



# BIBLIOTHÈQUE

CÉGEP DE L'ABITIBI-TÉMISCAMINGUE  
UNIVERSITÉ DU QUÉBEC EN ABITIBI-TÉMISCAMINGUE

## Mise en garde

La bibliothèque du Cégep de l'Abitibi-Témiscamingue et de l'Université du Québec en Abitibi-Témiscamingue (UQAT) a obtenu l'autorisation de la personne autrice de ce document afin de diffuser, dans un but non lucratif, une copie de son œuvre dans [Depositum](#), site d'archives numériques, gratuit et accessible à tous. L'autrice ou l'auteur conserve néanmoins ses droits de propriété intellectuelle, dont son droit d'auteur, sur cette œuvre.

## Warning

The library of the Cégep de l'Abitibi-Témiscamingue and the Université du Québec en Abitibi-Témiscamingue (UQAT) obtained the permission of the author to use a copy of this document for nonprofit purposes in order to put it in the open archives [Depositum](#), which is free and accessible to all. The author retains ownership of the copyright on this document.

Université du Québec en Abitibi-Témiscamingue

ROBUSTESSE DES MODÈLES SPATIAUX ET TEMPORELS DES ATTRIBUTS  
DENDROMÉTRIQUES FORESTIERS À L'AIDE DU BALAYAGE LASER  
AÉROPORTÉ MULTI-TEMPOREL

Mémoire  
présentée  
comme exigence partielle  
de la Maîtrise en Écologie et aménagement des écosystèmes forestiers

Par  
Carolina Villalobos

Mars 2026

© Carolina Villalobos, 2026

## REMERCIEMENTS

Je souhaite exprimer ma profonde gratitude à toutes les personnes qui ont contribué à la réalisation de ce mémoire.

Avant tout, je tiens à exprimer ma profonde gratitude à ma famille, en particulier à mon époux Alejandro et à mes enfants, Francisca et Nicolás. Leur amour, leur patience et leur soutien ont été une véritable source d'encouragement tout au long de ce parcours. Leur compréhension et leur appui, surtout lors du changement de pays et des nombreux défis qu'il a apportés, ainsi que durant les fins de semaine et les soirées consacrées à la recherche et à l'écriture, ont rendu cette aventure possible et encore plus significative.

J'exprime également ma sincère reconnaissance à mon directeur, Osvaldo Valeria (Université du Québec en Abitibi-Témiscamingue), et à mon codirecteur, Richard Fournier (Université de Sherbrooke), pour leur accompagnement, leurs conseils éclairés et leur rigueur scientifique. Leur expertise et leur bienveillance ont profondément enrichi ce travail et ont nourri ma progression académique et personnelle.

Je remercie Claude-Michel Bouchard pour son soutien essentiel dans la gestion, la mise à disposition et la compréhension des données de la FERLD, ainsi que pour sa disponibilité et ses échanges toujours constructifs.

Mes remerciements s'adressent également au personnel de soutien de l'Institut de recherche sur les forêts (IRF), en particulier Danielle, Marie-Hélène et Frédérique, pour leur efficacité, leur disponibilité et leur gentillesse.

Enfin, une pensée toute spéciale à mes amis chiliens, Carol, Oscar et Javiera, dont l'écoute, l'amitié et la motivation ont accompagné chaque étape de cette aventure scientifique, humaine et personnelle.

## **DÉDICACE**

*À mes enfants : Francisca et Nicolás pour leur amour, leur patience et leur lumière  
qui ont accompagné chaque page de ce travail.*

## RÉSUMÉ

L'inventaire forestier constitue un outil essentiel pour la gestion durable des forêts boréales. Cette étude évalue la performance et la transférabilité de modèles prédictifs basés sur des acquisitions multi-temporelles de balayage laser aéroporté (ALS) et des données de placettes d'inventaire dans la Forêt d'Enseignement et de Recherche du Lac Duparquet (FERLD). Quatre années d'acquisition (2003, 2007, 2017, 2023) ont été utilisées pour dériver des métriques structurales à partir de rasters (résolution de 1 m) selon l'approche par zones (ABA). Trois attributs dendrométriques d'intérêt ont été ciblés : la hauteur de Lorey ( $h_L$ ), la surface terrière (BA) et le volume marchand brut (V). Les modèles ont été calibrés par stratification selon le type de couvert forestier (résineux, feuillu, mixte) à l'aide de méthodes paramétriques (régression linéaire multiple, LR) et non paramétriques (forêts aléatoires, RF).

Les résultats montrent que l' $h_L$  est l'attribut le plus robuste et le plus transférable, particulièrement dans les couverts résineux où la stabilité structurelle renforce la performance prédictive. Les modèles RF surpassent généralement les modèles LR grâce à leur flexibilité, bien que ces derniers se révèlent compétitifs dans des contextes homogènes. Pour la BA et le V, la performance est restée plus limitée, reflétant la dépendance de ces attributs à la densité et au diamètre, paramètres que l'ALS ne mesure qu'indirectement. L'analyse spatiale a mis en évidence une transférabilité bidirectionnelle des modèles de  $h_L$  entre les secteurs d'origine 1760 et 1923, malgré leurs différences de composition spécifique. Cette stabilité des performances s'explique par la structure verticale relativement homogène et la prédominance des essences conifériennes dans les deux secteurs, soulignant le rôle déterminant de la stabilité structurelle dans la robustesse des prédictions.

Sur le plan opérationnel, les résultats suggèrent que les inventaires forestiers québécois peuvent optimiser leurs stratégies temporelles en privilégiant des intervalles d'acquisition intermédiaires, ce qui permet d'atteindre un compromis efficace entre précision des modèles et coûts d'acquisition. En ce qui concerne la composition forestière, les acquisitions réalisées dans les couverts résineux apparaissent plus prometteuses, car la stabilité structurelle de ces forêts favorise une meilleure transférabilité des modèles. Pour les forêts feuillues, caractérisées par une variabilité structurelle élevée, il convient d'adopter des approches adaptatives, telles que la stratification par classes d'âge ou l'intégration de données supplémentaires.

Mots-clés : télédétection, balayage laser aéroporté (ALS), l'Approche par zones (ABA), modélisation forestière, transférabilité spatiale, analyse multi-temporelle, forêt boréale, FERLD.

## ABSTRACT

Forest inventory is an essential tool for the sustainable management of boreal forests. This study evaluates the performance and transferability of predictive models based on multi-temporal airborne laser scanning (ALS) acquisitions and field plot data in the Forêt d'Enseignement et de Recherche du Lac Duparquet (*FERLD*). Four acquisition years (2003, 2007, 2017, 2023) were used to derive structural metrics from rasters (1 m resolution) using the Area-Based Approach (ABA). Three dendrometric attributes of interest were targeted: Lorey's height ( $h_L$ ), basal area (BA), and gross merchantable volume (V). Models were calibrated by stratification according to forest cover type (coniferous, broadleaf, mixed) using parametric methods (multiple linear regression, LR) and non-parametric methods (Random Forest, RF).

The results indicate that  $h_L$  is the most robust and transferable attribute, particularly in coniferous stands where structural stability enhances predictive performance. RF models generally outperformed LR due to their flexibility, although LR remained competitive in more homogeneous contexts. For BA and V, predictive performance was more constrained, reflecting their reliance on stem density and diameter, parameters that are only indirectly captured by ALS. The spatial analysis revealed a bidirectional transferability of Lorey's height models between the 1760 and 1923 origin sectors, despite their contrasting species composition. This stability in model performance is explained by the relatively homogeneous vertical structure and the dominance of conifer species in both sectors, highlighting the key role of structural stability in ensuring prediction robustness.

From an operational perspective, the results suggest that Québec's Forest inventory programs can optimize temporal strategies by prioritizing intermediate acquisition intervals, thereby achieving an effective balance between model accuracy and acquisition costs. Regarding forest composition, ALS acquisitions in coniferous stands appear more promising, as the structural stability of these forests favors better model transferability. In contrast, for broadleaf forests, characterized by higher structural variability, adaptive strategies such as stratification by age class or the integration of complementary data should be considered.

Keywords: remote sensing, airborne laser scanning (ALS), area-based approach (ABA), forest modelling, spatial transferability, multi-temporal analysis, boreal forest, FERLD.

## TABLE DES MATIÈRES

REMERCIEMENTS .....	III
DÉDICACE .....	IV
RÉSUMÉ .....	V
ABSTRACT .....	VI
TABLE DES MATIÈRES .....	VII
LISTE DES FIGURES .....	X
LISTE DES TABLEAUX .....	XI
LISTE DES SIGLES ET DES ABRÉVIATIONS .....	XII
INTRODUCTION .....	1
1. MISE EN CONTEXTE .....	2
1.1 Inventaires forestiers au Québec .....	2
1.2 Balayage laser aéroporté et inventaires forestiers améliorés .....	2
1.3 Modélisation des attributs forestiers à l'aide de l'ALS .....	4
1.3.1 Approche par zones .....	4
1.3.2 Applications de l'ALS pour l'étude de la dynamique forestière .....	7
1.4 Modélisation paramétrique et non paramétrique .....	8
1.5 Approches multi-temporelles et transférabilité spatiale .....	12
1.6 Perturbations naturelles et dynamiques forestières .....	15
1.7 Objectifs et hypothèses .....	17
2. ROBUSTESSE DES MODÈLES SPATIAUX ET TEMPORELS DES ATTRIBUTS DENDROMÉTRIQUES FORESTIERS À L'AIDE DU BALAYAGE LASER AÉROPORTÉ MULTI-TEMPOREL .....	19
2.1 Introduction .....	19
2.2 Matériels et méthodes .....	21
2.2.1 Aire d'étude .....	21
2.2.2 Données LiDAR aéroportées .....	23
2.2.3 Placettes d'inventaire .....	23
2.2.4 Analyse temporelle et spatiale .....	26
2.3 Résultats .....	30
2.3.1 Modélisation multi-temporelle .....	30

2.3.1.1	Hauteur de Lorey .....	30
2.3.1.2	Surface terrière.....	34
2.3.1.3	Volume marchand brut .....	38
2.3.2	Modélisation spatiale .....	44
2.3.2.1	Hauteur de Lorey.....	44
2.3.2.2	Surface terrière.....	46
2.3.2.3	Volume marchand brut .....	48
2.4	Discussion.....	50
2.4.1	Impact des configurations ALS multi-temporelles sur la capacité prédictive .....	50
2.4.2	Transférabilité spatiale.....	53
2.4.3	Comparaison des méthodes paramétriques et non paramétriques ....	54
2.4.4	Limitations et perspectives d'amélioration.....	55
2.5	Conclusion .....	56
	CONCLUSION GÉNÉRALE.....	59
	ANNEXE A – COMPOSITION SPÉCIFIQUE SELON LE TYPE DE COUVERT ET L'ANNÉE D'ORIGINE DU FEU .....	61
	ANNEXE B – DESCRIPTION DES VARIABLES EXPLICATIVES ET NOMENCLATURE DES MÉTRIQUES (MHC, MNT, PENTE, TPI, TWI).....	62
	ANNEXE C – PERFORMANCE DES MODÈLES DE RÉGRESSION LINÉAIRE POUR LA PRÉDICTION DE LA HAUTEUR DE LOREY SELON LES TYPES DE COUVERT FORESTIER.....	64
	ANNEXE D – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA HAUTEUR DE LOREY, SELON LES DIFFÉRENTES COMBINAISONS TEMPORELLES.....	66
	ANNEXE E – PERFORMANCE DES MODÈLES DE RÉGRESSION LINÉAIRE DANS POUR LA PRÉDICTION DE LA SURFACE TERRIÈRE SELON LES TYPES DE COUVERT FORESTIER.....	69
	ANNEXE F – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA SURFACE TERRIÈRE, SELON LES DIFFÉRENTES COMBINAISONS TEMPORELLES.....	71

ANNEXE G – PERFORMANCE DES MODÈLES DE RÉGRESSION LINÉAIRE POUR LA PRÉDICTION DU VOLUME MARCHAND BRUT SELON LES TYPES DE COUVERT FORESTIER .....	74
ANNEXE H – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DU VOLUME MARCHAND BRUT, SELON LES DIFFÉRENTES COMBINAISONS TEMPORELLES.....	76
ANNEXE I – IMPORTANCE DE VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA HAUTEUR DE LOREY ( $H_L$ ) .....	79
ANNEXE J – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA SURFACE TERRIÈRE (BA).....	80
ANNEXE K – IMPORTANCE DE VARIABLES (RANDOM FOREST) POUR LA PREDICTION DE VOLUME MARCHAND BRUT (V) .....	81
LISTE DE REFERENCES .....	82

## LISTE DES FIGURES

Figure 1 Répartition spatiale des 577 placettes d'inventaire dans la FERLD (Forêt d'Enseignement et de Recherche du Lac Duparquet). .....	22
Figure 2 Performance des modèles Random Forest (RF) pour la prédiction de la hauteur de Lorey ( $h_L$ ) dans le couvert forestier résineux. ....	31
Figure 3 Performance des modèles Random Forest (RF) pour la prédiction de la hauteur de Lorey ( $h_L$ ) dans le couvert forestier feuillu. ....	32
Figure 4 Performance des modèles Random Forest (RF) pour la prédiction de la hauteur de Lorey ( $h_L$ ) dans le couvert forestier mixte. ....	33
Figure 5 Performance des modèles Random Forest (RF) pour la prédiction de la surface terrière (BA) dans le couvert forestier résineux. ....	35
Figure 6 Performance des modèles Random Forest (RF) pour la prédiction de la surface terrière (BA) dans le couvert forestier feuillu. ....	36
Figure 7 Performance des modèles Random Forest (RF) pour la prédiction de la surface terrière (BA) dans le couvert forestier mixte. ....	37
Figure 8 Performance des modèles Random Forest (RF) pour la prédiction du volume marchand brut (V) dans le couvert forestier résineux. ....	40
Figure 9 Performance des modèles Random Forest (RF) pour la prédiction du volume marchand brut (V) dans le couvert forestier feuillu. ....	41
Figure 10 Performance des modèles Random Forest (RF) pour la prédiction du volume marchand brut (V) dans le couvert forestier mixte. ....	42

## LISTE DES TABLEAUX

Tableau 1 Résumé des attributs structuraux des placettes d'inventaire pour chaque année d'acquisition ALS. ....	26
Tableau 2 Description des variables explicatives utilisées pour la calibration des modèles et de leurs sources de données correspondantes. ....	29
Tableau 3 Performance des modèles pour la prédiction de la hauteur de Lorey ( $h_L$ ). .....	45
Tableau 4 Performance des modèles pour la prédiction de la surface terrière (BA).46	
Tableau 5 Performance des modèles pour la prédiction du volume marchand brut (V). .....	49

## LISTE DES SIGLES ET DES ABRÉVIATIONS

ABA	Area Based Approach (Approche par zones)
ACP	Principal component analysis (Analyse en composantes principales)
ALS	Airborne Laser Scanning (balayage laser aéroporté)
BA	Basal Area (m <sup>2</sup> /ha) (Surface terrière)
DAP	Digital Aerial Photogrammetry (Photogrammétrie aérienne numérique)
DHP	Diamètre à Hauteur de poitrine (1,3 m)
FERLD	Forêt d'Enseignement et de Recherche du Lac Duparquet
h <sub>L</sub>	Hauteur de Lorey (m)
IFA	Inventaire Forestier Amélioré
K-NN	K-nearest neighbours (k-plus proches voisins)
LiDAR	Light Detection and Ranging (Téledétection par laser)
LR	Multiple Linear Regression (Régression linéaire multiple)
ME	Erreur moyenne
ME%	Erreur moyenne relative
MHC	Modèle de hauteur de canopée
MNT	Modèles numériques de terrain
OLS	Ordinary Least Squares (Régression des moindres carrés ordinaires)
R <sup>2</sup>	Coefficient de détermination
RF	Random Forest (Forêt aléatoire)
RMSE	Erreur quadratique moyenne
RMSE%	Erreur quadratique moyenne relative
SUR	Seemingly Unrelated Regression (Régression apparemment non corrélée)
TPI	Indice topographique de position
TWI	Indice topographique d'humidité
V	Volume marchand brut (m <sup>3</sup> /ha)

## INTRODUCTION

L'analyse de la structure et de la dynamique des forêts boréales constitue un enjeu majeur pour la gestion durable et la planification des ressources. Au cours des vingt dernières années, la technologie de balayage laser aéroporté (ALS) s'est rapidement imposée comme une solution essentielle pour la réalisation d'inventaires forestiers. Intégré à l'Inventaire forestier amélioré (IFA), il permet de produire des estimations spatialisées de variables dendrométriques clés. L'Approche par zones (ABA), qui relie des métriques ALS à des mesures de terrain issues de placettes géoréférencées, est la méthode la plus couramment utilisée.

Ce mémoire s'inscrit dans la nécessité d'évaluer la stabilité temporelle et la transférabilité spatiale des modèles ABA dans des paysages boréaux marqués par des perturbations contrastées. L'objectif général est de développer et d'évaluer des modèles paramétriques (régression linéaire multiple) et non paramétriques (forêts aléatoires) pour la prédiction de la hauteur de Lorey, de la surface terrière et du volume marchand brut, à partir de données ALS multi-temporelles dérivées de rasters. Trois objectifs spécifiques guident cette recherche : analyser l'effet de différentes combinaisons temporelles sur la performance des modèles ; évaluer la transférabilité spatiale entre les secteurs établis après les incendies de 1760 et de 1923 ; comparer la performance des approches paramétriques et non paramétriques.

La méthodologie repose sur quatre campagnes ALS (2003, 2007, 2017, 2023) acquises dans la Forêt d'Enseignement et de Recherche du Lac Duparquet (FERLD, Abitibi, Québec), accompagnées de données provenant de placettes d'inventaire.

La structure du mémoire est organisée comme suit : le chapitre 1 expose l'état de la question ainsi que le cadre conceptuel de l'étude; le chapitre 2 décrit l'aire d'étude, les données, la méthodologie employée et les résultats; la discussion analyse la performance et la transférabilité des modèles. Enfin, la conclusion générale présente une synthèse ainsi que des recommandations opérationnelles.

## 1. MISE EN CONTEXTE

### 1.1 Inventaires forestiers au Québec

Les forêts constituent une ressource naturelle majeure au Québec, couvrant environ 906 168 km<sup>2</sup>, et représentent ainsi un élément structurant du paysage provincial. Une grande partie de cette superficie se situe dans la zone boréale, qui s'étend entre les latitudes 48° et 58° nord et se divise écologiquement en trois sous-zones distinctes. La sous-zone de la forêt fermée, la plus méridionale (48° à 52°N), se caractérise par des couverts denses dominés par le sapin baumier (*Abies balsamea* (L.) Mill.) et l'épinette noire (*Picea mariana* (Mill.) Britton, Sterns & Poggenb.). Cette zone comprend également des espèces feuillues, notamment le bouleau blanc (*Betula papyrifera* Marshall) et le peuplier faux-tremble (*Populus tremuloides* Michx), en particulier dans le domaine bioclimatique du sapin baumier et du bouleau blanc.

À des fins d'inventaire forestier et d'aménagement, le territoire québécois comprend une zone d'inventaire intensif, qui couvre une superficie de 547 404 km<sup>2</sup> dans la partie méridionale de la zone boréale (Gouvernement du Québec, 2024). Depuis 1970, un inventaire écoforestier décennal est réalisé dans les forêts publiques et privées du sud du Québec. Ce dispositif de suivi systématique constitue une base essentielle pour l'évaluation des dynamiques forestières et l'élaboration de stratégies de gestion durable à l'échelle du territoire (Gouvernement du Québec, 2025a). Les méthodes d'inventaire ont considérablement évolué, passant de l'interprétation de photographies aériennes à l'échelle 1:15 000 à l'intégration, à partir de 2015, du balayage laser aéroporté (LiDAR aéroporté, *Airborne Laser Scanning*, ALS), utilisant la technologie connue sous le nom de LiDAR (Light Detection and Ranging), ainsi que de l'imagerie satellitaire (Direction des inventaires forestiers, 1975, 1984, 2009; Gouvernement du Québec, 2025b). Entre 2016 et 2023, des données ALS ont été collectées sur l'ensemble du territoire forestier aménagé, et une seconde campagne d'acquisition a été lancée en 2024 (Gouvernement du Québec, 2025a).

### 1.2 Balayage laser aéroporté et inventaires forestiers améliorés

Le LiDAR constitue une technologie de télédétection active basée sur l'émission d'impulsions laser et la mesure du temps de retour de celles-ci après réflexion sur les

surfaces; ce principe permet d'obtenir des mesures de distance d'une grande précision. Déployé à partir de plateformes aéroportées, spatiales ou terrestres, le LiDAR permet de générer des représentations tridimensionnelles détaillées du relief ainsi que de la structure de la végétation. Contrairement aux images optiques, les données LiDAR ne sont pas soumises à des distorsions géométriques et ne nécessitent pas de processus de géoréférencement complexe, dans la mesure où elles enregistrent directement les coordonnées spatiales en trois dimensions. L'un des principaux avantages de cette technologie dans le contexte forestier réside dans sa capacité à pénétrer des canopées denses, permettant ainsi de caractériser simultanément la structure verticale de la végétation et l'élévation du sol (Dubayah *et al.*, 2000; Lefsky *et al.*, 2002; Dong *et al.*, 2017).

Dans le secteur forestier, le LiDAR offre deux avantages principaux : il permet l'extraction d'informations structurales détaillées, impossibles à obtenir par les méthodes conventionnelles sur le terrain, et facilite une acquisition de données rapide, précise et rentable à grande échelle (Nelson, 2013). Parmi les différentes plateformes LiDAR, l'ALS se révèle particulièrement bien adapté aux applications forestières. Il permet l'estimation directe de variables structurales essentielles telles que la hauteur de la canopée, la fraction de couverture, l'élévation du terrain, la pente et l'orientation (Wulder *et al.*, 2008). Il permet également la production de produits compatibles avec les systèmes d'information géographique, tels que des modèles de hauteur de la végétation, des délimitations de couronnes, des localisations de tiges, des cartes d'indice de surface foliaire et de biomasse (Lim *et al.*, 2003; Goodbody *et al.*, 2024). La possibilité de réaliser des acquisitions ALS répétées permet de suivre l'évolution temporelle des écosystèmes forestiers, notamment en ce qui concerne la croissance et les perturbations. Sa capacité d'adaptation à différentes échelles spatiales autorise des analyses allant du niveau de la parcelle à celui du paysage régional. Toutefois, l'ALS présente certaines limites : il fournit peu d'informations sur la structure interne du couvert forestier et son acquisition reste coûteuse, nécessitant généralement des levés aériens spécialisés réalisés dans des conditions météorologiques favorables, ainsi qu'une coordination étroite avec les opérations de terrain (Béland *et al.*, 2019).

Les avancées technologiques associées à l'utilisation de l'ALS s'inscrivent dans les principes de l'inventaire forestier amélioré (IFA), qui combine des technologies de télédétection de pointe avec des mesures terrestres afin de produire des informations détaillées sur les attributs structuraux des forêts (White *et al.*, 2017; Stinson *et al.*, 2018). Les IFA intègrent des estimations dérivées de l'ALS pour des variables structurelles clés (hauteur, surface terrière et volume), offrant une résolution spatiale et structurelle significativement supérieure à celle des systèmes d'inventaire conventionnels. Cette amélioration résulte de la finesse spatiale des données et de la richesse des informations structurelles produites par l'ALS. Les IFA sont spatialement explicites et assurent une couverture continue et exhaustive de vastes superficies forestières (Goodbody *et al.*, 2024). À ce titre, l'IFA constitue une avancée majeure dans les méthodes d'inventaire forestier, en améliorant à la fois la précision des évaluations de la structure forestière et leur valeur opérationnelle pour le suivi et la gestion des ressources (White *et al.*, 2016). Au Québec, un produit IFA mur à mur a été développé pour les forêts résineuses, tandis que les forêts mixtes et boréales du sud de la province ne disposent pas encore d'équivalents (White *et al.*, 2025).

### 1.3 Modélisation des attributs forestiers à l'aide de l'ALS

#### 1.3.1 Approche par zones

Dans le cadre des IFA, l'approche par zones, ou Area-Based Approach (ABA) est largement reconnue comme la méthode principale pour l'estimation des attributs structuraux des forêts à grande échelle spatiale. ABA a été largement appliquée dans divers écosystèmes forestiers, soutenue par un corpus croissant de travaux attestant de sa robustesse méthodologique et de sa faisabilité opérationnelle. Cette approche repose sur des modèles statistiques qui établissent des liens entre les métriques dérivées de l'ALS et les mesures de terrain recueillies dans des placettes géoréférencées (Næsset *et al.*, 2001; Næsset, 2002). Des travaux récents ont également montré que l'utilisation de métriques dérivées de rasters constitue une alternative robuste aux métriques calculées directement à partir des nuages de points, en simplifiant le flux de travail et en accélérant la production des résultats, sans perte notable de précision (Mauro *et al.*, 2021).

L'ABA utilise à la fois des méthodes de modélisation paramétriques et non paramétriques (Means *et al.*, 2000; Næsset, 2002, 2004; Penner *et al.*, 2013). Ces approches établissent des relations statistiques entre les métriques dérivées de l'ALS et les variables mesurées sur le terrain, permettant la prédiction spatiale de caractéristiques dendrométriques essentielles telles que la hauteur des arbres, la surface terrière, le volume et la biomasse. La distinction fondamentale entre ces deux types de méthodes réside dans la manière dont elles représentent la relation entre les variables explicatives et les variables à prédire. Les méthodes paramétriques reposent sur des formes fonctionnelles prédéfinies et des hypothèses statistiques sous-jacentes, tandis que les méthodes non paramétriques n'imposent pas de telles contraintes. Bien que les modèles non paramétriques offrent une flexibilité accrue pour saisir des relations complexes (Gareth *et al.*, 2013), ils nécessitent généralement des échantillons de plus grande taille afin de garantir des prédictions stables et fiables.

La calibration des modèles s'effectue à l'échelle de la placette (au Québec, les placettes d'inventaire présentent généralement une superficie de 400 m<sup>2</sup>, correspondant à un rayon de 11,28 m), où les nuages de points ALS sont spatialement découpés afin de correspondre à chaque unité d'échantillonnage. Diverses métriques structurelles sont dérivées de ces données, incluant les percentiles de hauteur de canopée, les densités de retours ainsi que d'autres statistiques descriptives reflétant la structure de la végétation. Afin d'améliorer la précision des prédictions, Bouvier *et al.* (2015) ont proposé d'intégrer des métriques supplémentaires, telles que la hauteur moyenne de la canopée, la variation de hauteur, la proportion de trouées et le coefficient de variation du profil de densité foliaire, qui se sont révélées particulièrement pertinentes dans les peuplements forestiers à structure complexe.

Une fois validés, les modèles sont appliqués à l'ensemble du territoire couvert par les données ALS, à l'aide d'une grille continue de petites unités spatiales, généralement comprises entre 200 et 625 m<sup>2</sup>. Pour les provinces du Québec et du Nouveau-Brunswick cette grille correspond à des cellules de 400 m<sup>2</sup> (20 m × 20 m) (White *et al.*, 2025), ce qui est conforme aux dimensions standard des placettes de terrain. Chaque cellule est associée à des valeurs prédites calculées à partir des métriques

dérivées de l'ALS, permettant la production de cartes à haute résolution des attributs forestiers. Ce niveau de détail spatial constitue un avantage majeur par rapport aux inventaires traditionnels réalisés à l'échelle du peuplement, puisqu'il permet de saisir la variabilité intra-peuplement et de prendre des décisions plus précises et mieux informées en matière de gestion et de planification forestières (White *et al.*, 2017).

L'effet de la densité des points de l'ALS sur la précision des prédictions des attributs forestiers a été particulièrement étudié. Treitz *et al.* (2012) ont exploré cette problématique en simulant trois densités de points : 3,2 ; 1,6 et 0,5 impulsions/m<sup>2</sup>, à partir d'un même jeu de données LiDAR acquis en Ontario. En maintenant constants les paramètres du capteur et du vol, ils ont pu isoler l'effet de la densité d'échantillonnage. Leurs résultats ont montré que les estimations de variables telles que la hauteur dominante, la hauteur moyenne, le volume et la surface terrière demeuraient remarquablement stables, même à la densité la plus faible. Cela suggère que, pour de nombreuses applications opérationnelles, des densités LiDAR modérées peuvent suffire à produire des résultats fiables tout en maintenant des coûts d'acquisition raisonnables. Des résultats similaires ont été observés dans les forêts scandinaves et finlandaises. Par exemple, Gobakken *et al.* (2008) ont montré qu'une réduction de la densité des points, d'environ 1,13 à 0,25 impulsions/m<sup>2</sup>, n'entraînait qu'une légère augmentation des erreurs de prédiction pour des attributs tels que la hauteur moyenne des arbres, la surface terrière et le volume, à condition que la taille des placettes et la stratégie d'échantillonnage soient bien optimisées. De même, Maltamo *et al.* (2006) ont constaté une dégradation minimale de la précision des estimations du volume des tiges pour des densités aussi faibles que 0,13 impulsions/m<sup>2</sup>, tandis que Strunk *et al.* (2012) ont rapporté des estimations cohérentes de la surface terrière et de la hauteur de Lorey même à 0,05 impulsions/m<sup>2</sup>. Enfin, Socha *et al.* (2017) ont confirmé la robustesse des modèles de hauteur dominante dérivés de l'ALS, en montrant des écarts inférieurs à un mètre entre deux acquisitions (2007 et 2012, avec respectivement 4 et 6 impulsions/m<sup>2</sup>), validant ainsi la stabilité temporelle des prédictions de l'ABA malgré les variations de la densité de points et des dates d'acquisition.

### 1.3.2 Applications de l'ALS pour l'étude de la dynamique forestière

Au Québec, l'ALS est devenue un outil essentiel pour l'analyse de la structure forestière, de la succession écologique et de la dynamique des perturbations, en particulier dans les écosystèmes boréaux et les forêts anciennes.

