

Inventaires successifs en forêt : intérêt théorique et limites pratiques de l'échantillonnage partiellement renouvelé

F. HOULLIER (*)

*I.N.R.A., Station de Sylviculture et de Production
Centre de Recherches forestières de Nancy
Champenois, F 54280 Seichamps
Laboratoire de Biométrie, Université Claude Bernard
Lyon I, F 69622 Villeurbanne Cedex*

Résumé

L'échantillonnage partiellement renouvelé s'applique lors d'inventaires successifs. Cette méthode de sondage, développée en Amérique du Nord, s'appuie sur un plan d'échantillonnage mixte, composé de placettes permanentes et de placettes temporaires, et sur un modèle statistique structurel de la variable étudiée : le modèle linéaire.

Après avoir rappelé les bases théoriques de cette méthode, le gain de précision des estimateurs SPR relativement aux estimateurs usuels est présenté.

On insiste alors sur les limites de cette méthode en discutant les hypothèses initiales. On observe ainsi que les estimateurs SPR peuvent être généralisés dans le cas où on dispose d'un modèle d'évolution de la population. L'absence de prise en compte de toute structure spatiale peut être palliée en utilisant la théorie des Variables Régionalisées. Le problème de l'estimation pratique est posé : la non connaissance a priori des caractéristiques de la population (variances et covariances) introduit un biais dont l'amplitude est majorée.

La recherche d'un plan d'échantillonnage optimal est résolue dans un cas particulier : minimisation d'un critère global de précision sous une contrainte de coût linéaire.

Deux exemples fournissent des indications sur l'intérêt pratique du SPR ; ils indiquent que des gains de précision sensibles, surtout pour l'estimation du changement, sont possibles, tant pour des inventaires régionaux que d'aménagement.

On discute enfin les problèmes liés à l'intégration du SPR dans un plan d'échantillonnage global, tel celui de l'Inventaire Forestier National.

1. Introduction

Les inventaires forestiers répondent, à différentes échelles (parcelle, forêt, département), à trois grands types de questions posées par les différents acteurs forestiers ou industriels : 1) estimation, locale ou globale, à un instant donné, de l'état des peup-

(*) Stagiaire.

ments, du « stock sur pied » : volume, nombre de tiges, biomasse, conditions de milieu ou d'exploitation,... ; 2) Evaluation des modifications intervenues dans un passé récent : changement d'usage du sol, accroissement des peuplements, prélèvements et mortalité,... ; 3) Prévision de l'évolution future des peuplements, du stock, de l'accroissement, des ressources disponibles,... (BALLEYDIER et BERTRAND, 1982).

Ces questions ne sont pas indépendantes les unes des autres ; leur résolution fait cependant appel à des outils ou méthodes statistiques et mathématiques différents mais complémentaires :

— la question 1) relève de la théorie de l'échantillonnage. Les méthodes utilisées diffèrent principalement par la façon dont les caractéristiques de la population, en particulier sa structure spatiale, sont prises en compte par le plan d'échantillonnage et les estimateurs associés (HOULLIER, 1983a). On sait ainsi que l'échantillonnage aléatoire simple ne permet pas cette prise en compte de la structure spatiale. C'est pourquoi des plans d'échantillonnage et des estimateurs plus sophistiqués ont été introduits dans les inventaires forestiers : échantillonnage stratifié, à plusieurs degrés, échantillonnage systématique,... (voir les ouvrages classiques : LOETSCH et HALLER, 1973 ; DUPLAT et PERROTTE, 1983, p. 17-78 ; FRONTIER, 1983) et plus récemment la théorie des Variables Régionalisées (MARBEAU, 1976 ; BOUCHON, 1979 ; DUPLAT et PERROTTE, p. 205-232).

— la question 2) introduit une dimension supplémentaire : le temps. Deux grandes méthodes ont été traditionnellement utilisées en forêt :

- la comparaison d'inventaires successifs, indépendants ou non ;
- l'échantillonnage direct de la variable à estimer ; exemple : estimation de l'accroissement radial par sondage à la tarière (DUPLAT et PERROTTE, p. 91-123).

— la question 3) peut être considérée comme le prolongement des deux précédentes : il s'agit, à partir d'un état supposé connu, d'extrapoler des tendances passées. Le temps joue donc ici un rôle central et cette question fait appel à la fois aux méthodes d'échantillonnage déjà mentionnées et à des techniques de modélisation de la dynamique des peuplements.

On se limitera ici aux questions 1) et 2) et on les abordera au travers d'une méthode d'échantillonnage, spécialement conçue en vue de leur résolution simultanée : l'échantillonnage partiellement renouvelé (en anglais : SPR, *sampling with partial replacement*). Le SPR fait partie des techniques de sondage classiques (DESABIE, 1966 ; SCHERRER, 1983, p. 149-162) et a été l'objet d'un grand intérêt de la part des forestiers, principalement Nord-Américains, ce dont témoigne une abondante bibliographie. WARE et CUNIA (1962) ont présenté le plan d'échantillonnage et les estimateurs SPR dans le cas le plus simple : estimation d'une seule variable lors de deux inventaires successifs. Plusieurs améliorations ou compléments ont ensuite été proposés :

- estimation simultanée de plusieurs variables (NEWTON, BICKFORD et CUNIA, 1974) ;
- application au cas d'un nombre quelconque d'inventaires successifs (CUNIA et CHEVROU, 1969) ;
- application à l'estimation des surfaces par comptage de points (HAZARD, 1977 ; CHEVROU, 1982) ;
- compatibilité du SPR avec l'échantillonnage stratifié (CUNIA, 1965) et avec l'échantillonnage à plusieurs degrés (OMULE, 1981) ;
- optimisation du plan d'échantillonnage SPR (HAZARD et PROMNITZ, 1974 ; OMULE et WILLIAMS, 1982 ; KILPATRICK, 1981).

Le SPR a été appliqué en Amérique du Nord (FRAYER, 1978) et commence à diffuser hors de son « aire d'origine » (NETTO, 1983). Sa mise en œuvre est par ailleurs envisagée en Europe dans le cadre des inventaires nationaux (BENGTSSON, 1978), et en particulier en France par l'Inventaire Forestier National (I.F.N.) qui a entrepris depuis plusieurs années de repérer des placettes afin de les remesurer lors du troisième cycle ; le problème du futur traitement de ce nouveau type de données se pose donc à l'I.F.N. On pourrait d'autre part envisager l'application du SPR dans le cadre des inventaires d'aménagement.

