



**HAL**  
open science

# Quel est le niveau de détail pertinent pour modéliser la croissance d'une forêt mélangée ? Comparaison d'une famille de modèles et application aux peuplements mélangés chêne sessile - pin sylvestre

Thomas Perot

## ► To cite this version:

Thomas Perot. Quel est le niveau de détail pertinent pour modéliser la croissance d'une forêt mélangée? Comparaison d'une famille de modèles et application aux peuplements mélangés chêne sessile - pin sylvestre. Sciences de l'environnement. Doctorat en Sciences agronomiques et forestières, Institut des Sciences et Industries du Vivant et de l'Environnement (Agro Paris Tech), 2009. Français. NNT : 2009AGPT0033 . tel-02592299

**HAL Id: tel-02592299**

**<https://hal.inrae.fr/tel-02592299v1>**

Submitted on 15 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° / / / / / / / / / / / / / / / /

**Thèse**

Pour obtenir le grade de

**Docteur**  
de

**l'Institut des Sciences et Industries du Vivant et de l'Environnement  
(Agro Paris Tech)**

Spécialité : *Sciences agronomiques et forestières*

présentée et soutenue publiquement par

**Thomas PEROT**

Le 26/06/2009

---

**Quel est le niveau de détail pertinent pour modéliser la croissance d'une  
forêt mélangée ?**

**Comparaison d'une famille de modèles et application aux peuplements  
mélangés chêne sessile - pin sylvestre**

---

Directeur de thèse : *Jean-François DHÔTE*

Codirecteurs de thèse : *Christian GINISTY* et *François GOREAUD*

Travail réalisé au *Cemagref, Unité Ecosystèmes Forestiers, 45 290 Nogent-sur-Vernisson.*

Devant le jury :

Benoit COURBAUD	:	Rapporteur
Sebastian HEIN	:	Rapporteur
Meriem FOURNIER	:	Présidente
Isabelle VINKLER	:	Examineur
Jean-François DHÔTE	:	Directeur de thèse
Christian GINISTY	:	Codirecteur de thèse



## Remerciements

Quand on commence à rédiger cette partie ce n'est évidemment que du plaisir. On se rend compte également que tout ce travail n'a pu voir le jour que grâce à de nombreuses personnes et grâce aux résultats de travaux précédents.

Merci à l'école doctorale ABIES pour les formations de qualité qui y sont proposées. Je pense notamment aux formations assurées par Eric Birlouez et Donald White. Merci également d'avoir financé une partie de mon voyage au Japon pour le colloque IUFRO “uneven-aged silviculture”.

Merci à l'Office National des Forêts pour les échanges très réguliers sur nos travaux, pour nous avoir autorisé à sonder plus de 500 arbres en forêt domaniale d'Orléans et pour son soutien financier dans le cadre de la convention Cemagref ONF.

Merci aux responsables de l'unité EFNO du Cemagref de Nogent-sur-Vernisson (Bernard Abrial, Michel Denis et Jean-Jacques Tolron) de m'avoir soutenu dans la réalisation de ce projet.

Merci à Pierre Auger, Thomas Cordonnier, Philippe Dreyfus et Quentin Ponette d'avoir accepté de faire partie de mon comité de thèse.

Merci à Sebastian Hein, Benoit Courbaud, Isabelle Vinkler et Meriem Fournier d'avoir accepté de faire partie de mon jury de thèse.

Merci à Jean-François Dhôte d'avoir accepté d'être mon directeur de thèse. Nos réunions ont souvent marqué un tournant dans l'évolution de ce travail.

Merci à Christian Ginisty de m'avoir accueilli au Cemagref en 2002. Merci de m'avoir proposé des sujets enrichissants entre 2002 et 2005 qui ont certainement contribué à me mettre en condition pour ce projet de thèse. Enfin, à chaque étape de ce travail tes remarques et tes interrogations m'ont amené à me poser de nouvelles questions et à approfondir l'argumentation de mes résultats.

Merci à François Goreaud pour avoir cru en moi et pour m'avoir proposé ce sujet de thèse très riche. Merci d'avoir accepté de continuer à m'encadrer après ton départ du Cemagref. Merci enfin d'avoir toujours su trouver les mots pendant les moments difficiles.

Merci à François de Coligny pour sa disponibilité et son aide dans l'implémentation des modèles sous Capsis en particulier pour le modèle utilisant les distributions de voisinage.

Merci à Christine Deleuze pour l'intérêt qu'elle a porté à ce travail, pour ses éclairages et les discussions toujours agréables que nous avons eues.

Merci à Nicolas Picard pour ses travaux qui ont été précieux pour cette thèse. Merci aussi pour sa disponibilité dans les quelques échanges que nous avons eu.

Merci à tous les membres de l'équipe "Forêt hétérogène" pour leurs contributions à ce travail. Je pense aux membres actuels mais aussi aux membres passés et particulièrement à Jean-Noël Piroche et Régis Allain qui sont à l'origine des premiers travaux sur le mélange chêne-pin réalisés dans l'équipe.

Merci à Marie Ange et Noémie mes voisines doctorantes successives de bureau. Merci à Marie Ange d'avoir ouvert la voie des doctorants de l'équipe sur le projet chêne-pin et à Noémie qui reprend le flambeau. Merci aussi pour leur bonne humeur communicative.

Merci à Franck Milano qui s'est occupé des contacts avec l'INRA de Champenoux pour la mise au point de la "carotteuse" thermique et sa participation lors de la première campagne de récolte des carottes en 2006. Merci à Grégory Décelière pour son aide lors de l'échantillonnage de la première campagne de terrain. Merci à Sébastien Marie et Romain Vespierre pour leur aide dans la préparation des carottes et la mesure des largeurs de cerne (1020 carottes et 35 161 cernes). Merci à Sandrine Perret pour son aide dans la définition des protocoles de mesure, pour la réalisation des mesures de hauteur, pour la réalisation des inventaires et pour sa gestion de la base de données. Merci à Yann Dumas pour les analyses pédologiques sur l'ensemble des dispositifs.

Merci à Gilles Bergevin et Philippe Guillemard pour leurs coups de main informatiques rapides et efficaces.

Merci à Josette Baptiste et Sylvie Bourgeois pour leur efficacité dans les tâches administratives et pour leur bonne humeur.

Merci à Sonia Launay pour son aide et sa veille pour la biblio.

Merci à tous les habitants des "Pommiers" pour leur contribution au climat chaleureux de ce bâtiment.

Merci à tous ceux qui ont pris le temps de relire ma thèse ou une partie. Les nombreuses corrections et remarques ont permis d'améliorer le manuscrit.

Merci à mes parents et à mon frère pour leur soutien tout particulièrement le jour de la soutenance et pour l'organisation de la partie festive.

Enfin merci à Sonia, Etienne et Raphaël pour leur patience et leur soutien.

## Résumé

Face à l'intérêt grandissant pour les forêts mélangées, des outils et des modèles adaptés à leur gestion sont nécessaires. L'objectif de cette thèse est de montrer comment la construction et la comparaison de modèles ayant différents niveaux de détail peuvent nous aider à choisir le niveau le plus approprié pour modéliser la croissance d'un peuplement mélangé dans un contexte d'application donné.

A partir de données récoltées dans des peuplements mélangés chêne sessile (*Quercus petraea* L.) pin sylvestre (*Pinus sylvestris* L.), nous avons construit une famille de modèles à différents niveaux de détail : un modèle arbre indépendant des distances (MAID), un modèle arbre dépendant des distances (MADD), trois modèles peuplement et un modèle intermédiaire entre le MAID et le MADD utilisant des distributions de voisinage. Pour que la comparaison de ces modèles soit pertinente, nous avons assuré une cohérence entre les modèles en utilisant plusieurs approches.

Ces modèles nous ont permis d'acquérir des connaissances sur la croissance et la dynamique de ces forêts en particulier sur les interactions spatiales et temporelles entre le chêne et le pin. Ainsi le MAID a permis de montrer qu'il peut y avoir des phénomènes de compensation de croissance entre les deux espèces. Le MADD a permis de montrer que dans ces peuplements la compétition intraspécifique est supérieure à la compétition interspécifique. Un modèle peuplement obtenu à partir du MADD a permis d'étudier l'influence du taux de mélange sur la production.

Pour évaluer la qualité prédictive des modèles, nous avons utilisé un jeu de données indépendant obtenu en partageant nos données avant la construction des modèles. Nous avons ainsi montré que le MAID était plus performant que le MADD pour prédire les accroissements individuels. Les modèles ont aussi été comparés sur des exemples d'applications mettant en œuvre des simulations à court ou moyen terme.

La démarche proposée présente un intérêt aussi bien pour la compréhension du phénomène étudié que pour sa modélisation dans un but prédictif. En regroupant l'ensemble des résultats acquis, ce travail nous a permis d'apprécier la pertinence d'un type de modèle en fonction du contexte d'utilisation. Cette démarche très générale pourrait être appliquée à la modélisation d'autres processus comme la mortalité ou la régénération.

**Mots-Clés** : forêt mélangée ; niveau de détail ; modélisation ; modèle arbre indépendant des distances ; modèle arbre dépendant des distances ; modèle peuplement ; distribution de voisinage ; comparaison de modèles ; évaluation de modèles ; pin sylvestre ; chêne sessile ; interaction entre espèces ; croissance en diamètre.

---

## Abstract

Appropriate tools and models are needed for the management of mixed forests. The aim of this thesis is to show how the construction and the comparison of models with different levels of detail can help us to choose the most appropriate level to model the growth of a mixed stand.

We developed a family of models at different levels of detail from data collected in mixed stands of sessile oak (*Quercus petraea* L.) and Scots pine (*Pinus sylvestris* L.) : a tree distance independent model (MAID), a tree distance dependent model (MADD), three stand models and an intermediate model bridging the MAID and the MADD. We ensured consistency between models using several approaches in order to make relevant comparisons.

These models have given us some knowledge on the growth and dynamics of these forests, in particular on the spatial and temporal interactions between oaks and pines. Thus, we showed a compensatory growth phenomenon between the two species using the MAID. The MADD made it possible to show that, in these stands, the intraspecific competition was stronger than interspecific competition. A stand model developed from the MADD helped us to study the influence of mixing rate on production.

To assess the quality of models predictions, we used an independent data set obtained by splitting our data. For example, we have shown that the MAID was more efficient than the MADD to predict individual increments. The models were also compared on examples of applications with short or medium term simulations.

The proposed approach is of interest for both understanding the studied phenomenon and developing predictive tools. The different results of this work, allowed us to assess the relevance of a type of model for different contexts of use. This very general approach could be applied to the modeling of other processes such as mortality or regeneration.

**Keywords** : mixed stand; level of detail; modeling; tree distance independant model; tree distance dependant model; stand model; neighbourhood distribution; model comparison; model evaluation; Scots pine; sessile oak; species interaction; growth.



## Sommaire général

<b>1. Introduction .....</b>	<b>1</b>
1.1 Les forêts mélangées : un intérêt croissant qui entraîne un besoin d'outil d'aide à la gestion .....	1
1.2 Les modèles de croissance peuvent faciliter la gestion des forêts mélangées.....	2
1.3 Comment choisir le niveau de détail d'un modèle pour une forêt mélangée ? .....	5
1.4 Objectif de la thèse et stratégie .....	7
<b>2. Comment assurer une cohérence entre des modèles ayant des niveaux de détail différents ? .....</b>	<b>10</b>
2.1 Du modèle peuplement au modèle arbre dépendant des distances.....	11
2.2 Aperçu des méthodes utilisées pour assurer une cohérence entre des modèles .....	18
2.3 Assurer une cohérence en ajustant les modèles sur les mêmes données .....	20
2.4 Assurer une cohérence par agrégations ou simplification d'un modèle détaillé.....	21
2.5 Assurer une cohérence par des outils de changement d'échelle .....	26
2.6 Bilan.....	31
<b>3. Données récoltées dans les forêts mélangées chêne sessile - pin sylvestre.....</b>	<b>32</b>
3.1 Les forêts de chêne sessile et pin sylvestre : un bon exemple de peuplement mélangé.....	33
3.2 Modèles existants pour la croissance du chêne sessile et du pin sylvestre.....	34
3.3 Choix des arbres et mesures dendrométriques .....	36
3.4 Mesure de la croissance individuelle .....	44
3.5 Partage des données pour ajuster et évaluer les modèles .....	48
3.6 Bilan.....	49
<b>4. Construction et ajustement d'une famille de modèles de croissance.....</b>	<b>50</b>
4.1 Le modèle arbre indépendant des distances .....	51
4.2 Le modèle arbre dépendant des distances .....	74
4.3 Le modèle arbre basé sur des distributions de voisinage .....	86
4.4 Construction de modèles peuplement par agrégation des modèles individus-centrés.....	93
4.5 Bilan.....	109
<b>5. Evaluation et comparaison des modèles pour prédire la croissance .....</b>	<b>112</b>
5.1 Critères et méthodes utilisables pour évaluer des modèles de croissance forestiers .....	114
5.2 Prédire les accroissements individuels, leur variabilité et leur distribution .....	129
5.3 Prédire l'accroissement du peuplement .....	142
5.4 Prédire l'évolution du peuplement dans des contextes appliqués.....	154
5.5 Bilan.....	170
<b>6. Discussion générale et perspectives.....</b>	<b>171</b>
6.1 Bilan des principaux résultats .....	171
6.2 Synthèse sur les critères permettant de choisir un type de modèle .....	176
6.3 Perspectives .....	183
<b>7. Conclusion.....</b>	<b>187</b>

<b>Bibliographie .....</b>	<b>188</b>
<b>ANNEXE A : Agrégation du MAID en un modèle peuplement.....</b>	<b>203</b>
<b>ANNEXE B : Evolution de la circonférence au cours du temps dans les 9 dispositifs.....</b>	<b>205</b>
<b>ANNEXE C : Implémentation du MADV : le module Oakpine2 dans Capsis .....</b>	<b>206</b>
<b>ANNEXE D : Calcul du taux de mélange optimal à partir du modèle peuplement .....</b>	<b>219</b>
<b>ANNEXE E : Indices et tests statistiques utilisés pour évaluer des modèles de croissance .....</b>	<b>221</b>
<b>ANNEXE F : Présentation des données d'évaluation par dispositif pour le chêne. ....</b>	<b>225</b>
<b>ANNEXE G : Présentation des données d'évaluation par dispositif pour le pin.....</b>	<b>228</b>
<b>ANNEXE H : Evaluation du MADD et du MAID .....</b>	<b>231</b>

## 1. Introduction

### 1.1 *Les forêts mélangées : un intérêt croissant qui entraîne un besoin d'outil d'aide à la gestion*

#### 1.1.1 **Les forêts mélangées suscitent un intérêt croissant**

Depuis le début des années 1990 et la volonté au niveau mondial de lutter contre l'érosion de la biodiversité (ONU, 1992 ; ONU, 1993), les forêts mélangées font l'objet d'un intérêt croissant. Récemment la question de l'adaptation des écosystèmes face aux changements climatiques a encore renforcé cet intérêt (voir par exemple Sevrin *et al.*, 2008). Selon Legay *et al.* (2008), le mélange d'essences pourrait permettre aux forestiers de faire face à des changements qui vont affecter le fonctionnement et la production des forêts dans les prochaines décennies. Ce mélange vise également à satisfaire une demande sociale toujours bien présente pour une meilleure prise en compte de l'environnement dans les pratiques de gestion. La composition en essences, fait ainsi partie des critères de gestion durable adoptés lors de la conférence de Vienne en 2003 sur la protection des forêts en Europe (MCPFE, 2003). En outre, de plus en plus d'études scientifiques montrent l'avantage de gérer des peuplements mélangés par rapport à des peuplements purs. Ainsi, le mélange peut permettre de réduire les dégâts face à l'attaque d'insectes phytophages (Jactel et Brockerhoff, 2007) et peut permettre d'augmenter la productivité du peuplement par une meilleure occupation de l'espace, une meilleure exploitation des ressources et parfois grâce à des phénomènes de facilitation entre essences (Kelty, 2006).

#### 1.1.2 **Savoir gérer les forêts mélangées : un enjeu pour la filière forestière**

Même si certains avantages restent discutables, comme l'intérêt du mélange pour la conservation de la biodiversité (Barbier, 2007), et que d'autres sont subjectifs comme l'intérêt paysager, les forêts mélangées offrent de nombreux bénéfices (Bristow *et al.*, 2006a). Malgré cela, au niveau mondial, la grande majorité des forêts plantées sont encore des monocultures (Kelty, 2006 ; FAO, 2007 ; Jactel *et al.*, 2008). Ceci résulte essentiellement d'habitudes d'industriels qu'il est difficile de faire évoluer (Nichols *et al.*, 2006). Néanmoins, dans certains pays et notamment en France, le mélange d'essences est bien ancré dans les pratiques (Nichols *et al.*, 2006) et fait partie intégrante des orientations de gestion (Sardin *et al.*, 2008). De plus, les récentes statistiques de l'Inventaire Forestier National font apparaître que les forêts

mélangées<sup>1</sup> représentent 49% des forêts de production de la France métropolitaine soit 7 200 000 ha (Morneau *et al.*, 2008). Savoir gérer les peuplements mélangés est donc un enjeu important pour la filière forestière.

### **1.1.3 Les forestiers ont besoin d'outils pour gérer ces peuplements**

Aujourd'hui, gérer les forêts de façon durable pose des difficultés car les demandes de la société, comme la demande croissante en bois énergie (Nabuurs *et al.*, 2007), peuvent entrer en conflit avec d'autres enjeux de la gestion durable des forêts : conservation de la biodiversité (Jonsell, 2007), conservation des ressources du sol (Cacot *et al.*, 2006). Dans un tel contexte, les forestiers ont besoin d'outils d'aide à la gestion capables d'intégrer des enjeux de nature variée (Von Teuffel *et al.*, 2006). D'un point de vue sylvicole, il existe un savoir-faire non négligeable concernant la gestion des peuplements mélangés mais le mélange d'essences pose encore de nombreuses difficultés aux forestiers (Sardin *et al.*, 2008). En effet, un peuplement mélangé présente beaucoup plus d'options de gestion qu'un peuplement pur à cause de ses nombreuses sources de variation. Cela renforce l'intérêt d'avoir des outils d'aide à la gestion pour ce type de peuplements (Pukkala *et al.*, 1998).

## ***1.2 Les modèles de croissance peuvent faciliter la gestion des forêts mélangées***

### **1.2.1 Les modèles de croissance et de dynamique forestière : des outils prometteurs encore peu appliqués aux peuplements mélangés**

Les modèles de croissance et de dynamique forestière (voir Houllier *et al.*, 1991 pour une définition de ces modèles) peuvent contribuer à répondre aux questions que l'on se pose sur la gestion des forêts mélangées (Dhôte *et al.*, 2005). En effet, ces outils peuvent apporter aux gestionnaires de nombreux éléments de réponse ou de réflexion (Pretzsch *et al.*, 2002a ; Goreaud *et al.*, 2005a) :

- ils permettent d'acquérir des connaissances qualitatives et quantitatives sur le fonctionnement des peuplements ;

---

<sup>1</sup> Au sens de l'Inventaire Forestier National (IFN) une forêt est mélangée lorsque le taux de couvert relatif de l'essence principale ne dépasse pas 75%. Le taux de couvert relatif d'une espèce est le rapport entre la surface projetée de la partie des houppiers de l'espèce accédant à la lumière et la somme des surfaces projetées de toutes les espèces. Pendant les campagnes d'inventaire de l'IFN, ces taux de couvert sont estimés sur des placettes de 25 m de rayon selon une grille qui couvre l'ensemble des zones forestières françaises.

- ils permettent de prédire l'évolution des peuplements dans des cas de figure complexes, en intégrant par exemple les changements de productivité (Bontemps, 2006) ou en considérant des peuplements irréguliers ou mélangés ;
- ils rendent possible une comparaison rapide de plusieurs scénarios sylvicoles, ce qui aurait demandé auparavant de longues années d'expérimentation. Cette utilisation est tout particulièrement indiquée lors de la rédaction des guides de sylviculture ;
- la modélisation permet également d'intégrer des phénomènes à plus large échelle de temps et d'espace, comme la colonisation successive d'un espace par différentes espèces, qu'il nous est difficile d'appréhender autrement ;
- ce sont d'excellents outils pédagogiques, pour la formation initiale ou continue, car ils facilitent la représentation des peuplements, par exemple dans des expériences de marteloscope, ou en permettant aux étudiants de tester et de comparer plusieurs itinéraires sylvicoles.

Au niveau international, il existe maintenant de nombreux exemples de modèles adaptés aux peuplements mélangés (Porté et Bartelink, 2002). En ce qui concerne la France métropolitaine, l'effort de construction de modèles de croissance a jusqu'à présent surtout porté sur les peuplements réguliers (Perot et Ginisty, 2004). Concernant les peuplements mélangés, quatre mélanges différents ont fait l'objet ou font l'objet de travaux de modélisation : le mélange hêtre, sapin et épicéa dans le Jura (Bruciamacchie *et al.*, 1991 ; Buongiorno *et al.*, 1995), le mélange sapin et épicéa dans les Alpes (Courbaud, 2004), le mélange frêne et hêtre dans le Nord-est (Ottorini et Le Goff, 2002), et le mélange hêtre, pin noir et sapin dans l'arrière-pays méditerranéen (Dreyfus, 2008). Ces travaux sont loin de couvrir l'ensemble des types de mélange que l'on peut rencontrer en France (Morneau *et al.*, 2008) et ne couvrent pas l'ensemble des régions géographiques. Il est donc nécessaire de poursuivre la construction de modèles adaptés aux peuplements mélangés pour mieux répondre aux demandes des gestionnaires ou des pouvoirs publics.

### **1.2.2 Une tendance qui va vers des modèles de plus en plus détaillés**

Les modèles de croissance et de dynamique forestière que nous venons de présenter peuvent avoir des caractéristiques très différentes. En particulier, le peuplement forestier peut y être décrit de manière plus ou moins détaillée. Nous proposons d'utiliser l'expression "niveau de détail" d'un modèle pour caractériser la façon dont le peuplement est décrit dans le modèle. On peut aussi parler de "niveau de description" (Picard, 1999 p. 2), d'"échelle de modélisation" ou de "degré de précision" (Edwards, 2004). On peut ainsi définir plusieurs

grands types de modèles correspondant à des niveaux de détail différents (Munro, 1974 ; Houllier *et al.*, 1991 ; Porté et Bartelink, 2002) : les modèles "peuplement" (ou "arbre moyen") où le peuplement est décrit par des variables globales (exemple : surface terrière totale) ou moyennes (exemple : diamètre moyen), les modèles de distribution où le peuplement est décrit par une distribution par classe de taille des arbres (exemple : distribution par classe de diamètre) et les modèles individus-centrés où le peuplement est décrit par des caractéristiques concernant tous les arbres du peuplement (exemple : liste des diamètres individuels). Parmi ce dernier type de modèles, on distingue généralement les modèles individus-centrés non spatialisés ou Modèles Arbre Indépendants des Distances (MAID) pour lesquels les positions spatiales des arbres dans le peuplement ne sont pas connues, et les modèles individus-centrés spatialement explicites ou Modèles Arbre Dépendants des Distances (MADD) qui utilisent la position spatiale de chaque arbre dans le peuplement.

L'étude réalisée en 2002 par Porté et Bartelink a recensé environ 130 modèles publiés entre 1947 et 2000 portant sur des peuplements mélangés. La description de ces modèles, disponible dans cet article, nous a permis de les répartir par type de modèles et par année de parution. La figure 1.1 montre que les premiers modèles construits pour des peuplements mélangés étaient des modèles peuplement et que les autres types de modèles sont apparus à partir des années 1970. L'émergence des modèles arbre dépendants des distances est récente, plutôt début des années 1990 ; ils deviennent majoritaires dans la dernière période 1998-2000. On observe donc une tendance à la construction de modèles de plus en plus détaillés (voir aussi Picard, 2007 p. 27).

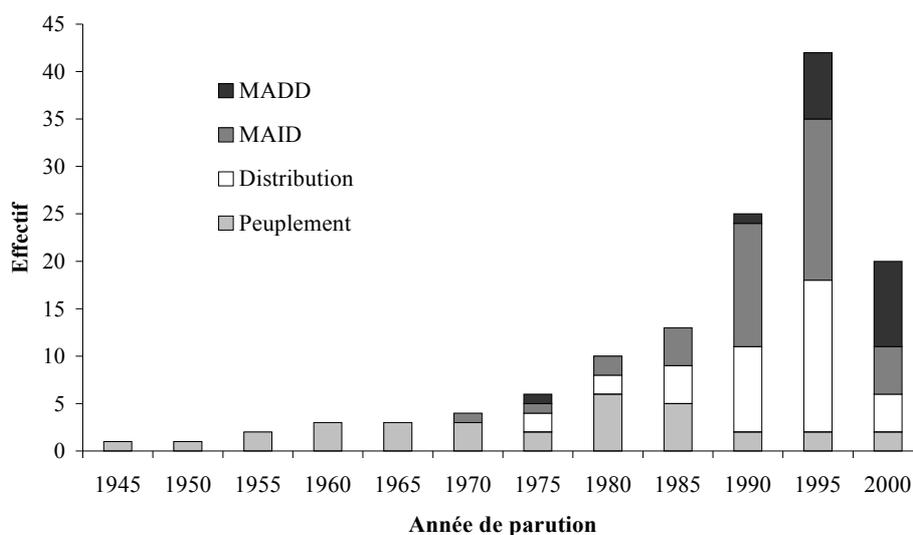


Figure 1.1 : Répartition des travaux de modélisation sur les peuplements mélangés par type de modèles et par année de parution d'après Porté et Bartelink (2002). Les années ont été regroupées par classe de 5 ans sauf pour la classe 2000 qui comprend 3 années (1998, 1999 et 2000).

### ***1.3 Comment choisir le niveau de détail d'un modèle pour une forêt mélangée ?***

Si les questions posées par la société ont pu évoluer, le développement des connaissances mathématiques et surtout l'amélioration de l'informatique expliquent en partie l'évolution que nous venons de présenter. Aujourd'hui, nous supposons que, pour un objectif et des contraintes donnés, il est possible de construire des modèles ayant des niveaux de détail différents. On peut alors légitimement se demander lequel de ces modèles est le plus pertinent pour la question posée.

#### **1.3.1 Une question générale en modélisation**

La pertinence d'un modèle individus-centré par rapport à un modèle plus agrégé est une question générale et d'actualité en modélisation (Edwards *et al.*, 2005 ; Picard, 2007). Dans de nombreux domaines, après une période de développement intense de modèles individus-centrés ou de systèmes multi-agents (voir Bousquet, 2001 p. 23), les modélisateurs se sont rendus compte que ces modèles répondaient parfois moins bien aux objectifs des utilisateurs que des modèles plus frustes (Ek et Monserud, 1979 ; Iwasa *et al.*, 1987 ; Grimm, 1999).

Dans le domaine forestier, les modèles peuplement qui utilisent des données simples sont bien adaptés aux contraintes de la gestion forestière. Ces modèles sont relativement efficaces dans le cadre des peuplements purs et réguliers (même âge pour tous les arbres). En effet, comme les arbres y ont des conditions de croissance très proches, des variables moyennes ou globales, comme la densité ou la surface terrière totale, peuvent être utilisées pour estimer l'intensité de la compétition qui s'exerce sur les individus (Houllier *et al.*, 1991). En revanche, dans les peuplements complexes (mélange d'espèces ou irrégularité en âge), on observe généralement une forte variabilité des individus et de leurs conditions de croissance (environnement local), si bien que l'approche peuplement semble moins pertinente. Des travaux ont par exemple montré que la structure spatiale du mélange pouvait jouer un rôle essentiel dans l'évolution de la composition du peuplement (Goreaud *et al.*, 2002). Ceci plaide plutôt pour la construction d'un modèle à l'échelle de l'arbre, et spatialement explicite. Plusieurs applications ont pu illustrer l'intérêt de ces modèles, notamment pour simuler l'effet local de différents scénarios sylvicoles (Courbaud *et al.*, 2001). Cependant, ils posent de réels problèmes de transfert vers les gestionnaires, à cause du grand volume de données qu'ils demandent. En effet, ces modèles nécessitent la cartographie d'un peuplement qui est une

information coûteuse inaccessible dans un cadre de gestion courante (Munro, 1974 ; Wimberly et Bare, 1996).

Le choix du type de modèle à construire pour les peuplements mélangés est donc une question difficile car, la volonté de rendre compte des phénomènes de croissance et de dynamique de façon réaliste, peut entrer en conflit avec la possibilité d'utilisation du modèle dans un cadre appliqué.

### **1.3.2 Quelques principes utilisés pour choisir le niveau de détail d'un modèle**

Choisir un type de modèle est une première étape délicate du travail de modélisation qui doit tenir compte à la fois des spécificités du peuplement considéré, des données disponibles, et de l'objectif du modèle (Pavé, 1994). Le niveau de détail du modèle peut être défini de manière experte. Ainsi, d'après Pretzsch (2000), l'objectif du modèle et la connaissance existante sur le système observé permettent de déterminer la complexité de l'approche de modélisation. Dans la pratique, le choix du niveau de détail du modèle est souvent dicté par les données disponibles. Il existe aussi des principes permettant de guider ce choix (Burkhart, 2003). L'un d'entre eux consiste à construire le modèle à un niveau plus détaillé que le niveau souhaité pour les prédictions. D'après Burkhart (2003), cette recommandation incite à la construction de modèles individus-centrés dans le cadre de la modélisation des peuplements forestiers. Or, selon lui, même si les modèles individus-centrés ont prouvé leur flexibilité pour la description de la structure des peuplements et pour la simulation des interventions sylvicoles, ils peuvent ne pas être aussi précis que des modèles agrégés pour estimer l'évolution de variables descriptives du peuplement comme le volume ou la surface terrière par hectare. Ce principe présente donc des limites. Un autre principe est basé sur l'échelle de temps des projections que l'on souhaite réaliser avec le modèle : pour conserver un même niveau de précision, plus le temps de projection est important plus le modèle doit être agrégé (voir Burkhart, 2003 pour plus de détails).

### **1.3.3 La comparaison de modèles peut faciliter le choix du niveau de détail**

Les principes évoqués plus haut sont des guides intéressants mais ils laissent une part de subjectivité importante dans le choix du niveau de détail. Une autre approche, plus objective, consiste à construire des modèles à différents niveaux de détail puis à comparer les résultats obtenus par ces modèles, et leur adéquation à l'objectif retenu et aux contraintes d'utilisation (Lett *et al.*, 1999 ; Dieckmann *et al.*, 2000 ; Edwards *et al.*, 2005 ; Gourlet-Fleury *et al.*, 2005). La comparaison des modèles peut alors être effectuée sur des critères très variés, allant de la précision des prédictions à la facilité d'appropriation par les utilisateurs. C'est cette démarche que nous avons choisi de mettre en œuvre dans cette thèse.

## **1.4 Objectif de la thèse et stratégie**

**L'objectif de cette thèse est de montrer comment la construction et la comparaison de modèles ayant différents niveaux de détail peuvent nous aider à choisir le niveau le plus approprié pour modéliser la croissance d'un peuplement mélangé dans un contexte d'utilisation donné.**

Pour atteindre cet objectif, nous avons construit une stratégie en quatre points que nous présentons ci-dessous en même temps que le plan de ce mémoire.

### **1.4.1 Assurer la cohérence entre des modèles ayant des niveaux de détail différents**

Pour qu'elle soit valide, la comparaison de modèles doit être effectuée entre des modèles cohérents. Par exemple, la comparaison d'un modèle peuplement et d'un modèle individus-centré construits pour un même type de peuplement mais ajustés sur des jeux de données différents, ne permettra pas de choisir le bon niveau de détail. En effet, il ne sera pas possible de savoir si les différences observées sur un résultat de simulation proviennent de la différence des jeux de données d'ajustement ou du type de modèle. Cela signifie que l'on ne peut pas comparer directement des modèles préexistants mais qu'il est nécessaire de les adapter au problème considéré. Assurer la cohérence entre les modèles est l'une des questions importantes que nous aborderons dans la **deuxième partie** de ce mémoire.

### 1.4.2 Application aux peuplements mélangés chêne sessile – pin sylvestre

Cette thèse s'inscrit dans le cadre d'un projet scientifique du Cemagref de Nogent-sur-Vernisson, en partenariat avec l'Office National des Forêt, visant à construire un modèle de dynamique forestière des peuplements mélangés chêne sessile - pin sylvestre de la région Centre (Allain *et al.*, 2005). Nous avons donc mis en place un plan d'échantillonnage dans des peuplements mélangés chêne-pin de la forêt domaniale d'Orléans pour obtenir des données précises sur plusieurs parcelles forestières, à la fois en termes de croissance (prélèvement de carottes et étude rétrospective), de taille des arbres (diamètre et hauteur) et de répartition spatiale (position des arbres sur la placette). L'acquisition de ces données est présentée dans la **troisième partie**.

### 1.4.3 Construction d'une famille de modèles

Dans ce travail, afin de simplifier le problème, nous nous sommes restreints au phénomène de croissance en diamètre pour des arbres dépassant 7.5 cm de diamètre. De plus, nous avons choisi une approche dite "phénoménologique"<sup>2</sup> pour modéliser la croissance des peuplements mélangés. A partir des données de croissance, nous avons donc construit une famille de modèles incluant trois grands types : les modèles arbre indépendants des distances, les modèles arbre dépendants des distances et les modèles peuplement. Etant donné les contraintes relatives à l'utilisation des modèles arbre dépendants des distances, nous souhaitons évaluer la possibilité et l'intérêt de construire un modèle intégrant des informations sur le voisinage des arbres mais ne nécessitant pas une cartographie complète du peuplement. Nous avons donc construit un quatrième type de modèle, intermédiaire entre un MAID et un MADD, répondant à cet objectif. La construction de ces modèles et les connaissances qu'ils ont permis d'acquérir sur le mélange chêne-pin sont présentées dans la **quatrième partie**.

---

<sup>2</sup> Méthode visant à rendre compte de phénomènes observés (croissance, déformation de houppier) à partir de données d'expérimentation, de placettes permanentes ou de données d'inventaire. Cette approche intègre des concepts biologiques ou écologiques (compétition, fertilité) mais n'intègre pas explicitement les mécanismes et les processus élémentaires qui génèrent les phénomènes observés (cycles physico-chimiques par exemple). D'après Houllier (1991) et Dhôte (2008).

#### 1.4.4 Evaluation et comparaison des modèles pour prédire la croissance

Les modèles que nous avons construits peuvent être utilisés comme outils prédictifs. Pour évaluer et comparer l'efficacité des modèles dans ce domaine, nous avons mis en œuvre trois approches complémentaires. Nous avons utilisé un jeu de données indépendant des données de calibration, pour évaluer la capacité des modèles à prédire la croissance individuelle. Nous avons également évalué les modèles pour prédire la croissance du peuplement. Enfin, nous avons évalué et comparé les modèles sur des exemples d'application faisant intervenir des simulations à court et moyen termes. L'évaluation et la comparaison des modèles sont présentées dans la **cinquième partie**.

La **sixième partie** de ce mémoire sera consacrée à une discussion générale où nous reprendrons les principaux résultats et où nous ferons une synthèse sur leurs apports pour permettre de choisir un type de modèle en fonction du contexte d'utilisation. Nous y présenterons également les perspectives que ces travaux permettent d'envisager.

## 2. Comment assurer une cohérence entre des modèles ayant des niveaux de détail différents ?

### - Sommaire de la partie -

2.1	Du modèle peuplement au modèle arbre dépendant des distances.....	11
2.1.1	<i>Qu'appelons-nous modèle de dynamique forestière ?</i> .....	11
2.1.2	<i>Les modèles peuplement</i> .....	12
2.1.3	<i>Les modèles de distribution</i> .....	13
2.1.4	<i>Les modèles arbre indépendants des distances (MAID)</i> .....	15
2.1.5	<i>Les modèles arbre dépendants des distances (MADD)</i> .....	16
2.2	Aperçu des méthodes utilisées pour assurer une cohérence entre des modèles .....	18
2.2.1	<i>Ajuster les modèles sur les mêmes données</i> .....	18
2.2.2	<i>Expliciter un lien mathématique entre les modèles</i> .....	18
2.2.3	<i>Utiliser un modèle détaillé pour générer des données agrégées</i> .....	19
2.2.4	<i>Utiliser des outils de changement d'échelle</i> .....	19
2.3	Assurer une cohérence en ajustant les modèles sur les mêmes données .....	20
2.3.1	<i>Exemple d'un MAID</i> .....	20
2.3.2	<i>Passage à un modèle spatialement explicite par ajout d'un indice de compétition local</i> .....	20
2.4	Assurer une cohérence par agrégations ou simplification d'un modèle détaillé.....	21
2.4.1	<i>Présentation de deux méthodes permettant de simplifier des modèles</i> .....	21
2.4.2	<i>Agrégation d'un MAID</i> .....	23
2.4.3	<i>Simplifications d'un MADD</i> .....	25
2.5	Assurer une cohérence par des outils de changement d'échelle .....	26
2.5.1	<i>Génération de peuplements virtuels</i> .....	26
2.5.2	<i>Construction d'un modèle basé sur des distributions de voisinage</i> .....	27
2.6	Bilan.....	31

Dans cette thèse, nous souhaitons comparer un ensemble de modèles ayant des niveaux de détail différents. Lorsque l'on compare deux modèles quelconques, ces modèles ne diffèrent pas uniquement par leur niveau de détail mais aussi par tous les choix faits lors de leur construction : variables et processus pris en compte, forme des équations, jeu de données servant à estimer les paramètres, etc. La comparaison porte alors sur tous ces éléments. Dans notre cas, nous souhaitons si possible que la seule chose qui diffère entre nos modèles soit le niveau de détail afin d'évaluer son effet. Pour cela, nous souhaitons construire une série de modèles qui ont une structure commune. Nous utiliserons donc le terme de "cohérence" pour décrire le lien de structure qui existe entre nos modèles et nous appellerons "famille de modèles" un ensemble de modèles cohérents.

Dans cette partie nous allons dans un premier temps préciser les caractéristiques des grands types de modèles rencontrés en foresterie (sous-partie 2.1). Nous verrons ensuite comment on peut assurer une cohérence entre des modèles ayant des niveaux de détail différents (sous-parties 2.2 à 2.5). Enfin, nous terminerons par un bilan où nous présenterons les choix que nous avons faits dans cette thèse en matière de type de modèle et de méthodes utilisées pour assurer la cohérence (sous-partie 2.6).

## ***2.1 Du modèle peuplement au modèle arbre dépendant des distances***

Dans les paragraphes qui suivent, nous allons décrire quatre grands types de modèles utilisés en sciences forestières : les modèles peuplement, les modèles de distribution, les modèles arbre indépendants des distances et les modèles arbre dépendants des distances. Ces types de modèles ont déjà été largement décrits dans la littérature (voir par exemple Munro, 1974 ; Houllier *et al.*, 1991 ; Porté et Bartelink, 2002). L'objectif ici est de présenter la structure générale de ces différents modèles en nous appuyant sur des exemples quand cela est nécessaire. Nous préciserons également les caractéristiques des données permettant d'estimer les paramètres de ces modèles. Ces informations nous seront utiles par la suite lorsque nous étudierons comment créer une cohérence entre les modèles. Nous allons commencer cette sous-partie par préciser la définition des modèles auxquels nous nous intéressons.

### **2.1.1 Qu'appelons-nous modèle de dynamique forestière ?**

Un modèle est une représentation symbolique de certains aspects d'un objet ou d'un phénomène du monde réel (Pavé, 1994). L'objet que nous étudions ici est un peuplement forestier et le phénomène réel qui nous intéresse est la dynamique forestière, c'est-à-dire l'évolution dans le temps du peuplement forestier. Il existe plusieurs systèmes symboliques

qui permettent de représenter cet objet et ce phénomène (ex : équations, schéma ...). Dans cette thèse nous utiliserons le langage mathématique. Pour être plus précis, nous dirons donc qu'un modèle est un ensemble de relations mathématiques (issues d'une analyse statistique ou mécaniste) qui représente l'évolution dans le temps d'un ensemble de variables descriptives du peuplement forestier (Houllier *et al.*, 1991). De tels modèles sont aussi appelés modèles dynamiques ou systèmes dynamiques.

## 2.1.2 Les modèles peuplement

### *Définition*

Un modèle peuplement est un modèle qui décrit l'évolution dans le temps de grandeurs caractéristiques du peuplement comme le nombre total d'individus, la surface terrière totale ou le diamètre moyen. Par exemple, l'accroissement en surface terrière totale  $\Delta G$  peut être exprimé de la façon suivante :

$$\Delta G = f(G, N, Site)$$

avec  $G$  la surface terrière totale,  $N$  le nombre de tiges et  $Site$  une ou plusieurs variables caractéristiques du milieu. Dans un peuplement mélangé, soit le modèle ne distingue pas les espèces soit il y a une relation par espèce. L'influence des autres espèces sur la dynamique peut être incluse explicitement dans le modèle en utilisant par exemple des proportions d'espèces (Puettmann *et al.*, 1992). Les modèles peuplement sont des modèles intéressants car ils permettent de prédire l'évolution d'un peuplement avec des variables facilement accessibles et peu nombreuses. Historiquement, ce sont les premiers modèles construits pour les peuplements mélangés (voir paragraphe 1.2.2) et ils sont encore utilisés aujourd'hui pour modéliser la croissance et la dynamique de forêts mélangées (voir par exemple Payandeh et Wang, 1996 ; Jogiste, 1998 ; Gobakken et Naasset, 2002).

### *Données nécessaires pour estimer les paramètres*

Généralement, pour obtenir des données de croissance à l'échelle du peuplement, il faut suivre sur plusieurs années la croissance des arbres sur une surface donnée ou reconstituer la croissance de tous les arbres sur cette surface. Pour que ce modèle soit efficace, des mesures dans un grand nombre de peuplements sont nécessaires de façon à prendre en compte notamment les effets de la sylviculture et de la fertilité.

### 2.1.3 Les modèles de distribution

#### Définition

Un modèle de distribution est un modèle pour lequel le peuplement est décrit par une distribution des individus selon leur taille. Le modèle décrit alors l'évolution de cette distribution au cours du temps. Comme la distribution et le temps peuvent être soit continus soit discrets, quatre grands types de modèles de distribution peuvent être distingués (Picard, 1999). On s'intéressera ici aux modèles ayant à la fois une distribution et un temps discrets. Si de plus, les relations entre les variables prises en compte sont linéaires, on parle alors de modèles matriciels. Ces modèles sont couramment utilisés pour modéliser la dynamique forestière et correspondent bien aux types de données que l'on peut récolter dans un peuplement forestier. Pour illustrer à quoi correspond un modèle matriciel, nous allons prendre l'exemple d'un modèle matriciel simple : le modèle de "Usher" (voir Favrichon, 1995 pour une présentation approfondie de ce modèle). Dans ce modèle, les arbres sont regroupés dans des classes de taille, par exemple des classes de circonférence. L'état du peuplement à un temps  $t$  peut être représenté par un vecteur des effectifs d'arbres par classes de circonférence :

$$N_t = (n_{1,t}, n_{2,t}, \dots, n_{k,t})$$

Où  $n_{i,t}$  est le nombre d'arbres du peuplement appartenant à la classe  $i$  au temps  $t$ . En général on définit plusieurs paramètres pour passer d'un état  $t$  à un état  $t+1$  :

- la probabilité  $p_{i \rightarrow i+1}$  pour un individu de la classe  $i$  de passer à la classe adjacente  $i+1$  entre  $t$  et  $t+1$  ;
- la probabilité  $p_{i \rightarrow i}$  pour un individu de la classe  $i$  de rester dans le classe  $i$  entre  $t$  et  $t+1$  ;
- si on souhaite modéliser la mortalité, on définit une probabilité  $m_i$  de mourir pour un arbre de la classe  $i$ . On a alors :  $p_{i \rightarrow i} + p_{i \rightarrow i+1} + m_i = 1$  ;
- si de plus on souhaite modéliser le recrutement, on définit un paramètre  $f_i$  qui est le nombre d'individus recrutés dans la classe 1, entre  $t$  et  $t+1$ , issus d'un individu présent dans la classe  $i$ .

L'effectif de la classe  $i$  au temps  $t+1$  s'écrit alors (exceptée la première classe pour laquelle il faut ajouter le recrutement) :

$$n_{i,t+1} = p_{(i-1 \rightarrow i)} n_{i-1,t} + p_{(i \rightarrow i)} n_{i,t}$$

Nous pouvons également écrire cette relation sous forme matricielle. En effet, les probabilités  $p_{i \rightarrow j}$  appelées "probabilités de transition" sont représentées classiquement dans une matrice carrée dite matrice de transition ou de passage  $P$  de dimension  $(k,k)$  :

$$P = \begin{pmatrix} p_{1 \rightarrow 1} + f_1 & f_2 & f_3 & \dots & f_k \\ p_{1 \rightarrow 2} & p_{2 \rightarrow 2} & 0 & \dots & 0 \\ 0 & p_{2 \rightarrow 3} & p_{3 \rightarrow 3} & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & p_{k \rightarrow k} \end{pmatrix}$$

On a alors :

$$N_{t+1} = PN_t$$

Contrairement aux modèles peuplement, les modèles matriciels prennent en compte une partie de l'hétérogénéité d'un peuplement à travers la distribution des arbres en classes de taille. En revanche, ils sont basés sur des hypothèses fortes (Favrichon, 1995 page 62). Ainsi, les arbres sont supposés être indépendants les uns des autres. Malgré cela ils ont été largement utilisés pour modéliser la dynamique de peuplements mélangés (voir par exemple Vanclay, 1989 ; Bruciamacchie *et al.*, 1991 ; Volin et Buongiorno, 1996 ; Alder et Silva, 2000 ; Gourlet-Fleury et Houllier, 2000 ; Orois et Soalleiro, 2002 ; Hao *et al.*, 2005 ; Zhao *et al.*, 2005).

#### *Données nécessaires pour estimer les paramètres*

Dans ce type de modèles, il faut attacher une attention particulière au choix du pas de temps et au découpage en classes de taille (Favrichon, 1995 p. 76) car ils conditionnent l'intérêt et l'efficacité du modèle. Une fois ces choix réalisés, pour définir les probabilités de transition, on se base généralement sur des inventaires en classes de taille réalisés successivement sur un même peuplement. On peut ensuite connaître pour chaque classe l'effectif qui n'a pas changé de classe  $n_{i \rightarrow i, t+1}$  et celui qui est passé dans la classe suivante  $n_{i \rightarrow i+1, t+1}$ . On peut alors estimer la valeur des probabilités de transition  $p_{i \rightarrow i} = n_{i \rightarrow i, t+1} / n_{i, t}$  et  $p_{i \rightarrow i+1} = n_{i \rightarrow i+1, t+1} / n_{i, t}$  (Il s'agit d'estimateurs du maximum de vraisemblance Favrichon, 1995 p. 67).

## 2.1.4 Les modèles arbre indépendants des distances (MAID)

### Définition

Un modèle arbre indépendant des distances (MAID) ou modèle individus-centré non spatialisé, est un modèle qui décrit l'évolution dans le temps de chaque individu dans le peuplement (Porté et Bartelink, 2002 ; Grimm et Railsback, 2004). Soit  $y_i$  une grandeur caractéristique d'un arbre  $i$  dans le peuplement (son diamètre par exemple), son accroissement au cours du temps sera prédit en fonction de différentes variables, qui peuvent être soit au niveau de l'arbre soit propre à l'ensemble du peuplement. Plus précisément, ces variables peuvent correspondre à :

- des grandeurs descriptives de l'arbre, que l'on notera  $\tilde{y}_i$  et qui sont généralement associées au potentiel de croissance de l'individu dans le peuplement (exemples : diamètre, circonférence, surface terrière, hauteur totale) ;
- des variables qui caractérisent la compétition subie par l'arbre sujet. Elles peuvent être globales (exemples : surface terrière totale, densité du peuplement) et on les notera dans ce cas  $IC$ . Elles peuvent également dépendre des caractéristiques de l'individu (exemple : somme des surfaces terrières des arbres plus gros que l'arbre sujet) et dans ce cas, on les notera  $IC_i$  ;
- des variables qui caractérisent le milieu, que l'on notera  $Site$ , qui permettent de prendre en compte l'influence de facteurs qui agissent sur la croissance de l'ensemble des arbres du peuplement (exemples : indice de fertilité, type de station, pluviométrie moyenne).

A partir de ces éléments, nous pouvons écrire une équation générale pour un MAID :

$$\frac{dy_i}{dt} = f(\tilde{y}_i, IC_i, IC, Site)$$

Ces modèles intègrent donc un grand nombre d'informations et permettent d'améliorer la connaissance des facteurs qui influencent la croissance individuelle. En agrégeant les croissances individuelles, ces modèles permettent également d'obtenir des informations à des niveaux de détail intéressants pour la gestion (distribution des arbres en classes de taille, variables caractéristiques du peuplement). De nombreux modèles arbre indépendants des distances ont été construits pour des peuplements mélangés (voir par exemple Lejeune, 1996 ; Monserud et Sterba, 1996 ; Jogiste, 2000 ; Andreassen et Tomter, 2003 ; Trasobares *et al.*, 2004 ; Zhang *et al.*, 2004 ; Zhao *et al.*, 2004 ; Hein et Dhote, 2006).

Les relations de croissance sont parfois relativement complexes comme le montre cet exemple (Monserud et Sterba, 1996) :

$$\ln\left(\frac{\Delta g_i}{\Delta t}\right) = [a + b_1 \times \ln(d_i) + b_2 \times d_i^2 + b_3 \times \ln(\text{CrownRatio})] + [c_1 \times ic1_i + c_2 \times ic2_i] + s \times \text{Site}$$

où  $\Delta g_i$  est l'accroissement en surface terrière de l'arbre  $i$ ,  $d_i$  le diamètre de l'arbre  $i$ ,  $ic1_i$  est la somme de la surface terrière des arbres plus gros que l'arbre  $i$ ,  $ic2_i$  est un indice de compétition basé sur le houppier (Crown Competition Factor) et  $\text{Site}$  est un ensemble de variables descriptives de la station (topographie, type d'humus, type de sol). Ces modèles font cependant un certain nombre de simplifications (c'est le cas de l'exemple ci-dessus) :

- la plupart du temps, le modèle ne change pas d'une espèce à l'autre, ce sont les valeurs des paramètres qui changent ;
- bien souvent, toutes les espèces du peuplement sont considérées comme équivalentes pour la compétition.

#### *Données nécessaires pour estimer les paramètres*

Pour construire ce type de modèles, il est nécessaire de récolter des données individuelles sur la croissance des arbres mais aussi sur leurs caractéristiques dendrométriques. Ces données peuvent être récoltées sur l'ensemble des arbres d'un peuplement ou sur un échantillon.

### **2.1.5 Les modèles arbre dépendants des distances (MADD)**

#### *Définition*

Un modèle arbre dépendant des distances (MADD) ou modèle individus-centré spatialement explicite, est un modèle individus-centré pour lequel la position de chaque individu dans le peuplement est connue. Ces positions sont utilisées pour calculer des informations sur l'environnement local de chaque arbre de façon à prendre en compte la compétition exercée par le voisinage pour exploiter les ressources aériennes ou souterraines. Cependant, si l'on définit le voisinage d'un arbre comme l'ensemble des arbres en interaction avec lui (Goreaud, 2000), il est très difficile d'identifier quels arbres appartiennent au voisinage. C'est pourquoi, pour simplifier cette identification, on utilise des règles qui permettent de définir un espace relativement simple autour d'un arbre sujet et pour lequel on dira que tous les arbres appartenant à cet espace font partie du voisinage. Le voisinage peut par exemple être défini par un disque de rayon 10 m centré sur l'arbre sujet. On peut de cette façon donner un nombre illimité de définitions du voisinage. Ensuite, pour quantifier

l'influence des arbres du voisinage sur la croissance d'un arbre sujet, on peut utiliser des indices de compétition locaux qui sont intégrés dans le modèle (voir par exemple Biging et Dobbertin, 1992 ; Prévosto, 2005). Ces indices sont généralement calculés à partir des caractéristiques des arbres du voisinage préalablement défini et parfois de l'arbre sujet. Ainsi, la surface terrière des arbres présents dans un disque de rayon 10 m autour d'un arbre sujet est un exemple d'indice de compétition. En fait, ces indices peuvent traduire des interactions autres que la compétition selon la manière de les introduire dans les modèles. Ils peuvent par exemple avoir un effet positif sur la croissance et mettre ainsi en évidence une relation de coopération entre essences. Dans un peuplement mélangé, il est possible de distinguer les différentes espèces présentes dans le voisinage de façon à prendre en compte les différences d'interaction entre espèces.

Les MADD se sont fortement développés à partir des années 1990 (Porté et Bartelink, 2002) et ils ont permis de faire des progrès sur l'étude de la dynamique des peuplements mélangés (voir par exemple Lett *et al.*, 1999 ; Bartelink, 2000 ; Gourlet-Fleury et Houllier, 2000 ; Pretzsch *et al.*, 2002a ; Phillips *et al.*, 2003 ; Lee *et al.*, 2004 ; Uriarte *et al.*, 2004b ; Zhao *et al.*, 2006). Cependant, l'utilisation des indices de compétition présente des limites pour rendre compte de phénomènes aussi complexes que les interactions entre espèces (Goreaud, 2000 page 21). En particulier, la performance d'un indice dépend généralement fortement des données sur lesquelles il a été testé (Prévosto, 2005) et leur apport sur le pouvoir prédictif du modèle est parfois remis en question (Biging et Dobbertin, 1995).

#### *Données nécessaires pour estimer les paramètres*

Pour ce type de modèles, il est nécessaire d'avoir des informations sur le voisinage des arbres dont la croissance est connue ou sera mesurée. Pour cela, deux stratégies sont possibles :

- soit on définit a priori un indice de compétition (exemple : somme des surfaces terrières des arbres du voisinage). Dans ce cas il suffit de mesurer cet indice sur le terrain autour des arbres dont la croissance est connue ou sera mesurée ;
- soit on souhaite définir a posteriori l'indice qui donnera les meilleurs résultats. Dans ce cas, le plus souvent, il est nécessaire de faire une cartographie exhaustive du peuplement de façon à pouvoir faire varier la taille du voisinage et la nature de l'indice au moment de l'ajustement du modèle.

## ***2.2 Aperçu des méthodes utilisées pour assurer une cohérence entre des modèles***

Après avoir présenté les grands types de modèles que l'on peut rencontrer en foresterie, nous allons montrer qu'il existe plusieurs manières d'assurer une cohérence entre des modèles ayant des niveaux de détail différents (Picard et Franc, 2004). Nous allons les décrire brièvement puis nous illustrerons trois d'entre elles à travers des exemples qui nous serviront par la suite (paragraphe 2.3, 2.4 et 2.5).

### **2.2.1 Ajuster les modèles sur les mêmes données**

Le principe de cette méthode consiste à ajuster les différents types de modèles sur un même jeu de données de façon indépendante. Le fait d'utiliser un même jeu de données assure déjà une certaine cohérence. Cependant, pour aller plus loin, il est préférable que les modèles soient emboîtés : le modèle le moins détaillé doit pouvoir être obtenu en simplifiant le modèle le plus détaillé, par exemple en fixant l'un des paramètres du modèle détaillé (Pinheiro et Bates, 2000 ; Astrup *et al.*, 2008). On peut alors évaluer la pertinence d'intégrer des variables ou une complexité supplémentaires à l'aide de critères prenant en compte la qualité de l'ajustement et le nombre de paramètres (exemples : critères AIC et BIC). Cette méthode peut cependant s'avérer difficile à appliquer lorsque l'on souhaite ajuster des modèles qui n'ont pas les mêmes niveaux de description du peuplement. Par exemple, les données permettant d'ajuster un modèle individus-centré ne sont pas forcément adaptées pour ajuster un modèle peuplement. Nous illustrerons cette méthode à travers un exemple dans la sous-partie 2.3.

### **2.2.2 Expliciter un lien mathématique entre les modèles**

Une autre manière de créer une cohérence entre les modèles consiste à expliciter un lien mathématique entre deux modèles fonctionnant à des niveaux de détail différents. Ce lien peut être plus ou moins complexe. Le cas le plus simple consiste à effectuer la somme de la dynamique du modèle le plus détaillé, on parle alors d'agrégation. Ces opérations ne sont pas toujours réalisables, en particulier dès que la dynamique étudiée n'est pas linéaire. De plus, pour passer d'un modèle spatialisé à un modèle non spatialisé, il est nécessaire de faire des approximations sur la position des arbres. Il existe ainsi plusieurs méthodes permettant de simplifier un modèle spatialisé. Certaines sont détaillées dans Picard (1999) qui présente également des exemples issus de la littérature et des applications au cas de la forêt tropicale guyanaise. Dans la sous-partie 2.4, nous illustrerons par des exemples les cas suivants :

agrégation d'un modèle MAID en un modèle peuplement, agrégation d'un MAID en un modèle de distribution et simplification d'un MADD.

### **2.2.3 Utiliser un modèle détaillé pour générer des données agrégées**

Une autre méthode consiste à utiliser un modèle détaillé comme générateur de données pour caler ensuite un modèle plus grossier. Par exemple, à l'aide d'un MADD, on peut reconstruire l'évolution de la surface terrière totale au cours du temps en faisant des simulations. Ces données reconstruites peuvent alors servir pour ajuster un modèle peuplement. Cette approche est bien adaptée lorsque le modèle détaillé est complexe et que les méthodes précédentes ne peuvent pas être appliquées. Elle peut également être complémentaire à la méthode précédente lorsque cette dernière ne peut pas être menée jusqu'au bout (Verzelen *et al.*, 2006). Dans cette étude, nous avons choisi des formes de modèle relativement simples, c'est pourquoi nous ne développerons pas davantage cette méthode.

### **2.2.4 Utiliser des outils de changement d'échelle**

On peut également passer d'un type de modèle à un autre en développant des outils de changement d'échelle. Ainsi, les outils permettant de générer des peuplements virtuels peuvent être considérés comme des outils de changement d'échelle. En effet, à partir de données simples (distribution en classes de taille et densité du peuplement), ils permettent par exemple de générer une liste d'arbres qui peut être utilisée dans un modèle individus-centré. On peut également utiliser le terme de désagrégation pour désigner l'opération qui consiste à générer des données détaillées à partir de données plus simples (Picard, 2007 page 50). Le couplage entre l'outil de changement d'échelle et le modèle de croissance est alors une autre forme de modèle. Ces passerelles peuvent être particulièrement intéressantes quand les données du modèle de croissance sont contraignantes à acquérir, ce qui est le cas des données spatialisées. C'est pourquoi, dans la sous-partie 2.5, nous verrons deux exemples illustrant le couplage entre un MADD et des outils de changement d'échelle : l'utilisation d'un modèle de structure pour générer des peuplements virtuels réalistes et l'utilisation de distributions de voisinage pour reconstruire le voisinage des arbres.

## 2.3 Assurer une cohérence en ajustant les modèles sur les mêmes données

Il est possible d'assurer une cohérence entre deux modèles en les ajustant sur les mêmes données à condition que les deux modèles aient une structure commune. Pour illustrer cette méthode, nous allons prendre l'exemple d'un MAID et d'un MADD.

### 2.3.1 Exemple d'un MAID

En anticipant un peu sur la partie 4, nous allons considérer l'exemple du modèle suivant qui relie l'accroissement en rayon à la taille de l'arbre :

$$\Delta r_i = \alpha_{\text{espèce}_j} + \beta_{\text{espèce}_j} c_i \quad (2.1)$$

Où  $\Delta r_i$  est l'accroissement en rayon de l'arbre  $i$  appartenant à l'espèce  $j$  entre les dates  $t$  et  $t+\Delta t$ ,  $\alpha_{\text{espèce}_j}$  et  $\beta_{\text{espèce}_j}$  sont les paramètres du modèle pour l'espèce  $j$  et  $c_i$  est la circonférence de l'arbre  $i$  à la date  $t$ . Pour simplifier l'expression du modèle, nous n'avons pas fait intervenir de variables liées au milieu, comme la fertilité, ou liées au peuplement comme la densité (ces variables pourraient être introduites au niveau des paramètres  $\alpha$  et  $\beta$ ). Ce modèle est non spatialisé car il n'utilise pas la position des arbres du peuplement pour calculer des informations sur le voisinage.

### 2.3.2 Passage à un modèle spatialement explicite par ajout d'un indice de compétition local

Nous pouvons faire l'hypothèse qu'une partie de la variabilité non expliquée par le modèle non spatialisé (équation 2.1) est liée à des différences de compétition locale ou à des différences de composition du voisinage qui s'exercent au niveau de chaque arbre. Pour prendre en compte cette hypothèse, nous pouvons ajouter un indice de compétition local comme variable explicative supplémentaire. On obtient alors un modèle spatialement explicite de la forme suivante :

$$\Delta r_i = \alpha'_{\text{espèce}_j} + \beta'_{\text{espèce}_j} c_i + \gamma IC_i \quad (2.2)$$

En fixant  $\gamma = 0$  dans ce modèle, on retrouve alors la forme du MAID. Pour aller jusqu'au bout de la démarche, ces modèles doivent être ajustés en utilisant les mêmes arbres. En conséquence, les valeurs des indices de compétition locaux doivent être connues pour tous les individus. En pratique, il est alors nécessaire de mesurer les indices pour tous les arbres ou de disposer d'une cartographie du peuplement.

## 2.4 Assurer une cohérence par agrégations ou simplification d'un modèle détaillé

Les modèles auxquels nous nous intéressons sont des modèles mathématiques. La cohérence entre deux modèles peut également être assurée de façon mathématique. Dans cette sous-partie, nous allons commencer par présenter deux méthodes utilisées en écologie pour passer d'un modèle détaillé à un modèle moins détaillé. Nous verrons que l'application de ces méthodes passe par la formulation d'hypothèses simplificatrices. Ces méthodes seront illustrées dans les paragraphes qui suivent à travers différents exemples mais, comme nous le verrons également, elles ne sont pas toujours nécessaires pour assurer un lien mathématique entre deux modèles.

### 2.4.1 Présentation de deux méthodes permettant de simplifier des modèles

Nous avons choisi de présenter deux méthodes mathématiques permettant de simplifier des modèles : la méthode des moments et la méthode de l'approximation du champ moyen. Nous commençons par la méthode des moments car il s'agit de la méthode la plus générale.

#### 2.4.1.1 La méthode des moments

On considère ici que le système (le peuplement forestier) est décrit par l'ensemble des rayons individuels des arbres ( $r_i$ ) ainsi que par la position de chaque arbre dans le peuplement ( $x_i$ ). L'esprit de la méthode des moments, tel qu'elle est décrite par Dieckmann *et al.* (2000) et Law *et al.* (2000), est de projeter ce système dans un espace de plus petite dimension sans déformer sa dynamique. Pour cela, on recherche les variables suffisantes ainsi que leurs équations d'évolution. On appelle cette méthode la méthode des moments car les variables choisies pour décrire la dynamique du système dans l'espace réduit sont en général les moments d'ordre 1 et 2 des variables décrivant le système dans l'espace d'origine. Ainsi le moment d'ordre 1 des  $r_i$  sera le rayon moyen des arbres. Le moment d'ordre 2 des  $r_i$  sera la variance des  $r_i$ . Si on considère maintenant que la position des arbres est la réalisation d'un processus ponctuel alors le moment spatial d'ordre 1 de la position des arbres est la densité  $N$  du peuplement (Law et Dieckmann, 2000). Le moment spatial d'ordre 2 est la fonction de densité des paires de points.

Pour rendre cette méthode applicable, il faut parfois faire des hypothèses simplificatrices. Ainsi, si on fait l'hypothèse que le processus ponctuel est homogène et isotrope<sup>3</sup>, son moment d'ordre 2 est directement relié à la fonction  $K(r)$  de Ripley (Goreaud, 2000 page 49) :

$$\lambda K(r) = E(\text{Nombre d'arbres à une distance } \leq r \text{ d'un arbre sujet})$$

où  $\lambda$  est la densité (tiges/ha),  $E$  l'espérance et  $r$  le rayon du disque centré sur l'arbre sujet. La fonction  $K(r)$  permet alors de prendre en compte les relations de voisinage entre les individus et donc de tenir compte de la structure spatiale du peuplement sans avoir besoin de la position de tous les arbres.

La principale difficulté de cette méthode est de décrire l'évolution dans le temps des différents moments choisis pour décrire le système.

#### 2.4.1.2 L'approximation du champ moyen

Cette méthode est particulièrement adaptée pour simplifier la dépendance spatiale dans les modèles individus-centrés spatialement explicites. Nous considérons donc un MADD, c'est-à-dire un modèle qui fait intervenir des variables locales qui dépendent du voisinage de chaque arbre. De plus, on considère que la cartographie des arbres est une réalisation d'un processus ponctuel. La méthode de l'approximation du champ moyen consiste à simplifier l'expression du voisinage dans ce modèle en considérant que tous les arbres subissent la même influence de leur voisinage. Dans le modèle, on peut donc remplacer l'expression exacte de l'indice de compétition par son espérance. Cependant, on ne peut pas estimer l'espérance du processus ponctuel sur une seule réalisation. On fait donc l'hypothèse que le processus est ergodique, c'est-à-dire que l'espérance sur plusieurs réalisations est égale à la moyenne spatiale sur une réalisation (Goreaud, 2000 page 38). Dans le modèle, on peut donc remplacer l'espérance de l'indice de compétition par sa moyenne spatiale, c'est-à-dire la moyenne de l'indice de compétition sur l'ensemble des arbres du peuplement. Pour simplifier davantage l'expression de la moyenne spatiale, il faut parfois faire des hypothèses supplémentaires sur le processus ponctuel : indépendance entre les diamètres et la localisation des arbres, structure spatiale aléatoire. Cette méthode permet d'aboutir à un modèle arbre indépendant des distances qui prend plus ou moins en compte la structure spatiale en fonction des hypothèses simplificatrices retenues (voir par exemple Verzelen *et al.*, 2006 pour une application de cette méthode à un modèle forestier). L'approximation du champ moyen est la

---

<sup>3</sup> Voir Goreaud (2000), page 36 pour la définition d'un processus homogène et isotrope.

technique la plus simple utilisée en écologie pour simplifier les modèles spatialisés mais, comme nous venons de le voir, elle repose sur des hypothèses fortes.

## 2.4.2 Agrégation d'un MAID

La cohérence mathématique entre deux modèles peut être assurée en agrégeant plus ou moins l'un des modèles. Nous allons voir successivement l'agrégation d'un MAID en un modèle peuplement puis l'agrégation d'un MAID en un modèle matriciel.

### 2.4.2.1 Du MAID vers un modèle peuplement

Nous considérons le MAID suivant :  $\Delta r_i = \alpha_{\text{espèce}_j} + \beta_{\text{espèce}_j} c_i$  (équation 2.1), dans lequel le système (le peuplement forestier) est représenté par l'ensemble des rayons individuels des arbres ( $r_i$ ) ainsi que par la position de chaque arbre dans le peuplement ( $x_i$ ). Le nombre d'arbre d'une espèce  $j$  dans le peuplement est noté  $N_{\text{espèce}_j}$  et la surface du peuplement est notée  $A$ . Nous souhaitons agréger ce modèle en un modèle peuplement. Pour cela, nous allons appliquer la méthode des moments (voir paragraphe 2.4.1.1). Nous allons donc définir les variables qui vont servir à décrire le système dans un espace réduit. Nous nous intéressons tout d'abord aux moments d'ordre 1. Le moment d'ordre 1 des  $r_i$  est la moyenne des  $r_i$  soit :

$$\bar{r}_{\text{espèce}_j} = \frac{1}{N_{\text{espèce}_j}} \sum_{i=1}^{N_{\text{espèce}_j}} r_i$$

De même, le moment d'ordre 1 du processus ponctuel est la densité moyenne soit  $N_{\text{espèce}_j} / A$  que nous pouvons remplacer par  $N$  pour simplifier. Nous avons donc choisi de décrire le système uniquement par deux moments. Il nous faut maintenant décrire l'évolution de ces variables au cours du temps. D'après l'équation du modèle et en considérant qu'il n'y a pas de mortalité nous pouvons écrire (voir annexe A) :

$$\begin{aligned} \Delta \bar{r}_{\text{espèce}_j} &= \alpha + Cte(\beta) \bar{r}_{\text{espèce}_j} \\ \Delta N_{\text{espèce}_j} &= 0 \end{aligned}$$

où  $Cte(\beta)$  est une constante exprimée en fonction du paramètre  $\beta$  du MAID. Cependant, les forestiers utilisent en général le volume ou la surface terrière pour rendre compte de la production d'un peuplement. Ici, nous allons choisir la surface terrière, notée  $G_{\text{espèce}_j}$ , car le volume nécessite d'avoir des informations sur la hauteur des arbres. On peut alors montrer sans faire d'hypothèses supplémentaires (annexe A) que l'évolution de la surface terrière s'écrit de la façon suivante :

$$\Delta G_{\text{espèce}_j} = Cte(\alpha) N_{\text{espèce}_j} + Cte(\alpha, \beta) N_{\text{espèce}_j} \bar{r}_{\text{espèce}_j} + Cte(\beta) G_{\text{espèce}_j} \quad (2.3)$$

où  $Cte(\alpha)$  et  $Cte(\alpha, \beta)$  sont des constantes exprimées en fonction des paramètres du modèle MAID (équation 2.1). Le modèle peuplement obtenu par agrégation du MAID permet donc de décrire l'évolution du système avec un nombre limité de variables :  $N_{\text{espèce}_j}$ ,  $G_{\text{espèce}_j}$  et  $\bar{r}_{\text{espèce}_j}$ .

#### 2.4.2.2 Du MAID vers un modèle matriciel

Cette fois, nous considérons toujours le MAID :  $\Delta r_i = \alpha_{\text{espèce}_j} + \beta_{\text{espèce}_j} c_i$ , mais nous souhaitons l'agréger en un modèle matriciel (pour simplifier nous ne ferons pas intervenir l'espèce dans les notations). Pour cela, nous cherchons à calculer les valeurs des paramètres, c'est-à-dire les probabilités de transition qui assurent que les deux modèles décrivent bien la même croissance. Nous allons d'abord discrétiser la variable circonférence en un certain nombre de classes. Ensuite, nous allons transformer l'accroissement en rayon, qui est continu, en une probabilité de passage de la classe  $i$  à la classe  $i+1$ . Picard (1999 pages 301 et 302) relève dans la littérature différentes façons de faire cette transformation. Nous allons utiliser ici une approche dite déterministe en considérant tous les individus contenus dans les classes de circonférence. Nous considérons que les deux modèles sont cohérents si la somme des accroissements dans chaque classe de circonférence du modèle matriciel est identique à celle obtenue avec le MAID. En prenant l'équation de croissance du MAID, on peut calculer la somme des accroissements de chaque classe de la façon suivante :

$$\sum_{c_j \in \text{classe } i} \frac{\Delta r_j}{\Delta t} = \sum_{c_j \in \text{classe } i} \alpha + \beta c_j = \alpha n_i + \beta n_i c_{m,i}$$

où  $c_{m,i}$  est la circonférence moyenne pour la classe  $i$ <sup>4</sup>. Puis nous pouvons calculer cette somme en utilisant le modèle matriciel décrit au paragraphe 2.1.3. Avec ce modèle, un arbre a une probabilité  $p_{i \rightarrow i}$  de rester dans la même classe (dans ce cas sa croissance peut être considérée comme nulle), et une probabilité  $p_{i \rightarrow i+1}$  de passer dans la classe  $(i+1)$ , dans ce cas on peut considérer que sa croissance est proportionnelle à  $r_{m,i+1} - r_{m,i}$ , où  $r_{m,i}$  et  $r_{m,i+1}$  sont les rayons moyens des classes  $i$  et  $i+1$ . Finalement, la somme des accroissements de la classe  $i$  est donnée par les arbres qui changent de classe, soit :

$$\sum_{c_j \in \text{classe } i} \frac{\Delta r_j}{\Delta t} = n_i p_{i \rightarrow i+1} (r_{m,i+1} - r_{m,i})$$

Pour que les deux modèles soient cohérents, les probabilités de transition doivent donc vérifier l'équation suivante :

---

<sup>4</sup> Dans le cas particulier où, à l'intérieur de la classe  $i$ , la distribution des circonférences est uniforme, et si on note  $l_1, l_2, \dots, l_{k+1}$  les bornes définissant les classes de circonférence, alors on peut montrer que :  $c_{m,i} = (l_{i+1} + l_i) / 2$

$$p_{i \rightarrow i+1} = \frac{\alpha + \beta c_{m,j}}{(r_{m,i+1} - r_{m,i})} = 2\pi \frac{\alpha + \beta c_{m,i}}{(c_{m,i+1} - c_{m,i})}$$

Le modèle matriciel obtenu par agrégation du MAID permet donc de décrire l'évolution du système avec un nombre moins important de variables : les effectifs par classes de taille.

### 2.4.3 Simplifications d'un MADD

Les MADD sont des modèles qui nécessitent beaucoup de données. Il est donc tentant de vouloir les simplifier. Cependant, la simplification de la dépendance spatiale d'un modèle n'est pas toujours possible et elle nécessite des approximations. A travers des exemples simples, nous allons présenter le passage d'un MADD à un MAID puis le passage d'un MADD à un modèle peuplement.

#### 2.4.3.1 Du MADD vers un MAID

Nous considérons que le peuplement forestier est décrit par l'ensemble des rayons individuels des arbres ( $r_i$ ) ainsi que par la position de chaque arbre dans le peuplement ( $x_i$ ). Nous allons considérer maintenant que la croissance des arbres est décrite par un MADD :  $\Delta r_i = \alpha'_{\text{espèce}_j} + \beta'_{\text{espèce}_j} c_i + \gamma IC_i$  (équation 2.2). A partir de ce modèle, nous souhaitons obtenir un modèle individus-centré non spatialisé (MAID). Pour cela, nous allons utiliser l'approximation du champ moyen (paragraphe 2.4.1.2) c'est-à-dire considérer que tous les arbres subissent une compétition locale identique. En remplaçant  $IC_i$  par sa moyenne spatiale, on aura alors :

$$\Delta r_i = \alpha'_{\text{espèce}_j} + \beta'_{\text{espèce}_j} c_i + \gamma \langle IC_i \rangle \quad (2.4)$$

Il est possible de connaître  $\langle IC_i \rangle$  en faisant des mesures spécifiques dans les peuplements mais nous ne pouvons pas écrire le système complet car nous ne connaissons pas l'évolution de  $\langle IC_i \rangle$  au cours du temps. Pour cela, il faut préciser la nature de l'indice de compétition. Nous verrons dans la partie 4 que, dans certains cas, il est possible d'aller plus loin lorsque l'on connaît l'expression de cet indice.

#### 2.4.3.2 Du modèle MADD vers un modèle peuplement

Nous considérons toujours que le peuplement forestier est décrit par l'ensemble des rayons individuels ( $r_i$ ) ainsi que par la position de chaque arbre dans le peuplement ( $x_i$ ). La croissance des arbres est décrite par un MADD :  $\Delta r_i = \alpha'_{\text{espèce}_j} + \beta'_{\text{espèce}_j} c_i + \gamma IC_i$  (équation 2.2). Comme dans l'exemple du paragraphe 2.4.2.1, nous pouvons appliquer la méthode des moments mais cela implique d'explicitier la somme des indices de compétition. Dans le cas

particulier où l'indice est le nombre de voisins, il est possible d'exprimer cette somme en fonction de  $K(r)$  (moment d'ordre 2 du processus ponctuel). Néanmoins, expliciter cette somme n'est pas toujours possible. Nous pouvons donc choisir une autre approche en commençant par faire l'approximation du champ moyen et on retrouve alors l'équation 2.4. Comme  $\langle IC_i \rangle$  est une constante propre au peuplement, il suffit ensuite de poser  $\alpha = \alpha' + \lambda \langle IC_i \rangle$  et  $\beta = \beta'$  dans l'équation 2.3 obtenue en agrégeant le MAID (paragraphe 2.4.2.1) pour obtenir le modèle peuplement. Malheureusement, de la même façon que pour le passage au modèle individus-centré non spatialisé (paragraphe 2.4.3.1), il nous faut connaître l'évolution de  $\langle IC_i \rangle$ . A nouveau, nous verrons dans la partie 4 qu'il est parfois possible d'aller plus loin lorsque l'on connaît l'expression de l'indice de compétition.

## 2.5 Assurer une cohérence par des outils de changement d'échelle

### 2.5.1 Génération de peuplements virtuels

Les modèles MADD que nous avons présentés plus haut nécessitent de connaître la position de chaque arbre dans le peuplement. En conséquence, comme nous l'avons déjà évoqué en introduction, ces modèles posent des difficultés d'utilisation car la cartographie d'un peuplement est une information coûteuse. Pour utiliser les modèles spatialisés, on a alors recours à des peuplements virtuels (Pretzsch, 1997 ; Kokkila *et al.*, 2002 ; Pommerening, 2006) générés à l'aide de processus ponctuels (Tomppo, 1986 ; Stoyan et Penttinen, 2000 ; Comas et Mateu, 2007). Néanmoins, peu de travaux de ce type se sont attachés à ajuster des modèles de génération de peuplement sur des données réelles ou à évaluer le réalisme des structures générées (Goreaud *et al.*, 2004 ; Ngo Bieng, 2007). En effet, si la structure spatiale a un rôle dans le fonctionnement de l'écosystème, il est nécessaire que le peuplement simulé ait des propriétés spatiales proches d'un peuplement réel. En particulier, dans le cas de peuplements mélangés, il faut non seulement respecter la répartition spatiale de chaque espèce mais également respecter les interactions spatiales entre espèces. Ainsi, Ngo Bieng (2007) a construit un modèle de structure basé sur les processus ponctuels permettant de simuler des peuplements mélangés chêne sessile-pin sylvestre (à partir d'une liste d'arbres ou de données sur les distributions en taille des arbres).

Ce modèle de structure peut être considéré comme un outil de changement d'échelle (ou changement de niveau de description) permettant le passage de données non spatialisées à des données spatialisées. Le couplage entre le modèle de structure et le MADD peut alors être

considéré comme une forme de modèle intermédiaire entre un modèle individus-centré non spatialisé et un modèle de croissance spatialement explicite.

### 2.5.2 Construction d'un modèle basé sur des distributions de voisinage

La cartographie d'un peuplement est en fait une information très riche et qui est bien souvent sous-utilisée dans les modèles spatialement explicites décrivant les processus écologiques. On peut donc poser la question : ne peut-on pas se contenter d'une information moins coûteuse ? Ainsi, les modèles obtenus dans le paragraphe 2.4.3 en simplifiant un modèle spatialement explicite intègrent une partie de l'information spatiale qui ne nécessite pas une cartographie complète du peuplement. Ces modèles peuvent donc être considérés comme intermédiaires entre les modèles spatialisés et les modèles non spatialisés. Cependant, ils intègrent une information spatiale moyenne et ne prennent donc pas en compte la variabilité de cette information dans le peuplement. Or, cette variabilité peut avoir un rôle important dans la survie d'une espèce et sur la coexistence des espèces (Vieilledent, 2009). Phillips *et al.* (2004) proposent de modéliser des distributions d'indices de compétition pour pouvoir utiliser un modèle de croissance spatialisé. Chaque arbre a alors un indice de compétition attribué de manière aléatoire mais respectant des distributions normales dont les paramètres dépendent de l'espèce et de la taille du sujet. Cependant, pour que cette approche reste pertinente (pour tenir compte en particulier du processus de croissance), Phillips *et al.* (2004) modélisent également l'évolution dans le temps des paramètres de ces distributions.

Dans le même esprit, nous proposons de construire un modèle qui utilise des distributions de nombre de voisins dans un disque de rayon  $R$  pour reconstruire l'environnement local des arbres du peuplement. On appellera ces distributions des "distributions de voisinage". Pour cela, nous nous plaçons dans le cas général d'un modèle arbre dépendant des distances utilisant des indices de compétition locaux pour prédire la croissance :

$$\frac{dy_i}{dt} = f(\tilde{y}_i, IC_i, Site)$$

où  $IC_i$  correspond à un ou plusieurs indices de compétition locaux. Pour calculer les indices de compétition locaux, il est nécessaire de connaître explicitement le voisinage de chaque arbre, et généralement on utilise pour cela une cartographie du peuplement. Le modèle que nous proposons ici utilise des distributions de voisinage à la place de cette cartographie pour reconstruire le voisinage des arbres. Les voisinages reconstruits permettent ensuite le calcul des indices de compétition locaux qui sont utilisés dans le modèle de croissance. L'avantage de cette méthode est qu'elle ne nécessite pas de modéliser l'évolution de la distribution au

cours du temps. En effet, comme nous le verrons, chaque arbre du voisinage est lui-même un arbre du peuplement soumis au processus de croissance. En revanche, les individus n'ont pas de position explicite, le modèle ainsi construit est intermédiaire entre un modèle arbre indépendant des distances et un modèle arbre dépendant des distances.