St-Onge *et al.* (2004) ont utilisé des données ALS multi-temporelles pour analyser la dynamique des trouées de canopée dans la Forêt d'enseignement et de recherche du lac Duparquet (FERLD). En comparant des modèles de hauteur de canopée sur un intervalle de cinq ans, les auteurs ont documenté l'ouverture, l'expansion et la fermeture de trouées, apportant ainsi l'une des premières démonstrations concrètes de la capacité de la technologie ALS à détecter des variations structurelles fines dans des forêts non aménagées. Par la suite, Martin *et al.* (2021) ont intégré des données ALS et des images satellitaires afin de classifier les types de vieilles forêts au sein de paysages boréaux soumis à diverses perturbations. Les résultats de l'étude démontrent que la télédétection permet de différencier de manière efficace les changements structurels induits par le feu, les épidémies d'insectes et les activités de récolte. Toutefois, la précision de classification diminue dans les peuplements en transition et dans les secteurs aux limites de perturbations diffuses, ce qui souligne la nécessité de stratégies de calibration améliorées dans des environnements écologiquement complexes.

Plus récemment, Martin *et al.* (2022) ont appliqué des modèles de hauteur de canopée dérivés de l'ALS à une région boréale de 3 006 km<sup>2</sup> afin de cartographier la variabilité structurale. À l'aide d'un modèle Random Forest calibré sur des peuplements de référence, ils ont identifié des patrons structuraux associés aux stades de succession, dépassant les explications fondées uniquement sur l'âge des arbres. Leurs résultats plaident en faveur d'un passage des classifications forestières basées sur l'âge vers des classifications fondées sur la structure, en particulier dans un contexte de conservation. Ils mettent également en évidence l'importance des données ALS accessibles publiquement pour l'identification des peuplements de fin de succession, souvent oubliés dans les inventaires traditionnels.

Enfin, Leboeuf et al. (2022) ont développé et validé des modèles paramétriques à grande échelle afin d'estimer le volume de bois marchand en mobilisant des données de placettes historiques et des acquisitions ALS asynchrones. En se concentrant sur une superficie de 440 000 km<sup>2</sup> de la forêt aménagée du Québec, ils ont identifié la hauteur dominante et un indice de densité relative comme prédicteurs robustes, pouvant être systématiquement dérivés des données de terrain et de l'ALS. Les modèles obtenus ont permis d'estimer le volume à une résolution de 10 × 10 m. La validation des modèles a été réalisée à partir de sites de coupe sélectionnés en fonction de leur proximité temporelle avec les acquisitions ALS, c'est-à-dire récoltés moins de cinq saisons de croissance après le vol LiDAR. Ces sites disposaient de mesures précises des volumes récoltés et résiduels, calculées à l'aide d'équations masse volume et d'inventaires post-récolte. Cet exemple illustre l'intérêt d'approches ne nécessitant pas d'acquisitions synchronisées, ouvrant ainsi la voie à des alternatives évolutives et économiquement avantageuses pour les programmes d'inventaire opérationnel.

Collectivement, ces travaux démontrent comment l'ALS est en train de transformer l'évaluation des attributs dendrométriques au Québec. En facilitant la détection des changements temporels, la reconstitution des historiques de perturbations et l'identification des secteurs à haute valeur de conservation, l'ALS apporte à la fois une contribution scientifique et une valeur opérationnelle pour la gestion des paysages forestiers complexes.

#### *1.4 Modélisation paramétrique et non paramétrique*

La littérature scientifique documente un ensemble de techniques de modélisation couramment mobilisées dans le cadre de l'ABA. Les méthodes paramétriques s'appuient fréquemment sur des approches telles que la régression des moindres carrés ordinaires (Ordinary Least Squares, OLS) et la régression apparemment non corrélée (Seemingly Unrelated Regression, SUR). À titre d'exemple, Næsset (2002) a employé des modèles de régression multiplicatifs afin d'estimer divers attributs de peuplement, notamment la hauteur moyenne de Lorey, la hauteur dominante, le diamètre moyen (pondéré par la surface terrière), la densité de tiges, la surface

terrière et le volume. Les modèles ont été stratifiés selon la classe d'âge et la qualité de station, calibrés à partir de métriques dérivées de l'ALS, puis appliqués à l'échelle de la cellule d'analyse, avec agrégation des résultats au niveau du peuplement. Les meilleures performances ont été obtenues pour la hauteur moyenne de Lorey, la surface terrière et le volume, tandis que des coefficients de détermination ( $R^2$ ) plus faibles ont été observés pour la densité de tiges et le diamètre moyen.

Dans une étude ultérieure, Næsset (2004) a collecté des données ALS sur une superficie de 6 500 ha dans le sud-est de la Norvège, lesquelles ont été analysées à l'aide de régressions multiplicatives linéaires afin d'estimer six attributs de peuplement, tant dans les forêts jeunes que matures. L'étude a également examiné l'influence de la topographie, des paramètres de vol et de la composition en essences. Les résultats ont indiqué que la hauteur moyenne des arbres, la hauteur dominante, le diamètre moyen, la surface terrière et le volume étaient estimés avec une précision supérieure à celle des valeurs correspondantes obtenues par mesures terrestres, tandis que la densité de tiges présentait un niveau de précision plus faible. Bien que les variations topographiques aient eu un impact limité sur la précision des estimations, les peuplements mixtes, conduisaient à une diminution de la performance des modèles.

Les techniques non paramétriques, telles que les forêts aléatoires (Random Forest, RF) (Breiman, 2001) et les k-plus proches voisins (k-nearest neighbours, k-NN), ont également démontré de fortes capacités prédictives dans le cadre de l'ABA. Par exemple, dans le nord du Minnesota, McRoberts *et al.* (2017) ont appliqué une technique de k-NN multivariée afin d'estimer simultanément six variables d'inventaire forestier : la superficie forestière, la surface terrière, le volume sur pied, la biomasse aérienne, la densité de tiges et la hauteur moyenne. Les prédictions ont été intégrées dans un cadre de régression généralisée assistée par modèle et comparées à des estimateurs post-stratifiés. Alors, que la post-stratification assurait la cohérence entre variables, elle se révélait moins précise pour les métriques continues, tandis que la méthode fondée sur k-NN produisait des prédictions multivariées cohérentes avec des

erreurs standard significativement plus faibles (23% à 35%), ce qui souligne son efficacité pour l'inventaire forestier à grande échelle.

Une étude comparative conduite par Penner et al. (2013) dans les forêts boréales de l'Ontario visait à évaluer la performance des modèles paramétriques et non paramétriques utilisant des données ALS. L'objectif était de prédire des attributs clés de l'inventaire forestier, incluant la hauteur dominante, la surface terrière marchande et le volume marchand brut. Les modèles paramétriques étaient fondés sur SUR, tandis que les méthodes non paramétriques incluaient RF et k-NN. La calibration des modèles a été réalisée à partir de 442 placettes d'inventaire couvrant huit types forestiers distincts, et la validation a été effectuée sur un jeu indépendant de 64 placettes. Les résultats ont démontré que RF ainsi que la SUR stratifiée atteignaient les niveaux de précision les plus élevés pour l'ensemble des attributs. Par exemple, le volume marchand brut a été prédit avec des RMSE de 6,92 m<sup>3</sup>/ha pour le RF et de 6,37 m<sup>3</sup>/ha pour la SUR stratifiée. De même, les prédictions de la surface terrière marchande présentaient des RMSE de 2,48 m<sup>2</sup>/ha et de 0,71 m<sup>2</sup>/ha pour RF et la SUR stratifiée, respectivement. Dans le cas de la hauteur dominante, les deux méthodes produisaient des estimations très précises, avec des RMSE compris entre 0,19 et 0,21 m. À l'inverse, k-NN affichait des erreurs de prédiction nettement plus élevées, en particulier pour le volume (10,34 m<sup>3</sup>/ha) et la hauteur dominante (0,32 m), ainsi qu'une variabilité globale accrue. Penner *et al.* (2013) ont conclu que RF constitue une alternative fiable et flexible aux méthodes paramétriques pour les inventaires forestiers opérationnels. Ses performances tirent parti de l'utilisation de la validation croisée, qui améliore la stabilité et la capacité de généralisation des modèles. Toutefois, les auteurs ont également souligné que l'efficacité de toute méthode de modélisation dépend étroitement de la disponibilité de données de calibration bien distribuées et représentatives, condition indispensable pour garantir la précision des estimations dans une diversité de contextes forestiers.

Le choix d'une méthode de modélisation, qu'elle soit paramétrique ou non paramétrique, de même que des techniques spécifiques mises en œuvre, est conditionné par une combinaison de facteurs, incluant la structure forestière, la

configuration des placettes, la qualité des données ALS ainsi que les exigences propres aux utilisateurs finaux (White *et al.*, 2017). L'un des principaux défis dans le développement de modèles consiste à sélectionner les variables prédictives, en raison des corrélations souvent élevées entre les métriques dérivées des données ALS. Dans les modèles linéaires, la multicolinéarité peut entraîner une distorsion des estimations des paramètres et compromettre leur interprétabilité. Afin d'y pallier, des techniques de réduction de dimensionnalité, telles que l'analyse en composantes principales (ACP), sont couramment utilisées (White *et al.*, 2013b).

La validation des modèles constitue également une étape fondamentale pour garantir la fiabilité des prédictions. Lorsque la quantité de données le permet, une partie des placettes d'inventaire est systématiquement réservée à la validation. Dans les cas où les données disponibles sont limitées, des techniques de rééchantillonnage, telles que la validation croisée ou le bootstrapping, sont couramment utilisées pour évaluer la performance des modèles (Efron, 1994). Parmi ces approches, la validation croisée en K segments (K-fold cross-validation) s'impose comme l'une des méthodes les plus couramment utilisées pour estimer l'erreur de prédiction. Elle consiste à partitionner l'ensemble de données en K sous-ensembles, à entraîner le modèle sur certains d'entre eux tout en l'évaluant sur d'autres, puis à calculer la moyenne des erreurs obtenues (Hastie *et al.*, 2009). Plus récemment, la foresterie opérationnelle a commencé à intégrer des données réelles de récolte dans le processus de validation (White *et al.*, 2013b; Leboeuf *et al.*, 2022), ce qui renforce à la fois la précision et la pertinence opérationnelle des résultats de modélisation.

L'exactitude des modèles est généralement évaluée au moyen de plusieurs indicateurs de performance standardisés, parmi lesquels le coefficient de détermination ( $R^2$ ), l'erreur quadratique moyenne (RMSE), sa forme relative (RMSE%), ainsi que l'erreur moyenne (ME) et l'erreur moyenne relative (ME%). Le  $R^2$  exprime la proportion de la variance de la variable dépendante expliquée par le modèle. Le RMSE fournit une estimation de l'écart-type des résidus du modèle et constitue une mesure couramment mobilisée pour apprécier la précision, bien qu'il soit sensible aux valeurs aberrantes (Weiskittel *et al.*, 2011). Le ME et le ME%

permettent quant à eux de caractériser le biais systématique en indiquant si le modèle tend à surestimer ou à sous-estimer les valeurs (White *et al.*, 2017).

### *1.5 Approches multi-temporelles et transférabilité spatiale*

Avec la disponibilité croissante des données ALS, l'intérêt pour l'exploitation d'acquisitions multi-temporelles afin de suivre l'évolution des attributs forestiers et de caractériser les dynamiques de croissance a connu un essor marqué. En Alberta, Tompalski *et al.* (2018) ont démontré le potentiel de la combinaison des données ALS et de la photogrammétrie aérienne numérique (Digital Aerial Photogrammetry, DAP) pour estimer des attributs forestiers à deux moments distincts sur une superficie de 700 000 hectares. Les résultats ont permis de générer des courbes de rendement à l'échelle de la cellule, montrant que la précision des estimations issues de l'ALS constituait le principal levier d'amélioration des modèles de rendement. Cette étude a mis en évidence les avantages d'intégrer les données ALS à des simulateurs de croissance pour soutenir une planification forestière spatialisée et à long terme. Dans une étude ultérieure, Tompalski *et al.* (2019a) ont exploré l'intégration de l'ALS et de la DAP pour le suivi de la croissance dans des peuplements perturbés. L'analyse visait à estimer les variations de la hauteur dominante ( $\Delta H$ ), de la surface terrière par hectare ( $\Delta BA$ ) et du volume total par hectare ( $\Delta V$ ) sur une période de huit ans dans un paysage fortement perturbé. Deux stratégies de modélisation ont été évaluées : la méthode directe, consistant à calculer les différences entre des métriques structurelles communes extraites de jeux de données acquis à deux moments distincts; et la méthode indirecte, qui déduit la croissance à partir de la différence entre les sorties de deux modèles calibrés indépendamment aux temps T1 et T2, mobilisant éventuellement des ensembles de variables explicatives distincts. L'influence de la mortalité au niveau des placettes sur la précision des modèles a également été examinée. Les données de 38 placettes, combinées à des acquisitions ALS de 2006 et 2008 ainsi qu'à des données DAP de 2015, ont servi à l'élaboration de modèles de régression. Les résultats ont montré que la modélisation directe produisait les estimations de croissance les plus précises pour l'ensemble des attributs. La mortalité affectait significativement la précision des prédictions dans les deux approches, mettant en évidence les difficultés associées à l'utilisation de données bi-temporelles

issues de technologies différentes, en particulier lorsque les intervalles d'acquisition sont courts, que le développement des peuplements est lent ou que les taux de mortalité sont élevés. Si des intervalles plus longs peuvent accroître la précision, ils réduisent en contrepartie la fréquence du suivi forestier.

La transférabilité spatiale se définit comme l'application d'un modèle calibré dans une région à une autre région, tout en préservant sa fiabilité prédictive (Thomas *et al.*, 1993; Yates *et al.*, 2018). Cette propriété permet d'étendre l'utilisation de modèles déjà établis à des territoires où les données de terrain sont limitées ou absentes, améliorant ainsi l'applicabilité et l'efficacité opérationnelle des inventaires fondés sur l'ALS. En réduisant la nécessité de réaliser des inventaires de terrain exhaustifs dans chaque région cible, les modèles transférables offrent la possibilité d'une mise en œuvre plus étendue et plus économiquement efficiente de l'ABA (Tompalski *et al.*, 2019b).

Plusieurs études ont examiné la faisabilité du transfert de modèles basés sur l'ALS. Fekety *et al.* (2018) ont évalué la transférabilité spatiale de modèles RF entraînés sur six jeux de données ALS à retour discret et 360 placettes d'inventaire dans la région de l'Inland Northwest en Idaho (États-Unis). Ces modèles, développés pour prédire la surface terrière et la densité de tiges, ont été appliqués à des régions ALS non incluses dans l'entraînement initial. Malgré des différences dans les paramètres d'acquisition et les protocoles d'inventaire, les résultats ont montré des performances cohérentes. Trois des six modèles de surface terrière et deux des six modèles de densité de tiges se sont avérés statistiquement équivalents aux modèles originaux en termes de biais et de structure d'erreur. De plus, la plupart des valeurs de RMSE% restaient comprises dans un écart-type des métriques de validation croisée, ce qui soutient l'idée que la transférabilité spatiale est réalisable lorsque les données d'entraînement représentent de manière adéquate la variabilité structurelle des peuplements forestiers, entendue comme l'organisation verticale et horizontale des arbres, incluant la stratification de la canopée, la densité et la distribution des diamètres (Oliver *et al.*, 1997). Cette étude était toutefois circonscrite à une région présentant une similarité

écologique, c'est-à-dire des essences arborées et des structures forestières comparables (Bailey, 1983).

Dans une autre étude, Tompalski *et al.* (2019b) ont examiné la transférabilité spatiale dans des forêts pluviales tempérées côtières à l'aide de modèles ABA construits à partir de données ALS. Les modèles prédisaient la hauteur de Lorey, le diamètre quadratique moyen et le volume marchand brut en utilisant trois techniques : OLS, RF et k-NN. Ces modèles ont été testés selon trois configurations: (1) leur application à des secteurs voisins présentant des spécifications ALS similaires, (2) l'utilisation de données ALS aux caractéristiques différentes au sein d'une même région, et (3) leur application à des régions entièrement distinctes. Les résultats ont montré que la performance des modèles était davantage influencée par la variable cible et la technique de modélisation que par les différences dans les caractéristiques des nuages de points. Plus précisément, pour la hauteur de Lorey, les modèles OLS et RF ont produit des valeurs de RMSE relatif inférieures à 15% dans l'ensemble des scénarios de transfert, tandis que le modèle k-NN a généré des erreurs plus élevées, avec des RMSE% approchant 23%. À l'inverse, les prédictions du diamètre quadratique moyen et du volume marchand brut se sont révélées moins transférables, les valeurs de RMSE relatif dépassant systématiquement 20%. Parmi les méthodes testées, RF a obtenu les meilleures performances pour le diamètre quadratique moyen et le volume, malgré des variations spatiales et des conditions ALS contrastées. Bien que la complexité structurelle et la dominance du Pruche de l'Ouest (*Tsuga heterophylla*, western hemlock) dans les peuplements étudiés, l'étude a montré que la similarité écologique entre les secteurs de calibration et d'application constitue un facteur déterminant. Les auteurs ont recommandé de compléter les modèles transférés par quelques placettes locales de calibration afin d'améliorer la précision, offrant ainsi une alternative pragmatique à une nouvelle modélisation complète.

Dans l'ensemble, ces études mettent en évidence à la fois le potentiel et les limites de la modélisation multi-temporelle et de la transférabilité spatiale dans le cadre de l'inventaire forestier. Le succès de ces approches dépend de plusieurs facteurs : la

variable d'intérêt, la méthode de modélisation, la similarité écologique et les caractéristiques des données ALS. La combinaison stratégique de modèles existants avec un volume limité de données locales peut accroître l'efficacité du suivi tout en maintenant la précision sur de vastes paysages diversifiés. Toutefois, peu de recherches abordent de manière approfondie la question de la transférabilité des modèles, ce qui souligne la nécessité de tester ce concept à une échelle plus large et sur un plus grand nombre de sites.

### *1.6 Perturbations naturelles et dynamiques forestières*

Les perturbations naturelles constituent des moteurs fondamentaux de la structure et du fonctionnement écologiques des forêts boréales du Canada. Parmi celles-ci, les feux de forêt constituent l'agent le plus fréquent et le plus influent, influençant la composition forestière, les trajectoires de succession et l'hétérogénéité spatiale à grande échelle. Entre 2013 et 2022, la superficie annuelle moyenne brûlée au Canada était d'environ 2,76 millions d'hectares, dont une moyenne de 372 000 hectares par an au Québec (Canadian Council of Forest Ministers, 2023). La foudre est systématiquement reconnue comme la cause prédominante des feux de forêt en termes de superficie brûlée. La saison des feux de 2023 s'est classée parmi les plus sévères jamais enregistrées, sous l'effet de conditions de sécheresse étendues et de vagues de chaleur sans précédent (Barnes *et al.*, 2023).

Au Québec, les régimes de feu sont structurés selon un système de protection zonale comprenant une zone de protection intensive au sud, où les politiques de suppression sont rigoureusement mises en œuvre, et une zone nordique où l'intervention demeure limitée, permettant aux dynamiques naturelles du feu de prévaloir (Forêts, 2025).

Les analyses historiques ont mis en évidence des transformations temporelles majeures dans le régime des feux au Québec. À partir de données issues des domaines bioclimatiques *sapin baumier-bouleau blanc* (forêts mixtes) et *épinette noire-mousses* (forêts résineuses), Bergeron *et al.* (2004) ont rapporté un allongement marqué du cycle des feux au cours des deux derniers siècles. Avant 1850, l'intervalle moyen entre deux feux s'élevait à 83 ans dans les forêts mixtes et à 101 ans dans les

peuplements résineux. Ces intervalles se sont prolongés à 111 et 135 ans, respectivement, durant la seconde moitié du XIX<sup>e</sup> siècle et le début du XX<sup>e</sup> siècle, puis se sont encore accrus depuis 1920 pour atteindre 326 ans dans les forêts mixtes et 398 ans dans les forêts résineuses. Cet allongement progressif traduit les effets conjugués de la variabilité climatique, des politiques de suppression des feux et de la fragmentation accrue du paysage, qui a contribué à limiter la propagation des incendies et à espacer les cycles de feu. En outre, un gradient ouest-est significatif de la fréquence des feux a été observé, caractérisé par l'allongement des cycles de feu et une prévalence accrue des forêts anciennes dans la partie orientale de la province (Bergeron *et al.*, 2012).

Les dynamiques de succession post-feu dans les forêts boréales du Québec présentent une forte variabilité et dépendent du type de forêt, de la sévérité du feu et des conditions édaphiques. Dans les écosystèmes résineux, en particulier dans la région de la Ceinture d'argile dominée par l'épinette noire, les feux de forte intensité peuvent engendrer des trajectoires de succession précoces et divergentes. Les peuplements peuvent être dominés par l'épinette noire ou co-dominés avec le pin gris (*Pinus banksiana* Lamb.), en fonction de l'épaisseur de la couche organique résiduelle et des conditions de station. En l'absence de feux subséquents, une convergence structurale tend généralement à s'opérer en l'espace de deux siècles, reflétant les tendances auto-organisatrices des systèmes forestiers boréaux (Bergeron *et al.*, 1988; Lecomte *et al.*, 2006).

Dans les forêts mixtes, la régénération post-feu se caractérise généralement par l'établissement de cohortes équiennes dominées par des espèces pionnières telles que le peuplier faux-tremble, le bouleau blanc, le pin gris et l'épinette noire (Bergeron *et al.*, 1993). À mesure que le peuplement évolue, ces espèces sont progressivement remplacées par des conifères plus tolérants à l'ombre, notamment le sapin baumier, l'épinette blanche (*Picea glauca*) et le thuya occidental (*Thuja occidentalis*), conduisant à la formation de peuplements structurellement complexes et verticalement stratifiés. Dans les écosystèmes exempts de feu depuis plus de deux siècles, des perturbations secondaires telles que les épidémies d'insectes ou les

dynamiques de trouées contribuent au maintien des conditions de vieilles forêts, dominées principalement par des conifères de fin de succession (Bergeron *et al.*, 1988; Bergeron, 2000; Gauthier *et al.*, 2000).

### 1.7 Objectifs et hypothèses

Cette étude vise à contribuer aux recherches en cours dans le domaine de la télédétection forestière en examinant la performance prédictive et l'applicabilité des modèles basés sur les données ALS dans différents contextes temporels et spatiaux. L'objectif général de ce projet de recherche est de développer et d'évaluer des modèles paramétriques et non paramétriques pour la prédiction d'attributs dendrométriques clés, la hauteur de Lorey ( $h_L$ , m), la surface terrière (BA,  $m^2/ha$ ) et le volume marchand brut (V,  $m^3/ha$ ), à partir de données LiDAR aéroportées multi-temporelles, dans le cadre de l'Approche par zones (ABA) basée sur des rasters. Afin d'opérationnaliser cet objectif général, trois objectifs spécifiques ont été définis, chacun étant associé à une hypothèse de recherche

**Objectif spécifique 1 :** Identifier la combinaison multi-temporelle de données ALS qui maximise la performance prédictive pour l'estimation des attributs dendrométriques.

Cet objectif spécifique est associé à l'hypothèse suivante : des intervalles temporels plus courts entre les acquisitions ALS sont susceptibles de réduire la performance prédictive des modèles, en comparaison avec ceux construits à partir de données couvrant des intervalles plus longs (Hopkinson *et al.*, 2008).

**Objectif spécifique 2 :** Évaluer la transférabilité spatiale des modèles ALS multi-temporelles entre secteurs forestiers distincts.

Cet objectif spécifique est associé à l'hypothèse suivante : les modèles prédictifs de la hauteur de Lorey présentent une transférabilité spatiale plus élevée entre des secteurs forestiers caractérisés par des histoires de perturbations et des conditions écologiques distinctes (secteurs établis après les feux de 1760 et de 1923) (Tompalski *et al.*, 2019b).

**Objectif spécifique 3 :** Comparer les performances prédictives des modèles paramétriques et non paramétriques.

Cet objectif spécifique est associé à l'hypothèse suivante : en raison de leur flexibilité, les modèles non paramétriques devraient présenter une performance prédictive supérieure à celle des modèles paramétriques (Penner *et al.*, 2013).

## 2. ROBUSTESSE DES MODÈLES SPATIAUX ET TEMPORELS DES ATTRIBUTS DENDROMÉTRIQUES FORESTIERS À L'AIDE DU BALAYAGE LASER AÉROPORTÉ MULTI-TEMPOREL

### 2.1 Introduction

Les forêts de la zone boréale sont structurées par une interaction complexe entre les processus écologiques et les régimes de perturbations qui déterminent leur composition, leur structure et leur dynamique à long terme. Au Québec, où les forêts boréales couvrent une grande partie du territoire, la caractérisation de cette variabilité structurelle est indispensable à la mise en œuvre de stratégies de gestion durable. Ces forêts constituent des systèmes dynamiques, fortement influencés par des perturbations naturelles récurrentes, en particulier les feux de forêt et les épidémies d'insectes, qui orientent les trajectoires de succession et contribuent à l'hétérogénéité à l'échelle du paysage (Gauthier *et al.*, 2000; Bergeron *et al.*, 2004; Lecomte *et al.*, 2006).

Au cours des deux dernières décennies, les progrès de la télédétection ont considérablement amélioré les méthodes d'inventaire forestier. Le LiDAR aéroporté (ALS) permet une caractérisation tridimensionnelle détaillée de la structure de la végétation sur de vastes superficies (Nelson *et al.*, 2003; Wulder *et al.*, 2008; Wulder *et al.*, 2012). Dans le cadre de l'inventaire forestier amélioré (IFA), l'Approche par zones (ABA) s'est imposée comme une méthodologie centrale (White *et al.*, 2013a), en reliant des statistiques résumées issues des nuages de points ALS aux mesures de placettes d'inventaire pour estimer des attributs dendrométriques clés tels que la hauteur de Lorey, la surface terrière et le volume marchand brut. Comparée aux méthodes conventionnelles, l'ABA offre une résolution spatiale plus élevée, une efficacité opérationnelle renforcée et la possibilité d'un suivi répété (Næsset, 2002; White *et al.*, 2017).

Des études récentes ont mis en évidence plusieurs défis associés à la mise en œuvre de l'ABA, notamment en ce qui concerne la généralisation et la stabilité des modèles dans le temps et dans l'espace. La précision des prédictions est influencée par divers facteurs, dont la configuration temporelle des acquisitions ALS, la densité des nuages

de points (Gobakken *et al.*, 2008; Socha *et al.*, 2017), la technique de modélisation (Penner *et al.*, 2013), ainsi que les conditions forestières et les spécifications d'acquisition. Par ailleurs, des questions subsistent quant à la mesure dans laquelle des modèles développés dans un contexte spatial ou écologique particulier peuvent être transférés efficacement vers d'autres secteurs, un enjeu majeur pour les applications d'inventaire à grande échelle (Fekety *et al.*, 2018; Tompalski *et al.*, 2019b), particulièrement dans des territoires caractérisés par des histoires de perturbations contrastées et des structures de la végétation hétérogènes.

La présente recherche avait pour objectif d'évaluer la performance prédictive et la stabilité temporelle des modèles ABA dans des forêts boréales appartenant au sous-domaine bioclimatique de la sapinière à bouleau blanc de l'ouest. L'objectif général était de développer et d'évaluer des modèles paramétriques et non paramétriques pour la prédiction d'attributs dendrométriques clés, hauteur de Lorey ( $h_L$ , m), surface terrière (BA,  $m^2/ha$ ) et volume marchand brut ( $V$ ,  $m^3/ha$ ), à partir de données ALS multi-temporelles. Plus précisément, nous avons analysé l'effet de combinaisons temporelles d'acquisitions ALS sur la précision des prédictions en formulant l'hypothèse que des intervalles plus courts entre acquisitions diminueraient la performance relativement à des périodes plus longues (Hopkinson *et al.*, 2008).

L'étude visait également à examiner la transférabilité spatiale des modèles construits à partir de secteurs forestiers distincts, en postulant que la transférabilité des modèles de  $h_L$  serait moins affectée que celle des modèles de BA et de  $V$  par les contrastes liés aux perturbations, notamment entre les secteurs établis après les feux de 1760 et de 1923 (Tompalski *et al.*, 2019b).

Enfin, nous avons comparé les performances de modèles paramétriques et non paramétriques, en postulant que la flexibilité des approches non paramétriques, telles que les forêts aléatoires, leur conférerait une supériorité en termes de précision et de capacité de généralisation (Penner *et al.*, 2013).

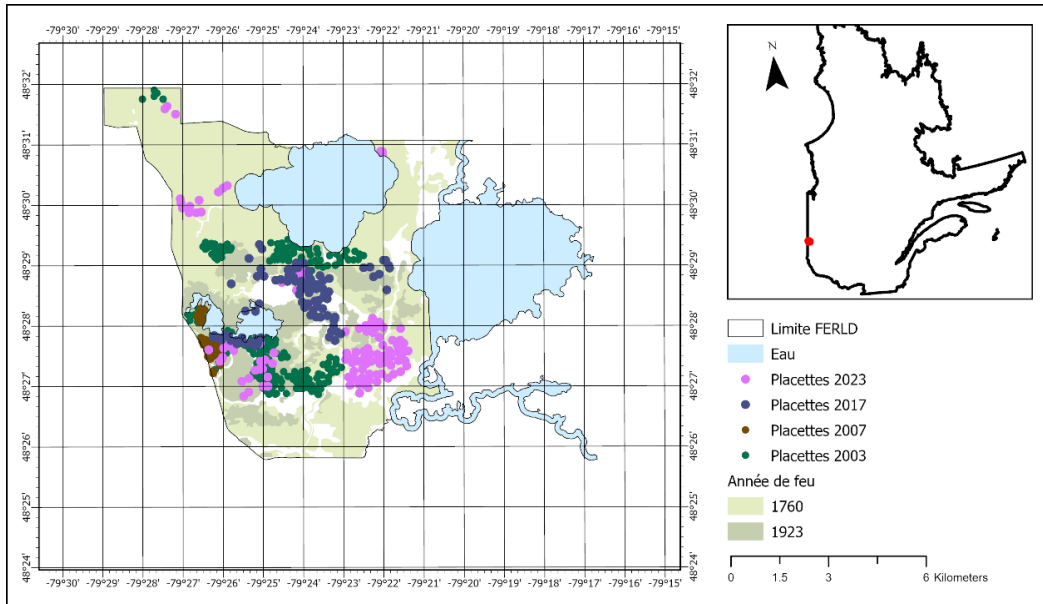
## 2.2 Matériels et méthodes

### 2.2.1 Aire d'étude

L'étude a été réalisée dans la Forêt d'Enseignement et de Recherche du Lac Duparquet (FERLD), localisée entre les latitudes 48°25'80" et 48°32'00" N ainsi que les longitudes 79°17'00" et 79°28'00" O (Figure 1). La FERLD est située dans la zone de végétation boréale, plus précisément dans la sous-zone de la forêt boréale continue, et appartient au sous-domaine bioclimatique de la sapinière à bouleau blanc de l'ouest (Bergeron *et al.*, 1992). Elle se trouve dans la région écologique la plus occidentale de ce sous-domaine, désignée comme région écologique 5a (Grondin, 1999). La dynamique forestière y est principalement gouvernée par le cycle des feux (Bergeron *et al.*, 2006).

