

MATHIEU FORTIN

ANALYSE DE LA DYNAMIQUE DES PEUPLEMENTS MIXTES  
DE SAPIN BAUMIER ET D'ÉPINETTE ROUGE  
APRÈS COUPE PARTIELLE

CONTRAINTES ET MÉTHODOLOGIES STATISTIQUES

Thèse présentée

à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de doctorat en sciences forestières  
pour l'obtention du grade de *Philosophiae Doctor* (Ph.D.)

Département des sciences du bois et de la forêt  
FACULTÉ DE FORESTERIE ET GÉOMATIQUE

UNIVERSITÉ LAVAL

QUÉBEC

AVRIL 2006

## Résumé

Des éléments de la dynamique des peuplements mixtes de sapin baumier et d'épinette rouge après coupe partielle ont été étudiés sur une base statistique, afin de déterminer dans quelle mesure ils pourraient contribuer au phénomène de raréfaction de l'épinette rouge. L'étude visait principalement à identifier les changements structurels et à quantifier la croissance tant à l'échelle du peuplement que des tiges individuelles. Pour ce faire, les données de deux dispositifs de suivi ont été utilisées : l'Aire d'observation de la rivière Ouareau (Parc national du Mont-Tremblant) et la Forêt expérimentale du lac Édouard (Parc national de la Mauricie). Chacun de ces dispositifs disposent d'un réseau de placettes permanentes dont les mesures s'étendent sur un horizon de plus de 50 ans après une coupe partielle.

Des méthodologies statistiques ont été proposées afin de formuler des inférences statistiques adaptées à la nature des données disponibles. L'analyse de la croissance s'est faite à partir d'un modèle linéaire incluant des effets aléatoires (à l'échelle de la tige individuelle) et d'un modèle non linéaire comportant une structure de covariance (à l'échelle du peuplement) afin de tenir compte de l'hétéroscédasticité des données et de l'autocorrélation des erreurs. Les changements structurels ont été analysés à l'aide d'un modèle linéaire généralisé puisque la variable dépendante n'était pas une variable continue, mais plutôt une fréquence. Pour tenir compte de l'effet dû aux placettes, des effets aléatoires ont été ajoutés au modèle. Finalement, une matrice de transition à deux niveaux a été construite sur la base de distributions discrètes dont les paramètres sont estimés par la méthode du maximum de vraisemblance. L'utilisation de distributions discrètes offre ainsi des résultats plus cohérents par rapport à l'approche traditionnelle.

À l'échelle des tiges individuelles, les résultats de ces analyses démontrent que les différences quant à la croissance en diamètre sont relativement faibles. Le recrutement et la mortalité sont en fait des facteurs beaucoup plus importants dans l'évolution de ces peuplements après coupe partielle. À l'échelle du peuplement, ils induisent une variabilité importante dans l'évolution de la surface terrière du sapin

baumier, de sorte que l'épinette rouge apparaît comme une essence beaucoup plus stable. La durée et l'importance de l'ouverture de la canopée après coupe partielle sont des éléments critiques qui déterminent l'abondance du recrutement du sapin baumier. Pour maintenir la proportion d'épinette rouge et une distribution diamétrale irrégulière, les coupes partielles de faible intensité comme le jardinage par pied d'arbre sont indiquées.

## Avant-propos

Cette étude s'inscrit dans la continuité de travaux de recherche entrepris au cours des années 1950 par le Service canadien des forêts (SCF). Je ne peux que vouer une profonde admiration aux chercheurs de cette époque qui ont consacré de nombreuses heures à comprendre la complexité des peuplements mixtes de sapin baumier et d'épinette rouge. Le présent ouvrage est en grande partie le fruit de leurs efforts. Merci au SCF pour le prêt des données et à ses chercheurs : MM. R.J. Hatcher, R.G. Ray et C.C. Heimbürger.

Je tiens à adresser des remerciements à M. Jean Bégin (Université Laval) et M. Louis Archambault (SCF-CFL), qui ont respectivement été le directeur et le codirecteur de cette thèse de doctorat. Un merci tout spécial à M. Chhun-Huor Ung (SCF-CFL). À eux trois, ils ont su me guider au travers de ce périple de plus de quatre ans. Merci aux coauteurs et collaborateurs qui ont participé aux manuscrits qui composent cette thèse : M. Gaétan Daigle (Université Laval), Mme Michèle Bernier-Cardou (SCF-CFL), M. Bernard Parresol (USDA) et quatre réviseurs anonymes pour l'aspect scientifique ainsi que Mme Pamela Cheers (SCF-CFL) et M. William F.J. Parsons (Université de Sherbrooke) pour la correction des deux manuscrits de langue anglaise. Merci aux collègues du laboratoire de dendrométrie, Claude Delisle (SCF-CFL) et Mme Marie-Josée Verreault pour leur soutien quotidien. Merci à M. Daniel Kneeshaw (Université du Québec à Montréal) et M. Louis-Paul Rivest (Université Laval) pour leur appréciation de cette thèse.

Finalement, je tiens à remercier la Compagnie Abitibi-Consolidated, division Mauricie (Jean Girard), le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), le Centre de foresterie des Laurentides (SCF), le Fonds québécois de la recherche sur la nature et les technologies (FQRNT), le Programme de mise en valeur de la ressource forestière (Volet 1), le Parc national de la Mauricie (Albert van Dijk) et le Parc national du Mont-Tremblant (Claire Ducharme) pour leur soutien financier tout au long de ce projet de doctorat.

## Table des matières

Résumé.....	i
Avant-propos.....	iii
Table des matières.....	iv
Liste des figures .....	viii
Liste des tableaux .....	ix
Liste des tableaux .....	ix
Introduction générale .....	1
CHAPITRE 1 - Étude rétrospective de la croissance en diamètre du sapin baumier ( <i>Abies balsamea</i> (L.) Mill.) et de l'épinette rouge ( <i>Picea rubens</i> Sarg.) en peuplements mixtes après une coupe à diamètre limite .....	4
Résumé .....	5
Mots-Clés .....	5
Introduction.....	6
Matériel.....	7
Cadre écologique .....	7
Historique des perturbations .....	9
Dispositif de suivi .....	10
Récolte et compilation des données .....	11
Méthode.....	14
Résultats.....	15
Discussion .....	19
Conclusion.....	23

CHAPITRE 2 - A variance-covariance structure to take into account repeated measurements and heteroscedasticity in growth modeling.....	24
Abstract .....	25
Key words.....	25
Introduction.....	26
Statistical developments.....	28
Statistical theory.....	28
Estimation method and model diagnostics.....	30
The predictions .....	32
The case study .....	33
The study site and the database .....	33
The model.....	35
Results.....	38
Discussion .....	42
Conclusions .....	45
CHAPITRE 3 - Calibrating a generalized diameter distribution model with mixed effects. ....	47
Abstract .....	48
Key words.....	48
Introduction.....	49
Material and method.....	51
The study site.....	51
Statistical developments .....	52
Model specifications.....	56
Results.....	58

Discussion .....	61
Conclusions .....	65
CHAPITRE 4 - Une matrice de transition à deux niveaux basée sur des distributions de probabilités : l'exemple de l'épinette rouge ( <i>Picea rubens</i> Sarg.) et du sapin baumier ( <i>Abies balsamea</i> (L.) Mill.) en peuplements mixtes. ....	68
Résumé .....	69
Mots-clés .....	69
Introduction.....	70
Développements du modèle .....	72
Aire d'étude .....	77
Cadre écologique.....	77
Historique des perturbations .....	78
Dispositif de suivi et base de données .....	79
Méthode.....	80
La calibration.....	80
La validation.....	83
Résultats.....	84
Discussion .....	89
L'approche de modélisation .....	89
La dynamique des espèces .....	91
Conclusion.....	94
Conclusion générale .....	96
Applicabilité des méthodologies développées.....	96
Conséquences pour l'aménagement .....	98
Bibliographie .....	101

Annexe 1 – Relation entre le diamètre avec écorce et le diamètre sans écorce .....	113
Relations $d_{hpse}$ vs $d_{hpae}$ .....	113
Annexe 2 – Équations pour calculer la croissance diamétrale des tiges de sapin baumier et d'épinette rouge en fonction du type écologique .....	113
Équations tirées du modèle [1] du chapitre 1 .....	113
Type écologique riche sur dépôt très mince .....	113
Type écologique riche sur dépôt épais .....	113
Type écologique pauvre sur dépôt très mince .....	114

## Liste des figures

Figure 1 – Accroissements annuels courants (AAC) du diamètre à hauteur de poitrine (dhp) des tiges de sapin baumier (SAB) et d'épinette rouge (EPR) entre 1954 et 1998 sur l'Aire d'observation de la rivière Ouareau. ....	10
Figure 2 – Prédictions de la croissance en diamètre des tiges de sapin baumier (SAB) et d'épinette rouge (EPR) après coupe en fonction de leur diamètre au moment de la récolte (dhpinit) et du type écologique (SAB : sapin baumier; EPR : épinette rouge).....	18
Figure 3 – Observed correlations among the within-plot normalized residuals and 1%-confidence intervals around the null hypothesis $H_0: \text{Corr}(\varepsilon_{ij}, \varepsilon_{j'})=0$ (the triangles indicate the observed correlations). ....	39
Figure 4 – Predictions and their 5%-confidence intervals for the different models of each species. ....	41
Figure 5 - Analysis of model residuals. ....	58
Figure 6 – The three typical patterns related to the fit of the model. ....	60
Figure 7 – Prédictions du recrutement, de l'accroissement et de la survie sur une période de 10 ans en fonction de trois différents niveaux de surface terrière. ....	86
Figure 8 – Distributions des biais et des ETE résultant de la validation du modèle de l'épinette rouge. ....	87
Figure 9 – Distributions des biais et des ETE résultant de la validation du modèle du sapin baumier. ....	88

## Liste des tableaux

Tableau 1 – Compositions et densités des trois types écologiques en 1953 (les valeurs entre parenthèses indiquent l'écart type).....	12
Tableau 2 – Sommaire des données utilisées pour la calibration du modèle (les valeurs entre parenthèses indiquent l'écart type).....	13
Tableau 3 – Résultats de la calibration du modèle.....	17
Table 4 – Summary of the 1956-57 measurement for the 96 permanent plots of the spruce-fir-hardwood forest type. ....	36
Table 5 – Comparison of the goodness of fit for models with different variance-covariance (VC) features. ....	38
Table 6 – Parameter estimates for model (13) calibrated with VC structure (5) (asymptotic standard errors are shown between parentheses).....	40
Table 7 – Density and basal area of the unthinned and the thinned groups (stems larger than 1.3 cm (0.5 in) in diameter at breast height). ....	52
Table 8– Parameter estimates resulting from the calibration of model (8) (probability of significance shown between parentheses).....	59
Tableau 9 – Sommaire des données de calibration et de validation (toutes les tiges de plus de 1,3 cm en dhp).....	81
Tableau 10 – Résultats de la calibration de la matrice de transition à deux niveaux pour l'épinette rouge et le sapin baumier. ....	85

## Introduction générale

Les peuplements mixtes de sapin baumier (*Abies balsamea* (L.) Mill.) et d'épinette rouge (*Picea rubens* Sarg.) sont présents sur l'ensemble de la zone tempérée du Québec méridional (Blum 1990; Grondin *et al.* 1996; Robitaille et Saucier 1998). Cette association végétale est également présente dans les provinces maritimes du Canada et dans les états américains du Nord de la Nouvelle-Angleterre. En termes de volume, l'épinette rouge et le sapin baumier constituent les principales essences forestières de la région Nord-Est des États-Unis (Seymour 1995).

L'historique d'exploitation de ces peuplements est relativement long pour le contexte nord-américain. Dès le début du 19<sup>ème</sup> siècle, l'industrie du sciage s'est approvisionnée dans ces forêts. Les coupes réalisées étaient alors d'intensité légère puisque l'on ne récoltait que les tiges de forts diamètres (Seymour 1992). L'essor de l'industrie des pâtes et papiers au début du 20<sup>ème</sup> siècle a profondément changé le mode d'exploitation de ces peuplements. Comme la dimension des bois ne constituait plus un obstacle à la transformation, les tiges d'épinette et de sapin étaient récoltées à des diamètres beaucoup plus faibles (Ray 1941; Seymour 1992). Les premières coupes à blanc ont été réalisées très tôt dans les années 1900.

Ces changements radicaux ont été perçus avec crainte (Heimburger 1941; Ray 1941). Déjà à cette époque, on doutait que les coupes à blanc et les coupes partielles de forte intensité puissent régénérer des peuplements d'une qualité équivalente à ceux récoltés. Les effets néfastes de la coupe à blanc seront rapidement confirmés : envahissement par la végétation compétitrice, « ensapinage » des peuplements et conversion de structure (Ray 1941; Westveld 1953; Frank et Bjorkbom 1973; Seymour 1992, 1995). Vers 1910-20, au moment même où l'on constatait ces problèmes d'aménagement, une épidémie de tordeuse des bourgeons de l'épinette (*Choristoneura fumiferana* Clem.) sévissait et provoquait des ravages sans précédent au Maine. C'est le mode d'exploitation dans son ensemble qui sera alors remis en question.

Dès la fin des années 1930, on a cherché à développer des méthodes d'aménagement mieux adaptées à la dynamique de ces peuplements. L'objectif était d'élaborer des traitements qui permettraient d'obtenir une croissance acceptable tout en minimisant les impacts sur la composition et la structure des peuplements forestiers. Les coupes partielles de faible intensité sont alors envisagées comme une alternative intéressante (Heimburger 1941; Ray 1941; Westveld 1953; Hatcher 1959a). Au milieu du 20<sup>ème</sup> siècle, plusieurs dispositifs permanents de grande envergure seront implantés dans le but de suivre et de comprendre la dynamique des peuplements mixtes de sapin baumier et d'épinette rouge. Différents types de coupes partielles y sont expérimentés. Les différents chapitres de cette thèse utilisent des données provenant de deux de ces dispositifs : la Forêt expérimentale du lac Édouard, située dans le Parc national de la Mauricie, et l'Aire d'observation de la rivière Ouareau, localisée à l'intérieur du Parc national du Mont-Tremblant.

Au Québec, l'aménagement des peuplements mixtes de sapin baumier et d'épinette rouge constitue toujours un enjeu important. Le suivi des dispositifs du lac Édouard et de la rivière Ouareau fournit un historique de la dynamique de ces peuplements à la suite de certaines interventions sur une période de plus de 50 ans. Ces données permettent aujourd'hui de mieux comprendre la dynamique de ces peuplements après coupe partielle.

Cette thèse est présentée en quatre chapitres, qui ont tous été rédigés sous forme de manuscrit scientifique. Chaque chapitre utilise une partie des données de l'un ou l'autre des dispositifs dans le but d'explorer certains éléments de la dynamique forestière dans le contexte mentionné. Même si certains chapitres portent davantage sur la méthodologie que les résultats, il n'en demeure pas moins que tous les manuscrits visaient à l'origine à déterminer dans quelle mesure l'élément de dynamique abordé pourrait contribuer au phénomène de raréfaction de l'épinette rouge.

Le premier chapitre aborde le sujet sous l'angle de la croissance en diamètre des tiges individuelles de sapin baumier et d'épinette rouge après une coupe partielle de forte intensité. Traitant également de la croissance de ces deux essences, le second

chapitre présente une analyse à l'échelle du peuplement. Les problèmes statistiques liés à l'analyse de mesures répétées y sont abordés. Des pistes de solution sont notamment proposées pour obtenir des prédictions plus cohérentes à partir de modèles de croissance. L'analyse des changements structurels est présentée dans le troisième chapitre. Ces changements sont abordés par le biais des distributions diamétrales. On y propose une méthodologie adaptée à la nature des données considérées dans l'analyse, données qui s'expriment sous forme de fréquences par classe de diamètre. L'accroissement et la mortalité des tiges individuelles ainsi que le recrutement sont analysés dans le dernier chapitre. Cette analyse est faite par le biais d'une matrice de transition à deux niveaux. Ces matrices sont en fait des modèles de croissance qui permettent d'obtenir des prédictions à une échelle plus fine que celle du peuplement. Comparée à l'approche traditionnellement utilisée pour calibrer ces matrices, la méthodologie suggérée dans le dernier chapitre repose sur l'utilisation de distributions de probabilités connues (binomiale, de Bernoulli et de Poisson) et permet d'obtenir des résultats plus cohérents.

La conclusion générale porte à la fois sur les méthodologies développées et sur les résultats obtenus à partir de celles-ci. D'une part, on y discute de l'applicabilité de ces méthodologies à d'autres contextes et, d'autre part, on y rassemble les résultats obtenus en un tout cohérent pour définir les conséquences sur l'aménagement de ces peuplements. Les recommandations énoncées contribuent à l'élaboration d'une stratégie de maintien de cet écosystème. Des décennies se sont écoulées depuis l'implantation des dispositifs du lac Édouard et de la rivière Ouareau, mais l'objectif demeure le même : apporter des éléments de connaissance qui permettront de développer une sylviculture mieux adaptée à ces peuplements.

## CHAPITRE 1

CHAPITRE 1 - Étude rétrospective de la croissance en diamètre du sapin baumier (*Abies balsamea* (L.) Mill.) et de l'épinette rouge (*Picea rubens* Sarg.) en peuplements mixtes après une coupe à diamètre limite

Copie du manuscrit publié  
dans l'édition novembre-décembre 2005  
de la revue Forestry Chronicle

Auteur : Mathieu Fortin<sup>a</sup>

<sup>a</sup>Département des sciences du bois et de la forêt, Université Laval

## Résumé

La croissance en diamètre à hauteur de poitrine de 341 tiges de sapin baumier et d'épinette rouge a été reconstituée sur un horizon de 48 à 49 ans après une coupe à diamètre limite. L'information disponible a été analysée à l'aide d'un modèle statistique dans lequel les effets du temps, de l'essence (2 niveaux), du type écologique (3 niveaux) et du diamètre au moment de la récolte ont été testés.

Les résultats révèlent que les gaules de sapin baumier ont des croissances diamétrales supérieures à celles des gaules d'épinette rouge, mais uniquement sur les types écologiques riches. Ces écarts entre les deux espèces s'atténuent chez les tiges dont le diamètre était marchand ( $d_{hp} > 9,0$  cm) au moment de la récolte. Sur le type pauvre, la croissance en diamètre de l'épinette rouge est supérieure à celle du sapin baumier et ce, peu importe le diamètre au moment de la récolte. Par ailleurs, il ressort de l'analyse que le diamètre de la tige au moment de l'intervention est une variable importante dans la prédiction de la croissance en diamètre chez l'épinette rouge, alors que l'effet de cette même variable est beaucoup moins marqué dans le cas du sapin baumier.

Bien que significatifs, les écarts de croissance entre les deux espèces sont modestes. La compétition exercée par la régénération de sapin baumier sur les tiges d'épinette rouge n'apparaît donc pas comme un facteur important dans le phénomène de raréfaction de l'épinette rouge.

## Mots-Clés

Épinette rouge (*Picea rubens* Sarg.), sapin baumier (*Abies balsamea* (L.) Mill.), croissance en diamètre, coupe partielle, peuplements mixtes, modèle linéaire, modèle mixte.

## Introduction

Les peuplements mixtes de sapin baumier (*Abies balsamea* (L.) Mill.) et d'épinette rouge (*Picea rubens* Sarg.) sont typiques de la région forestière acadienne (Rowe 1972). Par conséquent, ils occupent une proportion importante du territoire forestier dans les provinces maritimes du Canada et les états américains du Nord de la Nouvelle-Angleterre. Au Québec, on retrouve également cette association végétale dans les Basses-Laurentides et les Cantons de l'Est, deux régions qui correspondent au sous-domaine bioclimatique de l'érablière à bouleau jaune de l'Est (Grondin *et al.* 1996; Gosselin *et al.* 2001).

Depuis le début du 19<sup>ème</sup> siècle, ces peuplements ont joué un rôle prépondérant dans l'approvisionnement en matière ligneuse. Offrant un potentiel élevé pour le sciage de fortes dimensions, ils ont d'abord été exploités pour en récolter les tiges de pin blanc et les épinettes de fort diamètre (Ray 1941; Gaudreau 1986; Seymour 1992). À cette époque, les coupes étaient sélectives et l'intensité du prélèvement était légère (Ray 1941; Hatcher 1959a). Au tournant du 20<sup>ème</sup> siècle, l'essor de l'industrie des pâtes et papiers a provoqué une augmentation de la superficie des coupes et de l'intensité du prélèvement (Heimburger 1941; Ray 1956; Seymour 1992).

À partir des années 1940, on s'interroge sur la durabilité de l'aménagement forestier alors pratiqué dans les peuplements mixtes de sapin baumier et d'épinette rouge (e.g. Heimburger 1941). La diminution des proportions d'épinette rouge au profit du sapin baumier après certaines interventions est identifiée comme un problème alarmant, particulièrement sur les stations les plus fertiles (Heimburger 1941; Ray 1941; Hatcher 1959a; Seymour 1992; Gordon 1996). Gordon (1996) signale que le taux d'occupation actuel de l'épinette rouge ne représente plus que le cinquième de ce qu'il était dans la forêt précoloniale.

Plusieurs facteurs semblent contribuer à la raréfaction de l'épinette rouge. Bien que cette espèce soit très tolérante à l'ombre et que sa régénération puisse persister pendant plusieurs décennies sous couvert (Blum 1990; Seymour 1992), les semis et les gaules d'épinette rouge sont rarement aussi abondants que ceux de sapin

baumier (Ray 1941; Brissette 1996; Fortin *et al.* 2003a). De plus, les semis d'épinette rouge semblent sensibles aux changements de conditions environnementales et, par le fait même, plus vulnérables que ceux de sapin (Seymour 1992; Alexander *et al.* 1995). En plus d'être abondante, la régénération du sapin baumier réagit fortement au dégagement (Bakuzis et Hansen 1965; Frank 1990; McCarthy 2001). Après une ouverture du couvert, la régénération de cette espèce pourrait donc induire une forte compétition sur les semis et les gaules d'épinette rouge et ralentir leur croissance (Blum 1990; Seymour 1992).

Cette étude vise à mieux documenter ce problème de compétition, lequel pourrait s'avérer un facteur déterminant dans le phénomène de raréfaction de l'épinette rouge. Pour ce faire, la croissance diamétrale de 341 tiges d'épinette rouge et de sapin baumier a été reconstituée sur un horizon de 50 ans à la suite d'une coupe à diamètre limite. À partir de l'information recueillie, les trois hypothèses suivantes ont été vérifiées à l'aide d'un modèle statistique : (i) à la suite de la coupe, la croissance en diamètre des gaules de sapin baumier est supérieure à celle des gaules d'épinette rouge, (ii) compte tenu que les tiges marchandes d'épinette rouge conservent une capacité de réaction à l'éclaircie jusqu'à des âges avancés (Ray 1956; Blum 1990; Seymour 1992), cette différence de croissance s'atténue chez les tiges marchandes et, enfin, (iii) les écarts de croissance sont plus importants sur les stations les plus fertiles. Les résultats obtenus ont été mis en perspective par rapport aux connaissances actuelles quant à l'autécologie des deux espèces concernées. L'objet de la discussion était de constater dans quelle mesure ce problème de compétition entre les deux espèces contribue au phénomène de raréfaction de l'épinette rouge.

## **Matériel**

### ***Cadre écologique***

En 1953, le dispositif de l'Aire d'observation de la rivière Ouareau (AORO) (46°26'N, 74°10'W) a été implanté sur une concession forestière de la Consolidated Paper Corporation. Il est aujourd'hui inclus dans les limites du Parc national du Mont-Tremblant, lequel se situe à environ 100 km au nord de Montréal. La température

annuelle moyenne dans la localité la plus proche, Saint-Donat-de-Montcalm, est de 3,2°C pour la période 1971-2000 (Environnement Canada 2002). Juillet est le mois le plus chaud, avec une température moyenne de 18,1°C, tandis que janvier est le mois le plus froid, avec une valeur moyenne de -13,6°C. Les précipitations annuelles moyennes sont de 1128 mm, dont 295 mm tombent sous forme de neige.

L'AORO a une superficie totale de 13 km<sup>2</sup>. L'altitude moyenne du dispositif est de 515 m. Les dépôts de surface qui la caractérisent sont majoritairement des tills d'une épaisseur variable. Le relief typique de cette région se compose d'une succession de collines de basse altitude aux sommets arrondis (Robitaille et Saucier 1998).

Le cadre écologique général des Basses-Laurentides a été amplement décrit par plusieurs auteurs (Heimburger 1941; Gagnon et Marcotte 1980; Bélanger 1985; Lalumière et Thibault 1988; Grondin *et al.* 1996; Gosselin *et al.* 2001). Les peuplements mixtes de sapin baumier et d'épinette rouge se retrouvent généralement sur les stations à sols minces et sur les stations mésiques situées au bas des pentes et en terrains plats. Selon la classification écologique actuellement utilisée dans la province de Québec, ces stations correspondent aux types écologiques de la bétulaie jaune à sapin (MJ20, MJ22) et de la sapinière à épinette rouge (RS50) (*cf.* Gosselin *et al.* 2001). Dans une perspective plus large, la bétulaie jaune à sapin et la sapinière à épinette rouge peuvent être respectivement associées aux types forestiers « Spruce-fir-hardwood » et « Spruce-fir » définis par Westveld (1953).

Aucune donnée d'indice de qualité de station (IQS) n'est disponible pour documenter la richesse des stations sur l'AORO. La structure irrégulière des peuplements mixtes (Hatcher 1954; Fortin *et al.* 2003a), la tolérance à l'ombre des espèces considérées (Blum 1990; Frank 1990) et l'oppression en bas âge à laquelle elles sont souvent soumises (Ray 1956; Seymour 1992) sont autant de facteurs qui viennent compliquer le calcul de cette variable. Il existe toutefois un lien étroit entre certaines variables écologiques et l'IQS (Bélanger *et al.* 1995), de sorte qu'une classification écologique peut s'avérer efficace pour discriminer la productivité des stations forestières (Ménard 1999). Sur des types écologiques semblables à ceux de cette étude, Bédard (2002) a

démonstré que les IQS du sapin baumier et de l'épinette rouge variaient significativement d'un type à l'autre.

La bétulaie jaune à sapin (« Spruce-fir-hardwood ») possède une richesse relative plus élevée que la sapinière à épinette rouge (« Spruce-fir ») (Heimburger 1941; Gosselin 2001). La richesse relative se définit ici comme l'addition de plusieurs variables écologiques, notamment la présence de drainage oblique, le pH de l'humus, le type d'humus et la diversité floristique (*cf.* Gosselin *et al.* 2001). Pour alléger le texte, les références ultérieures aux types écologiques ont été faites en fonction de leur richesse relative, la bétulaie jaune à sapin étant le type riche et la sapinière à épinette rouge étant le type pauvre.

Le type riche est caractérisé par des peuplements mélangés dont les essences principales sont le sapin baumier, l'épinette rouge, le bouleau jaune (*Betula alleghaniensis* Britton) et le bouleau à papier (*Betula papyrifera* Marsh.) (Heimburger 1941; Grondin *et al.* 1996; Gosselin *et al.* 2001). Les essences qui les accompagnent sont l'érable à sucre (*Acer saccharum* Marsh.), l'érable rouge (*Acer rubrum* L.) et le thuya (*Thuja occidentalis* L.). Les peuplements qui croissent sur le type pauvre se caractérisent par une dominance des essences résineuses, soit l'épinette rouge et le sapin baumier (Heimburger 1941; Grondin *et al.* 1996; Gosselin *et al.* 2001). Le bouleau à papier et le thuya y sont présents, mais de façon sporadique.

### ***Historique des perturbations***

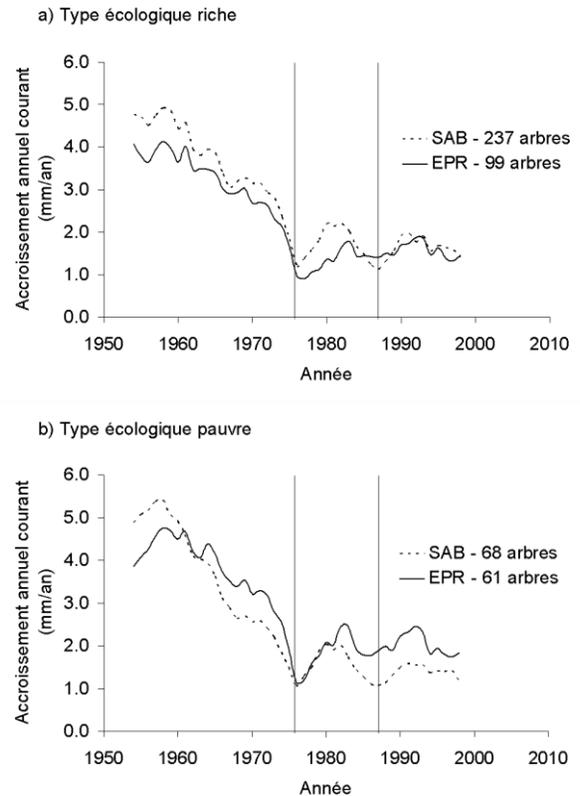
Depuis 1938, deux épidémies de tordeuse des bourgeons de l'épinette (*Choristoneura fumiferana* Clem.) ont eu lieu dans ce secteur, la première durant les décennies 1940-50 et la seconde durant celles de 1970-80 (Hardy *et al.* 1987). Il n'y a pas eu d'infestation sévère à l'intérieur du dispositif lors de la première épidémie (Hatcher 1954). Par ailleurs, les travaux d'inventaires à l'été 1998 et l'été 1999 n'ont pas permis d'identifier des zones de mortalité sévère, ce qui suggère que les effets de la seconde épidémie ont été limités en termes de mortalité. Des carottes de croissance prélevées sur des tiges de sapin baumier et d'épinette rouge ont permis d'obtenir une estimation de l'accroissement annuel courant (AAC) en diamètre à hauteur de poitrine (*dhp*, mesuré à 1,30 m) pour la période 1954-98 (Figure 1). Chez

les deux espèces, les effets de la seconde épidémie se sont traduits par une diminution d'accroissement pour une période d'environ 5 ans au milieu des années 1970. On observe également une diminution d'accroissement chez le sapin baumier pour une deuxième période de 5 ans durant les années 1980. La diminution des années 1980 est moins marquée que celle de la décennie précédente. De manière générale, les défoliations n'ont pas été critiques au point d'inhiber la croissance. Au pire de l'épidémie, les deux espèces ont en effet maintenu des AAC supérieurs ou équivalents à 1 mm par année.

En ce qui a trait à l'exploitation forestière, une récolte de bois à pâte a été réalisée durant les saisons hivernales entre l'automne 1947 et le printemps 1956. Le traitement alors en vigueur était la récolte des résineux dont le diamètre à la souche était équivalent ou supérieur à 18 cm (7 po). Les tiges ont été abattues à la sciote et au godendart, tandis que le débardage s'est effectué à cheval et à traîneau (Hatcher 1954). Entre 50 et 60% du volume total des tiges de plus de 8,9 cm (3,5 po) au *dhp* ont été récoltés lors de cette intervention (Hatcher 1954).

