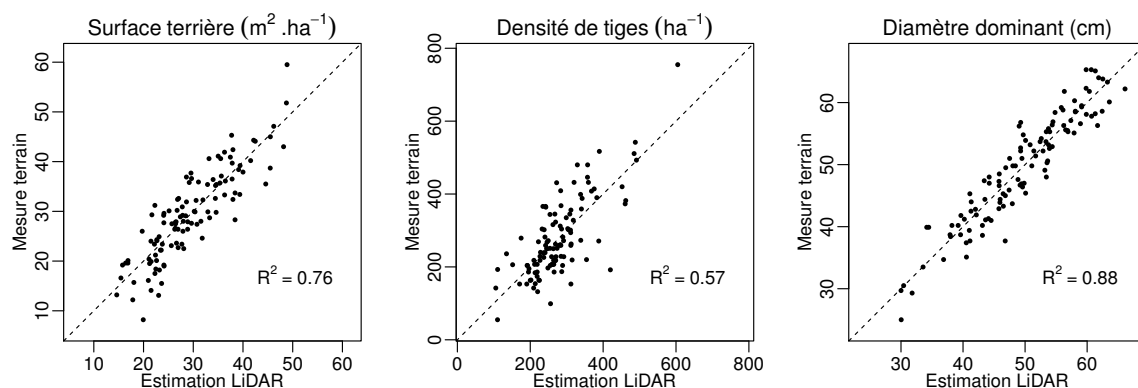
## 4. Résultats

### 4.1. Modèles d'estimation

Les modèles d'estimation obtenus après calibration sont respectivement pour la surface terrière (G), la densité de tiges (N) et le diamètre dominant (D) :

$$\begin{aligned} \hat{G} &= 1.05 + 2.56g_1 - 1.28g_2 + 0.0123g_3 \\ \hat{N} &= 531 + 25.4n_1 - 2.37n_2 - 3.29n_3 \\ \hat{D} &= 25.8 + 0.0118d_1 + 0.00107d_2 + 0.917d_3 \end{aligned} \quad (6)$$

avec les métriques suivantes dans les modèles retenus :  $g_1$  la hauteur moyenne des derniers échos (hors sol),  $g_2$  l'écart moyen entre les derniers échos et la hauteur moyenne,  $g_3$  le quatrième décile du volume des houppiers segmentés,  $n_1$  le cinquième percentile de hauteur des premiers retours (hors sol),  $n_2$  le sixième décile de surface des houppiers segmentés,  $n_3$  le coefficient de variation de la hauteur de derniers retours de hauteur supérieure à 1 m,  $d_1$  le neuvième décile de la hauteur



**FIGURE 4:** Valeurs des paramètres forestiers mesurées sur le terrain en fonction des valeurs prédites par le modèle, pour les 113 placettes permanentes. De gauche à droite : surface terrière, densité de tiges et diamètre dominant.

			G	N	D
			m <sup>2</sup> .ha <sup>-1</sup>	ha <sup>-1</sup>	cm
Modèle	R <sup>2</sup>		0.76	0.57	0.88
	<i>eqm</i>		4.5	75.0	3.3
Validation parcelle	SRE	R <sup>2</sup>	0.85	0.76	0.91
		<i>Biais</i>	-0.4	-7.4	-1.2
		<i>EQM</i>	1.9	19.8	1.7
	GREG	R <sup>2</sup>	0.66	0.46	0.77
		<i>Biais</i>	0.0	-0.9	-1.0
		<i>EQM</i>	3.8	47.0	2.1

**Tableau 3:** Évaluation de la précision des modèles d'estimation LiDAR pour la surface terrière (G), la densité de tiges (N) et le diamètre dominant (D), à l'échelle de la placette et de la parcelle. Pour l'agrégation par parcelle, les critères retenus sont ceux du paragraphe 4.2.2, SRE est la moyenne des estimations LiDAR et GREG est égal à SRE corrigé du biais constaté sur les placettes de terrain.

