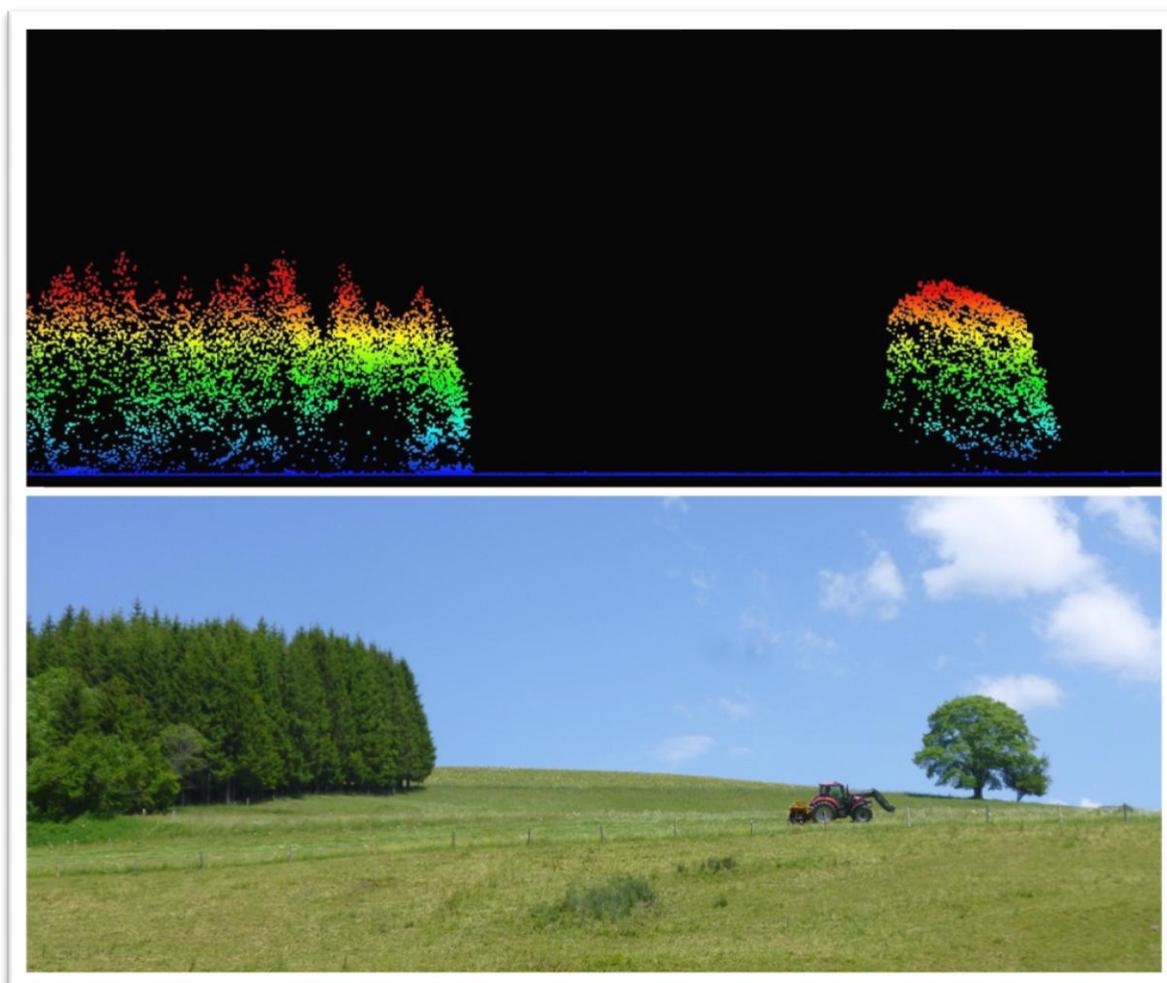


## **Prédiction de la structure des peuplements forestiers à partir de données LiDAR aéroporté**



Mémoire de fin d'études  
Dominante d'approfondissement Gestion Forestière

**Illustration de couverture :**

*Lisière forestière dans le massif du Vercors (38)*

Visualisation de données LiDAR : logiciel MARS 7 - Source de la photographie : Claire Bastick

## FICHE SIGNALÉTIQUE D'UN TRAVAIL D'ÉLÈVE AgroParisTech

Formation des ingénieurs forestiers d'AgroParisTech	<b>TRAVAUX D'ÉLÈVES</b>
TITRE : Prédiction de la structure des peuplements forestiers à partir de données LiDAR aéroporté	Mots clés : LiDAR aérien, structure des peuplements, indice de Gini, classification supervisée, kMSN
AUTEUR(S) : Claire BASTICK	Promotion : 2014-2015
Caractéristiques : 1 volume ; 75 pages ; 34 figures ; 11 tableaux ; 7 annexes ; bibliographie.	

### CADRE DU TRAVAIL

**ORGANISME PILOTE OU CONTRACTANT : Office National des Forêts – Pôle RDI de Chambéry**

**Nom du responsable : Alain MUNOZ**

**Fonction : Spécialiste SIG, Géomatique, Télédétection**

**Nom du correspondant AgroParisTech : Eric LACOMBE**

Option  
D. d'approfondissement  
CEI

Stage 2A  
Stage fin d'études

Autre

Date de remise : octobre 2015

**SUITE À DONNER (réservé à la Direction des études)**

- Consultable et diffusable
- Confidentiel de façon permanente
- Confidentiel jusqu'au     /     /     , puis diffusable

## **Avertissement**

Le service RDI de l'ONF s'est rendu compte tardivement, après la fin du stage de Claire Bastick et la rédaction du présent rapport, que certaines placettes terrain du site feuillu du « PNF » n'avaient pas été inventoriées en plein. Bien que ces erreurs, imputables uniquement au service RDI de l'ONF, ne modifient en rien la grande qualité du travail réalisé par Claire Bastick, les résultats concernant le site feuillu du PNF (chapitre III.2 du rapport) risquent d'être erronés. Le service RDI de l'ONF présente ses excuses à Claire Bastick ainsi qu'aux lecteurs du présent rapport.

*Le service RDI de l'ONF*

## Résumé

Le LiDAR aéroporté est une technologie prometteuse pour déterminer les caractéristiques des peuplements forestiers. Cette étude s'est penchée sur le potentiel du LiDAR aéroporté pour prédire la structure des peuplements. Deux grands contextes forestiers français ont été choisis pour réaliser et tester les prédictions : les résineux de montagne (Alpes) et les feuillus de plaine (Bourgogne). La démarche a consisté en : (1) la caractérisation de la structure sur des placettes terrain de référence, (2) la sélection d'indicateurs LiDAR liés à la structure des peuplements, (3) la calibration et la validation de modèles permettant de distinguer les peuplements réguliers et irréguliers, (4) la réalisation d'une cartographie prédictive des types structuraux.

En fonction des méthodes utilisées et des contextes forestiers, ces prédictions ont permis de classer correctement entre 75% et 90% de la surface en peuplements réguliers ou irréguliers. Un gain significatif de précision et de robustesse serait possible, notamment par l'amélioration des méthodes d'inventaire de calibration. Compte-tenu de l'importance de la connaissance de la structure des peuplements, les cartes ainsi obtenues pourraient à l'avenir trouver des applications en aménagement et en gestion forestière.

## Abstract

Airborne laser scanning (ALS) is a promising technology to determine various parameters of forest stands. This study focused on the potential of ALS to predict forest structure. Two major forest contexts have been chosen to produce and test the predictions: mountain coniferous forest (Alps) and lowland deciduous forest (Burgundy). The approach consisted in: (1) characterizing the structure of reference plots, (2) selecting ALS indicators related to the forest structure, (3) calibrating and validating the models that discriminate even-aged and uneven-aged stands, (4) mapping the predictions of the structural types.

Depending on the methods and the forest contexts, between 75% and 90% of the surface was correctly classified in even-aged or uneven-aged stands by these predictions. A significant gain in accuracy and robustness would be possible, particularly thanks to an improvement of the inventory methods for the calibration plots. Considering the importance of the information about the forest structure, the resulting maps of these predictions could find applications for forest management issues in the future.

## Remerciements

Si ce stage a été pour moi enrichissant, c'est en grande partie grâce aux personnes que j'y ai rencontrées. Je tiens ici à les remercier pour leur aide et leurs conseils.

En premier lieu, un très grand merci à Alain Munoz, mon maître de stage, qui m'a guidée et m'a fait confiance tout au long de ce stage. Merci pour le temps passé à m'apprendre les techniques de traitement de données LiDAR et à discuter de l'avancement de mon travail, et ce malgré les autres projets en cours.

Merci également à Jérôme Bock pour ses remarques avisées et ses conseils réguliers, ainsi qu'à Jacques Fay et Emilie Dubois pour leur aide sur le terrain. Plus généralement, je tiens à remercier l'équipe et les stagiaires du pôle R&D de Chambéry pour leur accueil, leur convivialité et bonne humeur.

Mes remerciements vont également aux chercheurs du site de Nancy, avec qui les discussions ont été très intéressantes. Merci particulièrement à Alexandre Piboule pour ses explications sur *Segma* et *Computree* et à Jean-Pierre Renaud pour ses nombreux conseils bibliographiques et statistiques.

Enfin, je souhaite remercier Eric Lacombe, mon tuteur de stage, pour son suivi, ses recommandations et sa grande disponibilité.

# Sommaire

<b>Remerciements</b> .....	<b>1</b>
<b>Sommaire</b> .....	<b>2</b>
<b>Index des abréviations</b> .....	<b>4</b>
<b>Introduction</b> .....	<b>5</b>
<b>I. Eléments de contexte</b> .....	<b>6</b>
I.1. Le LiDAR et son utilisation en forêt.....	6
I.1.1. Principes de fonctionnement du LiDAR aéroporté.....	6
I.1.2. Applications forestières du LiDAR .....	7
I.2. La structure des peuplements forestiers .....	8
I.2.1. Définitions et intérêts de la structure .....	8
I.2.2. Indices de structure des peuplements .....	8
I.2.3. Typologies et structure des peuplements .....	9
I.3. Revue des techniques de détermination de la structure à partir de données LiDAR aéroporté.....	10
I.3.1. Segmentation a priori.....	11
I.3.2. Classification supervisée.....	11
I.3.3. Prédiction de variables structurales quantitatives .....	11
I.3.4. Prédiction de la répartition des diamètres.....	12
I.4. Bilan : démarche de travail .....	14
<b>II. Matériel et méthodes</b> .....	<b>15</b>
II.1. Données et sites d'étude .....	15
II.1.1. Description des sites étudiés .....	15
II.1.2. Informations sur les vols LiDAR .....	16
II.1.3. Données de référence de terrain .....	17
II.2. Caractérisation de la structure sur les données terrain .....	17
II.2.1. Modes de la répartition.....	18
II.2.2. Utilisation du coefficient de Gini .....	18
II.2.3. Typologie du Massif Vosgien .....	19
II.2.4. Jeu de placettes de calibration.....	19
II.3. Calcul de métriques LiDAR.....	20
II.3.1. Métriques issues du nuage de points.....	20
II.3.2. Métriques issues du MNH .....	21
II.3.3. Métriques issues de la segmentation des couronnes .....	22
II.4. Sélection des métriques LiDAR .....	23
II.4.1. Sélection visuelle et par régression linéaire .....	23
II.4.2. Sélection de variables par l'algorithme VSURF.....	24
II.4.3. Sélection finale .....	24
II.5. Méthodes de prédiction de la structure .....	25
II.5.1. Les k plus proches voisins (kNN).....	25
II.5.2. L'analyse discriminante linéaire (LDA) .....	25
II.5.3. Support vector machine (SVM) .....	26
II.5.4. Random Forest (RF) .....	27
II.5.5. Les k voisins les plus similaires (kMSN) .....	27

II.6. Production de cartes de prédiction .....	28
II.7. Validation des modèles de prédiction .....	29
II.7.1. Validation croisée .....	29
II.7.2. Tests de robustesse sur des sites résineux de montagne .....	30
II.7.3. Validation indépendante sur le terrain .....	30
II.8. Exemple d'application des modèles : vers une cartographie opérationnelle des types structuraux .....	31
<b>III. Résultats et analyses .....</b>	<b>34</b>
III.1. Résineux de montagne .....	34
III.1.1. Structure des placettes de terrain .....	34
III.1.2. Métriques sélectionnées .....	34
III.1.3. Validation et erreurs de prédiction .....	37
III.1.4. Prédications sur l'ensemble du site d'étude .....	40
III.1.5. Vers une cartographie opérationnelle des types structuraux .....	42
III.2. Feuillus de plaine .....	42
III.2.1. Structure des placettes de terrain .....	42
III.2.2. Métriques sélectionnées .....	43
III.2.3. Validation et erreurs de prédiction .....	45
III.2.4. Prédications sur l'ensemble du site d'étude .....	46
<b>IV. Discussion et perspectives .....</b>	<b>47</b>
IV.1. Limites de l'étude .....	47
IV.1.1. Surface de perception de la structure des peuplements .....	47
IV.1.2. Répartition des placettes de calibration .....	47
IV.1.3. Segmentation des apex pour les feuillus .....	48
IV.2. Evolutions envisageables .....	49
IV.2.1. Amélioration des métriques LiDAR .....	49
IV.2.2. Utilisation de photos aériennes ou satellitaires .....	49
IV.2.3. Application de la méthode kMSN pour la prédiction de la répartition des tiges .....	50
IV.3. Perspectives d'utilisation .....	51
IV.3.1. Intérêts pour l'aménagiste et le gestionnaire .....	51
IV.3.2. Amélioration d'autres modélisations .....	51
<b>Conclusion .....</b>	<b>53</b>
<b>Bibliographie .....</b>	<b>54</b>
<b>Liste des figures et des tableaux .....</b>	<b>57</b>
<b>Annexes .....</b>	<b>59</b>

## Index des abréviations

**CG 73** : Conseil Général de la Savoie

**G** : Surface terrière

**GNSS** : Global Navigation Satellite System (système de positionnement par satellites)

**IGN** : Institut national de l'information géographique et forestière

**IRSTEA** : Institut national de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture

**kMSN** : k-Most Similar Neighbors (k voisins les plus similaires)

**kNN** : k-Nearest Neighbors (k plus proches voisins)

**LDA** : Linear Discriminant Analysis (analyse discriminante linéaire)

**LiDAR** : Light Detection And Ranging

**MNH (CHM, MHC)** : Modèle Numérique de Hauteur

**MNT** : Modèle Numérique de Terrain

**ONF** : Office National des Forêts

**OOB** : Out-Of-Bag

**PNF** : Parc National des Forêts de Champagne et de Bourgogne

**R<sup>2</sup>** : Coefficient de détermination

**R&D** : Recherche et Développement

**RF** : Random Forest (forêts aléatoires)

**RMSE** : Root Mean Square Error (erreur quadratique moyenne)

**SVM** : Support Vector Machine (machines à vecteurs de support)

**TDS** : Terminal de saisie

**TIN** : Triangulated Irregular Network (réseau de triangles irréguliers)

**VSURF** : Variable Selection Using Random Forest (sélection de variables avec la méthode des forêts aléatoires)

**PB** : Petits Bois (de 17,5 à 27,5 cm de diamètre)

**BM** : Bois Moyens (de 27,5 à 47,7 cm de diamètre)

**GB** : Gros Bois (plus de 47,5 cm de diamètre)

## Introduction

La structure des peuplements est un élément essentiel dans la compréhension et la gestion des écosystèmes forestiers. Classiquement décrite de façon empirique à dire d'expert ou à l'aide de typologies sur des inventaires ponctuels, cette structure oriente fortement les choix sylvicoles. La cartographie des types de peuplements est ainsi un élément primordial des inventaires et des aménagements forestiers. La photo-interprétation permet déjà d'aider à la délimitation de peuplements variés, mais l'apport de données LiDAR pourrait bien préciser et faciliter cette cartographie.

Depuis une dizaine d'années, les recherches ont montré tout le potentiel et l'intérêt du LiDAR aéroporté pour la caractérisation des milieux forestiers. Fournissant une information tridimensionnelle de la forêt, cette technologie permet de prédire un certain nombre de paramètres dendrométriques utiles au forestier. Ainsi, des modèles précis et des cartographies exhaustives ont déjà été obtenus pour des variables comme la hauteur dominante ou la surface terrière. En outre, dans certains pays de plus en plus d'inventaires et d'aménagements se basent sur ces données.

En lien avec ce développement récent, ce stage a pour objectif de distinguer et de prédire des grands types structuraux (notamment structure irrégulière et structure régulière) à partir de données LiDAR. Ce travail nécessite donc d'une part la caractérisation de ces types structuraux, dont la terminologie n'est pas universelle, et d'autre part l'utilisation de méthodes statistiques adaptés à ce type de prédiction.

L'intérêt opérationnel de cette étude est de fournir une cartographie des structures de peuplements, qui à terme pourrait avoir des applications en aménagement forestier, tant au niveau de la stratification d'inventaires que de la planification des travaux sylvicoles et des récoltes. C'est d'ailleurs dans cette optique de développement d'outils répondant aux besoins des gestionnaires que s'inscrit le pôle R&D de l'ONF, structure d'accueil de ce stage.

Ce rapport présente le travail réalisé en commençant tout d'abord par la recherche bibliographique qui a permis de mettre en place une démarche de prédiction de la structure. Le détail de la méthode suivie et l'analyse des résultats des modèles statistiques y sont exposés par la suite.

# I. Eléments de contexte

## I.1. Le LiDAR et son utilisation en forêt

### I.1.1. Principes de fonctionnement du LiDAR aéroporté

Le LiDAR aéroporté (Light Detection and Ranging) est un système de télédétection actif, c'est à-dire basé sur l'émission et la réception d'un signal. Cette technologie récente permet de représenter les éléments de la portion de surface terrestre scannée sous forme d'un nuage de points en trois dimensions (3D).

Embarqué à bord d'un aéronef, l'émetteur envoie une impulsion laser qui est réfléchiée par les différents objets rencontrés sur sa trajectoire (branches, sol...). Les retours ou échos de cette impulsion sont alors enregistrés par le récepteur et le temps mis par chaque écho pour faire l'aller-retour permet de calculer la distance entre l'objet et l'appareil. L'avion étant équipé d'un récepteur GNSS et d'une centrale inertielle qui prend en compte les mouvements de l'avion, les coordonnées en x, y et z de l'impact du signal avec l'objet peuvent être obtenues (cf. Figure 1). Les précisions de ces positions varient alors entre 20 et 30 cm en planimétrie (xy) et entre 15 et 20 cm en altimétrie (z) (Bonnet et al., 2013).

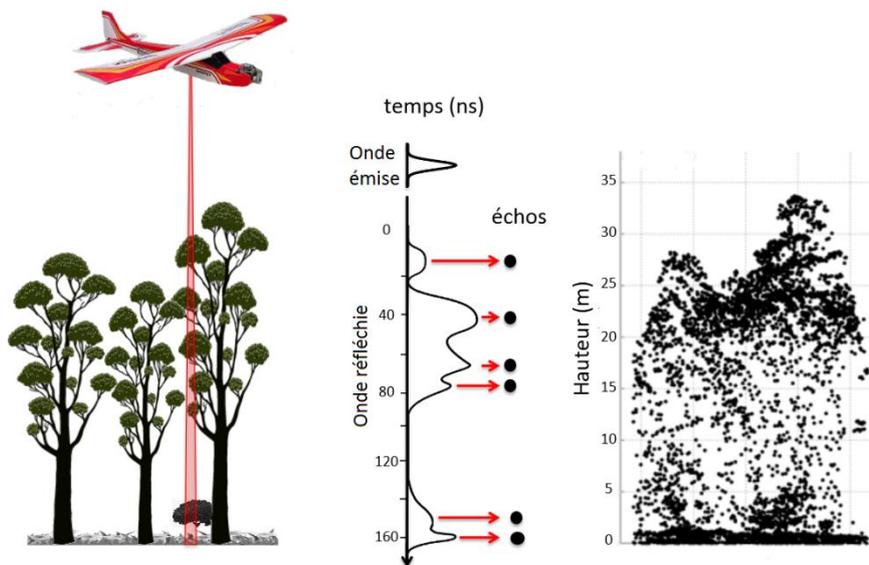


FIGURE 1 : FONCTIONNEMENT DU LiDAR AEROPORTE (D'APRES ANTONIO FERRAZ)

Le nuage de points ainsi obtenu est ensuite traité afin de différencier les points en fonction de leur nature (sol, végétation, bâtiments, etc.). Cette classification permet notamment de générer un Modèle Numérique de Terrain (MNT) représentant le relief exact à partir des points « sol ». Les données du nuage de points sont généralement fournies sous forme d'un fichier au format « las » contenant au moins les coordonnées des points, leurs classifications, leurs intensités, leurs numéros de retour (White et al., 2013). De ces données, peut également être déduit un Modèle Numérique de Hauteur (MNH) représentant la différence d'altitude entre la surface des objets (canopée) et le sol.

A la différence du système LiDAR à retour discret présenté ci-dessus, certains capteurs enregistrent l'onde retour complète avec la variation d'intensité lumineuse associée. Ce système, globalement

moins utilisé pour des applications forestières, a généralement une plus large empreinte que les systèmes à retour discret (Lefsky et al., 2009).

### I.1.2. Applications forestières du LiDAR

Depuis quelques années, l'utilisation du LiDAR s'est considérablement développée dans différents domaines. Les applications forestières de cette nouvelle technologie sont particulièrement nombreuses. La capacité du LiDAR à donner une information précise sur la hauteur des peuplements et la répartition verticale du couvert forestier offre des possibilités importantes en termes de modélisation des paramètres forestiers. Dans certains pays, notamment scandinaves, le LiDAR est ainsi couramment utilisé dans le cadre d'inventaires forestiers (Maltamo et al., 2014).

Deux approches sont généralement adoptées pour prédire ces paramètres : **l'approche « arbre » et l'approche « surfacique »**. L'approche « arbre » consiste à identifier chaque individu à l'aide d'un algorithme, puis, en appliquant une relation allométrique, à déterminer les paramètres forestiers. L'inconvénient de cette méthode est que le risque d'oubli d'arbres peut être important, notamment en sous-étage, et qu'elle requiert une densité d'émission supérieure à 4 pt/m<sup>2</sup> (White et al., 2013). L'approche surfacique, quant à elle, met en relation une valeur mesurée sur le terrain avec des indicateurs LiDAR (ou métriques) calculés sur la même emprise que les placettes de référence terrain. Ces métriques peuvent être dérivées directement du nuage de points LiDAR ou calculées à partir d'une individualisation préalable des arbres sur le MNH. Dans cette approche, des informations dendrométriques de terrain sont nécessaires à la calibration (Monnet, 2010).

Il est ainsi possible d'obtenir de bonnes prédictions pour des paramètres forestiers comme la hauteur dominante, très directement liée aux données brutes, mais également pour la surface terrière, le volume, le diamètre moyen ou dominant et dans une moindre mesure la densité. Dans la littérature, les erreurs de prédiction varient de 2 à 7 % pour la hauteur dominante, de 10 à 30 % pour la surface terrière et de 15 à 50 % pour la densité de tiges (Heurich et Thoma, 2008; Næsset, 2004; Treitz et al., 2012; Woods et Lim, 2008).

Ces prédictions sont alors meilleures avec le LiDAR qu'avec d'autres systèmes de télédétection (Lefsky et al., 2001a), et même parfois qu'avec certains inventaires « terrain » à l'exemple de la hauteur ou de la surface terrière (Bock et al., 2014).

En plus de la prédiction de ces paramètres forestiers, d'autres applications du LiDAR ont été développées pour la gestion forestière. L'utilisation du MNT LiDAR d'un pas très fin, permet entre autre de réaliser des inventaires de la desserte forestière (cf. Figure 2) ou d'aider à l'implantation de lignes de câble ou de routes forestières (Munoz et al., 2013). Par ailleurs, des études ont été menées sur le LiDAR par rapport à d'autres enjeux forestiers, comme par exemple la biodiversité avec la détection d'arbres morts ou encore les risques avec la modélisation de la biomasse inflammable.

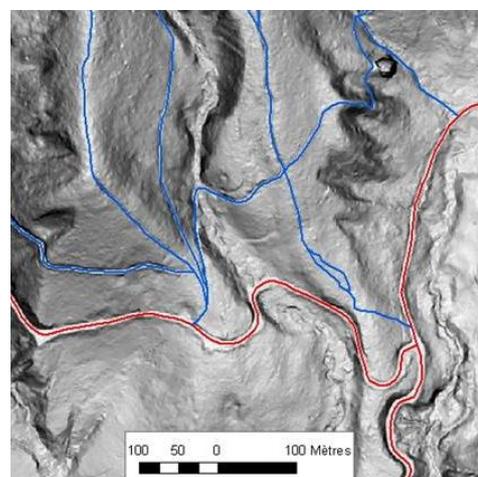


FIGURE 2 : EXEMPLE DE DETECTION DE DESSERTE FORESTIERE A PARTIR DU MNT (SOURCE : ONF R&D)

## I.2. La structure des peuplements forestiers

### I.2.1. Définitions et intérêts de la structure

La structure des écosystèmes forestiers est généralement appréhendée sous trois aspects différents : la distribution spatiale des individus, la diversité des essences et la répartition en diamètre ou hauteur des arbres (Pommerening, 2002). Cette structure de la végétation est un facteur déterminant du fonctionnement écologique des peuplements forestiers. Ainsi une grande hétérogénéité structurale est généralement liée à une diversité d'habitats et d'espèces (Pommerening, 2002). La connaissance de la structure des peuplements permet également de prédire la dynamique des peuplements ou encore d'évaluer la fonction protectrice de la forêt.

Le type structural d'un peuplement est aussi une information essentielle à la gestion forestière. En effet, de ce type structural et des itinéraires sylvicoles choisis dépendront les règles de gestion appliquées au peuplement. Dans ce domaine, c'est classiquement la **répartition des tiges par catégorie de diamètre** qui définit la structure (Bouchon, 1979). C'est donc sur cette définition de la structure qu'est basée cette étude. Dans ce cadre, on distingue deux grands types de structure. La structure régulière, où la répartition des diamètres correspond à une courbe en cloche (unimodale) et la structure irrégulière, où la répartition n'est au contraire pas unimodale (Bastien et Gauberville, 2011).

### I.2.2. Indices de structure des peuplements

Afin de décrire de façon quantitative la structure forestière et son hétérogénéité, des indices variés ont été développés. De nombreuses études (Cordonnier et al., 2012; Lexerød et Eid, 2006; Valbuena et al., 2012) ont eu pour objectif de les classer et de les comparer. Certains de ces indices sont plus influencés par la gamme des diamètres, d'autres par l'abondance des différentes classes de diamètre.

La plupart de ces indices sont issus de l'écologie des espèces et ont été adaptés à la mesure de l'hétérogénéité des diamètres en y remplaçant les espèces par les classes de diamètre, et le nombre d'individus par le nombre de tiges ou la surface terrière. L'un des plus populaires est l'indice de Shannon. D'autres indices sont issus de mesures statistiques comme l'écart-type, le coefficient de variation des diamètres ou encore l'indice *Stand Variance Index* (STVI) basé sur la variance de la distribution en surface terrière (LeMay et Staudhammer, 2005). Enfin, le coefficient de Gini est à l'origine utilisé en économie, où il permet d'évaluer les inégalités de revenus.

En gestion forestière, le plus souvent appliqué à la surface terrière, le **coefficient de Gini est un bon indicateur de la structure des peuplements**. En effet, il discrimine particulièrement bien les peuplements et se révèle très peu sensible à la taille de l'échantillon (Lexerød et Eid, 2006; Valbuena et al., 2012). Cet indice est basé sur la courbe de Lorenz, qui correspond à la courbe de la proportion cumulée de surface terrière en fonction de la proportion cumulée du nombre d'arbres, lorsque ces arbres sont ordonnés par surface terrière croissante (cf. Figure 3).

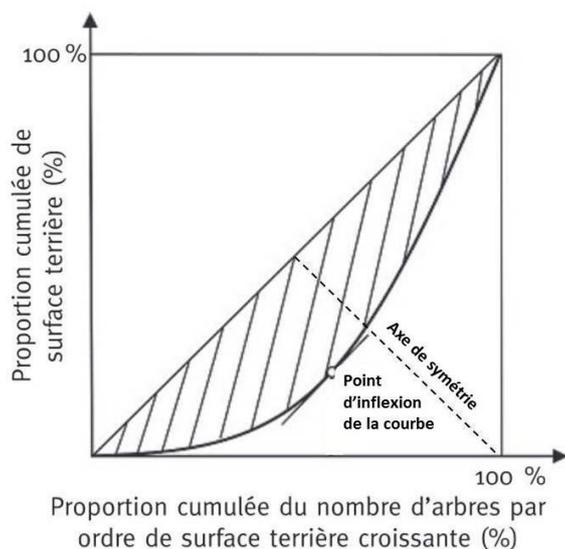


FIGURE 3 : COURBE DE LORENZ ET DETERMINATION DE L'INDICE DE GINI (D'APRES CORDONNIER ET AL., 2012)

La première bissectrice correspond à la diagonale d'égalité parfaite (cas où tous les arbres auraient exactement la même surface terrière). **Plus un peuplement est hétérogène**, plus la courbe de Lorenz s'éloigne de cette diagonale et **plus son coefficient de Gini est élevé**. En effet, le coefficient de Gini correspond à deux fois la surface entre la diagonale d'égalité parfaite et la courbe de Lorenz (soit deux fois la surface hachurée sur la Figure 3). Mathématiquement, l'expression du coefficient de Gini appliqué à la surface terrière est alors :

$$Gini = 2 \frac{\sum_{i=1}^N i g_i}{NG} - \frac{N + 1}{N}$$

avec  $g_i$  correspondant à la surface terrière de l'arbre de rang  $i$ ,  $N$  le nombre de tiges du peuplement et  $G$  la surface terrière totale du peuplement (Cordonnier et al., 2012). Le coefficient de Gini reste donc théoriquement borné entre 0 et 1.

Cependant, pour des peuplements à l'hétérogénéité identique et au coefficient de Gini égal, différentes structures et différentes formes pour la courbe de Lorenz peuvent exister. Pour les différencier, un deuxième coefficient, l'asymétrie de Lorenz, a été développé. Ce coefficient correspond à la position relative du point d'inflexion de la courbe de Lorenz (le point où la pente de la courbe de Lorenz est parallèle à la diagonale d'égalité parfaite) par rapport à l'axe de symétrie (en pointillé sur la Figure 3). Si ce point est au-dessus de l'axe de symétrie, l'asymétrie de Lorenz est supérieure à 1, ce qui peut être causé par quelques gros arbres présents dans le peuplement. Dans le cas contraire, l'asymétrie est inférieure à 1, ce qui peut être le cas s'il y a un nombre important de petits arbres dans le peuplement (Damgaard et Weiner, 2000).

### I.2.3. Typologies et structure des peuplements

En gestion forestière, la question des types de peuplements est souvent appréhendée via une typologie. Ces typologies, nées du besoin de clarifier certains termes forestiers et d'adopter un langage commun pour décrire des peuplements, ne sont cependant pas universelles, variant selon les régions et les forêts (Bruciamacchie, 2001). Dans ces classifications, la structure du peuplement reste la plupart du temps la clé d'entrée dans la différenciation des types de peuplements, avant la composition en essences ou même le capital.

Dans cette diversité de typologies, la majorité se base sur la répartition en catégorie de grosseur des peuplements (pourcentage de petits bois PB, de bois moyens BM, de gros bois GB). Cette entrée par catégorie de grosseur peut être exprimée en pourcentage du nombre de tiges (Abt, 2014; Gaudin et Jenner, 2001) ou en pourcentage de la surface terrière (Asael et al., 1999; Bachelet et al., 2002). Une représentation graphique de ces typologies est le triangle des structures (cf. Figure 4), dans lequel les peuplements irréguliers se situent vers le centre.

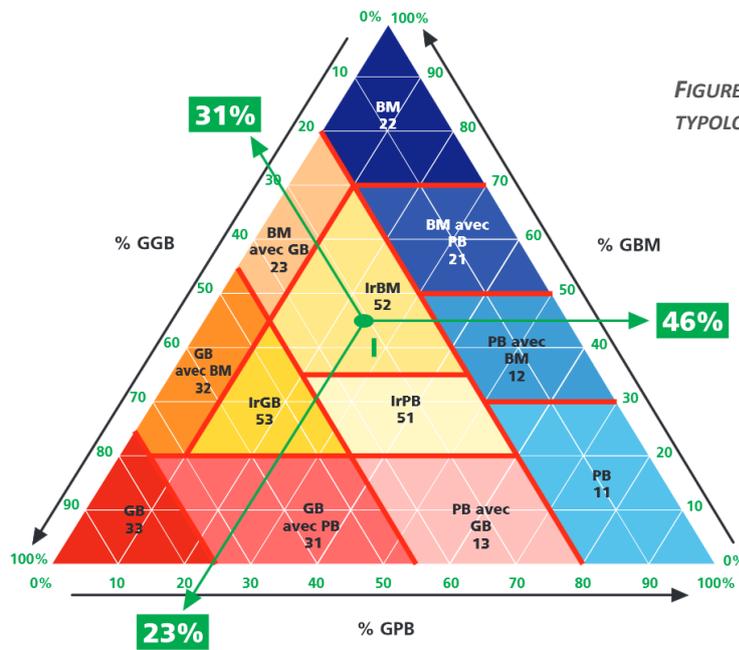


FIGURE 4 : TRIANGLE DES STRUCTURES DE LA TYPOLOGIE DU PLATEAU LORRAIN (SOURCE : CRPF LORRAINE-ALSACE)

Les typologies basées sur la structure verticale des peuplements sont rares. Du fait de l'importance de la fonction de protection des forêts de montagne, seul le massif alpin est doté d'une typologie basée sur la stratification en hauteur (Cemagref et al., 2006).

