Table des matières

Remerciements

Je souhaite remercier Mr Jean Pierre Renaud, chargé de recherche au département R&D de l'ONF de m'avoir donné l'opportunité de faire ce stage.

Je remercie spécialement mes maîtres de stage, Mr Nicolas Picard , Mr Mathieu Fortin, Mr Jean Pierre Renaud et Mme Christine Deleuze, pour tous les précieux conseils qu'ils m'ont apportés tout au long de mon stage aussi bien dans le domaine des statistiques que dans le domaine forestier.

Je tiens également à remercier toute l'équipe de l'INRA pour leur accueil et leur gentillesse.

1. Introduction

1.1 Présentation de la structure d'accueil

L'Office National des Forêts (ONF) est un établissement public à caractère industriel et commercial. Il a été créé en 1964 et assure la gestion durable des forêts publiques françaises. Il gère 4,7 Mha de forêts et espaces boisés en métropole (27 % de la forêt française), ainsi que près de 6 Mha dans les départements d'Outre-mer (DOM). Il est donc le premier gestionnaire d'espaces naturels en France. Il commercialise environ chaque année 40 % des bois mis sur le marché en France. En 2010 par exemple, l'ONF a mobilisé 6.4 Mm3 de bois en forêts domaniales (hors DOM), ce qui représente un chiffre d'affaires de 221 M€. En forêts communales, ce sont 8.3 Mm3 qui ont été mobilisés, pour un chiffre d'affaires équivalent.

L'état confie à l'ONF quatre grandes missions d'intérêt général :

-La production de bois en conjuguant des exigences économiques, écologiques et sociales.

-L'accueil du public, l'information et la sensibilisation à l'environnement.

-La protection du territoire par la gestion des risques naturels.

-La protection de la forêt par la création de réserves naturelles et biologiques.

Ainsi, dans le cadre de ses missions, l'ONF mobilise du bois pour la filière et assure le renouvellement des forêts publiques. Il gère ces forêts de façon durable pour préserver la biodiversité et crée également des réserves biologiques. En région de montagne, l'ONF assure des missions de service public pour la prévention et la gestion des risques naturels. Enfin, dans le contexte actuel de changements globaux, l'ensemble des missions qu'assure l'ONF est réalisé dans un souci de dynamiser le rôle de la forêt au service de la lutte contre les changements climatiques

Au centre de ses outils de gestion, l'ONF élabore des **guides de sylviculture** qui servent de références aux gestionnaires dans leurs différentes interventions (Deleuze *et al*. 1996). Ils s'appuient généralement sur des **modèles de croissance** et de productivité forestière issus de placettes expérimentales de grandes dimensions établies à travers la France dans des peuplements homogènes. Un partenaire majeur de l'ONF dans l'élaboration de ces modèles est l'INRA qui a établi de telles placettes de suivi de la croissance dans différentes forêts depuis plus de 50 ans.

1.2 Quelques notions forestières

1.2.1 Peuplement forestier

Un peuplement forestier se décrit comme une superficie continue sur laquelle la forêt a des caractéristiques relativement homogènes en matière de composition et de dimension d'arbres. Le plus souvent, la dénomination du peuplement reprend ces éléments. On parlera donc d'un peuplement mélangé de hêtre et de chêne de 90 ans ou d'un peuplement pur de sapin de 120 ans. Les peuplements sont souvent identifiés à partir de photographies aériennes.

1.2.2 Surface terrière et diamètre

Pour un arbre, la surface terrière représente la section transversale de son tronc à 1,3 m de hauteur, une hauteur qui représente une norme dans le domaine forestier et à laquelle on se réfère souvent sous le vocable « hauteur de poitrine ». Ainsi, le diamètre à hauteur de poitrine (DHP) est le diamètre de l'arbre à 1,3 m de hauteur ou son diamètre à hauteur à hauteur de poitrine (DHP) est le diamètre de l'arbre à 1,3 m de hauteur.

A l'échelle du peuplement, la surface terrière est la surface que représenteraient toutes les sections transversales des arbres de ce peuplement si on les coupait à 1,3 m de hauteur. Cette variable donne une indication de la densité du peuplement. Comme les peuplements ont des superficies variables, la surface terrière est souvent exprimée sur la base d'un hectare de forêt. On parle donc de surface terrière en m^2/ha. D'un point de vue statistique, et en assimilant la forme des troncs d'arbres à un cylindre, la surface terrière est proportionnelle au moment non-centré d'ordre 2 de la distribution des diamètres des arbres (le coefficient de proportionnalité valant $N\pi/4$, où N est la densité d'arbres à l'hectare).

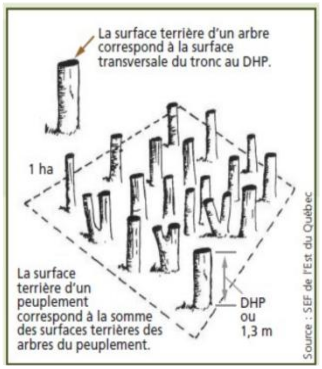

FIG. 1 : graphique illustrant la surface terrière d'une placette-échantillon

1.2.3 Inventaire forestier

L'inventaire forestier permet d'estimer certains paramètres des peuplements forestiers ainsi que d'en suivre l'évolution. Dans la mesure où les forêts sont très vastes pour être inventoriées en plein, l'inventaire forestier repose généralement sur un plan d'échantillonnage aléatoire ou aléatoire stratifié.

Dans le plan d'échantillonnage aléatoire stratifié, la population est partitionnée en sous-populations appelées strates dans lesquelles des placettes-échantillons sont distribuées aléatoirement et de façon indépendante.

1.2.4 Strate

Dans un contexte de gestion forestière, il est impensable que le peuplement soit à la base de la stratification forestière. On regroupe donc les peuplements semblables en un nombre plus réduits d'entité que l'on nomme « strates forestières ». Les strates sont définies comme des regroupements de peuplements semblables en ce qui a trait à leur composition, leur densité de couvert et leur âge. Par exemple, on parlera d'une strate forestière de peuplements jeunes de chêne et de hêtre en mélange avec une forte densité de couvert. La surface d'une strate forestière n'est pas nécessairement continue. Il s'agit de l'unité de base dans la gestion des forêts à l'échelle régionale.

1.2.5 Placette-échantillon

Une placette est une unité d'observation dont la surface est généralement inférieure à 0,1 ha et dans laquelle on dénombre les arbres et on mesure certaines de leurs caractéristiques comme le diamètre à hauteur de poitrine DHP. Ces placettes peuvent être temporaires lorsqu'elles ne sont mesurées qu'une fois ou permanentes lorsqu'elles font l'objet de remesures.

1.2.6 Modèles de croissance forestière

En fonction de leur composition, du climat, ou de la richesse des sols, les forêts croissent à des rythmes différents. Ces rythmes dictent la fréquence des interventions que doivent effectuer les gestionnaires forestiers (comme par exemple les éclaircies ou les récoltes). Pour les aider dans cette tâche, les forestiers ont conçu des guides sylvicoles, leur permettant de planifier ces travaux. Pour élaborer ces guides, l'ONF s'appuie sur des modèles de croissance ajustés aux comportements des principales essences forestières.

Un modèle de croissance forestière est un ensemble d'équations mathématiques correspondant aux relations entre les caractéristiques dendrométriques (caractéristiques physiques et quantifiables des arbres) d'une placette-échantillon et les dimensions (ainsi que l'âge, le plus souvent) des arbres qui le constituent. Un tel modèle permet de représenter et de prévoir l'évolution des arbres et du peuplement au cours du temps. Différents types de modèles de croissance peuvent être distingués selon leurs fonctionnements et l'échelle à laquelle ils s'appliquent.

1.2.7 Plate-forme Capsis

Capsis est une plate-forme de développement de modèles de croissance et de dynamique forestière. Il permet de construire et d'évaluer des scénarios sylvicoles en s'appuyant sur un modèle pour une essence et une région donnée (fertilité du site, densité initiale, intensité, type et nature des éclaircies, élagage…). Il est utilisé par des chercheurs pour évaluer leurs modèles et par des gestionnaires forestiers pour aider à l'élaboration de guides sylvicoles et pour l'enseignement.

1.3 Contexte et objectifs

Les peuplements forestiers évoluent sur de larges pas de temps et sur de grandes surfaces qui rendent difficiles l'expérimentation et la prévision de la production au cas par cas. La modélisation et la simulation sont donc souvent mises à contribution pour prédire la production et optimiser la gestion des forêts (*Vanclay, 1994 ; Pretzsch, 2009*). Dans la plupart des pays où la forêt est gérée durablement, les gestionnaires utilisent des modèles de croissance pour définir les lignes directrices

des guides de sylviculture des différentes essences ou établir les plans d'aménagement des forêts exploitées (*Favrichon, 1996*).

Les propriétés statistiques (biais, précision) des modèles de croissance sont importantes pour assurer la fiabilité des prédictions, donc de la gestion forestière. L'ajustement des modèles de croissance se fait le plus souvent par des techniques de régression qui assurent l'absence de biais de prédiction pour les données ayant servi à calibrer le modèle. Cependant, pour des données indépendantes de celles ayant servi à l'ajustement du modèle de croissance, un biais de prédiction a souvent été observé. Ce biais peut atteindre 20 % et remettre en cause l'intérêt de la modélisation pour la prévision de la production forestière (*Salas González, 1995*). Jusqu'à présent, les forestiers ont corrigé ce biais de manière empirique, en appliquant un coefficient multiplicateur correcteur *a posteriori*. Il serait cependant plus judicieux de comprendre l'origine de ce biais pour le corriger dès la phase d'ajustement du modèle de croissance.

Le biais de prédiction des modèles de croissance peut avoir plusieurs origines. Lorsque le modèle de croissance se compose de plusieurs sous-modèles qui sont ajustés séparément, un biais peut résulter du couplage des sous-modèles. Lorsque les données entrées dans le modèle ne respectent pas le protocole de celles qui ont servi à étalonner le modèle, un biais de prédiction peut aussi apparaître. Dans ce travail, nous nous focaliserons sur ce second phénomène. D'un point de vue statistique, le problème consiste à confronter des distributions différentes (mais de même espérance) pour les variables explicatives du modèle de croissance.

Dans la plupart des cas, les modèles de croissance ont des composantes non linéaires. En raison de ces relations non linéaires, des placettes de dimensions plus petites ou plus grandes que celles utilisées pour l'ajustement des modèles peuvent engendrer un biais de prédiction (*Salas Gonzales et al., 1993, 2001*). Changer la surface de la placette revient à changer la distribution des variables explicatives du modèle de croissance sans en modifier l'espérance. Lorsque le modèle de croissance est linéaire vis-à-vis de ses variables explicatives, cela n'a aucune incidence sur l'espérance de la prédiction. Mais dès lors que le modèle est non linéaire, cela entraîne un biais de prédiction. D'un point de vue biologique, l'effet de la surface de la placette sur la prédiction a une interprétation naturelle et renvoie à deux phénomènes biologiques : la variabilité stationnelle car une surface en gestion n'est pas toujours homogène, et la compétition entre arbres qui dépend de la densité locale (*Picard et al.,2009*).

Le changement de la surface des placettes peut renvoyer à deux cas de figure différents : (*i*) les placettes sont disjointes et distantes les unes des autres. D'un point de vue statistique, cela revient à considérer que les différentes placettes sont indépendantes entre elles. (*ii*) Les placettes de surface $|A'|$ sont obtenues par partition en *k* parts égales d'une grande placette de surface $|A| = k|A'|$. Dans ce cas, les placettes sont dépendantes, la structure de dépendance résultant de l'auto-corrélation spatiale des variables explicatives dans la grande placette.

L'objectif de cette étude est de comprendre l'origine du biais de prédiction dû à un changement de la surface des placettes dans un modèle de croissance forestière, et de fournir une expression (au moins de façon approchée) de ce biais. On parlera d'utilisation « correcte » (respectivement « incorrecte ») du modèle lorsque la surface des placettes est égale à (respectivement différente de) la surface des placettes utilisée pour la calibration du modèle. Le biais sera calculé comme la différence entre la prédiction dans le cas d'une utilisation « incorrecte » et la prédiction dans le cas

d'une utilisation « correcte ». Les résultats théoriques seront appliqués à deux modèles de croissance forestière :

(i) ARTEMIS, un modèle utilisé dans la province de Québec au Canada pour prévoir la croissance des forêts publiques *(Fortin et Langevin 2010, 2012)* et

(ii) le modèle Favrichon, un modèle utilisé pour prévoir la croissance de forêts tropicales en Guyane française (*Favrichon 1998*).

Alors que le modèle ARTEMIS utilise des placettes indépendantes, le modèle Favrichon utilise des placettes corrélées spatialement.

Le modèle ARTEMIS est utilisé pour prévoir l'évolution des strates forestières dans les forêts publiques du Québec. Les prédictions sont effectués par strate, une strate étant représentée par un échantillon de placettes de 400 m^2.

Traditionnellement, pour des raisons pratiques, les gestionnaires utilisent le modèle en agrégeant les placettes au sein d'une strate afin de créer une placette moyenne. Cet exercice que nous appellerons ici « agrégation *a priori* » revient à créer une grande placette dont la dimension est évidemment supérieure à celle initialement prescrite. Une utilisation correcte du modèle consisterait à effectuer une « agrégation *a posteriori* », soit à effectuer des projections pour chacune des placettes d'une strate et à en agréger les prédictions après la simulation. Cependant, on observe un biais entre les deux utilisations du modèle.

Le modèle matriciel Favrichon est utilisé en Guyane française. Il a été ajusté à partir de placettes d'une surface de $1{,}56625\ ha$. En pratique, les inventaires forestiers pratiqués par l'ONF utilisent des placettes de surface plus réduite (p .ex 0.07 ha), ce changement d'échelle pourrait entraîner des biais dans les prédictions du modèle.

Plus spécifiquement, nous avons tenté de quantifier le biais associé à une utilisation $\ll incorrecte \gg$ de ces deux modèles dans un premier temps, puis de déterminer, dans un second temps, les sources de biais, afin de pouvoir éventuellement les corriger et proposer des outils de simulation plus réalistes, comportant des intervalles de confiance de prédiction. Notre objectif n'était pas de décrire les modèles en détails, mais plutôt de quantifier les biais engendrés par des placettes de dimensions différentes (plus petites ou plus grandes) que prescrites originalement.

2. Estimation du biais de prédiction

2 .1 Définition d'une grandeur surfacique

Soit Z un processus ponctuel marqué dans $\mathbb{R}^2$, c'est-à-dire un processus aléatoire générant des nuages de points marqués $\{(x_1,\ m_1), (x_2,\ m_2), \dots, (x_N, m_n)\}$ dans tout sous-ensemble *A* de $\mathbb{R}^2$, où *N* désigne le nombre de points présents dans *A*, $x_i \in A$ désigne la position du i^e point, et m_i désigne la marque position du i^e point. A la fois le nombre *N* de points, les positions x_i des points et les marques m_i des points sont aléatoires. Dans l'application qui nous intéresse, *A* représentera une placette forestière, *N* le nombre d'arbres présents dans la placette, x_i les coordonnées spatiales de l'arbre i, et m_i le diamètre à hauteur de poitrine de l'arbre i .

Par définition, on appellera grandeur surfacique, notée *S(A),* une variable aléatoire positive calculée à partir de la restriction de *Z* à *A*.

Par exemple, la densité d'arbres présents dans la placette *A* est une variable surfacique définie par : $S(A) = \frac{N}{|A|}$ où $|A|$ représente la superficie de la placette *A*.