On rappellera dans un premier temps les fondements théoriques du SPR sous sa forme la plus simple. Les généralisations qui en découlent compliquent le formalisme mais ne changent rien aux principes de base. On insistera donc plutôt sur les hypothèses sous-jacentes et sur la complémentarité du plan d'échantillonnage (phase d'acquisition des données) et des estimateurs (phase de traitement de données). On examinera ensuite un certain nombre de limites de la méthode, liées au non respect de certaines hypothèses. Dans la même optique, le lien avec les modèles dynamiques sera fait à partir du modèle sous-jacent. Le problème de l'optimisation d'un plan d'échantillonnage partiellement renouvelé fera l'objet d'un paragraphe et sera traité dans un cas particulier. Des indications pratiques sur les gains en précision seront fournies à partir de données réelles. On évoquera alors la question de l'intégration du SPR dans un plan d'échantillonnage global tel celui de l'I.F.N.

2. La théorie de base

2.1. Position du problème

On s'intéresse :

- à une population de limites supposées connues et fixes dans le temps (exemple : une forêt),
- à la valeur moyenne d'un paramètre de cette population (c'est-à-dire une densité, exemple : volume moyen/Ha) ; on supposera que ce paramètre, noté X , peut être défini ponctuellement.

On suppose que cette population est inventoriée à deux dates successives, t_1 et t_2 . On note X_1 et X_2 , le paramètre étudié à t_1 et t_2 . On cherche à estimer simultanément les valeurs moyennes de X_1 , X_2 et $D = X_2 - X_1$ (l'état à t_1 , à t_2 et la variation de cet état entre t_1 et t_2), notées $E(X_1)$, $E(X_2)$ et $E(D)$. Dans la pratique, l'estimation a posteriori de $E(X_1)$, à la date t_2 , est peu intéressante et la prévision de $E(X_2)$, à la date t_1 , ne relève pas des seules méthodes d'échantillonnage. On se placera donc à la date t_2 et on cherchera à estimer $E(X_2)$ et $E(D)$, sachant que l'estimation de $E(X_1)$ reste cependant possible.

2.2. Les fondements du SPR

Le SPR repose sur :

- une idée très simple : en forêt, les changements sont suffisamment progressifs pour que l'information collectée à un instant donné garde une « certaine valeur » quelques années plus tard (typiquement, l'intervalle entre deux inventaires successifs est

de 10 à 25 ans) ; il existe donc un lien non totalement aléatoire entre deux mesures successives faites au même point ;

— deux principes : 1) pour accéder à ce lien, il faut disposer de données appropriées : ce sont des couples de mesures espacées dans le temps (cf. 2.3) ; 2) pour utiliser ce lien et améliorer la précision des estimations (du stock et du « changement »), il faut un modèle statistique de ce lien et des estimateurs adaptés : dans le cas du SPR, c'est le modèle linéaire (cf. 2.4 et 3.1).

2.3. *Le plan d'échantillonnage partiellement renouvelé*

Le plan d'échantillonnage SPR généralise deux plans classiques : inventaires successifs indépendants (placettes temporaires) et inventaire continu (placettes permanentes). Aux dates t_1 et t_2 on tire respectivement N_1 et N_2 échantillons (en pratique des placettes) sur lesquels on mesure $x_1(i)$ et $x_2(j)$, valeurs respectives de X_1 et X_2 pour les échantillons i et j . Parmi ces échantillons, m sont communs à t_1 et t_2 . Il en reste donc u et n qui n'ont été mesurés qu'une seule fois, respectivement à t_1 et t_2 (cf. figure 1). Le cas particulier des inventaires successifs indépendants est obtenu pour $m = 0$; celui de l'inventaire continu pour $u = n = 0$.

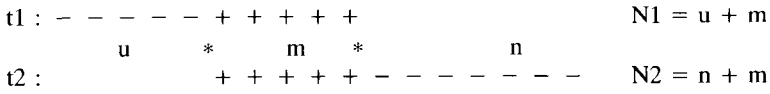


Fig. 1

*Plan d'échantillonnage SPR : nombre d'échantillons à t_1 et t_2 .
SPR sampling design : number of sampling units at dates t_1 and t_2 .*

On note $\overline{X1m}$ et $\overline{X2m}$ les moyennes respectives des $x_1(i)$ et $x_2(j)$ pour les m placettes mesurées à t_1 et t_2 ; on définit de façon analogue $\overline{X1u}$, $\overline{X2n}$. On fait les hypothèses supplémentaires suivantes :

- (H1) les tirages sont obtenus par échantillonnage aléatoire simple ;
- (H2) les échantillons sont ponctuels (c'est-à-dire que les mesures faites sur une surface, la placette, sont affectées à un point, le centre de cette placette) ;
- (H3) la population est de taille infinie [(H3) et (H2) sont en fait reliées].

2.4. *Le modèle statistique, les estimateurs SPR*

On fait des hypothèses sur la nature des paramètres X_1 et X_2 :

- (H4) X_1 et X_2 sont deux variables aléatoires, de variances et covariance supposées définies, et dont les valeurs observées $x_1(i)$ et $x_2(j)$ sur les échantillons sont des réalisations ;
- (H5) on connaît a priori les variances respectives de X_1 et X_2 , S_{11} et S_{22} , et leur covariance, S_{12} . On note ρ le coefficient de corrélation entre X_1 et X_2 : $\rho = S_{12}/\sqrt{S_{11}.S_{22}}$.

L'hypothèse (H4) définit un modèle statistique structurel de la population et justifie a posteriori les notations $E(X1)$ et $E(X2)$. L'hypothèse (H5) permet de calculer les estimateurs SPR sous le modèle (H4).