maximum des segments,  $d_2$  le volume total des houppiers segmentés et  $d_3$  l'écart-type de la hauteur maximum des segments. Pour la surface terrière, le  $R^2$  du modèle est 0.76 et l'erreur donnée par la validation croisée est de 4.5 m<sup>2</sup>.ha<sup>-1</sup>, soit 15% de la valeur moyenne (tableau 3). Pour la densité de tiges, le  $R^2$  est de 0.57 et l'erreur est de 72 ha<sup>-1</sup> (26%). Pour le diamètre dominant, le  $R^2$  est de 0.88 et l'erreur de 3.3 cm (6.5%). La figure 4 présente les mesures de terrain en fonction des valeurs prédites par le modèle.

## 4.2. Validation à l'échelle parcelle

### 4.2.1. Effet des paramètres de spatialisation et d'agrégation

Dans cette section, les valeurs d'erreur sont calculées à partir du « synthetic regression estimator » (équation 3). Pour la surface terrière, le biais est compris en -0.2 et -0.5 m<sup>2</sup>.ha<sup>-1</sup> pour toutes les combinaisons, sauf celles basées sur des pixels de 20 m, pour lesquelles le biais est entre -0.8 et -1.0. Dans tous les cas, l'erreur quadratique moyenne est comprise entre 1.9 et 2.3 m<sup>2</sup>.ha<sup>-1</sup>. Si l'on exclut les pixels de 20 m, la tendance pour le biais est la suivante : il est le plus faible en valeur absolue pour

un tampon de 20 m, et le plus élevé pour 10 m (figure 5). En ce qui concerne l'erreur, elle est plus élevée avec un tampon de 20 m. L'erreur est également plus faible pour les supports circulaires espacés de 10 ou 20 m.

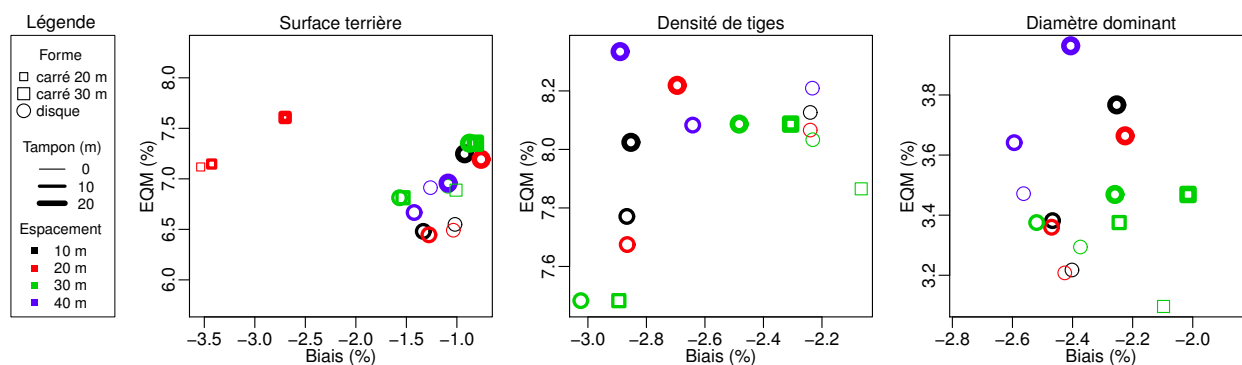
Pour la densité de tiges, le biais est compris entre -5.3 et -7.8 ha<sup>-1</sup>, sauf pour les combinaisons basées sur des pixels de 20 m (-27.4 à -28.7). Dans ce cas, l'erreur est supérieure à 34.5 ha<sup>-1</sup>, alors qu'elle est comprise entre 19.3 et 21.2 dans les autres cas. Pour ces derniers, on remarque également un biais absolu légèrement plus faible quand l'agrégation se fait sans tampon négatif (-5.3 à -5.5 contre -6.0 à -7.5). L'erreur est par contre légèrement plus faible dans le cas des tampons de 10 m. Qu'il s'agisse du biais ou de l'erreur, l'espacement de 30 m semble donner les meilleurs résultats.