2.2 Espérance d'une prédiction

Soient A une placette-échantillon et $S_1(A), \dots\dots, S_p(A)$ des grandeurs surfaciques caractérisant une surface forestière, et $f: \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$ un modèle qui prédit une grandeur Y (la croissance par unité de surface par exemple) en fonction des variables explicatives $S_1(A), \dots\dots, S_p(A)$, et à l'aide de q paramètres collectivement notés θ :

$$Y = f(S_1(A), \dots\dots, S_p(A); \theta) \qquad (1)$$

Dans le cadre de cette étude, on considérera le modèle f comme connu, c'est-à-dire que l'on considérera le vecteur de paramètres θ comme fixe. Toute la variabilité des prédictions découlera donc de la distribution des variables explicatives $S_i(A)$. Ainsi, l'espérance de la prédiction

$$E(Y) = E(f(S_1(A), \dots\dots, S_p(A); \theta))$$

se calcule par rapport à la distribution des variables explicatives $S_i(A)$. Lorsque *A* varie, l'espérance de $S_i(A)$ ne change pas (c'est une propriété des variables surfaciques) tandis que la distribution de $S_i(A)$ change. Par conséquent, lorsque le modèle *f* est linéaire vis-à-vis de ses variables explicatives : $Y = a_1 S_1(A) + \cdots + a_p S_p(A)$ et $\theta = (a_1, \dots., a_p)$, l'espérance de *Y* ne change pas lorsque *A* varie. La question du biais de prédiction ne se pose donc que lorsque le modèle *f* est non linéaire.

On veut déterminer le biais de prédiction lorsque le calcul des grandeurs surfaciques $S_1(A), \dots\dots, S_p(A)$ ne respecte pas les règles d'utilisation du modèle, ce qui se produit lorsque les placettes ont des superficies différentes de celles ayant servi à l'ajustement ou lorsqu'on utilise une placette dite «moyenne».

Dans tous les cas, le biais de prédiction peut être estimé de façon exacte en utilisant une méthode de Monte Carlo : pour *M* tirages successifs, on échantillonne les variables explicatives $S_1(A), \dots\dots, S_p(A)$ dans leur distribution et, pour chaque tirage *m*, on calcule la prédiction Y_m du modèle de croissance selon l'équation (1). Lorsque *M* tend vers l'infini, la moyenne empirique $\sum_{m=1}^{M} Y_m/M$ converge vers l'espérance de *Y*. Dans le cas de placettes auto-corrélées spatialement (c'est-à-dire résultant de la partition d'une grande placette), l'échantillonnage des variables explicatives correspondra à un tirage des placettes au sein de la grande placette. On pourra alors réaliser un tirage avec ou sans remise.

Si le modèle f est dérivable, le biais de prédiction induit par un changement d'échelle entre les placettes expérimentales et les placettes de gestion peut être déterminé de manière approchée en utilisant un développement de Taylor au second ordre.

2.3 Expression analytique approchée du biais

On veut déterminer le biais de prédiction induit par un changement de la surface de la placette-échantillon, c'est-à-dire lorsque l'on passe d'une placette-échantillon de référence A à une placette A' comprise dans A (ou contenant A). D'après les propriétés des variables surfaciques (cf. § 2.1), les espérances $E(S_i(A))$ et $E(S_i(A'))$ sont égales

$$E\big[(S_1(A), \dots, S_p(A))\big] = E\big[(S_1(A'), \dots, S_p(A'))\big] \equiv \mathbf{a}$$

On a d'après le développement de Taylor au second ordre :

$\forall\, \mathbf{x} = (S_1(A), \dots, S_p(A))$ et $\mathbf{y} = (S_1(A'), \dots, S_p(A'))$:

$$Y = f(\mathbf{y}) \simeq f(\mathbf{x}) + (\mathbf{y}-\mathbf{x})\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^T} + \frac{1}{2}(\mathbf{y}-\mathbf{x})\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^T \partial \mathbf{x}}(\mathbf{y}-\mathbf{x})^\top$$

En utilisant la δ-méthode, l'espérance de Yest alors approximativement :

$$E(Y) \simeq f(\mathbf{x}) + \frac{1}{2}\, E[(\mathbf{y}-\mathbf{x})\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^T \partial \mathbf{x}}(\mathbf{y}-\mathbf{x})^\top]$$

$$E(Y) \simeq f(\mathbf{x}) + \frac{1}{2}\sum_{i=1}^{p} \mathrm{Var}[S_i(A)]\frac{\partial^2 f(\mathbf{x})}{\partial^2 x_i} + \sum_{i=2}^{p}\sum_{j<i} \mathrm{Cov}\big[S_i(A), S_j(A)\big]\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}$$

$$E(Y) \simeq f(\mathbf{x}) + \tfrac{1}{2}\mathbf{1}^T(\mathbf{\Sigma} \odot \mathbf{H}_Y)\mathbf{1} \qquad (2)$$

$$E(Y) - f(\mathbf{x}) \simeq \tfrac{1}{2}\mathbf{1}^T(\mathbf{\Sigma} \odot \mathbf{H}_Y)\mathbf{1} \qquad (3)$$

où $\mathbf{\Sigma}$ est la matrice de variance-covariance de $\mathbf{y}$, $\mathbf{H}_Y$ est la matrice hessienne de Y calculée en $\mathbf{x}$, $\mathbf{1}$ un vecteur colonne de longueur p dont les éléments sont tous des 1 et $\odot$ le produit de Hadamard, soit le produit terme à terme des matrices (*cf. Magnus et Neudecker, 2007, p.53*).

Pour obtenir une expression approchée du biais, il faut donc calculer la matrice hessienne $\mathbf{H}_Y$ et de Y ainsi que la matrice de variance-covariance $\mathbf{\Sigma}$ de $(S_1(A), \dots \dots \dots S_p(A))$. La matrice hessienne se calcule en dérivant deux fois la fonction f. Le calcul de Σ peut se faire de plusieurs façons, selon que l'on modélise ou non la distribution de $S_i(A)$ (approche paramétrique ou non paramétrique). Dans le cadre de cette étude, on remplacera Σ par la matrice de variance-covariance empirique de ($S_1(A)$, ..., $S_p(A)$). En la confrontant à l'estimation exacte du biais obtenue par Monte Carlo, on évaluera à quel point l'expression approchée (3) constitue une estimation satisfaisante du biais, et on identifiera les différentes sources de biais.

2.4 Estimation empirique du biais

Lorsque la fonction f n'est pas dérivable, le biais de prédiction pourrait être modélisé de façon empirique. Il s'agit ici de réaliser un grand nombre de simulation et de mettre le biais ainsi simulé en relation avec des caractéristiques des placettes. Toutefois, les données de croissance forestière ont toujours des problèmes de dépendance (par exemple lorsque les observations proviennent de

différentes dates successives $t \in \{1, \ldots . T\}$) et d'hétéroscédasticité (on peut penser que la variance des biais augmente avec le temps, par exemple).

Dans ce cas, un estimateur basé sur la méthode des moindres carrés ordinaires n'est pas approprié car la variance des estimations est biaisée de sorte que les tests de significativité sont faussés et on peut faire une mauvaise interprétation quant à la significativité des régresseurs. Cependant, il existe d'autres estimateurs qui permettent de tenir compte de la dépendance entre les observations. Pour prendre en compte à la fois la dépendance et l'hétéroscédasticité, on peut utiliser les moindres carrés généralisés.

2.4.1 Modèle linéaire

Un modèle linéaire est une hypothèse statistique qui peut s'écrire sous la forme $Y = \sum_{\mathbf{j=1}}^{\mathbf{k}} \mathbf{X_j}\, \boldsymbol{\beta_j} + \varepsilon$.

Y est une variable aléatoire que l'on observe et qu'on souhaite prédire, on l'appelle variable à expliquer.

Les k variables $X_1, \ldots, X_k$ sont des variables non aléatoires également observées, on les appelle variables explicatives.

Les β_j sont les paramètres associés aux variables explicatives, ils sont non observés et donc à estimer.

ε est le terme dans le modèle, c'est une variable aléatoire non observé.

Les hypothèses à vérifier pour valider le modèle sont les suivantes :

Les termes d'erreurs sont centrés
Les termes d'erreurs suivent une loi normale
Les termes d'erreurs sont indépendants
Les termes d'erreurs ont une variance constante

L'estimation des paramètre de ce modèle est basé un échantillon de n individus, on mesure $y_i, x_{i1}, \ldots, x_{ik}$ pour $i = 1, \ldots . n$ qui sont n observations des variables Y, $X_1, \ldots, X_k$.

Pour la i^e observation, le modèle s'écrit $y_i = \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i$

Les hypothèses à vérifier pour valider le modèle sont les suivantes :

Les termes d'erreur ε_i sont centrés
Les termes d'erreurs ε_i suivent une loi normale
Les termes d'erreurs ε_i sont indépendants
Les termes d'erreurs ε_i ont une variance constante

Posons $\mathbf{y} = \begin{pmatrix} y_1 \\ . \\ . \\ . \\ y_n \end{pmatrix}$, $\mathbf{X} = \begin{pmatrix} x_{11} & \dots & x_{1k} \\ & \vdots & \\ x_{n1} & \dots & x_{nk} \end{pmatrix}$ la matrice (n,k) de rang k, contenant les valeurs observés des k variables explicatives, $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ . \\ . \\ . \\ \beta_k \end{pmatrix}$ et $\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}$

On peut écrire le modèle sous la forme matricielle suivante :

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$$

Un estimateur de β est obtenu en utilisant la méthode des moindres carrés ordinaires qui consiste à minimiser la somme des carrés résiduels SCR.

$$\underset{\beta}{Min}[\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]^\top[\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]$$

$$SCR = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1} - \quad \dots \quad - \beta_k x_{ik})^2 = \|\mathbf{y} - \mathbf{X}\beta\|^2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

$$= \mathbf{y}^\top\mathbf{y} - \mathbf{2}\boldsymbol{\beta}^\top\mathbf{X}^\top\mathbf{y} + \boldsymbol{\beta}^\top\mathbf{X}^\top\mathbf{X}\boldsymbol{\beta}$$

La dérivée par rapport à β est alors égale à :

$$-2\,\mathbf{X}^\top\mathbf{y} + 2\,\mathbf{X}^\top\mathbf{X}\boldsymbol{\beta}$$

Afin de minimiser la SCR, on cherche la valeur $\hat{\beta}$ pour laquelle la dérivée est égale à 0. Donc on doit résoudre l'équation suivante :

$$\mathbf{X}^\top\mathbf{y} = \mathbf{X}^\top\mathbf{X}\boldsymbol{\beta}$$

Si $\mathbf{X}^\top\mathbf{X}$ est inversible, c'est-à-dire qu'aucune des colonnes qui compose cette matrice n'est proportionnelle aux autres alors, l'estimateur de β est donné par :

$$\hat{\beta} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y}$$

Si $\boldsymbol{\varepsilon}$ suit une loi multinormale qui vérifie les hypothèses :

$$E(\boldsymbol{\varepsilon}) = 0$$

$$Var(\boldsymbol{\varepsilon}) = \sigma^2 I_n$$

Alors :

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$$

$$Var(\hat{\beta}) = \sigma^2(\mathbf{X}^\top\mathbf{X})^{-1}$$

On dit que $\widehat{\boldsymbol{\beta}}$ est le meilleur estimateur non biaisé (BLUE) pour $\boldsymbol{\beta}$.

2.4.2 Moindres carrés généralisés

Lorsque la variance des observations n'est pas constante, il suffit, pour se ramener au modèle standard de régression linéaire, de minimiser la quantité $SCR_w = (\mathbf{y} - \mathbf{X}\beta)^\top(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/\sigma_i^2$ où σ_i^2 est la variance de l'observation i.

Cette méthode de calcul des paramètres du modèle s'appelle la méthode des moindres carrés pondérés. Elle consiste donc encore à minimiser la somme des carrés résiduels, mais où chacun des termes de la somme est pondéré par l'inverse de la variance correspondante de l'erreur.

Lorsque l'hypothèse d'indépendance des erreurs est à son tour violée dans le modèle $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, mais que l'on connait la structure de covariance des termes d'erreurs, alors l'estimation par la méthode des moindres carrés ordinaires n'est pas appropriée. En effet, soit la matrice définie positive **V**, telle que :

$Var(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{V}$ **,** $\boldsymbol{V}$ connue, $\sigma^2 > 0$ pas nécessairement connue.

Si on utilise $\hat{\beta} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y}$ l'estimateur par la méthode des moindres carrés ordinaires alors

$E(\widehat{\boldsymbol{\beta}}|\mathbf{X}) = \boldsymbol{\beta}$ et $Var(\widehat{\boldsymbol{\beta}}|\mathbf{X}) = \sigma^2(\mathbf{X}^\top\mathbf{X})^{-1}V\mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}$

On aura un estimateur sans biais de β mais dont la variance explose, ce qui entraînerait une perte d'efficacité, des intervalles de confiances trop larges et une mauvaise évaluation de la précision de $\widehat{\boldsymbol{\beta}}$.

On introduit un autre estimateur appelé estimateur des moindres carrés généralisés qui est un cas particulier des estimateurs par la méthode du maximum de vraisemblance.

Comme **V** est définie positive cela implique qu'il existe une matrice $\mathbf{V^{1/2}}$ telle que

$$\mathbf{V^{1/2}V^{1/2}} = \mathbf{V}$$

On a le modèle

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Ce qui implique que : $\mathbf{V^{-1/2}y} = \mathbf{V^{-1/2}X}\boldsymbol{\beta} + \mathbf{V^{-1/2}}\boldsymbol{\varepsilon}$ où $\mathbf{V^{-1/2}} = (\mathbf{V^{1/2}})^{-1}$

En posant $\mathbf{y}_V = V^{-1/2}\mathbf{y}$, $\mathbf{X}_V = V^{-1/2}\mathbf{X}$ et $\boldsymbol{\varepsilon}_\mathbf{V} = \mathbf{V^{-1/2}}\boldsymbol{\varepsilon}$, On obtient le modèle suivant :

$\mathbf{y_V} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_\mathbf{V}$ Avec $Var(\boldsymbol{\varepsilon}_\mathbf{V}) = \sigma^2 I_n$

Utilisons les moindres carrés ordinaires dans ce modèle

On a $\widehat{\boldsymbol{\beta}}_\mathbf{G} = (\mathbf{X}_V^\top\mathbf{X}_V)^{-1}\mathbf{X}_V^\top\mathbf{y}_V = (\mathbf{X}^\top\mathbf{V^{-1}X})^{-1}\mathbf{X}^\top\mathbf{V^{-1}y}$

On peut vérifier que $E(\widehat{\boldsymbol{\beta}}_\mathbf{G}|\,\mathbf{X}) = E((\mathbf{X}^\top\mathbf{V^{-1}X})^{-1}\mathbf{X}^\top\mathbf{V^{-1}y}|\mathbf{X}) = (\mathbf{X}^\top\mathbf{V^{-1}X})^{-1}\mathbf{X}^\top V^{-1}E(\mathbf{y}|\mathbf{X}) = \beta$ et que

$$Var(\widehat{\boldsymbol{\beta}}_{\mathbf{G}}|\mathbf{X}) = Var((\mathbf{X}^{\top}V^{-1}\mathbf{X})^{-1}\,\mathbf{X}^{\top}V^{-1}\mathbf{y}\,|\,\mathbf{X}) =$$
$$\left(\mathbf{X}^{\top}\mathbf{V^{-1}}\mathbf{X}\right)^{-1}\mathbf{X}^{\top}V^{-1}\,Var(\mathbf{y}|\mathbf{X})V^{-1}\mathbf{X}^{\top}(\mathbf{X}^{\top}V^{-1}\mathbf{X})^{-1} = \sigma^2(\mathbf{X}^{\top}V^{-1}\mathbf{X})^{-1}$$

L'estimateur par la méthode des moindres carrés généralisés est donc sans biais et de variance minimale parmi tous les estimateurs non biaisés qui sont linéaires en les observations. Il est également possible d'obtenir mathématiquement l'expression de son niveau de précision, de sa matrice de covariance : $Var(\hat{\beta}_G|\mathbf{X}) = \sigma^2(\mathbf{X}^{\top}V^{-1}\mathbf{X})^{-1}$.

L'enjeu consiste donc à paramétrer la matrice de covariance des erreurs par le biais d'une structure de corrélation et/ou une fonction de variance. Il existe plusieurs façons de modéliser cette corrélation.