Les estimateurs SPR de $E(X2)$ et $E(D)$, notés respectivement $\widehat{X2}$ et \widehat{D} , sont, par définition, les estimateurs linéaires, non biaisés, et de moindre variance, formés à partir des $x1(i)$ et $x2(j)$; sous les hypothèses (H1) à (H5) on obtient (WARE et CUNIA, 1962) :

$$\widehat{X2} = \overline{X2n} + a \cdot (\overline{X2m} - \overline{X2n}) + b \cdot (\overline{X1u} - \overline{X1m}) \quad (1)$$

$$\widehat{D} = \overline{X2n} - \overline{X1u} + e \cdot (\overline{X2m} - \overline{X2n}) + f \cdot (\overline{X1u} - \overline{X1m}) \quad (2)$$

$$\text{avec } a = \frac{m \cdot (m + u)}{(m + u) \cdot (m + n) - u \cdot n \cdot \rho^2} \quad (3)$$

$$b = \frac{m \cdot u}{(m + u) \cdot (m + n) - u \cdot n \cdot \rho^2} \cdot \frac{S12}{S11} \quad (4)$$

$$e = \left(\frac{m \cdot (m + u)}{(m + u) \cdot (m + n) - u \cdot n \cdot \rho^2} \right) \cdot \left(1 + \frac{n}{m + u} \cdot \frac{S12}{S22} \right) \quad (5)$$

$$f = \left(\frac{m \cdot (m + n)}{(m + u) \cdot (m + n) - u \cdot n \cdot \rho^2} \right) \cdot \left(1 + \frac{u}{m + n} \cdot \frac{S12}{S11} \right) \quad (6)$$

La variance de ces estimateurs est :

$$V(\widehat{X2}) = S22 \cdot \left(m + n + u \cdot \frac{m \cdot \rho^2}{m + u - u \cdot \rho^2} \right)^{-1} \quad (7)$$

$$V(\widehat{D}) = (1 - e) \cdot S22/n + (1 - f) \cdot S11/u \quad (8)$$

2.5. L'efficacité théorique des estimateurs SPR

Sous les hypothèses (H1) à (H5), cette efficacité est mesurée relativement aux estimateurs classiques de $E(X2)$ et $E(D)$ par le rapport des variances des estimateurs. Les estimateurs classiques sont :

$$\bullet \widehat{X2}_o = (n \cdot \overline{X2n} + m \cdot \overline{X2m})/N2 \quad (9)$$

$$V(\widehat{X2}_o) = S22/N2 \quad (10)$$

$$\bullet \widehat{D}_o = \widehat{X2}_o - \widehat{X1}_o \quad (11)$$

$$V(\widehat{D}_o) = S22/N2 + S11/N1 - 2 \cdot m \cdot S12/[(m + u) \cdot (m + n)] \quad (12)$$

et dans le cas où $m = 0$, $u = N1$ et $n = N2$ (inventaires successifs indépendants) :

$$\bullet \widehat{D}'_o = \overline{X2n} - \overline{X1u} \quad (13)$$

$$V(\widehat{D}'_o) = S22/N2 + S11/N1 \quad (14)$$

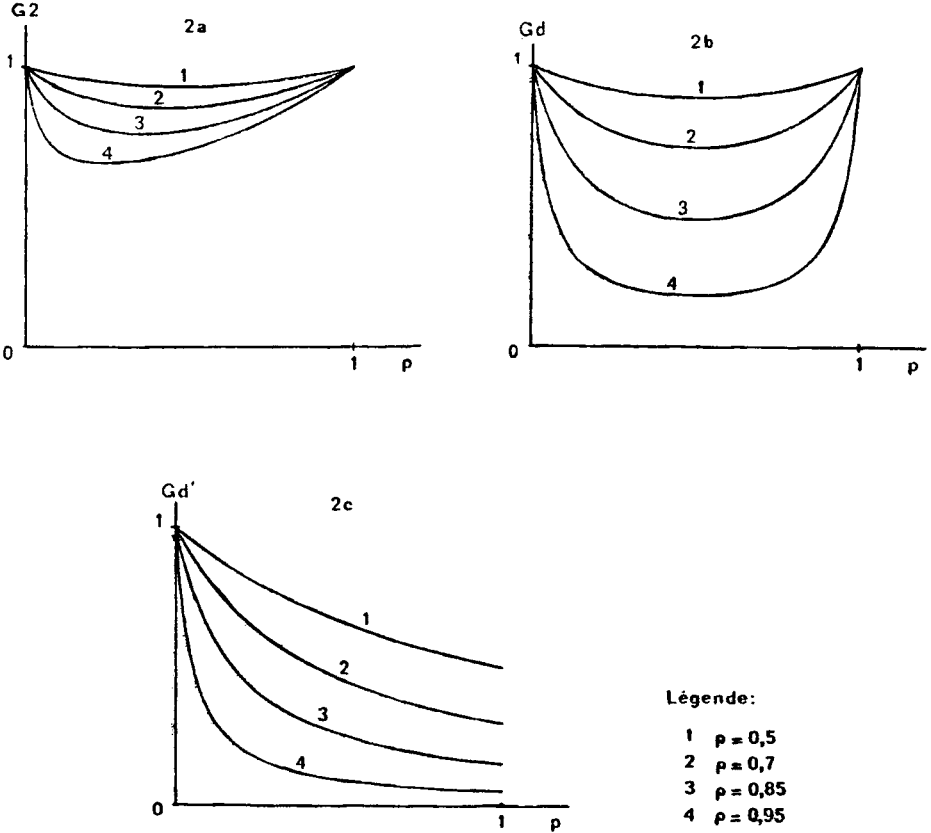


Fig. 2

Gain en précision des estimations SPR pour différentes valeurs de la corrélation ρ .
 On suppose $S_{11} = S_{22}$ et $u = n$. Voir formules (15) à (17) pour la définition de G_2 , G_d , $G_{d'}$.
 Relative efficiency of classical estimators compared to SPR estimators, assuming that $S_{11} = S_{22}$
 and $u = n$; ρ = temporal autocorrelation; p = remeasurement ratio.

- 2a) G_2 value (15) : Comparison of \widehat{X}_{2_0} (9) (estimator independent of t_1) and SPR estimator \widehat{X}_2 of $E(X_2)$ (1).
- 2b) G_d value (16) : Comparison of classical estimator \widehat{D}_0 (11) and SPR estimator \widehat{D} of $E(D)$ (2).
- 2c) $G_{d'}$ value (17) : Comparison of \widehat{D}'_0 (13) (successive independent inventories with $m = 0$) and \widehat{D} (2).