### I.3. Revue des techniques de détermination de la structure à partir de données LiDAR aéroporté

La structure forestière étant un élément important pour la gestion forestière et la compréhension des écosystèmes forestiers, la détermination de celle-ci à partir de données LiDAR a été étudiée par plusieurs chercheurs dans différents contextes. Cette partie, complétée par le tableau de synthèse en Annexe 1, fournit un aperçu des études menées sur ce sujet. Les outils statistiques évoqués sont détaillés dans la suite du rapport (cf. § II.5.).

Globalement 4 grandes approches sont utilisées pour déterminer cette structure :

- La **segmentation a priori** de types de structures à partir du nuage de points 3D
- La **classification supervisée** basé sur un apprentissage des structures issues du terrain
- La **prédiction de variables de structure** quantitatives à partir d'une calibration sur le terrain
- La prédiction complète de l'histogramme de **répartition des diamètres**

### **I.3.1. Segmentation a priori**

Le LiDAR fournissant une représentation 3D des peuplements, certains indicateurs issus du nuage de points peuvent directement donner une idée de la structure, ou du moins de la stratification verticale des peuplements. Ainsi, des indices tels que la rugosité de la canopée issue du MNH ou la densité de points à différentes strates de hauteur peuvent être directement issus du LiDAR et permettre de différencier, grâce à un seuillage, certains types de peuplements (Lesur, 2011).

Des méthodes statistiques de classification non supervisée peuvent également être utilisées sur des indicateurs pertinents pour différencier les structures forestières. Dans cette optique, Dickinson et al. (2014) ainsi que Pascual et al. (2008) ont réalisé des classifications en utilisant la méthode des *k-means*, méthode de partitionnement courante permettant de répartir automatiquement les individus en *k* groupes.

La segmentation peut également être réalisée grâce à des algorithmes spécifiques, permettant de définir des frontières entre zones homogènes selon des variables issues du LiDAR ou combinées avec de la photo aérienne (Leppänen et al., 2008).

### **I.3.2. Classification supervisée**

Afin d'obtenir des types structuraux correspondant aux classifications utilisées sur le terrain, une méthode d'apprentissage peut ensuite être appliquée à ces premières segmentations (Tiede et al., 2004). Sullivan (2008) a ainsi utilisé le logiciel *Spring* pour réaliser une classification orientée objet suivie d'un apprentissage afin de délimiter six grands types de peuplement aux Etats-Unis, obtenant des taux de bonne classification compris entre 80 et 90%.

Plus couramment, la classification supervisée est directement appliquée sur les données sans passer par une segmentation préalable. En utilisant une méthode de classification supervisée classique, l'analyse discriminante linéaire (LDA), Zhang et al. (2011) ont obtenu une précision de classement de 91.4% sur 5 grands types de forêts australiennes.

Les apprentissages basés sur des arbres de décision sont également des méthodes de classification courantes. Chirici et al. (2013) ont appliqué ces méthodes à la prédiction de types de biomasse inflammable dans les forêts de Sicile et ont obtenu des taux de bonnes classification entre 45 % (méthode *Random Forest*) et 84 % (méthode *Stochastic gradient boosting*).

Dans un contexte de forêt alpine et avec une typologie basée sur la répartition par catégorie de grosseur des arbres, Torresan et al. (2014) ont aussi utilisé des méthodes de classification supervisées pour classer les peuplements de leur zone d'étude selon les 11 groupes de cette typologie. Contrairement aux études précédentes, qui utilisaient des métriques issues de l'ensemble du nuage de points LiDAR, Torresan et al. se sont servis exclusivement de métriques issues du MNH. Que ce soit avec une classification par LDA ou par *Random Forest*, les précisions des classifications étaient alors moyennes, respectivement 41% et 37%.

### **I.3.3. Prédiction de variables structurales quantitatives**

#### **I.3.3.1. Prédiction du pourcentage par catégorie de grosseur**

Un autre volet de l'étude de Torresan et al. (2014) a consisté à prédire la surface terrière par catégorie de grosseur pour en déduire ensuite le type de peuplement. Différentes méthodes de

prédiction ont été utilisées, basées d'une part sur la moyenne des k plus proches voisins (kNN) et d'autre part sur des régressions linéaires et non linéaires. Avec des taux de bonnes classifications entre 26 et 37 %, les résultats étaient alors légèrement inférieurs aux précédents.

### **I.3.3.2. Prédiction du coefficient de Gini**

La prédiction de la structure peut également passer par la prédiction d'indices structuraux. Etant donné que ces indices sont des variables quantitatives, une variété de méthodes statistiques existe pour les prédire (régressions linéaires, plus proches voisins, *Random Forest*). Ainsi, pour l'exemple du calcul de l'indice de Gini, Valbuena et al. (2014) ont utilisé plusieurs stratégies comme la prédiction directe de l'indice, la prédiction de centiles de la courbe de Lorenz, ou encore la prédiction d'une liste des diamètres des arbres. Ces stratégies combinées aux deux approches de traitement des données LiDAR (arbre ou surfacique) ont révélé une meilleure performance de l'approche surfacique. En fonction de cette valeur prédite du coefficient de Gini, les peuplements peuvent ensuite être classés en types structuraux (Valbuena et al., 2013b).

### **I.3.3.3. Métriques LiDAR utilisées**

De ces études, certaines métriques LiDAR se sont révélées être de bons prédicteurs des variables structurales. Ainsi, que ce soit avec le MNH (Torresan et al., 2014) ou directement avec le nuage de points (Valbuena et al., 2014), le coefficient de variation des hauteurs, les pourcentages de points (ou de pixels) à différentes tranches de hauteur et les centiles de hauteurs sont les variables ayant servi à ces prédictions. Une autre étude a également montré le rôle primordial de ces trois types de métriques LiDAR dans la prédiction du coefficient de Gini (Valbuena et al., 2013a).

## **I.3.4. Prédiction de la répartition des diamètres**

La prédiction de la répartition en diamètre permet non seulement de décrire la structure du peuplement, mais également de déduire différentes caractéristiques dendrométriques (diamètre moyen, volume, densité...). Ce type de modélisation a fait l'objet de nombreuses études (Maltamo et al., 2014). De nouveau, différentes méthodes existent et peuvent être comparées, certaines se basant sur des lois de distribution, d'autres prédisant directement la répartition (Kangas et Maltamo, 2000a).

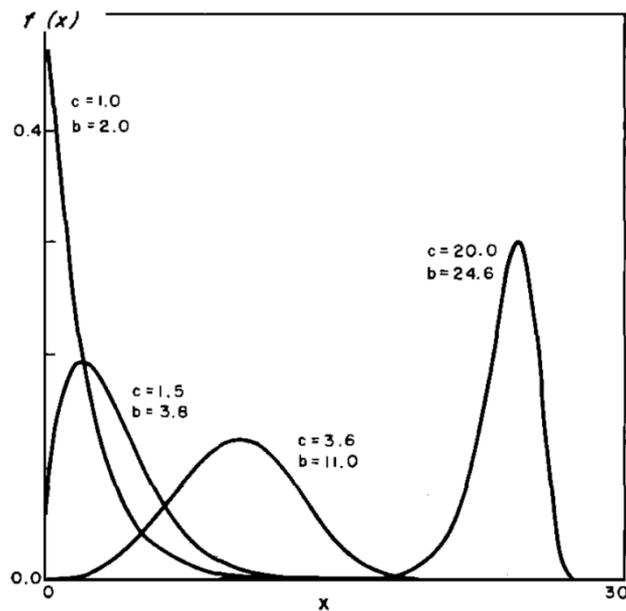
### **I.3.4.1. Lois de distribution**

Certaines fonctions de densité de probabilité peuvent s'ajuster aux répartitions diamétrales des peuplements forestiers. La loi de Weibull est souvent utilisée, du fait de sa forme flexible (Bailey et Dell, 1973). En effet, suivant la valeur de son paramètre de forme, elle peut prendre une forme « en J-inverse » convenant bien à certains peuplements irréguliers ou une forme en « cloche », convenant mieux aux peuplements réguliers. A deux paramètres, la densité de probabilité de la loi de Weibull s'écrit :

$$f(x) = \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} \exp\left[-\left(\frac{x}{b}\right)^c\right]$$

avec  $c$  le paramètre de forme et  $b$  le paramètre d'échelle (cf. Figure 5)

FIGURE 5 : DENSITE DE PROBABILITE DE LA DISTRIBUTION DE WEIBULL A DEUX PARAMETRES (D'APRES BAILEY ET DELL, 1973)



Cette loi ne s'adaptant pas aux répartitions bimodales, certains auteurs utilisent un modèle de mélange fini de distributions pour modéliser ce genre de répartition (Liu et al., 2014; Zhang et al., 2001). Ce modèle combinant deux densités de probabilité de Weibull s'écrit alors :  $g(x) = a_1f(x) + a_2f(x)$

Passer par ces lois de distribution pour prédire la répartition des diamètres, nécessite de déterminer les paramètres de la loi. Deux méthodes paramétriques de prédiction sont alors classiquement utilisées : on peut directement prédire ces paramètres (Thomas et al., 2008) ou bien prédire certains centiles ou moments de la répartition des diamètres pour ensuite retrouver les paramètres de la loi correspondant à cette répartition (Burk et Newberry, 1984; Maltamo et al., 2014).

#### I.3.4.2. Distribution libre

Afin de s'affranchir de l'ajustement d'une loi de distribution, il est possible de prédire des centiles de la répartition des diamètres à partir des données LiDAR, puis d'en interpoler la répartition complète (Kangas et Maltamo, 2000b). Cette prédiction peut alors s'effectuer via des régressions, mais la méthode des  $k$  voisins les plus similaires (kMSN) donne de meilleurs résultats (Bollandsås et al., 2013).

Cette méthode kMSN permet également d'attribuer directement à une placette cible une répartition diamétrale (ou une liste d'arbres et de diamètres) à partir des listes d'arbres des placettes de référence (placettes de calibration) les plus similaires (Maltamo et al., 2009).

#### I.3.4.3. Approche « surfacique » et approche « arbre »

Toutes ces prédictions sont basées sur une approche surfacique et sur des métriques « classiques » issues du nuage de points LiDAR (statistiques simples de la répartition en hauteur : centiles, moyennes, coefficients de variation, densité de points à différentes hauteurs, pourcentage de points de végétation).

Cependant, grâce aux relations allométriques qui existent entre diamètre et hauteur des arbres, l'approche « arbre » et la détection des apex peuvent fournir quasi directement une estimation de la répartition des diamètres (Peuhkurinen et al., 2007). Cette approche entraînant une très nette sous-détection des arbres du sous-étage, la répartition obtenue sera donc biaisée. Pour remédier à ce problème, Maltamo et al. (2004) ont combiné cette approche allométrique pour la prédiction des diamètres des gros arbres à une prédiction via une loi de Weibull pour les plus petits arbres en-dessous d'un seuil de diamètre. De façon semblable et en se basant plutôt sur le rayon des couronnes détectées, Vauhkonen et Mehtätalo (2014) ont prédit cette répartition diamétrale grâce au lien existant entre la fonction cumulative des diamètres et la fonction cumulative des couronnes (celle-ci étant corrigée pour tenir compte de l'oubli de petits arbres).

#### **I.4. Bilan : démarche de travail**

Que ce soit pour la caractérisation de la structure de peuplements ou pour sa prédiction à partir de données LiDAR, nous venons de voir qu'il existait une variété de techniques disponibles. Les approches utilisées par les différents chercheurs, ont permis d'orienter la méthode utilisée dans cette étude.

En effet, l'objectif de cette étude étant d'utiliser les données LiDAR pour **distinguer et prédire des grands types structuraux (irrégulier et régulier dans un premier temps)**, certaines de ces techniques ont été appliquées dans deux contextes forestiers français contrastés : **des forêts feuillues de plaine et des forêts résineuses de montagne**.

Afin de s'ajuster sur des types traditionnellement utilisés et définis en gestion forestière classique, l'approche utilisée est une **calibration sur des données de terrain**. La démarche consiste alors à caractériser la structure des placettes de calibration, puis à les utiliser pour construire un modèle de prédiction basé sur des données LiDAR, et enfin à réaliser une cartographie de ces types structuraux.

## II. Matériel et méthodes

Le schéma de la Figure 6 présente les différentes étapes de la démarche suivie. Les méthodes utilisées aux différentes étapes sont détaillées dans les paragraphes qui suivent. Les traitements ont été pour l'essentiel effectués grâce aux logiciels *R* (R Core Team, 2014) et *ArcGis* (ESRI, 2010).

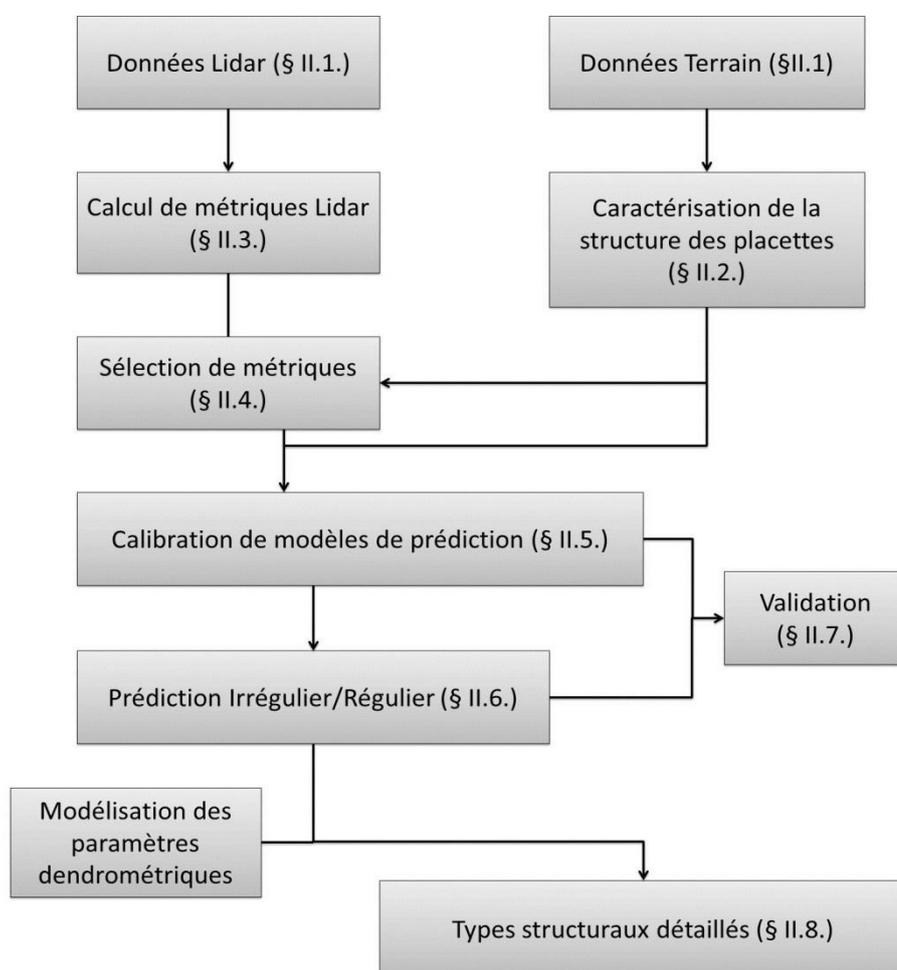


FIGURE 6 : DEMARCHE SUIVIE POUR LA PREDICTION DE LA STRUCTURE (RENOI AUX PARAGRAPHES CORRESPONDANTS)

### II.1. Données et sites d'étude

#### II.1.1. Description des sites étudiés

Deux grands contextes forestiers contrastés ont été étudiés (cf. Figure 7) :

- **1. Les forêts résineuses de montagne**, représentées par les sites des Quatre-Montagnes dans le **Vercors** (départ. 38), de la Combe d'**Aillon** dans le massif des Bauges (départ. 73) et de **Prénoval** dans le Jura (départ. 39). Ces forêts, souvent traitées en futaie irrégulière ou jardinée, sont essentiellement composées de sapin (*Abies alba*) et d'épicéa (*Picea abies*) parfois mélangés au hêtre (*Fagus sylvatica*). La pente et l'altitude y jouent un rôle important. Le site du Vercors, où la diversité des structures est la plus importante, a été utilisé pour calibrer des modèles de prédiction qui ont ensuite été appliqués aux sites de la Combe d'Aillon et de Prénoval permettant ainsi de tester la robustesse des prédictions.

- **2. Les forêts feuillues de plaine**, représentées par la forêt de Châtillon-sur-Seine située dans le futur Parc National des forêts de Champagne et Bourgogne (PNF, dép. 21). Cette forêt issue d'anciens taillis sous futaie est essentiellement composée de réserves de chêne (*Quercus sp.*) et de hêtre (*Fagus sylvatica*) et d'un sous-étage de charme (*Carpinus betulus*).



FIGURE 7 : CARTE DE LOCALISATION DES SITES D'ETUDE

Compte-tenu des différences majeures entre ces deux contextes forestiers, la méthode de modélisation a été appliquée indépendamment dans chacun de ces deux types de forêts.

### II.1.2. Informations sur les vols LiDAR

Le Tableau 1 ci-dessous récapitule les informations des vols LiDAR ayant eu lieu sur les sites d'études. Les sites de montagne ont des caractéristiques semblables en termes de hauteur de vol ou de densité d'émission, ce qui est important pour l'application des modèles lors des tests de robustesse.

TABLEAU 1 : INFORMATIONS DES VOLS LiDAR

Site	Vercors	Aillons	Prénovel	PNF
Date de vol	Août 2011 et septembre 2010	Août 2011	Septembre 2012	Mars 2012
Propriétaire des données	© IRSTEA	©CG 73	© IRSTEA	© PNF
Opérateur	Sintegra	Sintegra	Altoa	Sintegra
Surface du vol LiDAR	5653 ha	2500 ha	915 ha	38100 ha
Hauteur de vol	595 m	550 m	500 m	550 m
Densité d'émission LiDAR	11,3 pts/m <sup>2</sup>	11,7 pts/m <sup>2</sup>	9,3 pts/m <sup>2</sup>	14,0 pts/m <sup>2</sup>

### II.1.3. Données de référence de terrain

Dans ces différents sites, des placettes circulaires d'inventaire dendrométrique ont été réalisées à un ou deux ans d'intervalle du vol LiDAR. Ces placettes réalisées dans le cadre de révisions d'aménagements forestiers ou de projets de recherche, sont basées sur des protocoles d'inventaires qui diffèrent un peu les uns des autres. Afin de permettre la modélisation de la structure, les placettes ont donc été rendues comparables en termes de surface et d'arbres pris en compte. Les calculs se sont ainsi portés sur les **arbres vivants compris dans les 15 m de rayon et dont le diamètre était supérieur ou égal à 17,5 cm**. Pour les placettes imbriquées, le poids de chaque arbre a permis de recalculer le bon nombre de tiges sur la surface d'une placette de 15m de rayon. Les placettes ayant connu une perturbation importante (exploitation, chablis...) entre la date d'inventaire et la mission LiDAR ont été exclues des analyses. Le nombre total de placettes par site est disponible dans la présentation des résultats (cf. Tableau 3 et Tableau 8)

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 au mieux la position des arbres relevés sur le terrain avec les couronnes des arbres visibles sur le MNH LiDAR. Le recalage évite que le mauvais géoréférencement n'induisse des erreurs dans la modélisation (Monnet et Mermin, 2014).

## II.2. Caractérisation de la structure sur les données terrain

L'**objectif prioritaire** de cette étude étant de **prédire la structure régulière (taillis compris) ou irrégulière (jardinée et bimodale comprises) des peuplements**, il est nécessaire de distinguer les peuplements irréguliers et réguliers dans les placettes d'inventaire terrain qui serviront à la calibration.

Comme nous l'avons vu dans la bibliographie, il existe différentes approches pour appréhender la structure des peuplements. De même, il n'existe pas de frontière nette et bien définie entre des structures régulières ou irrégulières. Afin d'être le plus proche de la réalité possible, une combinaison de différentes approches a donc été utilisée dans cette étude :

- Les **modes de la répartition** des diamètres
- Le **coefficient de Gini** et l'asymétrie de Lorenz
- Une **typologie** basée sur la répartition par catégorie de grosseur

Une dernière approche de différenciation des peuplements irréguliers et réguliers, basée sur les paramètres de forme et d'échelle de la densité de probabilité de Weibull a finalement été abandonnée. En effet, l'ajustement de la loi ne convenait pas aux peuplements bimodaux et déformait certaines répartitions non typiques (cf. § I.3.4.1.).

## II.2.1. Modes de la répartition

Selon la définition du *Vocabulaire Forestier* (Bastien et Gauberville, 2011), la structure régulière correspond à une répartition des diamètres unimodale, au contraire de la structure irrégulière. Afin d'appliquer directement cette définition, un script R utilisant le package *segmag* (Papenmeier, 2014) a permis de détecter les maxima locaux de la répartition lissée des tiges par classe de diamètre, puis de supprimer les maxima ne correspondant pas à des modes (maximum engendré par un seul arbre ou maximum contraire à une tendance plus globale de la courbe). Les modes ainsi détectés ont permis d'en déduire le type de répartition (régulière, irrégulière, bimodale) (cf. Figure 8).

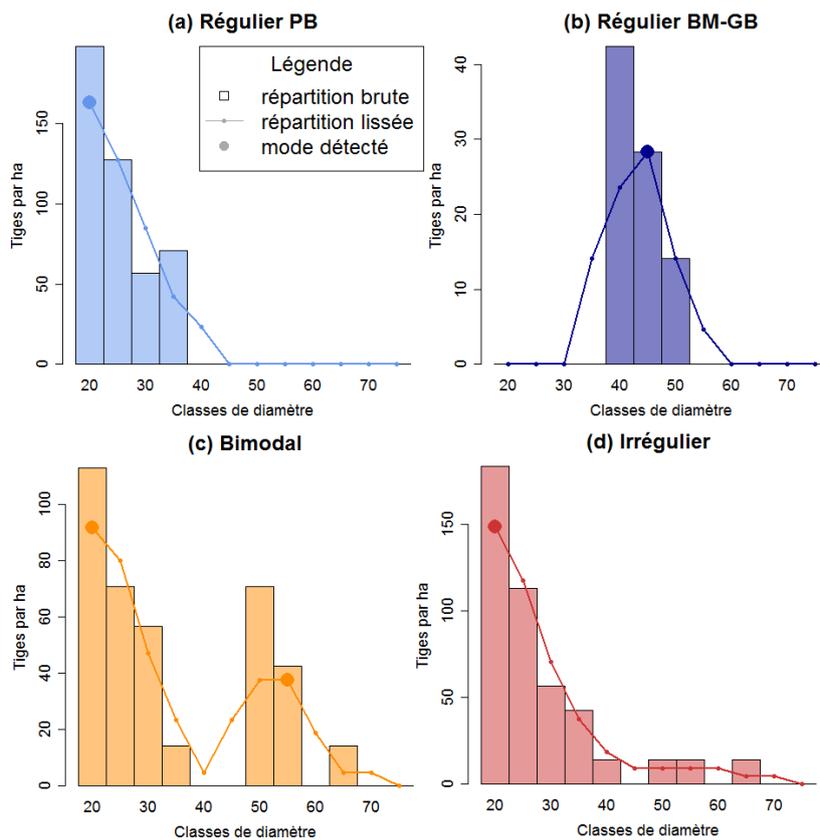


FIGURE 8 : REPARTITION DES DIAMETRES ET DETECTION DES MODES SUR 4 PLACETTES DU VERCORS

### Principes de conversion des modes détectés en type structural :

- (a) mode dans les PB et prédominance de PB : régulier à PB
- (b) mode dans les BM-GB : régulier à BM-GB
- (c) deux modes : bimodal
- (d) mode dans les PB sans prédominance de PB : irrégulier

## II.2.2. Utilisation du coefficient de Gini

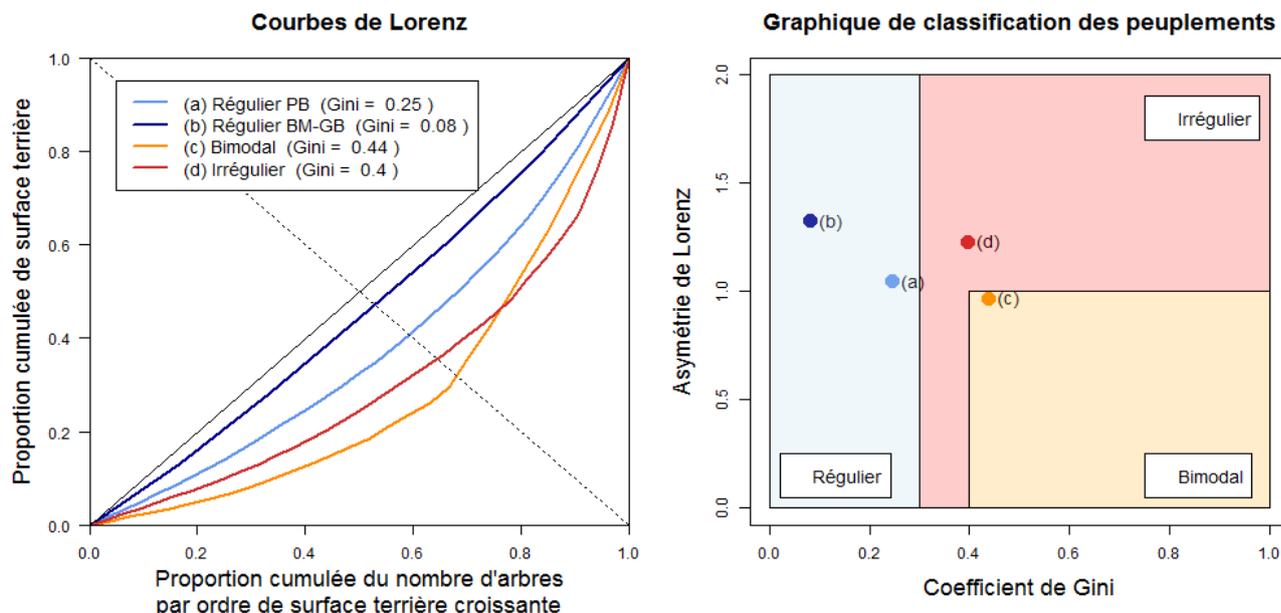
Le coefficient de Gini permet de caractériser l'hétérogénéité de structure des peuplements (cf. § I.2.2.), on peut donc admettre qu'au-delà d'un certain seuil du coefficient de Gini, les peuplements deviennent irréguliers.

Selon les études et les contextes forestiers, ce seuil permettant de séparer les peuplements réguliers des irréguliers a été estimé entre 0,3 et 0,5 (Duduman, 2011; Lexerød et Eid, 2006; Valbuena et al., 2012). Dans ces études, le diamètre précomptable varie de 0 à 12 cm, ce qui influe naturellement sur ce seuil (plus le diamètre précomptable est élevé, plus la frontière est basse). Par ailleurs, Valbuena et al. (2012) ont distingué les peuplements à deux étages ou bimodaux par un coefficient de Gini un peu plus fort et surtout une asymétrie de Lorenz inférieure à 1.

Le seuil correspondant le mieux aux peuplements étudiés et au diamètre précomptable de cette étude a été trouvé manuellement. L'appréciation s'est faite à partir de graphiques comparant le

coefficient de Gini de peuplements typiquement réguliers ou irréguliers. Ainsi, la **frontière entre régulier et irrégulier a été fixée à un coefficient de Gini de 0,3** et les peuplements ont été considérés comme irréguliers bimodaux au-delà d'un coefficient de Gini de 0,4 et en-deçà d'une asymétrie de 1 (cf. Figure 9). Le calcul du coefficient de Gini et de l'asymétrie de Lorenz ont été effectués à l'aide du package *ineq* (Zeileis, 2014) de l'outil R.

FIGURE 9 : COURBES DE LORENZ ET CRITERES DE CLASSEMENT DES PEUPELEMENTS POUR LES 4 PLACETTES DU VERCORS DE LA FIGURE 8



Les placettes (a), (b), (c) et (d) correspondent aux peuplements représentés à la Figure 8. Des courbes de Lorenz sont issues le coefficient de Gini et l'asymétrie de Lorenz dont les valeurs permettent de déduire le type structural d'après les seuils représentés dans le graphique de classification des peuplements.

### II.2.3. Typologie du Massif Vosgien

L'approche « typologie » a été utilisée via la typologie des peuplements du Massif Vosgien (Asael et al., 1999). Bien qu'adaptée aux peuplements de résineux et de hêtre de montagne, elle est proche d'autres typologies comme celle du Plateau Lorrain construites pour des futaies feuillues et d'anciens taillis sous futaie. Cette typologie utilise une méthode assez transversale basée sur les proportions de surface terrière de chaque catégorie de grosseur (PB, BM, GB) et a permis de caractériser les peuplements des deux contextes forestiers étudiés. D'après cette typologie, dont la clé de détermination est disponible en Annexe 2, les types 50 à 54 correspondent à des peuplements irréguliers, le type 31 a été considéré comme bimodal et les autres types comme des peuplements réguliers.

Bien que la pratique soit différente en montagne, la **limite inférieure des gros bois a été fixée à 47,5 cm** dans cette étude, afin de respecter les critères de construction de la typologie Massif Vosgien. La limite entre les petits bois et les bois moyens reste à 27,5 cm de diamètre.

### II.2.4. Jeu de placettes de calibration

Les placettes de référence terrain ont donc été classées en irrégulier ou régulier selon les trois différentes approches. Les différentes approches ne donnant pas toujours le même résultat, un

groupe restreint a été constitué à partir des placettes dont le classement est identique, quel que soit l'approche utilisée. Ce groupe de placettes est donc a priori **formé de peuplements typiquement réguliers ou irréguliers** et permet d'exclure les peuplements difficilement interprétables et modélisables. Les peuplements clairs et peu denses, d'une surface terrière < 10m<sup>2</sup>/ha ou d'une densité < 80 tiges/ha, ont également été retirées de ce groupe (cf. § III.1.1. et III.2.1. pour le détail de la répartition des placettes).

Dans la suite de l'étude, les modélisations ont d'abord été réalisées à partir de ce jeu de données restreint, **nommé jeu de placettes « typiques »**. Des tests complémentaires ont ensuite été menés en utilisant l'ensemble des placettes dont la structure a été déterminée à l'aide du coefficient de Gini. L'emploi de ce **jeu de placettes nommé « Gini »** permet d'avoir un nombre plus important de données de calibration.

### II.3. Calcul de métriques LiDAR

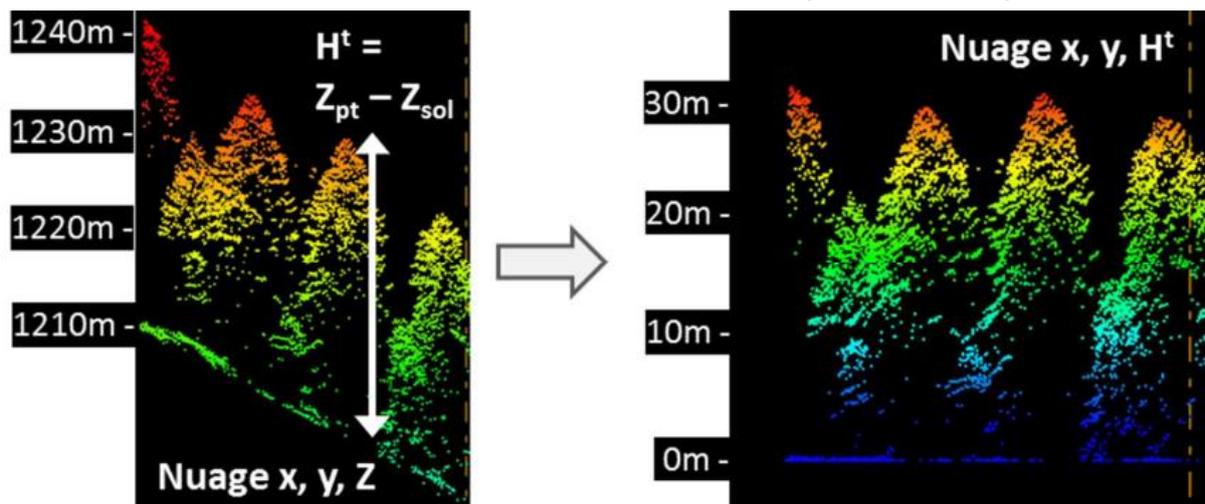
Les métriques ou indicateurs LiDAR sont les variables indépendantes qui seront utilisées pour la prédiction. Il est nécessaire que celles-ci soient fortement liées et corrélées à la structure des peuplements. Les métriques utilisées dans cette étude sont issues de trois types de données dérivées du nuage de points LiDAR : le nuage de points normalisé en hauteur, le modèle numérique de hauteur (MNH) et les apex détectés sur ce MNH. Ces métriques sont calculées sur la même emprise que celle des placettes de référence terrain, c'est-à-dire sur des **cercles de 15m de rayon**.