Par exemple lorsqu'on a des données en cluster ou données groupées, on peut considérer que les erreurs sont distribuées suivant une moyenne mobile. Le terme d'erreur à la date t subit donc l'influence de la réalisation à la période précédente du bruit blanc ε et lorsqu'on a des données longitudinales, la méthode la plus fréquemment utilisée consiste à supposer que les erreurs sont distribuées suivant un processus autorégressif. Le terme d'erreur ε_t à la date t subit alors l'influence de sa propre réalisation à la période précédente, ε_{t-1} .

2.5 Discussion

Le fait de changer la surface de la placette dans les modèles de croissance forestière induit bien un biais de prédiction dès lors que le modèle de croissance est non linéaire. De façon intuitive, ce biais peut se comprendre comme le fait que l'espérance de $f(X)$ où f est une fonction non linéaire et X une variable aléatoire n'est pas égale à l'image par f de l'espérance de X. L'expression approchée du biais de prédiction obtenue dans le paragaphe 2.3 a plusieurs intérêts. Premièrement, elle permet de prédire le signe du biais, c'est-à-dire si la croissance est sous-estimée ou surestimée lorsque la surface des placettes change. Le signe du biais dépendra de la dérivée seconde de f, c'est-à-dire de sa convexité. Deuxièmement, l'expression approchée du biais permet de décomposer ce biais en différentes contributions liées aux variances et covariances des variables explicatives surfaciques. On pourra ainsi identifier, parmi les différentes variables explicatives, lesquelles contribuent le plus au biais de prédiction. Troisièmement, l'expression approchée du biais montre que l'écart entre l'espérance de la prédiction et $f(X)$ est proportionnel aux variances et covariances des variables explicatives. On s'attend donc à ce que cette expression approchée du biais soit d'autant meilleure que les variances et covariances des variables explicatives sont faibles. Comme ces variances et covariances décroissent avec la surface des placettes, on s'attend donc à ce que l'expression approchée du biais soit d'autant meilleure que les placettes sont grandes.

Le travail qui a été fait dans cette étude pourrait être étendu de diverses manières. Tout d'abord, on pourrait chercher à calculer une meilleure expression approchée du biais en poussant le développement de Taylor de la fonction f au-delà du terme de degré 2. En pratique, les calculs deviennent cependant vite inextricables.

Ensuite, on pourrait remplacer l'estimation empirique de la matrice de variance-covariance Σ de $S_1(A), \ldots\ldots, S_p(A)$ par une expression paramétrique. Cela revient à modéliser la variance-covariance des variables explicatives. Cette modélisation entraînerait un écart par rapport aux observations, mais ferait l'économie de ces observations. Un cadre mathématique approprié pour

développer un modèle pour la variance-covariance de $S_1(A), \ldots\ldots, S_p(A)$ est celui de la théorie des processus ponctuels (*Cressie, 1993 ; Stoyan and Stoyan, 1994 ; van Lieshout, 2000 ; Møller and Waagepetersen, 2004 ; Illian et al., 2008).* Lorsque le processus ponctuel qui génère les positions des arbres au sein des placettes est connu de façon explicite, on peut généralement calculer l'expression explicite de $Var(S_i(A)$ (*Stoyan and Stoyan, 1994* ; *Picard & Favier, 2011*). Les calculs peuvent toutefois s'avérer laborieux. De plus, dans la plupart des applications pratiques, le processus ponctuel générateur du nuage de points observé n'est pas connu. Les agronomes et les forestiers, depuis Fairfield Smith (1938), ont alors utilisé un modèle empirique de la forme puissance :

$$Var\big(S_i(A')\big) = \left(\frac{|A|}{|A'|}\right)^{\beta_i} \times Var(S_i(A))$$

qui s'est avérée être une bonne approximation de la vraie variance (Wagner et al., 2010). Dans la plupart des applications pratiques en foresterie, on observe de l'autocorrélation spatiale positive, ce qui correspond à $0 < \beta_i < 1$. Le cas $\beta_i = 1$ correspondrait à l'absence d'autocorrélation spatiale.

Une troisième extension du présent travail consisterait à considérer des variables explicatives $S_i(A)$ qui ne sont pas forcément surfaciques. Le diamètre maximal des arbres dans une placette, ou plus généralement tout quantile calculé sur les arbres de la placette, constituent des variables non surfaciques. Ces variables explicatives non surfaciques ne constituent pas un exercice de style mais sont réellement utilisées par les forestiers. La hauteur dominante, définie comme la hauteur moyenne des 100 plus grands arbres du peuplement, est ainsi fréquemment utilisée comme variable explicative dans les modèles de croissance forestière. Ces variables non surfaciques posent une difficulté supplémentaire car, contrairement aux variables surfaciques, leur espérance varie lorsque la surface $|A|$ de la placette varie (*Pierrat et al., 1995*). Elles induisent donc directement un biais de prédiction, même lorsque le modèle de croissance *f* est linéaire vis-à-vis de ses variables explicatives.

Enfin, une quatrième extension du présent travail consisterait à considérer le vecteur θ des paramètres du modèle de croissance non plus comme fixe, mais comme un vecteur aléatoire résultant de l'ajustement du modèle aux données de calibration. Remplacer dans l'équation (1) θ par son estimateur $\hat{\theta}$ permettrait de régler directement la question de la correction du biais au niveau de l'ajustement du modèle, plutôt qu'*a posteriori* sur les prédictions.

3. Application à la dynamique forestière

En application de la question théorique développée dans la section 2, on va utiliser deux modèles différents.

Ces modèles sont les suivants :

(i)ARTEMIS, un modèle utilisé dans la province de Québec au Canada pour prévoir la croissance des forêts publiques et

(ii) le modèle Favrichon, un modèle utilisé pour prévoir la croissance de forêts tropicales en Guyane française.

3.1 Modèle ARTEMIS

3.1.1 Description du modèle

Le modèle comprend quatre sous-modèles qui permettent de prédire l'évolution d'une placette de 400 m2 sur une période de 10 ans. Les deux premiers sous-modèles concernent les arbres présents au début de l'intervalle. Ils permettent de prédire la probabilité de mortalité de chaque arbre et son accroissement en diamètre s'il ne meurt pas. Les deux autres sous-modèles concernent le recrutement de nouveaux arbres au cours de l'intervalle. Ils permettent de prédire le nombre de nouveaux arbres de chaque essence qui franchissent les 9,1 cm à 1,3 m de hauteur et leurs diamètres. Pour obtenir des prédictions sur plus de 10 ans, les projections sont réinsérées dans le modèle.

3.1.2 Description des données

La forêt publique québécoise est constituée d'unité d'aménagement forestier (UAF). Chaque UAF comprend un nombre de strates dans lesquelles on établit des placettes-échantillons afin d'en évaluer les paramètres, notamment la surface terrière à l'hectare et le nombre d'arbres à l'hectare. Ces placettes sont distribuées selon un plan de sondage aléatoire stratifié. Elles sont circulaires et chacune couvre une surface de $400m^2$. A l'intérieur des limites de la placette, tous les arbres dont le DHP est plus grand ou égal à 9.1 cm sont numérotés et identifiés. L'essence de ces arbres est notée et leur DHP est mesuré.

Dans le cas présent, nous utiliserons les placettes échantillons de l'UAF 6152 dont les caractéristiques sont brièvement présentées dans le tableau suivant.

Strate	Nbres de placettes	Surface terrière (m²/ha)		
		Min	Max	Moyenne
1	32	15.8	42.0	28.8
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
251	15	8.0	15.1	20.0

Tableau1 récapitulatif des données

Ces placettes sont en tous points conformes à celles qui ont servi à étalonner le modèle ARTEMIS. Toutefois, lors des simulations, on compte parfois plus de 600 strates pour un territoire, chacune de ces strates comptant au minimum 5 placettes. Réaliser des simulations pour chacune des placettes peut prendre un certain temps. Par le passé, les gestionnaires forestiers ont pris l'habitude de ne réaliser qu'une seule simulation par strate en utilisant la moyenne des placettes en entrée du

modèle. Cette agrégation a priori est susceptible d'engendrer des biais dans l'estimation des paramètres d'une strate comparativement à la méthode originale qui consiste à agréger les prédictions a posteriori.

3.1.3 Agrégation *a priori*

L'échantillon de chaque strate comprend généralement un minimum de 5 placettes-échantillons.

Dans ce rapport, faire une agrégation *a priori* d'une variable sur une strate donnée consiste à mesurer la variable sur toutes les placettes-échantillons de cette strate, et à en calculer la moyenne. La valeur attribuée à la strate est donc la valeur moyenne de ses placettes-échantillons pour cette variable.

Comme l'illustre la figure 2, on applique la simulation à une placette moyenne de la strate.

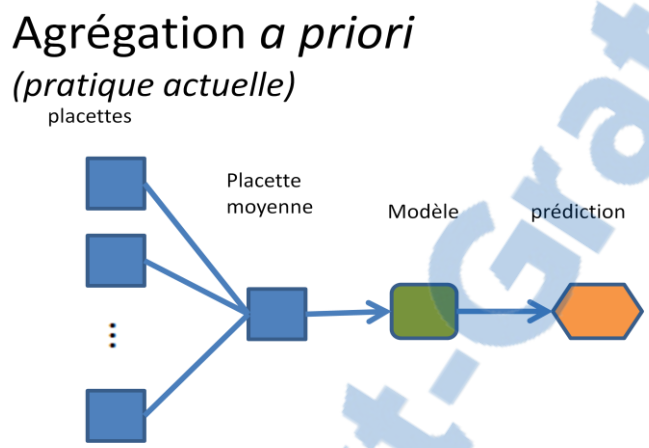

FIG. 2 Représentation schématique d'une agrégation dite « a priori ».

3.1.4 Agrégation *a posteriori*

Faire une agrégation *a posteriori* revient à utiliser les valeurs individuelles des placettes comme variables de départ aux simulations, puis de calculer la moyenne des prédictions obtenues. Contrairement à l'agrégation *a priori*, la moyenne est effectuée ici après que la simulation ait été effectuée sur chacune des placettes. La figure 3 illustre cette façon de faire.

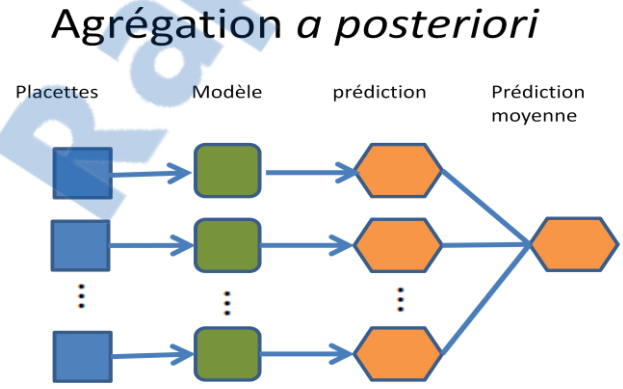

FIG. 3 Représentation schématique de l'agrégation « a posteriori »

3.1.5 Estimation du biais

On va essayer d'estimer le biais de prédiction du nombre d'arbres à l'hectare qui est égal à la différence entre la prédiction obtenue avec une agrégation *a priori* des données et la prédiction obtenue avec une agrégation *a posteriori*.

Pour cela, à l'aide du logiciel Capsis, on fait une simulation *a priori* et une simulation *a posteriori* de l'accroissement du nombre d'arbres par hectare au cours du temps dans chaque strate, on en déduit le biais observé qui est égale à la prédiction *a priori* moins la prédiction *a posteriori*.

Comme le montre la *figure 4* suivante pour la majorité des strates, il existe un biais négatif assez important, donc on a tendance à sous-estimer le nombre d'arbres en faisant des agrégations *a priori*.

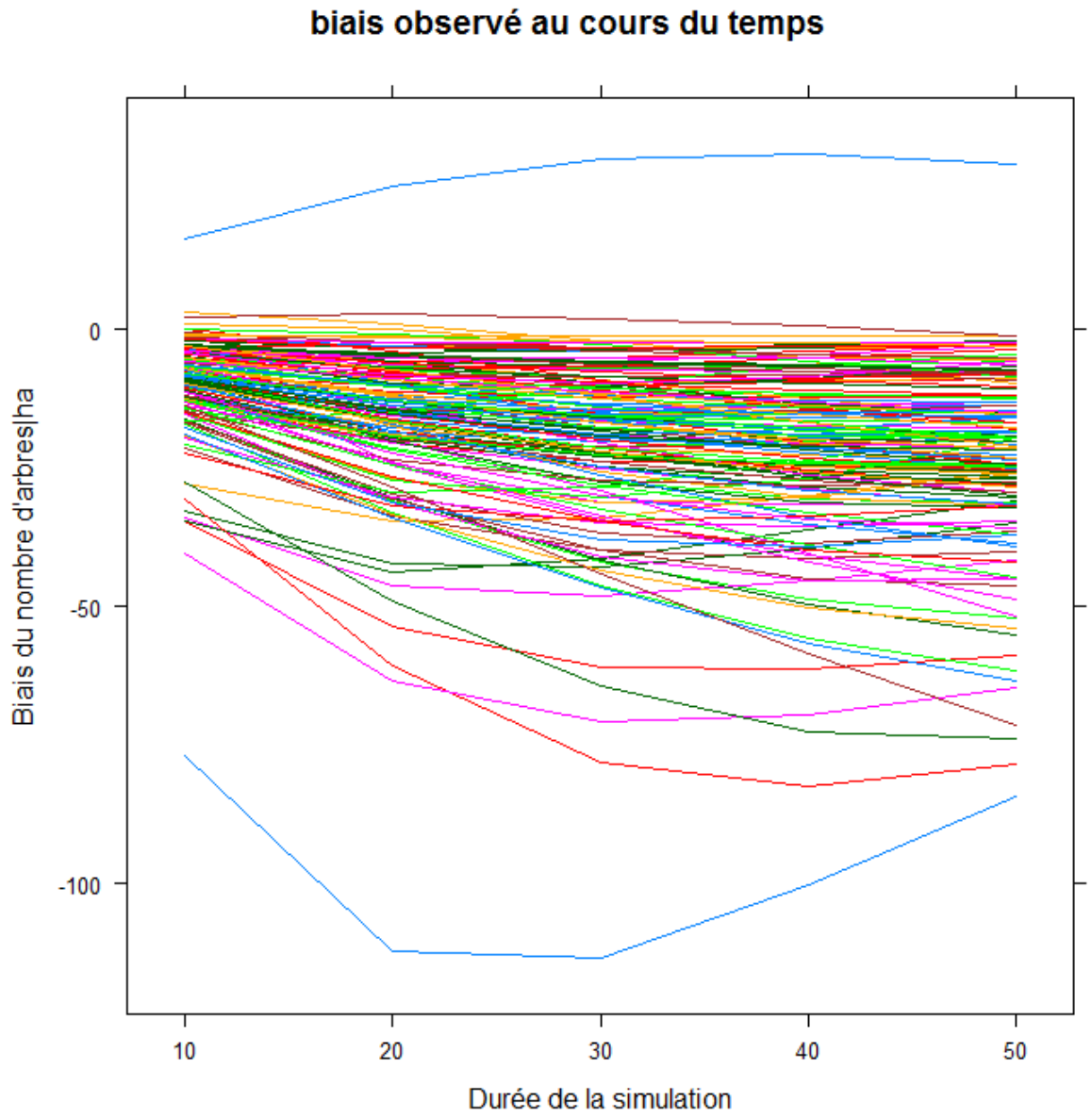

FIG. 4 – Biais observé pour chaque strate au cours du temps

Une fois qu'on a déterminé le biais de prédiction, il devient intéressant de pouvoir le prévoir.

On désigne par Y_{ij} le biais observé entre les deux prédictions pour la strate i au bout d'un pas de temps *j*. L'un des modèles à effets fixes les plus connus est le modèle linéaire.