### 2.5.2.1 Distributions de voisinage : définition et intérêt

Si nous comptons le nombre de voisins dans un disque de rayon  $R$  pour plusieurs arbres d'un peuplement, nous obtenons en général plusieurs valeurs. Pour rendre compte de cette variabilité, on peut utiliser une distribution du nombre de voisins (comme on utilise une distribution de la circonférence des arbres dans les modèles par classes de taille). Dans un peuplement mélangé, plusieurs distributions doivent être définies en fonction de l'espèce de l'arbre sujet et de l'espèce des arbres voisins (figure 2.1). Nous noterons ces distributions  $D_{ss,ns}$  où  $ss$  est l'espèce de l'arbre sujet et  $ns$  l'espèce des arbres voisins.  $D_{ss,ns}(x)$  est alors la probabilité qu'un arbre de l'espèce  $ss$  ait  $x$  arbres de l'espèce  $ns$  dans son voisinage.

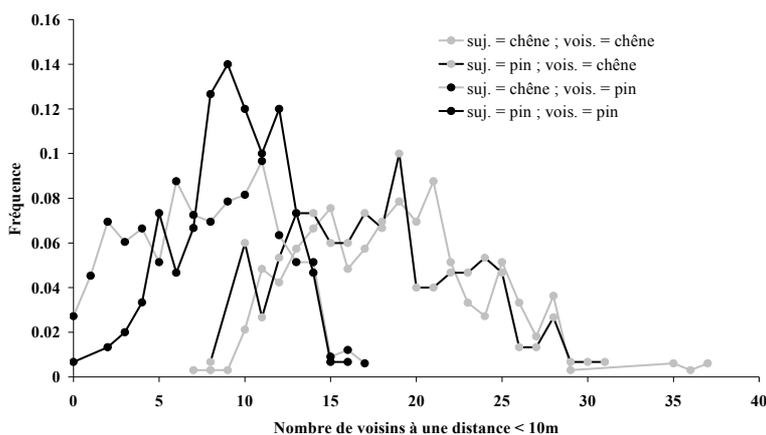


Figure 2.1 : Exemple de 4 distributions de voisinage dans un peuplement mélangé chêne sessile – pin sylvestre (dispositif D49). suj. = espèce de l'arbre sujet ; vois. = espèce des arbres voisins. Les distributions ont été obtenues à partir de l'ensemble des arbres du dispositif.

Il existe un lien direct entre certaines caractéristiques de ces distributions et certaines statistiques utilisées pour caractériser la structure spatiale d'un peuplement. En effet, dans un peuplement mélangé, la structure spatiale interspécifique entre deux espèces peut être caractérisée à l'aide de la fonction Intertype (Lotwick et Silverman, 1982) définie de la façon suivante :

$$\lambda_{ns} K_{ss,ns}(r) = E(\text{Nombre d'arbres de l'espèce } ns \text{ à une distance } \leq r \text{ d'un arbre de l'espèce } ss)$$

où  $\lambda_{ns}$  est la densité de l'espèce  $ns$ ,  $E$  est l'espérance et  $r$  le rayon du disque centré sur l'arbre sujet. Si on note  $\mu_{ss,ns}$  l'espérance de la distribution  $D_{ss,ns}$  alors on a :

$$\lambda_{ns} K_{ss,ns}(R) = \mu_{ss,ns}$$

On peut montrer que  $K_{ss,ns}(r) = K_{ns,ss}(r)$  (Lotwick et Silverman, 1982). Ainsi, si  $ss \neq ns$ ,  $\mu_{ss,ns}/\lambda_{ns}$  caractérise les interactions spatiales entre les deux espèces à une distance inférieure à  $R$ . Si  $ss = ns$ , alors  $K_{ss,ss}(r)$  est la fonction de Ripley (Ripley, 1977) et  $\mu_{ss,ss}/\lambda_{ss}$  caractérise la structure spatiale d'une espèce à une distance inférieure à  $R$ .

Les distributions de voisinage permettent également de rendre compte de la variabilité des voisinages. La figure 2.2 représente trois exemples de distributions du nombre de voisins dans trois peuplements virtuels de même densité caractérisés par des distributions spatiales différentes (les structures spatiales ont été caractérisées à l'aide de la fonction de Ripley, données non présentées). Cette figure montre que les trois peuplements ont, comme on pouvait s'y attendre, des valeurs moyennes différentes mais qu'ils ont également des variabilités très différentes. Le peuplement caractérisé par une structure régulière (semis de points en répulsion) présente une très faible variabilité alors que le peuplement caractérisé par une structure agrégée présente une très forte variabilité. Dans le premier cas, la moyenne de la distribution peut permettre d'obtenir un bon estimateur de la compétition locale alors que, dans le second cas, la moyenne n'a pas beaucoup de sens. Les distributions de voisinage permettent donc de rendre compte des différences de structures spatiales entre peuplements mais également de la variabilité des voisinages dans le peuplement.

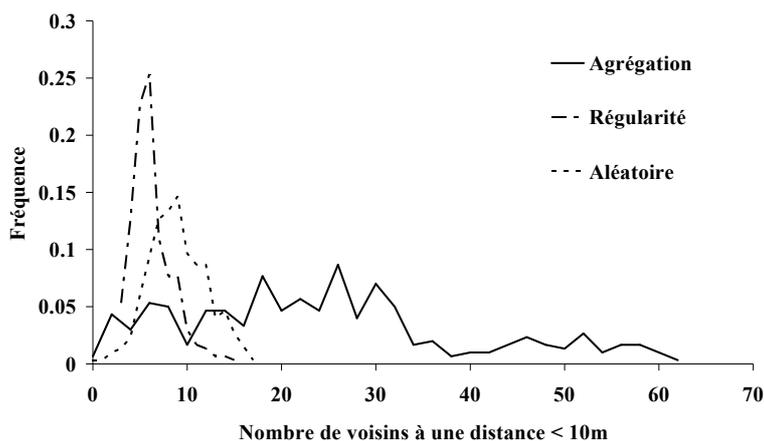


Figure 2.2 : Distribution des fréquences du nombre de voisins dans un disque de rayon 10 m pour trois peuplements caractérisés par des structures spatiales différentes.

### 2.5.2.2 Reconstruction du voisinage

Si l'on suppose que l'on connaît les distributions de voisinage ainsi que la liste des arbres du peuplement, il est alors possible de reconstruire les voisinages dans un peuplement mélangé, c'est-à-dire d'attribuer à chaque arbre une liste de voisins compatible avec les distributions de voisinage. La première étape pour reconstruire les voisinages est d'affecter de manière aléatoire, à chaque arbre, un nombre de voisins conforme aux distributions de voisinage. Dans une deuxième étape, les arbres voisins sont tirés dans la liste des arbres du peuplement (de manière aléatoire ou non selon les objectifs et les données disponibles). De cette façon, chaque arbre possède une liste de voisins de chaque espèce ce qui permet de calculer les indices de compétition locaux utilisés dans le modèle de croissance et de pouvoir faire évoluer le voisinage comme indiqué ci-après.

### 2.5.2.3 Mise à jour du voisinage au cours du temps

Le voisinage d'un arbre en forêt évolue au cours du temps à cause des processus de régénération, de croissance, de mortalité mais aussi par le biais des interventions humaines. Dans ce travail, nous ne prenons pas en compte les processus de mortalité et de régénération. En conséquence, le voisinage est uniquement influencé par le processus de croissance et par les éclaircies. Comme nous l'avons dit dans le paragraphe précédent, chaque arbre possède une liste de voisins appartenant eux-mêmes au peuplement. Etant donné que la croissance d'un arbre entre deux dates dépend de l'état du voisinage au début de la période de croissance, nous pouvons mettre à jour les voisinages après le processus de croissance lorsque tous les arbres ont poussé et que leurs nouvelles caractéristiques sont connues. De la même façon, les voisinages sont mis à jour après une éclaircie. Nous pouvons ainsi assurer que les voisinages reconstruits changent au cours du temps en fonction de la croissance et des éclaircies.

Dans ce paragraphe, nous avons présenté le principe général du modèle utilisant des distributions de voisinage. Dans la partie 4, nous verrons une application de ce modèle aux peuplements mélangés chêne sessile – pin sylvestre.

## 2.6 *Bilan*

Cette partie nous a permis de présenter l'étendue des types de modèles de croissance que l'on peut rencontrer en foresterie en fonction de leur niveau de détail : les modèles peuplement, les modèles de distribution, les modèles arbre indépendants des distances et les modèles arbre dépendants des distances. En réalité, il existe un continuum entre ces différents types de modèles. Il n'y a pas de limite formelle au sens mathématique. Par exemple, on peut considérer un modèle peuplement comme un modèle de distribution avec une seule classe, et un modèle arbre comme un modèle de distribution avec autant de classes que d'arbres. Pour des raisons matérielles, nous avons dû restreindre le nombre de modèles étudiés dans cette thèse. Nous avons choisi de construire un modèle individus-centré spatialement explicite car ce type de modèle est réputé bien adapté aux peuplements hétérogènes. Cependant, dans cette thèse nous souhaitons évaluer si ce type de modèle était vraiment justifié. Nous avons donc choisi de construire également un modèle individus-centré non spatialisé. De plus, nous souhaitons évaluer la possibilité et l'intérêt de construire un modèle intermédiaire entre un MAID et un MADD. Nous avons donc construit un modèle utilisant des distributions de voisinage plutôt qu'une cartographie du peuplement. Enfin, nous avons choisi d'explorer l'intérêt des modèles peuplement plutôt que les modèles de distribution. Nous avons considéré que le modèle peuplement permettait de franchir une marche supplémentaire dans la simplification des variables utilisées pour prédire la croissance d'un peuplement mélangé. Finalement, nous avons choisi de construire des modèles appartenant aux types suivants :

- les modèles arbre dépendants des distances (MADD) ;
- les modèles arbre avec distributions de voisinage (MADV) ;
- les modèles arbre indépendants des distances (MAID) ;
- les modèles peuplement ;

Dans cette partie, nous avons également montré comment on pouvait assurer une cohérence entre des modèles ayant des niveaux de détail différents. Nous avons illustré trois méthodes avec des exemples : ajuster les modèles sur les mêmes données, expliciter un lien mathématique entre les modèles et utiliser des outils de changement d'échelle. Dans la partie 4, nous verrons comment nous avons appliqué ces méthodes pour assurer un lien entre les modèles que nous avons construits.

### 3. Données récoltées dans les forêts mélangées chêne sessile - pin sylvestre

#### - Sommaire de la partie -

3.1	Les forêts de chêne sessile et pin sylvestre : un bon exemple de peuplement mélangé .....	33
3.1.1	<i>Les mélanges chênes-pins</i> .....	33
3.1.2	<i>Le mélange chêne sessile - pin sylvestre en forêt domaniale d'Orléans</i> .....	34
3.2	Modèles existants pour la croissance du chêne sessile et du pin sylvestre.....	34
3.2.1	<i>La croissance dans le modèle Fagacées</i> .....	35
3.2.2	<i>La croissance dans le modèle Sylvestris</i> .....	35
3.2.3	<i>Enseignements issus de ces modèles</i> .....	35
3.3	Choix des arbres et mesures dendrométriques .....	36
3.3.1	<i>Choix et caractéristiques des dispositifs</i> .....	36
3.3.2	<i>Choix des arbres à l'intérieur des dispositifs</i> .....	40
3.3.3	<i>Mesure de la hauteur totale des arbres</i> .....	43
3.4	Mesure de la croissance individuelle .....	44
3.4.1	<i>Méthode utilisée</i> .....	44
3.4.2	<i>Informations sur les âges</i> .....	45
3.4.3	<i>Estimations d'indices de fertilité par essence</i> .....	46
3.5	Partage des données pour ajuster et évaluer les modèles .....	48
3.6	Bilan.....	49

Dans cette thèse, nous souhaitons construire et comparer des modèles de croissance ayant des niveaux de détail différents. Nous avons choisi d'appliquer notre problématique aux peuplements mélangés chêne sessile – pin sylvestre de la Forêt domaniale d'Orléans. L'objectif de cette partie est de présenter les peuplements que nous avons étudiés ainsi que la manière dont nous avons récolté les données de croissance. Nous souhaitons également présenter des informations qui peuvent nous permettre de mieux comprendre le phénomène étudié ou de mieux interpréter les résultats que nous obtiendrons par la suite.

Dans la première sous-partie, nous présenterons le mélange chêne sessile – pin sylvestre (sous-partie 3.1). Ensuite, nous présenterons les travaux de modélisation concernant le chêne et le pin en peuplements purs afin de dégager des informations nous permettant d'aborder l'étude du mélange chêne-pin (sous-partie 3.2). Nous présenterons ensuite la manière dont nous avons acquis les données de croissance (sous-parties 3.3 et 3.4) et comment nous les avons partagées pour assurer la calibration et la validation des modèles (sous-partie 3.5). Nous terminerons par un bilan présentant les principaux choix faits dans cette partie (sous-partie 3.6).

### ***3.1 Les forêts de chêne sessile et pin sylvestre : un bon exemple de peuplement mélangé***

#### **3.1.1 Les mélanges chênes-pins**

Les forêts composées de chênes (genre *Quercus*) et de pins (genre *Pinus*) sont présentes dans de nombreux endroits dans le monde. En France, les mélanges chênes-pins représentent 420 000 ha pour 60 millions de m<sup>3</sup> faisant intervenir principalement trois espèces du genre *pinus* (Morneau *et al.*, 2008). Au centre de la France, ces mélanges résultent principalement des reboisements en pin sylvestre (*Pinus sylvestris*) effectués dans ces régions depuis le 19<sup>ème</sup> siècle. Cette espèce se trouve alors associée au chêne sessile (*Quercus petraea*) et au chêne pédonculé (*Quercus robur*). A l'ouest, le pin sylvestre est remplacé par le pin maritime (*Pinus pinaster*). Enfin, lorsque l'on se rapproche de la zone méditerranéenne, on trouve le pin d'alep (*Pinus halepensis*) qui s'associe principalement avec le chêne vert (*Quercus ilex*). Ces mélanges nous paraissent intéressants à étudier pour deux raisons principales. D'une part, ils sont très répandus et d'autre part, d'un point de vue écologique, ils associent des espèces feuillues tolérantes à l'ombrage à croissance lente appartenant au genre *Quercus* à des espèces résineuses intolérantes à l'ombrage et à croissance rapide appartenant au genre *Pinus* (Guirado *et al.*, 2008).

### 3.1.2 Le mélange chêne sessile - pin sylvestre en forêt domaniale d'Orléans

Notre étude porte sur les peuplements mélangés chêne sessile - pin sylvestre de la forêt domaniale d'Orléans (47°51'N, 2°25'E), un massif de 35 000 ha situé dans la région Centre de la France. Dans ce massif, le pin sylvestre a été introduit de façon massive à partir du 19<sup>ème</sup> siècle. L'objectif principal de cette introduction était de constituer une ambiance forestière dans des zones de landes résultant de la surexploitation en bois et du pâturage des animaux. Aujourd'hui, le mélange chêne sessile – pin sylvestre est très présent : par exemple sur le massif de Lorris (7774 ha), 25% des peuplements ne présentent pas d'essence prépondérante (ONF, 2004). De ce fait, ce mélange a acquis une valeur patrimoniale ; il est maintenant représentatif des paysages de nombreux massifs des régions Centre et Ile-de-France. De plus, il est souvent bien adapté aux stations très contraignantes fréquentes dans la région : terrains sableux avec une hydromorphie forte plus ou moins proche de la surface. Depuis plusieurs années, les gestionnaires forestiers se posent des questions sur la conduite du chêne sessile et du pin sylvestre en mélange. Lors du dernier aménagement du massif de Lorris, cela s'est traduit par la création d'une série irrégulière expérimentale de 340 ha visant à rechercher des modes de gestion adaptés au mélange chêne-pin sur stations pauvres. En 2003, le Cemagref et l'ONF ont construit un projet en partenariat sur l'étude du mélange chêne sessile – pin sylvestre. Ce projet comprend des travaux sur la structure des peuplements, sur la croissance et sur la régénération. Le travail sur la croissance que nous présentons dans cette thèse a débuté en 2006.

### 3.2 Modèles existants pour la croissance du chêne sessile et du pin sylvestre

Actuellement, il n'existe pas de modèles de croissance adaptés aux peuplements mélangés chêne-pin pour la région Centre. En revanche, il existe un modèle de croissance pour les peuplements purs de chêne appelé "Fagacées" (Dhôte *et al.*, 2000) et un modèle de croissance pour les peuplements purs de pin sylvestre appelé "Sylvestris" (Perot *et al.*, 2007). Ces deux modèles sont des modèles arbre indépendants des distances. Dans cette sous-partie, nous allons présenter comment la croissance est modélisée dans ces deux modèles et voir quels enseignements ils peuvent apporter pour aborder l'étude de la croissance des peuplements mélangés chêne-pin. Nous ne détaillerons pas toutes les équations qui sont décrites dans les articles cités ci-dessus.

### 3.2.1 La croissance dans le modèle Fagacées

Dans le modèle Fagacées, la croissance individuelle est modélisée de la façon suivante :

$$\frac{\Delta g_i}{\Delta t}(c_i) = \max[0, \gamma(c_i - \sigma)]$$

où  $\Delta g_i$  est l'accroissement individuel en surface terrière,  $c_i$  est la circonférence individuelle à 1.30 m.  $\gamma$  et  $\sigma$  sont les paramètres du modèle.  $\gamma$  est une fonction de l'indice de fertilité :

$$\gamma(SI) = \gamma_1 + \gamma_2 SI$$

où  $SI$  est l'indice de fertilité du chêne. La valeur du seuil de circonférence  $\sigma$  dépend de l'accroissement en surface terrière du peuplement, ce qui fait la particularité de ce modèle. En effet,  $\sigma$  est fixé de manière à ce que la somme des accroissements individuels soit égale à l'accroissement en surface terrière du peuplement. L'accroissement en surface terrière du peuplement  $\Delta G$  est une fonction de l'accroissement en hauteur dominante  $\Delta H_0$  (qui dépend de l'âge du peuplement), de la hauteur dominante  $H_0$  et de l'indice de densité de Reineke  $RDI$  qui permet d'exprimer l'intensité de la compétition dans le peuplement (Dhôte *et al.*, 2000) :

$$\frac{\Delta G}{\Delta t} = f(H_0, \Delta H_0, RDI)$$

### 3.2.2 La croissance dans le modèle Sylvestris

Dans le modèle Sylvestris, la croissance individuelle est modélisée de la façon suivante :

$$\frac{\Delta d_i}{\Delta t} = \left[ 1 - e^{-q \frac{\Delta H_0}{\Delta t}} \right] \times \left[ \alpha(G) + \beta(G) \frac{d_i}{H_0} \right]$$

Où  $\Delta d_i$  est l'accroissement en diamètre de l'arbre  $i$ ,  $d_i$  est le diamètre à 1.30 m de l'arbre  $i$ . L'accroissement en hauteur dominante  $\Delta H_0$  dépend de l'âge et de l'indice de fertilité. Contrairement au modèle Fagacées, il n'y a pas de lien entre la croissance du peuplement et la croissance individuelle.

### 3.2.3 Enseignements issus de ces modèles

Lorsque l'on considère un peuplement donné, les deux modèles utilisent une relation entre l'accroissement individuel et la taille des individus. Dans le cas du chêne, c'est l'accroissement en surface terrière qui est modélisé, et dans le cas du pin c'est l'accroissement en diamètre. De plus, ces deux modèles intègrent les effets de facteurs qui agissent sur cette relation. Il s'agit de l'indice de fertilité, de la hauteur dominante, de l'âge et d'une variable

relative à la densité du peuplement (dans le cas du chêne, c'est l'indice de densité de Reineke et dans le cas du pin, c'est la surface terrière totale).

Ces modèles fournissent donc des informations sur la croissance des deux espèces en peuplement pur et régulier qui sont évidemment des bases intéressantes pour aborder l'étude de la croissance de peuplements mélangés faisant intervenir ces essences. Ainsi, le modèle Fagacées a été utilisé et adapté pour modéliser la croissance du chêne dans des peuplements mélangés de chênes et de hêtres (Hein et Dhote, 2006). Cependant, certains concepts fondamentaux comme la hauteur dominante sont difficiles à transposer au cas des peuplements mélangés.

### **3.3 Choix des arbres et mesures dendrométriques**

Entre 2004 et 2007, 34 dispositifs d'une surface comprise entre 0.5 et 1.25ha ont été mesurés et cartographiés en forêt domaniale d'Orléans pour étudier de manière approfondie la structure spatiale horizontale des peuplements mélangés chêne sessile - pin sylvestre (Ngo Bieng, 2007). Nous nous sommes basés sur ce réseau de dispositifs pour étudier la croissance de ces peuplements. Comme nous n'avons pas de suivi de croissance de ces peuplements, nous avons choisi de mesurer la croissance en prélevant des carottes et en mesurant les largeurs de cerne. Cette méthode est particulièrement coûteuse en temps de mesure. Nous avons donc limité le nombre total d'échantillons à mesurer. Plutôt que de mesurer la croissance de tous les arbres sur un ou deux dispositifs, nous avons préféré multiplier le nombre de dispositifs en mesurant uniquement un échantillon d'arbres de chaque essence.

Dans cette sous-partie, nous allons présenter comment nous avons sélectionné les dispositifs (paragraphe 3.3.1) et comment nous avons sélectionné les arbres à l'intérieur de ces dispositifs (paragraphe 3.3.2) pour mesurer la croissance. Nous présenterons également les mesures de hauteur réalisées sur les arbres échantillonnés.

#### **3.3.1 Choix et caractéristiques des dispositifs**

Parmi les dispositifs cartographiés et inventoriés, ceux qui avaient subis des éclaircies ou de fortes perturbations dans les 10 dernières années ont été écartés. De plus, nous avons retenu des dispositifs homogènes aussi bien par rapport à la station que par rapport à la structure spatiale (Ngo Bieng, 2007). Pour 30 de ces dispositifs, 4 grands types de structure de la canopée ont été décrits à l'aide des fonctions de Ripley pour étudier la structure intraspécifique et de la fonction Intertype pour étudier les interactions entre les deux espèces (Ngo Bieng *et al.*, 2006). Ces 4 types correspondent à un gradient allant du "mélange pied à

ped" au "mélange par bouquet" (tableau 3.1). Nous avons donc choisi 9 dispositifs sur 34 en essayant de couvrir ce gradient de structure spatiale de façon à disposer de conditions de compétition très variables entre les arbres d'une même espèce mais également entre les arbres d'espèces différentes. Malheureusement, les contraintes imposées pour le choix des dispositifs (absence de perturbations récentes et homogénéité du dispositif) ne nous ont pas permis de retenir des dispositifs appartenant au type 2 de la typologie. Il s'agit aussi du type le moins représenté (4 sur 30). En revanche, le type 3 est très représenté dans notre échantillon.

Tableau 3.1 : Définition des types de structure spatiale d'après (Ngo Bieng, 2007) et répartition des dispositifs sélectionnés pour l'étude de la croissance selon le type de structure. Type = type de structure spatiale de la canopée ; Chêne = structure spécifique des chênes ; pin = structure spécifique des pins ; Chêne/Pin = structure interspécifique ; Nb disp. = nombre de dispositifs classés dans le type ; Disp. croissance = dispositifs sélectionnés pour les mesures de croissance.

Type	Chêne	Pin	Chêne/Pin	Nb disp.	Disp. croissance
1	Forte agrégation	Forte agrégation	Répulsion à petites et moyennes distances (2 à 24 m)	5	D02 ; D20
2	Agrégation	Aléatoire ou légère agrégation	Répulsion à petites distances (2 à 12 m)	4	
3	Aléatoire	Aléatoire	Indépendance ou légère répulsion	13	D108 ; D42 ; D49 ; D534 ; D563 ; D78
4	Aléatoire ou légère agrégation	Agrégation	Indépendance ou légère répulsion	8	D27

### 3.3.1.1 Caractéristiques dendrométriques

Pour les 9 dispositifs sélectionnés, nous avons mesuré les circonférences à 1.30 m de tous les arbres. Cet inventaire a eu lieu pendant l'hiver 2005-2006 pour les dispositifs D02, D20, D42 et D49 et pendant l'hiver 2006-2007 pour les dispositifs D108, D27, D534, D563 et D78. Le tableau 3.2 présente les caractéristiques dendrométriques des 9 dispositifs. D'autres espèces feuillues peuvent être présentes avec le chêne et le pin (*Carpinus betulus* L., *Betula pendula* R. et *Sorbus torminalis* L. principalement) mais elles sont très minoritaires et représentent environ 4% de la surface terrière totale (tableau 3.2). De façon générale, la distribution des diamètres des chênes est plutôt comparable à celle d'un peuplement irrégulier alors que celle des pins est comparable à celle d'un peuplement régulier (figure 3.1).

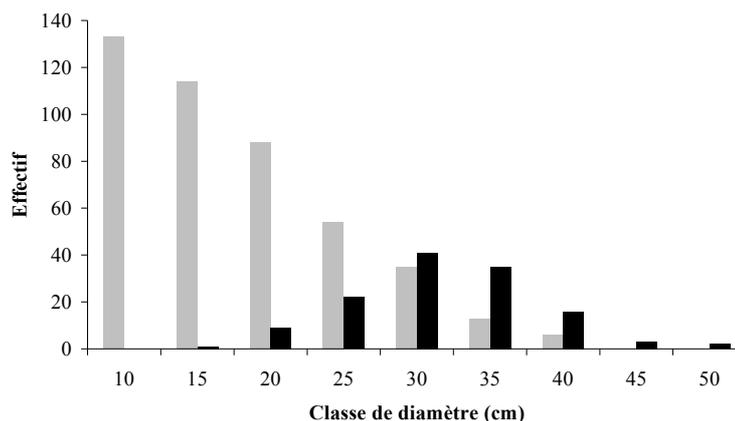


Figure 3.1 : Répartition des tiges par classe de diamètre dans le dispositif D49. Les chênes sont représentés en gris et les pins en noir.

### 3.3.1.2 Caractéristiques édaphiques

Les propriétés du sol (réserve utile, composition chimique) ont une forte influence sur la croissance des arbres. Afin d'évaluer les différences entre les dispositifs et l'hétérogénéité au sein des dispositifs, nous avons effectué des sondages pédologiques (au moins 1 par quart d'hectare) et des descriptions de la flore (espèces indicatrices du milieu) sur chacun des dispositifs. Ces descriptions nous ont permis de déterminer le type de station à chaque point de sondage d'après le catalogue des stations de l'orléanais (Brethes, 1993). Les types de station que nous avons identifiés sont relativement proches. Ainsi, les sols sont tous caractérisés par une texture à dominante sableuse avec un plancher argileux se situant entre 35 et 66 cm de profondeur et une flore caractéristique des milieux acides. Les dispositifs se différencient essentiellement par la profondeur d'apparition du plancher argileux et par la profondeur d'apparition de l'hydromorphie. Ces stations sont relativement contraignantes puisqu'il y a un engorgement l'hiver lié à la présence du plancher argileux proche de la surface et une sécheresse l'été à cause de la texture sableuse des sols. Néanmoins, les connaissances sur l'exploitation des ressources par les espèces ligneuses dans ce type de sol restent partielles (Lebourgeois *et al.*, 2008).

Tableau 3.2 : Surface terrière (G), densité (N), diamètre quadratique moyen (Dg) pour le chêne, le pin et les autres espèces feuillues dans les 9 dispositifs (Disp.). Parcelle = numéro de parcelle forestière à laquelle appartient le dispositif.

Disp.	Parcelle	Surf. (ha)	Chêne			Pin			Autres espèces feuillues			G total (m <sup>2</sup> /ha)	N total (tiges/ha)
			G (m <sup>2</sup> /ha)	N (tiges/ha)	Dg (cm)	G (m <sup>2</sup> /ha)	N (tiges/ha)	Dg (cm)	G (m <sup>2</sup> /ha)	N (tiges/ha)	Dg (cm)		
D02	178	0.95	17.5	395	23.8	10.6	97	37.3	1.7	95	15.1	29.8	587
D108	108	0.80	9.8	355	18.8	23.4	233	35.8	1.5	99	13.9	34.8	686
D20	57	1.02	11.7	488	17.5	15.1	164	34.3	0.5	41	11.9	27.2	693
D27	184	0.63	14.1	469	19.6	14.4	133	37.1	3.0	242	12.5	31.4	843
D42	216	0.50	12.0	472	18.0	15.8	280	26.8	1.0	80	12.6	28.8	832
D49	255	0.99	13.5	496	18.6	19.2	241	31.9	1.3	65	15.6	34.0	803
D534	534	0.50	12.5	494	17.9	19.5	172	38.0	1.0	64	13.8	32.9	730
D563	563	0.50	13.7	242	26.9	19.7	214	34.3	0.2	22	11.5	33.7	478
D78	78	0.70	14.8	409	21.4	16.5	113	43.1	1.0	57	14.8	32.2	579

### 3.3.2 Choix des arbres à l'intérieur des dispositifs

Nous avons vu que nous souhaitons en particulier construire un modèle de type arbre dépendant des distances. Dans le cadre des peuplements purs, nous avons vu que la taille de l'arbre est une variable importante à prendre en compte (sous-partie 3.2). Dans les peuplements hétérogènes, les différences de compétition entre individus (que l'on peut estimer à l'aide d'un indice de compétition local) peuvent également être importantes à prendre en compte (paragraphe 2.1.5). Si on se place au niveau d'un peuplement, nous avons donc considéré trois variables pour diriger notre échantillonnage individuel : l'essence, la taille de l'arbre et un indice de compétition local.

Pour avoir un domaine de validité le plus large possible, nous souhaitons que le modèle de croissance soit capable de prédire avec la même précision l'évolution d'arbres ayant des circonférences et des environnements locaux très variables. Nous devons donc faire en sorte que les arbres échantillonnés se répartissent le plus uniformément possible dans les différentes classes de circonférence et les différentes classes d'indices de compétition : c'est le principe d'un échantillonnage stratifié (Tomassone *et al.*, 1993 page 57).

Pour choisir les indices de compétition, nous nous sommes basés sur une première analyse réalisée à partir des premiers dispositifs sondés dans l'hiver 2005-2006 (Perot *et al.*, 2008). Dans le cas des chênes, l'indice de compétition qui a été choisi pour la stratification est la somme des rapports entre le diamètre des voisins ( $dbh_j$ ) et la distance à l'arbre sujet ( $l_{ij}$ ) :

$$CI_i = \sum_{j=1}^n \frac{dbh_j}{l_{ij}}$$

Pour les pins, nous avons choisi la somme des surfaces terrières des voisins pins ( $g_{j,pin}$ ) :

$$CI_i = \sum_{j=1}^{n_{pin}} g_{j,pin}$$

Les indices ont été calculés dans un rayon de 10 m. Ce rayon se situe dans la gamme des distances obtenues dans d'autres études (Canham *et al.*, 2004 ; Uriarte *et al.*, 2004a ; Stadt *et al.*, 2007). Nous verrons dans la partie 4 que ces indices ne sont pas forcément ceux qui seront retenus dans le modèle final car les cartographies nous permettent de tester d'autres indices et d'autres rayons de voisinage.

Nous disposons de 3 variables de stratification au niveau des arbres : l'essence, la taille et une variable caractérisant la compétition locale. Nous allons voir maintenant comment nous avons procédé pour sélectionner les individus dans les peuplements pour constituer l'échantillon final.

Pour chaque dispositif retenu, nous avons restreint l'effectif par essence à 30 individus. De plus, nous avons choisi de ne pas sélectionner d'arbres à moins de 10 m de la limite d'un dispositif. Nous souhaitons, de cette façon, limiter les corrections dues aux effets de bord dans le calcul des indices de compétition qui seront retenus dans le modèle de croissance. Nous avons donc un échantillon de 30 arbres par essence à construire dans la zone centrale. En respectant ces conditions, les individus de l'échantillon final ont été sélectionnés suivant une stratégie d'échantillonnage individuel stratifiée et optimisée. Cette stratégie a été choisie après comparaison de différents types d'échantillonnage (Goreaud *et al.*, 2006b).

Pour les raisons que nous avons évoquées plus haut, l'objectif de cet échantillonnage est de faire en sorte que la distribution des valeurs des différentes variables de stratification soit la plus uniforme possible dans l'échantillon<sup>5</sup>. Nous appellerons "échantillon optimal" tout échantillon qui se rapprochera le plus possible d'une distribution uniforme, compte tenu des contraintes du peuplement. Pour obtenir cet "échantillon optimal", nous avons tout d'abord choisi de représenter la distribution des variables continues de stratification à l'aide de 10 classes de valeurs. Pour une essence donnée, l'échantillon parfait comprendrait donc 3 individus pour chacune des classes et ceci pour chaque variable de stratification. Afin de comparer l'échantillon avec cet échantillon théorique, nous avons utilisé un indice de comparaison. Cet indice est défini, pour chaque essence et pour chaque variable retenue, comme la somme des carrés des écarts entre les effectifs des deux échantillons, pour chaque classe (équations 3.1 et 3.2) : c'est une distance euclidienne entre les deux distributions. L'intérêt principal de cet indice est qu'il défavorise l'accumulation dans une même classe de plusieurs arbres supplémentaires.

$$D_{circ} = \sum_{i=1}^{10} (\text{Effectif}(\text{ClasseCirc}_i) - 3)^2 \quad (3.1)$$

$$D_{CI} = \sum_{i=1}^{10} (\text{Effectif}(\text{ClasseCI}_i) - 3)^2 \quad (3.2)$$

où  $\text{ClasseCirc}_i$  est la classe  $i$  de la distribution en classe de circonférence et  $\text{ClasseCI}_i$  est la classe  $i$  de la distribution en classe d'indice de compétition. Pour obtenir l'échantillon optimal nous avons utilisé un algorithme d'optimisation qui fait intervenir les indices de comparaison dans une fonction de coût :

$$\text{Coût} = D_{circ} + D_{IC}$$

---

<sup>5</sup> Pour un peuplement réel, il est parfois impossible de trouver un échantillon avec une distribution tout à fait uniforme, parce que certaines classes de valeurs ne correspondent pas à assez d'arbres dans le peuplement. Nous devons donc nous contenter de nous approcher le plus possible de cette distribution uniforme.

Plus cette fonction de coût est faible, plus l'échantillon se rapproche de l'échantillon optimal (ici la stratification sur la compétition locale et la stratification sur la taille de l'arbre ont le même poids). L'algorithme d'optimisation consiste à tirer au sort un échantillon dans la population puis à remplacer un individu de l'échantillon par un individu de la population en appliquant les règles suivantes :

- a) le coût est calculé pour l'échantillon ;
- b) un individu de l'échantillon est tiré au sort ;
- c) il est remplacé par un individu de la population qui n'a pas déjà été pris dans l'échantillon ;
- d) le coût est calculé avec ce nouvel individu ;
- e) si le nouveau coût est inférieur à l'ancien, le nouvel individu est conservé dans l'échantillon.

Cette boucle est répétée 10 000 fois afin d'obtenir un échantillon pour lequel le coût est minimum au regard des contraintes imposées et des caractéristiques du peuplement. L'obtention de l'échantillon a été réalisée à l'aide du programme informatique PASS (Goreaud et Ngo Bieng, 2005). La figure 3.2 présente un exemple d'arbres sélectionnés selon cette procédure dans le dispositif D108. La figure 3.3 présente sur ce même dispositif la distribution des pins dans le peuplement et dans l'échantillon selon les deux variables de stratification.

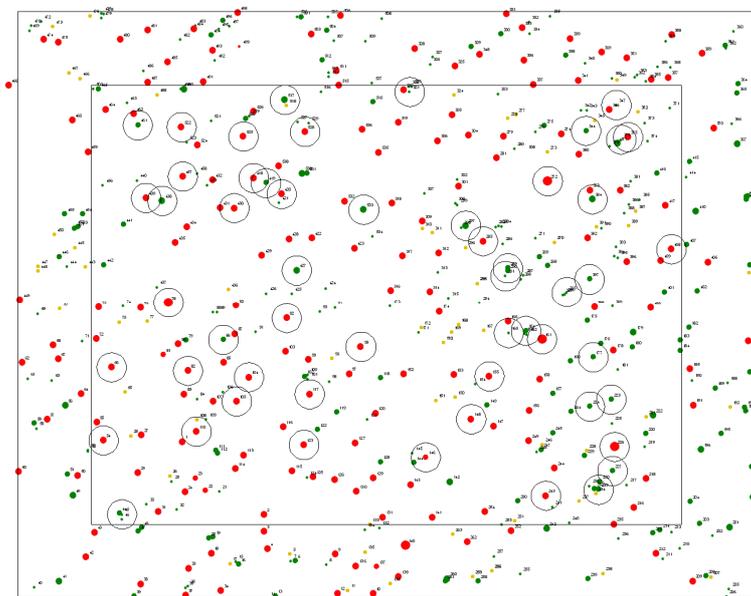


Figure 3.2 : Résultat de l'échantillonnage dans le dispositif D108 (0.8 ha). Les arbres sélectionnés sont entourés. Les chênes sont en vert et les pins en rouge.

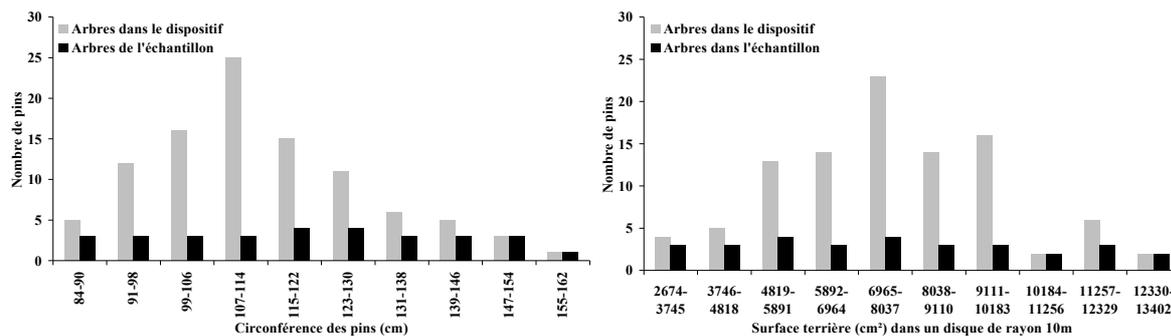


Figure 3.3 : Exemple d'échantillonnage pour les pins du dispositif D108. A gauche : distribution de la circonférence dans le dispositif et dans l'échantillon. A droite : distribution de l'indice de compétition dans le dispositif et dans l'échantillon.

Il est important de préciser une nouvelle fois que nous avons pu réaliser ce type d'échantillonnage parce qu'un inventaire et une cartographie de tous les dispositifs avaient été réalisés dans un travail précédent (Ngo Bieng, 2007). De plus, nous rappelons que cet échantillonnage a pour objectif de construire un modèle ayant le domaine de validité le plus large possible par rapport aux variables de stratification. Les arbres échantillonnés ne sont donc pas représentatifs du peuplement. Par exemple, sur la figure 3.2 les arbres de petit diamètre sont surreprésentés dans l'échantillon par rapport aux arbres du peuplement.

### 3.3.3 Mesure de la hauteur totale des arbres

Dans cette thèse, nous avons uniquement travaillé sur la croissance en diamètre. Cependant, en dendrométrie, la hauteur des arbres est une variable importante car elle permet le calcul des volumes (Vallet *et al.*, 2006). La hauteur totale est également une variable qui est fréquemment utilisée dans les modèles relatifs à la branchaison (prédiction de caractéristiques des branches ou des caractéristiques plus globales du houppier) ainsi que dans le calcul de nombreux indices : indice de compétition, critère de stabilité, position sociale de l'arbre (Perot *et al.*, 2007). C'est pourquoi, en complément de la circonférence à 1.30 m, nous avons mesuré la hauteur totale de tous les arbres échantillonnés (figure 3.4). Ces mesures seront utilisées de façon plus approfondie en dehors de cette thèse. Comme nous le verrons un peu plus loin, nous avons utilisé ces données pour estimer une hauteur dominante par dispositif et par essence.

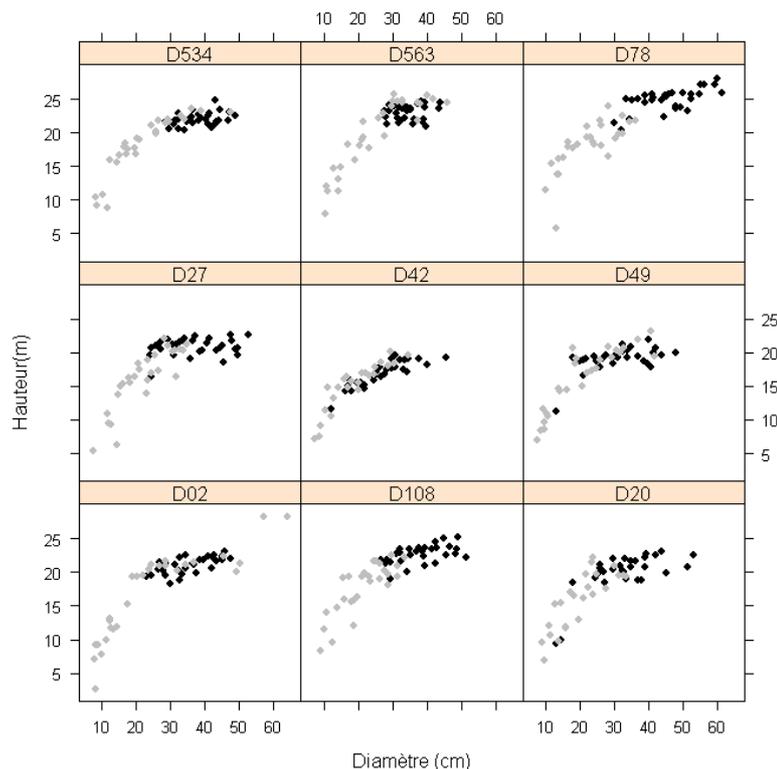


Figure 3.4 : Hauteur en fonction du diamètre à 1.30 m pour les arbres échantillonnés. Les points gris représentent les chênes et les points noirs les pins.

### 3.4 Mesure de la croissance individuelle

#### 3.4.1 Méthode utilisée

Sur chaque arbre échantillonné, deux carottes ont été prélevées à 1.30 m suivant deux directions perpendiculaires. Les arbres ont été sondés jusqu'au cœur de façon à obtenir des informations sur leur âge. Après préparation, les carottes ont été scannées puis analysées avec le logiciel Windendro version 2005a (Regent, 2005). Les largeurs de cernes ont été mesurées avec une précision de 0.01 mm. Le logiciel COFECHA (Grissino-Mayer, 2002) a été utilisé pour détecter les erreurs de datation. Comme certains arbres n'ont pu être sondés et que certaines carottes étaient illisibles, il restait au final 230 chênes et 269 pins pour les analyses. L'annexe B présente les courbes d'évolution de la circonférence pour chacun des dispositifs. Après une étude sur l'historique des parcelles, la période de croissance 2000-2005 a été retenue pour construire les modèles de croissance car cette période ne présentait pas de perturbation naturelle (tempête) ou anthropique (éclaircies). La circonférence en début de période (année 1999) a été reconstituée pour chaque arbre échantillonné à partir des largeurs de cerne et de la circonférence mesurée lors de l'inventaire (2005 ou 2006 suivant les dispositifs).

### 3.4.2 Informations sur les âges

Les données rétrospectives de croissance nous ont permis de connaître plus précisément l'âge des arbres. On constate d'une part que les arbres à l'intérieur d'une espèce ont tous à peu près le même âge, et d'autre part, qu'il peut exister un décalage plus ou moins important entre l'âge des chênes et l'âge des pins (figure 3.5 et tableau 3.3).

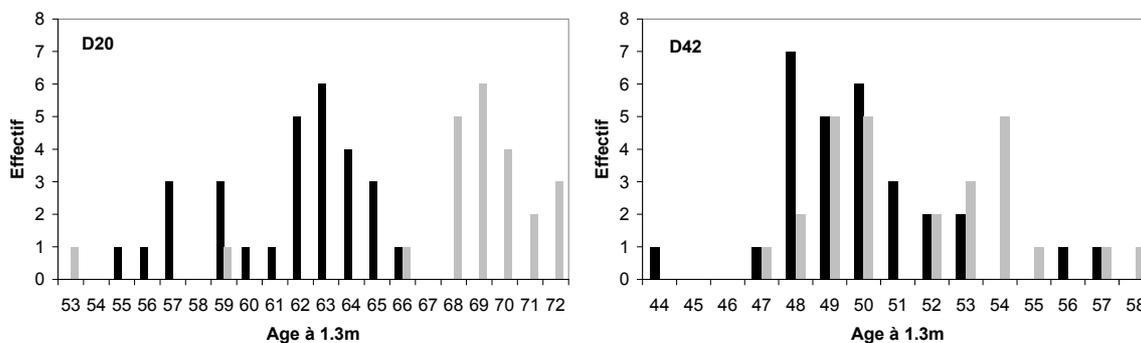


Figure 3.5 : Répartition des âges à 1.30 m pour les arbres sondés dans deux dispositifs. En gris les chênes et en noir les pins. A gauche : un dispositif où les pins ont environ 10 ans de moins que les chênes. A droite : un dispositif où les chênes et les pins ont pratiquement le même âge.

Tableau 3.3 : Données sur les âges obtenus sur les arbres échantillonnés. \* = statistiques sur les âges du dispositif D02 en enlevant les 5 chênes beaucoup plus âgés. \*\* = statistiques sur les âges du dispositif D78 en enlevant les 3 pins beaucoup plus jeunes.

Dispositif	essence	Nombre	Age Moyen	Age Median	Age min	Age max	Écart-type
D02	Chêne	30	91.3	80.0	61	174	31.96
D02*	Chêne	25	77.6	80.0	61	84	5.56
D108	Chêne	25	68.8	70.0	61	72	3.10
D20	Chêne	24	67.1	69.0	41	72	6.94
D27	Chêne	24	72.0	73.0	35	109	15.26
D42	Chêne	26	51.6	51.0	47	58	2.90
D49	Chêne	25	67.5	67.0	59	86	6.11
D535	Chêne	25	59.2	59.0	55	66	2.25
D563	Chêne	27	69.7	70.0	62	74	3.13
D78	Chêne	25	61.5	63.0	47	68	5.69
D02	Pin	30	77.3	77.5	72	81	2.16
D108	Pin	30	66.1	66.0	61	70	2.33
D20	Pin	29	61.6	62.0	55	66	3.00
D27	Pin	30	68.2	69.5	51	72	4.17
D42	Pin	29	49.9	50.0	44	57	2.62
D49	Pin	27	61.7	62.0	44	70	5.36
D535	Pin	31	83.6	84.0	75	93	3.30
D563	Pin	30	69.2	69.0	65	76	2.29
D78	Pin	29	112.3	118.0	60	121	17.55
D78**	Pin	26	118.1	118.0	116	121	1.37

Néanmoins, quelques dispositifs se distinguent. Dans le dispositif D02, nous avons mesuré 5 chênes beaucoup plus âgés que les autres. Ces arbres sont des témoins d'un mode de gestion ancien : le taillis sous-futaie. Ces arbres ont eu des conditions de croissance complètement différentes des autres individus. C'est pourquoi par la suite ces arbres ont été retirés des analyses sur la croissance. De même, dans le dispositif D78, 3 pins beaucoup plus jeunes appartenant à une autre population que les autres pins du dispositif ont été mesurés. De la même façon que pour les chênes issus du taillis sous-futaie, ils ont été retirés des analyses sur la croissance. Enfin, la variabilité des âges des chênes dans le dispositif D27 est élevée mais aucun critère ne nous a permis d'expliquer cette variabilité. Nous avons donc conservé l'ensemble des chênes de ce dispositif pour les analyses sur la croissance.

### 3.4.3 Estimations d'indices de fertilité par essence

La connaissance du niveau de fertilité du site est un élément essentiel pour prévoir la production d'un peuplement. Dans le cadre des peuplements purs et réguliers, on utilise généralement un indice de fertilité correspondant à la hauteur dominante atteinte à un âge donné (Duplat et Tran-Ha, 1986). Par exemple, pour le chêne, il s'agit de la hauteur dominante atteinte à 100 ans (Dhôte *et al.*, 2000) et pour le pin sylvestre, c'est la hauteur dominante atteinte à 50 ans (Perot *et al.*, 2007). Cependant, la notion de hauteur dominante et les concepts qui en découlent sont valables pour des peuplements purs réguliers. En peuplement mélangé, les effets de la compétition entre les essences peuvent jouer sur la croissance en hauteur dominante et modifier la valeur de l'indice de fertilité. En particulier dans le dispositif D78, les chênes sont largement dominés par les pins (figure 3.4). Nous ne pouvons donc pas utiliser ces indices pour estimer directement la fertilité dans nos peuplements mélangés. Malgré tout, ces indices donnent une idée de la performance des deux essences en mélange. De plus, dans nos peuplements, les âges des espèces sont plutôt homogènes ce qui les rapproche des peuplements réguliers. Nous avons donc tenté d'appliquer cette méthode à nos peuplements. Pour cela, nous avons estimé pour chaque essence et chaque dispositif une hauteur dominante et un âge correspondant. Ensuite, nous avons utilisé les modèles existants pour le chêne sessile (Dhôte *et al.*, 2000) et pour le pin sylvestre (Perot *et al.*, 2007) pour calculer les indices de fertilité. Pour estimer la hauteur dominante, nous avons établi une relation entre la circonférence et la hauteur des arbres échantillonnés. Cette relation nous a permis de calculer la hauteur correspondant à la circonférence dominante que nous avons retenue comme estimation de la hauteur dominante. Pour estimer l'âge correspondant à cette hauteur dominante, nous avons calculé l'âge moyen des arbres

échantillonnés ayant participé au calcul de la circonférence dominante. Les modèles de croissance en hauteur dominante présentés ci-dessus utilisent l'âge total. Pour le chêne comme pour le pin, nous avons donc ajouté 7 années (temps estimé pour atteindre 1.30 m) à l'âge à obtenu à 1.30 m pour avoir un âge total. Le tableau 3.4 et la figure 3.6 montrent que pour le chêne l'indice de fertilité varie de 22.8 m à 28 m ce qui correspondrait, d'après le guide de sylviculture du chêne (Jarret, 2004) à des indices compris entre la classe 3 (faible fertilité) et la classe 2 (bonne fertilité). Pour le pin, l'indice de fertilité varie de 15.3 m à 18.1 m ce qui correspondrait à la classe 3 (faible fertilité) d'après le guide de sylviculture du pin (Chabaud et Nicolas, Sous presse). Ces informations seront utilisées dans la partie 4 et dans la partie 5.

Tableau 3.4 Estimation des hauteurs dominantes (Hdom) et des indices de fertilité (SI) par dispositif (Dispo) et par essence. SI = indice de fertilité, pour le chêne c'est la hauteur dominante à 100 ans et pour le pin c'est la hauteur dominante à 50 ans. Cdom = circonférence dominante.

Dispo	Essence	Cdom (cm)	Hdom (m)	Age total	SI (m)
D02	Chêne	128.0	21.1	88	23.76
D108	Chêne	106.1	22.3	76	25.96
D20	Chêne	93.8	20.4	78	23.5
D27	Chêne	117.7	21.9	93	22.82
D42bis	Chêne	95.4	18.8	61	25.06
D49	Chêne	102.0	20.1	79	23
D534	Chêne	103.0	22.1	66	27.86
D563	Chêne	126.3	24.5	78	28.01
D78	Chêne	107.6	21.8	72	26.18
D02	Pin	139.7	22.1	85	15.93
D108	Pin	131.5	23.0	73	18.11
D20	Pin	125.4	21.2	69	17.2
D27	Pin	138.5	20.8	75	16.06
D42bis	Pin	119.9	19.0	58	17.17
D49	Pin	121.3	19.7	69	15.95
D534	Pin	140.1	22.5	92	15.56
D563	Pin	124.6	23.0	75	17.82
D78	Pin	158.3	25.6	125	15.32

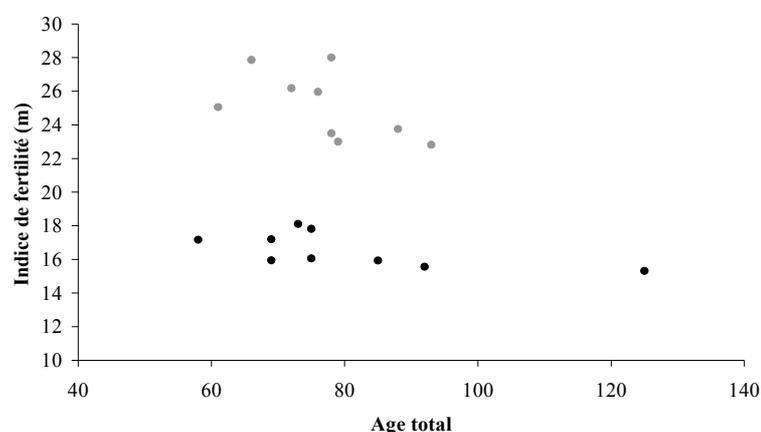


Figure 3.6 : Indice de fertilité en fonction de l'âge total à 1.30 m des arbres dominants.

### 3.5 Partage des données pour ajuster et évaluer les modèles

Lorsque l'on souhaite évaluer la qualité prédictive des modèles, il est fortement conseillé d'utiliser un jeu de données indépendant<sup>6</sup> des données de calibration (Vanclay et Skovsgaard, 1997 ; Astrup *et al.*, 2008). Plusieurs stratégies sont possibles pour constituer ce jeu de données (Meredieu, 1998) : acquérir de nouvelles données, partager les données déjà acquises ou utiliser des techniques de rééchantillonnage. Nous avons écarté les techniques de rééchantillonnage car elles ne nous sont pas apparues adaptées pour évaluer, selon une procédure commune, l'ensemble de nos modèles. Etant donné le coût d'acquisition des données de croissance, nous avons donc choisi de partager les données acquises en deux parties : une partie pour la calibration des modèles et une partie pour leur évaluation.

Dans notre cas, au moins trois options sont possibles pour partager les données : partager les données au niveau des dispositifs, partager les données au niveau des arbres sans tenir compte des dispositifs ou partager les données au niveau des arbres en tenant compte des dispositifs. Une première analyse réalisée sur les premiers dispositifs mesurés (D02, D20, D42 et D49) a mis en évidence un effet du facteur dispositif sur la croissance (Perot *et al.*, 2008). Etant donné le faible nombre de dispositifs, nous avons jugé qu'il n'était pas possible de modéliser cet effet (nous y reviendrons dans la partie 4). La connaissance du dispositif est donc nécessaire pour prédire la croissance. Pour cette raison, nous avons partagé les données au niveau des arbres en prenant en compte le niveau dispositif. Ainsi, nous avons réalisé le partage en effectuant un tirage aléatoire sans remise par dispositif en conservant 2/3 des données pour l'ajustement et 1/3 pour l'évaluation (Steyerberg, 2009 page 301) (tableau 3.5). L'indépendance entre notre jeu de données d'évaluation et notre jeu de données de calibration est évidemment discutable puisque nous avons sélectionné des arbres dans les mêmes dispositifs. Cependant, d'une manière générale, la constitution d'un jeu de données indépendant pour évaluer les modèles est une question difficile (Vanclay et Skovsgaard, 1997) et nous verrons dans la partie 5 l'intérêt de nos données d'évaluation.

---

<sup>6</sup> Voir aussi partie 5.

Tableau 3.5 : Répartition des données entre le jeu d'ajustement et le jeu d'évaluation par dispositif et par essence.

Dispositif	Chêne		Pin	
	Ajustement	Validation	Ajustement	Validation
D02	14	7	23	11
D108	17	8	19	10
D20	18	9	19	10
D27	16	8	20	10
D42	18	9	21	10
D49	20	10	20	10
D534	17	8	21	10
D563	18	9	19	10
D78	16	8	17	9
<b>Total</b>	<b>154</b>	<b>76</b>	<b>179</b>	<b>90</b>

### 3.6 Bilan

Dans cette partie, nous avons présenté les peuplements sur lesquels nous avons appliqué notre questionnement et nous avons présenté les méthodes utilisées pour l'acquisition des données de croissance en peuplement mélangé. Sur la base de données acquises (inventaires et cartographies de peuplement), de connaissances a priori, et de contraintes, nous avons mis en œuvre un échantillonnage individuel stratifié et optimisé. Cette méthode nous a permis de faire des mesures de croissance dans 9 dispositifs en sélectionnant dans chaque dispositif 30 arbres par essence. Les mesures rétrospectives de la croissance des arbres nous ont permis d'acquérir des connaissances sur l'âge des arbres qui ont été utilisées pour estimer des indices de fertilité par espèces. L'étude sur l'historique des parcelles nous a amené à choisir la période de croissance 2000-2005 pour construire les modèles. Enfin, nous avons présenté comment nous avons partagé nos données de croissance afin d'assurer la calibration et l'évaluation de nos modèles. Ce partage a été réalisé en conservant 2/3 des données pour l'ajustement des modèles et 1/3 pour leur évaluation. Dans la partie suivante, nous allons utiliser les données que nous venons de présenter pour construire une famille de modèles de croissance adaptés aux peuplements chêne-pin.

## 4. Construction et ajustement d'une famille de modèles de croissance

### - Sommaire de la partie -

4.1	Le modèle arbre indépendant des distances .....	51
4.1.1	<i>La taille de l'arbre traduit son potentiel de croissance</i> .....	51
4.1.2	<i>Résultats de l'effet taille de l'arbre sur la croissance</i> .....	57
4.1.3	<i>L'hypothèse reliant le potentiel de croissance et la taille est-elle pertinente ?</i> .....	59
4.1.4	<i>A quoi peut-on attribuer l'effet dispositif ?</i> .....	61
4.1.5	<i>Les paramètres de croissance évoluent au cours du temps</i> .....	63
4.1.6	<i>Le chêne et le pin ne poussent pas de manière indépendante : mise en évidence d'une compensation de croissance</i> .....	68
4.2	Le modèle arbre dépendant des distances .....	74
4.2.1	<i>Prise en compte de l'influence du voisinage</i> .....	74
4.2.2	<i>Résultats sur l'effet du voisinage sur la croissance</i> .....	76
4.2.3	<i>L'intérêt d'introduire un indice de compétition local dépend de l'espèce considérée</i> .....	80
4.2.4	<i>Comment séparer le potentiel de croissance de la compétition locale ?</i> .....	81
4.2.5	<i>La compétition intraspécifique est supérieure à la compétition interspécifique</i> .....	85
4.3	Le modèle arbre basé sur des distributions de voisinage .....	86
4.3.1	<i>Estimation des distributions de voisinage</i> .....	86
4.3.2	<i>Fonctionnement du modèle</i> .....	88
4.3.3	<i>Utilisation du modèle dans un contexte de gestion</i> .....	90
4.3.4	<i>Limites du modèle</i> .....	90
4.3.5	<i>Conclusion</i> .....	92
4.4	Construction de modèles peuplement par agrégation des modèles individus-centrés.....	93
4.4.1	<i>Agrégation du modèle non spatialisé</i> .....	94
4.4.2	<i>Agrégation du modèle spatialisé</i> .....	94
4.4.3	<i>Influence du taux de mélange sur la production du peuplement</i> .....	99
4.4.4	<i>A propos de l'erreur d'agrégation</i> .....	107
4.5	Bilan.....	109

Dans la partie 2, nous avons vu que la croissance d'un peuplement pouvait être modélisée de façon plus ou moins détaillée. Nous avons également montré comment on pouvait assurer une certaine cohérence entre les modèles pour pouvoir les comparer. Dans cette partie, nous allons présenter la construction de quatre grands types de modèles adaptés à la croissance des peuplements mélangés chêne-pin de la région Centre (voir partie précédente) et ayant des niveaux de détail différents. Nous verrons tout d'abord la construction du modèle arbre indépendant des distances et du modèle arbres dépendant des distances (sous-parties 4.1 et 4.2). Nous verrons également les connaissances qu'ils permettent d'acquérir et les questions qu'ils soulèvent par rapport à la croissance et à la dynamique des peuplements mélangés chêne-pin. Nous verrons ensuite la construction d'un modèle basé sur des distributions de voisinage (sous partie 4.3) qui permet d'utiliser l'équation de croissance du modèle arbre dépendant des distances sans utiliser une cartographie du peuplement. Nous verrons enfin la construction de modèles peuplement obtenus en agrégeant les modèles individus-centrés (sous partie 4.4).

## ***4.1 Le modèle arbre indépendant des distances***

### **4.1.1 La taille de l'arbre traduit son potentiel de croissance**

La croissance d'un arbre est liée à sa capacité à exploiter les ressources aériennes et souterraines de son environnement. Nous faisons donc l'hypothèse classique que dans un peuplement, plus un arbre est gros, plus son potentiel de croissance est important (Wykoff, 1990). C'est également l'hypothèse qui a été retenue pour les modèles de croissance du chêne et du pin en peuplements purs (Dhôte *et al.*, 2000 ; Perot *et al.*, 2007). Pour rendre compte de la taille des arbres, nous avons utilisé la circonférence à 1.30 m<sup>7</sup>. A partir de là, plusieurs choix restent possibles, en particulier ceux concernant la variable à expliquer et la forme du modèle.

#### **4.1.1.1 Quelle variable pour rendre compte de l'accroissement ?**

Dans les modèles de croissance forestiers, deux types de variables sont généralement utilisés pour rendre compte de l'accroissement individuel : les variables liées à l'accroissement en surface terrière et les variables liées à l'accroissement en rayon. Les données d'accroissement en volume sont rarement disponibles. D'un point de vue biologique, certains auteurs préfèrent utiliser l'accroissement en surface terrière plutôt que l'accroissement en

---

<sup>7</sup> Nous aurions pu prendre la hauteur totale mais dans ce cas il aurait fallu reconstituer la hauteur en début de période de croissance.

rayon car ils considèrent que cette variable est plus proche de l'accroissement en volume et donc plus proche du phénomène de croissance à l'échelle de l'arbre. Cependant, en termes de prédiction, West (1980) montre que la précision du diamètre futur est la même quelle que soit la variable utilisée pour décrire l'accroissement individuel (accroissement en surface terrière ou accroissement sur le rayon). Néanmoins, pour ce qui concerne l'étude du phénomène observé, certaines variables peuvent être plus intéressantes que d'autres.

La figure 4.1 présente les accroissements en surface terrière et les accroissements en rayon pour la période 2000-2005 en fonction de la circonférence en 1999. Nous laissons de côté l'accroissement en volume car nous n'avons pas accès à cette variable.

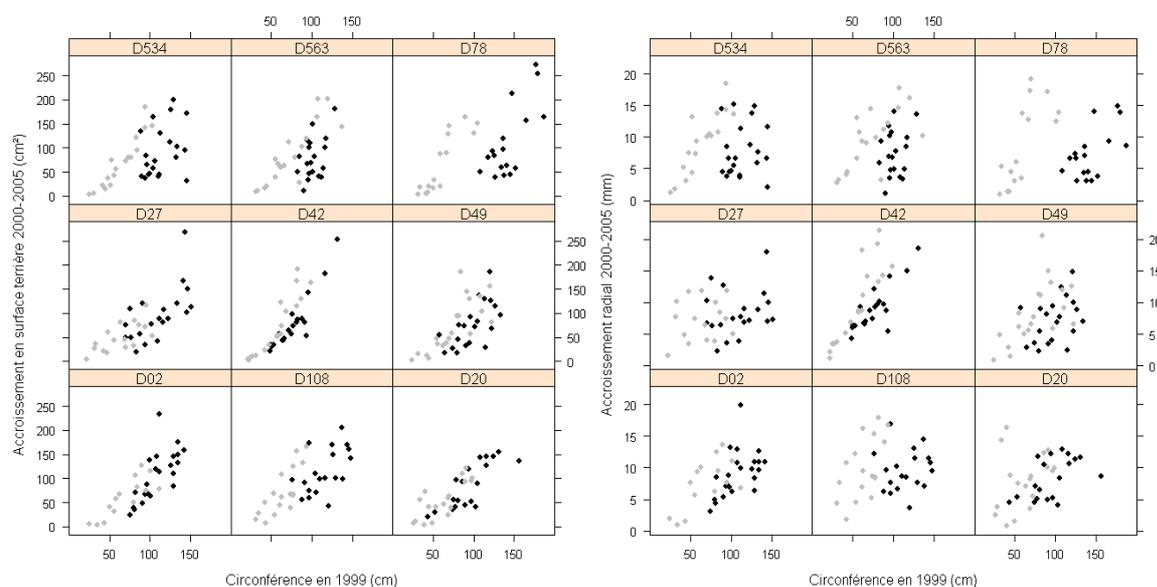


Figure 4.1 : Accroissements sur la période 2000-2005 en fonction de la circonférence 1999 (cm). A gauche, accroissement en surface terrière (cm<sup>2</sup>), R<sup>2</sup> moyen\* pour le chêne = 0.71, R<sup>2</sup> moyen\* pour le pin = 0.52. A droite, accroissement en rayon (mm) en fonction de la circonférence, R<sup>2</sup> moyen\* pour le chêne = 0.45, R<sup>2</sup> moyen\* pour le pin = 0.25. Les chênes sont représentés par les points gris et les pins par les points noirs. \*Les R<sup>2</sup> moyen correspondent à la moyenne des R<sup>2</sup> obtenus par régression linéaire simple pour chaque essence et chaque dispositif (l'éventuelle non linéarité de la relation n'est pas prise en compte).

Nous voyons qu'il y a généralement une bonne relation entre l'accroissement en surface terrière et la circonférence en début de période de croissance. Globalement, la relation est meilleure pour le chêne que pour le pin. Pour le chêne, la relation n'est pas tout à fait linéaire et la variabilité de l'accroissement augmente quand la circonférence augmente (présence d'une hétéroscédasticité). Lorsque l'on observe l'accroissement en rayon en fonction de la circonférence en début de période, la relation est moins nette (voir les R<sup>2</sup> moyens indiqués dans la légende de la figure 4.1) et la variabilité est plus importante. En revanche la relation entre l'accroissement en rayon et la circonférence est plutôt linéaire, en particulier pour le chêne. En conclusion, le choix de prendre l'accroissement en surface terrière ou l'accroissement en rayon n'apparaît pas ici de façon évidente, et peut dépendre des objectifs

que l'on se fixe. D'autres critères que nous allons préciser dans les prochains paragraphes peuvent nous aider à faire ce choix.

#### 4.1.1.2 Quelle forme de modèle ?

Sur la base des observations précédentes, nous avons évalué 4 types de modèles pouvant théoriquement s'appliquer à l'accroissement en rayon ou à l'accroissement en surface terrière :

- un modèle segmenté avec un accroissement nul en dessous d'un seuil de circonférence  $A$  et un accroissement qui suit une droite de pente  $P$  au dessus du seuil. Ce modèle, présenté par l'équation 4.1, a été utilisé pour l'accroissement en surface terrière du hêtre et du chêne en peuplement pur (Dhôte, 1991) ;
- un modèle segmenté avec un accroissement constant en dessous d'un seuil de circonférence  $A$  et un accroissement qui suit une droite au dessus du seuil. Dans le cas du chêne, l'accroissement constant en dessous du seuil pourrait correspondre à l'accroissement minimum permettant la survie de l'arbre (Dhôte, 1999). Ce modèle correspond à l'équation 4.2 ;
- un modèle continu de type hyperbolique avec trois paramètres, le seuil  $A$ , la pente  $P$  et un paramètre de forme  $m$  qui fait la transition entre l'accroissement avant et après le seuil. Ce modèle, présenté par l'équation 4.3, a été utilisé pour modéliser plusieurs espèces en peuplement pur (Deleuze, 2000 ; Deleuze *et al.*, 2004) ;
- un modèle segmenté avec deux portions de droite qui se rejoignent au niveau d'un seuil  $A$ . La pente de la droite en dessous du seuil est largement inférieure à la pente de la droite au dessus du seuil. Il s'agit d'une forme intermédiaire aux trois autres modèles. Ce modèle correspond à l'équation 4.4 ;

$$\frac{\Delta x}{\Delta t} = P(circ_{i,t} - A) \text{ si } circ_{i,t} \geq A, 0 \text{ sinon} \quad (4.1)$$

$$\frac{\Delta x}{\Delta t} = P(circ_{i,t} - A) + k \text{ si } circ_{i,t} \geq A, k \text{ sinon} \quad (4.2)$$

$$\frac{\Delta x}{\Delta t} = P \left( \frac{circ_{i,t} - mA + \sqrt{(mA + circ_{i,t})^2 - 4Acirc_{i,t}}}{2} \right) \quad (4.3)$$

$$\frac{\Delta x}{\Delta t} = P(circ_{i,t} - A) + k \times A \text{ si } circ_{i,t} \geq A, k \times circ_{i,t} \text{ sinon} \quad (4.4)$$

où  $\Delta x$  représente l'accroissement en rayon ou en surface terrière,  $\Delta t$  l'intervalle de temps,  $circ_{i,t}$  la circonférence de l'arbre  $i$  au temps  $t$ ,  $P$ ,  $A$ ,  $k$  et  $m$  les paramètres des modèles.

Ces 4 modèles se différencient essentiellement par la façon de prendre en compte l'accroissement en dessous du seuil de circonférence  $A$  (accroissements nuls ou faibles). D'après les données présentées dans la figure 4.1, ils semblent donc plutôt adaptés à la relation entre l'accroissement en surface terrière et la circonférence. Pour l'accroissement en rayon, l'équation 4.1 paraît suffisante. Cependant, dans le cadre de ce paragraphe, nous avons également évalué les autres équations pour modéliser la relation entre l'accroissement en rayon et la circonférence. Afin d'évaluer la pertinence de ces différents modèles, nous avons comparé leur capacité à prédire à la fois l'accroissement en rayon et l'accroissement en surface terrière de manière cohérente. En effet, nous souhaitons que la forme des accroissements prédits corresponde aux données présentées au paragraphe 4.1.1.1. Par exemple, si on utilise le modèle de l'équation 4.2 pour prédire l'accroissement en surface terrière, comment l'accroissement en rayon est-il prédit ? Pour passer d'un accroissement à l'autre à partir d'un modèle, nous avons utilisé la relation suivante entre l'accroissement en rayon et l'accroissement en surface terrière :

$$\frac{dg}{dt} = 2\pi r \frac{dr}{dt} \quad (4.5)$$

Nous pouvons donc étudier la forme de l'accroissement en rayon lorsqu'on utilise un modèle pour prédire l'accroissement en surface terrière et réciproquement. Les résultats sous forme graphique sont présentés dans la figure 4.2.

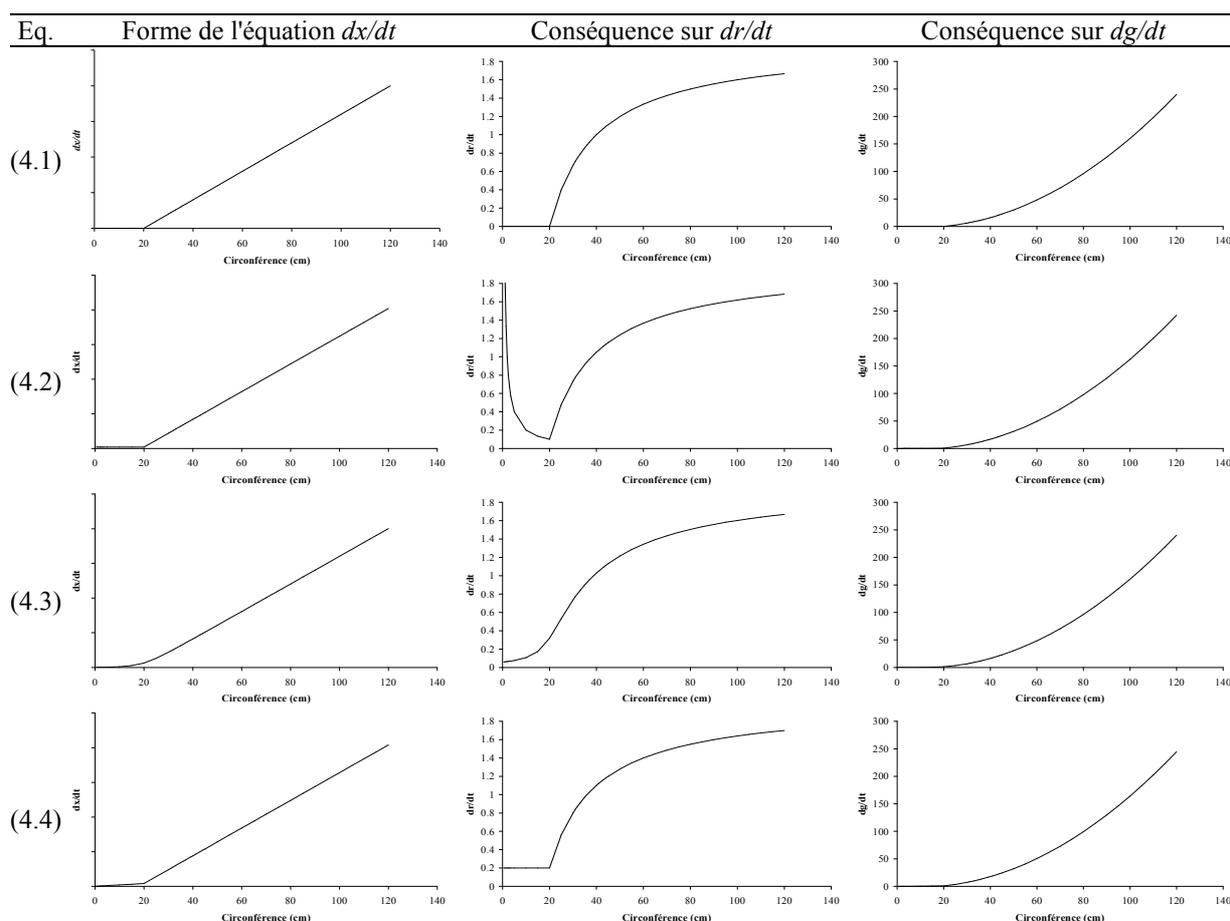


Figure 4.2 : L'accroissement en surface terrière ou en rayon est modélisé avec l'une des 4 équations 4.1 à 4.4 (figures de gauche). Au milieu : conséquence du choix du modèle d'accroissement en surface terrière sur la forme de l'accroissement en rayon. A droite : conséquence du choix du modèle d'accroissement en rayon sur la forme de l'accroissement en surface terrière. Eq. = équation

En conclusion, le choix du modèle d'accroissement en surface terrière peut avoir un impact important sur la forme des accroissements en rayon. Cet impact n'est pas toujours cohérent avec le phénomène biologique étudié. Les modèles d'accroissement en surface terrière les plus cohérents avec les données d'accroissement en rayon (voir figure 4.1) sont les équations 4.1 et 4.3. Les équations 4.2 et 4.4 sont à rejeter. Le modèle hyperbolique est certainement celui qui est le plus satisfaisant mais l'observation des données d'accroissement en rayon ne justifie pas vraiment l'utilisation de ce modèle. Dans les modèles que nous avons testés, le choix du modèle d'accroissement en rayon a assez peu d'impact sur la forme des accroissements en surface terrière. De plus, l'accroissement en surface terrière obtenu à partir de ces modèles est cohérent avec les données observées.

#### 4.1.1.3 Bilan et modèle retenu

D'un point de vue statistique, l'observation des données fait apparaître pour chacune des variables à expliquer (accroissement en surface terrière et accroissement en rayon) des avantages et des inconvénients. La relation entre l'accroissement en surface terrière et la circonférence est plus nette mais la variance n'est pas homogène<sup>8</sup> et le lien n'est pas linéaire. La relation entre l'accroissement en rayon et la circonférence est plus linéaire mais la variabilité est plus importante.

Nous avons également vu qu'il existe un lien entre l'accroissement en surface terrière et l'accroissement en rayon (équation 4.5). Ainsi on peut considérer que l'utilisation de l'accroissement en surface terrière dans un modèle prédictif induit une augmentation "artificielle" de la liaison entre la variable dépendante et une des variables explicatives potentielles, la circonférence (Lejeune, 1996). Les modèles basés sur l'accroissement en surface terrière obtiennent donc des indices de performance (comme le R<sup>2</sup>) meilleurs que ceux basés sur l'accroissement en rayon (West, 1980 ; Stadt *et al.*, 2002).

L'étude sur le choix de la forme du modèle que nous avons faite au paragraphe précédent, montre que le choix de la forme du modèle d'accroissement en surface terrière, peut avoir un impact important sur la forme de l'accroissement en rayon. Utiliser un modèle d'accroissement en rayon (équations 4.1 à 4.4) nous paraît plus cohérent par rapport au phénomène biologique étudié et observé. Lorsqu'on observe l'accroissement en rayon en fonction de la circonférence (figure 4.1), nous n'avons pas de données en dessous d'un seuil théorique de circonférence (accroissements nuls ou très faibles). La droite correspondant à une croissance non nulle de l'équation 4.1 est donc suffisante par rapport à nos données.

Ces différents éléments nous conduisent donc à retenir l'accroissement en rayon pour exprimer la croissance individuelle des arbres et à choisir une droite pour décrire la relation entre cet accroissement et la circonférence (équation 4.6). Ce modèle (MAID) a été ajusté séparément pour le chêne et pour le pin par la méthode des moindres carrés ordinaire en utilisant le jeu de données de calibration présenté dans la partie 3.

$$\Delta r_{k,s,i} = \alpha_{k,s} + \beta_{k,s} \text{circ}_{k,s,i} + \varepsilon_{k,s,i} \quad (4.6)$$

avec  $\Delta r_{k,s,i}$  est l'accroissement en rayon (mm) sur 6 ans (période 2000-2005) de l'arbre  $i$  appartenant au dispositif  $k$  et à l'espèce  $s$ ,  $\text{circ}_{k,s,i}$  est la circonférence (cm) en 1999 de l'arbre  $i$ ,  $\varepsilon_{k,s,i}$  est l'erreur résiduelle,  $\alpha_{k,s}$  et  $\beta_{k,s}$  sont les paramètres du modèle pour le dispositif  $k$  et

<sup>8</sup> Le modèle de croissance en surface terrière génère automatiquement de l'hétéroscédasticité sur les résidus lorsque l'on suppose un résidu additif de variance constante dans le modèle Bontemps (2002).

l'espèce *s*. L'effet dispositif permet de prendre en compte les effets des facteurs agissant au niveau du peuplement, comme les effets stationnels, la densité totale ou le stade de développement (peuplement âgé ou jeune). Pour tester l'influence de l'effet dispositif sur la croissance, nous avons comparé 4 modèles : le modèle complet donné par l'équation 4.6, un modèle avec l'effet dispositif uniquement sur le coefficient  $\alpha$ , un modèle avec l'effet dispositif uniquement sur le coefficient  $\beta$  et un modèle sans effet dispositif. Comme critère de sélection, nous avons utilisé la statistique F qui compare les carrés moyens de deux modèles en fonction de leurs degrés de liberté. Lorsque deux modèles avaient les mêmes degrés de liberté, celui ayant la somme des carrés résiduels la plus faible a été choisi.

#### 4.1.2 Résultats de l'effet taille de l'arbre sur la croissance

Pour les deux essences, le modèle le plus parcimonieux est celui pour lequel l'effet dispositif est uniquement sur le coefficient  $\beta$  (tableau 4.1). Ce modèle explique 43.3% de la variabilité totale pour le chêne contre 15.5% pour le pin.

Tableau 4.1 : Tableau de l'ANOVA, estimation des paramètres et statistiques du modèle arbre indépendant des distances (équation 4.6). SC = somme des carrés ; d.d.l. = degré de liberté.

Source	Chêne				Pin			
	SC	d.d.l.	F	Pr>F	SC	d.d.l.	F	Pr>F
Constante	1.75	1	0.126	0.722	23.6	1	2.116	0.147
<i>circ</i> 1999	482.65	1	34.941	<.0001	332.4	1	29.808	<.0001
<i>circ</i> 1999:dispositif	434.25	8	3.929	<.0001	310.7	8	3.482	<.0001
Résiduelle	1989.06	144			1884.9	169		

Coefficients	Chêne		Pin	
	Estimation	Ecart-type ( $\sigma$ )	Estimation	Ecart-type ( $\sigma$ )
$\alpha_s$	0.3075	0.8651	1.7717	1.2178
$\beta_{k,s}$	0.0984 < $\beta_{k,chêne}$ < 0.1793		0.0396 < $\beta_{k,pin}$ < 0.0982	
	0.0169 < $\sigma$ < 0.0199		0.0088 < $\sigma$ < 0.0125	

<b>R<sup>2</sup> ajusté</b>	<b>0.433</b>	<b>0.155</b>
<b>Ecart-type résiduel</b>	<b>3.717</b>	<b>3.340</b>

Comme cela était attendu, les résultats montrent un effet positif de la taille de l'arbre sur la croissance avec une valeur pour le paramètre  $\beta$  comprise entre 0.0984 et 0.1793 mm/cm pour le chêne et entre 0.0396 et 0.0982 mm/cm pour le pin (tableau 4.2). Quels que soient les dispositifs, à taille égale, les accroissements en rayon des chênes sont supérieurs à ceux des pins. Ces résultats sont bien sûr valables dans le cadre de nos données en particulier pour la gamme d'âge échantillonnée (voir partie 3).