#### II.3.1. Métriques issues du nuage de points

##### II.3.1.1. Normalisation du nuage de points

Le nuage de points brut donnant des altitudes, il convient de le normaliser par rapport au modèle numérique de terrain (MNT) afin d'obtenir des hauteurs (cf. Figure 10). Cette opération a été réalisée à l'aide de l'outil *Lastools* (Rapidlasso GmbH, 2013).

FIGURE 10 : NORMALISATION DU NUAGE DE POINTS EN HAUTEUR (SOURCE : ONF R&D)



A l'altitude de chaque point LiDAR (à gauche) est soustraite l'altitude du sol modélisée par un réseau de triangles irréguliers (TIN) réalisé à partir des points LiDAR classés « sol ». On obtient ainsi un nuage de points normalisé en hauteur (à droite).

### II.3.1.2. Calcul des métriques « LAS »

A partir de ce nuage de points normalisé en hauteur, différentes métriques ont été calculées à l'aide de l'outil *Lasmanager* (© ONF). Des écart-types et centiles de la répartition des points en hauteur, des densités relatives de points par tranches de hauteur et des taux de pénétration du signal ont ainsi été calculés.

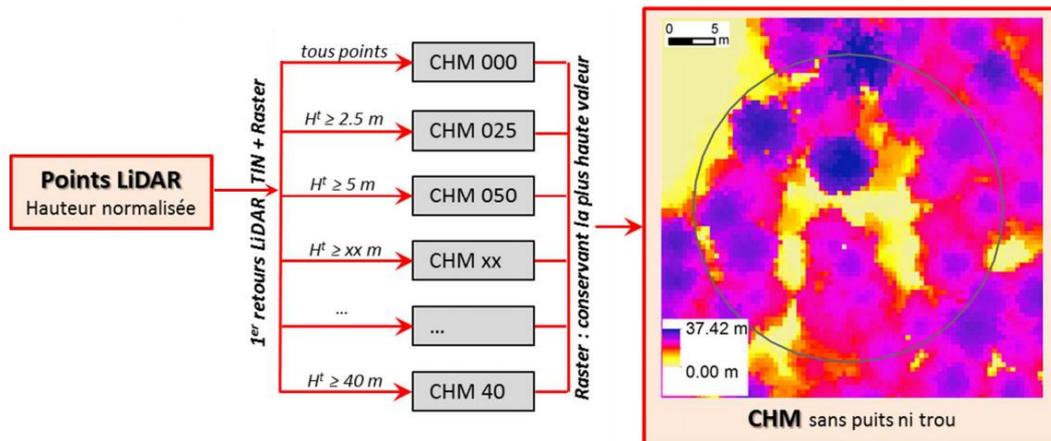
La forme de l'histogramme de la répartition en hauteur des points apporte également des informations sur la stratification des peuplements. Comme pour les diamètres, cette forme peut être caractérisée par des indices comme le coefficient de Gini, des modes, des moments, ou encore par les paramètres de lois s'y ajustant (Coops et al., 2007; Jaskierniak et al., 2011). Ces métriques ont été calculées à l'aide de l'outil R.

### II.3.2. Métriques issues du MNH

#### II.3.2.1. Obtention du MNH

Le MNH est obtenu via un script *Lastools* qui suit la méthode de Khosravipour et al. (2013) basée sur le nuage de points normalisé en hauteur. Ce MNH, créé au pas de 50 cm dans l'étude, est la résultante d'une combinaison de MNH partiels réalisés à différents seuils de hauteur (cf. Figure 11). Cette procédure permet d'éviter les « trous » et « puits » traditionnellement obtenus dans d'autres méthodes de création de MNH.

FIGURE 11 : SCHEMA DE LA METHODE DE CREATION DU MNH (SOURCE : ONF R&D)



Pour différents seuils de hauteur, les premiers retours LiDAR supérieurs à ces seuils permettent de réaliser plusieurs TIN. Les TIN sont ensuite convertis en rasters, créant ainsi un jeu de MNH. Le MNH final résulte de la combinaison de ce jeu de MNH partiels : à chaque pixel du MNH final est affectée la valeur de hauteur la plus élevée des MNH partiels.

#### II.3.2.2. Calcul de métriques « MNH »

De la même façon que pour le nuage de points, des statistiques de base ont été calculées sur les hauteurs des pixels du MNH (maximum, minimum, écart-type, quantiles...). Le MNH a également permis de calculer un pourcentage de trouées ou encore un volume de canopée à différentes hauteurs.

Enfin un indice de rugosité de la canopée (rumple index) a été calculé grâce au package *insol* de R (Corripio, 2014). Cet indice étant le ratio entre la surface drapée du MNH et la surface réelle au sol, il

offre une bonne image de l'hétérogénéité 3D de la canopée (Kane et al., 2010). Ainsi des peuplements dont le couvert est très hétérogène ont un indice de rugosité plus fort.

### II.3.3. Métriques issues de la segmentation des couronnes

#### II.3.3.1. Segmentation par la méthode des « bassins versants »

La méthode de segmentation des « bassins versants » utilise les outils hydrologiques d'*ArcGis*. En effet, en inversant le MNH, les couronnes des arbres peuvent être considérées comme des cuvettes ou bassins versants dont les frontières et les minima locaux sont détectables. Ces frontières correspondent alors aux limites des couronnes et les minima locaux à la localisation des apex, dont la hauteur est récupérée sur le MNH de départ (cf. Annexe 3).

Cette segmentation peut être effectuée sur le MNH brut, où des sur-détections (détections de branches...) sont possibles, ou bien sur un MNH lissé, où des sous-détections sont possibles suivant les peuplements.

Pour tous les **sites de montagne**, c'est cette **méthode des « bassins versants »** qui a été utilisée. Les taux de bonnes prédictions étant meilleurs avec la segmentation basée sur le **MNH non lissé** que sur le MNH lissé, c'est donc cette segmentation qui a été conservée. En revanche, cette technique détectant beaucoup trop de tiges dans les peuplements **feuillus de plaine**, un autre **algorithme (Segma)** a donc été testé. En effet, pour ces peuplements feuillus, la segmentation *Segma* est visuellement bien meilleure (cf. Annexe 3) même si elle n'augmente que très légèrement le taux de bonnes prédictions.

#### II.3.3.2. Segmentation par le logiciel *Segma*

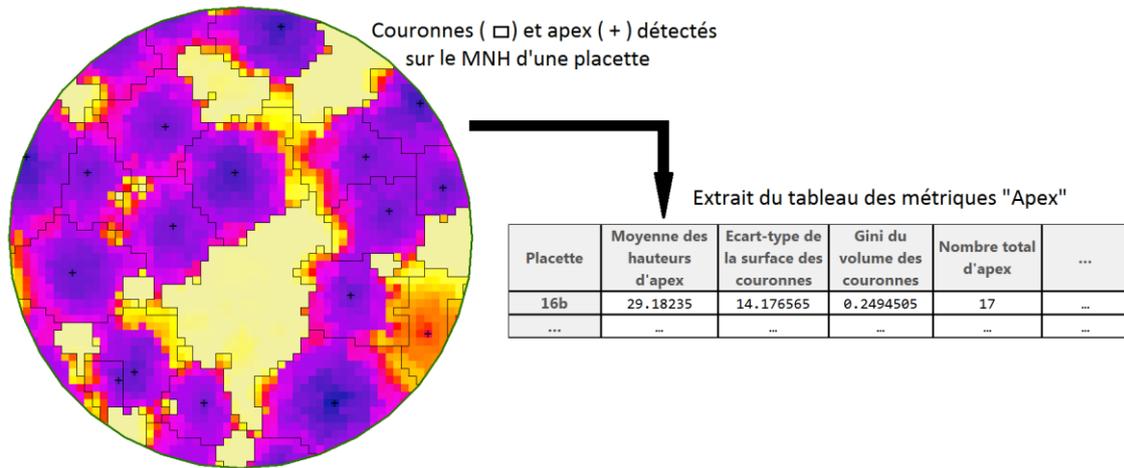
Cet algorithme développé par St-Onge et al. (2015) se base sur la croissance de régions pour segmenter des couronnes d'arbres. L'entrée est toujours un MNH, qui va être lissé par un filtre gaussien adaptatif, c'est-à-dire que les arbres hauts seront plus lissés que les petits. A partir des maxima locaux détectés sur ce MNH lissé, les couronnes vont être formées par agrégation de pixels jusqu'à ce qu'une condition d'arrêt stoppe cette croissance (« vallée » trop importante, forme non circulaire de la couronne, rayon de couronne trop important...). Tous ces critères ainsi que le lissage sont paramétrables ; il convient donc de les tester selon les peuplements (cf. Annexe 3).

Par ailleurs, certaines incohérences ayant été relevées dans les résultats de cette segmentation, un script Python a été écrit afin de récupérer les bonnes informations pour chaque couronne et apex (hauteur, surface...). De même, de toutes petites couronnes sans correspondance au terrain étant créées lors de cette segmentation, des filtres ont été utilisés lors du calcul de métriques, permettant de supprimer les couronnes détectées d'une surface inférieure à 0,25 ou 0,5 fois la hauteur de l'apex.

#### II.3.3.3. Calcul de métriques « Apex »

Une fois les couronnes et les apex détectés, il est possible d'obtenir un certain nombre de variables liées directement aux arbres du peuplement et de fait assez parlantes pour un forestier. Par ailleurs ces métriques améliorent généralement les modèles prédictifs (André et al., 2015). Ainsi, des statistiques de base (cf. Figure 12) ont été calculées sur les hauteurs d'apex, les surfaces de couronne ou encore les volumes de couronne ( $volume = \sum_{\text{la couronne}} \text{pixels de } Hauteur_{\text{pixel}} \cdot Surface_{\text{pixel}}$ ).

FIGURE 12 : SCHEMA DU PRINCIPE DE CALCUL DE METRIQUES "APEX"



## II.4. Sélection des métriques LiDAR

Une fois toutes ces métriques calculées, environ 500 variables sont disponibles pour la prédiction. Dans ce grand nombre de variables, beaucoup sont corrélées entre elles et beaucoup n'ont aucun lien avec la structure des peuplements. L'étape de sélection des variables, réalisée sur les placettes de calibration, est donc importante pour construire des modèles compréhensibles avec de bons taux de prédiction. En sélectionnant un petit jeu de métriques importantes pour la structure et peu corrélées entre elles, cette étape permet d'appliquer le principe de parcimonie, d'éviter le sur-apprentissage et de respecter les hypothèses de validité des modèles.

### II.4.1. Sélection visuelle et par régression linéaire

Tout d'abord, une analyse graphique de la répartition des métriques LiDAR en fonction du coefficient de Gini a permis de détecter des relations, linéaires ou non, entre la structure des peuplements et certaines métriques. Puis, afin de déterminer les variables ayant la meilleure corrélation (meilleur  $R^2$  ajusté) avec l'indice de structure ainsi que celles le prédisant avec le moins d'erreur (faible RMSE), des régressions linéaires entre chaque variable et l'indice de Gini ont été réalisées (cf. Figure 13).

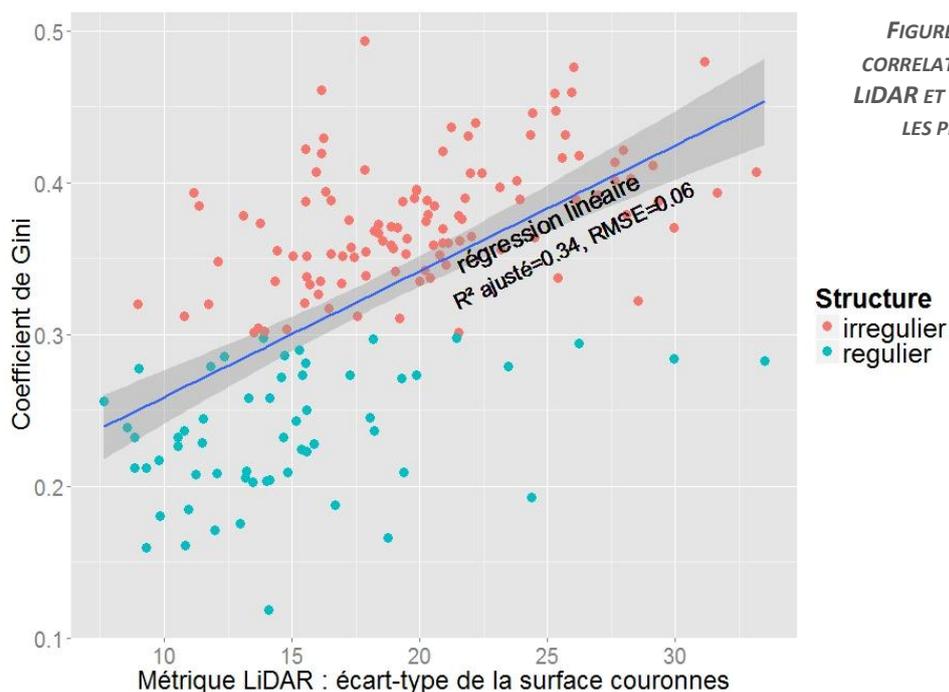


FIGURE 13 : GRAPHIQUE DE LA CORRELATION ENTRE UNE METRIQUE LIDAR ET LE COEFFICIENT DE GINI SUR LES PLACETTES DU VERCORS

#### II.4.2. Sélection de variables par l'algorithme VSURF

Dérivé de la méthode statistique *Random Forest* (ou forêts aléatoires), le package *VSURF* de R (Genuer et al., 2014) permet de sélectionner les variables les plus importantes dans un grand jeu de données. *Random Forest* (RF) est une méthode non paramétrique de prédiction pour la régression ou la classification, dont le principe est l'agrégation d'un grand nombre d'arbres de décision (cf. § II.5.4.). Une erreur OOB (out-of-bag) de prédiction  $y$  est automatiquement calculée. A partir de cette erreur, un test de permutation permet de déterminer l'importance d'une variable : c'est alors la différence entre l'erreur OOB de l'échantillon perturbé (où les valeurs de la variable sont permutées) et l'erreur OOB de l'échantillon normal.

C'est à partir de l'importance des variables qu'une sélection peut être effectuée. L'algorithme *VSURF* consiste en trois étapes :

- *Etape 1 (classement et seuillage)* : les variables sont classées par ordre d'importance décroissante. En dessous d'un seuil d'importance, les variables sont supprimées. Il reste alors un jeu de  $r$  variables.
- *Etape 2 (sélection pour l'interprétation)* : une série de RF est construite en partant d'une RF avec la variable ayant la plus grande importance, puis en ajoutant les  $r$  variables par ordre d'importance décroissante. Les variables du modèle avec l'erreur OOB la plus faible sont conservées.
- *Etape 3 (sélection pour la prédiction)* : afin d'éliminer la redondance entre les variables, des modèles sont alors construits en ajoutant les variables pas à pas. Seules les variables qui font plus décroître l'erreur OOB qu'une variable de bruit sont conservées.

*Random Forest* fonctionnant pour la prédiction de variables quantitatives comme qualitatives, l'algorithme *VSURF* a permis de sélectionner les variables prédisant le mieux la structure (type régulier/irrégulier) ou l'indice de Gini. La corrélation forte entre variables pouvant « masquer » l'importance de certaines de ces variables, l'algorithme a également été appliqué en ne conservant que quelques variables par famille de corrélation (une famille regroupant des variables corrélées à plus de 90% entre elles).

#### II.4.3. Sélection finale

De ces processus de sélection, plusieurs jeux de variables importantes se sont démarqués. Le jeu de variables finalement conservé pour la prédiction est celui qui engendrait la plus faible erreur de prédiction. Un second jeu de variables issu uniquement du MNH et de la segmentation des apex a également été conservé. En effet, le MNH est une source de données qui, à l'avenir, grâce à la photogrammétrie et au MNT rénové de l'IGN, pourra éventuellement être obtenue sur de plus grandes zones et de façon moins onéreuse que le nuage de points LiDAR. Enfin, de façon à respecter les conditions de validité des méthodes de prédiction et afin d'éviter le fléau de la dimensionnalité, la colinéarité entre variables sélectionnées a été diminuée en supprimant les variables les moins importantes et les plus corrélées (à plus de 80%).

## II.5. Méthodes de prédiction de la structure

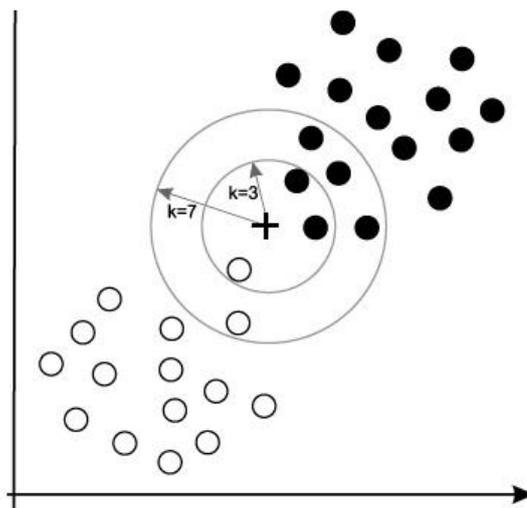
En se basant sur les résultats de la bibliographie (cf. § I.3.), deux grandes approches de prédiction de la structure avec chacune différentes méthodes statistiques ont été appliquées :

- La **classification supervisée** pour la prédiction directe du type structural : k-Nearest Neighbor (kNN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Random Forest (RF)
- La **prédiction du coefficient de Gini et de l'asymétrie de Lorenz** pour définir par la suite le type structural : Random Forest (RF), k-Most Similar Neighbour (kMSN)

### II.5.1. Les k plus proches voisins (kNN)

La classification par la méthode des k plus proches voisins, réalisable grâce au package *class* sous R (Venables et Ripley, 2002), permet d'assigner un individu à un groupe selon sa proximité avec les observations du jeu de calibration. Dans un espace défini par les métriques LiDAR, la distance euclidienne entre l'observation à classer (+ sur la Figure 14) et les observations de référence (● et ○ sur la Figure 14) est calculée. Un vote majoritaire sur les k plus proches voisins permet alors de classer la nouvelle observation.

FIGURE 14 : SCHEMA DU PRINCIPE DE CLASSIFICATION PAR LA METHODE DES K PLUS PROCHES VOISINS



Dans cette étude k a été fixé à 7, ce nombre de voisins donnant la plus faible erreur de prédiction.

### II.5.2. L'analyse discriminante linéaire (LDA)

L'analyse discriminante linéaire (ou Linear Discriminant Analysis) est une méthode permettant de discriminer des groupes d'individus à partir de variables indépendantes. Cette méthode peut être utilisée de manière descriptive afin de connaître les variables qui départagent le mieux les groupes entre eux, mais également de façon prédictive afin de classer automatiquement des individus en fonction de leurs caractéristiques. Dans le cas de cette étude, la fonction LDA, disponible dans le package *MASS* (Venables et Ripley, 2002), a été utilisée en tant qu'outil de classification supervisée pour prédire les types structuraux en fonction des métriques LiDAR.

Cette analyse discriminante linéaire consiste en la recherche de l'espace vectoriel permettant d'expliquer au mieux la variabilité interclasse (à la différence d'une analyse en composantes

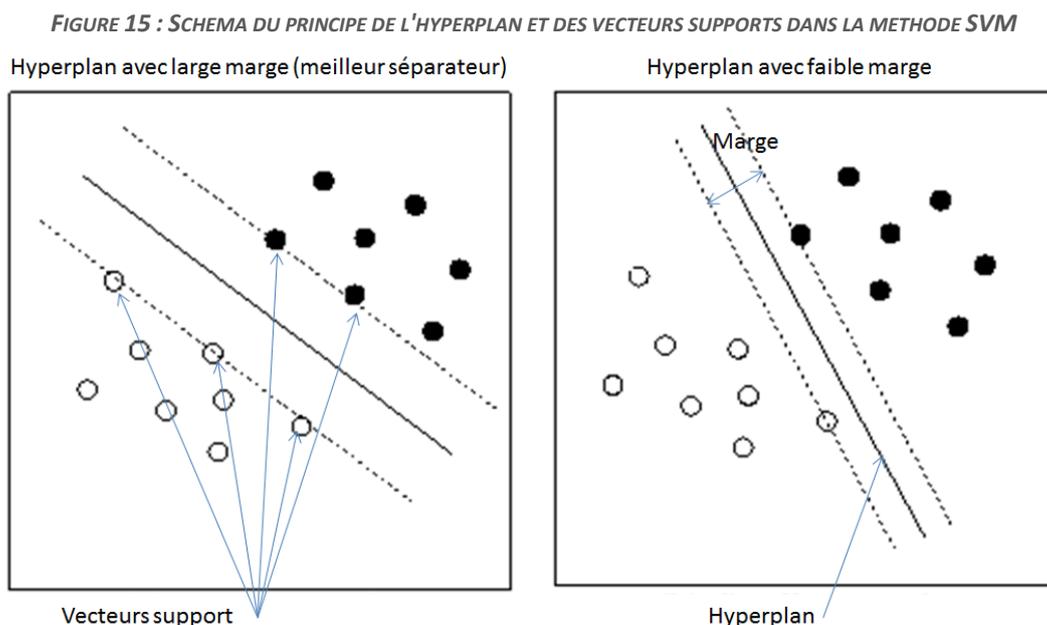
principales permettant d'expliquer la variabilité entre individus). Les dimensions de cet espace seront donc des combinaisons linéaires des différentes variables indépendantes (métriques LiDAR) maximisant la séparation entre groupes tout en minimisant l'hétérogénéité au sein d'un groupe.

La classification LDA se base sur une règle bayésienne de décision qui cherche à trouver quelle classe  $k$  (parmi les  $n$  classes) maximise la probabilité que l'individu appartienne à  $k$ , sachant que les métriques sont égales à  $x$  :

$$\arg \max_k P(k|x) = \frac{P(x|k) \cdot P(k)}{\sum_{i=1}^n P(x|i) \cdot P(i)}$$

### II.5.3. Support vector machine (SVM)

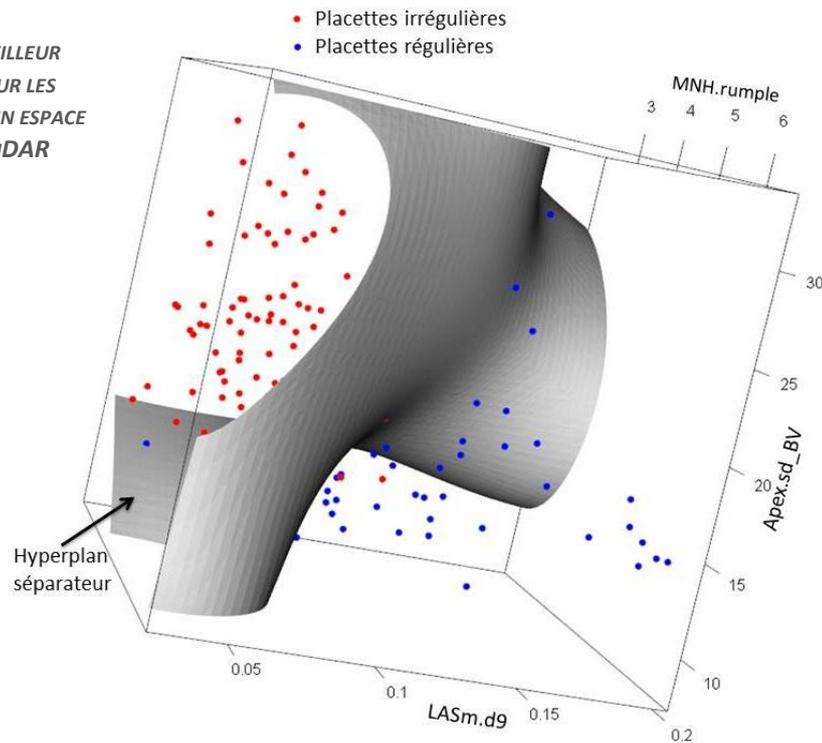
Support Vector Machine (SVM) est également une méthode de classification supervisée, contenue dans le package *e1071* (Meyer et al., 2014). A partir des données d'apprentissage, SVM détermine un « hyperplan optimal » dans l'espace constitué par les variables explicatives. Pour être optimal, cet hyperplan doit séparer au mieux deux classes avec la plus grande marge possible. Cette frontière se base sur des points proches appelés vecteurs supports (cf. Figure 15).



L'hyperplan n'est pas obligatoirement linéaire. Pour mieux séparer les classes, une fonction noyau (polynomiale, sigmoïde ou de base radiale) peut être utilisée. Lors de la prédiction, les individus seront classés selon leur position par rapport à cette frontière.

Les hyperplans linéaires et de base radiale (cf. Figure 16) donnant les meilleurs taux de prédiction, ce sont eux qui ont été utilisés dans cette étude.

**FIGURE 16 : EXEMPLE DU MEILLEUR HYPERPLAN SEPARATEUR POUR LES PLACETTES DU VERCORS DANS UN ESPACE FORME PAR 3 METRIQUES LIDAR**



#### II.5.4. Random Forest (RF)

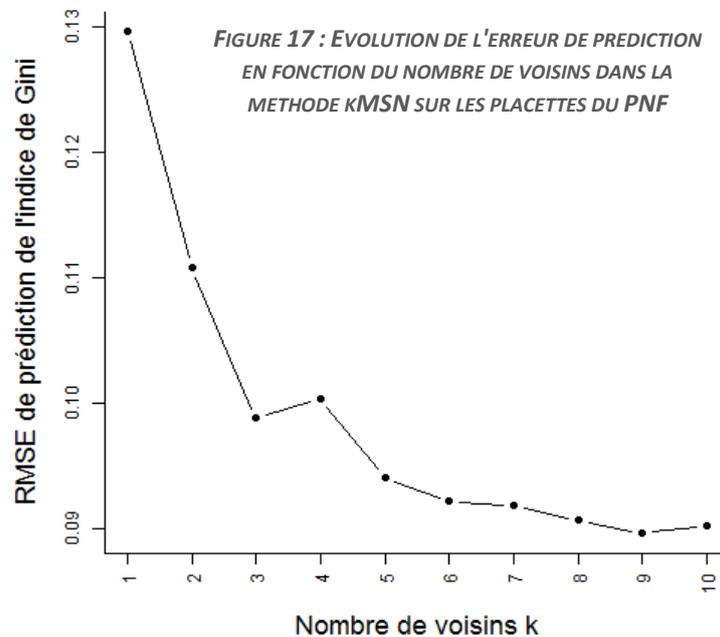
Comme évoqué au paragraphe II.4.2., *Random Forest* se base sur une multitude d'arbres de décisions afin de prédire une variable quantitative (régression) ou qualitative (classification). La construction d'un arbre de décision basique consiste à trouver pour chaque nœud le meilleur critère, sur les métriques LiDAR, permettant de séparer les observations en deux par rapport à la variable à prédire. *Random Forest* est donc une agrégation d'un grand nombre de ces arbres de décision construits par la méthode du *Bagging*, c'est-à-dire que chaque arbre est construit à partir d'un sous-échantillon différent de l'échantillon d'apprentissage et que chaque nœud est déterminé à partir d'une sélection aléatoire de variables explicatives. L'ensemble des arbres est utilisé pour la prédiction, le résultat est alors le groupe majoritaire en classification et la moyenne en régression. Dans cette étude, cette méthode a été utilisée pour la prédiction du type de structure (irrégulier ou régulier), mais aussi des coefficients de Gini et de l'asymétrie de Lorenz.

Le côté aléatoire fait que cette méthode ne donne pas les mêmes résultats d'une fois sur l'autre et que le mécanisme de prédiction est difficilement interprétable. Mais c'est justement cette variabilité qui permet à la méthode d'être très robuste et d'avoir de faibles erreurs de prédiction. Cet outil a été utilisé sur R via le package *randomForest* (Liaw et Wiener, 2002).

#### II.5.5. Les k voisins les plus similaires (kMSN)

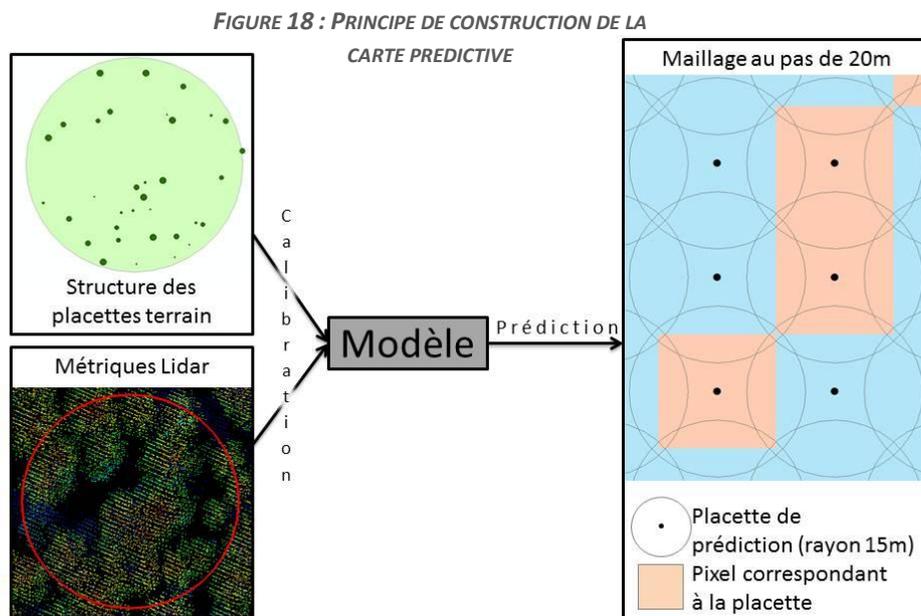
La méthode des k voisins les plus similaires (ou k-Most Similar Neighbour) permet de prédire pour un individu « cible » une ou plusieurs variables quantitatives (ici le coefficient de Gini et l'asymétrie de Lorenz) en fonction des k individus de « référence » les plus similaires. La valeur prédite est alors la moyenne pondérée par la distance des valeurs des k voisins. Le calcul de la distance entre les individus de référence et les individus cibles est basé sur une analyse canonique des corrélations entre les métriques LiDAR et les coefficients structuraux. Le package *yalmpute* (Crookston et Finley, 2007) permet de réaliser ce type de prédiction sur R.

Si cette méthode permet l'estimation réaliste de plusieurs variables d'intérêts, il est nécessaire pour cela de trouver le bon nombre de voisins (k). En effet, en augmentant la valeur de k, il y a une diminution de l'erreur de prédiction mais également une diminution de la variabilité des prédictions. Dans notre cas d'étude, les graphiques comme celui de la Figure 17 ont permis de fixer le nombre de voisins à 5.



## II.6. Production de cartes de prédiction

Une fois la calibration des différents modèles réalisée, il est possible de les appliquer sur l'ensemble des zones d'étude. Pour cela, un maillage de placettes de prédiction a été réalisé. Ces placettes de prédiction sont de la même taille que les placettes de calibration, soit 15 m de rayon, afin que les modèles restent valides (White et al., 2013). Le maillage a été réalisé au pas de 20 m pour le site du Vercors et au pas de 30 m pour le site du PNF, cette distance correspondant *in fine* à la résolution du raster résultat des prédictions (cf. Figure 18).



Dans la plupart des méthodes de classification supervisée utilisées (SVM, LDA et RF), en plus de la classe prédite, une probabilité d'appartenance à chaque classe est calculée. Cette probabilité est donc une façon de qualifier l'incertitude de la prédiction pour chaque pixel. Si la différence entre la probabilité d'avoir une structure régulière et celle d'avoir une structure irrégulière est faible, le peuplement est certainement à la limite entre les deux types de structure (cf. § III.1.4.).

Ainsi trois grands types de cartes (ou rasters) ont été produits :

- La prédiction du **type de structure** (régulier ou irrégulier)
- La **probabilité** de régularité ou irrégularité du peuplement
- La prédiction de **l'indice de Gini et de l'asymétrie de Lorenz**

## II.7. Validation des modèles de prédiction

### II.7.1. Validation croisée

L'erreur de prédiction des modèles a été calculée par une validation croisée (ou cross-validation). Le principe consiste à partitionner les données en un échantillon d'apprentissage permettant de calibrer le modèle et un échantillon test sur lequel les modèles sont appliqués pour calculer l'erreur. Dans le cadre de cette étude, **l'échantillon test a été fixé à 1/3 du jeu de données** et l'erreur de prédiction a été calculée par la **moyenne des erreurs sur 100 validations croisées** avec des échantillonnages aléatoires. Pour certaines méthodes comme *Random Forest* ou *kMSN*, cette erreur est automatiquement générée comme un résultat du modèle, c'est donc cette dernière qui a été conservée.

Pour la classification supervisée, l'erreur est liée à la **précision de la classification**, qui est le taux de bonnes classifications :

$$\text{précision} = \frac{\text{nombre d'observations correctement classées}}{\text{nombre total d'observations}}$$

Par ailleurs, un second critère, le coefficient **Kappa**, couramment utilisé pour évaluer la qualité des classifications, a également été calculé. Il correspond au pourcentage de bonnes classifications n'étant pas dues au hasard. La précision et le coefficient Kappa ont été calculés grâce au package *caret* (Kuhn, 2015) de l'outil R.