Pour le diamètre dominant, le biais est compris entre -1.0 et -1.3 cm sauf pour les combinaisons basées sur des pixels de 20 m (5.8 à 6.0). Pour ces supports spatiaux, l'erreur est supérieure à 6.0 cm, alors qu'elle est comprise entre 1.5 et 2.0 dans les autres cas. Pour ces derniers, on remarque également un biais absolu légèrement plus faible quand l'agrégation se fait avec 20 m de tampon négatif. Par contre, l'erreur est plus faible lorsqu'aucun tampon n'est appliqué.

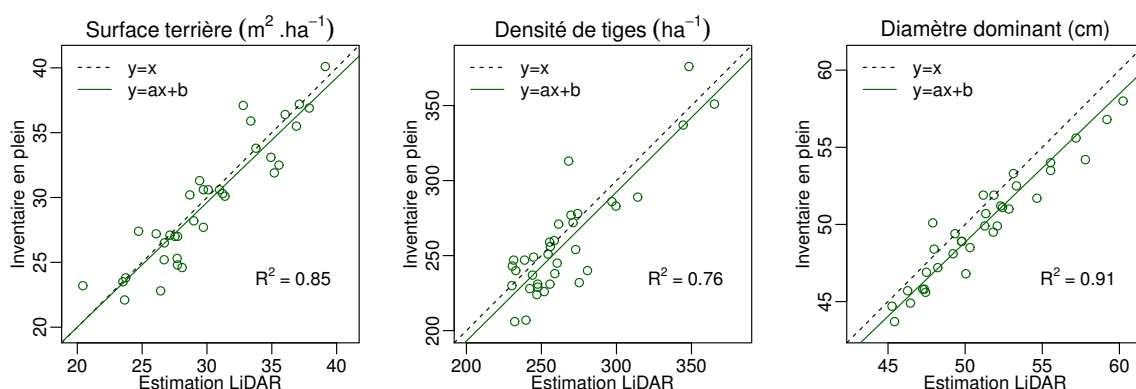
Dans tous les cas, les résultats pour les pixels de 20 m sont les plus mauvais. Les supports spatiaux carrés et circulaires d'espacement 30 m obtiennent des valeurs similaires, ainsi que les cercles d'espacements 10 et 20 m. Le cercle d'espacement 40 m paraît moins pertinent, sauf pour la surface terrière.

### 4.2.2. Correction du biais

Dans ce paragraphe, les critères retenus pour la spatialisation et l'agrégation sont : cercle de 17 m de rayon, espacement de 20 m et agrégation des supports spatiaux dont le centre est à plus de 10 m de la bordure. Le biais et l'erreur à l'échelle de la parcelle pour la surface terrière, la densité de tiges et le diamètre dominant sont respectivement -0.4 et 1.9 m<sup>2</sup>.ha<sup>-1</sup>, -7.4 et 19.8 ha<sup>-1</sup>, -1.2 et 1.7 cm pour l'estimateur SRE (tableau 3). Les coefficients de corrélation entre les valeurs prédites et celles inventoriées pour ces paramètres sont respectivement 0.85, 0.76 et 0.91, et sont supérieurs à ceux



**FIGURE 5:** Erreur quadratique moyenne en fonction du biais pour la validation à l'échelle parcelle, des estimations des paramètres forestiers (SRE). Les valeurs sont exprimées en pourcentage de la valeur moyenne mesurée par l'inventaire en plein. L'apparence des symboles est fonction des paramètres de spatialisation et d'agrégation. Pour la densité de tiges et le diamètre dominant, les valeurs obtenues avec les pixels de 20 m sont hors figure car le biais est très supérieur en valeur absolue.