On note G_2 , G_d , G_d' l'inverse de l'efficacité théorique des estimateurs pour des valeurs données de N_1 et N_2 :

$$G_2 = V(\widehat{X_2})/V(\widehat{X_2}_o) \quad (15)$$

$$G_d = V(\widehat{D})/V(\widehat{D}_o) \quad (16)$$

$$G_d' = V(\widehat{D})/V(\widehat{D}'_o) \quad (17)$$

Les formules (7) et (10) montrent clairement l'apport du SPR par rapport aux traditionnels inventaires successifs indépendants (cf. figure 2) : l'utilisation de la corrélation linéaire entre X_1 et X_2 , mesurée sur les placettes permanentes, permet d'augmenter le nombre d'échantillons, N_2 , mesurés à t_2 , d'un nombre fictif d'échantillons, $u \cdot k(u, m, \rho)$, qui représente la « quantité d'information » apportée par les u placettes non remesurées pour l'estimation de $E(X_2)$:

où $0 \leq k(u, m, \rho) = (m \cdot \rho^2)/(m + u - u \rho^2) \leq 1$.

De la même manière, la corrélation entre X_1 et X_2 permet de mieux utiliser l'information contenue dans les placettes temporaires pour l'estimation de $E(D)$. Les paramètres dont dépendent les gains de précision sont de deux types :

— paramètres relatifs à la population : il n'est possible d'influer sur eux qu'indirectement, par exemple en stratifiant la population échantillonnée ; ce sont : S_{11} , S_{22} , et S_{12} ou ρ ;

— paramètres relatifs au plan d'échantillonnage : on peut directement agir sur eux ; ce sont : N_1 ; $p = m/N_1$, le taux de remesure ; $q = N_2/N_1$, le rapport des taux d'échantillonnage à t_1 et t_2 .

3. Les limites du SPR classique

3.1. Le modèle sous-jacent

Le SPR ne fait pas d'hypothèses sur un lien fonctionnel entre X_1 et X_2 , mais il utilise le lien structurel linéaire entre X_1 et X_2 (cf. (H4), voir TRANCHEFORT, 1974 pour la nuance entre modèles structurels et fonctionnels). On montre ainsi que les estimateurs SPR peuvent être obtenus en utilisant l'estimation par régression (CUNIA, 1965). Le modèle dynamique sous-jacent au SPR est donc le modèle linéaire classique :

$$X_2 = A \cdot X_1 + B + e \quad (18)$$

où e est un terme d'erreur, d'espérance supposée nulle et de variance $V(e)$. L'estimation SPR de $E(X_2)$ s'obtient en ajustant ce modèle par les moindres carrés et en faisant une moyenne pondérée par l'inverse de leur variance des deux estimateurs indépendants suivants :

$$\overline{X_2}_n \text{ et } \overline{X_2}_m + (u/(m + u)) \cdot (\widehat{X_2}_u - \widehat{X_2}_m) \text{ avec}$$

$$\widehat{X_2}_u = (1/u) \cdot \sum_{i=1}^u A \cdot x_1(i) + B, \quad \widehat{X_2}_m = (1/m) \cdot \sum_{j=u+1}^{u+m} A \cdot x_1(j) + B,$$

$$A = S_{12}/S_{11} \text{ et } B = E(X_2) - A \cdot E(X_1)$$

Plusieurs publications ont porté directement ou indirectement sur l'amélioration de ce modèle ou sur l'étude de sa validité :

— pour améliorer le lien statistique (la corrélation) entre les observations à t1 et t2, ce modèle univariable a été généralisé au cas multivariable : X1 et X2 ne sont alors plus des variables aléatoires scalaires mais vectorielles. Les notations et formules sont un peu plus compliquées, mais le principe reste identique (NEWTON, BICKFORD et CUNIA, 1974). Le problème de la sélection des meilleures variables explicatives se pose alors de la même façon qu'en régression (TOMASSONE, LESQUOY et MILLIER, 1983, p. 158-168) ;

— FRAYER (1966) a étudié la validité de l'ajustement du modèle par les moindres carrés (hypothèse d'homogénéité de V(E)) et proposé les moindres carrés pondérés pour les cas où cette hypothèse n'est pas vérifiée ;

— de façon plus générale, si on dispose d'un modèle déterministe de l'évolution des peuplements, il est possible de construire des estimateurs SPR à partir d'un plan d'échantillonnage partiellement renouvelé ; si on note ce modèle :

$$X2 = F(X1) + e \quad (19)$$

On obtient un estimateur de E(X2) en faisant une moyenne pondérée par l'inverse de leur variance des deux estimateurs suivants :

$$\overline{X2n} \text{ et } \overline{X2m} + u/(m + u) \cdot (\widehat{X2u} - \widehat{X2m})$$

$$\text{où } \widehat{X2u} = (1/u) \cdot \sum_{i=1}^u F(x1(i)) \text{ et } \widehat{X2m} = (1/m) \cdot \sum_{j=u+1}^{u+m} F(x1(j))$$

Dans les calculs, on remplace alors ρ^2 par le coefficient de détermination du modèle noté R^2 : $R^2 = 1 - V(e)/S22$.

3.2 La prise en compte de l'espace dans le SPR

Cette prise en compte est nécessaire du fait que les hypothèses (H1), (H2) et (H3) ne sont pas vérifiées dans la pratique :

— le dispositif d'échantillonnage est plus fréquemment systématique qu'aléatoire ; comme, par ailleurs, la structure (horizontale) des peuplements n'est pas totalement aléatoire (BOUCHON, 1979), les estimateurs SPR ne sont pas réellement optimaux et la variance de ces estimateurs n'est qu'approximative ;

— les échantillons sont mesurés sur un support qui régularise la variable observée (DUPLAT et PERROTTE, p. 49-50). La population sondée n'est ainsi infinie que si l'on s'intéresse à la variable régularisée supposée définie ponctuellement.