Pour les méthodes prédisant directement le coefficient de Gini et l'asymétrie de Lorenz, c'est **l'erreur quadratique moyenne (RMSE)** qui a été calculée :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (v_{pred} - v_{obs})^2}$$

avec  $n$  le nombre d'observations,  $v_{pred}$  la valeur prédite,  $v_{obs}$  la valeur observée. Pour ces méthodes, la précision et le Kappa ont aussi pu être calculés, une fois le coefficient de Gini et l'asymétrie de Lorenz convertis en type de structure.

## II.7.2. Tests de robustesse sur des sites résineux de montagne

Afin de tester la robustesse des modélisations, les modèles de prédiction, issus de la calibration sur le site du Vercors, ont été appliqués aux sites de la combe d'Aillon et de Prénovel. Ainsi après avoir déterminé la structure des placettes de référence grâce aux données de terrain, les types structuraux ont été prédits à partir des métriques issues de ces placettes et des modèles calibrés sur le site du Vercors. La qualité de la prédiction a été calculée grâce à la précision des classifications et la RMSE des prédictions.

Ce type de test n'a pas été effectué pour les feuillus de plaine du fait de l'absence de données facilement disponibles sur des sites proches et similaires du site du PNF.

## II.7.3. Validation indépendante sur le terrain

Cette étape de validation n'a été réalisée que pour le site du Vercors. Elle consiste en la réalisation d'un échantillonnage sur le terrain. 51 placettes réparties en 17 grappes ont ainsi été réalisées. Ces grappes de trois placettes ont été positionnées à partir des cartes prédictives de façon à ce que des situations variées soient représentées (cf. Tableau 2).

TABLEAU 2 : REPARTITION DES GRAPPES DE PLACETTES DE VALIDATION INDEPENDANTE (SITE DU VERCORS)

	Type prédit par <i>Random Forest</i>		
	Irrégulier	Régulier	Limite régulier/irrégulier
Forêt privée	2	4	2
Forêt publique	4	2	3
<b>Total</b>	<b>6</b>	<b>6</b>	<b>5</b>

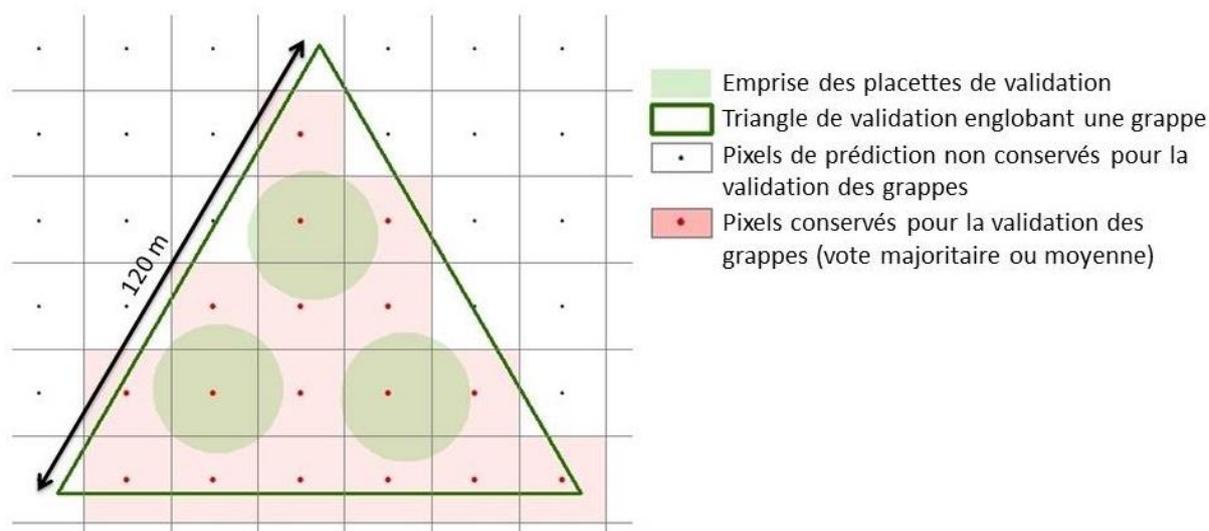
Ces placettes ont été positionnées à l'aide d'un récepteur GNSS mais, contrairement aux placettes de calibration, elles n'ont pas été recalées par rapport au MNH LiDAR. En effet, cette étape est coûteuse en temps puisqu'elle nécessite de prendre la position de plusieurs arbres sur le terrain. La position de nos placettes de validation a donc une imprécision de l'ordre de quelques mètres.

Le protocole de mesure appliqué sur ces placettes de terrain a été réalisé de façon à correspondre à celui des placettes de calibration. Il est disponible en Annexe 4. Sur des placettes de 15 m de rayon, le diamètre de tous les arbres vivants et précomptables (diamètre supérieur à 17,5 cm) a été mesuré. Par ailleurs, l'écart entre le vol LiDAR et la réalisation des placettes étant de 4-5 ans, les diamètres au niveau des souches des arbres visiblement récemment exploités ont été convertis en diamètres à 1,30 m et ajoutés au calculs (formule de Schaeffer (1953)). De même, les chablis récents ont été pris en compte. Cependant, les autres formes de mortalité et l'accroissement durant cette période ont été négligés dans les calculs.

Ces mesures ont permis de déterminer la structure de chaque placette. Celle-ci a été comparée au type de structure prédit à partir des modèles calibrés précédemment (les modèles ont été appliqués aux métriques extraites sur chaque placette).

La structure étant généralement appréhendée à plus large échelle qu'une placette de 15 m de rayon, le type structural de chaque grappe calculé sur l'ensemble des trois placettes a été comparé avec la prédiction issue du raster global (cf. Figure 19).

FIGURE 19 : PRINCIPE DE VALIDATION DE LA PREDICTION DE LA STRUCTURE AU NIVEAU DES GRAPPES DE PLACETTES



De même que pour la validation croisée, la précision, le Kappa et la RMSE ont été utilisés pour qualifier la qualité de la prédiction.

## II.8. Exemple d'application des modèles : vers une cartographie opérationnelle des types structuraux

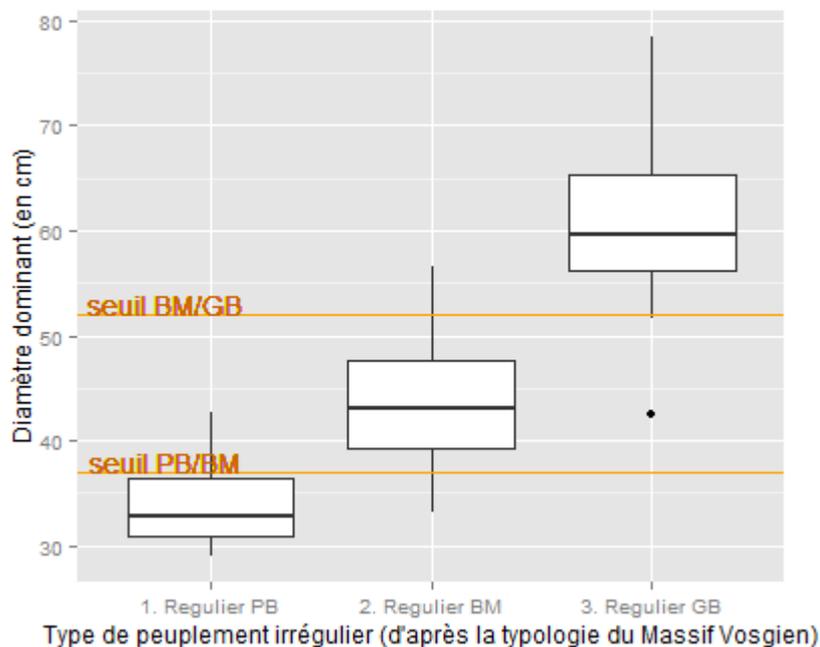
Les méthodes présentées dans les paragraphes précédents permettent de distinguer et de prédire les structures plutôt régulières ou irrégulières. Cependant, de façon opérationnelle, le gestionnaire ou l'aménagiste se contente rarement de cette simple information pour décrire des types de peuplements. Du point de vue de la structure, des distinctions sont apportées entre les différents stades de développement des peuplements réguliers ou encore les variétés de peuplements irréguliers (bimodaux ou à dominance de petits bois, bois moyens ou gros bois). De **façon exploratoire** dans le cadre du stage, une cartographie de **types détaillés** a été réalisée sur le site du Vercors.

Depuis quelques années, des modèles de prédiction de différentes variables dendrométriques ont été développés à la R&D de l'ONF. L'idée est ici de tirer parti des informations issues de ces modèles prédictifs pour détailler la cartographie des types de structures obtenue précédemment. Par ailleurs certaines métriques LiDAR peuvent aussi directement apporter des indications sur certains peuplements. La carte détaillée a donc été obtenue à partir de ces données multiples (métriques et modèles) via la calculatrice raster d'*ArcGis* (cf. Figure 21).

Ainsi le 90<sup>ème</sup> centile des hauteurs et le taux de trouées ont tout d'abord permis de distinguer les **zones non forestières et les peuplements « clairs »** ou peu denses. Ces peuplements « clairs » sont généralement définis par une surface terrière faible. Le modèle de prédiction de la surface terrière aurait donc pu servir. Cependant, ces modèles sont souvent peu fiables en dehors ou à la limite des valeurs de calibration, ce qui est le cas pour de faibles surfaces terrières. Le taux de trouées étant une variable directement issue du LiDAR et permettant de différencier nettement les peuplements clairs, il a donc été préféré au modèle prédictif.

*Cas des peuplements réguliers* : les préconisations des itinéraires sylvicoles pour les peuplements réguliers se basent classiquement sur des seuils de hauteur pour la phase d'amélioration et les premières éclaircies, puis de diamètre d'exploitabilité pour la récolte et la phase de régénération. Les modèles prédisant la hauteur dominante et le diamètre quadratique moyen seraient donc bien adaptés. N'étant pas disponibles sur le site du Vercors, c'est le modèle de prédiction du diamètre dominant qui a été utilisé. Ce modèle a une RMSE de prédiction de 8 % (Munoz et al., 2015). Des seuils de diamètre dominant ont donc été définis empiriquement de façon à différencier les peuplements réguliers à gros bois, bois moyen ou petit bois. (cf. Figure 20).

**FIGURE 20 : DIAMETRE DOMINANT ET SEUILS EN FONCTION DES TYPES DE PEUPELEMENTS REGULIERS POUR LES PLACETTES DU VERCORS**



*Cas des peuplements irréguliers* : mis à part les taillis sous futaie et autres structures bimodales, les peuplements irréguliers peuvent se différencier par la prépondérance d'une strate ou d'une catégorie de diamètre. Ainsi des peuplements irréguliers à gros bois pourront avoir un objectif de récolte plus important que dans d'autres peuplements irréguliers où l'enjeu d'amélioration sera prépondérant. Les peuplements irréguliers bimodaux ont été distingués par la valeur de leur asymétrie de Lorenz (cf. § II.2.2.), alors que les autres peuplements irréguliers ont été différenciés par leur pourcentage de gros bois, calculé à partir des modèles de G (RMSE de prédiction de 14 %) et de G gros bois (RMSE de prédiction de 30 %) (Munoz et al., 2015). Ces modèles étant basés sur la limite des gros bois traditionnellement utilisée en zone de montagne (42,5 cm et non 47,5 cm comme dans le reste de l'étude), le pourcentage-seuil différenciant les peuplements à gros bois des autres peuplements a cette fois été repris de la typologie locale (Cemagref et al., 2006).



## III. Résultats et analyses

### III.1. Résineux de montagne

#### III.1.1. Structure des placettes de terrain

Sur les 317 placettes de calibration du Vercors, la grande majorité représente des peuplements irréguliers souvent de type jardiné. Le jeu de placettes « typiques », c'est-à-dire suffisamment denses et caractérisées avec la même structure par les trois approches, est constitué de 169 placettes (cf. Tableau 3). Seule une très faible part des placettes sont « claires » (environ 2% des placettes).

Les autres sites de montagne sont semblables à celui du Vercors en termes de structure. Cependant, leurs caractéristiques dendrométriques ne sont pas identiques, ce qui pourrait influencer les résultats des tests de robustesse. Par rapport au site du Vercors, la forêt de la combe d'Aillon est très capitalisée avec de nombreux très gros bois et la forêt de Prénovel est légèrement plus riche en gros bois et un peu moins dense (cf. Tableau 3).

**TABLEAU 3 : REPARTITION DES PLACETTES ET CARACTERISTIQUES DENDROMETRIQUES PAR SITE D'ETUDE**

	Vercors (calibration)		Vercors (validation indépendante)		Aillons (test de robustesse)		Prénovel (test de robustesse)	
	Placettes « Gini »	Placettes « typiques »	Placettes « Gini »	Placettes « typiques »	Placettes « Gini »	Placettes « typiques »	Placettes « Gini »	Placettes « typiques »
Placettes à structure irrégulière	219 (69 %)	111 (66 %)	33 (65 %)	19 (58 %)	49 (91 %)	14 (78 %)	93 (82 %)	35 (74 %)
Placettes à structure régulière	98 (31 %)	58 (34 %)	18 (35 %)	14 (42 %)	5 (9 %)	4 (22 %)	20 (18 %)	12 (26 %)
<b>Nombre total de placettes</b>	<b>317</b>	<b>169</b>	<b>51</b>	<b>33</b>	<b>54</b>	<b>18</b>	<b>113</b>	<b>47</b>
Surface terrière* (m <sup>2</sup> /ha)	32,0 (± 11,7)		37,5 (± 11,7)		43,1 (± 13,8)		29,9 (± 9,4)	
Densité* (tiges/ha)	363 (± 152)		408 (± 131)		345 (± 185)		282 (± 108)	
Diamètre moyen* (cm)	34,6 (± 7,3)		34,5 (± 4,6)		42,0 (± 8,6)		37,8 (± 7,2)	

\* Valeur ± écart-type

#### III.1.2. Métriques sélectionnées

En suivant le processus de sélection, deux lots de métriques LiDAR ont été conservés (cf. Tableau 4) :

- Un lot avec les variables LiDAR donnant les meilleurs résultats de cross-validation (appelé Sélection de métriques « LiDAR »)
- Un lot de métriques issues uniquement du MNH et de la segmentation des apex (appelé Sélection de métriques « Apex-MNH »)

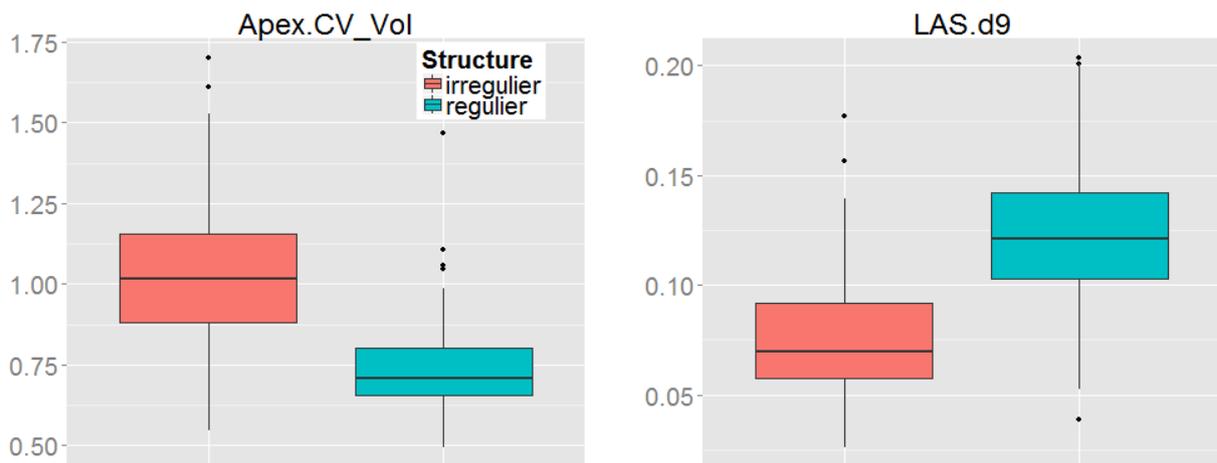
TABLEAU 4 : JEUX DE METRIQUES SELECTIONNES SUR LE SITE DU VERCORS

Sélection de métriques LiDAR	
Apex.CV_Vol	Coefficient de variation des volumes de couronnes issues de la segmentation
LAS.d9	Densité relative de tous les points entrés dans la tranche 10 (10 tranches de 2 m au 95 <sup>ème</sup> centile des hauteurs)
Apex.sd_BV	Ecart-type de la surface des couronnes (ou bassins versants) détectées
LAS.dHmean_fst	Moyenne des écarts à la moyenne sur les premiers retours du nuage : $moyenne_{retours}( h - h_{moy} )$
MNH.max	Hauteur maximale du MNH
Sélection de métriques « Apex-MNH »	
Apex.CV_Vol	Coefficient de variation des volumes de couronnes issues de la segmentation
Apex.sd_BV	Ecart-type de la surface des couronnes (ou bassins versants) détectées
MNH.rumple	Ratio de la surface de la canopée (MNH) par rapport à celle du sol (rumple index)
Apex.q50_HT	Médiane des hauteurs des apex détectés

Lors de cette sélection VSURF de métriques, trois grands types de métriques LiDAR se sont révélés particulièrement importants pour la prédiction :

- Les mesures de dispersion (coefficient de variation, écart-type et coefficient de Gini) des volumes ou surfaces de couronnes (*par ex. Apex.CV\_Vol*) : de façon logique ces variables ont des valeurs plus importantes pour les peuplements irréguliers où l'hétérogénéité des couronnes est plus grande (cf. Figure 22).
- La densité relative de points LiDAR dans les strates de hauteur les plus élevées (*par ex. LAS.d9*) : ces variables représentent en quelque sorte la quantité relative de végétation dans le haut de la canopée, cette quantité est donc supérieure pour des peuplements réguliers, où les arbres ont globalement tous la même hauteur (cf. Figure 22).

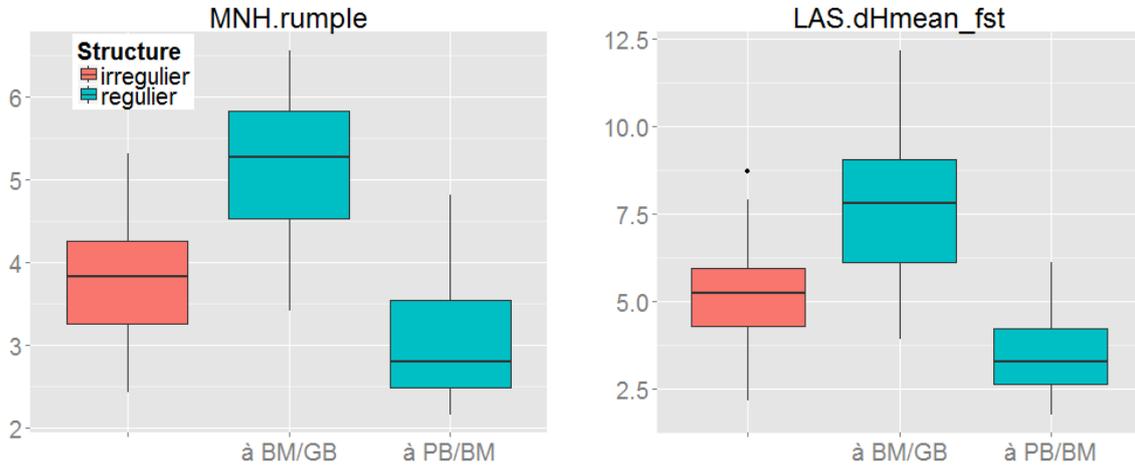
FIGURE 22 : VALEUR DES METRIQUES APEX.CV\_VOL ET LAS.D9 SUR LES PLACETTES "TYPIQUES" DU VERCORS EN FONCTION DE LEUR STRUCTURE



- L'indice de rugosité du couvert (*MNH.rumple*) : cette variable augmente avec l'hétérogénéité du couvert (cf. § II.3.2.2.). Cependant, elle ne discrimine pas directement les peuplements réguliers des peuplements irréguliers. Certains peuplements réguliers, notamment à bois moyens ou gros bois, n'ont pas de continuité de la canopée et leur couvert est assez hétérogène (forte valeur du *rumple index*) : les arbres y sont trop espacés et le MNH est alors constitué d'une série de creux et de bosses (cf. Figure 23 à gauche et Figure 24 en bas). Cet

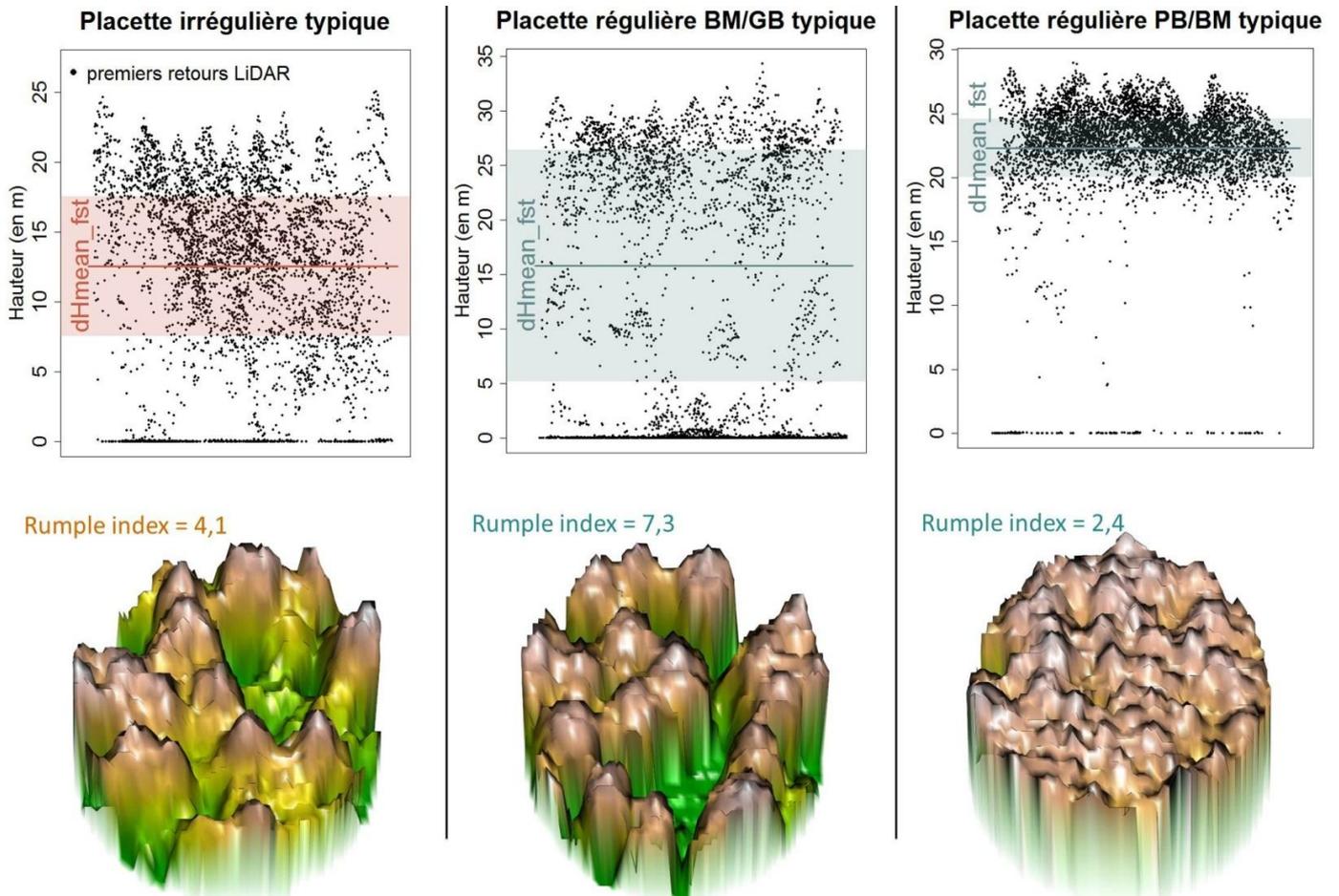
indice est par ailleurs très corrélé et proche des variables mesurant des écarts à la moyenne de la hauteur des points LiDAR (*par ex. LAS.dHmean\_fst*). Le principe est alors le même : dans des peuplements réguliers peu denses souvent à gros bois, certains points passent à travers la végétation, faisant baisser la hauteur moyenne et donc augmenter les écarts à cette moyenne (cf. Figure 23 à droite et Figure 24 en haut).

**FIGURE 23 : VALEUR DES METRIQUES MNH.RUMPLE ET LAS.DHMEAN\_FST SUR LES PLACETTES "TYPQUES" DU VERCORS EN FONCTION DE LEUR STRUCTURE**



**FIGURE 24 : NUAGE DE POINTS LIDAR ET REPRESENTATION 3D DU MNH POUR 3 PLACETTES DE TYPE STRUCTURAL DISTINCT DU VERCORS**

*La ligne horizontale correspond à la moyenne des hauteurs des premiers retours et la zone de couleur représente la valeur de la métrique LAS.dHmean\_fst*



### III.1.3. Validation et erreurs de prédiction

Les tableaux complets des résultats de validation sont disponibles en Annexe 5. Les tableaux 5, 6 et 7 présentent les résultats obtenus à partir de la sélection de métriques « LiDAR ».

#### III.1.3.1. Validation croisée

##### Classification supervisée

La validation croisée a tout d'abord été réalisée sur le jeu de placettes « typiques », puis sur le jeu de placettes « Gini ». Dans ce second cas, la précision des classifications est inférieure en moyenne de 5%, quelle que soit la méthode de modélisation. Cela étant sans doute dû aux placettes à la limite entre l'irrégulier et le régulier (cf. Tableau 5).

Hormis la méthode des plus proches voisins (kNN), qui donne les moins bonnes précisions, ce sont les méthodes de classification qui ne sont pas fondées sur des séparations linéaires (SVM à noyau radial et RF) qui classent le mieux la structure des placettes (cf. Tableau 5). Ces méthodes, plus souples, s'adaptent mieux aux métriques liées de façon non linéaire à la structure.

TABLEAU 5 : RESULTATS DE LA VALIDATION CROISEE DES METHODES DE CLASSIFICATION SUPERVISEE SUR LE VERCORS

Méthodes de classification supervisée	Validation croisée sur le jeu de placettes « typiques »		Validation croisée sur le jeu de placettes « Gini »	
	Précision	Kappa	Précision	Kappa
kNN (k=7)	81,5 %	57,7 %	77,7 %	44,4 %
LDA	88,6 %	73,8 %	81,4 %	54,5 %
SVM (hyperplan linéaire)	88,8 %	74,0 %	81,4 %	54,1 %
SVM (noyau radial)	90,8 %	79,5 %	84,6 %	62,4 %
Random Forest	89,7 %	75,4 %	84,7 %	61,2 %

##### Prédiction des coefficients de Gini et de l'asymétrie de Lorenz

La qualité des prédictions quantitatives du coefficient de Gini et de l'asymétrie de Lorenz a été également évaluée par validation croisée.

TABLEAU 6 : RESULTATS DE LA VALIDATION CROISEE DES METHODES DE PREDICTION QUANTITATIVE SUR LE VERCORS

Méthodes de prédiction quantitative	Validation croisée sur le jeu de placettes « typiques »				Validation croisée sur le jeu de placettes « Gini »			
	Prédiction des indicateurs de Lorenz		Classification à partir des coefficients prédits		Prédiction des indicateurs de Lorenz		Classification à partir des coefficients prédits	
	RMSE Gini*	RMSE asymétrie*	Précision	Kappa	RMSE Gini*	RMSE asymétrie*	Précision	Kappa
Random Forest	0,047	0,110	89,9 %	77,5 %	0,050	0,121	85,2 %	63,8 %
kMSN (k=5)	0,049	0,116	88,3 %	73,0 %	0,053	0,131	82,5 %	57,6 %

\* Le coefficient de Gini varie entre 0 et 1 en théorie (plutôt entre 0.1 et 0.6 en pratique) et l'asymétrie de Lorenz entre 0 et 2. En relatif, ces RMSE sont donc de l'ordre de 15% pour le coefficient de Gini et de 12% pour l'asymétrie de Lorenz.

La classification à partir des prédictions des coefficients de Lorenz donne globalement les mêmes résultats que la classification supervisée directe (cf. Tableau 5 et Tableau 6).

### Métriques « Apex-MNH »

Les classifications et prédictions réalisées à partir de métriques « Apex-MNH » ne font que légèrement baisser la qualité des résultats. Selon les méthodes, la précision est alors entre **1 et 3% plus faible** qu'avec la sélection de métriques « LiDAR ».

### Influence du jeu de calibration

La répartition des placettes dans le jeu de calibration influence nettement la qualité des prédictions. Le site du Vercors comptant un nombre plus important de placettes irrégulières que régulières (cf. Tableau 3), les résultats sont meilleurs pour la structure irrégulière, et ce quelles que soient les méthodes statistiques employées. En moyenne pour la validation croisée, la précision de classement des placettes **irrégulières est de 92%**, alors qu'elle est de **72% pour le régulier** (cf. Annexe 5).

Le risque de ce biais est qu'au niveau de la cartographie prédictive un certain nombre de zones soient classées en irrégulier, alors qu'elles sont en réalité régulières.

#### **III.1.3.2. Validation indépendante sur le site du Vercors**

La validation indépendante sur des placettes réalisées dans le Vercors donne des résultats très proches de ceux de validation croisée (cf. Tableau 7 et Tableau 5). Les prédictions sur le site du Vercors sont donc relativement stables.

Au niveau des **grappes** de placettes, les précisions restent du même ordre avec **76% de bonnes classifications et une RMSE de 0,05** pour le coefficient de Gini.

### Influence du jeu de calibration du Vercors (« Gini » et « typique »)

Les résultats de la validation indépendante et des tests de robustesse présentés dans le Tableau 7 ont été réalisés avec des modèles calibrés avec le jeu de placettes « Gini » du Vercors. Pour la validation indépendante sur le Vercors et les tests de robustesse sur les Aillons et sur Prénovel, la calibration sur le jeu de placettes « typiques » du Vercors donne en moyenne des résultats légèrement plus bas qu'avec le jeu de calibration « Gini » (cf. Annexe 5). Cela est sans doute dû à une quantité moindre de placettes de calibration dans le groupe « typique » et donc une moins bonne représentation de la diversité des cas.

**TABLEAU 7 : RESULTATS DE LA VALIDATION INDEPENDANTE ET DES TESTS DE ROBUSTESSE DES MODELES CALIBRES SUR LE JEU DE PLACETTES "GINI" DU VERCORS**

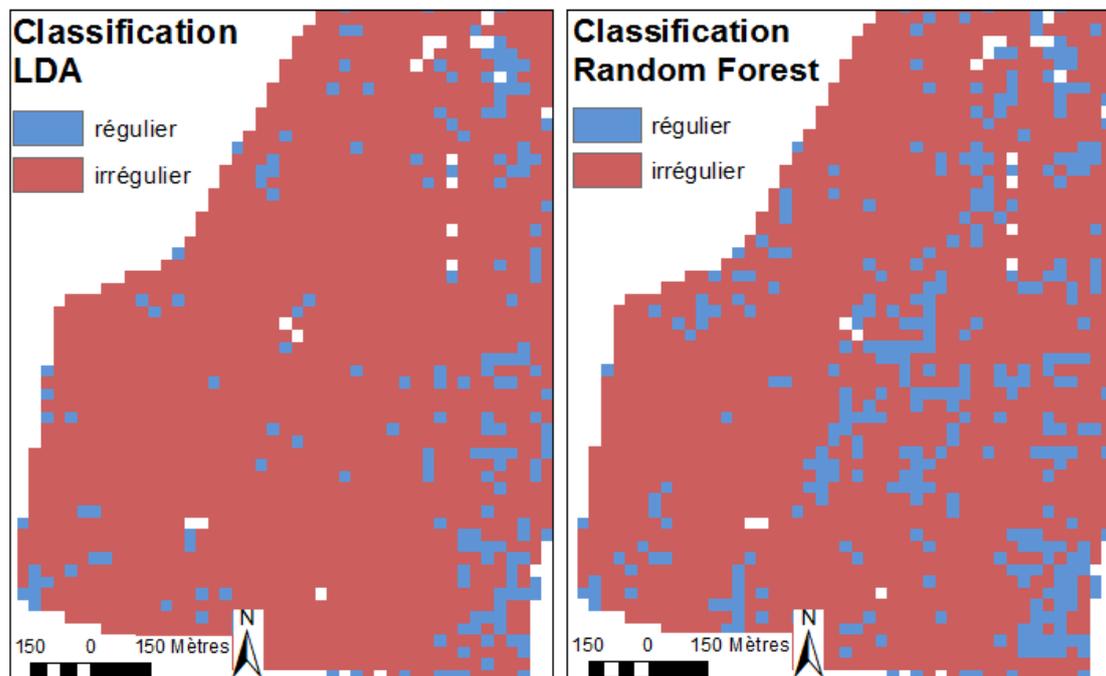
<i>Méthode de classification supervisée</i>	Validation indépendante sur le Vercors		Test de robustesse sur les Aillons		Test de robustesse sur Prénovel	
	<b>Précision</b>	<i>Kappa</i>	<b>Précision</b>	<i>Kappa</i>	<b>Précision</b>	<i>Kappa</i>
kNN (k=7)	<b>72,5 %</b>	42,8 %	<b>87,0 %</b>	29,2 %	<b>82,3 %</b>	34,1 %
LDA	<b>84,3 %</b>	66,5 %	<b>85,2 %</b>	42,7 %	<b>86,7 %</b>	41,9 %
SVM (hyperplan linéaire)	<b>84,3 %</b>	66,5 %	<b>90,7 %</b>	56,6 %	<b>86,7 %</b>	35,4 %
SVM (noyau radial)	<b>82,4 %</b>	63,7 %	<b>81,5 %</b>	35,9 %	<b>80,5 %</b>	20,7 %
Random Forest	<b>78,4 %</b>	56,6 %	<b>85,2 %</b>	35,1 %	<b>79,6 %</b>	25,8 %
<i>Méthode de prédiction quantitative</i>	RMSE Gini	RMSE asymétrie	RMSE Gini	RMSE asymétrie	RMSE Gini	RMSE asymétrie
Random Forest	0,050	0,116	0,081	0,126	0,069	0,146
kMSN (k=5)	0,054	0,124	0,070	0,130	0,070	0,153

### III.1.3.3. Test de robustesse sur le site des Aillons et de Prénovel

Entre la validation croisée réalisée sur le site du Vercors (cf. Tableau 5) et les tests de robustesse sur les Aillons et Prénovel (cf. Tableau 7), les précisions sont globalement du même ordre, mais les kappas et les taux de prédiction des placettes régulières sont bien moins bons dans le cas des tests de robustesse. Dans certains cas, le kappa peut ainsi descendre à 20% et le taux de bonnes classifications des peuplements réguliers à 30%. Les modèles de classification et de prédiction de la structure calibrés sur le Vercors ne sont donc pas totalement robustes à l'application sur des sites où les caractéristiques forestières sont légèrement différentes (cf. § III.1.1.).