Supposons que le biais observé dépende du temps (*i.e. de la durée de la période de croissance simulée*), du nombre de placettes et de la variance de la surface terrière initiale. En effet, on peut imaginer que pour une strate donnée, plus la variance de la surface terrière est grande, plus les placettes-échantillons qui la constituent sont hétérogènes. Or, plus il y a homogénéité entre les placettes-échantillons qui constituent une strate, plus les résultats des simulations entre placettes individuelles seront similaires. Ce qui devrait conduire à des estimations très proches de celles d'une placette moyenne et donc à de faible biais. On peut donc supposer qu'il existe un effet de la variance terrière initiale sur le biais de prédiction.

D'autre part, les agrégations étant faites sur les placettes-échantillons et les prédictions faites de manière itérative au cours du temps, on peut également suspecter que le biais de prédiction sur une strate donnée augmente avec le nombre de placettes-échantillons sur cette strate et avec le temps.

Le modèle linéaire stipule qu'il y a linéarité entre la variable à expliquer et les variables explicatives, avec un terme d'erreur $\mathcal{E}_{ij}$; ce qui se traduit par l'équation suivante :

$Y_{ij}=\beta_0+\beta \times t_j+ \gamma \times NbPl_i+ \delta \times Var_i +\mathcal{E}_{ij}$

Où Y_{ij} est le biais observé entre les deux prédictions pour la strate i au bout de *j* pas de temps,

$NbPl_i$: Le nombre de placette dans la strate i et

Var_i : La variance de la surface terrière au temps zéro de la strate *i*.

t_j =10, 20, 30, 40, 50 ans = durée de la période de simulation

$\mathcal{E}_{ij}$: Terme d'erreur aléatoire

α , β, γ , δ, σ^2 : 5 paramètres à estimer.

En supposant que les $\mathcal{E}_{ij}$ sont indépendants, identiquement distribués et $\mathcal{E}_{ij} \sim \boldsymbol{N(0, \sigma^2)}$

Une des limites du modèle à effets fixes est qu'il suppose une indépendance entre les Y_{ij}. Or ici les Y_{ij} d'un même individu présentent un certain degré de corrélation (ou dépendance). En effet, pour chaque strate i, on détermine le biais observé pour cinq pas de temps. On a ainsi des mesures répétées pour chaque strate.

Dans ce cas, l'estimation par la méthode des moindres carrés ordinaires n'est pas appropriée. Pour prendre en compte la dépendance entre certains individus, on a alors le choix entre plusieurs approches possibles, comme nous l'avons déjà énoncé à la partie (**2.2**).

Nous allons d'abord tenter d'ajuster un modèle linéaire par la méthode des moindres carrés généralisés. De cette façon, nous prenons en compte la structure de dépendance associée aux données. Puisqu'on a des données longitudinales, on va supposer que les résidus ont une structure de corrélation autorégressif d'odre 1 (AR1). On suppose que la matrice de correlations associée aux observations pour une strate *i* est donnée par :

$$R_i(\alpha) = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^5 \\ \rho & 1 & \rho^2 & \dots & \rho^4 \\ & & \ddots & & \\ \rho^5 & \rho^4 & \rho^3 & \dots & 1 \end{bmatrix}$$

Durbin et Watson ont mis en place un test en 1950 pour détecter l'autocorrélation d'ordre 1. Les hypothèses sont très simples à établir : en effet si on suppose que le bruit est issu d'un processus AR(1), la dépendance des variables ne dépend que de ρ et on sait que l'indépendance est caractérisée par le fait que ρ = 0. D'où les hypothèses suivantes :

H0: ρ = 0

contre H1 : $\rho \neq 0$

La statistique de test appelée statistique de Durbin-Watson est calculée comme suit :

$$d = D - W = \frac{\sum_{t=1}^{n-1}\left(\hat{\varepsilon}_{t+1}^{MCO} - \hat{\varepsilon}_t^{MCO}\right)^2}{\sum_{t=1}^{n}(\hat{\varepsilon}_t^{MCO})^2}$$

Où le vecteur ε^{MCO} est le vecteur des résidus obtenus par la méthode des M.C.O et n le nombre d'observation .

La distribution de la statistique **d** oscille entre deux distributions $\boldsymbol{d_L}$ et $\boldsymbol{d_U}$, la règle de décision est la suivante :

$0 < d < \boldsymbol{d_L}$: autocorrélation positive

$\boldsymbol{d_L} < d < \boldsymbol{d_U}$: indétermination

$\boldsymbol{d_U} < d < \boldsymbol{4} - \boldsymbol{d_U}$: pas d'autocorrélation

$\boldsymbol{4} - \boldsymbol{d_U} < d < 4$: autocorrélation négative

Où $\boldsymbol{d_L}$ et $\boldsymbol{d_U}$ sont les valeurs critiques du test . Elles sont tabulées pour différentes valeurs de k (nombre de variables explicatives constante exclue) et de n (nombre d'observations).

Nous allons utiliser la fontion durbin.watson(…) dans le package *car* pour réaliser le test. Les résultats du test donné dans le tableau suivant nous permettent pas d'accepter l'hypothèse H0.

lag	Autocorrelation	D-W Statistic	p-value
1	0.05751394	1.884699	0.048

En effet, pour notre jeu de données, on a 1065 observations et 3 variables donc d'après le tableau 2 ci-dessous d=1.8844699 < $\boldsymbol{d_L}$.

Nombre de variables k=3		
n	d_L	d_U
1050	1.89475	1.90238
1100	1.89725	1.90454

Tableau2: Valeurs Critiques pour le Test de Durbin-Watson : Niveau de Significativité de 5 % (n=1050, 1100)

Le test de Durbin Watson ne permet pas d'accepter l'hypothèse H0, nous concluons qu'il y a une autocorrélation positive.

Nous pouvons alors ajuster notre modèle en utilisant la fonction gls du package nlme avec une structure de corrélation AR1, et à l'aide de la fonction summary(), nous avons les estimations des paramètres du modèle et leurs erreurs-type estimées dans le tableau ci-dessous.

	Value	Std.Error	t-value	p-value
intercept	-2.2578	0.5306	-4.2553	0.0000
NbPl	-0.0053	0.0183	-0.2894	0.7723
Var	-0.0415	0.0117	-3.5497	0.0004
Annee	-0.3754	0.0195	-19.2097	0.0000

Les tests de Student permettant de juger de l'influence de la variance initiale de la surface terrière Var et du temps t étant tous significatifs au seuil α=0.05. Nous décidons donc, avec un risque de première espèce de α=0.05 que les coefficients de régression associés aux variables Var et t sont non nuls. Leurs estimations sont : $\hat{\beta} = -0.3753699$, $\hat{\delta} = -0.0415229$

Le test de Student nous permet également de conclure de la significativité de la constante β_0 au seuil α=0.05. Donc avec un risque de première espèce de α=0.05 on peut dire que la constante β_0 est non nulle, elle est estimée par : $\widehat{\beta_0} = -2.2578150$

Par contre le test de Student concernant la variable *NbPl* est non significatif. On peut considérer que le paramètre associé à cette variable est nul avec un risque d'erreur de seconde espèce.

L'estimateur du paramètres σ^2 est : $\widehat{\sigma_\varepsilon} = 1.102194$.

Pour tester l'effet des variables explicatives *Var*, *t*, et *NbPl*, nous avons estimé un premier modèle appelé modele1 avec la matrice de corrélation identité (en supposant l'indépendance des mesure répétées) et un second modèle appelé modele3 avec une structure de corrélation AR1. Comme l'effet de la variable *NbPl* est non significatif dans les deux modéles, nous avons repris ces deux modèles sans la variable *NbPl* , les modèles obtenus sont respectivement appelés modele4 et modele6.

Pour selectionner le modèle avec lequel nous allons travailler par la suite, nous allons étudier leurs citères AIC et BIC donnés dans le tableau suivant.

	AIC	**BIC**
modele1	8174.034	8198.869
modele4	8191.945	8211.817
modele3	6081.291	6116.060
modele6	6074.290	6104.097

Les critéres AIC et BIC du modele6 étant plus petits, nous décidons de travailler avec le modele6.

Grâce à la fonction summary, Nous avons donc les estimations des paramètres du modèle6 et leurs erreurs-type estimées.

	Value	Std.Error	t-value	p-value
intercept	-2.3670	0.4624	-5.1188	0.0000
Var	-0.0406	0.0117	-3.4790	0.0005
t	-0.3758	0.0196	-19.1821	0.0000

	(Intercept)	Var
Var	-0.7260	
t	0.1620	0.0560

Les tests de Student permettent de conclure avec un risque d'erreur de première espèce de α=0.05 que les paramètres sont tous significatifs. Leurs estimations sont : $\hat{\beta} = -0.3757853$, $\widehat{\beta_0} = -2.3670$ et $\hat{\delta} = -0.0405544$

Nous devons également examiner les intervalles de confiance des paramètres $\beta_0, \beta, \delta \; et \; \sigma_\varepsilon$ en utilisant la fonction $intervals$.

Intervals(modele6)

Effets fixes

	lower	est.	upper
(Intercept)	-3.2743	-2.3670	-1.4596
Var	-0.0634	-0.0406	-0.0177
t	-0.4142	-0.3758	-0.3373

Ecart type résiduel

	lower	est.	upper
$\widehat{\sigma_\varepsilon}$	0.8807	1.0998	1.3734

Structure de correlation

	lower	est.	upper
phi	0.9443	0.9549	0.9636

Variance fonction

	lower	est.	upper
puissance	0.7922	0.8675	0.9427

Nous pouvons voir qu'il n'y a pas d'imprécision considérable dans les estimations des paramètres $\beta_0, \beta, \delta, \sigma_\varepsilon$ du modèle.

Procédons à présent à la vérification des hypothèses du modèle c'est-à-dire la normalité des résidus et l'homogénéité de la variance des résidus.

On va vérifier ces hypothèses graphiquement, pour la normalité des résidus, on va utiliser la fonction qqnorm et pour l'homogénéité de la variance des résidus on l'interprétera à l'aide du graphique représentant les résidus en fonction des valeurs prédites.

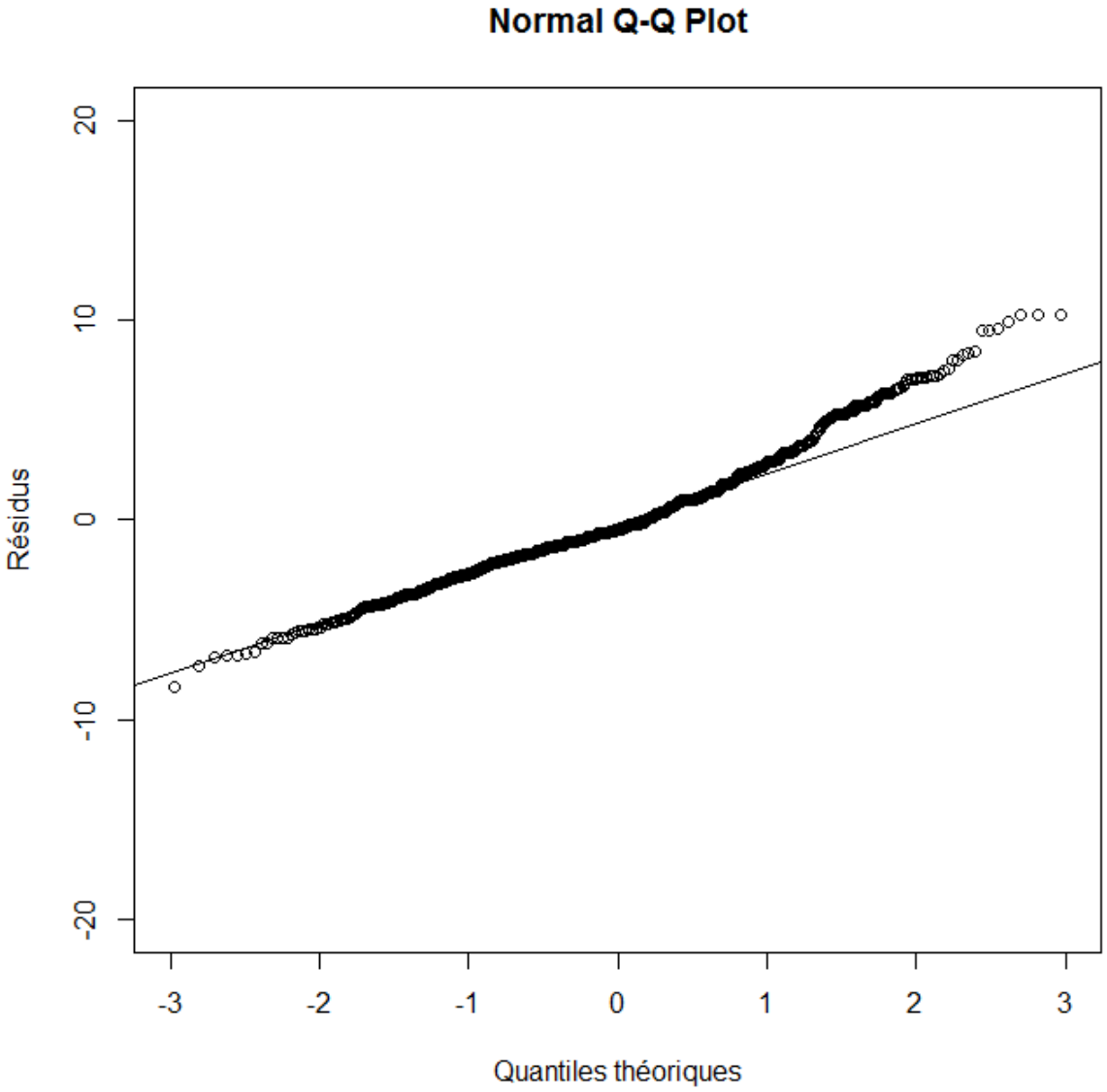

FIG 5 : Représentation graphique des résidus en fonction des quantiles théoriques d'une loi normale

L'examen des résidus (*FIG 5*) ne montre pas une forte violation de la normalité.

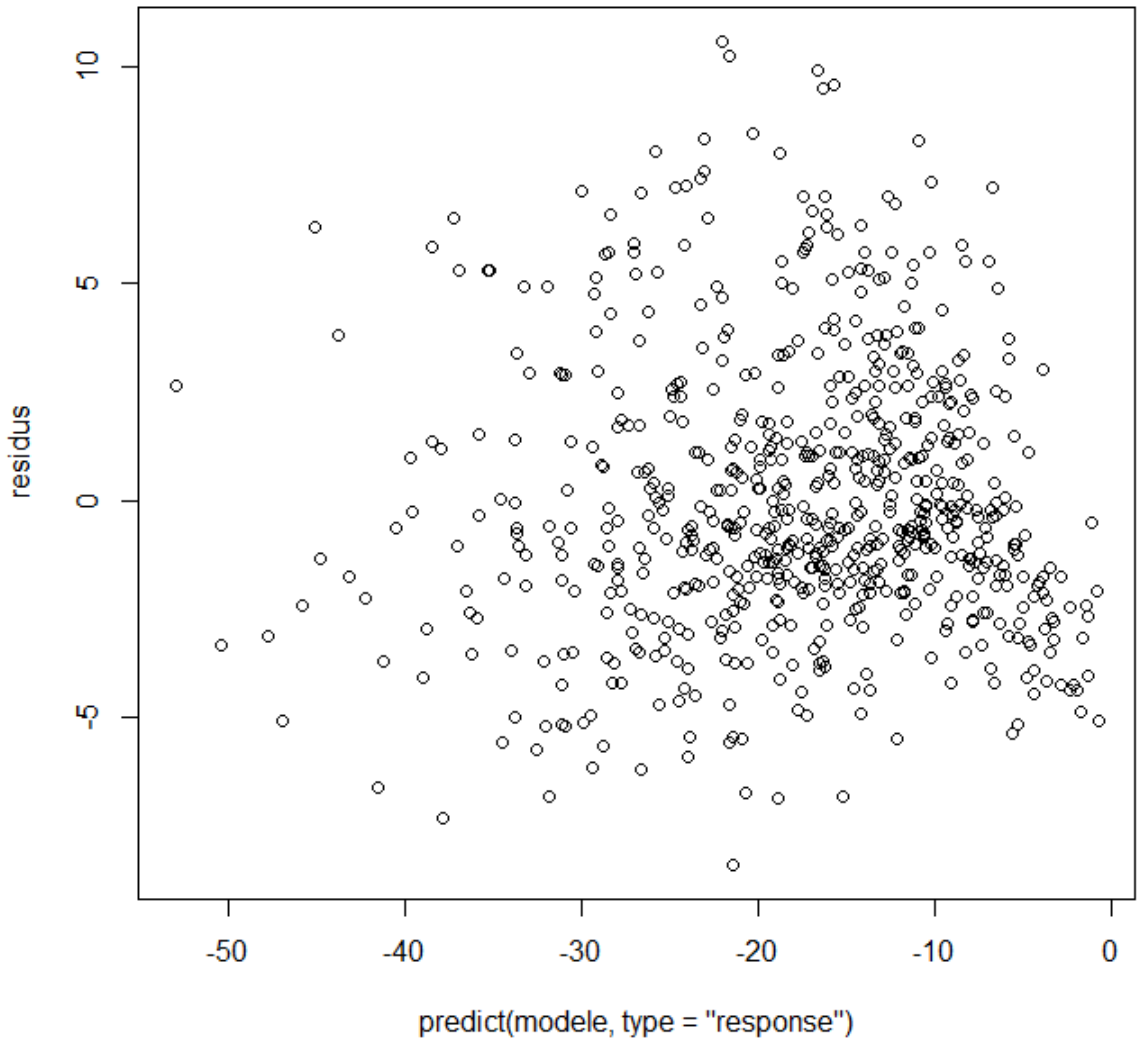

FIG 6 : Représentation graphique des résidus en fonction des valeurs prédites

Les résidus semblent sans structure apparente et relativement homogènes en termes de variabilité. L'hypothèse d'homoscédasticité de la variance des ε_{ij} ne semble pas déraisonnable.