### **Dispositif de suivi**

Le dispositif de suivi de l'AORO a été établi en deux blocs à l'été 1953. Au total, il compte 288 placettes permanentes réparties systématiquement dans chacun des blocs selon une grille dont les maillons sont espacés de 201 m. Chaque placette a



**Figure 1 – Accroissements annuels courants (AAC) du diamètre à hauteur de poitrine (*dhp*) des tiges de sapin baumier (SAB) et d'épinette rouge (EPR) entre 1954 et 1998 sur l'Aire d'observation de la rivière Ouareau.**

une superficie de 405 m<sup>2</sup> (0,1 acre). À l'intérieur de chacune d'elles, toutes les tiges dont le *dhp* était supérieur à 1,3 cm (0,5 po) ont été dénombrées par classe de 2,5 cm (1 po). Le dispositif a été mesuré à nouveau en 1963 et entre 1974 et 1977. En 1998 et 1999, 59 placettes ont été mesurées pour une quatrième fois.

### ***Récolte et compilation des données***

Au printemps 1999, 13 placettes du dispositif ont fait l'objet d'une étude détaillée de la croissance en diamètre. Ces placettes ont été choisies parmi celles (i) coupées lors de l'intervention de 1947-56, (ii) retrouvées lors de l'inventaire de 1998-99 et (iii) correspondant à l'un ou l'autre des deux types écologiques retenus. À l'intérieur de ces 13 placettes, toutes les tiges de sapin baumier et d'épinette rouge dont le *dhp* était supérieur ou équivalent à 63 mm (2,5 po) en 1999 ont été échantillonnées pour en connaître le diamètre, l'épaisseur de l'écorce et la croissance en diamètre depuis la coupe de 1947-56. Le diamètre a été mesuré au millimètre près à l'aide d'un ruban diamétrique. Trois mesures de l'épaisseur de l'écorce ont été effectuées à hauteur de poitrine à l'aide d'une jauge à écorce. Ces mesures ont été prises au millimètre près selon un angle de 120 degrés l'une par rapport à l'autre. Trois carottes ont ensuite été prélevées à hauteur de poitrine à la sonde de Pressler selon une configuration semblable à celle des mesures d'écorce, c'est-à-dire à un angle de 120 degrés l'une par rapport à l'autre. Au total, 465 tiges ont été échantillonnées.

En 2000, tous les échantillons ont été séchés et sablés pour ensuite en mesurer la largeur des cernes annuels à l'aide du logiciel WinDendro. Trois années repères ont été utilisées pour dater les échantillons. Cette mesure visait à éviter l'omission de cernes manquants ou peu apparents en raison des défoliations résultant des épidémies de tordeuse. Pour estimer la largeur des cernes à l'état vert, des retraits de 2,7% et 4,0% ont été respectivement utilisés pour le sapin baumier et l'épinette rouge (Jessome 1977).

À l'été 2003, les 13 placettes étudiées ont été visitées et classifiées à nouveau en raison de précisions apportées au système de classification écologique provincial (*cf.* Gosselin 2001). Cette nouvelle classification a permis d'obtenir trois groupes homogènes à partir des 13 placettes initialement retenues : le type riche sur dépôt

très mince (4 placettes), le type riche sur dépôt épais (6 placettes) et le type pauvre sur dépôt très mince (3 placettes). Le seuil de 25 cm est utilisé dans cette classification pour qualifier l'épaisseur du dépôt (Gosselin 2001). Le Tableau 1 présente un portrait sommaire des placettes en fonction des trois types écologiques retenus. Ce sommaire a été compilé à partir des données de l'inventaire de 1953, soit l'inventaire réalisé dans les premières années suivant la coupe.

**Tableau 1 – Compositions et densités des trois types écologiques en 1953 (les valeurs entre parenthèses indiquent l'écart type)**

<b>Caractéristiques</b>	<b>Type riche très mince</b>	<b>Type riche épais</b>	<b>Type pauvre très mince</b>
Nombre de placettes	4	6	3
<b>Bouleau jaune</b>			
Surface terrière (m <sup>2</sup> ·ha <sup>-1</sup> )	0,3 (0,5)	2,2 (1,7)	0,0 (0,0)
Densité (ti·ha <sup>-1</sup> )	30 (29)	41 (33)	9 (15)
<b>Bouleau à papier</b>			
Surface terrière (m <sup>2</sup> ·ha <sup>-1</sup> )	2,8 (1,0)	1,4 (1,3)	0,2 (0,2)
Densité (ti·ha <sup>-1</sup> )	868 (339)	261 (210)	193 (80)
<b>Épinette rouge</b>			
Surface terrière (m <sup>2</sup> ·ha <sup>-1</sup> )	7,6 (3,5)	4,0 (2,8)	5,3 (2,5)
Densité (ti·ha <sup>-1</sup> )	1044 (265)	427 (290)	764 (239)
<b>Sapin baumier</b>			
Surface terrière (m <sup>2</sup> ·ha <sup>-1</sup> )	9,5 (7,9)	8,3 (3,6)	2,7 (1,5)
Densité (ti·ha <sup>-1</sup> )	3213 (2608)	1878 (585)	837 (249)
<b>Autres essences</b>			
Surface terrière (m <sup>2</sup> ·ha <sup>-1</sup> )	0,8 (1,6)	2,3 (5,5)	0,0 (0,0)
Densité (ti·ha <sup>-1</sup> )	75 (151)	12 (30)	0 (0)
<b>Toutes essences confondues</b>			
Surface terrière (m <sup>2</sup> ·ha <sup>-1</sup> )	21,1 (8,6)	18,1 (7,1)	8,3 (4,1)
Densité (ti·ha <sup>-1</sup> )	5230 (3040)	2618 (914)	1802 (465)

Afin de retracer l'évolution du diamètre avec écorce (*dhpae*) à partir des carottes de croissance, les mesures d'épaisseur d'écorce ont d'abord été utilisées pour établir une relation entre le diamètre avec écorce et le diamètre sans écorce (*dhpse*). Les équations obtenues par régression linéaire sont présentées en annexe. Les accroissements annuels en *dhpse* ont ensuite été calculés en prenant la moyenne arithmétique de la largeur des cernes d'une année donnée pour les carottes d'une même tige et en multipliant cette moyenne par deux. L'année 1998 a été considérée comme la dernière année de croissance complète. Le *dhpse*<sub>1997</sub> a été estimé en soustrayant l'accroissement annuel de 1998 du *dhpse*<sub>1998</sub>. Cette démarche a été répétée autant de fois que nécessaire pour obtenir tous les *dhpse* entre l'année de la coupe et 1998. Les *dhpae* correspondant à chacune de ces années ont été calculés en utilisant la relation préalablement établie entre le *dhpae* et le *dhpse*.

**Tableau 2 – Sommaire des données utilisées pour la calibration du modèle (les valeurs entre parenthèses indiquent l'écart type)**

Caractéristiques	Type riche très mince	Type riche épais	Type pauvre très mince
Épinette rouge			
Nombre de tiges	50	44	40
dhpinit (mm) <sup>†</sup>	59 (104)	48 (102)	46 (81)
Sapin baumier			
Nombre de tiges	67	101	39
dhpinit (mm) <sup>†</sup>	51 (64)	48 (69)	25 (55)

<sup>†</sup>*dhpinit* : *dhp* avec écorce au moment de la coupe

Les tiges pour lesquelles il était impossible d'estimer le *dhpae* au moment de la coupe ont été retirées de l'analyse. Parmi les 124 arbres rejetés, 80 tiges (21 épinettes et 59 sapins) étaient des recrues, c'est-à-dire des tiges qui n'avaient pas de *dhp* au moment de la coupe. Pour les 44 autres tiges (5 épinettes et 39 sapins), les carottes étaient cariées ou trop endommagées pour prendre des mesures. Au total, 341 tiges ont été retenues pour l'analyse, chacune d'elles comptant en moyenne 48 à 49

mesures. Le Tableau 2 présente la répartition de ces tiges selon les différents types écologiques.

## Méthode

Les données ont d'abord été soumises à un examen préliminaire afin de faciliter l'élaboration d'un modèle. À partir du patron général d'accroissement, le modèle suivant a été retenu :

$$[1] \quad \Delta dh_{pae} = \beta_1 \ln(t+1) + \beta_2 \ln^2(t+1)$$

où  $\Delta dh_{pae}$  représente la croissance en  $dh_p$  avec écorce après la coupe (mm),  $t$ , le nombre d'années écoulées depuis la récolte et  $\beta_1$  et  $\beta_2$ , les paramètres du modèle. Le modèle [1] est linéaire, polynomial du second degré et ne comporte pas d'ordonnée à l'origine. Pour tester les hypothèses de départ, chacun des deux paramètres du modèle ( $\beta_1$  et  $\beta_2$ ) a été remplacé par une fonction linéaire exprimant trois effets, soit l'essence, le type écologique et le diamètre au moment de la coupe ( $dh_{pinit}$ ) ainsi que les interactions possibles entre ceux-ci.

Au point de vue statistique, les données servant à l'analyse posent certains problèmes. D'une part, les placettes constituent des blocs expérimentaux dans le plan d'expérience, ce qui représente un effet aléatoire. D'autre part, les mesures d'accroissement sont des mesures répétées. Le fait d'avoir échantillonné le même arbre plus d'une fois laisse supposer que les erreurs issues de ces mesures sont corrélées entre elles. Les analyses qui font abstraction de cette corrélation fournissent des estimés non biaisés des paramètres (Sullivan et Clutter 1972; Vanclay 1994). Par contre, elles biaisent les variances qui sont associées à ces estimés, affectant ainsi le choix des variables explicatives à inclure dans le modèle (Gregoire *et al.* 1995).

Les effets aléatoires et les effets des mesures répétées peuvent être pris en compte en estimant la structure de la matrice de variance-covariance des erreurs (Littell *et al.* 1996). Pour les effets aléatoires, on ajoute à la variance du modèle une composante pour tenir compte de chaque effet aléatoire (Gregoire 1987; Littell *et al.* 1996). Cette composante de la variance permet de contrôler des effets qui ne sont pas inclus dans

le modèle sous forme de variables explicatives (Gregoire 1987). Les effets aléatoires peuvent être associés à l'ordonnée à l'origine ou à certaines variables indépendantes sous forme d'interactions (Gregoire *et al.* 1995; Littell *et al.* 1996). En ce qui concerne l'effet des mesures répétées, plusieurs structures de covariance permettent de définir la relation qui existe entre ces mesures. Parmi les plus connues, la structure autorégressive définit une corrélation qui diminue en fonction de la distance entre les mesures. À l'inverse, la structure symétrique (« compound symmetry ») conserve une corrélation constante, peu importe le nombre d'unités de temps séparant les mesures répétées.

Dans le cas présent, deux composantes ont été ajoutées à la variance du modèle, chacune correspondant à un effet aléatoire. Ces effets ont été inclus dans chacune des deux fonctions remplaçant les deux paramètres du modèle [1] et ont pris la forme d'une interaction entre l'essence et la placette. Cette interaction apparaissait logique, les effets incontrôlables dus à la placette pouvant être différents d'une espèce à l'autre. En ce qui concerne les mesures répétées, une matrice de variance-covariance non structurée a d'abord été utilisée pour identifier la structure de covariance la plus appropriée, conformément à la méthode décrite dans Littell *et al.* (1996). Toutes les analyses ont été effectuées à l'aide de la procédure MIXED du progiciel SAS (SAS Institute Inc. 2002).

## Résultats

Une calibration préliminaire du modèle a permis de détecter une structure de covariance autorégressive de premier ordre dans la matrice de variance-covariance des erreurs. Les résidus ont également affiché une hétéroscédasticité marquée. La structure autorégressive à variance hétérogène (*cf.* Littell *et al.* 1996) a été retenue pour les essais subséquents. Toutefois, l'utilisation de cette structure de covariance implique l'estimation d'une variance propre à chaque intervalle de temps en plus d'un paramètre établissant la corrélation entre les intervalles successifs. Dans le cas présent, cette structure supposait l'estimation de 52 paramètres. Pour alléger le processus, seuls des intervalles de cinq ans ont été conservés. Le nombre de paramètres est ainsi passé de 52 à 11.

Le modèle a été calibré de nouveau en incluant cette nouvelle structure de variance-covariance et en ne conservant que les effets et interactions significatifs à un seuil  $\alpha=5\%$ . Les résidus standardisés de cette nouvelle calibration affichaient les caractéristiques requises pour assurer la validité des inférences statistiques, soit une variance homogène, une distribution normale et des corrélations très faibles. Le Tableau 3 présente les résultats de la calibration.

La corrélation entre deux mesures successives de cinq ans est très élevée et atteint près de 99%. L'erreur type augmente avec le temps, démontrant l'hétéroscédasticité des données. L'erreur passe ainsi d'une valeur de 20,9 mm, 10 ans après coupe, à une valeur de 43,5 mm, 50 ans après coupe. Les effets aléatoires dus aux placettes ont des variances respectives de 6,56 et 1,96.

Les probabilités associées aux effets fixes démontrent que la croissance diamétrale du sapin baumier est significativement différente de celle de l'épinette rouge, l'essence étant présente dans le modèle sous forme d'interaction avec le temps ( $\text{Essence} \times \ln(t+1)$  et  $\text{Essence} \times \ln^2(t+1)$ ). De plus, cette différence de croissance entre les deux essences varie en fonction du type écologique ( $\text{Type écologique} \times \text{Essence} \times \ln(t+1)$  et  $\text{Type écologique} \times \text{Essence} \times \ln^2(t+1)$ ).

Le diamètre des tiges au moment de l'intervention ( $dhpinit$ ) a également un effet significatif sur la croissance diamétrale à la suite de l'intervention, cette variable explicative étant, elle aussi, présente sous forme d'interaction avec le temps ( $dhpinit \times \ln(t+1)$  et  $dhpinit \times \ln^2(t+1)$ ). Cette interaction entre le diamètre initial et le temps varie significativement d'une espèce à l'autre ( $dhpinit \times \text{Essence} \times \ln(t+1)$  et  $dhpinit \times \text{Essence} \times \ln^2(t+1)$ ) et d'un type écologique à l'autre ( $dhpinit \times \text{Type écologique} \times \ln(t+1)$ ).

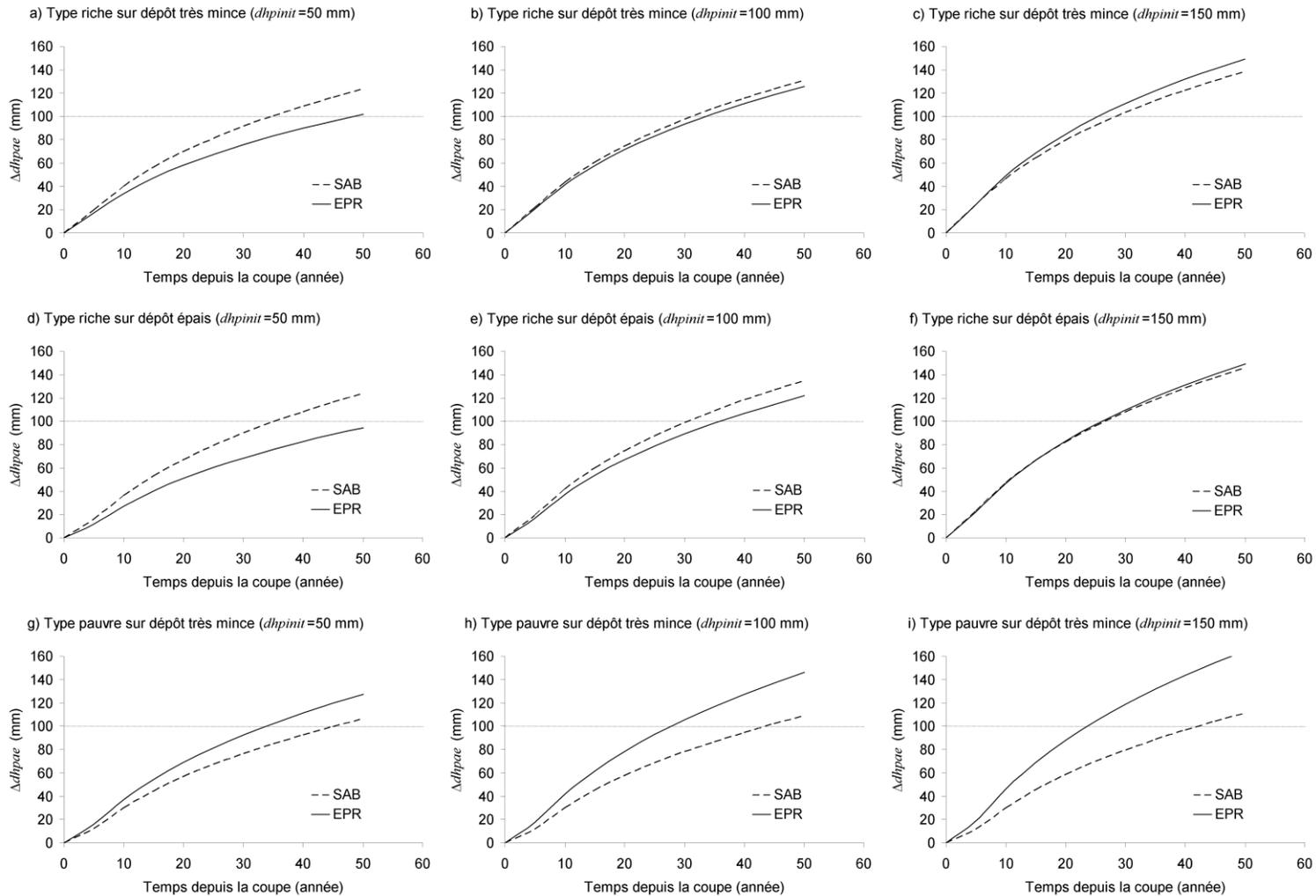
Les croissances diamétrales prédites par le modèle [1] sont présentées à la Figure 2 en fonction de l'essence, du type écologique et de trois diamètres initiaux ( $dhpinit$ ). Les équations à partir desquelles sont tirées ces courbes sont présentées en annexe. Les croissances atteignent entre 95 et 165 mm sur un horizon de 50 ans, ce qui correspond à des accroissements annuels moyens (AAM) variant entre 1,9 et 3,3

mm·an<sup>-1</sup>. Les patrons de croissance des types riches sont relativement semblables, alors que celui du type pauvre diffère.

**Tableau 3 – Résultats de la calibration du modèle**

Éléments de l'analyse	Valeur
Logarithme de la vraisemblance (-2Loglikelihood)	19962
Nombre d'observations	3191
Corrélation entre deux mesures successives	0,9879
Écart type de l'erreur résiduelle (mm)	
10 ans après la coupe	20,9
30 ans après la coupe	33,9
50 ans après la coupe	43,5
Variances des effets aléatoires	
Essence × Placette × ln( <i>t</i> +1)	6,56
Essence × Placette × ln <sup>2</sup> ( <i>t</i> +1)	1,96
Niveau de signification des effets fixes (Pr>F)	
ln( <i>t</i> +1)	<0,0001
Essence × ln( <i>t</i> +1)	0,0470
Type écologique × Essence × ln( <i>t</i> +1)	0,0115
<i>dhpinit</i> × ln( <i>t</i> +1) <sup>†</sup>	0,0118
<i>dhpinit</i> × Essence × ln( <i>t</i> +1) <sup>†</sup>	<0,0001
<i>dhpinit</i> × Type écologique × ln( <i>t</i> +1) <sup>†</sup>	0,0007
ln <sup>2</sup> ( <i>t</i> +1)	<0,0001
Essence × ln <sup>2</sup> ( <i>t</i> +1)	0,0099
Type écologique × Essence × ln <sup>2</sup> ( <i>t</i> +1)	0,0477
<i>dhpinit</i> × ln <sup>2</sup> ( <i>t</i> +1) <sup>†</sup>	<0,0001
<i>dhpinit</i> × Essence × ln <sup>2</sup> ( <i>t</i> +1) <sup>†</sup>	<0,0001

<sup>†</sup>*dhpinit* : *dhp* avec écorce au moment de la coupe



**Figure 2 – Prédications de la croissance en diamètre des tiges de sapin baumier (SAB) et d'épinette rouge (EPR) après coupe en fonction de leur diamètre au moment de la récolte ( $d_{hpinit}$ ) et du type écologique.**

Sur les types riches, les tiges de sapin baumier d'un diamètre initial (*d<sub>hpinit</sub>*) de 50 mm affichent une croissance moyenne plus rapide que celles d'épinette rouge (Figure 2a, d). Chez les deux espèces, l'augmentation du diamètre initial se traduit par une augmentation de la croissance. Cette augmentation est toutefois plus importante chez l'épinette rouge, de sorte que la croissance diamétrale de cette espèce en vient à rattraper et même à dépasser légèrement celle du sapin baumier chez les tiges dont le diamètre initial est de 150 mm (Figure 2c, f).

Sur le type pauvre, l'épinette rouge affiche une croissance moyenne supérieure à celle du sapin baumier et ce, pour chacun des trois exemples présentés (Figure 2g, h, i). Quant à l'augmentation du diamètre initial, elle se traduit par une augmentation de la croissance, mais seulement dans le cas de l'épinette rouge. En effet, les courbes du sapin baumier sur le type pauvre sont pratiquement identiques dans les trois cas.

Pour un même diamètre initial, la croissance en diamètre des tiges d'épinette rouge sur le type pauvre est plus rapide que celle sur les types riches. Chez le sapin baumier, la tendance inverse est observée : la croissance moyenne des types riches est plus rapide que celle du type pauvre.

## **Discussion**

Le sapin baumier est une espèce opportuniste qui profite à la fois des petites et des grandes trouées pour réagir fortement à la mise en lumière (Bakuzis et Hansen 1965; Frank 1990; McCarthy 2001). De façon générale, les gaules d'épinette rouge sont moins nombreuses, plus susceptibles d'être affectées par les changements de conditions environnementales et leur réaction à l'éclaircie n'est pas aussi forte (Ray 1956; Frank et Bjorkbom 1973; Seymour 1992). On présume donc que les gaules d'épinette souffriraient de la compétition exercée par le sapin baumier (Blum 1990; Seymour 1992).

À la suite d'une coupe à diamètre limite, les gaules de sapin baumier ont effectivement une croissance diamétrale moyenne supérieure à celle des gaules d'épinette rouge sur les types écologiques riches. Bien que significatives, les

différences sont toutefois modestes. Sur un horizon de 10 ans, soit durant les premières années suivant la récolte, la différence entre les deux espèces est en moyenne inférieure à 10 mm pour des tiges dont le *dhp* au moment de la coupe était de 50 mm (Figure 2a, d). Les résultats de la présente étude portent donc à conclure que la compétition exercée par le sapin baumier sur les types écologiques riches après une coupe à diamètre limite a une incidence limitée sur la croissance diamétrale des gaules d'épinette rouge. Non seulement les écarts de croissance entre les deux espèces sont relativement faibles, mais les erreurs types du modèle sont beaucoup plus élevées que ces écarts (Tableau 3). Si l'on sélectionnait une gaule de chaque espèce au hasard, il ne serait donc pas improbable que la gaule d'épinette rouge ait une croissance plus rapide que celle de sapin baumier, bien que la tendance générale favorise le sapin.

Par ailleurs, les différences de croissance observées entre les gaules des deux espèces s'atténuent lorsque le diamètre initial des tiges augmente. Il en ressort que le diamètre de la tige au moment de l'intervention est un facteur d'importance dans la prédiction de la croissance en diamètre chez l'épinette rouge, alors que l'effet de cette même variable est beaucoup moins marqué dans le cas du sapin baumier. Bien que le patron de croissance soit différent sur le type pauvre, on remarque ce même effet « diamètre » chez l'épinette rouge.

La capacité de réaction du sapin baumier à la suite d'une éclaircie tend à diminuer, puis à se stabiliser, en fonction de la hauteur initiale de la tige (Hatcher 1964; Riopel 1999). Un dégagement favorise donc les tiges de sapin de faible diamètre (Bakuzis et Hansen 1965). Après une coupe partielle, Riopel (1999) estime que les tiges de 6 m de hauteur sont celles qui affichent un accroissement maximum. Alors que la longévité du sapin baumier n'excède pas 200 ans (Bakuzis et Hansen 1965; Frank 1990), l'épinette rouge peut atteindre des âges supérieurs à 400 ans (Ray 1956; Blum 1990; Busing et Wu 1990), ce qui en fait l'une des espèces les plus longévives du Québec. Même lorsqu'elle atteint un âge très avancé, l'épinette rouge conserve une étonnante capacité de réaction à l'éclaircie (Robertson 1942; Ray 1956; Seymour 1995). En effet, même les arbres très âgés peuvent s'accroître significativement, du moment que leurs cimes ont l'espace nécessaire pour s'étendre (Seymour et Kenefic

2002). Par ailleurs, même si les tiges d'épinette rouge sont souvent soumises à des périodes d'oppression (Blum 1990; Seymour 1992), la croissance à la suite du dégagement s'en trouve peu affectée (Seymour et Fajvan 2001). En somme, l'effet « diamètre » qui s'observe chez l'épinette rouge s'explique par le fait que cette espèce peut maintenir une efficacité de croissance relativement bonne en dépit de l'âge et de l'historique du peuplement.

Alors que l'on attendait, dans une certaine mesure, un patron semblable à celui des types riches, la croissance diamétrale des gaules de sapin baumier est en moyenne inférieure à celle des gaules d'épinette rouge sur le type pauvre. Toutefois, la différence de croissance entre les deux espèces est peu importante, soit de l'ordre de 20 mm sur un horizon de 50 ans (Figure 2g) pour des tiges dont le *dhp* était de 50 mm au moment de l'intervention. À cause de l'effet du diamètre initial chez l'épinette rouge, cet écart de croissance tend à augmenter en fonction du diamètre des tiges au moment de la récolte.

Si le patron de croissance des gaules sur le type pauvre est surprenant d'un point de vue écologique, il est néanmoins conforme à ce que l'on connaît de la productivité des types écologiques étudiées. Sur des stations comparables à celles de cette étude, Bédard (2002) a en effet démontré que l'IQS de l'épinette rouge augmente légèrement sur le type pauvre, passant de 16,2 m à 17,2 m en prenant 50 ans comme âge de référence. À l'inverse, l'IQS du sapin baumier tend à diminuer fortement, passant de 19,2 m sur les types riches à une valeur de 16,4 m sur le type pauvre pour le même âge de référence (Bédard 2002). En outre, à des densités équivalentes, la croissance en volume de l'épinette rouge à l'échelle du peuplement ne varie pas de façon significative d'un type écologique à l'autre, alors que celle du sapin baumier est beaucoup plus faible sur le type écologique pauvre (Fortin *et al.* 2003b). Les résultats de la présente étude vont donc dans le même sens que ceux des études de Bédard (2002) et Fortin *et al.* (2003b) : par rapport aux types écologiques riches, la croissance diamétrale de l'épinette rouge connaît une légère augmentation sur le type pauvre, alors que celle du sapin baumier diminue de façon significative.

L'ajout d'une mesure de densité parmi les effets fixes du modèle aurait permis d'améliorer la précision de l'analyse de la croissance en diamètre des deux espèces. Plusieurs essais de modélisation ont été réalisés en utilisant les surfaces terrières et les nombres de tiges à l'hectare calculés à partir de l'inventaire de 1953. Dans tous les cas, ces variables se sont avérées non significatives. En théorie, les effets aléatoires prennent compte des variables de site qui ne sont pas incluses comme effets fixes dans un modèle mixte (Gregoire 1987). Toutefois, dans le cas présent, les densités sont associées aux types écologiques (Tableau 1). En effet, une analyse de variance (ANOVA) réalisée sur les surfaces terrières des types écologiques indique un lien étroit entre les deux variables ( $Pr > F = 0,0348$ ), les surfaces terrières les plus faibles se retrouvant essentiellement sur le type pauvre. Les deux variables étant fortement associées, la capacité de la régression à ne pas les confondre peut raisonnablement être mise en doute. L'effet du type écologique reflète donc une interaction indissociable entre cette variable et la densité. Un plan d'expérience offrant une meilleure dispersion des densités initiales permettrait de mieux distinguer ces deux effets l'un de l'autre. En l'absence de telles données, les résultats obtenus par rapport aux types écologiques doivent être interprétés avec prudence, même s'ils semblent conformes à ceux obtenus par d'autres auteurs dans des contextes semblables (e.g. Bédard 2002, Fortin *et al.* 2003b).

Par ailleurs, cette étude n'aborde que la croissance des tiges qui ont survécu sur un horizon de près de 50 ans. L'effet des deux épidémies de tordeuse sur la croissance de ces tiges est plutôt faible et ne représente, tout au plus, qu'un retard de quelques années. Les tiges mortes durant cet intervalle sont probablement celles qui affichaient les accroissements les plus faibles et qui ont été les plus touchées par la tordeuse. Dans le cadre de cette étude, aucune donnée ne permettait d'évaluer la croissance de ces tiges. Cette lacune incite à croire qu'une utilisation des équations issues du modèle [1] à des fins de prédiction sur un horizon inférieur à 50 ans pourrait mener à surestimer la croissance des tiges des deux espèces concernées.

## Conclusion

Sur l'Aire d'observation de la rivière Ouareau, les proportions d'épinette rouge ont fortement diminué sur le type écologique de la bétulaie jaune à sapin après une coupe à diamètre limite (Fortin *et al.* 2003a). Cette raréfaction de l'épinette rouge ne semble pas être le résultat d'une compétition exercée par les gaules de sapin baumier. Dans les faits, les écarts de croissance en diamètre entre les deux espèces sont en moyenne relativement faibles. Sur le type écologique de la sapinière à épinette rouge, certains auteurs ont signalé que le problème de raréfaction de l'épinette rouge n'était pas alarmant (Ray 1956; Hatcher 1959a; Fortin *et al.* 2003b). Les résultats de cette étude vont dans le même sens en démontrant que, sur ce même type écologique, les gaules d'épinette rouge réagissent plus fortement à l'éclaircie que celles de sapin baumier.

L'abondance de la régénération de chacune des deux espèces est sans doute un facteur qui contribue plus fortement aux changements de composition que la croissance elle-même. Après une coupe partielle, la régénération du sapin baumier est généralement plus abondante que celle de l'épinette rouge (Heimburger 1941; Ray 1956; Brissette 1996; Archambault *et al.* 2003). Comme la régénération de l'épinette rouge demeure relativement constante dans le temps, la raréfaction de l'espèce semble plus le résultat d'un prélèvement trop important que d'une régénération déficiente (Fortin *et al.* 2003a). Or, on connaît peu les facteurs qui régissent l'établissement de la régénération d'épinette rouge. Aussi, cette avenue de recherche pourrait se révéler un domaine clé dans la sylviculture de cette espèce.

## CHAPITRE 2

CHAPITRE 2 - A variance-covariance structure to take into account repeated measurements and heteroscedasticity in growth modeling.

Manuscript submitted to  
the European Journal of Forest Research

Authors: Mathieu Fortin<sup>a</sup>, Gaétan Daigle<sup>b</sup>, Chhun-Huor Ung<sup>c</sup>,  
Jean Bégin<sup>a</sup>, and Louis Archambault<sup>c</sup>

<sup>a</sup>Department of Wood and Forest Sciences, Laval University

<sup>b</sup>Department of Mathematic and Statistics, Laval University

<sup>c</sup>Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre

## Abstract

This study proposes a variance-covariance (VC) structure to take into account repeated measurements and heteroscedasticity in a context of growth modeling. The VC structure integrates a variance function and a continuous autoregressive covariance structure. It was tested on a nonlinear growth model calibrated with permanent-plot data. Using a whole stand approach, basal area growth was independently modeled for red spruce (*Picea rubens* Sarg.) and balsam fir (*Abies balsamea* (L.) Mill.) in mixed stands.