Contrairement à la validation croisée, les méthodes de classification basées sur des délimitations linéaires (SVM à hyperplan linéaire et LDA) donnent des résultats légèrement meilleurs que les autres méthodes. Ces méthodes, moins souples, sont donc plus robustes. Les différences entre ces deux grands types de méthodes sont également visibles au niveau de la prédiction à l'échelle cartographique (cf. Figure 25).

FIGURE 25 : COMPARAISON DES CARTES PREDICTIVES LDA ET RANDOM FOREST SUR UNE ZONE DE LA FORET DE PRENOVEL



#### Comparaison avec une re-calibration des modèles sur les sites tests

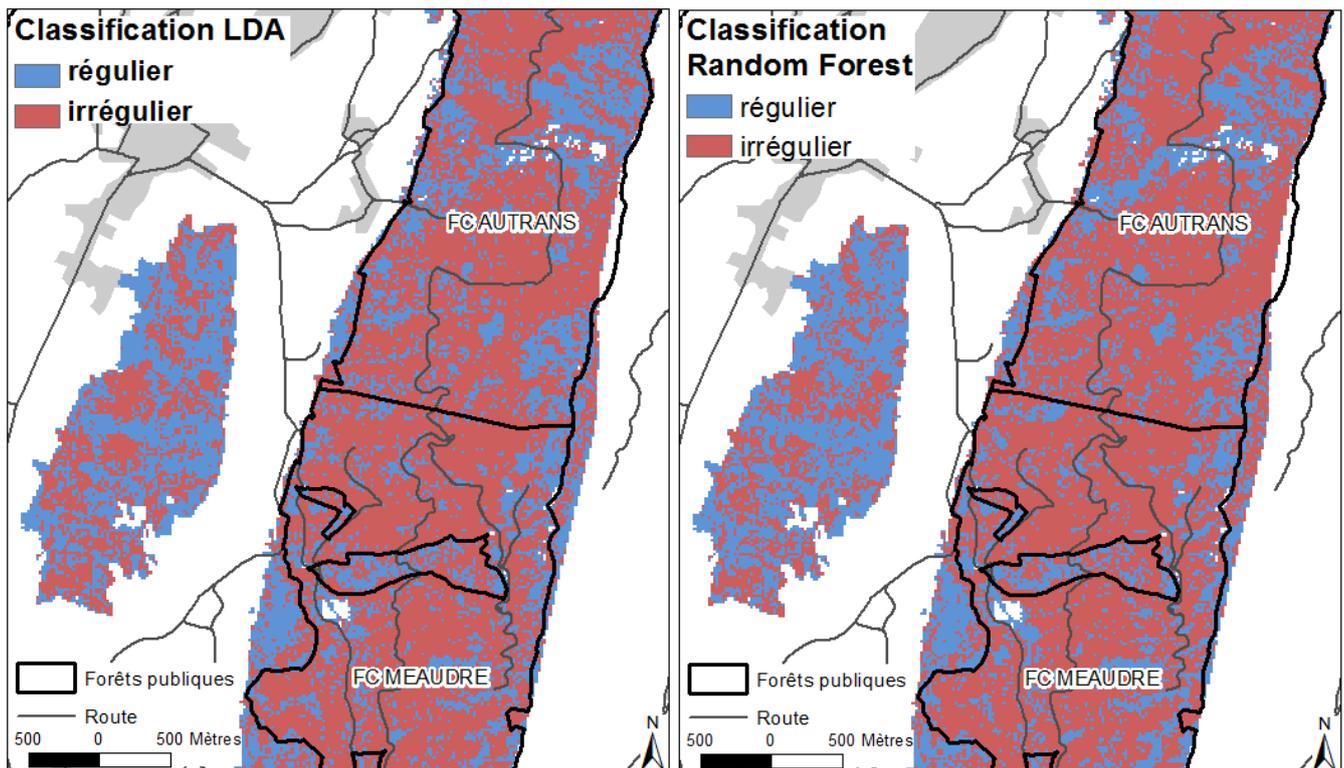
Il est possible de comparer les résultats de ces tests de robustesse avec une validation croisée sur chacun des deux sites tests. Celle-ci a été réalisée en re-calibrant les modèles sur les placettes des Aillons puis de Prénovel avec les métriques définies sur le Vercors. Sur les Aillons, la re-calibration améliore la classification puisqu'en validation croisée la précision moyenne est de 94 % et le kappa moyen de 48 % (contre respectivement 86 % et 40 % pour les tests de robustesse). Pour la forêt de Prénovel, la validation croisée donne des résultats proches de ceux des tests de robustesse (84 % contre 83 % pour la précision, et 29 % contre 28 % pour le kappa).

### III.1.4. Prédications sur l'ensemble du site d'étude

Trois grands types de cartes prédictives ont été réalisés sur le site du Vercors :

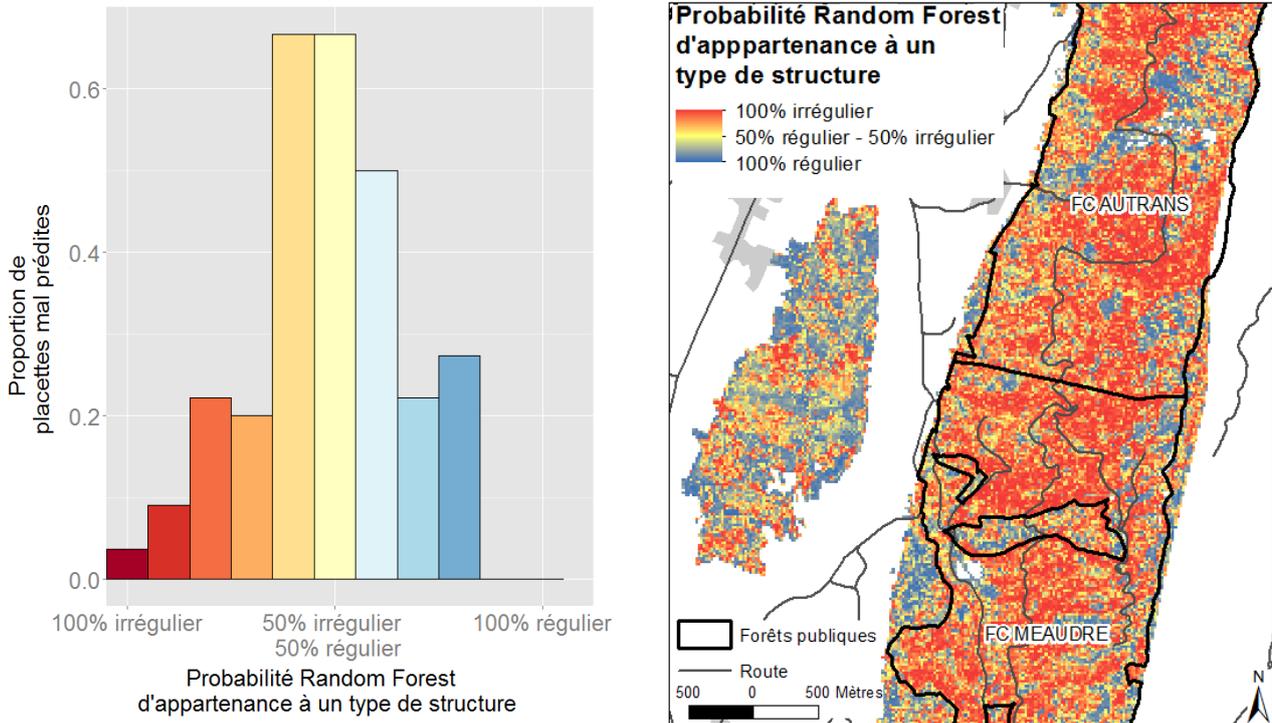
**1. La carte des structures prédites par classification supervisée :** quelles que soient les méthodes de classification utilisées, le résultat est visuellement très proche (cf. Figure 26). La forêt privée se distingue avec une prédominance de la structure régulière, en lien avec un type de gestion fréquent dans ces propriétés, alors que les forêts publiques, majoritairement traitées en futaies irrégulières (Demoulin, 2015), n'ont que de faibles surfaces prédites en régulier.

FIGURE 26 : CARTES PREDICTIVES ISSUES DES CLASSIFICATIONS LDA ET RANDOM FOREST



**2. La cartographie de la probabilité d'appartenance à un type structural :** elle permet de tenir compte de l'incertitude de la prédiction et de représenter des zones qui sont à la frontière entre le régulier et l'irrégulier. En effet, il existe un lien entre erreur de prédiction et probabilité d'appartenance à une classe (cf. graphique de la Figure 27). La carte de la probabilité d'appartenance à un type structural (cf. carte de la Figure 27) permet donc de nuancer la classification directe en deux groupes avec des zones intermédiaires (en jaune) qui représentent souvent des peuplements à la limite des structures régulières et irrégulières.

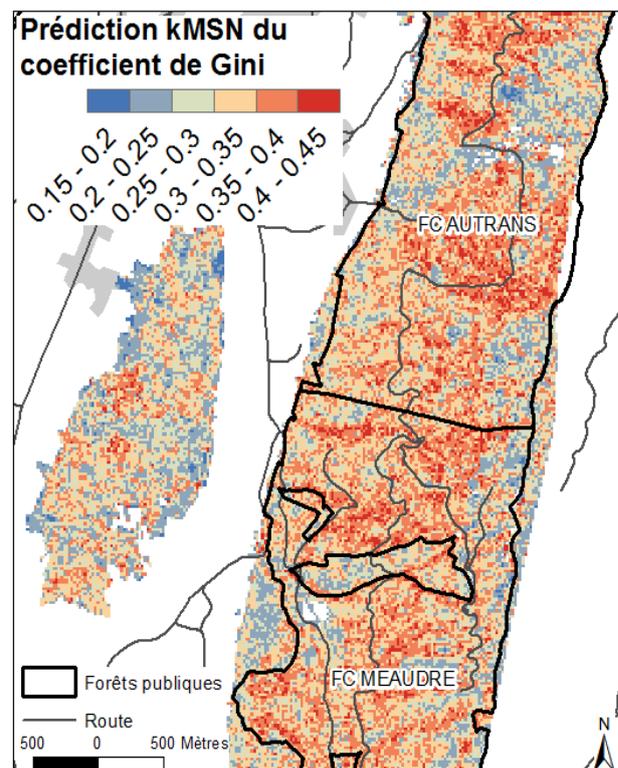
**FIGURE 27 : GRAPHIQUE DU LIEN ENTRE ERREURS ET PROBABILITES DE PREDICTION ET CARTE DE LA PROBABILITE D'APPARTENANCE A UN TYPE DE STRUCTURE**



Le graphique de gauche, réalisé sur un échantillon test de validation croisée dans le Vercors, montre que pour des peuplements dont la prédiction donne une probabilité d'être régulier égale à la probabilité d'être irrégulier (probabilité proche de 50%), le taux d'erreur de classification est relativement plus fort que pour des peuplements dont la prédiction donne une probabilité forte d'appartenir à une classe (probabilité proche de 100%).

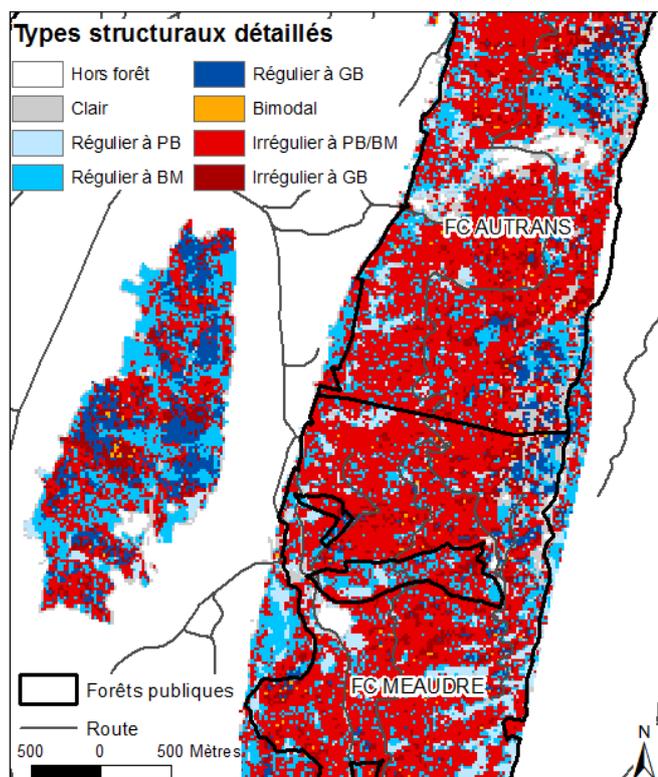
**FIGURE 28 : CARTE DE LA PREDICTION KMSN DU COEFFICIENT DE GINI**

**3. La cartographie des coefficients de Lorenz :** l'indice de Gini permet également de représenter la variabilité au sein des peuplements réguliers et irréguliers. Les peuplements à la frontière entre les deux types ont un indice assez proche de 0,3, alors que des peuplements typiques ont des valeurs de Gini souvent très fortes ou très faibles. Les liens entre ce type de cartographie et la cartographie de probabilité d'appartenance sont donc très forts (cf. Figure 27 et Figure 28). L'association de cette cartographie avec celle de la prédiction de l'asymétrie de Lorenz permet aussi de distinguer les peuplements bimodaux (cf. § II.8.), type de peuplement rarement présent sur le site du Vercors.



### III.1.5. Vers une cartographie opérationnelle des types structuraux

FIGURE 29 : CARTE DETAILLÉE DES TYPES STRUCTURAUX



La carte détaillée des types structuraux (cf. Figure 29), obtenue par combinaison de différents modèles prédictifs et de métriques LiDAR, apporte davantage d'informations sur les peuplements présents sur le site du Vercors. En haut de versant, les peuplements semblent souvent être peu denses ou majoritairement à petits bois, ce qui correspond aux conditions de croissance dans ces zones. D'après cette carte, la forêt privée est composée d'une mosaïque de peuplements essentiellement réguliers à bois moyens ou à gros bois, alors que la forêt publique possède une grande majorité de peuplements irréguliers, dont une petite partie est riche en gros bois.

La comparaison des types obtenus par cette cartographie exploratoire et des types correspondants sur les placettes de validation indépendante a permis d'estimer une **précision de l'ordre de 63 %**. La comparaison avec des cartographies classiques des structures de

peuplements est délicate, celles-ci ne se basant pas sur les mêmes clés de détermination (cf. Annexe 6 et Annexe 7).

## III.2. Feuillus de plaine

Voir avertissement au début du rapport.

### III.2.1. Structure des placettes de terrain

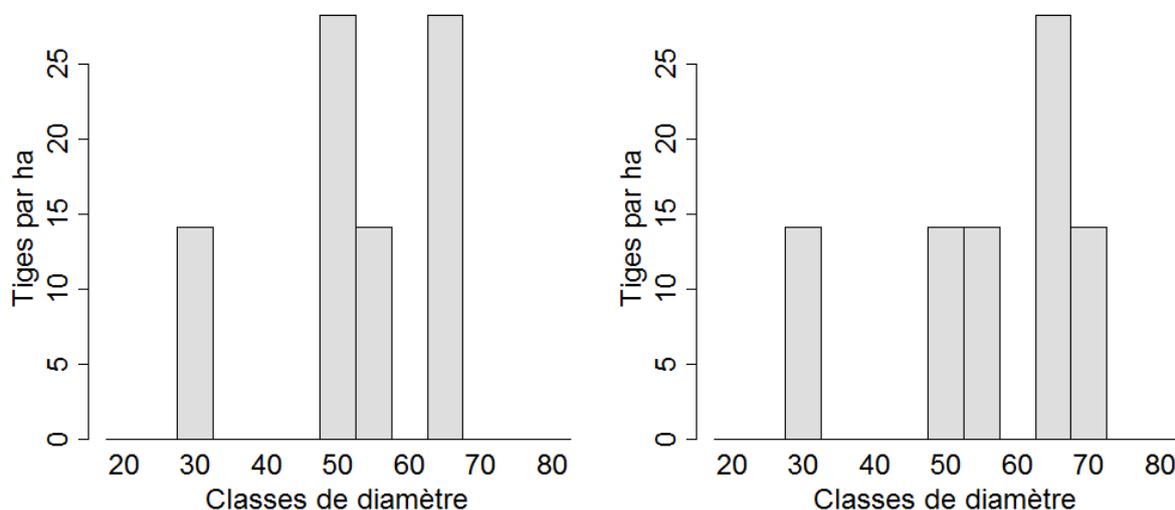
TABLEAU 8 : REPARTITION DES PLACETTES ET CARACTERISTIQUES DENDROMETRIQUES DU SITE DU PNF

	PNF (calibration)	
	Placettes « Gini »	Placettes « typiques »
Placettes à structure <b>irrégulière</b>	21 (21 %)	11 (17 %)
Placettes à structure <b>régulière</b>	80 (79 %)	55 (88 %)
<b>Nombre total de placettes</b>	<b>101</b>	<b>66</b>
Surface terrière* ( <i>m<sup>2</sup>/ha</i> )	32,0 ( $\pm$ 11,7)	
Densité* ( <i>tiges/ha</i> )	363 ( $\pm$ 152)	
Diamètre moyen* ( <i>cm</i> )	34,6 ( $\pm$ 7,3)	

Les 101 placettes du site du PNF sont majoritairement de structure régulière. Par rapport aux sites de montagne, la densité et la surface terrière de ce type de forêt sont beaucoup plus faibles. Avec un rayon de 15m, beaucoup de placettes ne comptabilisent que très peu d'arbres et la proportion de

placettes « claires » atteint ici 19%. De plus, de nombreuses placettes se situent juste au-dessus du seuil (55% des placettes ont une densité comprise entre 80 tiges/ha et 90 tiges/ha) mais ont été conservées dans le jeu de placettes « typiques » afin d’avoir un échantillon de calibration suffisant malgré une classification structurale pas toujours représentative (cf. Figure 30)

FIGURE 30 : EXEMPLE DE DEUX PLACETTES DE STRUCTURE DIFFICILEMENT CARACTERISABLE



Ces deux exemples sont des placettes avec seulement 6 arbres (85 tiges/ha). Elles ont été classées en régulier par les différentes approches de caractérisation, cependant il pourrait s’agir de taillis sous futaie vieillis à structure plutôt bimodale.

Sur ces placettes de faible rayon par rapport à la densité, la caractérisation de la structure présente donc des incertitudes majeures ayant des répercussions sur la modélisation.

### III.2.2. Métriques sélectionnées

La segmentation des apex ne correspond visuellement pas à la réalité de terrain : malgré l’optimisation des paramètres de segmentation, les gros arbres ont tendance à être sur-segmentés alors que les petits arbres ou les brins de taillis sont agglomérés entre eux (cf. Annexe 3).

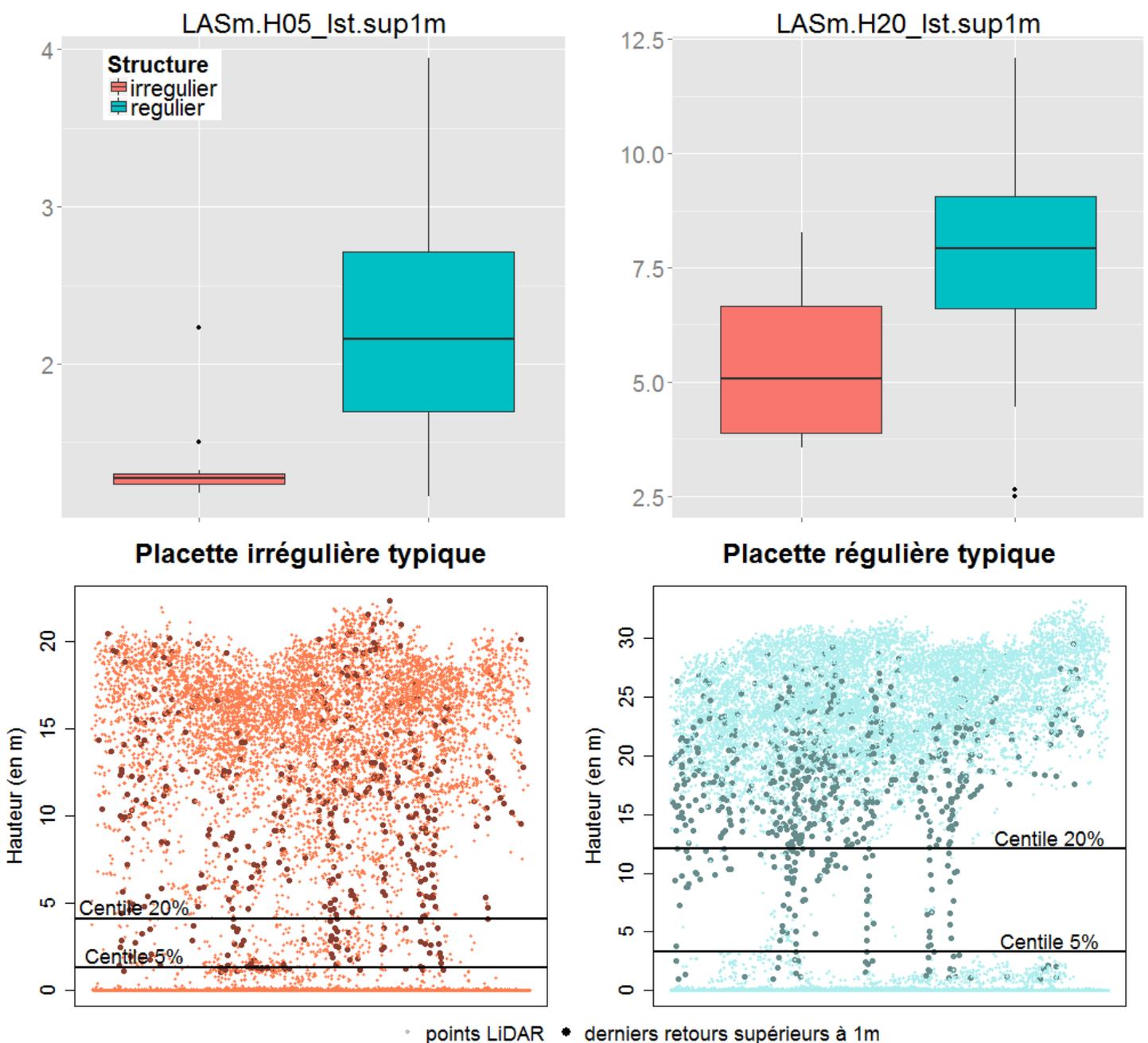
TABLEAU 9 : JEUX DE METRIQUES SELECTIONNES SUR LE SITE DU PNF

Sélection de métriques LiDAR	
LAS.H05_Ist.sup1m	Centile 5 % des hauteurs des derniers retours supérieurs à 1 m des points LiDAR
LAS.H20_Ist.sup1m	Centile 20 % des hauteurs des derniers retours supérieurs à 1 m des points LiDAR
LAS.TxPen.int	Taux de pénétration dans la couche intermédiaire ( $\frac{n_1}{n_2}$ où $n_1$ est le nombre de points inférieur à 0.5 fois le centile 99 % et $n_2$ le nombre de points inférieur à 0.8 fois le centile 99 %)
Sélection de métriques « Apex-MNH »	
MNH.cov.med	Pourcentage de pixels du MNH dont la hauteur est intermédiaire (comprise entre 15 et 25 m)
Apex.seuil0.25_min_HT	Minimum des hauteurs d’apex (filtre supprimant les couronnes de surface inférieure à 0.25 fois la hauteur de l’apex)
Apex.seuil0.5_sd_Vol	Ecart-type des volumes de couronnes (filtre supprimant les couronnes de surface inférieure à 0.5 fois la hauteur de l’apex)

En conséquence, les métriques « Apex » ne ressortent que très peu dans les sélections VSURF. Un jeu de métriques « Apex-MNH » a tout de même été conservé et testé (cf. Tableau 9).

Les premiers centiles de hauteur des derniers retours LiDAR (*ex LAS.H05\_Ist.sup1m et LAS.H20\_Ist.sup1m*) se sont révélés être les variables les plus importantes pour la prédiction de la structure dans le site du PNF : les derniers retours supérieurs à 1m (sans le sol) représentant généralement assez bien le sous-étage, les valeurs des centiles sont plus faibles dans le cas de l'irrégulier et du taillis sous futaie où le sous-étage est relativement important (cf. Figure 31). Cependant, cette distinction ne peut être efficace que si les peuplements réguliers sont relativement mûrs, ce qui est le cas du jeu de placettes de calibration (seules deux placettes sont de structure taillis ou régulière à petits bois).

FIGURE 31 : VARIATION DES VALEURS DES CENTILES 5% ET 20% DES PREMIERS RETOURS SUR LES PLACETTES "TYPIQUES" DU PNF EN FONCTION DE LEUR STRUCTURE



### III.2.3. Validation et erreurs de prédiction

Les tableaux complets des résultats de validation sont disponibles en Annexe 5. Les tableaux 10 et 11 présentent les résultats obtenus à partir de la sélection de métriques « LiDAR ».

#### III.2.3.1. Validation croisée

##### Classification supervisée

Les résultats de validation croisée sur le PNF sont globalement moins bons que sur le site du Vercors. En effet, si les taux de bonnes classifications sont proches de ceux du Vercors (autour de 85%), les kappas sont eux beaucoup plus faibles (environ 35% contre 60%)(cf. Tableau 10).

**TABEAU 10 : RESULTATS DE LA VALIDATION CROISEE DES METHODES DE CLASSIFICATION SUPERVISEE SUR LE SITE DU PNF**

	Validation croisée sur les placettes « typiques »		Validation croisée sur le jeu de placettes « Gini »	
	Précision	Kappa	Précision	Kappa
kNN (k=7)	<b>81,6%</b>	18,0%	<b>78,7%</b>	18,8%
LDA	<b>84,3%</b>	30,1%	<b>83,9%</b>	40,3%
SVM (hyperplan linéaire)	<b>85,5%</b>	40,4%	<b>84,0%</b>	40,5%
SVM (noyau radial)	<b>85,1%</b>	25,1%	<b>82,9%</b>	31,1%
Random Forest	<b>91,9%</b>	54,3%	<b>89,7%</b>	55,7%

##### Prédiction des coefficients de Gini et de l'asymétrie de Lorenz

En prédiction des coefficients de Lorenz, les RMSE de validation croisée sont également moins bonnes que sur le site du Vercors (cf. Tableau 11).

**TABEAU 11 : RESULTATS DE LA VALIDATION CROISEE DES METHODES DE PREDICTION QUANTITATIVE SUR LE SITE DU PNF**

	Validation croisée sur les placettes « typiques »				Validation croisée sur le jeu de placettes « Gini »			
	Prédiction des indicateurs de Lorenz		Classification à partir des coefficients prédits		Prédiction des indicateurs de Lorenz		Classification à partir des coefficients prédits	
	RMSE Gini*	RMSE asymétrie*	Précision	Kappa	RMSE Gini*	RMSE asymétrie*	Précision	Kappa
Random Forest	0,079	0,196	<b>86,8%</b>	46,1%	0,087	0,192	<b>86,8%</b>	50,6%
kMSN (k=5)	0,076	0,198	<b>86,0%</b>	48,4%	0,094	0,196	<b>82,7%</b>	39,9%

\* Le coefficient de Gini varie entre 0 et 1 en théorie (plutôt entre 0.1 et 0.6 en pratique) et l'asymétrie de Lorenz entre 0 et 2. En relatif, ces RMSE sont donc de l'ordre de 25% pour le coefficient de Gini et de 20% pour l'asymétrie de Lorenz.

##### Influence du jeu de calibration

Avec seulement 11 placettes « typiquement » irrégulières, le jeu d'apprentissage est très déséquilibré, entraînant une très mauvaise prédiction de la structure irrégulière (cf. Tableau 8). Ainsi, en validation croisée, seulement **42% des placettes irrégulières** sont bien prédites contre **95% pour**

**les placettes régulières** (cf. Annexe 5). Ce net sur-apprentissage sur la structure régulière entraîne un biais de prédiction risquant de masquer la présence de peuplements irréguliers.

Si le nombre de placettes était suffisant, une solution serait de calibrer le modèle sur un échantillon équilibré constitué des placettes irrégulières et de certaines placettes régulières tirées au sort.

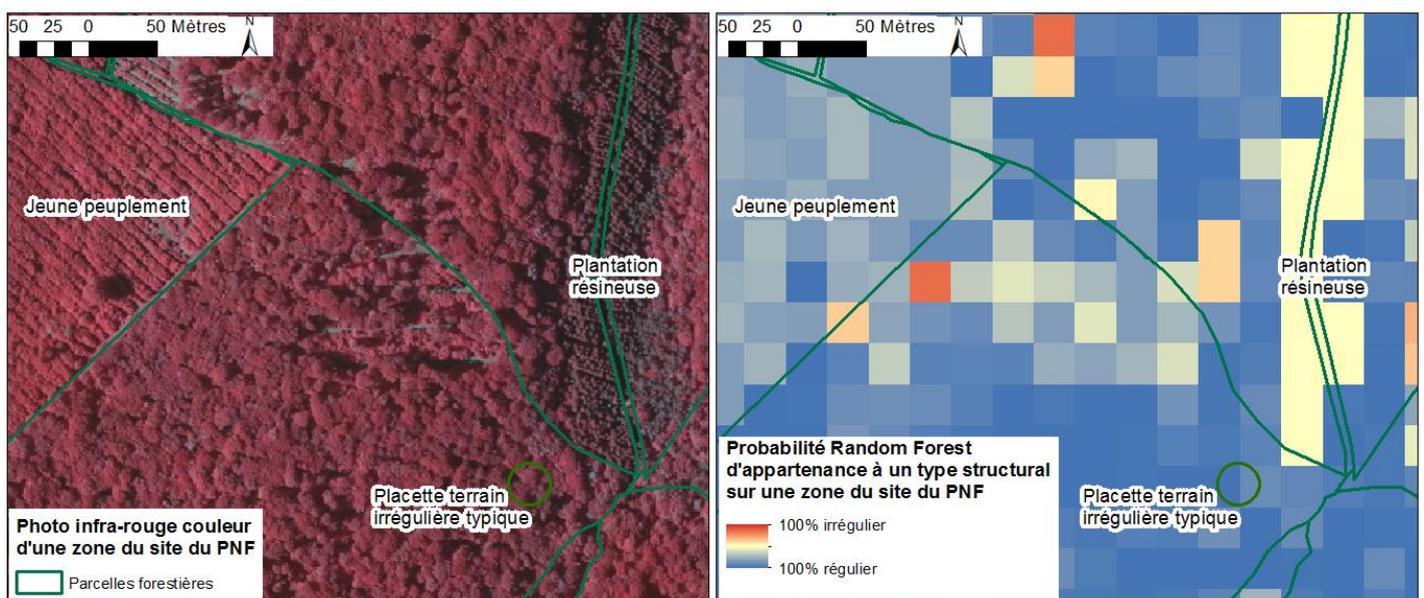
#### Métriques « Apex-MNH »

En lien avec la délicate segmentation des apex sur ce site feuillu, les résultats de validation croisée n'utilisant que des métriques « Apex-MNH » sont décevants (précision entre 75 et 80 % et kappa entre 0 et 25 %) (cf. Annexe 5). Dans ce type de site volé hors feuille, les algorithmes actuels de délimitation de couronnes à partir des MNH ne sont donc pas adaptés à une prédiction uniquement basée sur des métriques « Apex-MNH ».

### III.2.4. Prédictions sur l'ensemble du site d'étude

Comme pressenti avec l'analyse des résultats de validation croisée, la prédiction cartographique ne semble pas correspondre à la réalité de terrain, classant des zones probablement irrégulières en peuplements réguliers (cf. Figure 32).