La dynamique écologique post-feu dans cette région conduit généralement à la formation de peuplements équiennes, purs ou mixtes, dominés par le peuplier faux-tremble (*Populus tremuloides*), le bouleau blanc (*Betula papyrifera*), le pin gris (*Pinus banksiana* Lamb.) ou l'épinette noire (*Picea mariana* (Mill.) Britton, Sterns & Poggenb.). À mesure que ces peuplements vieillissent, généralement plus de 100 ans après l'incendie, ils sont progressivement remplacés par une seconde cohorte dominée par des conifères tolérants à l'ombre, tels que le sapin baumier (*Abies balsamea* (L.) Mill.), le thuya occidental (*Thuja occidentalis* L.), l'épinette blanche (*Picea glauca*) et l'épinette noire. La structure forestière devient alors plus hétérogène, souvent caractérisée par des peuplements irréguliers et un couvert multi-étagé. Dans les vieilles forêts qui se développent généralement plus de 225 ans après un incendie, des perturbations naturelles telles que les trouées ou les épidémies d'insectes contribuent au maintien de peuplements hétérogènes dominés par les conifères (Bergeron, 1998; Harvey, 1999; Gauthier *et al.*, 2000; Bergeron *et al.*, 2001). Ce schéma de succession, qui correspond aux forêts matures et anciennes, est observé dans l'aire d'étude, en particulier dans les peuplements originaires des importants incendies de 1760 et de 1923, illustrés à la Figure 1 et délimités respectivement par des zones vert clair et vert foncé. Ces deux événements de feu ont conduit à des trajectoires de développement forestier distinctes. Le secteur de 1760 (en vert clair) se caractérise par une dominance de conifères de fin de succession, principalement

le sapin baumier, l'épinette noire et l'épinette blanche. Le secteur de 1923 (en vert foncé) présente, quant à lui, une composition plus mixte, dominée par des espèces pionnières telles que le pin gris, le peuplier faux-tremble et le bouleau blanc (voir ANNEXE A, pour plus de détails).



**Figure 1**  
**Répartition spatiale des 577 placettes d'inventaire dans la FERLD (Forêt d'Enseignement et de Recherche du Lac Duparquet).**

Les secteurs brûlés lors des deux grands incendies historiques sont représentés en vert foncé (1760) et en vert pâle (1923). Les placettes sont associées à quatre acquisitions ALS : 2003 (points verts), 2007 (points bruns), 2017 (points bleu foncé) et 2023 (points violets). Les plans d'eau apparaissent en bleu. La carte en encadré situe l'aire d'étude dans l'ouest du Québec, au Canada.

Cette région est caractérisée par un climat subpolaire, subhumide et continental. La température annuelle moyenne est de 2,5 °C, avec des précipitations annuelles comprises entre 800 et 900 mm, et une saison de croissance d'environ 160 jours (Blouin, 2002). La topographie est majoritairement plane, avec une altitude moyenne de 322 m. La composition géologique de l'aire comprend des roches cristallines d'origine granitique et volcanique, ainsi que des roches sédimentaires. De minces dépôts de till recouvrent les sommets et les versants supérieurs, tandis que des

dépôts marins plus épais, composés d'argile et de limon et parfois intercalés de blocs glaciaires, occupent les versants moyens et inférieurs (Grondin, 1999).

### 2.2.2 Données LiDAR aéroportées

Les données ALS ont été recueillies durant la période foliaire (leaf-on), de juin à septembre, en 2003, 2007, 2017 et 2023. Les acquisitions de 2003, 2017 et 2023 couvraient la totalité de l'aire d'étude, tandis que celle de 2007 présentait une couverture spatiale plus limitée. Les données ALS de 2017 ont été obtenues dans le cadre du programme d'acquisition du gouvernement du Québec, tel qu'archivé sur le site web en libre accès *Forêt Ouverte* (Gouvernement du Québec, 2023), alors que les autres campagnes ont été réalisées conjointement par l'Université du Québec en Abitibi-Témiscamingue et l'Université du Québec à Montréal dans le cadre d'études spécifiques (St-Onge *et al.*, 2004; Véga *et al.*, 2008; Vepakomma *et al.*, 2008). Les densités moyennes de retours de végétation (retours/m<sup>2</sup>) enregistrées pour 2003, 2007, 2017 et 2023 étaient respectivement de 3, 6, 7 et 15. Des modèles de hauteur de canopée (MHC) ont été générés à une résolution de 1 mètre pour chaque acquisition et alignés par la suite à l'aide de dix points de référence dans ArcGIS Pro 3.3.0 (Esri, 2024), en prenant comme référence les données du vol de 2017. Les MHC provenant de chacune des acquisitions ALS (2003, 2007, 2017 et 2023) ont été analysés de manière indépendante afin de caractériser la structure verticale. De plus, un modèle numérique de terrain (MNT), ainsi que des variables topographiques telles que la pente, l'indice topographique d'humidité (TWI) et l'indice topographique de position (TPI), ont été établis à partir des données de 2017 fournies par le Gouvernement du Québec, avec une résolution spatiale uniforme de 1 mètre (Direction des inventaires forestiers, 2021).

### 2.2.3 Placettes d'inventaire

Des placettes d'inventaire géoréférencées pour les inventaires pré-récolte (i.e. forêt mature) étaient disponibles à la FERLD de 2003 à 2022 (Figure 1). Ces données de terrain sont asynchrones (Leboeuf *et al.*, 2022) par rapport aux acquisitions ALS, puisqu'elles n'ont pas été collectées simultanément. Chaque placette couvre une superficie de 400 m<sup>2</sup>. Dans chacune d'elles, le diamètre à hauteur de poitrine (DHP,

en centimètres) ainsi que l'essence de chaque arbre mesuré ont été enregistrés. Les essences commerciales considérées dans cette analyse étaient le sapin baumier, le bouleau blanc, l'épinette blanche, l'épinette noire, le pin gris et le peuplier faux-tremble. Les placettes ont été catégorisées selon le type de couvert en fonction de la surface terrière ( $m^2/ha$ ) prédominante. Plus précisément, elles ont été désignées comme résineuses lorsque les essences résineuses représentaient plus de 75% de la surface terrière; à l'inverse, lorsque les essences résineuses représentaient moins de 25%, elles ont été classées comme feuillues. Les placettes présentant une proportion comprise entre 25 % et 75 % ont été catégorisées comme des forêts mixtes (Direction des inventaires forestiers, 2024b).

Comme des mesures systématiques directes de la hauteur totale des arbres n'ont pas été réalisées, un modèle non linéaire à effets mixtes (Mehtätalo *et al.*, 2015) a été développé afin d'estimer cette variable à partir des mesures de DHP (cm). L'espèce a été intégrée comme facteur aléatoire, et le modèle a été calibré à l'aide de 1 028 arbres provenant de placettes d'inventaire du gouvernement du Québec situées dans un rayon de 10 km autour de la FERLD, ce qui garantit que les prédictions de hauteur soient cohérentes avec les conditions forestières locales.

Les placettes d'inventaire ont été organisées selon le type de couvert et l'année de mesure. Pour le vol de 2003, les placettes mesurées entre 2003 et 2005 ont été incluses (nombre de placettes = 287). Le vol de 2007 a intégré les placettes de 2006 à 2008 (nombre de placettes = 64), tandis que celui de 2017 a utilisé des placettes mesurées entre 2016 et 2018 (nombre de placettes = 68). Enfin, le vol de 2023 a pris en compte les placettes mesurées entre 2021 et 2022 (nombre de placettes = 158). Chaque ensemble de placettes a été apparié à la campagne ALS correspondante afin d'assurer la cohérence temporelle entre les données de terrain et les acquisitions ALS.

Un total de 577 placettes a été analysé, comprenant 44% de forêts mixtes, 37% de forêts résineuses et 19% de forêts feuillues (Figure 1). Pour chaque placette, la hauteur de Lorey ( $h_L$ ), la surface terrière (BA) et le volume marchand brut (V) ont été estimés à partir des données dendrométriques individuelles, incluant l'essence et le

diamètre des arbres. L' $h_L$  a été calculée comme la moyenne des hauteurs pondérées par la surface terrière, tandis que BA a été calculée en  $m^2/ha$  à partir des DHP. Le volume a été estimé à l'aide de modèles définissant le V en fonction du DHP et de la hauteur totale, développés par essence selon la méthodologie établie par Perron (2003). L'ensemble des variables ( $h_L$ , BA et V) a été exprimé à l'échelle de l'hectare.

La composition spécifique des placettes d'inventaire a révélé des différences marquées selon le type de couvert. Dans les couverts résineux, le pin gris dominait la structure en 2003 (plus de 50% de la surface terrière), mais sa proportion a diminué progressivement à 30% en 2017 et 16% en 2023, au profit du sapin baumier et de l'épinette noire. Les couverts feuillus étaient largement composés d'essences pionnières, principalement le peuplier faux-tremble et le bouleau blanc, qui représentaient 70% de la surface terrière. Enfin, les couverts mixtes présentaient une composition plus équilibrée, avec un pourcentage de conifères d'environ 55% en 2003 et 59% en 2023.

Les statistiques descriptives des placettes d'inventaire permettent de caractériser l'évolution structurale des peuplements selon les années d'acquisition ALS : 2003, 2007, 2017 et 2023 (Tableau 1). L' $h_L$  a montré une variabilité significative entre les types de couverts et les années de mesure. Dans les forêts feuillues, elle présentait systématiquement les valeurs moyennes les plus élevées, variant de 20,26 mètres lors du vol de 2003 à 21,59 mètres en 2023. Les forêts mixtes présentaient des valeurs intermédiaires, avec des hauteurs moyennes comprises entre 18,60 et 18,75 mètres, ce qui indique une stabilité structurelle dans le temps. Les couverts résineux, quant à eux, présentaient des hauteurs moyennes de canopée plus faibles, oscillant entre 17,81 et 18,24 mètres, pour atteindre 18,00 mètres lors du vol de 2023.

Des tendances similaires ont été observées pour BA, qui variait également selon le type de couvert et l'année d'inventaire. Le couvert feuillu présentait les valeurs moyennes de BA les plus élevées, bien qu'une légère diminution ait été constatée au fil du temps, passant de 28,2  $m^2/ha$  en 2003 à 25,3  $m^2/ha$  lors du vol de 2023. Les types de couverts résineux et mixtes montraient des valeurs de BA plus faibles,

respectivement de 22,9 m<sup>2</sup>/ha et 25,1 m<sup>2</sup>/ha, avec des variations minimales entre les années de mesure.

**Tableau 1**  
**Résumé des attributs structuraux des placettes d'inventaire pour chaque année d'acquisition ALS.**

Les valeurs correspondent à l'intervalle (minimum-maximum) et à la moyenne de la hauteur de Lorey ( $h_L$ ), de la surface terrière (BA) et du volume marchand brut (V), regroupées par types de couvert forestier (résineux, feuillus et mixtes).

Variables	Résineux			Feuillus			Mixtes		
	Min	Max	Moy.	Min	Max	Moy.	Min	Max	Moy
<b>2003</b>									
<b><math>h_L</math> (m)</b>	11,68	23,06	18,11	15,44	24,05	20,26	13,06	23,99	18,60
<b>BA(m<sup>2</sup>/ha)</b>	6,6	51,7	27,2	6,3	52,3	28,2	6,9	48,2	25,1
<b>V (m<sup>3</sup>/ha)</b>	56,3	380,8	208,3	54,6	445,6	243,2	54,2	415,9	196,7
<b>2007</b>									
<b><math>h_L</math> (m)</b>	13,82	21,29	18,24	16,21	22,98	20,75	16,27	23,45	19,45
<b>BA(m<sup>2</sup>/ha)</b>	9,8	44,6	25,5	10,5	46,0	30,2	8,2	42,7	25,3
<b>V (m<sup>3</sup>/ha)</b>	78,1	332,1	190,3	79,1	448,4	275,1	69,2	348,1	209,1
<b>2017</b>									
<b><math>h_L</math> (m)</b>	12,83	20,71	17,81	17,96	23,70	20,67	14,83	22,42	18,81
<b>BA(m<sup>2</sup>/ha)</b>	9,7	43,0	23,6	6,4	51,4	28,6	12,3	51,1	27,5
<b>V (m<sup>3</sup>/ha)</b>	64,0	351,5	176,1	60,7	449,2	251,0	91,3	437,6	220,2
<b>2023</b>									
<b><math>h_L</math> (m)</b>	11,94	22,60	18,00	17,29	24,85	21,59	13,28	22,16	18,75
<b>BA(m<sup>2</sup>/ha)</b>	7,7	40,8	22,9	7,8	44,5	25,3	7,6	45,8	25,1
<b>V (m<sup>3</sup>/ha)</b>	55,0	310,2	199,2	57,4	420,2	243,3	61,4	436,2	199,2

#### 2.2.4 Analyse temporelle et spatiale

En suivant l'approche proposée par Mauro *et al.* (2021), l'ABA a été utilisée pour traiter les données dérivées de l'ALS sous forme de couches rastérisées, à partir desquelles ont été dérivées les métriques structurales à l'échelle des placettes d'inventaire. À partir de la géoréférenciation du centre de chaque placette d'inventaire, des unités carrées de 400 m<sup>2</sup> ont été générées à l'aide d'ArcGIS Pro 3.3.0. Celles-ci ont ensuite

été intersectées avec les rasters disponibles à une résolution de 1 mètre : modèle de hauteur de canopée (MHC), modèle numérique de terrain (MNT), pente, altitude, indice topographique de position (TPI) et indice topographique d'humidité (TWI). Le TPI quantifie la position topographique relative d'un point donné en comparant son altitude à l'altitude moyenne d'un voisinage défini, ce qui permet d'identifier des entités géomorphologiques telles que les crêtes, les dépressions et les terrains plans (Weiss, 2001). Le TWI est un indicateur dérivé du terrain qui combine la superficie contributive en amont et la pente locale afin d'estimer la manière dont l'eau s'accumule à l'échelle du paysage, permettant ainsi d'identifier les secteurs présentant un potentiel d'humidité du sol plus élevé (Sørensen *et al.*, 2006). Afin de faciliter la gestion et l'exploitation des données, les valeurs extraites pour chaque placette ont été systématiquement intégrées dans une base de données PostgreSQL (Group, 2024), garantissant la conservation de l'identifiant original associé à chaque placette.

Afin d'obtenir une caractérisation précise de la structure forestière et des conditions stationnelles, un ensemble complet de métriques (White *et al.*, 2013a) a été calculé pour chaque placette. Les métriques de hauteur de canopée comprenaient la moyenne, l'intervalle interquartile, l'écart-type, la variance, l'asymétrie (skewness) et l'aplatissement (kurtosis). De plus, les percentiles 5e, 25e, 50e, 75e, 90e, 95e et 99e ont été calculés afin de représenter la distribution verticale de la végétation. Des indices structurels, incluant l'indice de complexité verticale (VCI), le rapport de relief de la canopée et le couvert de canopée aux niveaux de hauteur moyenne et modale, ont également été intégrés à l'analyse. En outre, la proportion d'éléments de la canopée appartenant à des strates de hauteur définies (par ex. <0,15 m, 0,15-1,37 m, 1,37-5 m, 5-10 m, 10-20 m et >20 m) a été calculée afin de caractériser la stratification verticale des peuplements forestiers (Tableau 2) (Variables explicatives détaillées dans l'ANNEXE B).

Les variables topographiques ont été dérivées du MNT et incluaient la pente, le TPI et le TWI. Pour chacune de ces variables, les percentiles du 5e au 99e ont été calculés, ainsi que des indicateurs de tendance centrale et de variabilité, comprenant l'intervalle interquartile, l'écart-type, la variance, l'asymétrie et l'aplatissement. En

outre, les placettes ont été enrichies par des attributs écologiques issus de la carte écoforestière. Cette carte repose sur l'inventaire forestier réalisé tous les dix ans à partir de photographies aériennes et de points de contrôle terrain. Elle rend compte des caractéristiques forestières et écologiques du sud du Québec, en reflétant les conditions consignées à chaque cycle d'inventaire (Direction des inventaires forestiers, 2024a). Dans le cadre de ce projet, les placettes ont été intégrées à la carte écoforestière (Direction des inventaires forestiers, 2024b) afin de les caractériser selon les dépôts de surface, apportant ainsi un contexte supplémentaire aux variations observées dans la structure et la composition forestières.

Pour l'analyse temporelle, l'ensemble de données a été regroupé en six combinaisons de deux périodes d'acquisition ALS, quatre combinaisons de trois périodes et une combinaison couvrant les quatre périodes. Cette approche connue sous le terme de *pooling*, dans Fekety *et al.* (2015), est conceptuellement cohérente avec la stratégie de regroupement proposée par ces auteurs, qui vise à améliorer la calibration des modèles en intégrant la variabilité structurelle provenant de plusieurs périodes d'acquisition. Chaque combinaison temporelle a ensuite été stratifiée selon le type de couvert forestier afin de mieux saisir les différences structurelles au fil du temps et entre types de végétation.

Pour l'analyse spatiale, la combinaison des quatre périodes d'acquisition ALS (2003-2007-2017-2023) a été utilisée, de manière similaire à Fekety *et al.* (2018). Les données ont été stratifiées à la fois par type de couvert forestier et par année de feu. Cette analyse a été réalisée exclusivement pour les couverts résineux et mixtes, en raison de la disponibilité limitée de placettes de terrain dans le couvert feuillu. Les modèles ont été calibrés à partir des placettes situées dans des secteurs affectés par les incendies historiques de 1760 ou de 1923 (Figure 1). Chaque modèle a été validé en utilisant 30% des placettes du même secteur d'origine de feu, puis appliqué à l'autre secteur afin d'évaluer la transférabilité spatiale.

Afin d'évaluer la performance des différentes méthodes de modélisation, la Forêt aléatoire (RF) et la Régression linéaire multiple (LR) ont été utilisées pour construire

des modèles prédictifs. La RF, méthode non paramétrique par ensembles, a permis de capter les interactions complexes et les relations non linéaires, et s'appuie sur l'analyse de l'importance des variables par permutation pour la sélection des variables explicatives.

## Tableau 2

### Description des variables explicatives utilisées pour la calibration des modèles et de leurs sources de données correspondantes.

Les variables dérivées du Modèle de hauteur de canopée (MHC) incluent les percentiles de hauteur, les métriques de structure verticale et les indicateurs de couvert de canopée. Les variables liées au terrain ont été extraites du Modèle numérique de terrain (MNT), de la pente, de l'Indice topographique de position (TPI) et de l'Indice topographique d'humidité (TWI), et comprennent un ensemble de statistiques descriptives telles que les percentiles, l'intervalle interquartile, l'écart-type, la variance, l'asymétrie, l'aplatissement et le coefficient de variation.

<b>Description des variables explicatives</b>
<p><b>Variabes dérivées du MHC:</b>  Hauteur moyenne des cellules &gt; 2 m  5e, 25e, 50e, 75e, 90e, 95e et 99e percentiles des hauteurs de cellules &gt; 2 m  Intervalle interquartile des hauteurs de cellules &gt; 2 m  Écart type des hauteurs de cellules &gt; 2 m  Variance des hauteurs de cellules &gt; 2 m  Coefficient d'asymétrie des hauteurs de cellules &gt; 2 m  Coefficient d'aplatissement des hauteurs de cellules &gt; 2 m  Coefficient de variation des hauteurs de cellules &gt; 2 m  Pourcentage de hauteurs de cellules &lt; 0,15 m  Pourcentage de hauteurs de cellules &gt; 0,15 m et &lt; 1,37 m  Pourcentage de hauteurs de cellules &gt; 1,37 m et &lt; 5 m  Pourcentage de hauteurs de cellules &gt; 5 m et &lt; 10 m  Pourcentage de hauteurs de cellules &gt; 10 m et &lt; 20 m  Pourcentage de hauteurs de cellules &gt; 20 m  Indice de complexité verticale (VCI)  Rapport de relief de la canopée  Pourcentage de couvert à la hauteur moyenne de canopée  Pourcentage de couvert à la hauteur modale de canopée</p>
<p><b>Variabes dérivées du MNT, Pente, TPI et TWI:</b>  5e, 25e, 50e, 75e, 90e, 95e et 99e percentiles  Intervalle interquartile  Écart type  Variance  Coefficient d'asymétrie  Coefficient d'aplatissement  Coefficient de variation</p>

Les hyperparamètres ont été optimisés par validation croisée k-fold en minimisant le RMSE. Pour LR, la sélection des variables a été réalisée en deux étapes: (i) contrôle de la colinéarité à l'aide du VIF, avec retrait itératif des variables présentant un VIF > 5; (ii) réduction et sélection des prédicteurs numériques par ACP (analyse en composantes principales), en retenant des composantes expliquant jusqu'à 80 % de la variance cumulée. Pour LR, la performance prédictive a également été évaluée par validation croisée k-fold.

## 2.3 Résultats

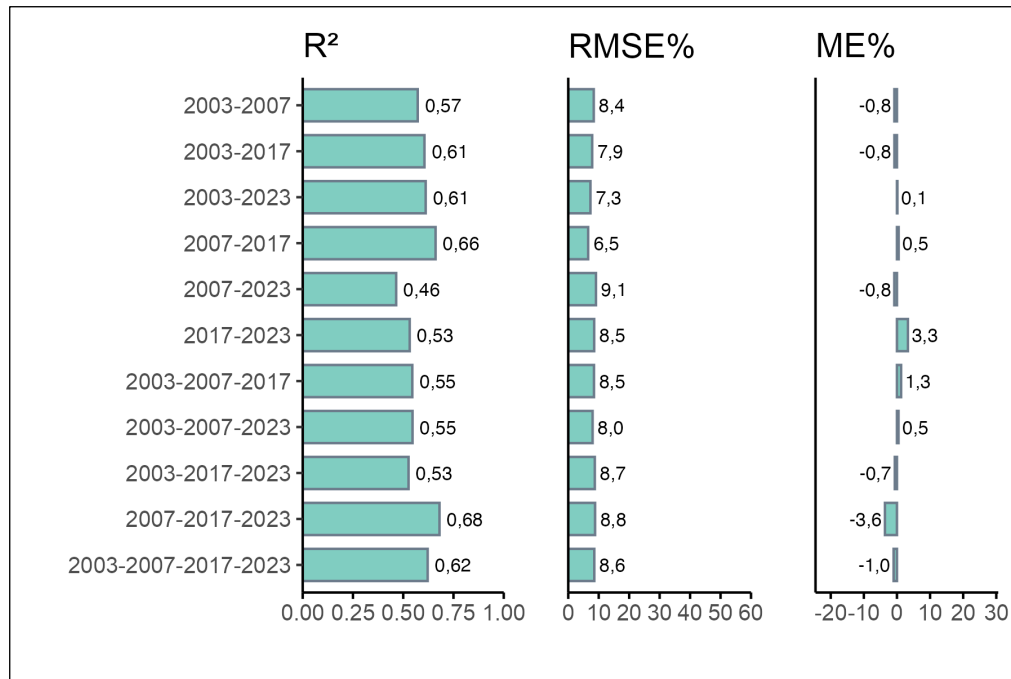
### 2.3.1 Modélisation multi-temporelle

#### 2.3.1.1 Hauteur de Lorey

Dans le couvert résineux, les modèles RF ont démontré une performance prédictive robuste. La valeur de  $R^2$  la plus élevée, soit 0,68, a été observée pour l'agrégation des données des années 2007, 2017 et 2023, comme l'illustre la Figure 2. Cependant, cette combinaison présentait également une erreur moyenne relative (ME%) de -3.6. Les modèles calibrés à partir des acquisitions de 2003-2017 et de 2003-2023 ont également montré de bonnes performances, ces deux modèles présentant des  $R^2$  supérieurs à 0,60 et affichant un biais plus faible.

Par contraste, les modèles LR ont généralement présenté une précision prédictive inférieure. Néanmoins, le modèle LR fondé sur la combinaison 2007-2017 a obtenu des résultats compétitifs, avec un  $R^2$  de 0,63 et un ME% faible (voir ANNEXE C). Ces résultats mettent en évidence la robustesse des modèles RF, tout en révélant que, dans certaines circonstances, la LR peut fournir des résultats comparables.

Dans l'estimation de  $h_L$  pour le couvert feuillu, les modèles RF ont obtenu de meilleures performances que les modèles LR. Le modèle RF fondé sur la combinaison 2007-2017 a atteint la performance la plus élevée, avec un  $R^2$  de 0,83 et un RMSE% remarquablement faible, bien qu'accompagné d'un ME% élevé, montrant un biais important. Le modèle intégrant les quatre périodes d'acquisition (2003-2007-2017-2023) a également montré une performance robuste, avec des valeurs de  $R^2$  supérieures à 0,51 et des RMSE% et ME% considérablement réduits (Figure 3).



**Figure 2**

**Performance des modèles Random Forest (RF) pour la prédiction de la hauteur de Lorey ( $h_L$ ) dans le couvert forestier résineux.**

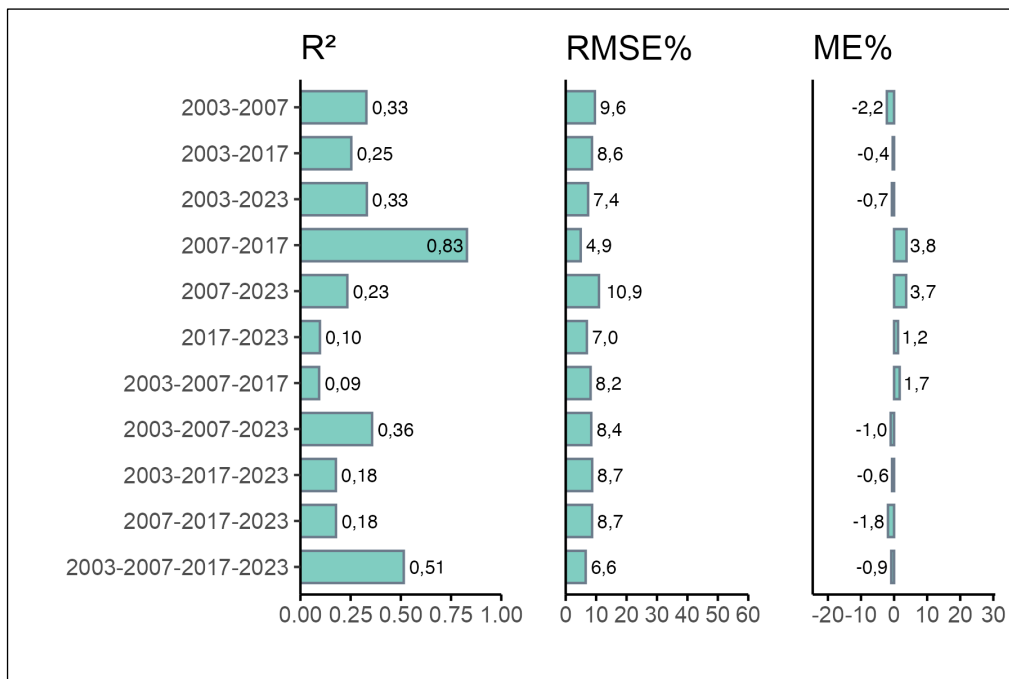
Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

En ce qui concerne les modèles LR, ils ont démontré une capacité prédictive limitée, avec des valeurs de  $R^2$  généralement inférieures à 0,30. Plusieurs combinaisons, en particulier celles intégrant les données de 2003 et 2017, n'ont pas permis de produire de modèles LR valides, probablement en raison d'une taille d'échantillon insuffisante ou d'une variabilité structurelle trop élevée (voir ANNEXE C).

Dans le couvert mixte, les modèles RF ont constamment surpassé les modèles LR dans l'estimation de  $h_L$ , et ce pour l'ensemble des combinaisons d'acquisitions ALS. Le modèle RF fondé sur l'ensemble de données 2007-2017 a obtenu la performance la plus élevée, avec un  $R^2$  de 0,62 et un RMSE% de 6,6; tout en maintenant un biais modéré (ME% = -1,2). D'autres modèles RF, incluant les combinaisons 2003-2007-2017 et 2007-2017-2023, ont également présenté une capacité prédictive élevée, se

traduisant par des valeurs de  $R^2$  supérieures à 0,55 et par des erreurs relativement faibles (Figure 4).

Par contraste, les modèles LR ont affiché des performances prédictives plus limitées (voir ANNEXE C). Le meilleur résultat a été observé avec la combinaison 2003-2017 ( $R^2 = 0,46$ ), tandis que la performance a décroît avec la combinaison 2003-2023 ( $R^2 = 0,20$ ), malgré une augmentation de la taille de l'échantillon. Une tendance analogue a été constatée pour les modèles RF, la combinaison 2003-2023 produisant des valeurs de  $R^2$  plus faibles et un RMSE% légèrement accru par rapport à la période 2003-2017.