For both species, the implementation of the VC structure significantly improved the maximum likelihood of the model. In both cases, it efficiently accounted for heteroscedasticity and autocorrelation, since the normalized residuals no longer exhibited departures from the assumptions of independent error terms with homogeneous variances. Moreover, models calibrated with this VC structure provided more accurate predictions when compared with traditional nonlinear least squares (NLS) models.

This case study demonstrates that the implementation of a VC structure may provide parameter estimates that are consistent with asymptotically unbiased variances in a context of nonlinear growth modeling using a whole stand approach. Since the variances are no longer biased, the hypothesis tests performed on the estimates are valid when the number of observations is large.

## Key words

Nonlinear modeling, variance modeling, covariance structure, predictions, red spruce (*Picea rubens* Sarg.), balsam fir (*Abies balsamea* (L.) Mill.).

## Introduction

Historically, growth and yield models have been used either to bring inventory data to a common year or to provide growth forecasts. When using a whole-stand approach (*cf.* Vanclay 1994, §2), growth and yield models calibrated with permanent-plot data explicitly predict recruitment and mortality. However, from a statistical standpoint, repeated measurements taken on the same subject, for instance a plot, invalidate the basic assumption of independent error terms, which is underlying the traditional least squares methods (Steel *et al.* 1997, §7.10). Moreover, the variances of the error terms often tend to increase along with the predicted values. The resulting heteroscedasticity represents a violation of another basic assumption of these traditional methods (Steel *et al.* 1997, §7.10). In cases of departures from the assumptions of independent error terms with homogeneous variances, the ordinary least squares (OLS) and the nonlinear least squares (NLS) methods provide consistent estimators for the vector of parameters (Sullivan and Clutter 1972; Gregoire *et al.* 1995). However, the estimated variances of the parameter estimates are both biased and inconsistent (LeMay 1990).

Although some statistical tools are now available, many growth and yield models are still calibrated without considering these two statistical issues (*e.g.* Dhôte and Hervé 2000; Pretzsch *et al.* 2002; Fortin *et al.* 2003b; Deleuze *et al.* 2004). As outlined in Gregoire *et al.* (1995), this is probably due to the complexity of error modeling and the traditional assumption that OLS and NLS estimators are unbiased. However, the biased variances resulting from such regressions hinder the selection of the appropriate explanatory variables (Gregoire *et al.* 1995). Somehow, the calibration of a growth model with permanent-plot data should take into account repeated measurements and heteroscedasticity in order to obtain unbiased statistical inferences.

Some authors have already addressed the matter of repeated measurements in growth and yield modeling by using either random effects and/or direct error modeling (*e.g.* Sullivan and Clutter 1972; Gregoire 1987; Goelz and Burk 1992;

Gregoire *et al.* 1995; Beaumont *et al.* 1999; Fang and Bailey 2001; Hall and Bailey 2001). Both methods consist in parameterizing the variance-covariance (VC) matrix of the error terms. To limit the number of covariance parameters to be estimated, the design of the variance-covariance structure is typically assumed to be the same across the individuals (Vonesh and Carter 1992). Many VC structures have already been documented, e.g. compound symmetric, first-order autoregressive, and the moving average structures (*cf.* Littell *et al.* 1996, p.93; Pinheiro and Bates 2000, §5.3.3). With permanent-plot data, the main concern is often to choose a covariance structure that fits unequally spaced and unbalanced repeated measurements, which are common with this type of data. Gregoire *et al.* (1995) clearly demonstrated that it is possible through a continuous-time autoregressive structure. However, this structure alone does not account for heterogeneous variances. Indeed, a transformation can be performed on the dependent variable to overcome the problem of heteroscedasticity (Steel *et al.* 1997, §9.16). Nevertheless, conversion problems arise when the resulting estimates and their inferences have to be retransformed to the original scale (Duan 1983). A weighted regression could also be performed as an alternative to transformations. However, choosing the appropriate weight requires some subjectivity and modeling the variance seems to be preferable, especially if the purpose of the model is to obtain predictions (Gregoire and Dyer 1989; Parresol 1993). In fact, unequally spaced repeated measurements and heteroscedasticity are common in forest growth modeling but rarely simultaneously considered in the common VC structures.

This study proposes a VC structure to take into account both problems in a context of growth modeling. This structure was tested on a nonlinear model. Basal area growth following a partial cutting was modeled with a Bertalanffy-Richards equation (*cf.* Richards 1959). The model was calibrated with permanent-plot data for two species in mixed stands: red spruce (*Picea rubens* Sarg.) and balsam fir (*Abies balsamea* (L.) Mill.). The discussion focused on the effectiveness of the tested structure as well as its potential and limitations. The Bertalanffy-Richards equation has been largely discussed in previous studies (e.g. Pienaar and Turnbull 1973).

## Statistical developments

### Statistical theory

According to the previously described context, a statistical nonlinear growth model calibrated with permanent-plot data can be expressed as follows:

$$y_{ij} = f(\mathbf{x}_{ij}; \boldsymbol{\beta}) + \varepsilon_{ij} \quad \text{for } i=(1,2,\dots,q) \text{ and } j=(1,2,\dots,m_i)$$

$$\boldsymbol{\varepsilon}_{i\bullet} = (\varepsilon_{i,1}, \varepsilon_{i,2}, \dots, \varepsilon_{i,m_i})^T \quad (1)$$

$$\boldsymbol{\varepsilon}_{i\bullet} \sim N_{m_i}(0, \mathbf{V}_i)$$

where  $y_{ij}$  is the dependent variable evaluated at measurement  $j$  of plot  $i$ ,  $\mathbf{x}_{ij}$  is vector of explanatory variables also evaluated at measurement  $j$  of plot  $i$ ,  $\boldsymbol{\beta}$  is a column vector of unknown parameters,  $\boldsymbol{\varepsilon}_{i\bullet}$  is a column vector of  $m_i$  within-plot error terms associated with plot  $i$ , and  $T$  denotes a matrix transposition. The elements of  $\boldsymbol{\varepsilon}_{i\bullet}$  are assumed to be normally distributed with mean 0 and variance-covariance  $\mathbf{V}_i$ .

Since many measurements were taken on the same plot, a serial correlation can reasonably be expected between the errors associated with the  $i$ th plot. Many structures already exist for modeling the within-subject covariance structure as previously mentioned. Among them, the time-continuous autoregressive structure is a generalization of the first-order autoregressive structure (Littell *et al.* 1996, p. 127; Pinheiro and Bates 2000, p. 236):

$$Cov(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma^2 \rho^{|t_{ij} - t_{ij'}|} \quad (2)$$

where  $\sigma^2$  is the residual variance,  $\rho$  is the correlation between two measurements spaced one unit of time apart, and  $|t_{ij} - t_{ij'}|$  is the absolute distance in time units between measurements  $j$  and  $j'$  of plot  $i$ . The more distant from each other the measurements, the less correlated they are. This structure and its derivatives have been largely used in forest growth and yield modeling (e.g. Goelz and Burk 1992; Gregoire *et al.* 1995; Beaumont *et al.* 1999; Nigh *et al.* 2002). Being continuous in time, covariance structure (2) no longer requires the measurements to be equally

spaced and balanced. Its efficiency with permanent-plot data has already been demonstrated in Gregoire *et al.* (1995). However, due to the assumption of homogeneous variances through parameter  $\sigma^2$ , this structure does not account for heteroscedasticity under its current form.

To take heteroscedasticity into account, the variance can be modeled through a function involving either the predicted values or some explanatory variables (Parresol 1993; Davidian and Giltinan 1995, §2.2.3; Littell *et al.* 1996, §8.2.4; Pinheiro and Bates 2000, §5.2). The power-of-the-mean function is a well-known example of a variance function based on predicted values:

$$\text{Var}(\varepsilon_{ij}) = \sigma^2 |f(\mathbf{x}_{ij}; \boldsymbol{\beta})|^{2\theta} \quad (3)$$

where  $\sigma^2$  is the residual variance,  $f(\mathbf{x}_{ij}; \boldsymbol{\beta})$  is the predicted value for measurement  $j$  of plot  $i$ , and  $\theta$  is a parameter to be estimated. Equation (3) is a generalization of the constant coefficient of variation model, for which  $\theta = 1$ , and the Poisson-like variance structure, for which  $\theta = 0.5$  (Davidian and Giltinan 1995, p. 23).

To take into account unequally spaced repeated measurements and heteroscedasticity, the continuous-time autoregressive covariance structure (2) and the variance function (3) can be merged. The general form of the resulting function expresses the within-plot covariance:

$$\text{Cov}(\varepsilon_{ij}, \varepsilon_{i'j'}) = v_{ijj'} = \sigma^2 (f(\mathbf{x}_{ij}; \boldsymbol{\beta}))^\theta (f(\mathbf{x}_{i'j'}; \boldsymbol{\beta}))^\theta \rho^{|t_{ij} - t_{i'j'}|} \quad (4)$$

where  $\sigma^2$ ,  $\boldsymbol{\beta}$ ,  $\theta$ , and  $\rho$  are assumed to be constant across the plots.

When  $j = j'$ , equation (4) reduces to the original variance function (3). From (4), the VC matrix of within-plot error terms ( $\mathbf{V}_i$ ) is obtained as follows:

$$\mathbf{V}_i = \begin{bmatrix} v_{i,1,1} & v_{i,1,2} & \cdots & v_{i,1,m_i} \\ v_{i,2,1} & v_{i,2,2} & \cdots & v_{i,2,m_i} \\ \cdots & \cdots & \cdots & \cdots \\ v_{i,m_i,1} & v_{i,m_i,2} & \cdots & v_{i,m_i,m_i} \end{bmatrix} \quad (5)$$

Assuming a generalization of (5) to all the plots, the overall VC matrix of the error terms ( $\mathbf{V}$ ) encompasses  $q$  diagonal blocks, with each block  $\mathbf{V}_i$  being the VC matrix of plot  $i$ . Usually, the plots are assumed to be independent from each other and the off-block elements within  $\mathbf{V}$  are set to 0.

### ***Estimation method and model diagnostics***

Estimation methods for models with parameterized VC matrices are mostly based on a likelihood approach (Gallant 1987, §5.5; Davidian and Giltinan 1995, p. 36; Littell *et al.* 1996, p. 498; Pinheiro and Bates 2000, §2.2). To optimize the likelihood function with respect to the generalized VC structure, a generalized least squares (GLS) regression can be performed (Davidian and Giltinan 1995, §2.3.4). The generalized least squares algorithm consists in generating a preliminary OLS or NLS estimator for the vector of parameters. This is done by calibrating the model with no VC structure. Then, the resulting parameter estimates are kept constant in order to estimate the VC parameters. Model parameters are estimated again by keeping the VC parameter estimates constant, and so forth until the convergence of the likelihood function on a maximum. At convergence, a GLS estimator is obtained. Most subalgorithms for the optimization of model and VC parameters are tedious. We will not describe them here. Readers can refer to Gallant (1987, §1.4) for further details about the algorithms used for estimating model parameters. The maximization of the likelihood function with respect to the VC parameters is usually performed with the expectation-maximization and/or the Newton-Raphson algorithms (*cf.* Lindstrom and Bates 1988; Wolfinger *et al.* 1994). These algorithms can be coded in any matrix languages such as IML in SAS system (SAS Institute Inc. 2002). The *gnls* function available in the S-PLUS or R software also performs such regressions (*cf.* Pinheiro and Bates 2000, §8.3.3; Pinheiro *et al.* 2004).

Once the VC parameters are estimated, the inversion of  $\mathbf{V}$  provides the marginal contribution of each observation to the likelihood of the model. The inversed matrix  $\mathbf{V}$ , *i.e.*  $\mathbf{V}^{-1}$ , is required to obtain the maximum likelihood estimator of  $\boldsymbol{\beta}$  and its VC matrix (Davidian and Giltinan 1995, §2.3.5):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{Y} \quad (6)$$

$$\text{var}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \hat{\boldsymbol{\Omega}} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1}$$

where  $\mathbf{X}$  is the Jacobian matrix, *i.e.* a matrix of model partial derivatives with respect to  $\boldsymbol{\beta}$ ,  $\mathbf{Y}$  is a column vector whose elements are  $y_{ij}$ ,  $\boldsymbol{\Omega}$  is the VC matrix of the parameter estimates, and the hat denotes the maximum likelihood estimator. A robust estimator of  $\boldsymbol{\Omega}$  can also be obtained through sandwich estimators (Hall and Bailey 2001; McCulloch and Searle 2001, p. 212).

In fact, matrix  $\mathbf{V}$  can be seen as a generalization for statistical models working with the assumptions of additive and normally distributed error terms. Under the assumptions of independent errors and homoscedasticity, matrix  $\mathbf{V}$  is set to  $\sigma^2 \mathbf{I}$ , with  $\mathbf{I}$  being the identity matrix. A calibration with  $\mathbf{V} = \sigma^2 \mathbf{I}$  reduces to a OLS or NLS regression according to the context.

Once the GLS estimator is obtained, the adequacy of the VC structure can be assessed with normalized residuals. A Cholesky decomposition of matrix  $\mathbf{V}$  is performed to calculate the vector of normalized residuals ( $\mathbf{r}_{norm}$ ) as follows (Pinheiro and Bates 2000, p. 239):

$$\mathbf{r}_{norm} = (\hat{\mathbf{C}}^{-1})^T (\mathbf{Y} - \hat{\mathbf{Y}}) \quad (7)$$

where  $\hat{\mathbf{Y}}$  is a column vector, whose elements are the predicted values  $f(\mathbf{x}_{ij}; \hat{\boldsymbol{\beta}})$ , and  $\hat{\mathbf{C}}$  is the upper triangle of the Cholesky decomposition of  $\hat{\mathbf{V}}$ . If matrix  $\mathbf{V}$  is properly parameterized, the normalized residuals should be independently and normally distributed with mean 0 and variance 1. Variance homogeneity is easily checked by plotting the normalized residuals against predicted values. The assumption of independent errors can be verified by calculating the empirical correlations among the normalized residuals. Then, the hypothesis that correlations are null is tested with confidence intervals. The semi-variogram technique also provides a measure of similarity among the within-plot normalized residuals (Gregoire *et al.* 1995).

In addition to diagnostic plots, likelihood ratio tests (LRT) can be performed to compare nested models that were fitted using the maximum likelihood method. A first model is considered nested in a second one, if this second model is identical to the first excepted for the inclusion of a new explanatory variable or a new VC parameter. The statistic, which is computed as  $2\log(L_2/L_1)$  with  $L$  being the likelihood of the model, follows a  $\chi^2$  distribution with  $k_2-k_1$  degrees of freedom, with  $k$  being the number of parameters. LRT are useful for assessing the significance of VC parameters (e.g. Fang and Bailey 2001; Hall and Bailey 2001). A significant probability associated with the statistic indicates that the additional parameter significantly improves the likelihood of the model.

Models calibrated with the same data can also be compared with the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (Littell *et al.* 1996, §3.2.2; Pinheiro and Bates 2000, §2.4.1). Both statistics are based on log-likelihood values penalized for the number of parameters. AIC and BIC can be considered as “parsimony” criteria: the smaller the value of the statistics, the better the model in terms of fit and simplicity. Since the penalty for parameters is greater in the BIC, this criterion tends to be more conservative than the AIC.

### ***The predictions***

Once calibrated, the purpose of a model is usually to provide predictions. A general prediction theorem is presented in Valliant *et al.* (2000, §2.2). First, let us define the indices  $r$  and  $s$  as the unobserved and the observed part of their respective matrices or vectors:

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{ss} & \mathbf{V}_{sr} \\ \mathbf{V}_{rs} & \mathbf{V}_{rr} \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_s \\ \mathbf{X}_r \end{bmatrix} \tag{8}$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_s \\ \mathbf{Y}_r \end{bmatrix}$$

$$\hat{\mathbf{Y}} = \begin{bmatrix} \hat{\mathbf{Y}}_s \\ \hat{\mathbf{Y}}_r \end{bmatrix}.$$

Information provided by the known part of  $\mathbf{V}$ ,  $\mathbf{Y}$ , and  $\hat{\mathbf{Y}}$  is used to enhance the predicted values calculated with the fixed part of the models. Let  $\tilde{\mathbf{Y}}_r$  be the empirical best predictor of  $\mathbf{Y}_r$ . This vector can be obtained as follows:

$$\tilde{\mathbf{Y}}_r = \hat{\mathbf{Y}}_r + \hat{\mathbf{V}}_{rs} \hat{\mathbf{V}}_{ss}^{-1} (\mathbf{Y}_s - \hat{\mathbf{Y}}_s) \quad (9)$$

where the last term represents the improvement computed with the information that is already available. The prediction error associated with vector  $\tilde{\mathbf{Y}}_r$  has a variance that originates from two different sources of variability, the variability due to the error terms, which is defined in  $\mathbf{V}$ , and the variability due to the parameter estimates, which is encompassed in  $\mathbf{\Omega}$ . A third source of variability, which is the variability due to VC parameter estimates, has not been integrated into the calculation of prediction variances yet. The error variance of the empirical best predictor is:

$$\text{var}(\tilde{\mathbf{Y}}_r - \mathbf{Y}_r) = \hat{\mathbf{V}}_{rr} - \hat{\mathbf{V}}_{rs} \hat{\mathbf{V}}_{ss}^{-1} \hat{\mathbf{V}}_{sr} + (\mathbf{X}_r - \hat{\mathbf{V}}_{rs} \hat{\mathbf{V}}_{ss}^{-1} \mathbf{X}_s) \hat{\mathbf{\Omega}} (\mathbf{X}_r - \hat{\mathbf{V}}_{rs} \hat{\mathbf{V}}_{ss}^{-1} \mathbf{X}_s)^T \quad (10)$$

where the last term represents the variance due to the parameter estimates, whereas the first two originate from the variance of the error terms. With the variance (10), confidence intervals can be calculated around vector  $\tilde{\mathbf{Y}}_r$ .

## The case study

### *The study site and the database*

The 16-km<sup>2</sup> Lake Édouard Experimental Forest (LEEF) (46°45'N, 72°56'W) is located in the Laurentian section (L.4a) of the Great Lakes-St. Lawrence River Region (*cf.* Rowe 1972). Created in 1918 for monitoring purposes, the LEEF has been part of La Mauricie National Park, Québec, Canada, since 1970. Its climate is cool and humid. The nearby Shawinigan locality has a 1971-2000 mean annual temperature of 4.5°C (Environment Canada 2002). January is the coldest month with an average daily temperature of -13.2°C. The warmest month is July with an average

of 19.6°C. The mean annual precipitation is 1069 mm and the mean annual snowfall averages 249 cm. The LEEF average elevation is about 350 m above sea level. The topography is a succession of low-altitude hills interspaced with large areas of flat land (Heimburger 1941; Robitaille and Saucier 1998). The predominant soil type is a glacial till.

The ecological features of the Laurentian section are similar to those of the well-known Acadian Forest Region (Rowe 1972). The spruce-fir-hardwood type described in Westveld (1953) is the most abundant forest type within the LEEF (Heimburger 1941; Ray 1956). It is usually located on lower slopes and well-drained flat lands (Heimburger 1941; Ray 1956; Grondin *et al.* 1996; Gosselin *et al.* 2001). Red spruce, balsam fir, yellow birch (*Betula alleghaniensis* Britton), and white birch (*Betula papyrifera* Marsh.) are the dominant species. Minor but common species are white spruce (*Picea glauca* (Moench) Voss.), eastern hemlock (*Tsuga canadensis* (L.) Carr.), eastern white pine (*Pinus strobus* L.), white cedar (*Thuja occidentalis* L.), red maple (*Acer rubrum* L.) and sugar maple (*Acer saccharum* Marsh.).

Between 1950 and 1957, an experimental diameter-limit cutting was conducted at the LEEF. The treatment aimed (i) to avoid over-cutting the softwoods, (ii) to favour spruce over fir, and (iii) to reduce the cutting cycle from 40 to 20-30 years (Hatcher 1959a). To meet the management goals, the diameter limits were set at 20 cm (8 in) for balsam fir and 41 cm (16 in) for red spruce and yellow birch at a stump height of 30 cm (12 in) (Ray 1956). Harvested volumes were estimated to 30-35% of the total volume of all stems larger than 8.9 cm (3.5 in) in diameter at breast height (*dbh*) (Ray 1956; Hatcher 1959a).

The growth following this partial cutting was monitored with previously established permanent plots. These 405-m<sup>2</sup> (0.1-acre) plots had been systematically distributed all over the LEEF territory using a 201-m (660-ft) grid design. Within each plot, all the trees greater than 1.3 cm (0.5 in) in *dbh* had been tallied by 2.5-cm (1-in) diameter class. After the 1950-57 partial cutting, successive measurements were done in 1956-57, 1967, 1994-96, and 2001-04.

A subsample of these permanent plots was used as a calibration data set in order to model the basal area growth following the 1950-57 cutting. The plots were

selected among those measured since 1994 (203 plots). To limit the ecological variability as much as possible, the study focused on the spruce-fir-hardwood forest type for which the largest number of plots was available. The data set included 96 plots for a total of 239 observations. Some characteristics of the selected plots are summarized in Table 4.

### ***The model***

In southern Quebec, partial cutting is mostly regulated by merchantable basal area growth. Merchantable basal area is here defined as the basal area of all trees greater than 8.9 cm in diameter at breast height (*dbh*, 1.3 m). This variable was selected as the dependent variable for the modeling. Merchantable basal area was first plotted against time since logging to obtain a general trend of the data, which supported the choice of a Bertalanffy-Richards equation (*cf.* Richards 1959) as model. Expressed as a function of time since logging, the original equation is defined as follows:

$$y_{ij} = A_i (1 - e^{-B_i t_{ij}})^{C_i} + \varepsilon_{ij} \quad (11)$$

where  $y_{ij}$  is the merchantable basal area at the measurement  $j$  of plot  $i$  ( $\text{m}^2 \cdot \text{ha}^{-1}$ ),  $t_{ij}$  is the time since logging at measurement  $j$  of plot  $i$  (yr),  $A_i$ ,  $B_i$ , and  $C_i$  are the general parameters of the function, and  $\varepsilon_{ij}$  is the error term. Parameters may either be constant or vary according to some explanatory variables common to the measurements of plot  $i$ . Parameter  $A_i$  sets the asymptote of the equation, whereas parameters  $B_i$  and  $C_i$  change the rate of growth as well as the inflection point.

Equation (11) must be modified to suit the partial cutting context. Actually, immediately after logging, *i.e.* at  $t_{ij}=0$ , the merchantable basal area is likely to be different from 0. For the purpose of this study, an additional parameter was included in the model so that the intercept would be different from 0. Basically, two alternatives exist. This additional parameter makes it possible to move the curve either along the time axis ( $x$ -axis) or along the merchantable basal area axis ( $y$ -axis). In this case study, the first option was preferred. The displacement along the time axis means that merchantable basal area growth reduces to a displacement of general yield curves

driven by other explanatory variables, such as stand and species densities. This additional parameter ( $D_i$ ) was included in the model as follows:

$$y_{ij} = A_i(1 - e^{B_i(t_{ij} + D_i)})^{C_i} + \varepsilon_{ij} \quad (12)$$

Previous works successfully related shifts of general yield curves to advance regeneration characteristics a few years after logging (Riopel 1999; Fortin *et al.* 2003b).

**Table 4 – Summary of the 1956-57 measurement for the 96 permanent plots of the spruce-fir-hardwood forest type.**

Species and characteristics	Minimum	Mean	Maximum	Standard deviation
Red spruce				
Merchantable basal area <sup>†</sup> (m <sup>2</sup> ·ha <sup>-1</sup> )	0.0	3.5	9.0	2.0
Merchantable density <sup>†</sup> (stems·ha <sup>-1</sup> )	0	236	1062	162
<i>mqd250</i> <sup>††</sup> (cm)	3.4	15.6	26.7	4.8
Density index ratio <sup>†††</sup>	0.01	0.31	0.78	0.16
Balsam fir				
Merchantable basal area <sup>†</sup> (m <sup>2</sup> ·ha <sup>-1</sup> )	0.2	2.2	7.1	1.6
Merchantable density <sup>†</sup> (stems·ha <sup>-1</sup> )	25	216	767	155
<i>mqd250</i> <sup>††</sup> (cm)	3.6	13.1	23.4	4.3
Density index ratio <sup>†††</sup>	0.02	0.21	0.57	0.12
All species				
Merchantable basal area <sup>†</sup> (m <sup>2</sup> ·ha <sup>-1</sup> )	3.9	12.4	31.2	4.9
Merchantable density <sup>†</sup> (stems·ha <sup>-1</sup> )	222	664	1408	276
Time since logging	1	3	6	1.9

<sup>†</sup>Merchantable: all trees greater than 8.9 cm (3.5 in) in diameter at breast height (*dbh*)

<sup>††</sup>*mqd250*: mean quadratic diameter of the 250 largest stems per hectare

<sup>†††</sup>Density index ratio (DIR): as defined in equation (14).

After a few preliminary trials, the general parameters of equation (12) were set to:

$$\begin{aligned}
 A_i &= e^{b_1 + b_2 DIR_{i,1}} \\
 B_i &= b_3 \\
 C_i &= b_4 \\
 D_i &= b_5 mqd250_{i,1} - t_{i,1} + 1
 \end{aligned}
 \tag{13}$$

where  $DIR_{i,1}$  is a density index ratio evaluated at first measurement of plot  $i$ ,  $mqd250_{i,1}$  is the mean quadratic diameter of the 250 largest stems per hectare at the first measurement of plot  $i$  (cm), and  $t_{i,1}$  is the time after logging for the first measurement of plot  $i$ . This latter variable acts upon the shift of the curve as a correction factor since the first measurements were not done at the same time after logging (Table 4).

The density index ratio ( $DIR$ ) we used in this case study was computed as a ratio between the stand density index ( $SDI$ ) calculated for the selected species and the all-species  $SDI$ . For convenience, Reineke's (1933) rule was used for calculating the different  $SDI$ , although there might be some variations at the species level (Pretzsch and Biber 2005). From this rule, the variable  $DIR_{i,1}$  is obtained as follows:

$$DIR_{i,1} = \frac{DSP_{i,1}}{DST_{i,1}} \left( \frac{mqdsp_{i,1}}{mqdst_{i,1}} \right)^{1.605}
 \tag{14}$$

where  $DSP$  and  $DST$  are respectively the species and the all-species densities (stems·ha<sup>-1</sup>), and  $mqdsp$  and  $mqdst$  are respectively the species and the all-species mean quadratic diameter (cm), with all these variables evaluated at the first measurements of plot  $i$ . Equation (12) with its parameters as defined in (13) was independently calibrated for red spruce and balsam fir.

To take into account repeated measurements and heteroscedasticity, VC structure (5) was included in the regression. Likelihood ratio tests (LRT) were performed to check the significance of VC parameters. Normalized residuals were also plotted to ensure the structure was adequate. Moreover, GLS predictions were compared with NLS predictions in order to illustrate the consistency and the reliability

of the approach. Predictions and their confidence intervals were computed with equations (9) and (10). The predictions were calculated *a priori*, *i.e.* assuming the subsequent measurements were unknown. For instance, predictions for the first measurement were computed with no additional information to enhance the prediction. The observed prediction error of the first measurement was used to enhance the prediction for the second measurement, and so forth.

## Results

Results of the likelihood ratio tests (LRT) and both AIC and BIC statistics are shown in Table 5. For both species, the AIC and the BIC clearly indicated that the best model was the one including VC structure (5). LRT demonstrated the significance of the VC parameters  $\theta$  and  $\rho$ . Adding variance function (3) to a NLS-calibrated model significantly improved the likelihood of the model. Moreover, adding covariance structure (2) over variance function (3) also resulted in a significant improvement of the likelihood of the model.

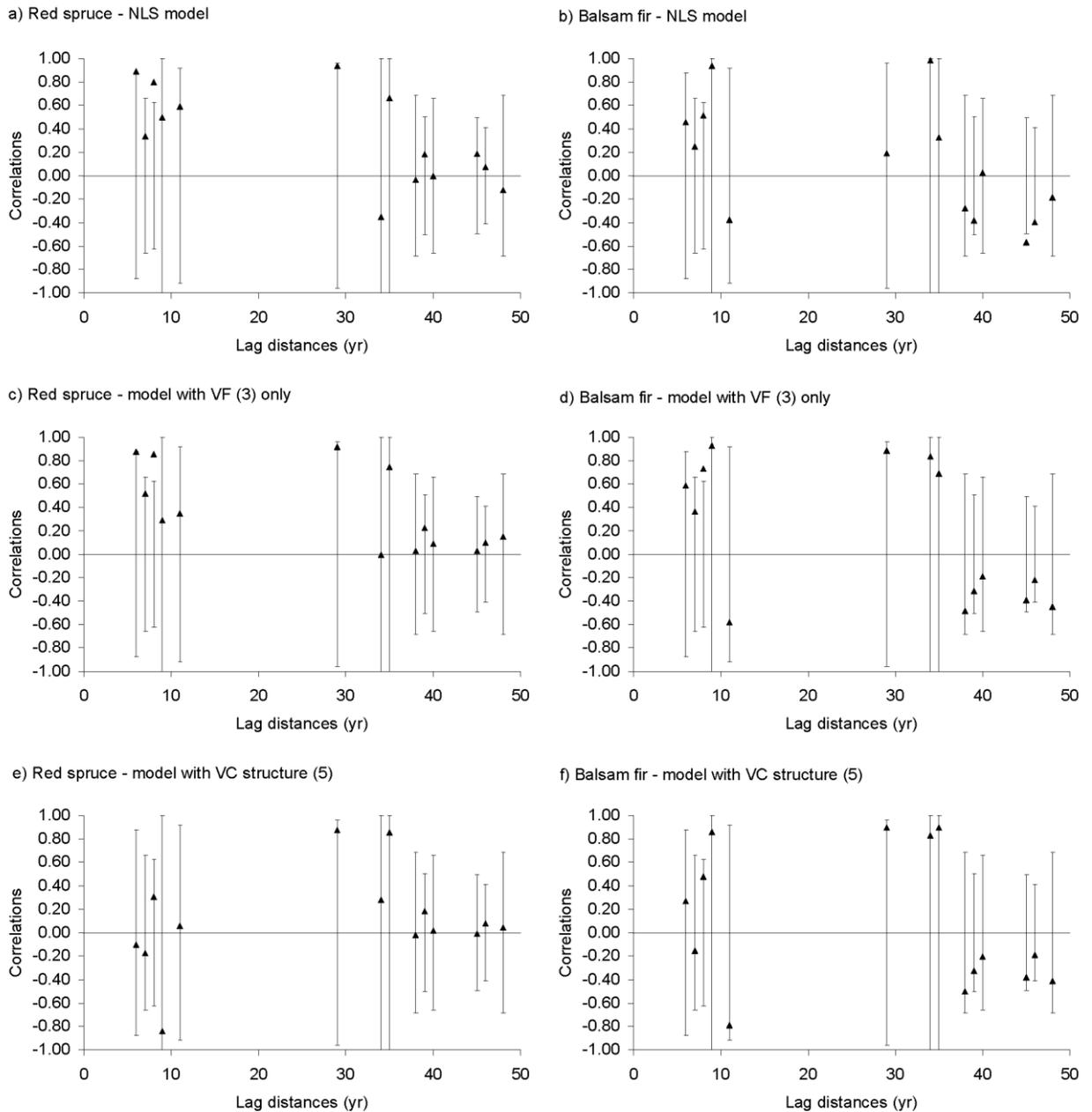
**Table 5 – Comparison of the goodness of fit for models with different variance-covariance (VC) features.**

Model	VC features	$k^\dagger$	AIC	BIC	Log-likelihood	LRT <sup>††</sup>	$Pr > \chi^2$
Red spruce							
1	None (NLS)	6	1031.04	1051.90	-509.52		
2	VF <sup>†††</sup> (3) only	7	913.40	937.73	-449.70	1 vs 2	<0.0001
3	VC structure (5)	8	882.92	910.73	-433.46	2 vs 3	<0.0001
Balsam fir							
1	None (NLS)	6	931.95	952.81	-459.97		
2	VF <sup>†††</sup> (3) only	7	799.83	824.17	-392.92	1 vs 2	<0.0001
3	VC structure (5)	8	790.15	817.96	-387.08	2 vs 3	0.0006

<sup>†</sup> $k$ : number of parameters

<sup>††</sup>LRT: likelihood ratio test

<sup>†††</sup>VF: variance function



**Figure 3 – Observed correlations among the within-plot normalized residuals and 1%-confidence intervals around the null hypothesis  $H_0: \text{Corr}(\varepsilon_{ij}, \varepsilon_{ij'})=0$  (the triangles indicate the observed correlations).**

For both species, the normalized residuals did not exhibit major departures from the assumption of normally distributed error terms with homogeneous variances. The observed correlations among the normalized residuals as well as the 1%-

confidence intervals around the null hypothesis  $H_0: \text{Corr}(\varepsilon_{ij}, \varepsilon_{ij'}) = 0$  are illustrated in Figure 3 for the three models of both species, *i.e.*, the NLS model, the model with VF (3) only, and the model with VC structure (5). Normalized residuals of both the NLS model and the model with VF (3) only exhibited a few significant correlations for some lags (Figure 3a, b, c, d). VC structure (5) seemed adequate since no observed correlations were significantly different from 0 at 1% after including the VC structure (Figure 3e, f).