La théorie des variables régionalisées (MATHERON, 1965) permet de prendre en compte la structure des peuplements et celle du plan d'échantillonnage (localisation des échantillons) ; elle généralise ainsi le SPR (HOULLIER, 1983b) :

— les hypothèses (H1), (H2), (H3) sont alors supprimées (on tient compte explicitement des caractéristiques du dispositif d'échantillonnage) ;

— (H4) est généralisée par une hypothèse plus forte sur la nature de la variable étudiée : X1 et X2 ne sont plus des variables aléatoires, mais des fonctions aléatoires, ou processus spatiaux, intrinsèques (voir MATHERON, 1970, p. 53) ;

— la connaissance a priori de S_{11} , S_{22} , S_{12} (hypothèse (H5)) est remplacée par la connaissance de deux fonctions qui décrivent la structure spatio-temporelle de X : la dérive et le demi-variogramme ; cette hypothèse est, elle aussi, plus forte que l'hypothèse classique, puisque la dérive et le variogramme contiennent plus d'informations sur X que les simples variances et covariances.

La formulation, très abstraite, de cette généralisation ne sera pas développée ici ; les principes exposés au 2.2 restent valables : le plan d'échantillonnage, par sa propre structure fournit des informations sur les variations spatio-temporelles de X ; cette information est extraite puis utilisée par les estimateurs grâce à sa formalisation par un modèle structurel. Cette méthode reste jusqu'à présent plus utile d'un point de vue théorique que pratique : la prise en compte des structures spatio-temporelles de la population et du plan d'échantillonnage induit en effet des calculs longs et compliqués ; aucune application forestière de cette généralisation n'est, à ce jour, connue. Sur ce sujet, on peut aussi consulter MATERN et RANNEY (1983).

L'influence de la taille du support de mesure (non respect de (H2)) sur la valeur observée de l'autocorrélation temporelle, c'est-à-dire ρ , a été étudiée de façon expérimentale (HOULLIER, 1983b). Dans deux peuplements de Chêne sessile, servant à la comparaison de deux normes d'éclaircie, cartographiés pied par pied et mesurés à plusieurs reprises, on a simulé des placettes rectangulaires de surface croissante et on a calculé la valeur de ρ pour chaque surface. Il est apparu que selon le type d'éclaircie pratiqué et selon le paramètre étudié (nombre de tiges, surface terrière), l'influence de la taille du support sur ρ n'est pas identique (tableau 1). Comme pour les inventaires usuels, on observe donc que le choix de la taille des placettes est un acte essentiel de l'échantillonneur. Faute de modèles décrivant les structures spatiale et temporelle des peuplements, cette observation reste qualitative (DUPLAT et PERROTTE, p. 51 et 61).

3.3. Estimations pratiques - Le biais du SPR

On est resté jusqu'ici dans un domaine très théorique, en supposant en particulier que (H5) était vérifiée. Dans la pratique, S_{11} , S_{22} et S_{12} ne sont pas connues a priori. On doit donc les remplacer par leurs estimations \widehat{S}_{11} , \widehat{S}_{22} , \widehat{S}_{12} dans les formules (1) à (8). Par souci de cohérence, NEWTON, BICKFORD et CUNIA recommandent de n'utiliser que les m échantillons remesurés pour ces estimations. \widehat{S}_{11} , \widehat{S}_{22} , \widehat{S}_{12} , $\hat{\rho}$, \hat{a} , \hat{b} , \hat{e} , \hat{f} sont alors considérées comme des variables aléatoires définies sur l'ensemble des plans d'échantillonnage possibles pour (u, m, n) donné. Il s'ensuit que \widehat{X}_2 et \widehat{D} sont alors des estimateurs non optimaux (au sens de la variance minimale). Ils sont même biaisés, à moins que le modèle liant X_1 à X_2 ne soit effectivement linéaire (SCOTT, 1984).

SCHERRER souligne que ni le biais, ni la consistance de \widehat{X}_2 , de \widehat{D} et des estimations de leur variance ne sont connus dans le cas général. La valeur théorique des biais est :

$$B(\widehat{X}_2) = \text{Cov}(\hat{a}, \overline{X_{2m}}) - \text{Cov}(\hat{b}, \overline{X_{1m}}) \quad (20)$$

$$B(\widehat{D}) = \text{Cov}(\hat{e}, \overline{X_{2m}}) - \text{Cov}(\hat{f}, \overline{X_{1m}}) \quad (21)$$

Deux approches ont été développées pour étudier l'importance de ce biais. La première consiste à simuler des tirages SPR sur un peuplement connu de façon exhaustive. Dans l'exemple traité par HOULLIER (1983b), l'écart entre les valeurs exactes et les moyennes observées des estimations SPR ne s'est pas avéré significatif. En utilisant les formules (20) et (21), on a pu reconstituer des estimations du biais relatives inférieures à 1 % et 5 % (respectivement pour \widehat{X}_2 et \widehat{D}). Il n'est cependant pas possible de déduire des considérations générales, à partir de cet unique exemple.

TABLEAU 1

Influence conjointe de la sylviculture et de la surface, S, des placettes de lever sur la valeur du coefficient d'autocorrélation temporelle, ρ à l'intérieur d'une jeune futaie de Chêne sessile (données C.N.R.F.).

Influence of silviculture and plot size, S, on the temporal autocorrelation, ρ in a young Oak stand.

S Ares	DT ans	Surface terrière/ha			Nombre de tiges/ha		C
		0	9	18	9	18	
Eclaircies fortes ⁽¹⁾							
1/8		1,00	0,99		1,00		0
		0,90	0,89	0,88	0,80	0,80	1
			0,80	0,77		0,64	2
1/2		0,89	0,87	0,85	0,80	0,80	1
			0,82	0,79		0,68	2
1		0,90	0,89	0,90	0,83	0,83	1
			0,83	0,81		0,70	2
3 125		1,00	0,99		1,00		0
		0,88	0,87	0,88	0,86	0,86	1
			0,83	0,81		0,80	2
Eclaircies faibles ⁽²⁾							
1/16		1,00	1,00		1,00		0
		0,97	0,96	0,96	0,87	0,87	1
			0,92	0,90		0,75	2
1/4		0,96	0,95	0,94	0,84	0,84	1
			0,90	0,87		0,71	2
1		0,93	0,90	0,88	0,80	0,80	1
			0,82	0,77		0,63	2
2		1,00	0,99		1,00		0
		0,93	0,91	0,88	0,76	0,76	1
			0,80	0,73		0,51	2

DT = intervalle entre deux mesures successives - *Time lap between two inventories.*

C = nombre d'éclaircies pendant DT - *number of thinnings during DT.*

(1) heavy thinnings

(2) light thinnings

(d'après HOULLIER, 1983b)