On peut donc conclure que le biais de prédiction s'exprime en fonction de la variance de la surface terrière initiale et de l'intervalle de temps sous la formule suivante :

$$biais_{ij} = -2.3669520 - 0.0405544 * Var_i - 0.3757853 * t_j + \varepsilon_{ij}$$

Le biais de prédiction est lié à la durée de la période de croissance simulée et la variabilité de la surface terrière initiale des placettes sur une strate donnée. En effet nous remarquons que plus la variance de la surface terrière initiale et la durée de la simulation augmentent, plus le biais de prédiction augmente.

3.2 Modèle de FAVRICHON

3.2.1 Description des données

Le deuxième site d'étude est le dispositif expérimental de Paracou en Guyane française. Mis en place en 1984, ce dispositif comporte 12 placettes carrées de 250 m de coté (6,25 ha de surface). Dans chaque placette, tous les arbres de plus de 10cm de diamètre à hauteur de poitrine (soit 1,30m au dessus du sol) ont été identifiés au niveau de l'espèce (essence), étiquetés et suivi individuellement. Les morts sont notés, ainsi que les individus nouvellement recrutés au dessus du seuil de 10 cm de diamètre. Chaque année, le diamètre des arbres a été mesuré.

3.2.2 Description du modèle

Le modèle de Favrichon (*1998*) est un modèle matriciel de dynamique de population (*Caswell, 2001*), de type modèle de Usher (*1966, 1969*), c'est-à-dire pour des populations structurées en taille. Pour tenir compte de la diversité spécifique de la forêt, les essences forestières ont été réparties en 5 groupes fonctionnels ayant un comportement supposé identique d'un point de vue de leur croissance. Le temps est discret et est indexé par les entiers naturels. Le pas de temps est de 2 ans (donc le temps t correspond à $2t$ années). Au temps t, les arbres sont répartis dans k classes de diamètre (k = 11 pour le modèle de Favrichon). Le peuplement forestier au temps t est donc décrit par un vecteur d'effectifs $\mathbf{n_s(t)} = (n_{s1}(t), n_{s2}(t), \ldots\ldots, n_{si}(t) \ldots\ldots, n_{sk}(t))$ où est $n_{si}(t)$ le nombre d'arbres dans la i^e classe de diamètre pour le groupe d'essences ***s.***

Connaissant l'état du peuplement au temps t, le modèle matriciel de Favrichon permet de prédire l'état du peuplement au temps t+1 en faisant intervenir les variables comme la croissance et la mortalité du peuplement considéré sur une surface de référence de $|A_0|$ = 1,5625 ha (125 m × 125 m). Le modèle est donné par :

$$\mathbf{n}_s(\mathbf{t}+\mathbf{1}) = \mathbf{U_s(t)} * \mathbf{n_s(t)} + \mathbf{r_s(t)}$$

$$\mathbf{n}_s(\mathbf{t}) = \begin{pmatrix} n_{s1} \\ \\ n_{sk} \end{pmatrix}_t, \mathbf{U_s(t)} = \begin{pmatrix} q_{s1} & 0 & \ldots & & 0 \\ p_{s1} & q_{s2} & 0 & \ldots & 0 \\ & & & & \\ 0 & \ldots & p_{s10} & & q_{s11} \end{pmatrix}_t, \mathbf{r_s(t)} = \begin{pmatrix} r_{s1} \\ \\ r_{s11} \end{pmatrix}_t$$

$$\begin{pmatrix} n_{s1} \\ \\ n_{s11} \end{pmatrix}_{t+1} = \begin{pmatrix} q_{s1} & 0 & \ldots & & 0 \\ p_{s1} & q_{s2} & 0 & \ldots & 0 \\ & & & & \\ 0 & \ldots & p_{s10} & & q_{s11} \end{pmatrix}_t \times \begin{pmatrix} n_{s1} \\ \\ n_{s11} \end{pmatrix}_t + \begin{pmatrix} r_{s1} \\ \\ r_{s11} \end{pmatrix}_t$$

La matrice de transition $\mathbf{U_s(t)}$ est une matrice stochastique qui décrit les transitions entre les différentes classes de diamètre. Toutes les transitions entre classes de diamètre ne sont pas permises. Étant donné un arbre du groupe s dans la classe de diamètre i au temps t, trois transitions entre les temps t et $t+1$ sont possibles :

Sachant qu'un arbre du groupe s dans la classe de diamètre i au temps t peut, au temps $t+1$

. Soit il reste dans cette classe avec la probabilité q_{si}.

. Soit il passe dans la classe $i+1$ avec la probabilité p_{si}.

. Soit il meurt avec la probabilité $m_{is} = 1 - q_{si} - p_{si}$.

Comme on ne considère que les arbres de plus de 10 cm de diamètre à 1,30m, le peuplement est tronqué et l'on doit introduire également le recrutement $r_{si}(t)$ c'est-à-dire le nombre d'arbres qui atteignent le diamètre de précomptage entre les temps *t* et *t+1* dans la classe *i* pour le groupe d'essences *s*.

Le modèle de Favrichon est un modèle densité-dépendant, c'est-à-dire que les probabilités de transition et le taux de recrutement au temps *t* sont eux-mêmes fonction de l'état du peuplement au temps *t*, selon les relations suivantes :

$$p_{si} = p_0[s] + q_0[s] + p_1[s] \times D_i + p_2[s] \times D_i^2 + p_3[s] \times D_i^3 - q_1[s] \times Bc$$

$$m_{si} = d_1[s] + d_2[s] \times D_i + d_3[s] \times D_i^2$$

$$p_{si} + q_{si} + m_{si} = 1$$

$$r_s = c_1[s] - c_2[s] \times Nc - c_3[s] \times Bc$$

$r_{si} = r_s \times \frac{|A|}{|A_0|}$ $\quad pour\ s = 1,2,3,4 \quad et\ i = 1$

où $|A|\ est\ la\ surface\ de\ la\ placette - échantillon\ \ et\ |A_0|\ la\ surface\ de\ référence$

$r_{si} = \exp(r_s) \times \frac{|A|}{|A_0|} \quad pour\ s = 5\ et\ i = 1$

$r_{si} = 0 \quad pour\ i \neq 1$

où ($p_0[s]$, $q_0[s]$, $p_1[s]$, $p_2[s]$, $p_3[s]$, $q_1[s]$, $d_1[s]$, $d_2[s]$, $d_3[s]$, $c_1[s]$, $c_2[s]$, $c_3[s]$) = θ est l'ensemble des paramètres estimés du modèle, D_i est le diamètre médian de la classe *i*,

$$Bc = \frac{\sum_{s=1}^{5} \sum_{i=1}^{11} n_{si} B_i}{B_0}$$

$$Nc = \frac{\sum_{s=1}^{5} \sum_{i=1}^{11} n_{si}}{N_0}$$

Où $B_i = (\pi/4) D_i^2$ est la surface terrière d'un arbre de la classe *i*, B_0 = 48,95 m^2 est la surface terrière de référence pour une placette de surface $|A_0| = 1{,}56625\ ha$, et N_0 = 968,25 est l'effectif de référence pour une placette de surface $|A_0| = 1{,}56625\ ha$. Les ratios *Bc* et *Nc* sont des indices de compétition qui régulent la dynamique : comme les paramètres $q_1[s]$, $c_2[s]$ et $c_3[s]$ sont positifs, la croissance et le recrutement diminuent lorsque la surface terrière ou la densité du peuplement augmentent.

3.2.3 Estimation du biais par Monte Carlo

Dans cette partie, on va calculer par la méthode de Monte Carlo le biais de prédiction induit par un changement de la surface de la placette-échantillon. On s'intéresse à ce qui se passe à la prédiction du modèle lorsque l'on passe d'une placette A à une placette A' incluse dans A ou contenant A .

Pour cela, on va déterminer le biais de prédiction quand les placettes utilisées pour faire les prédictions sont une partition d'une grande placette, et sont donc corrélées spatialement. On va également calculer la variance des variables explicatives qui résulte de cette autocorrélation spatiale.

On va donc prédire l'évolution temporelle des placettes de Paracou pour des surfaces de grain variables allant de d'environ 0,0625 ha (soit 100 placettes par parcelle) à 6,25 ha (une seule placette-échantillon).

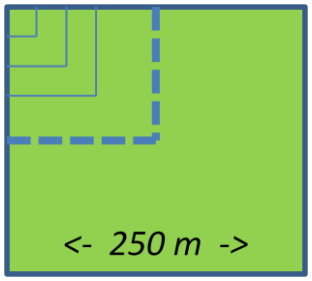

FIG. 7 Parcelle de 6,25 ha partitionnée en k^2 placettes. Cette partition s'obtient en divisant chaque côté de 250 m de la parcelle en k parties égales (pour k allant de 1 à 10).

Soit $P = [0, 250] \times [0, 250]$ une parcelle de 6,25 ha de Paracou, n_t le nombre total d'arbres qui s'y trouvent à l'année t, et (x_i, y_i, D_i, s_i) les coordonnées, le diamètre et le groupe d'essence du i^e arbre $i = 1, \ldots\ldots, n_t$. On fait des tirages de placettes-échantillon de surface inférieure à 6,25 ha dans cette parcelle P . Pour chaque placette-échantillon, on calcule le nombre d'arbres à l'hectare prédite par le modèle matriciel de Favrichon au bout d'un nombre de pas de temps fixé. Cette densité est, par définition, la prédiction *Y* dont on veut estimer le biais.

On peut faire deux sortes de tirages :

tirage sans remise : la parcelle P est divisée en $k \times k$ placettes carrées de même surface, qui forment une grille régulière de coté $a = \frac{250}{k}(m)$;
tirage avec remise : on tire au hasard dans P les coordonnées du coin inférieur gauche de la placette carrée de coté a .

Les deux modes de tirage donnent des résultats très semblables. Pour cette étude, on utilisera un tirage sans remise.

Soit M_k le nombre de tirages d'une placette A_k de coté $250/k$ avec $M_k = k^2$ pour le tirage sans remise (on aurait M_k aussi grand que l'on veut pour le tirage avec remise).

On définit le nombre total d'arbres sur une placette comme étant la somme de tous les effectifs d'arbres par classe de diamètre et par groupe d'essences sur la placette.

Soit Y_{kj} (avec $j = 1, \ldots\ldots, M_k$) le nombre total d'arbres obtenu par une prédiction du modèle de Favrichon utilisant le j^e tirage de A_k et $\mathbf{n}_{\mathrm{A_k}}(\mathbf{t})$ comme état initial. La moyenne empirique $\bar{Y}_k = (\sum_{j=1}^{M_k} Y_{kj})/M_k$ est un estimateur de $E(Y_k)$. On s'intéresse aux variations de $\bar{Y}_k$ avec la surface de la placette A_k .

S'il n'y avait pas de biais de prédiction lié à la surface de la placette, $\bar{Y}_k$ ne varierait pas avec k.

A titre d'exemple, *la figure 8* montre les variations de $\bar{Y}_k$ avec la surface de A_k pour la parcelle 1 de Paracou lorsque la prédiction Y_k représente le nombre d'arbres à l'hectare prédite par le modèle matriciel au bout de 10 pas de temps (soit 20 ans). Il y a bien un biais de prédiction (voir figure 8). On a tendance à surestimer le nombre de tiges pour des placettes de surface inférieure à la surface de référence (1.5625ha) ayant servie à établir le modèle. Toutefois ce biais ne dépasse par 1,5 % de la valeur correcte (534,6 arbres ha^{-1} pour la surface de référence de 1,5625 ha).

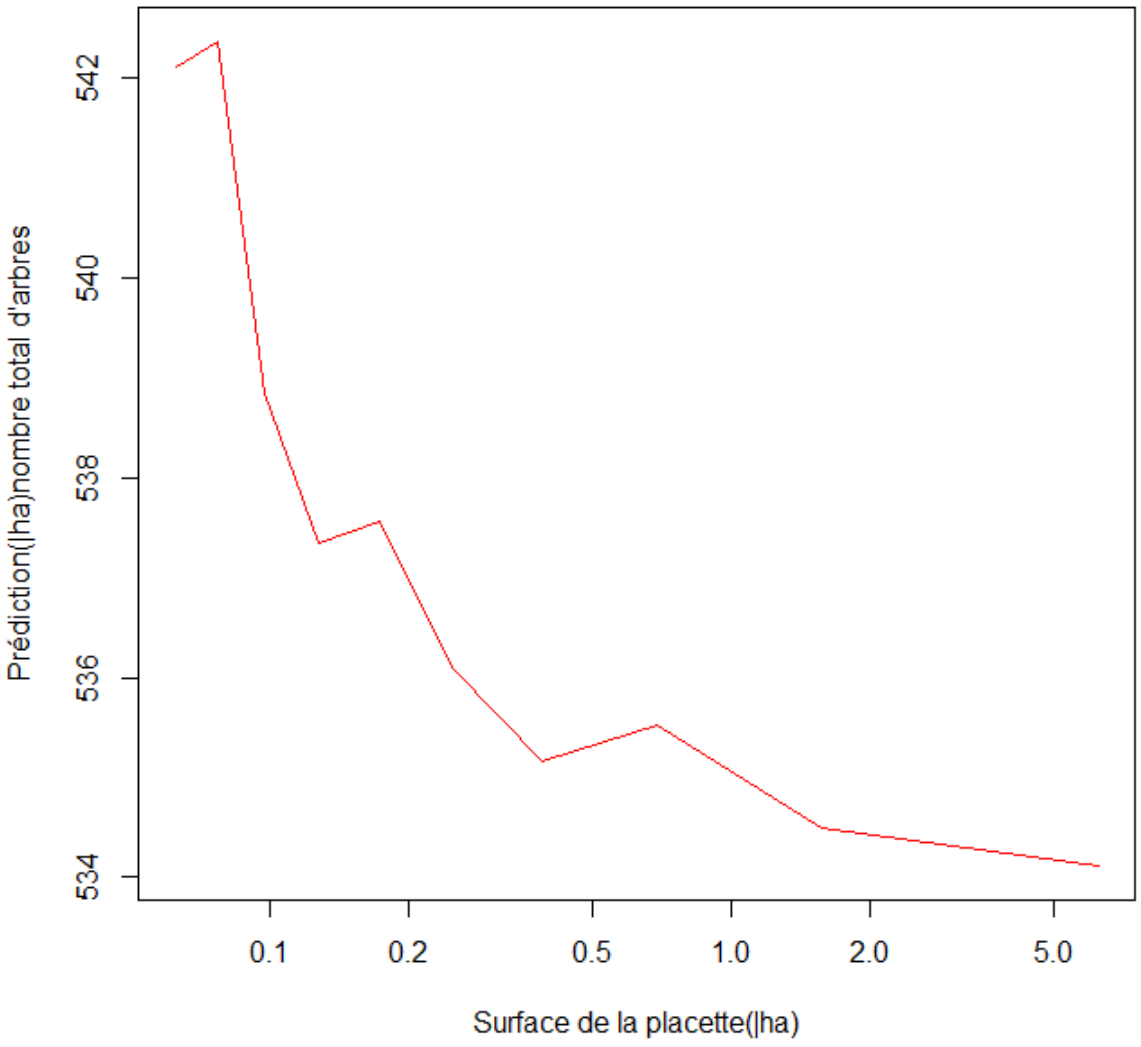

FIG. 8 –Impact de la surface des placettes sur la prédiction du nombre d'arbres (ha^{-1}) *prédite par le modèle matriciel de Favrichon au bout de 10 pas de temps (20 ans)*

3.2.4 Estimation analytique approchée du biais

On se propose dans cette partie de calculer une estimation approchée du biais induit par un changement d'échelle en utilisant l'expression analytique approchée du biais.