**FIGURE 32 : COMPARAISON ENTRE LA PHOTO AERIEENNE ET LA PROBABILITE PREDITE D'APPARTENANCE A UN TYPE STRUCTURAL SUR UNE ZONE DU SITE DU PNF**



Très peu de zones sont prédites en irrégulier sur l'ensemble de la forêt. Par ailleurs, l'absence de placettes de calibration dans les parcelles en régénération et dans les plantations résineuses font que ces peuplements ne sont prédits en régulier qu'avec une relativement faible probabilité (cf. Figure 32). Ces zones hors calibration seraient à masquer de la prédiction (assez facilement grâce à la hauteur LiDAR des peuplements).

## IV. Discussion et perspectives

### IV.1. Limites de l'étude

#### IV.1.1. Surface de perception de la structure des peuplements

Comme nous l'avons vu, notamment pour le site feuillu du PNF, la surface d'inventaire des placettes de calibration influence nettement la caractérisation et la modélisation de la structure des peuplements.

Les typologies, entre autres celle du Massif Vosgien, se basent généralement sur des données issues d'inventaires relascopiques (Asael et al., 1999). Le passage de placettes à angle fixe (relascopiques) à des placettes circulaires de rayon de 15 m (comme dans l'étude) engendre quelques incertitudes. En effet, sur des placettes à angle fixe, les gros bois sont pris en compte sur une surface beaucoup plus importante (jusqu'à 25 m de rayon pour un arbre de 50cm de diamètre). Dans des anciens taillis sous futaie avec prédominance de gros bois comme dans le PNF, la différence avec les placettes de 15 m de rayon est importante.

De même, le coefficient de Gini calculé sur une placette n'est qu'un estimateur de la diversité des diamètres au niveau du peuplement. Pour de petits échantillons, celui-ci devient alors biaisé. Dixon et al. (1987) ont estimé que le coefficient de Gini devenait suffisamment précis s'il était calculé à partir d'au moins 50 individus. La majorité des placettes du PNF comprenant moins de six arbres, le risque d'erreur sur la caractérisation du peuplement à partir de ce coefficient est fort.

L'utilisation de placettes circulaires de rayon plus important ou de grappes de placettes permettrait certainement d'améliorer la caractérisation de la structure des placettes de calibration et, par enchaînement, la qualité des prédictions. Cependant, cela induirait également une augmentation des temps et des coûts d'inventaire des placettes de référence terrain.

#### IV.1.2. Répartition des placettes de calibration

La répartition des placettes de calibration entre types structuraux joue un rôle essentiel lors de la phase d'apprentissage. Une surreprésentation d'une classe dans le jeu de calibration engendre un biais en faveur de cette classe lors de la prédiction. Or, en pratique, c'est souvent le type de structure minoritaire que l'on cherche à détecter et à cartographier.

Afin d'avoir un jeu de données équilibré, l'idéal serait de stratifier l'échantillonnage. Cela nécessite néanmoins une connaissance a priori de la cartographie des structures de peuplements alors que c'est le but même de la prédiction. L'usage de la photo-interprétation ou de cartographies préexistantes pourrait alors être une solution. Si le jeu contient un nombre suffisant de placettes, une autre possibilité consiste à réduire le nombre de placettes dont le type est surreprésenté.

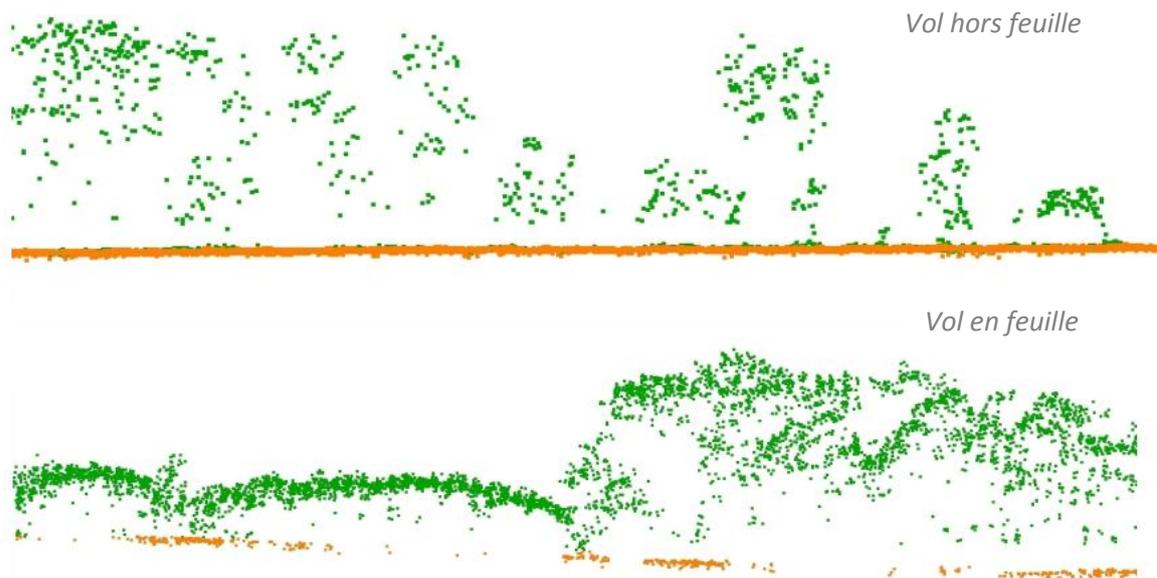
### IV.1.3. Segmentation des apex pour les feuillus

La moins bonne qualité des modèles réalisés sur le site feuillu du PNF est en partie liée à la difficulté de détection des couronnes sur le MNH LiDAR, avec une nette sur-segmentation (cf. Annexe 3).

Cette difficulté est principalement liée à :

- L'architecture des feuillus par rapport aux conifères : alors que les houppiers de conifères sont coniques et rarement fourchus, ceux de feuillus sont beaucoup plus arrondis et se divisent en plusieurs parties issues des différentes charpentières. Les ruptures de pentes entre les houppiers feuillus sont donc plus faibles et se distinguent parfois difficilement des irrégularités au sein même du houppier.
- La saison du vol LiDAR : les vols réalisés hors saison de végétation (cas du vol sur le site du PNF) ont des taux de pénétration du signal beaucoup plus importants que les vols en feuille (cf. Figure 33). Un avantage de ces vols hors feuille est de fournir des MNT de bonne précision. Cependant, au niveau de la canopée, les points sont relativement en faible densité et ne représentent que les branches et non la totalité du houppier. Les MNH qui en sont issus sont donc plus difficiles à segmenter.

FIGURE 33 : COMPARAISON DE DEUX VOLS LIDAR DE MEME DENSITE D'EMISSION (4 PTS/M<sup>2</sup>) REALISES EN HIVER ET EN ETE (D'APRES CLEMENT MALLET)



L'usage de MNH dérivés de vols en feuille est une piste pour améliorer la détection des arbres et par conséquent la qualité des modèles sur des sites feuillus. D'autres pistes consistent en l'utilisation directe du nuage de points des vols hors feuille. En effet, si la canopée est relativement mal représentée, les informations sur la partie « basse » de l'arbre sont plus importantes. Des algorithmes comme PTree utilisant l'ensemble du nuage de points (Vega et al., 2014) ou des méthodes de recherche des troncs dans le nuage (avec la densité d'émission du vol LiDAR sur le PNF, les troncs se distinguent nettement, cf. Figure 31) pourraient améliorer la détection des arbres.

## **IV.2. Evolutions envisageables**

### **IV.2.1. Amélioration des métriques LiDAR**

Trois grands types de métriques LiDAR, issues du nuage de points normalisé en hauteur, du MNH et de la segmentation des apex, ont été utilisés dans cette étude. Certains chercheurs ont également expérimenté d'autres pistes pour calculer des métriques reflétant la structure des peuplements.

Compte-tenu de la pénétration du signal LiDAR, la densité de points au niveau du sous-étage dépend de la quantité de biomasse au niveau supérieur de la canopée. L'information sur le sous-étage est donc variable, ce qui impacte la déduction de la stratification du peuplement. Afin de déduire des profils de végétation plus réalistes à partir du nuage de points, une correction ou normalisation de ce dernier est envisageable. En utilisant la loi de Beer-Lambert d'absorption de la lumière et la probabilité de vide au-dessus d'une hauteur  $z$ , il est possible de déterminer l'indice foliaire à la hauteur  $z$ , puis, en le cumulant sur toutes les hauteurs, le profil de végétation complet. De ce profil corrigé, différentes métriques peuvent ensuite être calculées (Bouvier et al., 2015; Coops et al., 2007).

Les systèmes LiDAR enregistrant l'onde retour complète ont aussi été utilisés afin de créer des métriques reliées à la stratification. Lefsky et al (1999) ont ainsi calculé des volumes de canopée à partir des ondes retour puis des fractions de volumes pour différentes zones de la végétation (trouées couvertes, zone oligophotique, euphotique et trouées ouvertes). Pour appliquer ces métriques avec des systèmes LiDAR à retour discret, une fonction d'interpolation permettant de convertir les points en onde est utilisable (Muss et al., 2011). Il est également possible d'utiliser des voxels (pixels en trois dimensions) qui seront classés dans les différentes zones (Coops et al., 2007).

L'ajout de ces métriques issues d'une modification du nuage de points permettrait éventuellement d'améliorer les modèles de prédiction, même si les statistiques classiques du nuage de points restent celles qui sont le plus liées à la structure (Nijland et al., 2015). Par ailleurs, il n'existe pas toujours de méthode ou d'outil simple et utilisable sur des grandes surfaces pour calculer ces métriques.

### **IV.2.2. Utilisation de photos aériennes ou satellitaires**

Si les données LiDAR aéroporté offrent une information tridimensionnelle de la forêt, elles restent relativement coûteuses. A l'avenir, d'autres types de données pourraient être accessibles de façon beaucoup moins onéreuse et sur de plus grandes surfaces. Par exemple, il serait possible d'obtenir un MNH grâce à la combinaison de données de photogrammétrie et du MNT rénové de l'IGN. Les tests réalisés au cours de ce stage avec la sélection de métriques « Apex-MNH » explorent cette éventualité, même si le MNH de cette étude est issu du nuage de points et non de données externes.

En dehors de la photogrammétrie, les photos aériennes ou satellitaires peuvent être utilisées pour la prédiction des paramètres structuraux comme unique source d'information (Ozdemir et Karnieli, 2011) ou en complément des informations LiDAR (Packalén et Maltamo, 2008). Dans ce cas, ce sont essentiellement des variables texturales issues de différentes bandes spectrales qui sont utilisées (écart-types des niveaux de gris, contraste, entropie...). Ozdemir et Karnieli (2011) ont ainsi prédit le coefficient de Gini avec une RMSE de prédiction de l'ordre de 0,05 (similaire aux erreurs de cette étude). Cependant, malgré le faible coût et la relative disponibilité de ces photos, certains freins existent à leur utilisation, comme l'absence d'information sur l'intérieur du peuplement ou les

phénomènes d'ombrages parfois importants, surtout en montagne. Le LiDAR semble prédire de façon plus performante les paramètres structuraux (Lefsky et al., 2001b).

#### IV.2.3. Application de la méthode kMSN pour la prédiction de la répartition des tiges

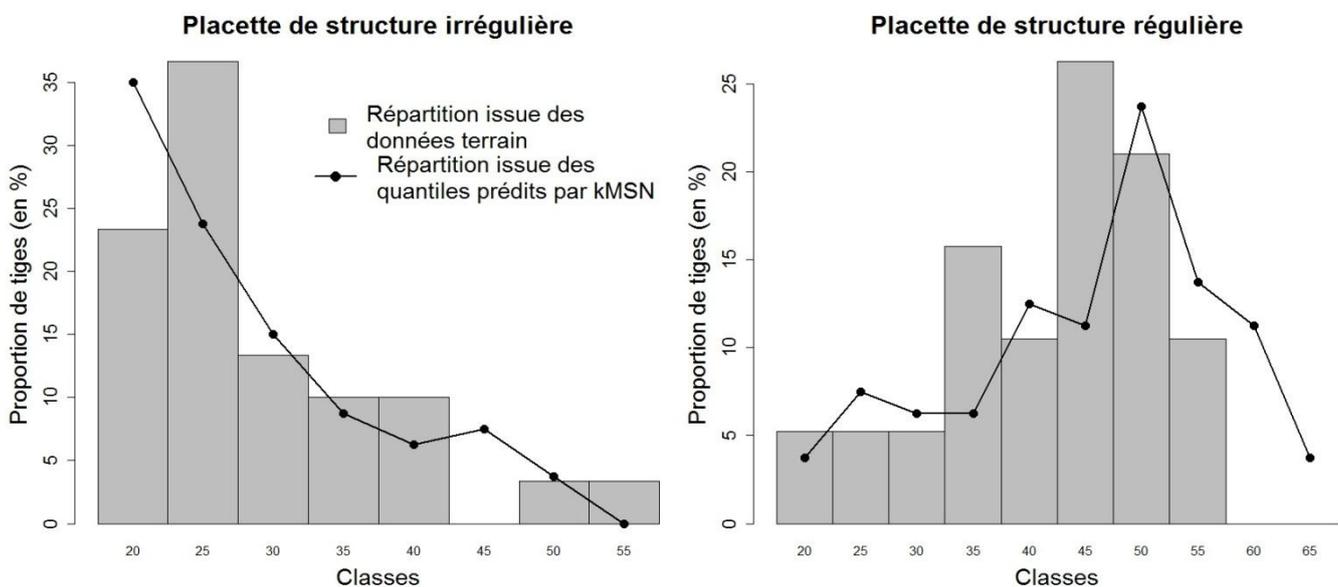
Pour l'estimation simultanée de plusieurs variables pouvant être corrélées, la méthode kMSN présente un certain intérêt. Elle a donc été largement employée dans le cadre de la prédiction de la structure et plus particulièrement pour la répartition des tiges (cf. § I.3.4. et Annexe 1). Prédire la répartition des tiges plutôt qu'un type structural présente plusieurs avantages :

- Il n'est pas nécessaire de déterminer la structure des placettes de calibration, la seule répartition des tiges suffit
- N'importe quelle typologie peut être appliquée à la répartition prédite des tiges, s'adaptant ainsi au contexte local
- Un nombre plus important d'informations peut être extrait de cette prédiction, étant donné que l'ensemble de la répartition est disponible

Cette prédiction de la répartition des diamètres par la méthode kMSN peut être abordée par différentes approches : la prédiction directe de la liste des diamètres, la prédiction de quantiles de la répartition, la prédiction de paramètres de lois de distribution (cf. § I.3.4. et Annexe 1).

La prédiction kMSN de quatre quantiles de la répartition des tiges à partir de métriques LiDAR a été expérimentée sur les sites du Vercors et du PNF. De ces quantiles prédits (quantiles 25, 50, 75 et 99 %), la répartition complète peut être interpolée et un type de structure peut être déduit (cf. Figure 34).

FIGURE 34 : EXEMPLES DE PREDICTION KMSN DE LA REPARTITION DES TIGES SUR DEUX PLACETTES "CIBLES" DU VERCORS



Une fois cette répartition convertie en type régulier ou irrégulier, les résultats sont légèrement moins bons que dans le reste de l'étude (respectivement les précisions et les coefficients kappa sont de 87 % et 70 % en validation croisée sur le site du Vercors, de 67 % et 52 % en validation indépendante sur le site du Vercors, et de 73 % et 14 % en validation croisée sur le site du PNF).

Même si des améliorations et des analyses plus approfondies seraient nécessaires (plus de quantiles prédits, application à la répartition en surface terrière...), cet exemple reflète l'intérêt de la méthode kMSN et de la prédiction de la répartition des diamètres. Comme le montre cet exemple et pour les raisons évoquées précédemment, cette approche pourrait être une piste intéressante de prédiction des types structuraux.

### **IV.3. Perspectives d'utilisation**

#### **IV.3.1. Intérêts pour l'aménagiste et le gestionnaire**

D'un point de vue opérationnel, il existe deux grands intérêts à la prédiction et à la cartographie des types structuraux :

- La stratification d'inventaire : cette stratification permet d'adapter l'échantillonnage à chaque bloc d'inventaire ou type de peuplement, selon leur surface et leur variance. Dans le cas du site du PNF par exemple, l'utilisation de blocs d'inventaires est une solution pour la description précise des grands types de peuplements présents (ancien taillis sous futaie, peuplement irrégulier et stades de développement des peuplements réguliers). En effet, les peuplements irréguliers étant minoritaires et les taillis sous futaie présentant une grande variance, l'effort d'échantillonnage nécessaire à une bonne description de ces peuplements est plus important que dans les peuplements réguliers. La modélisation de la structure issue d'une étape de calibration nécessitant déjà des placettes d'inventaire, l'adaptation de l'échantillonnage à partir de cette modélisation consisterait alors plutôt en l'ajout de placettes dans des zones relativement peu inventoriées.
- L'application de règles de gestion adaptées : de la connaissance du type structural et des objectifs sylvicoles découlent les préconisations d'intervention. Classiquement, les documents de gestion et les aménagements incluent la description de chaque type de peuplement, les règles de gestion correspondantes, et la cartographie de ces différentes zones. Dans le cas du site du Vercors par exemple, un recensement des « patchs » de peuplements réguliers est en cours afin d'adapter les mesures de gestion courante à ces peuplements.

La cartographie prédictive pourrait répondre à ces enjeux. Cependant, la résolution fine des rasters de prédiction (pixels de 20 ou 30 m dans cette étude) ne concorde pas avec l'échelle de description courante de la structure qui est généralement supérieure au quart d'hectare (Gauquelin, 2015, *comm. pers.*). L'agrégation ou le lissage de ces rasters paraît donc nécessaire afin de rendre la cartographie plus lisible et plus exploitable. Même si ces opérations modifient plus ou moins légèrement les résultats de la prédiction, elles permettent de mieux faire ressortir les zones de structures différentes en atténuant les pixels divergeant de la tendance globale de la zone (cf. Annexe 6).

#### **IV.3.2. Amélioration d'autres modélisations**

La connaissance de la structure des placettes de prédiction pourrait également permettre d'améliorer d'autres modélisations issues de données LiDAR.

La structure est un facteur influençant de nombreux autres paramètres forestiers comme la densité, la surface terrière, le diamètre moyen ou dominant. La calibration de différents modèles par grand

type structural (un modèle pour l'irrégulier et un pour le régulier par exemple) permettrait éventuellement de réduire les erreurs de prédiction par rapport à un modèle unique. L'ajout d'un indice de structure en variable explicative des modèles peut jouer un rôle semblable.

L'information sur la structure pourrait également être utile pour améliorer les relations allométriques utilisées dans l'approche « arbre ». En effet, la relation entre la hauteur des arbres et leur diamètre dépend non seulement de l'essence, de la station, mais aussi du type de peuplement en question. En tenant compte de cette structure dans les relations allométriques, un gain de précision sur la prédiction des diamètres en fonction des hauteurs d'apex détectés pourrait éventuellement être obtenu.

## Conclusion

L'intérêt de cette étude est d'expérimenter la prédiction de types structuraux à partir de données LiDAR. Selon les objectifs de prédiction, différentes méthodes sont disponibles, allant de la segmentation a priori pour stratifier un massif en zones de structure homogène, à la prédiction complète de la répartition des tiges. Afin de correspondre aux définitions classiques des types de structures utilisées sur le terrain, le choix méthodologique s'est ici porté sur une calibration de modèles à partir de données de référence. Ces modèles ont permis de distinguer les peuplements irréguliers des peuplements réguliers, information de base sur la structure pouvant facilement être complétée pour obtenir des types structuraux plus précis.

Dans chacun des deux grands contextes forestiers étudiés, les modèles ont, en moyenne, permis de prédire correctement la structure régulière ou irrégulière de 85 % des placettes de validation croisée. Ce résultat varie légèrement en fonction des approches de caractérisation de la structure et des méthodes statistiques de prédiction utilisées. Dans le contexte des résineux de montagne, la validation indépendante et les tests de robustesse montrent une certaine fiabilité et répliquabilité des modèles (environ 83 % des placettes sont alors bien prédites). S'il reste à réaliser ces tests dans le contexte des feuillus de plaine, la cartographie prédictive ainsi que les coefficients kappa laissent présager une moins bonne qualité des modèles pour ce type de forêt.

Des évolutions, notamment sur les inventaires des placettes de référence, permettraient certainement d'améliorer les résultats des modèles de prédiction. En effet, une meilleure répartition de l'échantillon de calibration entre grands types éviterait le biais important des modèles, et une surface plus importante d'inventaire faciliterait la caractérisation de la structure. Les cartes prédictives de la structure, associées à d'autres types de prédictions issues de données LiDAR aéroporté, pourraient alors compléter efficacement les inventaires terrain réalisés dans le cadre d'aménagements et constituer une source d'information notable pour le gestionnaire.

## Bibliographie

- Abt, D., 2014. Guide des sylvicultures - Arc jurassien - Sapin et épicéa, ONF. ed.
- 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 aéroporté pour l'estimation des variables dendrométriques de placettes. Rev. Fr. Photogrammétrie Télédétection in press.
- Asael, S., Ancel, P., Lacombe, E., Wilhelm, M.-E., 1999. Peuplements forestiers du massif Vosgien - Typologie et sylvicultures.
- Bachelet, F., Ancel, P., Leclerc, D., 2002. Peuplements forestiers feuillus du Plateau Lorrain - Typologies et sylvicultures.
- Bailey, R.L., Dell, T.R., 1973. Quantifying Diameter Distributions with the Weibull Function. For. Sci. 19, 97–104.
- Bastien, Y., Gauberville, C., 2011. Vocabulaire forestier. Écologie, gestion et conservation des espaces boisés. Institut pour le développement forestier, Paris.
- Bock, J., Munoz, A., Fay, J., 2014. Précision des mesures en forêt.
- Bollandsås, O.M., Maltamo, M., Gobakken, T., Næsset, E., 2013. Comparing parametric and non-parametric modelling of diameter distributions on independent data using airborne laser scanning in a boreal conifer forest. Forestry 86, 1–9.
- Bollandsås, O.M., Næsset, E., 2007. Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. Scand. J. For. Res. 22, 33–47.
- Bonnet, S., Toromanoff, F., Bauwens, S., Michez, A., Dedry, L., Lejeune, P., 2013. Principes de base de la télédétection et ses potentialités comme outil de caractérisation de la ressource forestière—II. LiDAR aérien. For. Wallonne 28–41.
- Bouchon, J., 1979. Structure des peuplements forestiers. Ann. Sci. For. 36, 175–209.
- 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 Sens. Environ. 156, 322–334.
- Bruciamacchie, M., 2001. Les typologies de peuplements, 20 ans après. Rev. For. Fr. 53, 3–43.
- Burk, T.E., Newberry, J.D., 1984. Notes: A Simple Algorithm for Moment-Based Recovery of Weibull Distribution Parameters. For. Sci. 30, 329–332.
- Cemagref, CRPF Rhône-Alpes, ONF, 2006. Guide des sylvicultures de montagne - Alpes du Nord françaises.
- Chirici, G., Scotti, R., Montagni, A., Barbati, A., Cartisano, R., Lopez, G., Marchetti, M., McRoberts, R.E., Olsson, H., Corona, P., 2013. Stochastic gradient boosting classification trees for forest fuel types mapping through airborne laser scanning and IRS LISS-III imagery. Int. J. Appl. Earth Obs. Geoinformation 25, 87–97.
- Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J.A. (Tony), 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. Trees 21, 295–310.
- Cordonnier, T., Dreyfus, P., Trouvé, R., 2012. Quelles dimensions et quels indices d'hétérogénéité privilégier pour l'expérimentation dans les peuplements forestiers mélangés ou irréguliers? Rev. For. Fr. LXIV, 773–787.
- Corripio, J.G., 2014. insol: Solar Radiation.
- Crookston, N.L., Finley, A.O., 2007. yalmp: An R Package for kNN Imputation. J. Stat. Softw. 23.
- Damgaard, C., Weiner, J., 2000. Describing inequality in plant size or fecundity. Ecology 81, 1139–1142.
- Demoulin, G., 2015. Aménagement forestier - Forêt communale de Méaudre. ONF, Isère.
- Dickinson, Y., Zenner, E.K., Miller, D., 2014. Examining the effect of diverse management strategies on landscape scale patterns of forest structure in Pennsylvania using novel remote sensing techniques. Can. J. For. Res. 44, 301–312.
- Dixon, P.M., Weiner, J., Mitchell-Olds, T., Woodley, R., 1987. Bootstrapping the Gini Coefficient of Inequality. Ecology 68, 1548–1551.
- Duduman, G., 2011. A forest management planning tool to create highly diverse uneven-aged stands. Forestry 84, 301–314.
- ESRI, 2010. ArcGIS Desktop. Environmental Systems Research Institute, Redlands, Canada.
- Gachet, G., Junod, P., 2008. Apports et perspectives de la technologie LiDAR aéroporté pour la pratique forestière | Contributions and perspectives of airborne LiDAR technology for forestry. Schweiz. Z. Forstwes. 159, 19–30.
- Gaudin, S., Jenner, X., 2001. Typologie des peuplements feuillus et IFN. Rev. For. Fr. LIII, 459–467.
- Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2014. VSURF: Variable Selection Using Random Forests.

- 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.
- Jaskierniak, D., Lane, P.N., Robinson, A., Lucieer, A., 2011. Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. *Remote Sens. Environ.* 115, 573–585.
- Kane, V.R., McGaughey, R.J., Bakker, J.D., Gersonde, R.F., Lutz, J.A., Franklin, J.F., 2010. Comparisons between field- and LiDAR-based measures of stand structural complexity. *Can. J. For. Res.* 40, 761–773.
- Kangas, A., Maltamo, M., 2000a. Performance of percentile based diameter distribution prediction and Weibull method in independent data sets. *Silva Fenn.* 34, 381–398.
- Kangas, A., Maltamo, M., 2000b. Percentile based basal area diameter distribution models for Scots pine, Norway spruce and birch species. *Silva Fenn.* 34.
- 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, in: *Silvilaser International Conference on Lidar Applications for Assessing Forest Ecosystems*, Beijing, China. pp. 125–128.
- Kuhn, M., 2015. *caret: Classification and Regression Training*.
- Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., 1999. Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. *Remote Sens. Environ.* 70, 339–361.
- Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2009. Lidar Remote Sensing for Ecosystem Studies. *BioScience* 19–30.
- Lefsky, M.A., Cohen, W.B., Spies, T.A., 2001a. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. For. Res.* 31, 78–87.
- Lefsky, M.A., Cohen, W.B., Spies, T.A., 2001b. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. For. Res.* 31, 78–87.
- LeMay, V., Staudhammer, C., 2005. Indices of stand structural diversity: Mixing discrete, continuous, and spatial variables, in: *Proceedings of the IUFRO Sustainable Forestry in Theory and Practice: Recent Advances in Inventory & Monitoring, Statistics, Information & Knowledge Management, and Policy Science Conference*. p. 4.
- Leppänen, V.J., Tokola, T., Maltamo, M., Mehtätalo, L., Pusa, T., Mustonen, J., 2008. Automatic delineation of forest stands from LIDAR data. *GEOBIA* 05–08.
- Lesur, D., 2011. Rapport d'étude sur l'utilisation des données Lidar (Light detection and ranging) pour la caractérisation des milieux forestiers de Mayotte. Conseil Général de Mayotte.
- Lexerød, N.L., Eid, T., 2006. An evaluation of different diameter diversity indices based on criteria related to forest management planning. *For. Ecol. Manag.* 222, 17–28.
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News* 2, 18–22.
- Liu, F., Li, F., Zhang, L., Jin, X., 2014. Modeling diameter distributions of mixed-species forest stands. *Scand. J. For. Res.* 29, 653–663.
- Maier, B., Tiede, D., Dorren, L., 2008. Characterising mountain forest structure using landscape metrics on LiDAR-based canopy surface models, in: Blaschke, T., Lang, S., Hay, G.J. (Eds.), *Object-Based Image Analysis, Lecture Notes in Geoinformation and Cartography*. Springer Berlin Heidelberg, pp. 625–643.
- Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyypä, J., Vehmas, M., 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sens. Environ.* 90, 319–330.
- Maltamo, M., Næsset, E., Bollandsås, O.M., Gobakken, T., Packalén, P., 2009. Non-parametric prediction of diameter distributions using airborne laser scanner data. *Scand. J. For. Res.* 24, 541–553.
- Maltamo, M., Næsset, E., Vauhkonen, J. (Eds.), 2014. *Forestry Applications of Airborne Laser Scanning, Managing Forest Ecosystems*. Springer Netherlands, Dordrecht.
- Maltamo, M., Packalén, P., Yu, X., Eerikäinen, K., Hyypä, J., Pitkänen, J., 2005. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *For. Ecol. Manag.* 216, 41–50.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2014. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien.
- Monnet, J.-M., 2010. Développement et utilisation des outils d'analyse LiDAR pour la cartographie des gisements forestiers et l'évaluation des volumes sur pied en zone de montagne. CEMAGREF.

- 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, 2307–2326.
- Munoz, A., Bock, J., Monnet, J.-M., Renaud, J.-P., Jolly, A., Riond, C., 2015. Evaluation par validation indépendante des prédictions des paramètres forestiers réalisés à partir de données LiDAR aéroporté. *Rev. Fr. Photogrammétrie Télédétection* in press.
- Munoz, A., Viard-Cretat, G., Fay, J., 2013. Implantation d'un projet de desserte forestière en forêt de montagne : les apports du Lidar aérien. *Rendez-Vous Tech.* 3–6.
- Muss, J.D., Mladenoff, D.J., Townsend, P.A., 2011. A pseudo-waveform technique to assess forest structure using discrete lidar data. *Remote Sens. Environ.* 115, 824–835.
- Næsset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand. J. For. Res.* 19, 164–179.
- Nijland, W., Coops, N.C., Macdonald, S.E., Nielsen, S.E., Bater, C.W., 2015. Comparing patterns in forest stand structure following variable harvests using airborne laser scanning data. *For. Ecol. Manag.* in press.
- Ozdemir, I., Karnieli, A., 2011. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *Int. J. Appl. Earth Obs. Geoinformation* 13, 701–710.
- Packalén, P., Maltamo, M., 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Can. J. For. Res.* 38, 1750–1760.
- Papenmeier, F., 2014. segmag: Determine Event Boundaries in Event Segmentation Experiments.
- Pascual, C., García-Abril, A., García-Montero, L.G., Martín-Fernández, S., Cohen, W.B., 2008. Object-based semi-automatic approach for forest structure characterization using lidar data in heterogeneous *Pinus sylvestris* stands. *For. Ecol. Manag.* 255, 3677–3685.
- Peuhkurinen, J., Maltamo, M., Malinen, J., 2008. Estimating species-specific diameter distributions and saw log recoveries of boreal forests from airborne laser scanning data and aerial photographs: a distribution-based approach. *Silva Fenn.* 42, 625–641.
- Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J., Packalén, P., 2007. Preharvest Measurement of Marked Stands Using Airborne Laser Scanning. *For. Sci.* 53, 653–661.
- Pommerening, A., 2002. Approaches to quantifying forest structures. *Forestry* 75, 305–324.
- Rapidlasso GmbH, 2013. LAsTools Rapid LiDAR Processing. Gilching, Germany.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Schaeffer, L., 1953. Estimation des peuplements forestiers après réalisation.
- St-Onge, B., Hernandez, D., Kiepura, J., 2015. La cartographie forestière par arbres individuels à l'aide de SEGMA.
- Sullivan, A., 2008. LIDAR based delineation in forest stands (Master of Science). University of Washington, Washington.
- Thomas, V., Oliver, R.D., Lim, K., Woods, M., 2008. LiDAR and Weibull modeling of diameter and basal area. *For. Chron.* 84, 866–875.
- Tiede, D., Blaschke, T., Heurich, M., 2004. Object-based semi automatic mapping of forest stands with Laser scanner and Multi-spectral data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* 36, 328–333.
- Torresan, C., Strunk, J., Zald, H.S.J., Zhiqiang, Y., Cohen, W.B., 2014. Comparing statistical techniques to classify the structure of mountain forest stands using CHM-derived metrics in Trento province (Italy). *Eur. J. Remote Sens.* 75–94.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D., Etheridge, D., 2012. LiDAR Sampling Density for Forest Resource Inventories in Ontario, Canada. *Remote Sens.* 4, 830–848.
- Valbuena, R., Maltamo, M., Martín-Fernández, S., Packalén, P., Pascual, C., Nabuurs, G.-J., 2013a. Patterns of covariance between airborne laser scanning metrics and Lorenz curve descriptors of tree size inequality. *Can. J. Remote Sens.* 39, S18–S31.
- Valbuena, R., Packalén, P., Martín-Fernández, S., Maltamo, M., 2012. Diversity and equitability ordering profiles applied to study forest structure. *For. Ecol. Manag.* 276, 185–195.
- Valbuena, R., Packalén, P., Mehtätalo, L., García-Abril, A., Maltamo, M., 2013b. Characterizing forest structural types and shelterwood dynamics from Lorenz-based indicators predicted by airborne laser scanning. *Can. J. For. Res.* 43, 1063–1074.
- Valbuena, R., Vauhkonen, J., Packalén, P., Pitkänen, J., Maltamo, M., 2014. Comparison of airborne laser scanning methods for estimating forest structure indicators based on Lorenz curves. *ISPRS J. Photogramm. Remote Sens.* 95, 23–33.