La seconde approche consiste à rechercher une majoration théorique du biais à partir des formules (20) et (21) :

$$\text{si } S_{11} = S_{22}, B(\widehat{X}_2) = \sqrt{S_{22}/m} \cdot (\sigma(\hat{a}) \cdot \text{Cor}(\hat{a}, \overline{X_{2m}}) + \sigma(\hat{b}) \cdot \text{Cor}(\hat{b}, \overline{X_{1m}})) \quad (22)$$

$$\text{d'où } |B(X_2)| \leq \sqrt{S_{22}/m} \cdot (\sigma(\hat{a}) + \sigma(\hat{b})) \quad (23)$$

avec $\sigma(\hat{a}) = \text{écart-type de la variable aléatoire } \hat{a}$;

Cor (\hat{a} , $\overline{X2m}$) = corrélation des variables aléatoires \hat{a} et $\overline{X2m}$. Si on suppose de plus que $u = n$ (soit $N1 = N2$), que ρ est compris entre 0 et 1 et qu'on note $p = m/(m + u)$ on déduit des valeurs extrêmes de \hat{a} et \hat{b} une majoration exacte mais très grossière de $B(\widehat{X2})$ (on procède de même pour $B(\widehat{D})$) ; on a par exemple :

$$\text{Min}(\hat{a}) = m/(m + n) \leq \hat{a} \leq (m \cdot (m + u))/((m + u) \cdot (m + n) - u \cdot n) = \text{Max}(\hat{a})$$

$$\text{et } V(\hat{a}) \leq (1/4) \cdot (\text{Max}(\hat{a}) - \text{Min}(\hat{a}))^2$$

$$\text{d'où } |B(\widehat{X2})| \leq \sqrt{S22/m} \cdot (1 - p)/2 \quad (24)$$

$$|B(\widehat{D})| \leq \sqrt{S22/m} \cdot (1 - p) \quad (25)$$

Cette majoration reste peu satisfaisante. Elle peut être affinée moyennant des hypothèses supplémentaires : on suppose que $S11 = S22$ et que la valeur estimée de ρ , notée $\hat{\rho}$, est proche de ρ , soit $|\hat{\rho} - \rho| \ll \rho$. Dans la pratique, cela ne sera vérifié que si m est assez grand ; en utilisant la transformation de Fisher (KENDALL et STUART, 1969, p. 390-391) on voit que pour $\rho = 0,8$ il faut avoir environ $m \geq 50$. Si ces hypothèses sont vérifiées on peut approcher $\sigma(\hat{a})$ et $\sigma(\hat{b})$ en faisant des développements limités de \hat{a} et \hat{b} au premier ordre en $\rho - \hat{\rho}$ (\hat{a} et \hat{b} sont des fonctions de $\hat{\rho}$) ; on obtient alors (KENDALL et STUART, p. 231-232) :

$$E(\hat{a}) = a \text{ au deuxième ordre près,}$$

$$V(\hat{a}) = (da/d\rho)^2 \cdot V(\hat{\rho}) \text{ au troisième ordre près } (V(\hat{\rho}) = \text{variance de } \hat{\rho}).$$

On procède de même pour \hat{b} , \hat{e} , \hat{f} . On trouve alors :

$$|B(\widehat{X2})| \leq \sqrt{\frac{S22}{m}} \cdot \sigma(\hat{\rho}) \cdot \frac{p - (1 - p)}{1 - p - (1 - p)} \quad (26)$$

$$|B(\widehat{D})| \leq 2 \cdot \sqrt{\frac{S22}{m}} \cdot \sigma(\hat{\rho}) \cdot \frac{p - (1 - p)}{1 - p - (1 - p)} \quad (27)$$

Des exemples de ces valeurs sont donnés au tableau 2. Il faut en retenir que :

- le biais est borné par une quantité qui décroît plus vite que $1/\sqrt{m}$;
- si $m \geq 50$ et $\rho \geq 0,8$, le biais est certainement négligeable ;
- ces résultats sont des majorations (exactes pour (24) et (25) ou approchées pour (26) et (27)) et qu'ils sont donc, par définition, pessimistes.

Par ailleurs, si on suppose que $S11 = S22$, il est possible d'éviter tout biais en fixant a priori, et indépendamment du plan d'échantillonnage, la valeur de $\hat{\rho}$, $\hat{\rho} = r$. En faisant cela on enlève aux estimateurs SPR leur caractère optimal ; il faut cependant noter que l'augmentation de variance est du second ordre en $(r - \rho)$. Pour le vérifier, on effectue un développement limité de \hat{a} , \hat{b} , \hat{e} , \hat{f} , $V(\widehat{X2})$ et $V(\widehat{D})$ en $(r - \rho)$; on note $\widehat{X2}'$ et \widehat{D}' les estimateurs obtenus en remplaçant ρ par r dans (1) à (8) (avec $S11 = S22$) :

TABLEAU 2

Majoration du biais des estimateurs SPR exprimé en valeur relative de l'écart-type d'estimation théorique, et en fonction du taux de remesure ρ ; d'après les formules (26) et (27); avec : $S11 = S22$, $u = n$ et $\rho = 0,8$

$$\sigma(\widehat{X2}) = \sqrt{V(\widehat{X2})} \text{ et } \sigma(\widehat{D}) = \sqrt{V(\widehat{D})}$$

Bias majoration (formulas (26) (27)) for the SPR estimators. Bias values are relative to the standard deviation of the SPR estimators (7), (8). $\rho =$ remeasurement ratio.

ρ	$\sigma(\hat{\rho}) = 0,1$		$\sigma(\hat{\rho}) = 0,05$	
	$B(\widehat{X2})/\sigma(\widehat{X2})$	$B(\widehat{D})/\sigma(\widehat{D})$	$B(\widehat{X2})/\sigma(\widehat{X2})$	$B(\widehat{D})/\sigma(\widehat{D})$
0,1	0,39	0,61	0,19	0,30
0,3	0,22	0,42	0,11	0,21
0,5	0,11	0,24	0,05	0,12
0,7	0,05	0,12	0,02	0,06
0,9	0,02	0,03	0,01	0,02

$V(\widehat{X2}) = V(\widehat{X}) + [(\hat{a} - a)^2 \cdot (m + n)/(m \cdot n) + (\hat{b} - b)^2 \cdot (m + u)/(m \cdot u) - 2(\hat{a} - a) \cdot (\hat{b} - b)/m]$. S11 avec, par exemple : $(\hat{a} - a) = (da/d\rho) \cdot (r - \rho)$ au deuxième ordre près.