**FIGURE 6:** Valeurs des paramètres forestiers calculées à partir de l'inventaire en plein, en fonction des estimations obtenues par agrégation des estimations LiDAR (SRE), pour les 35 parcelles forestières. De gauche à droite : surface terrière, densité de tige et diamètre dominant.

observés à l'échelle de la placette. Les courbes de régression linéaire illustrent la sur-estimation par les prédictions LiDAR (figure 6). Le biais est significatif pour la densité de tiges (test de Student apparié,  $p < 0.05$ ) et pour le diamètre dominant ( $p < 0.001$ ). La correction par le biais constaté sur les placettes de calibration (estimateur GREG) annule quasiment le biais pour la surface terrière ( $-0.01 \text{ m}^2 \cdot \text{ha}^{-1}$ ) et pour la densité de tiges ( $-0.9 \text{ ha}^{-1}$ ). Le biais pour le diamètre dominant change peu ( $-1.0 \text{ cm}$ ) et reste significatif ( $p < 0.01$ ). Par contre, l'erreur quadratique est doublée pour la surface terrière et la densité de tiges, alors qu'elle n'augmente que légèrement pour le diamètre dominant. Les corrélations entre les estimations GREG et les mesures de l'inventaire en plein sont inférieures à celles constatées à l'échelle de la placette (tableau 3).

## 5. Discussion

L'écart temporel existant entre le vol LiDAR et les données de terrain peut être une source d'erreur malgré la correction apportée concernant l'exploitation sylvicole. En effet, c'est un vol de 2012 qui est utilisé pour estimer des valeurs mesurées un an plus tôt. L'accrois-

sement moyen ainsi que le passage à la futaie (arbres dont le diamètre devient supérieur à la limite de recensement) sont *a priori* intégrés par les modèles. Par contre, les changements dus à l'exploitation ne sont pas traités de la même manière : exclusion pour les placettes ayant subi des changements, mais suppression des arbres effectivement coupés pour les parcelles exploitées entre la date des relevés de terrain et celle du vol LiDAR. Dans la mesure où les coupes ne sont pas effectuées de manière aléatoire, mais préférentiellement dans les peuplements à forte surface terrière ou fort diamètre dominant, il est possible que les placettes restantes soient un échantillon non représentatif vis à vis de l'ensemble des forêts communales. De plus, les placettes situées sur Les Piards ont été utilisées pour la calibration. Ainsi les modèles LiDAR, bien que théoriquement non biaisés pour l'échantillon de placettes utilisées, pourraient engendrer un biais lors de leur application sur Prénovel.

Les forêts publiques des communes de Prénovel et Les Piards restent un massif assez particulier du fait de l'homogénéité du milieu et du traitement sylvicole en futaie jardinée. La comparaison de la précision des modèles d'estimation avec celles obtenues dans d'autres contextes doit donc se faire avec précaution. Pour des

forêts résineuses et mélangées de montagne majoritairement traitées en futaie irrégulière, Munoz et al. (2015) ont obtenu avec la même méthodologie des précisions de 14 à 16% pour la surface terrière, 17 à 21% pour la densité de tiges et 7 à 8% pour le diamètre dominant. Les valeurs obtenues ici sont très proches, sauf pour la densité de tiges où la précision est légèrement moins bonne avec 26%. Avec 116 placettes réparties en trois types de peuplements, Næsset (2004) a obtenu pour des forêts de Norvège à dominante résineuse des erreurs de prédiction avec un écart-type de 14.8 à 23.3% pour la surface terrière et de 25.3 à 26.2% pour la densité de tiges, les biais étant non significatifs. Ces résultats ont cependant été obtenus pour des placettes plus petites en taille (232.9 contre 907.9 m<sup>2</sup>) et avec un diamètre de recensement plus faible (4 cm dans les jeunes peuplements et 10 cm dans les peuplements matures). Dans cette étude scandinave, les estimations LiDAR ont été agrégées pour 57 grandes placettes d'environ 3736 m<sup>2</sup>. L'écart-type des erreurs était de 8.7 à 13.6% pour la surface terrière et de 14.3 à 29.3% pour la densité de tiges, selon le type de peuplement. Les grandes placettes ont été choisies arbitrairement, mais il est notable que pour la surface terrière, l'écart-type de l'erreur a été divisé par deux pour une augmentation de la surface d'un facteur 16. Lorsque leurs modèles sont évalués pour 18 grappes de quatre placettes situées dans le Vercors (Isère, France), Munoz et al. (2015) ont constaté que l'erreur passait de 14 à 13% pour la surface terrière et de 17 à 13% pour la densité de tiges, tandis que l'erreur pour le diamètre dominant restait à 8%. L'amélioration est moins nette, mais l'augmentation de surface n'est que d'un facteur trois.