**Table 6 – Parameter estimates for model (13) calibrated with VC structure (5) (asymptotic standard errors are shown between parentheses).**

Parameter estimates	Red spruce	Balsam fir
$b_1$	3.703 (0.556)**	1.092 (0.152)**
$B_2$ (density index ratio effect)	0.4463 (0.1752)*	1.451 (0.364)**
$b_3$	$-7.429 \times 10^{-3}$ ( $3.249 \times 10^{-3}$ )*	$-4.874 \times 10^{-2}$ ( $1.874 \times 10^{-2}$ )**
$b_4$	2.376 (0.281)**	2.788 (0.382)**
$b_5$ (mdq250 <sup>†</sup> effect)	3.348 (0.274)**	2.199 (0.660)**
$\theta$ (variance parameter)	0.804	1.275
$\rho$ (correlation parameter)	0.944	0.900
$\sigma^2$ (residual variance)	0.158	0.135

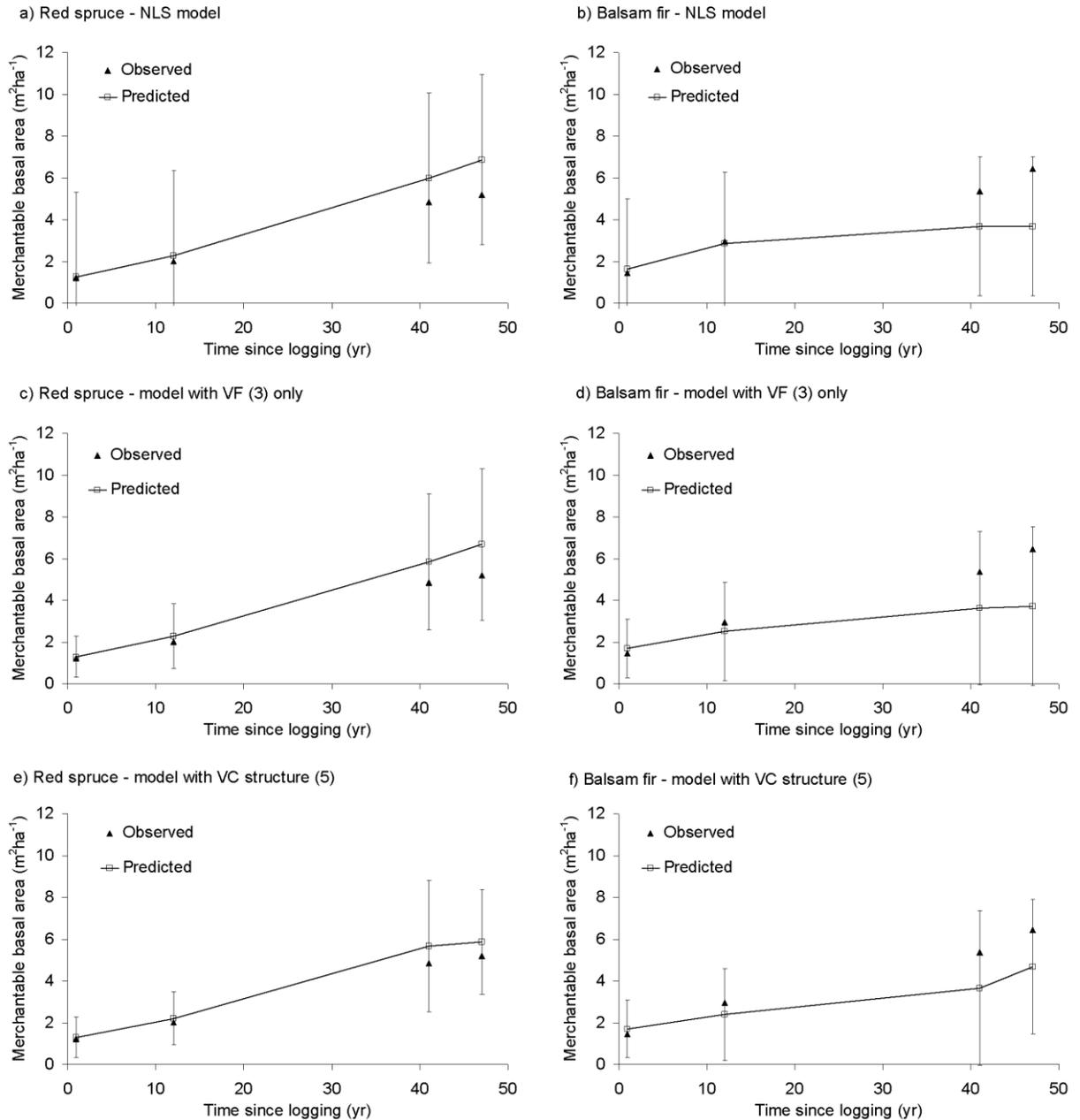
\*significant at 5%

\*\*significant at 1%

<sup>†</sup>mdq250: mean quadratic diameter of the 250 largest stems per hectare

Parameter estimates of the “best” model, *i.e.* the model with VC structure (5), are shown in Table 6. All model parameter estimates were significantly different from 0 at 5%, at least. Also, estimates of  $b_4$  were significantly different from 1, meaning that the models could not be simplified by omitting this parameter. For both species, the estimates of  $b_2$  indicated a positive effect of the density index ratio (*DIR*) on basal area growth. Increases in species density index with respect to the all-species density index resulted in a higher maximum basal area, *i.e.* a higher model asymptote. The estimates of  $b_5$  also indicated a positive effect of mean quadratic diameter of the 250

largest stems per hectare ( $mdq250$ ) on the shift of the yield curve along the time axis. As expected, the greater the mean quadratic diameter of the 250 largest stems per hectare, the greater the displacement of the curve.



**Figure 4 – Predictions and their 5%-confidence intervals for the different models of each species.**

The predictive abilities of each model are illustrated in Figure 4 for both species. Plot 1 was selected as an example. Obviously, the NLS confidence intervals are the widest intervals, since the regression assumes homogeneous variances (Figure 4a,b). On the other hand, models with variance function (3) produce confidence intervals more consistent with the heteroscedastic pattern of the data (Figure 4c,d). Note that OLS models and models with variance function (3) lead to similar predicted values. The relevance of VC structure (5) is clearly demonstrated in Figure 4e,f. The structure provides more accurate predictions and narrower confidence intervals, especially for the fourth measurement of each species. This indicates that previous error predictions can be used to enhance the current prediction and reduce its error.

## Discussion

This case study clearly demonstrates that VC structure (5) may adequately account for heteroscedastic and correlated error terms in a context of growth modeling using a whole stand approach. The goodness of fit of the model can be assessed through the normalized residuals, which are expected to be normally and independently distributed with homogeneous variances. For both species, there were no major departures from these assumptions and, consequently, the GLS estimator provided parameter estimates that were consistent and asymptotically normal with unbiased variances. If the number of observations is large, the approach makes it possible to select the explanatory variables according to their true significance level, a common problem with NLS regressions (Gregoire *et al.* 1995). Moreover, as the estimates and their standard errors are calculated from the original scale of the dependent variable, there is no need for a reverse transformation, which facilitates the interpretation of the inferences.

The parameterization of a VC matrix of the error terms ( $\mathbf{V}$ ) is a solution to departures from the assumptions of homogeneous variances and independent error terms. The approach proposed in this study is not the only way to parameterize this matrix. In fact, some simultaneously calibrated equations (*cf.* Gallant 1987, §6) and mixed-effects (both random and fixed) models (*cf.* Pinheiro and Bates 2000, §5.1.2)

are additional statistical tools that make parameterization possible. In forestry, the hierarchical approach based on mixed models has been largely promoted in growth modeling (Gregoire *et al.* 1995; Fang and Bailey 2001; Fang *et al.* 2001; Hall and Bailey 2001). Basically, the hierarchical approach consists in specifying random effects associated with each level of grouping within the data set. The random effects of a given level are considered nested in the next level of grouping. In the context of permanent-plot data, inclusion of random effects in a growth model relies on the assumption that some unobserved influences act upon plot  $i$  in a constant fashion over time (Gregoire 1987). This constant influence is usually associated with a fixed-effect parameter and is expressed as an error varying at the plot level. Many examples of multilevel mixed models can be found in Pinheiro and Bates (2000, §2.1.2).

In practice, a positive correlation can persist even after including random effects (*e.g.* Gregoire *et al.* 1995; Garber and Maguire 2003). In such cases, a continuous-time autoregressive covariance structure added to random effects gives better results (Gregoire *et al.* 1995; Fang and Bailey 2001; Garber and Maguire 2003). This study clearly demonstrates that a VC structure alone may be sufficient to address the matters of heteroscedasticity and repeated measurements. In fact, the addition of random effects on the parameters of model (12) was tested with and without VC structure (5) and the resulting models exhibited lower likelihood values, nonsignificant LRT, as well as higher AIC and BIC statistics. These results were not surprising since there was no evidence of constant trends due to some unobserved plot factors among the residuals in Figure 3. In this case study, models with VC structure (5) only were better models than those with random effects.

Actually, the two ways of parameterizing the VC matrix, *i.e.* VC structures and random effects, are not so different from each other. Even if random effects are referred to as between-subject errors, their inclusion in a mixed model indirectly leads to the specification of a VC structure for the within-subject error terms (Littell *et al.* 1996, §3.2.2; Pinheiro and Bates 2000, §5.1.2). For linear models, Littell *et al.* (1996, p. 93) demonstrated that a random effect on the intercept and a compound symmetry VC structure yield exactly the same parameterization, with the same results and

inferences. In practice, random effects and VC structures may interfere with each other in the model specification and the resulting VC matrix may be overparameterized (Pinheiro and Bates 2000, p. 204). Even if the likelihood converges on a maximum value, the model might not be the best one. For the lowest level of grouping, it is strongly recommended that this more complex model be compared with simpler models calibrated either with a VC structure only or random effects only. The comparison might reveal that a VC structure only is better than random effects or random effects with a VC structure. As outlined by Littell *et al.* (1996, p. 93), the contribution of the subject may be negligible in some repeated measures experiments.

Even if a model exhibits the lowest AIC and BIC statistics, there is no certainty that its VC parameterization adequately accounts for heteroscedasticity and correlated error terms. The only way to know whether the correction is appropriate is to plot the observed variances and correlations among the normalized residuals. Indeed, the lower the AIC and BIC statistics, the more likely the assumptions of independent error terms and homogeneous variances are to be expected. However, the selection of the final model should not be based on these statistics only. Moreover, the choice of the VC parameterization should take into account other factors besides the information criteria, such as the interpretability of the VC parameters (Pinheiro and Bates 2000, p. 409).

Gregoire (1987) interpreted plot random effects as influences due to variables mostly of an ecological nature, such as drainage and soil fertility. In this case study, the lack of constant correlations among the normalized residuals can be explained by the low variability of the ecological conditions, since all the plots belong to the same ecological group. Actually, random effects due to ecological variables are likely to be embedded in the effects of mortality and recruitment, which are thought to be far more important at the stand level. Except for the density index ratio (*DIR*), which can partially account for these effects, there is no variable that directly controls mortality and recruitment at the stand level. The main effect of having no control on these components is the disruption of the correlation between two measurements as the distance between them increases. As time goes on, basal area underestimations are

likely to be offset by recruitment whereas mortality may compensate for overestimations. The continuous autoregressive covariance structure in VC structure (5) clearly fits the disruption of correlation along with time.

In addition to correction for heteroscedasticity and correlated error terms, a VC structure can be used to improve model predictions. With mixed models, a similar enhancement is obtained by calculating a best unbiased linear predictor (BLUP) of random effects (*cf.* McCulloch and Searle 2001, §9). In forestry, some studies have already demonstrated the improvement of the predictions by reinserting a BLUP into a mixed model (*e.g.* Fang and Bailey 2001; Hall and Bailey 2001). Compared with previous work based on the hierarchical approach, the enhanced predictions in this case study are not as accurate. In fact, the accuracy of the predictions depends on the purpose of the model. Usually, growth models are used to predict future forest conditions. Therefore, the future basal area measurements are unknown and the enhancement must be done *a priori*, *i.e.* only with the observations prior to the prediction. This approach is the one we used in this case study. Indeed, more accurate enhancements would have been obtained if we had assumed that future measurements were available. Most studies based on mixed models performed the enhancement using this latter assumption. The BLUP are calculated with all the available observations, even observations subsequent to the prediction.

## Conclusions

Progress in understanding the relationships between growth and environmental factors is often hindered by the lack of reliable and consistent statistical models. Most of the existing growth models ignore the impact of heteroscedasticity and unequally spaced repeated measurements on the reliability of prediction errors. In this study, a VC structure based on a variance function and a covariance structure has been proposed to take into account both statistical issues.

Although the hierarchical approach is strongly recommended in forestry, this case study demonstrates that a VC structure alone may provide the flexibility for simultaneously considering heteroscedasticity and correlated error terms. As a

consequence, we recommend checking the assumption of constant influences due to the subjects before using random effects as proposed by the hierarchical approach. This can be done by comparing the AIC and BIC statistics for a mixed model and a model calibrated with a VC structure only. Plotting the normalized residuals can also give some hints about the appropriate correlation structure.

In addition to the flexibility it provides, the parameterization of a VC matrix also enhances the predictions. In forestry, this enhancement has already been demonstrated with mixed models (e.g. Fang and Bailey 2001; Fang *et al.* 2001; Hall and Bailey 2001). In this study, we demonstrated it is also possible to improve the predictions through a VC structure.

## CHAPITRE 3

CHAPITRE 3 - Calibrating a generalized diameter distribution model with mixed effects.

Manuscript accepted in  
Forest Science

Authors: Mathieu Fortin<sup>a</sup>, Chhun-Huor Ung<sup>b</sup>,  
Louis Archambault<sup>b</sup>, and Jean Bégin<sup>a</sup>

<sup>a</sup>Department of Wood and Forest Sciences, Laval University

<sup>b</sup>Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre

## Abstract

This paper proposes a mixed model approach as an alternative to the traditional parameter prediction method in diameter distribution modeling. Unthinned and thinned plots established in mixed stands were used for calibrating a generalized diameter distribution model. The model was based on a two-parameter Weibull cumulative density function (*cdf*). This *cdf* was linearized through a complementary log-log link function. Plot random effects were included to account for autocorrelation and the dependent variable, *i.e.*, cumulative stem frequency, was assumed to follow a binomial distribution.

With respect to the parameter prediction method, this mixed model approach may generate a consistent estimator with unbiased variance for the vector of parameters when the variance-covariance matrix of the error terms is properly parameterized. Moreover, the approach enables a better assessment of the different variance components. For the whole group of plots, it provides a predicted average diameter distribution. At the plot level, the random effects can be considered as a departure from this average distribution. As long as the diameter distributions of the individual plots do not exhibit major departures from unimodality, the method proposed in this study should be used to calibrate generalized diameter distribution models.

## Key words

Two-parameter Weibull function, random effects, link function, generalized linear mixed models.

## Introduction

The diameter distribution of trees is an important factor in forest management as it is often used to characterize forest structure (Smith *et al.* 1997, p.24). Harvest options are evaluated by projecting the diameter distributions into the future and comparing the resulting forest structures (*e.g.* Bowling *et al.* 1989). To address the concern for ecosystem diversity and stability, new forest policies also require information on forest structure at different scales (Lähde *et al.* 1999). As a consequence, an effort is needed to develop diameter distribution models suitable for different levels of analysis.

Until now, three major trends have distinguished investigations of diameter distributions. The first approach has focused on developing and testing different, mostly nonlinear, statistical models to fit the diameter distributions of individual plots (*e.g.* Bailey and Dell 1973; Cao and Burkhart 1984; Borders *et al.* 1987; Zhang *et al.* 2001, 2003). The second approach consists in comparing different groups of plots through indices such as skewness, kurtosis, or the Shannon Index (*cf.* Shannon and Weaver 1949, p.19). These indices are initially calculated for each individual plot. Then, an analysis of variance (ANOVA) is performed on the individual values to identify significant differences among groups of plots (*e.g.* Groot and Horton 1994; Kuuluvainen *et al.* 1996; Linder *et al.* 1997; Linder 1998; Lähde *et al.* 2001). The third approach attempts to link the previous two by generalizing a single population diameter distribution model to a group of plots (*e.g.* Magnussen 1986; Bowling *et al.* 1989, Maltamo *et al.* 1995, 2000; Maltamo 1997; Siipilehto 1999, 2000; Robinson 2004). These generalized diameter distribution models aim at providing information at two levels, namely the individual plot and the group of plots.

From a statistical standpoint, calibrating a generalized diameter distribution model from a sample of plots is often hindered by autocorrelated data. This problem is discussed extensively by García (1992). In most regular forest surveys, the plot is the sampling unit and many diameter measurements are taken within the plot. As a result, we can expect this kind of grouping to induce plot random effects in the model.

Even though they are often considered as negligible, the effects of autocorrelated data on statistical inferences and predictions may be important. The variances estimated by traditional methods, such as ordinary least squares, are biased (Sullivan and Clutter 1972; Gregoire 1987; Gregoire *et al.* 1995). Hence, the selection of the best covariates is hampered (Gregoire *et al.* 1995) and the fit of the model cannot be correctly assessed. Consequently, the model predictions are inefficient, as their confidence intervals are biased.

Over the last two decades, the forestry literature has provided some examples of the flexibility of mixed-effects models for considering the departures from the assumption of independent errors (*e.g.* Gregoire 1987; Gregoire *et al.* 1995; Hall and Bailey 2001; Garber and Maguire 2003). Gregoire *et al.* (1995) pointed out that mixed models were still infrequently used in growth modeling due to their complexity and the previous lack of software suitable for estimation. The current situation in diameter distribution modeling is even worse. The stem frequencies that are inherent to diameter distributions follow a discrete distribution, whereas most statistical tools for mixed model regressions are designed to work with normally distributed variables.

A mixed model approach applied to a generalized diameter distribution model would make it possible to consider different levels of uncertainty due to the data structure. Moreover, this approach would enable valid statistical inferences since it accounts for autocorrelated data. The purpose of this study was to demonstrate the feasibility of a mixed model approach applied to diameter distribution modeling. Thinned and unthinned plots that were located in mixed stands of red spruce (*Picea rubens* Sarg.), balsam fir (*Abies balsamea* (L.) Mill.), and yellow birch (*Betula alleghaniensis* Britton.) were used to calibrate a two-parameter Weibull function with a mixed model approach. The discussion focuses on the advantages and limitations of the approach.

## Material and method

### *The study site*

The Ouareau River Observation Area (OROA) (46°26'N, 74°10'W) was established in 1953 by the Department of Northern Affairs and National Resources of the Government of Canada. The study site, which is 13-km<sup>2</sup> in area, is located within the Laurentian Section (L.4a) of the Great Lakes-St.Lawrence Forest Region (*cf.* Rowe 1972, p.96). The nearby municipality of Saint-Donat-de-Montcalm has an average daily temperature of 3.2°C (running average for 1971-2000, Environment Canada 2002). January is the coldest month whereas July is the warmest with respective average temperatures of -13.6 and 18.1°C. Total annual precipitation averages 1128 mm, with a mean annual snowfall of about 295 cm. The OROA has an average elevation of 515 m above sea level. It is characterized by a succession of low-altitude hills. Most soils are glacial tills. The red spruce-balsam fir-yellow birch association is the most abundant forest type (Hatcher 1954). It is usually located on well-drained flat lands and lower slopes (Heimbürger 1941).

Between 1948 and 1956, a diameter-limit cutting was conducted over the entire OROA territory. All spruce and fir greater than 17.8 cm (7 in) in diameter at stump height (30 cm) were logged for pulpwood. As early as 1953, permanent plots were established within the observation area for monitoring purposes. In these 404.7-m<sup>2</sup> (0.1-ac) plots, all trees with a diameter at breast height (*dbh*) greater than 1.3 cm (0.5 in) were tallied in 2.5-cm (1-in) classes. At that time, about one-third of the OROA was still uncut and the plots established within that area could be considered as a representative sample of the conditions prior to thinning (Hatcher 1954). By 1956, the cutting had encompassed almost the entire study site, with the uncut patches being mostly inaccessible or uninteresting from a commercial standpoint (Hatcher 1954). During the period of 2002-2004, only the thinned plots located in the red spruce-balsam fir-yellow birch forest type were visited and remeasured.

The data set used in this case study included two groups of plots, the first of which was formed from the 1953 unthinned plots located in the red spruce-balsam fir-

yellow birch association (44 plots). The 2002-04 survey provided an additional 22 plot measurements, which represented mature forest conditions 50 years following thinning for the same association. The average densities and basal areas of both groups are summarized in Table 7.

**Table 7 – Density and basal area of the unthinned and the thinned groups (stems larger than 1.3 cm (0.5 in) in diameter at breast height).**

	Minimum	Mean	Maximum
Unthinned group (44 plots)			
Basal area (m <sup>2</sup> ·ha <sup>-1</sup> )	10.3	29.3	58.1
Density (stems·ha <sup>-1</sup> )	445	2182	8970
Mean quadratic diameter (cm)	6.5	14.3	20.5
Thinned group (22 plots)			
Basal area (m <sup>2</sup> ·ha <sup>-1</sup> )	26.4	35.1	49.4
Density (stems·ha <sup>-1</sup> )	766	3007	8525
Mean quadratic diameter (cm)	6.3	14.0	24.0

### ***Statistical developments***

Over the last several decades, the Weibull function has been one of the most, if not the most, popular models used in diameter distribution modeling (Johnson 2000, p.260). An extensive description of the model can be found in Bailey and Dell (1973). The three-parameter version of this probability density function (*pdf*) is very flexible and suits a vast array of random variables such as *dbh* (Vanclay 1994, p.23). As diameters are often measured in finite diameter classes, the cumulative density function (*cdf*) is used instead of the original *pdf*. Let *i* and *j* be, respectively, the plot and diameter class indices such that *i*=(1, 2,..., *n*) and *j*=(1, 2,..., *m<sub>i</sub>*). The three-parameter Weibull *cdf* can be expressed as follows:

$$\Pr(dbh \leq d_{ij}) = \pi_{ij} = 1 - e^{-\left(\frac{d_{ij} - \lambda_i}{\lambda_i}\right)^{\alpha_i}} \quad (1)$$

where  $\pi_{ij}$  is the probability of having a  $dbh$  equal or lower than  $d_{ij}$ , which is defined as the upper limit of the  $j^{\text{th}}$  diameter class of the  $i^{\text{th}}$  plot, and  $\alpha_i$ ,  $\lambda_i$ , and  $\chi_i$  are respectively the shape, the scale, and the location parameters of the  $cdf$  associated with the  $i^{\text{th}}$  plot. When the lower bound of the diameter distribution is already known, let this limit be defined as  $d_{\min,i}$ . The limit value can be substituted, in turn, for the location parameter ( $\chi_i$ ), thereby reducing equation (1) to a two-parameter Weibull  $cdf$  (e.g. Magnussen 1986; Zhang *et al.* 2001).

Thus, for any tree of the  $i^{\text{th}}$  plot, the probability of having  $dbh \leq d_{ij}$  would be  $\pi_{ij}$ . Likewise, the probability of having  $dbh > d_{ij}$  would be the complement of the probabilities, namely  $1 - \pi_{ij}$ . Consequently, given the number of trees with  $dbh \leq d_{ij}$ , let  $s_{ij}$  be this variable, which is expected to follow a binomial distribution:

$$\Pr(s_{ij}; \pi_{ij}, r_i) = \binom{r_i}{s_{ij}} \pi_{ij}^{s_{ij}} (1 - \pi_{ij})^{r_i - s_{ij}} \quad (2)$$

where  $r_i$  is the total number of trees in the  $i^{\text{th}}$  plot. To refer to more classical terminology,  $s_{ij}$  and  $r_i$  are the number of successes and trials, respectively.

Discrete variables are usually modeled through generalized linear models. An extensive description of these models is found in McCulloch and Searle (2001, §5). In addition to the specification of a distribution for the dependent variable (binomial in our case), these models require a transformation of the dependent variable in order to obtain a projection that can be modeled in a linear fashion. As the probability  $\pi_{ij}$  is assumed to follow a Weibull  $cdf$ , we expect the transformation to provide a link between equation (1) and its derived linear form. Such a linearization is obtained through a complementary log-log transformation (*cf.* Collett 2003, p.57):

$$\begin{aligned} \ln(-\ln(1 - \pi_{ij})) &= \beta_{0,i} + \beta_{1,i} \ln(d_{ij} - d_{\min,i}) \\ \beta_{0,i} &= -\alpha_i \ln(\lambda_i) \\ \beta_{1,i} &= \alpha_i \end{aligned} \quad (3)$$

where  $d_{\min,i}$  has been substituted for the location parameter ( $\chi_i$ ). Let the function  $g(\pi_{ij})$  define the complementary log-log transformation of probability  $\pi_{ij}$  as shown in (3).

All the observations of the data set would be independent if each plot had a single  $d_{ij}$ . However, the data typically provide much more information. Actually, the upper limit of any diameter class can be used as  $d_{ij}$ . Besides, considering more than one  $d_{ij}$  per plot implies serially correlated data. To take into account this possible correlation, plot random effects can be integrated into the model. The resulting generalized linear mixed model (*glmm*) can be expressed as follows:

$$g(\boldsymbol{\pi}_{i\bullet}) = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{u}_i$$

$$\boldsymbol{\pi}_{i\bullet} = (\pi_{i,1}, \pi_{i,2}, \dots, \pi_{i,m_i})^T \quad (4)$$

$$\mathbf{u}_i \stackrel{iid}{\sim} N(0, \mathbf{G})$$

where  $\mathbf{X}_i$  is a matrix of explanatory variables, including an intercept and  $\ln(d_{ij} - d_{\min,i})$ ,  $\boldsymbol{\beta}$  is a vector of unknown parameters,  $\mathbf{Z}_i$  is a design matrix indicating the variables with which the random effects are associated,  $\mathbf{u}_i$  is a vector of unobserved random effects due to the  $i^{\text{th}}$  plot, and  $T$  denotes a transposition. We expect  $\mathbf{u}_i$  to be normally distributed with mean 0 and variance-covariance  $\mathbf{G}$ .

With linear mixed models, the final estimator of  $\boldsymbol{\beta}$  is provided by the maximization of a marginal likelihood function (Pineiro and Bates 2000, p.62). This marginal likelihood function is obtained by integrating the original likelihood function involving the random variables with respect to the distribution of these variables (Collett 2003, p.272). Once the model has converged, empirical best linear unbiased predictors (*eBLUP*) of the random effects are calculated from the design matrix  $\mathbf{Z}_i$ , the vector of residuals, and the variance-covariance matrix associated with this vector (McCulloch and Searle 2001, p.255). Nevertheless, with a *glmm*, the link function hinders the integration of the likelihood function analytically (McCulloch and Searle 2001, p.231).

An alternative consists of using a quasi-likelihood function instead of a marginal likelihood. With simple generalized linear models, the quasi-likelihood approach has several attractive features, including model robustness and less

restrictive assumptions (McCulloch and Searle 2001, p.281). Furthermore, according to this approach, the vector of mean probabilities  $\boldsymbol{\pi}_{i\cdot}$  is conditional on the random effects:

$$E[\mathbf{p}_{i\cdot} | \mathbf{u}_i] = \boldsymbol{\pi}_{i\cdot}$$

$$\mathbf{p}_{i\cdot} = (p_{i,1}, p_{i,2}, p_{ij}, \dots, p_{i,m_i})^T \quad (5)$$

$$p_{ij} = \frac{s_{ij}}{r_i} .$$

Consequently, maximization of the quasi-likelihood function requires that both the cumulative vector of random effects ( $\mathbf{u}$ ) and the vector of parameters ( $\boldsymbol{\beta}$ ) be estimated during optimization. Solutions for  $\boldsymbol{\beta}$  and  $\mathbf{u}$  are obtained from the quasi-likelihood function by iteratively solving the equations (Wolfinger and O'Connell 1993):

$$\begin{bmatrix} \mathbf{X}^T \mathbf{W} \mathbf{X} & \mathbf{X}^T \mathbf{W} \mathbf{Z} \\ \mathbf{Z}^T \mathbf{W} \mathbf{X} & \mathbf{Z}^T \mathbf{W} \mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T \mathbf{W} \mathbf{p}^* \\ \mathbf{Z}^T \mathbf{W} \mathbf{p}^* \end{bmatrix}$$

$$\mathbf{W} = \mathbf{D}^T \mathbf{R}^{-1} \mathbf{D}$$

$$\mathbf{p}^* = g(\hat{\boldsymbol{\pi}}) + (\mathbf{p} - \hat{\boldsymbol{\pi}}) \mathbf{D}^{-1} \quad (6)$$

$$\mathbf{D} = [\partial \boldsymbol{\pi} / \partial g(\boldsymbol{\pi})]$$

$$\mathbf{R} = \mathbf{R}_\pi^{1/2} \mathbf{A} \mathbf{R}_\pi^{1/2}$$

where  $\mathbf{R}_\pi$  is a diagonal matrix whose elements are the variances associated with the vector of predicted probabilities ( $\boldsymbol{\pi}$ ),  $\mathbf{A}$  is a covariance matrix, and the hat denotes the estimate obtained from the optimization. Since we assume a binomial distribution, the variance  $\text{var}(\pi_{ij})$  depends on the value of  $\pi_{ij}$  as follows:

$$\text{var}(\pi_{ij}) = \pi_{ij}(1 - \pi_{ij}) / r_i . \quad (7)$$

Matrix  $\mathbf{A}$  is usually set to  $\phi \mathbf{I}$ , with  $\mathbf{I}$  being the identity matrix and  $\phi$  being a parameter equivalent to the residual variance. To refer to traditional mixed models, matrix  $\mathbf{A}$

corresponds to the within-plot variance-covariance matrix of the error terms. Since the plots are assumed to be independent from each other, matrix  $\mathbf{A}$  is diagonally blocked, with each block  $\mathbf{A}_i$  being the variance-covariance matrix of plot  $i$ . In cases of persistent correlations, a covariance structure may be generalized to all the blocks  $\mathbf{A}_i$  (Vonesh and Carter 1992). Common covariance structures, such as the compound symmetry, the autoregressive, and the moving average, are extensively described in Pinheiro and Bates (2000, §5.3.3).