Dans un article récent, SCOTT (1984) développe des estimateurs généraux, non biaisés, tenant compte du non respect de (H5), mais il suppose a priori que $X1$ et $X2$ sont liées par un modèle linéaire classique (cf. (18)).

4. Optimisation d'un plan d'échantillonnage SPR

4.1. Position du problème

On se place sous les hypothèses (H1) à (H5); on se limite donc au SPR classique (cf. 2) sans aucune référence à la théorie des Variables Régionalisées, par exemple. Dans ce cadre le SPR fournit les estimateurs optimaux (de moindre variance) pour un plan d'échantillonnage et une population fixés a priori ($u, m, n, S11, S22$ et $S12$ sont supposés connus).

Le problème de l'échantillonneur est plus compliqué : il évolue en effet dans un contexte économique où s'opposent classiquement des objectifs de précision et des contraintes de coût. En pratique :

- il doit choisir lui-même le plan d'échantillonnage, soit en particulier (u, m, n);
- il ne connaît pas exactement a priori $S11, S22, S12$;
- son budget est limité;
- il s'intéresse simultanément à plusieurs variables;
- il souhaite obtenir une certaine précision sur ses estimations.

On observe aussi que le taux optimal de remesure n'est pas le même selon que l'on veut estimer $E(X2)$ ou $E(D)$ (cf. figure 2).

On peut donc poser deux problèmes d'optimisation dont l'un généralise l'autre :

(P1) rechercher les valeurs optimales de u , m , n pour une population de paramètres supposés connus ;

(P2) rechercher les valeurs optimales de u , m , n pour une population de paramètres inconnus mais représentés par des variables aléatoires ((H5) n'est plus vérifiée). (P2) revient à considérer un modèle de superpopulation (CASSEL, SARNDAL et WRETMAN, 1977) et généralise (P1).

4.2. Définition et résolution d'un problème d'optimisation (P1)

Bien que (P2) soit plus proche du problème concret de l'échantillonneur, on se limitera ici à (P1) et au cas particulier suivant :

(i) on se place à la date t_2 et on suppose que $N_1 = m + u$ est connu ; en considérant $u = N_1 - m$, on élimine u de la suite des calculs ;

(ii) le coût de l'inventaire à la date t_2 s'exprime par une fonction linéaire en m et n : $C(m,n) = C_0 + m \cdot C_m + n \cdot C_n$ (28)

où C_m incorpore éventuellement le coût actualisé de repérage des placettes m à la date t_1 (en plus des coûts effectifs à la date t_2) ;

(iii) ce coût est soumis à une contrainte budgétaire CF : $CF \geq C(m,n)$;

(iv) l'échantillonneur s'intéresse à L variables X_k , pour lesquelles on définit les variances S_{11k} , S_{22k} , et la corrélation ρ_k ; il estime simultanément $E(X_{2k})$ et $E(D_k)$ par les formules (1) à (8) pour chacune des variables X_k ($1 \leq k \leq L$) ;

(v) l'échantillonneur est capable de formaliser ses objectifs de précision pour les estimations \widehat{X}_{2k} et \widehat{D}_k par un critère global J du type :

$$J(m,n) = \sum_{k=1}^L (w_k \cdot V(\widehat{X}_{2k}) + z_k \cdot V(\widehat{D}_k)) \quad (29)$$

La supériorité d'un plan (u, m, n) sur un autre (u', m', n') du point de vue de la précision se traduit alors par : $J(m,n) \leq J(m',n')$ (cf. figure 3).

Dans ces conditions (P1) consiste à rechercher (m,n) tel que :

$$\left[\begin{array}{l} CF \geq C(m,n) \\ 0 \leq n \leq N_1 \\ 0 \leq m \\ J(m,n) \text{ soit minimal} \end{array} \right.$$

L'approche duale — minimisation d'une fonction de coût sous des contraintes de précision — est traitée par HAZARD et PROMNITZ (1974). Très satisfaisante d'un point de vue intellectuel, cette approche ne correspond sans doute pas aux contraintes économiques d'un inventaire.

La linéarité de la contrainte et la convexité de J assurent l'existence et l'unicité de la solution de (P1) (HOULLIER, 1983b). Pour le calcul de cette solution on peut utiliser une méthode graphique tout à fait adaptée (cf. figure 4). L'optimum (m,n) trouvé dépend des paramètres suivants :