Les données d'inventaire en plein des forêts de Pré-novel permettent de quantifier l'erreur à l'échelle de la parcelle, pour une augmentation de surface bien supérieure, d'un facteur 30 à 200 selon la superficie des parcelles. Les résultats montrent une diminution de l'erreur de 15 à 6.4% pour la surface terrière, de 26 à 7.7% pour la densité de tiges et de 6.5 à 3.4% pour le diamètre dominant, entre le modèle à la placette et l'agrégation par parcelle (SRE). L'amélioration de la précision est d'un facteur deux pour la surface terrière et le diamètre dominant, et de plus de trois pour la densité de tiges. Ce paramètre forestier est souvent mal estimé à l'échelle de la placette par télédétection LiDAR, avec des erreurs de l'ordre de 30%, mais il s'avère donc que l'erreur est fortement réduite à l'échelle de la parcelle. Lors d'un inventaire en plein, les erreurs vont plutôt dans le sens d'un oubli de tiges que d'un double compte (Duplat et Perrotte, 1981), ce qui pourrait expliquer que les estimations LiDAR pour la surface terrière et la densité de tiges soient légèrement supérieures à celles de l'inventaire en plein.

La correction par les résidus constatés sur les placettes de calibration (GREG) donne un biais moyen sur l'ensemble des parcelles très faible pour ces deux paramètres. Ces résultats confirment que les inventaires basés sur des modèles peuvent engendrer un biais (Mc-Roberts, 2010). Pour le diamètre dominant le biais persiste, ce qui suggère l'existence d'autres causes. No-

tamment, ce ne sont pas les mêmes définitions qui sont appliquées pour le calcul à partir des entités LiDAR et pour les données d'inventaire en plein. Les modèles estiment un diamètre dominant défini comme la moyenne des diamètres des huit plus gros arbres sur des entités de 907.9 m<sup>2</sup>. L'estimation LiDAR pour la parcelle est ensuite un moyenne pour les supports spatiaux situés dans la parcelle. À partir des données d'inventaire, le diamètre dominant est calculé comme la moyenne des diamètres des  $N$  plus gros arbres où  $N$  est la partie entière de la surface de la parcelle en ares. Les définitions sont convergentes si les gros arbres sont répartis de manière homogène sur la parcelle. Sinon, la valeur estimée pour l'ensemble de la parcelle sera supérieure à celle obtenue en moyennant les estimations obtenues pour les supports qui la composent. Du fait de cette différence de mode de calcul, les estimations LiDAR devraient être plus faibles que celles de l'inventaire en plein, or les résultats montrent le contraire. Dans le cas d'une futaie jardinée, il est donc probable que cette différence de définition n'ait qu'un effet mineur. Les causes pourraient être liées à une différence de protocole pour la mesure des diamètres entre les placettes et les parcelles.

La correction du biais s'accompagne pour tous les paramètres d'une augmentation importante de l'erreur, liée au faible nombre de placettes par parcelle, entre 0 et 6. La correction est pertinente à l'échelle du massif, mais plutôt dommageable au niveau de la parcelle. L'estimateur SRE semble donc le plus adapté dans le cas où le produit final est l'estimation par parcelle. Le calcul théorique de l'erreur de cet estimateur reste complexe (Mc-Roberts, 2010), et comprend un terme correspondant à la corrélation spatiale des erreurs. L'erreur serait *a priori* plus élevée pour les parcelles les plus petites en surface. Pour la surface terrière et le diamètre dominant, la corrélation entre l'erreur quadratique et la surface des parcelles est non significative (coefficient de corrélation de Spearman  $\rho=-0.6$  et  $\rho=-0.12$  respectivement), alors qu'elle est de 0.32 pour la densité de tiges ( $p=0.06$ ), là où on attendrait une valeur négative. Le traitement en futaie jardinée fait que la structure du peuplement est localement hétérogène, mais homogène sur l'ensemble du massif, ce qui devrait limiter la portée de la corrélation spatiale de l'erreur pour des parcelles faisant entre 2.5 et 18 hectares, ce qui représente entre 28 et 200 pixels à résolution 30 m.

