

Représentativité locale des placettes d'inventaire en vue de l'estimation de variables dendrométriques de peuplement

R Salas González ¹, F Houllier ¹, B Lemoine ², JC Pierrat ¹

avec la collaboration technique de F Jean ²

¹ INRA-ENGREF, laboratoire de recherches en sciences forestières, unité dynamique des systèmes forestiers, 14, rue Girardet, 54042 Nancy cedex ;

² INRA, station de recherches forestières, laboratoire croissance et production, domaine l'Hermitage, Pierroton, 33610 Cestas Principal, France

(Reçu le 3 novembre 1992; accepté le 10 mai 1993)

Résumé — L'étude de la représentativité locale des petites placettes d'inventaire au sein d'un peuplement est abordée en simulant des placettes du type de celles de l'inventaire forestier national (IFN) français au sein de peuplements cartographiés de pin maritime (*Pinus pinaster*). Les variables dendrométriques étudiées sont le nombre de tiges à l'hectare N , la surface terrière G , la circonférence moyenne C_g , et la circonférence dominante C_0 . Pour chacune d'entre elles, on estime la variance et le biais systématique d'estimation dus à la taille des placettes ainsi qu'à la structure des peuplements. Pour les jeunes peuplements, la variance de N et G est élevée, elle est plus modérée pour C_g et C_0 , mais il y a une sous-estimation systématique de C_0 . Dans le cas des peuplements plus âgés et de placettes simulées de grande surface (0,05 à 0,10 ha), il semble que la taille supérieure des placettes d'inventaire ainsi que la régularisation de la structure du peuplement expliquent l'amplitude moindre des erreurs aléatoires et du biais. Ces résultats suggèrent qu'il n'est pas possible de projeter directement les données issues d'un inventaire régional au moyen d'un modèle de croissance sans avoir vérifié préalablement que le protocole d'inventaire est compatible avec celui utilisé pour l'acquisition des données ayant servi à la construction du modèle.

***Pinus pinaster* : pin maritime / échantillonnage / inventaire / modélisation / recensabilité**

Summary — **Effect of the design and the size of inventory plots on the estimation of stand characteristics.** This paper aims at estimating the variance and bias of the estimation of characteristics of *Pinus pinaster* stands that are due to survey design and especially to plot size. Small survey plots (0.01–0.07 ha) similar to those used by the French National Forest Inventory (IFN) are simulated inside 39 large experimental plots, located in the Landes of Gascogne region in France. The stand characteristics studied are: N (number of stems/ha), G (basal area), C_g (average girth) and C_0 (dominant girth). The results show that small plots are not representative of the local stand condi-

tions: the variance of N and G is high and much larger than that of C_g and C_0 ; nevertheless, only C_0 is systematically biased (underestimated). When the stands are older and when the simulated plots are larger (0.05–0.10 ha), our results suggest that the larger plot size and the regularization of the structure of the stand induced by silvicultural practices diminish both the variance and the bias of the random error. These results indicate that before using a growth model in order to project regional inventory data it is necessary to check the compatibility between the data used to build the model and the data gathered in an operational survey.

maritime pine / sampling method / inventory / modelling / census

INTRODUCTION

Les qualités d'un modèle de croissance dépendent conjointement des connaissances biologiques, des méthodes mathématiques et statistiques employées ainsi que de la nature et de la qualité des données (Houllier, 1990). Cet article concerne l'origine des données et les procédures d'échantillonnage : nous nous intéressons, d'une part, à la définition de l'unité statistique — c'est-à-dire à l'échelle à laquelle sont rapportées les mesures utilisées pour construire, valider et/ou appliquer un modèle — et, d'autre part, aux problèmes générés par l'existence d'un seuil de recensabilité.

Ces problèmes sont en effet apparus importants dans le cadre d'un projet qui vise à simuler l'évolution du massif des Landes de Gascogne à partir du modèle de peuplement construit par Lemoine (1991) pour le pin maritime (*Pinus pinaster*). Le modèle de Lemoine a initialement été conçu pour représenter l'évolution d'un peuplement (d'une parcelle) installé sur une station plus ou moins fertile et soumis à différents traitements. Il a été construit à partir d'un réseau de placettes semi-permanentes, de surface comprise entre 0,10 et 0,50 ha, complété par des dispositifs expérimentaux spécifiques (expériences sur la concurrence, la fertilisation, l'entretien du sol et comparaison de matériels génétiques différents) et par des analyses de tige. Il se compose de 3 «lois» de

croissance non linéaires qui permettent de prédire les accroissements de la hauteur dominante (I_{H0}), de la surface terrière moyenne de l'arbre dominant (I_{g0}) et de la surface terrière de l'arbre moyen (I_g) en fonction de l'âge, de la fertilité et du nombre de tiges. Il est complété par des relations dendrométriques qui permettent d'estimer la hauteur moyenne (H_g), le volume moyen et la circonférence moyenne (C_g).

La méthode envisagée pour représenter l'évolution du massif consiste à projeter les données de l'inventaire forestier national (IFN) français au moyen du modèle de Lemoine, selon une procédure qui a déjà été employée par Maugé (1979, 1981). Cela revient à simuler différents scénarios de gestion et à utiliser les équations du modèle de croissance en les appliquant [les scénarios et les équations], placette par placette, aux données de l'IFN considérées comme des conditions initiales. Or les données de l'IFN (IFN, 1985) diffèrent de celles utilisées pour la construction du modèle par plusieurs aspects. Elles sont notamment basées sur des placettes dont la surface est sensiblement inférieure à celle des placettes utilisées par Lemoine. De plus, les individus ne sont considérés que si leur circonférence dépasse le seuil de recensabilité (tableau I).

La question posée est donc de savoir si un modèle de croissance ajusté à l'échelle du peuplement peut être appliqué à une échelle plus grande, ou encore si chacune

Tableau I. Placettes de l'IFN et dimension des arbres mesurés.

Rayon (m)	Domination	Circonférence à 1,5 m	Surface (ha)
–	Non recensable (NR)	< 24,5 cm	–
6	Petit bois (PB)	24,5 – 54,4 cm	0,0113
9	Moyen bois (MB)	54,5 – 94,4 cm	0,0254
15	Gros bois (GB)	> 94,5 cm	0,0707

des placettes de l'IFN, prise isolément, est localement représentative du peuplement qui l'entoure. Cette question est différente du problème qui est usuellement traité dans les manuels d'inventaire, à savoir la représentativité globale d'un ensemble de placettes vis-à-vis d'une forêt ou d'un massif. Elle a notamment été étudiée par Häggglund (1982) qui signale que les valeurs des caractéristiques observées sur des placettes de 0,1 ha sont plus variables que les valeurs de ces mêmes caractéristiques vues à l'échelle du peuplement (1 ha). Cette question est aussi abordée par Chevrou et al (1988) à propos du diagnostic de la richesse des taillis sous futaie à partir des données de l'IFN : il s'avère que des données acquises sur une surface comprise entre 0,01 et 0,07 ha ne permettent pas de juger avec fiabilité de la densité des réserves dans le peuplement avoisinant.

Cette question peut être formalisée dans le cadre de la théorie des variables régionalisées (Bouchon, 1979). Matheron (1965, p 129-133) définit en effet la notion de variance d'extension comme la variance de l'erreur faite en «étendant» une valeur observée localement à la zone qui l'entoure. La valeur de cette variance dépend de la structure spatiale de la variable étudiée et en conséquence son calcul nécessite de disposer d'un modèle probabiliste de cette structure (ie le demi-

variogramme). Cette approche n'a pas été développée car nous ne disposons pas de modèles généraux du demi-variogramme ponctuel des variables dendrométriques usuelles (nombre de tiges, surface terrière...).