In order to check the validity of the assumptions of independently and binomially distributed errors with homogeneous variances, a vector of scaled Pearson residuals ( $\mathbf{e}_n$ ) can be derived from mixed model theory (*cf.* Pinheiro and Bates 2000, p.239):

$$\mathbf{e}_n = (\mathbf{C}^{-1})^T (\mathbf{p} - \boldsymbol{\pi}) \quad (8)$$

where  $\mathbf{C}$  is the upper triangle resulting from the Cholesky decomposition of matrix  $\mathbf{R}$  as shown in (6). If a covariance structure is specified within matrix  $\mathbf{A}$ , the calculation of the scaled residuals through matrix  $\mathbf{C}$  takes into account this structure. Consequently, it provides the marginal contribution of each observation to the error of the model. If the covariance structure is appropriate, the scaled residuals should not be correlated.

### ***Model specifications***

There is no limit to the number of diameter classes in a plot, since each one of them can be a potential  $d_{ij}$ , which could be used for calibration. However, large numbers of diameter classes for which no trees were tallied may be cumbersome during the optimization of the quasi-likelihood function. In this case study, the 40.6-cm (16-in) diameter class represented the 99<sup>th</sup> percentile of the observed average diameter distribution. Larger diameter classes were not considered. Consequently, 16 distinct  $d_{ij}$  were used per plot, each one corresponding to the upper limit of one of the first 16 diameter classes. During the model calibration, all trees were considered without distinguishing among species.

Model (3) was first simplified by setting  $d_{min,i}$  to 1.3 cm (0.5 in) for all plots. Preliminary calibrations were performed on each individual plot to obtain estimates of the intercept ( $\beta_{0,i}$ ) and the slope ( $\beta_{1,i}$ ) of the model. Then, these estimates were plotted against several covariates to identify the explanatory variables, which were the most likely to be significant.

After a few trials, the following explanatory variables and random effects were selected and tested:

$$g(\pi_{ij}) = \beta_{0,i} + \beta_{1,i} \ln(d_{ij} - 1.3)$$

$$\beta_{0,i} = b_0 + b_1 density_i + b_2 group_i + u_{0,i} \quad (9)$$

$$\beta_{1,i} = b_3 + b_4 basalarea_i + b_5 group_i + u_{1,i}$$

$$\begin{bmatrix} u_{0,i} \\ u_{1,i} \end{bmatrix} \stackrel{iid}{\sim} N_2(0, \mathbf{G})$$

where  $d_{ij}$  is a diameter limit as defined above (cm),  $density_i$  is the density of plot  $i$  (stems·ha<sup>-1</sup>),  $group_i$  a dummy variable taking a value of either 0 or 1 for the respective unthinned and the thinned groups,  $basalarea_i$  is the basal area of plot  $i$  (m<sup>2</sup>·ha<sup>-1</sup>),  $b_0$ ,  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$ , and  $b_5$  are the model parameters, and  $u_{0,i}$  and  $u_{1,i}$  are the plot random effects, which follow a bivariate normal distribution with mean 0 and variance-covariance  $\mathbf{G}$  as previously mentioned.

Likelihood ratio tests were performed to test the significance of the random effects. Scaled residuals were plotted to verify the assumptions of binomial distribution, homoscedasticity, and independent error terms. Also, the predictors of the random effects were plotted in order to check for normality. The analyses were performed with the GENMOD procedure and the GLIMMIX macro available in SAS (Littell *et al.* 1996, §11; SAS Institute Inc. 2002).

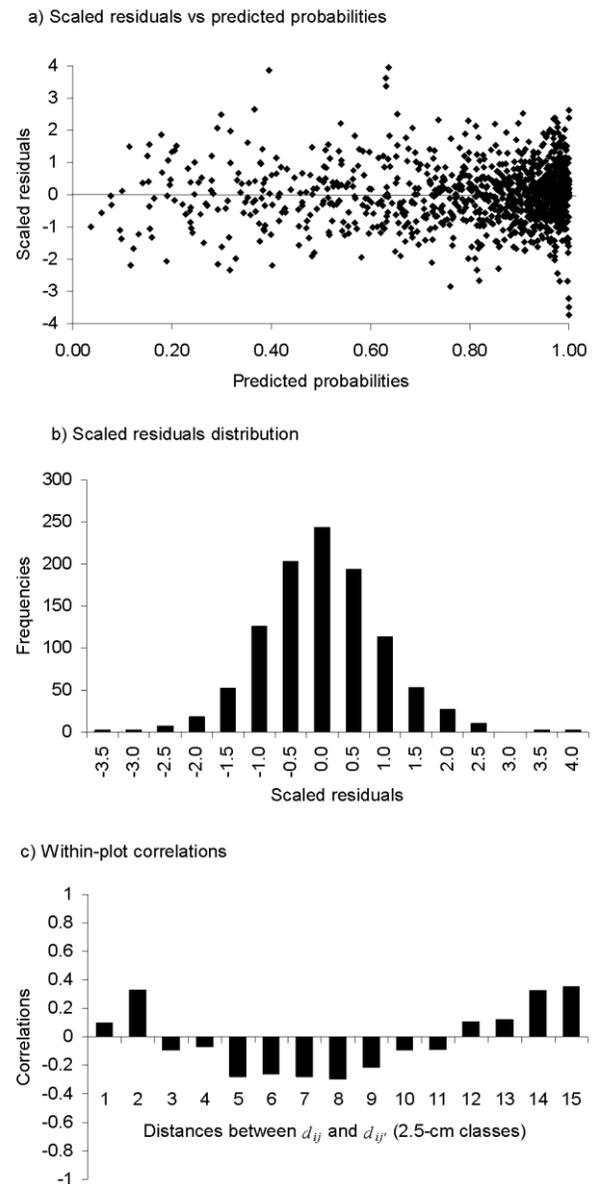
## Results

Even when plot random effects were included in the model, within-plot residuals still exhibited persistent correlations. The two-band Toeplitz covariance structure (*cf.* SAS Institute Inc. 2002) seemed to be the most appropriate structure. So, the structure of matrix  $\mathbf{A}$  was modified according to the Toeplitz covariance structure as follows:

$$\mathbf{A} = \mathbf{I}_n \otimes \phi \begin{bmatrix} 1 & \rho & 0 & 0 & \dots & 0 & 0 \\ \rho & 1 & \rho & 0 & \dots & 0 & 0 \\ 0 & \rho & 1 & \rho & \dots & 0 & 0 \\ 0 & 0 & \rho & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 & \rho \\ 0 & 0 & 0 & 0 & \dots & \rho & 1 \end{bmatrix}_{16 \times 16} \quad (10)$$

where  $\rho$  is the Toeplitz parameter and  $\otimes$  denotes the Kronecker product.

The *glmm* was recalibrated with the inclusion of this new feature. The residuals exhibited homogeneity of variances (Figure 5a), and there was no departure from the assumption of binomially distributed errors (Figure 5b). Moreover, none of the correlations was greater than 0.40 in terms of their absolute values (Figure 5c). Both random effects ( $u_{0,i}$  and  $u_{1,i}$ ) did not exhibit major departures from the assumption of normality.



**Figure 5 - Analysis of model residuals.**

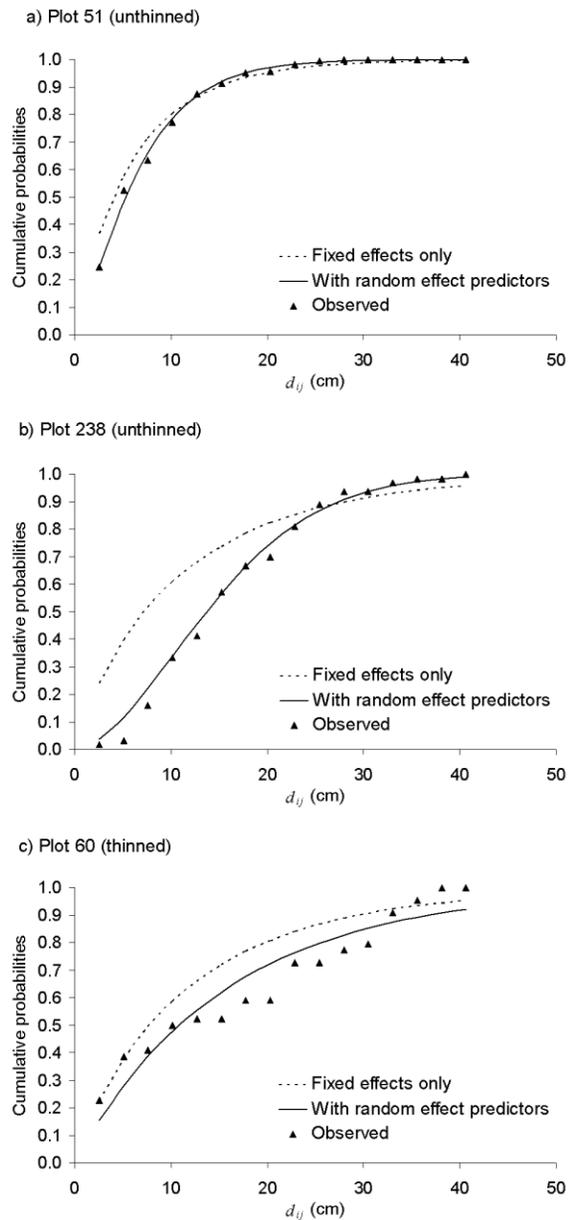
The estimates resulting from the final calibration of model (8) are shown in Table 8. Except for the single effect of the group and the interaction  $\ln(d_{ij}-1.3) \times \text{group}$ , the parameter estimates related to the fixed effects all exhibited probabilities close to or less than 2%. Referring to equation (3), a single effect is related to both the scale and the shape parameters ( $\lambda_i$  and  $\alpha_i$ ) of the Weibull *cdf* as shown in (1). Interactions with  $\ln(d_{ij}-1.3)$  indicate an effect on the shape ( $\alpha_i$ ) of the *cdf* only. Consequently, the plot density had a significant effect on both the scale and the shape of the Weibull function. On the other hand, the basal area had a significant effect on the shape of the *cdf* only. With regards to random effects and the covariance structure, the probabilities associated with the Toeplitz parameter estimate and the standard deviations of both  $u_{0,i}$  and  $u_{1,i}$  were all less than 0.01%. Therefore, including the random effects and the correlation structure in the model significantly improved its likelihood, justifying these additional features.

**Table 8– Parameter estimates resulting from the calibration of model (8) (probability of significance shown between parentheses).**

Parameter and effect	Estimate
Fixed effects and covariates	
$b_0$	-2.621 (<0.0001)
$b_1$ – density	$2.320 \times 10^{-4}$ (0.0014)
$b_2$ – group	-0.2970 (0.2470)
$b_3$ – $\ln(d_{ij}-1.3)$	0.6411 (<0.0001)
$b_4$ – $\ln(d_{ij}-1.3) \times \text{basal area}$	0.01089 (0.0205)
$b_5$ – $\ln(d_{ij}-1.3) \times \text{group}$	-0.02362 (0.7724)
Plot random effects	
Standard deviation of $u_{0,i}$	0.9185 (<0.0001)
Standard deviation of $u_{1,i}$	0.2843 (<0.0001)
Covariance parameter	
Two-band Toeplitz parameter	0.4581 (<0.0001)

Each random effect can be interpreted as the variability of a fixed effect at the plot level. Actually,  $u_{0,i}$  and  $u_{1,i}$  express this variability around the estimates of  $b_0$  and  $b_3$ , respectively. A 95% confidence interval around the fixed effect is easily computed by multiplying  $\pm 1.96$  by the standard deviation of the appropriate random effect. In this case study, the contribution of the random effects to the fit of the model was not negligible, since their 95% confidence intervals represented a variability of  $\pm 69\%$  and  $\pm 87\%$  around their respective fixed effects.

The fit of the model was also assessed by plotting the observed and predicted probabilities for each plot. Predictions based solely on fixed effects usually were not accurate. This result was expected since the variability at the plot level was important. Basically, three typical patterns were observed. These are illustrated at Figure 6. First, plot 51 illustrates one of the few examples of reliable predictions based on fixed effects only (Figure 6a). Second, in most cases, the inclusion of the empirical predictors of the random effects (*eBLUP*) resulted in more accurate predictions, with the fit of the model being nearly perfect (Figure 6b). The improvement was more important for the smallest  $d_{ij}$ , for which the probabilities were the most variable. The third pattern is represented by a few plots for which the fit



**Figure 6 – The three typical patterns related to the fit of the model.**

with the random effect predictors still exhibited non-negligible errors (Figure 6c). The observed cumulative probabilities of these plots had a very irregular shape that the Weibull *cdf* was unable to fit perfectly.

## Discussion

Ordinary least squares (OLS) regressions are based on the assumptions of homogeneous, normally, and independently distributed errors (Steel *et al.* 1997, §7.10). In this context, diameter distribution modeling is often hindered by data that are measured on a discrete scale, such as diameter class. Even when tree diameters are precisely measured, the observation remains a frequency, which cannot be negative. Consequently, the assumption of normally distributed errors about the mean probability is irrelevant, as it may result in confidence intervals including inconsistent values, namely probabilities outside of the range (0,1).

Even if the OLS regressions produce in a model with good fit, the errors are no longer entirely independent of the predicted probabilities. For predicted probabilities close to 1, the distribution of the residuals tends to be left-skewed. Likewise, right-skewed residuals are associated with predicted probabilities close to 0. Consequently, regular statistical analyses are likely to provide biased inferences since the assumptions of normality and independent errors no longer hold.

Generalized linear models enable the analysis of a response variable whose distribution belongs to the exponential family of distributions, such as the binomial, the Poisson, and the normal distribution (McCulloch and Searle 2001, §5.2). This analysis requires the specification of an appropriate distribution and a link function. The link function makes it possible to model the transformed probabilities as a linear function of explanatory variables over an interval  $(-\infty, +\infty)$  (Collett 2003, p.56). Regular statistical tests that assume normality, homogeneous variances, and independent errors are then suitable for such transformed probabilities. Although the calibration of *glmm* is more complex than regular mixed models, statistical tools such as the GLIMMIX macro available in SAS (Littell *et al.* 1996, §11) perform relatively well with one level of grouping within the data structure. These tools facilitate the application of

mixed models theory to diameter distribution modeling. This study is an example of this application.

The complementary log-log link function allows the Weibull *cdf* to be linearized. The linearized form results in an intercept and a slope related to the natural logarithm of the diameter at breast height. To calibrate the *glmm*, the probability generated by the Weibull *cdf* is used as a probability of success in a binomial distribution for which the number of trees in the plot represents the number of trials. To our knowledge, no other study has addressed the matter of generalized diameter distribution models in this fashion.

Since there is more than one successes/trials ratio per plot, random effects have to be included in the model. To ensure that all appropriate random effects are present, Hall and Bailey (2001) recommended that all basic fixed-effect parameters should be allowed to vary across the different levels of grouping in the data set. The conversion of the linearized Weibull *cdf* into a mixed model form is relatively easy to perform, since the model only has required two random effects, *i.e.*, one associated with the intercept ( $\beta_{0,i}$ ) and one related to the slope parameter ( $\beta_{1,i}$ ).

Even though the appropriate random effects were included, the within-plot residuals still exhibited significant correlations in our case study. This was essentially due to the nature of the response variable. As cumulative frequencies are used, a deficit in a particular diameter class is likely to induce a deficit in the following diameter classes. A persistent correlation between the residuals of a mixed model is not an entirely new phenomenon, as it has been already observed in growth and stem taper modeling (Gregoire *et al.* 1995; Garber and Maguire 2003). To overcome this problem, Gregoire *et al.* (1995) and Garber and Maguire (2003) used an autoregressive structure, which assumes a positive correlation decreasing as the distance increases between the observations (*cf.* Pinheiro and Bates 2000, p.228). In this case study, we attempted to specify this structure, without success. The negative correlations observed for some distances between  $d_{ij}$  and  $d_{ij'}$  (Figure 5c) may explain why the model failed to converge. A two-band Toeplitz covariance structure accounts for the correlation between a given observation, namely the ratio of successes/trials,

and its nearest neighbour in terms of  $d_{ij}$ . Adding this structure to the *glmm* significantly improved the likelihood and decreased the correlations between the residuals, although a pattern of persistent correlations could still be observed (Figure 5c). The implementation of a covariance structure enabling a sinusoidal pattern might solve the problem of persistent correlations. Multinomial models with cumulative ordinal responses use covariance structures (*cf.* McCullagh and Nelder 1989, §5.3.2), which might be adapted to the context of diameter distribution modeling. However, the gain of reliability might be rather small since the persistent correlations resulting from the Toeplitz structure are already low.

The scaled residuals of the model provided strong evidence that the assumptions of binomial distribution, homogeneous variances, and independent errors were globally respected (Figure 5). Consequently, the estimator for the vector of parameters can be considered as consistent and asymptotically normal (Pinheiro and Bates 2000, p.81) with nearly unbiased variance. This feature is particularly useful because it enabled more reliable statistical comparisons among different groups of plots. In this case study for instance, there were no significant differences between the unthinned and the thinned groups. Besides, two continuous explanatory variables, *i.e.*, plot basal area and plot density, had significant effects on the slope and intercept of the model. These two variables accounted for most of the variability between the two groups.

Until now, the classical approach for the calibration of generalized diameter distribution models has been two-stage regression<sup>1</sup> (*cf.* Steimer *et al.* 1984), an early method for integrating random effects into nonlinear models (Vonesh and Carter 1992). The model is first calibrated for each individual plot. Then, the resulting parameter estimates are regressed against several plot covariates. In the forestry literature, two-stage regression is referred to as the parameter prediction or

---

<sup>1</sup> Two-stage regression should not be confused with two-stage least squares. The first is an early approach in nonlinear mixed-effects modeling, whereas the second is a common statistical method in econometrics for solving simultaneous equations.

parameter recovery approach (*cf.* Vanclay 1994, p.23). Different fitting techniques for the parameter prediction approach are discussed by Robinson (2004).

Although it accounts for plot random effects, two-stage regression has two major flaws in addition to the assumption of normally distributed error terms. First, the regression performed during the second stage does not take into account the original data, which were used during the first-stage individual calibrations. Consequently, there is no guarantee of convergence on a maximum for the likelihood function. Secondly, it requires that the diameter distribution model converge for each individual plot during the first stage. As a consequence, it is possible that data from some subjects may be ignored, if the first stage regression fails to converge for certain plots (Vonesh and Carter 1992).

A *glmm* offers a solution to these problems. Basically, it combines the aforementioned two stages into a single step. The variability between the individual plots is no longer assessed with a second regression, but rather through random effects. As the plot random effects and the fixed-effect parameters are simultaneously estimated, the *glmm* ensures the convergence of the likelihood function on a local or a global maximum, whenever convergence is reached. Moreover, there is no possibility of omitting data from certain plots as no preliminary calibrations are required. In fact, the *glmm* provides a predicted average diameter distribution for one or many groups of plots, depending on the model specifications. The random effects allow a departure from this average distribution at the plot level. Therefore, the model enables the analysis of diameter distributions at two different levels, namely the individual plot and the group of plots. This assessment of different error components is a direct response to the issue of variance estimation outlined by García (1992).

A major limitation of the method is the subjectivity associated with selecting the number of  $d_{ij}$  to be used for the model calibration. Actually, an infinite number of diameter limits is available. However, problems of convergence arise when the ratio of successes/trials approaches 1. In order to avoid such problems, we recommend using evenly-spaced limits up to the 99<sup>th</sup> percentile of the observed average diameter distribution as we did in this case study. Using such a boundary does not limit the

ability of the model to predict the stem frequency above the 99<sup>th</sup> percentile. Actually, the modeling approach considers the number of stems above this limit as the difference between the total number of trees and the number of trees below the 99<sup>th</sup> percentile. Although it provides a stem frequency for the trees above this threshold, the prediction of their diameters relies on the assumption that the diameter distribution of larger trees follows the trend observed in smaller diameters. If this assumption does not hold, the results are limited to the distribution below the selected boundary. Modelers should keep in mind that this recommendation about a threshold is a trade-off between convergence and model accuracy as regards to the diameter predictions of largest stems. Even though the predicted diameter distributions of different groups of plots are similar, differences might be observed with respect to the distribution of the largest stems, as reported by Fortin *et al.* (2003a).

In this case study, the predictions based on fixed effects only tended to be inaccurate at the plot level (Figure 6b,c). We did not investigate all possible relationships between parameters of the Weibull *cdf* and the explanatory variables, nor did we include all possible explanatory variables in the model. To demonstrate the feasibility of the mixed model approach, we tried to keep the analysis as simple as possible. Indeed, nonlinear modeling and additional covariates might improve the accuracy of the predictions based on fixed effects only, but this remains to be investigated. The results of past studies relying on two-stage regression might prove a good framework for future developments for generalized diameter distribution modeling calibrated with mixed effects.

## Conclusions

The application of mixed model theory to diameter distribution modeling is possible through a generalized linear mixed model (*glmm*). If the random effects and/or the covariance structure provide the appropriate correction for correlation, a *glmm* generates an estimator for the vector of parameters that is consistent and asymptotically normally distributed with unbiased variance. This feature enables not only valid statistical tests on parameter estimates, but also coherent diameter distribution predictions compatible with stand variables. Thereby, it is possible to

compare different groups of plots on a statistical basis by including dummy variables in the model. A *glmm* also makes it possible to obtain an average diameter distribution for a group of plots. The random effects can be seen as departures from this average distribution at the plot level.

In this case study, the covariance structure was quite complex and none of the available predefined structures could fit the observed pattern perfectly. As a result, the within-plot residuals still exhibited low persistent correlations, even after specifying random effects and a covariance structure. Note that the residuals of *glmm* are conditional on the random effects. Therefore, the correlations between the within-plot residuals do not consider the correlation absorbed by the random effects. For instance, if the random effects account for most of the within-plot variability, the correlations calculated with the conditional residuals may be negligible overall. In this case study, the random effects clearly absorbed a large part of variability (Figure 6b) and the specification of a covariance structure was a matter of fine fitting. Indeed, the implementation of a covariance structure that fits the observed sinusoidal pattern would provide a consistent estimator with truly unbiased variance for the vector of parameters. Multinomial models theory might prove useful for the specification of such a structure. Meanwhile, the variance of the estimator should be considered as “nearly” unbiased.

On the other hand, the mixed model approach is subject to convergence problems when many ratios of successes/trials approach 1, *viz.*, when no tree has a diameter greater than the limit  $d_{ij}$ . As a consequence, a threshold must be set to avoid repeating such ratios unnecessarily. In our case study, we demonstrated that the 99<sup>th</sup> percentile of the empirical average diameter distribution may be an appropriate upper value. The method also is limited by the link function. This constraint may be important, especially if observed diameter distributions are irregular. Although a finite mixture of two Weibull functions may fit a bimodal distribution (Zhang *et al.* 2001), problems arise when this model has to be linearized. A complex *pdf* cannot be linearized with regular link functions. Developments in both irregular distribution modeling and complex link functions might prove a significant advance. Meanwhile, as long as the individual distributions do not exhibit major departures from

unimodality, the proposed approach should be used as it ensures compatibility between plot scale and landscape scale.

## CHAPITRE 4

CHAPITRE 4 - Une matrice de transition à deux niveaux basée sur des distributions de probabilités : l'exemple de l'épinette rouge (*Picea rubens* Sarg.) et du sapin baumier (*Abies balsamea* (L.) Mill.) en peuplements mixtes.

Manuscrit à soumettre à la  
Revue canadienne de la recherche forestière

Auteurs : Mathieu Fortin<sup>a</sup>, Louis Archambault<sup>b</sup>,  
Jean Bégin<sup>a</sup> et Chhun-Huor Ung<sup>b</sup>

<sup>a</sup>Département des sciences du bois et de la forêt, Université Laval

<sup>b</sup>Ressources naturelles Canada, Service canadien des forêts, Centre de foresterie  
des Laurentides

## Résumé

Cet article propose une nouvelle approche pour calibrer des matrices de transition à deux niveaux. La méthode suppose que les probabilités de la matrice sont dictées par trois distributions de probabilités (de Bernoulli, binomiale et de Poisson) dont les paramètres sont des fonctions de variables explicatives mesurées à l'échelle de l'arbre et de la placette. Cette approche a été testée à partir des données de placettes permanentes établies dans des peuplements mixtes de sapin baumier (*Abies balsamea* (L.) Mill.) et d'épinette rouge (*Picea rubens* Sarg.) après une coupe partielle. La matrice a été calibrée avec la méthode du maximum de vraisemblance pour chacune des deux essences individuellement. Pour les deux essences, une validation du modèle a été réalisée.

Bien que la méthode proposée soit plus complexe, elle comporte plusieurs avantages par rapport à l'approche classique, à savoir une cohérence de probabilités à l'intérieur de la matrice, un nombre restreint de paramètres à estimer et une plus grande flexibilité quant à la durée de la projection. Par ailleurs, l'exercice de validation démontre que les biais du modèle sont faibles et les estimés de paramètres sont en partie conformes à ce que l'on connaît de l'autécologie des deux espèces en cause.

Dans le cas présent, les prédictions de la matrice de transition démontrent que le recrutement, l'accroissement et la survie du sapin baumier sont favorisés par des coupes partielles de forte intensité. Pour éviter des changements de composition dans les peuplements de seconde venue, les coupes partielles de faible intensité sont préconisées car elles permettraient de contrôler l'expansion de la régénération de sapin baumier.

## Mots-clés

Modèle de croissance, maximum de vraisemblance, coupe partielle, distributions de probabilités.

## Introduction

Dans le domaine forestier, les modèles de croissance sont souvent utilisés pour anticiper les conséquences de certaines décisions d'aménagement sur les futures caractéristiques des peuplements forestiers. L'approche traditionnelle en modélisation forestière, dite par peuplement entier (*cf.* Vanclay 1994), est basée sur des variables mesurables à l'échelle du peuplement comme la surface terrière ou le volume par exemple (*e.g.* Sullivan et Clutter 1972; Pothier et Savard 1998; Fortin *et al.* 2003b). Bien que cette approche soit attrayante par sa simplicité (Vanclay 1994), il n'en demeure pas moins que les prédictions obtenues ne sont valables qu'à l'échelle du peuplement.

Depuis quelques décennies, de nouvelles méthodes ont été expérimentées dans le but d'obtenir l'information à l'échelle de l'arbre individuel. Des modèles dépendants ou indépendants des distances ont été développés pour prédire l'accroissement en diamètre de chacune des tiges (*e.g.* Lessard *et al.* 2001; Ricker et del Río 2004). Ces modèles nécessitent cependant une information détaillée sur chacune des tiges d'une placette, une information qui n'est pas toujours disponible lorsque les arbres sont dénombrés par classe de diamètre.

Face à cette contrainte quant à la nature de données recueillies, des approches intermédiaires ont également été développées. Les matrices de transition (*e.g.* Bruner et Moser 1973), les systèmes d'équations (*e.g.* Ek 1974) et les fonctions de densité de probabilités (*e.g.* Bowling *et al.* 1989) constituent un compromis entre les deux approches précédentes. Ces techniques de modélisation sont moins exigeantes en termes de données et fournissent l'information par classe de dimension, en l'occurrence par classe de diamètre. Ce manuscrit est consacré au développement des matrices de transition dans un contexte où l'information disponible ne permet pas la modélisation de la croissance à l'aide d'une approche par arbre individuel. En foresterie, les premiers modèles de croissance basés sur des matrices de transition ont été développés vers la fin des années 1960 (*e.g.* Usher 1966, 1969).

Les matrices de transition ont longtemps été critiquées pour leur manque de flexibilité du fait qu'elles ne peuvent s'adapter à différents contextes (Vanclay 1994). Pour surmonter ce problème, les matrices de transition à deux niveaux (« two-stage transition matrices ») ont été développées (e.g. Solomon *et al.* 1986). Au lieu de présumer que les probabilités de la matrice sont constantes, on pose l'hypothèse qu'elles sont plutôt le résultat d'équations incluant des variables explicatives à l'échelle de l'arbre et de la placette. Ainsi, une matrice de transition à deux niveaux peut s'adapter à différentes conditions initiales.

D'un point de vue statistique, la calibration d'une telle matrice est complexe. Elle requiert des données de placettes permanentes à l'intérieur desquelles les tiges sont individuellement identifiées (Vanclay 1994). Chacune des probabilités de la matrice peut alors être estimée à partir de la proportion de tiges répondant à un événement particulier tel que l'accroissement d'un certain nombre de classes de diamètre ou la mortalité par exemple. Cependant, dans la majorité des cas, les données disponibles ne permettent pas d'estimer tous les éléments de la matrice. On préfère alors limiter le nombre de probabilités à estimer en utilisant des intervalles de projection courts et constants. La calibration d'une matrice de transition avec des intervalles constants est également compliquée du fait que les mesurages de placettes permanentes qui fournissent les données nécessaires à la calibration sont souvent inégalement espacés dans le temps (Gregoire *et al.* 1995).

Un autre problème vient du fait que les probabilités de la matrice sont souvent estimées de façon indépendante. Ce faisant, il devient possible que la somme des probabilités pour les tiges appartenant à la même classe de diamètre en début de projection soit différente de 1. Ceci implique que des arbres peuvent disparaître ou apparaître sans raison apparente durant la projection. Dans une telle situation, Solomon *et al.* (1986) suggèrent un rééchelonnage des probabilités. Cependant, les effets de cette technique sur les prédictions du modèle n'ont toujours pas été évalués.

L'objectif de cette étude est de proposer une méthode pour calibrer des matrices de transition à deux niveaux (i) qui permettent de prédire la croissance sur des

intervalles de temps inégaux et (ii) dont les probabilités de transition sont cohérentes. L'approche suggérée suppose que les probabilités de la matrice de transition ne sont pas dictées par des équations ou des constantes, mais plutôt par des distributions de probabilités. Ces distributions de probabilités sont régies par des paramètres qui peuvent être exprimés à leur tour comme des fonctions de variables mesurées à l'échelle de l'arbre et de la placette. En somme, trois équations seraient suffisantes pour calculer toutes les probabilités d'une matrice de transition à deux niveaux. Cette approche a été testée avec un cas d'étude. Des données de placettes permanentes situées dans des peuplements mixtes de sapin baumier (*Abies balsamea* (L.) Mill.) et d'épinette rouge (*Picea rubens* Sarg.) ont été utilisées pour calibrer le modèle. La discussion traite des avantages et des limites de l'approche suggérée ainsi que des processus écologiques qui sous-tendent la croissance des deux essences en peuplements mixtes.