Pour la surface terrière, Duplat et Perrotte (1981) estiment par ailleurs que la valeur calculée à partir d'un inventaire en plein a 95% de chances de se situer dans l'intervalle [-15%, +10%] de la vraie valeur. En faisant l'hypothèse d'une distribution normale, les erreurs relatives suivraient une loi normale de moyenne -2.5 et d'écart-type 6.4. Si les estimations SRE sont prises comme référence, les valeurs de l'inventaire en plein présentent un biais de -1.2%, les 35 valeurs sont dans l'intervalle [-14%, +14%] et leur distribution ne peut pas être considérée comme significativement différente de la loi normale centrée en -2.5 et d'écart-type 6.4 ( $p=0.54$ , test de Kolmogorov-Smirnov). Ces valeurs montrent que les es-



timations LiDAR atteignent à l'échelle de la parcelle un niveau de précision comparable à celui d'un inventaire en plein.

Les étapes de spatialisation et d'agrégation ont une influence non négligeable sur la précision finale à l'échelle de la parcelle. Il est notamment crucial de respecter lors de la spatialisation la surface des placettes de calibration. À surface similaire (environ 900 m<sup>2</sup>), les résultats obtenus avec des pixels semblent légèrement meilleurs que pour les supports circulaires, du point de vue du biais. Cela pourrait être dû au fait que l'ensemble du nuage de points est utilisé par les pixels, alors que pour les cercles certaines parties sont utilisées en redondance tandis que d'autres ne le sont pas du tout. La question de l'espacement de la grille finale ne semble pas jouer un rôle prépondérant. L'influence de la taille du tampon semble dépendre du paramètre forestier, probablement car les métriques qui composent les différents modèles sont plus ou moins affectées par les zones de bordures.

Les effets de bord dus à la « pixelisation » de l'emprise des parcelles devraient être réduits en utilisant des supports plus petits pour le calcul et une grille moins espacée, mais l'erreur des estimations et son auto-corrélation seraient probablement augmentées. Pour des supports plus grands, l'exclusion d'une zone tampon devrait éviter que ceux qui se trouvent en partie sur la parcelle voisine ne soient intégrés au calcul. Par contre, cela risque également de se traduire par une sous-représentation des bordures du peuplement dans le calcul, alors que les arbres y sont fréquemment différents de ceux qui se trouvent à l'intérieur du peuplement, notamment dans le cas des lisières. Le cas des bordures est particulièrement problématique pour les petites parcelles et pour les parcelles très allongées, où le ratio périmètre sur surface est élevé. Afin de mieux les prendre en compte, une stratégie différente est envisageable pour les supports spatiaux en bordure. Un support partagé entre deux parcelles pourrait être réparti au prorata de la surface, et celui partagé entre une parcelle et une zone non forestière, entièrement affecté à la parcelle. Une autre possibilité serait de contraindre l'étape de spatialisation par l'emprise des parcelles, en faisant en sorte qu'aucun support n'intersecte une bordure. Cependant cette prise en compte des limites de parcelles ne pourra être efficace que si les fichiers vectoriels de l'emprise des parcelles ont une précision suffisante, c'est à dire largement inférieure à la taille des supports spatiaux.

## 6. Conclusion

Tirant parti d'un jeu de données d'inventaire en plein sur 380 ha, cette étude montre que pour trois paramètres forestiers usuels, surface terrière, diamètre dominant et densité de tiges, la télédétection par LiDAR aéroporté atteint une précision d'estimation meilleure que 8% à l'échelle de la parcelle. Ce niveau de précision rejoint celui des références terrain disponibles et il sera probablement difficile d'aller plus loin dans l'évaluation des limites de cette technologie à cette échelle spatiale.