- Vauhkonen, J., Mehtätalo, L., 2014. Matching remotely sensed and field-measured tree size distributions. *Can. J. For. Res.* 45, 353–363.
- Vega, C., Hamrouni, A., Mokhtari, S. El, Morel, J., Bock, J., Renaud, J.-P., Bouvier, M., Durrieu, S., 2014. PTrees: A point-based approach to forest tree extraction from lidar data. *Int. J. Appl. Earth Obs. Geoinformation* 33, 98–108.
- Venables, W.N., Ripley, B.D., 2002. *Modern Applied Statistics with S*, Fourth. ed. Springer, New York.
- White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D.G., Woods, M., 2013. Guide des meilleures pratiques pour générer des attributs d'inventaire forestier provenant de données obtenues par balayage laser aéroporté en utilisant une approche par zones.
- Woods, M., Lim, K., 2008. Predicting forest stand variables from LiDAR data in the Great Lakes – St. Lawrence forest of Ontario. *For. Chron.* 84, 827–839.
- Zeileis, A., 2014. *ineq: Measuring Inequality, Concentration, and Poverty*.
- Zhang, L., Gove, J.H., Liu, C., Leak, W.B., 2001. A finite mixture of two Weibull distributions for modeling the diameter distributions of rotated-sigmoid, uneven-aged stands. *Can. J. For. Res.* 31, 1654–1659.
- Zhang, Z., Liu, X., Peterson, J., Wright, W., 2011. Statistical analysis of airborne LiDAR data for forest classification in the Strzelecki Ranges, Victoria, Australia, in: Chan, F., Marinova, D., Anderssen, R.S. (Eds.), *Proceedings of the 19th International Congress on Modelling and Simulation (MODSIM 2011)*. Modelling and Simulation Society of Australia and New Zealand Inc., Canberra, Australia, pp. 1937–1943.

## Liste des figures et des tableaux

Figure 1 : Fonctionnement du LiDAR aéroporté (d'après Antonio Ferraz) .....	6
Figure 2 : Exemple de détection de desserte forestière à partir du MNT (source : ONF R&D) .....	7
Figure 3 : Courbe de Lorenz et détermination de l'indice de Gini (d'après Cordonnier et al., 2012) .....	9
Figure 4 : Triangle des structures de la typologie du Plateau Lorrain (source : CRPF Lorraine-Alsace) .....	10
Figure 5 : Densité de probabilité de la distribution de Weibull à deux paramètres (d'après Bailey et Dell, 1973) .....	13
Figure 6 : Démarche suivie pour la prédiction de la structure (renvoi aux paragraphes correspondants) .....	15
Figure 7 : Carte de localisation des sites d'étude .....	16
Figure 8 : Répartition des diamètres et détection des modes sur 4 placettes du Vercors .....	18
Figure 9 : Courbes de Lorenz et critères de classement des peuplements pour les 4 placettes du Vercors de la figure 8 .....	19
Figure 10 : Normalisation du nuage de points en hauteur (source : ONF R&D) .....	20
Figure 11 : Schéma de la méthode de création du MNH (source : ONF R&D) .....	21
Figure 12 : Schéma du principe de calcul de métriques "apex" .....	23
Figure 13 : Graphique de la corrélation entre une métrique LiDAR et le coefficient de Gini sur les placettes du Vercors .....	23
Figure 14 : Schéma du principe de classification par la méthode des k plus proches voisins .....	25
Figure 15 : Schéma du principe de l'hyperplan et des vecteurs supports dans la méthode SVM .....	26
Figure 16 : Exemple du meilleur hyperplan séparateur pour les placettes du Vercors dans un espace formé par 3 métriques LiDAR .....	27
Figure 17 : Evolution de l'erreur de prédiction en fonction du nombre de voisins dans la méthode kMSN sur les placettes du PNF .....	28
Figure 18 : Principe de construction de la carte prédictive .....	28
Figure 19 : Principe de validation de la prédiction de la structure au niveau des grappes de placettes .....	31

Figure 20 : Diamètre dominant et seuils en fonction des types de peuplements réguliers pour les placettes du Vercors.....	32
Figure 21 : Schéma de construction d'une carte détaillée des structures à partir de données et de modèles LiDAR.....	33
Figure 22 : Valeur des métriques Apex.CV_Vol et LAS.d9 sur les placettes "typiques" du Vercors en fonction de leur structure .....	35
Figure 23 : Valeur des métriques MNH.rumple et LAS.dHmean_fst sur les placettes "typiques" du Vercors en fonction de leur structure .....	36
Figure 24 : Nuage de points LiDAR et représentation 3D du MNH pour 3 placettes de type structural distinct du Vercors.....	36
Figure 25 : Comparaison des cartes prédictives LDA et Random Forest sur une zone de la forêt de Prénovel.....	39
Figure 26 : Cartes prédictives issues des classifications LDA et Random Forest.....	40
Figure 27 : Graphique du lien entre erreurs et probabilités de prédiction et carte de la probabilité d'appartenance à un type de structure.....	41
Figure 28 : Carte de la prédiction kMSN du coefficient de Gini .....	41
Figure 29 : Carte détaillée des types structuraux.....	42
Figure 30 : Exemple de deux placettes de structure difficilement caractérisable .....	43
Figure 31 : Variation des valeurs des centiles 5% et 20% des premiers retours sur les placettes "typiques" du PNF en fonction de leur structure.....	44
Figure 32 : Comparaison entre la photo aérienne et la probabilité prédite d'appartenance à un type structural sur une zone du site du PNF .....	46
Figure 33 : Comparaison de deux vols LiDAR de même densité d'émission (4 pts/m <sup>2</sup> ) réalisés en hiver et en été (d'après Clément Mallet) .....	48
Figure 34 : Exemples de prédiction kMSN de la répartition des tiges sur deux placettes "cibles" du Vercors .....	50
Tableau 1 : Informations des vols LiDAR .....	16
Tableau 2 : Répartition des grappes de placettes de validation indépendante (site du Vercors) .....	30
Tableau 3 : Répartition des placettes et caractéristiques dendrométriques par site d'étude .....	34
Tableau 4 : Jeux de métriques sélectionnés sur le site du Vercors.....	35
Tableau 5 : Résultats de la validation croisée des méthodes de classification supervisée sur le Vercors .....	37
Tableau 6 : Résultats de la validation croisée des méthodes de prédiction quantitative sur le Vercors .....	37
Tableau 7 : Résultats de la validation indépendante et des tests de robustesse des modèles calibrés sur le jeu de placettes "Gini" du Vercors.....	38
Tableau 8 : Répartition des placettes et caractéristiques dendrométriques du site du PNF .....	42
Tableau 9 : Jeux de métriques sélectionnés sur le site du PNF.....	43
Tableau 10 : Résultats de la validation croisée des méthodes de classification supervisée sur le site du PNF .....	45
Tableau 11 : Résultats de la validation croisée des méthodes de prédiction quantitative sur le site du PNF .....	45

## Annexes

Annexe 1 : Tableau de synthèse bibliographique .....	1
Annexe 2 : Clé de détermination des types de Structure d'après la typologie du Massif Vosgien.....	4
Annexe 3 : Segmentation des couronnes et apex (méthodes et paramètres) .....	5
Annexe 4 : Protocole d'inventaire des placettes de validation du Vercors .....	8
Annexe 5 : Résultats des validations des modèles de prédiction .....	9
Annexe 6 : Exemples de cartes prédictives des types structuraux sur le site du Vercors.....	10
Annexe 7 : Carte des types de peuplements sur le site du Vercors issue de données terrain et de photo-interprétation .....	11

**ANNEXE 1 : TABLEAU DE SYNTHÈSE BIBLIOGRAPHIQUE**

(se référer à la fin du tableau pour les abréviations et définitions)

	Article	Site d'étude	Inventaire de terrain	Variables explicatives (ou entrées)	Variables prédites (ou sorties)	Méthode	Erreur
Segmentation a priori	(Lesur, 2011)	<b>Mayotte</b>	Pas de terrain	<b>Surface</b> : Rumble, indice de complexité de la structure verticale, Hmoy, Hsd, Tx Pen.	8 types de peuplements	Classification par la méthode des <b>k-means</b>	Pas de validation
	(Pascual et al., 2008)	<b>Espagne</b> (pins)	10 pl. entre 1260 et 2400 m <sup>2</sup>	<b>MNH</b> : MNH brut pour la segmentation, Hsd et Hmean pour la classification	5 types de peuplement	<b>Segmentation</b> du MNH (logiciel eCognition) puis classification des segments par la méthode des <b>k-means</b>	Vérification par comparaison au terrain
	(Maltamo et al., 2005)	<b>Finlande</b> (épicéas)	28 pl. de 30x30m Diam. préc. : 5cm	<b>Surface</b> : distribution des points LiDAR en hauteur	Etagement du peuplement (mono/ multistrate)	<b>Détection des modes</b> de la répartition (algorithme HistMod)	24/28 peuplements bien classés
	(Leppänen et al., 2008)	<b>Finlande</b> (pins, épicéas et feuillus)	683 pl. de 50m <sup>2</sup>	<b>MNH</b>	Zones de structure homogène	<b>Segmentation</b> de régions à partir du MNH (logiciel eCognition)	Comparaison des méthodes via le coefficient de corrélation entre N, G, V terrain et la segmentation
				<b>Surface</b> : H85 et VEG <b>Données spectrales</b> : photo IRC	Zones de structure homogène	<b>Croissance de régions homogènes</b> à partir de la photo IRC et des couches LiDAR H85 et VEG	
	(Tiede et al., 2004)	<b>Allemagne, Bavière</b> (épicéas et feuillus)	Carte préexistante des stades de développement	<b>Surface</b> : classe de hauteur... <b>Données spectrales</b> : essences	4 stades de développement	<b>Segmentation</b> automatique fine puis <b>classification</b> selon des critères de structure puis agrégation des segments identiques	Overall Accuracy = 62%
	(Gachet and Junod, 2008)	<b>Suisse</b> (région de montagne)	Carte préexistante des types de peuplements	<b>MNH</b>	Zones de structure homogène	<b>Segmentation</b> du MNH (logiciel Definiens Professional) puis classification des segments	Comparaison visuelle
	(Maier et al., 2008)	<b>Autriche</b> (région de montagne, épicéas)	33 pl. de 314m <sup>2</sup>	<b>Etape 1</b> : MNH et MNT <b>Etape 2</b> : densité de canopée (% de surface d'arbres d'une classe de hauteur dans le patch), distribution des trouées, rugosité de la canopée	<b>Etape 1</b> : patchs de peuplements <b>Etape 2</b> : types de structure	<b>Etape 1</b> : Double <b>segmentation</b> : échelle arbre et échelle peuplement. <b>Etape 2</b> : Calcul d'indices par patch de peuplement permettant la <b>classification</b> en types	Overall Accuracy = 82% (pour les types mono/ multistrate)
(Dickinson et al., 2014)	<b>Pennsylvanie</b> , (mélange feuillus et résineux)	21 blocs de 0.8ha et 4 sites de démonstration Diam. préc. : 5cm	<b>Surface</b> : Hmoy, H95, Hsd, Hcv, Hkurt, Hskew, VEG, rumple <b>Données spectrales</b> : essences	14 types de peuplement	Séparation en 4 groupes selon la hauteur H95 puis classification <b>k-means</b> par groupe. (Régressions logistiques multinomiales pour assigner des zones n'ayant pas servi à la calibration dans des classes)	Overall Accuracy = 96%	
Classification supervisée	(Sullivan, 2008)	<b>Washington</b>	108 pl. de 809m <sup>2</sup> Diam. préc. : 0cm	<b>Arbre</b> : nombre et hauteurs d'apex <b>Surface</b> : VEG	6 types de peuplement	Classification orientée objet	Overall Accuracy = 84-88%
	(Chirici et al., 2013)	<b>Italie, Sicile</b> (biomasse inflammable)	Carte préexistante des types (basée sur photo-interprétation et 935 placettes)	<b>Surface</b> : Hmoy, Hmax, Hmin, Hcentiles, Hcv... (sur tout le nuage et par strate de hauteur) <b>Données spectrales</b> : 4 bandes spectrales	9 types de biomasse inflammable	Arbres de decision CART, Random Forest, stochastic gradient boosting	Overall Accuracy = resp. 47%, 45% et 84%

	Article	Site d'étude	Inventaire de terrain	Variables explicatives (ou entrées)	Variables prédites (ou sorties)	Méthode	Erreur
Classification supervisée	(Zhang et al., 2011)	<b>Australie</b> (forêt humide et forêt tropicale tempérée froide)	Carte préexistante des types (basée sur photo-interprétation et terrain)	<b>Surface</b> : Hmax, Hmoy, Hsd, Dens, Tx. Pen. (par strate de hauteur du nuage)	5 grands types de forêts	Linear Discriminant Analysis	Overall Accuracy = 91.4%
	(Torresan et al., 2014)	<b>Nord Italie</b> (région de montagne, pins et épicéas)	90 pl. de 531 à 2827m <sup>2</sup> Diam. préc. : 7.5cm	<b>MNH</b> : Hcv, Cov	9 types issus d'une typologie en G	Linear Discriminant Analysis et Random Forest	Overall Accuracy : LDA=41% RF=37%
Prédiction variables quantitatives	(Valbuena et al., 2013b)	<b>Espagne</b> (pins de structure variée)	37 pl. de 20m Diam. préc. : 0cm	<b>Surface</b> : Hmax, Hsd, H25, VEG, Lratios	Coefficient de Gini et asymétrie de Lorenz permettant la classification en 4 types	Régression bêta pour estimer les coefficients	Overall Accuracy : 73% rRMSE sur Gini : 0.084
	(Valbuena et al., 2014)	<b>Finlande</b> (pins, épicéas et quelques feuillus)	79 pl. de 20x20 à 30x30m Diam. préc. : 5cm	<b>Surface</b> : Hcentiles, Lmoments, Lratios, Hskew, VEG, Dens.	Indicateurs (Gini, asymétrie de Lorentz, %G>Dqm et %N>Dqm)	Estimation paramétrique, Random Forest, KNN basé sur Random Forest et kMSN	rRMSE sur Gini : entre 0.076 et 0.130
					Quantiles de la courbe de Lorenz puis indicateurs (Gini...)	kNN basé sur Random Forest et kMSN	rRMSE sur Gini : 0.080
				<b>Surface</b> : idem <b>Arbres</b> : hauteurs et couronnes	Liste d'arbres puis indicateurs (Gini...)	Détermination de la liste d'arbres à partir des arbres détectés ou à partir de métriques LiDAR de surface	rRMSE sur Gini : entre 0.072 et 0.213
Prédiction de la répartition des diamètres	(Thomas et al., 2008)	<b>Ontario</b> (peuplements mixtes, feuillus et résineux)	115 pl. de 400 à 2500m <sup>2</sup> Diam. préc. : 9 cm	<b>Surface</b> : Hmoy, Hsd, Hcv, H10, H20, H90, Dens.	Coefficients de la loi de Weibull (et mélange fini de lois de Weibull pour les bimodaux) puis répartition des diamètres	Régressions linéaires multiples pour la prédiction des coefficients	Coefficients de détermination des modèles entre 0.65 et 0.88
	(Packalén and Maltamo, 2008)	<b>Finlande</b> (pins, épicéas et bouleaux)	463 pl. de 9m Diam. préc. : 5cm	<b>Surface</b> : Hcentiles, Dens. <b>Données spectrales</b> : valeurs spectrales et variables texturales	Répartition des diamètres	Liste d'arbres par kMSN	rRMSE sur V : 10.4%
						Prédiction de G, Dg, Hg pour ajuster une loi de Weibull	rRMSE sur V : 12.7%
	(Maltamo et al., 2004)	<b>Finlande</b> (épicéas)	10 (+23) pl. de 30x30m Diam. préc. : 5cm	<u>Gros arbres</u> : <b>Arbre</b> : hauteur et surface de couronne <u>Petits arbres</u> : <b>Surface</b> : mode, Hmax	Répartition des diamètres <u>Gros arbres</u> : Diamètre directement <u>Petits arbres</u> : paramètres de la loi de Weibull	<u>Gros arbres</u> : Régression logarithmique <u>Petits arbres</u> : régression linéaire pour les paramètres de la loi de Weibull	rRMSE sur V : 16% rRMSE sur N : 49.2%
	(Bollandsås et al., 2013)	<b>Norvège</b> (épicéas, pins et bouleaux)	201pl. de 200 à 400m <sup>2</sup> Diam. préc. : 4cm	<b>Surface</b> : Hcentiles, Dens, VEG, Hmoy, Hsd, Hcv.	Répartition des diamètres	Prédiction par kMSN	Reynolds : 59%
Prédiction des centiles de la répartition (11 centiles) par des régressions						Reynolds : 55%	

Prédiction de la répartition des diamètres	Article	Site d'étude	Inventaire de terrain	Variables explicatives (ou entrées)	Variables prédites (ou sorties)	Méthode	Erreur
	(Maltamo et al., 2009)	<b>Norvège</b> (épicéas, pins et bouleaux)	50 pl. x4 de 200 à 400m <sup>2</sup> Diam. préc. : 0cm	<b>Surface</b> : Hcentiles, Dens, VEG, Hmoy, Hsd, Hcv.	Répartition des diamètres	Liste d'arbres par KMSN	Reynolds : 83.2% rRMSE sur V : 15.4% rRMSE sur N : 35.1%
	(Bollandsås and Næsset, 2007)	<b>Norvège</b> (épicéas et pins)	20 (+18) pl. de 1000m <sup>2</sup> Diam. préc. : 4cm	<b>Surface</b> : Hcentiles, Dens, Hmoy, Hsd, Hcv.	Répartition des diamètres	Prédiction des centiles de la répartition (10 centiles) par des régressions « partial least square »	Reynolds : 78.1%
	(Peuhkurinen et al., 2008)	<b>Finlande</b> (pins, épicéas et feuillus)	474 pl. de 9m, Diam. préc. : 5cm	<b>Surface</b> : métriques de hauteur <b>Données spectrales</b> : essences	Répartition des diamètres par essence	KMSN avec distance de Minkowski	rRMSE sur V : 21.8-22.8%
	(Peuhkurinen et al., 2007)	<b>Finlande</b> (épicéas)	457 pl. de 9m + données de récolte Diam. préc. : 5cm	<b>Arbre</b> : hauteur des apex et diamètre des couronnes	Répartition des diamètres	Relations allométriques entre hauteurs et diamètres (notamment)	Reynolds (!) formule différente pour l'erreur : de 32 à 65
	(Vauhkonen and Mehtätalo, 2014)	<b>Finlande</b> (pins, forêt boréale gérée typique)	40 pl. de 400m <sup>2</sup> Diam. préc. : 0cm	<b>Arbre</b> : rayons des couronnes (variable principale) <b>Surface</b> : Hcentiles, Hmoy, Hsd (variables secondaire pour prédire l'effet aléatoire)	Répartition des diamètres	Distribution Matching : fonction de transformation entre la fonction cumulative des couronnes (corrigée pour prendre en compte la sous-détection des petits arbres) et la fonction cumulative des diamètres	rRMSE sur N : entre 27% et 84% selon les équations et configurations

#### Définitions et abréviations

**pl.** : placettes  
**Diam. préc.** : diamètre précomptable  
**G** : surface terrière  
**N** : nombre de tiges  
**V** : volume  
**Dg** : diamètre quadratique moyen

**Surface** : métriques issues du nuage de points sur la surface d'une placette

**Hmoy** : hauteur moyenne des points  
**Hmax, Hmin** : hauteur maximale, minimale  
**Hsd** : écart-type des hauteurs  
**Hcv** : coefficient de variation des hauteurs  
**Tx Pen** : taux de pénétration du signal LiDAR

**Hcentiles ou H05, H10,..., H99** : centiles de la répartition en hauteur  
**Lmoments, Lratios** : moments de la répartition des hauteurs et ratios des moments  
**Hkurt** : kurtosis (coefficient d'aplatissement de la répartition des hauteurs)  
**Hskew** : skewness (asymétrie de la répartition des hauteurs)  
**VEG** : pourcentage de points végétation (ou points supérieurs à 1m)  
**Dens** : ratio du nombre de retours dans une strate sur le nombre total de retours  
**Mode** : nombre de modes de la répartition des hauteurs

**MNH** : modèle numérique de hauteur (utilisé en donnée d'entrée ou pour calculer des métriques)

**Cov** : nombre de pixels d'une strate de hauteur sur le nombre total de pixels  
**Rumple** : ratio de la surface de la canopée par rapport à la surface au sol

**Arbre** : variables issues d'une segmentation des apex (approche « arbre »)

#### Erreurs

**Overall Accuracy** : précision globale de la classification (% bien classé)  
**rRMSE** : erreur quadratique moyenne en relatif calculée sur une variable dérivée de la répartition des diamètres

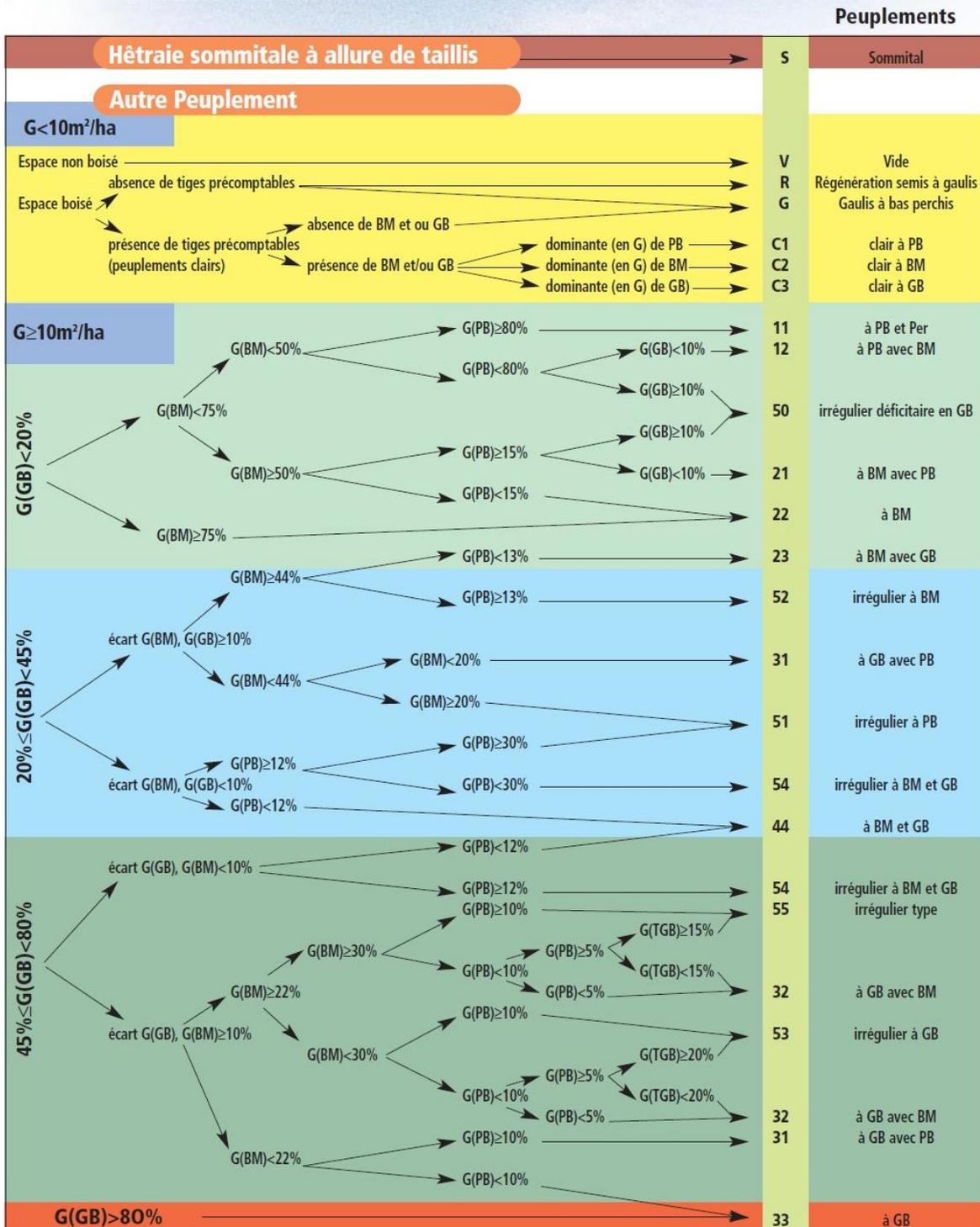
**Reynolds** : erreur de Reynolds,  $e = 100 \frac{\sum_{j=1}^k |n_{jpred} - n_{jobs}|}{N}$  (avec  $N$  nombre de tiges,  $k$  nombre de classes de diamètre,  $n_{jpred}$  nombre de tiges prédit de la classe de diamètre  $j$ ,  $n_{jobs}$  nombre de tiges observé de la classe de diamètre  $j$ )

#### Sources

cf. bibliographie générale p.54

**Catégories de bois :**  
 - Perches : 10-15 cm  
 - PB (Petits Bois) : 20-25 cm  
 - BM (Bois Moyens) : 30 à 45 cm  
 - GB (Gros Bois) : 50 cm et plus comprenant la sous-catégorie des Très Gros Bois (70 cm et plus)

Les surfaces terrières sont mesurées uniquement sur les arbres d'un diamètre supérieur à 17,5 cm.  
 G(...) correspond au pourcentage de la surface terrière dans la catégorie de diamètre désignée.



**Richesse en perches**

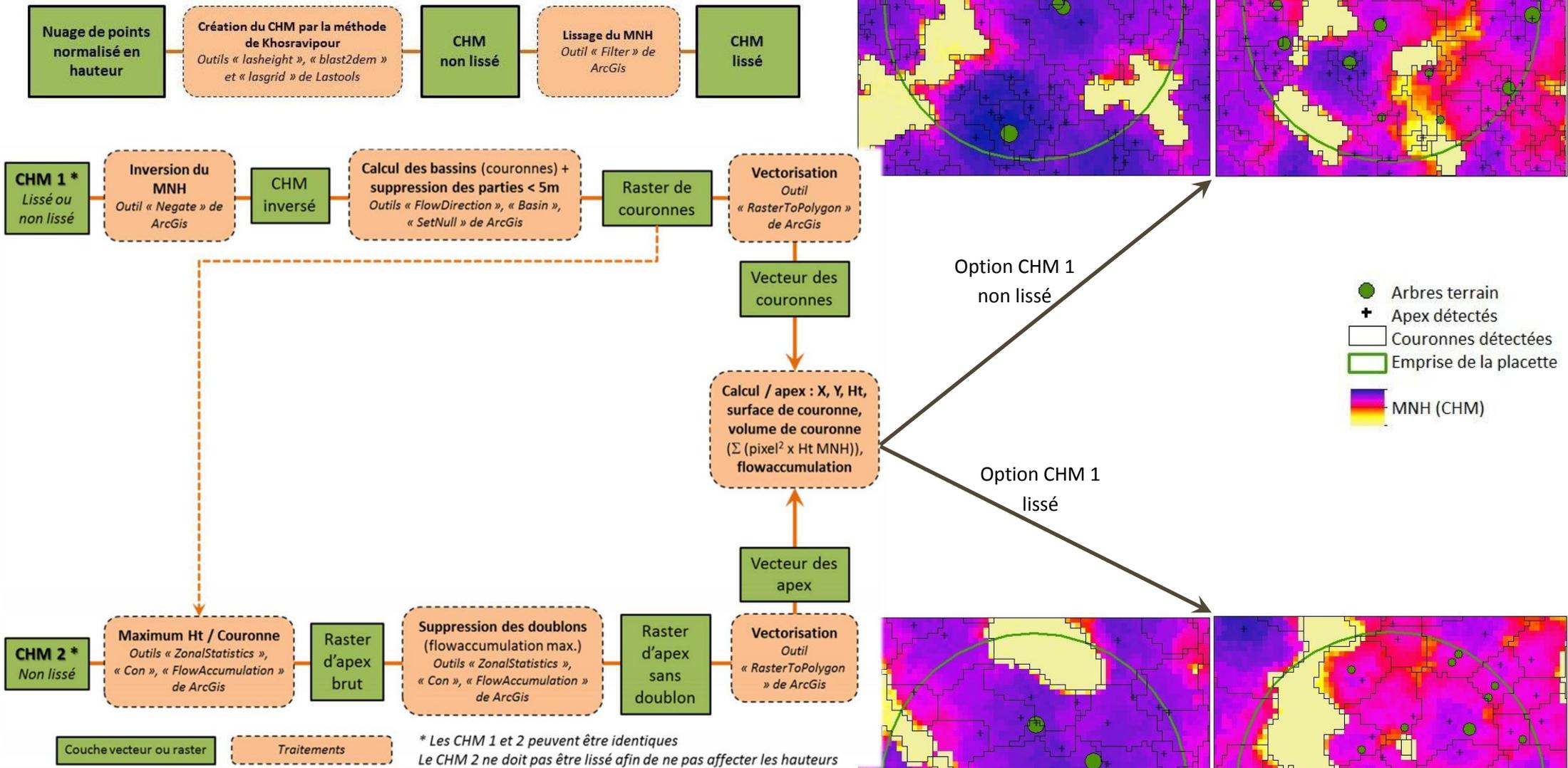
La structure sera dite riche en perches si dans un cercle de 15 m de rayon (jugé à l'œil), on trouve plus de :

- 6 perches pour les types 21-22-23-44-32-33
- 10 perches pour les types 50-52-53-54-55

Par exemple : Type 33 "pauvre en perches" 330  
 "riche en perches" 331

Remarque : Les types 11, 12, 31 et 51 sont toujours riches en perches.