Tableau 4.2 : Estimation des paramètres du modèle arbre indépendant des distances pour chaque dispositif.

<b>Dispositif</b>	$\alpha_{chêne}$	$\beta_{k.chêne}$	$\alpha_{pin}$	$\beta_{k.pin}$
D02	0.3075	0.1026	1.7717	0.06833
D108	0.3075	0.1507	1.7717	0.06583
D20	0.3075	0.1150	1.7717	0.06793
D27	0.3075	0.0984	1.7717	0.05829
D42	0.3075	0.1793	1.7717	0.09825
D49	0.3075	0.1058	1.7717	0.05508
D534	0.3075	0.1426	1.7717	0.05246
D563	0.3075	0.1227	1.7717	0.05624
D78	0.3075	0.1457	1.7717	0.03961

### 4.1.3 L'hypothèse reliant le potentiel de croissance et la taille est-elle pertinente ?

Sur la période de croissance 2000-2005, nous avons établi un lien entre la croissance individuelle et la taille des arbres. Ce résultat est classique que ce soit dans les peuplements purs ou mélangés (voir par exemple Wykoff, 1990 ; Monserud et Sterba, 1996). Dans notre étude, la relation est beaucoup plus forte pour le chêne ( $R^2 = 43.3\%$ ) que pour le pin (15.5%). Ce résultat est cohérent avec d'autres études comme celle de Monserud et Sterba (1996). Avec leur modèle arbre indépendant des distances, l'effet taille explique 42.2% de la variabilité pour le chêne (*Quercus robur*, *Quercus petraea* et *Quercus cerris*) et seulement 24.5% pour le pin sylvestre. D'autres études en revanche trouvent des  $R^2$  similaires pour le pin sylvestre et les autres espèces en mélange avec lui : *Picea abies* (Pukkala *et al.*, 1998) ; *Picea abies* et autres espèces feuillues (Andreassen et Tomter, 2003) ; *Pinus nigra* (Trasobares *et al.*, 2004). Plusieurs explications peuvent être avancées pour expliquer la faible relation entre l'accroissement et la taille pour le pin :

- il peut y avoir des différences micro-stationnelles à l'intérieur du dispositif qui masquent la relation entre la taille et l'accroissement ;
- il peut y avoir des différences génétiques forte entre les individus liées à différentes phases de colonisation (Andreassen et Tomter, 2003) ou à l'introduction de provenances d'origines différentes (Valadon et Voccia, 2006) ;
- en peuplement régulier, l'effet de la taille sur la croissance a été observé lorsqu'une compétition intense pour la lumière se met en place (Bi et Turvey, 1996). Or pour la période de croissance considérée, la compétition pour la lumière n'est peut-être pas le facteur limitant pour lequel les pins entrent en compétition (Garcia-Abril *et al.*, 2007) ;
- ce résultat peut être lié au tempérament de l'espèce. Selon Jogiste (2000), la taille et l'âge sont des bons prédicteurs de la croissance pour les espèces tolérantes à l'ombrage ce qui n'est pas le cas du pin sylvestre. Cette explication peut être en lien avec l'explication précédente car selon Stadt *et al.* (2007) dans un peuplement mélangé les arbres survivants des espèces intolérantes à l'ombre tendent à occuper une position dominante au niveau de la canopée et sont donc moins concernés par la compétition pour la lumière ;
- les pins ont peut-être atteint un stade de développement qui les rend moins compétitifs. Ils sont peut-être en phase de sénescence (Garcia-Abril *et al.*, 2007) et poussent de la même façon quelle que soit leur taille ;



#### 4.1.4 A quoi peut-on attribuer l'effet dispositif ?

Nous avons mis en évidence que la valeur du paramètre  $\beta$  du modèle arbre indépendant des distances dépendait du dispositif (tableau 4.1 et tableau 4.2). Le modèle peut donc s'appliquer uniquement sur les 9 dispositifs utilisés pour le construire. Pour rendre le modèle plus général, il faut pouvoir exprimer l'effet dispositif en fonction d'autres variables plus générales. Cet effet dispositif peut être dû à différents facteurs quantitatifs. Les études sur le chêne et le pin en peuplement pur montrent que ce paramètre  $\beta$  dépend de la fertilité, de l'âge et de la surface terrière totale (Dhôte *et al.*, 2000 ; Perot *et al.*, 2007). Ainsi, pour ce qui concerne nos dispositifs, on constate qu'il y a une relation entre l'âge des populations (âge estimé sur les arbres échantillonnés) et  $\beta$  (figure 4.4). En revanche on n'observe pas de relation avec la surface terrière totale (figure 4.4).

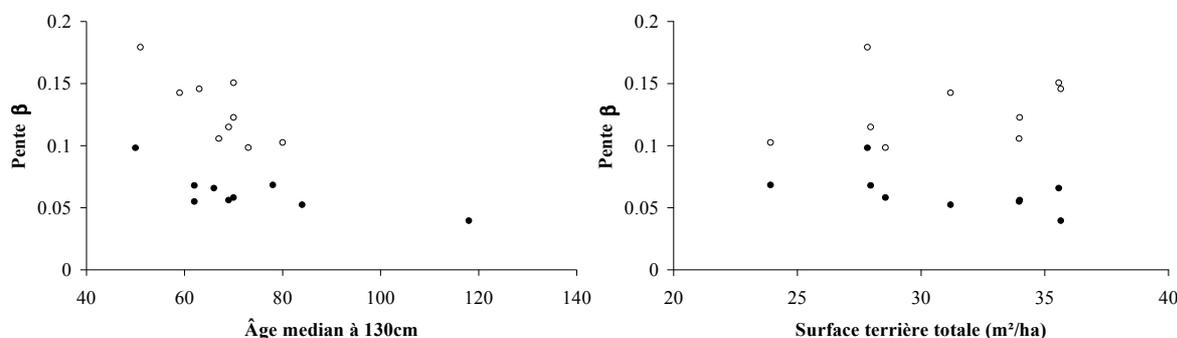


Figure 4.4 : A gauche, relation entre la pente  $\beta$  du modèle arbre indépendant des distances et l'âge. A droite relation entre la pente  $\beta$  et la surface terrière totale. Les points blancs sont les valeurs de la pente pour le chêne. Les points noirs sont les valeurs de la pente pour le pin. La surface terrière est calculée sur la zone échantillonnée.

La relation entre  $\beta$  et la surface terrière des chênes ou la surface terrière des pins est faible (figure 4.5). Dans le cas du chêne, on peut noter une relation légèrement positive entre la pente et la surface terrière des pins.

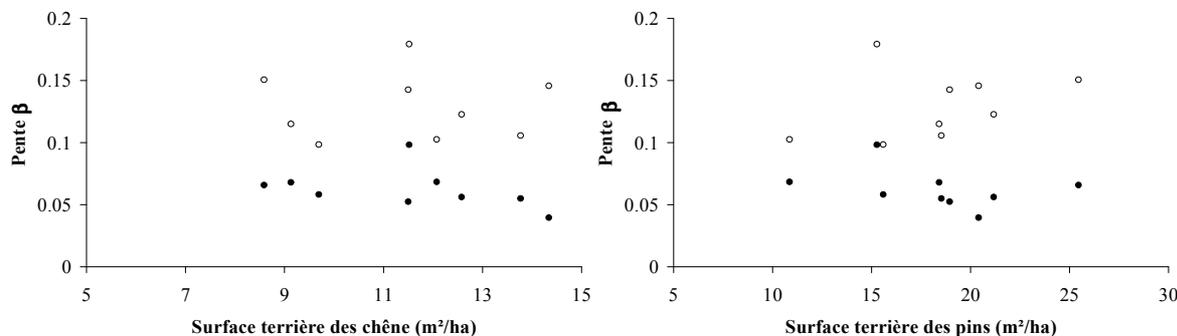


Figure 4.5 : A gauche, relation entre la pente  $\beta$  du modèle arbre indépendant des distances et la surface terrière des chênes. A droite relation entre la pente  $\beta$  et la surface terrière des pins. Les points blancs sont les valeurs de la pente pour le chêne. Les points noirs sont les valeurs de la pente pour le pin. Les surfaces terrières sont calculées sur la zone échantillonnée.

Dans la partie 3 (paragraphe 3.4.3), nous avons estimé des indices de fertilité pour chaque dispositif et pour chaque essence. Nous avons aussi présenté les limites de l'utilisation de ces indices dans le cadre de peuplements mélangés. Nous pouvons constater qu'il existe une relation positive entre  $\beta$  et l'indice de fertilité bien que celle-ci ne soit pas très prononcée (figure 4.6).

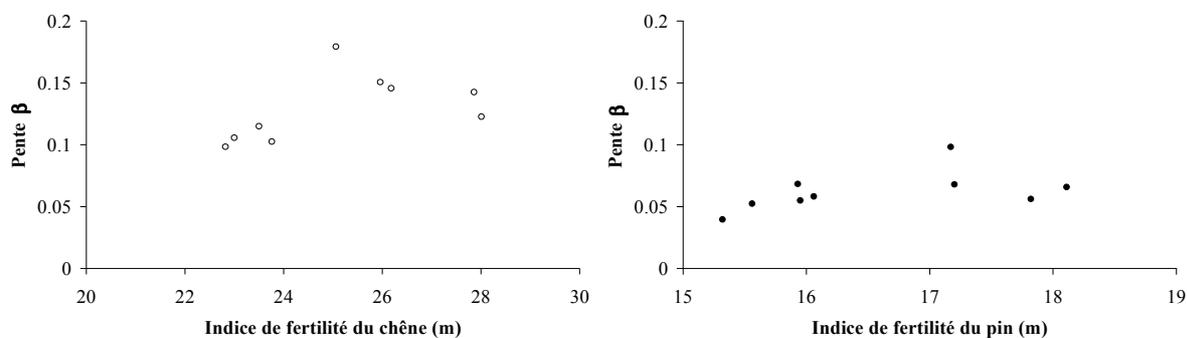


Figure 4.6 : A gauche, relation entre la pente  $\beta$  du modèle chêne arbre indépendant des distances et l'indice de fertilité des chênes (hauteur dominante à 100 ans). A droite relation entre la pente  $\beta$  du pin et l'indice de fertilité des pins (hauteur dominante à 50 ans). Les points blancs sont les valeurs de la pente pour le chêne. Les points noirs sont les valeurs de la pente pour le pin.

Dans ces résultats, c'est la relation entre l'âge et le paramètre  $\beta$  qui est la plus marquée. Plus les populations sont âgées, plus le paramètre diminue. Cependant, il est difficile de tirer des conclusions claires de ces observations car à l'échelle du dispositif, beaucoup de facteurs peuvent être confondus. De plus, nous avons considéré les facteurs séparément ce qui ne permet pas de mettre en évidence d'éventuelles interactions. Pour aller plus loin, il serait intéressant d'introduire les variables que nous avons testées directement dans l'équation 4.6 pour voir si elles peuvent expliquer l'effet dispositif obtenu sur le paramètre  $\beta$ . Cependant, étant donné le faible nombre de dispositifs et l'objectif de cette thèse, nous avons choisi de conserver le facteur dispositif dans nos modèles.

#### 4.1.5 Les paramètres de croissance évoluent au cours du temps

Le modèle arbre indépendant des distances (MAID) a été construit sur la période de croissance 2000-2005. Comme nous venons de le voir, les paramètres du modèle peuvent dépendre de facteurs qui évoluent au cours de temps (la surface terrière, l'âge). L'étude rétrospective de la croissance des arbres donne la possibilité d'étudier l'évolution des paramètres du modèle au cours du temps. Pour cela, nous avons ajusté le modèle de l'équation 4.6 sur des périodes de croissance de 6 ans en utilisant la fonction *lmList* du package *nlme* du logiciel R (Pinheiro *et al.*, 2008). Chaque période de croissance contient au moins 10 arbres d'une espèce. Nous avons commencé par ajuster le modèle avec uniquement le paramètre  $\alpha$  ce qui revient à calculer la moyenne des accroissements par dispositif, par essence et par période de croissance. Nous avons ensuite ajusté le modèle complet avec les paramètres  $\alpha$  et  $\beta$ . Comme nous le verrons par la suite, nous avons choisi de ne pas intégrer l'évolution des paramètres au cours du temps dans le modèle arbre indépendant des distances qui sera évalué dans la partie 5. Pour ces analyses, nous avons donc utilisé l'ensemble des arbres disponibles (jeu de données de calibration et jeu de données d'évaluation présentés dans la partie 3).

Pour 7 des 9 dispositifs, les pins et les chênes ont pratiquement le même âge à 1.30 m (tableau 4.3). Pour ces dispositifs, comme le montre la figure 4.7, une première différence apparaît entre le chêne et le pin au niveau de l'accroissement dans le jeune âge. Dans le jeune âge, les pins ont un accroissement beaucoup plus fort que les chênes. L'autre différence concerne l'évolution de l'accroissement en rayon avec le temps. Pour les pins, l'accroissement diminue avec le temps. Cette diminution est très forte dans le jeune âge. Les chênes, sur les périodes considérées, ont un accroissement sur le rayon qui reste stable dans le temps. Les accroissements en rayon des chênes et des pins se rejoignent plus ou moins rapidement selon les dispositifs.

Tableau 4.3 : Age médian à 1.30 m des chênes et des pins dans les 9 dispositifs.

Dispositif	Age médian à 1.30 m	
	Chêne	Pin
D02	80	78
D108	70	66
D20	69	62
D27	73	70
D42	51	50
D49	67	62
D534	59	84
D563	70	69
D78	63	118

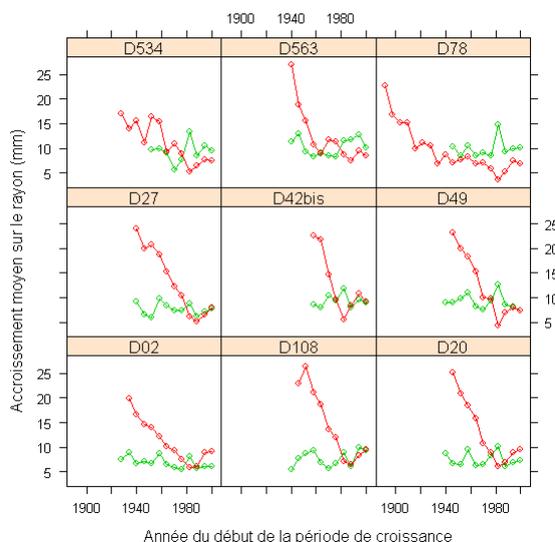


Figure 4.7 : Accroissement moyen en rayon sur 6 ans (mm) en fonction de la première année de la période de croissance pour chaque dispositif et pour chaque espèce. Le chêne est en vert et le pin en rouge.

Si on regarde maintenant l'évolution des paramètres  $\alpha$  et  $\beta$  (figure 4.8), on retrouve pour le pin les mêmes tendances sur le paramètre  $\alpha$  que sur l'accroissement moyen : de fortes valeurs dans le jeune âge puis une diminution au cours du temps. Pour le chêne, on constate également une diminution du paramètre  $\alpha$  au cours du temps mais elle est en général plus modérée que pour le pin. Concernant le paramètre  $\beta$ , il est plus difficile de donner un résultat général, les variations dépendant fortement du dispositif.

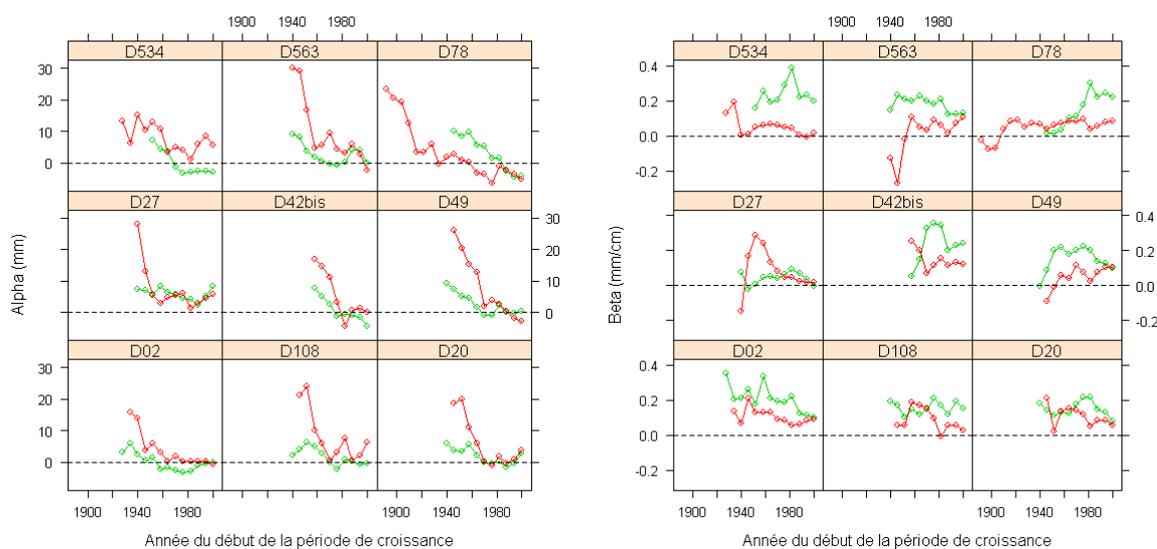


Figure 4.8 : Evolution des paramètres  $\alpha$  et  $\beta$ . Le chêne est en vert et le pin en rouge.

Nous avons vu que les deux essences ont des croissances très différentes, en particulier dans le jeune âge. Ces résultats sont comparables à ceux obtenus pour des peuplements purs : la diminution de l'accroissement en rayon avec l'âge est plus marquée pour le pin que pour le chêne (Lebourgeois, 1997 pages 41 et 73). Dans des peuplements spontanés de pin sylvestre dans le Massif Central, Prevosto (1999) montre que la tendance biologique de la croissance radiale avec l'âge n'est pas affectée par le statut compétitif pour les arbres en peuplement. La largeur de cerne est maximale à 3 ans, puis la croissance diminue, d'abord rapidement puis plus régulièrement et se stabilise entre 30 et 35 ans. Ces résultats sont également cohérents avec les connaissances sur l'écologie de ces espèces. Le pin est une espèce pionnière colonisatrice (Prevosto, 1999) alors que le chêne est une espèce postpionnière caractérisée par une grande longévité (Rameau *et al.*, 1989). Cette différence de comportement semble bien prise en compte par le paramètre  $\alpha$  qui diminue avec le temps de façon plus ou moins forte selon l'espèce. Ce paramètre est donc lié à l'âge de la population ce qui rejoint les résultats obtenus pour le chêne en peuplement régulier (Dhôte, 1999).

L'évolution du paramètre  $\beta$  est plus difficile à interpréter. Cette difficulté vient en partie du fait que la valeur de ce paramètre dépend de l'intensité de la compétition entre les individus. L'apparition de cette compétition se traduit par une hiérarchisation du peuplement : des arbres dominés et des arbres dominants apparaissent. Ce résultat semble bien s'appliquer à certains de nos dispositifs. Ainsi, la hiérarchisation est très visible dans le cas du dispositif D42 aussi bien pour le chêne que pour le pin (figure 4.9). Dans ce dispositif, les chênes et les pins ont le même âge à 1.30 m (environ 50 ans, voir tableau 4.3) et se sont certainement installés en même temps à la faveur d'une grande ouverture (coupe rase ou incendie). Dans ce cas, on peut penser que dans les premiers stades de développement du peuplement, la lumière n'était pas un facteur limitant. Les individus poussent alors de manière identique ce qui se traduit par une faible différenciation entre individus (pour la première période de croissance de la figure 4.9 les circonférences sont très regroupées pour les deux essences). A mesure que les individus grandissent, la compétition pour la lumière s'accroît, la domination des arbres les plus gros sur les plus petits s'accroît, les circonférences s'écartent et la relation entre l'accroissement et la taille est bien marquée : il y a hiérarchisation dans le peuplement. Si on regarde l'évolution de la circonférence au cours du temps, cela correspond à des courbes qui sont d'abord relativement parallèles puis qui divergent (voir figure 4.9 partie droite). Logiquement, cette hiérarchisation doit être plus marquée chez le chêne que chez le pin qui est connu pour être moins tolérant à l'ombrage. Pour cette raison, comme nous l'avons dit au paragraphe 4.1.3, la taille est considérée comme un bon prédicteur de la croissance pour les

espèces tolérante à l'ombrage (Jogiste, 2000). Dans le cas du dispositif D42, on observe soit une stabilisation soit une diminution du paramètre  $\beta$  avec l'âge.

Le paramètre  $\beta$  peut être fortement influencé par des facteurs climatiques ou biotiques variant d'une période de croissance à l'autre (Dhôte, 1999). Les opérations sylvicoles, en modifiant la compétition entre les individus, doivent également avoir une influence sur les valeurs de ce paramètre. Nous voyons ici les limites de notre approche car, comme nous n'avons pas de données historiques sur les caractéristiques des peuplements, nous ne pouvons pas étudier l'influence de la sylviculture sur ce paramètre.

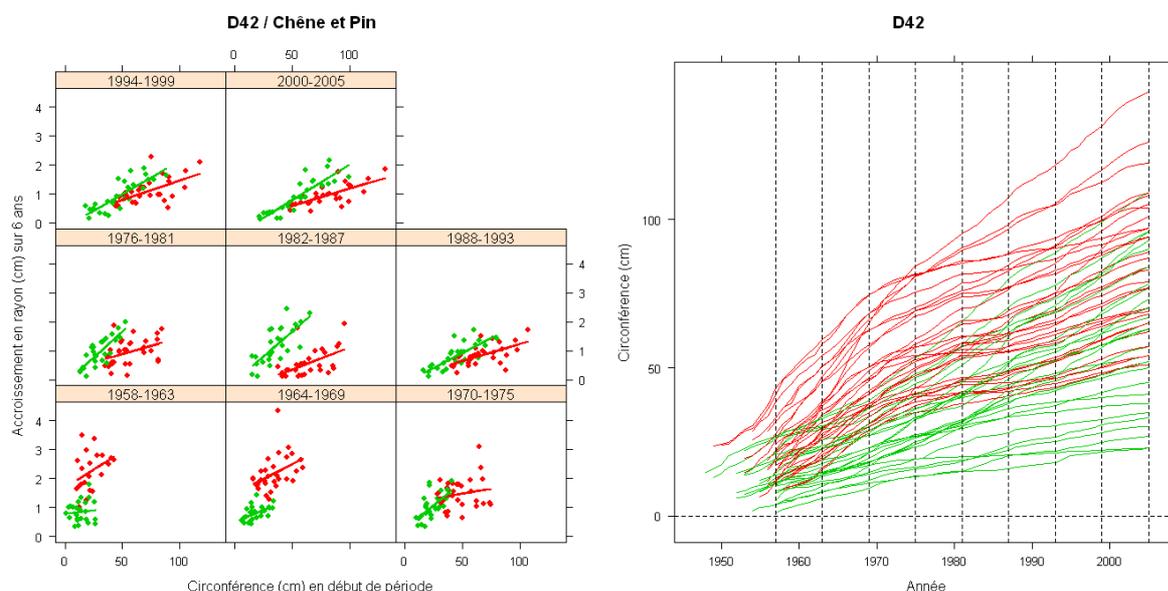


Figure 4.9 : A gauche, accroissement en rayon en fonction de la circonférence pour le dispositif D42 sur 8 périodes de croissance de 6 ans allant de la période 1958-1963 à la période 2000-2005. Le chêne est en vert et le pin en rouge. A droite, circonférence à 1.30 m en fonction du temps pour tous les arbres du dispositif D42. Les lignes verticales en pointillés délimitent les périodes de croissance représentées dans la figure de gauche. Les courbes vertes sont des chênes et les courbes rouges des pins.

Les résultats que nous obtenons ici ne sont pas contradictoires avec ceux du paragraphe 4.1.4 (A quoi peut-on attribuer l'effet dispositif ?) dans lequel nous avons montré que le paramètre  $\beta$  diminuait avec l'âge des populations. En effet, dans le modèle arbre indépendant des distances que nous avons construit, le paramètre  $\alpha$  est le même pour tous les dispositifs. La variabilité liée à l'effet dispositif est donc entièrement reportée sur le paramètre  $\beta$ . L'effet de l'âge est probablement également reporté sur ce paramètre.

Avec les données actuelles, nous pourrions reproduire les tendances à long terme liées à l'âge sur chaque dispositif. On pourrait pour cela utiliser des modèles à effets mixtes (Pinheiro et Bates, 2000) qui permettraient de prendre en compte les différents niveaux de notre échantillonnage (dispositif, année, arbre) et de tester la meilleure façon d'intégrer l'effet temporel. Cependant, nous pouvons difficilement intégrer de manière déterministe les fluctuations d'une période de croissance à l'autre qui peuvent être importantes et dépendantes de la sylviculture. L'évolution des paramètres au cours du temps ne sera donc pas prise en compte dans les modèles que nous évaluerons et comparerons par la suite. Cette prise en compte est une étape indispensable pour garantir un comportement cohérent des modèles dans les simulations à long terme. Ce travail fera donc l'objet de perspectives à la suite de cette thèse.

## 4.1.6 Le chêne et le pin ne poussent pas de manière indépendante : mise en évidence d'une compensation de croissance

### 4.1.6.1 Introduction

Dans le début des années 1980, le lophyre du pin (*Diprion pini* L.) un hyménoptère phytophage spécifique du pin sylvestre a fait de très gros dégâts de la Sologne jusqu'au bassin parisien (Geri et Goussard, 1984). Dans le massif de Lorris, les attaques massives sont survenues essentiellement en 1981 (figure 4.10). Des défeuillaisons totales ont été observées à l'automne 1981 sur le massif de Lorris. La gradation a été violente mais très brève puisqu'à l'automne 1982 la présence de l'insecte était pratiquement nulle. Ce schéma est caractéristique de l'espèce dans les plaines de l'Europe atlantique (Geri et Goussard, 1984). D'après l'étude réalisée par Geri et Goussard (1984), lors de cette gradation, les peuplements où le pin sylvestre était en mélange avec des feuillus ou d'autres conifères n'ont pas été moins attaqués que les peuplements purs de pin sylvestre. De plus, l'espèce attaque aussi bien les jeunes que les vieux peuplements.

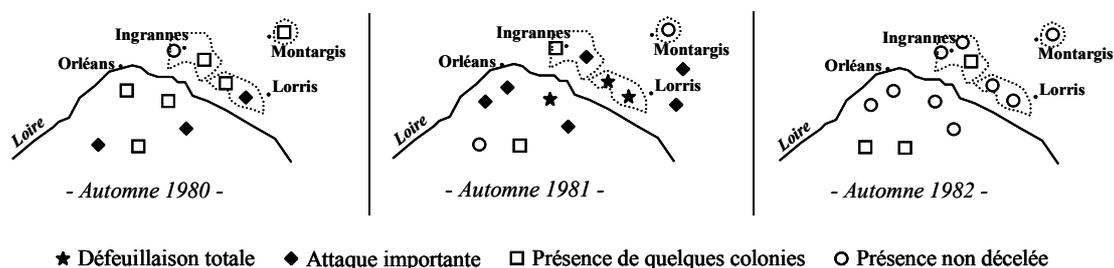


Figure 4.10 : Evolution des populations de *Diprion pini* dans le secteur de la forêt domaniale de Lorris d'après Geri et Goussard (1984).

La gradation dont nous venons de parler a eu lieu environ 35 ans avant la récolte de nos données. Ceci signifie que les pins sylvestres de notre échantillon sont tous contemporains de cet événement et qu'ils ont survécu aux attaques de l'insecte. Les défeuillaisons ont eu lieu principalement à l'automne 1981, on peut donc supposer que les arbres touchés ont eu une croissance plus faible en 1982. Parallèlement, les chênes ont probablement été favorisés par cet événement car après les défeuillaisons, la compétition exercée par les pins pour les ressources aériennes et souterraines devait être beaucoup plus faible.

L'objectif de cette sous-partie est de montrer comment l'attaque d'un insecte sur le pin se traduit sur les paramètres de croissance du pin mais également indirectement sur ceux du chêne.

#### 4.1.6.2 Méthode

Pour observer l'évolution conjointe de la croissance du chêne et du pin, nous avons analysé l'évolution des paramètres du modèle de croissance au cours du temps. Contrairement paragraphe 4.1.5, nous avons pris un pas de temps annuel pour pouvoir observer les variations des paramètres d'une année sur l'autre. Nous avons retenu la période s'étalant de 1970 à 1995 de façon à couvrir largement la gradation du lophyre. Nous avons fait une première tentative en ajustant le modèle de l'équation 4.6 pour chaque année et chaque dispositif. Pour cela nous avons utilisé la fonction *lmList* du package *nlme* du logiciel R (Pinheiro *et al.*, 2008). Cependant, l'interprétation des résultats était difficile car les deux paramètres de l'équation étaient concernés par les variations annuelles. Pour simplifier l'analyse, nous avons choisi de fixer l'un des paramètres. Pour choisir entre  $\alpha$  et  $\beta$ , nous avons représenté les paramètres ainsi que leur intervalle de confiance pour chaque année et chaque dispositif. Pour la période considérée, l'analyse montre que c'est le paramètre  $\alpha$  qui a les intervalles de confiance qui se recouvrent le plus. Nous avons donc choisi de fixer  $\alpha$  plutôt que  $\beta$  dans le modèle (Pinheiro et Bates, 2000 page 143). De plus, la valeur de  $\alpha$  est proche de zéro pour les deux essences (pour le chêne : moyenne de  $\alpha = -0.005$ , écart-type de  $\alpha = 0.045$  ; pour le pin : moyenne de  $\alpha = 0.021$ , écart-type de  $\alpha = 0.066$ ). Dans un deuxième temps, nous avons donc ajusté le modèle en fixant la valeur du paramètre  $\alpha$  à zéro. De cette façon nous pouvons observer les changements au cours du temps avec un seul paramètre. Nous avons ensuite calculé pour chaque année le paramètre  $\beta$  relatif ( $\beta_{cart_{s,k,n}}$ ) qui permet de comparer plus facilement le comportement des deux espèces sur un même dispositif :

$$\beta_{cart_{s,k,n}} = \frac{\beta_{s,k,n} - \bar{\beta}_{s,k}}{\bar{\beta}_{s,k}}$$

avec  $\beta_{s,k,n}$  l'accroissement par cm de circonférence de l'année  $n$  pour les arbres de l'espèce  $s$  appartenant au dispositif  $k$  et  $\bar{\beta}_{s,k}$  le paramètre  $\beta$  moyen sur l'ensemble de la période allant de 1970 à 1995 pour les arbres de l'espèce  $s$  appartenant au dispositif  $k$ . Nous avons ensuite représenté l'évolution de  $\beta_{cart_{s,k,n}}$  pour chaque essence et pour chaque dispositif. Pour avoir des résultats quantitatifs plus parlants, nous avons également calculé l'accroissement moyen sur le rayon des arbres échantillonnés pour chaque essence et chaque année sur la période 1970 à 1995.

### 4.1.6.3 Résultats

Si on compare les valeurs de  $\beta$  des deux essences avant, pendant et après la gradation, nous observons le schéma suivant (figure 4.11 et figure 4.12 partie gauche) : avant l'année 1982, les courbes d'évolution relative de  $\beta$  pour le chêne et pour le pin sont relativement bien synchronisées. A partir de 1982, les courbes se désynchronisent. Elles sont à nouveau synchronisées après trois ou quatre années de croissance. L'année 1982 correspond pour le pin à une baisse importante de  $\beta$ . Elle est comprise entre -45% (dispositif D02) et -92% (dispositif D534) par rapport à la valeur du paramètre  $\beta$  en 1981. Dans le même temps la valeur de  $\beta$  pour le chêne augmente. Cette augmentation est comprise entre 2% (dispositif D27) et 41% (dispositif D534). Après l'année 1982, la valeur  $\beta$  du pin augmente progressivement alors que celle du chêne baisse progressivement. La figure 4.12 présente l'exemple du dispositif D534 sur la période 1970-1995 avec les accroissements de tous les arbres.

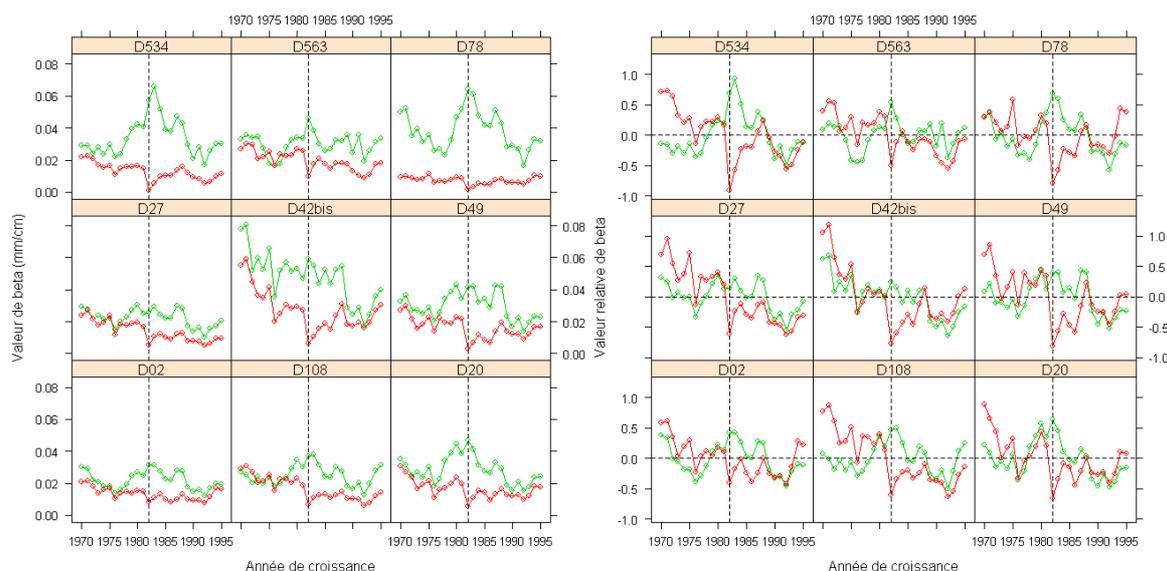


Figure 4.11 : A gauche, évolution de l'accroissement en rayon par cm de circonférence (paramètre  $\beta$ ) sur la période 1970-1995. A droite évolution de la valeur relative de  $\beta$ . La valeur moyenne de la pente  $\beta_{moy}$  a été calculée sur la période 1970-1995. L'écart relatif par rapport à la valeur moyenne est égal à  $(\beta_i - \beta_{moy}) / \beta_{moy}$ . La ligne en pointillés horizontale correspond à la valeur 0. La ligne en pointillés verticale matérialise l'année 1982. Le vert représente les données concernant le chêne et le rouge les données concernant le pin.

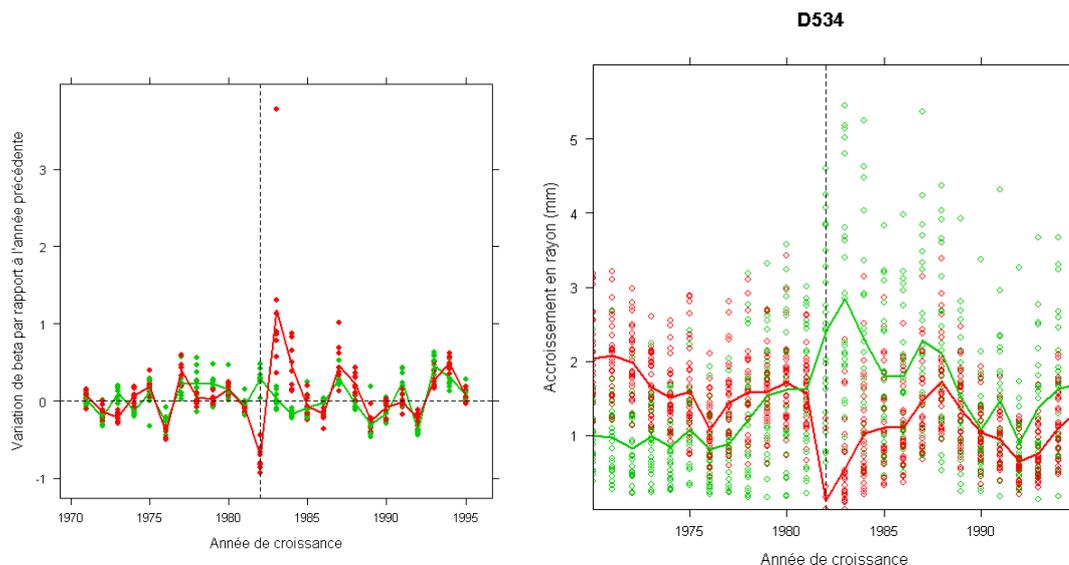


Figure 4.12 : A gauche, taux de variation du paramètre  $\beta$  par rapport à l'année précédente. Les lignes brisées correspondent aux moyennes par année sur l'ensemble des dispositifs. Le vert représente les données concernant le chêne et le rouge les données concernant le pin. A droite, accroissement en rayon pour chaque arbre et pour chaque année dans le dispositif D534. Les lignes brisées représentent les courbes moyennes pour chaque espèce. Les chênes sont en vert et les pins en rouge.

#### 4.1.6.4 Discussion

L'impact de l'attaque du lophyre sur la croissance du pin est très important. Dans notre cas, l'accroissement moyen sur le rayon des arbres échantillonnés diminue de -44% à -93% selon les dispositifs. En peuplement pur, cet insecte peut réduire la croissance en volume de 86% à 94% selon l'intensité des dégâts (Lyytikainen-Saarenmaa et Tomppo, 2002). De plus, la mortalité dans les peuplements défoliés atteint environ 30% après une gradation de l'insecte (Lyytikainen-Saarenmaa et Tomppo, 2002). Cependant, aucun seuil de défoliation au delà duquel il résulterait une mort inévitable du pin sylvestre n'a été mis en évidence (Augustaitis, 2007). Dans nos dispositifs, les arbres survivant semblent réagir bien et rapidement mais, d'après Langstrom *et al.* (2001), le retour à une croissance normale peut être très long. Nous pouvons donc faire plusieurs hypothèses sur la croissance des pins après l'attaque : a) les pins sont affaiblis, ils poussent moins bien qu'avant l'attaque ; b) les pins retrouvent une croissance comparable à celle qu'ils auraient eu sans attaque ; c) les pins ont une croissance supérieure à celle qu'ils auraient eu sans l'attaque car ils bénéficient de la mortalité survenue sur une partie de la population. Nos données pourraient nous permettre de vérifier ces hypothèses à condition de prendre en compte l'effet de l'âge sur l'évolution des paramètres du modèle (voir paragraphe 4.1.5).

Nous avons vu que dans la période suivant la gradation du lophyre, les chênes ont eu une croissance plus soutenue. On pourrait penser que cette réaction est liée à une année climatique particulièrement favorable pour le chêne. Cependant, l'année 1982 n'est pas une année

caractéristique positive pour le chêne poussant en peuplement pur dans le Centre de la France (Lebourgeois, 1997 page 45). Les chênes ont donc très probablement été favorisés par cet épisode et ont compensé la perte de production du peuplement due à la faible croissance des pins. Des données en peuplements purs de chêne permettraient d'évaluer réellement l'importance de la réaction des chênes. De plus, il serait intéressant de mettre en relation l'intensité de la réponse du chêne avec l'intensité de la réaction du pin suite l'attaque du Lophyre. Ce cas montre qu'il y a une interaction entre la croissance des deux espèces et que le comportement du peuplement mélangé ne peut pas être assimilé à la juxtaposition de deux peuplements purs. Les exemples de compensation de croissance en forêt mélangée sont peu nombreux (Cruickshank *et al.*, 2009). Au Canada, dans des peuplements mélangés de douglas (*Pseudotsuga menziesii* var. *glauca*) et de pin (*Pinus contorta*) Heath et Alfaro (1990) ont montré que la croissance du douglas avait augmenté suite à l'attaque des pins par un scolyte (*Dendroctonus ponderosae*). Au québec, dans des peuplements mélangés de bouleaux jaunes et de résineux, une augmentation de la croissance des bouleaux a été observée suite à une épidémie de la tordeuse des bourgeons de l'épinette (Saucier, 2001). On peut supposer que des phénomènes de compensation existent en dehors de l'évènement exceptionnel que nous venons de décrire mais ils sont certainement plus difficiles à mettre en évidence. De façon générale, il y a une corrélation positive entre les variations de croissance des deux espèces. Cette corrélation positive est probablement liée aux variations climatiques qui en général ont globalement les mêmes effets sur la croissance des chênes et sur la croissance des pins. Cependant, comme nous l'avons vu pour certaines années, les deux espèces peuvent réagir de façon opposée. Il serait intéressant d'identifier les années où les deux espèces réagissent significativement de façon opposée et de rechercher les facteurs susceptibles d'induire ce type de comportement : conditions climatiques particulières, intervention sylvicole, attaque d'un agent biotique.

Ce travail souligne l'intérêt d'utiliser un modèle arbre indépendant des distances pour étudier les peuplements mélangés de façon retrospective. Evidemment, nous ne disposons que de mesures sur des arbres survivants. Néanmoins, on peut faire l'hypothèse que les arbres morts ou éclaircis ont assez peu d'impact sur la relation entre l'accroissement et la taille. Avec un échantillonnage adéquat, Le MAID permet donc d'obtenir des résultats valables à l'échelle du peuplement sans mesurer tous les arbres. Les paramètres du modèle que nous avons obtenus sont des données synthétiques du comportement de chaque essence sur une période donnée.

Ce travail fait également ressortir une limite du MAID que nous avons construit : l'effet d'une espèce sur la croissance de l'autre n'apparaît pas de façon explicite dans le modèle. Cet effet est néanmoins intégré de façon implicite dans les paramètres. En conséquence, les conditions (taux de mélange, compétition) pour utiliser le modèle de manière prédictive devront être aussi proches que possible des conditions dans lesquelles il a été construit.

## 4.2 *Le modèle arbre dépendant des distances*

### 4.2.1 **Prise en compte de l'influence du voisinage**

Dans un peuplement forestier, la croissance d'un arbre est réduite par la compétition exercée par ses voisins pour exploiter les ressources aériennes et souterraines. Dans un peuplement mélangé, la nature des espèces présentes dans le voisinage d'un arbre peut également jouer sur l'accession de cet arbre aux ressources (voir par exemple Uriarte *et al.*, 2004b ; Zhao *et al.*, 2006). Nous faisons donc l'hypothèse qu'une partie de la variabilité non expliquée par le modèle arbre indépendant des distances (équation 4.6) est liée à des différences de compétition locale ou à des différences de composition du voisinage. Nous faisons également l'hypothèse que l'intensité de la compétition peut être estimée par un indice de compétition local.

Les indices de compétition sont calculés à partir des caractéristiques dendrométriques des arbres appartenant au voisinage d'un arbre sujet (Biging et Dobbertin, 1992 ; Prévosto, 2005). Si l'on veut pouvoir utiliser le modèle de manière prédictive, ces indices doivent être calculés en début de période de croissance (année 1999). Cependant, seul un échantillon d'arbres du peuplement a été sondé (voir partie 3). Pour estimer la valeur des indices de compétition en 1999, nous avons reconstitué la circonférence 1999 de tous les arbres à partir des données de croissance obtenues sur les arbres échantillonnés. Pour chaque dispositif, nous avons utilisé une relation reliant la circonférence mesurée lors de l'inventaire (2005 ou 2006 suivant les dispositifs) à la circonférence 1999. Une certaine imprécision sur le voisinage 1999 est introduite ici mais nous avons considéré qu'elle n'avait pas de conséquences importantes sur les résultats concernant les indices de compétition. En effet, en l'absence d'éclaircie<sup>9</sup>, il y a une forte corrélation entre le voisinage en début et en fin de période de croissance.

Dans notre travail, le voisinage a été défini par un disque de rayon R et trois rayons de voisinage ont été testés (R=5, R=10 et R=15 m) de façon à couvrir la gamme des rayons obtenus dans d'autres études (Canham *et al.*, 2004 ; Uriarte *et al.*, 2004a ; Stadt *et al.*, 2007) et pour minimiser l'influence des corrections dues aux effets de bord dans le calcul des indices de compétition<sup>10</sup>. Nous avons utilisé le *Crowing Index* (Canham *et al.*, 2004) comme variable prédictive dans le modèle de croissance. Cet indice très général suppose que l'effet des arbres

<sup>9</sup> Dans le cas particulier où l'indice de compétition est basé sur le nombre d'arbres voisins il n'y a alors aucune conséquence car le nombre d'arbres en début et en fin de période est identique.

<sup>10</sup> Pour corriger les effets bords nous avons utilisé la méthode proposée par Goreaud et Pelissier (1999).

du voisinage sur la croissance d'un arbre sujet varie comme une fonction de la taille du voisin et comme une fonction inverse de la distance au voisin :

$$CI_{i,s,v} = \sum_{j=1}^{n_v} \frac{dbh_{j,v}^a}{l_{ij}^b}$$

où  $CI_{i,s,v}$  est l'indice de compétition de l'arbre sujet  $i$  appartenant à l'espèce  $s$  et calculé sur les voisins appartenant à l'espèce  $v$ ,  $n_v$  est le nombre d'arbres de l'espèce  $v$  dans le voisinage de l'arbre  $i$ ,  $dbh_{j,v}$  est le diamètre à 1.30 m d'un voisin  $j$  de l'espèce  $v$ ,  $l_{ij}$  est la distance entre l'arbre sujet  $i$  et le voisin  $j$ .  $a$  et  $b$  sont des coefficients qui déterminent respectivement la forme des effets de la taille du voisin et de la distance sur  $CI$  (Uriarte *et al.*, 2004b). Nous avons choisi cet indice car, en fonction de la valeur des coefficients  $a$  et  $b$ , il permet de retrouver une grande partie des indices de compétition couramment utilisés. Ainsi, nous avons testé un nombre restreint de valeurs pour  $a$  et  $b$  (équation 4.7) de façon à couvrir les indices classiques de compétition, comme la densité locale ( $a = 0$  et  $b = 0$ ) ou la surface terrière locale ( $a = 2$  et  $b = 0$ ), mais également des indices intermédiaires. Comme nous l'avons dit dans la présentation des données, quelques espèces feuillues autres que le chêne peuvent éventuellement être présentes dans les voisinages. Comme elles sont très minoritaires (tableau 3.2), elles ont été regroupées avec le chêne pour le calcul des indices de compétition.

$$a \in \{0; 1; 2\} \times b \in \{0; 0.5; 1\} \quad (4.7)$$

Les indices de compétition de chaque espèce ont été ajoutés dans l'équation 4.6 (Stadt *et al.*, 2007) pour former le modèle arbre dépendant des distances (MADD) :

$$\Delta r_{k,s,i} = \alpha'_{k,s} + \beta'_{k,s} circ_{k,s,i} + \lambda_{s,chêne} CI_{i,s,chêne} + \lambda_{s,pin} CI_{i,s,pin} + \epsilon'_{k,s,i} \quad (4.8)$$

où  $\lambda_{s,chêne}$  et  $\lambda_{s,pin}$  sont les coefficients associés respectivement à l'indice de compétition calculé sur les compétiteurs chênes et à l'indice de compétition calculé sur les compétiteurs pins ( $s$  est l'espèce de l'arbre sujet). Nous avons considéré que, pour la croissance d'une essence,  $a$  et  $b$  étaient identiques pour  $CI_{i,s,chêne}$  et pour  $CI_{i,s,pin}$ . Cela a permis de comparer les pouvoirs compétiteurs des deux espèces et donc d'étudier les compétitions intra et interspécifiques à travers les coefficients  $\lambda_{chêne}$  et  $\lambda_{pin}$ . Le modèle (équation 4.8) a été ajusté séparément pour les chênes et pour les pins par la méthode des moindres carrés ordinaires sur le même jeu de données que celui utilisé pour le MAID. Au total, 27 indices ont été testés (3 rayons  $\times$  9 indices de compétition). Pour sélectionner le meilleur indice de compétition, nous avons classé les modèles à l'aide du  $R^2$  ajusté (Stadt *et al.*, 2007). Comme pour le modèle arbre indépendant des distances, pour tester l'influence de l'effet dispositif sur le modèle, nous avons comparé le modèle complet (équation 4.8), un modèle avec l'effet dispositif uniquement

sur le coefficient  $\alpha$ , un modèle avec l'effet dispositif uniquement sur le coefficient  $\beta$  et un modèle sans effet dispositif.

#### 4.2.2 Résultats sur l'effet du voisinage sur la croissance

La figure 4.13 montre les coefficients de détermination ajustés ( $R^2$  ajusté) obtenus avec le modèle de l'équation 4.8 pour chaque indice de compétition et pour chaque essence. Les indices sont classés en fonction de la valeur du  $R^2$  ajusté. Pour le chêne, les 4 premiers indices du classement sont calculés dans un disque de rayon 10 m. Pour cette espèce, le rayon de voisinage qui donne les meilleurs résultats est donc 10 m. Pour le pin, le résultat est encore plus net et correspond également à un rayon de 10 m. Concernant le coefficient  $a$ , ce sont les valeurs 0 et 1 qui donnent les meilleurs résultats pour le chêne alors que ce sont plutôt les valeurs 2 et 1 pour le pin. Enfin pour le coefficient  $b$ , ce sont les valeurs 0 et 0.5 qui donnent les meilleurs résultats pour le chêne alors qu'il n'y a pas de tendance nette pour le pin. Pour la croissance des chênes, la différence de performance entre le premier ( $a=0$  et  $b=0.5$ ) et le deuxième indice (la densité locale,  $a=0$  et  $b=0$ ) est très faible (0.5%) ; nous avons donc retenu la densité locale car cet indice ne nécessite pas la connaissance de la distance aux arbres voisins dans le disque de rayon 10 m. De même, pour la croissance des pins, la différence de performance entre le premier ( $a=2$  et  $b=0.5$ ) et le troisième (surface terrière locale,  $a=2$  et  $b=0$ ) est très faible (0.5%) ; nous avons donc retenu la surface terrière locale. Pour résumer, on peut donc conclure que l'indice qui donne les meilleurs résultats pour le chêne est plutôt lié à la densité locale (nombre de tiges dans un disque de rayon 10 m) alors que pour le pin, c'est plutôt la surface terrière locale. La connaissance de la position des arbres dans le disque de rayon 10 m n'améliore pas de façon importante les résultats.

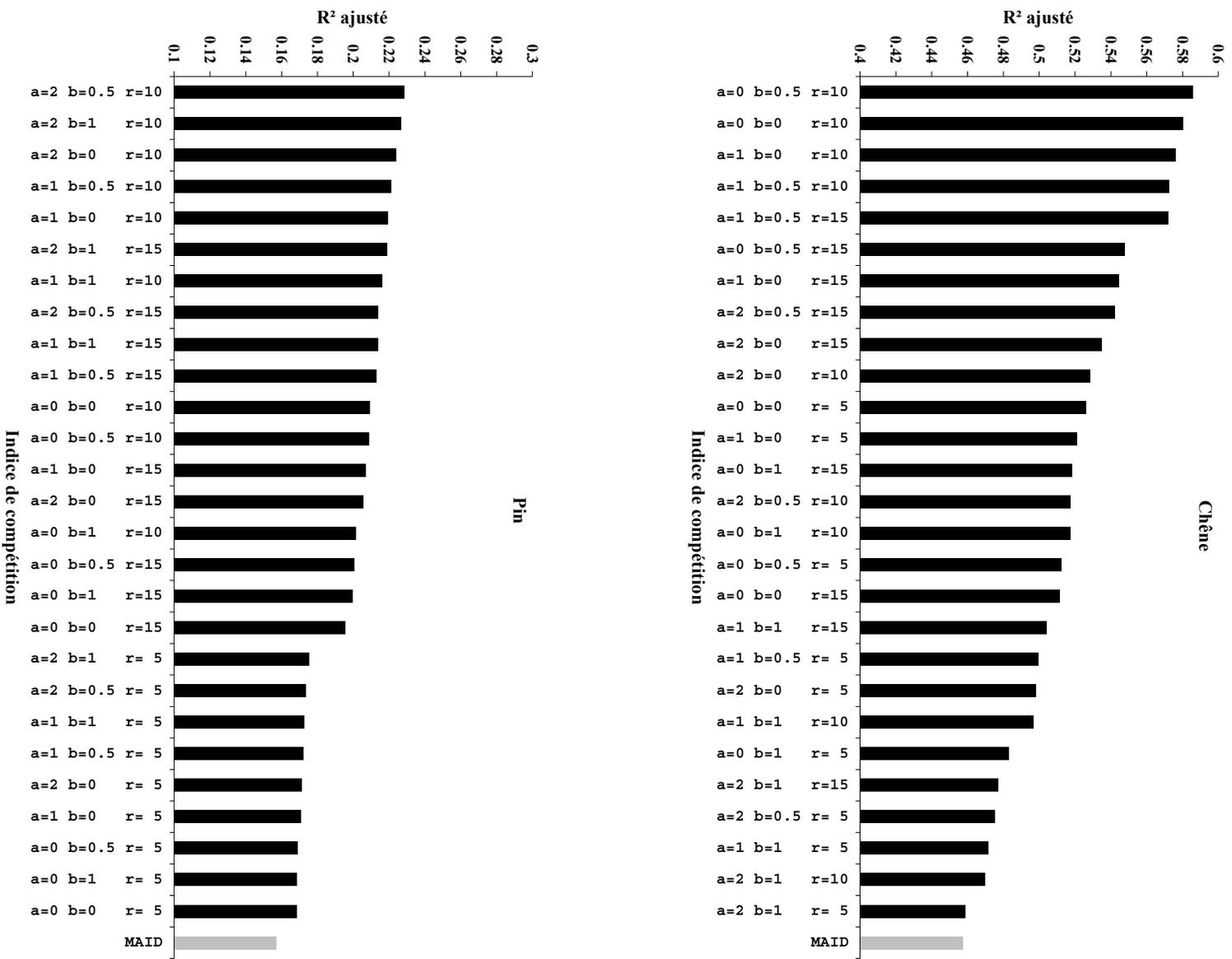


Figure 4.13 : R<sup>2</sup> ajusté pour les 27 indices de compétition (équation 4.8). Le résultat obtenu avec le modèle arbre indépendant des distances (équation 4.6) est montré en gris.

Le tableau 4.4 donne les valeurs des coefficients des 10 premiers indices de compétition ainsi que les valeurs de probabilité et les  $R^2$  ajustés associés.

Tableau 4.4 : Valeurs des coefficients pour les 10 premiers indices de compétition, valeurs des probabilités et  $R^2$  ajustés correspondant au modèle de l'équation 4.8.

Arbre sujet	Rang	<i>a</i>	<i>b</i>	<i>r</i>	$\lambda_{chêne}$	$Pr_{chêne}>F$	$\lambda_{pin}$	$Pr_{pin}>F$	$R^2$ ajusté
Chêne	1	0	0.5	10	-0.788	<.0001	-0.523	0.0196	0.585
<b>Chêne</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>10</b>	<b>-0.353</b>	<b>&lt;.0001</b>	<b>-0.242</b>	<b>0.0129</b>	<b>0.580</b>
Chêne	3	1	0	10	-2.431	<.0001	-1.123	0.0010	0.576
Chêne	4	1	0.5	10	-5.108	<.0001	-2.282	0.0035	0.572
Chêne	5	1	0.5	15	-4.310	<.0001	-1.897	0.0015	0.572
Chêne	6	0	0.5	15	-0.535	<.0001	-0.356	0.0375	0.547
Chêne	7	1	0	15	-1.417	<.0001	-0.619	0.0049	0.544
Chêne	8	2	0.5	15	-18.474	<.0001	-6.636	0.0008	0.542
Chêne	9	2	0	15	-26.885	<.0001	-9.616	0.0008	0.534
Chêne	10	2	0	10	-38.537	<.0001	-13.163	0.0025	0.528
Pin	1	2	0.5	10	-1.459	0.7729	-8.496	0.0002	0.228
Pin	2	2	1	10	-4.304	0.6662	-19.361	0.0001	0.226
<b>Pin</b>	<b>3</b>	<b>2</b>	<b>0</b>	<b>10</b>	<b>-0.750</b>	<b>0.9283</b>	<b>-12.779</b>	<b>0.0005</b>	<b>0.223</b>
Pin	4	1	0.5	10	0.582	0.5349	-2.488	0.0008	0.221
Pin	5	1	0	10	0.320	0.4174	-0.947	0.0019	0.219
Pin	6	2	1	15	-0.969	0.9203	-16.555	0.0005	0.218
Pin	7	1	1	10	0.612	0.7276	-5.684	0.0004	0.216
Pin	8	2	0.5	15	2.070	0.6338	-5.662	0.0037	0.213
Pin	9	1	1	15	0.688	0.6799	-4.977	0.0010	0.213
Pin	10	1	0.5	15	0.711	0.3621	-1.743	0.0044	0.212

Dans le cas du chêne, les coefficients  $\lambda_{chêne}$  et  $\lambda_{pin}$  sont tous les deux significatifs et négatifs. Si on regarde l'ensemble des 10 premiers indices, la valeur absolue de  $\lambda_{chêne}$  est entre 1.5 et 3 fois supérieure à celle de  $\lambda_{pin}$ . Dans le cas du pin, le coefficient  $\lambda_{chêne}$  n'est jamais significatif. On peut donc dire que, pour un chêne, la compétition exercée par les chênes voisins est entre 1.5 et 3 fois plus forte que celle exercée par les pins voisins, alors que pour un pin, seule la compétition exercée par les pins voisins joue de façon significative sur la croissance. Pour les deux essences et particulièrement pour le pin, la compétition intraspécifique est donc plus forte que la compétition interspécifique.

Pour le chêne, l'effet dispositif est significatif sur  $\alpha$  et  $\beta$  alors que pour le pin le modèle le plus parcimonieux est celui pour lequel l'effet dispositif est uniquement sur  $\beta$ . Le tableau 4.5 donne les résultats de l'ANOVA ainsi que les statistiques du modèle. Le modèle arbre dépendant des distances explique 58.1% de la variabilité pour le chêne ce qui fait un gain de 14.8% par rapport au modèle arbre indépendant des distances. Pour le pin, le modèle le plus parcimonieux explique 21.2% ce qui fait un gain de 5.7% par rapport au modèle arbre indépendant des distances.

Tableau 4.5 : Tableau de l'ANOVA, estimation des paramètres et statistiques du modèle arbre dépendant des distances (équation 4.8).

Source	Chêne (a=0 et b = 0 pour CI)				Pin (a=2 et b=0 pour CI)			
	SC	d.d.l.	F	Pr>F	SC	d.d.l.	F	Pr>F
Constante	50.48	1	4.939	0.0279	110.75	1	10.647	0.0013
circ1999	93.95	1	9.194	0.0029	250.38	1	24.071	<.0001
Dispositif	177.99	8	2.177	0.0329				
CI <sub>chêne</sub>	417.87	1	40.89	<.0001				
CI <sub>pin</sub>	64.83	1	6.344	0.0129	137.44	1	13.213	0.0003
circ1999:Dispositif	184.87	8	2.261	0.0267	224.02	8	2.692	0.0082
Résiduelle	1369.24	134			1747.46	168		

Coefficient	Estimation	Ecart-type ( $\sigma$ )	Estimation	Ecart-type ( $\sigma$ )
$\alpha_{k,s}$	$4.003 < \alpha_{k,chêne} < 13.014$	$2.270 < \sigma < 3.393$	4.602	1.410
$\beta_{k,s}$	$0.0361 < \beta_{k,chêne} < 0.1766$	$0.0256 < \sigma < 0.0434$	$0.0337 < \beta_{k,pin} < 0.0812$	$0.0097 < \sigma < 0.0175$
$\lambda_{chêne}$	-0.3539	0.0553		
$\lambda_{pin}$	-0.2424	0.0962	-11.364	3.126

<b>R<sup>2</sup> ajusté</b>	<b>0.5806</b>	<b>0.2117</b>
<b>Ecart-type résiduel</b>	<b>3.197</b>	<b>3.225</b>

### 4.2.3 L'intérêt d'introduire un indice de compétition local dépend de l'espèce considérée

Grâce à l'introduction d'un indice de compétition local, nous avons montré que la part de la variabilité de la croissance expliquée par le MADD était plus importante que celle expliquée par le MAID. Le gain par rapport au MAID est plus fort pour le chêne (+14.8% pour le  $R^2$  ajusté) que pour le pin (+5.7%). Stadt *et al.* (2007), en utilisant le même type d'indice de compétition, obtiennent un gain compris entre environ 10% et 30% suivant les essences, comparé à un modèle incluant uniquement un effet taille. Le gain pour le pin est assez faible et l'intérêt d'introduire un indice de compétition spatialisé pour cette essence peut être remis en question. Ce résultat pourrait être lié au comportement de l'espèce. Ainsi, dans l'étude de Stadt *et al.* (2007), l'effet de l'indice de compétition est plus faible avec les espèces intolérantes à l'ombrage (*Populus tremuloides*, *Populus balsamifera*, *Pinus contorta*) qu'avec les espèces tolérantes à l'ombrage (*Betula papyrifera*, *Picea glauca*). Selon eux, ce résultat pourrait être relié au fait que les arbres survivants des espèces intolérantes tendent à occuper une position dominante au niveau de la canopée et se trouveraient donc dans une gamme réduite d'intensité de compétition pour la lumière.

Le modèle arbre dépendant des distances nous a permis de sélectionner les indices de compétition les plus pertinents. Le *Crowding Index* (CI) est intéressant car il permet d'étudier les effets liés à la taille (paramètre  $a$ ) et à la distance (paramètre  $b$ ) des arbres du voisinage sur la croissance. Les meilleurs résultats ont été obtenus pour des indices calculés dans un disque de rayon 10 m. Ce rayon de voisinage donne les meilleurs résultats pour différents types d'indices et pour les deux espèces ce qui souligne la robustesse de cette valeur. Les indices les plus performants sont des indices liés à la densité locale ( $a = 0$  et  $b = 0$ ) pour le chêne et des indices liés à la surface terrière locale ( $a = 2$  et  $b = 0$ ) pour le pin. Dans d'autres études utilisant le *Crowding Index* (Canham *et al.*, 2004 ; Canham *et al.*, 2006 ; Stadt *et al.*, 2007), la valeur du paramètre  $a$  est très dépendante de l'espèce de l'arbre sujet ( $0 \leq a \leq 3.259$ ). En revanche, ces études aboutissent systématiquement à des coefficients  $b$  inférieurs à 1 ( $0 \leq b \leq 0.728$ ) ce qui montre que le déclin de l'effet d'un voisin avec la distance est moins important que ce qui est admis bien souvent lorsqu'on fixe  $b = 1$  dans les indices de compétition (Canham *et al.*, 2004).

Nous avons montré que la distance à l'arbre améliorerait de façon négligeable la performance des indices de compétition. Ainsi dans un voisinage donné, la présence et le

diamètre à 1.30 m d'un arbre sont des informations suffisantes pour estimer sa contribution à la compétition.

#### 4.2.4 Comment séparer le potentiel de croissance de la compétition locale ?

Dans les paragraphes précédents nous avons vu que les indices de compétition locaux expliquaient une faible part de la croissance individuelle. Une explication possible à ce résultat est que la circonférence de l'arbre, introduite dans le modèle pour exprimer le potentiel de croissance de l'arbre, intègre déjà l'influence du voisinage. L'objectif de ce paragraphe est de présenter des pistes qui permettraient de mieux séparer l'effet du potentiel de croissance de l'effet du voisinage.

##### 4.2.4.1 Rappel sur les avantages et les inconvénients de la circonférence à 1.30 m

On peut faire l'hypothèse que le potentiel de croissance d'un arbre dépend de la quantité de lumière qu'il reçoit. Or, nous n'avons pas accès à cette quantité. On utilise donc des variables sensées être reliées à cette quantité, par exemple des variables caractérisant la partie de l'arbre qui réceptionne la lumière : le houppier. Dans notre travail, nous avons utilisé la circonférence à 1.30 m. Cette variable présente deux avantages :

- dans le cas d'une croissance libre, elle est liée au potentiel de croissance car il existe une relation allométrique entre cette variable et les caractéristiques du houppier (voir par exemple Prevosto, 1999) ;
- elle est facilement accessible.

La circonférence présente également deux inconvénients :

- pour un arbre poussant en peuplement, elle n'est pas forcément bien reliée à la quantité de lumière reçue. En effet, deux arbres de même diamètre peuvent avoir eu une histoire différente, avoir des houppiers très différents et donc un accès à la lumière différent ;
- elle résulte de la compétition passée, donc elle est liée au voisinage passé qui est aussi très corrélé au voisinage actuel (Lorimer, 1983 ; Vanclay, 1994 ; Lee *et al.*, 2004). Ainsi certains auteurs considèrent la taille à 1.30 m comme un indice de compétition (Lorimer, 1983).

La circonférence ne permet donc pas de séparer complètement l'effet lié au potentiel de croissance et l'effet lié au voisinage. L'effet du voisinage local sur la croissance, que l'on a estimé avec un indice de compétition local, est donc déjà en partie intégré dans la circonférence de l'arbre. Ainsi, le fait que l'indice de compétition ne soit pas significatif dans

une analyse de variance ne veut évidemment pas dire qu'il n'y a pas d'effet de la compétition local mais que cet effet est peut-être déjà pris en compte par la circonférence.

#### 4.2.4.2 Utilisation d'une autre variable "accessible" pour exprimer le potentiel de croissance

Plutôt que d'utiliser la circonférence du tronc à 1.30 m pour exprimer le potentiel de croissance, une autre variable peut être utilisée mais il faut que celle-ci reste accessible. Par exemple, on peut utiliser la hauteur totale individuelle même si cette mesure est plus contraignante à acquérir que la circonférence à 1.30 m. En effet cette variable reflète mieux l'accès de l'arbre à la lumière et elle peut être moins liée au voisinage local comme cela a été montré pour le pin sylvestre (Prevosto, 1999 page 122). Cependant, pour que la démarche soit correcte, il serait préférable d'avoir la hauteur de l'arbre en début de période de croissance. Ceci nécessite de mettre en place un suivi de la hauteur ou de reconstituer la croissance en hauteur. On pourrait alors remplacer dans nos modèles la circonférence par la hauteur totale et vérifier si l'indice de compétition locale a plus de poids que dans le modèle utilisant la circonférence.

#### 4.2.4.3 Utilisation d'informations sur la croissance passée

Pour construire le modèle arbre indépendant des distances, nous avons utilisé les données de croissance uniquement sur la période 2000-2005, essentiellement parce que les interventions sylvicoles effectuées dans les dispositifs avant 2000 sont mal connues. Cependant, nous avons reconstitué toute la croissance des individus sondés. Ces données peuvent fournir des informations sur le rythme de croissance des arbres ou sur la croissance potentielle des chênes et des pins. En effet, dans les modèles de type "potentiel×réducteur" (Gourlet-Fleury et Houllier, 2000), le potentiel correspond au potentiel de croissance d'un arbre en croissance libre. Il peut être estimé de trois façons :

- avec un jeu de données sur des arbres en croissance libre (Hasenauer, 1997 ; Picard, 1999) ;
- par rapport aux données d'accroissement maximum observées sur le jeu de données (Picard, 1999 ; Manson *et al.*, 2006) ;
- en même temps que le réducteur, mais dans ce cas la partie potentiel et la partie réducteur perdent leur sens initial (Gourlet-Fleury et Houllier, 2000).

Nous proposons d'étudier ici comment nous pourrions exploiter nos données pour définir un potentiel de croissance propre à chaque essence à partir de la croissance passée d'arbres en croissance libre ou à partir d'accroissements maximums observés.

Dans notre jeu de données, nous pouvons penser que nous n'avons pas d'arbres ayant poussé en croissance libre. Cependant, pour les chênes, les arbres issus du traitement en taillis sous futaie appartenant au dispositif D02 (voir partie 3) pourraient se rapprocher d'arbres ayant poussé sans concurrence. Si on observe l'évolution du diamètre en fonction du temps dans ce dispositif (figure 4.14), on constate qu'au moins 2 individus ont une croissance en circonférence qui ne semble pas affectée par la croissance des autres arbres du peuplement.

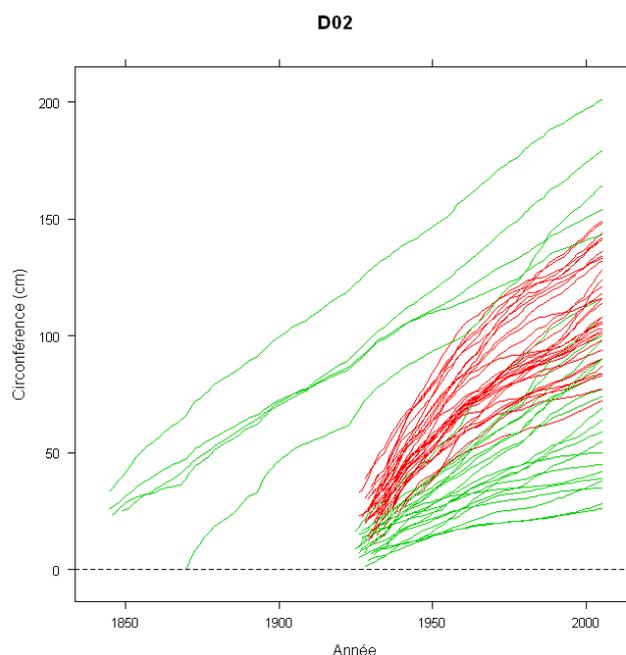


Figure 4.14 : Evolution de la circonférence au cours du temps pour tous les arbres sondés du dispositif D02. En rouge les pins et en vert les chênes.

De plus, nous voyons que pour ces arbres, la relation entre la circonférence et le temps est une droite. Cela peut se traduire par l'équation suivante :

$$c_i(t) = \alpha + \beta t + \varepsilon_i$$

où  $c_i(t)$  est la circonférence de l'arbre  $i$  au temps  $t$ . Si on considère que ces arbres sont des arbres en croissance libre, ils permettent d'exprimer l'accroissement potentiel en fonction du temps comme une constante :

$$\frac{\Delta c_i(t)}{\Delta t} = \beta_{pot} + \varepsilon_i'$$

Pour le chêne, on peut donc réécrire le modèle de croissance de la façon suivante :

$$\frac{\Delta c_i(t)}{\Delta t} = \beta_{pot} + \gamma \times IC_i + \varepsilon_i$$

avec  $IC_i$  un indice de compétition. Cependant, la constante  $\beta_{pot}$  est déterminée sur un dispositif et elle est probablement fautive si on change de site. Malheureusement, nous n'avons pas d'arbres permettant d'estimer la croissance potentielle des chênes pour tous les dispositifs.

D'autre part, même si cette approche semble intéressante pour le chêne, elle n'est pas applicable aux pins.

Enfin, dans une étude sur les boisements spontanés, Prevosto (1999) montre que la forme de la courbe de croissance d'arbres isolés n'est pas la même que pour les arbres en peuplement (figure 4.15). Il observe ce phénomène sur deux essences, le bouleau et le pin sylvestre. La croissance réduite des arbres isolés dans les premières années pourrait s'expliquer soit par un stress climatique, comme le manque de protection contre le vent par d'autres arbres, soit par la compétition par la végétation environnante, qui est dense et qui ralentit donc l'installation et la croissance initiale des jeunes arbres isolés. L'utilisation de la croissance d'arbres isolés pour estimer le potentiel de croissance d'arbres en peuplement présente donc des limites.

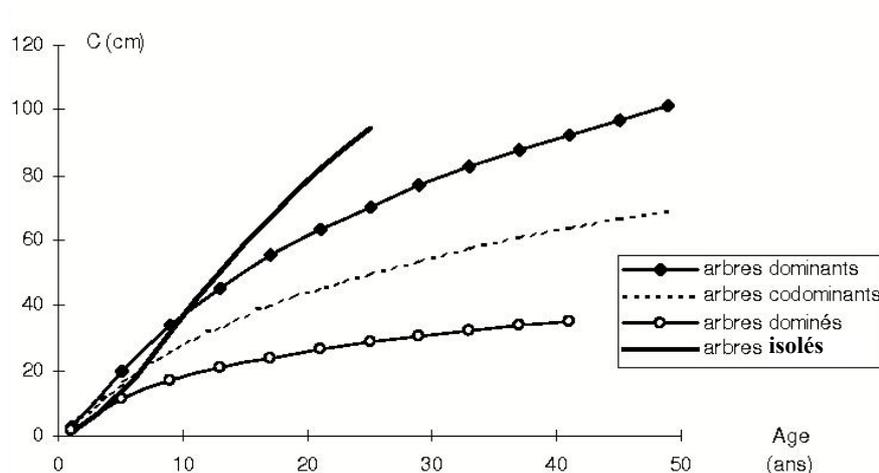


Figure 4.15 : Croissance moyenne en circonférence sous écorce calculée à partir des données de croissance radiale pour le pin sylvestre (figure extraite de Prevosto, 1999)

Une autre approche consisterait à sélectionner les arbres de notre jeu de données qui ont les croissances les plus élevées et considérer qu'elles reflètent la croissance d'arbres ayant poussé sans concurrence sur la période étudiée. Par exemple, Manson *et al.* (2006) calculent le potentiel d'accroissement maximum en effectuant la moyenne des cinq plus grandes valeurs observées. Cette fois, nous pouvons utiliser cette méthode aussi bien pour le chêne que pour le pin. L'inconvénient de cette approche est que le potentiel ainsi estimé n'est valable que pour la période considérée car il peut être fortement influencé par les variations climatiques. De plus, pour des peuplements denses, les arbres ayant la plus forte croissance représentent probablement mal la croissance potentielle.

Les différentes pistes évoquées ici pourraient permettre d'étudier plus finement l'effet de la compétition locale sur la croissance même si aucune n'est entièrement satisfaisante. Nous

voyons également ici les limites de l'utilisation des variables dendrométriques et des indices de compétition pour expliquer les différentes composantes de la croissance. Toutefois, étant donné les contraintes et les données disponibles dans nos dispositifs, nous avons choisi de garder la circonférence comme indicateur du potentiel de croissance.

#### **4.2.5 La compétition intraspécifique est supérieure à la compétition interspécifique**

Dans nos peuplements mélangés chêne-pin les analyses montrent que la compétition intraspécifique est supérieure à la compétition interspécifique. On retrouve ce résultat dans de nombreuses études sur les forêts mélangées (Tham, 1989 ; Pukkala *et al.*, 1998 ; Vettenranta, 1999 ; Stadt *et al.*, 2002 ; Bristow *et al.*, 2006b ; Piotto, 2008). Selon Bartelink (2000), les compétitions intra et interspécifiques sont cruciales dans la dynamique des peuplements mélangés. Elles pourraient notamment être impliquées dans les mécanismes de coexistence. Ainsi, le modèle théorique de compétition interspécifique de "Lotka-Volterra" montre que la compétition intraspécifique doit être supérieure à la compétition interspécifique pour que deux espèces coexistent (Begon *et al.*, 1996).

Nous avons également montré que la compétition exercée par les chênes sur les pins n'était pas significative. Ce résultat rejoint les travaux de Biondi *et al.* (1992) qui concluent que la compétition de chênes (*Quercus gambelii* Nutt) sur des pins (*Pinus ponderosa* Dougl ex Laws) est négligeable par rapport à la compétition des pins sur les pins quand le chêne reste principalement dans le sous-étage. Les houppiers des pins sont alors peu concurrencés par les houppiers des chênes. Dans notre étude, tous les chênes ne sont pas en sous-étage mais en général les pins sont plus gros que les chênes (voir partie 3, tableau 3.2), ce qui pourrait expliquer l'influence non significative des chênes sur les pins.

Nos résultats sont évidemment valables dans notre contexte stationnel et pour la période de croissance étudiée. En effet, les rapports de compétition entre les espèces changent notamment en fonction du site et du climat (Stadt *et al.*, 2002). Le mélange chêne-hêtre est un bon exemple de mélange pour lequel le rapport de force entre les espèces varie fortement d'une région de la France à l'autre (Hein et Dhote, 2006).

Le MADD nous a permis de préciser l'importance de la compétition intraspécifique par rapport à la compétition interspécifique. Ce résultat est important car, comme nous le verrons plus loin, elle peut conditionner le comportement du mélange à l'échelle du peuplement.

### 4.3 Le modèle arbre basé sur des distributions de voisinage

Dans la partie 2, nous avons vu que l'on pouvait construire un modèle de croissance qui utilise des distributions de voisinage (MADV). Ces distributions intègrent des informations sur la structure spatiale du peuplement et la variabilité des voisinages. Nous avons montré qu'elles permettaient de reconstruire des informations sur l'environnement local des arbres. Nous reprenons et précisons ici le fonctionnement de ce modèle appliqué à la croissance des peuplements mélangés chêne-pin. Pour ces peuplements, le modèle avec distributions de voisinage utilise la même équation de croissance que celle du MADD (équation 4.8) mais les indices de compétition sont calculés à partir des voisinages reconstruits et non à partir des voisinages réels (équation 4.9)<sup>11</sup>.

$$\Delta r_{k,s,i} = \alpha'_{k,s} + \beta'_{k,s} circ_{k,s,i} + \lambda_{s,chêne} CI^*_{i,s,chêne} + \lambda_{s,pin} CI^*_{i,s,pin} + \varepsilon'_{k,s,i} \quad (4.9)$$

Dans l'équation 4.9 correspondant au MADV, l'étoile ajoutée aux indices de compétition signifie qu'ils sont calculés à partir d'un voisinage reconstruit et non à partir du voisinage réel de l'arbre sujet.

#### 4.3.1 Estimation des distributions de voisinage

Pour estimer les distributions de voisinage, il faut connaître le nombre de voisins d'une partie des arbres du peuplement. Si on dispose d'une cartographie du peuplement, le nombre de voisins peut évidemment être calculé pour tous les arbres du peuplement ce qui permet d'obtenir une distribution de voisinage précise. Si on ne dispose pas d'une cartographie, on peut procéder par échantillonnage, par exemple en sélectionnant de façon aléatoire des arbres dans le peuplement pour lesquels le nombre de voisins est compté (figure 4.16).

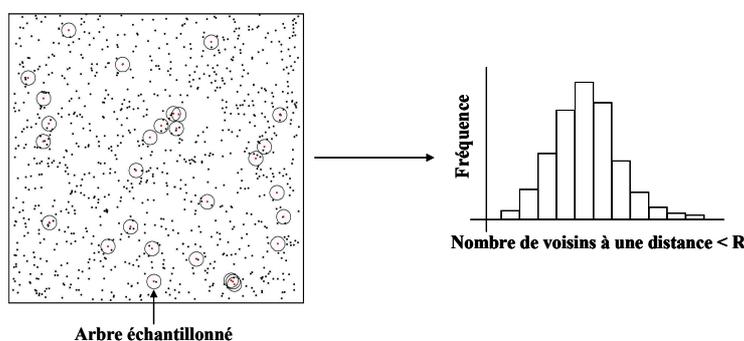


Figure 4.16 : Principe d'estimation d'une distribution de voisinage à partir d'un échantillon d'arbres. Pour chaque arbre échantillonné, on compte le nombre de voisins à une distance inférieure à R ce qui permet de construire une distribution du nombre de voisins.

<sup>11</sup> Voir l'équation 4.8 pour la définition des différents éléments de cette équation.

En fonction du taux d'échantillonnage, la distribution de voisinage sera estimée de façon plus ou moins précise. Dans cette partie, pour ne pas introduire un facteur de variation supplémentaire lié à l'estimation des distributions, nous avons opté pour la première solution.

Pour les peuplements mélangés chêne-pin, nous avons choisi de définir quatre distributions de voisinage de manière à prendre en compte les interactions spatiales spécifiques et interspécifiques entre les individus :

- la distribution du nombre de voisins chêne pour les sujets chêne ;
- la distribution du nombre de voisins pin pour les sujets chêne ;
- la distribution du nombre de voisins chêne pour les sujets pin ;
- la distribution du nombre de voisins pin pour les sujets pin.