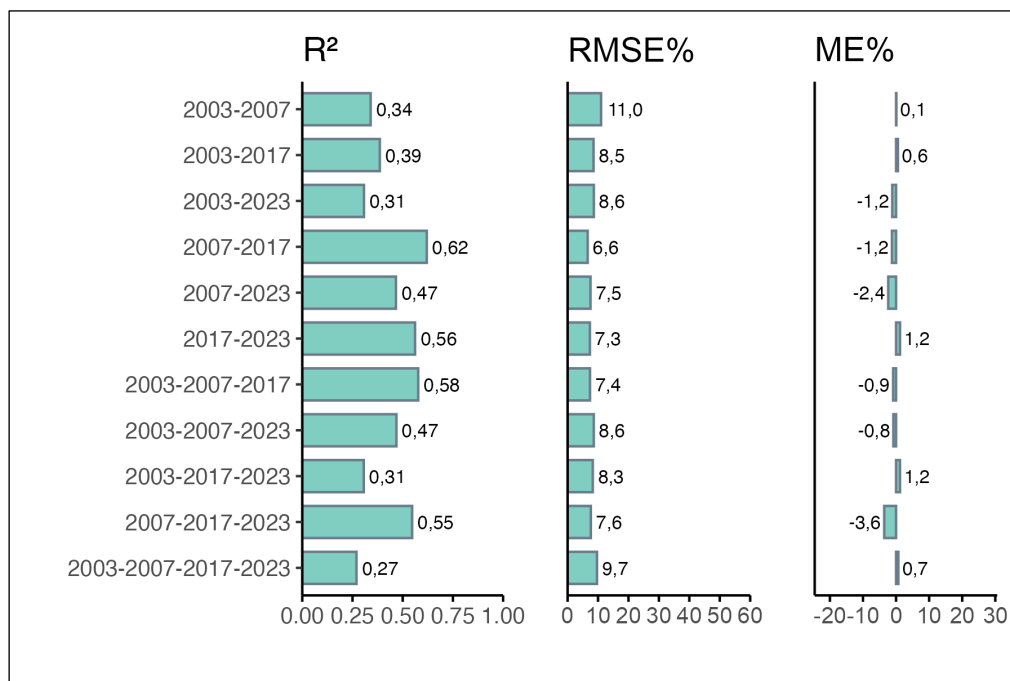


**Figure 3**  
**Performance des modèles Random Forest (RF) pour la prédiction de la hauteur de Lorey ( $h_L$ ) dans le couvert forestier feuillu.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

Ce déclin constant, observé dans les deux approches de modélisation par le couvert mixte, indique que des intervalles temporels prolongés entre acquisitions ALS peuvent

intensifier la variabilité structurelle ou réduire la cohérence temporelle, ce qui limite en conséquence la précision des modèles.



**Figure 4**  
**Performance des modèles Random Forest (RF) pour la prédiction de la hauteur de Lorey ( $h_L$ ) dans le couvert forestier mixte.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

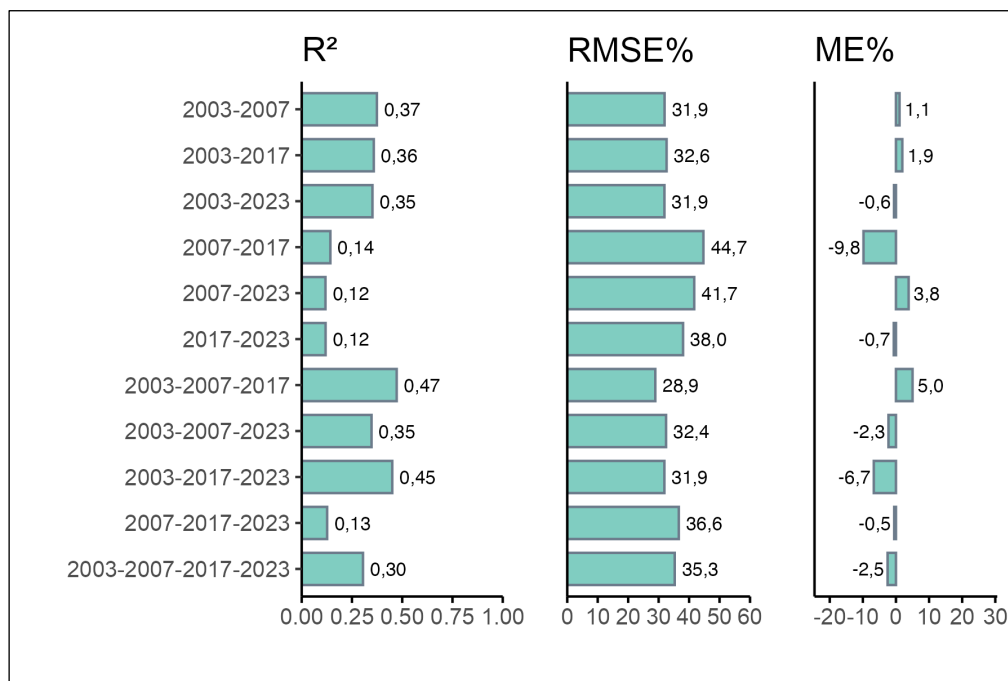
L'importance des variables du RF pour la prédiction de  $h_L$  (voir ANNEXE D), indique que les variables les plus influentes sont principalement des métriques de hauteur dérivées du MHC, avec une contribution récurrente des 99e et 95e percentiles (p99\_mhc, p95\_mhc) parmi les plus importantes. Pour le couvert résineux, p99\_mhc apparaît régulièrement comme la variable la plus importante, tandis que p95\_mhc, p90\_mhc et des indicateurs de dispersion (var\_mhc, sd\_mhc) présentent une importance secondaire. Pour le couvert feuillu, les percentiles élevés restent prédominants, mais certaines combinaisons temporelles mettent davantage en évidence des variables associées à l'occupation verticale du couvert, notamment les pourcentages de cellules MHC appartenant aux classes de hauteur de 10–20 m et >

20 m (perc\_10\_20, perc\_up\_20) ainsi que, de manière ponctuelle, des variables topographiques décrivant la position relative du terrain (moyenne de l'indice de position topographique (mean\_tpi)). Enfin, dans le couvert mixte, la même tendance se maintient : l'importance relative des percentiles varie davantage selon les combinaisons d'années, et les variables d'occupation verticale peuvent occuper une place prépondérante dans certaines configurations. Dans l'ensemble, la prédiction pour  $h_L$  repose prioritairement sur des descripteurs de hauteur, complétés par des indicateurs de structure verticale dont l'importance varie selon le couvert et la combinaison temporelle.

#### 2.3.1.2 Surface terrière

Dans le couvert résineux, tant les modèles RF que les modèles LR ont démontré une performance prédictive limitée pour l'estimation de BA, avec des valeurs de  $R^2$  systématiquement inférieures à 0,50 pour toutes les combinaisons d'acquisitions. Parmi les modèles RF, les combinaisons 2003-2007-2017 et 2003-2017-2023 se sont avérées relativement supérieures, atteignant respectivement des  $R^2$  de 0,47 et 0,45, tout en présentant les RMSE% les plus faibles observés dans l'ensemble du jeu de données. Néanmoins, ces modèles ont également montré un biais significatif, comme en témoigne des valeurs de ME% élevées, respectivement de 5,0 et -6,7 (Figure 5).

Dans le cas des modèles LR, les valeurs de  $R^2$  les plus élevées ont été obtenues avec les combinaisons 2003-2017 ( $R^2 = 0,46$ ) et 2003-2007-2017 ( $R^2 = 0,42$ ). Entre les deux, le modèle 2003-2017 présentait un ME% nettement plus faible (-0,2) comparé à 3,7 pour la combinaison à trois périodes, ce qui indique un biais réduit malgré des valeurs de  $R^2$  similaires (voir ANNEXE E).

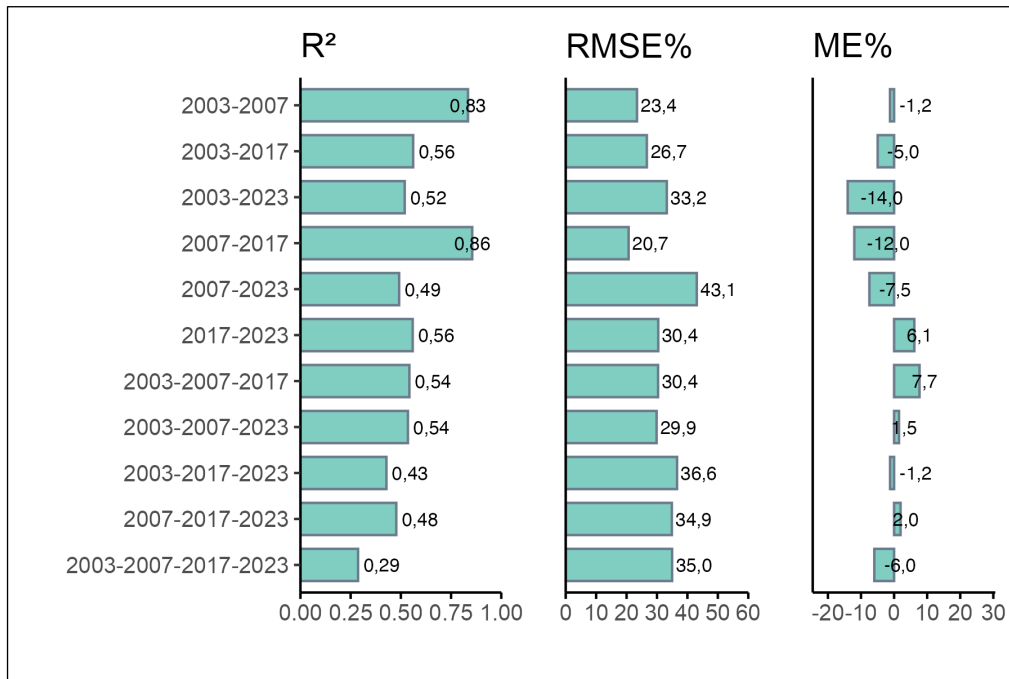


**Figure 5**

**Performance des modèles Random Forest (RF) pour la prédiction de la surface terrière (BA) dans le couvert forestier résineux.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

Dans le couvert feuillu, tant les modèles RF que les modèles LR ont obtenu des valeurs de  $R^2$  supérieures à celles observées pour les couverts résineux. Parmi les modèles RF, les combinaisons 2003-2007 et 2007-2017 ont produit les meilleurs résultats, avec des  $R^2$  de 0,83 et 0,86 respectivement, et des RMSE% relativement faibles. Toutefois, le modèle 2007-2017 présentait un biais important, avec un ME% de -12,0, ce qui limitait sa fiabilité malgré son pouvoir explicatif élevé. De même, la combinaison 2003-2023 a montré une performance modérée en termes de  $R^2$  (0,52), mais également un biais négatif marqué (ME% = -14,0). Parmi les modèles à trois périodes, 2003-2007-2017 et 2003-2007-2023 ont produit des  $R^2 = 0,54$ , le second présentant le compromis le plus équilibré entre performance prédictive et biais d'estimation (Figure 6).



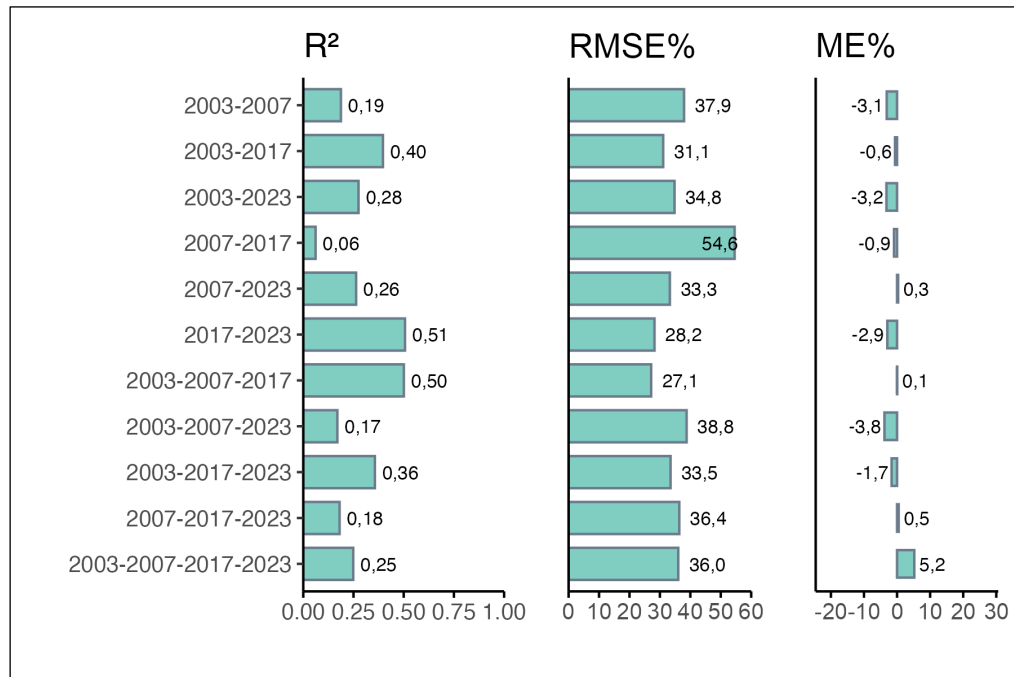
**Figure 6**  
**Performance des modèles Random Forest (RF) pour la prédiction de la surface terrière (BA) dans le couvert forestier feuillu.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

Par contraste, les modèles LR ont généralement montré des valeurs de  $R^2$  inférieures à celles des RF, bien que certaines combinaisons aient démontré une performance relativement robuste (voir ANNEXE E). Le modèle 2003-2007 a atteint le  $R^2$  le plus élevé (0,68) avec un ME% faible (-0,5), suggérant une calibration efficace et peu biaisée. Les combinaisons à trois périodes telles que 2003-2007-2017 et 2003-2007-2023 ont également obtenu des résultats satisfaisants, avec des  $R^2$  de 0,58 et 0,51 respectivement. Le modèle 2003-2007-2017, en particulier, présentait un ME% = 0,4. Il convient de noter que le modèle LR intégrant les quatre périodes d'acquisition (2003-2007-2017-2023) n'a pas convergé.

Dans le couvert mixte, tant les modèles RF que les modèles LR ont montré des performances prédictives relativement faibles pour l'estimation de BA, avec des

valeurs de  $R^2$  systématiquement inférieures à 0,5 pour la quasi-totalité des combinaisons d'acquisitions (Figure 7).



**Figure 7**  
**Performance des modèles Random Forest (RF) pour la prédiction de la surface terrière (BA) dans le couvert forestier mixte.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

Le meilleur résultat pour RF a été obtenu avec la combinaison 2003-2007-2017 ( $R^2 = 0,50$ ), qui a également présenté le RMSE% le plus faible (27,1) et un biais minimal (ME% = 0,1), indiquant un compromis équilibré entre précision et stabilité. Le modèle 2007-2017 a présenté les performances les plus faibles, avec un coefficient de détermination proche de zéro ( $R^2 = 0,06$ ) et une RMSE% = 54,6, comme l'illustre la Figure 7.

De même, les modèles LR ont également obtenu leurs meilleures performances avec la combinaison 2003-2007-2017 ( $R^2 = 0,50$ ), caractérisée par un RMSE% minimal (31,2) et un biais pratiquement nul (ME% = 0,9), ce qui indique une capacité prédictive

robuste dans ce contexte (voir ANNEXE E). D'autres combinaisons, incluant 2003-2017 et 2007-2023, ont produit des  $R^2$  modérés (0,39 et 0,35) et des biais comparables.

De manière générale, pour la prédiction du BA, les modèles LR ont présenté un biais plus faible que les modèles RF, mais se sont révélés moins flexibles pour modéliser la variabilité structurelle.

L'analyse de l'importance des variables du RF pour la prédiction de BA (voir ANNEXE F) montre que les variables les plus influentes combinent des métriques ALS décrivant la structure du couvert et sa distribution verticale, avec une hiérarchie plus variable que celle de  $h_L$ . Pour le couvert résineux, le pourcentage de couvert à la hauteur moyenne de canopée ( $ccr\_mhc$ ) figure fréquemment parmi les variables dominantes, complété par des métriques de hauteur (p. ex.  $p50\_mhc$ ,  $p90-p95\_mhc$ ) et, selon les combinaisons, par des variables de pente ( $max\_slope$ ,  $p95\_slope$ ,  $p99\_slope$ ). Pour le couvert feuillu, l'importance met davantage en évidence des descripteurs d'hétérogénéité ( $mhr$ ,  $rch$ ,  $cv\_mhc$ ) et des métriques de hauteur basses à intermédiaires ( $p05\_mhc$ ,  $p25-p50\_mhc$ ,  $mean\_mhc$ ), ainsi que des variables d'occupation verticale ( $perc\_up\_20$  : pourcentage de cellules > 20 m ;  $perc\_10\_20$  : pourcentage de cellules de 10–20 m ;  $perc\_5\_10$  : pourcentage de cellules de 5–10 m ;  $perc\_137\_5$  : pourcentage de cellules de 1,37–5 m). Enfin, pour le couvert mixte, l'importance est principalement portée par des métriques centrales ( $mean\_mhc$ ,  $p50\_mhc$ ,  $p75\_mhc$ ,  $p25\_mhc$ ), avec une contribution plus ponctuelle de  $perc\_up\_20$  (pourcentage de cellules > 20 m) et d'indices structuraux ( $crr\_mhc$ ,  $ccr\_mhc$ ). Dans l'ensemble, la prédiction de BA repose sur une combinaison de descripteurs de hauteur et de structure du couvert, dont l'importance varie selon le couvert et la combinaison temporelle.

### 2.3.1.3 Volume marchand brut

Dans le couvert résineux, les modèles RF ont démontré une performance prédictive modérée pour l'estimation du V, comme l'illustre la Figure 8. La combinaison 2003-2017 a présenté le  $R^2$  le plus élevé (0,68), le RMSE% le plus faible (25,6) et un biais

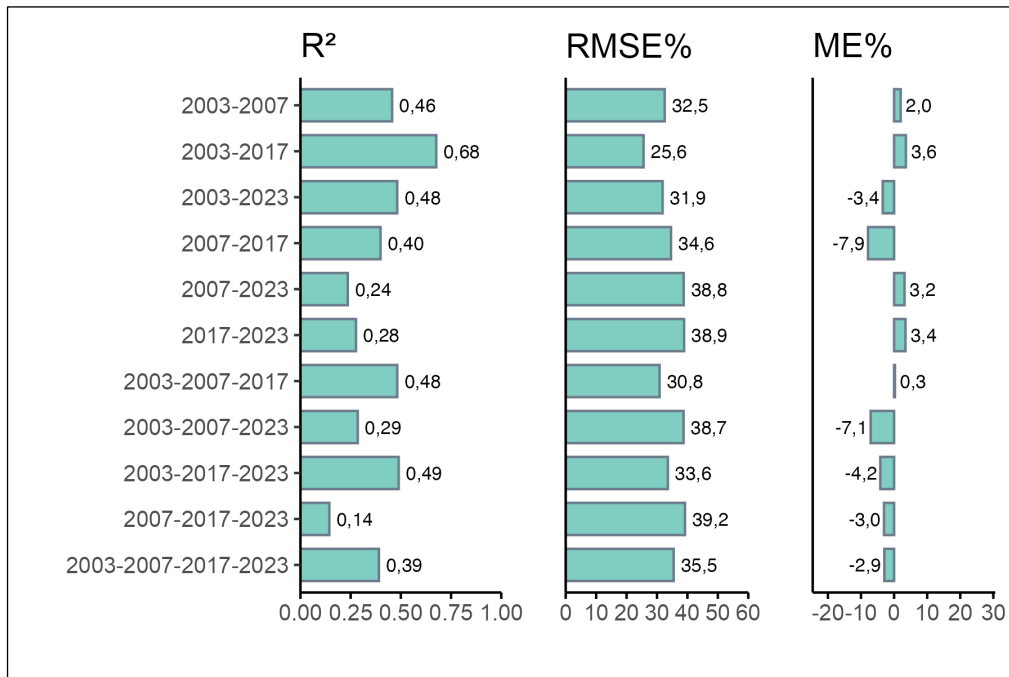
modéré ( $ME\% = 3,6$ ), représentant le résultat le plus favorable parmi l'ensemble des configurations RF. En outre, les combinaisons 2003-2023 et 2003-2007 ont présenté des performances solides ( $R^2 = 0,48$  et  $0,46$  respectivement), avec des  $RMSE\%$  de 31,9 et 32,5 et des biais compris entre -3,4 et 2,0. Parmi les modèles à trois périodes, la configuration 2003-2007-2017 a également présenté un  $R^2$  de 0,48, un  $RMSE\%$  de 30,8 et un biais minimal ( $ME\% = 0,3$ ).

Les modèles LR ont atteint des niveaux de  $R^2$  globalement comparables à ceux obtenus avec les RF dans plusieurs configurations. Par exemple, le modèle 2003-2017 ( $R^2 = 0,66$ ) s'est distingué par le  $RMSE\%$  le plus faible (25,9), bien qu'il ait présenté un biais supérieur à celui du RF ( $ME\% = 5,1$ ). Parmi les combinaisons à trois années, la configuration 2003-2007-2017 a représenté la meilleure performance pour la LR ( $R^2 = 0,50$ ), avec un  $RMSE\%$  de 30,2 et un biais quasi nul ( $ME\% = 0,3$ ), traduisant une forte convergence avec les résultats RF. Les modèles LR construits à partir de combinaisons temporelles plus longues, telles que 2003-2007-2017-2023, ont toutefois montré une performance diminuée ( $R^2 = 0,38$  ;  $RMSE\% = 35,2$  ;  $ME\% = -3,5$ ), tandis que le modèle 2007-2023 s'est avéré particulièrement limité ( $R^2 = 0,04$  ;  $RMSE\% = 43,3$  ;  $ME\% = 7,4$ ) (voir ANNEXE G).

Les résultats montrent que les modèles RF tendent à surpasser légèrement les modèles LR en termes de précision prédictive. La combinaison 2003-2017 se distingue comme la plus performante dans les deux approches de modélisation, offrant un compromis particulièrement robuste entre le pouvoir explicatif, le niveau d'erreur et le biais d'estimation.

Dans le couvert feuillu, les modèles RF ont produit une performance prédictive modérée pour l'estimation du V, comme l'illustre la Figure 9. Parmi les combinaisons à deux périodes, les modèles basés sur 2003-2023 et 2007-2017 ont présenté des  $R^2$  de 0,50 et 0,55 respectivement. Toutefois, les deux étaient associés à un biais élevé, tel qu'indiqué par des valeurs de  $ME\%$  de -9,9 et -5,8. La configuration RF la plus équilibrée a été observée pour le modèle 2003-2007-2023, qui a présenté un  $R^2$  de

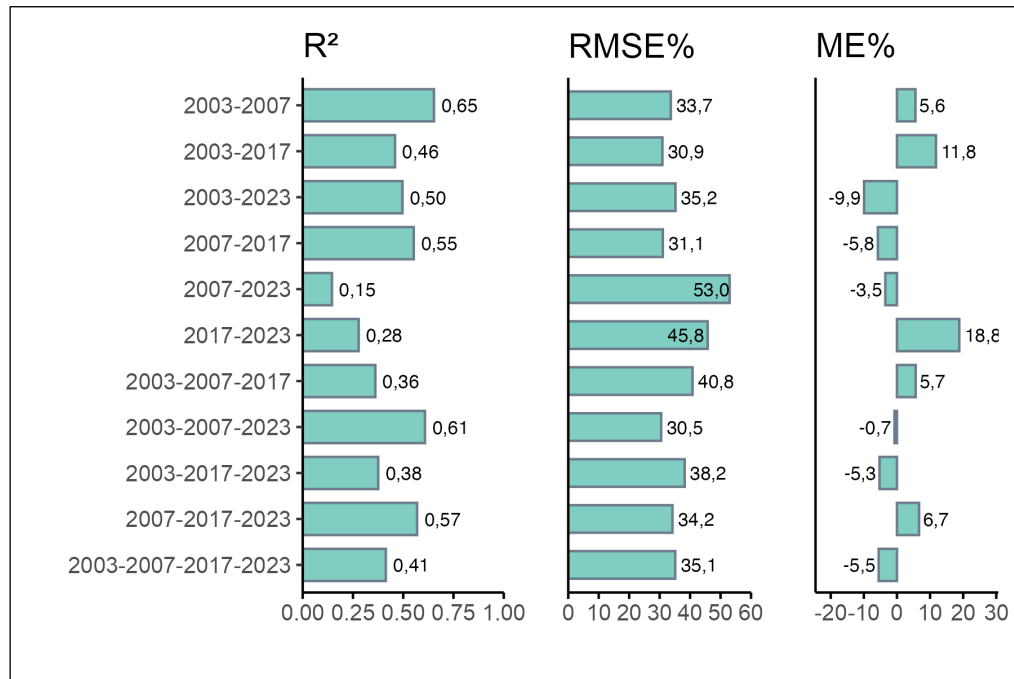
0,61, un RMSE% de 30,5 et un biais minimal ME% de -0,7 représentant le meilleur compromis entre performance prédictive et biais d'estimation dans ce type de couvert.



**Figure 8**  
**Performance des modèles Random Forest (RF) pour la prédiction du volume marchand brut (V) dans le couvert forestier résineux.**

Indicateurs de performance : coefficient de détermination (R<sup>2</sup>), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

Par contraste, les modèles LR ont atteint des valeurs de R<sup>2</sup> comparables. La combinaison 2003-2007-2023 a généré le R<sup>2</sup> = 0,58 et un RMSE% de 31,4 et un ME% de -3,5. De même, le modèle 2003-2007 a bien performé (R<sup>2</sup> = 0,61 ; RMSE% = 31,1 ; ME% = 3,5), confirmant sa fiabilité à la fois en termes d'ajustement du modèle et d'ampleur de l'erreur. D'autres configurations, telles que 2003-2017 et 2017-2023, ont été associées à des R<sup>2</sup> plus faibles et à une erreur d'estimation accrue, suggérant une pertinence limitée pour la modélisation opérationnelle (voir ANNEXE G).



**Figure 9**

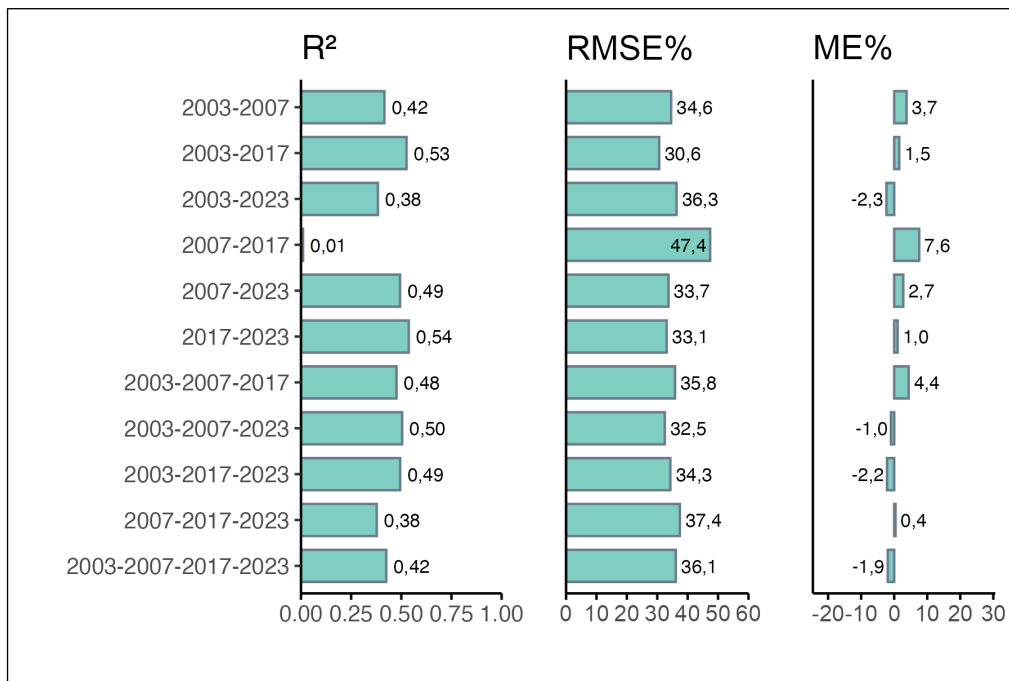
**Performance des modèles Random Forest (RF) pour la prédiction du volume marchand brut (V) dans le couvert forestier feuillu.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

De manière générale, ces résultats mettent en évidence que, bien que les modèles appliqués au couvert feuillu puissent atteindre des valeurs de  $R^2$  relativement élevées, ils demeurent caractérisés par une variabilité plus marquée des indicateurs d'erreur, notamment du ME% et du RMSE%, en comparaison avec les autres types de couvert forestier.

Dans le couvert mixte, les modèles RF ont montré une capacité prédictive modérée, avec des valeurs de  $R^2$  relativement stables entre la majorité des combinaisons de deux ou trois périodes d'acquisition ALS, à l'exception notable du modèle 2007-2017 qui a présenté une performance quasi nulle ( $R^2 = 0,01$  ; RMSE% = 47,4) (Figure 10). Parmi les combinaisons à deux périodes, le modèle 2017-2023 a produit le compromis le plus favorable entre pouvoir explicatif et erreur d'estimation ( $R^2 = 0,54$  ; RMSE% = 33,1 ; ME% = 1,0). Pour les combinaisons à trois périodes, le modèle 2003-2007-2023

s'est révélé particulièrement performant, avec un  $R^2$  de 0,50, un RMSE% de 32,5 et un biais négatif modéré ( $ME\% = -1,0$ ). Le modèle RF mobilisant les quatre périodes d'acquisition (2003-2007-2017-2023) a, pour sa part, présenté un  $R^2$  légèrement inférieur (0,42), mais ses valeurs de RMSE% (36,1) et de  $ME\%$  (-1,9) demeurent dans des seuils jugés acceptables pour des applications opérationnelles à grande échelle.



**Figure 10**  
**Performance des modèles Random Forest (RF) pour la prédiction du volume marchand brut (V) dans le couvert forestier mixte.**

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative ( $ME\%$ ), selon chaque combinaison de périodes d'acquisition ALS.