Dans une perspective de modélisation de la croissance, Hann et Azim (1991) ont noté que les prédictions obtenues avec un modèle d'arbre indépendant des distances sont très sensibles au plan d'échantillonnage (nombre, arrangement spatial et surface des placettes) utilisé pour caractériser un peuplement ; cette sensibilité provient surtout des indices de densité du peuplement ou de statut social des arbres. Hann et Azim introduisent ainsi la notion d'erreur différentielle de protocole («*differential design error*») qui est égale à la différence entre les estimations obtenues en appliquant au même point deux méthodes différentes d'échantillonnage (par exemple des placettes plus ou moins étendues) et qui ne doit donc pas être confondue avec l'erreur aléatoire usuelle (c'est-à-dire la différence entre la valeur observée sur l'unité échantillonnée et la valeur moyenne théorique).

Dans le même ordre d'idées, Shugart et West (1979) ont utilisé un modèle de croissance pour des peuplements hétérogènes et ont appliqué ce modèle à des petites placettes de taille variable. Ils observent que, si les placettes sont trop petites

(moins de 0,04 ha) ou trop grandes (plus de 0,20 ha), le comportement qualitatif des prédictions n'est pas satisfaisant car certains phénomènes de concurrence sont mal représentés. Ils en déduisent que la taille des placettes utilisées pour l'étude de la croissance des peuplements doit dépendre du «grain» du peuplement, c'est-à-dire de la répartition spatiale et de la taille moyenne des arbres qui les composent.

Dans 2 peuplements expérimentaux de chêne, Houllier (1986, p 50-53) a par ailleurs observé que la corrélation temporelle entre mesures successives de la surface terrière est sensiblement affectée par la taille des placettes de mesure. Il a aussi montré que la théorie des variables régionalisées permet, là aussi, de rendre compte de cette observation par le phénomène de régularisation.

Cet article vise à étudier la représentativité locale des placettes de l'IFN en évaluant empiriquement la variance d'extension et le biais éventuel dus à la taille réduite et à la forme des placettes. Notre démarche a consisté à utiliser des placeaux de pin maritime de l'Office national des forêts (ONF) et de l'Institut national de la recherche agronomique (INRA) comme références, puis à simuler des placettes de type IFN au sein de ces placeaux et à caractériser l'écart entre les variables dendrométriques estimées sur ces deux supports. Un deuxième aspect de notre travail a consisté à étudier l'influence du seuil de recensabilité adopté par l'IFN sur l'estimation de la densité, de la surface terrière et de la circonférence moyenne.

MATÉRIEL ET MÉTHODES

Afin de bien distinguer les problèmes liés à la taille des placettes de ceux associés au seuil de recensabilité, nous avons choisi d'aborder ces deux aspects successivement et indépendamment :

– au paragraphe «Étude des variables dendrométriques par simulation de placettes IFN», nous supposons que tous les arbres sont recensés et nous simulons des placettes IFN au sein de placeaux de grande surface ;

– au paragraphe «Effet du seuil de recensabilité», nous analysons l'effet du seuil de recensabilité sur les variables dendrométriques à l'échelle des placeaux (sans simuler de petites placettes IFN).

Étude des variables dendrométriques par simulation de placettes IFN

Placeaux INRA et ONF de référence

Nous avons utilisé un dispositif expérimental de l'INRA («parcelle U», 12 placeaux) et 27 placeaux cartographiés pied à pied, appartenant à des peuplements jeunes ou adultes dispersés dans les Landes de Gascogne. Ces placeaux ont une surface comprise entre 0,12 et 0,64 ha. Pour chaque placeau nous disposons du nombre de tiges à l'hectare (N), de l'âge du peuplement (A), de la liste et des coordonnées spatiales des arbres, ainsi que de leur circonférence à 1,30 m (C) et de leur hauteur (H) pour un sous-échantillon. Les peuplements ont une géométrie liée à leur mode de création : plantation en ligne, semis en ligne, semis en bande et régénération naturelle.

Plantation en ligne

Cette géométrie a été étudiée dans les 12 placeaux de la parcelle U. Il s'agit d'un peuplement de 14 ans (en 1988) où différents facteurs sont étudiés : l'espacement à la plantation (2 m x 2 m ou 4 m x 4 m), l'entretien du sol (présence ou absence) et l'origine génétique (deux variétés). Les placeaux ont une surface comprise entre 0,124 8 et 0,154 6 ha. Les mesures dendrométriques ont été effectuées en 1984 et 1988 (figs 1a et 1b).

Semis en ligne

Il s'agit de 20 placeaux appartenant à des peuplements âgés de 11 à 27 ans où la distance entre les lignes varie de 3,5 à 4,6 m. Cinq placeaux ont une surface de 0,36 ha, pour les autres elle vaut 0,16 ha (fig 2a).

Semis en bande

Il s'agit de 3 placeaux appartenant à des peuplements âgés de 20 ans où la distance entre les bandes varie de 5,7 à 7,1 m. Tous ces placeaux ont une surface de 0,36 ha (fig 2b).

Régénération naturelle

Les arbres sont distribués de manière quasi aléatoire. Les 4 placeaux ont une surface de 0,64 ha et appartiennent à des peuplements âgés de 40 à 60 ans (fig 3).

Simulation de placettes IFN dans les placeaux INRA et ONF

Les «points-terrain» de l'IFN sont constitués par un système de trois placettes circulaires concentriques de 6, 9 et 15 m de rayon, sur lesquelles sont mesurés des arbres de dimension différente (IFN, 1985 ; Chevrou, 1993). Elles servent à l'estimation des nombres de tiges, dimensions, volumes et accroissements des arbres et des peuplements (fig 4, tableau I).

Dans chacun des grands placeaux INRA-ONF nous avons simulé 100 points-terrain de type IFN (fig 3). Les coordonnées spatiales du centre de chaque placette ont été obtenues en

simulant 2 variables aléatoires de distribution uniforme. Lorsque la placette IFN débordait du placeau étudié nous avons utilisé la méthode de «l'effet miroir» qui revient à comptabiliser plusieurs fois (de 2 à 4) les arbres de bordure : cette technique permet d'obtenir des estimations non biaisées des caractéristiques moyennes des peuplements (Schmid-Haas, 1982) ; elle a néanmoins un effet indirect dans le cas de peuplements peu étendus puisqu'elle augmente la variance d'estimation.

Calcul des variables dendrométriques

Les variables étudiées sont le nombre de tiges/ha (N), la surface terrière à l'hectare (G), la circonférence quadratique moyenne (Cg) ($= \sqrt{4 \pi G/N}$) et la circonférence dominante ($C\phi$), définie comme la circonférence quadratique moyenne des 100 plus gros arbres par ha. Pour toutes les dates de mesure nous avons calculé sur chaque placeau INRA-ONF, la valeur de chacune de ces variables, notée μ (ie $\mu(N)$, $\mu(G)$, $\mu(Cg)$ et $\mu(C\phi)$). Pour les placettes simulées, nous avons estimé les mêmes caractéristiques dendrométriques en incluant la totalité des arbres présents (c'est-à-dire sans appliquer le seuil de recensabilité de l'IFN).

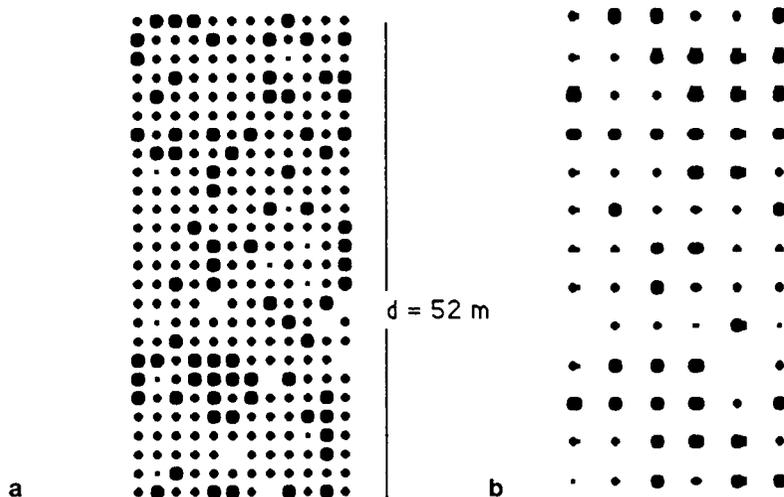


Fig 1. «Parcelle U» : peuplement de 14 ans, chaque arbre est représenté par un cercle de diamètre proportionnel à la circonférence mesurée à 1,30 m. **a.** placeau 1 : espacement 2 m x 2 m, surface = 0,12 ha, $N = 2\,444$ tiges/ha, $G = 25,3$ m²/ha, $Cg = 36,1$ cm, $C\phi = 47,1$ cm. **b.** placeau 2 : espacement 4 m x 4 m, surface = 0,12 ha, $N = 609$ tiges/ha, $G = 7,9$ m²/ha, $Cg = 40,3$ cm, $C\phi = 49,6$ cm.