Dans chaque dispositif, nous avons donc estimé ces distributions de voisinage à partir de l'ensemble des arbres du peuplement (voir figure 4.17 pour un exemple de distributions obtenues sur le dispositif D49). Lorsque cela était nécessaire, les autres espèces feuillues ont été regroupées avec le chêne (par exemple pour des simulations destinées à évaluer le modèle sur des données réelles, voir partie 5). Les distributions ont été construites de façon à avoir au moins 10 classes avec des limites de classe entières. Dans la discussion de cette partie (paragraphe 4.3.3), nous reviendrons sur l'estimation des distributions de voisinage dans un cadre de gestion.

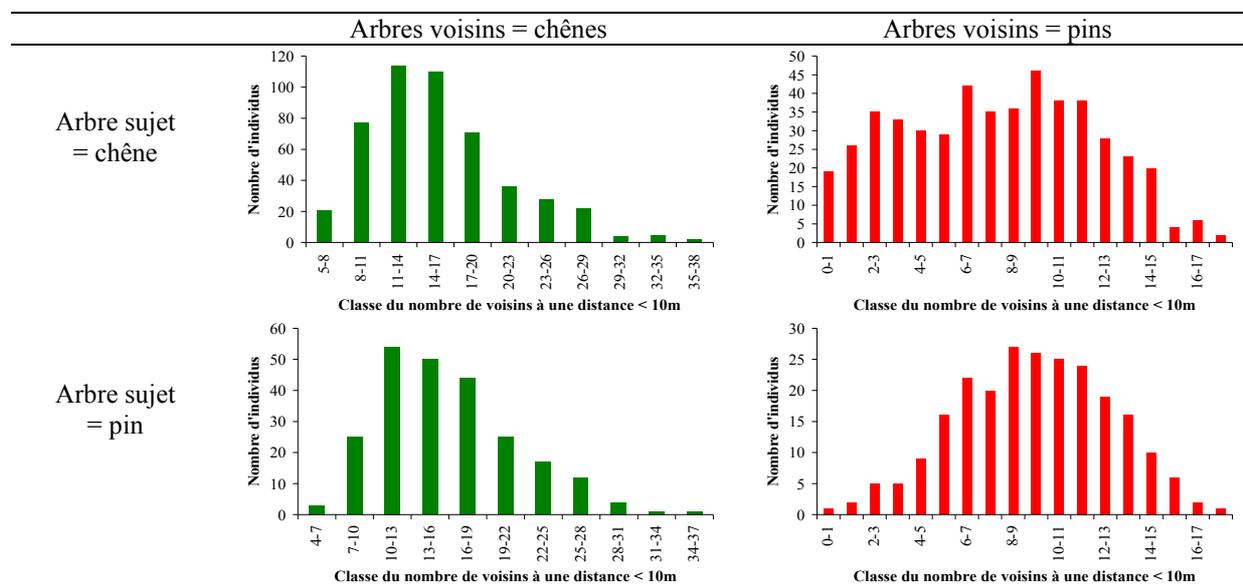


Figure 4.17 : Exemple des distributions de voisinage obtenues dans le dispositif D49.

### 4.3.2 Fonctionnement du modèle

Comme nous l'avons présenté dans le paragraphe 2.5.2, le modèle est basé sur la reconstruction des voisinages puis la mise à jour de ces voisinages au cours du temps :

- Reconstruction des voisinages : lors de l'étape initiale, chaque arbre reçoit une liste de voisins conforme aux distributions de voisinage et aux caractéristiques des arbres du peuplement (figure 4.18). Cette reconstruction des voisinages est donc réalisée une seule fois. Deux approximations sont réalisées à ce niveau. Premièrement : pour un individu, le nombre de voisins chêne et le nombre de voisins pin ne sont pas conditionnés par le nombre total de voisins. Deuxièmement : les voisins sont choisis indépendamment de la taille de l'arbre sujet ce qui revient à faire l'hypothèse d'indépendance spatiale entre un individu et sa taille.
- Mise à jour des voisinages : à chaque étape, les voisinages sont mis à jour pour tenir compte de la croissance et des éclaircies (voir figure 4.19 et figure 4.20).

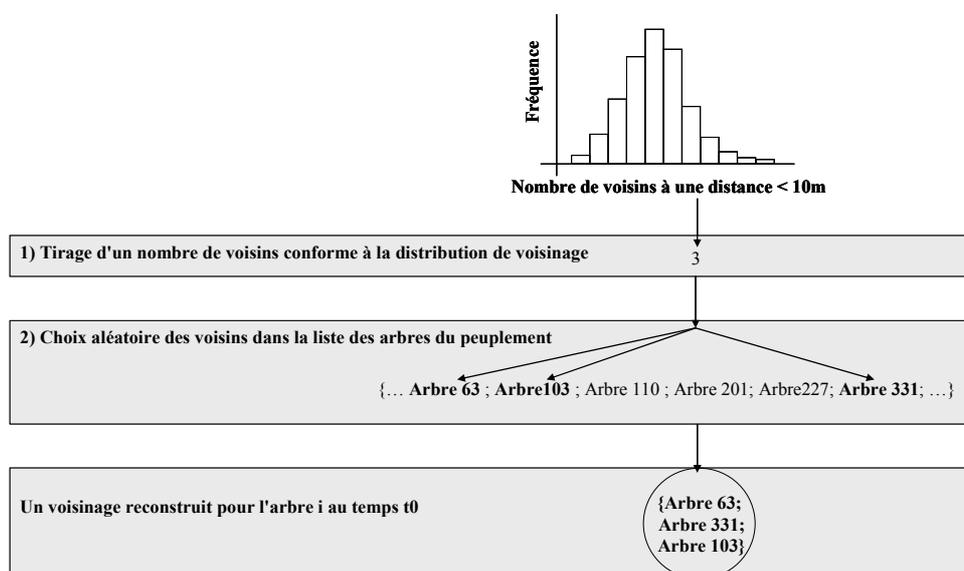


Figure 4.18 : Reconstruction du voisinage en début de période de croissance ( $t_0$ ) pour un arbre  $i$  du peuplement.

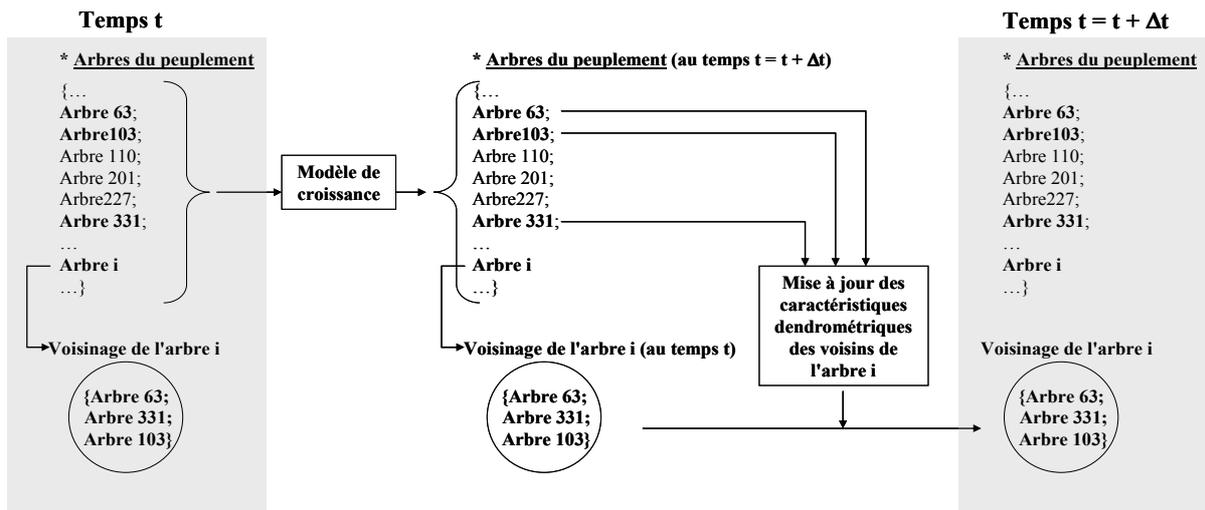


Figure 4.19 : Mise à jour du voisinage d'un arbre après la phase de croissance. Tous les arbres du peuplement ont une liste de voisins. Par exemple, l'arbre i a trois voisins numérotés 63, 331 et 103 appartenant au peuplement (en gras dans la liste). Le modèle de croissance fait pousser tous les arbres du peuplement entre les temps t et t+Δt en tenant compte de l'état des voisinages au temps t. Les caractéristiques des arbres des voisinages sont mises à jour après cette phase de croissance. Par exemple les arbres 63, 331 et 103 au temps t+Δt permettent de mettre à jour le voisinage de l'arbre i. Ce nouvel état du voisinage sera utilisé à l'étape suivante de croissance.

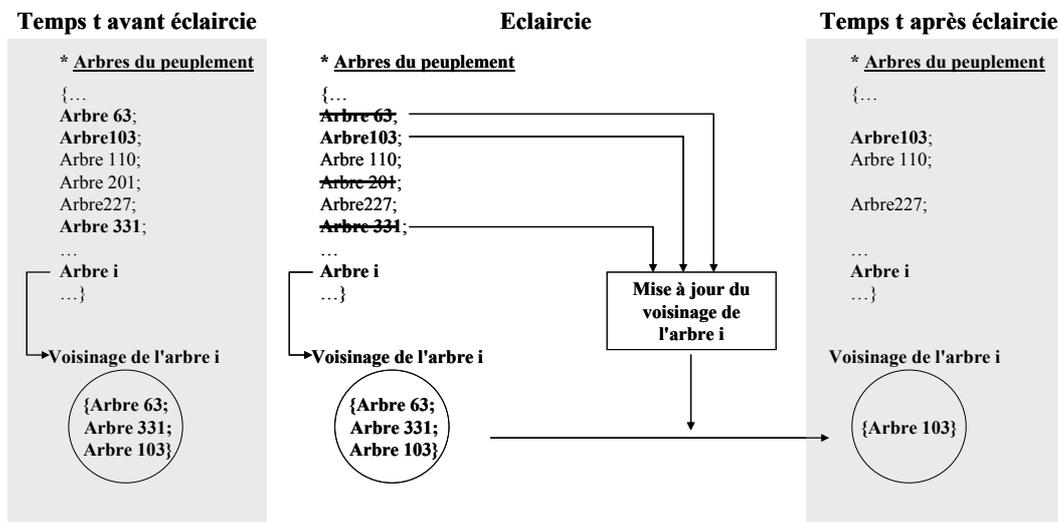


Figure 4.20 : Mise à jour du voisinage d'un arbre après une éclaircie. Les arbres rayés représentent des arbres enlevés lors de l'éclaircie.

### 4.3.3 Utilisation du modèle dans un contexte de gestion

Pour pouvoir utiliser le modèle que nous venons de présenter dans un cadre de gestion, il est nécessaire d'estimer les distributions de voisinage. Nous pensons que ces distributions pourraient être estimées à partir des données récoltées lors des aménagements. En France, lors de l'aménagement d'une forêt, les gestionnaires procèdent à des inventaires par points d'échantillonnage (Duplat et Perrote, 1981) avec, par exemple, une placette temporaire tous les quarts d'hectare (voir figure 4.21). Les placettes peuvent être des disques de rayon constant sur lesquelles les arbres sont inventoriés et leurs caractéristiques dendrométriques mesurées. Si on connaît l'espèce de l'arbre le plus proche du centre de la placette, on peut avoir pour cette espèce une estimation du nombre de voisins. En regroupant les informations recueillies sur chaque placette, il serait alors possible de reconstituer des distributions de voisinage à l'échelle d'une parcelle sans bouleverser de façon importante les pratiques des gestionnaires.

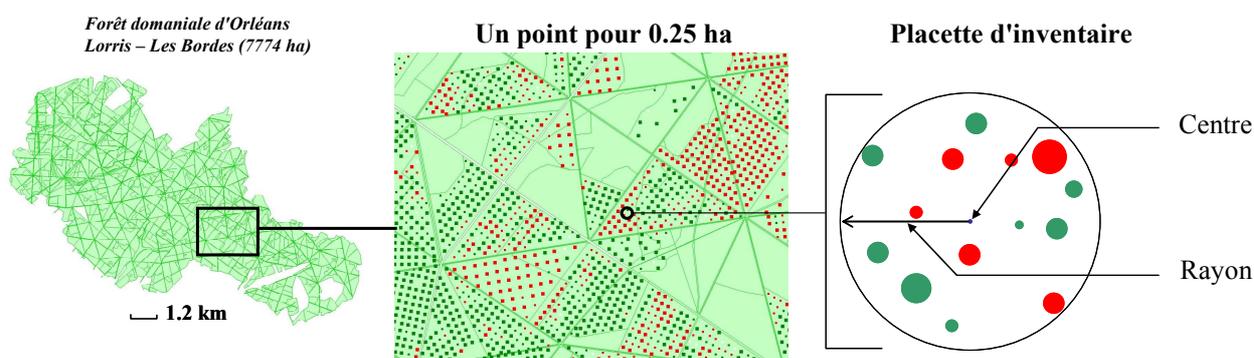


Figure 4.21 : Exemple d'un inventaire par point d'échantillonnage sur le massif de Lorris avec un point tous les 0.25 ha. Dans les parcelles, les points verts sont des placettes où le chêne est majoritaire et les points rouges des placettes où le pin est majoritaire. Sur le schéma à droite représentant une placette d'échantillonnage, les ronds verts représentent des chênes et les ronds rouges des pins.

### 4.3.4 Limites du modèle

Une première limite du modèle est que les voisinages reconstruits sont basés sur des approximations : le nombre de voisins chêne et le nombre de voisins pin ne sont pas conditionnés par le nombre total de voisins<sup>12</sup> et un voisin est choisi de façon aléatoire dans le peuplement indépendamment de la taille de l'arbre sujet. Ces approximations peuvent engendrer des différences entre les prédictions du modèle arbre dépendant des distances et celles du modèle utilisant les distributions de voisinage. Cependant, elles ne sont pas faciles à éviter. Par exemple, pour éviter la première approximation, il faudrait connaître les distributions conjointes qui donnent la probabilité qu'un arbre d'une espèce ait en même temps

<sup>12</sup> Le nombre total de voisins est en fait en partie contraint car pour chaque individu nous utilisons une distribution spécifique et une distribution interspécifique. Les distributions interspécifiques permettent également de prendre en compte la dépendance spatiale entre les deux espèces (voir paragraphe 2.5.2).

$n_{chêne}$  voisins chêne et  $n_{pin}$  voisins pin. Cette solution n'est pas envisageable car l'estimation de ces distributions nécessiteraient une grande quantité de données et aboutirait à une contrainte comparable à la cartographie d'un peuplement.

La deuxième limite vient de l'indice choisi pour estimer la compétition. La performance d'un indice de compétition dépend en général des données sur lesquelles il a été testé (Prévosto, 2005). Il est donc possible que, pour une autre étude, d'autres types d'indices soient plus pertinents que ceux que nous avons utilisés. Certains indices, comme ceux qui prennent en compte l'organisation spatiale des arbres dans le voisinage (Pukkala, 1989), pourraient être assez difficile à recalculer à partir de notre méthode, mais il est toujours possible d'attribuer une distance et un azimut à chaque voisin suivant des données de terrain ou des hypothèses sur la distribution des arbres dans le voisinage. On peut alors se demander comment se comporterait le modèle avec d'autres indices de compétition. En particulier, si les indices incluant la distance aux voisins étaient très performants, la différence avec le modèle spatialisé pourrait être importante. On pourrait alors travailler directement sur les distributions des indices de compétition (Phillips *et al.*, 2003), mais cette solution nécessite de modéliser l'évolution de ces distributions au cours du temps.

Enfin, une autre limite du modèle est que les processus de mortalité et de régénération ne sont pas pris en compte. Les peuplements chêne-pin de nos dispositifs sont assez denses et la régénération est en général absente. De plus, les éclaircies sont pratiquées environ tous les 10 ans ce qui limite l'importance de la mortalité par compétition. On peut donc considérer que le modèle que nous venons de présenter est utilisable dans ce contexte forestier. Cependant, on peut se demander s'il est possible d'intégrer les processus de mortalité et de régénération dans ce type de modèle. Pour la mortalité par compétition, on peut faire l'hypothèse que ce processus est décrit par un modèle individuel qui dépend de la taille des individus et de l'accroissement sur les dernières années (Kobe et Coates, 1997 ; Wyckoff et Clark, 2000 ; Wunder *et al.*, 2007). Dans ce cas, la prise en compte de la mortalité dans la mise à jour des voisinages peut se faire de la même manière que pour les éclaircies. En revanche, la prise en compte du processus de régénération nécessite en général des informations spatiales qui ne sont pas centrées sur les individus déjà en place. Par exemple, la quantité de lumière qui arrive en un point du sol (qui peut éventuellement être reliée à des caractéristiques dendrométriques du peuplement environnant) peut permettre d'estimer la probabilité pour un individu d'une espèce de s'installer. De plus, le voisinage pertinent pour le développement d'un jeune arbre peut être très différent du voisinage utilisé pour la croissance d'un arbre

adulte (Prévosto, 2005). Il est donc probable que le processus de régénération soit assez difficile à intégrer dans ce type de modèle sans apporter de nouvelles informations spatiales.

#### 4.3.5 Conclusion

Nous avons utilisé des distributions de voisinage pour construire un modèle intermédiaire entre un modèle arbre indépendant des distances et un modèle arbre dépendant des distances. Le MADV possède une part stochastique (Fox *et al.*, 2001) introduite au moment de la reconstruction des voisinages. La variabilité introduite dans ce modèle est liée à la structure spatiale du peuplement et à la variabilité des environnements locaux. Les voisinages reconstruits évoluent au cours du temps de manière à tenir compte de la croissance des arbres et des éclaircies. Ce modèle est assez proche d'un modèle arbre dépendant des distances couplé avec un modèle de génération de peuplements virtuels (Pretzsch, 2002) mais il n'utilise pas une position explicite pour chaque individu. Ceci peut présenter l'avantage d'être moins coûteux en temps de calcul et pourrait faciliter l'utilisation des résultats obtenus avec les modèles arbre dépendants des distances.

L'évaluation de ce modèle par rapport à la prédiction de la croissance sera présentée dans la partie 5. Nous y présenterons également la comparaison avec le MAID et le MADD.

#### ***4.4 Construction de modèles peuplement par agrégation des modèles individus-centrés***

Comme nous l'avons vu dans la partie 2, pour construire un modèle peuplement, il faut suivre ou reconstituer la croissance des arbres sur l'ensemble d'une surface. Dans notre travail, comme nous ne disposions pas de dispositifs de suivi pour les peuplements mélangés, nous avons opté pour une reconstitution de la croissance. Etant donné le coût d'acquisition de ces données, nous avons choisi d'échantillonner des arbres dans nos dispositifs plutôt que de reconstituer la croissance de l'ensemble des arbres (voir partie 3). Ceci nous a permis d'échantillonner plus de dispositifs et donc plus de situations différentes au niveau de la structure spatiale des peuplements mais aussi au niveau des caractéristiques dendrométriques ou stationnelles. Malheureusement, ces données individuelles sont mal adaptées à la construction d'un modèle peuplement et nous ne pouvons donc pas construire ce modèle directement à partir de nos données.

Il y a plusieurs façons de contourner ce problème. Par exemple, nous pourrions reconstruire la croissance passée des arbres non échantillonnés à partir d'un modèle individus-centré ad hoc et obtenir ainsi des données à l'échelle du peuplement. Une autre possibilité est de simuler des données peuplement à partir d'un modèle plus détaillé comme nous l'avons évoqué dans la partie 2. Cette solution peut être intéressante lorsque le modèle détaillé est complexe. Nous pourrions ainsi simuler l'évolution de variables caractéristiques du peuplement à partir du modèle individus-centré spatialement explicite. Enfin, le modèle peuplement peut être obtenu par agrégation à partir des équations d'un modèle plus détaillé comme nous l'avons montré dans la partie 2. Cette solution est envisageable lorsque le modèle détaillé est relativement simple, ce qui est notre cas. C'est cette solution que nous avons choisie et que nous développons dans cette sous-partie. Nous allons appliquer la méthode d'agrégation présentée dans la partie 2 pour construire un modèle peuplement à partir du modèle arbre indépendant des distances (paragraphe 4.4.1) puis à partir du modèle arbre dépendant des distances (paragraphe 4.4.2). Nous verrons ensuite comment un modèle peuplement peut permettre d'étudier l'influence du taux de mélange sur la production du peuplement (paragraphe 4.4.3). Enfin, nous discuterons du comportement de ces modèles par rapport aux modèles individus-centrés (paragraphe 4.4.4).

#### 4.4.1 Agrégation du modèle non spatialisé

En considérant que la croissance des arbres est décrite par le modèle arbre indépendant des distances (équation 4.6) et en appliquant directement le raisonnement présenté dans le paragraphe 2.4.2.1, on peut montrer que le modèle peuplement obtenu par agrégation pour un dispositif  $k$  et une espèce  $s$  est décrit par le système d'équations suivant (voir annexe A) :

$$\begin{cases} \Delta G_{k,s} = A(\alpha_{k,s})N_{k,s} + B(\alpha_{k,s}, \beta_{k,s})N_{k,s}\bar{r}_{k,s} + C(\beta_{k,s})G_{k,s} \\ \Delta \bar{r}_{k,s} = \alpha_{k,s} + D(\beta_{k,s})\bar{r}_{k,s} \\ \Delta N_{k,s} = 0 \end{cases} \quad (4.10)$$

avec  $\Delta \bar{r}_{k,s}$  en mm,  $\bar{r}_{k,s}$  en cm et  $\Delta G_{k,s}$  en cm<sup>2</sup>.  $\Delta N_{k,s} = 0$  car nous nous plaçons dans une situation sans mortalité ni coupe pendant la période de croissance. Les paramètres de ces équations s'expriment en fonction des paramètres de l'équation 4.6 de la façon suivante (l'équation 4.6 donne l'accroissement en rayon en mm en fonction de la circonférence en cm) :

$$\begin{cases} A(\alpha_{k,s}) = (\alpha_{k,s}/10)^2 \pi \\ B(\alpha_{k,s}, \beta_{k,s}) = (1 + 2\pi\beta_{k,s}/10)2\alpha_{k,s}\pi/10 \\ C(\beta_{k,s}) = (1 + \beta_{k,s}\pi/10)\beta_{k,s}4\pi/10 \\ D(\beta_{k,s}) = 2\pi\beta_{k,s} \end{cases}$$

L'accroissement en surface terrière de chaque espèce au niveau peuplement peut donc s'exprimer en fonction de variables agrégées relativement accessibles dans un cadre de gestion : la surface terrière, la densité et le rayon moyen (ou diamètre moyen).

#### 4.4.2 Agrégation du modèle spatialisé

Cette fois, nous considérons que la croissance individuelle est décrite par le modèle arbre dépendant des distances donné par l'équation 4.8. Nous pouvons appliquer l'approximation du champ moyen (voir paragraphe 2.4.1.2) de façon à transformer le modèle dépendant des distances en modèle indépendant des distances. Nous obtenons alors une nouvelle équation pour la croissance individuelle d'une espèce  $s$  appartenant au dispositif  $k$  :

$$\Delta r_{k,s,i} = \alpha'_{k,s} + \beta'_{k,s} \text{circ}_i + \lambda_{s,\text{chêne}} \langle CI_{i,s,\text{chêne}} \rangle + \lambda_{s,\text{pin}} \langle CI_{i,s,\text{pin}} \rangle \quad (4.11)$$

Comme nous connaissons l'expression des indices  $CI_{i,s,\text{chêne}}$  et  $CI_{i,s,\text{pin}}$ , nous pouvons, sous certaines hypothèses, simplifier l'expression de la moyenne spatiale de ces indices.

#### 4.4.2.1 Simplification du modèle arbre dépendant des distances pour la croissance du chêne

Dans le cas de la croissance du chêne, les indices retenus dans le modèle arbre dépendant des distances sont la densité locale des chênes (notée  $N_{i, \text{chêne}, \text{chêne}}$ ) et la densité locale des pins (notée  $N_{i, \text{chêne}, \text{pin}}$ ). Sous les hypothèses appropriées concernant le processus ponctuel (voir paragraphes 2.4.1.1 et 2.4.1.2), les moyennes spatiales de ces indices sont directement reliées à la fonction de Ripley et à la fonction Intertype :

$$\langle CI_{i, \text{chêne}, \text{chêne}} \rangle = \langle N_{i, \text{chêne}, \text{chêne}} \rangle = \frac{N_{\text{chêne}}}{A} K_{\text{chêne}, \text{chêne}}(R)$$

$$\langle CI_{i, \text{chêne}, \text{pin}} \rangle = \langle N_{i, \text{chêne}, \text{pin}} \rangle = \frac{N_{\text{pin}}}{A} K_{\text{chêne}, \text{pin}}(R)$$

Où  $A$  est la surface du dispositif,  $K_{\text{chêne}, \text{chêne}}(R)$  est la fonction de Ripley à une distance  $R$  pour les chênes et  $K_{\text{chêne}, \text{pin}}(R)$  est la fonction Intertype à une distance  $R$ . Le lien entre les indices de compétition et les fonctions de Ripley et Intertype permet de relier la croissance à la structure spatiale du peuplement. Le modèle de l'équation 4.11 peut donc s'écrire pour le chêne :

$$\Delta r_{i, \text{chêne}} = \alpha'_{k, \text{chêne}} + \beta'_{k, \text{chêne}} \text{circ}_i + \lambda_{\text{chêne}, \text{chêne}} \frac{N_{\text{chêne}}}{A} K_{\text{chêne}, \text{chêne}}(R) + \lambda_{\text{chêne}, \text{pin}} \frac{N_{\text{pin}}}{A} K_{\text{chêne}, \text{pin}}(R) \quad (4.12)$$

Si nous faisons l'hypothèse que la distribution des chênes est aléatoire alors on a :

$$\langle CI_{i, \text{chêne}, \text{chêne}} \rangle = \frac{N_{\text{chêne}}}{A} \pi R^2$$

Si de plus il y a indépendance entre la population des chênes et la population des pins alors on obtient :

$$\langle CI_{i, \text{chêne}, \text{pin}} \rangle = \frac{N_{\text{pin}}}{A} \pi R^2$$

Sous les deux hypothèses supplémentaires que nous venons de formuler, l'équation 4.11 pour les chênes peut donc s'écrire :

$$\Delta r_{i, \text{chêne}} = \alpha'_{k, \text{chêne}} + \beta'_{k, \text{chêne}} \text{circ}_i + \lambda_{\text{chêne}, \text{chêne}} \frac{N_{\text{chêne}}}{A} \pi R^2 + \lambda_{\text{chêne}, \text{pin}} \frac{N_{\text{pin}}}{A} \pi R^2 \quad (4.13)$$

Cette équation n'intègre plus la structure spatiale du peuplement mais elle intègre toujours les rapports de compétition entre espèces issus du MADD.

#### 4.4.2.2 Simplification du modèle arbre dépendant des distances pour la croissance du pin

Pour le pin, l'indice retenu dans le modèle arbre dépendant des distances est la surface terrière locale des pins. Si on note  $G_{i, \text{pin}, \text{pin}}$  la surface terrière des pins à une distance inférieure à 10 m d'un individu pin alors  $\langle CI_{i, \text{pin}, \text{pin}} \rangle = \langle G_{i, \text{pin}, \text{pin}} \rangle$ .

Pour exprimer l'espérance de la surface terrière locale, il faut prendre en compte les corrélations entre la surface terrière individuelle et la position des arbres. Pour cela, on pourrait utiliser l'autocorrélogramme ou le variogramme de la surface terrière individuelle (Goreaud, 2000 page 140), ou une fonction spécifique qui caractérise la structure spatiale de la surface terrière (Goreaud, 2000 page 166). Néanmoins, cette piste signifie que de nouvelles données seront nécessaires lors de l'utilisation du modèle pour estimer la corrélation entre la surface terrière individuelle et la position. Pour éviter cela, nous allons plutôt faire l'hypothèse que la surface terrière individuelle est indépendante de la position d'un arbre. On peut alors calculer la surface terrière locale moyenne autour d'un pin en multipliant la surface terrière individuelle moyenne d'un pin ( $\bar{g}_{pin}$ ) par le nombre moyen de voisins pins ( $\langle N_{i,pin,pin} \rangle$ ) :

$$\langle G_{i,pin,pin} \rangle = \bar{g}_{pin} \langle N_{i,pin,pin} \rangle = \bar{g}_{pin} \frac{N_{pin}}{A} K_{pin,pin}(R)$$

La surface terrière individuelle moyenne d'un pin est le rapport entre la surface terrière totale des pins et le nombre total de pins, on obtient donc finalement :

$$\langle G_{i,pin,pin} \rangle = \frac{G_{pin}}{N_{pin}} \frac{N_{pin}}{A} K_{pin,pin}(R) = \frac{G_{pin}}{A} K_{pin,pin}(R)$$

Le modèle de l'équation 4.11 peut donc s'écrire pour le pin :

$$\Delta r_{i,pin} = \alpha'_{k,pin} + \beta'_{k,pin} circ_i + \lambda_{pin,pin} \frac{G_{pin}}{A} K_{pin,pin}(R) \quad (4.14)$$

Comme pour le chêne, nous avons relié la croissance du pin à des informations sur la structure spatiale du peuplement. Si on fait l'hypothèse supplémentaire que la distribution des pins est aléatoire, on obtient :

$$\langle G_{i,pin,pin} \rangle = \frac{G_{pin}}{A} \pi R^2$$

Sous les deux hypothèses que nous venons de formuler, l'équation 4.11 pour les pins peut donc s'écrire :

$$\Delta r_{i,pin} = \alpha'_{k,pin} + \beta'_{k,pin} circ_i + \lambda_{pin,pin} \frac{G_{pin}}{A} \pi R^2 \quad (4.15)$$

#### 4.4.2.3 Croissance du mélange au niveau peuplement

L'approximation du champ moyen nous a permis de passer d'un modèle arbre dépendant des distances à un modèle arbre indépendant des distances. Nous pouvons donc utiliser les résultats de l'agrégation du modèle arbre indépendant des distances (équation 4.10) pour construire le modèle peuplement issu du modèle arbre dépendant des distances (cf partie 2).

En fonction du degré de simplification opéré, nous pouvons obtenir au moins deux versions du modèle peuplement.

Ainsi, en posant dans le système d'équations présenté au paragraphe 4.4.1 (équation 4.10):

$$\circ \beta_{k, \text{chêne}} = \beta'_{k, \text{chêne}} \quad \text{et} \quad \alpha_{k, \text{chêne}} = \alpha'_{k, \text{chêne}} + \lambda_{\text{chêne, chêne}} \frac{N_{\text{chêne}}}{A} K_{\text{chêne, chêne}}(R) + \lambda_{\text{chêne, pin}} \frac{N_{\text{pin}}}{A} K_{\text{chêne, pin}}(R) \quad \text{pour le chêne ;}$$

$$\circ \beta_{k, \text{pin}} = \beta'_{k, \text{pin}} \quad \text{et} \quad \alpha_{k, \text{pin}} = \alpha'_{k, \text{pin}} + \lambda_{\text{pin, pin}} \frac{G_{\text{pin}}}{A} K_{\text{pin, pin}}(R) \quad \text{pour le pin,}$$

on obtient une première version du modèle peuplement qui prédit l'accroissement en surface terrière du peuplement en fonction de la surface terrière des deux espèces, de la densité des deux espèces, du rayon moyen des deux espèces, des fonctions de Ripley des deux espèces et de la fonction Intertype.

Si on pose maintenant :

$$\circ \beta_{k, \text{chêne}} = \beta'_{k, \text{chêne}} \quad \text{et} \quad \alpha_{k, \text{chêne}} = \alpha'_{k, \text{chêne}} + \lambda_{\text{chêne, chêne}} \frac{N_{\text{chêne}}}{A} \pi R^2 + \lambda_{\text{chêne, pin}} \frac{N_{\text{pin}}}{A} \pi R^2 \quad \text{pour le chêne (hypothèses supplémentaires de distribution aléatoire des chênes et d'indépendance spatiale entre les deux espèces) ;}$$

$$\circ \beta_{k, \text{pin}} = \beta'_{k, \text{pin}} \quad \text{et} \quad \alpha_{k, \text{pin}} = \alpha'_{k, \text{pin}} + \lambda_{\text{pin, pin}} \frac{G_{\text{pin}}}{A} \pi R^2 \quad \text{pour le pin (hypothèse supplémentaire de distribution aléatoire des pins),}$$

on obtient une deuxième version du modèle peuplement qui prédit l'accroissement en surface terrière du peuplement sans utiliser les fonctions de caractérisation de la structure spatiale à 10 mètres.

Les deux versions du modèle peuplement que nous venons de présenter intègrent les rapports de compétition entre espèces issus du MADD. De plus, l'une des versions relie la croissance à la structure spatiale du peuplement.

#### 4.4.2.4 Modèles arbre issus du modèle arbre dépendant des distances

Notre méthode pour construire les modèles peuplement à partir du modèle arbre dépendant des distances nous a amené à définir des modèles arbre indépendants des distances. Ainsi, les équations 4.12 et 4.14 forment une première version d'un modèle arbre indépendant des distances. Ce modèle a été obtenu en appliquant l'approximation du champ moyen sur le modèle arbre dépendant des distances et en faisant l'hypothèse que la surface terrière

individuelle d'un pin est indépendante de sa position. De même, les équations 4.13 et 4.15 forment une autre version d'un modèle arbre indépendant des distances. Ce modèle fait des hypothèses supplémentaires : distribution aléatoire des chênes et des pins et indépendance spatiale entre les deux espèces. Contrairement au modèle indépendant des distances présenté dans la sous-partie 4.1 (équation 4.6), ces modèles explicitent une partie du paramètre  $\alpha$  comme une fonction de la compétition moyenne subie par les individus. Ces modèles individus-centrés ne seront pas utilisés dans le reste de cette thèse car ils sont assez redondants avec les modèles peuplement que nous avons construits. Cependant, ils pourraient permettre de tester l'importance des différentes hypothèses utilisées pour les construire au niveau de la croissance individuelle (Verzelen *et al.*, 2006).

### 4.4.3 Influence du taux de mélange sur la production du peuplement

#### 4.4.3.1 Introduction

Savoir si un peuplement mélangé est plus productif qu'un peuplement pur est une question qui intéresse les forestiers depuis longtemps (Pretzsch, 2005) et qui est toujours d'actualité (Piotto, 2008). D'un point de vue théorique, si la compétition interspécifique est supérieure à la compétition intraspécifique pour les deux espèces alors la production en peuplement mélangé sera inférieure à la production en monoculture (Forrester *et al.*, 2006). Cela peut correspondre à deux espèces ayant des niches écologiques très proches qui entrent en concurrence pour les mêmes ressources du milieu (Pretzsch, 2005). Si en revanche la compétition interspécifique est inférieure à la compétition intraspécifique pour les deux espèces alors la production en peuplement mélangé sera supérieure à la production en monoculture (Forrester *et al.*, 2006). Cette fois, cette situation peut correspondre à deux espèces qui occupent des niches écologiques différentes et qui utilisent les ressources de façon complémentaire<sup>13</sup>. Les phénomènes de facilitation conduisent également à une production supérieure en peuplement mélangé (Kelty, 2006). La figure 4.22 montre ainsi les trois grands types de réponse de la production en fonction du taux de mélange pour un peuplement composé de deux espèces (Forrester *et al.*, 2006).

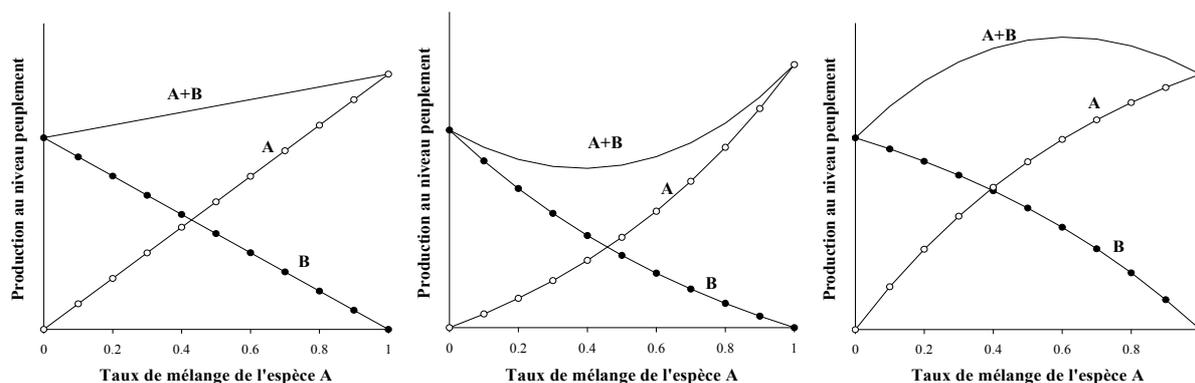


Figure 4.22 : Les trois grands types de réponse de la production d'un peuplement mélangé composé de deux espèces A et B en fonction du taux de mélange et de la nature des interactions entre espèces (d'après Harper, 1977 cité par Forrester *et al.*, 2006). A gauche les compétitions intra et interspécifiques sont identiques, il n'y a pas d'effet du mélange sur la production. Au milieu la compétition interspécifique est supérieure à la compétition intraspécifique, le mélange conduit à une réduction de la production. A droite la compétition interspécifique est inférieure à la compétition intraspécifique, le mélange conduit à une augmentation de la production.

<sup>13</sup> Il peut y avoir une ségrégation spatiale ou temporelle au niveau de l'acquisition des ressources propre à chaque espèce. C'est le cas par exemple de deux espèces qui ont leur système racinaire à des profondeurs différentes.

Le modèle arbre dépendant des distances nous a permis de montrer que, dans les peuplements mélangés chêne-pin de nos dispositifs, la compétition intraspécifique était supérieure à la compétition interspécifique (voir paragraphe 4.2). D'après ce que nous venons de dire, le mélange devrait donc apporter un gain de production par rapport à des peuplements purs. On peut alors se demander quel est le taux de mélange optimal du point de vue de la production. Pour répondre à cette question, nous pourrions faire des simulations à partir du modèle arbre dépendant des distances et étudier l'accroissement en surface terrière au niveau peuplement en fonction du taux de mélange. Cette première approche peut être lourde à entreprendre car elle nécessite de construire des peuplements virtuels ayant des taux de mélange variables pour servir d'états initiaux au modèle arbre dépendant des distances. De plus, n'ayant pas a priori d'idée sur la valeur de l'optimum, un grand nombre de simulations serait nécessaire pour le déterminer. L'objectif de cette sous-partie est de montrer que l'utilisation du modèle peuplement issu de l'agrégation du MADD, permet de répondre à la question de manière moins coûteuse, en exprimant directement le taux de mélange optimal en fonction des caractéristiques du peuplement.

#### 4.4.3.2 Méthode

Nous avons utilisé le modèle peuplement issu de l'agrégation du modèle arbre dépendant des distances. De plus, nous avons choisi la version qui utilise les fonctions caractérisant la structure spatiale à 10 m (version la moins simplifiée présentée au paragraphe 4.4.2.3). Ce modèle nécessite de connaître les caractéristiques dendrométriques du peuplement en début de période de croissance (année 1999). Nous avons donc reconstitué les circonférences 1999 de tous les arbres des dispositifs de la même façon que pour reconstituer les voisinages lors de la construction du modèle arbre dépendant des distances (voir paragraphe 4.2.1). De plus, nous n'avons pas pris en compte les espèces autres que le chêne et le pin (voir tableau 3.2).

Pour ce modèle, l'accroissement en surface terrière du peuplement sur la période 2000-2005 ( $\Delta G$  en  $\text{cm}^2$ ) s'écrit donc de la manière suivante (pour simplifier nous avons remplacé la notation  $\beta'$  par  $\beta$ ) :

$$\Delta G = A_c(\alpha_{k,c})N_c + B_c(\alpha_{k,c}, \beta_{k,c})N_c\bar{r}_c + C_c(\beta_{k,c})G_c + A_p(\alpha_{k,p})N_p + B_p(\alpha_{k,p}, \beta_{k,p})N_p\bar{r}_p + C_p(\beta_{k,p})G_p$$

avec :

$$\left\{ \begin{array}{l} A_j(\alpha_{k,j}) = (\alpha_{k,j}/10)^2 \pi \\ B_j(\alpha_{k,j}, \beta_{k,j}) = (1 + 2\pi\beta_{k,j}/10)2\alpha_{k,j}\pi/10 \\ C_j(\beta_{k,j}) = (1 + \beta_{k,j}\pi/10)\beta_{k,j}4\pi/10 \\ \alpha_{k,c} = \alpha'_{k,c} + \lambda_{c,c} \frac{N_c}{S} K_{c,c}(R) + \lambda_{c,p} \frac{N_p}{S} K_{c,p}(R) \\ \alpha_{k,p} = \alpha'_{k,p} + \lambda_{p,p} \frac{G_p}{S} K_{p,p}(R) \end{array} \right.$$

où  $k$  représente le dispositif,  $c$  l'espèce chêne,  $p$  l'espèce pin,  $j$  l'espèce,  $\alpha$ ,  $\beta$  et  $\lambda$  les paramètres du modèle arbre dépendant des distances,  $N_j$  la densité de l'espèce  $j$ ,  $G_j$  la surface terrière (en  $\text{cm}^2$ ) de l'espèce  $j$ ,  $\bar{r}_j$  le rayon moyen (en cm) de l'espèce  $j$ ,  $S$  la surface du dispositif (en  $\text{m}^2$ ),  $R$  le rayon de voisinage (10 m),  $K_{c,c}$  la fonction de Ripley pour les chênes,  $K_{c,p}$  la fonction Intertype entre les chênes et les pins et  $K_{p,p}$  la fonction de Ripley pour les pins. Pour cette version du modèle nous avons besoin de connaître les fonctions  $K_{c,c}$ ,  $K_{p,p}$  et  $K_{c,p}$  à une distance de 10 m. Nous avons donc calculé ces fonctions pour chaque dispositif (tableau 4.6).

Tableau 4.6 : Valeur des fonctions  $K_{c,c}$ ,  $K_{p,p}$  et  $K_{c,p}$  pour  $R = 10$  m pour les 9 dispositifs. Pour les indices  $K_{c,c}$  et  $K_{p,p}$  une valeur égale à  $\pi 10^2$  ( $\approx 314.16$ ) correspond à une structure aléatoire de l'espèce considérée, une valeur supérieure correspond à une structure agrégée et une valeur inférieure correspond à une structure régulière. Pour l'indice  $K_{c,p}$  une valeur égale à  $\pi 10^2$  correspond à une indépendance entre les deux espèces, une valeur supérieure correspond à une attraction et une valeur inférieure correspond à une répulsion.

Dispositif	D02	D108	D20	D27	D42	D49	D534	D563	D78
$K_{c,c}$	349.4	330.8	350.7	346.0	346.0	328.4	312.9	304.9	333.8
$K_{p,p}$	444.3	321.4	404.0	504.8	323.3	381.5	364.9	321.0	378.2
$K_{c,p}$	281.9	294.3	291.1	308.0	316.2	311.7	343.9	308.8	304.7

Pour définir la proportion de chacune des espèces dans un peuplement mélangé, il faut définir une variable de référence permettant de quantifier leur abondance dans le peuplement. On peut par exemple utiliser le nombre de tiges mais, dans ce cas, on accorde la même importance aux petits individus d'une espèce qu'aux gros individus d'une autre espèce. Pour

éviter ce genre de problème, on utilise des variables qui sont reliées au volume ou à la biomasse des populations comme par exemple la masse de matière sèche (Pretzsch, 2005). Ici nous allons utiliser la surface terrière qui permet de prendre en compte à la fois le nombre et le diamètre des individus. Pour un peuplement composé de deux espèces, la proportion ou taux de mélange d'une espèce  $j$  sera alors défini comme étant le rapport entre la surface terrière de l'espèce et la surface terrière totale :  $x_j = \frac{G_j}{G}$ . De plus nous allons introduire le rayon quadratique moyen  $\bar{r}_{G,j}$  qui permet de faire le lien entre la densité d'une espèce  $j$  et sa surface terrière :  $G_j = N_j \pi \bar{r}_{G,j}^2$ . Nous pouvons donc exprimer l'accroissement en surface terrière du peuplement en fonction par exemple du taux de mélange des chênes que nous noterons  $x$  pour simplifier :

$$\begin{aligned} \Delta G(x) = & A_c(\alpha_{k,c}) \frac{Gx}{\pi \bar{r}_{G,c}^2} + B_c(\alpha_{k,c}, \beta_{k,c}) \bar{r}_c \frac{Gx}{\pi \bar{r}_{G,c}^2} + C_c(\beta_{k,c}) Gx + \\ & A_p(\alpha_{k,p}) \frac{G(1-x)}{\pi \bar{r}_{G,p}^2} + B_p(\alpha_{k,p}, \beta_{k,p}) \bar{r}_p \frac{G(1-x)}{\pi \bar{r}_{G,p}^2} + C_p(\beta_{k,p}) G(1-x) \end{aligned} \quad (4.16)$$

avec

$$\alpha_{k,c} = \alpha'_{k,c} + \lambda_{c,c} \frac{1}{S} K_{c,c}(R) \frac{Gx}{\pi \bar{r}_{G,c}^2} + \lambda_{c,p} \frac{1}{S} K_{c,p}(R) \frac{G(1-x)}{\pi \bar{r}_{G,p}^2} \quad \text{et} \quad \alpha_{k,p} = \alpha'_{k,p} + \lambda_{p,p} \frac{1}{S} K_{p,p}(R) G(1-x)$$

Les modèles que nous avons construits mettent en relation l'accroissement et la taille des arbres. Dans un dispositif et pour une espèce, les paramètres du modèle sont valables pour une certaine gamme de tailles. Ainsi, nous avons vu qu'en général l'accroissement en rayon des pins était inférieur à celui des chênes mais que les pins étaient plus gros que les chênes. Nous avons également vu que lorsque ces mêmes pins avaient un diamètre inférieur et donc un âge inférieur, leurs accroissements étaient plus forts. Pour respecter le domaine de validité de nos modèles, nous considérons que les caractéristiques dendrométriques des espèces ne changent pas lorsque le taux de mélange varie. De cette façon, lorsque le taux de mélange  $x$  varie, la surface terrière totale, les rayons moyens de chaque espèce, ainsi que les rayons quadratiques moyens de chaque espèce restent constants. Seul le nombre de tiges change pour conserver une surface terrière constante.

Pour rechercher la valeur  $x$  qui maximise la production du peuplement, nous avons calculé la dérivée de la fonction  $\Delta G(x)$ . En effet, si la fonction  $\Delta G(x)$  atteint un minimum ou un maximum en  $x_0$  dans l'intervalle  $[0, 1]$  alors nécessairement sa dérivée est nulle pour  $x = x_0$ . Pour que cette valeur soit un maximum, il faut en plus que la dérivée soit positive ou nulle sur  $[0, x_0]$  et négative ou nulle sur  $[x_0, 1]$ . La dérivée de  $\Delta G(x)$  est une équation du second

degré en  $x$  (voir annexe D). Pour trouver la valeur optimale de  $x$ , nous avons donc déterminé les solutions de cette équation comprises entre 0 et 1.

#### 4.4.3.3 Résultats

Pour chaque dispositif, il existe bien un taux de mélange en chêne compris entre 0 et 1 qui maximise l'accroissement en surface terrière du peuplement (voir annexe D). Ce taux s'exprime directement en fonction des paramètres du modèle peuplement, des caractéristiques dendrométriques de chacune des espèces, de la surface terrière totale et des caractéristiques spatiales du peuplement (voir annexe D). Le tableau 4.7 présente, pour chaque dispositif, le taux de mélange en chêne qui maximise l'accroissement en surface terrière du peuplement. Ce taux est compris entre 38% et 74% pour une moyenne de 50%.

Tableau 4.7 : Taux de mélange optimal pour les différents dispositifs. Disp = dispositif ;  $S$  = surface du dispositif ;  $x_{max}$  = taux de chêne optimal ;  $\Delta G(x_{max})$  = accroissement du peuplement pour  $x = x_{max}$  ;  $x_{disp}$  = taux de chêne dans le dispositif ;  $\Delta G(x_{disp})$  = accroissement du peuplement pour  $x = x_{disp}$  ; MOY = moyenne des différentes variables sur l'ensemble des dispositifs. Gain = gain en pourcentage sur 6 années de croissance par rapport au taux de mélange observé dans le dispositif =  $100 * (\Delta G(x_{max}) - \Delta G(x_{disp})) / \Delta G(x_{disp})$ .

Disp	$S$ (m <sup>2</sup> )	$\bar{r}_{G,c}$ (cm)	$\bar{r}_{G,p}$ (cm)	$\bar{r}_c$ (cm)	$\bar{r}_p$ (cm)	$G$ (m <sup>2</sup> )	$x_{max}$	$\Delta G(x_{max})$ (m <sup>2</sup> /ha/an)	$x_{disp}$	$\Delta G(x_{disp})$ (m <sup>2</sup> /ha/an)	Gain
D02	9513	11.1	17.7	10.0	17.3	23.3	59.3%	0.477	59.2%	0.477	0.00%
D108	8000	8.5	16.9	8.0	16.7	28.9	43.1%	0.591	28.0%	0.561	4.91%
D20	10151	8.1	16.2	7.5	15.9	23.3	37.5%	0.555	42.6%	0.552	0.61%
D27	6250	8.2	17.7	7.3	17.3	21.0	46.5%	0.295	39.8%	0.292	0.96%
D42	5000	8.2	12.5	7.7	12.0	23.5	45.7%	0.374	41.9%	0.373	0.21%
D49	9940	8.7	15.2	8.0	14.8	29.0	46.2%	0.647	40.8%	0.642	0.64%
D534	5000	8.2	18.1	7.6	17.9	27.9	40.8%	0.291	37.0%	0.290	0.40%
D563	5000	12.3	16.2	11.4	16.1	28.9	73.6%	0.320	39.6%	0.291	9.05%
D78	7000	9.7	20.9	9.1	20.5	27.5	57.3%	0.401	43.6%	0.387	3.41%
MOY		9.2	16.8	8.5	16.5	25.9	50.2%	0.607	41.4%	0.599	1.30%

Pour illustrer ces résultats, nous avons représenté l'évolution de l'accroissement en surface terrière du peuplement en fonction du taux de mélange pour 4 dispositifs (figure 4.23).

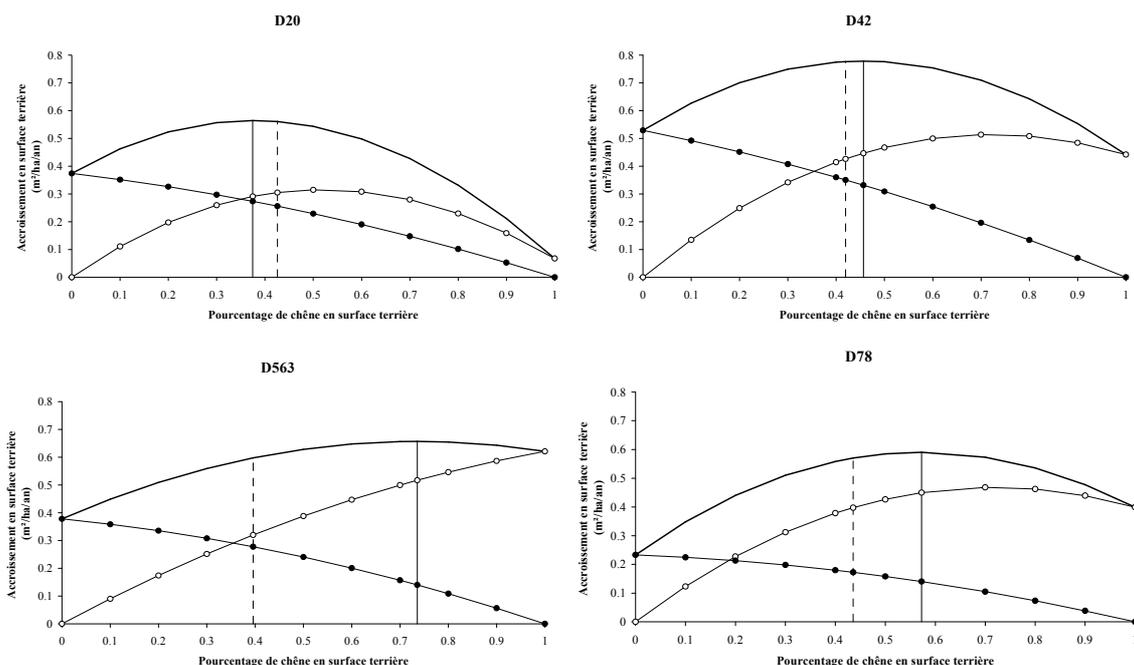


Figure 4.23 : Accroissement en surface terrière en fonction du taux de mélange pour les dispositifs D20, D42, D563 et D78. La courbe noire représente l'accroissement en surface terrière totale, les points noirs représentent l'accroissement en surface terrière des pins et les points blancs représentent l'accroissement en surface terrière des chênes. La droite pleine verticale représente le taux de mélange optimal, la droite verticale en pointillés représente le taux de mélange réel du dispositif.

#### 4.4.3.4 Discussion et conclusion

Nous pouvons constater qu'il y a une certaine variabilité sur la valeur de l'optimum (tableau 4.7). Nous pouvons expliquer de manière qualitative cette variabilité entre dispositifs en étudiant l'impact des différents paramètres sur l'optimum. Certains paramètres sont constants quels que soient les dispositifs, en particulier les paramètres liés aux indices de compétition qui conditionnent le rapport entre compétition intra et interspécifique. D'autres paramètres changent d'un dispositif à l'autre : des paramètres de croissance ( $\alpha'_{k,c}$ ,  $\beta_{k,c}$  et  $\beta_{k,p}$ ), le diamètre moyen des chênes et des pins, la surface terrière totale et enfin la structure spatiale. Nécessairement, tout changement d'un facteur qui favorise la production d'une espèce déplace l'optimum vers un taux de mélange où cette espèce est plus représentée. Ainsi, si on raisonne avec des paramètres de croissance constants, sachant que la production d'une espèce est liée positivement à sa taille moyenne, alors plus le rapport entre le diamètre moyen des chênes et celui des pins augmente, plus le taux optimal de mélange en chêne augmente. Si maintenant on raisonne avec des diamètres moyens constants, alors plus le rapport entre le rendement des chênes (paramètre  $\beta_{k,c}$ ) et celui des pins (paramètre  $\beta_{k,p}$ ) augmente, plus le taux optimal de mélange en chêne augmente. Si les paramètres de croissance sont constants et si les diamètres moyens sont constants, sachant que la surface terrière totale a un impact plus

négatif sur la production des chênes que sur la production des pins (à travers les paramètres  $\lambda_{c,c}$  et  $\lambda_{c,p}$ ), alors plus la surface terrière totale augmente plus le taux optimal de mélange en chêne diminue. Enfin, la structure spatiale joue également sur l'optimum. Plus les chênes sont agrégés (indice  $K_{c,c} > \pi 10^2$ ), plus le nombre de voisins à une distance inférieure à 10 m sera important et plus la croissance des chênes sera réduite. L'optimum se déplace alors vers un mélange avec plus de pins. Le même raisonnement peut être mené avec les pins. La structure interspécifique à 10 m intervient également sur la croissance des chênes. Plus il y a attraction entre les pins et les chênes (indice  $K_{c,p} > \pi 10^2$ ), plus il y a de pins à une distance inférieure à 10 m d'un chêne et plus la croissance des chênes est réduite. L'optimum se déplace alors vers un mélange avec plus de pins. Une analyse de sensibilité globale (voir partie suivante, paragraphe 5.1.2.1) permettrait de savoir quels facteurs ont le plus d'influence sur l'optimum.

Tableau 4.8 : Valeurs des paramètres du modèle peuplement pour chaque dispositif.

Dispositif	Chêne				Pin		
	$\alpha'_{k,c}$	$\beta_{k,c}$	$\lambda_{c,c}$	$\lambda_{c,p}$	$\alpha'_{k,p}$	$\beta_{k,p}$	$\lambda_{p,p}$
D02	5.99	0.103	-0.354	-0.242	4.603	0.0566	-0.000361
D108	12.09	0.090	-0.354	-0.242	4.603	0.0620	-0.000361
D20	12.46	0.043	-0.354	-0.242	4.603	0.0633	-0.000361
D27	12.74	0.036	-0.354	-0.242	4.603	0.0512	-0.000361
D42	9.62	0.163	-0.354	-0.242	4.603	0.0813	-0.000361
D49	13.01	0.060	-0.354	-0.242	4.603	0.0511	-0.000361
D534	7.73	0.136	-0.354	-0.242	4.603	0.0476	-0.000361
D563	6.00	0.107	-0.354	-0.242	4.603	0.0491	-0.000361
D78	4.00	0.177	-0.354	-0.242	4.603	0.0337	-0.000361

Même si la différence entre le taux de mélange optimal et le taux réel observé sur les dispositifs peut être importante (34% pour le dispositif D563, tableau 4.7), le gain de production sur 6 ans entre ces deux taux est relativement faible. Il est en moyenne de 2.2% avec un maximum de 9% (tableau 4.7).

Une autre question concerne le gain du peuplement mélangé par rapport à un peuplement pur. Il est difficile de répondre à cette question car le modèle donné par l'équation 4.16 a été construit pour des taux de mélange compris entre 28% et 59%. En dehors de cette gamme et en particulier pour les taux de mélange extrêmes, le comportement du modèle n'est pas garanti (voir par exemple le dispositif D20 sur la figure 4.23). De plus, comme nous l'avons dit en introduction, le gain de production du mélange par rapport à un peuplement pur dépend du rapport entre la compétition intraspécifique et la compétition interspécifique. Or nous avons vu dans la sous-partie 4.2 que ce rapport varie de 1.5 à 3 en fonction des indices de

compétition (paragraphe 4.2.2). Le choix de l'indice peut donc jouer sur la valeur de l'optimum. Enfin, il est possible que le rapport de compétition entre espèces varie d'un dispositif à l'autre ce que ne prend pas en compte le modèle.

Malgré les limites que nous avons précisées, le modèle peuplement issu de l'agrégation du modèle arbre dépendant des distances est un cadre intéressant, car il nous a permis d'étudier de manière très directe l'impact du taux de mélange sur la production. Nous avons ainsi montré qu'il était possible d'exprimer le taux de mélange qui maximise la production du peuplement en fonction des paramètres du modèle, des caractéristiques dendrométriques et spatiales du peuplement. Pour ce qui concerne nos dispositifs, nous avons montré que la production maximale était obtenue en moyenne pour un taux de mélange d'environ 50%. Ce travail permet d'illustrer, sur une question, ce que peut apporter le modèle peuplement par rapport aux autres modèles que nous avons présentés dans cette partie.

#### 4.4.4 A propos de l'erreur d'agrégation

Nous venons de voir que l'on pouvait construire des modèles de croissance à l'échelle du peuplement en agrégeant des modèles à l'échelle de l'individu. Nous allons voir qu'en fonction des cas rencontrés, la dynamique du peuplement décrite par le modèle peuplement peut être identique ou différente de la dynamique du peuplement obtenue en appliquant le modèle à l'échelle de l'individu sur l'ensemble des arbres du peuplement.

Jusqu'à maintenant, pour décrire l'accroissement individuel, nous avons utilisé un modèle reliant l'accroissement en rayon à la circonférence. Nous avons vu que pour décrire la croissance au niveau du peuplement, nous avons choisi l'accroissement en surface terrière car cette variable rend mieux compte de la production. Nous utiliserons donc ici l'accroissement individuel en surface terrière pour discuter du lien entre le niveau arbre et le niveau peuplement sachant que l'accroissement individuel en surface terrière peut être obtenu à partir du modèle d'accroissement en rayon (voir annexe A).

Dans le cas de l'agrégation du modèle arbre indépendant des distances, nous n'avons pas fait d'approximation pour obtenir le modèle peuplement. En conséquence, au niveau peuplement, le modèle arbre indépendant des distances et le modèle peuplement donneront les mêmes résultats (figure 4.24). Dans le cadre de la théorie de l'agrégation, on dit qu'il y a agrégation parfaite (Iwasa *et al.*, 1987 ; Picard, 1999 pages 238 à 241). Ainsi, si l'on souhaite des résultats à l'échelle du peuplement et si on dispose des variables agrégées adéquates, le modèle peuplement est suffisant.

$$\begin{array}{ccc}
 g_i & \xrightarrow{\gamma} & \frac{\Delta g_i}{\Delta t} \\
 \downarrow \Sigma & & \downarrow \Sigma \\
 G & \xrightarrow{\Gamma} & \frac{\Delta G}{\Delta t} \\
 \Gamma(\sum g_i) & = & \sum \gamma(g_i)
 \end{array}$$

Figure 4.24 : Schéma d'agrégation dans le cas du modèle arbre indépendant des distances. La fonction  $\gamma$  est le modèle qui relie la surface terrière individuelle à l'accroissement individuel en surface terrière. La fonction  $\Gamma$  est le modèle qui relie la surface terrière du peuplement à l'accroissement en surface terrière du peuplement. Dans ce cas, l'accroissement de la somme des surfaces terrières individuelles est égal à la somme des accroissements individuels, il y a agrégation parfaite.

Dans le cas de l'agrégation du modèle arbre dépendant des distances, nous avons été contraints de faire des approximations (approximation du champ moyen et différentes hypothèses simplificatrices). Dans ce cas, le modèle arbre dépendant des distances et le modèle peuplement ne donneront pas les mêmes résultats au niveau de l'accroissement en surface terrière du peuplement (figure 4.25). On peut alors estimer la différence entre les deux modèles. On appelle cette différence l'erreur d'agrégation (Picard, 1999 page 239). En agrégeant le modèle arbre dépendant des distances, nous avons vu que l'on pouvait construire deux versions du modèle peuplement correspondant à des degrés de simplification différents. L'erreur d'agrégation devrait donc être moins importante avec la version utilisant les fonctions caractérisant la structure spatiale à 10 m.

$$\begin{array}{ccc}
 g_i & \xrightarrow{\gamma'} & \frac{\Delta g_i}{\Delta t} \\
 \downarrow \Sigma & & \downarrow \Sigma \\
 G & \xrightarrow{\Gamma'} & \frac{\Delta G}{\Delta t} \\
 \Gamma'(\sum g_i) \neq \sum \gamma'(g_i) & & 
 \end{array}$$

Figure 4.25 : Schéma d'agrégation dans le cas du modèle arbre dépendant des distances. La fonction  $\gamma'$  est le modèle qui relie la surface terrière individuelle à l'accroissement individuel en surface terrière. La fonction  $\Gamma'$  est le modèle qui relie la surface terrière du peuplement à l'accroissement en surface terrière du peuplement. Dans ce cas l'accroissement de la somme des surfaces terrières individuelles est différent de la somme des accroissements individuels.

## 4.5 Bilan

Dans cette partie nous avons construit une famille de modèles. Nous avons construit 4 grands types de modèles mais au final nous avons été amenés à définir 8 modèles différents (figure 4.26) :

- 1 modèle arbre indépendant des distances (MAID, équation 4.6) ;
- 1 modèle arbre dépendant des distances (MADD, équation 4.8) ;
- 1 modèle arbre utilisant des distributions de voisinage (MADV, équation 4.9) ;
- 2 versions d'un modèle arbre obtenues à partir du modèle arbre dépendant des distances par approximation du champ moyen ( $MASS_{MADD}$  équations 4.12 et 4.14 et  $MA_{MADD}$ , équations 4.13 et 4.15) ;
- 2 versions d'un modèle peuplement obtenues par agrégation du modèle arbre dépendant des distances ( $MP_{MADD}$  et  $MPSS_{MADD}$ , paragraphe 4.4.2.3) ;
- 1 modèle peuplement obtenu par agrégation du modèle arbre indépendant des distances ( $MP_{MAID}$ , équation 4.10).

Comme nous l'avons dit dans le paragraphe 4.4.2.4, le  $MASS_{MADD}$  et le  $MA_{MADD}$  ne seront pas utilisés dans le reste de cette thèse

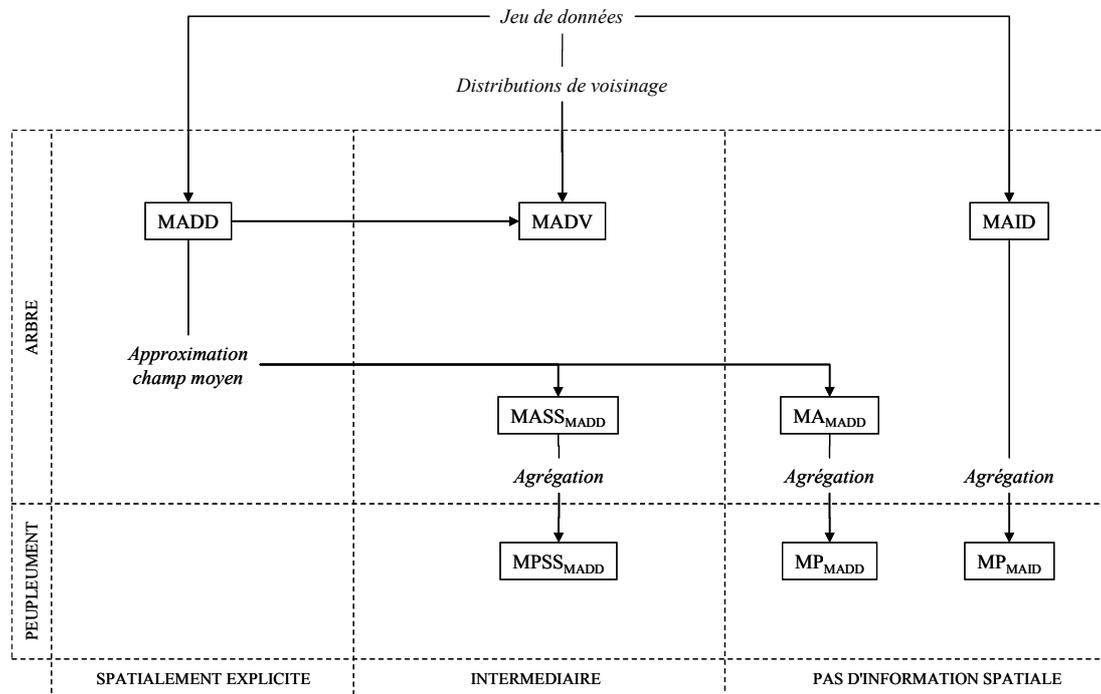


Figure 4.26 : Récapitulatif des modèles construits en fonction du niveau de description du peuplement (modèles arbre et peuplement) et en fonction de la dépendance spatiale (modèles spatialement explicites, modèles sans information spatiale explicite et modèles intermédiaires). Les expressions en italiques sont les différents moyens utilisés pour assurer la cohérence entre les modèles. MAID = Modèle Arbre Indépendant des Distances ; MADD = Modèle Arbre Dépendant des Distances ; MADV = Modèle Arbre utilisant des Distributions de Voisinage ; MA = Modèle Arbre ; MP = Modèle Peuplement ; SS = Structure Spatiale.

Ces modèles nous ont permis d'apporter des connaissances sur la croissance et la dynamique des peuplements mélangés chêne-pin. Le modèle arbre indépendant des distances a permis de montrer que la taille de l'arbre est une variable explicative de la croissance satisfaisante pour le chêne ce qui est moins vrai pour le pin. En utilisant les accroissements passés, une version simplifiée de ce modèle a également mis en évidence une compensation de croissance entre les deux essences suite à l'attaque d'un agent biotique. Ce phénomène montre que les deux espèces peuvent interagir de manière plus ou moins forte au cours du temps. Le modèle arbre dépendant des distances nous a permis de montrer que, dans ces peuplements, la compétition intraspécifique était supérieure à la compétition interspécifique ce qui peut être une condition permettant aux espèces de coexister. Ce résultat a également un impact sur la production du peuplement. Ainsi, nous avons vu qu'un modèle peuplement issu du modèle dépendant des distances permettait d'étudier de façon analytique l'influence du taux de mélange sur la production. Ainsi, le taux de mélange qui maximise la production n'est pas unique et se situe entre 38% et 74% de la surface terrière totale pour les dispositifs étudiés.

Nous avons vu que des distributions de voisinage permettaient d'utiliser l'équation de croissance du modèle arbre dépendant des distances sans utiliser une cartographie du peuplement. Ce type de modèle pourrait faciliter l'utilisation des résultats obtenus avec les modèles arbre dépendants des distances et pourrait se placer comme une alternative au couplage entre modèle dépendant des distances et modèle de génération de peuplements virtuels (Ngo Bieng, 2007).

Les modèles que nous avons construits sont des modèles simplifiés dont il faut rappeler les limites. Ils ont été ajustés sur une période de croissance de 6 ans allant de l'année 2000 à l'année 2005. Or, nous avons vu que les paramètres des modèles évoluent au cours du temps. Des simulations à long terme ne sont donc pas envisageables pour le moment. L'exploitation des données rétrospectives devrait permettre d'améliorer en partie ce point. De plus, ces modèles n'intègrent pas les processus de mortalité et de régénération. Cette simplification était recherchée pour répondre à l'objectif fixé dans cette thèse. Néanmoins, l'intégration de ces processus est importante pour pouvoir tester des sylvicultures variées. Enfin, nous n'avons parlé que de la croissance en diamètre et pas de la croissance en hauteur. Pour étudier les phénomènes de compétition, la croissance en diamètre reste la variable la mieux adaptée (Piotto, 2008) car la croissance en hauteur est fortement influencée par la productivité de la station (Oliver et Larson, 1996 page 60). Cependant, les données de hauteur concernant les arbres du voisinage pourraient permettre d'améliorer la définition des indices de compétition

(voir par exemple Pretzsch *et al.*, 2002a) surtout si on s'intéresse à la compétition pour la lumière.

Le MADD et le MADV ont été implémentés sur la plate-forme Capsis (Coligny *et al.*, 2003). Deux nouveaux modules ont été créés<sup>14</sup> : le module oakpine1 pour le MADD et le module oakpine2 pour le MADV. L'implémentation du module oakpine2 est présentée dans l'annexe C. Pour le MAID nous n'avons pas créé de module spécifique pour le moment. Les simulations du MAID que nous verrons dans la partie 5 ont été réalisées à partir d'oakpine2 en modifiant l'équation de croissance du module<sup>15</sup>. Les modèles peuplement ont été intégrés dans des tableurs Excel. Il serait intéressant de regrouper l'ensemble de ces modèles dans un même module ou de les rendre accessibles depuis une interface unique.

Pour assurer une cohérence entre ces modèles, nous avons utilisé plusieurs méthodes : ajustement des modèles sur le même jeu de données, utilisation des distributions de voisinage (moyen permettant le changement de niveau de détail), application de l'approximation du champ moyen et agrégation de modèles individus-centrés (figure 4.26). Dans la partie suivante, nous allons évaluer et comparer les modèles par rapport à la prédiction de la croissance. La cohérence entre les modèles nous permettra d'interpréter les différences entre les modèles comme des différences liées au niveau de détail choisi pour modéliser la croissance.

---

<sup>14</sup> Les codes de ces modules sont disponibles sur demande ou directement sur internet pour les membres du projet Capsis (voir <http://capsis.cirad.fr/>).

<sup>15</sup> Nous avons vérifié que le MAID (équation de croissance n'utilisant pas d'indice de compétition) donne exactement les mêmes résultats en utilisant oakpine1 ou oakpine2 ce qui est une façon de vérifier la cohérence de l'implémentation de ces modules.

## 5. Evaluation et comparaison des modèles pour prédire la croissance

### - Sommaire de la partie -

5.1	Critères et méthodes utilisables pour évaluer des modèles de croissance forestiers .....	114
5.1.1	<i>Evaluer la qualité des prédictions par rapport à des données d'évaluation</i> .....	114
5.1.2	<i>Sensibilité, robustesse et propagation d'erreurs</i> .....	118
5.1.3	<i>Critères informatiques : temps d'exécution et taille mémoire utilisée</i> .....	121
5.1.4	<i>Modélisation et évaluation participatives</i> .....	123
5.1.5	<i>Bilan et approches retenues pour évaluer et comparer les modèles</i> .....	124
5.2	Prédire les accroissements individuels, leur variabilité et leur distribution .....	129
5.2.1	<i>Présentation des données individuelles d'évaluation</i> .....	129
5.2.2	<i>Comment évaluer le modèle utilisant des distributions de voisinage ?</i> .....	132
5.2.3	<i>Indices et tests retenus pour évaluer et comparer les modèles</i> .....	134
5.2.4	<i>Résultats</i> .....	137
5.2.5	<i>Discussion</i> .....	140
5.3	Prédire l'accroissement du peuplement .....	142
5.3.1	<i>Comment estimer l'accroissement à l'échelle du peuplement ?</i> .....	142
5.3.2	<i>Résultats sur l'estimation des accroissements au niveau du peuplement</i> .....	145
5.3.3	<i>Méthodes retenues pour évaluer et comparer les modèles</i> .....	146
5.3.4	<i>Résultats</i> .....	148
5.3.5	<i>Discussion</i> .....	150
5.4	Prédire l'évolution du peuplement dans des contextes appliqués .....	154
5.4.1	<i>Estimation de la production et des caractéristiques dendrométriques à court terme</i> .....	154
5.4.2	<i>Evolution du peuplement à moyen terme</i> .....	157
5.4.3	<i>Estimation de l'impact du type de structure spatiale sur la production</i> .....	164
5.5	Bilan .....	170

Dans la partie précédente, nous avons construit une famille de modèles de croissance adaptés aux peuplements mélangés chêne sessile - pin sylvestre (tableau 5.1). Nous avons vu que chacun de ces modèles pouvait apporter des éléments de réponse ou de réflexion sur la croissance du mélange chêne-pin. Pour une même question, nous pouvons maintenant considérer qu'il est possible d'utiliser plusieurs modèles de croissance ayant des niveaux de détail différents. L'objectif de cette partie est d'évaluer et de comparer ces modèles en tant qu'outils prédictifs. Grâce à la cohérence que nous avons assurée entre les modèles, les différences que l'on observera en comparant les modèles proviendront essentiellement du niveau de détail choisi pour modéliser la croissance du peuplement mélangé. Cela nous permettra de déterminer le modèle le plus adapté à un objectif donné.

Tableau 5.1 : Abréviations et définitions des modèles évalués et comparés dans la partie 5.

<b>Modèle</b>	<b>Définition</b>
MAID	Modèle Arbre Indépendant des Distances
MADD	Modèle Arbre Dépendant des Distances
MADV	Modèle Arbre avec Distributions de Voisinage
MP <sub>MAID</sub>	Modèle Peuplement obtenu par agrégation du MAID
MP <sub>MADD</sub>	Modèle Peuplement obtenu par agrégation du MADD n'utilisant pas les fonctions de caractérisation de la structure spatiale du peuplement
MPSS <sub>MADD</sub>	Modèle Peuplement obtenu par agrégation du MADD utilisant les fonctions de caractérisation de la Structure Spatiale du peuplement

Nous commencerons cette partie par une présentation des différents critères et des différentes méthodes que l'on peut utiliser pour évaluer des modèles de croissance forestiers (sous-partie 5.1). A la fin de cette première sous-partie nous ferons un bilan concernant les critères et les méthodes d'évaluation retenus dans notre travail puis nous présenterons les approches que nous avons choisies de mettre en œuvre pour comparer nos modèles (paragraphe 5.1.5). Les sous-parties 5.2, 5.3 et 5.4 seront consacrées à l'évaluation et à la comparaison des modèles. Nous terminerons par un bilan récapitulatif des différents résultats obtenus dans cette partie (sous-partie 5.5).

## 5.1 Critères et méthodes utilisables pour évaluer des modèles de croissance forestiers

L'objectif de cette sous-partie est de présenter les critères et méthodes utilisés pour évaluer des modèles de croissance forestiers. Il existe de nombreux articles ou ouvrages concernant les procédures d'évaluation des modèles forestiers (Mayer et Butler, 1993 ; Soares *et al.*, 1995 ; Gourlet-Fleury, 1997 ; Vanclay et Skovsgaard, 1997 ; Huang *et al.*, 2002 ; Grimm et Railsback, 2004 ; Schmidt *et al.*, 2006). En effet l'évaluation d'un modèle fait partie intégrante de sa construction (Pavé, 1994 page 33). En nous appuyant sur ces travaux, nous avons distingué quatre grands types de critères :

- les critères et méthodes permettant d'évaluer la qualité des prédictions (paragraphe 5.1.1) ;
- les critères et méthodes permettant d'évaluer la sensibilité d'un modèle, sa robustesse ainsi que la propagation des erreurs (paragraphe 5.1.2) ;
- les critères et méthodes permettant d'évaluer les performances informatiques du simulateur qui accueille le modèle (paragraphe 5.1.3) ;
- les critères et méthodes permettant d'évaluer un modèle par rapport aux objectifs d'utilisation (paragraphe 5.1.4).

### 5.1.1 Evaluer la qualité des prédictions par rapport à des données d'évaluation

Confronter les données prédites par un modèle à d'autres données est certainement la démarche la plus simple pour évaluer un modèle. Différentes méthodes peuvent alors être utilisées pour apprécier la qualité prédictive du modèle. Nous avons distingué les méthodes basées sur des analyses graphiques et les méthodes basées sur l'utilisation d'indices et de tests statistiques. Nous commençons ce paragraphe par une présentation sur les données utilisées pour évaluer les modèles.

#### 5.1.1.1 Les données d'évaluation

Un des aspects de l'évaluation d'un modèle consiste à comparer les données prédites par ce modèle à des données provenant d'une autre source. Ces données sont appelées données d'évaluation. Les données d'évaluation peuvent être des données observées qui ont servi à la calibration des modèles mais il est préférable qu'elles soient différentes (Gourlet-Fleury, 1997 page 73) en particulier pour estimer la capacité prédictive des modèles (Vanclay et Skovsgaard, 1997 ; Astrup *et al.*, 2008). En effet, l'utilisation des données de calibration pour

l'évaluation d'un modèle conduit nécessairement à des conclusions optimistes<sup>16</sup>. Lorsque les données d'évaluation sont différentes des données de calibration, on dit qu'elles sont "indépendantes" (Vanclay et Skovsgaard, 1997). Néanmoins, il est parfois difficile de réunir suffisamment de données pour constituer plusieurs jeux de données. De plus, les modèles de croissance forestiers utilisent généralement des données sur le long terme qui sont plutôt rares. Il est alors possible d'utiliser des techniques de rééchantillonnage pour estimer les paramètres et la qualité prédictive du modèle à partir d'un seul jeu de données (Meredieu, 1998 page 190). Enfin, dans certains cas, les données d'évaluation peuvent être des données simulées par un modèle de référence (voir par exemple Gourlet-Fleury *et al.*, 2005).

#### 5.1.1.2 Evaluer la qualité des prédictions à l'aide de représentations graphiques

Les représentations graphiques permettent d'observer la différence entre des données d'évaluation observées et des données prédites par un modèle. Il existe un grand nombre de façons de représenter ces données. Nous allons en présenter quelques unes puis nous discuterons l'intérêt de l'approche graphique pour évaluer les modèles.

##### *Représentation des observations en fonction des prédictions et analyse des résidus*

Une façon très classique d'évaluer un modèle est de représenter les données observées en fonction des données prédites. Lorsque le modèle se comporte correctement, les points se situent autour de la bissectrice. L'analyse graphique des écarts entre les valeurs observées et les valeurs prédites est un autre moyen pour s'assurer du bon comportement du modèle. Lorsque les données d'évaluation sont les données de calibration on parle de résidus. Une première étape consiste à représenter les résidus du modèle en fonction des valeurs prédites. Cela peut permettre de mettre en évidence des anomalies comme un lien entre les résidus et les valeurs prédites (présence d'une relation linéaire ou non) ou une non-homogénéité de la variance (hétéroscédasticité). En fonction de l'importance des résultats de cette analyse, une modification du modèle peut être effectuée (changement de la forme du modèle, modélisation de la variance, etc.). Dans une deuxième étape, la représentation des résidus en fonction de toutes les variables prédictives peut permettre de détecter les auto-corrélations et les corrélations résiduelles (Meredieu, 1998).

---

<sup>16</sup> En particulier dans le cas d'un modèle linéaire ajusté par la méthode des moindres carrés, par construction la somme des écarts est nulle et la somme du carré des écarts est minimum.

### *Représentations propres aux modèles de croissance*

Les modèles auxquels nous nous intéressons sont des outils qui cherchent à prédire l'évolution d'un peuplement sur une durée plus ou moins longue. En conséquence ces modèles sont généralement évalués au bout de plusieurs années de simulation. On peut alors représenter l'évolution d'une variable au cours du temps et évaluer visuellement les résultats obtenus par le modèle. Cela permet parfois de mettre en évidence des comportements aberrants par rapport aux connaissances sur le système étudié (voir par exemple Gourlet-Fleury *et al.*, 2005) et suffit pour apporter des conclusions claires.