Par contraste, les modèles LR dans le couvert mixte ont présenté une variabilité plus marquée que les modèles RF en termes de performance prédictive. Le modèle 2003-2017 s'est affirmé comme la configuration la plus robuste, avec un  $R^2$  de 0,64, un RMSE% de 26,7 et un biais quasi nul ( $ME\% = -0,5$ ), constituant ainsi la référence parmi les modèles LR pour ce type de couvert. La configuration 2003-2017-2023 a également produit des résultats acceptables ( $R^2 = 0,43$  ; RMSE% = 36,3 ;  $ME\% = -0,7$ ), montrant que l'ajout d'une troisième année d'acquisition peut accroître la

capacité de généralisation du modèle sans introduire de biais significatif. Par ailleurs, le modèle LR mobilisant les quatre périodes d'acquisition a atteint un  $R^2$  modéré (0,49) et présenté le biais le plus faible de l'ensemble des modèles LR ( $ME\% = -0,3$ ) (voir ANNEXE G).

De manière générale, pour la prédiction du V, les résultats montrent que les modèles RF présentent une performance prédictive globalement stable sur la plupart des configurations d'acquisition, à l'exception notable de la période 2007-2017. Les modèles LR, pour leur part, se distinguent par leur capacité à atteindre, dans certaines configurations, un niveau de précision élevé accompagné d'un biais d'estimation réduit.

L'importance des variables du modèle RF pour la prédiction du V (voir ANNEXE H) montre que l'importance repose principalement sur des métriques de hauteur dérivées du MHC, complétées par des variables décrivant la structure et la distribution verticale du couvert. Pour le couvert résineux, les variables de structure et de distribution : coefficient d'asymétrie des hauteurs (*skewness\_mhc*), pourcentage de couvert à la hauteur moyenne de canopée (*ccmean\_mhc*), rapport de relief de la canopée (*ccr\_mhc*) et pourcentage de couvert à la hauteur moyenne de canopée (*crr\_mhc*) figurent régulièrement parmi les plus influentes, en complément des métriques de hauteur. Pour le couvert feuillu, l'importance est majoritairement associée aux métriques de hauteur dérivées du MHC, incluant la moyenne (*mean\_mhc*) et les 5<sup>e</sup>, 25<sup>e</sup>, 50<sup>e</sup> et 75<sup>e</sup> percentiles (*p05\_mhc*, *p25\_mhc*, *p50\_mhc*, *p75\_mhc*), ainsi qu'à des variables d'occupation verticale, avec l'apparition occasionnelle de descripteurs d'hétérogénéité (*mhr*, *rch*) et de variables topographiques, notamment la variance et écart-type du TPI (*var\_tpi*, *sd\_tpi*). Enfin, pour le couvert mixte, l'importance est principalement portée par des métriques de hauteur des 75<sup>e</sup>, 50<sup>e</sup>, 90<sup>e</sup>, 95<sup>e</sup> percentiles (*p75\_mhc*, *p50\_mhc*, *p90\_mhc*, *p95\_mhc*). Dans l'ensemble, les métriques de hauteur demeurent centrales pour la prédiction du V, tandis que les variables de structure et d'occupation verticale contribuent de manière complémentaire, avec une importance qui dépend du couvert et de la combinaison d'années.

## 2.3.2 Modélisation spatiale

### 2.3.2.1 Hauteur de Lorey

Pour la prédiction de l' $h_L$  dans le couvert résineux, tant les modèles RF que les modèles LR ont démontré une bonne capacité prédictive. Comme l'illustre le Tableau 3, le modèle RF calibré sur les placettes du secteur 1760 a présenté une performance prédictive robuste sur l'ensemble des tests correspondants ( $R^2 = 0,65$  ;  $RMSE\% = 8,6$  ;  $ME\% = 0,6$ ) et a conservé une précision acceptable lorsqu'il a été appliqué au secteur 1923 ( $R^2 = 0,47$ ). Le modèle RF entraîné sur le secteur 1923 a obtenu une précision légèrement inférieure sur son ensemble de test ( $R^2 = 0,43$  ;  $RMSE\% = 6,6$ ), mais a présenté une transférabilité spatiale comparable ( $R^2 = 0,50$  ;  $RMSE\% = 9,7$ ). En comparaison, le modèle LR calibré sur les placettes du secteur 1760 a atteint un  $R^2 = 0,50$  ( $RMSE\% = 8,5$  ;  $ME\% = 0,3$ ) sur l'ensemble de test et une meilleure transférabilité vers le secteur 1923 ( $R^2 = 0,52$  ;  $RMSE\% = 7,2$  ;  $ME\% = 0,2$ ), suggérant une performance plus stable avec un biais réduit.

Pour la prédiction de l' $h_L$  dans le couvert mixte, les modèles RF et LR ont présenté une capacité prédictive modérée. Les modèles RF ont atteint des valeurs de  $R^2$  légèrement supérieures lors de la calibration, tandis que les modèles LR ont généralement affiché un biais plus faible et une transférabilité plus stable dans certains cas (Tableau 3). Le modèle RF entraîné sur les placettes du secteur 1760 a obtenu la meilleure performance sur son ensemble de test ( $R^2 = 0,64$  ;  $RMSE\% = 6,9$  ;  $ME\% = 1,2$ ), mais la précision a diminué lors du transfert vers le secteur 1923 ( $R^2 = 0,21$  ;  $RMSE\% = 7,7$  ;  $ME\% = 0,5$ ). Le modèle RF calibré sur le secteur 1923 a montré une performance inférieure dans son ensemble de calibration ( $R^2 = 0,30$  ;  $RMSE\% = 8,2$  ;  $ME\% = -0,1$ ), mais une transférabilité légèrement meilleure ( $R^2 = 0,36$  ;  $RMSE\% = 9,3$  ;  $ME\% = 1,0$ ). Concernant les modèles LR, celui entraîné sur les placettes du secteur 1760 a affiché une bonne précision initiale ( $R^2 = 0,61$  ;  $RMSE\% = 6,8$  ;  $ME\% = 0,6$ ), mais sa transférabilité vers le secteur 1923 s'est nettement dégradée ( $R^2 = 0,23$  ;  $RMSE\% = 8,3$  ;  $ME\% = 2,9$ ). Enfin, le modèle LR construit sur les placettes du secteur 1923 a obtenu une précision plus limitée lors de la calibration ( $R^2 = 0,34$  ;  $RMSE\% = 7,9$  ;  $ME\% = 1,0$ ), mais a maintenu une performance relativement stable lors du transfert vers le secteur 1760 ( $R^2 = 0,28$  ;  $RMSE\% = 7,0$  ;  $ME\% = 3,8$ ).

**Tableau 3****Performance des modèles pour la prédiction de la hauteur de Lorey ( $h_L$ ).**

Indicateurs de performance des modèles Random Forest (RF) et Régression linéaire (LR) dans les couverts résineux et mixtes, selon le secteur d'origine du feu (1760 ou 1923). Les résultats sont présentés pour les données de test et de transférabilité spatiale.

Couvert / Indicateur	RF Feu 1760		RF Feu 1923		LR Feu 1760		LR Feu 1923	
	Test	Transfert Spatial	Test	Transfert Spatial	Test	Transfert Spatial	Test	Transfert Spatial
<b>Résineux</b>								
<b>R<sup>2</sup></b>	0,65	0,47	0,43	0,50	0,50	0,52	0,22	0,35
<b>RMSE%</b>	8,6	7,9	6,6	9,7	8,5	7,2	7,5	15,0
<b>ME%</b>	0,6	3,7	-0,3	-3,4	0,3	0,2	-2,1	-10,5
<b>Mixte</b>								
<b>R<sup>2</sup></b>	0,64	0,21	0,30	0,36	0,61	0,23	0,34	0,28
<b>RMSE%</b>	6,9	7,7	8,2	9,3	6,8	8,3	7,9	7,0
<b>ME%</b>	1,2	0,5	-0,1	1,0	0,6	2,9	1,0	3,8

Pour la prédiction de  $h_L$  à l'aide des modèles RF, l'analyse de l'importance des variables (voir Annexe I) montre que, dans le couvert résineux, le modèle calibré dans le secteur 1760 est principalement associé à `sd_mhc` et `var_mhc`, suivis de `p99_mhc` et `p95_mhc`, tandis que `perc_up_20` contribue plus faiblement. Dans le secteur 1923, `p99_mhc` constitue la variable prédominante, suivie de `perc_up_20`, `var_mhc` et `p95_mhc`, tandis que `mean_elev` présente une importance relative moindre. Dans le couvert mixte, le modèle du secteur 1760 se distingue par la prédominance de `p95_mhc`, suivie de `var_mhc`, `p99_mhc` et `perc_up_20`, avec `p90_mhc` comme la variable la moins contributive. Dans le secteur 1923, `p99_mhc` est la variable la plus importante, suivie de `p95_mhc` et `p90_mhc`, tandis que `p75_mhc` et `var_mhc` contribuent moins.

### 2.3.2.2 Surface terrière

Les résultats de la modélisation de BA ont révélé des performances plus faibles que celles observées pour l' $h_L$ . Dans l'ensemble, les modèles RF ont présenté une meilleure capacité prédictive que les modèles LR, mais les valeurs de  $R^2$  sont demeurées modestes, ce qui traduit la complexité de la représentation de cet attribut à partir des données ALS (Tableau 4).

**Tableau 4**

**Performance des modèles pour la prédiction de la surface terrière (BA).**

Indicateurs de performance des modèles Random Forest (RF) et Régression linéaire (LR) dans les couverts résineux et mixtes, selon le secteur d'origine du feu (1760 ou 1923). Les résultats sont présentés pour les données de test et de transférabilité spatiale.

Couvert / Indicateur	RF Feu 1760		RF Feu 1923		LR Feu 1760		LR Feu 1923	
	Test	Transfert Spatial	Test	Transfert Spatial	Test	Transfert Spatial	Test	Transfert Spatial
<b>Résineux</b>								
<b>R<sup>2</sup></b>	0,03	0,05	0,19	0,06	-0,52	0,12	0,05	0,05
<b>RMSE%</b>	48,3	35,7	31,8	42,3	60,0	37,8	33,3	43,3
<b>ME%</b>	4,0	21,5	1,1	-10,6	16,3	17,6	0,0	-17,8
<b>Mixte</b>								
<b>R<sup>2</sup></b>	0,43	0,06	0,18	0,17	0,27	0,10	0,09	0,35
<b>RMSE%</b>	30,8	35,8	33,4	45,6	34,6	32,0	34,5	43,7
<b>ME%</b>	-1,3	3,8	-5,6	-30,3	3,4	5,7	-0,2	31,0

Dans le couvert résineux, les performances des modèles RF et LR sont restées limitées. Le modèle RF calibré sur le secteur 1760 atteint un  $R^2 = 0,03$  et un  $RMSE\% = 48,3$ , tandis que le transfert vers le secteur 1923 présentait une performance comparable ( $R^2 = 0,05$  ;  $RMSE\% = 35,7$ ). Le modèle RF 1923 montrait une meilleure

calibration ( $R^2 = 0,19$  ;  $RMSE\% = 31,8$  ;  $ME\% = 1,1$ ) mais une baisse de précision lors du transfert vers le secteur 1760 ( $R^2 = 0,06$  ;  $RMSE\% = 42,3$  ;  $ME\% = -10,6$ ).

Les modèles LR affichaient des performances nettement inférieures à celles obtenues avec la RF. Pour le secteur 1760, le modèle atteignait un  $R^2 = -0,52$  et un  $RMSE\% = 60,0$ , indiquant une incapacité à reproduire la variance observée et une instabilité importante du modèle. Pour le secteur 1923, la calibration montrait une performance légèrement meilleure ( $R^2 = 0,05$  ;  $RMSE\% = 33,3$ ), mais la transférabilité demeurait faible ( $R^2 = 0,05$  ;  $RMSE\% = 43,3$ ), confirmant la faible robustesse de cette approche paramétrique pour la prédiction de la BA.

Dans le couvert mixte, les résultats obtenus étaient légèrement supérieurs. Le modèle RF calibré pour le secteur 1760 affiche un  $R^2$  de 0,43, un  $RMSE\%$  de 30,8, et un  $ME\%$  de -1,3, tout en conservant une transférabilité réduite vers le secteur 1923, avec un  $R^2$  de 0,06, un  $RMSE\%$  de 35,8, et un  $ME\%$  de 3,8. Le modèle RF 1923 présentait une calibration plus faible ( $R^2 = 0,18$ ;  $RMSE\% = 33,4$  ;  $ME\% = -5,6$ ) et une transférabilité réduite ( $R^2 = 0,17$  ;  $RMSE\% = 45,6$  ;  $ME\% = -30,0$ ). Les modèles LR montraient une variabilité importante : celui construit sur le secteur 1760 atteignait un  $R^2 = 0,27$  ;  $RMSE\% = 34,6$  ;  $ME\% = 3,4$ , tandis que le modèle LR 1923 affichait un  $R^2 = 0,09$  ;  $RMSE\% = 34,5$  ;  $ME\% = -0,2$ .

Pour la prédiction de BA à l'aide des modèles RF, l'importance des variables (voir ANNEXE J) varie selon le type de couvert et le secteur. Dans le couvert résineux, le modèle du secteur 1760 est principalement associé à `perc_up_20`, `mean_mhc` et `p90_mhc`, avec une contribution additionnelle de `max_slope`, tandis que `p99_tpi` présente une faible importance. Dans le secteur 1923, `ccr_mhc` constitue la variable la plus importante, suivie de `perc_10_20` et `ccmean_mhc`. Dans le couvert mixte, le secteur 1760 est dominé par `p75_mhc` et `p50_mhc`, suivis de `mean_mhc` et `perc_up_20`, tandis que `p99_mhc` contribue moins. Pour le secteur 1923, `max_elev` et `p25_elev` présentent les contributions les plus élevées, suivies de `perc_up_20` et `cv_twi`, tandis que `p75_elev` présente une importance relative moindre.

### 2.3.2.3 Volume marchand brut

Pour la prédiction du V dans le couvert résineux, la performance des modèles est demeurée généralement faible. Les approches RF et LR ont produit des coefficients de détermination similaires et limités, sans qu'aucune ne se démarque par une supériorité constante. Le modèle RF calibré sur les placettes du secteur 1760 a présenté une précision réduite ( $R^2 = 0,24$  ;  $RMSE\% = 46,0$  ;  $ME\% = 8,6$ ) et sa performance a encore décliné lors du transfert vers le secteur 1923 ( $R^2 = 0,11$  ;  $RMSE\% = 40,2$  ;  $ME\% = 21,3$ ). De même, le modèle RF issu du secteur 1923 a affiché une faible capacité prédictive, tant dans son ensemble de calibration ( $R^2 = 0,26$  ;  $RMSE\% = 32,6$  ;  $ME\% = 3,6$ ) que lors de son application au secteur 1760 ( $R^2 = 0,08$  ;  $RMSE\% = 48,2$  ;  $ME\% = -13,5$ ). Les modèles LR n'ont pas montré de gains notables ; celui développé à partir des placettes du secteur 1760 a obtenu un  $R^2$  de 0,13 ( $RMSE\% = 48,3$  ;  $ME\% = 8,6$ ), avec une légère amélioration lors de son transfert vers le secteur 1923 ( $R^2 = 0,18$  ;  $RMSE\% = 38,8$  ;  $ME\% = 21,5$ ). Enfin, le modèle LR calibré sur le secteur 1923 a produit les résultats les plus faibles ( $R^2 = -0,01$  ;  $RMSE\% = 35,9$  ;  $ME\% = 3,6$ ) et a conservé une précision très limitée lorsqu'il a été appliqué au secteur 1760 ( $R^2 = 0,11$  ;  $RMSE\% = 52,1$  ;  $ME\% = -25,6$ ). Dans l'ensemble, ces résultats soulignent la forte incertitude et la faible transférabilité des modèles pour l'estimation du V dans les couverts résineux (Tableau 5).

Dans le couvert mixte, les performances des modèles ont varié de manière notable selon l'approche retenue et le secteur de calibration (Tableau 5). Les modèles RF ont généralement produit des coefficients de détermination supérieurs à ceux des LR dans les ensembles de test, bien que leur transférabilité spatiale demeure limitée quelle que soit la configuration. Le modèle RF calibré sur les placettes du secteur 1760 a affiché la performance la plus élevée en test ( $R^2 = 0,67$  ;  $RMSE\% = 29,7$  ;  $ME\% = 0,9$ ), mais sa précision a fortement décliné lors du transfert vers le secteur 1923 ( $R^2 = 0,17$  ;  $RMSE\% = 38,0$  ;  $ME\% = 3,8$ ). À l'inverse, le modèle RF construit sur le secteur 1923 a montré une faible précision dès la calibration ( $R^2 = 0,15$  ;  $RMSE\% = 36,4$  ;  $ME\%$

= -2,4) et une sous-estimation marquée lorsqu'il a été appliqué au secteur 1760 ( $R^2 = 0,37$  ;  $RMSE\% = 47,1$  ;  $ME\% = -25,5$ ).

**Tableau 5**

**Performance des modèles pour la prédiction du volume marchand brut (V).**

Indicateurs de performance des modèles Random Forest (RF) et Régression linéaire (LR) dans les couverts résineux et mixtes, selon le secteur d'origine du feu (1760 ou 1923). Les résultats sont présentés pour les données de test et de transférabilité spatiale.

Couvert/ Indicateur	RF Feu 1760		RF Feu 1923		LR Feu 1760		LR Feu 1923	
	Test	Transfert Spatial	Test	Transfert Spatial	Test	Transfert Spatial	Test	Transfert Spatial
<b>Résineux</b>								
<b>R2</b>	0,24	0,11	0,26	0,08	0,13	0,18	-0,01	0,11
<b>RMSE%</b>	46,0	40,2	32,6	48,2	48,3	38,8	35,9	52,1
<b>ME%</b>	8,6	21,3	3,6	-13,5	8,6	21,5	3,6	-25,6
<b>Mixte</b>								
<b>R2</b>	0,67	0,17	0,15	0,37	0,51	0,17	0,18	0,39
<b>RMSE%</b>	29,7	38,0	36,4	47,1	36,0	34,2	34,6	54,1
<b>ME%</b>	0,9	3,8	-2,4	-25,5	-27,4	9,3	-37,9	-37,1

Concernant les modèles LR, celui calibré sur les placettes du secteur 1760 a atteint une performance modérée ( $R^2 = 0,51$  ;  $RMSE\% = 36,0$ ;  $ME\% = -27,4$ ), mais sa transférabilité vers le secteur 1923 a entraîné une baisse de précision et une inversion du biais ( $R^2 = 0,17$  ;  $RMSE\% = 34,2$ ;  $ME\% = 9,3$ ). Enfin, le modèle LR développé sur le secteur 1923 a fourni des résultats faibles tant en calibration ( $R^2 = 0,18$  ;  $RMSE\% = 34,6$ ;  $ME\% = -37,9$ ) qu'en transfert ( $R^2 = 0,39$  ;  $RMSE\% = 54,1$  ;  $ME\% = -37,1$ ), confirmant une forte tendance à la sous-estimation.

L'importance des variables pour la prédiction du V avec le RF (voir Annexe K) varie selon le type de couvert et le secteur. Dans le couvert résineux, le modèle du secteur 1760 est principalement associé à p90\_mhc et p95\_mhc, suivies de mean\_mhc,

tandis que `perc_up_20` et `var_mhc` présentent une contribution plus faible. Dans le secteur 1923, `skewness_mhc` constitue la variable la plus importante, suivie de `rch`, `ccmean_mhc`, `p75_elev`, alors que `mhr` affiche une importance relative moindre. Dans le couvert mixte, le secteur 1760 est dominé par `perc_up_20` et `p95_mhc`, suivies de `p50_mhc` et `p75_mhc`, tandis que `mean_mhc` contribue plus faiblement. Pour le secteur 1923, `perc_up_20` présente l'importance la plus élevée, suivie de `kurtosis_slope` et `p25_elev`, alors que `p90_mhc` et `p75_mhc` affichent des contributions plus limitées.

## 2.4 Discussion

### 2.4.1 Impact des configurations ALS multi-temporelles sur la capacité prédictive

Cette étude, fondée sur quatre acquisitions ALS réalisées sur une période de vingt ans dans un contexte boréal, met en évidence l'influence déterminante de la configuration temporelle des vols (c'est-à-dire la fréquence et la distance temporelle entre les acquisitions) sur la performance prédictive des modèles paramétriques et non paramétriques destinés à estimer les variables dendrométriques d'intérêt.

L'analyse multi-temporelle a montré que la combinaison de données issues de différentes campagnes ne conduit pas nécessairement à une amélioration de la performance des modèles, mais que certains intervalles temporels spécifiques favorisent des estimations plus stables et plus précises selon l'attribut étudié, le type de couvert forestier et la méthode de modélisation employée.

Cette influence de la configuration temporelle a déjà été abordée par plusieurs auteurs à partir de jeux de données bi-temporels. Par exemple, Yu *et al.* (2008) ont souligné que, dans le cas de forêts plus jeunes et à croissance rapide, des séries temporelles plus longues (10 à 15 ans) peuvent offrir des performances supérieures, car elles permettent de capturer des variations structurelles plus marquées et de disposer de données de référence plus précises. Ainsi, la durée optimale entre acquisitions dépend étroitement du stade de développement du peuplement et du taux de croissance, confirmant que la robustesse temporelle des modèles n'est pas uniforme à travers les types de couverts et les contextes écologiques.

Dans le cadre de cette étude, la modélisation de  $h_L$ , a montré les meilleures performances lorsqu'elle s'appuyait sur deux acquisitions ALS correspondant à la période 2007-2017. Cette combinaison s'est révélée la plus performante pour les trois types de couvert forestier, ce qui suggère qu'un intervalle d'environ une décennie constitue un compromis solide entre la représentativité des dynamiques de croissance propres aux forêts boréales et la stabilité des relations statistiques entre les métriques ALS et les attributs dendrométriques. Cet intervalle semble suffisamment long pour intégrer les changements structuraux liés à la croissance et à la compétition, tout en préservant la cohérence des corrélations observées.

En ce qui concerne l'estimation de  $h_L$  à l'aide d'un modèle prédictif, bien que tous les types de couvert présentent, avec le modèle RF, des valeurs de  $R^2$  supérieures à 0,6, un biais plus faible a été observé dans les peuplements résineux. Ces résultats confirment les observations de Næsset (2004) et Bouvier *et al.* (2015), qui soulignent que les forêts dominées par des conifères tolérants à l'ombre, telles que l'épinette noire et le sapin baumier, se distinguent par une forte stabilité structurale et une homogénéité spatiale réduisant la variance du signal et renforçant la qualité des relations statistiques entre les métriques ALS et la hauteur. Un facteur explicatif réside dans la morphologie des houppiers : les conifères à aiguilles favorisent une pénétration plus homogène du signal laser, limitant les variations de hauteur et renforçant la stabilité des modèles (Lefsky *et al.*, 1999).

À l'inverse, pour BA et V, les résultats se sont révélés plus variables d'une configuration temporelle à l'autre. Bien que les méthodes de modélisation utilisées ne permettent pas une comparaison directe avec les travaux existants, les tendances de variance expliquée concordent avec celles rapportées par Tompalski *et al.* (2019a), qui ont mis en évidence plusieurs défis liés à l'évaluation de la croissance forestière à partir de jeux de données bi-temporels acquis avec des technologies différentes (ALS et DAP (Digital Aerial Photogrammetry)). Ces difficultés sont liées non seulement à la nature et à la compatibilité des données, mais également à l'intervalle temporel, au taux de croissance et à la mortalité.

Ces résultats confirment également les dynamiques de croissance différenciées entre types de couverts, telles que décrites par Tompalski *et al.* (2019a), où la vitesse de croissance et le degré de perturbation du peuplement influencent directement la capacité des données ALS à détecter les changements structuraux au fil du temps. Dans les peuplements feuillus, souvent situés à des stades plus précoces de succession et caractérisés par une croissance rapide, les intervalles temporels courts (comme 2003-2007) ont permis d'obtenir les meilleures performances pour la surface terrière ( $R^2 = 0,83$ ). Ce comportement est conforme aux observations de Tompalski *et al.* (2019a), qui ont noté que, dans les forêts boréales mixtes et feuillues, les taux de croissance plus élevés favorisent la détection de changements structuraux même sur de courtes périodes. De manière similaire, les couverts mixtes ont également montré une bonne capacité de modélisation de la BA sur des intervalles récents (2017-2023), suggérant que la dynamique de renouvellement et de compétition y reste suffisamment active pour être captée par les métriques ALS.

À l'inverse, les couverts résineux, dominés par des espèces à croissance lente comme l'épinette noire et le sapin baumier, présentent une évolution structurelle plus graduelle, ce qui explique les faibles performances observées sur les périodes courtes ( $R^2 = 0,37$  pour 2003-2007). Concernant le volume marchand ( $V$ ), les meilleures performances observées dans les peuplements résineux ( $R^2 = 0,68$  pour 2003-2017) suggèrent qu'un intervalle plus long est nécessaire pour intégrer les effets cumulatifs de la croissance en diamètre et en hauteur, deux composantes directement liées à la variation du volume. En revanche, la variabilité accrue dans les couverts mixtes et feuillus pourrait refléter la complexité de leurs structures verticales et la sensibilité du volume aux erreurs de mesure et aux changements locaux de densité de tiges.

Enfin, d'un point de vue opérationnel, il est intéressant de noter que l'augmentation du nombre d'acquisitions ALS ne conduit pas nécessairement à une amélioration significative des performances prédictives. Ce résultat souligne qu'au-delà d'un certain seuil, la redondance temporelle n'apporte pas de gains substantiels en termes de capacité prédictive, tandis que les coûts logistiques et financiers des acquisitions supplémentaires demeurent élevés. Ainsi, la planification optimale des intervalles

temporels entre les vols apparaît plus déterminante que leur fréquence, ce qui constitue un point clé pour les applications de suivi forestier à long terme. Ce constat s'aligne directement avec le premier objectif de cette étude, visant à identifier la configuration temporelle la plus efficace des données ALS pour la modélisation multi-temporelle des attributs dendrométriques en forêt boréale, en conciliant la capacité prédictive des estimations avec la faisabilité opérationnelle.

#### 2.4.2 Transférabilité spatiale

L'analyse spatiale, basée sur la combinaison des quatre acquisitions ALS couvrant une période de vingt ans, a permis d'évaluer la transférabilité des modèles entre deux secteurs différenciés par l'année de feu. Les résultats montrent que la transférabilité dépend fortement de la variable dendrométrique modélisée, du type de couvert forestier et de la méthode de modélisation employée.

Les résultats de cette étude portant sur la transférabilité des variables dendrométriques indiquent que l' $h_L$  est davantage transférable que le BA ainsi que le volume. Ces conclusions concordent avec Tompalski *et al.* (2019b) qui a été évaluée la transférabilité spatiale à partir d'une seule acquisition LiDAR pour la modélisation de l' $h_L$ , du diamètre quadratique moyen et du volume marchand dans une forêt tempérée dominée par des conifères.

Un aspect supplémentaire mis en évidence par la présente étude est l'influence du type de couvert forestier sur la transférabilité des modèles. Les résultats montrent que les peuplements résineux présentent systématiquement une transférabilité plus élevée que les peuplements mixtes, et ce, même lorsqu'ils sont affectés par des perturbations historiques différentes. À notre connaissance, cet effet stabilisateur du couvert résineux sur la transférabilité entre secteurs soumis à des régimes de perturbations contrastés n'avait pas été documenté auparavant et constitue ainsi une contribution originale. Comme le montre l'analyse de composition (voir ANNEXE A), la composition spécifique varie entre les secteurs : le secteur 1760 est dominé par le sapin baumier, alors que le secteur 1923 est largement dominé par le pin gris. Ce résultat suggère que l'homogénéité de la structure de canopée et la dynamique de

croissance plus lente des forêts résineuses réduisent la variabilité spatiale des relations entre les métriques ALS et les attributs dendrométriques, favorisant ainsi une meilleure transférabilité des modèles.

Concernant les méthodes de modélisation, il est notable que les modèles RF, pour  $h_L$ , affichent une transférabilité bidirectionnelle robuste dans les couverts résineux, alors que les modèles LR montrent une capacité de généralisation plus limitée, particulièrement lorsque les modèles calibrés avec les données du secteur 1923, dominé par le pin gris, sont transférés vers le secteur 1760. Cette baisse de performance est cohérente avec Fekety *et al.* (2018), qui soulignent que l'extrapolation hors des conditions écologiques de calibration doit être prudente, car la transférabilité n'est assurée qu'entre peuplements présentant des signatures structurales comparables. Cependant, ce résultat diffère de celui obtenu par Tompalski *et al.* (2019b) qui ont observé un biais plus faible pour  $h_L$  avec les modèles paramétriques. Cette divergence pourrait s'expliquer par des différences méthodologiques, en particulier, Tompalski *et al.* (2019b) mentionnent que le pooling (utilisé dans cette étude) augmente le nombre d'échantillons disponibles, ce qui peut améliorer la performance des modèles, notamment des modèles non paramétriques, plus sensibles à l'étendue des données de calibration.

#### 2.4.3 Comparaison des méthodes paramétriques et non paramétriques

L'évaluation des modèles basés sur des acquisitions ALS multi-temporelles met en évidence des différences notables entre les approches paramétriques et non paramétriques. Dans l'ensemble, les modèles RF se sont distingués par une plus grande capacité à capturer la variabilité structurelle complexe des forêts boréales, avec des valeurs de  $R^2$  supérieures dans la majorité des combinaisons temporelles. Cette performance s'explique par la flexibilité des RF, qui ne reposent pas sur des formes fonctionnelles prédéfinies et s'adaptent mieux aux relations non linéaires entre les métriques ALS et les attributs dendrométriques (Breiman, 2001; Gareth *et al.*, 2013; White *et al.*, 2013b)

Les modèles LR ont montré une stabilité appréciable dans certains cas, notamment pour la prédiction de  $h_L$  dans les couverts résineux, où la structure homogène et les dynamiques lentes réduisent la complexité des relations à modéliser. Dans ce contexte, la simplicité et l'interprétabilité des modèles paramétriques constituent un avantage opérationnel, malgré une précision globale plus faible (Næsset *et al.*, 2005; Penner *et al.*, 2013).