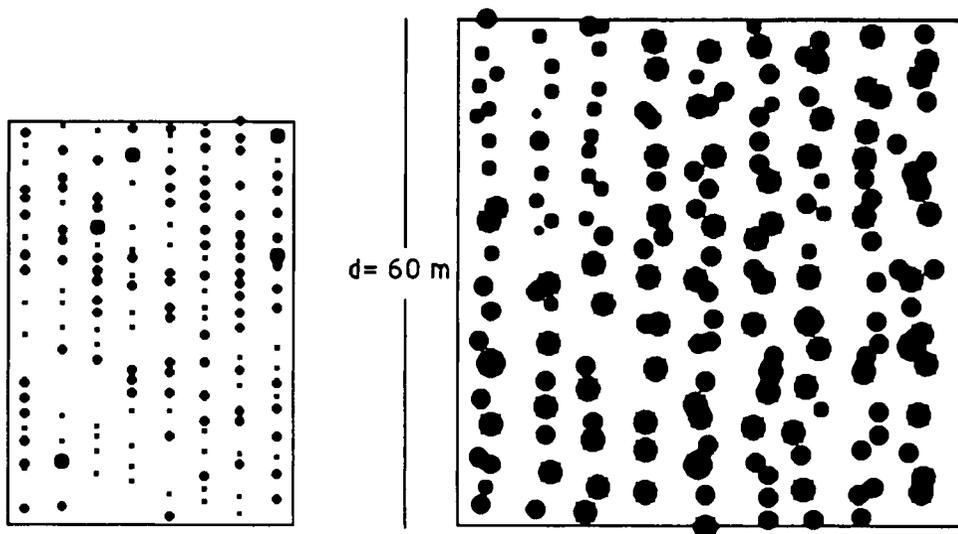


Fig 2a. Placeau «Commensacq II» : peuplement de 11 ans issu d'un semis en ligne, surface = 0,16 ha, $N = 956$ tiges/ha, $G = 7,5$ m²/ha, $C_g = 31,3$ cm, $C_\emptyset = 45,9$ cm. **2b.** Placeau «Le Bruillet» : peuplement de 20 ans issu d'un semis en bande, surface = 0,36 ha, $N = 492$ tiges/ha, $G = 24,5$ m²/ha, $C_g = 79,1$ cm, $C_\emptyset = 96,2$ cm.

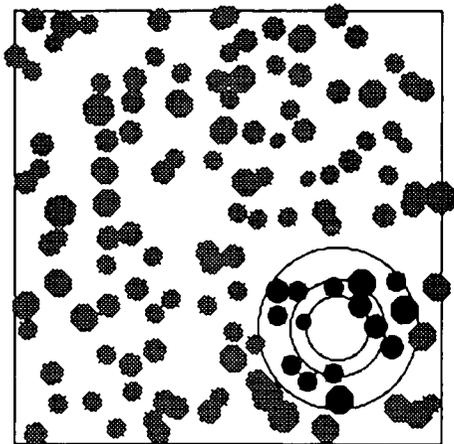


Fig 3. Placeau «Pissos 3» : peuplement de 60 ans issu d'une régénération naturelle, surface = 0,64 ha, $N = 200$ tiges/ha, $G = 30,3$ m²/ha, $C_g = 138,1$ cm, $C_\emptyset = 153,4$ cm, longueur du côté du placeau : 80 m. Placette de type IFN dessinée en bas à droite.

La circonférence dominante constitue un cas particulier, puisque cette variable est une entrée du modèle de Lemoine qui n'est pas directement calculée par l'IFN. Pour l'estimer sur un point-terrain IFN simulé, nous avons donc procédé comme suit : (i) les arbres ont été triés selon leur circonférence décroissante ; (ii) à chaque arbre, nous avons associé son poids (c'est-à-dire, le nombre de tiges qu'il représentait par ha, selon sa catégorie de circonférence), ainsi que le cumul de son poids et des poids des arbres plus gros ; (iii) nous avons ensuite retenu les p plus gros arbres tels que leur poids cumulé soit égal à 100 (pour que le cumul soit exactement 100, le poids donné au p^e arbre était le complément à 100 du cumul des poids des $p-1$ plus gros arbres) ; (iv) la circonférence dominante a ensuite été estimée comme la circonférence quadratique moyenne pondérée des p plus gros arbres.

À partir des 100 simulations de points-terrain IFN effectuées nous avons calculé les statistiques suivantes pour chaque variable y ($y = C_\emptyset, C_g, G$ et N) :

- la moyenne empirique : $m(y)$;
- l'écart type empirique de y et celui de $m(y)$: $s(y)$ et $s(m(y)) = s(y)/10$;

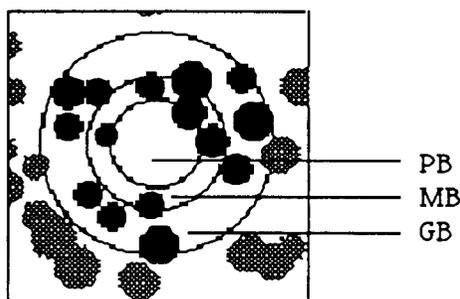


Fig 4. Placette de mesure de l'IFN avec les 3 sous-placettes utilisées pour le lever des petits, moyens et gros bois.

- le coefficient de variation empirique : $CV(y) = s(y)/\mu(y)$;
- les valeurs extrêmes simulées : $Min(y)$ et $Max(y)$;
- les biais empiriques absolu et relatif : $Ba(y) = m(y) - \mu(y)$ et $Br(y) = Ba(y)/\mu(y)$;
- la valeur $t(y) = (m(y) - \mu(y)) / s(m(y))$ qui suit approximativement une loi de Student à 98 degrés de liberté.

Pour deux variables, y_1 et y_2 , nous avons aussi calculé leur covariance et leur corrélation empiriques :

$$c(y_1, y_2) \text{ et } r(y_1, y_2) = \frac{c(y_1, y_2)}{s(y_1) s(y_2)}$$

Évolution du biais selon la surface de la taille de placette simulée

Nous avons utilisé le grand plateau Pissos3 de 0,64 ha, âgé de 60 ans et issu de régénération naturelle, pour analyser l'évolution du biais empirique d'estimation de $C\emptyset$ et de Cg lorsque la surface de la placette d'inventaire augmente de 0,01 à 0,10 ha.

Effet du seuil de recensabilité

Position du problème

L'IFN estime les caractéristiques des peuplements en ne tenant compte que des arbres re-

cenables (le seuil de recensabilité correspond à une circonférence de 24,5 cm, tableau 1) : les variables estimées sont donc biaisées. Ce problème qui n'est pas réellement de nature statistique n'est évidemment sensible que dans les jeunes peuplements de moins de 15 à 25 ans, selon la fertilité de la station et la densité du peuplement. Plus précisément, nous pouvons distinguer trois situations pour les futaies équiennes de pin maritime.

Lorsque les peuplements sont très jeunes (moins de 5 à 10 ans), aucun arbre n'est recensable et il est impossible de définir les variables dendrométriques de peuplement à partir du protocole de l'IFN.