A partir du modèle matriciel de Favrichon on définit le nombre total d'arbres à l'hectare au bout d'un pas de temps dans la placette A par :

$$Y_A = \frac{1}{|A|}\sum_{s=1}^{5}\sum_{i=1}^{11} n_{si}(t+1)$$

$$Y_A = \frac{1}{|A|}\sum_{s=1}^{5}\{\mathbf{1}^{\top}\{\mathbf{U_s(t)n_s(t)} + \mathbf{r_s(t)}\}\}$$

$$\mathbf{U_s(t)n_s(t)} + \mathbf{r_s(t)} = \begin{pmatrix} n_{s1} \\ \\ \\ n_{s11} \end{pmatrix}_{t+1} = \begin{pmatrix} q_{s1} & 0 & \dots & & 0 \\ p_{s1} & q_{s2} & 0 & \dots & 0 \\ & & & & \\ & & & & \\ 0 & \dots & & p_{s10} & q_{s11} \end{pmatrix}_t \times \begin{pmatrix} n_{s1} \\ \\ \\ n_{s11} \end{pmatrix}_t + \begin{pmatrix} r_{s1} \\ \\ \\ r_{s11} \end{pmatrix}_t$$

$$= \begin{pmatrix} q_{s1}n_{s1} + r_{s1} \\ p_{s1}n_{s1} + q_{s2}n_{s2} \\ p_{s2}n_{s2} + q_{s3}n_{s3} \\ . \\ . \\ p_{s9}n_{s9} + q_{s10}n_{s10} \\ p_{s10}n_{s10} + q_{s11}n_{s11} \end{pmatrix}_t \text{ car } r_{si} = 0 \; pour \; i \neq 1$$

Or $p_{si} + q_{si} + m_{si} = 1$ ce qui implique que $q_{si} = 1 - p_{si} - m_{si}$

On a donc :

$$\mathbf{U_s(t)n_s(t)} + \mathbf{r_s(t)} = \begin{pmatrix} (1 - p_{s1} - m_{s1})n_{s1} + r_{s1} \\ p_{s1}n_{s1} + (1 - p_{s2} - m_{s2})n_{s2} \\ p_{s2}n_{s2} + (1 - p_{s3} - m_{s3})n_{s3} \\ . \\ . \\ p_{s9}n_{s9} + (1 - p_{s10} - m_{s10})n_{s10} \\ p_{s10}n_{s10} + (1 - p_{s11} - m_{s11})n_{s11} \end{pmatrix}_t$$

$\mathbf{1}^{\top}\{\mathbf{U_s(t)n_s(t)} + \mathbf{r_s(t)}\} = \sum_{i=1}^{11}(1 - m_{si})\, n_{si}(t) - p_{s11}n_{s11}(t) + r_{s1}(t)$

D'où

$$Y_A = \frac{1}{|A|}\sum_{s=1}^{5}\{\sum_{i=1}^{11}(1 - m_{si})\, n_{si}(t) - p_{s11}n_{s11}(t) + r_{s1}(t)\}$$

Rappelons que pour chaque groupe d'essences s=1, …5, on a :

$\mathbf{n_s(t)} = \begin{pmatrix} n_{s1} \\ \\ n_{s11} \end{pmatrix}_t$ vecteur contenant les effectifs d'arbres par classe de diamètre du groupe d'essences *s* au temps *t*.

$\mathbf{r_s(t)} = \begin{pmatrix} r_{s1} \\ \\ r_{s11} \end{pmatrix}_t = \begin{pmatrix} r_{s1} \\ 0 \\ \\ 0 \end{pmatrix}_t$ Vecteur contenant les nombres d'arbres nouvellement recrus par classe de diamètre au temps *t*.

$\mathbf{p_s} = \begin{pmatrix} p_{s1} \\ \\ p_{s11} \end{pmatrix}$ où p_{si} est la probabilité pour un arbre du groupe *s* dans la classe de diamètre i au temps *t* de passer à la classe *i+1* au temps *t+1*.

$\mathbf{q_s} = \begin{pmatrix} q_{s1} \\ \\ q_{s11} \end{pmatrix}$ où p_{si} est la probabilité pour un arbre du groupe *s* dans la classe de diamètre i au temps *t* d'y rester au temps *t+1.*

$\mathbf{m_s} = \begin{pmatrix} m_{s1} \\ \\ m_{s11} \end{pmatrix}$ où m_{si} est la probabilité pour un arbre du groupe *s* dans la classe de diamètre i au temps *t* de mourir au temps *t+1.*

Posons: $\mathbf{n_A(t)} = \begin{pmatrix} \mathbf{n_1(t)} \\ \mathbf{n_2(t)} \\ \mathbf{n_3(t)} \\ \mathbf{n_4(t)} \\ \mathbf{n_5(t)} \end{pmatrix} = \begin{pmatrix} n_1 \\ . \\ . \\ . \\ n_{55} \end{pmatrix}$ vecteur de longueur 55 contenant les nombres d'arbres par classe de diamètre et par groupe d'essences dans la placette A.

$$\mathbf{r}\big(\mathbf{n_A(t)}\big) = \begin{pmatrix} \mathbf{r_1(t)} \\ \mathbf{r_2(t)} \\ \mathbf{r_3(t)} \\ \mathbf{r_4(t)} \\ \mathbf{r_5(t)} \end{pmatrix} = \begin{pmatrix} r_1 \\ . \\ . \\ . \\ r_{55} \end{pmatrix}$$

$$\begin{pmatrix}\mathbf{p_1}\\ \mathbf{p_2}\\ \mathbf{p_3}\\ \mathbf{p_4}\\ \mathbf{p_5}\end{pmatrix}=\begin{pmatrix}p_1\\ .\\ .\\ .\\ .\\ p_{55}\end{pmatrix}$$

$$\begin{pmatrix}\mathbf{q_1}\\ \mathbf{q_2}\\ \mathbf{q_3}\\ \mathbf{q_4}\\ \mathbf{q_5}\end{pmatrix}=\begin{pmatrix}q_1\\ .\\ .\\ .\\ .\\ q_{55}\end{pmatrix}$$

$$\begin{pmatrix}\mathbf{m_1}\\ \mathbf{m_2}\\ \mathbf{m_3}\\ \mathbf{m_4}\\ \mathbf{m_5}\end{pmatrix}=\begin{pmatrix}m_1\\ .\\ .\\ .\\ .\\ m_{55}\end{pmatrix}$$

$$\mathbf{U}\big(\mathbf{n}_{\mathrm{A}}(\mathbf{t})\big)=\begin{pmatrix}\mathbf{U_1}(\mathbf{t}) & 0 & \dots & & 0\\ 0 & \mathbf{U_2}(\mathbf{t}) & 0 & \dots & 0\\ & & & & \\ & & & & \\ 0 & \dots & 0 & & \mathbf{U_5}(\mathbf{t})\end{pmatrix}$$ Matrice (55×55) où $\mathbf{U_i}(\mathbf{t})$ est une matrice de transition.

$$\mathbf{1}=\begin{pmatrix}1\\ .\\ .\\ .\\ .\\ 1\end{pmatrix}$$

$$Y_A=\frac{1}{|A|}\{\mathbf{1}^t(\mathbf{U}[\mathbf{n}_{\mathrm{A}}(\mathbf{t})]\mathbf{n}_{\mathrm{A}}(\mathbf{t})+\mathbf{r}[\mathbf{n}_{\mathrm{A}}(\mathbf{t})])\}$$

Où $|A|$ est la surface de la placette A, $\mathbf{1}$ est le vecteur de longueur 55 ne contenant que des 1**,** **U($\mathbf{n}_{\mathrm{A}}(\mathbf{t})$)** est la matrice 55×55 de transition de Usher, $\mathbf{r}(\mathbf{n}_{\mathrm{A}}(\mathbf{t}))$ est le vecteur de longueur 55 du recrutement , et $\mathbf{n}_{\mathrm{A}}(\mathbf{t})$ le vecteur de longueur 55 contenant les effectifs par classe de diamètre et par groupe d'essences observés dans A.

$$Y_P=\frac{1}{|P|}\{\mathbf{1^t}(\mathbf{U}[\mathbf{n}_{\mathrm{P}}(\mathbf{t})]\mathbf{n}_{\mathrm{P}}(\mathbf{t})+\mathbf{r}[\mathbf{n}_{\mathrm{P}}(\mathbf{t})])\}$$

$$Y_A=\frac{1}{|A|}\{\sum_{i=1}^{55}(1-m_i)n_i-(p_{11}n_{11}+p_{22}n_{22}+p_{33}n_{33}+p_{44}n_{44}+p_{55}n_{55})+r_1+r_{12}+r_{23}+r_{34}+r_{45}\}$$

$$p_{si}=p_0[s]+q_0[s]+p_1[s]\times D_i+p_2[s]\times D_i^2+p_3[s]\times D_i^3-q_1[s]\times Bc$$

Pour $s=1,\dots,5$ et $i=1,\dots,11$ $p_0[s]+q_0[s]+p_1[s]\times D_i+p_2[s]\times D_i^2+p_3[s]\times D_i^3$ est connu et est indépendant de $\mathbf{n}_{\mathrm{A}}(\mathbf{t})$.

Pour simplifier les calculs, on va poser $p_0[s] + q_0[s] + p_1[s] \times D_i + p_2[s] \times D_i^2 + p_3[s] \times D_i^3 = b_{si}$

$$p_{si} = b_{si} - q_1[s] \times Bc$$

$$\begin{pmatrix} p_1 \\ . \\ . \\ p_{11} \\ . \\ . \\ . \\ p_{44} \\ . \\ p_{55} \end{pmatrix} = \begin{pmatrix} b_1 - q_1[1] \times Bc \\ . \\ . \\ b_{11} - q_1[1] \times Bc \\ . \\ . \\ . \\ b_{44} - q_1[5] \times Bc \\ . \\ b_{55} - q_1[5] \times Bc \end{pmatrix}$$

$$Y_A = \frac{1}{|A|}\{\sum_{i=1}^{55}(1 - m_i)n_i + \sum_{j \in \{(11,22,33,44,55)\}} \left(b_j - q_1\left[\frac{j}{11}\right] \times Bc\right) \times n_i + \sum_{s=1}^{4}(c_1[s] - c_2[s] \times Nc - c_3[s] \times Bc) \times \frac{A}{A_0} + \exp\left(c_1[5] - c_2[5] \times Nc - c_3[5] \times Bc\right) \times \frac{A}{A_0}\}$$

Avec

$Bc = \frac{\sum_{s=1}^{5}\sum_{i=1}^{11} n_{si}B_i}{B_0} = \frac{\sum_{i=1}^{55} n_i B_i}{B_0}$ où $B = \begin{pmatrix} B_1 \\ . \\ B_{11} \\ \vdots \\ B_1 \\ . \\ B_{11} \end{pmatrix}$

Et

$$Nc = \frac{\sum_{s=1}^{5}\sum_{i=1}^{11} n_{si}}{N_0} = \frac{\sum_{i=1}^{55} n_i}{N_0}$$

Pour A fixé, définissons une fonction g de $\mathbb{R}^{55}$ dans $\mathbb{R}$ qui au vecteur $\mathbf{n}_\mathrm{A}(\mathbf{t})$ associe :

$$g: \mathbb{R}^{55} \to \mathbb{R}$$
$$\mathbf{n}_\mathbf{A}(\mathbf{t}) \to Y = \frac{1}{|A|}\{\mathbf{1}^\top(\mathbf{U}[|\mathbf{A}|\mathbf{n}'_\mathbf{A}(\mathbf{t})] \times |A| \times \mathbf{n}'_\mathbf{A}(\mathbf{t}) + \mathbf{r}[|\mathbf{A}| \times \mathbf{n}'_\mathbf{A}(\mathbf{t})])\}$$

$$g(\mathbf{n}'_\mathrm{A}(\mathbf{t})) = Y_A = \frac{1}{|A|}\{\mathbf{1}^\top(\mathbf{U}[|\mathrm{A}|\mathbf{n}'_\mathrm{A}(\mathbf{t})] \times |A| \times \mathbf{n}'_\mathrm{A}(\mathbf{t}) + \mathbf{r}[|\mathrm{A}| \times \mathbf{n}'_\mathrm{A}(\mathbf{t})])\}$$

Où $\mathbf{n}'_\mathrm{A}(\mathbf{t}) = \frac{\mathbf{n}_\mathbf{A}(\mathbf{t})}{|A|}$ vecteur contenant le nombre d'arbres à l'hectare par classe de diamètre et par groupe d'essence dans la placette A.

$$g\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t})\right) = Y_A = \frac{1}{|A|}\{\sum_{i=1}^{55}(1-m_i)|A|n'_i + \sum_{j\,\epsilon\,\{(11,22,33,44,55)\}}\left(b_j - q_1\left[\frac{j}{11}\right]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0}\right)\times|A|n'_j + \sum_{s=1}^{4}(c_1[s] - c_2[s]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[s]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{|A|}{|A_0|} + \exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{|A|}{|A_0|}\}$$

$$g\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t})\right) = Y_A = \{\sum_{i=1}^{55}(1-m_i)n'_i + \sum_{j\,\epsilon\,\{(11,22,33,44,55)\}}\left(b_j - q_1\left[\frac{j}{11}\right]\times\frac{\sum_{i=1}^{55}n'_i|A|B_i}{B_0}\right)\times n'_j + \sum_{s=1}^{4}(c_1[s] - c_2[s]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[s]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|} + \exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|}\}$$

Un estimateur approché du biais peut être obtenu de manière analytique en faisant un développement limité de la prédiction Y_A par rapport aux éléments de $\mathbf{n}_{\mathrm{A}}(\mathbf{t})$. Pour éviter que les calculs ne soient trop complexes, on ne considère ici que la prédiction du modèle matriciel au bout d'un seul pas de temps de 2 ans. Mais, dans le principe, le calcul serait généralisable à n'importe quel pas de temps :

$$g(\mathbf{n}'_{\mathrm{A}}(\mathbf{t})) = Y_A = \frac{1}{|A|}\{1^t(\mathbf{U}[|A|\mathbf{n}'_{\mathrm{A}}(\mathbf{t})]|A|\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) + \mathbf{r}[|A|\mathbf{n}'_{\mathrm{A}}(\mathbf{t})])\}$$

$$g(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})) = Y_P = \frac{1}{|P|}\{1^t(\mathbf{U}[|P|\mathbf{n}'_{\mathrm{P}}(\mathbf{t})]|A|\mathbf{n}'_{\mathrm{P}}(\mathbf{t}) + \mathbf{r}[|P|\mathbf{n}'_{\mathrm{P}}(\mathbf{t})])\}$$

$$g\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t})\right) \simeq g\left(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right) + \left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right)g'\left(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right) + \frac{1}{2}(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t}))g''(\mathbf{n}'_{\mathrm{P}}(\mathbf{t}))(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t}))^{\top}$$