Cependant, dans les forêts mixtes et feuillues, caractérisées par une structure verticale plus complexe et une plus grande variabilité spatiale, les modèles linéaires ont montré des limites significatives en termes de généralisation et de robustesse.

Ces résultats confirment que les approches non paramétriques tendent à surpasser les méthodes paramétriques en termes de performance prédictive, mais que la pertinence des modèles LR demeure dans des contextes spécifiques, caractérisés par une structure relativement simple ou une faible hétérogénéité spatiale (Cosenza *et al.*, 2021).

#### 2.4.4 Limitations et perspectives d'amélioration

En ce qui concerne les limites de cette étude, il est primordial d'examiner en premier lieu la variabilité entre les différentes campagnes d'acquisition LiDAR, notamment en ce qui concerne l'altitude de vol, les paramètres instrumentaux et les méthodes de traitement. Ces différences peuvent avoir une incidence sur la précision des métriques dérivées à partir des rasters. Bien que tous les modèles de hauteur aient été normalisés à une résolution spatiale d'un mètre, les différences de conditions d'acquisition peuvent entraîner de légères divergences dans la représentation du couvert forestier. Ces écarts, bien que modérés, peuvent influencer la comparaison directe des métriques structurelles entre les années et, par conséquent, la stabilité des modèles multi-temporelles.

Deuxièmement, une limite significative concerne les placettes d'inventaire, notamment l'absence de mesures directes de la hauteur totale des arbres. Dans le cadre de ce travail, les hauteurs ont été estimées à l'aide de modèles allométriques basés sur le DHP. Bien que ces modèles soient statistiquement robustes et largement

employés en foresterie, leur utilisation introduit une marge d'erreur supplémentaire, particulièrement dans les peuplements hétérogènes présentant une variabilité interspécifique prononcée. Cette erreur peut affecter la précision des attributs dérivés tels que la hauteur de Lorey, la surface terrière et le volume marchand.

Troisièmement, la désynchronisation temporelle entre les acquisitions ALS et les mesures de terrain constitue une autre source potentielle d'erreur. Les placettes utilisées n'ont pas toujours été mesurées la même année que le vol correspondant, mais dans un intervalle de deux à trois ans autour de celui-ci. Bien que cet écart temporel puisse être considéré comme acceptable compte tenu du rythme de croissance des forêts boréales, il peut toutefois réduire la cohérence entre les conditions réelles du couvert au moment du vol et les données dendrométriques utilisées pour la calibration, notamment dans les peuplements à croissance rapide, tels que les forêts feuillues.

Enfin, le plan d'échantillonnage des placettes de terrain pourrait être amélioré afin d'éviter la sous-représentation de certains types de couvert, notamment les forêts feuillues. Un échantillonnage plus équilibré entre les différents types de peuplements renforcerait la robustesse des modèles et la fiabilité des analyses de transférabilité. Il convient également de souligner que cette étude s'applique principalement à des forêts matures ; des recherches complémentaires seraient donc nécessaires pour évaluer la performance des modèles dans des peuplements plus jeunes, où la structure verticale et la dynamique de croissance diffèrent. De plus, dans l'analyse spatiale, la différenciation entre secteurs a principalement été interprétée à partir du régime de perturbation par le feu, sans intégrer explicitement d'autres facteurs susceptibles de structurer la variabilité spatiale (p. ex. des facteurs édaphiques tels que le type de sol et la fertilité, ainsi qu'un indice de qualité de station). L'intégration de telles covariables permettrait de mieux distinguer l'effet du feu de celui des conditions du site, et ainsi d'interpréter plus clairement les écarts de performance et de transférabilité observés entre secteurs.

## *2.5 Conclusion*

Ce mémoire a permis d'évaluer la performance prédictive ainsi que la transférabilité spatiale des modèles ABA dérivés de rasters, fondés sur des données ALS multi-temporelles, une approche encore peu documentée dans le contexte boréal québécois. L'analyse conjointe de quatre campagnes ALS (2003, 2007, 2017 et 2023) a permis de quantifier, pour la première fois à l'échelle de la FERLD, l'effet de différentes combinaisons temporelles sur la stabilité et la précision des modèles dendrométriques. Les résultats ont mis en évidence l'influence déterminante de la configuration temporelle des acquisitions sur la performance des modèles, permettant d'identifier les intervalles les plus efficaces pour la prédiction des attributs dendrométriques selon le type de couvert forestier.

Cette recherche met en lumière trois résultats fondamentaux qui contribuent à une meilleure compréhension du comportement des modèles ABA en contexte boréal. Premièrement,  $h_L$  s'est révélée l'attribut le plus robuste et le plus transférable, particulièrement dans les forêts résineuses ( $R^2 = 0,46 - 0,68$ ), où la structure verticale homogène et la dynamique de croissance plus lente favorisent la stabilité des relations entre les métriques ALS et les mesures de terrain. Deuxièmement, les modèles RF ont surpassé LR, confirmant leur aptitude à représenter les relations non linéaires complexes entre les variables LiDAR et les attributs forestiers. Troisièmement, les intervalles temporels intermédiaires (environ 10 à 15 ans) ont offert le meilleur compromis entre précision et stabilité pour  $h_L$  (2007-2017 :  $R^2 = 0,66$  ;  $RMSE\% = 6,5$ ), tandis que des intervalles trop courts engendrent une perte de performance prédictive (2003-2007 :  $R^2 = 0,57$  ;  $RMSE\% = 8,4$  ; 2017-2023 :  $R^2 = 0,53$  ;  $RMSE\% = 8,5$ ).

En ce qui concerne la transférabilité spatiale, les résultats indiquent que les modèles RF appliqués à  $h_L$  en couverts résineux ont présenté une performance robuste et bidirectionnelle ( $R^2 \approx 0,5$ ). Cette capacité de généralisation s'explique par la similarité structurelle et compositionnelle entre les deux secteurs, dominés par des conifères à croissance lente. À l'inverse, la transférabilité s'est avérée plus limitée dans les couverts mixtes ( $R^2 < 0,3$ ), où l'hétérogénéité structurelle et la dynamique rapide de régénération accentuent les écarts entre secteurs. La BA et le V demeurent les attributs les plus difficiles à transférer, car ils dépendent de paramètres horizontaux

(densité et diamètre) moins bien captés par le signal laser. Ces résultats mettent en évidence que la robustesse des modèles repose avant tout sur la similarité écologique entre les secteurs de calibration et d'application (Fekety *et al.*, 2018; Tompalski *et al.*, 2019b).

Les résultats obtenus démontrent que l'utilisation de modèles prédictifs appliqués à des couches raster de données forestières, combinée à des intervalles temporels intermédiaires et à des méthodes d'apprentissage flexibles comme RF, constitue une voie prometteuse pour l'estimation multi-temporelle des attributs forestiers et la mise en œuvre de systèmes d'inventaire efficaces, reproductibles et économiquement viables dans les forêts boréales.

## CONCLUSION GÉNÉRALE

Ce mémoire a permis d'approfondir la compréhension du comportement des modèles ABA appliqués à des données ALS multi-temporelles dans les forêts boréales du Québec. En intégrant quatre campagnes ALS réparties sur vingt ans, cette recherche s'est inscrite dans un cadre novateur qui a exploré simultanément les dimensions temporelles et spatiales de la modélisation dendrométrique. L'approche adoptée, fondée sur des métriques dérivées de rasters à 1 m de résolution et sur des modèles d'apprentissage flexibles, a apporté une contribution méthodologique importante à l'amélioration des inventaires forestiers à grande échelle.

Sur le plan scientifique, les résultats ont démontré que la configuration temporelle des acquisitions ALS avait exercé une influence déterminante sur la stabilité et la précision des modèles. Les intervalles intermédiaires, de l'ordre de dix à quinze ans, se sont révélés les plus efficaces pour représenter la dynamique de croissance forestière tout en limitant les effets d'erreur liés aux variations d'acquisition ou à la structure du peuplement. Parmi les attributs étudiés, la hauteur de Lorey s'est distinguée comme la variable la plus robuste et la plus transférable, tandis que la surface terrière et le volume marchand se sont avérés plus sensibles aux différences structurelles entre les types de couvert.

Sur le plan méthodologique, l'utilisation de métriques dérivées de rasters ALS s'est révélée particulièrement prometteuse. Cette approche a réduit les contraintes computationnelles, a assuré une meilleure homogénéité des données et a facilité la reproductibilité des modèles à grande échelle. Combinée à des modèles non paramétriques multi-temporelles, elle a constitué une solution opérationnelle fiable et transférable pour la prédiction d'attributs dendrométriques clés dans les inventaires forestiers contemporains. L'analyse de l'importance des variables a montré que les percentiles supérieurs de hauteur (95e et 99e percentiles des hauteurs des cellules > 2 m) étaient les plus performants pour la prédiction de  $h_L$ , ce qui traduit fidèlement la structure dominante du couvert. Pour BA et V, les métriques de dispersion et de variabilité verticale (écart-type et variance des hauteurs de cellules > 2 m, indice de complexité verticale, pourcentage de hauteurs de cellules > 10 m et < 20 m) se sont

révélées plus explicatives, soulignant l'importance des contrastes verticaux et de la distribution des hauteurs pour représenter la densité et la structure du couvert forestier. Ces constats ont démontré la complémentarité des métriques structurelles issues des rasters ALS pour la prédiction des attributs dendrométriques clés.

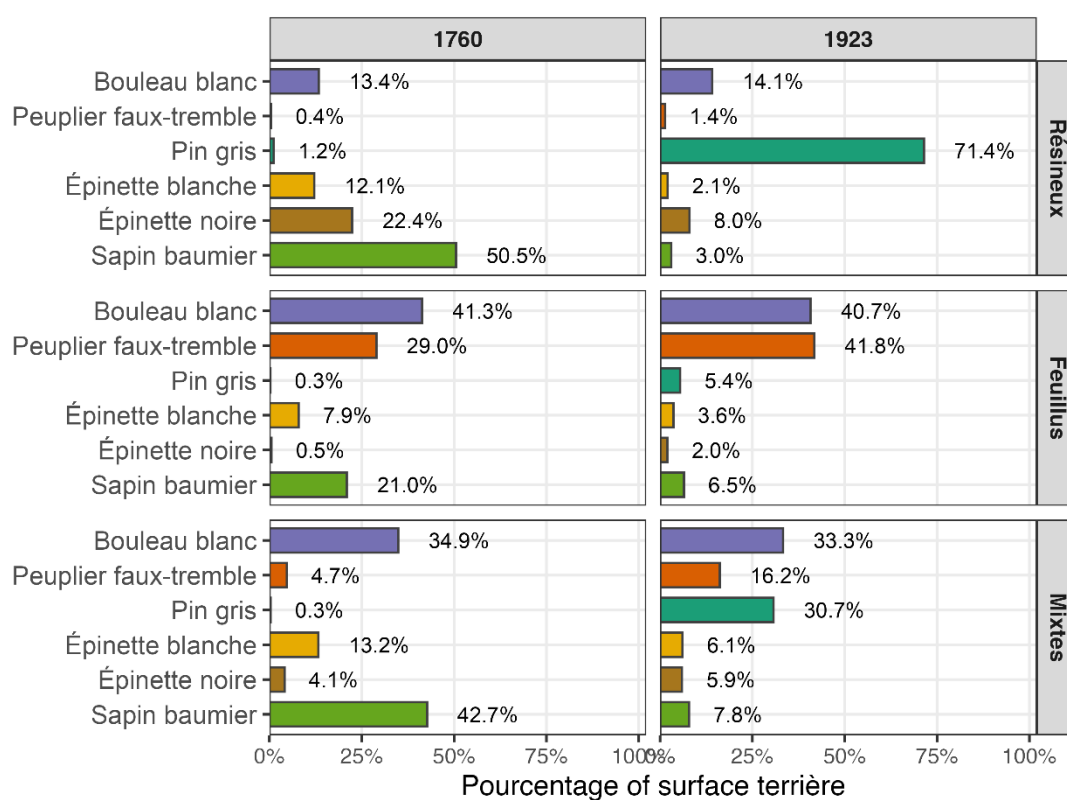
Sur le plan opérationnel, cette étude a fourni des balises concrètes pour l'optimisation des inventaires forestiers au Québec. Elle a suggéré que la planification des acquisitions ALS devait viser des intervalles intermédiaires (10-15 ans), mieux adaptés aux rythmes de croissance observés dans les forêts boréales. Elle a également mis en avant l'intérêt de concentrer les calibrations sur les couverts résineux, dont la stabilité structurelle a permis une extrapolation plus robuste vers d'autres types de peuplements. À l'inverse, dans les couverts feuillus et mixtes, plus dynamiques et hétérogènes, l'intégration de données complémentaires, telles que la cartographie écoforestière, l'indice de qualité de station ou la stratification par classes d'âge, aurait pu améliorer la précision des estimations.

L'application de ces recommandations dans les inventaires forestiers du Québec offrirait plusieurs avantages tangibles. Elle permettrait de fiabiliser les estimations dendrométriques, de réduire la dépendance à des placettes de terrain coûteuses et de valoriser pleinement les acquisitions ALS déjà disponibles. Elle contribuerait également à la mise en place d'inventaires forestiers de nouvelle génération, capables d'intégrer les dimensions temporelles et spatiales de manière cohérente. Une telle évolution favoriserait à la fois la planification opérationnelle et la prise de décision stratégique, en fournissant des informations mieux adaptées aux exigences de l'aménagement durable et à la surveillance des ressources dans un contexte de changements écologiques et climatiques. De plus, la possibilité de transférer des modèles entre secteurs écologiquement similaires représente une opportunité pour étendre la couverture des inventaires vers des territoires éloignés ou difficiles d'accès.

En résumé, les résultats de ce mémoire démontrent que l'ALS multi-temporelle, intégrée dans un cadre ABA et appuyée par des modèles RF, offre une voie crédible pour renforcer les inventaires et soutenir la gestion durable des forêts boréales.

## ANNEXE A – COMPOSITION SPÉCIFIQUE SELON LE TYPE DE COUVERT ET L'ANNÉE D'ORIGINE DU FEU

Pourcentage de surface terrière par essence pour deux années d'origine (1760, 1923) et trois types de couverts (résineux, feuillus, mixtes). Les peuplements résineux sont dominés par l'épinette noire et le sapin baumier, tandis que les feuillus présentent une forte proportion de bouleau blanc et de peuplier faux-tremble. Les formations mixtes montrent une composition plus équilibrée entre essences.



## ANNEXE B – DESCRIPTION DES VARIABLES EXPLICATIVES ET NOMENCLATURE DES MÉTRIQUES (MHC, MNT, PENTE, TPI, TWI)

### a. Variables dérivées du MHC (CHM)

Variable	Description
mean_mhc	Hauteur moyenne des cellules > 2 m
p05_mhc	5e percentile des hauteurs de cellules > 2 m
p25_mhc	25e percentile des hauteurs de cellules > 2 m
p50_mhc	50e percentile (médiane) des hauteurs de cellules > 2 m
p75_mhc	75e percentile des hauteurs de cellules > 2 m
p90_mhc	90e percentile des hauteurs de cellules > 2 m
p95_mhc	95e percentile des hauteurs de cellules > 2 m
p99_mhc	99e percentile des hauteurs de cellules > 2 m
iqr_mhc	Intervalle interquartile des hauteurs de cellules > 2 m
sd_mhc	Écart type des hauteurs de cellules > 2 m
var_mhc	Variance des hauteurs de cellules > 2 m
cv_mhc	Coefficient de variation des hauteurs de cellules > 2 m
skewness_mhc	Coefficient d'asymétrie des hauteurs de cellules > 2 m
kurtosis_mhc	Coefficient d'aplatissement (kurtosis) des hauteurs de cellules > 2 m
perc_less015	Pourcentage de hauteurs de cellules < 0,15 m
perc_015_137	Pourcentage de hauteurs de cellules > 0,15 m et < 1,37 m
perc_137_5	Pourcentage de hauteurs de cellules > 1,37 m et < 5 m
perc_5_10	Pourcentage de hauteurs de cellules > 5 m et < 10 m
perc_10_20	Pourcentage de hauteurs de cellules > 10 m et < 20 m
perc_up_20	Pourcentage de hauteurs de cellules > 20 m
vci_mhc	Indice de complexité verticale
crr_mhc	Rapport de relief de la canopée
ccmean_mhc	Pourcentage de couvert à la hauteur moyenne de canopée
ccmode_mhc	Ratio moyenne/max (cellules > 2 m).
mhr	Ratio moyenne/max (cellules ≥ 0 m).
rch	Ratio moyenne/max (cellules > 2 m).

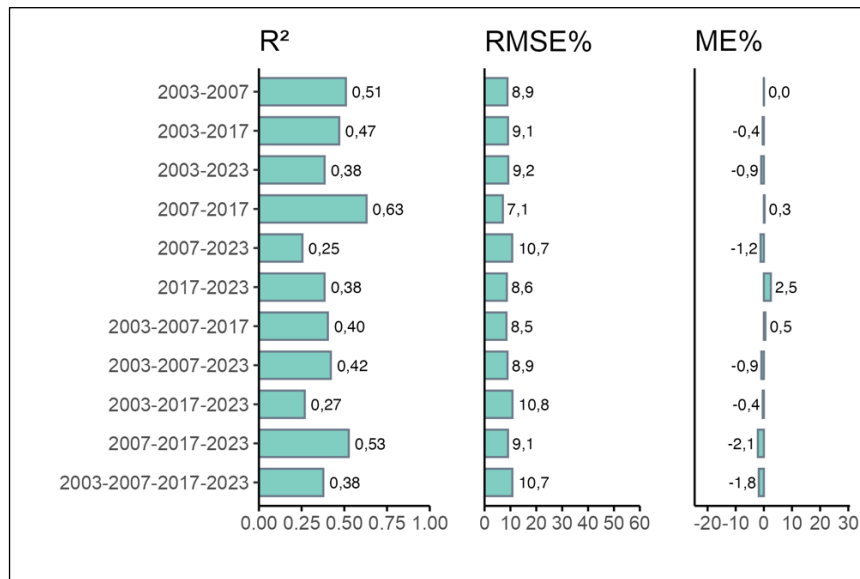
## b. Variables dérivées du MNT, Pente, TPI et TWI

Variable	Description
p05_mnt(elev), p05_slope, p05_tpi, p05_twi	5e percentile
p25_mnt(elev), p25_slope, p25_tpi, p25_twi	25e percentile
p50_mnt (elev), p50_slope, p50_tpi, p50_twi	50e percentile
p75_mnt (elev), p75_slope, p75_tpi, p75_twi	75e percentile
p90_mnt (elev), p90_slope, p90_tpi, p90_twi	90e percentile
p95_mnt (elev), p95_slope, p95_tpi, p95_twi	95e percentile
p99_mnt (elev), p99_slope, p99_tpi, p99_twi	99e percentile
iqr_mnt (elev), iqr_slope, iqr_tpi, iqr_twi	Intervalle interquartile
sd_mnt (elev), sd_slope, sd_tpi, sd_twi	Écart type
Var mnt(elev), var_slope, var_tpi, var_twi	Variance
skewness_dtm, skewness_slope, skewness_tpi, skewness_twi	Coefficient d'asymétrie
kurtosis_dtm, kurtosis_slope, kurtosis_tpi, kurtosis_twi	Coefficient d'aplatissement
cv_dtm, cv_slope, cv_tpi, cv_twi	Coefficient de variation

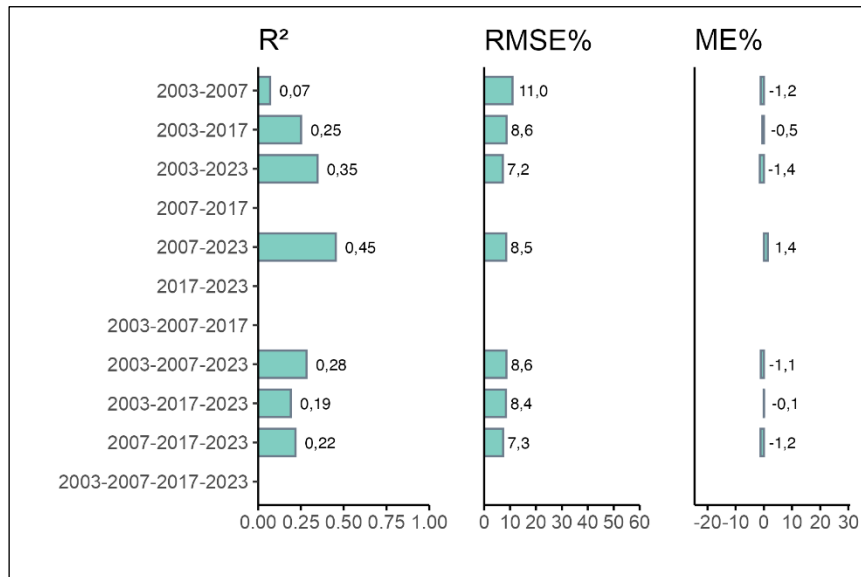
### ANNEXE C – PERFORMANCE DES MODÈLES DE RÉGRESSION LINÉAIRE POUR LA PRÉDICTION DE LA HAUTEUR DE LOREY SELON LES TYPES DE COUVERT FORESTIER

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

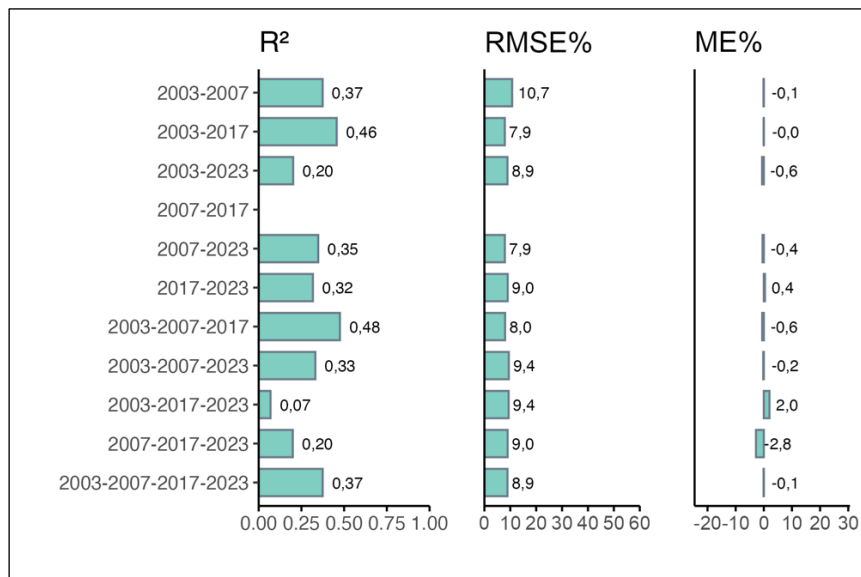
#### a. Couvert résineux



#### b. Couvert feuillu

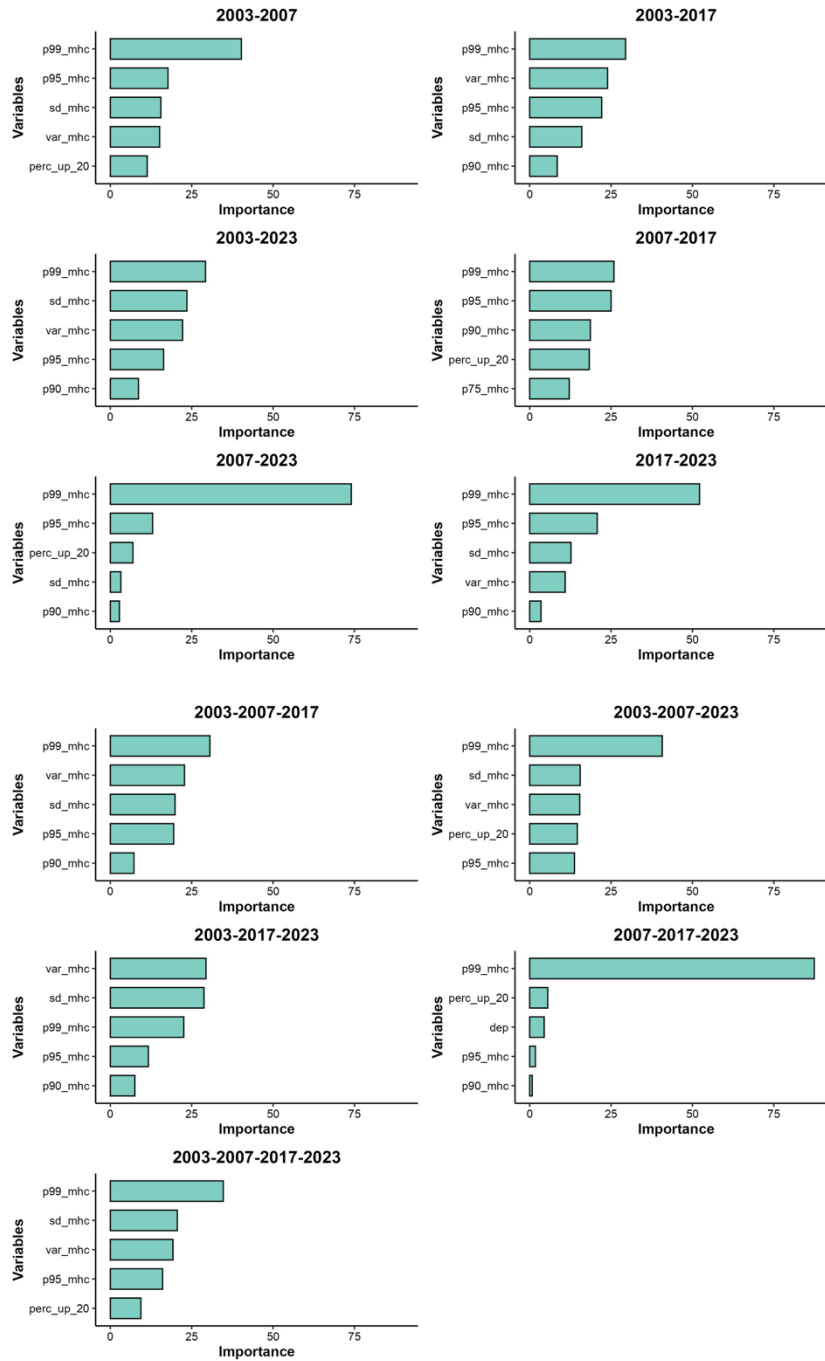


### c. Couvert mixte

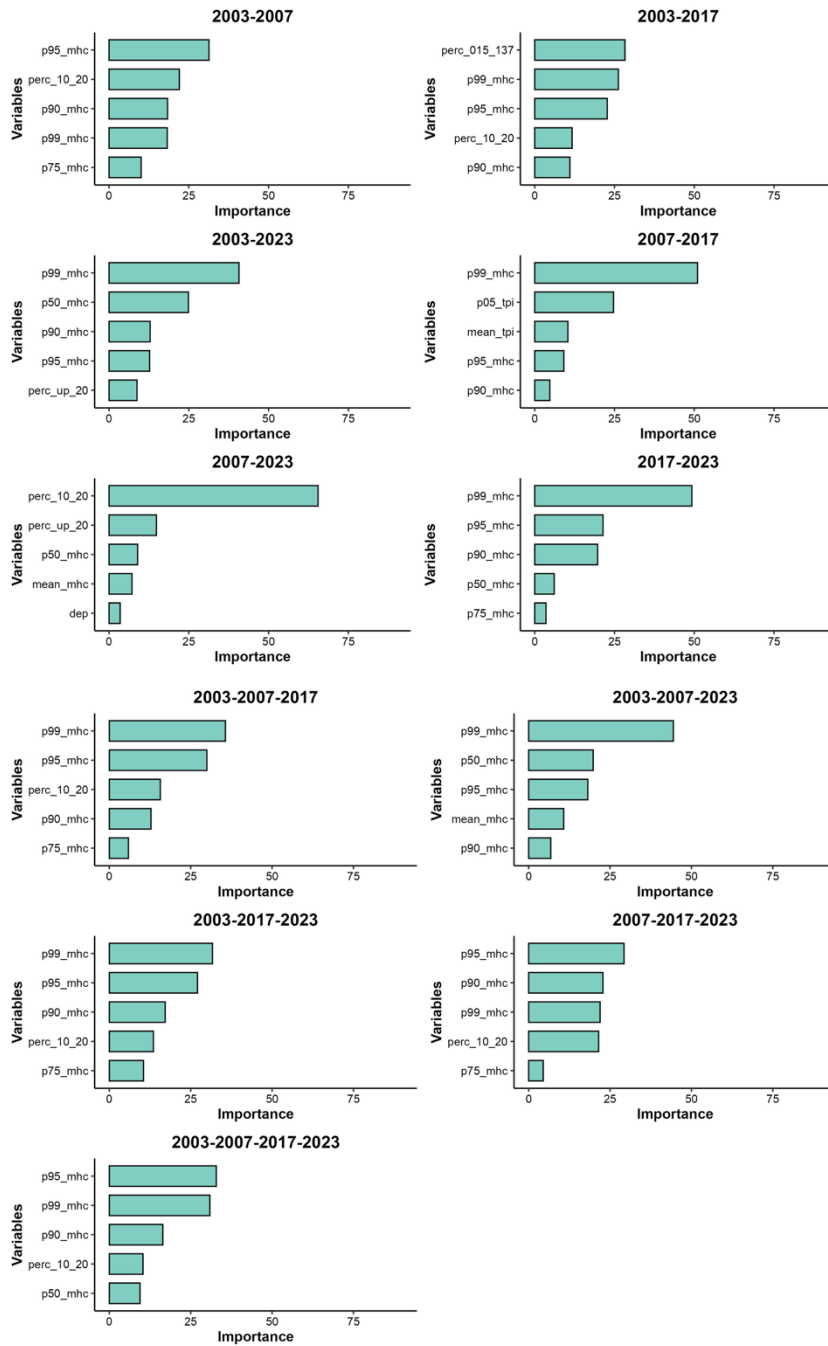


## ANNEXE D – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA HAUTEUR DE LOREY, SELON LES DIFFÉRENTES COMBINAISONS TEMPORELLES