Vient ensuite une période où une fraction du peuplement est recensable : il est alors possible de définir les variables N_r , G_r , Cg_r et $C\emptyset_r$ relatives aux seuls arbres recensables. Du fait de la non mesure des petits arbres, ces variables vérifient les relations suivantes : $N_r \leq N$, $G_r \leq G$, $Cg_r \geq Cg$; quant à $C\emptyset_r$, nous pouvons considérer que $C\emptyset_r = C\emptyset$, puisque les arbres dominants sont les premiers à franchir le seuil de recensabilité et que la plus petite placette de l'IFN a une surface d'environ 0,01 ha.

Enfin, au-delà de 15-25 ans, la totalité du peuplement est recensable : $N_r = N$, $G_r = G$ et $Cg_r = Cg$.

Pour simuler ultérieurement l'évolution du massif aquitain avec le modèle de Lemoine, il est nécessaire d'estimer les caractéristiques N , G et Cg et $C\emptyset$ alors que seules N_r , G_r , Cg_r et $C\emptyset_r$ sont, éventuellement, connues. Nous nous sommes intéressés ici au deuxième cas énoncé ci-dessus en cherchant à prédire les valeurs N , G et Cg à partir des valeurs de N_r , G_r , Cg_r et $C\emptyset_r (= C\emptyset)$.

Principe de la méthode d'analyse

Nous avons défini des «taux de recensabilité» N_r/N , G_r/G et Cg_r/Cg et nous avons analysé leur variation en fonction de la circonférence dominante ($C\emptyset$). En effet, $C\emptyset$ est définie par l'IFN dès qu'au moins un arbre est recensable ; de plus, cette variable permet d'intégrer implicitement divers facteurs tels l'âge, la fertilité et la vigueur génétique (la hauteur dominante aurait sans doute aussi été une variable pertinente, mais elle n'était pas disponible dans tous les placetages étudiés).

Afin de ne pas interférer avec les problèmes de surface des placettes déjà évoqués, l'étude de l'influence du taux de recensabilité a été menée sur les grands placeaux de surface supérieure à une dizaine d'ares et nous n'avons pas simulé les 3 placettes concentriques de l'IFN : les variables dendrométriques ont été calculées à l'échelle du placeau en considérant tous les arbres (y compris les arbres non recensables au sens de l'IFN).

Placeaux supplémentaires

Nous avons utilisé les placeaux INRA et ONF déjà présentés ainsi que des données complémentaires, fournies par Lemoine et venant de 37 placeaux non cartographiés situés à Pissos et Rousset.

Parcelles de Pissos

Elles incluent :

– 2 vieux placeaux âgés de plus de 40 ans, issus de régénération naturelle, ayant une surface de 0,64 ha et une densité proche de 200 tiges/ha ;

– 25 placeaux, de moins de 20 ans, issus, pour quelques-uns, de semis en bande ou, pour la plupart, de semis en ligne, et où la densité varie entre 1 189 et 3 060 tiges/ha ; leur surface varie entre 0,09 et 0,19 ha et ils ont été mesurés plusieurs fois entre 4 et 16 ans.

Parcelle de Rousset

Cette parcelle de 2,618 ha est issue de plantation. Elle contient 10 placeaux dont la surface varie entre 0,10 et 0,40 ha et dans lesquels 4 modalités de plantation (1 275 et 2 500 tiges/ha, maille rectangulaire ou carrée) et divers matériels génétiques ont été testés. Les mesures dendrométriques ont été effectuées à 3, 8, 12, 16 et 19 ans.

RÉSULTATS

Éléments théoriques sur le biais et la variance des estimations IFN

Plaçons-nous dans un peuplement (un placeau). Soit y une variable dendrométri-

que quelconque ($C\emptyset$, Cg , G ou M). La valeur moyenne de y évaluée sur l'ensemble du peuplement est notée $\mu(y)$. Soit y_s l'estimation de y obtenue sur une placette de faible surface S . Notons $\mu_s(y) = E[y_s]$ l'espérance de y_s et $\sigma_s^2(y) = \text{Var}[y_s]$ sa variance. Notons aussi $C_s(y_1, y_2) = \text{Cov}[y_{1_s}, y_{2_s}]$ la covariance des deux variables dendrométriques y_{1_s} et y_{2_s} .

Quelle que soit la surface S de la placette, les estimations de G et N sont théoriquement non biaisées : $\mu_s(y) = \mu(y)$ (pour $y = G, M$). Seules $C\emptyset$ et Cg sont susceptibles d'être biaisées.

Pour $y = Cg$, on a :

$$Cg_s = \sqrt{4\pi} \frac{G_s}{N_s} = \mu(Cg) \left\{ 1 + \frac{G_s - \mu(G)}{\mu(G)} \right\}^{1/2} \left\{ 1 + \frac{N_s - \mu(N)}{\mu(N)} \right\}^{-1/2}$$

Il est possible d'obtenir une évaluation approximative du biais de Cg et de la variance d'estimation en faisant quelques hypothèses complémentaires. Notons : $\varepsilon_N = N_s - \mu(N)$ et $\varepsilon_G = G_s - \mu(G)$. Supposons que $\varepsilon_N / \mu(N)$ et $\varepsilon_G / \mu(G)$ sont petits devant 1. Par un développement limité au deuxième ordre, il vient alors :

$$Cg_s \approx \mu(Cg) \left\{ 1 + \frac{\varepsilon_G}{2\mu(G)} - \frac{\varepsilon_N}{2\mu(N)} - \frac{\varepsilon_G^2}{8\mu(G)^2} + \frac{3\varepsilon_N^2}{8\mu(N)^2} - \frac{\varepsilon_G\varepsilon_N}{4\mu(G)\mu(N)} \right\}$$

d'où $\mu_s(Cg) \approx \mu(Cg)$

$$\left\{ 1 - \frac{\sigma_s^2(G)}{8\mu(G)^2} + \frac{3\sigma_s^2(N)}{8\mu(N)^2} - \frac{C_s(N, G)}{4\mu(G)\mu(N)} \right\}$$

Sous ces hypothèses, il apparaît donc que le biais de Cg est du deuxième ordre en ε_N et ε_G . De plus, si on admet que G_s et N_s sont approximativement proportionnels, on observe que les trois termes du biais tendent à s'annuler (le biais s'exprime alors en fonction des moments d'ordre 3 ou plus de ε_N et ε_G).

On obtient une approximation de la variance de Cg_s en procédant de la même manière :

$$\begin{aligned}\sigma_s^2(Cg) &\approx \mu(Cg)^2 \left\{ \frac{\sigma_s^2(N)}{\mu(N)^2} - \frac{C_s(N,G)}{\mu(G)\mu(N)} \right\} \\ &= \mu(Cg)^2 \frac{\sigma_s^2(N) b_{0'G_s/N_s}}{\mu(N)^2 \mu(G)}\end{aligned}$$

où $b_{0'G_s/N_s}$ est la constante obtenue quand on prédit G_s à partir de N_s par régression linéaire simple. Remarquons qu'en général le terme $b_{0'G_s/N_s} / \mu(G)$ est petit devant 1.

De ces considérations théoriques, on déduit (i) que la variance d'estimation de Cg est en valeur relative sensiblement plus faible que celle de N ou G et (ii) que le biais d'estimation de Cg existe, mais que son ordre de grandeur est faible. En fait, comme nous le verrons au paragraphe «Variabilité des estimations IFN autour de la moyenne estimée», les hypothèses nécessaires à ces calculs ne sont pas totalement vérifiées : $\varepsilon_G / \mu(G)$ et $\varepsilon_N / \mu(N)$ ne sont pas toujours petits devant 1, si bien que ces résultats théoriques ne sont qu'indicatifs.

La situation est différente pour $C\emptyset$ car il ne s'agit pas d'une statistique moyenne mais d'une statistique basée sur les plus gros arbres échantillonnés.