Certains auteurs ont également développé des représentations graphiques particulières pour évaluer les résultats des modèles de croissance. Par exemple, afin de visualiser la dynamique de croissance des arbres, Matala *et al.* (2003) représentent le rapport entre l'accroissement en volume d'un arbre et l'accroissement maximum observé dans le même peuplement en fonction d'une variable indiquant la taille relative de l'arbre (par exemple le diamètre individuel sur le diamètre maximum).

### *Discussion sur l'approche graphique*

Huang *et al.* (2002) insistent sur l'importance de l'évaluation des modèles à l'aide de graphiques en particulier lorsque l'on cherche à évaluer des modèles complexes constitués de plusieurs sous-modèles. En effet, selon ces auteurs, les représentations graphiques peuvent permettre de montrer rapidement la qualité des prédictions et sont plus faciles à interpréter que des indices. Le principal inconvénient de l'approche graphique est sa subjectivité (Huang *et al.*, 2002). Sur un même jeu d'évaluation, des personnes différentes n'arriveront pas aux mêmes conclusions. Il peut donc être nécessaire de combiner l'approche graphique avec des méthodes statistiques.

#### 5.1.1.3 Evaluer la qualité des prédictions à l'aide d'indices et de tests statistiques

Les indices numériques et les tests statistiques sont les méthodes les plus utilisées pour évaluer les modèles. Comme les représentations graphiques, les indices permettent d'évaluer la distance entre des valeurs observées et des valeurs prédites par un modèle. Contrairement aux graphes, ils donnent une mesure numérique de cette distance et des tests ont été développés pour indiquer si cette distance est statistiquement significative ou non. Encore une fois, il est préférable que les données d'évaluation soient indépendantes des données de calibration.

### *Evaluer la qualité de prédiction d'une variable*

Les indices et les tests statistiques permettant d'évaluer les modèles par rapport à la prédiction d'une variable sont très nombreux (Balci et Sargent (1984) cité par Huang *et al.*, 2002). On peut distinguer les indices et les tests applicables dans le cas d'un jeu de données indépendant et ceux utilisables directement sur le jeu de données de calibration (annexe E).

### *Evaluer la qualité de prédiction de la distribution d'une variable*

La bonne représentation de la variabilité du phénomène observé est une dimension importante à prendre compte pour évaluer les modèles. L'annexe E présente des indices et des tests statistiques permettant d'évaluer les modèles par rapport à la prédiction de la distribution d'une variable.

Schmidt *et al.* (2006) notent le fait que les tests standards sur les distributions attribuent une importance équivalente aux erreurs effectuées dans les différentes classes de diamètre. Par exemple l'erreur faite sur les arbres de 10 cm a autant de poids que l'erreur faite sur les arbres de 40 cm bien que ces derniers soient plus importants d'un point de vue forestier. Ainsi certains tests ont donc été développés pour résoudre ce genre de problème (voir par exemple Zucchini *et al.*, 2001).

### *Discussion sur les indices et tests statistiques*

Selon Huang *et al.* (2002) l'utilité des tests statistiques pour la validation des modèles est parfois limitée. Pour illustrer cela, ils évaluent les performances d'un modèle en comparant 9 tests sur 6 jeux de données indépendants. Ils montrent qu'en fonction des tests et en fonction des jeux de données les résultats sont différents et peuvent être contradictoires. Ils en concluent qu'aucun test n'est suffisamment générique pour convenir dans toutes les situations. Il nous paraît donc important de choisir un test ou un critère en fonction de l'utilisation future du modèle.

## 5.1.2 Sensibilité, robustesse et propagation d'erreurs

### 5.1.2.1 Evaluer la sensibilité d'un modèle

D'après Saltelli (2006) le processus de construction d'un modèle est soumis à quatre sources d'incertitude<sup>17</sup> (voir figure 5.1) : (1) incertitude sur le niveau de résolution choisi pour modéliser le système, (2) incertitude lors des mesures sur le terrain, (3) incertitude lors de l'estimation des paramètres et (4) incertitude sur le choix des relations décrivant la dynamique de croissance. Un modèle peut se comporter plus ou moins bien par rapport à ces différentes sources d'incertitudes. Des méthodes existent pour évaluer ce comportement. Les plus utilisées sont les analyses de sensibilité et les analyses d'incertitude.

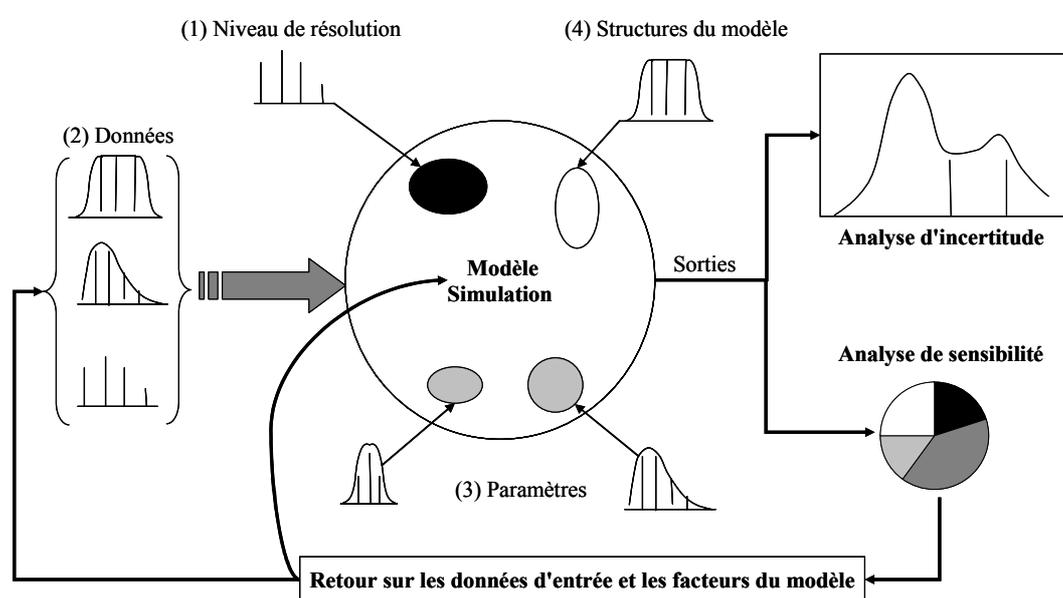


Figure 5.1 : Les différentes sources d'incertitude dans le processus de construction d'un modèle (1 à 4). Représentation schématique de l'analyse d'incertitude et de l'analyse de sensibilité. Figure extraite de Saltelli (2006).

### *Définition de la sensibilité et intérêt des analyses de sensibilité*

Les résultats d'un modèle peuvent être influencés par des variations sur les données d'entrée ou sur les paramètres du modèle. Dans les analyses de sensibilité, les données d'entrée ainsi que les paramètres du modèle sont appelés des facteurs. Si on fait des modifications faibles sur les facteurs et que cela entraîne de grosses différences au niveau des prédictions alors on dit que le modèle est sensible.

Les analyses de sensibilité étudient comment des perturbations sur les facteurs du modèle engendrent des perturbations sur la réponse (Jacques, 2005). Elles permettent de vérifier ou d'explorer comment les sorties du modèle varient en fonction des facteurs. Elles permettent

<sup>17</sup> Ici nous utilisons le mot incertitude avec la définition suivante : ce qui ne peut être établi avec exactitude, qui laisse place au doute.

également d'identifier les facteurs importants sur lesquels une attention particulière devra être portée au moment de leur estimation ou lors d'une phase d'amélioration du modèle. A l'opposé, elles permettent d'identifier les facteurs peu influents sur les sorties et permettent donc d'envisager des simplifications du modèle. Enfin, elles permettent d'identifier d'éventuelles interactions entre les facteurs ce qui peut permettre de mieux appréhender et comprendre le phénomène modélisé (Jacques, 2005).

### *Analyse de sensibilité et analyse d'incertitude*

Parmi les analyses de sensibilité, on peut distinguer les analyses de sensibilité au sens strict (par la suite nous les nommerons simplement analyse de sensibilité) et les analyses d'incertitude. Ces deux types d'analyse sont très proches et sont souvent associés (Saltelli *et al.*, 2006). Les analyses d'incertitude quantifient la variabilité de la sortie du modèle engendrée par des incertitudes sur les paramètres ou les données d'entrée (Cariboni *et al.*, 2007). Autrement dit, elles cherchent à répondre à la question (Martinez *et al.*, 2006) : Quelle est l'incertitude sur la sortie du modèle connaissant les incertitudes sur les entrées ? (voir la représentation schématique de la figure 5.1). Les analyses de sensibilité quant à elles cherchent à identifier les facteurs qui ont le plus d'influence sur la variabilité de la sortie et à quantifier cette influence. Autrement dit, elles cherchent à répondre à la question (Martinez *et al.*, 2006) : Quelle est la contribution d'une source d'incertitude en entrée sur l'incertitude de la sortie ? (voir la représentation schématique de la figure 5.1).

### *Les différents types d'analyse de sensibilité*

Il existe une littérature très abondante sur les méthodes d'analyse de sensibilité. Cariboni *et al.* (2007) donnent un aperçu rapide des différentes méthodes utilisables. Ces méthodes peuvent être classées en trois grandes catégories (Campolongo et Braddock, 1999 ; Jacques, 2005) :

- les analyses de sensibilité locales qui étudient la réponse de la variable de sortie à de petites variations d'un facteur autour d'une valeur donnée. Les autres facteurs sont fixés à leurs valeurs nominales. L'étude se fait facteur par facteur ;
- les analyses de sensibilité globales qualitatives, dites "screening", dont l'objectif est d'isoler, parmi un grand nombre de facteurs, ceux qui sont les plus influents sur la variable de sortie. Dans l'analyse globale on étudie l'effet d'un facteur en faisant varier également l'ensemble des autres facteurs ;

- les analyses de sensibilité globales quantitatives, qui permettent non seulement de hiérarchiser l'influence des facteurs, mais également d'attribuer à chaque facteur une part de la variabilité observée sur la variable de sortie. Ce dernier type d'analyse met en œuvre des techniques statistiques proches de l'analyse de variance (Saltelli *et al.*, 2005).

Si on se trouve face à un modèle soumis à 2 facteurs l'analyse sera plus simple que face à un modèle soumis à 100 facteurs. Les techniques d'analyse mises en œuvre ne seront probablement pas les mêmes. De même si le modèle est un ensemble de sous-modèles ou s'il s'agit d'une régression linéaire simple. Certains auteurs proposent des critères permettant de choisir parmi les différentes méthodes d'analyse de sensibilité (Saltelli *et al.*, 2005 ; Iooss, 2006 ; Cariboni *et al.*, 2007). Ces critères dépendent :

- du temps nécessaire pour faire une simulation ;
- du temps d'analyse nécessaire pour mettre en œuvre l'étude de sensibilité.
- du nombre de facteurs ;
- du degré de complexité du modèle (certaines méthodes font des hypothèses sur le modèle) ;

#### 5.1.2.2 Evaluer la robustesse d'un modèle et son comportement en extrapolation

##### *Notion de robustesse et analyse de robustesse*

En statistiques, la robustesse d'un estimateur est sa capacité à ne pas être modifiée par une petite modification dans les données ou dans les paramètres du modèle choisi pour l'estimation (voir Ripley, 2004 pour la notion de statistiques robustes ; voir également Huber (2004) cité par Schmidt *et al.*, 2006). Pour illustrer cette notion, on utilise souvent l'exemple de l'estimation de l'espérance d'une variable aléatoire. Pour estimer cette espérance, on peut utiliser la moyenne empirique (ou moyenne arithmétique). Ce modèle ne donnera pas une estimation robuste de l'espérance car si on lui donne une valeur erronée très grande parmi plusieurs autres valeurs, le résultat sera très influencé. En revanche, si on utilise la médiane, les résultats seront beaucoup moins influencés par les valeurs aberrantes. La médiane est donc plus robuste que la moyenne empirique. Cependant, selon Grimm (2004 pages 319 à 322) la robustesse est une notion plus large qui relève plus de la stratégie de modélisation que d'une technique spécifique.

### *Comportement du modèle en dehors du domaine de calibration*

Les modèles auxquels nous nous intéressons sont utilisés pour faire des prédictions. Dans la pratique, le modèle est parfois utilisé en dehors de son domaine de calibration. C'est pourquoi, la capacité d'un modèle à faire de bonnes prédictions en dehors de son domaine de calibration, est une qualité recherchée. Cette propriété fait partie de la robustesse d'un modèle. C'est généralement la forme du modèle choisie sur la base de connaissances écologiques qui permet de garantir un comportement acceptable en conditions d'extrapolation (Dhôte *et al.*, 2007).

#### 5.1.2.3 Etude du cumul des erreurs

Lorsque l'on fait fonctionner un modèle sur le pas de temps pour lequel il a été construit, la prédiction du modèle est entachée d'une erreur. Lorsque l'on fait fonctionner le modèle plusieurs fois de suite, par exemple si on utilise un modèle qui prédit la croissance par pas de 5 ans sur une période de 30 ans, au terme de la simulation l'erreur peut soit s'amplifier, soit s'éteindre, soit se stabiliser. L'étude de l'évolution de l'erreur au cours du temps de simulation est ce qu'on appelle l'étude de la propagation de l'erreur ou étude du cumul de l'erreur. Ce cumul d'erreurs peut aussi être étudié à travers une chaîne de modèle (Daquitaine, 2003).

L'erreur du modèle sur le premier pas de temps peut être soit l'erreur inhérente au modèle soit une erreur provoquée. Comme dans le cadre de l'analyse de sensibilité, on peut volontairement introduire une source de variation en début de simulation et observer l'impact de cette source de variation sur plusieurs étapes de simulation (Goreaud *et al.*, 2006a). On peut ensuite comparer les résultats entachés d'erreurs avec les résultats du modèle sans erreur introduite.

#### **5.1.3 Critères informatiques : temps d'exécution et taille mémoire utilisée**

Les modèles que l'on utilise sont implémentés dans un langage informatique permettant d'obtenir, à partir de données que l'on fournit, des résultats du modèle. Outre les performances statistiques des modèles, il peut être important d'évaluer la performance des modèles en termes de temps d'exécution et d'espace mémoire utilisé. Un modèle très performant, mais demandant plusieurs jours pour réaliser une simulation sera certainement peu utilisable dans la pratique. Il en sera de même pour un modèle demandant énormément de mémoire informatique, par exemple un modèle qui représente chaque feuille d'un arbre. Ces deux notions, temps d'exécution et taille mémoire nécessaires pour exécuter une simulation

sont appelés respectivement complexité en temps et en espace. Nous allons brièvement développer ces deux aspects.

### 5.1.3.1 Temps d'exécution et complexité d'un algorithme

Le temps d'exécution d'un programme va dépendre de plusieurs choses :

- du nombre de données ;
- de la taille du code ;
- du type d'ordinateur utilisé ;
- de la complexité de l'algorithme.

Le temps de calcul dépend notamment de la complexité de l'algorithme<sup>18</sup>. La complexité d'un algorithme est le nombre d'opérations élémentaires qu'il effectue, en fonction de la taille de son entrée (Casteys, 2000). Il existe différentes règles que nous n'allons pas détailler pour estimer la complexité d'un algorithme. Si on considère que toutes les opérations élémentaires prennent le même temps, par exemple  $10^{-10}$  secondes, qui dépend de la machine. Si de plus, le nombre d'opérations élémentaires est proportionnel à la taille des données d'entrée, alors le temps d'exécution sera proportionnel à la taille des données d'entrée quelle que soit la machine utilisée. Deux algorithmes qui effectuent la même opération peuvent avoir des temps d'exécution différents. Par exemple, le temps d'exécution peut être proportionnel au nombre d'arbres pour l'un alors que pour l'autre il peut être proportionnel au nombre d'arbres au carré. Ainsi, si on double le nombre d'arbres simulés, dans le premier cas le temps sera simplement doublé mais dans le second il sera quadruplé. Bien sûr, les capacités de calcul augmentent d'année en année mais la demande sur les performances des ordinateurs augmente également. Il est donc préférable d'avoir un algorithme le plus simple possible. Le temps que l'on perd lors des simulations avec un modèle plus complexe est alors à mettre en relation avec ce que l'on gagne.

### 5.1.3.2 Taille mémoire utilisée et complexité en espace de l'algorithme

L'analyse de la complexité en place mémoire revient à évaluer, en fonction de la taille de la donnée, la place mémoire nécessaire pour l'exécution de l'algorithme, en dehors de l'espace occupé par les instructions du programme et des données. C'est généralement la complexité en temps des algorithmes que l'on cherche à minimiser. Cependant, en fonction de la structure des données utilisées et du type de mémoire utilisé, un algorithme ayant une complexité faible peut avoir des temps d'exécution très importants. En effet l'accès à des informations sur le

---

<sup>18</sup> Un algorithme est une séquence finie d'opérations élémentaires exécutables directement par une machine.

disque dur de l'ordinateur demande beaucoup plus de temps que l'accès à des informations dans la mémoire vive.

D'autres critères informatiques peuvent être définis. Des critères liés à l'utilisation du simulateur peuvent notamment être important à prendre en compte dans le cadre de l'évaluation d'un modèle (Pretzsch *et al.*, 2002b) : facilité d'utilisation, possibilité d'intégration dans les pratiques de gestion, documentation du simulateur, etc. Nous en reparlerons dans la discussion générale.

#### **5.1.4 Modélisation et évaluation participatives**

Depuis les années 1990, les acteurs de la société qui ne sont ni scientifiques ni techniciens professionnels s'impliquent de plus en plus dans la recherche et l'innovation (Bonneuil, 2005). Cette participation peut prendre différentes formes et elle peut notamment être intégrée dans les processus de construction et d'évaluation d'un modèle. C'est pourquoi nous allons définir deux concepts qui peuvent avoir un intérêt dans notre démarche: la modélisation participative et l'évaluation participative.

##### 5.1.4.1 Modélisation participative

Tous les types de dialogues possibles entre modélisateurs et acteurs peuvent être considérés comme des processus de modélisation participative plus ou moins formalisés (Antona *et al.*, 2005). La modélisation participative et l'usage de modèles au sein d'approches participatives ont une popularité croissante. Dans tous les cas, il s'agit de mettre en interaction un groupe d'acteurs avec un modèle, soit pour sa conception, soit pour son usage. Ce modèle peut être à construire, en cours de construction, ou stabilisé. La construction du modèle ou son évolution se produit alors dans l'interaction avec les acteurs participants (Barreteau et Castellás, 2006).

Il existe deux points de vue par rapport à l'utilité de la modélisation participative (Barreteau, 2007). Le premier point de vue, qui concerne la majorité des travaux s'intéressant à la modélisation participative, met la participation au service de la modélisation. La participation est supposée améliorer la qualité du modèle du point de vue des attentes d'un client qui aura ensuite à l'utiliser. Pour l'autre point de vue, qui est par exemple celui développé au sein du réseau ComMod (Antona *et al.*, 2005), l'objectif est la compréhension ou la facilitation d'un processus de décision collective, la modélisation en interaction avec les acteurs de ce processus a pour but de mieux atteindre cet objectif.

#### 5.1.4.2 Evaluation participative

Il s'agit d'un concept essentiellement utilisé pour l'évaluation des projets de développement. D'après Jackson et Kassam (2000), il s'agit d'un processus d'auto-évaluation et d'action coopérative qui met à contribution les parties intéressées à un projet de développement. Les intervenants participent activement à la détermination des questions d'évaluation pertinentes, à la conception du schéma d'évaluation, à l'analyse et à la collecte de données ainsi qu'aux mesures de suivi qui seront prises en fonction des résultats de l'évaluation. L'exercice permet également aux intervenants de se doter des moyens voulus pour entreprendre leurs propres recherches et évaluations dans d'autres domaines.

Les évaluations qui sont véritablement et entièrement participatives s'assurent que les bénéficiaires locaux participent à toutes les étapes du processus, de la planification et de la conception jusqu'à la collecte et à l'analyse des données.

L'évaluation participative a plusieurs avantages : elle renforce les compétences et les capacités des gens qui s'y engagent, elle favorise l'implication des acteurs en les valorisant, elle permet de recueillir des informations et de mieux comprendre le problème et enfin elle facilite l'appropriation du projet par les acteurs.

### 5.1.5 Bilan et approches retenues pour évaluer et comparer les modèles

#### 5.1.5.1 Bilan concernant les critères et les méthodes d'évaluation

Dans cette sous-partie nous avons vu quatre grandes catégories de critères permettant d'évaluer des modèles de croissance forestiers.

Nous avons vu que les représentations graphiques pouvaient être très intéressantes mais qu'elles restent subjectives. Les tests et les critères statistiques permettent d'éviter ce problème mais malheureusement ils sont nombreux et peuvent parfois apporter des résultats contradictoires. Il convient donc de choisir ces tests et ces critères statistiques en fonction de l'utilisation future des modèles. Dans le cadre d'une comparaison de modèles, les indices sont intéressants car ils ne vont pas forcément servir à rejeter un modèle mais à hiérarchiser les modèles. Il est préférable dans ce cas de conserver des indices que l'on peut interpréter comme par exemple le biais moyen, l'erreur absolue moyenne ou l'erreur quadratique moyenne. Il faut également souligner que ces tests et ces critères sont d'autant plus pertinents lorsque l'on dispose de données d'évaluation indépendantes des données de calibration.

L'analyse de sensibilité, l'étude de la robustesse et du cumul d'erreurs permettent une analyse poussée des modèles mais elles nécessitent un investissement plus important que le

simple calcul d'indices ou la réalisation de tests statistiques. Chacun des modèles que nous avons construits pourrait faire l'objet d'une ou plusieurs analyses de sensibilité, ce qui peut représenter un travail très lourd. Pour cette raison, nous n'avons pas mis en œuvre de véritable analyse de sensibilité dans cette thèse. Ce travail pourra être envisagé ultérieurement et nous en reparlerons dans les perspectives.

Si l'on souhaite intégrer les modèles dans des simulateurs, nous avons vu que les critères informatiques comme la complexité des algorithmes peuvent être importants à considérer. Néanmoins, dans notre cadre (nombre d'années et nombre d'arbres simulés), le temps d'exécution des simulations ne nous est pas apparu comme un critère majeur permettant de différencier les modèles.

Enfin, nous avons abordé la question de l'évaluation des modèles par les utilisateurs. Pour cela nous nous sommes inspirés des concepts de modélisation et d'évaluation participative. Ces deux concepts nous paraissent intéressants à examiner dans le cadre de notre démarche. Néanmoins, le degré de participation des acteurs dans la construction ou l'évaluation de modèles adaptés à la croissance des peuplements mélangés chêne-pin est potentiellement faible comparé à des projets sur la gestion et la consommation de ressource (voir par exemple Bousquet, 2001 ; Barreteau, 2007) ou à des projets de développement. L'évaluation des modèles en collaboration avec les utilisateurs est une étape importante mais, pour des raisons de temps, elle sera réalisée en dehors de ce travail de thèse.

Pour évaluer nos modèles, nous avons donc choisi des approches qui utilisent principalement des données indépendantes (ou partiellement indépendantes comme nous le verrons) ainsi que des critères et des tests statistiques.

#### 5.1.5.2 Points importants pour mettre en œuvre ces critères et méthodes dans une comparaison de modèles

Les critères et les méthodes que nous venons de présenter permettent d'évaluer un modèle. Nous souhaitons utiliser ces critères et ces méthodes pour comparer plusieurs modèles. Pour cela, nous avons analysé 10 articles portant sur la comparaison de modèles ayant des niveaux de détail différents (tableau 5.2). Il ne s'agit pas d'une recherche exhaustive, néanmoins ces articles nous ont permis de définir 5 points importants pour mettre en œuvre nos comparaisons : détermination de l'objectif de la comparaison, choix des données permettant de comparer les modèles, production des données à partir des différents modèles, choix des données d'évaluation permettant d'arbitrer entre les différents modèles et choix des méthodes permettant de comparer les modèles. Nous allons décrire ces cinq points dans les

paragraphes qui suivent puis nous présenterons les approches retenues pour comparer nos modèles.

#### *Détermination de l'objectif de la comparaison*

Parmi les articles analysés, nous avons dégagé trois grands types d'objectifs : comparer la qualité des prédictions des différents modèles au niveau individuel et au niveau de la distribution des individus en classes de taille ; comparer la qualité des prédictions des différents modèles au niveau peuplement ; comparer la qualité des prédictions des différents modèles sur le long terme. Il s'agit d'une liste limitée et bien d'autres objectifs pourraient être définis.

#### *Choix des données permettant de comparer les modèles*

Les différentes études utilisent des données individuelles, des données sur la distribution des individus en classes de taille et des données agrégées. Les données individuelles sont utilisées uniquement dans deux études. Les données agrégées (nombre de tiges, surface terrière totale, ...) sont préférées d'une part parce qu'elles peuvent être obtenues avec des modèles ayant des niveaux de détail différents mais également parce que ce sont les données les plus fréquemment utilisées par les forestiers.

#### *Production des données à partir des différents modèles*

Les données prédites peuvent s'obtenir directement à partir des équations de croissance et d'un jeu de données (Wimberly et Bare, 1996 ; Sanchez-Gonzalez *et al.*, 2006). Cependant, dans les études analysées, la production des données à partir des modèles nécessite le plus souvent la mise en œuvre de simulations (8 articles sur 10). Pour que la comparaison soit valide, il est important de veiller à ce que les données d'entrée des différents modèles soient cohérentes.

#### *Choix des données d'évaluation permettant d'arbitrer entre les différents modèles*

Quatre approches permettant d'arbitrer entre les modèles peuvent être distinguées : utilisation de données indépendantes ; utilisation de données simulées par un modèle de référence ; utilisation de données d'ajustement ; comparaison des résultats donnés par les différents modèles sans données d'arbitrage. Parmi les études que nous avons analysées, peu d'entre elles utilisent un jeu de données indépendant pour comparer les modèles (3 sur 10). Cela vient certainement du fait que la plupart de ces études cherchent à évaluer le

comportement de différents modèles sur le long terme et que ce type de données est plutôt rare pour les écosystèmes forestiers.

#### *Choix des méthodes permettant de comparer les modèles.*

Parmi les méthodes utilisées, on peut distinguer les comparaisons graphiques sur les variables retenues, les comparaisons numériques sur des critères statistiques ou directement sur les variables retenues, les comparaisons qualitatives et l'utilisation de tests statistiques. Seulement deux études utilisent des tests statistiques. Les comparaisons s'appuient essentiellement sur des représentations graphiques ou des indices numériques.

#### 5.1.5.3 Approches retenues pour évaluer les modèles

En nous appuyant sur l'analyse que nous venons de faire, nous avons évalué les modèles de croissance en utilisant trois approches : l'évaluation et la comparaison des modèles individus-centrés pour prédire les accroissements individuels, leur variabilité et leur distribution en utilisant un jeu de données indépendant (sous-partie 5.2), l'évaluation et la comparaison des modèles pour prédire l'accroissement en surface terrière du peuplement en utilisant un jeu de données partiellement indépendant (sous-partie 5.3) et l'évaluation et la comparaison des modèles pour prédire l'évolution du peuplement dans des contextes appliqués incluant des simulations à court et moyen termes (sous-partie 5.4).

Tableau 5.2 : Description des études analysées présentant des comparaisons de modèles forestiers. \* voir la sous-partie 5.1 ou les articles pour la définition des indices et des tests statistiques mentionnés. Modèles = nombre de modèles comparés.

Auteur(s)	Modèles	Objet de la comparaison	Indicateur	Placettes ou cas	Projection	Données	Critères de comparaison*
(Astrup <i>et al.</i> , 2008)	5	Qualité des prédictions au niveau peuplement	Surface terrière totale	49	30 ans	Indépendantes	Biais moyen, erreur moyenne, comparaisons graphiques
(Ek et Monserud, 1979)	2	Qualité des prédictions au niveau peuplement et sur les distributions en taille	Surface terrière totale ; Nombre de tiges ; Distribution des diamètres	16	5 à 26 ans	Indépendantes	(1) test t par dispositif ; (2) "Freese's chisquare" sur l'ensemble des 16 placettes ; (3) test de Kolmogorov-smirnov pour les distributions des valeurs prédites et observées ; (4) Test de la pente égale à 1 et de l'ordonnée égale à 0 ; (5) Détection de tendances sur les résidus par régression ; (6) test de Kolmogorov-smirnov pour la distribution des diamètres
(Gourlet-Fleury <i>et al.</i> , 2005)	2	Comportement à long terme	Nb arbres > à 10 cm ; Nb arbres < à 60 cm	1	10*42 ans	Simulées par un modèle de référence	Comparaisons graphiques
(Kokkila <i>et al.</i> , 2006)	3	Comportement à long terme	Hauteur moyenne ; Diamètre moyen ; Densité	4	70 ans	Comparaisons entre modèles	Comparaisons graphiques
(Lett <i>et al.</i> , 1999)	2	Comportement à long terme	Surface terrière par espèce ; Structure spatiale ; Distribution des diamètres	1	1000 ans	Comparaisons entre modèles	Comparaisons graphiques, temps de calcul.
(Matala <i>et al.</i> , 2003)	2	Comportement à long terme	Densité ; Surface terrière ; Volume tige ; Production totale ; Hauteur dominante ; Distribution des diamètres ; Distribution des hauteurs	1	100 ans	Comparaisons entre modèles	Comparaisons graphiques
(Picard et Franc, 2001)	3	Qualité des prédictions sur les distributions en taille, comportement à long terme	Distribution des hauteurs ; Distribution des diamètres	1	>1000 ans	Indépendantes	Comparaisons graphiques, indice de dissimilarité
(Sanchez-Gonzalez <i>et al.</i> , 2006)	2	Qualité des prédictions au niveau individuel	Accroissements individuels en diamètre	72	9 à 11 ans	Ajustement	Biais, RMSE, PRESS, comparaisons graphiques, graphiques des résidus
(Verzelen <i>et al.</i> , 2006)	9	Comportement à long terme	Nombre de tiges ; Surface terrière cumulée ; Distribution des diamètres	1	>1000 ans	Simulées par un modèle de référence	Comparaisons graphiques, indice de dissimilarité
(Wimberly et Bare, 1996)	2	Qualité des prédictions au niveau individuel	Accroissements individuels en surface terrière	30	3 ans	Ajustement	R <sup>2</sup> , Coefficients de détermination partiels

## **5.2 Prédire les accroissements individuels, leur variabilité et leur distribution**

Nous avons construit trois types de modèles de croissance individus-centrés : un modèle arbre indépendant des distances (MAID), un modèle arbre dépendant des distances (MADD) et un modèle arbre utilisant des distributions de voisinage (MADV). L'objectif de cette sous-partie est d'évaluer la capacité de ces modèles à prédire les accroissements individuels, leur variabilité et leur distribution. Ces évaluations ont été réalisées à partir de données de croissance individuelle non utilisées lors de l'ajustement des modèles.

Nous allons d'abord présenter le jeu de données d'évaluation en le comparant au jeu de données de calibration (paragraphe 5.2.1) puis nous discuterons des données pertinentes pour évaluer le modèle utilisant les distributions de voisinage (paragraphe 5.2.2). Ensuite, nous présenterons les indices et les tests statistiques que nous avons utilisés pour comparer les modèles (paragraphe 5.2.3). Enfin, nous présenterons les résultats des comparaisons (paragraphe 5.2.4) et nous les discuterons en expliquant les différences obtenues entre les modèles (paragraphe 5.2.5).

### **5.2.1 Présentation des données individuelles d'évaluation**

Dans la partie 3, nous avons vu que nous avons partagé le jeu de données sur la croissance individuelle en deux parties : un jeu de données pour l'ajustement des modèles et un jeu pour leur évaluation. Nous avons réalisé le partage en effectuant un tirage aléatoire sans remise par dispositif et nous avons choisi de conserver 2/3 des données pour l'ajustement et 1/3 pour l'évaluation (voir le tableau 3.5).

Pour évaluer comment se situent les données d'évaluation par rapport aux données d'ajustement, nous avons représenté leur distribution en fonction des accroissements (période 2000-2005) et des variables retenues dans le MADD et dans le modèle arbre dépendant des distances. La surface terrière locale des feuillus est donnée pour information car elle n'a pas été retenue dans le MADD. Les mêmes graphiques pour chaque dispositif sont présentés dans l'annexe F pour les chênes et dans l'annexe G pour les pins. De plus nous avons représenté les accroissements en fonction de la circonférence par essence et pour chaque dispositif en distinguant les arbres retenus pour l'ajustement et ceux retenus pour l'évaluation des modèles.

Nous pouvons constater que les données d'évaluation sont bien réparties par rapport aux données d'ajustement (figure 5.2 pour les chênes et figure 5.3 pour les pins). Seules quelques valeurs extrêmes ne sont pas représentées ou mal représentées.

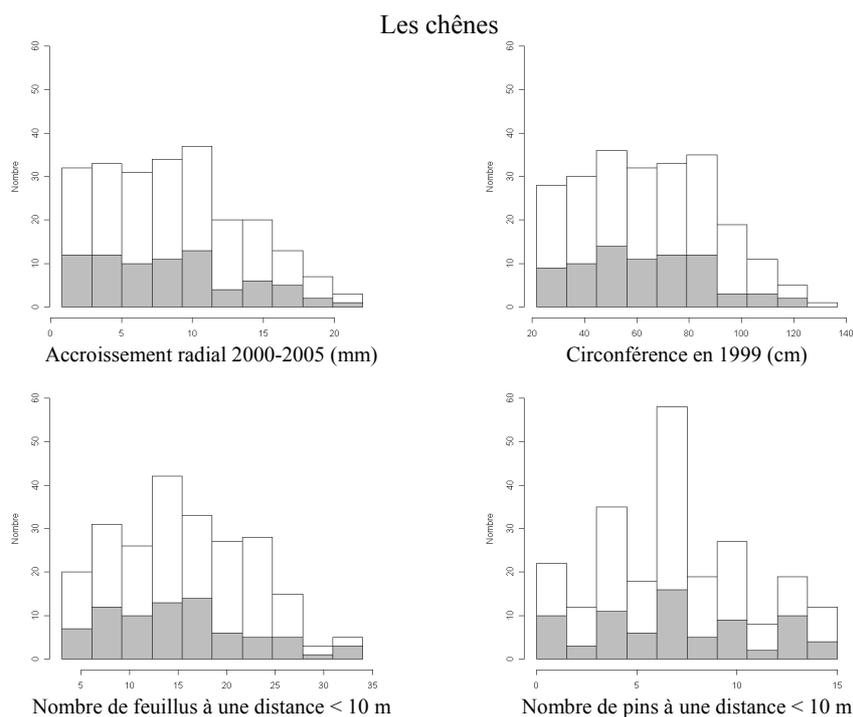


Figure 5.2 : répartition des données d'évaluation des chênes par rapport à l'ensemble des données disponibles en fonction des accroissements et des variables retenues dans les différents modèles. Les données sont réparties en 10 classes. Les données d'évaluation sont en gris.

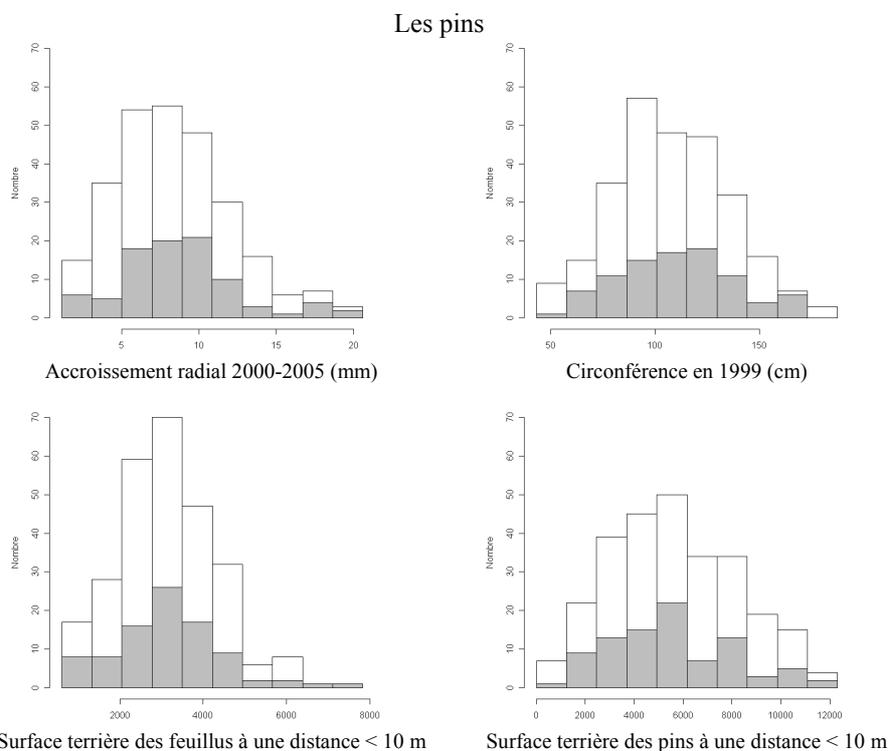


Figure 5.3 : répartition des données d'évaluation des pins par rapport à l'ensemble des données disponibles en fonction des accroissements et des variables retenues dans les différents modèles. Les données sont réparties en 10 classes. Les données d'évaluation sont en gris.

De même, la figure 5.4 montre que pour chaque dispositif les données d'évaluation sont plutôt bien réparties dans les nuages de points reliant l'accroissement à la circonférence. Les données d'évaluation et les données de calibration sont donc des données tout à fait comparables. Ces données vont notamment nous permettre de comparer la pertinence du MADD par rapport au MAID pour prédire les accroissements individuels. D'après les résultats de la partie 4, le MADD devrait être plus performant que le MAID grâce à l'utilisation d'indices de compétition locaux.

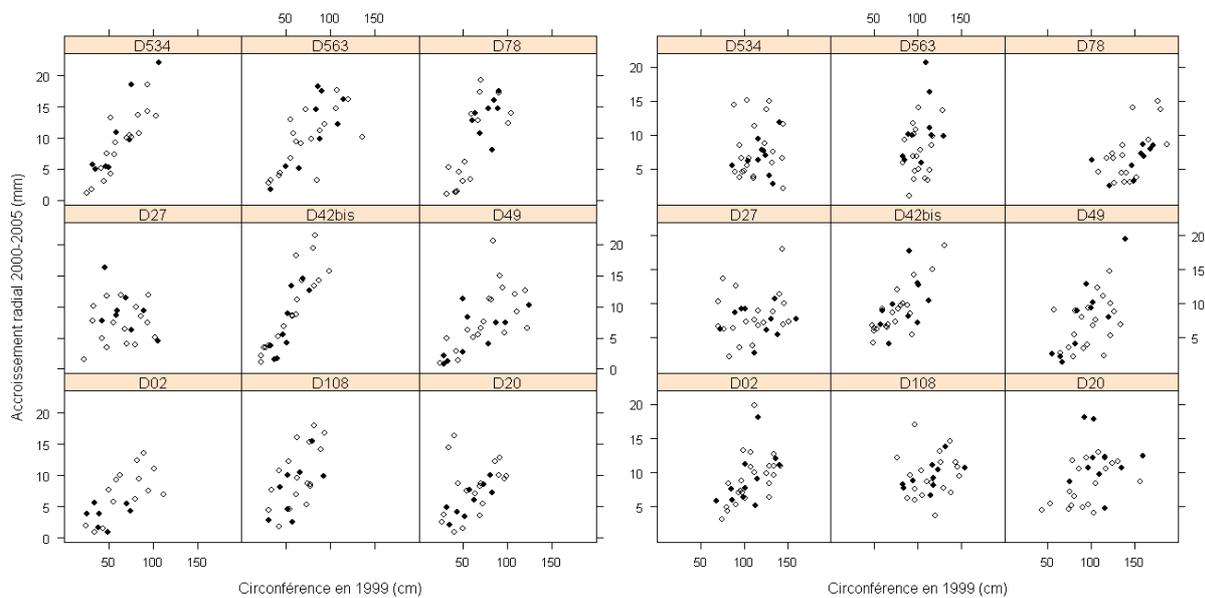


Figure 5.4 : Accroissements en fonction de la circonférence par essence (à gauche les chênes et à droite les pins) et par dispositif. Les points blancs sont les données d'ajustement et les points noirs sont les données d'évaluation.

## 5.2.2 Comment évaluer le modèle utilisant des distributions de voisinage ?

Nous disposons maintenant de données d'évaluation pour lesquelles la position des arbres est connue. Si l'on souhaite évaluer les modèles individus-centrés sur ces données nous avons trois façons d'approcher la croissance individuelle en rayon d'un arbre  $i$  (noté  $y_i$ ) :

- $\hat{y}_{i,MAID} = \alpha_{k,s} + \beta_{k,s} circ_{k,s,i}$  (équation 4.6) ;
- $\hat{y}_{i,MADD} = \alpha'_{k,s} + \beta'_{k,s} circ_{k,s,i} + \lambda_{s,chêne} CI_{i,s,chêne} + \lambda_{s,pin} CI_{i,s,pin}$  (équation 4.8) ;
- $\hat{y}_{i,MADV} = \alpha'_{k,s} + \beta'_{k,s} circ_{k,s,i} + \lambda_{s,chêne} CI^*_{i,s,chêne} + \lambda_{s,pin} CI^*_{i,s,pin}$  (équation 4.9).

où  $\hat{y}_{i,MAID}$ ,  $\hat{y}_{i,MADD}$  et  $\hat{y}_{i,MADV}$  sont les accroissements individuels estimés respectivement par le MAID, le MADD et le MADV. Nous rappelons que dans l'équation correspondant au MADV, l'étoile ajoutée aux indices de compétition signifie qu'ils sont calculés à partir d'un voisinage reconstruit et non à partir du voisinage réel. Dans un premier temps, nous allons estimer l'erreur commise par ces modèles au niveau individuel. Pour cela nous pouvons choisir comme critère de comparaison la quantité  $\sum (y_i - \hat{y}_{i,j})^2$  où  $j$  est le type de modèle. Le tableau 5.3 donne les résultats supposés d'un point de vue qualitatif.

Tableau 5.3 : Résultats supposés concernant l'erreur commise par les différents modèles au niveau individuel.

Type de modèle	Critère d'évaluation	Valeur attendue du critère
MAID	$\sum (y_i - \hat{y}_{i,MAID})^2$	Intermédiaire
MADD	$\sum (y_i - \hat{y}_{i,MADD})^2$	Minimum
MADV	$\sum (y_i - \hat{y}_{i,MADV})^2$	Maximum

Le modèle avec distributions de voisinage utilise bien les données individuelles sur la circonférence mais les données sur le voisinage sont tirées dans une distribution. A titre de comparaison, nous sommes dans le même cas de figure qu'un modèle arbre dépendant des distances utilisant un modèle de structure spatiale pour générer un peuplement virtuel (Ngo Bieng, 2007). Ainsi, l'accroissement individuel d'un arbre  $i$  sera estimé avec sa vraie circonférence mais avec un environnement différent de l'environnement réel. C'est pourquoi on peut s'attendre à ce que l'erreur commise par le modèle avec distribution de voisinage soit supérieure à celle commise par le MAID. L'erreur sera certainement minimum dans le cas du MADD car ce modèle, contrairement aux autres, prend en compte à la fois la circonférence individuelle des arbres et leur environnement réel pour prédire leur croissance.

Dans un deuxième temps nous allons estimer l'erreur commise sur la variabilité des accroissements. Pour cela nous allons considérer l'écart-type des  $y_i$ . A priori, les écart-types des différents modèles devraient être différents. Nous pouvons donc retenir le critère  $|\sigma(y_i) - \sigma(\hat{y}_{i,j})|$  pour comparer les modèles (tableau 5.4).

Tableau 5.4 : Résultats supposés concernant l'erreur commise par les différents modèles par rapport à la variabilité des accroissements individuels.

Type de modèle	Critère d'évaluation	Valeur attendue du critère
MAID	$ \sigma(y_i) - \sigma(\hat{y}_{i,MAID}) $	Maximum
MADD	$ \sigma(y_i) - \sigma(\hat{y}_{i,MADD}) $	Minimum
MADV	$ \sigma(y_i) - \sigma(\hat{y}_{i,MADV}) $	Minimum

Cette fois le critère retenu pour comparer les modèles évalue la capacité des modèles à rendre compte de la variabilité des  $y_i$ . On peut donc supposer que le MADD et le modèle avec distribution de voisinage se comportent environ de la même façon et que le MAID respecte moins bien la variabilité des accroissements observés.

Comme nous venons de le voir, le modèle avec distribution de voisinage n'est pas adapté pour respecter la valeur individuelle des accroissements. Son évaluation par rapport à la qualité des prédictions individuelles n'a donc pas beaucoup de sens. En revanche, il est sensé rendre compte de la variabilité des environnements locaux au sein du peuplement ainsi que des différences de structure spatiale entre peuplements. Il doit donc être évalué sur des données adaptées, comme par exemple la distribution des accroissements individuels, la variabilité des accroissements ou, comme nous le verrons plus loin, des variables à l'échelle du peuplement. Dans cette sous-partie nous allons donc insister sur l'évaluation de ce modèle par rapport à la distribution et la variabilité des accroissements individuels.

### 5.2.3 Indices et tests retenus pour évaluer et comparer les modèles

Nous avons utilisé des indices et des tests statistiques pour évaluer la capacité des modèles à prédire correctement les accroissements individuels. Outre la performance moyenne des modèles, nous avons également cherché à évaluer la capacité des modèles à prédire correctement la variabilité et la distribution des accroissements individuels. Tous les tests ont été réalisés avec le logiciel R (R Development Core Team, 2008)

#### 5.2.3.1 Indices et tests statistiques utilisés pour évaluer la performance des modèles par rapport à la prédiction de l'accroissement individuel

Comme nous l'avons précisé dans le paragraphe 5.2.2 ces indices et ces tests ne sont pertinents que pour le MADD et le MAID. Pour évaluer ces modèles par rapport à la prédiction de l'accroissement individuel nous avons retenu quatre indices couramment utilisés (tableau 5.5) : le biais moyen, l'erreur absolue moyenne, l'écart-type résiduel et le pseudo-R<sup>2</sup>. Par la suite nous utiliserons le mot "biais" pour désigner l'écart entre la valeur observée et la valeur prédite ( $y_i - \hat{y}_i$ ). Nous avons également utilisé des tests statistiques. Dans le cas des pins, un test de Shapiro-Wilk (R Development Core Team, 2008) met en évidence que les distributions des biais obtenus avec le MAID et le MADD ne suivent pas une loi normale (pour le MAID :  $W = 0.92$ ,  $p < 0.01$  ; pour le MADD :  $W = 0.94$ ,  $p < 0.01$ ). Nous avons donc utilisé pour le chêne et le pin, le test non paramétrique de Wilcoxon (Sprent, 1992 page 37) pour vérifier si la valeur du biais moyen était significativement différente de 0. Nous avons également testé si les biais moyens obtenus avec le MAID et le MADD étaient significativement différents (test de Wilcoxon). Enfin, nous avons testé si la valeur du biais dépendait du dispositif avec un test de Kruskal-Wallis (Sprent, 1992 page 130).

Tableau 5.5 : Critères statistiques retenus pour évaluer le MAID et le MADD.  $y_i$  = valeur observée sur l'individu  $i$ .  $\hat{y}_i$  = valeur estimée par le modèle sur l'individu  $i$ .  $\bar{y}$  = moyenne des valeurs observées.  $\bar{\hat{y}}$  = moyenne des valeurs prédites

Critère	Définition
Biais moyen	$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$
Erreur absolue moyenne (MAE)	$\frac{\sum_{i=1}^n  y_i - \hat{y}_i }{n}$
Ecart-type résiduel (RMSE)	$\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$
Pseudo-R <sup>2</sup> (EF)	$1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$

### 5.2.3.2 Relation entre les biais et les valeurs prédites et entre les biais et les variables explicatives

Pour mettre en évidence d'éventuelles anomalies dans les prédictions individuelles, nous avons également représenté les biais en fonction des valeurs prédites et en fonction des variables explicatives retenues dans le MAID et le MADD. A l'aide d'une régression, nous avons testé s'il y avait une relation significative entre les biais et les valeurs prédites et entre les biais et les variables explicatives.

### 5.2.3.3 Critères statistiques utilisés pour évaluer les modèles par rapport la prédiction de la distribution des accroissements individuels et par rapport à leur variabilité.

Pour évaluer les modèles par rapport à la prédiction de la distribution des accroissements individuels, nous avons utilisé l'indice "Absolute Discrepancy" (Aguirre *et al.*, 2003 ; Pommerening, 2006) qui permet d'évaluer l'écart entre deux distributions (tableau 5.6).  $AD$  est la proportion d'individus qu'il faut échanger entre les classes pour transformer une distribution en l'autre.  $I-AD$  est la proportion d'individus en commun pour les deux distributions. Cet indice nous a également permis de hiérarchiser les modèles. Pour vérifier si la distribution des accroissements observés et la distribution des accroissements prédits étaient identiques, nous avons utilisé un test de Kolmogorov-Smirnov (Sprenst, 1992 page 120).

Pour évaluer les modèles par rapport à la prédiction de la variabilité des accroissements, nous avons utilisé deux indices : le rapport des variances et la différence absolue d'écart-type (tableau 5.6). Enfin, pour vérifier l'égalité des variances entre les accroissements observés et les accroissements prédits nous avons utilisé un test de Fligner-Killeen (R Development Core Team, 2008).

Tableau 5.6 : Critères statistiques retenus pour évaluer les modèles par rapport la prédiction de la distribution des accroissements individuel et par rapport à leur variabilité.  $y_i$  = valeur observée sur l'individu  $i$ .  $\hat{y}_i$  = valeur estimée par le modèle sur l'individu  $i$ .  $\bar{y}$  = moyenne des valeurs observées.  $\bar{\hat{y}}$  = moyenne des valeurs prédites.  $f_i$  est la fréquence des arbres appartenant à la classe  $i$  de la distribution observée des accroissements individuels.  $f_i'$  est la fréquence des arbres appartenant à la classe  $i$  de la distribution des accroissements individuels issue d'un modèle

Indice	Définition
Absolute Discrepancy	$AD = \frac{1}{2} \sum_{i=1}^n  f_i - f_i' $
Rapport des variances	$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
Différence absolue d'écart type	$\left  \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} - \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2} \right $

#### 5.2.3.4 Production des accroissements individuels à partir des modèles

Pour le MAID et le MADD, les prédictions ont été obtenues directement à partir des équations de croissance (équations 4.6 et 4.8) et des données en début de période de croissance : circonférences 1999 et indices de compétition calculés à partir des voisinages 1999.

Pour le MADV nous avons estimé les distributions de voisinage pour chaque dispositif en considérant tous les arbres du dispositif. Nous avons ensuite utilisé le module *Oakpine2* dans *Capsis* (Coligny *et al.*, 2003) pour reconstruire les voisinages et simuler la croissance sur 6 années (2000-2005) pour tous les arbres de l'échantillon d'évaluation. Lors de la reconstruction des voisinages, il y a une part aléatoire qui intervient dans le choix du nombre de voisins ainsi que dans le choix des arbres du voisinage. Pour chaque dispositif nous avons donc réalisé 1000 répétitions en utilisant le mode *script* de *Capsis*. Pour chaque répétition nous avons calculé les indices présentés dans le paragraphe précédent. Pour comparer le MADV avec les autres modèles, nous avons utilisé la moyenne des indices sur les 1000 répétitions.

## 5.2.4 Résultats

### 5.2.4.1 Résultats concernant la prédiction des accroissements individuels

Les indices obtenus avec le MAID et le MADD sont proches mais nous constatons que l'erreur absolue moyenne, l'écart-type résiduel et le pseudo-R<sup>2</sup> sont moins bons dans le cas du MADD que dans le cas du MAID (tableau 5.7). Pour le chêne, on note que le biais moyen obtenu avec le MAID est plus proche de zéro que celui obtenu avec le MADD. Cependant les deux valeurs ne sont pas significativement différentes de 0 et ne sont pas significativement différentes l'une de l'autre (tableau 5.8). Pour le pin, les biais moyens obtenus avec les deux modèles sont très proches et non significativement différents de 0. Pour le MAID, il y a un effet significatif du dispositif sur la valeur du biais moyen pour les deux essences (tableau 5.8). Cela signifie que le MADD, contrairement au MAID, donne des résultats comparables quels que soient les dispositifs (figure 5.5 et figure 5.6).

Tableau 5.7 : Critères statistiques obtenus pour le chêne et le pin avec les trois modèles sur les 9 dispositifs. Dans le cas du MADV les valeurs indiquées sont calculées à partir de 1000 répétitions.

	Chêne			Pin		
	MAID	MADD	MADV	MAID	MADD	MADV
Nombre d'observations	76	76		90	90	
Biais moyen (mm)	-0.092	-0.455		0.381	0.342	
Erreur absolue moyenne (mm)	2.878	3.099		2.506	2.567	
Ecart-type résiduel (mm)	3.523	3.780		3.526	3.539	
Pseudo-R <sup>2</sup>	0.509	0.435		0.163	0.157	
Rapport des variances	0.393	0.548	0.449	0.181	0.206	0.227
Différence absolue d'écart type	1.874	1.304	1.663	2.212	2.105	2.019
AD	0.276	0.197	0.291	0.266	0.244	0.241

Tableau 5.8 : Hypothèses testées et résultats des tests pour le chêne et pour le pin. Si le test est non significatif au seuil de 5% (NS), l'hypothèse nulle n'est pas rejetée.

Hypothèse nulle du test	Chêne			Pin		
	MAID	MADD	MADV	MAID	MADD	MADV
Le biais moyen est nul	NS	NS		NS	NS	
Il n'y a pas d'effet dispositif sur le biais moyen du MAID	p = 0.037	NS		p = 0.045	NS	
Les données observées et les données prédites proviennent de la même distribution	NS	NS	NS	NS	NS	NS
Les données observées et les données prédites ont les mêmes variances	p < 0.001	p = 0.006	p < 0.001	p < 0.001	p < 0.001	p < 0.001
	Chêne			Pin		
Les biais moyen du MAID et du MADD sont identiques	NS			NS		

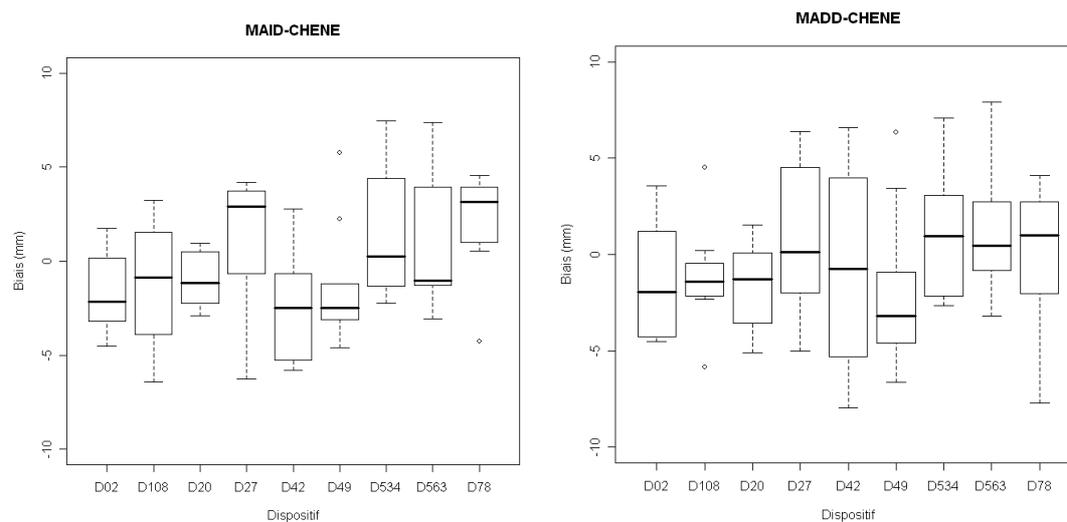


Figure 5.5 : boîtes à moustaches du biais pour chaque dispositif pour le chêne. A gauche biais obtenus avec le MAID et à droite biais obtenus avec le MADD.

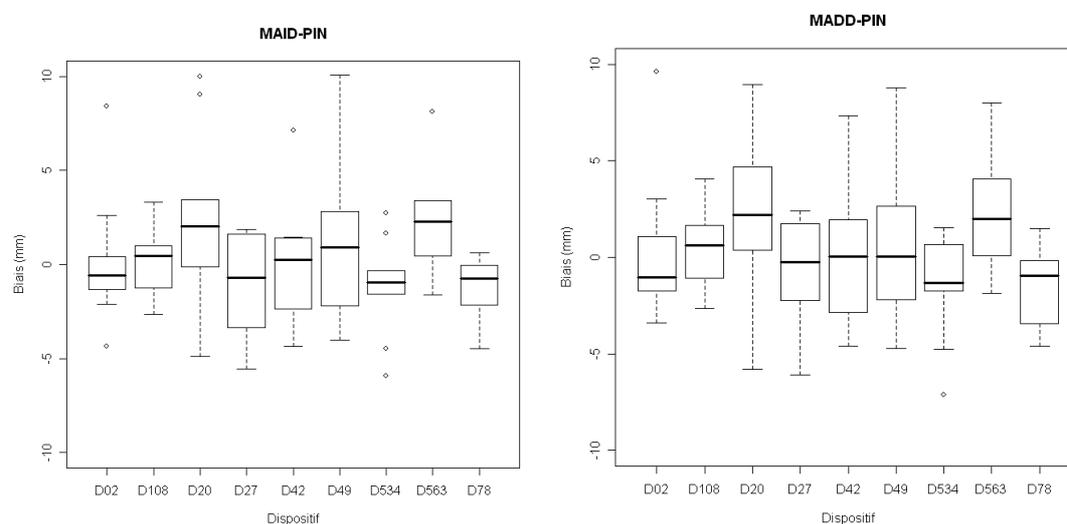


Figure 5.6 : boîtes à moustaches du biais pour chaque dispositif pour le pin. A gauche biais obtenus avec le MAID et à droite biais obtenus avec le MADD.

#### 5.2.4.2 Résultats concernant la prédiction de la distribution et de la variabilité des accroissements individuels

Le rapport des variances, la différence absolue des écart-types et l'indice  $AD$  sont meilleurs pour le MADD que pour le MAID (tableau 5.7). Concernant le MADV, le rapport des variances et la différence absolue des écart-types sont soit proches du MADD, soit intermédiaires entre le MAID et le MADD. La valeur de l'indice  $AD$  obtenue avec le MADV pour le chêne est plus proche de la valeur obtenue avec le MAID que de celle obtenue avec le MADD alors que c'est l'inverse pour le pin. Pour les trois modèles, la variance des données observées est significativement plus grande que la variance des données prédites. La figure

5.7 montre que, pour le chêne comme pour le pin, les différents modèles prédisent mal les accroissements extrêmes. Néanmoins, la distribution des valeurs observées et la distribution des valeurs prédites ne sont pas significativement différentes (tableau 5.8).

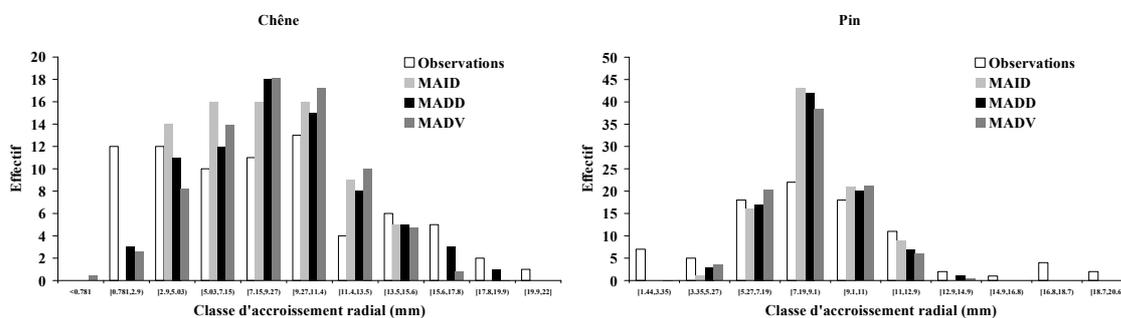


Figure 5.7 : Distribution observée des accroissements radiaux individuels sur l'ensemble des dispositifs et distributions obtenues avec les différents modèles. Dans le cas du modèle avec distributions de voisinage la distribution présentée est la moyenne de 1000 simulations.

### 5.2.4.3 Résultats sur le lien entre les biais et les valeurs prédites et entre les biais et les variables explicatives

L'ensemble des graphes entre les biais et les différentes variables est présenté dans l'annexe H. Pour le chêne, il y a un lien significatif entre les biais obtenus avec le MAID et le nombre de pins dans un rayon de 10 m (figure 5.8). Nous avons également détecté un lien significatif entre les biais obtenus avec le MADD et le nombre de feuillus dans un rayon de 10 m (figure 5.9). Pour le pin, nous n'avons pas détecté de tendance significative entre les biais et les différentes variables, que ce soit pour le MAID ou pour le MADD.

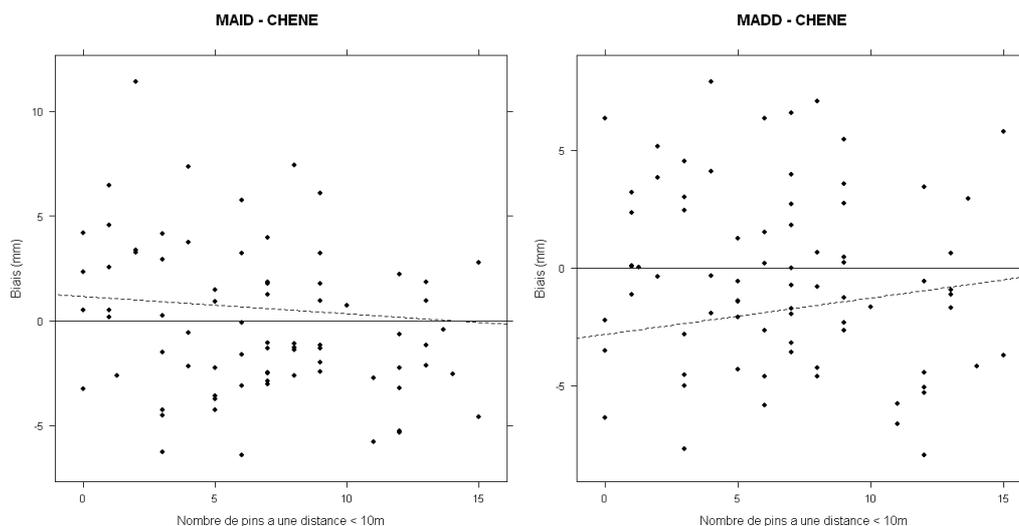


Figure 5.8 : A gauche, biais obtenus avec le MAID en fonction du nombre de pins dans un rayon de 10 m. A droite, biais obtenus avec le MADD en fonction du nombre de pins dans un rayon de 10 m.

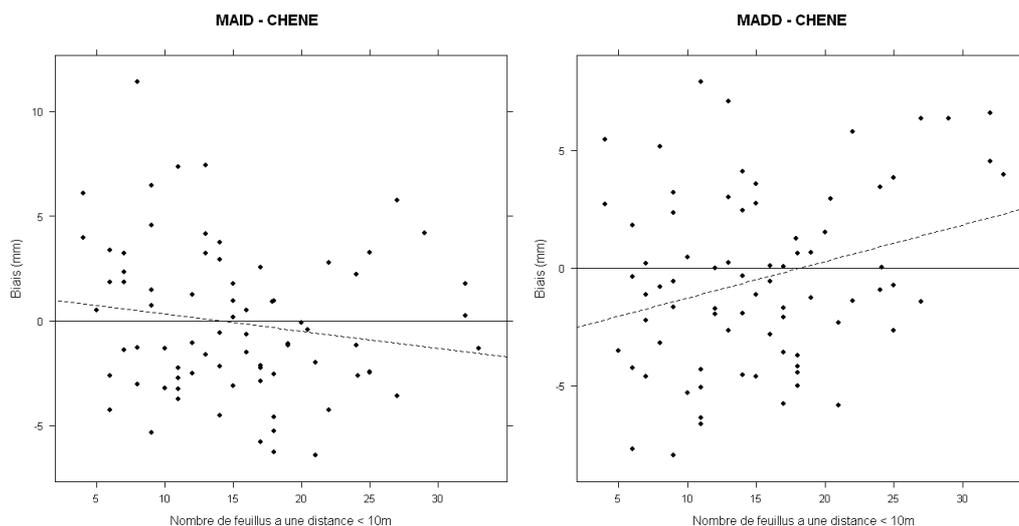


Figure 5.9 : A gauche, biais obtenus avec le MAID en fonction du nombre de feuillus dans un rayon de 10 m. A droite, biais obtenus avec le MADD en fonction du nombre de feuillus dans un rayon de 10 m.

## 5.2.5 Discussion

Globalement les statistiques que nous avons obtenues sur le jeu de données indépendant sont cohérentes avec celles obtenues sur les données de calibration (tableau 4.1 pour le MAID et tableau 4.5 pour le MADD). Ce résultat est logique puisque nous avons utilisé des arbres provenant des dispositifs utilisés pour l'ajustement des modèles. Cependant, ces données nous ont permis d'évaluer la pertinence du MADD par rapport au MAID pour prédire les accroissements individuels. De plus, elles nous ont permis d'évaluer le MADV par rapport à la variabilité des accroissements et par rapport à leur distribution.

### 5.2.5.1 Comparaison entre le MAID et le MADD

Contrairement à ce que l'on pouvait attendre, le MADD ne donne pas de meilleurs résultats que le MAID pour prédire les accroissements individuels. Les indices liés à la distribution ou à la variabilité des accroissements obtenus avec le MADD sont meilleurs mais les indices liés à la performance globale sont moins bons (pseudo  $R^2$ , écart-type résiduel). Ce résultat signifie que les paramètres liés aux indices de compétition estimés sur le jeu de données de calibration, ne sont pas transposables au jeu de données d'évaluation. Ces résultats montrent aussi l'importance d'utiliser un jeu de données indépendant pour évaluer la qualité prédictive des modèles. Il faut noter que ni le MAID ni le MADD ne parviennent à reproduire correctement la variabilité des accroissements observés. En particulier, les valeurs extrêmes ne sont pas bien prédites. Pour le chêne nous avons détecté un lien entre le biais obtenu avec le MAID et le nombre de pins dans un rayon de 10 m. Ce lien est négatif ce qui signifie que les

arbres qui ont peu de pins dans leur voisinage ont une croissance sous-estimée. Ce résultat est logique puisque ce modèle ne prend pas en compte le voisinage des arbres. Pour le nombre de feuillus on retrouve la même tendance (figure 5.9) mais le résultat n'est pas significatif. Toujours pour le chêne nous avons détecté un lien entre le biais obtenu avec le MADD et le nombre de feuillus dans un rayon de 10 m. Ce lien est positif ce qui signifie que les arbres qui ont beaucoup de feuillus dans leur voisinage ont une croissance sous-estimée. Cela voudrait dire que le paramètre de cet indice de compétition est un peu trop fort. Cela peut aussi être le résultat d'une relation non linéaire non prise en compte dans le modèle.

Si on considère le fait que les arbres appartiennent à des dispositifs différents, nous avons montré que le MADD avait un meilleur comportement que le MAID car il donne des résultats comparables quels que soient les dispositifs.

#### 5.2.5.2 Performance du MADV

Comme cela était attendu, le MADV reproduit mieux la variabilité de la croissance que le MAID. Cependant, pour le chêne, la distribution des accroissements individuels est plus éloignée de la distribution observée que la distribution des accroissements obtenue avec le MADD. La seule différence entre ces deux modèles vient du fait que, dans le cas du MADV, les indices de compétition sont calculés à partir de voisinages reconstruits et non à partir des voisinages réels. Les différents résultats obtenus sur la distribution des accroissements peut donc provenir :

- des hypothèses utilisées pour reconstruire les voisinages (voir sous-partie 4.3) : le nombre de voisins chêne et le nombre de voisins pin ne sont pas conditionnés par le nombre total de voisins et un voisin est choisi de façon aléatoire dans le peuplement indépendamment de la taille de l'arbre sujet ;
- de l'estimation des distributions de voisinage. En effet, elles sont estimées sur l'ensemble d'un dispositif pour tous les arbres présents sur cette surface. Ainsi les voisinages reconstruits reflètent les voisinages de l'ensemble du peuplement. Certains voisinages risquent donc d'être surreprésentés par rapport aux voisinages de l'échantillon d'évaluation.

Nous avons montré qu'il était possible d'évaluer le MADV à partir d'un jeu de données sur des accroissements individuels. Cependant, cette évaluation est délicate à cause de la part aléatoire introduite dans ce modèle et de la nécessité d'estimer correctement les distributions de voisinage.

### 5.3 *Prédire l'accroissement du peuplement*

Dans la sous-partie précédente, nous avons évalué la capacité des modèles individus-centrés à prédire les accroissements individuels, leur distribution ou leur variabilité. L'objectif de cette sous-partie est d'évaluer les modèles pour prédire des données à l'échelle du peuplement. Dans cette sous-partie nous n'avons pas inclus le modèle  $MP_{MAID}$  car nous avons montré dans la partie 4 qu'il donnait exactement les mêmes résultats que le MAID au niveau du peuplement. Nous avons tout de même vérifié que les résultats numériques donnés par ces deux modèles étaient identiques ce qui nous a permis de confirmer que notre agrégation était correcte. Nous allons d'abord préciser comment nous avons obtenu des données d'évaluation au niveau peuplement à partir d'un échantillon de données individuelles de croissance (paragraphe 5.3.1 et 5.3.2) et comment nous avons obtenu les données prédites par les différents modèles (paragraphe 5.3.3). Nous verrons ensuite les résultats de l'évaluation des modèles à l'échelle peuplement (paragraphe 5.3.4) et nous les discuterons (paragraphe 5.3.5).

#### 5.3.1 **Comment estimer l'accroissement à l'échelle du peuplement ?**

Comme nous l'avons précisé dans la partie 3, nous ne disposons pas des accroissements pour tous les arbres d'un dispositif mais seulement d'un échantillon. De plus, les arbres de cet échantillon ont été choisis de façon à respecter la variabilité des variables de stratification (taille des arbres et environnement local). En conséquence, l'échantillon n'est pas représentatif de la situation moyenne des arbres dans le peuplement, ce qui aurait été le cas avec un échantillonnage aléatoire non stratifié. Nous ne pouvons donc pas calculer l'accroissement au niveau du peuplement directement à partir de ces données en utilisant par exemple la moyenne des accroissements individuels de l'échantillon. Cependant, nous disposons de la circonférence de tous les arbres du peuplement et nous partons du principe que la circonférence est la meilleure variable dont nous disposons pour expliquer la croissance (voir les résultats de la partie 4). Pour obtenir des données à l'échelle du peuplement avec l'échantillon des données individuelles sans utiliser un modèle, nous pouvons imaginer deux solutions qui tiennent compte de la spécificité de notre échantillon :

- o Première solution : à partir des données de croissance individuelle, nous pouvons calculer un accroissement moyen par classe de circonférence. Il suffit ensuite de calculer les effectifs par classe de circonférence dans le peuplement, puis de multiplier chaque effectif par l'accroissement moyen de la classe de circonférence obtenu à partir de l'échantillon. On obtient ainsi une estimation de l'accroissement de l'ensemble du peuplement. Par exemple,

pour les chênes du dispositif D42 (figure 5.10), à partir de l'échantillon, on peut définir un accroissement moyen pour 3 classes de circonférence et obtenir ensuite un accroissement sur l'ensemble du peuplement ;

o Deuxième solution : on constitue toujours des classes de circonférence mais cette fois, pour chaque arbre du peuplement, on tire une valeur d'accroissement provenant de l'échantillon dans la classe qui correspond à sa circonférence. Si on reprend l'exemple des chênes du dispositif D42 (figure 5.10) et si on considère la classe des arbres les plus petits nous devons réaliser 57 tirages avec remise dans la classe correspondante des arbres échantillonnés qui contient 10 arbres. Si on recommence deux fois l'opération, on n'obtiendra pas tout à fait le même résultat à cause de la variabilité des accroissements dans la classe de taille. On peut alors réitérer l'opération un grand nombre de fois (cette méthode est comparable aux méthodes de rééchantillonnage de type bootstrap) et calculer la valeur moyenne de l'accroissement au niveau du peuplement ainsi que son intervalle de confiance.

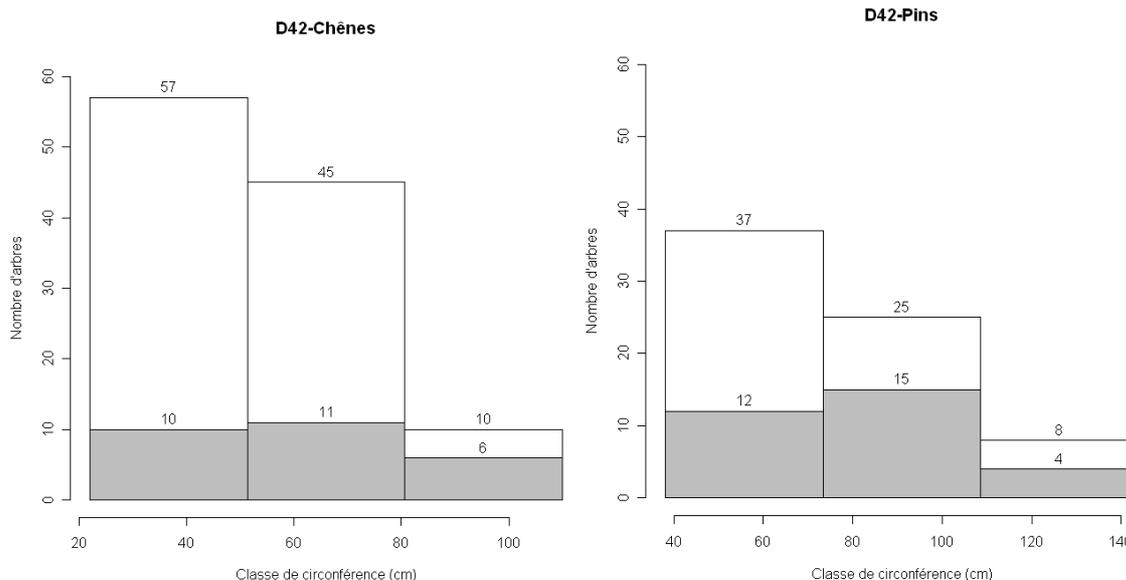


Figure 5.10 : Répartition des arbres en fonction des classes de circonférence retenues pour estimer l'accroissement en surface terrière sur le dispositif D42. En blanc l'ensemble des arbres sur la zone. En gris les arbres échantillonnés dans la zone. Les chiffres en haut des barres sont les effectifs des classes.

Dans cette sous-partie, nous avons opté pour la deuxième solution car elle permet d'apprécier l'erreur d'estimation de l'accroissement liée à la variabilité de la croissance individuelle. Cependant, dans notre cas nous avons peu de données d'accroissement par classe de taille. Si on se restreint à l'échantillon d'évaluation, le taux d'échantillonnage est de 6% pour le chêne et de 19% pour le pin. Si on prend l'ensemble des données, le taux d'échantillonnage passe à 15% pour le chêne et à 45% pour le pin. Pour avoir des résultats plus précis, nous avons choisi de prendre l'ensemble des arbres échantillonnés pour estimer

l'accroissement au niveau peuplement sur la période 2000-2005. Les données obtenues ne seront donc pas totalement indépendantes des données d'ajustement. La méthode proposée est d'autant plus pertinente que l'on définit un nombre de classes important. Cependant, étant donné le nombre d'individus échantillonnés disponibles, nous avons été contraints de prendre un faible nombre de classes. Pour chaque dispositif, nous avons donc défini trois classes de taille à partir de la circonférence mesurée lors de l'inventaire des dispositifs. Pour définir ces classes, nous nous sommes restreints à la zone centrale des dispositifs car aucun arbre n'a été échantillonné dans la zone périphérique (voir partie 3). Pour chaque dispositif et chaque essence, nous avons réalisé 10000 estimations de l'accroissement en surface terrière en utilisant la méthode décrite plus haut. La figure 5.11 présente les 10000 estimations obtenues pour le dispositif D42.

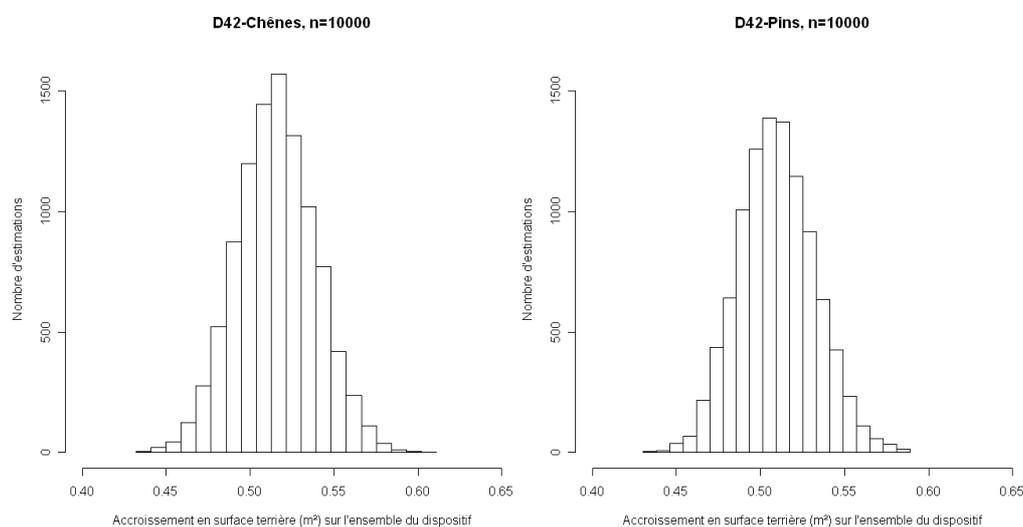


Figure 5.11 : Histogramme des 10000 estimations de l'accroissement en surface terrière sur la zone centrale du dispositif D42 pour le chêne et pour le pin.

### 5.3.2 Résultats sur l'estimation des accroissements au niveau du peuplement

Le tableau 5.9 donne les résultats concernant l'estimation des accroissements en surface terrière de chaque essence au niveau du peuplement.

Tableau 5.9 : Accroissements en surface terrière ( $\Delta G$  en  $m^2$ ) pour la période 2000-2005 par dispositif et par essence estimés à partir des accroissements individuels. IC- et IC+ sont les bornes inférieures et supérieures de l'intervalle de confiance à 95%.

Disp.	Surface (ha)	Chêne			Pin		
		$\Delta G$	IC-	IC+	$\Delta G$	IC-	IC+
D02	0.5653	0.7327	0.6835	0.7829	0.6183	0.5644	0.6770
D108	0.48	0.7086	0.6583	0.7587	1.2238	1.1645	1.2849
D20	0.5056	0.7733	0.7169	0.8326	1.1029	1.0221	1.1847
D27	0.315	0.4344	0.4036	0.4651	0.3974	0.3535	0.4461
D42	0.24	0.5156	0.4701	0.5619	0.5098	0.4672	0.5556
D49	0.6352	1.0412	0.9694	1.1144	1.0923	1.0017	1.1876
D534	0.36	0.7239	0.6876	0.7631	0.5338	0.4694	0.6007
D563	0.24	0.4755	0.4298	0.5219	0.4866	0.4355	0.5399
D78	0.37	1.0034	0.9331	1.0740	0.4210	0.3665	0.4824

Dans un peuplement mélangé, la production d'une espèce dépend évidemment de l'abondance de cette espèce dans le peuplement. Pour pouvoir comparer la production dans les différents dispositifs nous avons calculé la production d'une espèce en surface terrière rapportée à l'abondance de cette espèce dans le peuplement exprimée en surface terrière (figure 5.12). Ce rapport est un accroissement relatif et peut donc s'exprimer en pourcentage. Nous voyons que l'accroissement relatif du chêne est supérieur à celui du pin pour tous les dispositifs. Il varie sur la période considérée (années 2000 à 2005) de 2 à 3.9% par an pour le chêne et de 1 à 2.7% pour le pin. Ces résultats ne sont valables que dans le cadre de nos dispositifs car ils dépendent en particulier du stade de développement des arbres. Pour des populations plus jeunes le résultat serait très certainement inversé d'après ce que nous avons vu dans la partie 4. Ces données pourraient également être utilisées pour comparer l'accroissement relatif d'une essence entre un peuplement pur et un peuplement mélangé.