## Développements du modèle

Les matrices de transition à deux niveaux (« two-stage transition matrices ») se distinguent des matrices de transition classiques du fait que leurs probabilités ne sont pas constantes, mais plutôt définies par des équations comprenant des variables explicatives (Solomon *et al.* 1986). Ainsi, une matrice de transition à deux niveaux peut s'adapter à différentes conditions initiales selon la nature des variables explicatives. Celles-ci peuvent être mesurées à l'échelle de l'arbre individuel ou de la placette.

De façon générale, une matrice de transition  $\mathbf{M}$  permet de réaliser une projection d'un dénombrement de tiges dans le temps. La calibration d'un tel modèle requiert donc des couplets de données. Chacun de ces couplets est composé d'un dénombrement initial ( $\mathbf{Y}_0$ ) et d'un dénombrement final ( $\mathbf{Y}_1$ ) séparés par un intervalle de temps  $t$ . Les dénombrements  $\mathbf{Y}_0$  et  $\mathbf{Y}_1$  sont en fait des vecteurs dont les éléments représentent des fréquences observées à l'intérieur de certaines classes de diamètre. Définissons  $k$  comme un indice des classes de diamètre tel que  $k=(0,1,2,3,\dots, d)$ . La classe de diamètre  $k=0$  est ici une classe de diamètre fictive représentant les recrues,

soit les tiges qui sont absentes de  $Y_0$  et qui apparaissent dans  $Y_1$  au terme de l'intervalle. À partir de cet indice, une matrice de transition peut être définie comme suit :

$$[1] \quad Y_1^T = Y_0^T M$$

$$\begin{bmatrix} 0 \\ y_{1,1} \\ y_{1,2} \\ y_{1,k} \\ \dots \\ y_{1,d} \end{bmatrix}^T = \begin{bmatrix} r \\ y_{0,1} \\ y_{0,2} \\ y_{0,k} \\ \dots \\ y_{0,d} \end{bmatrix}^T \times \begin{bmatrix} 0 & Pr_{0,1} & Pr_{0,2} & Pr_{0,k} & \dots & Pr_{0,d} \\ 0 & Pr_{1,1} & Pr_{1,2} & Pr_{1,k} & \dots & Pr_{1,d} \\ 0 & 0 & Pr_{2,2} & Pr_{2,k} & \dots & Pr_{2,d} \\ 0 & 0 & 0 & Pr_{k,k} & \dots & Pr_{k,d} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & Pr_{d,d} \end{bmatrix}$$

où  $y_{0,k}$  et  $y_{1,k}$  sont les éléments de  $Y_0$  et  $Y_1$ , soit les fréquences observées dans la classe de diamètre  $k$  au début et à la fin de l'intervalle,  $r$  est le nombre de recrues observées,  $Pr_{k,k'}$  est la probabilité qu'une tige de la classe de diamètre  $k$  au début de l'intervalle passe à la classe de diamètre  $k'$  à la fin de l'intervalle et  $T$  indique la transposition du vecteur ou de la matrice.

Les valeurs sous la diagonale sont fixées à 0 puisque l'on suppose qu'un arbre ne peut passer à une classe de diamètre inférieure à celle observée au début de l'intervalle. Par ailleurs, la première colonne de la matrice  $M$  ne contient que des valeurs nulles, étant donné que le premier élément du vecteur  $Y_1$  doit être nul en fin de projection. Par définition, les recrues doivent en effet passer dans les classes de diamètre  $k \geq 1$  et ne peuvent demeurer dans la classe  $k=0$ .

Les matrices de transition peuvent être des modèles déterministes ou stochastiques selon l'utilisation que l'on fait des probabilités (Vanclay 1994). Si l'on considère une matrice de transition comme un modèle stochastique, les éléments de la matrice sont associés à des événements et ils représentent la probabilité que ces événements se produisent durant l'intervalle. En fait, une rangée de la matrice  $M$  représente les probabilités associées à différents événements qui peuvent se produire durant l'intervalle pour les tiges d'une classe de diamètre donnée. En théorie, la somme des

éléments d'une rangée ne peut être plus élevée que 1. Si tous les événements possibles sont représentés, cette somme doit être égale à 1.

Dans le cas présent, les événements d'une rangée de la matrice  $\mathbf{M}$  sont des accroissements qui s'expriment en termes de différences de classes de diamètre. Si les classes sont également espacées, l'accroissement  $g$  est  $g=k'-k$  et  $g \in (0,1,2,\dots,d)$  en supposant que  $k' \geq k$ . Toutefois, la probabilité d'observer un accroissement  $g$  pour une tige donnée sous-entend que la tige a préalablement survécu à l'intervalle. Définissons la survie ou la mort d'une tige par la variable binaire  $s$  en associant la valeur  $s=0$  aux tiges survivantes et  $s=1$  aux tiges mortes. Les probabilités de la matrice  $M$  sous la première rangée sont donc des probabilités composées :

$$[2] \quad \Pr_{k,k'} = \Pr(g = k' - k \mid s = 0) \cdot \Pr(s = 0) \quad \text{pour } k \geq 1 \text{ et } k' \geq k.$$

La règle [2] ne s'applique pas aux recrues parce que l'on n'observe aucune mortalité parmi ces tiges par définition. Par ailleurs, les recrues ne peuvent avoir un accroissement  $g=0$ . En conséquence, les probabilités de la première rangée de la matrice  $\mathbf{M}$  sont les probabilités conditionnelles au fait que  $k' \geq 1$  :

$$[3] \quad \Pr_{0,k'} = \frac{\Pr(g = k')}{1 - \Pr(g = 0)} \quad \text{pour } k=0 \text{ et } k' \geq 1.$$

On peut raisonnablement supposer que les distributions des variables  $g$  et  $s$  sont régies par des distributions de probabilités connues qui peuvent être généralisées à l'ensemble des classes de diamètre ainsi qu'à l'ensemble des couplets de données. De la même façon, on peut postuler que la distribution de la variable  $r$ , soit le nombre de recrues, suit une distribution de probabilités qui s'étend à l'ensemble des couplets de données. Il s'agit donc de définir les trois distributions de probabilités qui sont respectivement associées à l'accroissement ( $g$ ), à la survie ( $s$ ) et au recrutement ( $r$ ) et d'en estimer les paramètres.

La variable  $s$  ne peut suivre qu'une distribution de Bernoulli, une distribution de probabilités pour les variables binaires (Ghahramani 2000). En effet, cette variable ne peut prendre que deux valeurs possibles, soit 0 ou 1, ce qui signifie que  $\Pr(s=0)=1-\Pr(s=1)$ . La distribution de l'accroissement  $g$  est plus complexe étant donné que cette

variable peut prendre plusieurs valeurs comprises entre 0 et la classe de diamètre maximale ( $d$ ). On peut supposer qu'une certaine proportion des tiges d'une classe de diamètre donnée passe à la classe suivante au cours d'une année de croissance. Sur plusieurs années, la distribution de la variable  $g$  suivrait donc une distribution binomiale (cf. Ghahramani 2000), pour laquelle les nombres d'essais et de succès seraient respectivement le nombre d'années et le nombre de classes de diamètre franchies au cours de l'intervalle. En ce qui concerne le recrutement, on peut supposer que le nombre de recrues  $r$  suit une distribution de Poisson, une distribution pour laquelle la variance augmente proportionnellement à la valeur prédite (Ghahramani 2000). En effet, le recrutement et sa variabilité tendent à augmenter dans le temps, ce qui suggère que la distribution de Poisson serait appropriée.

Les trois distributions de probabilités, c'est-à-dire de Bernoulli, binomiale et de Poisson, sont définies par des paramètres pouvant être estimés par la méthode du maximum de vraisemblance. La fonction de vraisemblance représente la probabilité composée d'obtenir les réponses observées à partir des distributions mentionnées et d'un estimé des paramètres régissant ses distributions (Pinheiro et Bates 2000; McCulloch et Searle 2001). Pour des raisons pratiques, on préfère généralement utiliser le logarithme naturel de la vraisemblance (McCulloch et Searle 2001).

Pour le calcul de cette fonction, il est plus aisé de transformer les deux dénombrements d'un couplet de données en une liste d'arbres. Chaque élément d'une liste d'arbres constitue un arbre individuel pour lequel les valeurs d'accroissement  $g$  et de survie  $s$  sont connues. De plus, le recrutement ( $r$ ) peut être obtenue en comptabilisant le nombre de recrues à l'intérieur de la liste. Définissons  $i$  comme l'indice des listes d'arbres tel que  $i=(1,2,\dots, n)$  et  $j$  comme l'indice des tiges à l'intérieur de la liste d'arbres tel que  $j=(1,2,\dots, m_i)$  et trié de sorte que les  $r_i$  premières tiges soient les recrues. Ainsi, le logarithme naturel de la vraisemblance ( $\ell$ ) se définit :

$$[4] \quad \ell(\mathbf{R}, \mathbf{G}, \mathbf{S} | \boldsymbol{\theta}) = \sum_{i=1}^n \left( \ln A_i + \sum_{j=1}^{r_i} \ln B_{ij} + \sum_{j=r_i+1}^{m_i} \ln C_{ij} \right)$$

$$A_i = \text{Poisson}(r_i | \boldsymbol{\theta}_1),$$

$$B_{ij} = \frac{\text{Binomiale}(g_{ij} | \boldsymbol{\theta}_2)}{1 - \text{Binomiale}(g = 0 | \boldsymbol{\theta}_2)} \quad \text{pour } j \leq r_i$$

$$C_{ij} = \begin{cases} \text{Bernoulli}(s_{ij} = 1 | \boldsymbol{\theta}_3) \\ \text{Binomiale}(g_{ij} | \boldsymbol{\theta}_2, s = 0) \cdot \text{Bernoulli}(s_{ij} = 0 | \boldsymbol{\theta}_3) \end{cases} \quad \text{pour } j > r_i$$

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \boldsymbol{\theta}_3^T)^T$$

où  $\boldsymbol{\theta}$  est le vecteur de paramètres régissant les trois distributions discrètes et  $\mathbf{R}$ ,  $\mathbf{G}$  et  $\mathbf{S}$  sont des vecteurs dont les éléments sont respectivement  $r_i$ ,  $g_{ij}$  et  $s_{ij}$ .

À l'intérieur de la fonction [4], la composante  $A_i$  représente la vraisemblance associée au recrutement,  $B_{ij}$  représente la vraisemblance associée à l'accroissement des recrues et  $C_{ij}$  est la vraisemblance associée à la survie et l'accroissement des tiges présentes au début de l'intervalle. La maximisation de cette fonction par rapport au vecteur  $\boldsymbol{\theta}$  permet d'obtenir l'estimateur du maximum de vraisemblance de ce même vecteur (Pinheiro et Bates 2000; McCulloch et Searle 2001). Pour ce faire, plusieurs algorithmes existent. L'algorithme de Newton-Raphson donne généralement de bons résultats pour la maximisation de la vraisemblance de modèles mixtes (Lindstrom et Bates 1988; Wolfinger *et al.* 1994; Pinheiro et Bates 2000). Sommairement, la fonction [4] atteint son maximum lorsque les dérivées partielles de la fonction par rapport aux paramètres sont égales à 0. Avec la plupart des modèles complexes, cette équivalence n'a pas de solution analytique (Pinheiro et Bates 2000). En conséquence, les dérivées partielles sont estimées à l'aide par une série de Taylor de premier degré. L'approximation requiert donc les dérivées premières et secondes de la fonction de vraisemblance par rapport aux paramètres. Les matrices des dérivées premières et secondes sont donc évaluées en un point donné de la fonction de vraisemblance. Par convention, le gradient est la matrice des dérivées premières alors que la matrice hessienne est celle des dérivées secondes. On calcule un pas d'optimisation pour le vecteur de paramètres  $\boldsymbol{\theta}$  en multipliant moins l'inverse de la matrice hessienne par le gradient. L'inverse de la matrice hessienne fournit également un estimé des variances et des covariances associées aux estimés de paramètres (Wolfinger *et al.* 1994; SAS Institute Inc. 2002).

Si l'on présume qu'aucune variable explicative n'est associée aux distributions de probabilités, on obtient une matrice de transition traditionnelle, c'est-à-dire, une matrice dont les probabilités sont constantes. À l'inverse, si des variables explicatives sont incluses, les distributions de probabilités peuvent varier selon les valeurs de ces variables. Une matrice de transition à deux niveaux est ainsi obtenue. Les trois distributions telles que présentées en [4] peuvent être formulées comme suit en incluant des variables explicatives :

$$[5] \quad \text{Poisson}(r_i | \boldsymbol{\theta}_1, \mathbf{z}_i) = \frac{f_{recr}(\boldsymbol{\theta}_1, \mathbf{z}_i)^{r_i} e^{-f_{recr}(\boldsymbol{\theta}_1, \mathbf{z}_i)}}{r_i!}$$

$$\text{Binomiale}(g_{ij} | \boldsymbol{\theta}_2, \mathbf{x}_{ij}, \mathbf{z}_i, t_i) = \frac{t_i!}{g_{ij}!(t_i - g_{ij})!} f_{acc}(\boldsymbol{\theta}_2, \mathbf{x}_{ij}, \mathbf{z}_i)^{g_{ij}} (1 - f_{acc}(\boldsymbol{\theta}_2, \mathbf{x}_{ij}, \mathbf{z}_i))^{t_i - g_{ij}}$$

$$\text{Bernoulli}(s_{ij} | \boldsymbol{\theta}_3, \mathbf{x}_{ij}, \mathbf{z}_i, t_i) = \left( (f_{surv}(\boldsymbol{\theta}_3, \mathbf{x}_{ij}, \mathbf{z}_i))^{t_i} \right)^{1-s_{ij}} \left( 1 - (f_{surv}(\boldsymbol{\theta}_3, \mathbf{x}_{ij}, \mathbf{z}_i))^{t_i} \right)^{s_{ij}}$$

où  $\mathbf{z}_i$  est un jeu de variables explicatives communes aux arbres de la liste d'arbres  $i$ ,  $\mathbf{x}_{ij}$  est un jeu de variables explicatives mesurées à l'échelle des tiges individuelles et  $t_i$  est la durée de l'intervalle.

La calibration de la matrice de transition à deux niveaux proposée dans cette étude requiert donc la spécification de trois fonctions : une fonction de recrutement ( $f_{recr}(\boldsymbol{\theta}_1, \mathbf{z}_i)$ ), une fonction d'accroissement ( $f_{acc}(\boldsymbol{\theta}_2, \mathbf{x}_{ij}, \mathbf{z}_i)$ ) et une fonction de survie ( $f_{surv}(\boldsymbol{\theta}_3, \mathbf{x}_{ij}, \mathbf{z}_i)$ ). Ces trois fonctions sont les paramètres des trois distributions de probabilités qui régissent les probabilités de la matrice de transition. Étant donné que ces fonctions intègrent des variables explicatives, les probabilités de la matrice de transition peuvent donc varier selon les conditions initiales de la placette et des arbres de la placette.

## **Aire d'étude**

### **Cadre écologique**

L'Aire d'observation de la rivière Ouareau (AORO) (46°26'N, 74°10'W) couvre une superficie de 13 km<sup>2</sup>. La topographie de cette région est caractérisée par une

succession de collines aux sommets arrondis (Heimburger 1941; Robitaille et Saucier 1998). La plupart des dépôts de surface sont de type glaciaire. Les tills indifférenciés sont prédominants. Des dépôts fluvio-glaciaires et des dépôts organiques ont également été observés à certains endroits à l'intérieur du dispositif. L'altitude moyenne est de l'ordre de 515 m.

Un aperçu des normales et moyennes climatiques pour cette région nous est fourni par la station météorologique de la localité la plus proche, Saint-Donat-de-Montcalm (Environnement Canada 2002). Calculée à partir de la période 1971-2000, la température annuelle moyenne est de 3,2°C. Janvier est le mois le plus froid avec une moyenne de -13,6°C, alors que juillet est le plus chaud affichant une moyenne de 18,1°C. Les précipitations annuelles atteignent 1128 mm. En moyenne, les chutes de neige se chiffrent à 295 cm annuellement.

L'AORO se situe dans la section laurentienne (L.4a) de la région forestière des Grands Lacs et du Saint-Laurent (*cf.* Rowe 1972). Le cadre écologique de la section laurentienne a été largement décrit par plusieurs auteurs (Heimburger 1941; Gagnon et Marcotte 1980; Lalumière et Thibault 1988; Grondin *et al.* 1996; Gosselin *et al.* 2001). En fait, ses caractéristiques écologiques sont similaires à celles de la région forestière acadienne (Rowe 1972). Les peuplements mixtes de sapin baumier et d'épinette rouge sont surtout présents sur les tills mésiques au bas des pentes et en terrains plats (Heimburger 1941; Grondin *et al.* 1996; Gosselin *et al.* 2001). Le bouleau jaune s'associe également à ces deux espèces et peut représenter une proportion non négligeable des peuplements. Selon les différents auteurs, ces stations ont été classifiées comme appartenant au type forestier *Oxalis-Cornus* (*cf.* Heimburger 1941), au type forestier « Spruce-fir-hardwood » (*cf.* Westveld 1953) et au type écologique de la bétulaie jaune à sapin (*cf.* Gosselin *et al.* 2001). À l'intérieur du dispositif, ce type de station forestière est le plus répandu (Hatcher 1954).

### ***Historique des perturbations***

Une coupe à diamètre limite a été réalisée sur la presque totalité de l'aire d'étude entre 1948 et 1956 dans le but de constituer un approvisionnement de matière ligneuse destinée à la pâte. Un seul diamètre limite de 18 cm (7 po) à hauteur de

souche (30 cm) a alors été utilisé pour le sapin baumier et l'épinette rouge (Hatcher 1954). Quelques bouleaux jaunes de fortes dimensions ont également été récoltés pour la fabrication de placages. Toutefois, la récolte de cette espèce ne s'est pas faite de façon systématique sur l'ensemble du territoire. On estime l'intensité du prélèvement à 60% du volume total des tiges de plus de 8,9 cm (3,5 po) en diamètre à hauteur de poitrine (*dhp*) (Hatcher 1954).

En plus de cette coupe, deux épidémies de tordeuse des bourgeons de l'épinette (*Choristoneura fumiferana* Clem.) ont été recensées dans le secteur depuis 1938 (Hardy *et al.* 1987). La première a eu lieu durant les décennies de 1940-50, alors que la seconde s'est déroulée durant les décennies de 1970-80. Hatcher (1954) mentionne qu'il n'y a pas eu d'infestation sévère à l'intérieur du dispositif lors de la première épidémie. Par ailleurs, des carottes de croissance prélevées à la sonde de Pressler ont permis de caractériser les effets de la seconde épidémie sur la croissance en diamètre des tiges de sapin baumier et d'épinette rouge à l'intérieur de l'AORO (*cf.* Fortin *et al.* 2003a). Chez les deux espèces, on a observé des diminutions significatives de l'accroissement en diamètre à partir de 1976. Au maximum de l'effet de la défoliation, l'accroissement annuel moyen en diamètre s'est toutefois maintenu à un niveau équivalent ou supérieur à 1 mm·an<sup>-1</sup>. Les travaux d'inventaires à l'été 1998 et l'été 1999 n'ont pas permis d'identifier des zones de mortalité sévère, ce qui suggère que les effets de la seconde épidémie ont été limités en termes de mortalité.

### ***Dispositif de suivi et base de données***

En 1953, un réseau de placettes permanentes a été établi sur l'AORO en suivant un patron de distribution systématique. À l'intérieur de ces placettes de 404,7 m<sup>2</sup> (0,1 ac) chacune, toutes les tiges de plus de 1,3 cm (0,5 po) en *dhp* ont été dénombrées par classe de 2,5 cm (1 po). Des mesurages successifs ont été réalisés en 1963, en 1974-77, en 1998-99 et 2002-04. À partir de 1998, un inventaire écologique a été réalisé à l'intérieur des placettes. Par la suite, les placettes ont été classifiées selon la classification écologique de Gosselin *et al.* (2001).

La base de données utilisée dans cette étude a été constituée à partir des dénombrements de ces placettes. Seules les placettes dont le type écologique correspondait à la bétulaie jaune à sapin (*cf. Gosselin et al. 2001*) ont été retenues. Au total, 34 placettes ont été sélectionnées, chacune d'elles ayant quatre ou cinq dénombrements.

La matrice de transition a été calibrée de façon indépendante pour chacune des deux espèces principales, c'est-à-dire le sapin baumier et l'épinette rouge. Pour ce faire, des couplets de données ont été formés en associant les dénombrements de chaque espèce deux par deux pour une placette donnée. Les superpositions d'intervalles à l'intérieur de la même combinaison espèce-placette ont été évitées afin de limiter dans la mesure du possible l'autocorrélation des données (Vanclay 1994). De plus, les couplets couvrant la fin des années 1970 et le début des années 1980 ont été rejetés pour éviter d'inclure les effets directs de l'épidémie de tordeuse dans la calibration du modèle. Pour chacune des deux espèces, 30% des projections ont été mises de côté pour fins de validation. Le Tableau 9 présente un sommaire des données utilisées pour la calibration et la validation du modèle.

## Méthode

### *La calibration*

Le calcul de la fonction [4] requiert une liste d'arbres pour chaque couplet de données. Or, la construction de cette liste suppose que le diamètre de chaque arbre au début et à la fin de l'intervalle est connu. Dans le cas présent, les tiges n'ont pas été identifiées, de sorte que l'information requise n'est pas disponible. Certaines méthodes ont été proposées pour contourner ce problème. Elles consistent à ordonner les tiges en fonction de leur diamètre. Ensuite, on postule que le rang des tiges demeure le même, une prémisse dont on peut douter fortement (Vanclay 1994).

Définissons  $\mathbf{H}_{0,i}$  et  $\mathbf{H}_{1,i}$  comme deux vecteurs contenant respectivement les classes de diamètre de chacune des tiges au début et en fin d'intervalle à l'intérieur de la liste d'arbres  $i$ . Puisque l'on ne peut savoir avec certitude quelles tiges sont mortes et quelles tiges ont été recrutées durant l'intervalle, des éléments fictifs doivent être

ajoutés aux deux vecteurs de façon à couvrir l'ensemble des possibilités. En somme, toutes les tiges au début de l'intervalle peuvent mourir. Conséquemment, un nombre d'éléments équivalent doit être ajouté au vecteur  $\mathbf{H}_{1,i}$ . Dans un même ordre d'idée, toutes les tiges à la fin de l'intervalle peuvent être des recrues. Le même nombre d'éléments doit donc être ajouté au vecteur  $\mathbf{H}_{0,i}$ . Définissons  $\mathbf{H}'_{0,i}$  et  $\mathbf{H}'_{1,i}$  comme les versions modifiées des vecteurs originaux.

**Tableau 9 – Sommaire des données de calibration et de validation (toutes les tiges de plus de 1,3 cm en dhp).**

	Calibration (61 couplets)			Validation (27 couplets)		
	min.	moy.	max.	min.	moy.	max.
Épinette rouge						
Surf. terrière – placette ( $\text{m}^2 \cdot \text{ha}^{-1}$ )	3,7	22,9	43,2	4,9	24,5	39,3
Surf. terrière – espèce ( $\text{m}^2 \cdot \text{ha}^{-1}$ )	0,2	5,5	16,0	0,3	6,6	18,6
Durée de l'intervalle (années)	4	9	14	4	9	14
Densité – espèce ( $\text{ti} \cdot \text{ha}^{-1}$ )	49	500	1532	99	516	1359
Sapin baumier						
Surf. terrière – placette ( $\text{m}^2 \cdot \text{ha}^{-1}$ )	3,7	24,3	43,2	5,3	21,5	40,7
Surf. terrière – espèce ( $\text{m}^2 \cdot \text{ha}^{-1}$ )	0,7	10,8	28,3	0,2	9,1	23,2
Durée de l'intervalle (années)	4	9	14	4	10	14
Densité – espèce ( $\text{ti} \cdot \text{ha}^{-1}$ )	148	1251	4151	99	1657	6079

Si l'on aligne les deux vecteurs côte à côte pour former la liste d'arbres, la correspondance non observée entre les éléments des vecteurs  $\mathbf{H}'_{0,i}$  et  $\mathbf{H}'_{1,i}$  peut être vue comme une permutation des éléments à l'intérieur du vecteur  $\mathbf{H}'_{1,i}$ . Si  $permut_i^*$  est la permutation du vecteur  $\mathbf{H}'_{1,i}$  qui procure la véritable correspondance entre les éléments de  $\mathbf{H}'_{0,i}$  et  $\mathbf{H}'_{1,i}$ , on peut définir l'ensemble suivant :

$$[6] \quad \text{permut}_i^* \in \mathbf{W}_i$$

$$\mathbf{W}_i = (\text{permut}_{i,1}, \text{permut}_{i,2}, \text{permut}_{i,p}, \dots, \text{permut}_{i,q_i}).$$

N'importe quelle permutation  $\text{permut}_{i,p}$  appartenant à l'ensemble  $\mathbf{W}_i$  permet donc d'obtenir le même dénombrement en fin d'intervalle pour la liste d'arbres  $i$ . La différence entre les permutations  $\text{permut}_{i,p}$  réside dans la correspondance entre les tiges du début et de la fin de l'intervalle. Si l'on garde le vecteur de paramètres  $\theta$  constant, le résultat de la fonction de vraisemblance [4] peut être utilisé comme une mesure de l'ajustement du modèle en fonction de la permutation. La permutation affichant la vraisemblance la plus élevée est celle pour laquelle il est le plus probable d'obtenir le dénombrement observé en fin d'intervalle pour le vecteur de paramètres  $\theta$ . Définissons  $\text{permut}_{i,opt}$  comme étant la permutation qui affiche la plus forte vraisemblance. Il n'y a aucune indication que  $\text{permut}_{i,opt} = \text{permut}_i^*$  et cette supposition ne peut être vérifiée analytiquement, puisque la permutation  $\text{permut}_i^*$  n'a pas été observée. Toutefois, la validation devrait prouver si ce postulat est valable ou non.

Le nombre de permutations ( $q_i$ ) peut être relativement élevé pour certaines listes d'arbres. Des essais préliminaires basés sur une sélection au hasard se sont avérés infructueux dû au fait que le résultat de la fonction [4] dépend à la fois de l'estimé du vecteur de paramètres  $\theta$  et de la séquence des permutations associées à chacune des listes d'arbres. Il devient donc très difficile d'optimiser la vraisemblance en fonction du vecteur de paramètres et des permutations en une seule étape. Pour y parvenir, un algorithme basé sur l'optimisation successive des deux éléments a été utilisé. Cet algorithme s'inspire de l'algorithme des moindres carrés généralisés (cf. Davidian et Giltinan 1995, ch. 2), lequel est notamment utilisé pour optimiser la vraisemblance de modèles mixtes (SAS Institute Inc. 2002). La première étape de l'algorithme consiste à trouver les permutations  $\text{permut}_{i,opt}$  en maximisant la fonction de vraisemblance [4] à partir d'un estimé préliminaire du vecteur  $\theta$  que l'on garde constant. Lors de la deuxième étape, le vecteur  $\theta$  est estimé en maximisant à nouveau la fonction de vraisemblance à partir des permutations  $\text{permut}_{i,opt}$  obtenues lors de la première étape. À partir de ce nouvel estimé du vecteur  $\theta$ , on recalcule les

$permut_{i,opt}$  correspondant et ainsi de suite, jusqu'à convergence de la fonction de vraisemblance.

Cet algorithme a donc été testé et après quelques essais, les fonctions suivantes ont été retenues :

$$[7] \quad f_{recr}(\theta_1, st_i, t_i) = \beta_1 t_i e^{\beta_2 st_i}$$

$$f_{acc}(\theta_2, dinit_{ij}, st_i) = \frac{1}{1 + e^{\beta_3 + \beta_4 dinit_{ij} + \beta_5 dinit_{ij}^2 + \beta_6 st_i}}$$

$$f_{surv}(\theta_3, dinit_{ij}, st_i) = \frac{1}{1 + e^{\beta_7 + \beta_8 dinit_{ij} + \beta_9 st_i}}$$

où  $st$  est la surface terrière de la placette au début de l'intervalle ( $m^2 \cdot ha^{-1}$ ),  $t$  est la durée de l'intervalle (années),  $dinit$  est la valeur médiane de la classe de diamètre (cm) et  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8$  et  $\beta_9$  sont les paramètres du modèle, soit les éléments du vecteur  $\theta$ .

Le modèle a été indépendamment calibré pour chacune des deux espèces. Dans les deux cas, plusieurs calibrations ont été réalisées en utilisant différentes valeurs de départ pour les paramètres du modèle. Cette mesure visait à s'assurer que la fonction de vraisemblance converge en un maximum absolu et non pas en un maximum local. De plus, des tests d'hypothèses ont été réalisés sur les estimés des paramètres pour vérifier s'ils étaient significativement différents de 0. Les résidus de chacune des trois distributions de probabilités (de Bernoulli, binomiale et de Poisson) ont été observés pour s'assurer que ces distributions étaient adéquates. Les calibrations ont été effectuées à l'aide du langage matriciel IML disponible dans SAS (cf. SAS Institute Inc. 2002).

### **La validation**

Une fois l'estimateur du vecteur de paramètres  $\theta$  obtenu, la matrice de transition de chacune des deux espèces a été utilisée pour un exercice de validation. Pour ce faire, les probabilités de la matrice de transition ont été calculées en fonction des variables explicatives associées à chacune des listes d'arbres disponibles pour la

validation. Ensuite, des nombres aléatoires uniformément distribués entre 0 et 1 ont été générés. En se référant aux probabilités de la matrice, ces nombres ont servi à prédire (i) le nombre de recrues et leur accroissement et (ii) la survie et l'accroissement des tiges présentes au début de l'intervalle. Le dénombrement prédit en fin d'intervalle a été comparé à celui observé sur la base de la densité de tiges et de la surface terrière. Le biais et l'erreur type de l'estimé (ETE) ont été calculés dans le but de quantifier les différences entre les valeurs prédites et observées :

$$[8] \quad \text{biais} = \frac{\sum_{i=1}^n (w_i - \hat{w}_i)}{n}$$

$$\text{ETE} = \frac{\sum_{i=1}^n (w_i - \hat{w}_i)^2}{n}$$

où  $w$  est l'une des deux variables sélectionnées, soit la densité ou la surface terrière,  $n$  est le nombre de listes d'arbres et l'accent circonflexe dénote la valeur prédite par le modèle. Il importe de mentionner que le processus stochastique qui s'exprime par l'utilisation de nombres aléatoires implique que deux projections de la même liste d'arbres prédisent des dénombrements différents en fin d'intervalle. Conséquemment, les deux statistiques présentées en [8] sont sujettes à certaine variabilité. Une simulation Monte Carlo permet d'estimer cette variabilité en calculant les statistiques de manière itérative (Manly 1997, ch. 4). Il s'agit donc de calculer les statistiques plusieurs fois pour l'ensemble des listes d'arbres et d'en mesurer la variabilité. Dans le cas présent, la variabilité des deux statistiques, le biais et l'ETE, a été évaluée à partir de 1000 itérations.

## Résultats

Les résultats des calibrations sont présentés au Tableau 10. Les estimés de  $\beta_1$  indiquent un recrutement généralement plus faible chez l'épinette rouge (1.217 vs 6.226). Les valeurs négatives associées à  $\beta_2$  démontrent que l'augmentation de la surface terrière entraîne une diminution du recrutement. La diminution est plus importante chez le sapin baumier ( $-3.952 \times 10^{-2}$  vs  $-5.693 \times 10^{-2}$ ).