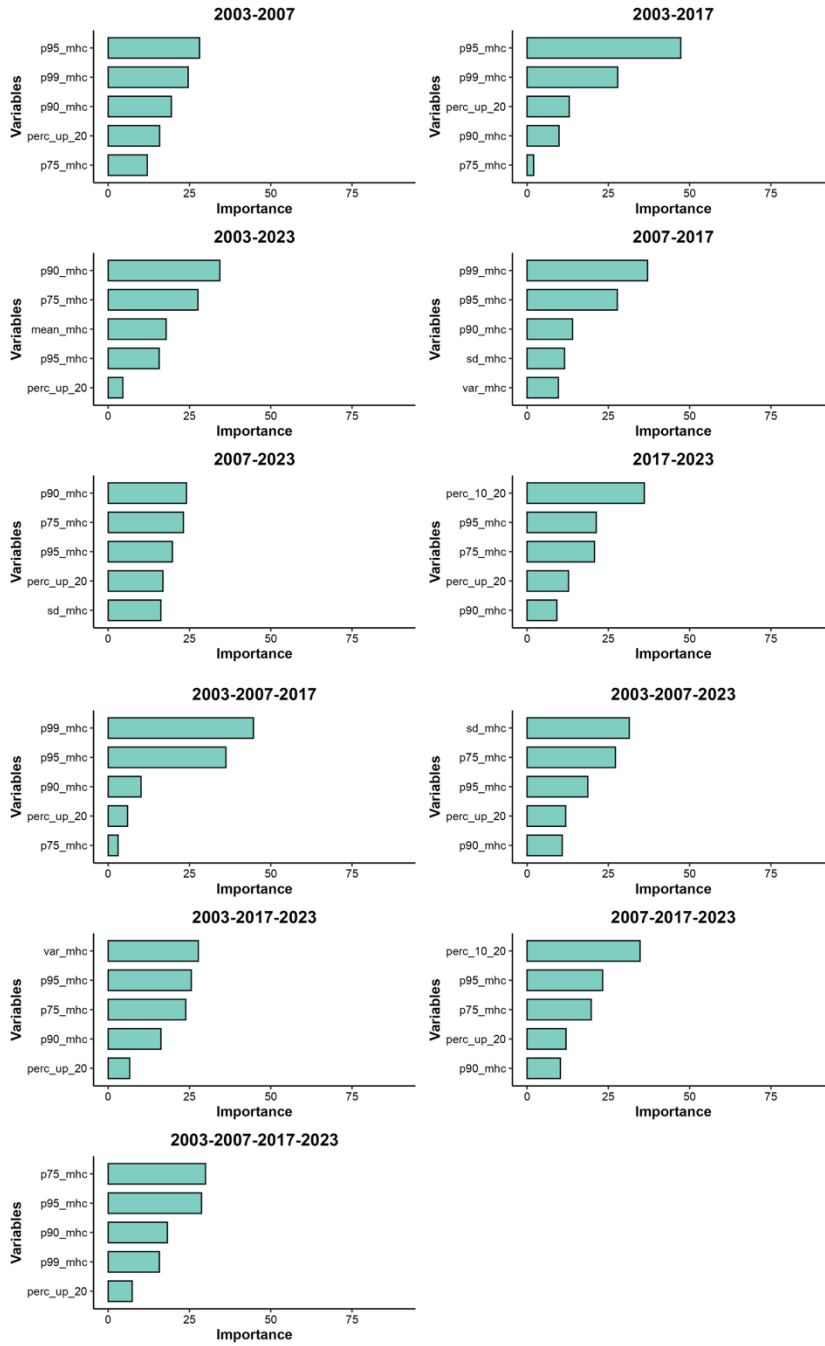
### a. Couvert résineux



## b. Couvert feuillu



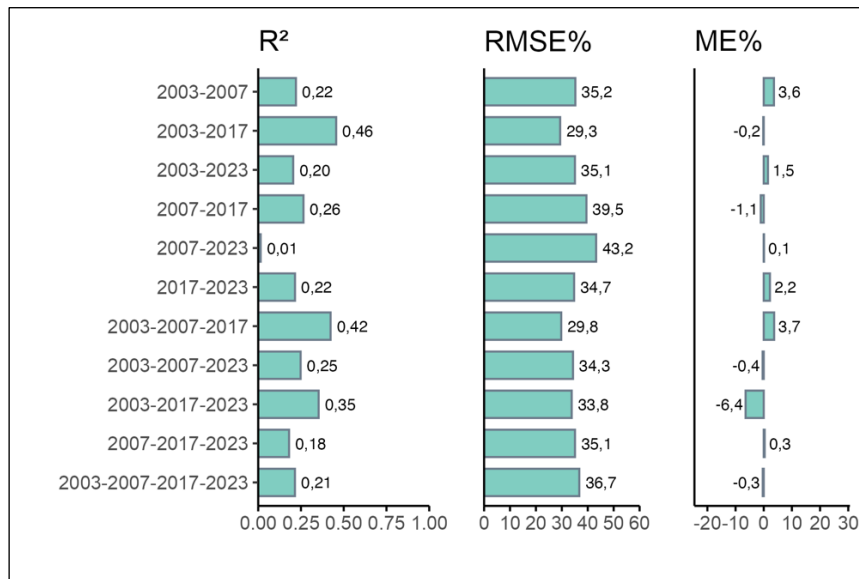
c. Couvert mixte



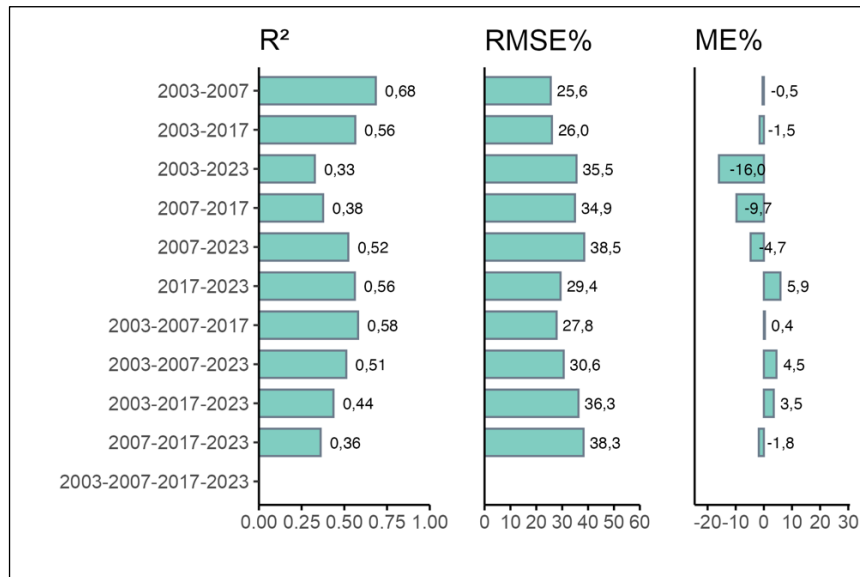
## ANNEXE E – PERFORMANCE DES MODÈLES DE RÉGRESSION LINÉAIRE DANS POUR LA PRÉDICTION DE LA SURFACE TERRIÈRE SELON LES TYPES DE COUVERT FORESTIER

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

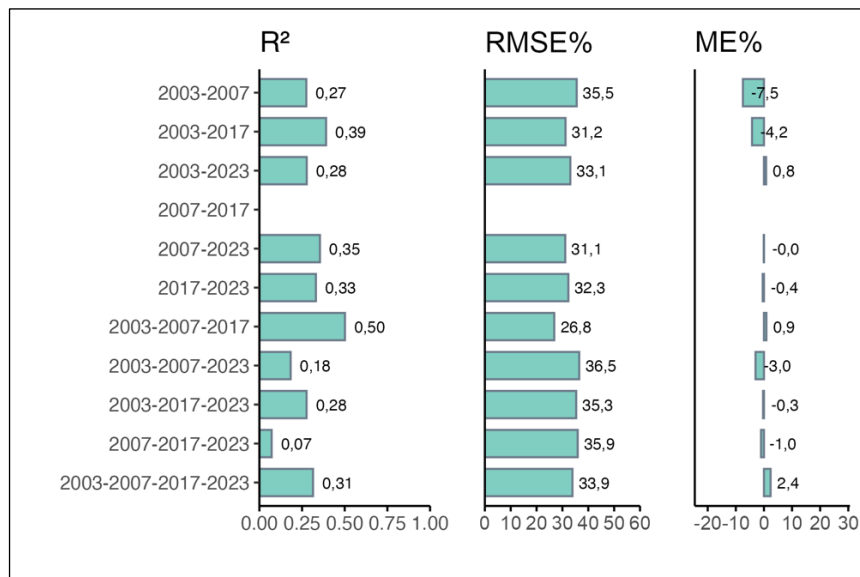
### a. Couvert résineux



### b. Couvert feuillu

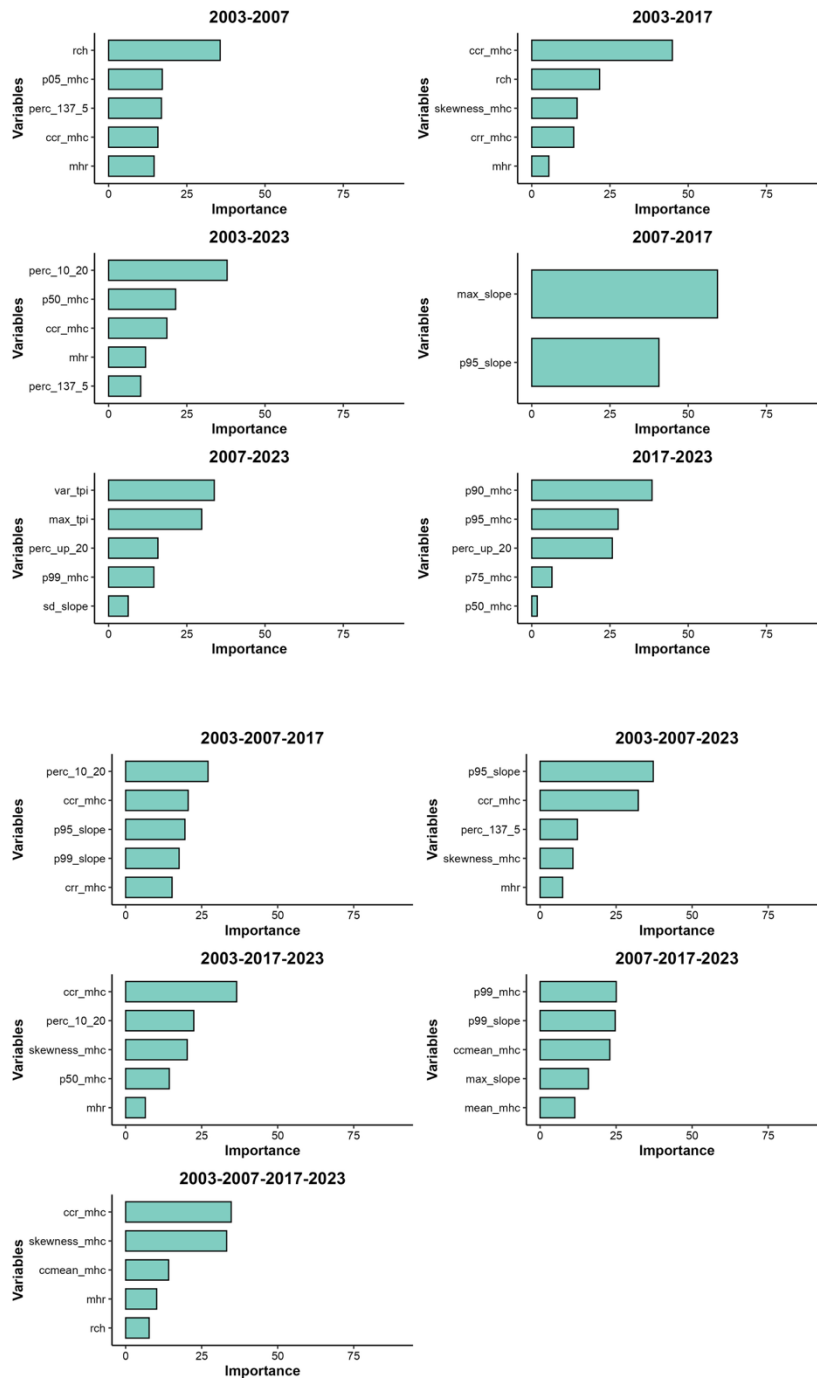


## c. Couvert mixte

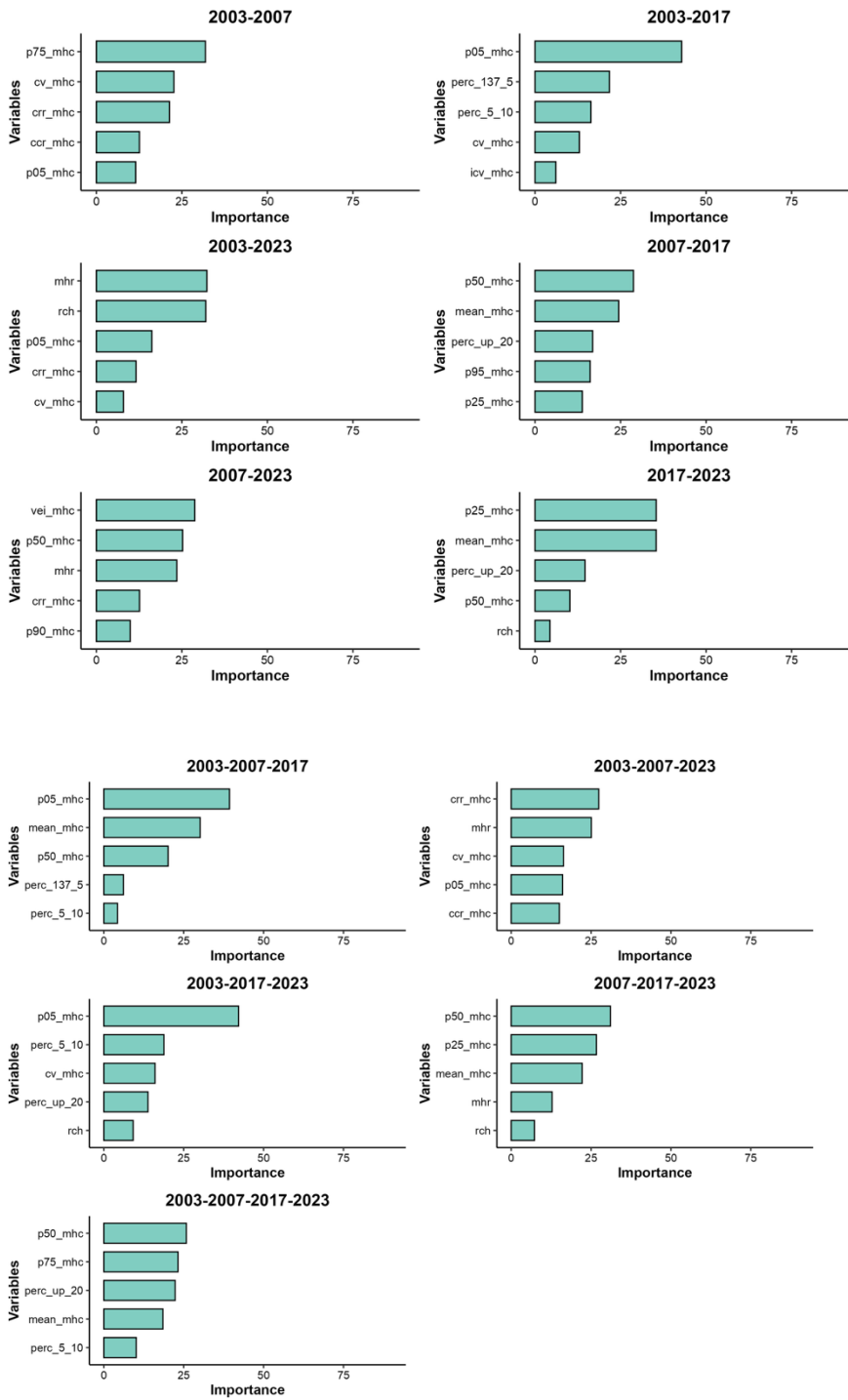


## ANNEXE F – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA SURFACE TERRIÈRE, SELON LES DIFFÉRENTES COMBINAISONS TEMPORELLES

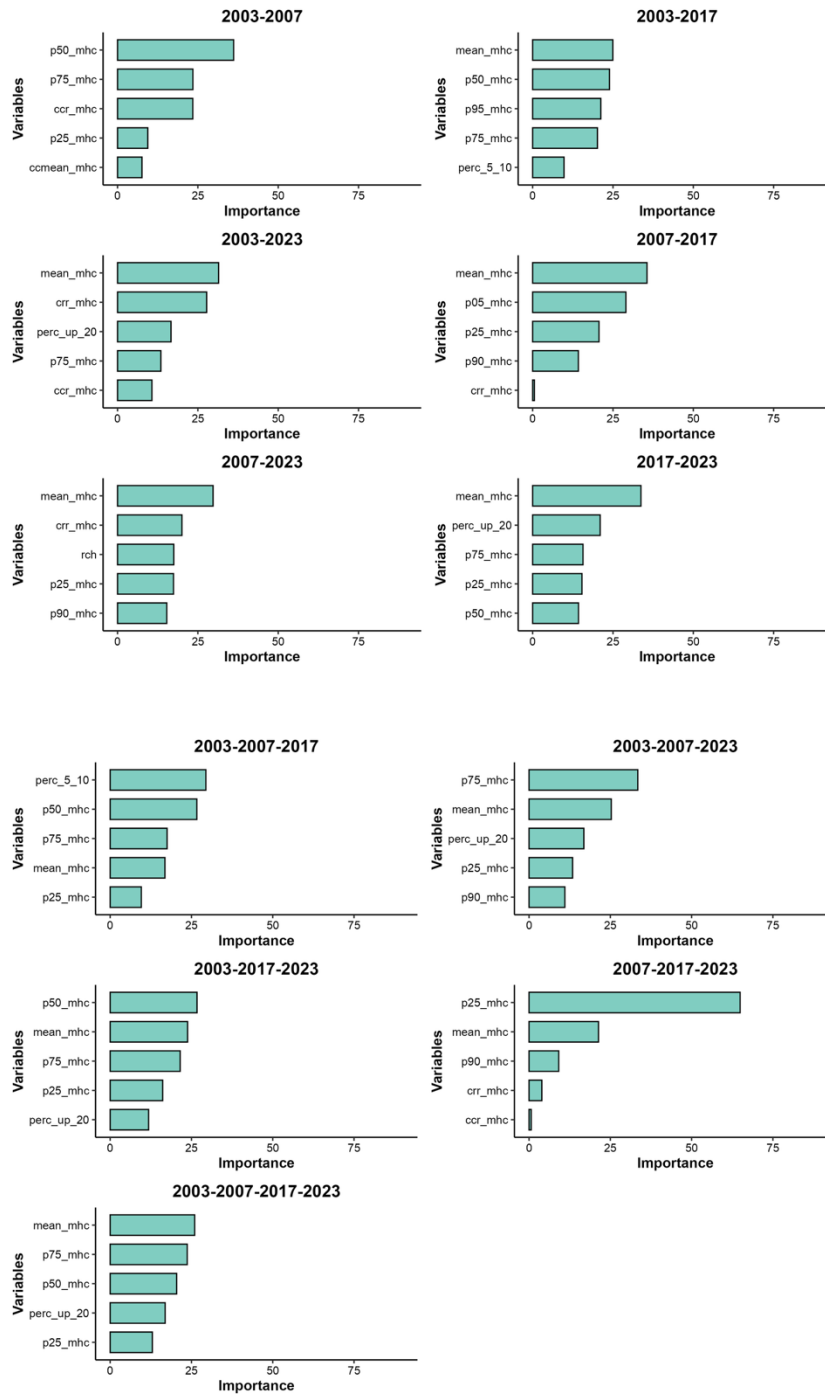
### a. Couvert résineux



## b. Couvert feuillu



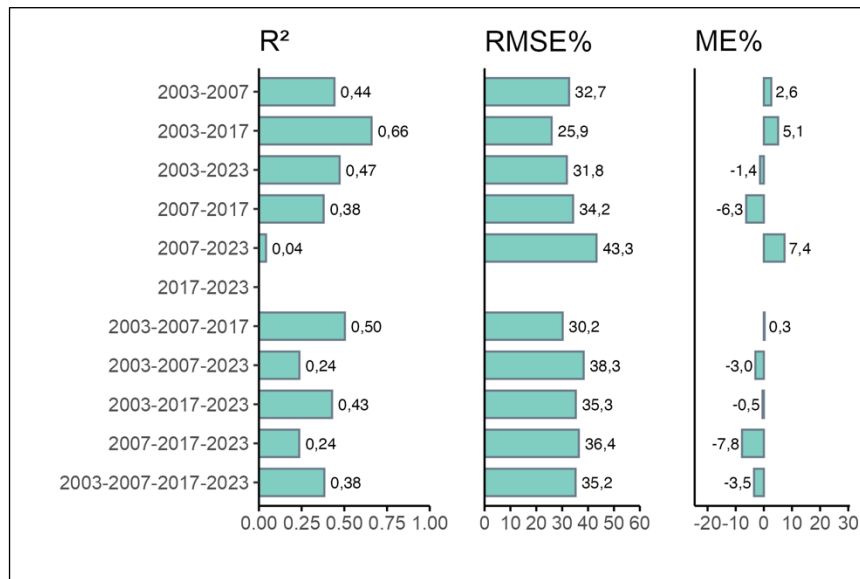
## c. Couvert mixte



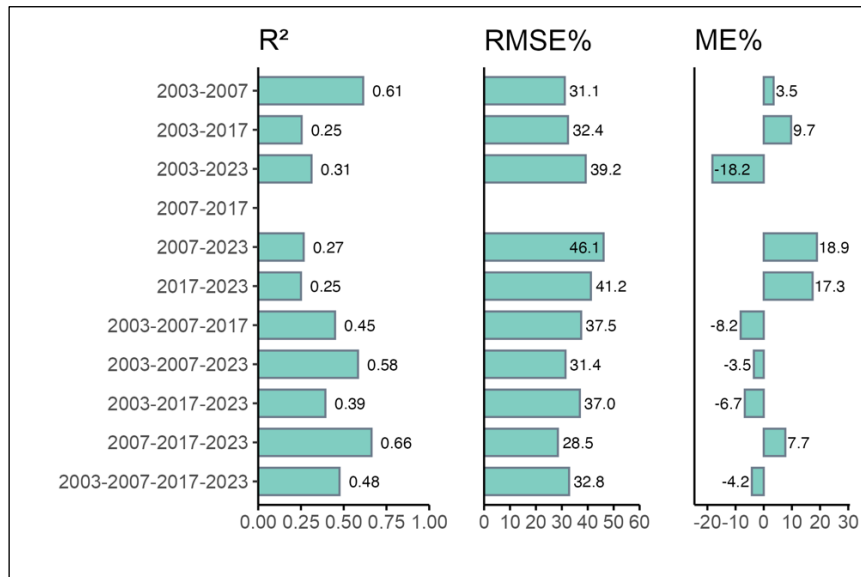
## ANNEXE G – PERFORMANCE DES MODÈLES DE RÉGRESSION LINÉAIRE POUR LA PRÉDICTION DU VOLUME MARCHAND BRUT SELON LES TYPES DE COUVERT FORESTIER

Indicateurs de performance : coefficient de détermination ( $R^2$ ), erreur quadratique moyenne relative (RMSE%) et erreur moyenne relative (ME%), selon chaque combinaison de périodes d'acquisition ALS.

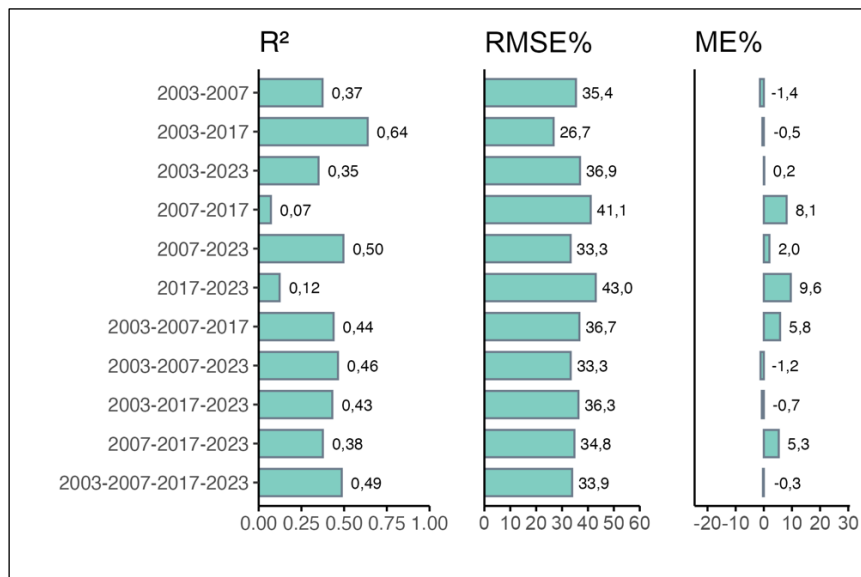
### a. Couvert résineux



### b. Couvert feuillu

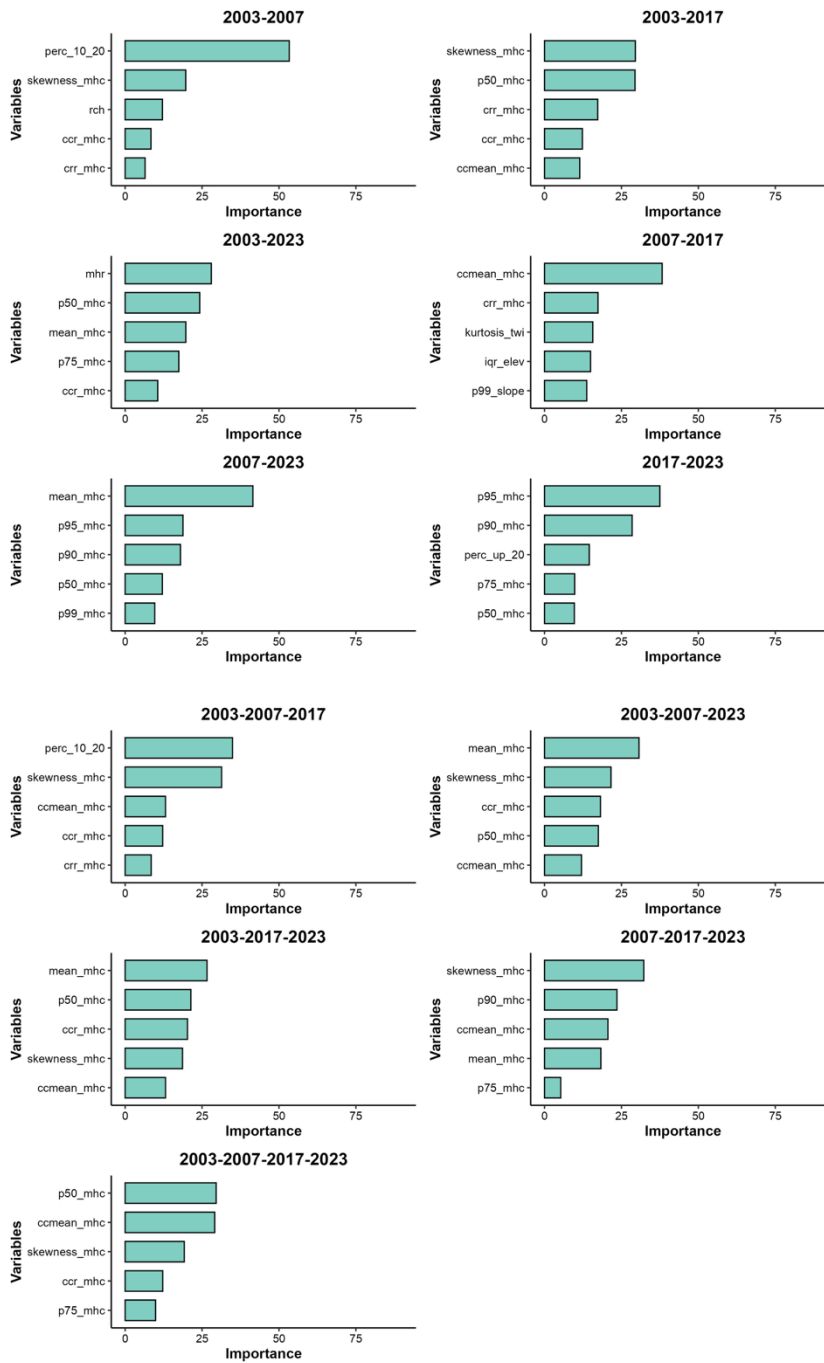


### c. Couvert mixte

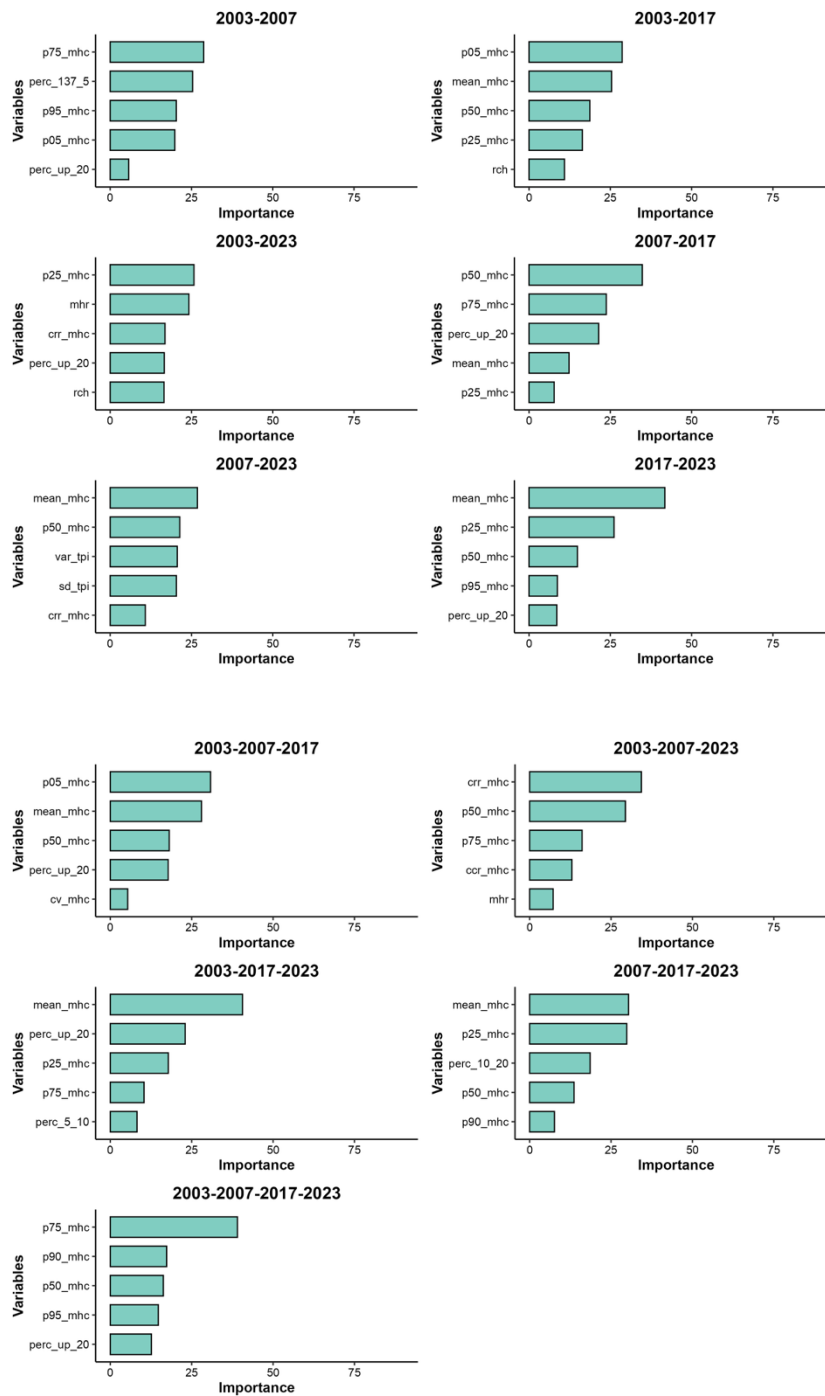


## ANNEXE H – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DU VOLUME MARCHAND BRUT, SELON LES DIFFÉRENTES COMBINAISONS TEMPORELLES