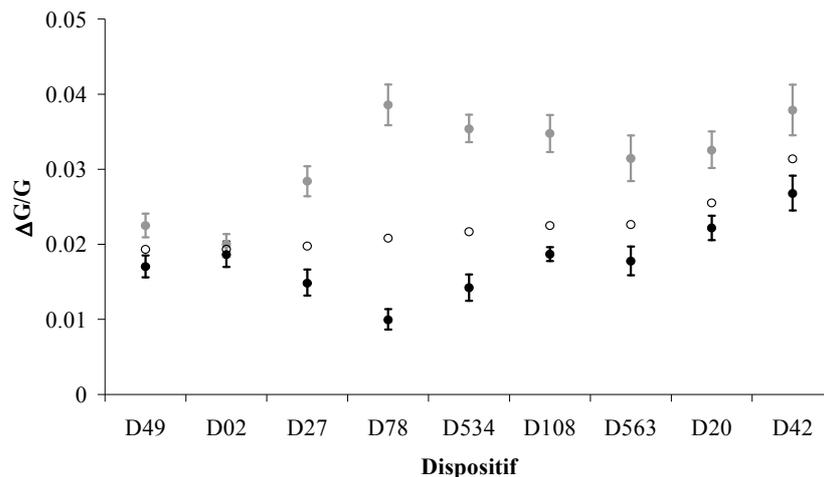


Figure 5.12 : Accroissement relatif du chêne et du pin dans les 9 dispositifs  $\Delta G/G$  (accroissement en surface terrière ( $m^2/an$ ) par  $m^2$  de surface terrière) calculé sur la période 2000-2005. Le chêne est en gris et le pin en noir. Les barres verticales représentent l'intervalle de confiance à 95%. Les ronds blancs représentent l'accroissement relatif global du peuplement avec les deux essences confondues. Les dispositifs sont classés de la gauche vers la droite selon accroissement relatif global croissant.

### 5.3.3 Méthodes retenues pour évaluer et comparer les modèles

#### 5.3.3.1 Méthodes retenues pour obtenir les données peuplements à partir des différents modèles

Les modèles que nous avons construits ont des spécificités différentes. Ils se distinguent notamment par les données d'entrée nécessaires pour lancer une simulation (tableau 5.10). Pour que les résultats des simulations soient comparables, il faut que les états initiaux de chacun des modèles soient cohérents. En particulier pour les données spatiales, les distributions de voisinage et les fonctions de caractérisation de la structure spatiale doivent être estimées à partir des arbres utilisés dans le modèle arbre dépendant des distances. Pour chaque dispositif et pour tous les modèles, nous avons donc utilisé comme état initial l'état du peuplement en 1999 sur la zone échantillonnée (les circonférences individuelles ont été reconstituées selon la méthode présentée dans le paragraphe 4.2.1). Les simulations des modèles individus-centrés ont été réalisées dans Capsis. Les simulations des modèles peuplement ont été réalisées dans un tableur Excel.

Tableau 5.10 : Données d'entrée nécessaires au lancement d'une simulation pour les différents modèles.

Modèle	Données individuelles	Données peuplement	Données spatiales	Remarques
MAID	Circonférence	Nom du dispositif	Aucune	
MADD	Circonférence	Nom du dispositif	Coordonnées de tous les arbres	Pour les arbres de bordure, nécessité de prendre en compte les effets de bord dans le calcul des indices de compétition
MADV	Circonférence	Nom du dispositif	4 distributions de voisinage	Pour les arbres de bordure, nécessité de prendre en compte les effets de bord dans le calcul du nombre de voisins.
MP <sub>MADD</sub>	Aucune	Nom du dispositif Pour chaque espèce : densité, surface terrière, rayon moyen	Aucune	
MPSS <sub>MADD</sub>	Aucune	Nom du dispositif Pour chaque espèce : densité, surface terrière, rayon moyen	Fonction de Ripley à 10 m pour chaque espèce. Fonction Intertype à 10 m	Pour les arbres de bordure, nécessité de prendre en compte les effets de bord dans le calcul des fonctions de Ripley et Intertype

### 5.3.3.2 Critères et tests retenus pour évaluer et comparer les modèles

Pour évaluer les modèles, nous avons comparé les données d'évaluation observées et les données prédites par les modèles. Nous avons utilisé un test des rangs signés sur données appariées en considérant les données observées et les données simulées au niveau d'un dispositif comme des données appariées. Pour chaque dispositif nous avons également regardé si la valeur donnée par le modèle se situait ou non dans l'intervalle de confiance à 95% de la valeur observée. Pour quantifier l'écart entre les données observées et les données simulées nous avons calculé le biais moyen :

$$B_s = \frac{1}{9} \sum_{i=1}^9 (\Delta G_{i,s,obs} - \Delta G_{i,s,pred})$$

où  $\Delta G_{i,s,obs}$  est l'accroissement en surface terrière observé au niveau peuplement pour l'espèce  $s$  du dispositif  $i$  et  $\Delta G_{i,s,pred}$  est l'accroissement en surface terrière prédit au niveau peuplement pour l'espèce  $s$  appartenant au dispositif  $i$ .

Nous avons également calculé la distance entre les valeurs observées et les valeurs prédites ( $D$ ) :

$$D_s = \sum_{i=1}^9 (\Delta G_{i,s,obs} - \Delta G_{i,s,pred})^2$$

Pour évaluer l'écart entre le MADD et les modèles qui en sont dérivés, nous avons calculé la distance  $D$  entre les valeurs données par le MADD et les valeurs données par les autres modèles (MADV, MP<sub>MADD</sub> et MPSS<sub>MADD</sub>). De plus, pour affiner ces résultats, nous

avons comparé les modèles MADD, MADV,  $MP_{MADD}$  et  $MPSS_{MADD}$  deux à deux avec un test des rangs signés sur données appariées.

Dans le cas du MADV, nous avons vu qu'il y a une variabilité introduite au moment de la reconstruction des voisinages. Comme dans la sous-partie 5.2, nous avons donc réalisé 1000 répétitions sur chaque dispositif avec ce modèle. Cependant, le tableau 5.11 montre que les prédictions du MADV concernant l'accroissement en surface terrière du peuplement sont peu variables d'une simulation à l'autre. Sur 1000 simulations, le coefficient de variation varie de 1.4% à 3.7% pour le chêne contre 0.9% à 2.1% pour le pin.

Tableau 5.11 : Ecart-type et coefficient de variation (CV) de l'accroissement en surface terrière prédit par le MADV pour chaque dispositif. Valeurs obtenues à partir des 1000 simulations réalisées par dispositif.

Dispositif	Chêne		Pin	
	Ecart-type (m <sup>2</sup> /ha)	CV	Ecart-type m <sup>2</sup> /ha	CV
D02	0.028	2.0%	0.011	1.0%
D108	0.039	2.4%	0.022	0.9%
D20	0.030	1.7%	0.019	1.0%
D27	0.046	3.7%	0.026	2.1%
D42	0.060	2.8%	0.016	0.8%
D49	0.038	1.8%	0.017	1.1%
D534	0.025	1.4%	0.030	2.0%
D563	0.026	1.6%	0.027	1.5%
D78	0.038	1.8%	0.026	2.1%

Pour simplifier les comparaisons avec les autres modèles, nous avons donc utilisé la moyenne des 1000 répétitions réalisées à le MADV sur chaque dispositif.

### 5.3.4 Résultats

Pour ce qui concerne le pin, nous voyons sur la figure 5.13 que les données prédites par les différents modèles sont proches des données observées. Le tableau 5.12 montre que les accroissements prédits se trouvant en dehors de l'intervalle de confiance des valeurs observées concernent peu de dispositifs (entre 2 et 4). Cependant, tous les modèles ont tendance à sous-estimer l'accroissement en surface terrière. Ainsi, comme le montre le tableau 5.12, les valeurs prédites qui sortent de l'intervalle de confiance des valeurs observées sont systématiquement inférieures aux valeurs observées. Pour le MADV cette différence est significative (tableau 5.12).

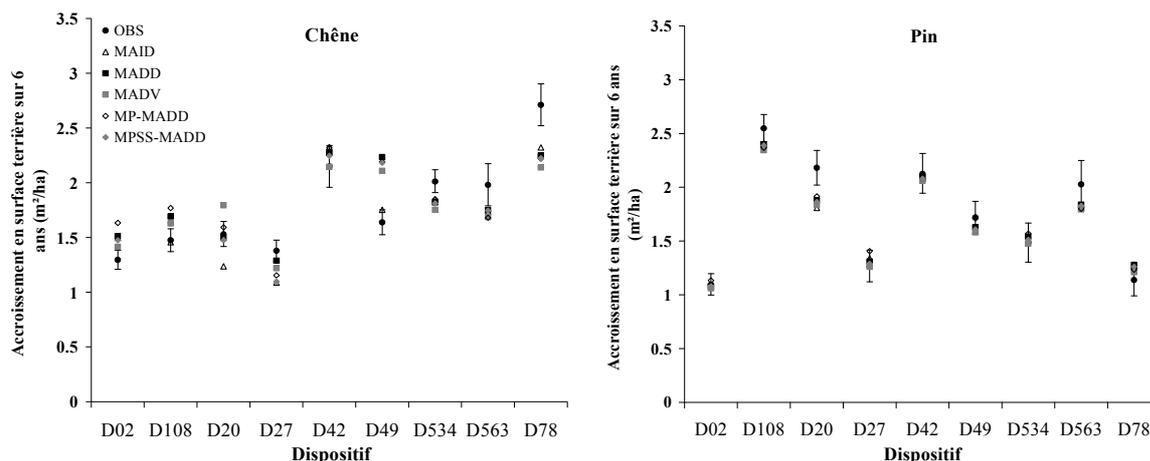


Figure 5.13 : Valeur observées de l'accroissement du peuplement en surface terrière et valeurs prédites par les différents modèles. OBS = valeurs observées ; MP-MADD =  $MP_{MADD}$  ; MPSS-MADD =  $MPSS_{MADD}$ . Les barres d'erreur sur les valeurs observées correspondent à l'intervalle de confiance à 95%.

Pour ce qui concerne le chêne, les résultats sont globalement moins bons que pour le pin puisque les accroissements prédits se trouvant en dehors de l'intervalle de confiance des valeurs observées concernent entre 6 et 8 dispositifs selon les modèles. Le MAID et les autres modèles ont un comportement légèrement différent. Le MAID obtient la valeur de  $D$  la plus faible. En revanche, sur les 9 dispositifs, il a tendance à sous-estimer l'accroissement en surface terrière du peuplement. Le MADD a une valeur de  $D$  plus forte mais ne montre pas une tendance systématique à la surestimation ou à la sous-estimation de l'accroissement en surface terrière. Les modèles MADD, MADV,  $MP_{MADD}$  et  $MPSS_{MADD}$  donnent des résultats relativement proches et pour ces modèles, ce sont globalement les mêmes dispositifs qui sont mal prédits (tableau 5.12). Le MADD est celui qui présente le moins de dispositifs mal prédits (6) alors que le MADV est celui qui en présente le plus (8).

Tableau 5.12 : Résultats de l'évaluation des modèles par rapport à la prédiction de l'accroissement en surface terrière au niveau peuplement. B = biais moyen ; D = somme des carrés des écarts entre valeurs observées et valeurs prédites ; Inf = liste des dispositifs dont la prédiction est inférieure à la borne inférieure de l'intervalle de confiance de la valeur observée ; Sup = liste des dispositifs dont la prédiction est supérieure à la borne supérieure de l'intervalle de confiance de la valeur observée.

Modèle	Essence	Test des rangs signés sur données appariés	B	D	Prédictions en dehors de l'intervalle de confiance	
					Inf	Sup
MAID	Chêne	NS	0.03111	0.07794	D20 D27 D534 D563 D78	D02
MADD	Chêne	NS	-0.03249	0.20735	D534 D563 D78	D02 D108 D49
MADV	Chêne	NS	-0.01642	0.17777	D27 D534 D563 D78	D02 D108 D20 D49
$MP_{MADD}$	Chêne	NS	-0.04167	0.22863	D27 D534 D563 D78	D02 D108 D49
$MPSS_{MADD}$	Chêne	NS	-0.01329	0.18520	D27 D534 D563 D78	D02 D108 D49
MAID	Pin	NS	0.03603	0.05255	D108 D20 D563	aucun
MADD	Pin	NS	0.02879	0.03747	D108 D20	aucun
MADV	Pin	p = 0.03906	0.04733	0.052603	D108 D20 D563	aucun
$MP_{MADD}$	Pin	NS	0.01729	0.033530	D108 D20 D563	aucun
$MPSS_{MADD}$	Pin	NS	0.03429	0.041246	D108 D20	aucun

Le tableau 5.13 montre que, dans le cas du chêne, le modèle le plus proche du MADD est le  $MPSS_{MADD}$ . Le plus éloigné est le MADV. Dans le cas du pin, le modèle le plus proche du MADD est le  $MPSS_{MADD}$  et le modèle le plus éloigné est le  $MP_{MADD}$ .

Tableau 5.13 : Distance  $D$  entre le MADD et les modèles qui en sont dérivés.

	Chêne				Pin		
	MADV	$MP_{MADD}$	$MPSS_{MADD}$		MADV	$MP_{MADD}$	$MPSS_{MADD}$
MADD	0.03667	0.01188	0.00623	MADD	0.003585	0.005568	0.000395

La comparaison deux à deux des modèles donne également des résultats différents suivant l'essence considérée (tableau 5.14). Dans le cas du chêne, le  $MPSS_{MADD}$  donne des prédictions significativement inférieures au MADD et au  $MP_{MADD}$ . Dans le cas du pin, le  $MPSS_{MADD}$  donne des prédictions significativement inférieures au MADD et le MADV donne des prédictions significativement inférieures aux modèles MADD,  $MPSS_{MADD}$  et  $MP_{MADD}$ .

Tableau 5.14 : Comparaisons deux à deux entre les prédictions des modèles MADD, MADV,  $MP_{MADD}$  et  $MPSS_{MADD}$  à l'aide d'un test des rangs signés sur données appariées. NS = test non significatif ; \* = test significatif au seuil de 5%.

	Chêne					Pin			
	MADD	MADV	$MP_{MADD}$	$MPSS_{MADD}$		MADD	MADV	$MP_{MADD}$	$MPSS_{MADD}$
MADD		NS	NS	*		*	NS	*	
MADV			NS	NS			*	*	
$MP_{MADD}$				*				NS	
$MPSS_{MADD}$									

### 5.3.5 Discussion

Pour le pin, nous avons montré que tous les modèles ont tendance à sous-estimer l'accroissement au niveau du peuplement. Ce résultat est cohérent avec l'évaluation du MAID et du MADD au niveau individuel car nous avons montré, dans le cas du pin, que le biais était positif donc que ces modèles sousestimaient l'accroissement au niveau individuel (tableau 5.7). Par contre, pour le chêne, l'évaluation au niveau individuel a montré plutôt une tendance à surestimer l'accroissement individuel, ce qui se retrouve bien dans l'évaluation du MADD au niveau peuplement, mais pas dans l'évaluation du MAID. En effet, dans le cas du chêne, le MAID a tendance à sous-estimer la croissance au niveau peuplement. Cette différence entre l'évaluation individuelle et l'évaluation au niveau du peuplement pourrait être liée au fait que la valeur du biais obtenu sur les données individuelles est proche de zéro.

Le MADV donne globalement des résultats cohérents avec les autres modèles. Cependant, il sous-estime l'accroissement en surface terrière du pin plus fortement que les autres. La comparaison avec le MADD montre que le MADV donne des résultats significativement inférieurs au MADD, mais uniquement dans le cas du pin. Ce résultat

provient certainement des méthodes et des hypothèses utilisées pour construire le MADV. En particulier, pour reconstruire les voisinages, nous avons fait l'hypothèse d'indépendance spatiale entre la taille d'un arbre et sa position. Cette hypothèse peut avoir des conséquences car, pour le pin, l'indice de compétition fait intervenir la taille de l'arbre. Nous constatons que cette approximation conduit à sousestimer l'accroissement au niveau du peuplement par rapport au MADD. Cela signifierait que la compétition estimée au niveau des pins dans le MADV est plus forte que la compétition estimée dans le MADD. Ainsi, un pin aurait autour de lui, en moyenne, des pins qui sont de plus petite dimension. Il y aurait donc une corrélation négative entre la taille d'un pin et la taille des pins voisins. Un autocorrélogramme (Goreaud, 2000 page 140) sur les circonférences du pin permettrait de vérifier cette hypothèse. Enfin, nous avons vu que lorsque nous réalisons plusieurs simulations sur un dispositif, les prédictions au niveau du peuplement sont très stables malgré la variabilité introduite au niveau individuel au moment de la reconstruction des voisinages. Ce résultat était prévisible car l'accroissement en surface terrière au niveau peuplement est une caractéristique globale. Pour ce modèle, cet accroissement dépend principalement de la circonférence des arbres et des indices de compétition. Si on considère un dispositif, les circonférences des arbres restent les mêmes d'une simulation à l'autre (on utilise la même liste d'arbres). Les caractéristiques des indices de compétition restent très stables (moyenne et variabilité) étant donné que, d'une simulation à l'autre, les nombres de voisins sont tirés dans les mêmes distributions et les voisins sont tirés dans la même liste d'arbres. D'un point de vue pratique, ceci signifie qu'il n'est pas obligatoire de faire beaucoup de simulations avec ce modèle pour obtenir des résultats sur la surface terrière totale.

Les comparaisons entre le MADD et les modèles peuplement qui en sont dérivés montrent que c'est le  $MPSS_{MADD}$  qui est le plus proche du MADD. C'est donc ce modèle qui a l'erreur d'agrégation la plus faible, ce qui est conforme à l'hypothèse formulée dans la partie 4 (paragraphe 4.4.4). Cependant, le  $MPSS_{MADD}$  donne des résultats significativement inférieurs au MADD à la fois pour le chêne et pour le pin. Nous pouvons essayer de relier ce résultat aux hypothèses formulées pour construire le  $MPSS_{MADD}$ . En particulier, nous avons utilisé l'approximation du champ moyen. Cette méthode a été appliquée sur le MADD et consiste à faire l'hypothèse que tous les arbres ont un même environnement moyen. En conséquence, la variabilité sur les accroissements en rayon prédits par le modèle résultant est réduite par rapport à celle prédite par le MADD. Comme on utilise la surface terrière pour comparer les

modèles, il est alors logique que le MADD donne des valeurs plus fortes que le  $MPSS_{MADD}$ <sup>19</sup>. Cette approximation expliquerait donc une partie du résultat concernant la différence entre le  $MPSS_{MADD}$  et le MADD. Sur cette évaluation, comme nous l'avons montré, les conséquences sont cependant faibles (tableau 5.13).

Il ressort également qu'il n'y a pas de différence significative entre le MADD et le  $MP_{MADD}$ . Encore une fois, ce résultat peut s'expliquer par les hypothèses formulées pour construire ce modèle. Dans ce modèle nous avons fait l'hypothèse de structure aléatoire des pins ce qui revient à considérer que la fonction de Ripley des pins ( $K(10)$ ) vaut 314.16. Dans le cas d'une structure agrégée ( $K(10) > 314.16$ ), cette hypothèse doit conduire à sousestimer la compétition donc à surestimer la croissance<sup>20</sup>. A l'inverse, dans le cas d'une structure régulière ( $K(10) < 314.16$ ), l'hypothèse doit conduire à surestimer la compétition donc à sousestimer la croissance. Le tableau 5.15 montre que les différences entre le  $MP_{MADD}$  et le MADD sont effectivement directement liées à cette hypothèse<sup>21</sup> (seul le dispositif D42 n'obtient pas le signe attendu). La différence entre le le  $MP_{MADD}$  et le MADD dépend donc de la structure spatiale. Dans nos dispositifs, la fonction de Ripley n'est pas toujours supérieure ou toujours inférieure à la valeur attendue dans le cas d'une structure aléatoire. Il n'y a donc pas une surestimation ou une sousestimation systématique du  $MP_{MADD}$  par rapport au MADD.

Tableau 5.15 : Fonction de Ripley à 10 m ( $K(10)$ ) pour les pins dans 9 dispositifs. Une valeur de  $K$  égale à  $\pi 10^2$  ( $\approx 314.16$ ) correspond à une structure aléatoire de l'espèce considérée, une valeur supérieure correspond à une structure agrégée et une valeur inférieure correspond à une structure régulière.  $MADD - MP_{MADD}$  = différence entre la prédiction du MADD et la prédiction du  $MP_{MADD}$  sur l'accroissement au niveau peuplement. Signe attendu = signe attendu de la différence. Valeur observée = valeur observée de la différence.

Dispo	K(10)	Valeur de K pour le $MP_{MADD}$	MADD- $MP_{MADD}$	
			Signe attendu	Valeur observée
D02	438	314.16	-	-0.0297
D108	305	314.16	+	0.0137
D20	354	314.16	-	-0.0186
D27	456	314.16	-	-0.0304
D42	316	314.16	-	0.0020
D49	388	314.16	-	-0.0533
D534	360	314.16	-	-0.0098
D563	299	314.16	+	0.0090
D78	292	314.16	+	0.0137

<sup>19</sup> En effet, réduire la variabilité sur les accroissements en rayon revient à réduire la variabilité sur les rayons individuels prédits. Or, lorsque l'on passe à la surface terrière totale, on fait intervenir la somme des carrés des rayons des arbres. La réduction de la variabilité entraîne nécessairement une réduction de cette somme de carrés.

<sup>20</sup> Pour le chêne l'analyse serait plus complexe car deux hypothèses interviennent : l'hypothèse d'indépendance spatiale entre les deux espèces et l'hypothèse d'une structure aléatoire des chênes. En fonction des dispositifs, ces deux hypothèses peuvent éventuellement se compenser. Ainsi, contrairement au pin, le résultat sur la croissance n'est pas immédiat.

<sup>21</sup> L'hypothèse d'indépendance spatiale entre la taille d'un arbre et sa position et l'approximation du champ moyen doivent également intervenir mais dans une moindre mesure.

Cette évaluation nous a permis de vérifier que les modèles que nous avons construits se comportent correctement à l'échelle du peuplement mais ne permet pas vraiment de départager les modèles. Les résultats obtenus avec les modèles individus-centrés sont cohérents avec l'évaluation réalisée sur les données individuelles. Cette évaluation nous a également permis de discuter les hypothèses formulées pour construire les différents modèles dérivés du MADD. C'est le modèle  $MPSS_{MADD}$  qui est le plus proche du MADD mais il donne des prédictions qui lui sont systématiquement inférieures. Les conclusions obtenues doivent cependant être relativisées à cause de la méthode utilisée pour obtenir les données observées et du nombre limité de dispositifs. Ces résultats constituent donc une première approche de l'évaluation des modèles au niveau peuplement. Des données indépendantes obtenues par suivi de croissance à l'échelle du peuplement seront bientôt disponibles dans le réseau de dispositifs installé par le Cemagref en forêt domaniale d'Orléans. Elles permettront de mettre en œuvre une seconde étape d'évaluation plus complète.

## **5.4 Prédire l'évolution du peuplement dans des contextes appliqués**

L'un des objectifs de la construction des modèles de croissance est leur utilisation dans un cadre appliqué. Dans cette sous-partie nous allons évaluer et comparer les modèles pour prédire l'évolution du peuplement dans des contextes appliqués. Pour cela nous allons présenter trois exemples correspondant à trois contextes appliqués. Le premier exemple concerne l'estimation de la production et des caractéristiques dendrométriques du peuplement après 6 années de croissance et vise principalement à rappeler les différentes informations que peuvent fournir les modèles que nous avons construits (paragraphe 5.4.1). Le deuxième exemple concerne la simulation de l'évolution du peuplement à moyen terme (paragraphe 5.4.2). Le troisième exemple concerne l'estimation de l'impact du type de structure spatiale sur la production (paragraphe 5.4.3).

### **5.4.1 Estimation de la production et des caractéristiques dendrométriques à court terme**

A différents niveaux dans la gestion forestière (révision d'aménagement, prévision de récolte, martelage), il est important de connaître la production des différentes parcelles d'une forêt et d'avoir quelques informations sur les caractéristiques dendrométriques des peuplements à court terme (< 10 ans). Un modèle de croissance est un bon outil pour répondre à ce besoin. Comme nous l'avons rappelé dans la partie 4, nos modèles ont un domaine de validité limité. C'est pourquoi nous nous limiterons dans ce paragraphe à des estimations après 6 années de croissance. Nous sommes dans la même configuration que lorsque nous avons évalué les différents modèles à l'échelle du peuplement. Nous insisterons donc ici plus sur les types de résultats que peuvent fournir les différents modèles que sur les différences entre les prédictions des modèles. De plus, nous n'avons pas fait de simulations sur l'ensemble des dispositifs mais nous avons pris quelques exemples de dispositif pour illustrer les propos.

A priori, si aucune contrainte sur les données d'entrée n'est imposée, tous les modèles que nous avons construits peuvent être utilisés pour estimer la production du peuplement. Les modèles de croissance nous permettent d'estimer la production totale en surface terrière du peuplement et de la répartir entre le chêne et le pin. Ainsi dans le cas du dispositif D02 la production totale sur la période 2006-2011 serait d'environ 0.61 m<sup>2</sup>/ha/an (entre 0.59 et 0.63 m<sup>2</sup>/ha/an selon les modèles). Le chêne représente 70% de cette production (entre 68 et 70% selon les modèles). Dans le dispositif D108 la production totale serait d'environ 0.75 m<sup>2</sup>/ha/ha (entre 0.68 et 0.78 m<sup>2</sup>/ha/an selon les modèles) et le chêne représente environ 50% de cette

production (entre 49 et 55% selon les modèles). Les modèles peuvent donner la surface terrière du peuplement et celle de chaque essence ainsi que les diamètres quadratiques moyens (tableau 5.16). Ces informations peuvent être utilisées comme critère de décision pour déclencher une éclaircie.

Tableau 5.16 : Evolution des caractéristiques dendrométriques du dispositif D20 d'après le MAID. N = nombre de tiges / ha ; G = surface terrière / ha ; Dg = diamètre quadratique moyen ; ΔG = accroissement en surface terrière.

Dispositif	Année	Chêne				Pin				Total	
		N (tiges/ha)	G (m <sup>2</sup> /ha)	Dg (cm)	ΔG (m <sup>2</sup> /ha)	N (tiges/ha)	G (m <sup>2</sup> /ha)	Dg (cm)	ΔG (m <sup>2</sup> /ha)	G (m <sup>2</sup> /ha)	ΔG (m <sup>2</sup> /ha)
D20	2005	448	11.3	17.9		119	11.3	34.8		22.6	
D20	2011	448	13.1	19.2	1.8	119	12.5	36.7	1.2	25.6	3.00

Les modèles individus-centrés peuvent apporter des résultats supplémentaires par rapport aux modèles peuplement. Ils peuvent notamment donner l'évolution de la distribution des individus par classe de diamètre (figure 5.14, nous avons pris l'exemple de la distribution en taille donnée par le modèle arbre indépendant des distances). Ce type d'information peut permettre par exemple de connaître la proportion d'arbres dans le peuplement qui a atteint le diamètre d'exploitabilité.

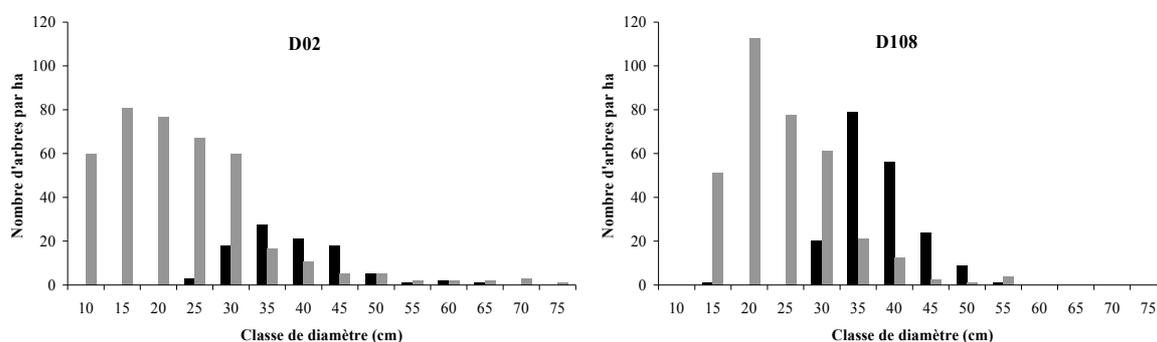


Figure 5.14 : Répartition des arbres par classe de diamètre après 6 années de croissance d'après le modèle arbre indépendant des distances. A gauche le dispositif D02 (état en 2011) et à droite le dispositif D108 (état en 2012). Les barres grises représentent le chêne et les barres noires représentent le pin.

Ces modèles permettent également de représenter les accroissements par classe de diamètre et par essence (figure 5.15). Nous pouvons ainsi constater que pour le dispositif D02 ce sont les chênes de la classe de diamètre 25 cm qui contribuent le plus à la croissance du peuplement alors que pour le dispositif D108 ce sont les pins de la classe de diamètre 35 cm.

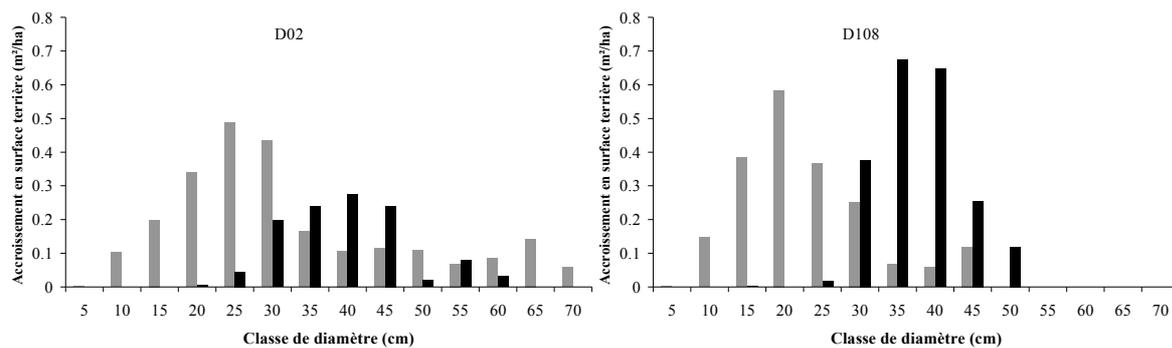


Figure 5.15 : Accroissement en surface terrière sur 6 années par classe de diamètre et par essence. Les barres grise représentent les données sur le chêne et les barres noires les données sur le pin.

Enfin, le MADD contrairement aux autres modèles, peut fournir des cartographies du peuplement. Cette information peut être intéressante pour mener des réflexions sur le martelage en forêt mélangée (figure 5.16). En effet, la phase de martelage dans les peuplements mélangés est délicate et demande une formation qui passe par des exercices pratiques de terrain (par exemple en utilisant des martélosopes). Le MADD peut servir de support à ces formations pour illustrer l'impact d'un martelage fictif sur le peuplement (Gauquelin *et al.*, 2008).

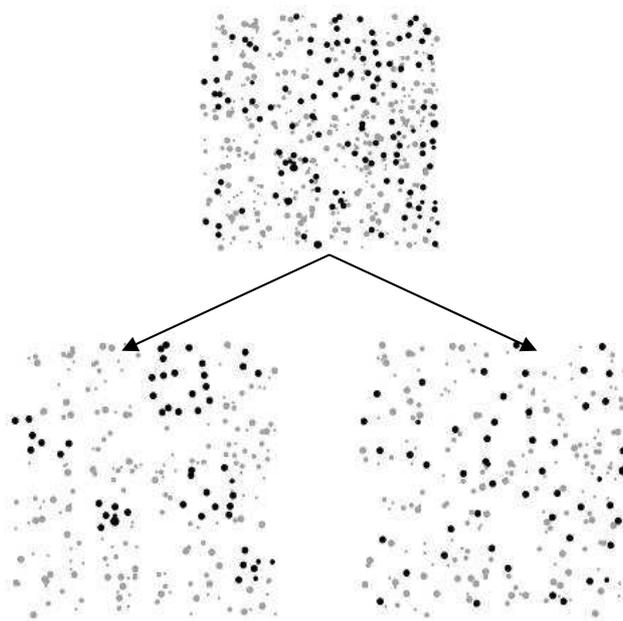


Figure 5.16 : Exemple de deux stratégies de martelage simulées dans le dispositif D49. Figure extraite de Perot *et al.* (2008). Les chênes sont représentés en gris et les pins en noir.

## 5.4.2 Evolution du peuplement à moyen terme

La gestion d'un peuplement mélangé pose un certain nombre de difficultés par rapport aux peuplements purs. L'une d'entre elles est liée au fait que les différentes espèces présentes dans un peuplement ont rarement le même âge d'exploitabilité ou le même diamètre d'exploitabilité. Par exemple, d'après le document d'aménagement du massif de Lorris (ONF, 2004) dans les stations sur lesquelles le mélange chêne-pin est préconisé, le diamètre d'exploitabilité du chêne est de 65 cm avec un âge total optimal fixé à 200 ans et le diamètre d'exploitabilité du pin est de 50 cm avec un âge total optimal fixé à 100 ans. Dans nos dispositifs, le diamètre moyen des pins est systématiquement supérieur à celui des chênes (tableau 3.2). On peut donc se douter que les pins seront exploités avant les chênes. La réalisation d'une simulation à moyen terme (entre 20 et 40 ans) peut nous permettre de préciser à partir de quand les pins seront exploitables et quel sera l'état du peuplement à ce moment là. Ces informations peuvent être utiles pour prendre des décisions sur l'avenir du peuplement ou peuvent servir de base de réflexion pour construire des tables de production en peuplement mélangé (ENGREF *et al.*, 1984).

Dans un premier temps nous allons présenter les dispositifs et les modèles que nous avons utilisés pour cette application (paragraphe 5.4.2.1). Nous expliquerons également comment nous avons réalisé les simulations (paragraphe 5.4.2.1). Ensuite, nous présenterons et discuterons la comparaison des modèles sur les simulations à moyen terme (paragraphe 5.4.2.2). Enfin, nous présenterons et discuterons les résultats quantitatifs obtenus à partir de ces simulations (paragraphe 5.4.2.3).

### 5.4.2.1 Choix des dispositifs et des modèles utilisés pour les simulations

Pour cette application, nous avons donc choisi d'utiliser deux dispositifs : le dispositif D108 où le pin est majoritaire en surface terrière, et le dispositif D02 où c'est le chêne qui est majoritaire (tableau 3.2). De plus, nous avons contrôlé que pour ces deux dispositifs, les paramètres  $\alpha$  et  $\beta$  du modèle de croissance arbre indépendant des distances étaient à peu près stabilisés (pour ces dispositifs, la figure 4.8 présentée dans la partie 4 montre que sur les dernières périodes de croissance, la valeur des paramètres fluctuent autour d'une valeur moyenne). Nous avons considéré que cela permettait d'envisager avec ce modèle des simulations à moyen terme (entre 20 et 40 ans). Nous avons également testé le MADD, le  $MP_{MAID}$  et le modèle  $MP_{MADD}$ . Pour cette application, nous n'avons pas testé le MADV car nous avons vu qu'il avait un comportement assez proche du MADD (paragraphe 5.2 et 5.3).

Nous n'avons pas non plus testé le modèle  $MPSS_{MADD}$  car ce modèle a un comportement proche du modèle  $MP_{MADD}$  et de plus il nécessite de connaître les valeurs des fonctions de caractérisation de la structure spatiale après une éclaircie<sup>22</sup>.

Pour faire les simulations, nous avons utilisé l'état des dispositifs (surface totale) lors du dernier inventaire (année 2005 pour le dispositif D02 et année 2006 pour le dispositif D108)<sup>23</sup>. Les arbres appartenant à d'autres espèces que le chêne et le pin n'ont pas été pris en compte dans les simulations. Dans la forêt domaniale d'Orléans les éclaircies sont réalisées environ tous les 10 ans. Afin de respecter le pas de temps de nos modèles (6 ans), nous avons réalisé une éclaircie tous les 12 ans prélevant l'accroissement en surface terrière de chaque essence. Nous avons fait l'hypothèse que ce régime d'éclaircie permettait d'être dans des conditions où la mortalité par compétition était négligeable. Enfin, afin de pouvoir comparer les éclaircies réalisées avec les différents modèles, nous avons utilisé le coefficient d'éclaircie  $kg$  qui est le rapport entre la surface terrière de l'arbre moyen enlevé lors de l'éclaircie ( $\bar{g}_e$ ) et la surface terrière de l'arbre moyen avant éclaircie ( $\bar{g}_{av}$ ) :

$$kg = \frac{\bar{g}_e}{\bar{g}_{av}} = \frac{G_e}{N_e} \times \frac{N_{av}}{G_{av}}$$

Ce coefficient peut s'exprimer en fonction du nombre de tiges enlevées lors de l'éclaircie ( $N_e$ ), de la surface terrière enlevée lors de l'éclaircie ( $G_e$ ), du nombre de tiges avant éclaircie ( $N_{av}$ ) et de la surface terrière avant éclaircie ( $G_{av}$ ). En fixant  $kg$  et connaissant  $G_e$ ,  $N_{av}$  et  $G_{av}$ , on en déduit  $N_e$ . Dans les simulations, nous avons fixé la valeur du  $kg$  à 0.85 pour les deux essences, ce qui correspond à des éclaircies mixtes qui prélèvent majoritairement des arbres plus petits que  $\bar{g}_{av}$  mais également des arbres plus gros.

A partir du moment où l'on réalise des éclaircies, les modèles peuplement que nous souhaitons tester ne sont pas facilement utilisables car ils utilisent le rayon moyen de chaque espèce. Pour s'affranchir de ce problème nous avons remplacé dans ces modèles le rayon moyen par le rayon quadratique moyen<sup>24</sup> (voir paragraphe 4.4.3.2 pour la définition). La moyenne quadratique est toujours supérieure à la moyenne arithmétique mais, cette approximation est raisonnable dans le cadre de nos dispositifs (figure 5.17). En comparant le MAID avec le modèle peuplement issu de son agrégation ( $MP_{MAID}$ ) nous pourrions apprécier l'impact de cette approximation.

<sup>22</sup> Néanmoins, si on fait l'hypothèse que les arbres enlevés lors de l'éclaircie sont choisis de manière aléatoire dans le peuplement alors on peut montrer que les fonctions de caractérisation de la structure spatiales sont inchangées.

<sup>23</sup> Le dispositif D108 a subi une éclaircie la même année que l'inventaire. Pour cette application, nous avons pris l'état du peuplement après éclaircie.

<sup>24</sup> On peut alors considérer qu'il s'agit de nouvelles versions de modèles peuplement incluant une approximation supplémentaire.

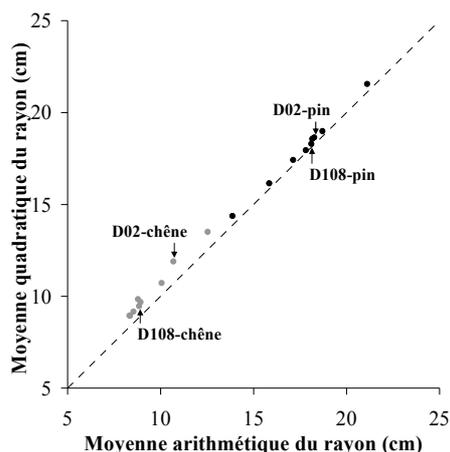


Figure 5.17 : Moyenne quadratique du rayon en fonction de la moyenne arithmétique. Les points gris sont les moyennes calculées sur les chênes et les points noirs sont les moyennes calculées sur les pins. La ligne en pointillés est la bissectrice.

#### 5.4.2.2 Comportement des modèles sur des simulations à moyen terme

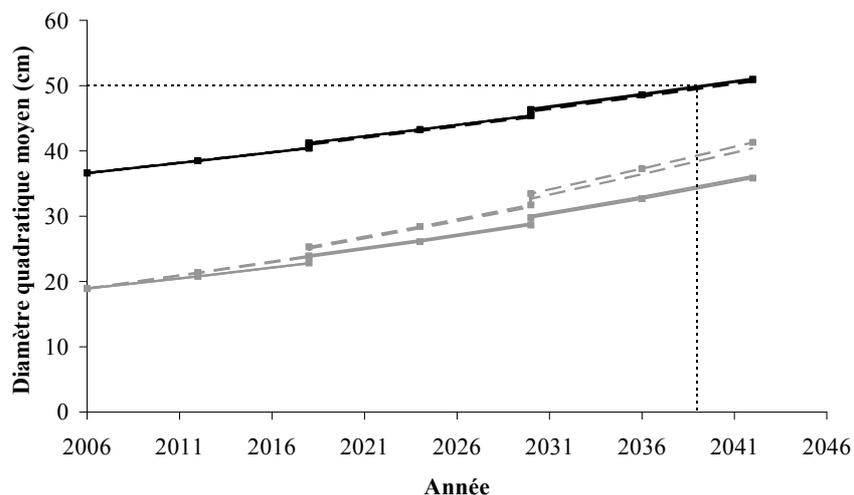


Figure 5.18 : Diamètre quadratique moyen prédit par les quatre modèles pour le dispositif D108. En gris les prédictions pour le chêne et en noir les prédictions pour le pin. Les lignes pleines représentent le MAID. Les lignes en pointillés représentent le MADD. Les lignes pleines avec un carré représentent le  $MP_{MAID}$ . Les lignes en pointillés avec un carré représentent le  $MP_{MADD}$ .

Pour ces comparaisons, nous nous sommes basés sur les résultats obtenus avec le dispositif D108. Le dispositif D02 donne les mêmes types de résultats. Pour le pin, les quatre modèles donnent les mêmes résultats pour ce qui concerne la prédiction du diamètre quadratique moyen (figure 5.18) ou la surface terrière (figure 5.19). Pour le chêne, nous pouvons distinguer deux groupes de modèles : le MADD et le  $MP_{MADD}$  d'un côté et le MAID et le  $MP_{MAID}$  de l'autre. Le MADD et le  $MP_{MADD}$  donnent des valeurs supérieures aux MAID et au  $MP_{MAID}$  pour le diamètre quadratique moyen et la surface terrière.

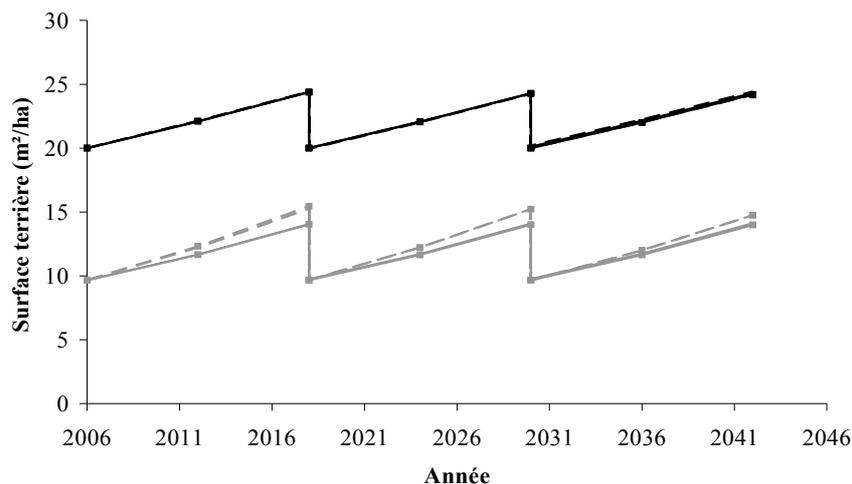


Figure 5.19 : Surface terrière prédite par les quatre modèles pour le dispositif D108. En gris les prédictions pour le chêne et en noir les prédictions pour le pin. Les lignes pleines représentent le MAID. Les lignes en pointillés représentent le MADD. Les lignes pleines avec un carré représentent le  $MP_{MAID}$ . Les lignes en pointillés avec un carré représentent le  $MP_{MADD}$ .

La comparaison que nous venons d'effectuer permet de dégager deux résultats principaux. Tout d'abord, le fait de remplacer le rayon moyen par le rayon quadratique dans les modèles peuplement n'a pas de conséquence importante sur les prédictions du diamètre quadratique et de la surface terrière. Ce résultat est intéressant car les modèles peuplement utilisent beaucoup moins de données que les modèles individus-centrés. Ensuite, le MADD (ainsi que le  $MP_{MADD}$ ) donne des résultats assez différents du MAID en ce qui concerne le chêne. La croissance du chêne prédite par le MADD est plus forte que celle prédite par le MAID. Ce résultat vient probablement de l'indice de compétition utilisé dans le MADD. Cet indice est basé sur le nombre de tiges. Il ne prend donc pas en compte la taille des individus, ce qui est gênant pour des simulations sur plusieurs pas de temps comme nous l'avons fait. De plus, le rayon de voisinage reste constant alors qu'il est probable qu'il évolue au cours du temps. Prévosto (2005) a par exemple montré que ce rayon augmentait en comparant des jeunes peuplements de pin sylvestre à des peuplements plus matures. Enfin, contrairement aux paramètres du MAID, nous n'avons aucune idée de l'évolution dans le temps des paramètres associés aux indices de compétition. Le MAID et le  $MP_{MAID}$  paraissent donc plus pertinent pour faire des simulations à moyen terme.

### 5.4.2.3 Résultats quantitatifs obtenus avec les simulations à moyen terme

Nous allons maintenant présenter et discuter les résultats quantitatifs obtenus à partir de ces simulations. Suite aux résultats que nous venons de présenter, nous allons nous focaliser sur les résultats donnés par le MAID ou le  $MP_{MAID}$ . La surface terrière totale varie entre 30 et 38  $m^2/ha$  pour le dispositif D108 et entre 28 et 36  $m^2/ha$  pour le dispositif D02. Ces niveaux sont assez proches des niveaux atteints dans les dispositifs sur la période de croissance utilisée pour construire les modèles (de 30 et 35  $m^2/ha$  pour le dispositif D108 et de 27 à 30  $m^2/ha$  pour le dispositif D02). Le guide de sylviculture du chêne en peuplement pur indique une surface terrière totale comprise entre 17 et 26  $m^2/ha$  pour la fertilité 2 (fertilité estimée sur nos dispositifs dans la partie 3) et pour un âge total compris entre 50 et 100 ans (Jarret, 2004 page 294). Pour le pin sylvestre, cette valeur varie entre 20 et 25  $m^2/ha$  pour une fertilité 3 et pour un âge total compris entre 50 et 100 ans (Chabaud et Nicolas, Sous presse). Nos simulations se situent donc à des niveaux plus élevés (tableau 5.17). Une éclaircie en début de simulation aurait été préférable si nous avions souhaité respecter ces chiffres.

Pour le dispositif D108, le diamètre d'exploitabilité du pin (50 cm) est atteint vers l'année 2039 soit après 33 années de croissance. Pour le dispositif D02, le diamètre d'exploitabilité du pin est atteint en 2035 soit après 30 années de croissance. En considérant l'âge médian des pins (tableau 4.3), ces dates correspondent à un âge de 99 ans à 1.30 m pour le dispositif D108 et à 108 ans pour le dispositif D02. Ces résultats sont cohérents avec l'âge d'exploitabilité espéré pour le pin dans l'aménagement (100 ans en âge total). Nous obtenons tout de même des âges supérieurs, ce qui peut être lié à la fertilité des dispositifs qui est plutôt faible pour le pin.

Dans le dispositif D108, lorsque les pins ont atteint le diamètre d'exploitabilité, les chênes ont une densité de 139 tiges/ha pour un diamètre quadratique de 35.8 cm. A ce stade, si tous les pins sont coupés, les chênes formeront un peuplement relativement clair (14  $m^2/ha$ ). En revanche, dans le dispositif D02, lorsque les pins ont atteint le diamètre d'exploitabilité, les chênes ont une densité de 211 tiges/ha pour un diamètre quadratique de 34.6 cm. Si tous les pins sont coupés, le peuplement de chêne restant aura une surface terrière de 19.9  $m^2/ha$  ce qui est dans la gamme des niveaux de surface terrière préconisés dans le guide de sylviculture du chêne. Nous voyons que dans un cas le passage du peuplement mélangé au peuplement pur de chêne peut se faire assez facilement alors que dans l'autre cas il peut poser problème à cause de la trop faible densité des chênes. Dans ce type de scénario, un mélange où le pin est très majoritaire n'est donc pas la meilleure situation à moins d'envisager un mode de gestion

alternatif permettant de gérer les deux essences en mélange de façon durable. On pourrait, par exemple, étaler la récolte du peuplement final de pin et rechercher des environnements favorables à la régénération du pin et du chêne.

Si l'ensemble des pins est coupé une fois le diamètre d'exploitabilité atteint, le MAID et le  $MP_{MAID}$  ne seront pas adaptés pour poursuivre la simulation car les espèces n'interviennent pas explicitement dans les équations de croissance, ce qui est une limite importante de ces modèles. De plus, la régénération qui peut intervenir suite à une éclaircie très forte n'est pas prise en compte dans ces modèles. Nous voyons ici l'importance de rester, lors de l'utilisation des modèles comme outils prédictifs, dans un contexte comparable à celui utilisé pour leur construction.

Tableau 5.17 : Table d'évolution des dispositifs D108 et D02 d'après le modèle MP<sub>MAID</sub>. Disp. = dispositif ; Age = âge médian à 1.30 m ; N = nombre de tiges ; G = surface terrière ; Dg = diamètre quadratique moyen.

Année	Disp.	Peuplement								Eclaircie					
		Age	Chêne			Age	Pin			N	G	Chêne		Pin	
			N (tiges/ha)	G (m <sup>2</sup> /ha)	Dg (cm)		N (tiges/ha)	G (m <sup>2</sup> /ha)	Dg (cm)			N (tiges/ha)	G (m <sup>2</sup> /ha)	N (tiges/ha)	G (m <sup>2</sup> /ha)
2006	D108	70	344	9.7	18.9	66	190	20.0	36.6	534	29.7				
2018	D108	82	344	14.0	22.8	78	190	24.4	40.4	534	38.4				
2018	D108	82	218	9.7	23.8	78	150	20.0	41.2	368	29.7	126	4.4	40	4.4
2030	D108	94	218	14.0	28.6	90	150	24.3	45.4	368	38.3				
2030	D108	94	139	9.7	29.8	90	119	20.0	46.3	258	29.7	79	4.3	31	4.3
2042	D108	106	139	14.0	35.8	102	119	24.2	50.9	258	38.2				
2005	D02	80	393	17.5	23.8	78	97	10.6	37.3	490	28.0				
2017	D02	92	393	22.7	27.1	90	97	12.9	41.3	490	35.6				
2017	D02	92	288	17.5	27.8	90	76	10.6	42.2	364	28.0	105	5.2	21	2.4
2029	D02	104	288	22.6	31.6	102	76	12.9	46.6	364	35.5				
2029	D02	104	211	17.5	32.5	102	60	10.6	47.4	271	28.0	77	5.1	16	2.3
2035	D02	110	211	19.9	34.6	108	60	11.7	49.8	271	31.5				
2041	D02	116	211	22.6	36.9	114	60	12.9	52.3	271	35.5				

### 5.4.3 Estimation de l'impact du type de structure spatiale sur la production

Dans un cadre de gestion des peuplements mélangés, il peut être intéressant d'estimer l'impact de la structure spatiale du peuplement sur la production par exemple pour guider les travaux dans les jeunes peuplements ou pour aider à la définition des consignes de martelage lors des éclaircies. L'objectif de ce paragraphe est de comparer les modèles dans leur capacité à estimer l'impact du type de structure spatiale sur la production en surface terrière du peuplement.

#### 5.4.3.1 Méthode

Nous avons choisi de comparer les résultats obtenus avec les modèles MADD, MADV et  $MPSS_{MADD}$ . Les autres modèles ne tiennent pas compte de la structure spatiale et ne peuvent donc pas être utilisés pour cette application.

Récemment, une typologie de la structure spatiale a été réalisée pour les peuplements mélangés chêne-pin de la forêt domaniale d'Orléans en utilisant les fonctions de Ripley et la fonction Intertype (Ngo Bieng *et al.*, 2006). Les types identifiés correspondent à un gradient allant du "mélange pied à pied" au "mélange par bouquet". De plus, des modèles de structure basés sur des processus ponctuels permettant de simuler la structure spatiale des différents types rencontrés ont été construits (Ngo Bieng, 2007). Pour répondre à l'objectif de ce paragraphe, nous avons choisi d'utiliser ces travaux. Nous avons donc comparé les modèles par rapport à deux types de structure spatiale simulés à l'aide des modèles de structure.

Pour étudier la structure spatiale des peuplements Ngo Bieng (2007) a distingué les arbres appartenant à la canopée et les arbres appartenant au sous-étage. Ainsi, il existe une typologie de la structure spatiale pour la canopée et une typologie pour le sous-étage. Dans ce travail nous nous sommes intéressés à deux types bien contrastés pour la canopée : le "mélange par bouquet" (type 1) et le "mélange pied à pied" (type 3). Pour obtenir des simulations réalistes nous sommes obligés de tenir compte du sous-étage. Pour simplifier, nous avons choisi un seul type de structure spatiale du sous-étage, appelé type 2 d'après la typologie construite par Ngo Bieng (2007). Nous avons donc comparé deux types de structure spatiale :

- Le type A "mélange par bouquet" qui associe le type 1 de la canopée au type 2 du sous-étage ;
- Le type B "mélange pied à pied" qui associe le type 3 de la canopée au type 2 du sous-étage ;

Pour avoir des données réalistes nous avons choisi l'un de nos dispositifs pour servir d'état initial aux simulations. Nous avons choisi un dispositif pour lequel les chênes et les pins appartenant à la canopée étaient suffisamment représentés<sup>25</sup>. Nous avons retenu le dispositif D42 qui répond à ces exigences. Nous avons utilisé la liste des arbres de ce dispositif puis nous avons simulé la structure spatiale de la canopée et du sous-étage à l'aide des modèles de structure sous Capsis (module oakpine1). Pour simuler un peuplement appartenant au type A, nous avons simulé la structure spatiale en utilisant les modèles de structure correspondant au type 1 de la canopée et au type 2 du sous-étage. Pour simuler un peuplement appartenant au type B, nous avons simulé la structure spatiale en utilisant les modèles de structure correspondant au type 3 de la canopée et au type 2 du sous-étage.

Comme les peuplements virtuels sont des réalisations de processus ponctuels utilisés par les modèles de structure, chaque peuplement simulé est différent. Dans le cas du MADD, nous avons donc simulé plusieurs peuplements virtuels pour estimer la production moyenne associée à un type. Nous avons simulé 50 peuplements virtuels pour chaque type de structure (figure 5.20). Pour chaque peuplement virtuel, nous avons ensuite simulé la croissance sur un pas de temps de 6 années (période 2000-2005) avec le MADD sous Capsis (module oakpine1).

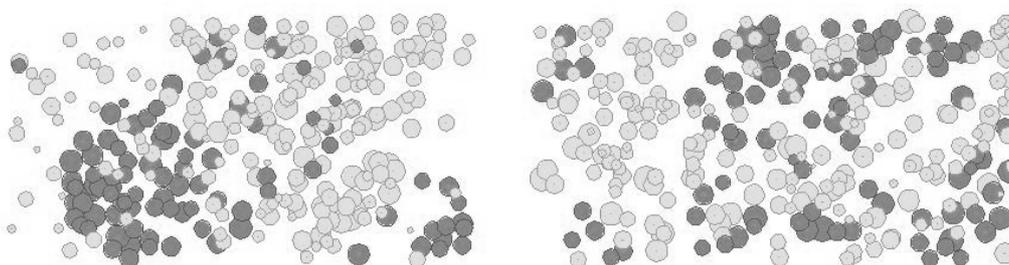


Figure 5.20 : Cartographies du dispositif D42 issues de Capsis. Les chênes sont en gris clair et les pins en gris foncé. A gauche un exemple de structure spatiale générée par le modèle de structure pour le type A ("mélange par bouquet"). A droite un exemple de structure spatiale générée par le modèle de structure pour le type B ("mélange pied à pied"). Les simulations sont réalisées avec la même liste d'arbres.

Pour utiliser le MADV, nous devons estimer les distributions de voisinage correspondant à chaque type de structure. Idéalement, il faudrait calculer les distributions théoriques correspondant aux modèles de structure du type A et du type B, mais ces calculs sont complexes et dépassent le cadre de cette thèse. Nous avons préféré estimer ces distributions sur des peuplements simulés avec les modèles de structure. Pour chaque type, nous avons donc simulé 1 peuplement de 10ha, 20 fois plus grand que le dispositif D42 mais respectant

<sup>25</sup> L'information concernant le statut des arbres (arbre de la canopée ou arbre du sous-étage) a été relevée pour tous les arbres de nos dispositifs lors du travail de thèse de Ngo Bieng (2007).

exactement les mêmes densités en chêne et en pin. grand nombre d'arbres simulés, cette méthode permet d'obtenir une bonne estimation des distributions de voisinage correspondant au type A et au type B (figure 5.21). Ces distributions nous ont permis de reconstruire les voisinages de chaque individu du dispositif D42 (voir sous-partie 4.3). Nous avons vu qu'il y avait une partie aléatoire lors de cette opération. Nous avons donc simulé 50 états initiaux pour chaque type de structure avec le MADV (module oakpine2 de Capsis). Pour chaque état initial nous avons ensuite simulé la croissance sur un pas de temps de 6 années.

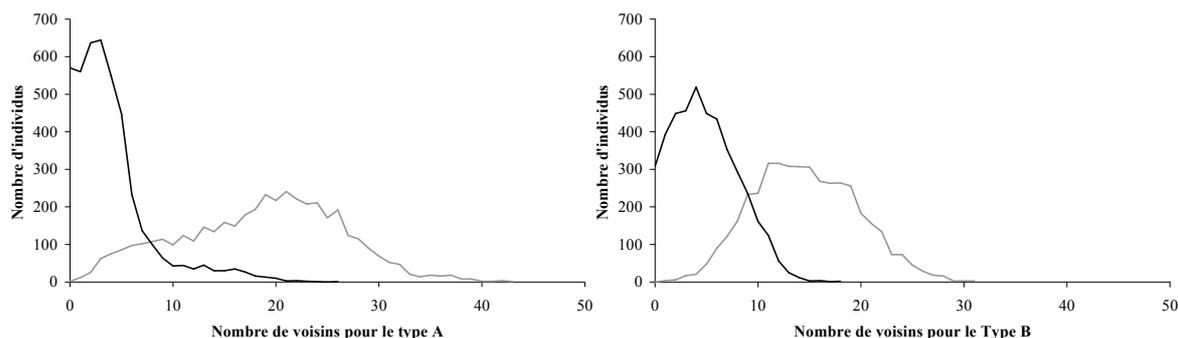


Figure 5.21 : Distributions du nombre de voisins dans un disque de rayon 10 m pour un arbre sujet appartenant à l'espèce chêne. Les courbes noires sont les voisins pins. Les courbes grises sont les voisins chênes. A gauche les distributions pour le type A ("mélange par bouquet"). A droite les distributions pour le type B ("mélange pied à pied"). Ces distributions ont été obtenues sur un dispositif de 10ha respectant les mêmes densités en chêne et en pin que le dispositif D42.

Pour utiliser le  $MPSS_{MADD}$ , nous devons estimer les fonctions de caractérisation de la structure spatiale du type A et du type B. Nous ne pouvons pas utiliser les valeurs théoriques des modèles de structure car les peuplements virtuels sont obtenus en appliquant successivement deux modèles de structure : un pour la canopée et un pour le sous-étage. Or le  $MPSS_{MADD}$  (comme les autres modèles) ne distingue pas les arbres de la canopée et du sous-étage. Nous avons donc estimé les fonctions de caractérisation de la structure spatiale sur les 2 peuplements utilisés précédemment pour estimer les distributions de voisinage du MADV (tableau 5.18). Ces paramètres étant fixés, pour ce modèle il n'y a qu'une seule prédiction de la production par type de structure. Cette prédiction a été simulée sur un pas de 6 années avec le  $MPSS_{MADD}$ .

Tableau 5.18 : Valeurs des fonctions de Ripley et de la fonction Intertype à 10 m pour les deux types de structure. Ces valeurs ont été obtenues sur un dispositif de 10ha respectant les mêmes densités en chêne et en pin que le dispositif D42.

Type	$K_{chêne}(10)$	$K_{pin}(10)$	$K_{chêne,pin}(10)$
Type A	431.7	544.1	597.9
Type B	341.1	381.1	434.9

Pour comparer les modèles dans leur capacité à estimer l'impact du type de structure spatiale du peuplement sur la production, nous avons représenté les prédictions des différents modèles pour chaque type de structure et pour chaque essence.

### 5.4.3.2 Résultats et discussion

La figure 5.22 montre qu'il y a effectivement une différence entre les prédictions des modèles selon le type de structure spatiale. C'est le "mélange pied à pied" (type B) qui a la production la plus forte à la fois pour le chêne et pour le pin. D'après les analyses réalisées par Ngo Bieng (2007), le "mélange pied à pied" se caractérise par une structure aléatoire ou légèrement agrégée des chênes et des pins<sup>26</sup>. La structure spatiale interspécifique y est caractérisée par une légère tendance à la répulsion ou quelquefois à une indépendance entre les 2 espèces. Le "mélange par bouquet" (type A) se caractérise par une structure fortement agrégée des chênes et des pins avec une répulsion interspécifique. Nous avons montré, pour le chêne comme pour le pin, qu'une structure spécifique agrégée est moins favorable du point de vue de la croissance qu'une structure aléatoire (voir paragraphe 4.4.3.4) car les indices de compétition prennent alors des valeurs plus fortes. En revanche, pour le chêne une structure en répulsion interspécifique est plus favorable pour la croissance qu'une structure d'indépendance. Comme la répulsion est présente dans les deux types, c'est essentiellement les structures spatiales spécifiques qui les différencient. Il est donc logique que le mélange pied à pied présente la production la plus forte. La figure 5.22 montre également que la variabilité des prédictions du MADD et du MADV est plus forte dans le type A que dans le type B. Ce résultat est également logique car la variabilité des environnements locaux est plus forte dans une structure agrégée que dans une structure aléatoire (voir paragraphe 2.5.2.1).

Pour le MADD et le MADV, il est intéressant de constater que les prédictions réalisées à partir des deux types de structure spatiale ne se recouvrent pratiquement pas (figure 5.22). Cela signifie que les productions prédites par ces modèles sont différentes d'un type à l'autre. Un test non paramétrique de comparaison de moyenne (test de Wilcoxon) montre une différence significative entre les deux types de structure spatiale pour les deux modèles (pour le MADD :  $W = 6$ ,  $p < 0.01$  ; pour le MADV :  $W = 0$ ,  $p < 0.01$ ). Pour le MADD et le MADV, L'augmentation de la production totale lorsque l'on passe du type 1 au type 3 est de 9% sur 6 ans. Elle est de 16% pour le  $MPSS_{MADD}$ .

---

<sup>26</sup> Ici nous ne parlons que des arbres de la canopée car les type A et B ont la même structure pour le sous-étage.

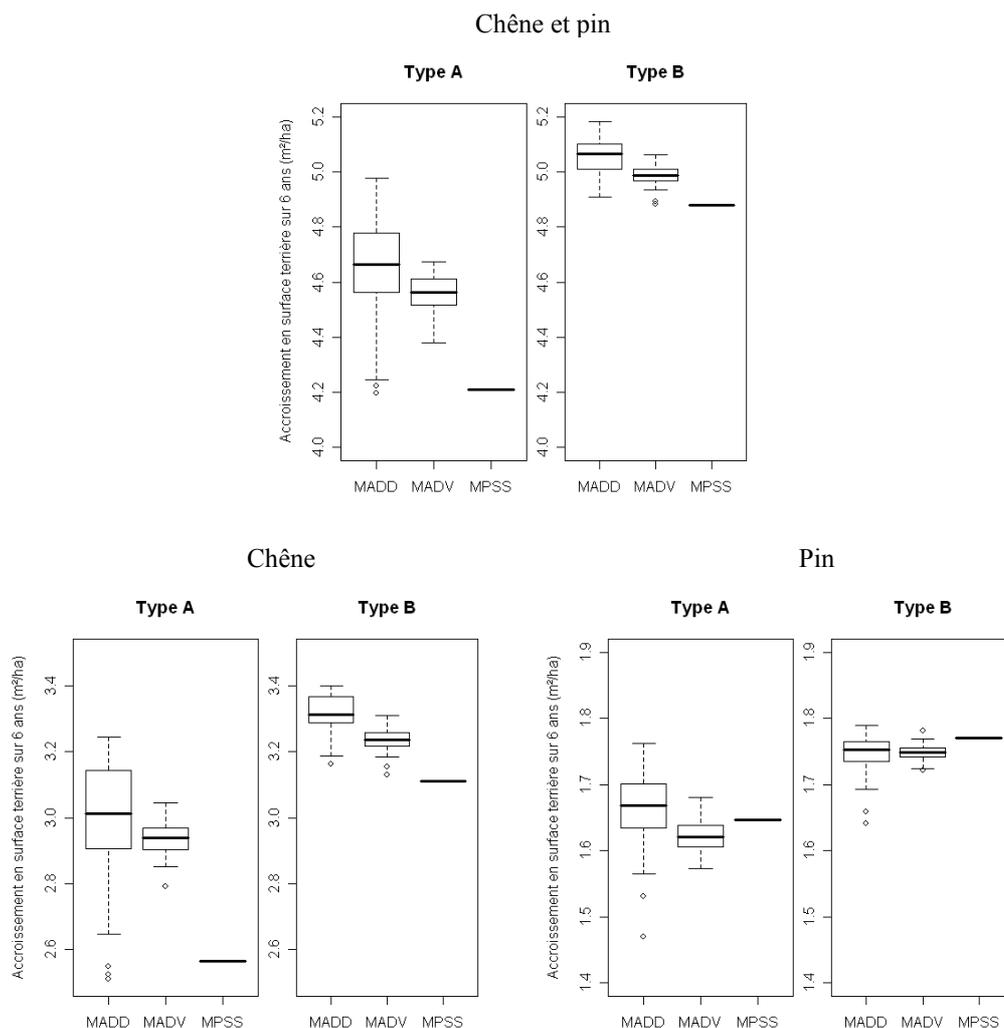


Figure 5.22 : Accroissement en surface terrière sur 6 années (m<sup>2</sup>/ha) pour la période 2000-2005 en fonction du type de structure spatiale pour le MADD, le MADV et le MPSS<sub>MADD</sub> (noté MPSS sur la figure).

Lorsque l'on compare les résultats des différents modèles, on constate que le MADD et le MADV donnent des résultats proches aussi bien pour les deux espèces que pour les types de structure. Le MPSS<sub>MADD</sub> donne des prédictions comparables aux autres dans le cas du pin. En revanche, pour le chêne il s'écarte des autres modèles et particulièrement dans le type 1. Dans le cas du pin, le résultat est probablement dû au fait que l'indice de compétition explique une faible part de la croissance. Les modèles donnent donc environ les mêmes résultats. Pour le chêne, une partie du résultat est certainement liée à l'application de l'approximation du champ moyen. En effet, nous avons vu que cela pouvait expliquer que le MPSS<sub>MADD</sub> donne des résultats inférieurs au MADD (voir paragraphe 5.3.5). D'après cette explication, la différence entre le MADD et le MPSS<sub>MADD</sub> devrait être d'autant plus grande que la variabilité des accroissements prédits par le MADD est importante. Or, comme on peut le voir sur la figure 5.22, le nombre de voisins chêne, qui intervient comme indice de compétition dans l'équation

de croissance du MADD (équation 4.8), est beaucoup plus variable dans le type A que dans le type B. La variabilité des accroissements prédits par le MADD est donc nécessairement plus importante dans le type A que dans le type B. Ceci expliquerait alors que la différence entre le MADD et le  $MPSS_{MADD}$  soit plus importante dans le type A que dans le type B. D'autres explications mériteraient cependant d'être recherchées.

#### 5.4.3.3 Conclusion

D'après les résultats que nous venons de présenter, le MADD le MADV et le  $MPSS_{MADD}$  donnent des prédictions différentes selon le type de structure spatiale. En ce sens, ce travail permet d'avoir une idée de la sensibilité de ces modèles à la structure spatiale (Goreaud *et al.*, 2005b). Cependant, le MADD et le MADV nous paraissent les plus pertinents pour estimer l'impact du type de structure sur la production du peuplement. En effet, selon le type de structure, le  $MPSS_{MADD}$  peut sousestimer plus ou moins fortement l'accroissement en surface terrière du peuplement.

La méthode que nous avons mise en œuvre présente l'intérêt de comparer des structures spatiales réalistes adaptées aux peuplements mélangés chêne-pin. En effet, ces structures ont été générées par un modèle de structure construit et évalué sur un réseau de dispositifs incluant l'ensemble des dispositifs que nous avons utilisés pour étudier la croissance (Ngo Bieng, 2007). L'impact sur la production est donc un impact possible et quantifiable qu'il peut être important de prendre en compte au niveau de la gestion de des peuplements mélangés. Néanmoins, les types que nous avons présentés sont des types très contrastés. De ce fait, il n'est probablement pas réaliste de passer d'un type de structure à l'autre en une seule éclaircie. D'autres types intermédiaires ont été définis (Ngo Bieng *et al.*, 2006) pour lesquels la même analyse pourrait être réalisée.

## 5.5 Bilan

Dans cette partie nous avons évalué et comparé une famille de modèles cohérents ayant des niveaux de détail différents par rapport à leurs capacités prédictives. Pour cela nous avons utilisé trois approches complémentaires dont nous allons reprendre les principales conclusions.

Contrairement à ce que l'on pouvait attendre, sur les données individuelles indépendantes d'évaluation, c'est le MAID qui donne les meilleurs résultats pour ce qui concerne les indices de performance globaux. Ce résultat met en évidence que le gain apporté par les indices de compétition pour la prédiction de la croissance individuelle est assez dépendant du jeu de données. Il souligne également l'importance de disposer de données indépendantes pour évaluer la capacité prédictive des modèles. En revanche, comme nous l'attendions, le MADD et le MADV donnent de meilleurs résultats en ce qui concerne la variabilité des prédictions de la croissance individuelle. Enfin, les trois modèles individus-centrés ont du mal à prédire des valeurs extrêmes d'accroissement.

L'évaluation et les comparaisons à l'échelle peuplement nous ont permis de montrer que les modèles peuplement que nous avons construits donnent des résultats tout à fait comparables aux modèles individus-centrés dont ils sont issus. A cette échelle, le  $MP_{MAID}$  donne, comme nous l'avons montré dans la partie 4, des résultats identiques au MAID. Parmi les modèles peuplement issus du MADD, c'est le  $MPSS_{MADD}$  qui en est le plus proche. Cette évaluation a également permis d'étudier l'impact des hypothèses formulées pour construire les modèles dérivés du MADD.

Enfin, l'évaluation par rapport à une utilisation dans un cadre appliqué a permis de faire ressortir certains atouts et certaines faiblesses des modèles. Ainsi les indices de compétition utilisés dans le MADD pour le chêne ne sont pas bien adaptés pour faire des simulations à moyen terme. En revanche, grâce à l'utilisation d'un modèle de structure permettant de générer des peuplements virtuels réalistes, nous avons montré que le MADD et le MADV permettent, sur le court terme, d'estimer l'impact d'un type de structure sur la production du peuplement. Le MAID et le  $MP_{MAID}$  apparaissent être les modèles les plus adaptés pour faire des simulations à moyen terme. Néanmoins, dans leur formulation actuelle, ils ne permettent pas de s'écarter trop fortement des conditions dans lesquelles ils ont été construits (composition du peuplement, niveau de surface terrière totale).

## 6. Discussion générale et perspectives

Dans cette dernière partie, nous allons reprendre les principaux résultats que nous avons acquis. Nous ferons ensuite une synthèse permettant de répondre à l'objectif principal de cette thèse qui était de *montrer comment la construction et la comparaison de modèles ayant différents niveaux de détail peuvent nous aider à choisir le niveau le plus approprié pour modéliser la croissance d'un peuplement mélangé, dans un contexte d'utilisation donné*. Nous terminerons par des perspectives évoquant des travaux à poursuivre ou des pistes que nous estimons intéressantes à explorer.

### 6.1 Bilan des principaux résultats

#### 6.1.1 Une famille de modèles cohérents ayant différents niveaux de détail

Pour modéliser la croissance des peuplements mélangés chêne sessile – pin sylvestre, nous avons construit une famille de modèles cohérents à différents niveaux de détail :

- **un modèle arbre indépendant des distances (MAID)** qui utilise la circonférence individuelle pour expliquer la croissance individuelle en rayon ;
- **un modèle arbre dépendant des distances (MADD)** qui intègre un indice de compétition dans l'équation utilisée par le MAID. Le MADD a été ajusté sur les mêmes données que le MAID, ce qui garantit une cohérence entre ces deux modèles ;
- **un modèle arbre qui utilise des distributions de voisinage (MADV)** pour reconstruire l'environnement local des arbres et calculer des indices de compétition. Les distributions de voisinage permettent d'utiliser l'équation de croissance du MADD sans utiliser une cartographie complète du peuplement. Ce modèle intègre une partie des informations spatiales et la variabilité des environnements locaux dans le peuplement. Pour reconstruire les voisinages, nous avons fait l'hypothèse principale d'indépendance spatiale entre la taille d'un arbre et sa position ;
- **un modèle peuplement en agrégeant le MAID ( $MP_{MAID}$ )**. Nous avons montré que ces deux modèles donnaient des résultats identiques à l'échelle du peuplement (agrégation parfaite) ;
- **un modèle peuplement issu du MADD utilisant des fonctions de caractérisation de la structure spatiale du peuplement ( $MPSS_{MADD}$ )**. Ce modèle a été obtenu en deux étapes : en simplifiant le MADD par approximation du champ moyen puis en agrégeant ce modèle simplifié. Là encore, des hypothèses ont été nécessaires pour obtenir ce modèle. Pour faire

l'approximation du champ moyen, nous avons fait l'hypothèse que la cartographie du peuplement était la réalisation d'un processus ponctuel ergodique (2.4). Pour intégrer les fonctions de caractérisation de la structure spatiale, nous avons de plus fait l'hypothèse que ce processus était homogène et isotrope. Dans le cas du pin, nous avons fait l'hypothèse d'indépendance spatiale entre la taille d'un pin et sa position ;

- **un modèle peuplement issu du MADD mais n'utilisant pas les fonctions de caractérisation de la structure spatiale ( $MP_{MADD}$ )**. Ce modèle implique des hypothèses supplémentaires par rapport au  $MPSS_{MADD}$  sur la structure spatiale du peuplement : structure aléatoire des chênes et indépendance spatiale des chênes et des pins dans le cas de la croissance du chêne ; structure aléatoire des pins dans le cas de la croissance du pin. Comme le  $MP_{MAID}$ , ce modèle utilise uniquement des variables globales du peuplement mais, contrairement à ce dernier, il intègre les rapports de compétition entre espèces provenant du MADD.

Ces modèles nous ont permis d'étudier différents aspects de la croissance des peuplements que nous allons détailler dans le paragraphe suivant.

## 6.1.2 Des modèles pour étudier la croissance

### 6.1.2.1 Le MAID permet d'étudier de façon rétrospective la croissance du chêne et du pin en mélange

Le MAID nous a permis de mettre évidence une relation classique entre l'accroissement individuel et la taille des arbres, ce qui était une hypothèse de départ dans notre travail (4.1). Grâce aux données rétrospectives, nous avons pu montrer que cette relation évoluait fortement au cours du temps et particulièrement dans le jeune âge dans le cas du pin (4.1.5). Ce modèle permet donc de faire ressortir les caractéristiques fortes des espèces par rapport à la croissance. Cependant, nous n'avons pas pu identifier précisément les facteurs qui agissent sur l'évolution de cette relation en particulier ceux qui sont responsables de la fluctuation des paramètres. Nous avons pu formuler des hypothèses sur la base des connaissances acquises dans les peuplements purs (Dhôte *et al.*, 2000 ; Perot *et al.*, 2007) mais il serait intéressant de pouvoir quantifier l'impact de ces facteurs sur les paramètres du modèle pour le rendre plus général.

Ce modèle a permis d'étudier des phénomènes d'interactions au cours du temps (4.1.6). Ainsi, nous avons mis en évidence un phénomène de compensation de croissance entre les deux espèces survenu après l'attaque d'un agent biotique : le lophyre du pin. Suite aux

défoliations subies par les pins, les chênes ont compensé la perte de production du peuplement par une augmentation de leur croissance. Ce phénomène serait intéressant à étudier par rapport à d'autres types de stress, comme les stress climatiques, qui peuvent avoir des effets sur la croissance différents suivant les espèces (Lebourgeois, 2006). Cela permettrait d'apporter d'autres éléments sur l'intérêt du mélange vis-à-vis des changements climatiques.

Le MAID est donc un outil intéressant pour étudier de façon rétrospective la croissance du chêne et du pin en mélange. En revanche, ce modèle, tel que nous l'avons construit, ne prend pas explicitement en compte l'effet d'une espèce sur la croissance de l'autre espèce.

#### 6.1.2.2 Le MADD permet d'étudier la compétition entre espèces

Le MADD nous a permis d'étudier l'influence du voisinage sur la croissance à l'aide d'indices de compétition locaux (4.2). Par rapport aux MAID, le MADD explique une plus grande part de la variabilité de la croissance. De plus, pour les deux essences étudiées, la compétition intraspécifique est plus forte que la compétition interspécifique. Ce résultat est fondamental puisqu'il conditionne le comportement du mélange à l'échelle du peuplement (Forrester *et al.*, 2006) et il peut être impliqué dans les mécanismes de coexistence dans les peuplements mélangés (Begon *et al.*, 1996). Cependant, les indices de compétition locaux sont plus ou moins pertinents selon l'espèce considérée (4.2.3). L'introduction d'un indice de compétition semble ainsi plus justifiée dans le cas du chêne que dans le cas du pin. L'utilisation des hauteurs individuelles dans les indices de compétition (Pretzsch, 2002) permettrait certainement d'affiner les résultats concernant la compétition entre ces deux espèces.

#### 6.1.2.3 Les modèles peuplement issus du MADD permettent d'étudier la croissance du peuplement en fonction du taux de mélange

A l'aide d'un modèle peuplement issu du MADD ( $MPSS_{MADD}$ ), nous avons vu que l'on pouvait étudier l'influence du taux de mélange sur la production du peuplement (4.4.3). Le MADD et MADV auraient pu être utilisés pour répondre à cette question mais ils nécessitent des simulations qui peuvent être lourdes à entreprendre. Le  $MPSS_{MADD}$  nous a permis d'exprimer le taux de mélange optimal en fonction des caractéristiques du peuplement. Il ressort qu'il n'existe pas un taux optimal unique. Il varie entre 38% et 74% de la surface terrière totale en fonction des dispositifs. Etant donné les résultats obtenus avec le MAID concernant l'évolution des paramètres de croissance au cours du temps, on peut penser que ce

taux évolue non seulement en fonction des caractéristiques dendrométriques des différentes populations mais également en fonction du temps. Les dispositifs que nous avons utilisés ne couvrent pas l'ensemble du gradient allant du peuplement pur de chêne au peuplement pur de pin. Pour étendre le domaine de validité de nos résultats, il serait intéressant de compléter ce gradient en installant de nouveaux dispositifs ou en faisant de l'expérimentation.

A cause des approximations que nous avons faites, nous avons vu que le  $MPSS_{MADD}$  n'a pas tout à fait le même comportement que le MADD au niveau peuplement. Cependant, les approximations ne concernent pas le rapport de compétition entre les deux espèces traduit par les paramètres associés aux indices de compétition. En conséquence, on peut penser que les résultats concernant le taux optimal de mélange ne changent pas trop en utilisant un modèle peuplement dérivé du MADD plutôt que le MADD lui-même. Il serait tout de même intéressant de le vérifier en réalisant des simulations à partir du MADD.

### 6.1.3 Des modèles pour prédire la croissance

La cohérence que nous avons assurée entre les modèles nous a permis de comparer leur capacité à prédire la croissance. Nous avons mis en œuvre trois approches complémentaires apportant chacune des éléments de réponse quant à la pertinence des modèles comme outils prédictifs. Nous avons d'abord utilisé un jeu de données indépendant pour évaluer la qualité de la prédiction individuelle. Nous avons ensuite évalué la qualité de la prédiction au niveau du peuplement en utilisant des données d'accroissement des peuplements estimées à partir de notre échantillon de données individuelles. Enfin, nous avons évalué les modèles sur des exemples d'applications nécessitant des simulations à court ou moyen terme.