**Tableau 10 – Résultats de la calibration de la matrice de transition à deux niveaux pour l'épinette rouge et le sapin baumier.**

	Épinette rouge	Sapin baumier
Nombre de listes d'arbres ( $n$ )	61	61
Logarithme de la vraisemblance	-2011	-4775
Paramètres de $f_{recr}$		
$\beta_1$	1.217	6.226
$\beta_2$ (effet de $st^\dagger$ )	$-3.952 \times 10^{-2}$	$-5.693 \times 10^{-2}$
Paramètres de $f_{acc}$		
$\beta_3$	2.829	2.028
$\beta_4$ (effet de $dinit^\dagger$ )	$-1.322 \times 10^{-1}$	$-1.345 \times 10^{-1}$
$\beta_5$ (effet du carré de $dinit^\dagger$ )	$2.383 \times 10^{-3}$	$2.334 \times 10^{-3}$
$\beta_6$ (effet de $st^\dagger$ )	$2.622 \times 10^{-2}$	$6.275 \times 10^{-2}$
Paramètres de $f_{surv}$		
$\beta_7$	-4.844	-3.512
$\beta_8$ (effet de $dinit^\dagger$ )	$-8.547 \times 10^{-2}$	$-3.066 \times 10^{-1}$
$\beta_9$ (effet de $st^\dagger$ )	$5.341 \times 10^{-2}$	$7.386 \times 10^{-2}$

$^\dagger dinit$  : médiane de la classe de diamètre,  $st$  : surface terrière de la placette.

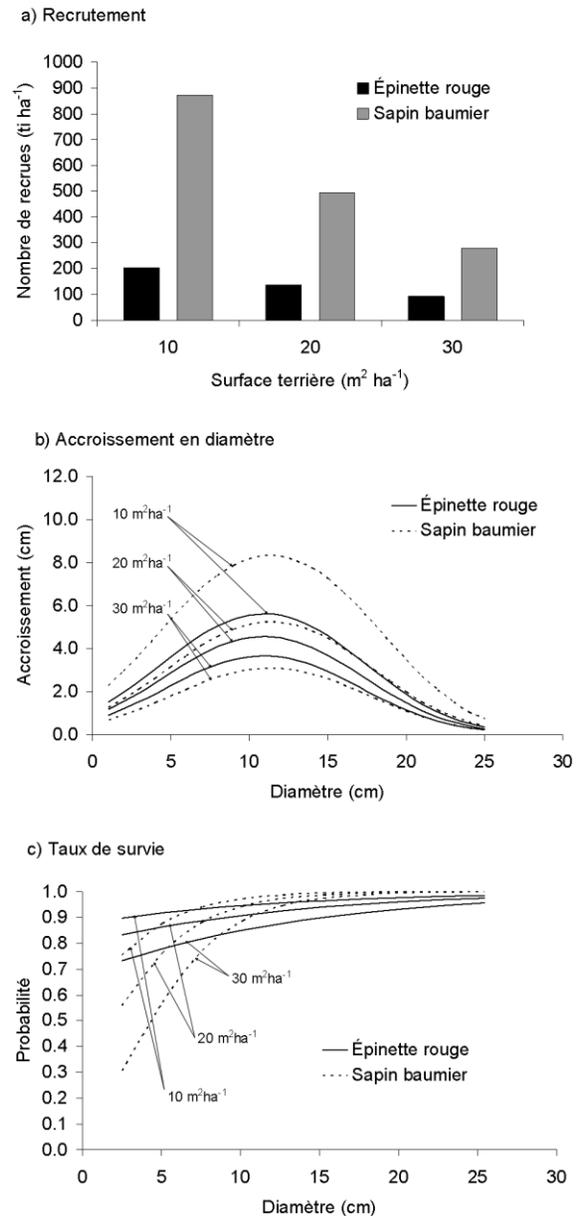
\*\*significatif à un seuil  $Pr < 0.01$ .

En raison de la forme des équations d'accroissement et de survie ( $f_{acc}$  et  $f_{surv}$ ) telles que présentées en [7], on doit interpréter les estimés de ces équations à l'inverse, c'est-à-dire que les valeurs négatives indiquent un effet positif et vice versa. Dans le cas présent, les estimés de  $\beta_3$  démontrent que l'accroissement de l'épinette rouge est généralement plus lent que celui du sapin baumier (2.829 vs 2.028). La classe de diamètre en début d'intervalle ( $dinit$ ) a un effet quadratique sur l'accroissement en diamètre des deux essences. Les valeurs négatives associées au paramètre  $\beta_4$  indiquent que l'accroissement en diamètre augmente avec l'augmentation de la classe de diamètre initiale ( $-1.322 \times 10^{-1}$  vs  $-1.345 \times 10^{-1}$ ). Toutefois, en raison de l'effet négatif lié au carré de cette même variable, cette augmentation tend à se résorber lorsque la classe de diamètre dépasse un certain seuil. Par ailleurs, les estimés du

paramètre  $\beta_6$  indiquent qu'une augmentation de la surface terrière ( $st$ ) induit une diminution de l'accroissement en diamètre chez les deux essences. La diminution est plus marquée chez le sapin baumier avec une valeur de  $6.275 \times 10^{-2}$ , comparativement à un estimé de  $2.622 \times 10^{-2}$  pour l'épinette rouge.

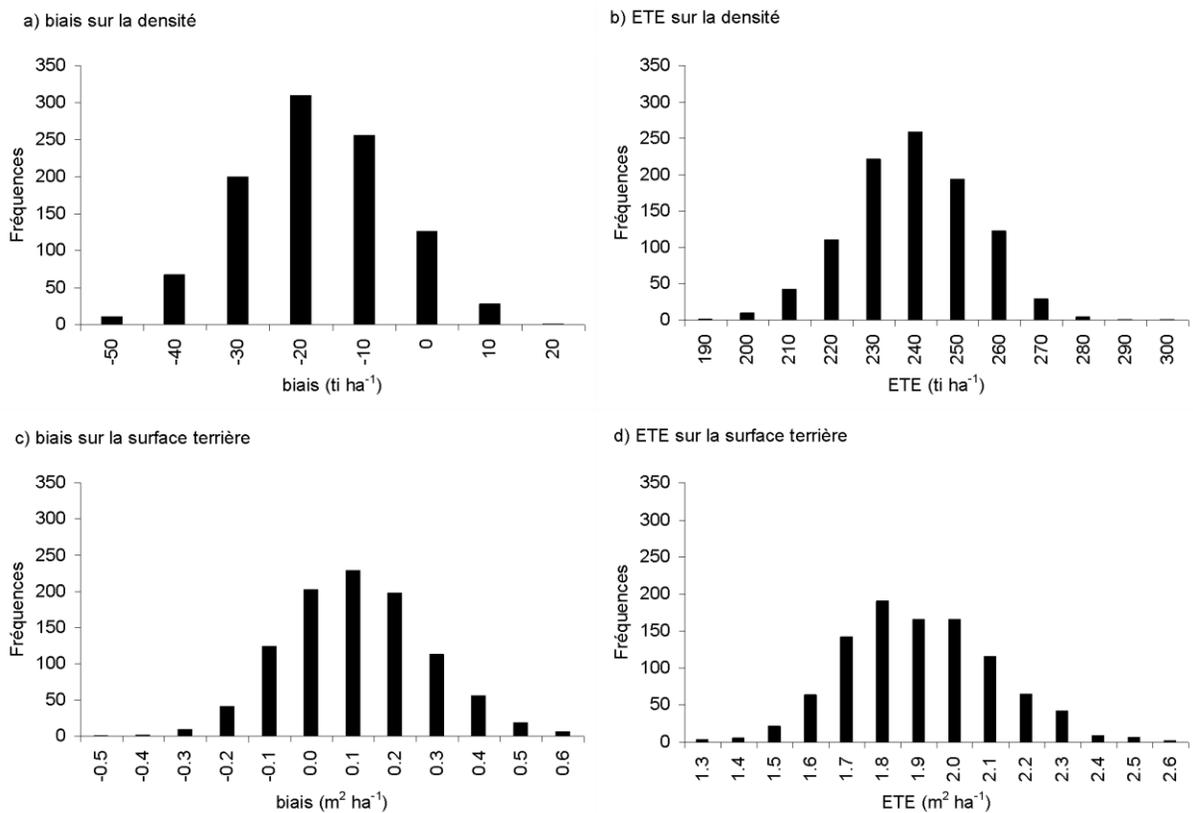
En ce qui concerne la survie des deux espèces, les estimés de  $\beta_7$  démontrent que le taux de survie est plus élevé chez l'épinette rouge (-4.844 vs -3.512). L'augmentation de la classe de diamètre se traduit par une augmentation du taux de survie chez les deux espèces, comme le démontrent les valeurs associées à  $\beta_8$ . L'effet positif est plus marqué pour les tiges de sapin baumier ( $-8.547 \times 10^{-2}$  vs  $-3.066 \times 10^{-1}$ ). L'augmentation de la surface terrière induit une diminution du taux de survie chez les deux espèces (paramètre  $\beta_9$ ). Le taux de survie du sapin baumier est plus sensible à l'augmentation de la surface terrière que celui de l'épinette rouge ( $5.341 \times 10^{-2}$  vs  $7.386 \times 10^{-2}$ ).

Les valeurs moyennes de recrutement, d'accroissement et de survie prédites par les matrices de transition ont été calculées sur une base de 10 ans pour illustrer l'effet des différentes variables explicatives (Figure 7). L'accroissement et la survie ont été



**Figure 7 – Prédiction du recrutement, de l'accroissement et de la survie sur une période de 10 ans en fonction de trois différents niveaux de surface terrière.**

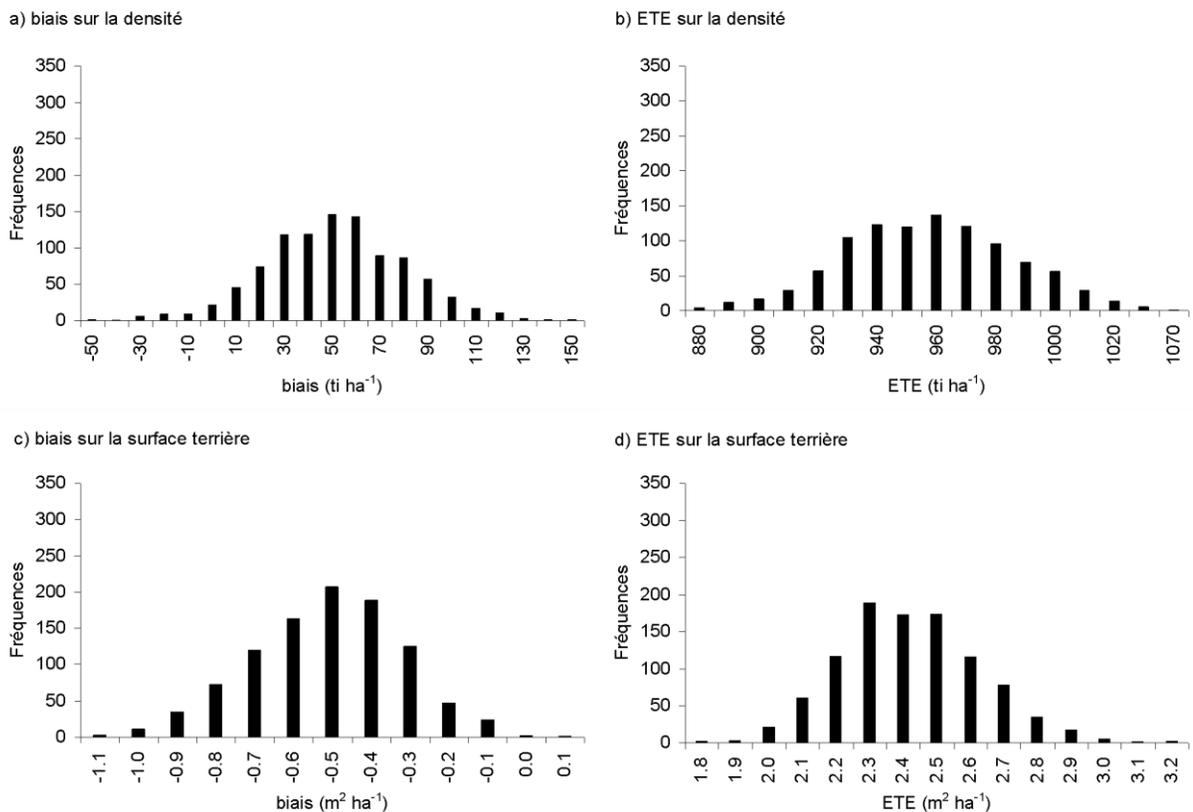
présentés comme des fonctions continues pour faciliter la lecture des graphiques. Tel que suggéré par Vanclay (1994), l'accroissement moyen a été estimé en effectuant le produit de l'accroissement en termes de nombre de classes de diamètre par la largeur des classes de diamètre.



**Figure 8 – Distributions des biais et des ETE résultant de la validation du modèle de l'épinette rouge.**

Le recrutement du sapin baumier se chiffre entre 300 to 900 ti·ha<sup>-1</sup> en fonction de la surface terrière, alors que celui de l'épinette rouge est généralement inférieur à 200 ti·ha<sup>-1</sup> (Figure 7a). Le patron d'accroissement des deux espèces est similaire. Les petites tiges marchandes, définies ici comme les tiges dont le *dhp* se situe entre 9 et 15 cm, sont celles qui affichent les plus forts accroissements (Figure 7b). Lorsque la surface terrière est relativement faible, à un niveau de 10 m<sup>2</sup>·ha<sup>-1</sup> par exemple, le

sapin baumier s'accroît plus rapidement que l'épinette rouge. Cependant, comme l'augmentation de la surface terrière induit des diminutions d'accroissement plus importantes chez le sapin, les deux espèces affichent des accroissements similaires lorsque la surface terrière dépasse les 20 m<sup>2</sup>·ha<sup>-1</sup>. Quant au taux de survie (Figure 7c), les gaules d'épinette rouge (*dhp* < 9 cm) affichent un taux de survie plus élevé que celles de sapin baumier. L'effet de l'augmentation de la surface terrière sur le taux de survie est beaucoup plus marqué chez le sapin. En effet, lorsque la surface terrière augmente à des niveaux supérieurs à 20 m<sup>2</sup>·ha<sup>-1</sup>, le taux de survie des gaules de sapin décroît rapidement et atteint des valeurs largement inférieures à ce que l'on observe chez les gaules d'épinette. L'effet de la surface terrière est relativement faible chez les tiges marchandes des deux espèces (*dbh* > 9.0 cm). Le taux de survie de ces tiges semble légèrement plus élevé chez le sapin baumier.



**Figure 9 – Distributions des biais et des ETE résultant de la validation du modèle du sapin baumier.**

Les résultats de la validation sont présentés aux Figure 8 et Figure 9. Les distributions des statistiques sont presque normales. Pour l'épinette rouge, il est relativement fréquent d'obtenir des biais presque nuls, autant pour la densité que pour la surface terrière (Figure 8a, c). Les erreurs types des estimés (ETE) sont beaucoup plus élevés que les biais en valeur absolue (Figure 8b, d). En ce qui concerne le sapin baumier, la matrice de transition tend à sous-estimer la densité et à surestimer la surface terrière (Figure 9a, c). Que ce soit pour la densité ou la surface terrière, un biais nul est en effet plutôt rare, ne représentant que quelques cas sur les 1000 itérations de la simulation Monte Carlo. Toutefois, les valeurs de ces biais demeurent faibles et les ETE sont largement supérieures à ces valeurs (Figure 9b, d).

## Discussion

### *L'approche de modélisation*

Les probabilités des matrices de transition sont traditionnellement déterminées à partir d'un jeu de constantes ou d'équations (e.g. Bruner et Moser 1973; Solomon *et al.* 1986, 1987). Au lieu d'utiliser directement des équations ou des constantes, l'approche préconisée dans cette étude suppose que les probabilités de la matrice de transition sont dictées par trois distributions de probabilités qui prédisent le recrutement, la survie et l'accroissement en diamètre. En exprimant les paramètres de ces trois distributions comme des fonctions incluant des variables explicatives, on obtient une matrice de transition à deux niveaux comparable à celle de Solomon *et al.* (1986). Bien que les bases statistiques de l'approche puissent sembler plus complexes, l'utilisation de distributions de probabilités comporte trois avantages par rapport à l'approche classique.

Premièrement, étant donné que les distributions de survie et d'accroissement sont liées entre elles par le biais d'une probabilité composée (équation [2]), la somme des probabilités ne peut être différente de 1. En d'autres termes, la somme des probabilités d'une rangée de la matrice  $M$  est toujours égale à 1 si l'on inclut la mortalité. Or, il s'agit d'un point faible de l'approche classique. Puisque rien ne lie les

probabilités d'une rangée de la matrice de transition sous l'approche classique, il est très probable d'obtenir une somme différente de 1. La méthode proposée dans cette étude évite d'avoir à effectuer un rééchelonnage des probabilités afin de corriger une somme des probabilités incohérente.

Deuxièmement, la méthode proposée n'est pas aussi limitée par les contraintes liées à la durée de l'intervalle et à l'accroissement en diamètre. Avec l'approche classique, le nombre de paramètres à estimer est proportionnel à l'accroissement des tiges. Ainsi, si l'on suppose que l'accroissement maximum est de deux classes de diamètre, trois paramètres doivent être estimés, soit un paramètre pour chacune des probabilités représentant les événements  $g=0$ ,  $g=1$  et  $g=2$ . Évidemment, plus l'intervalle est long, plus l'accroissement maximum est élevé et plus les paramètres sont nombreux. Face à ce problème, on recherche généralement un compromis par lequel on restreint la calibration du modèle aux intervalles les plus courts de façon à limiter le nombre de paramètres à estimer. La méthode préconisée dans cette étude ne requiert pas de paramètres additionnels et n'est donc pas limitée aux projections à court terme.

Troisièmement, comme la durée de l'intervalle est incluse dans chacune des trois distributions de probabilités, la matrice de transition peut s'adapter à différentes durées de projection. Avec l'approche classique, les projections sont limitées à des intervalles fixes. Même si l'on peut obtenir des résultats sur un intervalle plus long en effectuant plusieurs projections successives (*e.g.* Solomon *et al.* 1986), l'approche classique reste limitée à des intervalles dont la durée est un multiple de l'intervalle de base. Par exemple, on ne peut obtenir une projection sur 8 ans, si le modèle a été calibré à partir d'intervalles fixes de 5 ans. L'approche préconisée dans cette étude offre plus de latitude du fait que l'intervalle est ajustable.

Les seules contraintes en termes de durée de projection et d'accroissement sont liées à la distribution des probabilités associées à l'accroissement. L'accroissement des tiges est limité à une classe de diamètre par unité de temps. Dans la région Nord-Est de l'Amérique du Nord, cette contrainte n'est pas problématique aussi

longtemps que la largeur des classes de diamètre est de 2,5 cm et que les unités de temps sont exprimées en années.

L'absence d'identifiant pour les tiges individuelles constitue sans doute le plus grand obstacle dans cette étude. L'algorithme de calibration optimise les permutations de façon à rendre les données observées conformes aux prédictions du modèle. Ainsi, une certaine part de variabilité est artificiellement exclue du modèle. Les variances associées aux estimés des paramètres tendraient donc à être sous-estimées. L'évaluation de cette sous-estimation est complexe et dépasse le cadre de cette étude. Il importe de préciser que cette limitation n'affecte en rien l'approche de modélisation.

Par ailleurs, une certaine proportion de la variabilité des prédictions du modèle provient de la corrélation qui existe entre le recrutement, l'accroissement et la mortalité. Des accroissements au-dessus de la moyenne pourraient être dus à un taux de mortalité plus élevé que la moyenne, les arbres survivants ayant plus d'espace pour croître. De la même façon, une mortalité plus élevée que la moyenne peut favoriser le recrutement de nouvelles tiges. Certaines procédures statistiques permettent de calibrer plusieurs équations de façon simultanée en tenant compte de la corrélation entre elles. Les méthodes « three-stage least squares » et « general method of moments » en sont des exemples (*cf.* Gallant 1987). Ces méthodes présument toutefois que les variables dépendantes suivent des distributions semblables, ce qui n'est évidemment pas le cas lorsque la survie répond à une loi de Bernoulli, l'accroissement, à une loi binomiale et le recrutement, à une loi de Poisson. À notre connaissance, la calibration d'équations simultanées dont les variables dépendantes suivent différentes distributions reste à être développée.

### ***La dynamique des espèces***

Malgré ces limitations d'ordre statistique, la fiabilité du modèle est renforcée par la validation et les valeurs des estimés des paramètres, lesquelles semblent en partie conformes à l'autécologie des espèces en cause. Ces estimés démontrent clairement que le recrutement de sapin baumier est le plus élevé. Sur les stations étudiées, la

régénération de sapin baumier est généralement plus abondante que celle d'épinette rouge (Ray 1941; Brissette 1996).

Peu importe l'espèce, l'augmentation de la surface terrière induit une diminution du recrutement. En fait, la surface terrière représente une mesure de densité et il apparaît tout à fait normal qu'une densité élevée réduise le recrutement, l'espace de croissance (*cf.* Oliver et Larson 1996) étant moins disponible pour d'éventuelles recrues. Les estimés des paramètres indiquent toutefois que la diminution du recrutement est moins marquée chez l'épinette rouge. Bien que les semis d'épinette rouge soient peu abondants (Brissette 1996), l'espèce possède la capacité de croître dans des conditions de faible luminosité (Seymour 1992). Dans des conditions de faible intensité lumineuse, les semis des espèces de sapin préfèrent diminuer leur croissance en hauteur et privilégient une croissance latérale (Messier *et al.* 1999).

En ce qui concerne l'accroissement en diamètre, la réaction des gaules d'épinette rouge à l'éclaircie n'est pas aussi forte que celle de gaules de sapin baumier (Seymour 1992). Les résultats de cette étude vont dans le même sens. En effet, des surfaces terrières avoisinant les  $10 \text{ m}^2 \cdot \text{ha}^{-1}$  favorisent clairement l'accroissement des gaules de sapin baumier. Lorsque le peuplement est plus dense, c'est-à-dire lorsque l'on observe des surfaces terrières de plus de  $20 \text{ m}^2 \cdot \text{ha}^{-1}$ , les ouvertures de la canopée sont moins importantes, de sorte que l'accroissement moyen des deux espèces est similaire.

Bien que les gaules de sapin réagissent plus fortement à l'éclaircie que celles d'épinette, l'épinette rouge possède la capacité de réagir jusqu'à des âges très avancés (Ray 1956; Seymour 1992). Ainsi, même les tiges de fort diamètre peuvent réagir à l'éclaircie. Toutefois, les résultats de cette étude ne permettent pas de faire ressortir cet élément, puisque ce sont les petites tiges marchandes ( $9 \text{ cm} < d_{hp} < 15 \text{ cm}$ ) d'épinette rouge qui affichent l'accroissement diamétral maximum. En fait, si les tiges de fort diamètre avaient affiché cette réaction à l'éclaircie telle que suggérée dans la littérature, les courbes d'accroissement auraient pris une forme asymptotique.

L'absence d'une telle réaction chez les épinettes rouges de fort diamètre s'explique par la nature de données utilisées pour la calibration du modèle. En effet, les tiges d'épinette rouge de fort diamètre sont peu représentées dans la base de données. Cette situation est en grande partie due aux modalités du traitement réalisé dans les années 1950, au cours duquel les tiges d'épinette rouge et de sapin baumier de plus de 18 cm à la souche ont pratiquement toutes été récoltées (Hatcher 1954). La réaction attendue pourrait sans doute être observée si l'on disposait de plus d'observations parmi les tiges de fort diamètre. En l'absence de telles observations, l'utilisation du modèle à des fins de prédiction devrait se limiter à des contextes similaires à celui de cette étude.

En ce qui concerne le taux de survie, les diminutions abruptes que l'on observe chez les gaules de sapin baumier dans des conditions de surfaces terrières supérieures à  $20 \text{ m}^2 \cdot \text{ha}^{-1}$  sont plutôt surprenantes. En effet, le sapin baumier est une espèce qui s'adapte rapidement aux changements de conditions environnementales. Lors de la fermeture du couvert, le sapin favorise la croissance des branches latérales afin de maximiser son bilan énergétique (Messier *et al.* 1999). En l'absence d'études concernant la mortalité des gaules de sapin baumier et d'épinette rouge, les explications possibles à ce résultat demeurent hypothétiques. L'épinette rouge serait plus tolérante à l'ombre que le sapin baumier et les gaules de cette espèce continueraient de croître en hauteur même dans des conditions de faible luminosité (Seymour 1992). Ainsi, il est possible que cette espèce ait un meilleur taux de survie lors de la fermeture du couvert. Par ailleurs, l'historique des peuplements peut être un facteur important. En effet, les surfaces terrières supérieures à  $20 \text{ m}^2 \cdot \text{ha}^{-1}$  sont observées plusieurs années après le traitement, lorsque la phase de recrutement est terminée. On peut penser qu'un sous-étage de sapin se crée durant les premières années suivant le traitement. Lors de la fermeture du couvert, la phase d'exclusion des tiges s'enclenche (*cf.* Oliver et Larson 1996). La haute régénération du sapin étant abondante, il est possible que l'expansion latérale des branches soit limitée et que la stratégie de croissance en hauteur adoptée par l'épinette rouge soit gagnante.

En ce qui a trait aux tiges marchandes ( $d_{hp} > 9 \text{ cm}$ ), le sapin baumier est plus vulnérable que l'épinette rouge à la carie, au chablis et aux défoliations de la

tordeuse (Hatcher 1959b; Blum 1990; Seymour 1992). Cette vulnérabilité devrait se refléter par une diminution de la probabilité de survie chez les tiges de fort diamètre. Même chez l'épinette rouge, on devrait observer une diminution du taux de survie parmi les tiges de fort diamètre. Une telle relation implique que le carré de la classe de diamètre initial ( $d_{init}$ ) devrait être intégré dans la fonction de survie  $f_{surv}$  de façon à obtenir une expression quadratique plutôt qu'une tendance asymptotique. Dans le cas présent, les efforts visant à intégrer le carré de cette variable se sont avérés vains. Le faible nombre des tiges de fort diamètre dans la base de données pourrait en être responsable.

## Conclusion

Une approche basée sur une maximisation de la vraisemblance a permis de calibrer une matrice de transition à deux niveaux basée sur trois distributions de probabilités discrètes. En permutant les tiges à l'intérieur des listes d'arbres, il est possible d'obtenir une matrice de transition à partir de placettes permanentes dont les tiges ne sont pas identifiées. Il convient de mentionner que les effets de ces permutations ne sont pas connus et que les variances associées aux estimés des paramètres pourraient être sous-estimées. Bien que le processus de calibration ne tienne pas compte de l'autocorrélation entre les données, les résultats de la validation permettent d'affirmer que les prédictions du modèle sont relativement fiables. Par ailleurs, les estimés des paramètres sont en partie conformes à ce que l'on connaît de l'autécologie des espèces, particulièrement en ce qui concerne l'accroissement et le recrutement. En revanche, il apparaît évident que certaines caractéristiques bien connues des deux espèces ne sont pas prises en compte par le modèle. La réaction des épinettes de fort diamètre à l'éclaircie et le taux de mortalité plus élevée chez les sapins matures n'apparaissent pas clairement dans les prédictions du modèle en raison de la nature des données de cette étude. Par ailleurs, les taux de survie des gaules ne peuvent être comparés étant donné l'absence d'études à ce sujet.

Bien que l'utilisation du modèle soit limitée à son contexte d'origine, quelques tendances générales peuvent être mises en lumière. Les conditions de faibles surfaces terrières favorisent le recrutement, l'accroissement et la survie du sapin

baumier. Étant donné que la surface terrière augmente avec le temps, les différences d'accroissement devraient s'atténuer. À la suite d'une coupe partielle de forte intensité, les différences de croissance en diamètre entre les deux espèces sont très faibles sur un horizon de 50 ans (Fortin 2005). Toutefois, le fort recrutement de sapin baumier suggère que la composition de la nouvelle cohorte de tiges sera dominée par cette espèce si la fermeture du couvert tarde à se faire. Ainsi, on peut s'attendre à des changements de composition majeurs dans les peuplements de seconde venue sous des régimes d'aménagement préconisant des coupes partielles de forte intensité et des coupes à blanc. Les coupes partielles de faible intensité apparaissent comme une solution potentielle pour éviter l'« ensapinage » de ses peuplements. En effet, le faible taux de survie associé aux gaules de sapin baumier croissant dans des conditions de surfaces terrières supérieures à  $20 \text{ m}^2 \cdot \text{ha}^{-1}$  pourrait limiter l'expansion du sapin baumier. Il s'agit donc d'adopter une stratégie d'aménagement qui favorise une fermeture rapide du couvert avant que les recrues de sapin baumier n'aient atteint les dimensions marchandes ( $d_{hp} > 9 \text{ cm}$ ).

## **Conclusion générale**

Cette thèse aborde différents aspects d'une même problématique : la raréfaction de l'épinette rouge. Bien que certains chapitres soient principalement axés sur les développements méthodologiques, les objectifs de base demeurent les mêmes : identifier les éléments qui contribuent au processus de raréfaction. Cet ouvrage ne prétend en aucun cas répondre à l'ensemble de la problématique soulevée. En soi, il propose des méthodes statistiques pour l'analyse de la croissance et des distributions diamétrales. Mis en perspective, les résultats de ces méthodologies permettent de mieux comprendre les causes de la raréfaction de l'épinette rouge et de prévoir des interventions sylvicoles afin de l'éviter.

### ***Applicabilité des méthodologies développées***

Les problèmes d'hétéroscédasticité et d'autocorrélation sont communs en modélisation de la croissance forestière (Gregoire *et al.* 1995). La structure de variance-covariance présentée dans le chapitre 2 n'est pas une solution universelle. L'objectif de ce chapitre était essentiellement de démontrer qu'il est possible de corriger l'hétéroscédasticité et l'autocorrélation en fusionnant une fonction de variance et une structure de covariance. Le choix de la fonction de variance et de la structure de covariance dépend de la nature des données. Dans certains cas, une structure autorégressive avec moyenne mobile (ARMA) peut donner de meilleurs résultats (*cf.* Fang *et al.* 2001). Dans un même ordre d'idées, une fonction de variance exponentielle pourrait être appropriée. La sélection de la fonction de variance et de la structure de covariance devrait être dictée non seulement par les critères AIC et BIC, mais également par les caractéristiques des résidus normalisés. La structure appropriée devrait produire des résidus normalisés indépendants entre eux et dont la variance est homogène.

En termes d'applicabilité, le paramétrage des structures de variance-covariance complexifie le processus de calibration, ce qui peut paraître un obstacle. Toutefois, ce paramétrage permet une plus grande flexibilité de la structure. Ainsi, lorsque les paramètres de variance-covariance sont estimés, la structure s'adapte au meilleur de

sa capacité aux données de calibration, ce qui facilite l'application de cette méthodologie à différents contextes forestiers.

L'utilisation de modèles linéaires généralisés comportant des effets mixtes est plutôt rare en foresterie. Le chapitre 3 démontre que ce type de modèle peut être utilisé pour l'analyse des distributions diamétrales. Les limites de la méthodologie proposée reposent principalement sur le fait qu'elle ne tient pas compte des classes de diamètre excédant le 99<sup>ème</sup> percentile de la distribution et qu'elle ne s'applique qu'aux distributions unimodales. En effet, l'utilisation de la fonction de Weibull implique que les distributions diamétrales individuelles ont une forte tendance unimodale. Aussi longtemps que les distributions diamétrales individuelles n'affichent pas de divergence majeure par rapport à ce postulat, la méthodologie proposée peut être appliquée à n'importe quel contexte forestier. L'analyse des distributions diamétrales multimodales demeure une question complexe pour laquelle il n'existe toujours pas de réponse.