$Y_A \simeq Y_P + (\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t}))g'(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})) + \frac{1}{2}(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t}))g''(\mathbf{n}'_{\mathrm{P}}(\mathbf{t}))(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t}))^{\top}$

D'après la δ méthode, on a alors de manière approchée

$$E\left(g\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t})\right)\right) = E(Y_A) \simeq g\left(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right) + E\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right)g'\left(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right) + \frac{1}{2}E[\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right)g''\left(\mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right)\left(\mathbf{n}'_{\mathrm{A}}(\mathbf{t}) - \mathbf{n}'_{\mathrm{P}}(\mathbf{t})\right)^{\top}]$$

Or $E(\mathbf{n}'_{\mathrm{A}}(\mathbf{t})) = \mathbf{n}'_{\mathrm{P}}(\mathbf{t})$

Donc

$$E(Y_A) \simeq Y_P + \frac{1}{2}\mathbf{1}^{\top}(\mathbf{\Sigma}\odot\mathbf{H}_{\mathrm{Y}})\mathbf{1}$$

Où $\mathbf{\Sigma}$ est la matrice 55×55 de variance-covariance d,$N'_A(t)$, $\mathbf{H}_{\mathrm{Y}}$ est la matrice hessienne de Y calculée en $\mathbf{n}'_{\mathrm{P}}(\mathbf{t})$, et ʘ désigne le produit de Hadamard (ou produit terme à terme des matrices . la matrice de variance covariance de $\mathbf{n}_{\mathrm{A}}(\mathbf{t})$ est :

$$\Sigma = \begin{bmatrix} Var\left(n_{A,11}(t)\right) & Cov\left(n_{A,11}(t), n_{A,21}(t)\right) & \cdots \\ Cov\left(n_{A,21}(t), n_{A,11}(t)\right) & Var\left(n_{A,11}(t)\right) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Où $n_{A,ij}(t)$ est l'effectif d'arbres dans la i^e classe de diamètre $(1 \leq j \leq 11)$ et le j^e groupe d'essence$(1 \leq j \leq 5)$. Un estimateur $\widehat{\mathbf{\Sigma}}$ de $\mathbf{\Sigma}$ s'obtient comme précédemment en faisant B tirages avec ou sans remise de la placette A dans la parcelle P , et en prenant la variance empirique de ces B tirages.

Pour A fixé, l'application g est ici l'application de $\mathbb{R}^{55}$ dans $\mathbb{R}$ qui au vecteur $\mathbf{n}'$ associe :

$$g: \begin{matrix} \mathbb{R}^{55} \to \mathbb{R} \\ \mathbf{n}' \to Y = \frac{1}{|A|}\{1^t(U(|A|\mathbf{n}')|A|\mathbf{n}' + R[|A|\mathbf{n}'])\} \end{matrix}$$

Sa matrice jacobienne, J_Y , est une matrice 1×55 définie par :

$$J_Y \equiv \frac{\partial Y}{\partial \mathbf{n}'} = [\frac{\partial Y}{\partial n'_1} \frac{\partial Y}{\partial n'_2} \ldots . \frac{\partial Y}{\partial n'_{55}}]$$

Où n'_i est le i^e élément de $\mathbf{n}' = \frac{\mathbf{n}}{|A|}$.

La matrice jacobienne de l'application Y définit sa différentielle dY , puisque la différentielle de Y est l'application de $\mathbb{R}^{55}$ dans $\mathbb{R}^{55}$ qui à tout vecteur *u* de longueur 55 associe :

$$dY: \begin{matrix} \mathbb{R}^{55} \to \mathbb{R}^{55} \\ u \to J_Y u \end{matrix}$$

La différentielle d'ordre 2 de Y est par définition la différentielle de sa différentielle :$d^2Y = d(dY)$. Il s'agit d'une application de $\mathbb{R}^{55}$ dans $\mathbb{R}^{55}$ qui a tout vecteur *u* de longueur 55 associe

$$d^2Y: \begin{matrix} \mathbb{R}^{55} \to \mathbb{R}^{55} \times \mathbb{R}^{55} \\ u \to u^t J_Y u \end{matrix}$$

Où H_Y est la matrice hessienne de, qui n'est autre que la matrice jacobienne de dY. Il s'agit d'une matrice 55×55

$$\begin{bmatrix} \frac{\partial^2 Y}{\partial^2 n'_1} & \frac{\partial^2 Y}{\partial n'_1 \partial n'_2} & \cdots\cdots\cdots & \frac{\partial^2 Y}{\partial n'_1 \partial n'_{55}} \\ \frac{\partial^2 Y}{\partial n'_1 \partial n'_2} & \frac{\partial^2 Y}{\partial^2 n'_2} & \cdots\cdots\cdots & \frac{\partial^2 Y}{\partial n'_1 \partial n'_{55}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 Y}{\partial n'_1 \partial n'_{55}} & \frac{\partial^2 Y}{\partial n'_2 \partial n'_{55}} & \cdots\cdots\cdots & \frac{\partial^2 Y}{\partial^2 n'_{55}} \end{bmatrix}$$

$$g(\mathbf{n'_A(t)}) = Y_A = \{\sum_{i=1}^{55}(1-m_i)n'_i + \sum_{j\,\epsilon\,\{(11,22,33,44,55)\}}\left(b_j - q_1\left[\frac{j}{11}\right]\times\frac{\sum_{i=1}^{55}n'_i|A|B_i}{B_0}\right)\times n'_j + \sum_{s=1}^{4}(c_1[s] - c_2[s]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[s]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|} + \exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|}\}$$

Posons $R5 = c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0}$

Pour $\boldsymbol{i \notin \{(11,22,33,44,55)\}}$

$$\frac{\partial g}{\partial n'_i} = (1-m_i) + \sum_{j\epsilon\,\{(11,22,33,44,55)\}}\frac{q_1[\frac{j}{11}]|A|B_in'_j}{B_0} - \sum_{s=1}^{4}(\frac{c_2[s]|A|}{N_0} + \frac{c_3[s]|A|B_i}{B_0})\times\frac{1}{|A_0|} - (\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0})\exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|}$$

$$\frac{\partial^2 g}{\partial^2 n'_i} = (\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0})^2\exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|}$$

$$\frac{\partial^2 g}{\partial^2 N'_i} = (\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0})^2\exp\,(R5)\times\frac{1}{|A_0|}$$

Pour $\boldsymbol{i \notin \{(11,22,33,44,55)\}}$ $\boldsymbol{et\ k \notin \{(11,22,33,44,55)\}}$

$$\frac{\partial^2 g}{\partial n'_i\partial n'_k} = (\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0})\times(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_k}{B_0})\times exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|}$$

$$\frac{\partial^2 g}{\partial n'_i\partial n'_k} = (\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0})\times(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_k}{B_0})\times exp\,(R5)\times\frac{1}{|A_0|}$$

Pour $\boldsymbol{i \notin \{(11,22,33,44,55)\}}$ $\boldsymbol{et\ k\ \epsilon\ \{(11,22,33,44,55)\}}$

$$\frac{\partial^2 g}{\partial n'_i\partial n'_k} = \frac{q_1\left[\frac{k}{11}\right]|A|B_i}{B_0} + (\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0})\times(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_k}{B_0})\times exp\,(c_1[5] - c_2[5]\times\frac{\sum_{i=1}^{55}|A|n'_i}{N_0} - c_3[5]\times\frac{\sum_{i=1}^{55}|A|n'_iB_i}{B_0})\times\frac{1}{|A_0|}$$

$$\frac{\partial^2 g}{\partial n_i' \partial n_k'} = \frac{q_1\left[\frac{k}{11}\right]|A|B_i}{B_0} + \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0}\right) \times \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_k}{B_0}\right) \times exp\,(R5) \times \frac{1}{|A_0|}$$

Pour $\boldsymbol{i \in \{(11, 22, 33, 44, 55)\}}$

$$\frac{\partial g}{\partial n_i'} = (1 - m_i) + \sum_{j \in \{(11,22,33,44,55)\}} \frac{q_1[\frac{j}{11}]|A|B_i n_j'}{B_0} + \frac{q_1[\frac{i}{11}]|A|B_i n_i'}{B_0} - b_i + q_1[\frac{i}{11}] \sum_{l \neq i} \frac{|A|B_l n_l'}{B_0}$$
$$- \sum_{s=1}^{4} \left(\frac{c_2[s]|A|}{N_0} + \frac{c_3[s]|A|B_i}{B_0}\right) - \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0}\right) \exp\,(c_1[5] - c_2[5]$$
$$\times \frac{\sum_{i=1}^{55}|A|n_i'}{N_0} - c_3[5] \times \frac{\sum_{i=1}^{55}|A|n_i'B_i}{B_0}) \times \frac{1}{|A_0|}$$

$$\frac{\partial^2 g}{\partial^2 n_i'} = 2\frac{q_1[\frac{i}{11}]|A|B_i}{B_0} + \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0}\right)^2 \exp\,(c_1[5] - c_2[5] \times \frac{\sum_{i=1}^{55}|A|n_i'}{N_0} - c_3[5]$$
$$\times \frac{\sum_{i=1}^{55}|A|n_i'B_i}{B_0}) \times \frac{1}{|A_0|}$$

$$\frac{\partial^2 g}{\partial^2 n_i'} = 2\frac{q_1[\frac{i}{11}]|A|B_i}{B_0} + \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0}\right)^2 \exp\,(R5) \times \frac{1}{|A_0|}$$

Pour $\boldsymbol{i \in \{(11, 22, 33, 44, 55)\}\ et\ k \in \{(11, 22, 33, 44, 55)\}}$

$$\frac{\partial^2}{\partial n_i' \partial n_k'} = \frac{q_1\left[\frac{i}{11}\right]|A|B_k}{B_0} + \frac{q_1[\frac{k}{11}]|A|B_i}{B_0} + \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0}\right) \times \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_k}{B_0}\right)$$
$$\times exp\,(c_1[5] - c_2[5] \times \frac{\sum_{i=1}^{55}|A|n_i'}{N_0} - c_3[5] \times \frac{\sum_{i=1}^{55}|A|n_i'B_i}{B_0}) \times \frac{1}{|A_0|}$$

$$\frac{\partial^2}{\partial n_i' \partial n_k'} = \frac{q_1\left[\frac{i}{11}\right]|A|B_k}{B_0} + \frac{q_1[\frac{k}{11}]|A|B_i}{B_0} + \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_i}{B_0}\right) \times \left(\frac{c_2[5]|A|}{N_0} + \frac{c_3[5]|A|B_k}{B_0}\right)$$
$$\times exp\,(R5) \times \frac{1}{|A_0|}$$

Grâce au développement de Taylor au second ordre, on a pu estimer le biais de prédiction de façon analytique (mais approchée). La *figure 9* compare l'expression analytique approchée du biais (courbe bleue) et l'estimation du biais obtenue par Monte Carlo (courbe rouge).

On peut remarquer qu'il y'a une différence entre les deux courbes, qui est d'autant plus petite que la surface $|A|$ est grande. Le développement limité de Taylor n'est en effet qu'une approximation qui va être d'autant meilleure que la variabilité des effectifs est faible. Comme cette variabilité augmente quand la surface des placettes diminue, on s'attend à ce que le développement de Taylor donne un résultat de plus en plus approximatif au fur et à mesure que les placettes deviennent petites. C'est effectivement ce que l'on observe dans la figure 9. Pour améliorer l'estimation analytique approchée, on pourrait pousser le développement de Taylor à un ordre supérieur mais les calculs deviennent longs (tout en restant à la portée d'un logiciel de calcul formel dans le cas du modèle de Favrichon).

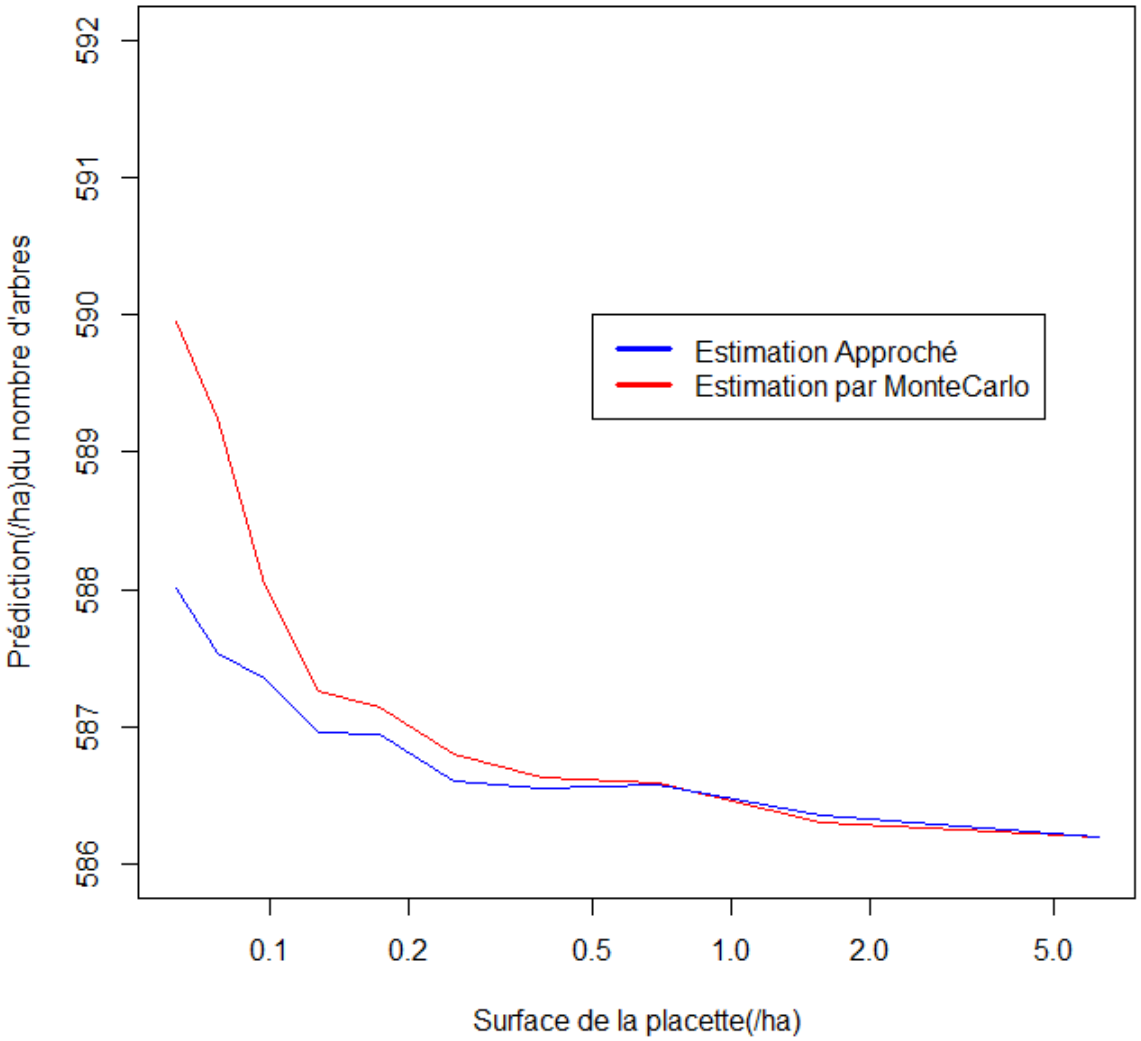

FIG. 9 - Densité totale (ha^{-1}) prédite par le modèle matriciel de Favrichon au bout d'un pas de temps (2 ans) lorsque l'état initial est une placette de surface variable placée dans la parcelle 1 de Parcou en 1992. En rouge, prédiction obtenue par la méthode de Monte Carlo (tirage sans remise). En bleu, approximation analytique basée sur un développement de Taylor d'ordre 2.