#### 6.1.3.1 Au niveau individuel

Sur la base des critères classiquement utilisés pour ce type d'évaluation, nous avons montré que sur le jeu de données indépendant, le MAID était plus performant que le MADD pour prédire les accroissements individuels (5.2). Ce résultat était plutôt contraire à nos hypothèses étant donné les résultats obtenus lors de la calibration des modèles. En revanche, conformément à nos attentes, nous avons montré que le MADD et le MADV étaient plus performants que le MAID pour prédire la variabilité des accroissements. Nous avons également montré que le MADD était le modèle le plus performant pour prédire la distribution des accroissements. Enfin, nous avons montré que les biais obtenus avec le MAID, dépendaient des dispositifs ce qui n'est pas le cas avec le MADD. Le MADD, grâce à

la prise en compte de l'environnement local, pourrait donc mieux "s'adapter" lorsque l'on change de peuplement.

### 6.1.3.2 Au niveau du peuplement

L'évaluation sur la prédiction de la croissance au niveau du peuplement nous a permis d'avoir une comparaison de l'ensemble des modèles (5.3). A cette échelle, il est difficile de départager les modèles car les résultats dépendent des critères et de l'essence considérée. Le MADV est le seul modèle présentant une différence significative avec les valeurs observées mais uniquement pour ce qui concerne le pin. Nous avons également vu que cette évaluation permettait de mesurer l'impact des approximations réalisées pour construire les modèles dérivés du MADD. Ainsi, conformément à ce que l'on attendait, le  $MPSS_{MADD}$  est plus proche du MADD que le  $MP_{MADD}$ . En confrontant les hypothèses formulées pour construire ces modèles aux données de structure spatiale réelles des peuplements, nous avons vu que l'on pouvait expliquer les différences obtenues entre le MADD et un modèle dérivé.

### 6.1.3.3 A court ou moyen terme

La comparaison des modèles dans des contextes appliqués (5.4) a permis de montrer que le MADD que nous avons construit n'était pas bien adapté à des simulations à moyen terme (de 20 à 40 ans), à cause notamment de l'indice utilisé pour la croissance du chêne. Cela pose la question du choix des indices de compétition au moment de l'ajustement des modèles mais cela montre également leurs limites et notamment le fait qu'ils sont statiques (Burton, 1993). En effet, nous ne connaissons pas l'évolution avec le temps du paramètre associé à cet indice. De plus, il est probable que le rayon de voisinage évolue au cours du temps. C'est pourquoi certains auteurs ont proposé d'utiliser des méthodes pour prendre en compte cette évolution. Par exemple, Lorimer (1983) propose d'utiliser la taille des couronnes pour déterminer le rayon de voisinage. En revanche, sur des simulations à court terme (< 10 ans) le MADD et le MADV permettent d'estimer l'impact d'un type de structure sur la production du peuplement. Ce résultat peut être intéressant pour mettre au point des stratégies de martelage ou, d'un point de vue pédagogique, pour visualiser l'impact d'un martelage sur l'évolution du peuplement (Courbaud *et al.*, 2001 ; Gauquelin *et al.*, 2008). Nous avons conclu que le MAID et le  $MP_{MAID}$  étaient les plus adaptés pour des simulations à moyen terme. Néanmoins, dans leur formulation actuelle, ce sont des modèles frustes qui ne permettent pas de tester tous les types de scénarios sylvicoles.

## 6.2 Synthèse sur les critères permettant de choisir un type de modèle

Les résultats que nous venons de récapituler peuvent nous aider à choisir un type de modèle en fonction du contexte d'utilisation ce qui était l'objectif principal de cette thèse (1.4). Dans le travail que nous avons réalisé, deux facteurs permettent de définir le contexte d'utilisation : les données nécessaires et l'objectif d'utilisation du modèle. Nous avons considéré un autre facteur qui peut influencer le choix d'un type de modèle : l'interface informatique nécessaire pour réaliser des simulations. Le tableau 6.1 récapitule la pertinence des modèles en fonction du contexte d'utilisation. Nous allons maintenant expliquer et discuter ce tableau de synthèse. Nous allons commencer par définir les contextes d'utilisation (paragraphe 6.2.1.1). Nous verrons ensuite comment l'interface informatique peut influencer le choix d'un type de modèle (paragraphe 6.2.1.2). Nous discuterons les résultats présentés dans le tableau de synthèse (paragraphe 6.2.1.3) et enfin, nous présenterons les limites de notre travail (paragraphe 6.2.1.4).

### 6.2.1.1 Les contextes d'utilisation sont définis par les données nécessaires et les objectifs d'utilisation

Les données nécessaires à l'utilisation des modèles peuvent être séparées en trois catégories : les données à l'échelle du peuplement, les données individuelles et les données spatiales. Ainsi, parmi les modèles que nous avons construits, certains nécessitent uniquement des données à l'échelle du peuplement ( $MP_{MAID}$  et  $MP_{MADD}$ ) alors que d'autres nécessitent de connaître la circonférence et la position de chaque arbre dans le peuplement (MADD).

Un modèle peut être utilisé comme outil d'explication ou comme outil de prédiction. Selon Ward (2008), l'utilisation d'un modèle comme outil d'explication répond plutôt à une problématique de recherche alors que son utilisation comme outil de prédiction répond plutôt à une problématique de gestion. Nous avons donc distingué les objectifs d'utilisation explicatifs et les objectifs d'utilisation prédictifs. Les objectifs d'utilisation explicatifs visent à mieux comprendre la croissance et la dynamique du mélange. Dans ces objectifs, nous incluons l'étude de la croissance rétrospective des espèces en peuplement mélangé, l'étude de la compétition entre espèces et l'étude de l'impact du taux de mélange sur la production du peuplement. Les objectifs d'utilisation prédictifs intègrent la prédiction des accroissements individuels, la prédiction de la distribution et de la variabilité des accroissements individuels, la prédiction de l'accroissement du peuplement, la prédiction de l'évolution du peuplement à court terme en fonction du type de structure et l'évolution du peuplement à moyen terme.

Nous nous sommes limités aux objectifs traités dans cette thèse mais on pourrait en définir beaucoup d'autres. A chaque croisement entre un modèle et un objectif, nous avons estimé la pertinence du modèle d'après les résultats que nous avons récapitulés dans le bilan précédent. Dans cette thèse, nous n'avons pas pu évaluer tous les modèles pour tous les objectifs. Dans le cas où un modèle n'a pas été évalué par rapport à un objectif, nous avons estimé sa pertinence par rapport à cet objectif en nous appuyant sur les résultats obtenus avec un ou plusieurs modèles similaires.

#### 6.2.1.2 L'interface informatique peut influencer le choix d'un type de modèle

Nous avons été amenés à implémenter les modèles que nous avons construits pour pouvoir faire des simulations. Deux modules oakpine1 (figure 6.1) et okpine2, correspondant respectivement au MADD et au MADV, ont été implémentés sur la plate-forme Capsis (Coligny *et al.*, 2003). Pour les simulations impliquant le MAID, nous avons provisoirement utilisé ces modules en adaptant les équations de croissance. Pour les simulations utilisant les modèles peuplement, nous avons simplement utilisé un tableur Excel. Nous estimons que les modèles individus-centrés nécessitent un simulateur "élaboré" alors qu'un outil informatique plus léger peu suffire pour les modèles peuplement. De plus, nous avons été amenés à simuler des éclaircies avec les différents types de modèles. Le degré de complexité de l'éclaircie que l'on peut réaliser avec un modèle peut également être un critère de sélection. Ainsi, le MADD permet de réaliser des éclaircies très fines puisque les arbres peuvent être sélectionnés à partir de la cartographie du peuplement. On peut, par exemple, simuler le détournement d'arbres ou simuler une éclaircie correspondant au martelage réalisé dans un martelloscope. Le MAID et le MADV permettent de réaliser des éclaircies fines mais moins élaborées que le MADD. On peut sélectionner les arbres en fonction de leur taille ou d'autres caractéristiques individuelles. Les modèles peuplement ( $MP_{MAID}$ ,  $MP_{MADD}$  et  $MPSS_{MADD}$ ) permettent également de réaliser des éclaircies mais de façon grossière en fixant une densité cible (surface terrière ou nombre de tiges) et un coefficient d'éclaircie. Les résultats concernant ces deux critères supplémentaires (complexité du simulateur de croissance et finesse des éclaircies simulées) n'ont rien de nouveau. Cependant, nous les avons ajoutés dans notre tableau de synthèse car ils peuvent intervenir comme critères de sélection, notamment lorsque l'on souhaite utiliser un modèle comme outil d'expérimentation virtuelle.

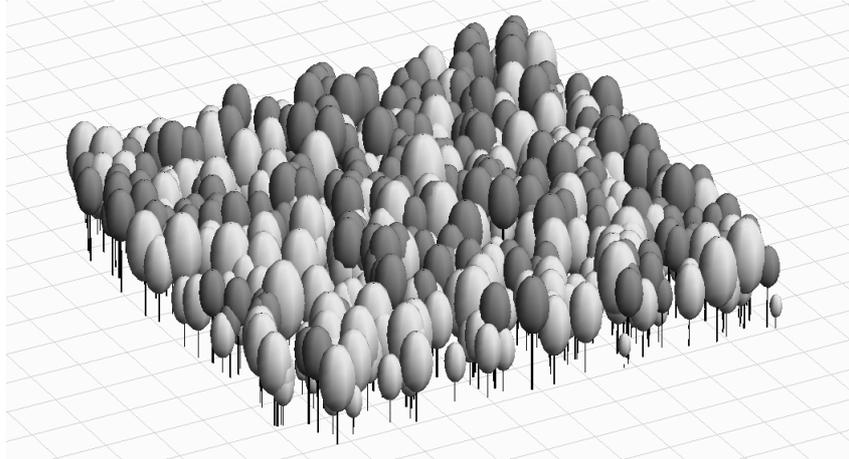


Figure 6.1 : Visualisation du dispositif D49 à l'aide du simulateur de croissance Capsis, module oakpine1 (MADD). Les chênes sont en gris clair et les pins en gris foncé

Tableau 6.1 : Tableau récapitulant les critères permettant de choisir un type de modèle. - = modèle inadapté pour cet objectif ; + = utilisation adaptée ; ++ = utilisation adaptée, modèle jugé le plus pertinent par rapport aux autres modèles pour cet objectif ; +- = utilisation possible mais modèle jugé non pertinent dans sa forme actuelle pour cet objectif. Les parenthèses signifient que l'évaluation du modèle pour cet objectif n'a pas été réalisée dans le cadre de cette thèse. Le signe entre parenthèse correspond au résultat attendu déduit de la connaissance des autres résultats obtenus avec un ou plusieurs modèles similaires. L'indice c désigne une variable concernant les chênes. L'indice p désigne une variable concernant les pins. N = densité en tiges/ha ; G = surface terrière en m<sup>2</sup>/ha ; R = rayon moyen en cm ; K (10) = fonction de Ripley à 10 m ; K<sub>c,p</sub>(10) = fonction Intertype à 10 m. {c<sub>i</sub>} = liste des circonférences individuelles ; {(x<sub>i</sub>,y<sub>i</sub>)} = coordonnées des arbres dans le peuplement ; D<sub>c,p</sub> = distribution de voisinage des voisins pins pour un arbre sujet chêne. Accr = accroissement. (\*) Prédiction à court terme (< 10 ans).

Modèle	Données nécessaires			Objectif d'utilisation explicatif			Objectif d'utilisation prédictif					Interface informatique	
	Peuplement	Individuelles	Spatiales	Croissance rétrospective des espèces en mélange	Compétition entre espèces	Impact du taux de mélange sur la production	(*) Accr. individuel	(*) Distribution / Variabilité des accr. individuels	(*) Accr. du peuplement	(*) Evolution en fonction de la répartition spatiale	Evolution à moyen terme (20 à 40 ans)	Simulateur élaboré	Simulation d'éclaircies
MP <sub>MAID</sub>	N <sub>c</sub> , G <sub>c</sub> , R <sub>c</sub> N <sub>p</sub> , G <sub>p</sub> , R <sub>p</sub>			-	-	-	-	- / -	+	-	+	Non	Grossière
MP <sub>MADD</sub>	N <sub>c</sub> , G <sub>c</sub> , R <sub>c</sub> N <sub>p</sub> , G <sub>p</sub> , R <sub>p</sub>			-	-	(++)	-	- / -	+	-	+-	Non	Grossière
MPSS <sub>MADD</sub>	N <sub>c</sub> , G <sub>c</sub> , R <sub>c</sub> N <sub>p</sub> , G <sub>p</sub> , R <sub>p</sub>		K <sub>c</sub> (10) K <sub>p</sub> (10) K <sub>c,p</sub> (10)	-	-	++	-	- / -	+	+	(+-)	Non	Grossière
MAID		{c <sub>i</sub> }		+	-	-	++	+ / +	+	-	+	Oui	Fine
MADV		{c <sub>i</sub> }	D <sub>c,c</sub> D <sub>c,p</sub> D <sub>p,c</sub> D <sub>p,p</sub>	-	-	(+)	+	+ / ++	+	++	(+-)	Oui	Fine
MADD		{c <sub>i</sub> }	{(x <sub>i</sub> ,y <sub>i</sub> )}	-	+	(+)	+	++ / ++	+	++	+-	Oui	Très fine

### 6.2.1.3 Quel modèle pour quel objectif ?

Si on lit le tableau 6.1 avec pour seule entrée l'objectif d'utilisation, sans contrainte sur les données et l'interface graphique, on peut conclure que les modèles les plus adaptés sont :

- le MAID pour étudier la croissance rétrospective du peuplement mélangé et pour prédire les accroissements individuels à court terme ;
- le MAID et le  $MP_{MAID}$  pour prédire l'évolution du peuplement à moyen terme ;
- le MADD pour étudier la compétition entre espèces ;
- le MADD et le MADV pour prédire à court terme la variabilité et la distribution des accroissements individuels et pour prédire l'impact à court terme du type de structure spatiale sur la production ;
- le  $MPSS_{MADD}$  et le  $MP_{MADD}$  pour étudier l'impact du taux de mélange sur la production ;

Il peut être difficile de relier directement ces résultats à ceux d'autres travaux car ils ne comparent pas toujours les mêmes types de modèles et ne considèrent pas forcément les mêmes objectifs. Néanmoins, nous faisons ressortir ci-après quelques points communs.

**Un modèle très détaillé, en particulier un MADD, n'est pas toujours celui qui a le meilleur pouvoir prédictif.** La performance d'un MAID par rapport à un MADD pour prédire la croissance d'un peuplement mélangé a été étudiée par plusieurs auteurs (Prévosto, 2005). Les conclusions diffèrent suivant les études. Ainsi, certains travaux concluent plutôt en faveur du MAID (par exemple Biging et Dobbertin, 1995 ; Wimberly et Bare, 1996) alors que d'autres concluent en faveur du MADD (par exemple Stadt *et al.*, 2007 ; Pukkala et Kolström, 1987 cité par Prévosto, 2005). Pour aller plus loin que la simple comparaison d'un MAID et d'un MADD, Astrup *et al.* (2008) ont comparé un MAID à différentes versions plus ou moins complexes d'un MADD pour prédire la croissance du peuplement. D'un point de vue prédictif, ils concluent que le MAID est moins performant car il fait trop d'approximations. Le MADD le plus détaillé ne convient pas non plus car, comme dans notre travail, il est trop dépendant du jeu de données sur lequel il a été ajusté (problème de surajustement). Par rapport à la prédiction de la croissance du peuplement, ils retiennent donc une version intermédiaire entre ces deux modèles. Ce modèle intermédiaire leur permet de faire un compromis entre la généralité du modèle et la précision des prédictions : c'est le modèle le plus parcimonieux. Ils précisent que le résultat de ce type de comparaison dépend de la quantité de données disponibles. Avec plus de données, les paramètres du modèle le plus détaillé auraient pu être

estimés de façon plus précise et ce modèle aurait certainement été retenu pour prédire la croissance du peuplement.

**Un modèle agrégé peut donner des résultats aussi bons voire meilleurs qu'un modèle individus-centré.** La comparaison entre un modèle individus-centré et des modèles plus agrégés est aussi une question qui a été traitée dans le domaine forestier. En effet, les modèles agrégés ont l'avantage d'être plus simples à construire et à utiliser que des modèles individus-centrés (Ek et Monserud, 1979 ; Gourlet-Fleury *et al.*, 2005), ce qui milite pour leur utilisation dans un cadre de gestion. Le modèle agrégé fournit parfois des résultats aussi bons que le modèle individus-centré. Ainsi, Gourlet-Fleury *et al.* (2005) ont comparé un MADD avec un modèle de distribution (modèle matriciel). Ils concluent que les deux types de modèles donnent des résultats comparables pour les critères choisis. De plus, dans certains cas, comme nous l'avons montré, l'agrégation parfaite entre un modèle individus-centré et un modèle peuplement est possible (Kokkila *et al.*, 2006). Le modèle agrégé peut aussi être plus performant que le modèle individus-centré. Après avoir comparé un MAID et un modèle peuplement, Burkhart (2003) conclut que le MAID est le modèle le plus fiable pour des prédictions allant jusqu'à 6 ans mais qu'au-delà le modèle peuplement donne les meilleurs résultats. Evidemment, ces résultats dépendent des critères utilisés et des objectifs considérés. Ainsi, Ek et Monserud (1979), en comparant un MAID avec un modèle de distribution, concluent que le MAID est plus adapté pour rendre compte de l'effet des éclaircies et pour prendre en compte la régénération.

**La simplification d'un MADD peut apporter des connaissances sur l'importance des effets spatiaux sur le système étudié et sur le fonctionnement du MADD.** D'autres travaux se sont intéressés à la comparaison de modèles issus de la simplification d'un MADD. Picard et Franc (2001) constatent que leurs modèles simplifiés donnent des résultats différents du MADD, ce qui permet de mettre en évidence l'importance des effets spatiaux dans le fonctionnement du MADD. De même, Verzelen *et al.* (2006) ont obtenu plusieurs versions d'un MADD en simplifiant les différents processus modélisés. Ils concluent que la comparaison du MADD avec ses versions simplifiées permet de mieux comprendre le fonctionnement du système étudié et, en particulier, les interrelations entre la distribution spatiale des arbres à l'échelle du peuplement et les processus dynamiques à l'échelle individuelle.

#### 6.2.1.4 Enjeux et limites de notre démarche

La question que nous avons abordée dans cette thèse ne concerne pas uniquement la recherche forestière. Ainsi, nous avons illustré l'intérêt de mener en parallèle des approches à différents niveaux de détail (de l'individu au peuplement) pour étudier un phénomène. Ce travail confirme également que la pertinence d'un type de modèle dépend des contraintes et de l'objectif d'utilisation. Par ailleurs, nous avons montré que l'on pouvait préciser, dans le cas du mélange chêne-pin, les avantages et les inconvénients des différents modèles que nous avons construits en les comparant sur la base d'un ensemble de critères.

Notre démarche présente cependant plusieurs limites :

- nous avons construit et évalué un nombre limité de modèles. En particulier, nous n'avons pas testé de modèle de distribution ;
- nous avons étudié la croissance en diamètre des arbres à une hauteur de 1.30 m, ce qui est une vision partielle du processus de croissance. De plus, nous n'avons pas pris en compte les processus de régénération et de mortalité ;
- nos modèles ont un domaine de validité limité. En particulier, leurs paramètres n'évoluent pas au cours du temps ce qui ne permet pas des projections à long terme ;
- certaines études, qui auraient pu apporter un nouvel éclairage sur la pertinence des modèles, n'ont pas été mises en œuvre dans cette thèse. C'est le cas des études de sensibilité ou de l'évaluation des modèles par les utilisateurs potentiels.

### 6.3 Perspectives

Suite à cette thèse, nous avons choisi de présenter trois axes pour de futures recherches : le premier portant sur l'évaluation et la comparaison des modèles, le deuxième sur l'amélioration des modèles construits dans cette thèse, et le troisième sur l'application de notre démarche aux processus de mortalité et de régénération.

#### 6.3.1 Evaluation et comparaison des modèles

Plusieurs travaux nous paraissent nécessaires pour enrichir une démarche d'évaluation et de comparaison de modèles de croissance comme celle que nous avons présentée.

Afin de compléter la gamme des modèles que nous avons construits, il serait intéressant de construire un modèle de distribution. En effet, ce type de modèle a montré son efficacité dans le cas des forêts tropicales. Ainsi, Picard (2007) estime que les modèles de distribution sont bien adaptés pour prédire la production à la fois par rapport à leurs performances mais également par rapport aux pratiques actuelles des gestionnaires de ces forêts. Ce modèle pourrait être construit en estimant les probabilités de transition à partir de l'équation de croissance du MAID comme nous l'avons proposé dans la partie 2.

Dans les méthodes utilisées pour évaluer les modèles, les études de sensibilité pourraient apporter de nouveaux éléments de connaissance sur les modèles. Ces études pourraient se faire à plusieurs niveaux : sur la forme des relations, sur les paramètres des équations et sur les données d'entrée. Par exemple, il serait intéressant d'étudier l'importance du paramètre lié à l'effet taille ( $\beta$ ) et de le comparer avec l'importance de l'effet dispositif. En effet, les imprécisions sur le paramètre  $\beta$  ont peut-être autant d'impact sur les prédictions que le fait de changer de dispositif. De même, comme nous l'avons évoqué dans la partie 4, il serait intéressant d'étudier l'impact de l'imprécision du rapport de compétition entre espèces sur la détermination du taux de mélange optimum du point de vue de la production.

L'évaluation participative a, comme nous l'avons présenté dans la partie 5, plusieurs avantages (voir aussi Pretzsch *et al.*, 2002b). Elle permet d'avoir une évaluation externe de nos modèles par des utilisateurs potentiels. En effet, les connaissances des experts sur le système étudié sont différentes de celles des modélisateurs. Des anomalies peuvent apparaître aux yeux des experts qu'il est important de comprendre ou de prendre en compte. Elle peut permettre de définir de nouveaux critères pour juger de la pertinence d'un modèle dans un contexte d'utilisation donné. Enfin, elle peut favoriser l'appropriation des modèles par les utilisateurs et contribuer au transfert des résultats que nous avons obtenus. C'est pourquoi,

nous pensons qu'il est important de réaliser ce type d'évaluation à la suite de notre travail. À ce titre, la construction d'un module ou d'une interface unique sous Capsis rassemblant l'ensemble des modèles pourrait faciliter leur évaluation par des utilisateurs potentiels.

Enfin, si on augmente le nombre de critères d'évaluation, il peut devenir difficile de choisir entre les différents types de modèles. Une comparaison basée sur une analyse multicritère (Pauwels *et al.*, 2007) peut alors être une perspective intéressante pour résoudre ce type de problème.

### 6.3.2 Améliorer ou compléter les modèles de croissance

Plusieurs travaux pourraient être envisagés pour améliorer les modèles que nous avons construits. Ces travaux d'amélioration contribueront à augmenter le domaine de validité des modèles et donc à pouvoir les comparer dans de nouvelles situations. Tout d'abord, la modélisation de l'effet dispositif et de l'évolution des paramètres au cours du temps nous paraissent les premiers aspects à améliorer dans nos modèles. En ce qui concerne l'effet temporel, le travail pourrait se baser sur l'exploitation des données rétrospectives. Dans un premier temps, ces données pourraient permettre d'intégrer l'effet de l'âge sur les paramètres du MAID (voir paragraphes 4.1.4 et 4.1.5). Néanmoins, l'âge n'est pas très satisfaisant pour des peuplements hétérogènes et d'autres moyens devront être recherchés. Concernant l'effet dispositif, l'influence de la fertilité ou des caractéristiques dendrométriques du peuplement comme la surface terrière totale devra être modélisée. La prise en compte de la fertilité pourrait se faire en utilisant des variables quantitatives caractéristiques du milieu et en les reliant au potentiel de croissance des arbres (Lejeune, 1996 ; Monserud et Sterba, 1996 ; Pretzsch, 2002 ; Trasobares *et al.*, 2004). Il nous semble également intéressant d'explorer la piste de l'utilisation des indices de fertilité (Puettmann *et al.*, 1992 ; Gobakken et Naesset, 2002 ; Andreassen et Tomter, 2003 ; Lee *et al.*, 2004 ; Hein et Dhote, 2006). Pour étudier correctement l'influence de ces facteurs, il pourrait être nécessaire d'intégrer de nouveaux dispositifs. Prochainement, le suivi du réseau de dispositifs installé par le Cemagref en forêt domaniale d'Orléans permettra de compléter nos données pour répondre à ce besoin.

Nous rappelons que la question que nous avons posée dans cette thèse se plaçait en amont d'un projet scientifique visant à construire un modèle de dynamique forestière (Allain *et al.*, 2005). Une prochaine étape de ce projet sera d'intégrer les processus de régénération et de mortalité. D'un point de vue scientifique, ces processus sont indispensables pour comprendre la dynamique d'un peuplement mélangé sur le long terme. D'un point de vue pratique, l'intégration de ces processus doit permettre de tester des sylvicultures plus variées

incluant des scénarios extrêmes. Ainsi, des travaux sont en cours au Cemagref de Nogent-sur-Vernisson sur la régénération et la mortalité du pin sylvestre (Gaudio *et al.*, 2008a ; Gaudio *et al.*, 2008b) et un module générique de régénération est en voie de construction sur la plateforme Capsis (Donès *et al.*, 2009).

### 6.3.3 Appliquer la démarche aux processus de mortalité et de régénération

Nous avons vu que la construction et la comparaison de modèles ayant différents niveaux de détail étaient riches pour appréhender la croissance d'une forêt mélangée aussi bien pour la compréhension du phénomène que pour sa modélisation dans un but prédictif. C'est pourquoi nous pensons que cette démarche pourrait être bénéfique pour la modélisation des processus de mortalité et de régénération.

Pour la mortalité, on peut envisager une approche "individus-centré" qui prendrait en compte la taille des individus et éventuellement leur croissance sur les dernières années (Kobe et Coates, 1997 ; Wyckoff et Clark, 2000 ; Wunder *et al.*, 2007) ou leur environnement local (Zhao *et al.*, 2006). Mais on peut également envisager une approche à l'échelle du peuplement, par exemple en étendant le concept des lois d'auto-éclaircie aux peuplements mélangés (Puettmann *et al.*, 1992 ; Rivoire, 2009).

De même, une des approches fines pour modéliser la régénération consiste à établir une relation entre la quantité de lumière disponible en sous-bois et le développement des jeunes plants. Contrairement aux peuplements réguliers (Sonohat *et al.*, 2004), dans les peuplements complexes cette quantité de lumière est difficile à mettre en relation avec des variables simples du peuplement comme la surface terrière (Piboule *et al.*, 2005). Pour prédire la lumière disponible en sous-bois dans ces peuplements, des modèles d'interception de la lumière par les arbres de la canopée sont alors nécessaires (Courbaud *et al.*, 2003). Des travaux récents ont cherché à évaluer l'impact du niveau de description du peuplement sur la précision de l'estimation de la lumière transmise à travers la canopée de peuplements hétérogènes (Da Silva, 2008). Ces travaux pourraient permettre de comparer des modèles de régénération plus ou moins détaillés. Des modèles plus simples pourraient également être évalués. On pourrait par exemple utiliser des modèles de recrutement pour prédire le nombre d'arbres qui atteint un seuil de taille en fonction du site et de caractéristiques du peuplement (Vanclay, 1994 ; Hasenauer, 2006).

La démarche pourra également être appliquée à des modèles de dynamique forestière complets qui couplent les processus de régénération, de croissance et de mortalité (Verzelen *et al.*, 2006). Comme pour le travail que nous avons réalisé sur la croissance, l'une des

difficultés pour mener à bien ces travaux est d'assurer la cohérence entre les différents types de modèles.

## 7. Conclusion

Les peuplements mélangés sont des systèmes complexes dont la gestion nécessite des outils adaptés comme les modèles de croissance. L'objectif de cette thèse était de montrer comment la construction et la comparaison de modèles ayant différents niveaux de détail pouvaient nous aider à choisir le niveau le plus approprié pour modéliser la croissance d'un peuplement mélangé dans un contexte d'utilisation donné. Nous avons appliqué ce travail à la croissance des peuplements mélangés chêne sessile – pin sylvestre de la forêt domaniale d'Orléans.

Les différents modèles que nous avons construits nous ont permis d'acquérir des connaissances sur la croissance et la dynamique de ces forêts. Nous avons montré que le modèle arbre indépendant des distances (MAID) était un outil efficace pour étudier de façon rétrospective la croissance du chêne et du pin en mélange. Nous avons vu que le modèle arbre dépendant des distances (MADD) permettait de préciser les phénomènes de compétition entre espèces au niveau de la croissance. Les modèles agrégés sont également apparus intéressants pour étudier le comportement du mélange à l'échelle du peuplement. D'un point de vue plus conceptuel, nous avons montré que l'on pouvait construire un modèle intermédiaire entre le MAID et le MADD en utilisant des distributions de voisinage. Ce modèle (MADV) pourrait être une alternative intéressante aux modèles spatialisés car il permet d'intégrer une partie de l'information sur la structure spatiale du peuplement ainsi que la variabilité des environnements locaux sans utiliser une cartographie du peuplement.

La cohérence que nous avons assurée entre les modèles nous a permis de les comparer par rapport à la prédiction de la croissance. Nous avons vu qu'il était important d'utiliser des données indépendantes pour évaluer la qualité prédictive des modèles. Nous avons par exemple montré que le MAID était plus performant que le MADD pour prédire la croissance des arbres sur le jeu de données indépendant. Nous avons vu que des exemples d'applications mettant en œuvre des simulations à court ou moyen terme étaient également un moyen efficace de comparer et d'évaluer les modèles.

La démarche proposée présente un intérêt aussi bien pour la compréhension du phénomène que pour sa modélisation dans un but prédictif. En regroupant l'ensemble des résultats acquis, ce travail nous a permis d'apprécier la pertinence d'un type de modèle en fonction du contexte d'utilisation. Cette démarche est très générale et pourrait être appliquée à la modélisation d'autres processus comme la mortalité ou la régénération.

## Bibliographie

- Aguirre, O., G. Hui, K. v. Gadow et J. Jiménez (2003). An analysis of spatial forest structure using neighbourhood-based variables. *Forest Ecology and Management* 183(1-3): 137-145.
- Alder, D. et J. N. M. Silva (2000). An empirical cohort model for management of Terra Firme forests in the Brazilian Amazon. *Forest Ecology and Management* 130(1-3): 141-157.
- Allain, R., Y. Dumas, A. Gama, C. Ginisty, F. Goreaud, F. Milano, M. A. Ngo Bieng, J. N. Piroche, T. Perot et S. Perret (2005). Projet scientifique de l'équipe "Forêts Hétérogènes : croissance, dynamique et diversité des structures". Etude et modélisation des peuplements mélangés Chêne - Pin sylvestre de la région Centre. Cemagref, Nogent-sur-Vernisson. 23p.
- Andreassen, K. et S. M. Tomter (2003). Basal area growth models for individual trees of Norway spruce, Scots pine, birch and other broadleaves in Norway. *Forest Ecology and Management* 180(1-3): 11-24.
- Antona, M., P. D'Aquino, S. Aubert, O. Barreteau, S. Boisseau, F. Bousquet, W. Daré, M. Etienne, C. Le Page, R. Mathevet, G. Trébuil et J. Weber (2005). La modélisation comme outil d'accompagnement. *Natures Sciences Sociétés* 13: 165-168.
- Astrup, R., K. D. Coates et E. Hall (2008). Finding the appropriate level of complexity for a simulation model: An example with a forest growth model. 6th North American Forest Ecology Workshop: From science to sustainability. *Forest Ecology and Management* 256(10): 1659-1665.
- Augustaitis, A. (2007). Pine sawfly (*Diprion pini* L.) - Related changes in Scots pine crown defoliation and possibilities of recovery. *Polish Journal of Environmental Studies* 16(3): 363-369.
- Barbier, S. (2007). Influence de la diversité, de la composition et de l'abondance des essences forestières sur la diversité floristique des forêts tempérées. Thèse d'écologie. Université d'Orléans, Cemagref, Nogent-sur-Vernisson. 273 p.
- Barreteau, O. (2007). Modèles et processus de décision collective : entre compréhension et facilitation de la gestion concertée de la ressource en eau. Rapport d'Habilitation à Diriger les Recherches en informatique. Université Paris Dauphine, Cemagref. 84 p.
- Barreteau, O. et J. C. Castellás (2006). Atelier "Modélisation interactive et usage interactif de modèles". Montpellier, Cirad.
- Bartelink, H. H. (2000). A growth model for mixed forest stands. *Forest Ecology and Management* 134(1-3): 29-43.
- Begon, M., J. L. Harper et C. R. Townsend (1996). *Ecology: individuals, populations and communities*, Blackwell Publishing. 1068 p.
- Bi, H. Q. et N. D. Turvey (1996). Competition in mixed stands of *Pinus radiata* and *Eucalyptus obliqua*. *Journal of Applied Ecology* 33(1): 87-99.
- Biging, G. S. et M. Dobbertin (1992). A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *Forest Science* 38(3): 695-720.

- Biging, G. S. et M. Dobbertin (1995). Evaluation of Competition Indexes in Individual Tree Growth-Models. *Forest Science* 41(2): 360-377.
- Biondi, F., J. O. Klemmedson et R. O. Kuehl (1992). Dendrochronological analysis of single-tree interactions in mixed pine oak stands of central Arizona, USA. *Forest Ecology and Management* 48(3-4): 321-333.
- Bonneuil, C. (2005). Les transformations des rapports entre sciences et société en France depuis la Seconde Guerre mondiale : un essai de synthèse. Actes du colloque Sciences, médias et société. 15-17 juin 2004. Lyon, ENS-LSH.
- Bontemps, J. D. (2002). Analyse de la croissance radiale dans les hêtraies régulières du nord-est de la France. Evolution de la vitesse de croissance au cours du XXème siècle. Rapport de DEA Biologie Forestière. LERFOB INRA-ENGREF, Nancy. 59 p.
- Bontemps, J. D. (2006). Evolution de la productivité des peuplements réguliers et monospécifiques de hêtre (*Fagus sylvatica* L.) et de chêne sessile (*Quercus petraea* Liebl.) dans la moitié Nord de la France au cours du XXème siècle. *Sciences Forestières*. ENGREF, Nancy. 405 p.
- Bousquet, F. (2001). Modélisation d'accompagnement, simulation multi-agents et gestion des ressources naturelles renouvelables. Mémoire pour l'obtention de l'Habilitation à Diriger les Recherches de l'Université de Lyon I. Cirad, Montpellier. 70 p.
- Brethes, A. (1993). Les types de station forestière de l'Orléanais. Paris, Office National des Forêts. 400 p.
- Bristow, M., J. D. Nichols et J. K. Vanclay (2006a). Improving productivity in mixed-species plantations - Preface. *Forest Ecology and Management* 233(2-3): 193-194.
- Bristow, M., J. K. Vanclay, L. Brooks et M. Hunt (2006b). Growth and species interactions of *Eucalyptus pellita* in a mixed and monoculture plantation in the humid tropics of north Queensland. *Forest Ecology and Management* 233(2-3): 285-294.
- Bruciamacchie, M., C. Groualle et P. Minot (1991). Evolution Model in Populations of Unevenly-Aged Stands. *Annales Des Sciences Forestières* 48(2): 215-224.
- Buongiorno, J., J. L. Peyron, F. Houllier et M. Bruciamacchie (1995). Growth and Management of Mixed-Species, Uneven-Aged Forests in the French Jura - Implications for Economic Returns and Tree Diversity. *Forest Science* 41(3): 397-429.
- Burkhardt, H. E. (2003). Suggestions for choosing an appropriate level for modelling forest stands. Modelling forest systems Workshop on the interface between reality, modelling and the parameter estimation processes, Sesimbra, Portugal, 2-5 June 2002., Wallingford UK: CABI Publishing.
- Burton, P. J. (1993). Some limitations inherent to static indices of plant competition. *Canadian Journal of Forest Research* 23(10): 2141-2152.
- Cacot, E., N. Eisner, F. Charnet, P. Leon, C. Rantien et J. Ranger (2006). La récolte raisonnée des rémanents en forêt. ADEME, Angers. 36 p.
- Campolongo, F. et R. Braddock (1999). The use of graph theory in the sensitivity analysis of the model output: a second order screening method. *Reliability Engineering & System Safety* 64(1): 1-12.
- Canham, C. D., P. T. LePage et K. D. Coates (2004). A neighborhood analysis of canopy tree competition: effects of shading versus crowding. *Canadian Journal of Forest Research* 34(4): 778-787.

- Canham, C. D., M. J. Papaik, M. Uriarte, W. H. McWilliams, J. C. Jenkins et M. J. Twery (2006). Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. *Ecological Applications* 16(2): 540-554.
- Cariboni, J., D. Gatelli, R. Liska et A. Saltelli (2007). The role of sensitivity analysis in ecological modelling. *Ecological Modelling* 203(1-2): 167-182.
- Casteyde, C. (2000). Cours de C/C++. 501 p.
- Chabaud, L. et L. Nicolas (Sous presse). Guide de sylvicultures pour les pineraies des plaines du Centre et du Nord-Ouest, Office National des Forêts.
- Coligny, F. d., P. Ancelin, G. Cornu, B. Courbaud, P. Dreyfus, F. Goreaud, S. Gourlet-Fleury, C. Meredieu et L. Saint-André (2003). CAPSIS : Computer-Aided Projection for Strategies In Silviculture : Advantages of a shared forest-modelling platform. International Workshop of IUFRO working party 4.01 "Reality, models and parameter estimation". 2-5 june. Sesimbra, Portugal, June 2-5, 2002, (eds) Modelling Forest Systems. CABI Publishing, Wallingford, UK.
- Comas, C. et J. Mateu (2007). Modelling forest dynamics: A perspective from point process methods. *Biometrical Journal* 49(2): 176-196.
- Courbaud, B. (2004). Intérêt des modèles pour l'ingénierie écologique : exemples à partir du modèle de dynamique des peuplements forestiers SAMSARA. *Ingénieries n° spécial Ingénierie écologique*: 49-56.
- Courbaud, B., F. d. Coligny et T. Cordonnier (2003). Simulating radiation distribution in a heterogeneous Norway spruce forest on a slope. *Agricultural and Forest Meteorology* 116(1-2): 1-18.
- Courbaud, B., F. Goreaud, P. Dreyfus et F. R. Bonnet (2001). Evaluating thinning strategies using a tree distance dependent growth model: some examples based on the CAPSIS software "uneven-aged spruce forests" module. *Forest Ecology and Management* 145(1-2): 15-28.
- Cruickshank, M. G., D. J. Morrison et A. Lalumière (2009). The interaction between competition in interior Douglas-fir plantations and disease caused by *Armillaria ostoyae* in British Columbia. *Forest Ecology and Management* 257(2): 443-452.
- Da Silva, D. (2008). Caractérisation de la nature multi-échelle des plantes par des outils de géométrie fractale, influence sur l'interception de la lumière. Thèse de doctorat en Informatique. Université de Montpellier II. 194.
- Daquitaine, R. (2003). Evaluation des chaînes de modèles de production ligneuse. Prédiction de la qualité des produits d'une ressource forestière sur pied en Epicéa commun (*Picea abies* Karst.). Thèse en Sciences forestières. ENGREF.
- Deleuze, C. (2000). Modélisation de la croissance individuelle en circonférence pour des peuplements équiennes monospécifiques d'Epicéa commun. Rapport final. Convention. DERF / AFOCEL. 83 p.
- Deleuze, C., O. Pain, J. F. Dhote et J. C. Herve (2004). A flexible radial increment model for individual trees in pure even-aged stands. *Annals of Forest Science* 61(4): 327-335.
- Deutschman, D. H., S. A. Levin et S. W. Pacala (1999). Error propagation in a forest succession model: The role of fine-scale heterogeneity in light. *Ecology* 80(6): 1927-1943.

- Dhôte, J. F. (1991). Modeling the Growth of Even-Aged Beech Stands - Dynamics of Hierarchical Systems and Yield Factors. *Annales Des Sciences Forestieres* 48(4): 389-416.
- Dhôte, J. F. (1999). Compétition entre classes sociales chez le chêne sessile et le hêtre. *Revue Forestière Française* LI(2): 309-325.
- Dhôte, J. F. (2008). Modèles phénoménologiques et changements environnementaux. *LERFoB Infos* 83.
- Dhôte, J. F., T. Cordonnier, P. Dreyfus et N. Le Goff (2005). Quelques enjeux autour des forêts hétérogènes tempérées. *RenDez-Vous techniques de l'ONF*. 10: 22-31.
- Dhôte, J. F., E. Hatsch et D. Rittie (2000). Forme de la tige, tarifs de cubage et ventilation de la production en volume chez le Chêne sessile. *Annals of Forest Science* 57(2): 121-142.
- Dhôte, J. F., P. Vallet, C. Deleuze, L. Saint-Andre et J. C. Herve (2007). Quelques réflexions sur les modèles : stratégies de conception, formes mathématiques, hypothèses théoriques, généralisation. *CAQ10*. Bordeaux.
- Dieckmann, R. et R. Law (2000). Relaxation projections and the method of moments. *The geometry of ecological interactions: Simplifying Spatial Complexity*, Cambridge University Press: 412-455.
- Dieckmann, U., R. Law et J. A. J. Metz (2000). *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*, Cambridge University Press. 579.
- Donès, N., P. Balandier, N. Gaudio et F. d. Coligny (2009). RReshar (Regeneration and Resource Sharing), un modèle pour la régénération et le partage des ressources sous Capsis. 12ème réunion Croissance Amélioration Qualité des bois. Nancy.
- Dreyfus, P. (2008). Dynamique du sapin, du hêtre et des pins dans l'arrière-pays méditerranéen : de la modélisation à l'aide à la gestion. *Revue Forestiere Francaise* LX(2): 233-249.
- Duplat, P. et G. Perrote (1981). Inventaire et estimation de l'accroissement des peuplements forestiers. Fontainebleau, ONF. 432 p.
- Duplat, P. et M. Tran-Ha (1986). Modèles de croissance en hauteur dominante pour le hêtre, le sapin pectiné, et le pin sylvestre (dans le massif de l'Aigoual). *Section technique - ONF*. 86: 34 p.
- Edwards, M. (2004). Intérêt d'un modèle agrégé pour étudier le comportement et simplifier la simulation d'un modèle individu-centré de consommation couplé à un modèle de ressource en eau. Thèse en Informatique. Université Paris VI Pierre et Marie Curie. Cemagref, Paris. 281p.
- Edwards, M., N. Ferrand, F. Goreaud et S. Huet (2005). The relevance of aggregating a water consumption model cannot be disconnected from the choice of information available on the resource. *Simulation Modelling Practice and Theory* 13(4): 287-307.
- Ek, A. R. et R. A. Monserud (1979). Performance and comparison of stand growth models based on individual tree and diameter-class growth. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 9: 23-244.
- ENGREF, INRA et ONF (1984). Tables de production pour les forêts françaises. - 2ème édition revue par B. Vannière. Nancy, Ecole Nationale du Génie Rurale, des Eaux et des Forêts. 160 p.

- FAO (2007). Situation des forêts du monde. Rome, Italie, FAO. 143 p.
- Favrillon, V. (1995). Modèle matriciel déterministe en temps discret - Application à l'étude de la dynamique d'un peuplement forestier tropical humide (Guyane française). LYON. Université Claude-Bernard, Lyon I, Lyon. 252 p.
- Forrester, D. I., J. Bauhus, A. L. Cowie et J. K. Vanclay (2006). Mixed-species plantations of Eucalyptus with nitrogen-fixing trees: A review. *Forest Ecology and Management* 233(2-3): 211-230.
- Fox, J. C., P. K. Ades et H. Q. Bi (2001). Stochastic structure and individual-tree growth models. *Forest Ecology and Management* 154(1-2): 261-276.
- Garcia-Abril, A., S. Martin-Fernandez, M. A. Grande et J. A. Manzanera (2007). Stand structure, competition and growth of Scots pine (*Pinus sylvestris* L.) in a Mediterranean mountainous environment. *Annals of Forest Science* 64(8): 825-830.
- Gaudio, N., E. Lancelle, S. Perret, P. Balandier, Y. Dumas, F. Vast, M. Bonin et C. Ginisty (2008a). Croissance et morphologie de pins sylvestres en réponse à la lumière sous couvert forestier. "La plasticité phénotypique des arbres", conférence du GEA (Groupe d'Etude de l'Arbre). Montpellier.
- Gaudio, N., S. Perret, P. Balandier, Y. Dumas, G. Philippe, M. Bonin et C. Ginisty (2008b). Interactions pour la lumière et l'eau entre les différentes composantes de l'écosystème forestier en peuplement mélangé : conséquences sur la régénération et modélisation. Application à la régénération du pin sylvestre en peuplement mélangé chêne - pin. Rapport d'avancement - Convention Cemagref / ONF 2008 - Action 1.5 Etude des conditions de régénération. Cemagref, Nogent-sur-Vernisson. 23 p.
- Gauquelin, X., B. Courbaud, J. Fay, F. Berger et E. Mermin (2008). Conduite de peuplements mélangés en forêt de montagne. Exemple d'un transfert chercheurs-gestionnaires. *Revue Forestière Française* LX: 207-214.
- Geri, C. et F. Goussard (1984). Evolution d'une nouvelle gradation de lophyre du pin (*Diprion pini* L.) dans le sud du Bassin Parisien. I. - Développement de la gradation jusqu'en 1982 et relation avec les facteurs du milieu. *Annales Des Sciences Forestières* 41(4): 376-403.
- Gobakken, T. et E. Naesset (2002). Spruce diameter growth in young mixed stands of Norway spruce (*Picea abies* (L.) Karst.) and birch (*Betula pendula* Roth B-pubescons Ehrh.). *Forest Ecology and Management* 171(3): 297-308.
- Goreaud, F. (2000). Apports de l'analyse de la structure spatiale en forêt tempérée à l'étude et la modélisation des peuplements complexes. Thèse en sciences forestières. ENGREF. 362p.
- Goreaud, F., I. Alvarez, B. Courbaud et F. de Coligny (2006a). Long-term influence of the spatial structure of an initial state on the dynamics of a forest growth model: A simulation study using the capsis platform. *Simulation-Transactions of the Society for Modeling and Simulation International* 82(7): 475-495.
- Goreaud, F., F. d. Coligny, B. Courbaud, J. F. Dhote, P. Dreyfus et T. Perot (2005a). La modélisation : un outil pour la gestion et l'aménagement en forêt. *VertigO* 6(2): 1-12.
- Goreaud, F., B. Courbaud et F. d. Coligny (2005b). How long does the spatial structure of an initial state influence the dynamics of a forest growth model ? A simulation study using the Capsis platform. Open International Conference on Modeling and Simulation - OICMS. 13-15 june 2005. Clermont Ferrand (France).

- Goreaud, F., G. Deceliere, C. Ginisty, F. Milano, M. A. Ngo Bieng, S. Perret et T. Perot (2006b). Comparaison de plusieurs stratégies d'échantillonnage des carottes de croissance pour l'ajustement de modèles de croissance des peuplements mélangés Chêne - Pin Sylvestre de la forêt d'Orléans. Cemagref. 34 p.
- Goreaud, F., M. Loreau et C. Millier (2002). Spatial structure and the survival of an inferior competitor: a theoretical model of neighbourhood competition in plants. *Ecological Modelling* 158(1-2): 1-19.
- Goreaud, F., B. Loussier, M. A. Ngo Bieng et R. Allain (2004). Simulating realistic spatial structure for forest stands: a mimetic point process. *Interdisciplinary Spatial Statistics Workshop*. Paris.
- Goreaud, F. et M. A. Ngo Bieng (2005). PASS : Programme d'Analyse de la Structure Spatiale. Clermont-Ferrand, Cemagref.
- Goreaud, F. et R. Pelissier (1999). On explicit formulas of edge effect correction for Ripley's K-function. *Journal of Vegetation Science* 10(3): 433-438.
- Gourlet-Fleury, S. (1997). Modélisation individuelle spatialement explicite de la dynamique d'un peuplement de forêt dense tropicale humide (dispositif de Paracou - Guyane française). Thèse de doctorat en sciences forestières. Université Claude Bernard Lyon I / CIRAD. 274 p.
- Gourlet-Fleury, S., G. Comu, S. Jesel, H. Dessard, J. G. Jourget, L. Blanc et N. Picard (2005). Using models to predict recovery and assess tree species vulnerability in logged tropical forests: A case study from French Guiana. *Forest Ecology and Management* 209(1-2): 69-86.
- Gourlet-Fleury, S. et F. Houllier (2000). Modelling diameter increment in a lowland evergreen rain forest in French Guiana. *Forest Ecology and Management* 131(1-3): 269-289.
- Grimm, V. (1999). Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling* 115(2-3): 129-148.
- Grimm, V. et S. F. Railsback (2004). *Individual-based modelling and ecology*. Princeton and Oxford, Princeton university press. 413 p.
- Grissino-Mayer, H. D. (2002). Research report evaluating crossdating accuracy: a manual and tutorial for the computer program COFECHA. *Tree-Ring Research* 57(2): 205-221.
- Guirado, M., J. Pino, F. Roda et C. Basnou (2008). Quercus and Pinus cover are determined by landscape structure and dynamics in peri-urban Mediterranean forest patches. *Plant Ecology* 194(1): 109-119.
- Hao, Q. Y., F. R. Meng, Y. P. Zhou et J. X. Wang (2005). A transition matrix growth model for uneven-aged mixed-species forests in the Changbai Mountains, northeastern China. *New Forests* 29(3): 221-231.
- Hasenauer, H. (1997). Dimensional relationships of open-grown trees in Austria. *Forest Ecology and Management* 96(3): 197-206.
- Hasenauer, H. (2006). Concepts within tree growth modelling. *Sustainable Forest Management: Growth Models for Europe*, Springer: 3-17.
- Heath, R. et R. I. Alfaro (1990). Growth response in a Douglas-fir/lodgepole pine stand after thinning of lodgepole pine by the mountain pine beetle: a case study. *Journal of the Entomological Society of British Columbia*(No. 87): 16-21.

- Hein, S. et J. F. Dhote (2006). Effect of species composition, stand density and site index on the basal area increment of Oak trees (*Quercus spec.*) in mixed stands with Beech (*Fagus sylvatica L.*) in northern France. *Annals of Forest Science* 63(5): 457-467.
- Houllier, F., J. Bouchon et Y. Birot (1991). Modélisation de la dynamique des peuplements forestiers : état et perspectives. *Revue Forestière Française* XLIII(2): 87-108.
- Huang, S., Y. Yang et Y. Wang (2002). A critical look at procedures for validating growth and yield models. Reality, models and parameter estimation - the forestry scenario. June 2-5, 2002. Sesimbra, Portugal.
- Iooss, B. (2006). Analyses de sensibilité globales de modèles complexes, quelques applications dans le nucléaire. Planification d'expériences et analyse d'incertitudes pour les gros codes numériques : Approches Stochastiques. 2 et 3 février 2006. Toulouse, 2 et 3 février 2006.
- Iwasa, Y., V. Andreassen et S. Levin (1987). Aggregation in model ecosystems. I. Perfect aggregation. *Ecological Modelling* 37(3-4): 287-302.
- Jackson, E. T. et Y. Kassam (2000). Entretien avec Edward T. Jackson et Yusuf Kassam auteurs de l'ouvrage "Knowledge Shared: Participatory Evaluation in Development Cooperation". Canada, CRDI.
- Jacques, J. (2005). Contribution à l'analyse de sensibilité et à l'analyse discriminante généralisée. Thèse en mathématiques appliquées. Université Joseph Fourier - Grenoble 1. 207 p.
- Jactel, H., E. Brockerhoff et D. Piou (2008). Le risque sanitaire dans les forêts mélangées. *Revue Forestière Française* LX(2): 168-180.
- Jactel, H. et E. G. Brockerhoff (2007). Tree diversity reduces herbivory by forest insects. *Ecology Letters* 10(9): 835-848.
- Jarret, P. (2004). Guide des sylvicultures de la chênaie atlantique, ONF. 335 p.
- Jogiste, K. (1998). Productivity of mixed stands of Norway spruce and birch affected by population dynamics: a model analysis. *Ecological Modelling* 106(1): 77-91.
- Jogiste, K. (2000). A basal area increment model for Norway spruce in mixed stands in Estonia. *Scandinavian Journal of Forest Research* 15(1): 97-102.
- Jones, H. E., N. McNamara et W. L. Mason (2005). Functioning of mixed-species stands: Evidence from a long-term forest experiment. *Forest Diversity and Function: Temperate and Boreal Systems* 176: 111-130.
- Jonsell, M. (2007). Effects on biodiversity of forest fuel extraction, governed by processes working on a large scale. *Biomass & Bioenergy* 31(10): 726-732.
- Kelty, M. J. (2006). The role of species mixtures in plantation forestry. Improving Productivity in Mixed-Species Plantations. *Forest Ecology and Management* 233(2-3): 195-204.
- Kobe, R. K. et K. D. Coates (1997). Models of sapling mortality as a function of growth to characterize interspecific variation in shade tolerance of eight tree species of northwestern British Columbia. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestière* 27(2): 227-236.

- Kokkila, T., A. Makela et A. Franc (2006). Comparison of distance-dependent and distance-independent stand growth models - Is perfect aggregation possible? *Forest Science* 52(6): 623-635.
- Kokkila, T., A. Makela et E. Nikinmaa (2002). A method for generating stand structures using Gibbs marked point process. *Silva Fennica* 36(1): 265-277.
- Lancelot, R. et M. Lesnoff (2005). Sélection de modèles avec l'AIC et critères d'information dérivés. CIRAD, Montpellier. 7 p.
- Langstrom, B., E. Annala, C. Hellqvist, M. Varama et P. Niemela (2001). Tree mortality, needle biomass recovery and growth losses in Scots pine following defoliation by *Diprion pini* (L.) and subsequent attack by *Tomicus piniperda* (L.). *Scandinavian Journal of Forest Research* 16(4): 342-353.
- Law, R. et R. Dieckmann (2000). Moment approximations of individual-based models. The geometry of ecological interactions: Simplifying Spatial Complexity, Cambridge University Press: 252-270.
- Lebourgeois, F. (1997). RENECOFOR - Etude dendrochronologique des 102 peuplements du réseau., Office National des Forêts, Département des Recherches Techniques. 307 p.
- Lebourgeois, F. (2006). Sensibilité au climat des chênes sessile et pédonculé dans le réseau RENECOFOR. Comparaison avec les hêtraies. *Revue Forestiere Francaise* LVIII(1): 29-44.
- Lebourgeois, F., F. Spicher et Y. Lefèvre (2008). Relations croissance du Chêne pédonculé et climat sur deux types de sol à nappe temporaire en lorraine (rédoxisol acide et pélosol différencié). *Revue Forestiere Francaise* 60(4): 411-424.
- Lee, W. K., K. von Gadow, D. J. Chung, J. L. Lee et M. Y. Shin (2004). DBH growth model for *Pinus densiflora* and *Quercus variabilis* mixed forests in central Korea. *Ecological Modelling* 176(1-2): 187-200.
- Legay, M. (2008). Des forêts mélangées pour composer avec les changements climatiques. *Revue Forestiere Francaise* LX(2): 181-190.
- Lejeune, P. (1996). Development of a tree model applied to the growth of irregular hardwood stands in southern Belgium. *Canadian Journal of Forest Research* 26(10): 1838-1848.
- Lett, C., C. Silber et N. Barret (1999). Comparison of a cellular automata network and an individual-based model for the simulation of forest dynamics. *Ecological Modelling* 121(2-3): 277-293.
- Lorimer, C. G. (1983). Tests of age-independent competition indices for individual trees in natural hardwood stands. *Forest Ecology and Management* 6(4): 343-360.
- Lotwick, H. W. et B. W. Silverman (1982). Methods for analysing spatial processes of several types of points. *Journal of the Royal Statistical Society B* 44(3): 406-413.
- Lyytikäinen-Saarenmaa, P. et E. Tomppo (2002). Impact of sawfly defoliation on growth of Scots pine *Pinus sylvestris* (Pinaceae) and associated economic losses. *Bulletin of Entomological Research* 92(2): 137-140.
- Manson, D. G., J. Harlan, M. Hunt, M. Bristow, P. D. Erskine, D. Lamb et S. Schmidt (2006). Modelling predicts positive and negative interactions between three Australian tropical tree species in monoculture and binary mixture. *Forest Ecology and Management* 233(2-3): 315-323.

- Martinez, J. M., S. Gazut, T. Crestaux et O. Le Maître (2006). Analyse de sensibilité par polynômes de chaos et planification active. Planification d'expériences et analyse d'incertitudes pour les gros codes numériques : Approches Stochastiques. Toulouse, 2 et 3 février 2006.
- Matala, J., J. Hynynen, J. Miina, R. Ojansuu, H. Peltola, R. Sievanen, H. Vaisanen et S. Kellomaki (2003). Comparison of a physiological model and a statistical model for prediction of growth and yield in boreal forests. *Ecological Modelling* 161(1-2): 95-116.
- Mayer, D. G. et D. G. Butler (1993). Statistical validation. *Ecological Modelling* 68(1-2): 21-32.
- MCPFE (2003). Improved pan-European indicators for sustainable forest management as adopted by the MCPFE Expert Level Meeting. Ministerial Conference on the Protection of Forests in Europe, Vienne. 6 p.
- Meredieu, C. (1998). Croissance et branchaison du Pin laricio (*Pinus nigra* Arn. ssp. *laricio* (Poiret) Maire) : élaboration et évaluation d'un système de modèles pour la prévision de caractéristiques des arbres et du bois. Thèse. Claude Bernard Lyon I. 250 p.
- Monserud, R. A. et H. Sterba (1996). A basal area increment model for individual trees growing in even- and uneven-aged forest stands in Austria. *Forest Ecology and Management* 80(1-3): 57-80.
- Morneau, F., C. Duprez et J. C. Hervé (2008). Les forêts mélangées en France métropolitaine. Caractérisation à partir des résultats de l'Inventaire Forestier National. *Revue Forestiere Francaise LX(2)*: 107-120.
- Munro, D. D. (1974). Forest growth models - a prognosis. Growth models for tree and stand simulation. Proceedings of the IUFRO congress S4-01-4. Stockholm, Department of forest yield research, Royal College of Forestry.
- Nabuurs, G. J., A. Pussinen, J. van Brusselen et M. J. Schelhaas (2007). Future harvesting pressure on European forests. *European Journal of Forest Research* 126(3): 391-400.
- Ngo Bieng, M. A. (2007). Construction de modèles de structure spatiale permettant de simuler des peuplements virtuels réalistes. Application aux peuplements mélangés Chêne sessile - Pin sylvestre de la région Centre. Thèse en sciences forestières. ENGREF-Cemagref, Nogent-sur-Vernisson. 183 p.
- Ngo Bieng, M. A., C. Ginisty, F. Goreaud et T. Perot (2006). A first typology of Oak and Scots pine mixed stands in the Orleans forest (France), based on the canopy spatial structure. *New Zealand Journal of Forestry Science* 36(2): 325-346.
- Nichols, J. D., M. Bristow et J. K. Vanclay (2006). Mixed-species plantations: Prospects and challenges. *Forest Ecology and Management* 233(2-3): 383-390.
- Oliver, C. D. et B. C. Larson (1996). Forest stand dynamics. New York, Etats Unis, John Wiley and sons. 520 p.
- ONF (2004). Aménagement forestier. Forêt domaniale d'Orléans. Massif de Lorris - Les Bordes. Office National des Forêts, Région Centre - Département du Loiret.
- ONU (1992). Sommet de la terre. Conférence des Nations unies sur l'environnement et le développement. 3 au 14 juin 1992. Brésil, Rio de Janeiro.
- ONU (1993). Convention sur la diversité biologique (avec annexes). Conclue a Rio de Janeiro le 5 juin 1992. Vol. 1760(I-30619).

- Orois, S. S. et R. R. Soalleiro (2002). Modelling the growth and management of mixed uneven-aged maritime pine-broadleaved species forests in Galicia, north-western Spain. *Scandinavian Journal of Forest Research* 17(6): 538-547.
- Ottorini, J. M. et N. Le Goff (2002). Modélisation et simulation de la croissance individuelle en peuplement du hêtre et du frêne. Convention GIP-ECOFOR / INRA. Rapport de convention. 30p.
- Palagos, B. (1997). Quelques estimateurs de l'erreur de prédiction. Document interne. Cemagref, Clermont-Ferrand. 6 p.
- Palahi, M., M. Tome, T. Pukkala, A. Trasobares et G. Montero (2004). Site index model for *Pinus sylvestris* in north-east Spain. *Forest Ecology and Management* 187(1): 35-47.
- Pauwels, D. (2003). Conception d'un système d'aide à la décision pour le choix d'un scénario sylvicole : Application aux peuplements de mélèze en Région wallonne. Thèse en Sciences agronomiques et ingénierie biologique. Faculté Universitaire des Sciences Agronomiques de Gembloux, Gembloux. 262 p.
- Pauwels, D., P. Lejeune, P. Balandier et J. Rondeux (2002). A silvicultural decision support system for pure larch stand : helping managers to meet their goals. Improvement of larch (*Larix sp.*) for better growth, stem form and wood quality. 16-21 september. Gap, France.
- Pauwels, D., P. Lejeune et J. Rondeux (2007). A decision support system to simulate and compare silvicultural scenarios for pure even-aged larch stands. *Annals of Forest Science* 64(3): 345-353.
- Pavé, A. (1994). Modélisation en biologie et en écologie. Lyon, Aléas éditeur. 559 p.
- Payandeh, B. et J. Wang (1996). Variable stocking version Plonski's yield tables formulated. *The Forestry Chronicle* 72(2): 181-184.
- Perot, T. et C. Ginisty (2004). Bilan et perspectives sur les modèles de croissance, de dynamique forestière et de qualité des bois. Rapport de convention Cemagref/DGFAR. Cemagref, Nogent-sur-Vernisson. 190p.
- Perot, T., C. Ginisty et F. Goreaud (2008). Quels modèles de croissance pour les peuplements mélangés ? Exemple du mélange chêne sessile - pin sylvestre. *Revue Forestiere Francaise* LX(2): 215-232.
- Perot, T., S. Perret, C. Meredieu et C. Ginisty (2007). Prévoir la croissance et la production du Pin sylvestre : le module Sylvestris sous Capsis 4. *Revue Forestiere Francaise* 59(1): 57-84.
- Phillips, P. D., T. E. Brash, I. Yasman, P. Subagyo et P. R. van Gardingen (2003). An individual-based spatially explicit tree growth model for forests in East Kalimantan (Indonesian Borneo). *Forest Ecology and Management* 159(1): 1-26.
- Phillips, P. D., I. S. Thompson, J. N. M. Silva, P. R. van Gardingen et B. Degen (2004). Scaling up models of tree competition for tropical forest population genetics simulation. *Ecological Modelling* 180(2-3): 419-434.
- Piboule, A., C. Collet, H. Frochot et J. F. Dhote (2005). Reconstructing crown shape from stem diameter and tree position to supply light models. I. Algorithms and comparison of light simulations. *Annals of Forest Science* 62(7): 645-657.

- Picard, N. (1999). Passage d'un modèle individuel à un modèle de distribution de la dynamique forestière. Application à une forêt dense tropicale humide de Guyane française. Thèse en sciences forestières. ENGREF - CIRAD, Paris. 457p.
- Picard, N. (2007). Modélisation de la dynamique de forêts naturelles tropicales : quel niveau de description. Mémoire de synthèse en vue d'une candidature à une Habilitation à Diriger des Recherches. Université de Montpellier II, Cirad, Montpellier. 99 p.
- Picard, N. et A. Franc (2001). Aggregation of an individual-based space-dependent model of forest dynamics into distribution-based and space-independent models. *Ecological Modelling* 145(1): 69-84.
- Picard, N. et A. Franc (2004). Approximating spatial interactions in a model of forest dynamics. *Forest Biometry, Modelling and Information Sciences* 1: 91-103.
- Pinheiro, J., D. Bates, S. DebRoy et D. Sarkar (2008). nlme: Linear and nonlinear mixed effects models. R package version 3.1-87.
- Pinheiro, J. C. et D. M. Bates (2000). *Mixed-effects models in S and S-PLUS*. New York, Springer. 528 p.
- Piotto, D. (2008). A meta-analysis comparing tree growth in monocultures and mixed plantations. *Forest Ecology and Management* 255(3-4): 781-786.
- Pommerening, A. (2002). Approaches to quantifying forest structures. *Forestry* 75(3): 305-324.
- Pommerening, A. (2006). Evaluating structural indices by reversing forest structural analysis. *Forest Ecology and Management* 224(3): 266-277.
- Porté, A. et H. H. Bartelink (2002). Modelling mixed forest growth: a review of models for forest management. *Ecological Modelling* 150(1-2): 141-188.
- Pretzsch, H. (1997). Analysis and modeling of spatial stand structures. Methodological considerations based on mixed beech-larch stands in Lower Saxony. *Forest Ecology and Management* 97(3): 237-253.
- Pretzsch, H. (2000). From yield tables to simulation models for pure and mixed stands. *Journal of Forest Science* 46(3): 97-113.
- Pretzsch, H. (2002). Application and evaluation of the growth simulator SILVA 2.2 for forest stands, forest estates and large regions. *Forstwissenschaftliches Centralblatt* 121: 28-51.
- Pretzsch, H. (2005). Diversity and productivity in forests: Evidence from long-term experimental plots. *Forest Diversity and Function: Temperate and Boreal Systems* 176: 41-64.
- Pretzsch, H., P. Biber et J. Dursky (2002a). The single tree-based stand simulator SILVA: construction, application and evaluation. *Forest Ecology and Management* 162(1): 3-21.
- Pretzsch, H., P. Biber, J. Dursky, K. von Gadow, H. Hasenauer, G. Kandler, G. Kenk, E. Kublin, J. Nagel, T. Pukkala, J. P. Skovsgaard, R. Sotke et H. Sterba (2002b). Recommendations for standardized documentation and further development of forest growth simulators. *Forstwissenschaftliches Centralblatt* 121(3): 138-151.
- Prevosto, B. (1999). Analyse fonctionnelle des boisements spontanés et modélisation de leur développement. Application aux peuplements de *Betula pendula* (Roth.) et *Pinus sylvestris* (L.) en moyenne montagne volcanique (la Chaîne des Puys, Massif Central). Thèse de doctorat en écologie. Université d' Aix-Marseille III. 197p.

- Prévosto, B. (2005). Les indices de compétition en foresterie : exemples d'utilisation, intérêts et limites. *Revue Forestiere Francaise* LVII(5): 413-430.
- Puettmann, K. J., D. E. Hibbs et D. W. Hann (1992). The dynamics of mixed stands of *Alnus rubra* and *pseudotsuga menziesii*: extension of size-density analysis to species mixture. *Journal of Ecology* 80(3): 449-458.
- Pukkala, T. (1989). Methods to describe the competition process in a tree stand. *Scandinavian Journal of Forest Research* 4: 187-202.
- Pukkala, T., J. Miina, M. Kurttila et T. Kolström (1998). A spatial yield model for optimizing the thinning regime of mixed stands of *Pinus sylvestris* and *Picea abies*. *Scandinavian Journal of Forest Research* 13(1): 31-42.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna, Austria, R Foundation for Statistical Computing.
- Rameau, J. C., D. Mansion et G. Dume (1989). *Flore forestière française : guide écologique illustré* : t. 1 Plaines et collines. Paris, IDF. 1785 p.
- Regent, I. (2005). *Windendro 2005a: manual for tree-ring analysis*, Université du Quebec à Chicoutimi. 132 p.
- Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the royal statistical society B* 39: 172-212.
- Ripley, B. D. (2004). Notes and references on robustness. Master of Science in applied statistics. University of Oxford. Department of statistics: 11 p.
- Rivoire, M. (2009). Modélisation générique de la loi d'auto-éclaircie pour des peuplements mélangés. 12ème réunion Croissance Amélioration Qualité des bois. 17 mars 2009. Nancy.
- Saltelli, A. (2006). Sensitivity analysis. Planification d'expériences et analyse d'incertitudes pour les gros codes numériques : Approches Stochastiques. Toulouse, 2 et 3 février 2006.
- Saltelli, A., M. Ratto, S. Tarantola et F. Campolongo (2005). Sensitivity analysis for chemical models. *Chemical Reviews* 105(7): 2811-2827.
- Saltelli, A., M. Ratto, S. Tarantola et F. Campolongo (2006). Sensitivity analysis practices: Strategies for model-based inference. *Reliability Engineering & System Safety* 91(10-11): 1109-1125.
- Sanchez-Gonzalez, M., M. del Rio, I. Canellas et G. Montero (2006). Distance independent tree diameter growth model for cork oak stands. *Forest Ecology and Management* 225(1-3): 262-270.
- Sardin, T., J. Bock et J. Becquey (2008). Les peuplements mélangés : enjeux et interrogations des gestionnaires. *Revue Forestiere Francaise* 60(2): 121-128.
- Saucier, J. (2001). L'effet de l'épidémie de la tordeuse des bourgeons de l'épinette sur le rendement des forêts. Tordeuse des bourgeons de l'épinette : l'appriivoiser dans nos stratégies d'aménagement. 27, 28, et 29 mars 2001. Shawinigan, Québec.
- Schmidt, M., J. Nagel et J. P. Skovsgaard (2006). Evaluating individual tree growth models. *Sustainable forest management: growth models for Europe*. Springer: 151-163.
- Sevrin, E., G. Landmann et T. Sardin (2008). Enjeux et scenarios pour des peuplements mélangés. *Revue Forestiere Francaise* LX(2): 251-255.

- Smith, E. P. et K. A. Rose (1995). Model Goodness-of-Fit Analysis Using Regression and Related Techniques. *Ecological Modelling* 77(1): 49-64.
- Soares, P., M. Tome, J. P. Skovsgaard et J. K. Vanclay (1995). Evaluating a Growth-Model For Forest Management Using Continuous Forest Inventory Data. *Forest Ecology and Management* 71(3): 251-265.
- Sonohat, G., P. Balandier et F. Ruchaud (2004). Predicting solar radiation transmittance in the understory of even-aged coniferous stands in temperate forests. *Annals of Forest Science* 61(7): 629-641.
- Sprent, P. (1992). *Pratique des statistiques non paramétriques*. Paris, INRA éditions. 294 p.
- Stadt, K. J., C. Huston, K. D. Coates, Z. Feng, M. R. T. Dale et V. J. Lieffers (2007). Evaluation of competition and light estimation indices for predicting diameter growth in mature boreal mixed forests. *Annals of Forest Science* 64: 477-490.
- Stadt, K. J., C. Huston et V. J. Lieffers (2002). A comparison of non-spatial and spatial, empirical and resource-based competition indices for predicting the diameter growth of trees in maturing boreal mixedwood stands. Final project report. SFM project: Spatial and non-spatial modeling of canopy tree dynamics in boreal forests. University of Alberta, Edmonton, Canada. 32 p.
- Steyerberg, E. W. (2009). *Clinical Prediction Models. A practical approach to development, validation, and updating*. New York, Springer. 500 p.
- Stoyan, D. et A. Penttinen (2000). Recent applications of point process methods in forestry statistics. *Statistical Science* 15(1): 61-78.
- Tham, A. (1989). Prediction of individual tree growth in managed stands of mixed picea abies (L.) Karst. and Betula pendula Roth & Betula pubescens. *Scandinavian Journal of Forest Research* 4: 491-512.
- Thornton, P. K. et J. W. Hansen (1996). A note on regressing real-world data on model output. *Agricultural Systems* 50(4): 411-414.
- Tomassone, R., C. Dervin et J. P. Masson (1993). *Biométrie : Modélisation de phénomènes biologiques*. Paris, Masson. 553 p.
- Tomppo, E. (1986). Models and methods for analysing spatial patterns of trees. *Communicationes Instituti Forestalis Fenniae*(No. 138).
- Trasobares, A., T. Pukkala et J. Miina (2004). Growth and yield model for uneven-aged mixtures of Pinus sylvestris L. and Pinus nigra Arn. in Catalonia, north-east Spain. *Annals of Forest Science* 61(1): 9-24.
- Uriarte, M., C. D. Canham, J. Thompson et J. K. Zimmerman (2004a). A neighborhood analysis of tree growth and survival in a hurricane-driven tropical forest. *Ecological Monographs* 74(4): 591-614.
- Uriarte, M., R. Condit, C. D. Canham et S. P. Hubbell (2004b). A spatially explicit model of sapling growth in a tropical forest: does the identity of neighbours matter? *Journal of Ecology* 92(2): 348-360.
- Valadon, A. et M. Voccia (2006). Retrouver l'origine des peuplements : un objectif encore trop souvent inaccessible à l'ONF ! *Rendez-vous techniques (ONF)* 14(automne): 9-16.

- Vallet, P., J. F. Dhote, G. Le Moguedec, M. Ravart et G. Pignard (2006). Development of total aboveground volume equations for seven important forest tree species in France. *Forest Ecology and Management* 229(1-3): 98-110.
- Vanclay, J. K. (1989). A growth model for north Queensland rainforests. *Forest Ecology and Management* 27(3-4): 245-271.
- Vanclay, J. K. (1994). Modelling forest growth and yield. Application to mixed tropical forests, CAB International, Wallingford UK. 312 p.
- Vanclay, J. K. et J. P. Skovsgaard (1997). Evaluating forest growth models. *Ecological Modelling* 98(1): 1-12.
- Verzelen, N., N. Picard et S. Gourlet-Fleury (2006). Approximating spatial interactions in a model of forest dynamics as a means of understanding spatial patterns. *Ecological Complexity* 3(3): 209-218.
- Vettenranta, J. (1999). Distance-dependent models for predicting the development of mixed coniferous forests in Finland. *Silva Fennica* 33(1): 51-72.
- Vieilledent, G. (2009). Structurer l'incertitude et la variabilité dans les modèles de dynamique forestière. Thèse en statistiques appliquées à l'écologie forestière. AgroParisTech, Cemagref de Grenoble. 259 p.
- Volin, V. C. et J. Buongiorno (1996). Effects of alternative management regimes on forest stand structure, species composition, and income: A model for the Italian Dolomites. *Forest Ecology and Management* 87(1-3): 107-125.
- Von Teuffel, K., S. Hein, M. Kotar, E. Pinto Preuhsler, J. Puumalainen et P. Weinfurter (2006). End user needs and requirements. *Sustainable Forest Management: Growth Models for Europe*, Springer.
- Ward, E. J. (2008). A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. *Ecological Modelling* 211(1-2): 1-10.
- West, P. W. (1980). Use of diameter increment and basal area increment in tree growth studies. *Canadian Journal of Forest Research* 10(1): 71-77.
- Wimberly, M. C. et B. B. Bare (1996). Distance-dependent and distance-independent models of Douglas-fir and western hemlock basal area growth following silvicultural treatment. *Forest Ecology and Management* 89(1-3): 1-11.
- Wunder, J., B. Reineking, J. F. Matter, C. Bigler et H. Bugmann (2007). Predicting tree death for *Fagus sylvatica* and *Abies alba* using permanent plot data. *Journal of Vegetation Science* 18(4): 525-534.
- Wyckoff, P. H. et J. S. Clark (2000). Predicting tree mortality from diameter growth: a comparison of maximum likelihood and Bayesian approaches. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 30(1): 156-167.
- Wyckoff, W. R. (1990). A basal area increment model for individual conifers in the northern Rocky Mountains. *Forest Science* 36(4): 1077-1104.
- Zhang, L. J., C. H. Peng et Q. L. Dang (2004). Individual-tree basal area growth models for jack pine and black spruce in northern Ontario. *Forestry Chronicle* 80(3): 366-374.
- Zhao, D., B. Borders et M. Wilson (2004). Individual-tree diameter growth and mortality models for bottomland mixed-species hardwood stands in the lower Mississippi alluvial valley. *Forest Ecology and Management* 199(2-3): 307-322.

- Zhao, D., B. Borders et M. Wilson (2005). A density-dependent matrix model for bottomland hardwood stands in the Lower Mississippi Alluvial Valley. *Ecological Modelling* 184(2-4): 381-395.
- Zhao, D., B. Borders, M. Wilson et S. L. Rathbun (2006). Modeling neighborhood effects on the growth and survival of individual trees in a natural temperate species-rich forest. *Ecological Modelling* 196(1-2): 90-102.
- Zucchini, W., M. Schmidt et K. von Gadow (2001). A model for the diameter-height distribution in an uneven-aged beech forest and a method to assess the fit of such models. *Silva Fennica* 35(2): 169-183.

## ANNEXE A : Agrégation du MAID en un modèle peuplement

Nous considérons l'équation de croissance suivante correspond à un modèle arbre indépendant des distances (MAID) :

$$\Delta r_{i,j} = \alpha_j + \beta_j c_{i,j}$$

avec  $\Delta r_{i,j}$  l'accroissement en rayon d'un arbre appartenant à l'espèce  $j$  entre  $t$  et  $t+\Delta t$ ,  $c_{i,j}$  la circonférence au temps  $t$  et  $r_{i,j}$  le rayon au temps  $t$ . Pour simplifier les choses, nous faisons l'hypothèse qu'il n'y a pas de mortalité sur la période de croissance étudiée. A partir de cette équation, nous allons construire un modèle peuplement pour chaque espèce. Nous allons commencer par exprimer l'accroissement du rayon moyen des arbres en fonction des caractéristiques du peuplement. Pour cela, on pose :

$$R_j(t) = \frac{1}{N_j} \sum_{i=1}^{N_j} r_{i,j}$$

où  $R_j(t)$  est le rayon moyen des arbres de l'espèce  $j$  au temps  $t$  et  $N_j$  est le nombre d'arbres appartenant à l'espèce  $j$ . L'accroissement du rayon moyen s'écrit en fonction des accroissements individuels en rayon :

$$\Delta R_j = R_j(t + \Delta t) - R_j(t) = \frac{1}{N_j} \sum_{i=1}^{N_j} r_{i,j}(t + \Delta t) - \frac{1}{N_j} \sum_{i=1}^{N_j} r_{i,j}(t) = \frac{1}{N_j} \sum_{i=1}^{N_j} \Delta r_{i,j}$$

Il faut donc exprimer la somme des accroissements individuels en fonction des variables déjà définies. En introduisant l'équation de la croissance individuelle, on obtient :

$$\sum_{i=1}^{N_j} \Delta r_{i,j} = \alpha_j N_j + 2\pi\beta_j N_j R_j$$

Le modèle peuplement correspond donc au système d'équation suivant :

$$\begin{cases} \Delta R_j = \alpha_j + 2\pi\beta_j R_j \\ \Delta N_j = 0 \end{cases}$$

Cependant, l'accroissement moyen en rayon ( $\Delta R$ ) n'est pas très utile pour le forestier car il ne donne pas une bonne idée de la production du peuplement. On aimerait plutôt connaître les accroissements en volume ou en surface terrière des différentes essences. Nous allons donc maintenant exprimer l'accroissement en surface terrière du peuplement ( $\Delta G$ ) pour chaque espèce en fonction de variables caractéristiques du peuplement.  $\Delta G$  peut s'écrire en fonction des accroissements individuels en surface terrière ( $\Delta g$ ) de la façon suivante :

$$\Delta G_j = G(t + \Delta t) - G(t) = \sum_{i=1}^{i=N_j} g_{i,j}(t + \Delta t) - \sum_{i=1}^{i=N_j} g_{i,j}(t) = \sum_{i=1}^{i=N_j} \Delta g_{i,j}$$

où  $g_{i,j}$  est la surface terrière individuelle d'un arbre de l'espèce  $j$ .