Variabilité des estimations IFN autour de la moyenne estimée

Quelques résultats détaillés au niveau du placeau sont présentés au tableau II. Les autres résultats sont fournis de manière plus synthétique sous forme de graphiques (figs 5 à 8).

La variabilité intra-peuplement des estimations IFN est très importante pour les grandeurs obtenues par sommation, N et G (tableau II). Leur coefficient de variation a tendance à être plus élevé dans les peuplements jeunes — notamment dans les peuplements issus de semis en ligne ($CV(y)$ compris entre 11 et 50%) et dans la plantation à espacement 4 m x 4 m ($CV(y)$ compris entre 10 et 44% selon les placeaux) — que dans les peuplements issus de semis en bande et âgés ($CV(y)$ compris entre 11 et 27%). D'autre part, les valeurs du coefficient de corrélation restent assez faibles pour les peuplements créés par plantation (de 0,32 à 0,40 ; fig 9a), alors que dans les peuplements ayant une autre géométrie la corrélation varie de 0,50 à 0,90 (figs 9b à 9d). Il est difficile d'attribuer ces phénomènes à un facteur particulier puisque la géométrie des peuplements est liée à la date de leur installation, donc à leur âge, et que la surface des placettes utilisées par l'IFN augmente quand les arbres grossissent.

Cette variabilité très forte est bien traduite par les valeurs extrêmes simulées dont le rapport $Max(y) / Min(y)$ varie de 2 (ex Château Jauge au tableau II) à 10 (ex Commensacq II). En revanche, la variabilité est beaucoup plus modérée pour Cg et $C\emptyset$ (tableau II) : dans les plantations en ligne, $CV(y)$ varie de 9 à 21% ; pour les autres placeaux $CV(y)$ est compris entre 4 et 11% et le rapport entre les valeurs extrêmes simulées varie autour de 1,5.

Tableau II. Propriétés statistiques des estimations basées sur les placettes IFN simulées au sein des placeaux INRA-ONF de référence (illustration des résultats obtenus pour différents placeaux et différentes variables).

Placeaux	Age (ans)	Géométrie	Variable (y)	$\mu(y)$	$m(y)$	Min(y)	Max(y)	s(y)	Cov (N,G)	Corr (N,G)
Parcelle U (6 placeaux en 1988)	14	Plantation densité 2 x 2 m	N	2 448,5	2 455,3	1 945,2	2 829,4	120,8	175,4	0,32
			G	22,6	22,5	10,8	33,5	4,4		
			Cg	34,0	33,8	23,5	40,5	3,3		
			Cø	45,9	43,7	32,9	51,5	3,6		
Parcelle U (6 placeaux en 1988)	14	Plantation densité 4 x 4 m	N	614,1	612,0	353,7	795,8	67,2	57,1	0,40
			G	6,5	6,6	1,9	12,9	2,1		
			Cg	36,3	36,2	22,7	48,5	5,5		
			Cø	45,8	43,7	30,8	53,9	5,5		
Commensacq II (1 placeau)	11	Semis en ligne	N	956,2	983,2	176,8	1 768,0	308,2	955,7	0,90
			G	7,5	7,7	0,9	20,4	3,5		
			Cg	31,3	30,9	21,5	38,1	3,5		
			Cø	45,9	42,7	30,4	50,6	4,8		
Le Bruillet (1 placeau)	20	Semis en bande	N	491,7	501,4	264,1	727,0	87,6	209,7	0,50
			G	24,5	25,3	12,5	35,5	4,7		
			Cg	79,1	79,8	55,3	89,2	6,9		
			Cø	96,2	94,7	66,2	101,2	6,4		
Château Jauge INRA (1 placeau)	40	Régé- nération naturelle	N	278,1	271,6	198,0	355,2	31,1	70,8	0,75
			G	27,9	27,3	19,8	33,8	3,0		
			Cg	112,3	112,5	103,0	122,5	4,5		
			Cø	124,1	123,3	114,0	134,6	5,0		

Dans le cas de la parcelle U, les résultats ont été obtenus par moyenne des estimations pour chacun des 6 placeaux (date = 1988).

Biais des estimations basées sur les placettes IFN

Conformément aux déductions faites au paragraphe «Éléments théoriques sur le biais et la variance des estimations IFN», il apparaît que seule la circonférence dominante est biaisée. Ce biais est constamment négatif (figs 6 à 8) et presque toujours significatif au seuil de 5%. Son amplitude varie entre -1 et -4 cm, soit de -1 à -10% selon les placeaux.

Afin de pouvoir, éventuellement, le corriger, nous avons cherché à prédire ce biais à partir de variables dendrométriques. Nous observons généralement que plus les peuplements sont jeunes, plus le biais est important en valeur absolue et surtout en valeur relative, notamment dans les peuplements semés ou plantés en ligne où la dispersion est la plus grande. La faiblesse apparente du biais pour les peuplements âgés peut être due aussi bien à une régularisation de la structure par les prati-

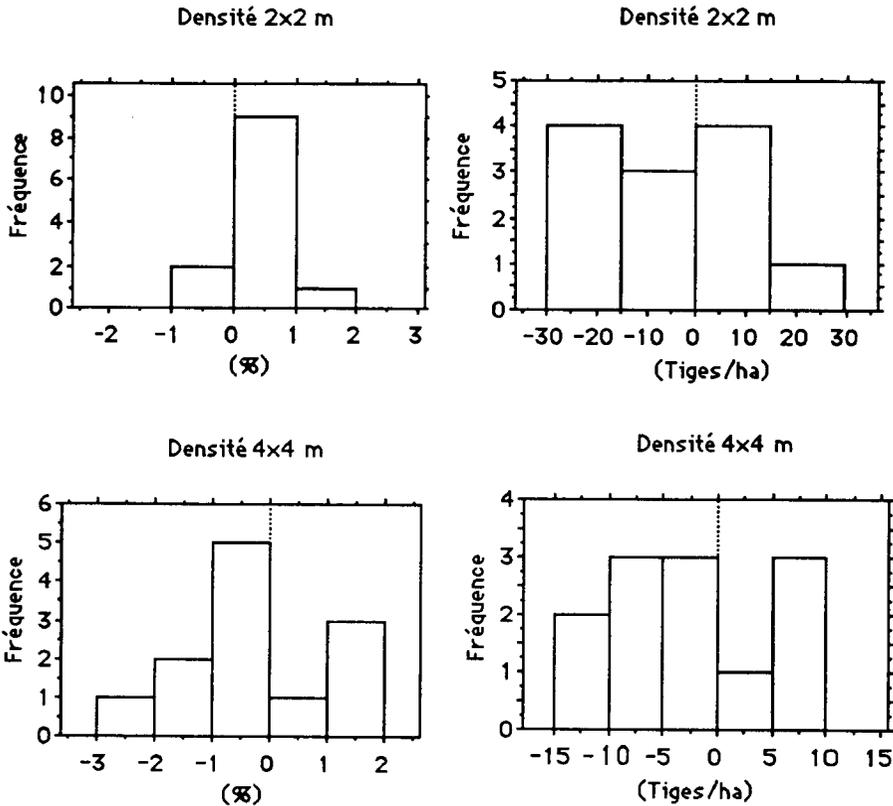


Fig 5. Distribution des écarts relatifs et absolus entre la moyenne empirique des estimations et la valeur calculée sur le placeau pour la densité (N). Simulations effectuées dans la parcelle U pour les deux espacements 2 m x 2 m ($N = 2448,5$ tiges/ha) et 4 m x 4 m ($N = 614,1$ tiges/ha). Histogrammes calculés à partir des résultats obtenus pour chacun des placeaux. Le biais est en moyenne faible. Les résultats sont qualitativement identiques pour la surface terrière (G) et la circonférence quadratique moyenne (Cg). La valeur de $t(y)$ n'est presque jamais significative au seuil de 5% (pour un placeau).

ques sylvicoles qu'à l'augmentation de la surface des placettes d'inventaire (fig 8). Néanmoins, ni l'effet de l'âge, ni celui de la géométrie ne se sont avérés statistiquement significatifs et la seule correction que nous pouvons proposer est de retrancher l'estimation de la valeur moyenne de $Ba(C\emptyset)$ obtenue sur une placette IFN : ce qui revient à ajouter environ 2 cm (puisque le biais moyen est de -2 cm).