4. Conclusion

Ce stage se place dans une démarche générale d'améliorer les modèles de croissance forestière afin de proposer une méthode permettant de tenir compte des biais de prédiction dans les simulateurs de croissance. Pour cela, il est nécessaire de connaître les sources de biais.

Ce travail nécessite des données relativement détaillées, comme c'est le cas des données issues du dispositif de la forêt tropicale en Guyane française et des forêts publiques du Québec. Grâce aux modèles de croissance utilisés dans ces deux forêts et grâce à des outils statistiques comme la modélisation, nous avons pu démontrer l'existence de biais de prédiction lorsque les modèles de croissance sont utilisés de façon à ce que la variabilité des paramètres d'entrée se trouve affectée.

Par exemple, pour le modèle ARTEMIS utilisé dans les forêts publiques du Québec, utiliser une valeur moyenne pour résumer un ensemble de placettes conduit à sous-estimer le nombre d'arbres à l'hectare de l'ensemble de ces placettes. En effet les nombres d'arbres à l'hectare obtenus après une agrégation *a priori* sont généralement inférieurs à ceux obtenus lorsque toutes les placettes sont utilisées pour les simulations. Nous avons établis une première étape dans la modélisation de ce biais. Par exemple, il semble que la variabilité de la surface terrière au sein d'une strate soit une source de biais importante.

Nous avons également démontré l'existence d'un biais associé à la surface des placettes échantillon. Lorsque cette surface est réduite par rapport à la surface des placettes de référence ayant servi à établir le modèle de croissance, on constate une augmentation du biais de prédiction. Ce résultat est obtenu avec le modèle de Favrichon, pour lequel on a même pu quantifier ce biais de façon analytique, à l'aide d'un développement de Taylor.

D'une façon générale, on constate que les biais tendent à augmenter avec la durée des périodes de simulation. Pour les forêts qui ont une durée de vie importante (par ex. >100 ans), il est important de bien déterminer les sources de biais associés à l'utilisation de modèles de croissance.

Les prédictions des modèles de croissance forestière seront complétées par des intervalles de confiance indiquant la variabilité des prédictions.

Des travaux supplémentaires pourront être faits pour intégrer ces différentes sources de biais dans l'utilisation des modèles.

Références

Caswell, H. (2001) Matrix Population Models: Construction, Analysis and Interpretation, 2e édition, Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, 722 p.

Deleuze, C., D. Blaudez et J.C. Hervé. 1996. Ajustement d'un modèle hauteur-circonférence pour l'épicéa commun. Effet de la densité. Annales des Sciences Forestières 53 : 93-111.

Favrichon, V. (1996) Modélisation en forêt naturelle : les modèles à compartiments comme outils d'aide à l'aménagement forestier. *Bois et Forêts des Tropiques* 249(3) : 23-32.

Favrichon, V., 1998. Modeling the dynamics and species composition of tropical mixed-species uneven-aged natural forest: effects of alternative cutting regimes. Forest Science 44 (1), 113–124.

Fortin, M. et L. Langevin. 2010. ARTÉMIS-2009 : un modèle de croissance basé sur une approche par tiges individuelles pour les forêts du Québec. Mémoire de recherche forestière n˚ 156. Gouvernement du Québec. Ministère des Ressources naturelles et de la Faune. 48p.

Fortin, M. et L. Langevin. 2012. Stochastic or deterministic single-tree models: is there any difference in growth predictions? Annals of Forest Science 69: 271-282.

Gourlet-Fleury, S., G. Cornu, H. Dessard, N. Picard, and P. Sist. 2004. Forest dynamics models for practical management issues. Bois et Forêets des Tropiques 280 : 41–52.

Magnus, J. R., Neudecker, H., 2007. Matrix Differential Calculus with Applications in Statistics and Econometrics, 3rd Edition. Wiley series in probability and statistics. John Wiley and sons, Chichester.

Picard, N., Y. Nouvellet et M.L. Sylla. 2004. Relationship between plot size and the variance of the density estimator in West African savannas. Journal Canadien de la Recherche Forestière 34 : 2018-2026.

Picard, N., Bar-hen, A., Mortier, F., Chadœuf, J., 2009. Understanding the dynamics of an undisturbed tropical rain forest from the spatial pattern of trees. Journal of Ecology 91 (1), 97–108.

Pierrat, J.C., F. Houllier, J.C. Hervé et R. Gonzalez. 1995. Estimation de la moyenne des valeurs les plus élevées d'une population finie : application aux inventaires forestiers. Biometrics 51 : 679-686.

Pretzsch, H. (2009) Forest Dynamics, Growth and Yield: From Measurement to Model, Springer-Verlag, Berlin, 664 p.

Salas Gonzalez, R. 1995. Modélisation de l'évolution de la ressource du massif du pin maritime (Pinus pinaster) des Landes de Gascogne. Thèse de doctorat. ENGREF, Paris.

Salas Gonzalez, R., F. Houllier, B. Lemoine, and J. C. Pierrat. 1993. Représentativité locale des placettes d'inventaire en vue de l'estimation de variables dendrométriques de peuplement. Annales des Sciences Forestières 50: 469–485.

Salas Gonzalez, R., F. Houllier, B. Lemoine, and G. Pignard. 2001. Forecasting wood resources on the basis of national forest inventory data. Application to Pinus pinaster Ait. in southwestern France. Annals of Forest Science 58:785–802.

Vanclay, J. K. (1994) Modelling Forest Growth and Yield - Applications to Mixed Tropical Forests, CAB International, Wallingford}, 312 p.

Annexes

Annexe1 : Modèle ARTEMIS

Annxe1.1: Gestion du fichier et calcul du biais

```
donnee1=read.table("simul_aposteriori.csv",dec=".",sep=";",quote="\"",header=T,na.string="*")

donnee2=read.table("simul_apriori.csv",dec=".",sep=";",quote="\"",header=T,na.string="*")

x=donnee_apriori1[order(donnee_apriori1[,1]),]

y=donnee_aposteriori1[order(donnee_aposteriori1[,1]),]

apriori_2012=subset(x,T==2012)

apriori_2022=subset(x,T==2022)

apriori_2032=subset(x,T==2032)

apriori_2042=subset(x,T==2042)

apriori_2052=subset(x,T==2052)

apriori_2062=subset(x,T==2062)

aposteriori_2012=subset(y,T==2012)

aposteriori_2022=subset(y,T==2022)

aposteriori_2032=subset(y,T==2032)

aposteriori_2042=subset(y,T==2042)

aposteriori_2052=subset(y,T==2052)

aposteriori_2062=subset(y,T==2062)

difference_2012$nbTi_HA=apriori_2012$nbTi_HA - aposteriori_2012$nbTi_HA

difference_2022$nbTi_HA=apriori_2022$nbTi_HA-aposteriori_2022$nbTi_HA

difference_2032$nbTi_HA=apriori_2032$nbTi_HA-aposteriori_2032$nbTi_HA

difference_2042$nbTi_HA=apriori_2012$nbTi_HA-aposteriori_2042$nbTi_HA

difference_2052$nbTi_HA=apriori_2052$nbTi_HA-aposteriori_2052$nbTi_HA

difference_2062$nbTi_HA=apriori_2062$nbTi_HA-aposteriori_2062$nbTi_HA
```

Annexe1.2 : Calcul de la variance de la surface terrière initiale

```
donnee=read.table("simul_aposterioriMod.csv",dec=".",sep=";",quote="\"",header=T,na.string="*")
```

```
donnee1_2012=subset(donneeM,T==2012)
i=1
v=rep(0,499)
while(i<500)
{
don=subset(donnee1_2012,StrateID==i)
if(length(don$ST_HA)==1)
v[i]=0
else
if(length(don$ST_HA)>1)
v[i]=var(don$ST_HA)
else
v[i]=1
i=i+1
}
Var=v[(v!=1)]
variance <- aggregate(donnee1_2012$ST_HA, by=list(donnee1_2012$StrateID), FUN=var)
diff_2022=append(as.data.frame(difference_2022),as.data.frame(Var))
diff_2022=as.data.frame(diff_2022)
diff_2032=append(as.data.frame(difference_2032),as.data.frame(Var))
diff_2032=as.data.frame(diff_2032)
diff_2042=append(as.data.frame(difference_2042),as.data.frame(Var))
diff_2042=as.data.frame(diff_2042)
diff_2052=append(as.data.frame(difference_2052),as.data.frame(Var))
diff_2052=as.data.frame(diff_2052)
diff_2062=append(as.data.frame(difference_2062),as.data.frame(Var))
diff_2062=as.data.frame(diff_2062)
```

```
donneefinal=as.data.frame(rbind(diff_2022,diff_2032,diff_2042,diff_2052,diff_2062))
View(donneefinal)
for(i in 1:1065)
{
if(donneefinal$T[i]==2022)
donneefinal$T[i]=10
else
if(donneefinal$T[i]==2032)
donneefinal$T[i]=20
else
if(donneefinal$T[i]==2042)
donneefinal$T[i]=30
else
if(donneefinal$T[i]==2052)
donneefinal$T[i]=40
else
if(donneefinal$T[i]==2062)
donneefinal$T[i]=50
}
```

Annexe1.3 : Visualisation graphique de toutes les strates

```
library(lattice)
plot=xyplot(nbTi_HA~T,data=d,type='l',groups=StrateID,main="Nombre d’arbres en fonction du temps")
print(plot)
```

Annexe1.4: Modélisation

```
modele=lm(nbTi_HA~ Var+T+NbPl,data=donneefinal)
summary(modele)
require(lmtest)
```

```
durbin.watson(modele)

modele=gls(nbTi_HA~ V+T+NbPlac,data=donneefinal)

modele2 <- update(modele,correlation=corAR1(form=~1|StrateID))

modele3 <- update(modele2, weights=varPower())

modele4=gls(nbTi_HA~ V+T,data=donneefinal)

modele5 <- update(modele,correlation=corAR1(form=~1|StrateID))

modele6 <- update(modele2, weights=varPower())

intervals(modele6)

residus=residuals(modele6)

qqnorm(residus,ylab="Résidus",xlab="Quantiles théoriques",xlim=c(-3,3),ylim=c(-20,20))

qqline(residus)
```

Annexe2 : Modèle de Favrichon

```
donnee=read.table("paracou1992.txt",dec=".",sep="",quote="\"",header=T,na.string="*")

table(donnee$group)
```

Annexe2.1 : Calcul de l'effectif par classe de diamètre et par groupe d'essence

```
echantillon=function(x0,y0,dim){

d=subset(donnee,xnew-x0>0 & ynew-y0>0 & xnew-x0<dim & ynew-y0<dim)

return(d)                              }

br<-c(seq(from=10, to=60, by=5),200)

Matrice=function(d)    {

N=matrix(0,11,5)

for(i in 1:5)                  {

eff1 <- hist(d$D1992[d$group==i & d$plot==1],breaks=br,plot=FALSE)

N[,i]=eff1$counts          }

return(N)                      }
```

Annexe2.2 :Simulation

```
favrichon <- function(tps,N=matrix(0,11,5), A,plot=FALSE)
```

```
{
}
```

Annexe2.3

```
x=donnee$x
tirage1=function()  {
dx=rnorm(length(x),0,0.01)
xnew=x+dx
for(i in 1:length(x))  {
while(xnew[i]<0 || xnew[i]>250)  {
ddx=rnorm(1,0,0.01)
xnew[i]=x+ddx                    }
                    }
xnew=as.data.frame(xnew)
return(xnew)     }
y=donnee$y
tirage2=function()  {
dy=rnorm(length(y),0,0.01)
ynew=y+dy
for(i in 1:length(y))  {
while(ynew[i]<0 || ynew[i]>250)  {
ddy=rnorm(1,0,0.01)
ynew[i]=y+ddy                    }
                    }
ynew=as.data.frame(ynew)
return(ynew)     }
```

Annexe2.4 :

```
m2=rep(0,n^2)
```

```
for(i in 0:n-1)
{
for(j in 0:n)
{
d=echantillon(i*250/n,j*250/n,250/n)
N=Matrice(d)
r=favrichon(1,(NT[1,1]/NT[1,2])*N,1.5625,plot=F)
k=n*i+j+1
m2[k]=sum(r)
}
}
```

Annex2.5 : CALCUL DE LA MATRICE HESSIENNEDE Y EN Np(t)

```
Bg <- Ng <- rep(0,5)
   Ng <- colSums(N)
   Bg <- colSums(N*B)
   Nc <- sum(Ng)/N0
   Bc <- sum(Bg)/B0
  R5 <- c1[5]-c2[5]*Nc-c3[5]*Bc
V=rep(B,5)
H=matrix(1,55,55)
for(i in 1:55)
{
for(j in 1:55)
{
H[i,j]=((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0
}
}
```

```
for(i in 1:55)
{
H[i,i]=(((c2[5]/N0)+c3[5]*V[i]/B0)^2)*exp(R5)*(A^2)/A0
}
for(i in 1:55)
{
for(j in c(11,22,33,44,55))
{
s=j/11
if((b[i]-q1[s]*Bc >=1 )||(b[i]-q1[s]*Bc<=0))
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)
else
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)+q1[s]*V[i]*A/B0
}
}
                    #========
for(i in c(11,22,33,44,55))
{
for(j in 1:55)
{
s=i/11
if((b[i]-q1[s]*Bc >=1 )||(b[i]-q1[s]*Bc<=0))
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)
else
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)+q1[s]*V[j]*A/B0
}
}
```

```
                #=========
for(i in c(11,22,33,44,55))
{
for(j in c(11,22,33,44,55))
{
s=i/11
l=j/11
if((b[i]-q1[s]*Bc >=1 ||b[i]-q1[s]*Bc<=0) & (b[j]-q1[l]*Bc >=1 ||b[j]-q1[l]*Bc<=0))
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)
else
if(b[i]-q1[s]*Bc >=1 ||b[i]-q1[s]*Bc<=0 &b[j]-q1[l]*Bc <1 & b[j]-q1[l]*Bc>0)
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)+(V[i]*q1[l])*A/B0
else
if(b[i]-q1[s]*Bc <1 & b[i]-q1[s]*Bc>0 & b[j]-q1[l]*Bc >=1 || b[j]-q1[l]*Bc<=0)
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)+(q1[s]*V[j])*A/B0
else
H[i,j]=(((c2[5]/N0)+c3[5]*V[i]/B0)*((c2[5]/N0)+c3[5]*V[j]/B0)*exp(R5)*(A^2)/A0)+((q1[s]*V[j])+(V[i]*
q1[l]))*A/B0
}
}
                #=========
for(i in c(11,22,33,44,55))
{
s=i/11
if(b[i]-q1[s]*Bc >=1 || b[i]-q1[s]*Bc<=0)
H[i,i]=((((c2[5]/N0)+c3[5]*V[i]/B0)^2)*exp(R5)*(A^2)/A0)
else
H[i,i]=((((c2[5]/N0)+c3[5]*V[i]/B0)^2)*exp(R5)*(A^2)/A0)+(2*q1[s]*V[i]*A/B0)
```

```
}
calcul=function(n,c,g)
{
Na=rep(0,n^2)
for(i in 0:n-1)
{
for(j in 0:n)
{
d=echantillon(i*250/n,j*250/n,250/n)
N=Matrice(d)
N=((NT[1,1]/NT[1,2])*N)/(((250/n)^2)/10000)
k=n*i+j+1
Na[k]=N[c,g]
}
}
return(Na[1:n^2])
}
```

Annexe2.6 : effectif d'arbre dans les 11 classes de diamètres

```
varcov=function(n,g)
{
MNa=matrix(0,n^2,11)
for(c in 1:11)
{
MNa[,c]=calcul(n,c,g)
}
return(MNa)
}
```

```
plot(eff~A,col="red",type="l",ylim=c(586,592),log="x",data=moy)

lines(approx~A,type="l",col="blue",data=tac)

legend(1,590,legend=c("Estimation Approché","Moyenne Simulée"),lwd=3,col=c("blue","red"))
```