### A. Méthode des « bassins versants »



Source du schéma : ONF R&D

## B. Méthode *SEGMA*

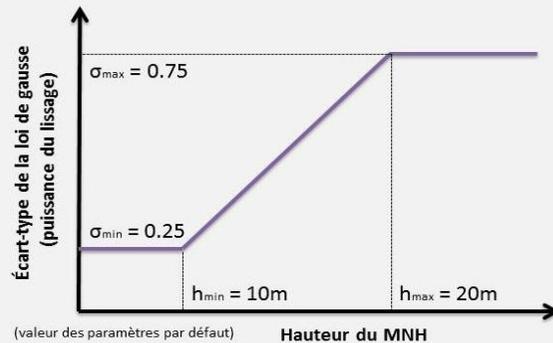
Principes de fonctionnement de *Segma* (valeur des paramètres par défaut)

### Intrants

- MNH (un ou plusieurs fichiers, éventuellement jointifs sans recouvrement) :
- format .flt obtenu avec Lastools
  - fichier d'entête .hdr transformé en version lisible par *Segma* via un script python

### Filtrage gaussien adaptatif du MNH

Lissage du MNH sur une fenêtre fixée (ici de rayon 4 pixels) plus fort pour le haut de la canopée que pour le sous-étage :



### Maxima locaux

Détection des maxima locaux dans la gamme des hauteurs possibles

### Identification des types

- Classification des maxima en deux grands types d'essences :
- conifère si la pente autour du maximum est supérieure à 75% sur le rayon d'analyse
  - feuillu sinon

### Croissance des couronnes, variable selon les essences

- Agrégation de pixels à partir des maxima locaux jusqu'à l'accomplissement d'une condition d'arrêt :
- forte vallée
  - vallée peu prononcée mais loin du centre de la couronne
  - hauteur < à 20% du maximum local
  - rayon de couronne maximal
  - hauteur d'un pixel > 110% de la hauteur moyenne de la périphérie de l'arbre
  - circularité affectée par l'ajout d'un pixel dans une direction donnée

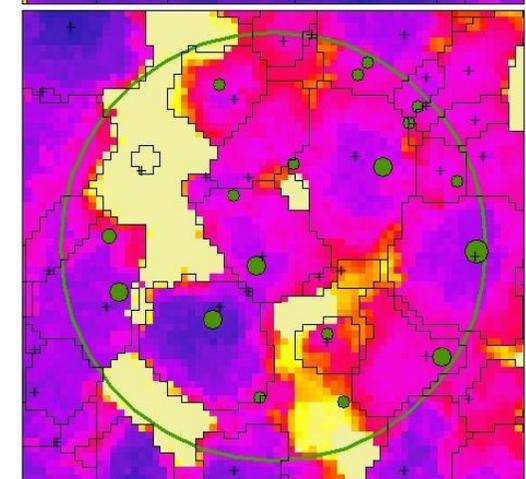
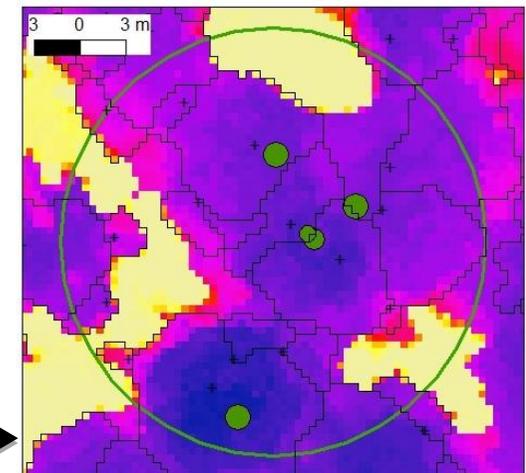
### Extrants

- Raster des couronnes ultérieurement traité par un script python-arcgis pour obtenir :
- vecteur des couronnes avec leur surface
  - la localisation et la hauteur des apex
  - volume des couronnes :  $\sum_{i=1}^n S_i \cdot h_i$  ( $n$  nombre de pixels,  $S_i$  surface d'un pixel,  $h_i$  hauteur du pixel)



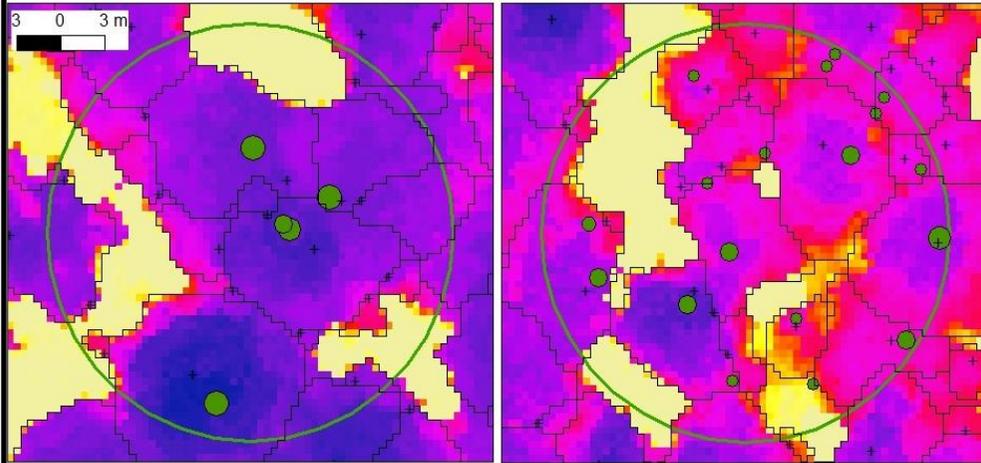
### Interface de *Segma* avec les paramètres par défaut

Intrants		Modèle de hauteur de couvert		D:\Claire\Tools\Segma\Test_PNF\MNH_fit\	
Filtrage gaussien adaptatif		Min	Max		
Seuil de hauteur du MHC		10.0	20.0	m	m
Écart-type du filtre pour hauteurs		0.25	0.75		
Rayon du filtre en pixels		4			
MHC filtré		D:\Claire\Tools\Segma\Test_PNF\Segma_d			
<input type="checkbox"/> Filtrage seulement					
Maxima locaux		Min	Max		
Seuil d'exclusion des valeurs		1.5	35.0	m	m
<input type="checkbox"/> Maxima locaux seulement					
Identification de types		Rayon d'analyse (% du maximum local) 20.0			
		Seuil de pente pour types 0.75			
Croissance des couronnes		Feuillu		Conifère	
Rayon maximal de croissance (% du maximum local)		40.0		30.0	
Distance de modulation pour vallées		1.0		2.3	
Seuil relatif pour vallées		0.5		0.5	
<input checked="" type="checkbox"/> Seuil absolu pour vallées		1.0			
Hauteur minimale (% du maximum local)		25.0			
Dilatation (pixels)		1			
Aire minimale (% du maximum local)		15.0			
<input checked="" type="checkbox"/> Fusion de petites couronnes					
Itérations		2			
Extrants		Position et hauteur des arbres D:\Claire\Tools\Segma\Test_PNF\Segma			
		Carte des couronnes D:\Claire\Tools\Segma\Test_PNF\Segma			
ID de départ		1			



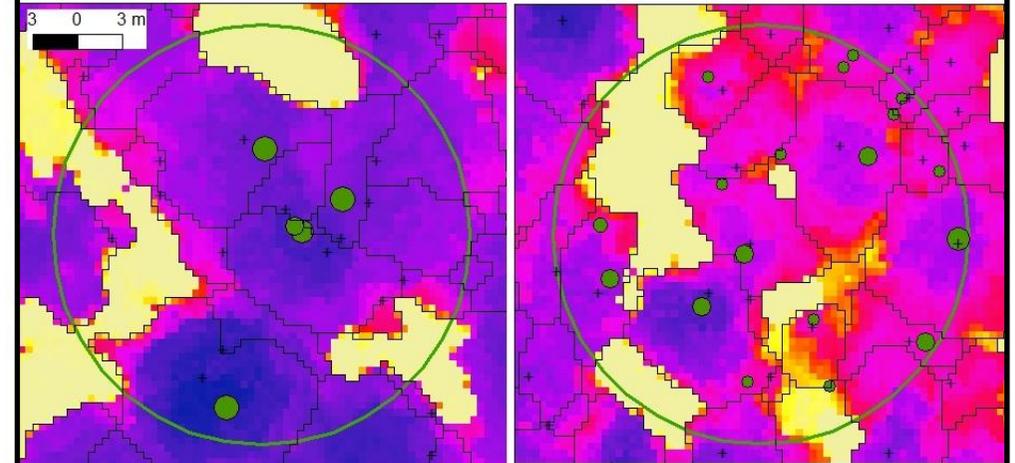
### Paramètres avec un filtrage fort

Filtrage gaussien adaptatif		Croissance des couronnes		Feuilleu	Conifère
Seuil de hauteur du MHC	Min 10.0 m	Max 20.0 m	Rayon maximal de croissance (% du maximum local)	40.0	40
Écart-type du filtre pour hauteurs	0.25	0.9	Distance de modulation pour vallées	1.0	1
Rayon du filtre en pixels	6		Seuil relatif pour vallées	0.5	0.5
Maxima locaux			<input checked="" type="checkbox"/> Seuil absolu pour vallées	1.0	
Seuil d'exclusion des valeurs	Min 5 m	Max 40 m	Hauteur minimale (% du maximum local)	25.0	
Identification de types			Dilatation (pixels)	1	
Rayon d'analyse (% du maximum local)	20.0		Aire minimale (% du maximum local)	15.0	
Seuil de pente pour types	0.75		<input checked="" type="checkbox"/> Fusion de petites couronnes		
			Itérations	2	



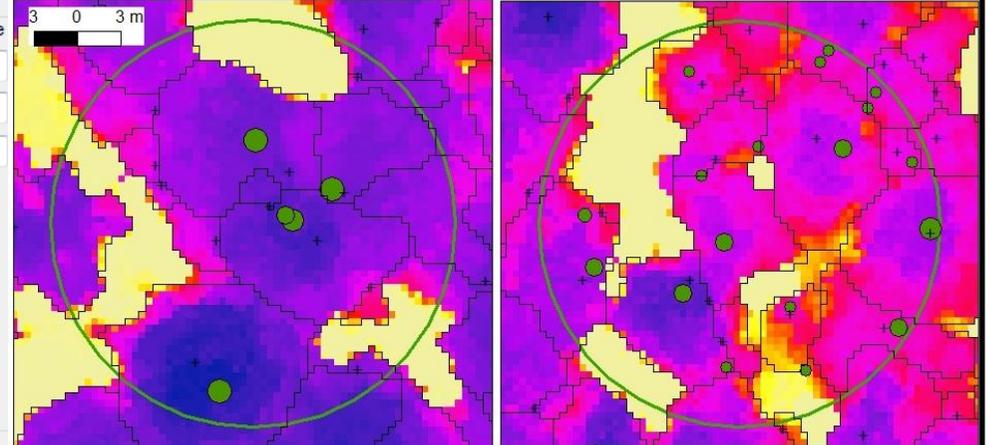
### Paramètres avec une croissance des couronnes plus large

Filtrage gaussien adaptatif		Croissance des couronnes		Feuilleu	Conifère
Seuil de hauteur du MHC	Min 10.0 m	Max 20.0 m	Rayon maximal de croissance (% du maximum local)	60	60
Écart-type du filtre pour hauteurs	0.25	0.75	Distance de modulation pour vallées	1.0	1
Rayon du filtre en pixels	4		Seuil relatif pour vallées	0.5	0.5
Maxima locaux			<input checked="" type="checkbox"/> Seuil absolu pour vallées	1.0	
Seuil d'exclusion des valeurs	Min 5 m	Max 40 m	Hauteur minimale (% du maximum local)	30	
Identification de types			Dilatation (pixels)	1	
Rayon d'analyse (% du maximum local)	20.0		Aire minimale (% du maximum local)	30	
Seuil de pente pour types	0.75		<input checked="" type="checkbox"/> Fusion de petites couronnes		
			Itérations	2	



### Paramètres avec filtrage et croissance plus grands (paramètres finaux utilisés dans l'étude pour le PNF)

Filtrage gaussien adaptatif		Croissance des couronnes		Feuilleu	Conifère
Seuil de hauteur du MHC	Min 10.0 m	Max 20.0 m	Rayon maximal de croissance (% du maximum local)	60	60
Écart-type du filtre pour hauteurs	0.25	0.9	Distance de modulation pour vallées	1.0	1
Rayon du filtre en pixels	6		Seuil relatif pour vallées	0.5	0.5
Maxima locaux			<input checked="" type="checkbox"/> Seuil absolu pour vallées	1.0	
Seuil d'exclusion des valeurs	Min 5 m	Max 40 m	Hauteur minimale (% du maximum local)	30	
Identification de types			Dilatation (pixels)	1	
Rayon d'analyse (% du maximum local)	20.0		Aire minimale (% du maximum local)	30	
Seuil de pente pour types	0.75		<input checked="" type="checkbox"/> Fusion de petites couronnes		
			Itérations	2	



**OBJECTIF :** placettes de validation par grappes. Ces placettes permettront de vérifier la prédiction à partir de données LiDAR de la structure des peuplements. L'échantillonnage est donc stratifié sur la zone pour inventorier des peuplements supposés irréguliers, réguliers ou limite, ainsi que des forêts publiques et privées.

**MATERIEL :**

- Récepteur GNSS Trimble
- TDS
- Trépied pour poser le récepteur GNSS
- Vertex et transpondeur
- Décamètre pour étalonner le vertex
- Piles de rechange pour le vertex
- Compas forestier
- Craies
- *Secondaire : boussole, stylos et papier, peinture*

**CHEMINEMENT VERS LA PLACETTE :**

Naviguer à l'aide du récepteur GNSS jusqu'au centre de la placette, puis positionner le récepteur sur le trépied et enregistrer le point GPS en prenant environ 300 enregistrements.

**INVENTAIRE SUR LA PLACETTE :**

Dans un nouveau relevé du TDS, enregistrer le nom de la placette puis les caractéristiques des arbres:

- **Diamètre** couvert (D130) issu de la moyenne de 2 diamètres perpendiculaires, au centimètre près.
- **Essence** (sapin, épicéa, hêtre, autres feuillus, autres résineux)
- **Etat de l'arbre** (Cod\_san) : **1** : Normal, **2** : Dépérissant, **3** : Mort avec branche, **4** : Mort sans branche
- Indiquer si la **cime** de l'arbre est cassée (Code\_cc) : **0** : Non, **1** : Oui
- Indiquer les fourches, les tiges « jumelles », les arbres penchés en observation.

Les arbres précomptables ont un diamètre supérieur ou égal à **17.5 cm** et se situent dans un rayon de **15 m** autour du centre de la placette (en distance horizontale, vérifier au vertex si besoin).

**CAS PARTICULIERS :**

Les vols LiDAR ayant eu lieu en 2010 et 2011, les informations sur les arbres ayant visiblement été vivants à cette date sont importantes. Par exemple, une coupe ayant eu lieu après cette date pourrait avoir modifié le type structural de la placette.

**Chablis récents (après 2010) :** inventorier les arbres visiblement récemment tombés (pas de mousse, pas pourris...), dont la souche tombe dans les 15 m de rayon. Indiquer « chablis » en observation.

**Placettes récemment exploitées (depuis 2010) :** indiquer « exploitation » en observation de la placette et inventorier les souches récentes dans les 15 m en notant « souche » en observation. Prendre deux diamètres perpendiculaires de la souche, ils seront convertis en D130 ultérieurement.

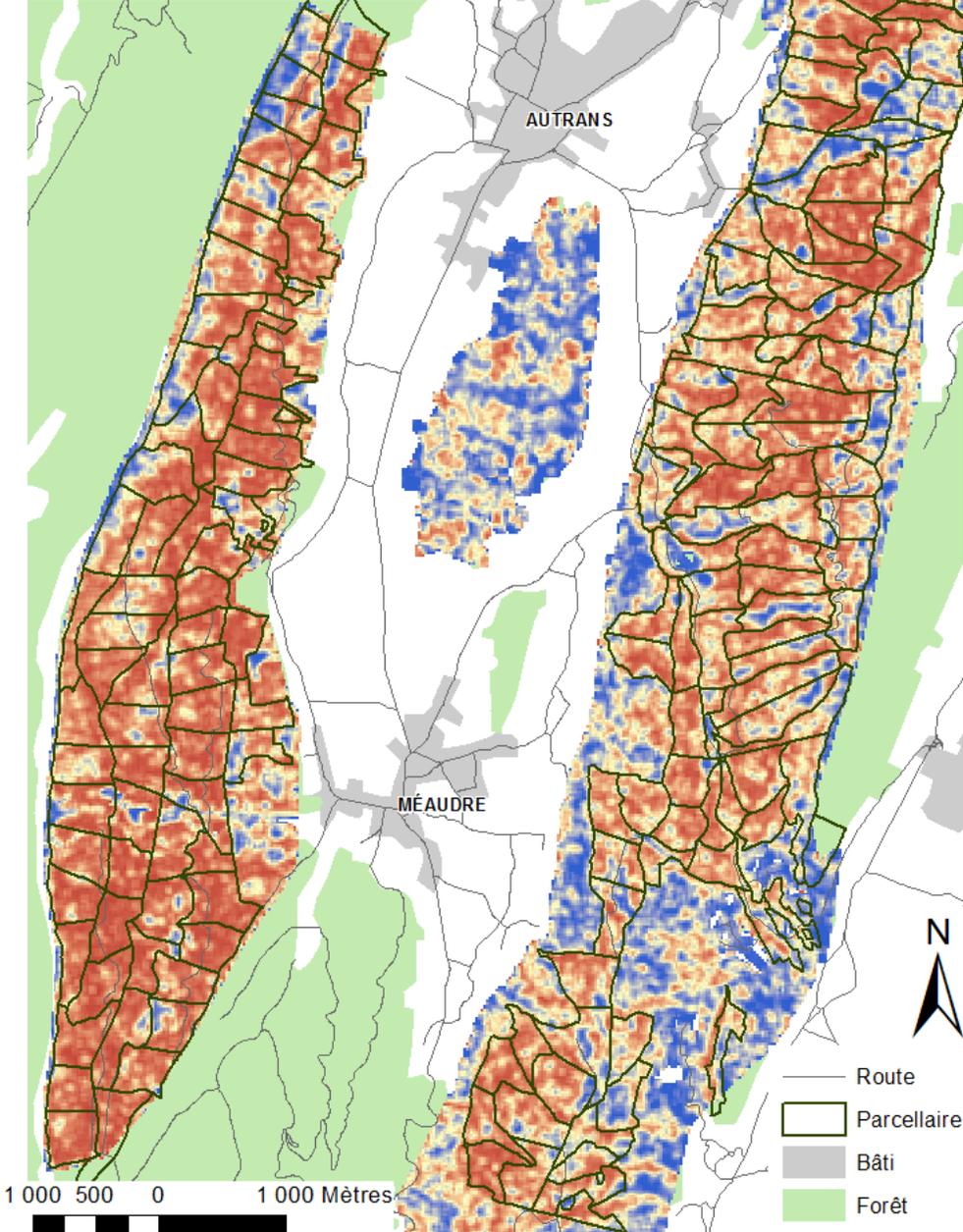
**ANNEXE 5 : RESULTATS DES VALIDATIONS DES MODELES DE PREDICTION**

		Calibration et validation sur les placettes "typiques"										Calibration et validation sur les placettes "Gini"									
		Validation croisée Vercors		Validation indépendante Vercors		Test de robustesse Aillons		Test de robustesse Prénovel		Validation croisée PNF		Validation croisée Vercors		Validation indépendante Vercors		Test de robustesse Aillons		Test de robustesse Prénovel		Validation croisée PNF	
Nb placettes de calibration		113		169		169		169		44		211		317		317		317		68	
Nb placettes de validation		56		33		18		47		22		106		51		54		113		33	
Jeu de métriques		"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"	"Lidar"	"Apex-MNH"
Nombre de métriques		5	4	5	4	5	4	5	4	3	3	5	4	5	4	5	4	5	4	3	3
kNN	<b>Précision</b>	<b>81,5%</b>	<b>83,5%</b>	<b>75,8%</b>	<b>87,9%</b>	<b>83,3%</b>	<b>72,2%</b>	<b>72,3%</b>	<b>76,6%</b>	<b>81,6%</b>	<b>80,1%</b>	<b>77,7%</b>	<b>79,2%</b>	<b>72,5%</b>	<b>76,5%</b>	<b>87,0%</b>	<b>77,8%</b>	<b>82,3%</b>	<b>77,9%</b>	<b>78,7%</b>	<b>77,3%</b>
	Kappa	57,7%	63,0%	49,4%	75,6%	55,7%	44,4%	32,8%	36,7%	18,0%	-3,6%	44,4%	48,9%	42,8%	49,8%	29,2%	30,3%	34,1%	22,5%	18,8%	1,9%
	<i>Irrégulier</i>	89,0%	88,6%	76,2%	94,1%	92,3%	100,0%	84,4%	83,3%	23,2%	0,0%	88,9%	88,8%	82,8%	83,9%	93,8%	97,4%	87,6%	86,2%	22,4%	5,9%
	<i>Régulier</i>	67,7%	74,3%	75,0%	81,3%	60,0%	44,4%	46,7%	54,5%	94,7%	96,3%	53,2%	58,0%	59,1%	65,0%	33,3%	26,7%	50,0%	36,8%	93,7%	95,7%
LDA	<b>Précision</b>	<b>88,6%</b>	<b>83,5%</b>	<b>87,9%</b>	<b>93,9%</b>	<b>77,8%</b>	<b>77,8%</b>	<b>85,1%</b>	<b>78,7%</b>	<b>84,3%</b>	<b>82,8%</b>	<b>81,4%</b>	<b>81,3%</b>	<b>84,3%</b>	<b>80,4%</b>	<b>85,2%</b>	<b>88,9%</b>	<b>86,7%</b>	<b>82,3%</b>	<b>83,9%</b>	<b>79,7%</b>
	Kappa	73,8%	62,5%	75,6%	87,8%	52,6%	45,5%	59,7%	37,2%	30,1%	6,8%	54,5%	53,7%	66,5%	60,1%	42,7%	43,9%	41,9%	24,4%	40,3%	1,9%
	<i>Irrégulier</i>	94,5%	89,9%	94,1%	100,0%	100,0%	91,7%	88,9%	82,1%	34,6%	7,5%	90,1%	90,8%	90,3%	92,6%	97,7%	95,7%	87,5%	85,4%	40,5%	1,8%
	<i>Régulier</i>	77,8%	72,1%	81,3%	87,5%	50,0%	50,0%	72,7%	62,5%	94,9%	98,3%	62,4%	60,4%	75,0%	66,7%	36,4%	42,9%	77,8%	50,0%	95,3%	99,5%
SVM (hyperplan linéaire)	<b>Précision</b>	<b>88,8%</b>	<b>85,0%</b>	<b>87,9%</b>	<b>93,9%</b>	<b>77,8%</b>	<b>72,2%</b>	<b>83,0%</b>	<b>78,7%</b>	<b>85,5%</b>	<b>82,9%</b>	<b>81,4%</b>	<b>81,5%</b>	<b>84,3%</b>	<b>82,4%</b>	<b>90,7%</b>	<b>87,0%</b>	<b>86,7%</b>	<b>81,4%</b>	<b>84,0%</b>	<b>79,8%</b>
	Kappa	74,0%	66,2%	75,6%	87,8%	52,6%	36,6%	46,4%	37,2%	40,4%	0,8%	54,1%	54,6%	66,5%	63,7%	56,6%	39,2%	35,4%	22,5%	40,5%	0,0%
	<i>Irrégulier</i>	95,0%	89,9%	94,1%	100,0%	100,0%	90,9%	82,9%	82,1%	46,8%	1,0%	90,5%	90,2%	90,3%	92,9%	97,8%	95,7%	86,1%	85,3%	39,9%	0,0%
	<i>Régulier</i>	77,0%	76,1%	81,3%	87,5%	50,0%	42,9%	83,3%	62,5%	94,6%	99,5%	61,2%	62,3%	75,0%	69,6%	50,0%	37,5%	100,0%	45,5%	95,9%	100,0%
SVM (noyau radial)	<b>Précision</b>	<b>90,8%</b>	<b>89,3%</b>	<b>81,8%</b>	<b>84,8%</b>	<b>66,7%</b>	<b>83,3%</b>	<b>83,0%</b>	<b>80,9%</b>	<b>85,1%</b>	<b>83,0%</b>	<b>84,6%</b>	<b>83,3%</b>	<b>82,4%</b>	<b>76,5%</b>	<b>81,5%</b>	<b>87,0%</b>	<b>80,5%</b>	<b>79,6%</b>	<b>82,9%</b>	<b>79,7%</b>
	Kappa	79,5%	76,0%	63,5%	69,3%	18,2%	55,7%	46,4%	37,7%	25,1%	0,3%	62,4%	59,0%	63,7%	52,1%	35,9%	39,2%	20,7%	19,0%	31,1%	-0,1%
	<i>Irrégulier</i>	93,5%	92,1%	88,2%	88,9%	83,3%	92,3%	82,9%	81,0%	22,7%	1,0%	92,0%	91,4%	92,9%	88,9%	97,6%	95,7%	85,1%	85,0%	28,1%	0,0%
	<i>Régulier</i>	86,1%	84,2%	75,0%	80,0%	33,3%	60,0%	83,3%	80,0%	98,7%	99,6%	68,3%	65,5%	69,6%	62,5%	30,8%	37,5%	41,7%	38,5%	97,5%	99,9%
RF classification	<b>Précision</b>	<b>89,7%</b>	<b>86,4%</b>	<b>81,8%</b>	<b>87,9%</b>	<b>72,2%</b>	<b>77,8%</b>	<b>85,1%</b>	<b>83,0%</b>	<b>91,9%</b>	<b>86,9%</b>	<b>84,7%</b>	<b>82,0%</b>	<b>78,4%</b>	<b>78,4%</b>	<b>85,2%</b>	<b>77,8%</b>	<b>79,6%</b>	<b>77,0%</b>	<b>89,7%</b>	<b>80,9%</b>
	Kappa	75,4%	69,7%	64,1%	75,6%	36,6%	45,5%	54,6%	42,7%	54,3%	31,3%	61,2%	55,0%	56,6%	55,6%	35,1%	23,0%	25,8%	10,5%	55,7%	24,3%
	<i>Irrégulier</i>	92,8%	88,2%	93,3%	94,1%	90,9%	91,7%	85,0%	81,4%	63,5%	28,3%	91,6%	89,9%	92,3%	89,3%	95,6%	95,1%	86,5%	83,8%	60,3%	31,1%
	<i>Régulier</i>	83,9%	83,0%	72,2%	81,3%	42,9%	50,0%	85,7%	100,0%	97,6%	98,7%	69,3%	64,3%	64,0%	65,2%	33,3%	23,1%	41,2%	28,6%	97,5%	94,0%
RF prédiction	RMSE Gini	0,047	0,050	0,052	0,060	0,075	0,077	0,058	0,066	0,079	0,108	0,050	0,053	0,050	0,056	0,081	0,097	0,069	0,076	0,087	0,104
	RMSE Asymétrie	0,110	0,108	0,120	0,119	0,110	0,100	0,133	0,131	0,196	0,186	0,121	0,121	0,116	0,113	0,126	0,120	0,146	0,147	0,192	0,178
	<b>Précision</b>	<b>89,9%</b>	<b>87,2%</b>	<b>87,9%</b>	<b>84,8%</b>	<b>66,7%</b>	<b>83,3%</b>	<b>83,0%</b>	<b>83,0%</b>	<b>86,8%</b>	<b>86,4%</b>	<b>85,2%</b>	<b>82,6%</b>	<b>80,4%</b>	<b>80,4%</b>	<b>85,2%</b>	<b>75,9%</b>	<b>77,9%</b>	<b>79,6%</b>	<b>86,8%</b>	<b>79,2%</b>
	Kappa	77,5%	71,4%	75,6%	69,8%	28,9%	62,0%	42,7%	42,7%	46,1%	31,0%	63,8%	58,0%	60,1%	58,1%	35,1%	27,9%	15,8%	28,7%	50,6%	12,6%
	<i>Irrégulier</i>	92,6%	90,7%	94,1%	93,8%	90,0%	100,0%	81,4%	81,4%	49,4%	29,9%	92,6%	90,1%	92,6%	87,1%	95,6%	97,4%	84,7%	87,2%	46,1%	14,0%
<i>Régulier</i>	85,1%	80,9%	81,3%	76,5%	37,5%	57,1%	100,0%	100,0%	94,9%	96,9%	68,9%	66,3%	66,7%	70,0%	33,3%	25,0%	33,3%	42,1%	97,5%	96,0%	
kMSN	RMSE Gini	0,049	0,057	0,059	0,058	0,079	0,084	0,056	0,069	0,076	0,111	0,053	0,057	0,054	0,055	0,070	0,085	0,070	0,071	0,094	0,115
	RMSE Asymétrie	0,116	0,120	0,132	0,128	0,116	0,124	0,150	0,133	0,198	0,208	0,131	0,132	0,124	0,120	0,130	0,137	0,153	0,153	0,196	0,195
	<b>Précision</b>	<b>88,3%</b>	<b>87,1%</b>	<b>87,9%</b>	<b>87,9%</b>	<b>77,8%</b>	<b>61,1%</b>	<b>83,0%</b>	<b>74,5%</b>	<b>86,0%</b>	<b>76,8%</b>	<b>82,5%</b>	<b>81,6%</b>	<b>80,4%</b>	<b>78,4%</b>	<b>88,9%</b>	<b>83,3%</b>	<b>81,4%</b>	<b>81,4%</b>	<b>82,7%</b>	<b>75,1%</b>
	Kappa	73,0%	71,1%	75,6%	75,6%	52,6%	22,2%	46,4%	24,6%	48,4%	3,2%	57,6%	55,8%	60,1%	55,6%	51,4%	31,5%	22,5%	22,5%	39,9%	3,7%
	<i>Irrégulier</i>	93,7%	90,9%	94,1%	94,1%	100,0%	88,9%	82,9%	79,5%	59,2%	14,9%	89,9%	88,7%	92,6%	89,3%	97,8%	95,5%	85,3%	85,3%	46,7%	13,1%
	<i>Régulier</i>	77,9%	80,3%	81,3%	81,3%	50,0%	33,3%	83,3%	50,0%	91,7%	88,1%	66,1%	66,0%	66,7%	65,2%	44,4%	30,0%	45,5%	45,5%	92,0%	90,5%

Précision : taux de placettes bien classées / Kappa : coefficient Kappa / Irrégulier : taux de placettes irrégulières bien classées / Régulier : taux de placettes régulières bien classées  
 RMSE : Root Mean Square Error (sur le coefficient de Gini et l'asymétrie de Lorenz)

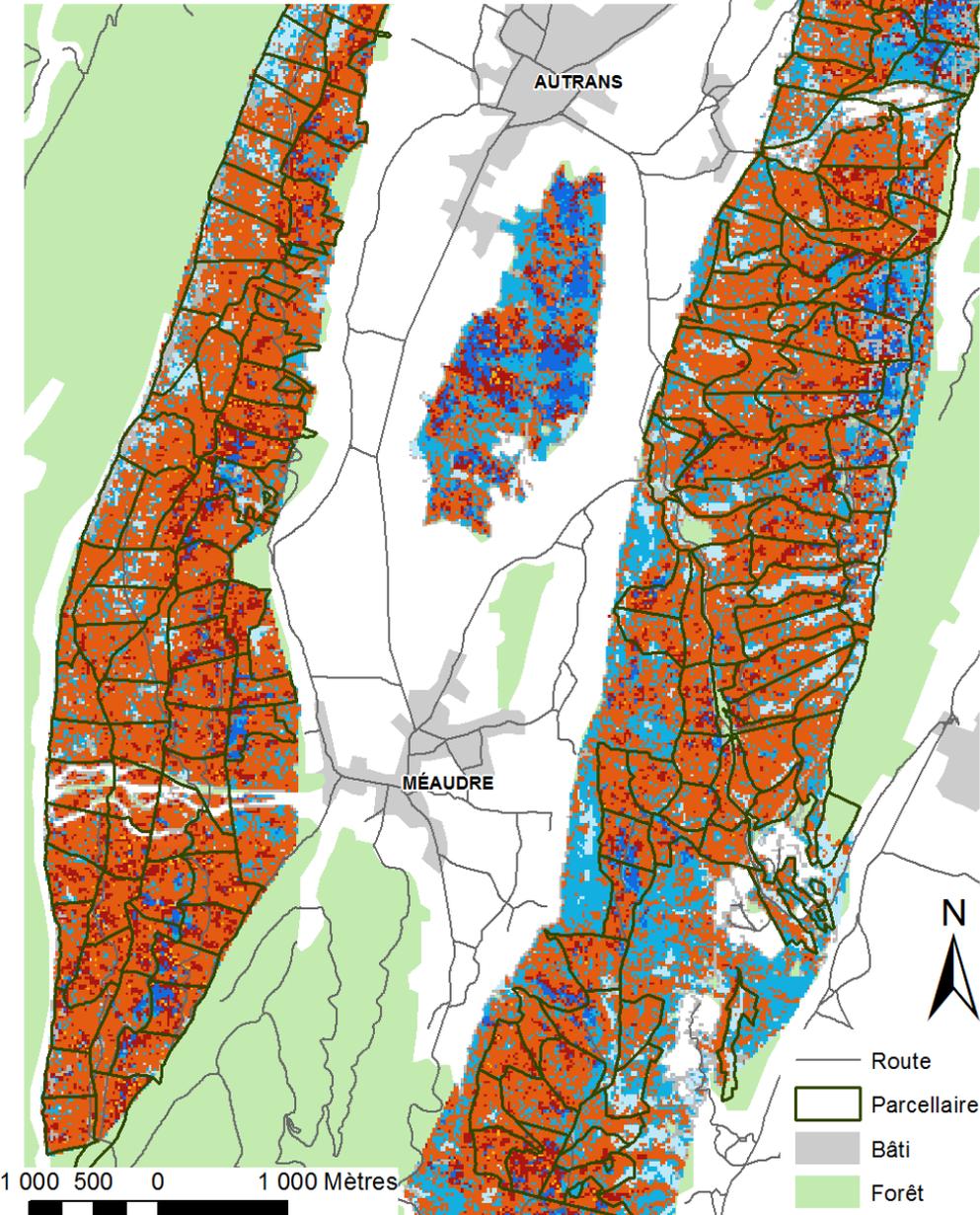
Carte lissée de la probabilité Random Forest d'appartenance à un type structural

■ 100% irrégulier  
■ 50% régulier - 50% irrégulier  
■ 100% régulier



Carte des types structuraux détaillés issus de modèles de prédiction LiDAR

■ Clair      ■ Régulier à BM      ■ Bimodal  
■ Régulier à PB      ■ Régulier à GB      ■ Irrégulier à PB/BM  
■ Irrégulier à GB



**Notes explicatives de la carte des peuplements réalisée par Mathilde Redon (2012) :**

Cette carte a été réalisée avec l'appui des cartes de peuplements disponibles dans les documents d'aménagement des différentes forêts concernées et par photo-interprétation de la BD ortho 2006 de l'IGN. Elle a été validée par un réseau de placettes, des visites de parcelles et par des agents de l'ONF.

Une typologie des peuplements simplifiée, regroupant de nombreuses appellations différentes a été constituée. Elle comprend 7 types de peuplements + une classe « autres » :

- **Futaie régulière monostrate** : une strate dominante, distribution des diamètres en « cloche », une ou deux catégories de diamètres contigües dominant.
- **Futaie régulière bistrates** : futaie à deux strates dominantes, distribution des diamètres avec deux modes distincts.
- **Futaie irrégulière** : au moins 3 strates distinctes, distribution des diamètres fortement asymétrique, se rapprochant d'un « J » inversé.
- **Taillis** : Une strate distincte, distribution des diamètres en « cloche », plus ou moins centrée sur les PB selon l'âge du taillis.
- **Mélange Taillis – Futaie** : Tout peuplement comportant des arbres de futaie et du taillis : taillis sous futaie mais aussi taillis avec réserve résineuse et taillis en reconversion en futaie sur souche. Distribution des diamètres avec un mode plus ou moins centré sur les PB et un ou deux autres modes moins étendus dans les BM, GB ou TGB.
- **Plantation** : Jeunes plantations (Diamètre quadratique moyen à l'hectare  $\leq 30$ ), dont l'agencement spatial est encore visible sur photographie aérienne (les plantations de  $Dq > 30$  sont confondues avec la futaie régulière monostrate). Distribution des diamètres en « cloche », une ou deux catégories de diamètres dominant.
- **Jeune peuplement** : Futaies jeunes ou à faible développement, non plantées (les jeunes plantations sont dans la catégorie « plantation »). Distribution des diamètres dominée par les PER et les PB.
- **Autre** : Zones non-boisées (rochers, éboulis, pelouses, etc.) et peuplements buissonnants ou très clairsemés, sans structure définie.

**Carte des types de peuplements réalisée par Mathilde Redon**