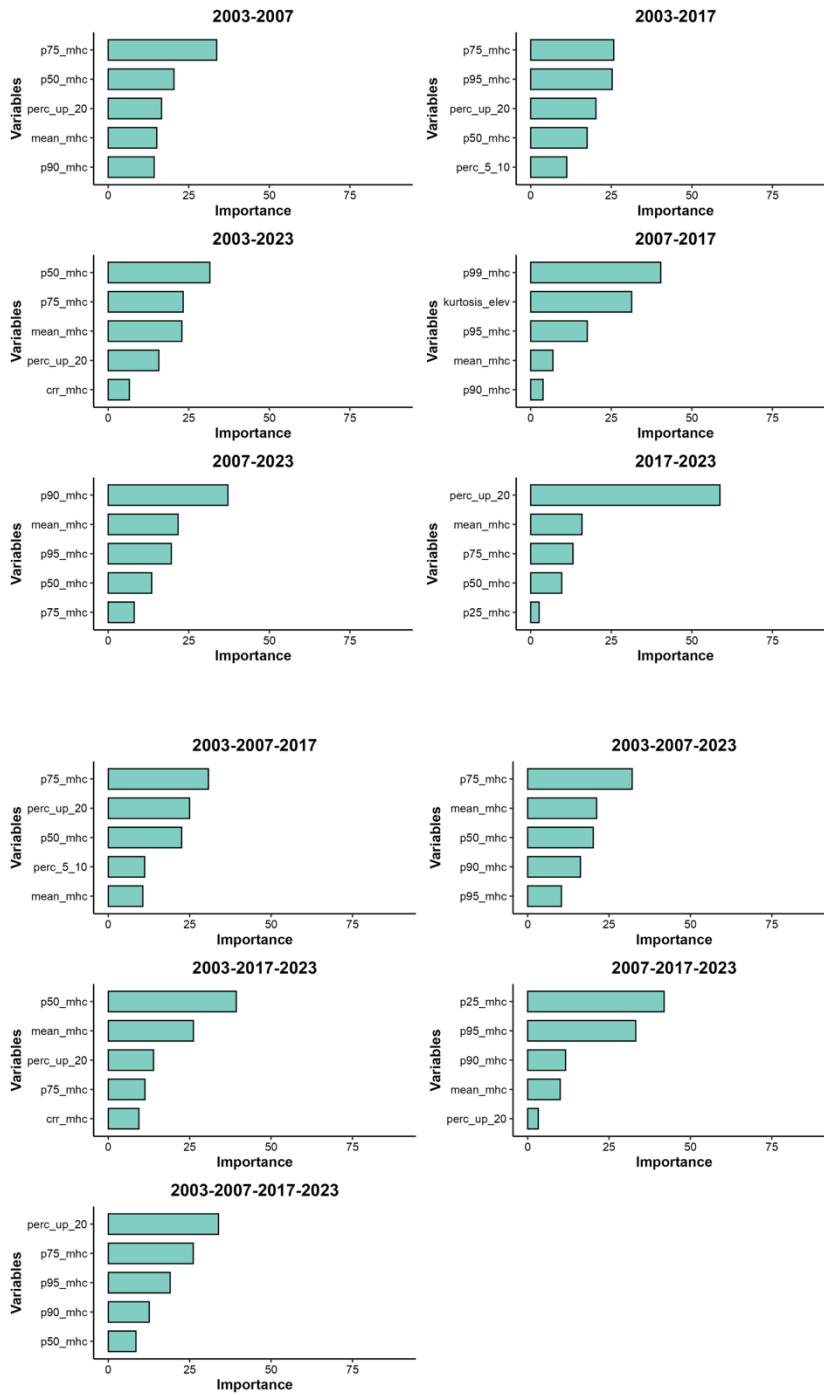
### a. Couvert résineux



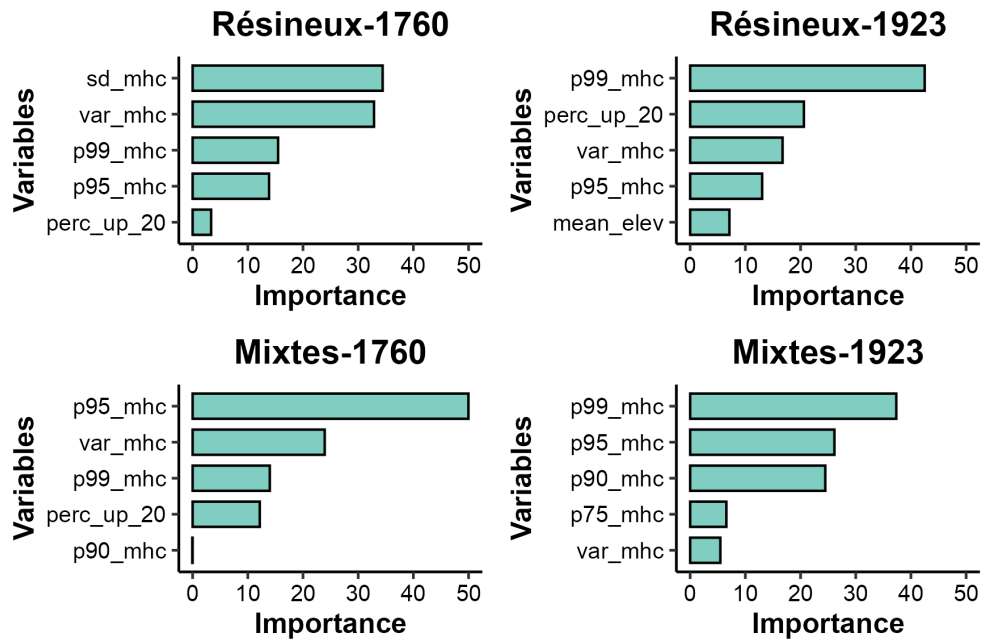
## b. Couvert feuillu

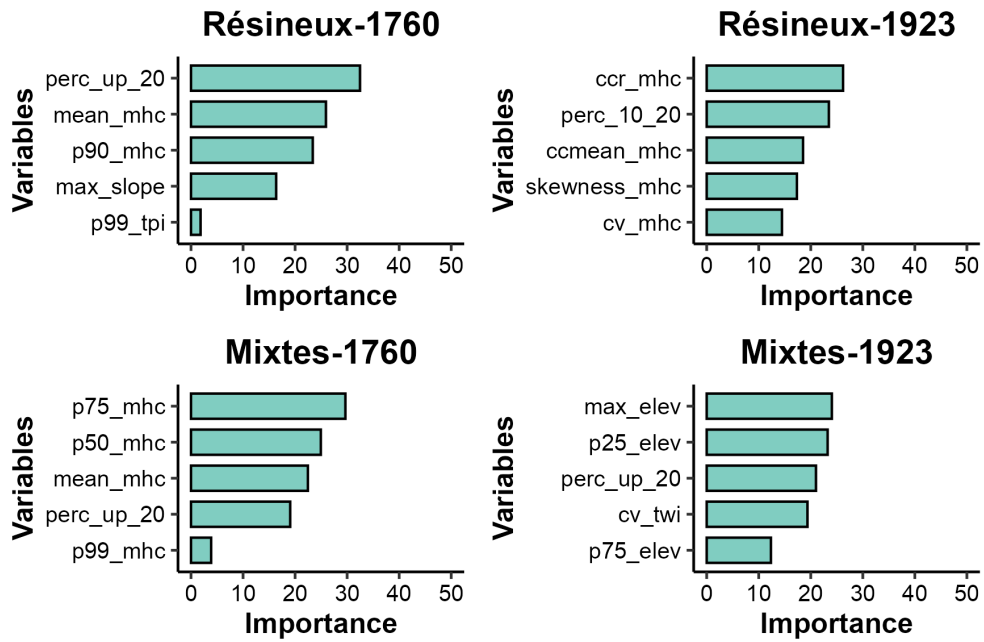


## c. Couvert mixte

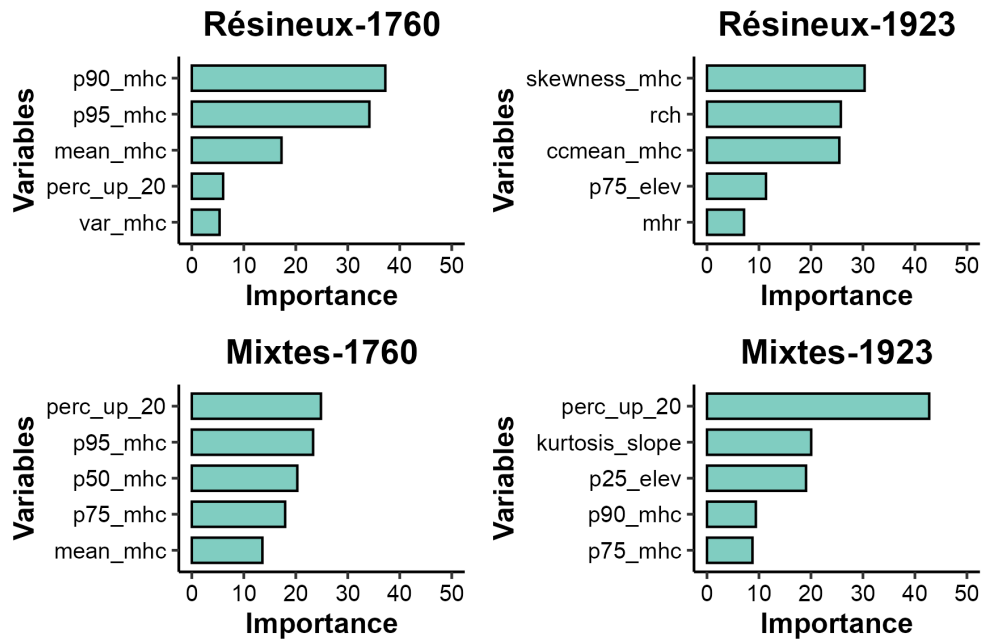


**ANNEXE I – IMPORTANCE DE VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA HAUTEUR DE LOREY ( $h_L$ )**



**ANNEXE J – IMPORTANCE DES VARIABLES (RANDOM FOREST) POUR LA PRÉDICTION DE LA SURFACE TERRIÈRE (BA)**

## ANNEXE K – IMPORTANCE DE VARIABLES (RANDOM FOREST) POUR LA PREDICTION DE VOLUME MARCHAND BRUT (V)



### LISTE DE REFERENCES

- Bailey, R. G. (1983, 1983/07/01). Delineation of ecosystem regions. *Environmental Management*, 7(4), 365-373. <https://doi.org/10.1007/BF01866919>
- Barnes, C., Boulanger, Y., Keeping, T., Gachon, P., Gillett, N., Haas, O., Wang, X., Roberge, F., Kew, S. et Heinrich, D. (2023). Climate change more than doubled the likelihood of extreme fire weather conditions in Eastern Canada.
- Béland, M., Parker, G., Sparrow, B., Harding, D., Chasmer, L., Phinn, S., Antonarakis, A. et Strahler, A. (2019). On promoting the use of lidar systems in forest ecosystem research. *Forest Ecology and Management*, 450, 117484.
- Bergeron, J.-F., Saucier, J.-P., Robert, D. et Robitaille, A. (1992). Québec forest ecological classification program. *The forestry chronicle*, 68(1), 53-63.
- Bergeron, Y. (1998). Les conséquences des changements climatiques sur la fréquence des feux et la composition forestière au sud-ouest de la forêt boréale québécoise. *Géographie physique et Quaternaire*, 52(2), 167-174.
- Bergeron, Y. (2000). *Ecology*, 81(6), 1500-1516.
- Bergeron, Y., Cyr, D., Drever, C. R., Flannigan, M., Gauthier, S., Kneeshaw, D., Lauzon, É., Leduc, A., Goff, H. L., Lesieur, D. et Logan, K. (2006). Past, current, and future fire frequencies in Quebec's commercial forests: implications for the cumulative effects of harvesting and fire on age-class structure and natural disturbance-based management. *Canadian Journal of Forest Research*, 36(11), 2737-2744. <https://doi.org/10.1139/x06-177>
- Bergeron, Y. et Dansereau, P.-R. (1993). Predicting the composition of Canadian southern boreal forest in different fire cycles. *Journal of Vegetation Science*, 4(6), 827-832. <https://doi.org/https://doi.org/10.2307/3235621>
- Bergeron, Y. et Dubue, M. (1988). Succession in the southern part of the Canadian boreal forest. *Vegetatio*, 79, 51-63.
- Bergeron, Y. et Fenton, N. J. (2012). Boreal forests of eastern Canada revisited: old growth, nonfire disturbances, forest succession, and biodiversity. *Botany*, 90(6), 509-523.
- Bergeron, Y., Gauthier, S., Flannigan, M. et Kafka, V. (2004). Fire regimes at the transition between mixedwood and coniferous boreal forest in northwestern Quebec. *Ecology*, 85(7), 1916-1932.

- Bergeron, Y., Gauthier, S., Kafka, V., Lefort, P. et Lesieur, D. (2001). Natural fire frequency for the eastern Canadian boreal forest: consequences for sustainable forestry. *Canadian Journal of Forest Research*, 31(3), 384-391.
- Blouin, J. (2002). *Guide de reconnaissance des types écologiques. Région écologique 5a - Plaine de l'Abitibi*.  
<https://numerique.banq.qc.ca/patrimoine/details/52327/2739120>
- Bouvier, M., Durrieu, S., Fournier, R. A. et Renaud, J.-P. (2015). Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote sensing of environment*, 156, 322-334.
- Breiman, L. (2001, 2001/10/01). Random Forests. *Machine Learning*, 45(1), 5-32.  
<https://doi.org/10.1023/A:1010933404324>
- Canadian Council of Forest Ministers. (2023). *Canada's National Forestry Database*  
<http://nfdp.ccfm.org/en/data/fires.php>
- Cosenza, D. N., Korhonen, L., Maltamo, M., Packalen, P., Strunk, J. L., Næsset, E., Gobakken, T., Soares, P. et Tomé, M. (2021). *Comparison of linear regression, k-nearest neighbour and random forest methods in airborne laser-scanning-based prediction of growing stock* (vol. 94, pp. 311-311-323).
- Direction des inventaires forestiers. (1975). *Norme d'inventaire forestier (premier inventaire)*.
- Direction des inventaires forestiers. (1984). *Norme d'inventaire forestier (deuxième inventaire)*.
- Direction des inventaires forestiers. (2009). *Norme de cartographie écoforestière – Troisième inventaire écoforestier*. Québec: Ministère des Ressources naturelles et de la Faune.
- Direction des inventaires forestiers. (2021). *Indice d'humidité topographique*: Ministère des Ressources naturelles et des Forêts.
- Direction des inventaires forestiers. (2024a). *Carte écoforestière à jour*. Dans M. d. R. n. e. d. Forêts (dir.), (p. 6): Gouvernement du Québec,.
- Direction des inventaires forestiers. (2024b). *Cinquième inventaire écoforestier du Québec meridional. Disponibilité de la Carte écoforestière*. Dans M. d. R. n. e. d. Forêts (dir.): Gouvernement du Québec,.
- Dong, P. et Chen, Q. (2017). *LiDAR remote sensing and applications*. CRC Press.
- Dubayah, R. O. et Drake, J. B. (2000). Lidar remote sensing for forestry. *Journal of forestry*, 98(6), 44-46.

- Efron, B. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC.
- Esri. (2024). *ArcGIS Pro (Version 3.3.0)*. Dans Environmental Systems Research Institute. <https://www.esri.com/en-us/arcgis/products/arcgis-pro>
- Fekety, P. A., Falkowski, M. J. et Hudak, A. T. (2015). Temporal transferability of LiDAR-based imputation of forest inventory attributes. *Canadian Journal of Forest Research*, 45(4), 422-435.
- Fekety, P. A., Falkowski, M. J., Hudak, A. T., Jain, T. B. et Evans, J. S. (2018). Transferability of lidar-derived basal area and stem density models within a northern Idaho ecoregion. *Canadian Journal of Remote Sensing*, 44(2), 131-143.
- Forêts, M. d. R. n. e. d. (2025). Insectes, Maladies et Feux Dans les Forêts du Québec en 2024.
- Gareth, J., Daniela, W., Trevor, H. et Robert, T. (2013). *An introduction to statistical learning: with applications in R*. Springer.
- Gauthier, S., De Grandpré, L. et Bergeron, Y. (2000). Differences in forest composition in two boreal forest ecoregions of Quebec. *Journal of Vegetation Science*, 11(6), 781-790.
- Gobakken, T. et Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 38(5), 1095-1109.
- Goodbody, T. R. H., Coops, N. C., Irwin, L. A. K., Armour, C. C., Saunders, S. C., Dykstra, P., Butson, C. et Perkins, G. C. (2024). Integration of Airborne Laser Scanning data into forest ecosystem management in Canada: Current status and future directions. *The forestry chronicle*, 1-21. <https://doi.org/10.5558/tfc2024-014>
- Gouvernement du Québec. (2023). Forêt ouverte. <https://www.foretouverte.gouv.qc.ca/>
- Gouvernement du Québec. (2024). Ressources et industries forestières du Québec, Portrait statistique 2023. Dans M. d. R. n. e. d. Forêts (dir.), (29 e édition, 2024 éd.). Québec: Ministère des Ressources naturelles et des Forêts.
- Gouvernement du Québec. (2025a). Inventaire écoforestier. <https://www.quebec.ca/agriculture-environnement-et-ressources-naturelles/forets/recherche-connaissances/inventaire-forestier/types/quebec-meridional>

- Gouvernement du Québec. (2025b). Technologies d'inventaire des forêts. Récupéré le 2025-03-15 de <https://www.quebec.ca/agriculture-environnement-et-ressources-naturelles/forets/recherche-connaissances/inventaire-forestier/technologies-utilisees#c228520>
- Grondin, P. (1999). Programme de connaissance des écosystèmes forestiers du Québec méridional. Rapport de classification écologique. Sapinière à bouleau blanc de l'ouest. <https://numerique.banq.qc.ca/patrimoine/details/52327/2748312>
- Group, P. G. D. (2024). *PostgreSQL (Version 16.2)*. Dans PostgreSQL Global Development Group. <https://www.postgresql.org>
- Harvey, B. (1999). The Lake Duparquet research and teaching forest: building a foundation for ecosystem management. *The forestry chronicle*, 75(3), 389-393.
- Hastie, T., Tibshirani, R., Friedman, J. H. et Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (vol. 2). Springer.
- Hopkinson, C., Chasmer, L. et Hall, R. (2008). The uncertainty in conifer plantation growth prediction from multi-temporal lidar datasets. *Remote sensing of environment*, 112(3), 1168-1180.
- Leboeuf, A., Riopel, M., Munger, D., Fradette, M.-S. et Bégin, J. (2022, 06/23). Modeling Merchantable Wood Volume Using Airborne LiDAR Metrics and Historical Forest Inventory Plots at a Provincial Scale. *Forests*, 13, 985. <https://doi.org/10.3390/f13070985>
- Lecomte, N., Simard, M. et Bergeron, Y. (2006). Effects of fire severity and initial tree composition on stand structural development in the coniferous boreal forest of northwestern Québec, Canada. *Ecoscience*, 13(2), 152-163.
- Lefsky, M. A., Cohen, W., Acker, S., Parker, G. G., Spies, T. et Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote sensing of environment*, 70(3), 339-361.
- Lefsky, M. A., Cohen, W. B., Parker, G. G. et Harding, D. J. (2002). Lidar remote sensing for ecosystem studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *BioScience*, 52(1), 19-30.
- Lim, K., Treitz, P., Wulder, M., St-Onge, B. et Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in physical geography*, 27(1), 88-106.

- Maltamo, M., Eerikäinen, K., Packalén, P. et Hyypä, J. (2006). Estimation of stem volume using laser scanning-based canopy height metrics. *Forestry*, 79(2), 217-229.
- Martin, M., Cerrejón, C. et Valeria, O. (2021, 2021/12/15/). Complementary airborne LiDAR and satellite indices are reliable predictors of disturbance-induced structural diversity in mixed old-growth forest landscapes. *Remote sensing of environment*, 267, 112746.  
<https://doi.org/https://doi.org/10.1016/j.rse.2021.112746>
- Mauro, F., Hudak, A. T., Fekety, P. A., Frank, B., Temesgen, H., Bell, D. M., Gregory, M. J. et McCarley, T. R. (2021). Regional Modeling of Forest Fuels and Structural Attributes Using Airborne Laser Scanning Data in Oregon. *Remote Sensing*, 13(2), 261. <https://www.mdpi.com/2072-4292/13/2/261>
- McRoberts, R. E., Chen, Q. et Walters, B. F. (2017). Multivariate inference for forest inventories using auxiliary airborne laser scanning data. *Forest Ecology and Management*, 401, 295-303.
- Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L. et Hendrix, C. J. (2000). Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering and Remote Sensing*, 66(11), 1367-1372.
- Mehtätalo, L., de-Miguel, S. et Gregoire, T. G. (2015). Modeling height-diameter curves for prediction. *Canadian Journal of Forest Research*, 45(7), 826-837.
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote sensing of environment*, 80(1), 88-99.
- Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19(2), 164-179.
- Næsset, E. et Bjerknes, K.-O. (2001). Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote sensing of environment*, 78(3), 328-340.
- Næsset, E. et Gobakken, T. (2005, 2005/06/30/). Estimating forest growth using canopy metrics derived from airborne laser scanner data. *Remote sensing of environment*, 96(3), 453-465.  
<https://doi.org/https://doi.org/10.1016/j.rse.2005.04.001>
- Nelson, R. (2013). How did we get here? An early history of forestry lidar. *Canadian Journal of Remote Sensing*, 39(sup1), S6-S17.

- Nelson, R., Valenti, M. A., Short, A. et Keller, C. (2003). A multiple resource inventory of Delaware using airborne laser data. *BioScience*, 53(10), 981-992.
- Oliver, C., Larson, B. et Ball, J. (1997). Forest Stand Dynamics. Update Edition. *Journal of Natural Resources and Life Sciences Education*, 26(1), 81.
- Penner, M., Pitt, D. et Woods, M. (2013). Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. *Canadian Journal of Remote Sensing*, 39(5), 426-443.
- Socha, J., Pierzchalski, M., Bałazy, R. et Ciesielski, M. (2017). Modelling top height growth and site index using repeated laser scanning data. *Forest Ecology and Management*, 406, 307-317.
- Sørensen, R., Zinko, U. et Seibert, J. (2006). On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrol. Earth Syst. Sci.*, 10(1), 101-112. <https://doi.org/10.5194/hess-10-101-2006>
- St-Onge, B. et Vepakomma, U. (2004). Assessing forest gap dynamics and growth using multi-temporal laser-scanner data. *Power*, 140, 173-178.
- Stinson, G. et White, J. (2018, 01/01). What's the difference between EFI and NFI? Demystifying current acronyms in forest inventory in Canada. *BC Forest Professional*, Jan/Feb 2018, 10-11.
- Strunk, J., Temesgen, H., Andersen, H.-E., Flewelling, J. P. et Madsen, L. (2012, 2012/11/20). Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing*, 38(5), 644-654. <https://doi.org/10.5589/m12-052>
- Thomas, J. A. et Bovee, K. D. (1993). Application and testing of a procedure to evaluate transferability of habitat suitability criteria. *Regulated Rivers: Research & Management*, 8(3), 285-294.
- Tompalski, P., Coops, N. C., Marshall, P. L., White, J. C., Wulder, M. A. et Bailey, T. (2018). Combining Multi-Date Airborne Laser Scanning and Digital Aerial Photogrammetric Data for Forest Growth and Yield Modelling. *Remote Sensing*, 10(2), 347. <https://www.mdpi.com/2072-4292/10/2/347>
- Tompalski, P., Rakofsky, J., Coops, N. C., White, J. C., Graham, A. N. et Rosychuk, K. (2019a). Challenges of multi-temporal and multi-sensor forest growth analyses in a highly disturbed boreal mixedwood forests. *Remote Sensing*, 11(18), 2102.

- Tompalski, P., White, J., Coops, N. et Wulder, M. (2019b). Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data. *Remote sensing of environment*, 227, 110-124.
- Treitz, P., Lim, K., Woods, M., Pitt, D., Nesbitt, D. et Etheridge, D. (2012). LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sensing*, 4(4), 830-848.
- Véga, C. et St-Onge, B. (2008). Height growth reconstruction of a boreal forest canopy over a period of 58 years using a combination of photogrammetric and lidar models. *Remote sensing of environment*, 112(4), 1784-1794.
- Vepakomma, U., St-Onge, B. et Kneeshaw, D. (2008). Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote sensing of environment*, 112(5), 2326-2340.
- Weiskittel, A. R., Hann, D. W., Kershaw, J. A., Jr. et Vanclay, J. K. (2011). *Forest Growth and Yield Modeling*. John Wiley & Sons, Incorporated.  
<http://ebookcentral.proquest.com/lib/udec-ebooks/detail.action?docID=819232>
- Weiss, A. (2001). Topographic position and landforms analysis. Poster presentation, ESRI user conference, San Diego, CA,
- White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T. et Tompalski, P. (2016). Remote sensing technologies for enhancing forest inventories: A review. *Canadian Journal of Remote Sensing*, 42(5), 619-641.
- White, J. C., Tompalski, P., Bater, C. W., Wulder, M. A., Fortin, M., Hennigar, C., Robere-McGugan, G., Sinclair, I. et White, R. (2025). Enhanced forest inventories in Canada: implementation, status, and research needs. *Canadian Journal of Forest Research*, 55, 1-37.
- White, J. C., Tompalski, P., Vastaranta, M. A., Wulder, M. A., Saarinen, N. P., Stepper, C. et Coops, N. C. (2017). A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach.
- White, J. C., Wulder, M. A., Varhola, A., Vastaranta, M., Coops, N. C., Cook, B. D., Pitt, D. et Woods, M. (2013a). A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *The forestry chronicle*, 89(6), 722-723.
- White, J. C., Wulder, M. A., Varhola, A., Vastaranta, M., Coops, N. C., Cook, B. D., Pitt, D. et Woods, M. (2013b). *A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based*

*approach*. (Information Report FI-X-10).

<http://cfs.nrcan.gc.ca/publications?id%26hairsp%3B=%26hairsp%3B34887>.

Wulder, M. A., Bater, C. W., Coops, N. C., Hilker, T. et White, J. C. (2008). The role of LiDAR in sustainable forest management. *The forestry chronicle*, 84, 807-826. <https://d1ied5g1xfqpx8.cloudfront.net/pdfs/29051.pdf>

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W. et Gobakken, T. (2012). Lidar sampling for large-area forest characterization: A review. *Remote sensing of environment*, 121, 196-209.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., Heikkinen, R. K., Heinänen, S., Jones, A. R., Krishnakumar, P. K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin, C., Mesgaran, M. B., Moreno-Amat, E., Mormede, S., Novaczek, E., Oppel, S., Ortuño Crespo, G., Peterson, A. T., Rapacciuolo, G., Roberts, J. J., Ross, R. E., Scales, K. L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L. G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M. J., Zharikov, Y., Zurell, D. et Sequeira, A. M. M. (2018). Outstanding Challenges in the Transferability of Ecological Models. *Trends in Ecology & Evolution*, 33(10), 790-802. <https://doi.org/10.1016/j.tree.2018.08.001>

Yu, X., Hyypä, J., Kaartinen, H., Maltamo, M. et Hyypä, H. (2008). Obtaining plotwise mean height and volume growth in boreal forests using multi-temporal laser surveys and various change detection techniques. *International journal of remote sensing*, 29(5), 1367-1386.