Effet de la variation de la surface de placette

Nous avons constaté que plus la surface de la placette simulée est grande, plus la valeur du biais est faible (tableaux III et IV). Pour Cg le problème d'imprécision ne se présente que pour la placette de 0,01 ha où le biais absolu est de 16,8 cm. Pour $C\emptyset$ il y a aussi une imprécision assez im-

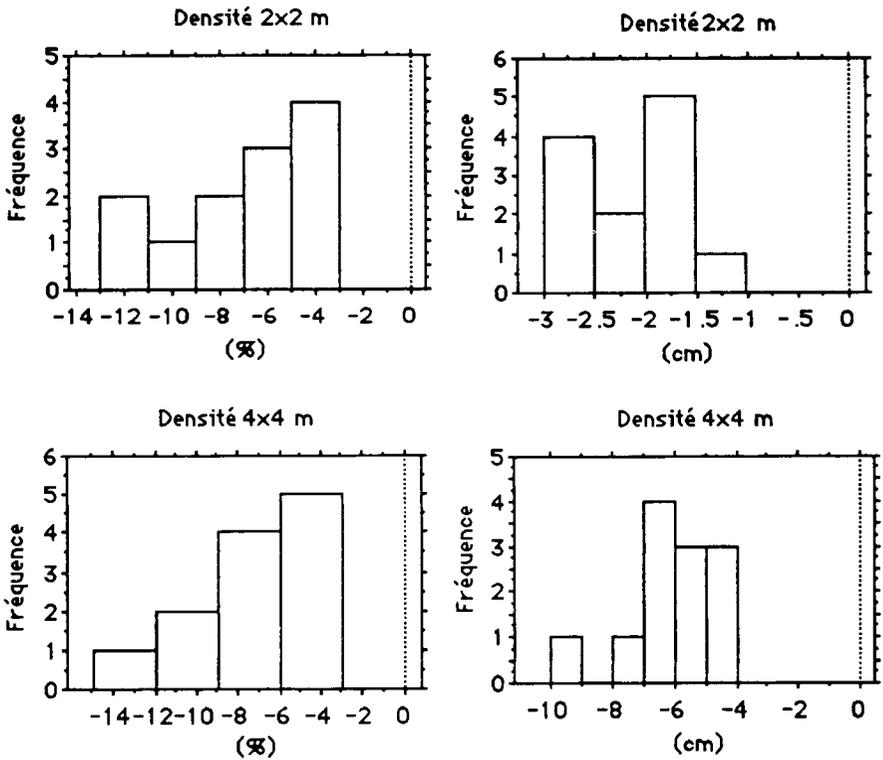


Fig 6. Distribution des biais empiriques relatifs et absolus pour la circonférence dominante ($C\emptyset$). Simulations effectuées dans la parcelle U pour les deux espacements 2 m x 2 m ($C\emptyset = 45,9$ cm) et 4 m x 4 m ($C\emptyset = 45,8$). Histogrammes calculés à partir des résultats obtenus pour chacun des placeaux. Le biais est toujours négatif. La valeur de $t(\gamma)$ est significative au seuil de 5% pour chacun des placeaux.

Tableau III. Évolution de l'écart-type et du biais empiriques de $C\emptyset$ en fonction de la surface des placettes simulées (placeau Pissos3 issu de régénération naturelle).

Surface (are)	$m(C\emptyset)$ (cm)	Min($C\emptyset$) (cm)	Max($C\emptyset$) (cm)	$s(m(C\emptyset))$ (cm)	$Br(C\emptyset)$	$Ba(C\emptyset)$ (cm)
1	128,6	0,0	175,0	49,16	-0,162	-24,8
2	148,3	111,6	180,5	14,92	-0,033	-5,1
3	149,9	117,7	176,1	12,17	-0,023	-3,5
4	150,7	119,2	173,3	10,92	-0,018	-2,7
5	150,7	128,4	170,6	10,02	-0,018	-2,7
6	150,2	128,6	167,0	8,54	-0,021	-3,2
7	150,5	126,7	167,4	8,07	-0,019	-2,9
8	150,3	128,0	165,5	7,56	-0,020	-3,1
9	150,2	128,1	164,4	7,28	-0,021	-3,2
10	150,5	129,2	164,3	6,88	-0,019	-2,9

Surface = 0,64 ha ; âge = 60 ans ; $C\emptyset = 153,4$ cm.

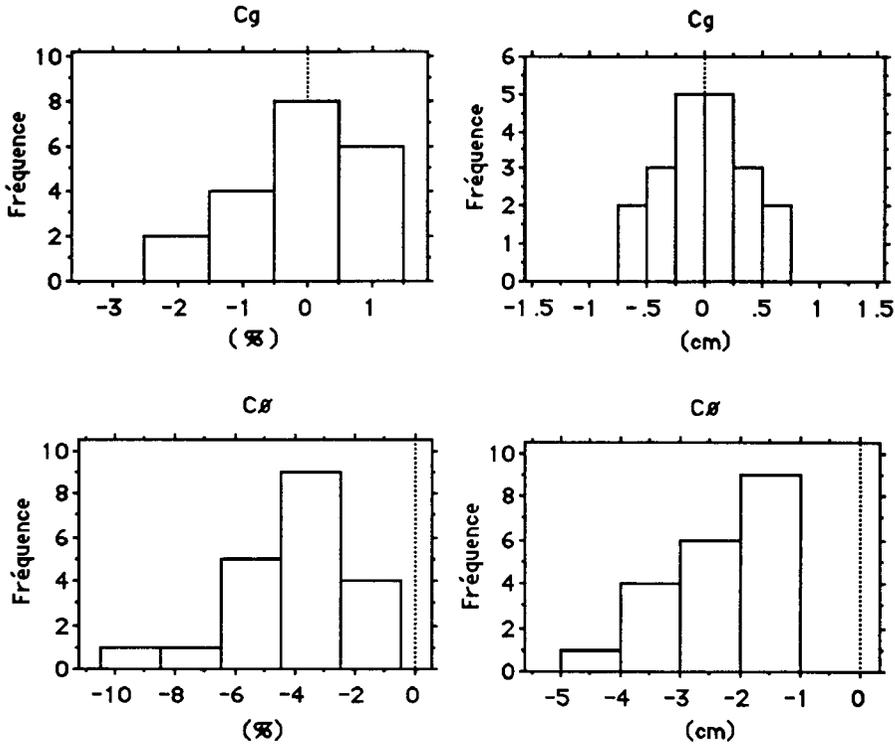


Fig 7. Distribution des biais empiriques relatifs et absolus pour la circonférence quadratique moyenne (Cg) et la circonférence dominante ($Cø$). Simulations effectuées dans les placeaux ONF-INRA (semis en ligne, âge compris entre 11 et 27 ans). Histogrammes calculés à partir des résultats obtenus pour chacun des placeaux. Le biais est en moyenne faible pour Cg . Les résultats sont qualitativement identiques pour la surface terrière (G) et la densité (N). La valeur de $t(y)$ n'est que rarement significative au seuil de 5% (pour un placeau). Le biais est toujours négatif et relativement important pour $Cø$. La valeur de $t(y)$ est significative au seuil de 5% pour chacun des placeaux.