La calibration de matrices de transition à deux niveaux est relativement complexe et la méthodologie proposée dans le chapitre 4 n'y fait pas exception. L'utilisation de distributions de probabilités et l'estimation de leurs paramètres par l'optimisation d'une fonction de vraisemblance permet d'obtenir des matrices dont les probabilités sont cohérentes. Cette cohérence est sans aucun doute le principal avantage de la méthodologie proposée. Toutefois, il importe de mentionner que la méthode est grandement limitée du fait qu'elle repose sur certains postulats qui peuvent être raisonnablement mis en doute. Par exemple, on considère que les composantes du modèle, c'est-à-dire le recrutement, l'accroissement et la mortalité, sont indépendantes les unes des autres. La dépendance spatiale entre les tiges d'une même placette n'est pas plus considérée. Étant donné que les tiges de la base de données n'étaient pas individuellement identifiées, ces postulats s'avéraient nécessaires. Dans un contexte où l'identification est disponible, la méthodologie pourrait être développée davantage. Par ailleurs, ce problème d'identification est contourné par un procédé de permutations visant à créer une correspondance fictive entre les tiges. Ce procédé repose sur une optimisation pure et simple de la fonction de vraisemblance et aucune mesure n'est prise pour évaluer l'impact sur les

inférences statistiques. Il n'y a donc aucune certitude que la méthodologie soit fonctionnelle dans un autre contexte forestier pour lequel l'identification des tiges serait manquante.

### ***Conséquences pour l'aménagement***

La problématique de la raréfaction de l'épinette rouge au profit du sapin baumier est soulignée par plusieurs auteurs (Hatcher 1954; Ray 1941, 1956; Seymour 1992; Gordon 1996). L'analyse des données des dispositifs de la rivière Ouareau et du lac Édouard permet de mieux comprendre ce phénomène et de suggérer une stratégie d'aménagement qui puisse maintenir des structures et des compositions typiques des peuplements mixtes de sapin baumier et d'épinette rouge de la forêt primitive (*cf.* Fortin *et al.* 2003a).

La croissance des tiges individuelles des deux espèces n'est pas un facteur majeur dans le phénomène de raréfaction de l'épinette rouge. Même si les gaules et les petites tiges marchandes de sapin baumier réagissent de façon plus dynamique lors du dégagement, la croissance diamétrale de cette espèce tend à ralentir lors de la fermeture du couvert. À long terme, les différences quant à la croissance diamétrale des deux espèces sont relativement faibles et ne peuvent expliquer à elles seules la diminution des proportions d'épinette rouge dans les peuplements de seconde venue.

Le recrutement et la mortalité apparaissent comme les facteurs déterminants dans la dynamique de ces peuplements. Le sapin baumier possède une abondante régénération qui constitue une banque de semis. Après une ouverture de la canopée, le recrutement est généralement important. Si le couvert reste ouvert suffisamment longtemps, ces nouvelles recrues auront l'opportunité d'atteindre le diamètre marchand ( $d_{hp} > 9$  cm). Toutefois, si le couvert se referme rapidement, une bonne proportion de ces nouvelles recrues pourrait ne pas survivre. À l'inverse, le recrutement d'épinette rouge est généralement peu abondant, mais ces quelques recrues ont un meilleur taux de survie lors de la fermeture de la canopée. Ceci serait en partie attribuable au fait que le sapin baumier est plus vulnérable que l'épinette rouge aux défoliations de la tordeuse, au chablis et à la carie (Heimburger 1941; Ray 1941; McLintock 1954; Hatcher 1959b; Seegrift and Arner 1982; Seymour 1992). À

l'échelle du peuplement, ces différences se traduisent par une plus grande variabilité de la croissance en surface terrière marchande chez le sapin baumier. En somme, l'épinette apparaît comme une essence stable qu'il est moins risqué de protéger lors d'une intervention par coupe partielle.

La possibilité d'une homogénéisation de la structure diamétrale après coupe partielle est soulevée par Seymour (1992). À la lumière des résultats de cette thèse, le maintien d'une distribution diamétrale irrégulière dans les peuplements mixtes de sapin baumier et d'épinette rouge n'est pas un élément problématique de l'aménagement de ces peuplements. En effet, à surfaces terrières et densités équivalentes, les peuplements matures issus d'une coupe partielle de forte intensité ont une distribution diamétrale similaire à celle des peuplements matures observés avant le traitement. Il n'en demeure pas moins que le suivi sur plus d'une rotation n'est pas disponible et que l'homogénéisation pourrait survenir à moyen ou à long terme.

En somme, les résultats présentés suggèrent que le maintien des proportions d'épinette rouge passe par un contrôle de la régénération de sapin baumier. Une fermeture rapide du couvert pourrait réduire la croissance des nouvelles recrues de sapin baumier et augmenter leur taux de mortalité. Les coupes partielles de faible intensité sont indiquées pour obtenir cette fermeture rapide du couvert. Une surface terrière résiduelle de  $20 \text{ m}^2 \cdot \text{ha}^{-1}$  après intervention semble être une limite à ne pas dépasser pour contenir la nouvelle cohorte de sapin baumier. Dans le cadre normatif québécois, le jardinage par pied d'arbre est le traitement le plus approprié pour maintenir ce seuil de  $20 \text{ m}^2 \cdot \text{ha}^{-1}$  et favoriser une fermeture rapide de la canopée. Des mesures additionnelles peuvent être envisagées pour contrôler la composition de la nouvelle cohorte, comme l'utilisation d'éclaircies précommerciales, mais ces alternatives dépassent le cadre de cette thèse.

La littérature forestière fournit très peu d'information quant aux processus qui régissent l'abondance de la régénération d'épinette rouge. Il est évident qu'une meilleure connaissance de ces processus pourrait permettre d'adopter des mesures visant à augmenter l'abondance des semis d'épinette rouge. Ces mesures pourraient

diminuer le risque d'un changement de composition au profit du sapin baumier après coupe partielle. Cette avenue de recherche mériterait d'être explorée.

## Bibliographie

- Alexander, J.D., J.R. Donnelly et J.B. Shane. 1995. Photosynthetic and transpirational responses of red spruce understory trees to light and temperature. *Tree Physiol.* 15: 393-398.
- Archambault, L., J. Bégin, C. Delisle et M. Fortin. 2003. Succession forestière après coupe partielle dans la Forêt expérimentale du Lac Édouard, Parc de la Mauricie, Québec. *For. Chron.* 79(3): 672-684.
- Bakuzis, E.V. et H.L. Hansen. 1965. Balsam fir. A monographic review. Univ. of Minnesota Press. Minneapolis. 445 p.
- Bailey, R.L. et T.R. Dell. 1973. Quantifying diameter distributions with the Weibull function. *For. Sci.* 26: 626-632.
- Beaumont, J.-F., C.H. Ung et M. Bernier-Cardou. 1999. Relating site index to ecological factors in black spruce stands: Tests of Hypotheses. *For. Sci.* 45(4): 484-491.
- Bédard, S. 2002. L'estimation du potentiel de croissance des stations forestières : exemple du sous-domaine de la sapinière à bouleau jaune de l'est du Québec. Ministère des Ressources naturelles du Québec, Direction de la recherche forestière. Mémoire de recherche forestière no 140. 36 p.
- Bélanger, L. 1985. Intégration de l'information écologique dans la gestion des terres de l'écoumène forestier par l'inventaire écologique des terres. Thèse de doctorat. Université Laval. 599 p.
- Bélanger, L., S. Paquette, S. Morel, J. Bégin, P. Meek, L. Bertrand, P. Beauchesne, S. Lemay et M. Pineau. 1995. Indices de qualité de station du sapin baumier dans le sous-domaine écologique de la sapinière à bouleau blanc humide. *For. Chron.* 71(3): 317-325.
- Blum, B.M. 1990. Red Spruce. *Dans Silvics of North America*, Vol. 1. USDA Agriculture Handbook no 654. p. 250-259

- Borders, B.E., R.A. Souter, R.L. Bailey et K.D. Ware. 1987. Percentile-based distributions characterize forest stand tables. *For. Sci.* 33(2): 570-576.
- Bowling, E.H., H.E. Burkhart, T.E. Burk et D.E. Beck. 1989. A stand-level multispecies growth model for Appalachian hardwoods. *Can. J. For. Res.* 19: 405-412.
- Brissette, J.C. 1996. Effects of intensity and frequency of harvesting on abundance, stocking and composition of natural regeneration in the Acadian forest of Eastern North America. *Silva Fennica* 30(2-3): 301-314.
- Bruner, H.D. et J.W. Moser Jr. 1973. A Markov chain approach to the prediction of diameter distributions in uneven-aged forest stands. *Can. J. For. Res.* 3: 409-417.
- Busing, R.T. et X. Wu. 1990. Size specific mortality, growth, and structure of a Great Smoky Mountains red spruce population. *Can. J. For. Res.* 20 : 206-210.
- Cao, Q.V. et H.E. Burkhart. 1984. A segmented distribution approach for modeling diameter frequency data. *For. Sci.* 30: 129-137.
- Collett, D. 2003. *Modelling Binary Data* (2<sup>nd</sup> ed.). Chapman & Hall/CRC. New York. 387 p.
- Davidian, M. et D.M. Giltinan. 1995. *Nonlinear models for repeated measurement data*. Chapman & Hall Ed. Monographs on Statistics and Applied Probability 62. 359 p.
- Deleuze, C., O. Pain, J.-F. Dhôte et J.-C. Hervé. 2004. A flexible radial increment model for individual trees in pure even-aged stands. *Ann. For. Sci.* 61: 327-335.
- Dhôte, J.-F. et J.-C. Hervé. 2000. Changements de productivité dans quatre forêts de chênes sessiles depuis 1930 : une approche au niveau du peuplement. *Ann. For. Sci.* 57: 651-680.
- Duan, N. 1983. Smearing estimate: a nonparametric retransformation method. *J. Am. Stat. Assoc.* 78: 605-610.
- Ek, A.R. 1974. Nonlinear models for stand table projection in northern hardwood stands. *Can. J. For. Res.* 4 : 23-27.

- Environnement Canada 2002. Normales et moyennes climatiques au Canada 1971-2000 [en ligne]. La Voie verte. Environnement Canada. Disponible à [http://climate.weatheroffice.ec.gc.ca/climate\\_normals/index\\_f.html](http://climate.weatheroffice.ec.gc.ca/climate_normals/index_f.html) [mis à jour le 10 mai 2004, cité le 27 juillet 2004].
- Fang, Z. et R.L. Bailey. 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *For. Sci.* 47: 287-300.
- Fang, Z., R.L. Bailey et B.D. Shiver. 2001. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects. *For. Sci.* 47: 550-562.
- Fortin, M. 2005. Étude rétrospective de la croissance en diamètre du sapin baumier (*Abies balsamea* (L.) Mill.) et de l'épinette rouge (*Picea rubens* Sarg.) en peuplements mixtes après une coupe à diamètre limite. *For. Chron.* 81: 791-800.
- Fortin, M., J. Bégin et L. Bélanger. 2003a. Évolution de la structure diamétrale et de la composition des peuplements mixtes de sapin baumier et d'épinette rouge de la forêt primitive après une coupe à diamètre limite sur l'Aire d'observation de la rivière Ouareau. *Can. J. For. Res.* 33: 691-704.
- Fortin, M., J. Bégin et L. Bélanger. 2003b. Les coupes partielles : une alternative à la coupe à blanc dans les peuplements mixtes de sapin baumier et d'épinette rouge en termes de rendement. *For. Chron.* 79(5): 948-956.
- Frank, R.M. 1990. Balsam fir. *Dans Silvics of North America*, Vol. 1. USDA Agriculture Handbook no 654. p. 26-35.
- Frank, R.M. et J.C. Bjorkbom. 1973. A silvicultural guide for spruce-fir in the Northeast. USDA For. Serv. Gen. Tech. Rep. NE-6. 29 p.
- Gagnon, G. et G. Marcotte. 1980. Description des types écologiques et de leur productivité dans la section forestière laurentienne de Rowe (L-4a). Gouvernement du Québec. Ministère de l'Énergie et des Ressources. Service de la recherche forestière. Mémoire no. 61. 456 p.

- Gallant, A.R. 1987. Nonlinear statistical models. John Wiley & Sons Ed. Applied Probability and Statistics Series. New York. 610 p.
- Garber, S.M. et D.A. Maguire. 2003. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. *For. Ecol. Manag.* 179: 507-522.
- García, O. 1992. What is a diameter distribution ? *Dans* Proceedings of Symposium on Integrated Forest Management Information Systems, Minowa, M., et S. Tsuyuki (eds.). Japan Society for Forest Planning Press. p.11-29.
- Gaudreau, G. 1986. L'exploitation des forêts publiques au Québec 1842-1905. Institut québécois de recherche sur la culture. Québec. 126 p.
- Ghahramani, S. 2000. Fundamentals of Probability. Second Edition. Prentice Hall Ed. Toronto. 511 p.
- Goelz, J.C.G. et T.E. Burk. 1992. Development of a well-behaved site index equation: jack pine in north central Ontario. *Can. J. For. Res.* 22: 776-784.
- Gordon, A.G. 1996. The sweep of the boreal in time and space, from forest formations to genes, and implications for management. *For. Chron.* 72(1) : 19-30.
- Gosselin, J. 2001. Guide de reconnaissance des types écologiques de la région écologique 3c – Hautes collines du Bas-Saint-Maurice. Forêt Québec. Direction des inventaires forestiers.
- Gosselin, J., P. Grondin et J.-P. Saucier. 2001. Rapport de classification écologique du sous-domaine bioclimatique de l'érablière à bouleau jaune de l'est. Ministère des Ressources naturelles du Québec, Direction de la gestion des stocks forestiers. 169 p.
- Gregoire, T.G. 1987. Generalized error structure for forestry yield models. *For. Sci.* 33(2): 423-444.
- Gregoire, T.G. et M.E. Dyer. 1989. Model fitting under patterned heterogeneity of variance. *For. Sci.* 35(1): 105-125.

- Gregoire, T.G., O. Schabenberger et J.P. Barrett. 1995. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Can. J. For. Res.* 25: 137-156.
- Grondin, P., C. Anseau, L. Bélanger, J.-F. Bergeron, Y. Bergeron, A. Bouchard, J. Brisson, L. De Grandpré, G. Gagnon, C. Lavoie, G. Lessard, S. Payette, P.J.H. Richard, J.-P. Saucier, L. Sirois et L. Vasseur. 1996. Écologie forestière. *Dans Manuel de foresterie*. Les Presses de l'Université Laval, Sainte-Foy. p. 133-279.
- Groot, A. et B.J. Horton. 1994. Age and size structure of natural and second-growth peatland *Picea mariana* stands. *Can. J. For. Res.* 24: 225-233.
- Hall, D.B. et R.L. Bailey. 2001. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *For. Sci.* 47(3): 311-321.
- Hardy, Y., M. Mainville et D.M. Schmitt. 1987. An atlas of spruce budworm defoliation in Eastern North America, 1938-80. USDA For. Serv. Misc. Pub. No. 1449. 52 p.
- Hatcher, R.J. 1954. A report on the establishment of observation area no. 12 on the limits of the Consolidated Paper Corporation Ouareau River, P.Q. 1953. Government of Canada. Department of Northern Affairs and National Resources. Forestry Branch. Forest Research Division. Project Q-54. 31 p.
- Hatcher, R.J. 1959a. Partial cutting with diameter limit control in the Lake Edward Experimental Forest, Quebec, 1950 to 1956. Government of Canada. Department of Northern Affairs and National Resources. Forestry Branch. Forest Research Division. Project Q-44. 8 p.
- Hatcher, R.J. 1959b. Mortality and regeneration following partial cutting of spruce-balsam fir-hardwood stands at Lake Edward, P.Q. Government of Canada. Department of Northern Affairs and National Resources. Forestry Branch. Forest Research Division. Project Q-44. 13 p.
- Hatcher, R.J. 1964. Balsam fir advance growth after cutting in Québec. *For. Chron.* 40(1): 86-92.
- Heimbürger, C.C. 1941. Forest site classification and soil investigation in the Lake Edward Forest Experimental Area. Government of Canada. Department of Mines

- and Resources. Lands, Parks and Forest Branch. Dominion Forest Service. Silvicultural Research Note No. 66. 60 p.
- Johnson, E.W. 2000. Forest Sampling Desk Reference. CRC Press. New York. 985 p.
- Jessome, A. P. 1977. Résistance et propriétés connexes des bois indigènes au Canada. Forintek Canada Corp. Publication SP-514F. Sainte-Foy. 37 p.
- Kuuluvainen, T., A. Penttinen, K. Leinonen et M. Nygren. 1996. Statistical opportunities for comparing stand structural heterogeneity in managed and primeval forests: An example from boreal spruce forest in southern Finland. *Silva Fenn.* 30(2-3): 315-328.
- Lähde, E., O. Laiho, Y. Novokorpi et T. Saksala. 1999. Stand structure as the basis of diversity index. *For. Ecol. Manage.* 104: 157-170.
- Lähde, E., O. Laiho et Y. Norokorpi. 2001. Structure transformation and volume increment in Norway spruce-dominated forests following contrasting silvicultural treatments. *For. Ecol. Manag.* 151: 133-138.
- Lalumière, R. et M. Thibault. 1988. Les forêts du Parc national de la Mauricie au Québec. Étude écologique no 11. Presses de l'Université Laval, Sainte-Foy. 495 p.
- LeMay, V.M. 1990. MSLS: a linear least squares technique for fitting a simultaneous system of equations with a generalized error structure. *Can. J. For. Res.* 20: 1830-1839.
- Lessard, V.C., R.E. McRoberts et M.R. Holdaway. 2001. Diameter growth models using Minnesota Forest Inventory and Analysis Data. *For. Sci.* 47(30): 301-310.
- Linder, P. 1998. Structural changes in two virgin boreal forest stands in central Sweden over 72 years. *Scand. J. For. Res.* 13: 451-461.
- Linder, P., B. Elfving et O. Zackrisson. 1997. Stand structure and successional trends in virgin boreal forest reserves in Sweden. *For. Ecol. Manag.* 98: 17-33.

- Lindstrom, M.J. et D.M. Bates. 1988. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *J. Am. Stat. Assoc.* 83(404): 1014-1022.
- Littell, R.C., G.A. Milliken, W.W. Stroup et R.D. Wolfinger. 1996. SAS System for mixed models. SAS Institute Inc., Cary, NC. 633 p.
- Magnussen, S. 1986. Diameter distributions in *Picea abies* described by the Weibull model. *Scand. J. For. Res.* 1: 493-502.
- Maltamo, M. 1997. Comparing basal area diameter distributions estimated by tree species and for the entire growing stock in a mixed stand. *Silva Fenn.* 31(1): 53-65.
- Maltamo, M., J. Puumalainen et R. Päivinen. 1995. Comparison of beta and Weibull functions for modelling basal area diameter distribution in stands of *Pinus sylvestris* and *Picea abies*. *Scand. J. For. Res.* 10: 284-295.
- Maltamo, M., A. Kangas, J. Uuttera, T. Torniainen et J. Saramäki. 2000. Comparison of percentile based prediction methods and the Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. *For. Ecol. Manag.* 133: 263-274.
- Manly, B.F.J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Second Edition. Chapman & Hall Ed. New York. 399 p.
- McCarthy, J. 2001. Gap dynamics of forest trees : A review with particular attention to boreal forests. *Environ. Rev.* 9: 1-59.
- McCullagh, P. et J.A. Nelder. 1989. Generalized linear models. 2<sup>nd</sup> ed. Monograph on statistics and applied probability 37. Chapman & Hall. New York. 511 p.
- McCulloch, C.E. et S.R. Searle. 2001. Generalized, Linear, and Mixed Models. John Wiley & Sons. New York. 325 p.
- McLintock, T.F. 1954. Factors affecting wind damage in selectively cut stands of spruce and fir in Maine and northern New Hampshire. USDA For. Serv. Northeastern Forest Exp. Stn., Sta. Pap. No. 70. 17 p.

- Ménard, B. 1999. Dynamique naturelle des forêts mixtes de la station expérimentale du lac Édouard, au Parc national de la Mauricie. Mémoire de maîtrise. Université Laval. 87 p.
- Messier, C., R. Doucet, J.-C. Ruel, Y. Claveau, C. Kelly et M.J. Lechowicz. 1999. Functional ecology of advance regeneration in relation to light in boreal forests. *Can. J. For. Res.* 29: 812-823.
- Nigh, G.D., P.V. Krestov et K. Klinka. 2002. Height growth of black spruce in British Columbia. *For. Chron.* 78(2): 306-313.
- Oliver, C.D. et B.C. Larson. 1996. *Forest Stand Dynamics*. Wiley Ed. New York. 520 p.
- Parresol, B.R. 1993. Modeling multiplicative error variance: an example predicting tree diameter from stump dimensions in baldcypress. *For. Sci.* 39(4): 670-679.
- Pienaar, L.V. et K.J. Turnbull. 1973. The Chapman-Richards generalization of von Bertalanffy's growth equation for basal area growth and yield in even-aged stands. *For. Sci.* 19: 2-22.
- Pinheiro, J.C. et D.M. Bates. 2000. *Mixed-effects models in S and S-PLUS*. Springer Ed. New York. 528 p.
- Pinheiro, J.C., D.M. Bates, S. DebRoy et D. Sarkar. 2004. *nlme*: Linear and nonlinear mixed effects models. R package version 3.1-53. The R Project for statistical computing [on line]. Available at <http://www.r-project.org/>.
- Pothier, D. et F. Savard. 1998. Actualisation des tables de production pour les principales espèces forestières du Québec. Gouvernement du Québec. Ministère des Ressources naturelles – Forêt Québec. Direction de la recherche forestière. 183 p.
- Pretzsch, H. et P. Biber. 2005. A re-evaluation of Reineke's rule and stand density index. *For. Sci.* 51: 304-320.
- Pretzsch, H., P. Biber et J. Ďurský. 2002. The single tree-based stand simulator SILVA: construction, application and evaluation. *For. Ecol. Manage.* 162: 3-21.

- Ray, R.G. 1941. Site-types and rate of growth. Government of Canada. Department of Mines and Resources. Lands, Parks and Forest Branch. Dominion Forest Service. Silvicultural Research Note No. 65. 63 p.
- Ray, R.G. 1956. Site-types, growth and yield at Lake Edward Experimental Area Quebec. Government of Canada. Department of Northern Affairs and National Resources. Forestry Branch. Forest Research Division. Technical Note No. 27. 53 p.
- Reineke, L.H. 1933. Perfecting a stand-density index for even-aged forests. J. Agric. Res. 46: 627-638.
- Richards, F.J. 1959. A flexible growth function for empirical use. J. Exp. Bot. 10: 290-300.
- Ricker, M. et R. del Río. 2004. Projecting diameter growth in tropical trees. A new modeling approach. For. Sci. 50(2): 213-224.
- Riopel, M. 1999. Potentiel de croissance de la haute régénération préétablie de sapin baumier : une analyse dendrométrique au niveau de la tige et du peuplement. Mémoire de maîtrise. Université Laval. 149 p.
- Robertson, W.M. 1942. Some growth characteristics of red spruce. Government of Canada. Department of Mines and Resources. Lands, Parks and Forest Branch. Dominion Forest Service. Silvicultural Research Note No. 70. 7 p.
- Robinson, A. 2004. Preserving correlation while modelling diameter distributions. Can. J. For. Res. 34: 221-232.
- Robitaille, A. et J.-P. Saucier. 1998. Paysages régionaux du Québec méridional. Ministère des Ressources naturelles du Québec. 213 p.
- Rowe, J.S. 1972. Les régions forestières du Canada. Ministère de l'environnement du Canada. Service canadien des forêts. Publication no 1300F. 172 p.
- SAS Institute Inc. 2002. SAS OnlineDoc 9 [en ligne]. SAS Institute inc. Cary, NC. Disponible à <http://v9doc.sas.com/sasdoc/>.

- Seegrist, D.W. et S.L. Arner. 1982. Mortality of spruce and fir in Maine in 1976-78 due to spruce budworm outbreak. USDA For. Serv. Res. Pap. NE-491. 3 p.
- Seymour, R.S. 1992. The red spruce-balsam fir forest of Maine : Evolution of silvicultural practice in response to stand development patterns and disturbances. *Dans* The ecology and silviculture of mixed-species forests. Kelty, M.J., B.C. Larson et C.D. Oliver (eds.). Kluwer Publishers, Norwell, MA. p. 217-244.
- Seymour, R.S. 1995. The northeast region. *Dans* Regional silviculture of the United States, 3<sup>rd</sup> Ed. Barrett, J.W. (Ed.). p. 31-79.
- Seymour, R.S. et M.A. Fajvan. 2001. Influence of prior growth suppression and soil on red spruce site index. North. J. Appl. For. 18(2): 55-62.
- Seymour, R.S. et L.S. Kenefic. 2002. Influence of age on growth efficiency of *Tsuga canadensis* and *Picea rubens* trees in mixed-species, multiaged northern conifer stands. Can. J. For. Res. 32: 2032-2042.
- Shannon, C.E. et W. Weaver. 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL. 117 p.
- Siipilehto, J. 1999. Improving the accuracy of predicted basal-area diameter distribution in advanced stands by determining stem number. Silva Fenn. 33(4): 281-301.
- Siipilehto, J. 2000. A comparison of two parameter prediction methods for stands structure in Finland. Silva Fenn. 34(4): 331-349.
- Smith, D.M., B.C. Larson, M.J. Kelty et P.M.S. Ashton. 1997. The practice of silviculture : Applied forest ecology (9<sup>th</sup> ed.). John Wiley & Sons. New York. 537 p.
- Solomon, D.S., R.A. Hosmer et H.T. Hayslett Jr. 1986. A two-stage matrix model for predicting growth of forest stands in the Northeast. Can. J. For. Res. 16: 521-528.
- Solomon, D.S., R.A. Hosmer et H.T. Hayslett Jr. 1987. Fiber handbook: A growth model for spruce-fir and northern hardwood forest types. USDA For. Serv. NE For. Exp. Stat. NE-RP-602. 19 p.

- Steel, R.G.D, J.H. Torrie et D.A. Dickey. 1997. Principles and Procedures of Statistics: A Biometrical Approach. McGraw-Hill Ed. New York. 666 p.
- Steimer, J.-L., A. Mallet, J.-L. Golmard et J.-F. Boisvieux. 1984. Alternative approaches to estimation of population pharmacokinetic parameters: Comparison with the nonlinear mixed-effect model. *Drug Metabolism Reviews* 15: 265-292.
- Sullivan, A.D. et J.L. Clutter. 1972. A simultaneous growth and yield model for loblolly pine. *For. Sci.* 18: 76-86.
- Usher, M.B. 1966. A matrix approach to the management of renewable resources, with special reference to selection forests. *J. Appl. Ecol.* 3: 355-367.
- Usher, M.B. 1969. A matrix model for forest management. *Biometrics* 25: 309-315.
- Valliant, R.L., A.H. Dorfman et R.M. Royall. 2000. Finite Population Sampling and Inference: A Prediction Approach. Wiley Ed. Toronto. 528 p.
- Vanclay, J.K. 1994. Modelling forest growth and yield, applications to mixed tropical forests. CAB International. 312 p.
- Vonesh, E.F. et R.L. Carter. 1992. Mixed-effects nonlinear regression for unbalanced repeated measures. *Biometrics* 48: 1-17.
- Westveld, M. 1953. Ecology and silviculture of the spruce-fir forests of eastern North America. *J. For.* 51: 422-430.
- Wolfinger, R. et M. O'Connell. 1993. Generalized linear models: a pseudo-likelihood approach. *J. Statist. Comput. Simul.* 48: 233-243.
- Wolfinger, R., R. Tobias et J. Sall. 1994. Computing Gaussian likelihoods and their derivatives for general linear mixed models. *SIAM J. Sci. Comput.* 15(6): 1294-1310.
- Zhang, L., J.H. Gove, C. Liu et W.B. Leak. 2001. A finite mixture of two Weibull distributions for modeling the diameter distributions of rotated-sigmoid, uneven-aged stands. *Can. J. For. Res.* 31: 1654-1659.

Zhang, L., K.C. Packard et C. Liu. 2003. A comparison of estimation methods for fitting Weibull and Johnson's  $S_B$  distributions to mixed spruce-fir stands in northeastern North America. *Can. J. For. Res.* 33: 1340-1347.

## Annexe 1 – Relation entre le diamètre avec écorce et le diamètre sans écorce

### ***Relations dhpse vs dhpae***

Pour le sapin baumier,

$$dhpse = 9,50 \times 10^{-1} \cdot dhpae + 3,71 \times 10^{-5} \cdot dhpae^2$$

Pour l'épinette rouge,

$$dhpse = 9,41 \times 10^{-1} \cdot dhpae + 5,68 \times 10^{-5} \cdot dhpae^2$$

où *dhpae* représente le diamètre à hauteur de poitrine avec écorce (mm) et *dhpse*, le diamètre à hauteur de poitrine sans écorce (mm).

## Annexe 2 – Équations pour calculer la croissance diamétrale des tiges de sapin baumier et d'épinette rouge en fonction du type écologique

### ***Équations tirées du modèle [1] du chapitre 1***

#### Type écologique riche sur dépôt très mince

Pour le sapin baumier,

$$\Delta dhpae = (-6,49 + 8,87 \times 10^{-3} dhpinit) \ln(t + 1) + (9,15 + 7,47 \times 10^{-3} dhpinit) \ln^2(t + 1),$$

Pour l'épinette rouge,

$$\Delta dhpae = (-2,93 - 5,18 \times 10^{-2} dhpinit) \ln(t + 1) + (5,80 + 3,74 \times 10^{-2} dhpinit) \ln^2(t + 1)$$

#### Type écologique riche sur dépôt épais

Pour le sapin baumier,

$$\Delta dhpae = (-1,14 \times 10^1 + 2,67 \times 10^{-2} dhpinit) \ln(t + 1) + (1,02 \times 10^1 + 7,47 \times 10^{-3} dhpinit) \ln^2(t + 1)$$

Pour l'épinette rouge,

$$\Delta dh_{pae} = (-7,81 - 8,39 \times 10^{-3} dh_{pinit}) \ln(t + 1) + (6,33 + 3,74 \times 10^{-2} dh_{pinit}) \ln^2(t + 1)$$

Type écologique pauvre sur dépôt très mince

Pour le sapin baumier,

$$\Delta dh_{pae} = (-9,32 - 1,67 \times 10^{-2} dh_{pinit}) \ln(t + 1) + (9,09 + 7,47 \times 10^{-3} dh_{pinit}) \ln^2(t + 1)$$

Pour l'épinette rouge,

$$\Delta dh_{pae} = (-8,05 - 5,18 \times 10^{-2} dh_{pinit}) \ln(t + 1) + (9,07 + 3,74 \times 10^{-2} dh_{pinit}) \ln^2(t + 1)$$

où  $\Delta dh_{pae}$  représente la croissance en diamètre à hauteur de poitrine avec écorce après la coupe (mm),  $t$  représente le nombre d'années écoulées depuis la coupe et  $dh_{pinit}$ , le diamètre au moment de la récolte (mm).