Pour calculer  $\Delta G$  il faut donc expliciter l'accroissement individuel en surface terrière. Par définition, on a  $\Delta g_{i,j} = g_{i,j}(t + \Delta t) - g_{i,j}(t)$  et  $g_{i,espèce}(t) = \pi(r_{i,espèce}(t))^2$ . Nous allons donc commencer par exprimer  $g_{i,j}(t + \Delta t)$  en fonction de l'accroissement individuel en rayon :

$$g_{i,j}(t + \Delta t) = \pi(r_{i,j}(t + \Delta t))^2 = \pi(r_{i,j}(t) + \Delta r_{i,j})^2 = \pi\left((r_{i,j}(t))^2 + 2r_{i,j}(t)\Delta r_{i,j} + (\Delta r_{i,j})^2\right)$$

On a donc :

$$\Delta g_{i,j} = g_{i,j}(t + \Delta t) - g_{i,j}(t) = 2\pi r_{i,j}(t)\Delta r_{i,j} + \pi(\Delta r_{i,j})^2$$

En introduisant l'équation de la croissance individuelle, on peut exprimer  $r_{i,j}(t)\Delta r_{i,j}$  et  $(\Delta r_{i,j})^2$  de la façon suivante :

$$r_{i,j}(t)\Delta r_{i,j} = \alpha_j r_{i,j}(t) + \beta_j 2\pi(r_{i,j}(t))^2 = \alpha_j r_{i,j}(t) + \beta_j 2g_{i,j}(t)$$

$$(\Delta r_{i,j})^2 = \alpha_j^2 + \alpha_j \beta_j 4\pi r_{i,j}(t) + \beta_j^2 4\pi^2 (r_{i,j}(t))^2 = \alpha_j^2 + \alpha_j \beta_j 4\pi r_{i,j}(t) + \beta_j^2 4\pi g_{i,j}(t)$$

On peut maintenant exprimer l'accroissement individuel en surface terrière en fonction de du rayon individuel, de la surface terrière individuel et des paramètres de l'équation de croissance :

$$\Delta g_{i,j} = \pi\alpha_j^2 + (2\pi\alpha_j + 4\pi^2\alpha_j\beta_j)r_{i,j}(t) + (4\pi\beta_j + 4\pi^2\beta_j^2)g_{i,j}(t)$$

Sachant que  $\sum_{i=1}^{N_j} 1 = N_j$ ,  $\sum_{i=1}^{N_j} r_{i,j} = N_j R_j(t)$  et  $\sum_{i=1}^{N_{espèce}} g_{i,j} = G_j(t)$ , on peut maintenant exprimer l'accroissement en surface terrière du peuplement en sommant les accroissements individuels :

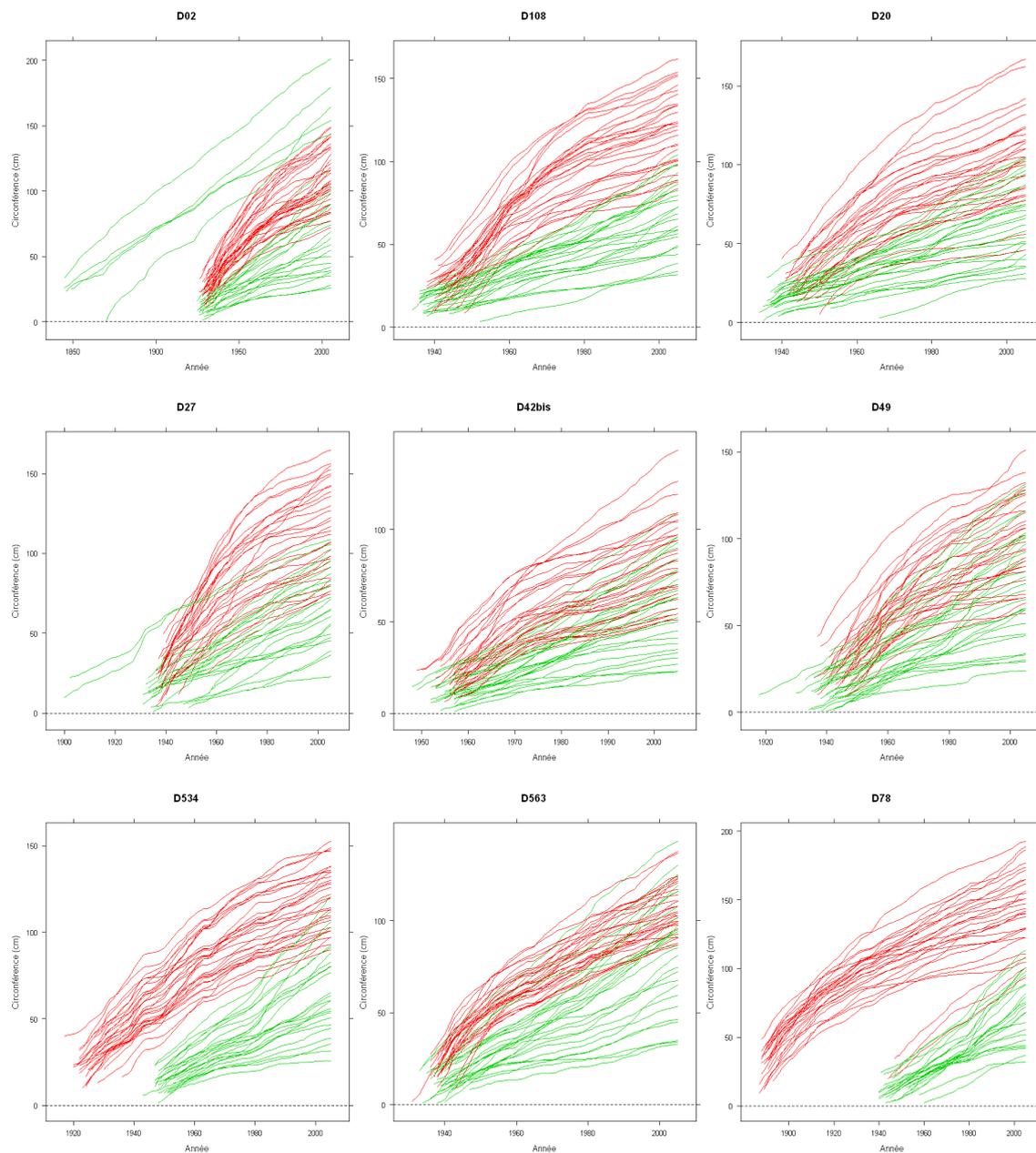
$$\Delta G_j = \pi\alpha_j^2 N_j + (2\pi\alpha_j + 4\pi^2\alpha_j\beta_j)N_j R_j + (4\pi\beta_j + 4\pi^2\beta_j^2)G_j$$

Le système d'équations du modèle peuplement que nous avons obtenu par agrégation du MAID est donc le suivant :

$$\begin{cases} \Delta G_j = \pi\alpha_j^2 N_j + (2\pi\alpha_j + 4\pi^2\alpha_j\beta_j)N_j R_j + (4\pi\beta_j + 4\pi^2\beta_j^2)G_j \\ \Delta R_j = \alpha_j + 2\pi\beta_j R_j \\ \Delta N_j = 0 \end{cases}$$

## ANNEXE B : Evolution de la circonférence au cours du temps dans les 9 dispositifs

Chaque courbe représente l'évolution de la circonférence au cours du temps pour un arbre. Les chênes sont en vert et les pins en rouge.



## ANNEXE C : Implémentation du MADV : le module Oakpine2 dans Capsis

Cette annexe présente l'implémentation du modèle basé sur les distributions de voisinage (MADV) sur la plate-forme Capsis (Coligny *et al.*, 2003). Le nom du module correspondant au MADV s'appelle Oakpine2. Ce travail a été réalisé en collaboration avec François de Coligny et a été présenté lors de la 10<sup>ème</sup> réunion Capsis (Perot et Goreaud, 2008).

### 1 Les arbres dans oakpine2

Dans oakpine2 un arbre a des caractéristiques supplémentaires par rapport à un arbre d'un modèle arbre non spatialisé (*public class Opi2Tree extends GmaidTree*). Il possède une liste de voisins chêne et une liste de voisins pin appartenant au même peuplement (*private Collection<Opi2Tree> oakNeighbours* et *private Collection<Opi2Tree> pineNeighbours*). Chaque arbre de ces listes a les mêmes caractéristiques qu'un arbre de Oakpine2. L'arbre possède également une table reliant les identifiants des voisins aux distances entre l'arbre et ses voisins (*private Map<Integer, Double> neighbourDistances*). Ces distances permettent ensuite de calculer des indices de compétition utilisant les distances séparant le sujet des voisins. Dans cette version, elles sont attribuées de manière aléatoire lors de l'initialisation du modèle après le chargement des données d'entrée. Dans le cas du mélange chêne-pin, les distances ne sont pas utilisées pour le calcul des indices de compétition.

#### 1.1 Le fichier des données d'entrée

Nous avons choisi d'utiliser un seul fichier pour les données d'entrée (figure 1). Dans ce fichier, les données utilisées par le module sont la date de l'inventaire du peuplement, la surface du peuplement inventorié, une liste d'arbres avec les circonférences à 1.30 m et 4 distributions :

- la distribution du nombre de voisins chêne pour les sujets chêne ;
- la distribution du nombre de voisins pin pour les sujets chêne ;
- la distribution du nombre de voisins chêne pour les sujets pin ;
- la distribution du nombre de voisins pin pour les sujets pin.

Pour charger ce fichier, nous avons créé une extension *Opi2Inventory* (contenue dans le package `capsis.extension.ioformat`) qui est appelé par `loadInitStand` dans la classe

Opi2Model et permet de créer le peuplement initial. L'extension *Opi2Inventory* reconnaît les mots clés *inventoryDate* et *inventoryArea*. Pour les données en tableau, elle construit des lignes d'enregistrement (record r) puis récupère les données de ces lignes pour construire les distributions :

```
Opi2Distribution d = new Opi2Distribution (intNumbers, r.minBound, r.classWidth);
settings.distribution[r.subjectSpecies][r.neighbourSpecies] = d ;
```

ou les arbres du peuplement :

```
Opi2Tree tree = new Opi2Tree (r.id, initStand, species, r.girth_cm, r.name);
initStand.addTree (tree);
```

```
# capsis4 - oakpine2 module

# Inventory file

#-----
#Stand
#-----
#Name of the plot : D108
#Date of the stand inventory
inventoryDate = 2006
#Area of the inventory (m²)
inventoryArea = 8081.72

#-----
#Neighbourhood
#-----
#Distributions of neighbours numbers in a disk of 10 m radius around a subject tree (0 : Oak ;
1 : Pine)
#4 distributions are needed
#subjectSpecies : species of the subject trees
#neighbourSpecies : species of the neighbour trees
#minBound : minimum value of neighbours number
#classWidth : Width of the distribution classes
#distributionNumbers = number of trees for each class of the distribution

#subjectSpecies      neighbourSpecies      minBound      classWidth      distributionNumbers
0      0      1      3      {4;17;70;52;57;35;27;14;3;4}
0      1      0      1      {10;14;37;43;50;45;35;35;8;6}
1      0      3      3      {5;24;49;35;30;20;12;6;2;2}
1      1      1      1      {5;10;12;13;52;26;25;20;12;10}

#-----
#Trees
#-----
#List of trees in the stand (0 : Oak ; 1 : Pine)
#girth_cm : girth of the tree at 1.30 m
#id      name      species girth_cm
1      1      1      128
2      2      1      120
3      3      1      132
```

Figure 1 : Exemple de fichier utilisé pour initialiser le module oakpine2

## 2 Reconstruction du voisinage des arbres

Dans le module `oakpine2`, nous reconstruisons un voisinage pour chaque arbre du peuplement une fois pour toute lors de la création du peuplement de départ. Le voisinage est défini par l'ensemble des arbres présents dans un disque de rayon 10 m. Pour reconstruire ce voisinage il nous faut donc le nombre de chênes et le nombre de pins. Ces nombres sont tirés selon les distributions fournies dans le fichier d'entrée. Ensuite on tire les arbres voisins à partir de la liste des arbres du peuplement.

### 2.1 Tirage d'une classe du nombre de voisins

Pour gérer les distributions et les tirages nous avons créé la classe `Opi2Distribution`. Comme nous l'avons vu plus haut (voir figure 1), la distribution d'une variable (ici le nombre d'arbres dans un disque de rayon 10 m autour d'un arbre) est caractérisé par la valeur minimum prise par la variable ( $a_0$ ), la largeur fixe des classes<sup>27</sup> et les effectifs de chaque classe. Dans l'exemple de la figure 1 les distributions contiennent 10 classes. Pour chaque classe on peut calculer une fréquence c'est-à-dire l'effectif d'une classe divisé par l'effectif total (figure 2).

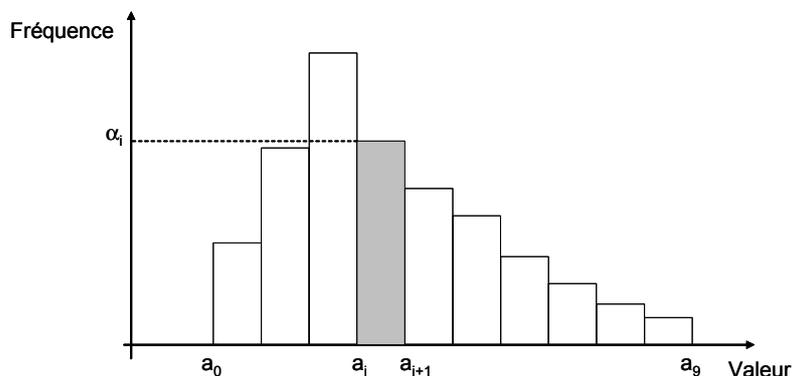


Figure 1 : Représentation de la distribution des fréquences des valeurs prises par une variable aléatoire découpée en 10 classes.

La fréquence d'une classe correspond à la probabilité que la valeur prise par la variable appartienne à cette classe :

$$P(x \in [a_i, a_{i+1}[) = \alpha_i$$

La variable `drawClassIndex` renvoie l'index d'une de ces classes en respectant les probabilités d'occurrence de chaque classe. Pour cela on calcule les fréquences cumulées correspondant à chaque classe :

<sup>27</sup> Il pourrait être intéressant de ne pas avoir une largeur constante, en particulier dans le cas de distribution avec des fréquences pouvant varier fortement sur certaines plages de valeur.

$$P(X \leq a_i) = \sum_{j=0}^{j=i} \alpha_j$$

Ensuite on tire un nombre aléatoire  $r$  dans l'intervalle  $[0, 1[$  selon une loi uniforme et on recherche l'index  $i$  tel que :

$$\sum_{j=0}^{j=i} \alpha_j \leq r < \sum_{j=0}^{j=i+1} \alpha_j$$

Si on prend l'exemple du tableau 1, l'index va de 0 à 9. Si  $r$  est dans  $[0, 0.014[$  alors `drawClassIndex` renvoie l'index 0. Si  $r$  est dans  $[0.505, 0.706 [$ , soit l'intervalle le plus large, alors elle renvoie l'index 4 qui est la classe ayant la fréquence la plus élevée. Pour la dernière classe, si  $r$  est supérieur ou égal à 0.985 alors elle renvoie l'index 9. Chaque classe a ainsi une probabilité d'être tirée conforme à sa fréquence (voir figure 3 pour le code correspondant à `drawClassIndex`).

<i>index</i>	<i>numbers</i>	<i>frequencies</i>	<i>cumulatedFrequencies</i>
0	4	0.014134276	0
1	17	0.060070671	0.014
2	70	0.247349823	0.074
3	52	0.183745583	0.321
4	57	0.201413428	0.505
5	35	0.123674912	0.706
6	27	0.09540636	0.830
7	14	0.049469965	0.925
8	3	0.010600707	0.975
9	4	0.014134276	0.985

Tableau 1 : Exemple de données relatives à une distribution

```

/**   Return a class index according to the frequencies of the distribution.
 */
public int drawClassIndex () {
    double v = random.nextDouble ();    // in [0, 1[
    int index = 0;
    for (int i = 0; i < cumulatedFrequencies.length; i++) {
        if (i >= cumulatedFrequencies.length-1 || cumulatedFrequencies[i+1] > v) {
            index = i;
            break;
        }
    }
    //~ System.out.println ("drawClassIndex: v="+v+" index="+index);
    return index;
}

```

Figure 3 : Code qui permet de tirer une classe de la distribution

## 2.2 Tirage d'une valeur entière du nombre de voisins

Comme nous venons de le voir, *drawClassIndex* renvoie une classe alors qu'il nous faut une valeur. En l'occurrence, ce qui nous intéresse c'est un nombre de voisins soit une valeur entière. La procédure *drawInt* de la classe *Opi2Distribution* permet de renvoyer un entier compris dans une classe de la distribution. Pour cela, connaissant les bornes de la classe, nous avons fait le choix de tirer de manière uniforme un nombre entre les bornes de cette classe. Puis nous avons arrondi ce nombre à l'entier le plus proche (voir figure 4 pour le code correspondant).

```

/**   Return a value for a class of the distribution, the value is drawn uniformly
 *    between the bounds of the class.
 */
public double drawDouble () {
    int index = drawClassIndex ();
    double classMin = minBound + index*classWidth;
    double v = random.nextDouble ();    // in [0, 1[
    double res = classMin + v*classWidth;
    return res;
}

/**   Return an int value for a class of the distribution, the value is drawn uniformly
 *    between the bounds of the class and rounded to the nearest int.
 */
public int drawInt () {
    int res = (int) Math.round (drawDouble ());
    return res;
}

```

Figure 4 : Code qui renvoie une valeur entière comprise dans une classe donnée

## 2.3 Les tirages respectent-ils bien les distributions ?

Afin de vérifier que les tirages effectués respectaient bien les distributions fournies dans le fichier d'entrée, nous avons construit un extracteur de données : l'extension *DEOpi2DistributionChecker* dans *capsis.extension.dataextractor*. Cet extracteur appelé "contrôleur de distribution" permet de visualiser la distribution d'un nombre de tirages donné. Nous pouvons par exemple visualiser la distribution des valeurs obtenues en effectuant 10000 tirages et comparer les fréquences obtenues avec les fréquences issues du fichier d'entrée. Les tirages sont effectués selon la procédure que nous avons décrite plus haut. Nous pouvons choisir de visualiser la distribution du tirage des index ou la distribution des valeurs entières. Nous pouvons également définir la distribution que nous souhaitons afficher en précisant l'espèce de l'arbre sujet et l'espèce des voisins (figure 5).

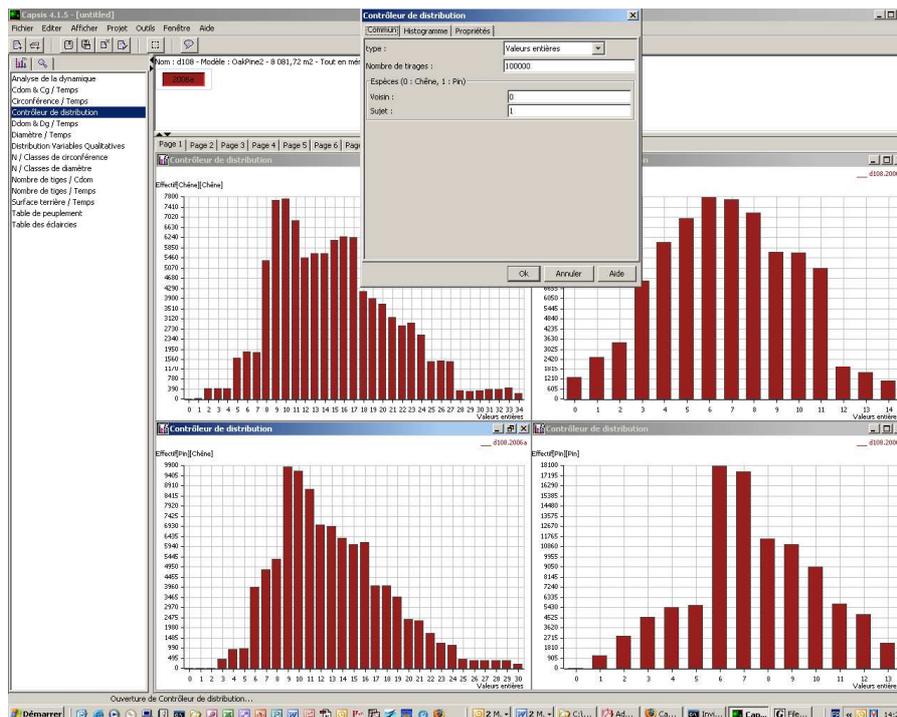


Figure 5 : Exemples de figures obtenues avec le contrôleur de distribution sous Capsis4. Visualisation de la distribution des valeurs entières obtenues en effectuant 10000 tirages selon les 4 distributions fournies dans le fichier d'entrée.

## 2.4 Attribution d'une distance entre un arbre et un voisin

On suppose que dans un disque de rayon  $R$  autour d'un arbre, les voisins sont distribués de manière complètement aléatoire. Sous cette hypothèse, nous pouvons envisager deux approches différentes pour attribuer une distance entre un arbre et son voisin : une approche statistique et une approche simulateur. Nous allons rapidement évoquer l'approche statistique puis nous développerons l'approche simulateur qui est celle utilisée dans Oakpine2.

### 2.4.1 Approche statistique

On va considérer la variable aléatoire  $d$ , dont chaque réalisation est une distance possible entre un arbre et un de ses voisins. Pour caractériser  $d$ , il nous faut identifier sa loi de répartition, c'est à dire la fonction :

$$f(r) = P(r \leq d \leq r + dr)$$

où  $r$  est un rayon centré sur l'arbre sujet. On peut calculer  $f(r)$  en utilisant les probabilité  $P(N(r)=n)$ , où  $N(r)$  est le nombre de voisins d'un arbre à distance  $r$ . En effet, pour qu'un arbre ait un voisin à distance  $d$ , il faut que le nombre de ses voisins augmente d'une unité quand on passe de  $d-dr$  à  $d+dr$ . On peut donc écrire :

$$f(r) = \sum_n P\left(N\left(d - \frac{dr}{2}\right) = n\right) \times P\left(N\left(d + \frac{dr}{2}\right) = n + 1\right)$$

et d'autre part on peut calculer proprement les  $P(N(r)=n)$  car  $N(r)$  suit une loi de Poisson sous l'hypothèse de distribution complètement aléatoire. Si on mène ce calcul à son terme, on doit être capable d'exprimer la loi probabiliste de la distance  $d$ . Mais on se retrouve avec un problème : comment simuler une telle loi ? Au final cette approche ne nous permet pas d'attribuer une distance entre un arbre et son voisin.

#### 2.4.2 Approche simulateur

Nous partons de ce qu'on a facilement accessible en simulation informatique : une loi uniforme. On peut facilement simuler une distribution équiprobable entre différentes classes de valeurs (voir paragraphe 2.1), il suffit de demander à l'ordinateur un nombre entier pseudo aléatoire compris entre 1 et  $p$ , si on a  $p$  classes. Pour utiliser cette méthode simulateur il faut donc construire  $p$  classes de distance équiprobables. Pour cela on peut construire des couronnes correspondant à des classes de distance (figure 6). La probabilité qu'un arbre appartienne à une couronne est proportionnelle à l'espérance du nombre de voisins dans cette couronne<sup>28</sup>, donc à sa surface. La surface totale, à la distance maximale  $R$ , est  $\pi R^2$ . Donc si on a  $p$  classes il faut que chaque couronne ait une surface égale à  $\frac{\pi R^2}{p}$ .

En particulier :  $\pi r_1^2 = \frac{\pi R^2}{p}$  donc  $r_1 = \sqrt{\frac{R^2}{p}} = \frac{R}{\sqrt{p}}$  et :  $\pi r_i^2 = i \frac{\pi R^2}{p}$  donc  $r_i = R \sqrt{\frac{i}{p}}$ .

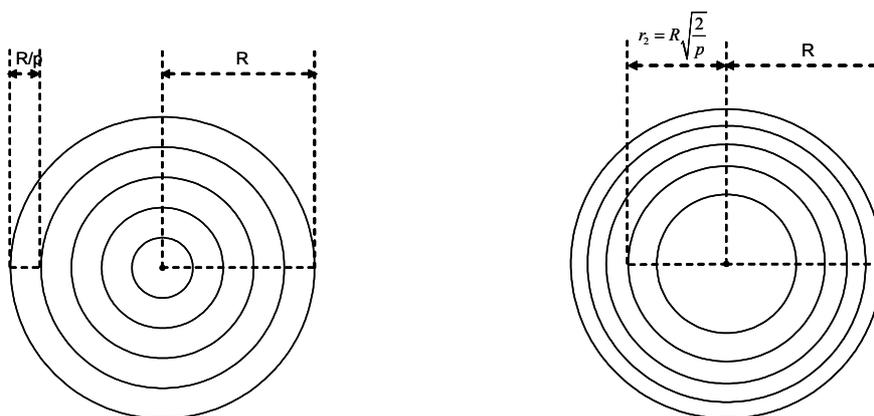


Figure 6 : exemple d'un disque de rayon  $R$ . A gauche le disque est partagé en 5 couronnes de largeurs identiques. Chaque couronne ne correspond pas à des classes de distance équiprobables. A droite, le disque est partagé en 5 couronnes de surfaces égales.  $r_2$  est le rayon du deuxième cercle. Chaque couronne correspond à des classes de distance équiprobables.

<sup>28</sup> Ceci est vrai dans le cas de l'hypothèse de distribution complètement aléatoire des arbres dans le disque de rayon  $R$ .

Nous nous mettons maintenant en position de simulation :

- si on tire un nombre entier  $i$  selon une loi uniforme entre 1 et  $p$ , alors on peut considérer que le voisin est dans la  $i$ ème couronne, c'est à dire qu'il correspond à une distance entre l'arbre et son voisin comprise entre  $r_{i-1}$  et  $r_i$  ;
- si  $p$  est assez grand, au vu de la précision de la position d'un arbre, on ne fait pas une grosse erreur en lui affectant la distance  $r_i$  ;
- si  $p$  tend vers l'infini, alors tirer un nombre entier  $i$  selon une loi uniforme entre 1 et  $p$  revient à tirer un nombre réel  $x$  selon une loi uniforme entre 0 et 1, et utiliser  $i = p \cdot x$ . Dans ce cas, on peut directement écrire :  $d = R\sqrt{x}$ .

Pour attribuer une distance  $d$  entre un arbre et un voisin nous tirerons donc un nombre aléatoire  $x$  entre 0 et 1 selon une loi uniforme et la distance attribuée sera égale à  $d = R\sqrt{x}$  où  $R$  est le rayon du disque. La figure 7 montre les positions des voisins obtenues en suivant cette procédure. Elle montre également ce qu'on obtiendrait si on effectuait un tirage de la distance de manière uniforme entre 0 et  $R$ .

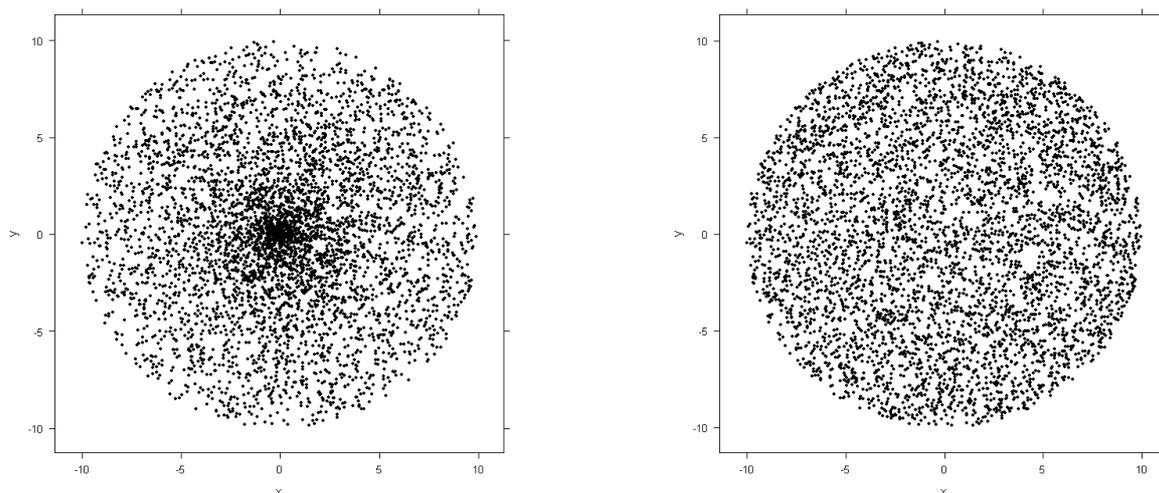


Figure 7 : Représentation de la position d'un arbre voisin dans un disque de rayon  $R$  après attribution d'une distance. 5000 tirages de distance sont représentés. A gauche les distances sont tirées de manière uniforme entre 0 et  $R$ . A droite les distances sont tirées en respectant l'hypothèse de distribution aléatoire des voisins (procédure utilisée dans oakpine2).

## 2.5 Tirage des voisins dans la liste des arbres du peuplement

Nous avons vu comment nous pouvons obtenir une valeur du nombre de voisins à partir d'une distribution connue. Nous avons également vu comment attribuer une distance entre un arbre et ses voisins. Nous allons voir maintenant comment on tire les arbres voisins à partir de la liste des arbres du peuplement. Dans ce module, les voisins d'un arbre sont déterminés au moment de l'initialisation du modèle (`initializeModel` dans la classe `Opi2Model`) par la

méthode `setNeighbourhoods` . Cette méthode commence par créer un tableau avec les chênes et un tableau avec les pins. Ensuite elle passe en revue tous les arbres du peuplement. Pour chaque arbre `t` elle récupère l'espèce :

```
for (Iterator i = stand.getTrees ().iterator (); i.hasNext ();) {
    Opi2Tree t = (Opi2Tree) i.next ();
    int subjectSpecies = t.getSpecies ().getValue (); // 0 OR 1
```

Ensuite elle tire le nombre de voisins chêne (`numberOfOakNeighbours`) dans la bonne distribution :

```
// Oak neighbours
int neighbourSpecies = 0;
Opi2Distribution distribution = sets.distribution[subjectSpecies][neighbourSpecies];
int numberOfOakNeighbours = distribution.drawInt ();
```

Ensuite la collection des chênes du peuplement est triée selon un ordre aléatoire. Les `numberOfOakNeighbours` premiers chênes sont ajoutés à la liste des voisins chêne de l'arbre `t`. Dans le même temps, pour chaque voisin, une distance en mètre est tirée au hasard (voir paragraphe précédent) dans l'intervalle `[minDistance, 10[` où `minDistance` est la distance minimum entre le sujet et son voisin que nous avons fixé à la somme des rayons en mètres des deux arbres. Cette distance est ajoutée à la table des distances de l'arbre :

```
Collections.shuffle (oaks);
Iterator j = oaks.iterator ();
Collection<Opi2Tree> oakNeighbours = new ArrayList<Opi2Tree> ();
for (int k = 0; k < numberOfOakNeighbours; k++) {
    Opi2Tree n = (Opi2Tree) j.next ();
    oakNeighbours.add (n);
    double neighbourTreeRadius_m = n.getDbh()/200; // tp 06.02.2008
    double minDistance=subjectTreeRadius_m + neighbourTreeRadius_m; //minimum distance
    double distance = Math.pow(random.nextDouble (),0.5)*(neighbourhoodRadius-
minDistance)+minDistance; // distance (m) in [minDistance, 10[ tp 22.04.2008.
    t.addNeighbourDistance (n.getId (), distance);
}
t.setOakNeighbours (oakNeighbours);
```

La même chose est réalisée pour les voisins pin de l'arbre `t`.

### 3 Prise en compte du processus de croissance et des éclaircies

Dans oakpine2 chaque arbre a une liste de voisins qui appartiennent au peuplement. Ces arbres sont donc soumis de la même façon au processus de croissance, à la mortalité et aux interventions sylvicole. Dans cette version du module il n'y a pas de mortalité. Nous allons donc voir ici comment on gère le voisinage de chaque arbre pour tenir compte des éclaircies et du processus de croissance.

#### 3.1 La croissance individuelle et calcul de l'indice de compétition

La croissance individuelle d'un arbre est réalisée par la procédure `processGrowth` de la classe `Opi2Tree`. En fonction de l'essence de l'arbre, elle utilise soit l'équation chêne soit l'équation pin. Dans les deux cas l'équation fait intervenir un indice de compétition qui est calculé par la procédure `getIC`. Cette procédure calcul l'indice de compétition d'un arbre en utilisant la liste des voisins et éventuellement les distances :

```

if (species.equals (Opi2Species.OAK)) {      // OAK
    for (Opi2Tree n : allNeighbours) {

        int localDensity=1;
        IC += localDensity;
    }
} else {      // PINE
    for (Opi2Tree n : pineNeighbours) {
        double basalArea = n.getGirth ()*n.getGirth () / (4*Math.PI) / 10000;
        IC += basalArea;
    }
}
return IC;

```

#### 3.2 Croissance du peuplement et mise à jour de la liste des voisins d'un arbre

Lors du processus de croissance du peuplement (méthode `processGrowth` dans la classe `Opi2model`) on fait d'abord pousser tous les arbres du peuplement entre l'années `n` et l'année `n+1` :

```

for (Iterator j = refStand.getTrees ().iterator (); j.hasNext ();) {
    Opi2Tree t = (Opi2Tree) j.next ();
    Opi2Tree newTree = t.processGrowth (random);
    newStand.addTree (newTree);
}

```

Ensuite pour chaque arbre *t* on met à jour les caractéristiques de ses voisins. On commence par récupérer les distances entre l'arbre *t* et ses voisins (elles n'ont pas changé). Ensuite, on récupère les identifiants des voisins de l'arbre *t* à l'année *n*. Ces identifiants nous permettent de récupérer les voisins mis à jour dans le peuplement à l'année *n+1*. On procède d'abord pour les voisins chêne puis on fait de même pour les voisins pin :

```
// Update new trees neighbourhoods: same tree ids than next year but trees of newStand
for (Iterator j = refStand.getTrees ().iterator (); j.hasNext ();) {
    Opi2Tree t = (Opi2Tree) j.next ();
    Opi2Tree newTree = (Opi2Tree) newStand.getTree (t.getId ());

    // SAME neighbour distances map : CAUTION: SAME REFERENCE,
    // TO BE CHECKED WHEN MORTALITY OR REGENERATION WILL BE ADDED
    newTree.setNeighbourDistances (t.getNeighbourDistances ());

    Collection<Opi2Tree> newOakNeighbours = new ArrayList<Opi2Tree> ();
    Collection<Opi2Tree> oakNeighbours = t.getOakNeighbours ();
    for (Opi2Tree n : oakNeighbours) {
        int nId = n.getId ();
        Opi2Tree newN = (Opi2Tree) newStand.getTree (nId);
        newOakNeighbours.add (newN);
    }
    newTree.setOakNeighbours (newOakNeighbours);
}
```

### 3.3 Modification de la liste des voisins lors d'une éclaircie

Après une éclaircie, on fait intervenir la méthode `processPostThinning` de la classe de la classe `Opi2Model`. Les arbres dans `oakpine2` ont un champ effectif. Dans les modèles individus-centrés, ce champ est égal à 1 lorsque l'arbre est vivant ou 0 lorsqu'il disparaît par mortalité ou éclaircie. La méthode commence par enlever tous les arbres qui ont été éclaircis, c'est à dire ceux donc la valeur du champ effectif est égale à 0 dans le peuplement après intervention :

```
// 1. Remove from the tree collection all the trees with number = 0
// fc - 1.2.2008
for (Iterator i = std.getTrees ().iterator (); i.hasNext ();) {
    Opi2Tree t = (Opi2Tree) i.next ();
    if (t.getNumber () == 0) {i.remove ();}
}
```

Ensuite on prend la liste des voisins d'un arbre et on les passe en revue. Si un voisin n'est pas dans le peuplement après éclaircie on le retire de la liste des voisins. Pour les distances, on crée une nouvelle table et on y ajoute toutes les distances pour lesquelles les identifiants sont bien dans le peuplement après éclaircie. Si un arbre a été enlevé, son identifiant n'est plus dans le peuplement, la distance ne figure plus dans la table des distances :

```

// 2. For each tree in the thinned stand, update its oak / pine neighbour collections:
// remove the ids of the cut neighbours
for (Iterator i = std.getTrees ().iterator (); i.hasNext ();) {
    Opi2Tree t = (Opi2Tree) i.next ();

    Collection<Opi2Tree> oaks = t.getOakNeighbours ();
    for (Iterator j = oaks.iterator (); j.hasNext ();) {
        Opi2Tree n = (Opi2Tree) j.next ();
        if (!std.contains (n)) {j.remove ();}
    }

    Collection<Opi2Tree> pines = t.getPineNeighbours ();
    for (Iterator j = pines.iterator (); j.hasNext ();) {
        Opi2Tree n = (Opi2Tree) j.next ();
        if (!std.contains (n)) {j.remove ();}
    }

    Map<Integer,Double> map = t.getNeighbourDistances ();
    Map<Integer,Double> newMap = new HashMap<Integer,Double> ();
    for (int nId : map.keySet ()) {

        if (std.getIds ().contains (nId)) {
            double distance = map.get (nId);
            newMap.put (nId, distance);
        }

    }
    t.setNeighbourDistances (newMap);
}

```

## Conclusion et perspectives

Dans ce document nous avons présenté comment nous avons implémenté un modèle arbre avec distribution de voisinage (Oakpine2) dans la plate-forme Capsis. Nous avons précisé les données d'entrée nécessaires au fonctionnement du module, la façon dont le voisinage d'un arbre est reconstruit et la prise en compte de la croissance et des éclaircies pour mettre à jour le voisinage au cours du temps.

Au cours de l'implémentation, nous avons réalisé à plusieurs reprises des vérifications. Nous avons en particulier vérifié que les distributions étaient bien respectées grâce à l'extracteur de données qui permet de visualiser les résultats des tirages sous la forme d'histogrammes. Nous avons également réalisé des comparaisons qualitatives entre les simulations réalisées par oakpine2 et celles réalisées par oakpine1 (module correspondant au modèle arbre dépendant des distances). Ces comparaisons ont permis de vérifier la cohérence du module oakpine2.

Nous avons été amenés à faire certains choix au cours de l'implémentation. Pour tirer une valeur entière du nombre de voisins, nous avons choisi de faire un tirage uniforme dans la classe sélectionnée. Nous aurions pu faire autrement par exemple renvoyer le milieu de la classe. Si les largeurs de classe sont importantes le choix que nous avons fait peut être gênant. D'autre part, nous avons attribué des distances entre un arbre et ses voisins. Dans cette version nous avons choisi d'attribuer ces distances en considérant que les voisins étaient distribués de manière complètement aléatoire dans le disque de voisinage. Nous pourrions attribuer ces distances de façon plus réaliste. En effet, on peut supposer par exemple qu'il y a une distance minimale entre un arbre et son voisin (distance de répulsion).

## Références citées

- Coligny, F. d., P. Ancelin, G. Cornu, B. Courbaud, P. Dreyfus, F. Goreaud, S. Gourlet-Fleury, C. Meredieu et L. Saint-Andre (2003). CAPSIS : Computer-Aided Projection for Strategies In Silviculture : Advantages of a shared forest-modelling platform. International Workshop of IUFRO working party 4.01 "Reality, models and parameter estimation". Sesimbra, Portugal, June 2-5, 2002. (eds) Modelling Forest Systems. CABI Publishing, Wallingford, UK.: 319-323.
- Perot, T. et F. Goreaud (2008). Oakpine2: A tree model using neighbourhood distributions to describe mixed forest growth. 10ème réunion Capsis. UMR AMAP Montpellier.

## ANNEXE D : Calcul du taux de mélange optimal à partir du modèle peuplement

L'accroissement en surface terrière du peuplement en fonction du taux de mélange en chêne s'écrit de la manière suivante :

$$\Delta G(x) = A_c(\alpha_{k,c}) \frac{Gx}{\pi \bar{r}_{G,c}^2} + B_c(\alpha_{k,c}, \beta_{k,c}) \bar{r}_c \frac{Gx}{\pi \bar{r}_{G,c}^2} + C_c(\beta_{k,c}) Gx +$$

$$A_p(\alpha_{k,p}) \frac{G(1-x)}{\pi \bar{r}_{G,p}^2} + B_p(\alpha_{k,p}, \beta_{k,p}) \bar{r}_p \frac{G(1-x)}{\pi \bar{r}_{G,p}^2} + C_p(\beta_{k,p}) G(1-x)$$

Avec :

$$A_j(\alpha_{k,j}) = (\alpha_{k,j}/10)^2 \pi, \quad B_j(\alpha_{k,j}, \beta_{k,j}) = (1 + 2\pi\beta_{k,j}/10) 2\alpha_{k,j}\pi/10, \quad C_j(\beta_{k,j}) = (1 + \beta_{k,j}\pi/10)\beta_{k,j} 4\pi/10,$$

$$\alpha_{k,c} = \alpha'_{k,c} + \lambda_{c,c} \frac{1}{S} K_{c,c}(R) \frac{Gx}{\pi \bar{r}_{G,c}^2} + \lambda_{c,p} \frac{1}{S} K_{c,p}(R) \frac{G(1-x)}{\pi \bar{r}_{G,p}^2} \text{ et } \alpha_{k,p} = \alpha'_{k,p} + \lambda_{p,p} \frac{1}{S} K_{p,p}(R) G(1-x)$$

Sachant que  $(\Delta G(x))' = (\Delta G_c(x))' + (\Delta G_p(x))'$  nous allons calculer séparément la dérivée de l'accroissement des chênes et la dérivée de l'accroissement des pins.

En posant pour les chênes :

$$m_1 = \frac{\pi}{10^2}, \quad m_2 = \frac{(2\pi + \beta_{k,c} 4\pi^2)}{10^2}, \quad m_3 = \frac{\beta_{k,c} 4\pi}{10^2} + \left(\frac{\beta_{k,c}}{10}\right)^2 4\pi^2,$$

$$n_1 = \alpha'_{k,c}, \quad n_2 = \lambda_{c,c} \frac{K_{c,c}(R)}{S} \text{ et } n_3 = \lambda_{c,p} \frac{K_{c,p}(R)}{S}$$

Et en posant pour les pins :

$$p_1 = \frac{\pi}{10^2}, \quad p_2 = \frac{(2\pi + \beta_{k,p} 4\pi^2)}{10^2}, \quad p_3 = \frac{\beta_{k,p} 4\pi}{10^2} + \left(\frac{\beta_{k,p}}{10}\right)^2 4\pi^2$$

$$q_1 = \alpha'_{k,p} \text{ et } q_2 = \lambda_{p,p} \frac{K_{p,p}(R)}{S}$$

On a alors pour les chênes :

$$(\Delta G_c(x))' = a_c x^2 + b_c x + c_c$$

avec

$$a_c = \left[ 3m_1 \left( n_2 \frac{G}{\pi \bar{r}_{G,c}^2} - n_3 \frac{G}{\pi \bar{r}_{G,p}^2} \right)^2 \frac{G}{\pi \bar{r}_{G,c}^2} \right]$$

$$b_c = \left[ 4m_1 \left( n_1 + n_3 \frac{G}{\pi \bar{r}_{G,p}^2} \right) \left( n_2 \frac{G}{\pi \bar{r}_{G,c}^2} - n_3 \frac{G}{\pi \bar{r}_{G,p}^2} \right) \frac{G}{\pi \bar{r}_{G,c}^2} + 2m_2 n_2 \left( \frac{G}{\pi \bar{r}_{G,c}^2} \right)^2 \bar{r}_c - 2 \frac{m_2 n_3 G^2 \bar{r}_c}{\pi^2 \bar{r}_{G,p}^2 \bar{r}_{G,c}^2} \right]$$

$$c_c = \left[ m_1 \left( n_1 + n_3 \frac{G}{\pi \bar{r}_{G,p}^2} \right)^2 \frac{G}{\pi \bar{r}_{G,c}^2} + \frac{m_1 m_2 G \bar{r}_c}{\pi \bar{r}_{G,c}^2} + \frac{m_2 n_3 G^2 \bar{r}_c}{\pi^2 \bar{r}_{G,p}^2 \bar{r}_{G,c}^2} + m_3 G \right]$$

Et pour les pins :

$$(\Delta G_p(x))' = a_p x^2 + b_p x + c_p$$

avec

$$a_p = \left[ -3p_1 q_2^2 \frac{G^3}{\pi \bar{r}_{G,p}^2} \right]$$

$$b_p = \left[ 4p_1 q_2 G^2 \frac{(q_1 + q_2 G)}{\pi \bar{r}_{G,p}^2} + 2p_1 q_2^2 \frac{G^3}{\pi \bar{r}_{G,p}^2} + 2p_2 q_2 \frac{G^2}{\pi \bar{r}_{G,p}^2} \bar{r}_p \right]$$

$$c_p = \left[ -p_1 (q_1 + q_2 G)^2 \frac{G}{\pi \bar{r}_{G,p}^2} - \frac{2p_1 q_2 G^2 (q_1 + q_2 G)}{\pi \bar{r}_{G,p}^2} - p_2 q_1 \bar{r}_p \frac{G}{\pi \bar{r}_{G,p}^2} - 2p_2 q_2 \frac{G^2}{\pi \bar{r}_{G,p}^2} \bar{r}_p - p_3 G \right]$$

On a donc :

$$(\Delta G(x))' = ax^2 + bx + c \text{ avec } a = a_c + a_p, b = b_c + b_p \text{ et } c = c_c + c_p$$

En posant  $(\Delta G(x))' = 0$  et en appliquant la résolution d'une équation du second degré on

obtient deux solutions :  $x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$  et  $x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$

Tableau 0.1 : Valeur des coefficients de la fonction  $(\Delta G(x))'$  et solutions de l'équation  $(\Delta G(x))' = 0$  pour les différents dispositifs.

Disp.	a	b	c	b <sup>2</sup> -4ac	x <sub>1</sub>	x <sub>2</sub>
D02	1493	-64785	37871	3971025509	0.593	43
D108	14688	-168693	69966	24346438131	0.431	11
D20	15403	-172769	62590	25992588522	0.375	11
D27	6466	-83284	37311	5971133257	0.465	12
D42	5225	-74980	33143	4929341608	0.457	14
D49	13152	-185208	82799	29946007476	0.462	14
D534	9357	-102027	40109	8908289845	0.408	10
D563	268	-31239	22837	951347822	0.736	116
D78	5457	-95660	52986	7994190952	0.573	17
MOY	7552	-133134	64898	15764319444	0.502	17

D'après le tableau ci-dessus, le coefficient  $a$  est toujours positif donc la fonction  $(\Delta G(x))'$  est convexe. Elle est donc négative entre  $x_1$  et  $x_2$  et positive pour  $x < x_1$  et  $x > x_2$ .  $x_1$  est donc un maximum pour la fonction  $\Delta G(x)$ .

## ANNEXE E : Indices et tests statistiques utilisés pour évaluer des modèles de croissance

*Quelques indices utilisés pour évaluer la performance d'un modèle par rapport la prédiction d'une variable à partir du jeu de données de calibration*

Indice	Formule	Définition des variables et commentaires
Erreur quadratique moyenne	$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p - 1}$	n est la taille du jeu de données de calibration. p est le nombre de paramètres estimés. Il s'agit d'un estimateur non biaisé de la variance résiduelle (Palagos, 1997).
R <sup>2</sup> ajusté	$R^2_{ajusté} = \frac{(n-1)R^2 - p}{n - p - 1}$	Le R <sup>2</sup> ajusté permet d'estimer le pouvoir explicatif du modèle
EPVC (estimation de l'erreur de prédiction par validation croisée)	$EPVC_k = \frac{\sum_1^{n_k} (y_i - \hat{y}_i)}{n_k} \quad EPVC = \frac{\sum_{k=1}^K EPVC_k}{K}$	L'échantillon de taille n est divisé en K parties de taille à peu près égale. On répète K fois la procédure suivante : on enlève la k <sup>ième</sup> partie et on ajuste le modèle sur le jeu de données restant. On calcule l'erreur de prédiction EPVC <sub>i</sub> sur les données de la k <sup>ième</sup> partie puis on calcule EPVC. L'EPVC permet d'estimer le pouvoir prédictif du modèle. Cet indice est basé sur les techniques de ré-échantillonnage, il est utilisé lorsque l'on ne dispose pas d'un jeu de données indépendant. Il est également possible d'estimer l'erreur de prédiction par bootstrap (Palagos, 1997).
PRESS	$PRESS = \sum_{i=1}^n (y_i - \hat{y}_{i-i})^2 \quad Mpress = \frac{\sum_{i=1}^n (y_i - \hat{y}_{i-i})}{n}$ $MApress = \frac{\sum_{i=1}^n  y_i - \hat{y}_{i-i} }{n}$	$\hat{y}_{i-i}$ représente l'estimation de la valeur $y_i$ à l'aide du modèle ajusté sur toutes les données sauf $y_i$ . PRESS est équivalent à EPVC lorsque l'on choisit K=n (voir un exemple d'application dans Sanchez-Gonzalez <i>et al.</i> , 2006).
Critère d'Akaiké (AIC)	$AIC = -2 \log L + 2k$	n est la taille du jeu de données de calibration, k est le nombre de paramètres estimés, L est la vraisemblance maximisée. Le critère d'Akaiké est une estimation de la divergence entre le vrai modèle et le modèle estimé. Ce critère est utilisé dans la sélection de modèle. Le modèle ayant la valeur d'AIC la plus faible est le meilleur. Lancelot <i>et al.</i> (2005) présente les conditions d'application de ce critère, en particulier il est nécessaire que les modèles comparés dérivent tous d'un même modèle "complet". Ce critère peut conduire à sélectionner des modèles ayant trop de paramètres surtout dans le cas de petits échantillons d'observation. Pour limiter ce risque on peut utiliser le critère d'Akaiké corrigé.
Critère d'Akaiké corrigé (AICc)	$AICc = AIC + \frac{2k(k+1)}{n-k-1}$	Le critère d'Akaiké corrigé est une estimation sans biais de la divergence entre le vrai modèle et le modèle estimé (Palagos, 1997).
Critère BIC (Bayesian Information Criterion)	$BIC = -2 \log L + k \log n$	n est la taille du jeu de données de calibration, k est le nombre de paramètres estimés, L est la vraisemblance maximisée. Le BIC a été initialement proposé pour sélectionner les modèles dans le cas de grands échantillons (plusieurs milliers d'observations) pour lesquels l'AIC et l'AICc ont tendance à sélectionner des modèles comportant de nombreuses variables explicatives : le BIC aboutit à des modèles plus parcimonieux. Cependant, les bases théoriques sous-tendant les deux approches (AIC vs. BIC) sont différentes, l'utilisation de l'AIC étant en premier lieu dans un objectif de prédiction, et non de décision vis-à-vis de la signification statistique des paramètres retenus dans le modèle (Lancelot et Lesnoff, 2005).

*Indices et tests statistiques utilisés pour évaluer la performance d'un modèle par rapport la prédiction d'une variable à l'aide d'un jeu de données indépendant. La plupart de ces indices sont extraits de Gourlet-Fleury (1997).*

Indice	Formule	Définition des variables, test(s) associé(s) à l'indice et commentaires
Biais moyen	$B = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$	<p><math>y_i</math> = valeur observée sur l'individu <math>i</math>, <math>\hat{y}_i</math> = valeur estimée par le modèle sur l'individu <math>i</math></p> <p>Hypothèse nulle <math>H_0 : B = 0</math>. <math>T = \frac{\sqrt{n} \times B}{s}</math> suit une loi normale centrée réduite. <math>s</math> est l'écart-type résiduel estimé sur l'échantillon ayant servi à calibrer le modèle.</p> <p><math>\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{s_u^2}</math> suit un <math>\chi^2</math> à <math>(n-2)</math> degrés de liberté. Ici <math>S_u^2</math> est défini par l'utilisateur et représente le degré de précision qu'il attend des prédictions.</p>
Biais relatif moyen	$B_{rel} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i) / y_i}{n}$	
Erreur absolue moyenne	$MAE = \frac{\sum_{i=1}^n  y_i - \hat{y}_i }{n}$	
Erreur absolue relative moyenne	$MAE_{rel} = \frac{\sum_{i=1}^n  y_i - \hat{y}_i  /  y_i }{n}$	Cet indice peut être intéressant pour comparer des modèles entre eux. Il est alors possible de fixer un seuil à cet indice, par exemple 10%. Cet indice est très sensible aux faibles valeurs de $y_i$ . De plus si $y_i = 0$ l'indice n'est pas défini. On peut alors le remplacer par $MAE/\bar{y}$ ou $RMSE/\bar{y}$ (Mayer et Butler, 1993).
Rapport des variances	$VR = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	Cet indice compare la variabilité des estimations et la variabilité des observations. Une valeur de 1 indique une bonne adéquation entre variabilité prédite et variabilité observée (Palahi <i>et al.</i> , 2004).
Erreur quadratique moyenne de validation	$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$	$n$ est la taille du jeu de données indépendant Il s'agit d'un des critères quantitatifs les plus couramment utilisés.
Ecart-type résiduel de validation	$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$	$n$ est la taille du jeu de données indépendant

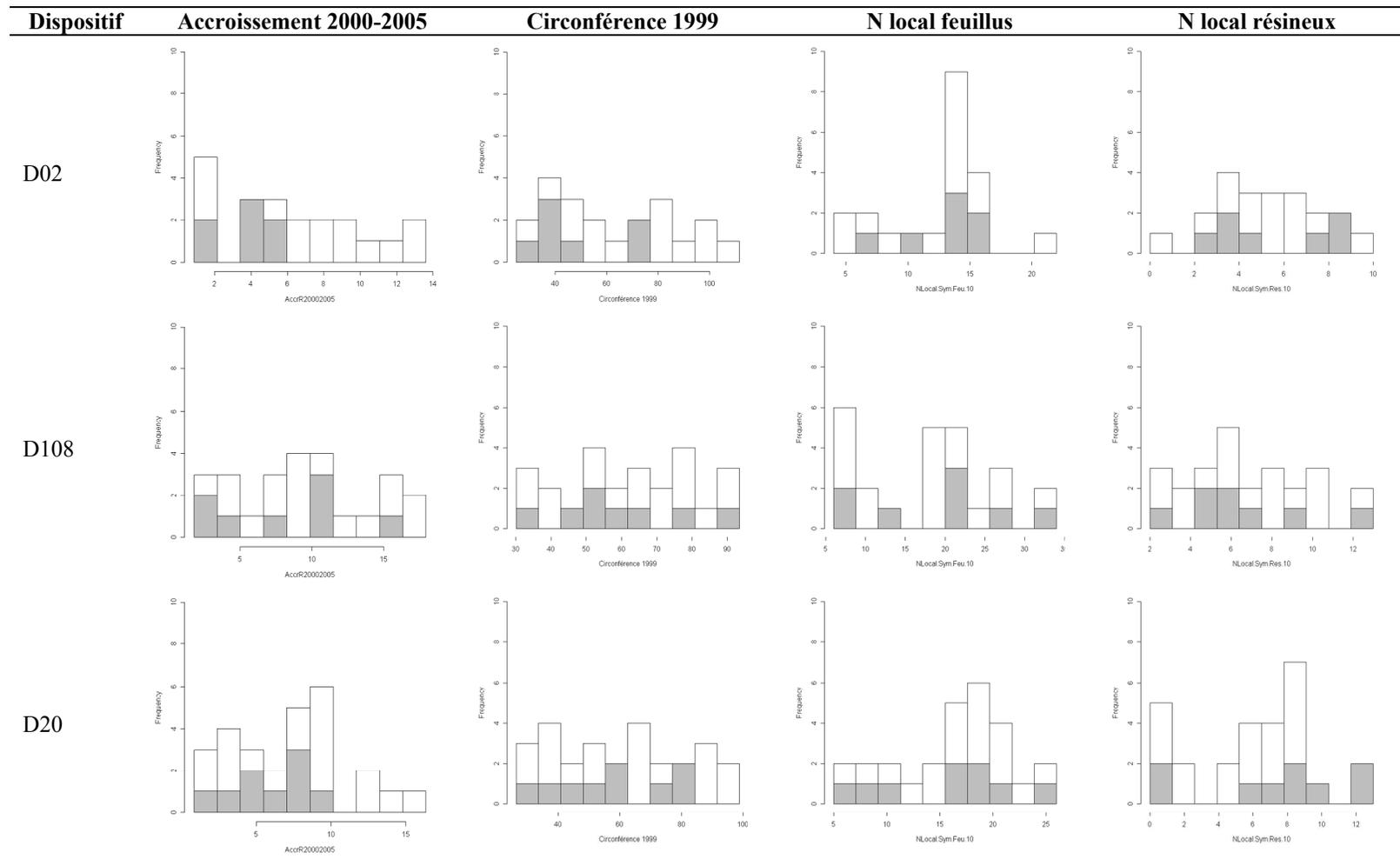
Indice	Formule	Définition des variables, test(s) associé(s) à l'indice et commentaires
Statistique Q	$Q = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n \times s^2}$	n est la taille du jeu de données indépendant, s est l'écart-type résiduel estimé sur l'échantillon ayant servi à calibrer le modèle. Hypothèse nulle H <sub>0</sub> : égalité des variances résiduelles estimées sur le jeu de données de calibration et sur le jeu de données indépendant. Q suit un F à (n, m-p) degrés de liberté avec p = nombre de paramètres du modèle ajusté et m = nombre d'individus ayant servi à ajuster le modèle. Cette statistique est très sensible aux points aberrants
Pseudo-R <sup>2</sup> ou "modelling efficiency"	$pseudoR^2 = EF = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	Il s'agit d'une extension du R <sup>2</sup> à la classe des modèles non linéaires (Mayer et Butler, 1993). Contrairement au R <sup>2</sup> , EF peut être négatif (cas d'un mauvais modèle). Des valeurs proches de 1 indiquent un bon modèle.
Test de la pente = 1 pour la régression $y_i = a \times \hat{y}_i + b$ (Pavé, 1994 page 55)	$t = \frac{a - a_{theo}}{s_a} \quad s_a = \frac{s}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \mu(\hat{y}))^2}} \quad s = \sqrt{\frac{\sum_{i=1}^n ((a\hat{y}_i + b) - y_i)^2}{n - 2}}$	a <sub>theo</sub> est la pente théorique 1. s <sub>a</sub> est l'estimation de l'écart-type de la pente. s est l'estimation de l'écart-type résiduel de la régression. Hypothèse nulle H <sub>0</sub> : a = 1. t suit une loi de Student à n-2 degré de liberté
Test de l'ordonnée à l'origine = 0 pour la régression $y_i = a \times \hat{y}_i + b$ (Pavé, 1994 page 55)	$t = \frac{b}{s \sqrt{\frac{1}{n} + \frac{\mu(\hat{y})^2}{\sum_{i=1}^n (\hat{y} - \mu(\hat{y}))^2}}}$	Hypothèse nulle H <sub>0</sub> : b = 0. t suit une loi de Student à n-2 degrés de liberté
Test F simultané sur la régression $y_i = a \times \hat{y}_i + b$	$FS = \frac{nb^2 + 2b(a-1)\left(\sum_{i=1}^n \hat{y}_i\right) + (a-1)\left(\sum_{i=1}^n \hat{y}_i^2\right)}{2s^2}$	Hypothèse nulle H <sub>0</sub> : a = 1 et b = 0. SF suit un F à (2, n-2) degrés de liberté. s = écart-type résiduel de la régression. D'après Thornton (1996) l'inconvénient majeur de ce test est qu'il dépend de la taille de l'échantillon : plus l'échantillon est grand plus le test rejette à tort l'hypothèse nulle.
La mesure d'association U	$U = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2 + \frac{1}{n} \sum_{i=1}^n \hat{y}_i^2}}$	U mesure une distance, elle varie entre 0 lorsqu'il y a adéquation parfaite des prédictions aux données observées, et 1 lorsque les valeurs observées et les valeurs prédites valent 0 (voir Smith et Rose, 1995 pour plus de détail). Contrairement au coefficient de corrélation cet indice prend en compte la valeur de la pente que l'on souhaite égale à 1.
Le coefficient de Janus	$J = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\frac{1}{m} \sum_{j=1}^m (y_j - \hat{y}_j)^2}$	n est la taille du jeu de données indépendant et m est la taille du jeu de données de calibration du modèle. Ce coefficient est un indice qui permet d'évaluer le changement dans l'exactitude du modèle entre le jeu de données de calibration et le jeu de données d'évaluation.

*Quelques indices et tests statistiques utilisés pour évaluer la performance d'un modèle pour la prédiction de la distribution des valeurs d'une variable.*

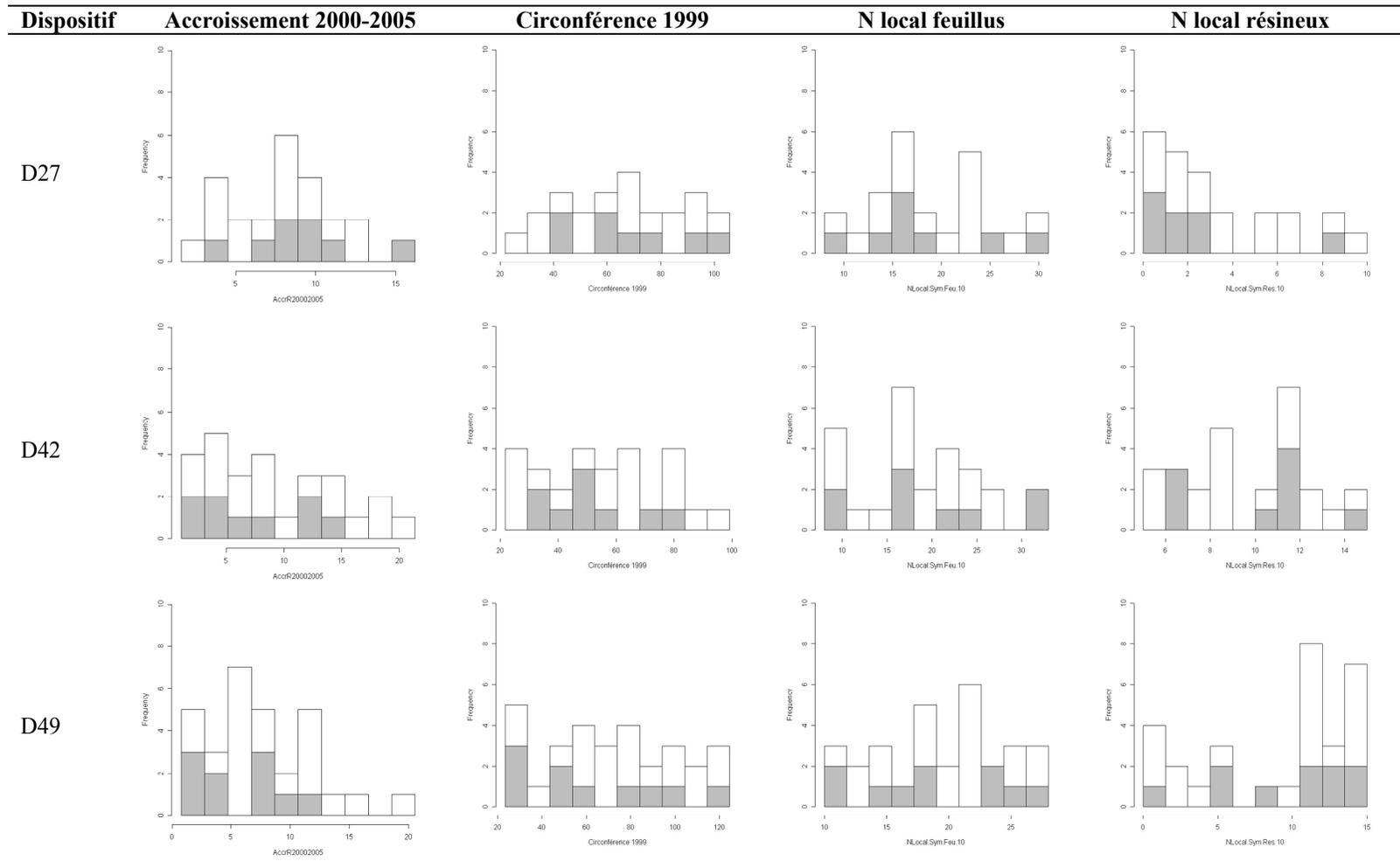
Indice	Formule	Définition des variables, test associé à l'indice et commentaires
Smirnov	$D = \max  G_M(x) - F_N(x) $	G est la fonction de distribution de l'échantillon 1 de taille M. F est la fonction de distribution de l'échantillon 2 de taille N H0 = Les deux échantillons proviennent de la même distribution. Sous H0 on a : $P\left(\sqrt{\frac{M \times N}{M + N}} \times D \geq z\right) = K(z)$ Les valeurs de K(z) sont obtenues dans une table en fonction de la valeur de z. Si K(z) est inférieur à 5% (ou 1%) on rejette H0. Il s'agit d'un test non paramétrique.
Kolmogorov	$T = \max  G(x) - S(x) $	G est la fonction de répartition théorique et S est la fonction de répartition empirique. H0 = l'échantillon est issu de la distribution théorique. Sous H0, les limites de la zone de rejet $T_{1-\alpha}$ sont tabulées. Le test non paramétrique de Kolmogorov est très proche du test de Smirnov mais il compare un échantillon à une distribution théorique quelconque. Il est possible de définir une région de confiance autour de la fonction de répartition empirique à l'intérieur de laquelle toute répartition contenue entièrement dans cette région vérifie l'hypothèse nulle (Sprent, 1992).
Indice de dissimilarité	$\delta(H, H') = \left[ \sum_{i=1}^n \left( \frac{H_i - H'_i}{H_i} \right)^2 \right]^{1/2}$	H est la distribution observée et H' est la distribution prédite par le modèle. $H_i$ est le nombre d'arbres dans la ième classe de diamètre de la distribution H et n est le nombre de classes de diamètre. Voir une application dans Verzelen <i>et al.</i> (2006).
Indice de différence absolue relative entre les distributions	$DA = \frac{1}{2} \sum_{i=1}^n  s_{1,i} - s_{2,i} $	$s_{1,i}$ représente la fréquence de la classe i dans la distribution 1 et $s_{2,i}$ représente la fréquence de la même classe i dans la distribution 2. Cet indice représente la proportion d'individus qui devraient être échangés entre les classes pour que la distribution 1 devienne identique à la distribution 2. Si DA = 1 les deux distributions n'ont aucune classe commune et si DA = 0 les deux distributions sont identiques (Pommerening, 2002 ; Aguirre <i>et al.</i> , 2003 ; Pommerening, 2006 ; Ngo Bieng, 2007).

# ANNEXE F : Présentation des données d'évaluation par dispositif pour le chêne.

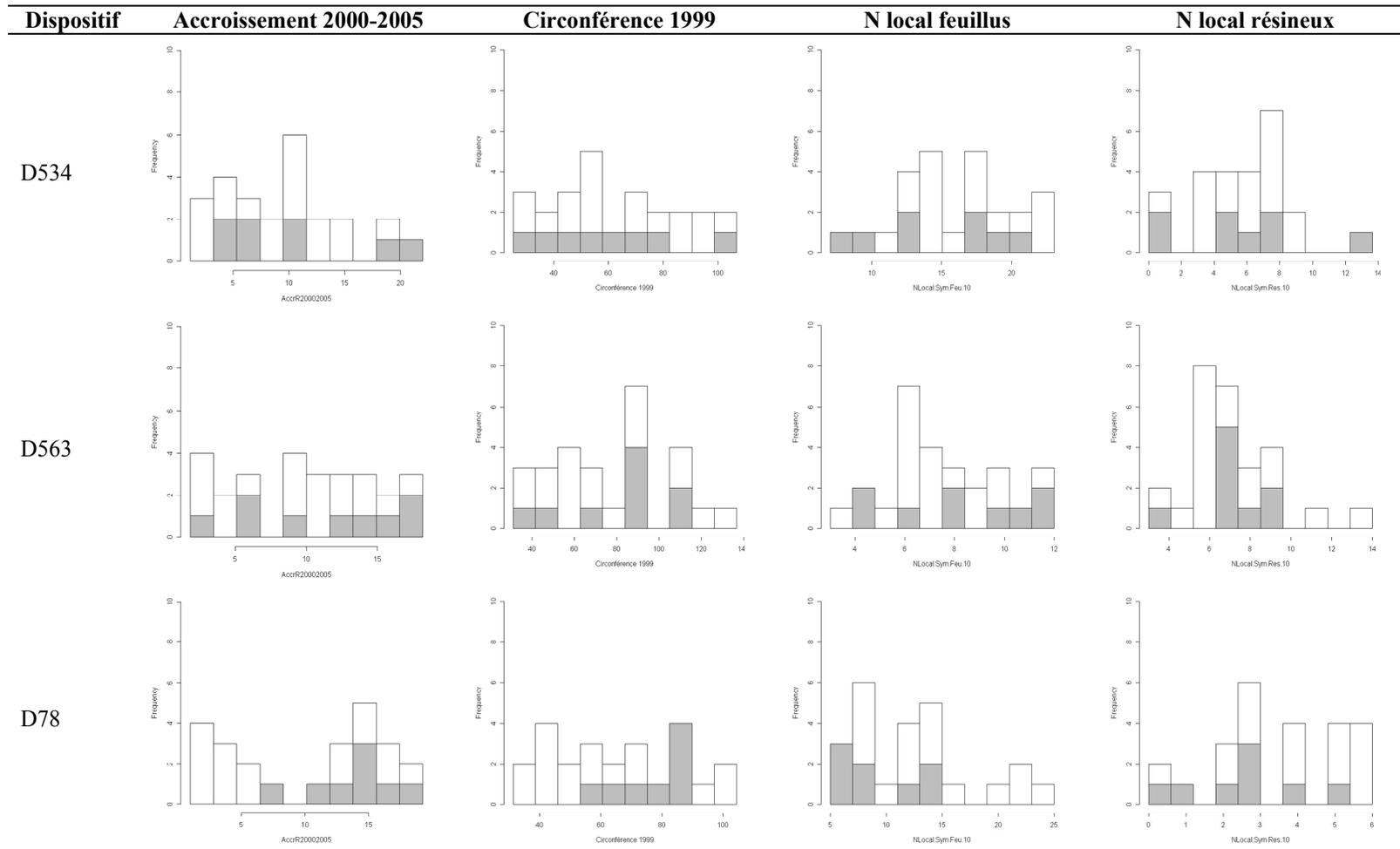
- Les données de calibration sont en blanc et les données d'évaluation sont en gris -



- Les données de calibration sont en blanc et les données d'évaluation sont en gris -

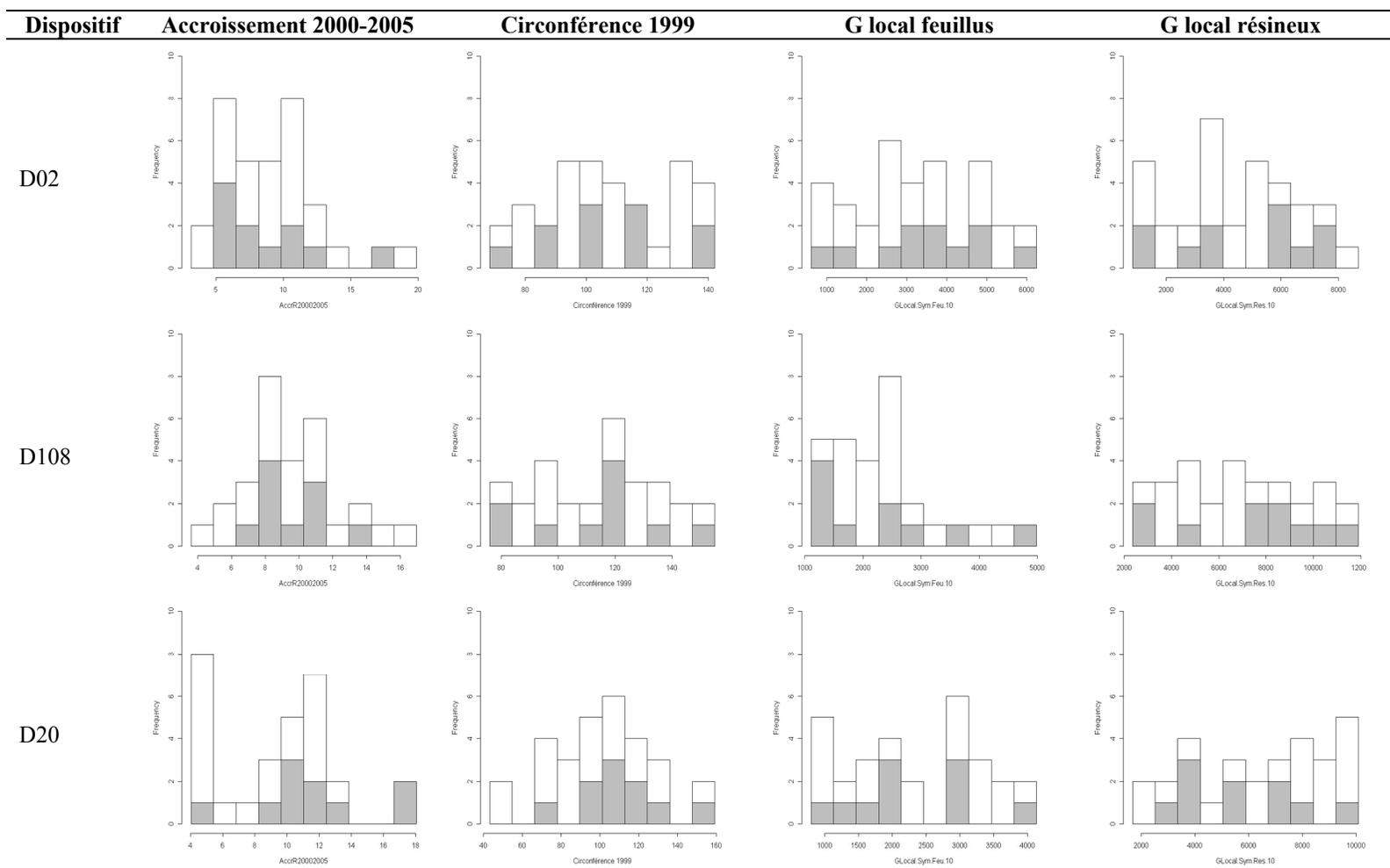


- Les données de calibration sont en blanc et les données d'évaluation sont en gris -

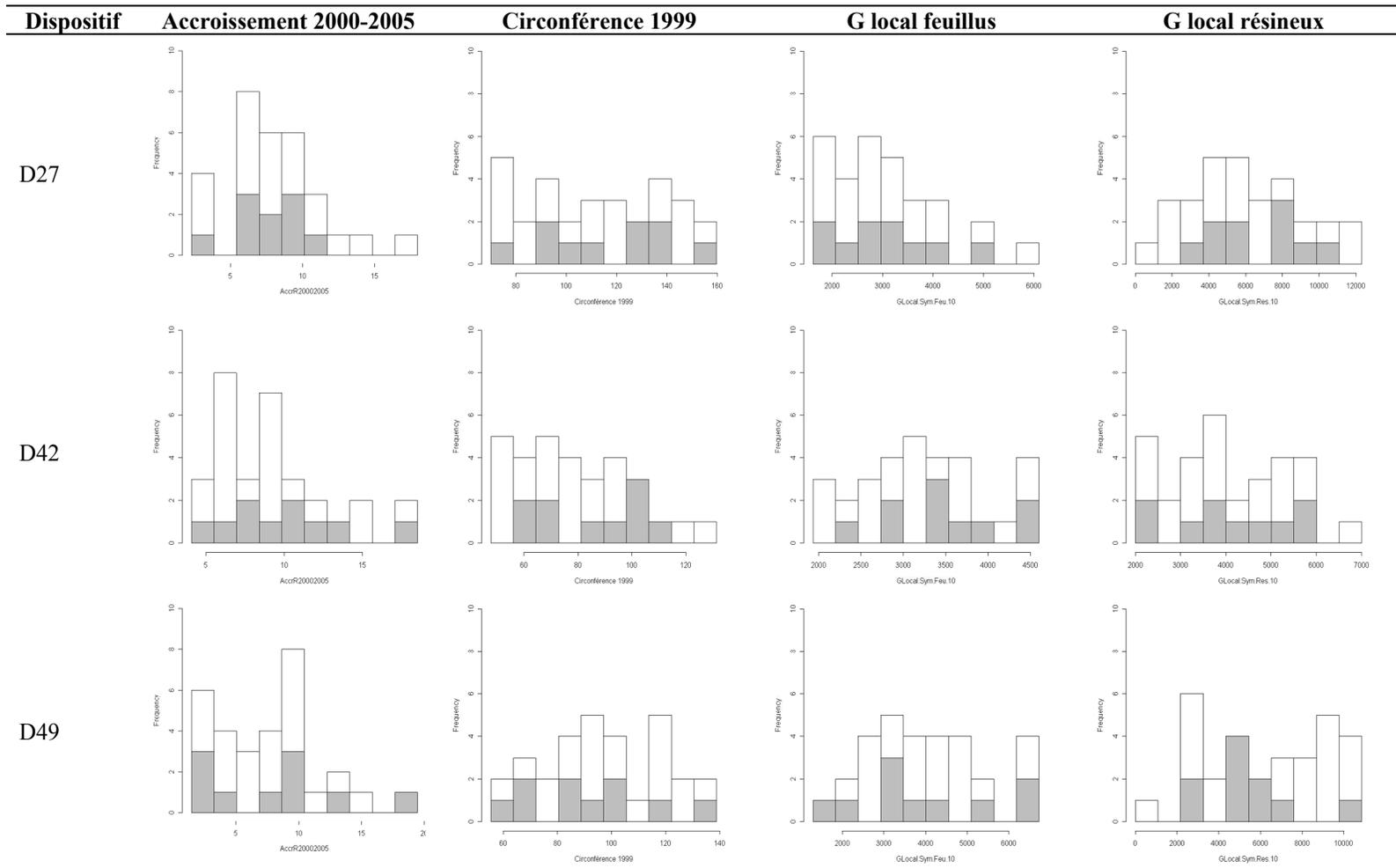


## ANNEXE G : Présentation des données d'évaluation par dispositif pour le pin.

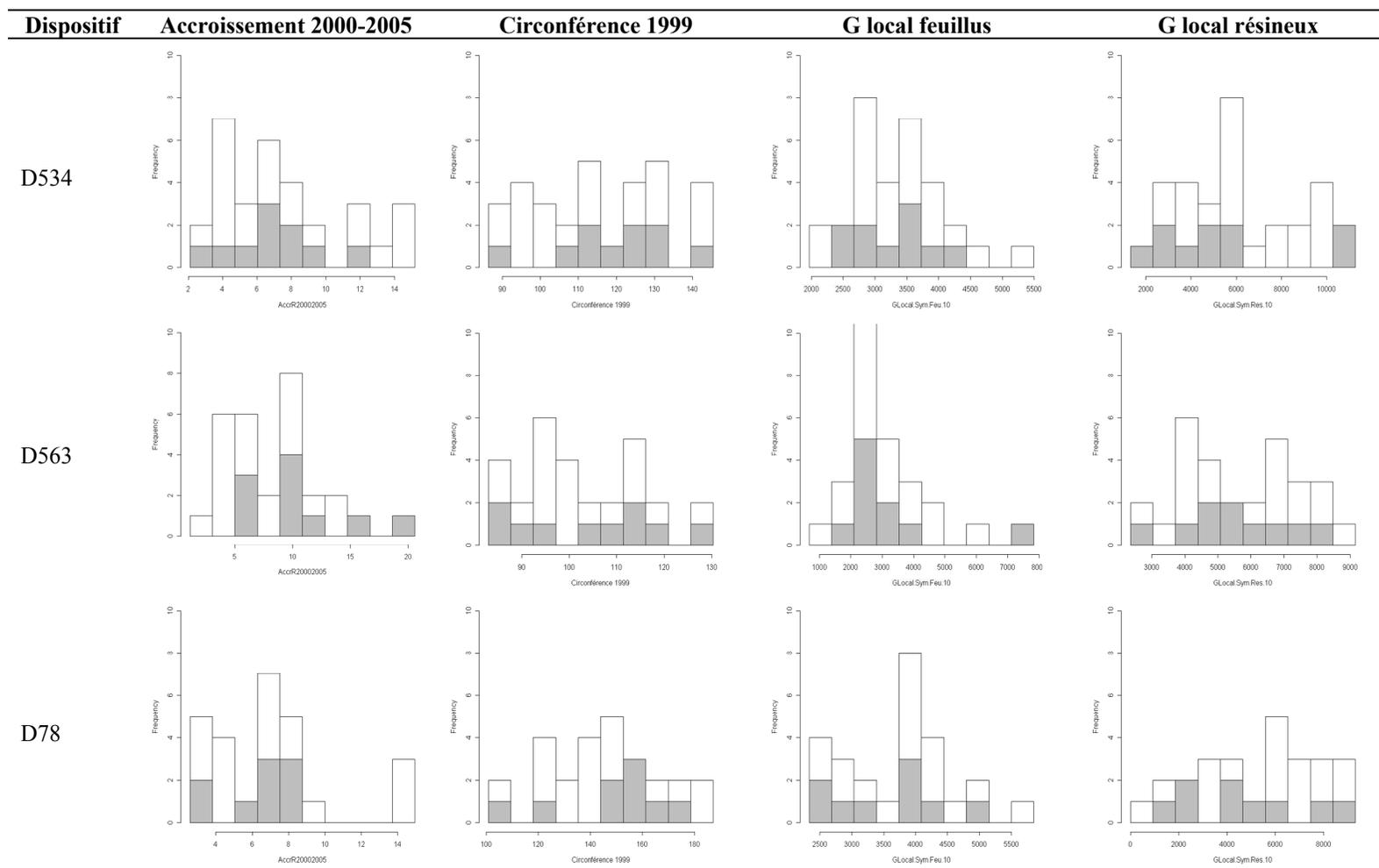
- Les données de calibration sont en blanc et les données d'évaluation sont en gris -



- Les données de calibration sont en blanc et les données d'évaluation sont en gris -



- Les données de calibration sont en blanc et les données d'évaluation sont en gris -



## ANNEXE H : Evaluation du MADD et du MAID

Relation entre les biais et les valeurs prédites et entre les biais et les variables explicatives.