Tableau IV. Évolution de l'écart type et du biais empiriques de Cg en fonction de la surface des placettes simulées (placeau Pissos3 issu de régénération naturelle)

Surface (are)	$m(Cg)$ (cm)	$Min(Cg)$ (cm)	$Max(Cg)$ (cm)	$s(m(Cg))$ (cm)	$Br(Cg)$	$Ba(Cg)$ (cm)
1	121,3	0,0	175,0	45,73	-0,122	-16,8
2	137,6	109,7	164,2	12,40	-0,0004	-0,5
3	138,1	114,4	157,3	8,98	0,000	0,0
4	138,4	115,0	150,4	9,15	0,002	0,3
5	138,4	120,0	157,4	8,23	0,002	0,3
6	137,8	122,1	155,0	7,27	-0,002	-0,3
7	138,0	123,7	159,0	7,27	-0,001	-0,1
8	137,9	123,7	157,2	6,68	-0,001	-0,2
9	137,8	124,7	153,7	6,57	-0,002	-0,3
10	137,9	127,3	153,7	6,04	-0,001	-0,2

Surface = 0,64 ha ; âge = 60 ans ; Cg = 138,1 cm.

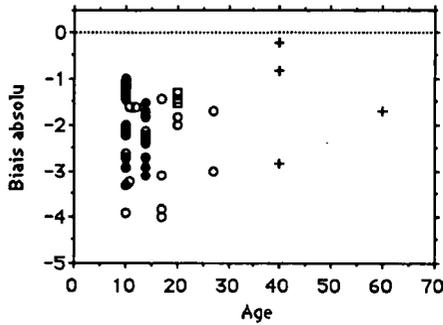


Fig 8. Représentation graphique de l'amplitude du biais empirique absolu de $C\varnothing$, en fonction de l'âge pour chaque géométrie de peuplement ($R^2 = 0,039$ non significatif au seuil de 5%). ● Plantation; ○ Semis en ligne; □ Semis en bande; + Rég naturelle.

portante dans la placette de 0,01 ha où le biais absolu est autour de 17 cm. Néanmoins, une tendance de diminution du biais se présente quand on augmente la surface de la placette ; de plus, dans des placettes avec une surface de 0,20 à 0,60 ha, il est au-dessous de 2,3 cm, ce qui nous montre que la surface de la placette dans laquelle on fait les estimations des caractéristiques dendrométriques du peuplement joue un rôle fondamental.

Influence du seuil de recensabilité de l'IFN

Dans cette section, les variables N , G , G_g , $C\varnothing$, N_r , G_r , Cg_r et $C\varnothing_r$ sont estimées sur

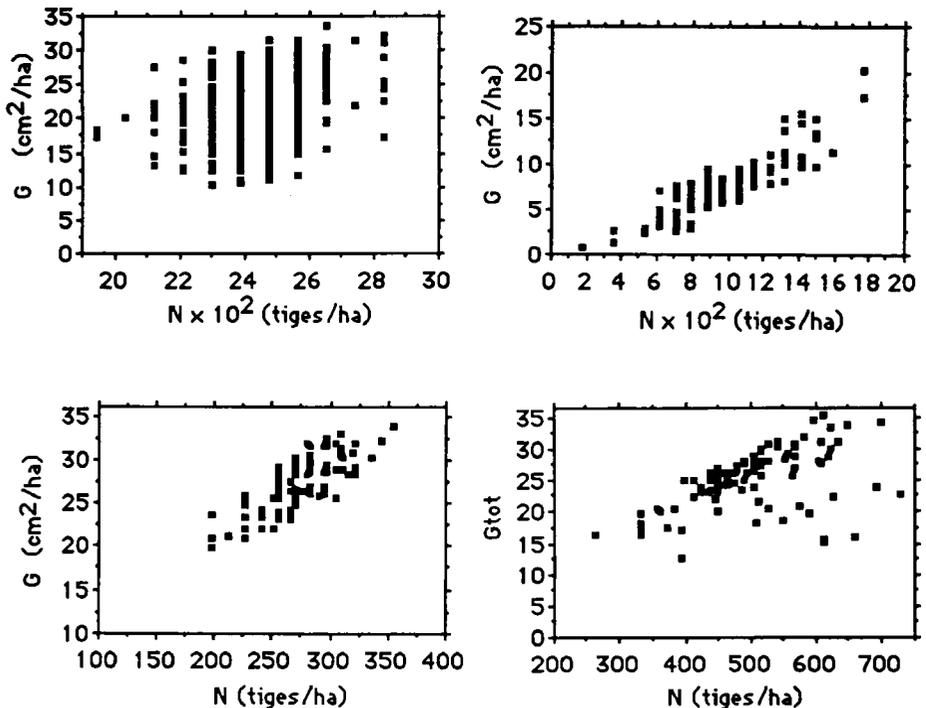


Fig 9. Représentation graphique de la corrélation entre N et G . a. Parcelle U, plantation à densité 2×2 m, 600 placettes simulées, $r = 0,32$. b. Placeau Commensacq II, semis en ligne, 100 placettes simulées, $r = 0,90$. c. Placeau Château Jaugé, régénération naturelle, 100 placettes simulées, $r = 0,75$. d. Placeau Le Bruillet, semis en bande, 100 placettes simulées, $r = 0,50$.

l'ensemble du plateau : il s'agit donc en toute rigueur de $\mu(N)$, $\mu(G)$, $\mu(Cg)$, $\mu(C\emptyset)$, $\mu(N_r)$, $\mu(G_r)$, $\mu(Cg_r)$ et $\mu(C\emptyset_r)$ et le problème du biais ne se pose donc pas. Nous observons que tous les arbres sont recensables dès que $C\emptyset$ dépasse 50 cm (fig 10). Nous avons ensuite représenté les taux de recensabilité N_r / N et G_r / G en fonction de $C\emptyset$. En nous appuyant sur l'allure du graphique, nous avons choisi le modèle suivant (équation dite de Chapman-Richards) :

$$Y = (1 - \exp[-\beta_1(C\emptyset - 24,5)])^{\beta_2}$$

où β_1 et β_2 sont des paramètres et $Y = N_r / N$ ou $Y = G_r / G$.

Nous avons ensuite ajusté les modèles suivants par régression non linéaire pondérée (avec la procédure NLIN du logiciel SAS ; voir résultats au tableau V) :

$$N = \frac{N_r}{(1 - \exp[-\beta_{1,N}(C\emptyset - 24,5)])^{\beta_{2,N}}} \quad [1]$$

$$G = \frac{G_r}{(1 - \exp[-\beta_{1,G}(C\emptyset - 24,5)])^{\beta_{2,G}}} \quad [2]$$

où $C\emptyset$ est en cm, N en nombre de tiges par ha et G en m^2/ha .

Tableau V. Ajustement des équations [1] et [2] permettant de prédire G et N à partir de G_r , N_r et $C\emptyset$.

Variable étudiée	Paramètre	Estimation	Écart-type
N	$\beta_{1,N}$	0,01505	0,00331
	$\beta_{2,N}$	1,6722	0,0856
G	$\beta_{1,G}$	0,03599	0,00525
	$\beta_{2,G}$	1,5126	0,0574

Les pondérations utilisées sont respectivement : $(1 - \exp[-0,015(C\emptyset - 24,5)])^{3,4}$ pour N et $(1 - \exp[-0,035(C\emptyset - 24,5)])^{3,4}$ pour G ($C\emptyset$ est en cm).

Les valeurs observées sont bien prédites par ces équations (fig 10). Dans le cas de la circonférence moyenne, nous avons utilisé les prédictions de N et G pour estimer Cg à partir de $Cg_r = \sqrt{4 \pi} G_r N_r$ et de $C\emptyset$: là aussi les prédictions sont satisfaisantes (fig 10).