À l'échelle de la parcelle, les estimations sont peu biaisées tant que les supports spatiaux pour la spatialisation respectent la surface des placettes de calibration. Pour des raisons pratiques, il semble donc judicieux d'utiliser des pixels, qui par ailleurs offrent une correspondance entre le support spatial et sa représentation raster.

La prise en compte des bordures via un tampon d'exclusion donne des résultats variables selon les paramètres forestiers. Pour aller plus loin dans l'analyse, il serait intéressant de passer d'une approche pixel à une approche orienté-objet, mais la précision des références existantes d'emprise des parcelles sera peut-être limitante dans l'évaluation de la qualité des résultats.

## Remerciements

Cette étude a été financée par la Commission Européenne (programme Alpine Space, projet 2-3-2-FR NEWFOR) et par l'Agence Nationale de la Recherche (projet ANR-2010-BIOE-008 FORESEE). Merci à l'Office National des Forêts et au projet BGF (Biodiversité et Gestion Forestière) pour les données de Prénovel-Les Piards ainsi qu'à Catherine Riond et Éric Mermin pour la collecte des données d'inventaire. Les auteurs remercient également deux relecteurs anonymes dont les remarques ont contribué à améliorer le contenu de cet article.

## Références

- Breidenbach, J., Astrup, R., 2012. Small area estimation of forest attributes in the norwegian national forest inventory. *European Journal of Forest Research* 131 (4), 1255–1267.
- Duplat, P., Perrotte, G., 1981. Inventaire et estimation de l'accroissement des peuplements forestiers. Office National des Forêts.
- Gonzalez, R. C., Woods, R. E., 2008. *Digital Image Processing*, 3rd Edition. Prentice Hall, Upper Saddle River, NJ.
- Hollaus, M., Dorigo, W., Wagner, W., Schadauer, K., Höfle, B., Maier, B., 2009. Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data. *International Journal of Remote Sensing* 30 (19), 5159–5175.
- Khosravipour, A., Skidmore, A., Isenburg, M., Wang, T., Hussin, Y., 2013. Development of an algorithm to generate a lidar pit-free canopy height model. Dans : *Proceedings of the 13th Silvilaser conference*, Beijing, China.
- Magnussen, S., Naesset, E., Gobakken, T., Frazer, G., 2012. A fine-scale model for area-based predictions of tree-size-related attributes derived from lidar canopy heights. *Scandinavian Journal of Forest Research* 27 (3), 312–322.
- McRoberts, R. E., mai 2010. Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. *Remote Sensing of Environment* 114 (5), 1017–1025.
- Monnet, J.-M., Chirouze, É., Mermin, É., 2015 (soumis). Estimation de paramètres forestiers par données lidar aéroporté et imagerie satellitaire rapideye : étude de sensibilité. *Revue Française de Photogrammétrie et Télédétection*.
- Monnet, J.-M., Mermin, É., 2014. Cross-correlation of diameter measures for the co-registration of forest inventory plots with airborne laser scanning data. *Forests* 5 (9), 2307–2326.

- Munoz, A., Bock, J., Monnet, J.-M., Renaud, J.-P., Jolly, A., Riond, C., 2015 (accepté). Évaluation par validation indépendante des prédictions des paramètres forestiers réalisées à partir de données LiDAR aéroporté. *Revue Française de Photogrammétrie et Télédétection*.
- Næsset, E., Apr. 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 (2), 164–179.
- Pardé, J., Bouchon, J., 1988. *Dendrométrie*, 2ème Edition. EN-GREF Nancy.
- Strunk, J., Temesgen, H., Andersen, H. E., Flewelling, J. P., Madsen, L., octobre 2012. Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing* 38 (5), 644–654.
- White, J., Wulder, M., Varhola, A., Vastaranta, M., Coops, N.C. and Cook, B., Pitt, D., Woods, M., 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre, Victoria, BC. Information Report FI-X-010.