DISCUSSION ET CONCLUSION

Les simulations réalisées montrent qu'un point-terrain IFN (en fait, le système des 3

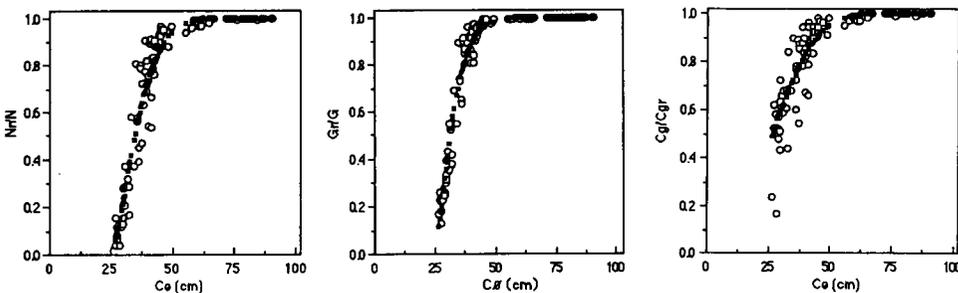


Fig 10. Prédiction des taux de recensabilité N_r/N , G_r/G et Cg/Cg_r en fonction de la circonférence dominante $C\emptyset$ (équations [1] et [2], tableau V) ; \blacksquare valeurs observées \circ valeurs prédites.

placettes concentriques) n'est pas localement représentatif du peuplement dans lequel il est situé : la faible surface des placettes, conjuguée à la structure du peuplement, induit une forte variabilité intra-peuplement des caractéristiques dendrométriques N et G . Cette variabilité est moindre pour Cg et $C\emptyset$, mais cette dernière variable est systématiquement sous-estimée de 1 à 4 cm. Des résultats voisins seraient sans doute obtenus pour la hauteur dominante : en valeur relative le biais serait sans doute plus faible car la dispersion des hauteurs individuelles est inférieure à la dispersion des circonférences.

Ces phénomènes sont surtout sensibles pour les jeunes peuplements où l'IFN n'échantillonne que sur la petite placette intérieure de 6 m de rayon. Dans le cas des peuplements âgés, nous avons montré sur un exemple que la surface supérieure de la placette conduit à une diminution de l'amplitude des erreurs aléatoires (et du biais pour $C\emptyset$). À cet effet se superpose la régularisation de la structure du peuplement induite par l'action des sylviculteurs (mise à distance des arbres par les éclaircies, concurrence inter-individuelle) et qui doit aussi contribuer à une réduction de la variabilité des estimations basées sur les données de l'IFN.

Nous avons par ailleurs montré que les problèmes liés à l'existence du seuil de recensabilité peuvent être partiellement résolus dès qu'une fraction du peuplement a effectivement franchi ce seuil : à condition de disposer d'un échantillon de placeaux où tous les arbres ont été recensés, il est possible de prédire les caractéristiques du peuplement à partir de la valeur de $C\emptyset$ et des caractéristiques de la fraction recensable du peuplement. Pour les très jeunes peuplements, pour lesquels aucun arbre n'est recensable, d'autres solutions restent à imaginer.

En combinant ces différents résultats, il est donc possible d'estimer $C\emptyset$, Cg , G et N à partir des données usuelles de l'IFN dès qu'une fraction du peuplement a franchi le seuil de recensabilité : dans un premier temps, on corrige le biais de $C\emptyset$ (paragraphe «Biais des estimations basées sur les placettes IFN»), puis on utilise les équations [1] et [2] pour prédire Cg , N et G à partir des seuls arbres recensables. Pour les peuplements plus jeunes, d'autres méthodes restent à développer.

Nos résultats confirment ceux de Azim (1990) et de Hann et Azim (1991), à savoir qu'il n'est pas possible d'utiliser directement un modèle de croissance ou une typologie de peuplements sans vérifier attentivement que la nature des données employées dans la phase d'application d'un modèle de croissance (respectivement d'une typologie) est compatible avec la nature des données utilisées pour construire ce même modèle (respectivement typologie). Il apparaît notamment que l'application directe du modèle de Lemoine – ou de tout autre modèle construit à partir de données observées sur des placeaux étendus – aux données dendrométriques élémentaires de l'IFN (ie au niveau de chaque placette) risque de poser de sérieux problèmes de biais et de précision :

- d'une part, les erreurs aléatoires et/ou systématiques induites par la faible surface des placettes vont en effet se propager et peuvent générer des biais du fait de la non linéarité des équations du modèle ;
- d'autre part, l'utilisation de scénarios de gestion (règles de déclenchement, intensité et nature des éclaircies) définis à l'échelle du peuplement et appliqués à celle de la placette d'inventaire peut contribuer à augmenter le biais et à amplifier les erreurs ;
- enfin, il est nécessaire de corriger les biais dus au seuil de recensabilité.

Pour résoudre ces problèmes, deux voies sont envisagées.

La première consiste à conserver l'échelle de la placette IFN comme niveau d'application du modèle de Lemoine, mais à recalibrer le modèle sur les données d'accroissement de l'IFN. Nous pourrions ainsi éliminer les biais dus à l'erreur différentielle de protocole, mais nous ne pourrions pas réduire la variabilité induite par la surface des placettes. Nous pourrions en revanche simuler l'effet de cette variabilité sur les projections à moyen terme et procéder à une analyse de sensibilité.

La deuxième voie consiste : (i) à stratifier les placettes selon des critères localement représentatifs tels l'âge, la hauteur ou la circonférence dominantes, la consistance (c'est-à-dire la fermeture du couvert estimée sur 20 ares autour du point d'inventaire) et (ii) à appliquer le modèle de Lemoine aux caractéristiques moyennes de chaque strate. Cette procédure permettra de diminuer la variabilité liée à la surface des placettes IFN, sans toutefois éliminer le biais sur C_0 .

REMERCIEMENTS

Les auteurs remercient vivement les sylviculteurs landais qui ont mis leurs peuplements à leur disposition ainsi que les deux lecteurs anonymes pour leurs remarques constructives.

RÉFÉRENCES

- Azim ZA (1990) Examining bias in estimating the response variable and assessing the effect of using alternative plot designs to measure predictor variables in diameter growth modelling. Ph D Thesis, Oregon State University, OR, USA, 75 p
- Bouchon J (1979) Structure des peuplements forestiers. *Ann Sci Forest* 36 (3), 175-209
- Chevrou RB, Guero MC, Houllier F (1988) Utilisation des résultats et des données brutes de l'Inventaire forestier national. Ministère de l'Agriculture, service de l'Inventaire forestier national, 187 p
- Chevrou RB (1993) La placette sol d'inventaire formée de plusieurs cercles concentriques. *Schwei Z Forstwes* 144 (4), 271-296
- Häggglund B (1982) Some remarks on the plot-stand problem in forest inventory. In: *Statistics in theory and practice: Essays in honour of Bertil Matérn*, Swedish University of Agricultural Sciences, Umea, Suède, 137-158
- Hann WD, Azim ZA (1991) Growth model predictions as affected by alternative sampling-unit designs. *For Sci* 37 (6), 1641-1655
- Houllier F (1986) Échantillonnage et modélisation de la dynamique des peuplements forestiers : application au cas de l'Inventaire forestier national. Thèse de doctorat, université Claude-Bernard (Lyon I), 267 p
- Houllier F (1990) Modèles de croissance en hauteur : remarques sur les relations entre la nature des données et le type de modèle ajusté. *Bull Rech Agron Gembloux* 25 (1), 65-75
- Inventaire forestier national (1985) Buts et méthodes de l'Inventaire forestier national. Ministère de l'Agriculture, Paris, 67 p
- Lemoine B (1991) Growth and yield of maritime pine (*Pinus pinaster*) : the average dominant tree of the stand. *Ann Sci For* 48, 593-611
- Matheron G (1965) *Les variables régionalisées et leur estimation*. Masson, Paris, 305 p
- Maugé JP (1979 et 1981) Études prospectives du massif landais et de ses ressources (pour la Gironde, 1979 et pour l'ensemble du massif, 1981). Centre de productivité forestière d'Aquitaine, 63 p
- Schmid-Haas P (1982) Sampling at the forest edge. In: *Statistics in theory and practice: Essays in honour of Bertil Matérn*, Swedish University of Agricultural Sciences, Umea, Suède, 263-276
- Shugart HH, West DC (1979) Size and pattern of simulated forest stands. *For Sci* 25 (1), 120-122