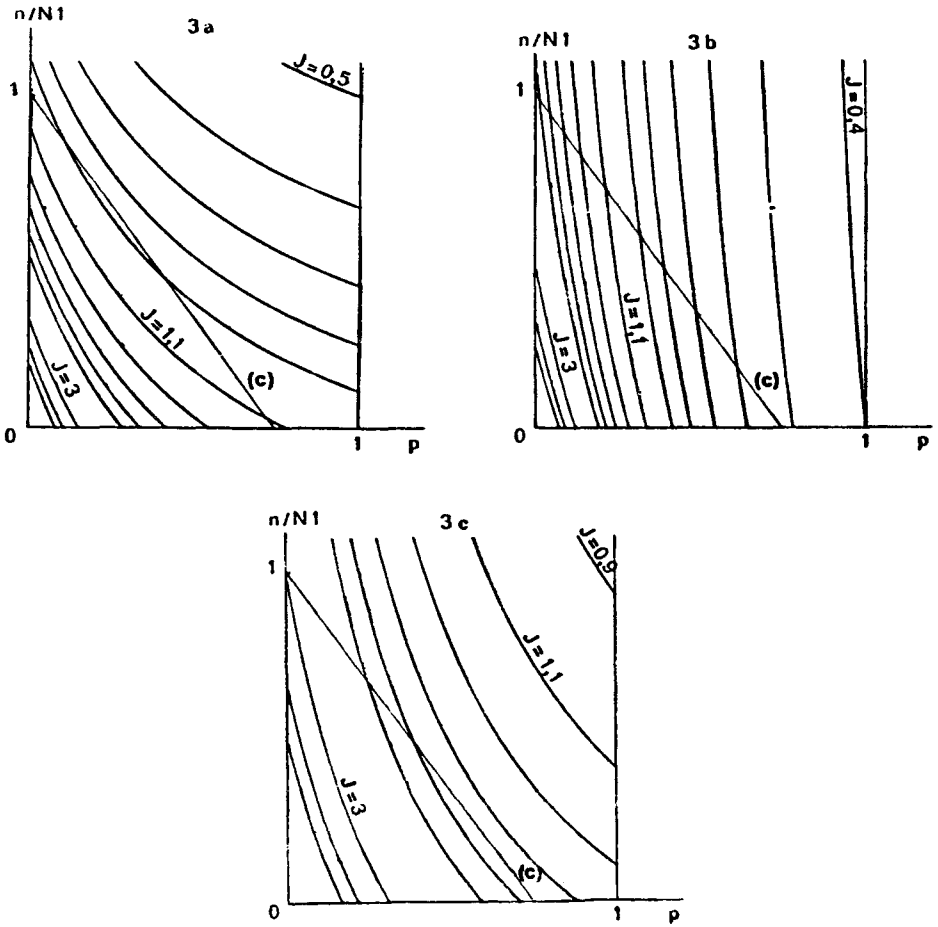


Fig. 3

Forme des courbes « d'isopréférence » pour le critère J (29) et position de la contrainte de précision (28), en fonction de p et $n/N1$, avec :

General form of the curve $J(m,n) = \text{constant}$ (29), and position of the linear cost constraint (28).

$L = 1, \rho = 0,8, S11 = S22,$
 $Cm = 4/3 \cdot Cn, Cn = (CF - C0)/N1,$

- 3a) $w_1 = 1 ; z_1 = 0$
- 3b) $w_1 = 0 ; z_1 = 1$
- 3c) $w_1 = 1 ; z_1 = 1$

p = remeasurement ratio = $m/N1$
 $n/N1$ = ratio of new plots at date $t2$ against total number of plots at $t1$
 L = number of studied variables $X_k, 1 \leq k \leq L$
 Cm = cost of remeasuring a plot
 Cn = cost of measuring a new plot
 $C0$ = fixed cost
 CF = cost constraint
 w_k = relative weight of $V(\widehat{X2}_k)$ in $J(m,n)$
 z_k = relative weight of $V(\widehat{D}_k)$ in $J(m,n)$

- $S11$, $S22$ et ρ ;
- Cm/Cn , $CF - C0$ et $(CF - C0)/N1$;
- les pondérations w_k et z_k .

Dans la pratique, on obtient des optima assez « plats » (près de l'optimum, J varie peu quand on modifie m et n) et on observe que :

- pour $(CF - C0)/(N1 \cdot Cn)$ inférieur ou égal à 1, n n'est supérieur à m que si w_k est nettement supérieur à z_k ou si Cm est nettement supérieur à Cn ;
- pour Cm peu supérieur à Cn , si $(CF - C0)/(N1 \cdot Cn)$ est inférieur ou égal à 1 et si $w_k \approx z_k$, alors m est nettement supérieur à n (cf. figure 4) ;
- pour Cm peu supérieur à Cn , m et n tendent très rapidement vers $(CF - C0)/Cm$ et 0 (respectivement) quand w_k/z_k tend vers 0.

Un ordre de grandeur de 50 % pour le taux de remesure semble raisonnable pour les cas usuels. Il faut cependant souligner que la difficulté majeure de la formulation choisie provient de ce qu'il n'est pas évident de fixer les pondérations, w_k et z_k , intervenant dans J . La forme même de J peut d'ailleurs prêter à discussion. Les résultats obtenus dépendent donc essentiellement de la pertinence du critère J .

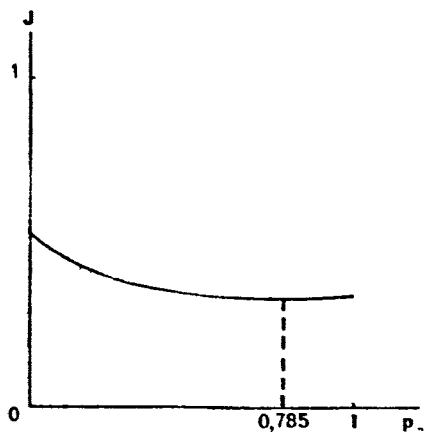


Fig. 4

Résolution graphique du problème d'optimisation (P1) pour :
Graphical resolution of the optimization problem (P1)

$L = 2$,	$S11_1 = S22_1 = 3,0$	$\rho_1 = 0,56$	$w_1 = 6$	$z_1 = 6$
	$S11_2 = S22_2 = 20,0$	$\rho_2 = 0,78$	$w_2 = 1$	$z_2 = 1$
$CF - C0 = 100$	$Cm = 0,55$	$Cn = 0,45$	$N1 = 200$	

Solution : $m = 157$, $n = 30$, $N2 = 187$.

$S11_k$ = variance of X_k at date $t1$

$S12_k$ = variance of X_k at date $t2$

ρ_k = temporal autocorrelation of X_k

5. Intérêt pratique du SPR

Pour chiffrer le gain potentiel du SPR par rapport aux inventaires successifs indépendants actuellement utilisés en France (I.F.N. et inventaires d'aménagement), on a calculé la valeur du coefficient de corrélation pour plusieurs variables dans deux conditions « expérimentales » différentes (HOULLIER, 1983b) :

— à l'échelle régionale (exemple : peuplements de conifères du Rhône) on observe que les valeurs de ρ dépendent fortement des variables étudiées ; que ces valeurs sont sensiblement affectées par l'existence « d'accidents majeurs » (coupe rase, chablis ; cf. tableau 3), comme on pouvait s'y attendre ; que l'introduction de variables auxiliaires augmente nettement la valeur de ρ ;

— à l'échelle de la parcelle, on observe pour $t_2 = t_1 \approx 20$ ans des valeurs de ρ qui sont de l'ordre de 0,8 pour la surface terrière et de 0,6 pour le nombre de tiges/ha (tableau 1). Dans la perspective d'inventaires d'aménagement, ces valeurs autorisent des gains de précision assez sensibles : de l'ordre de 10-15 % pour l'estimation de $E(X_2)$ et de 25-60 % pour l'estimation de $E(D)$ (gains sur les variances d'estimation).

TABEAU 3

Exemples de valeur de l'autocorrélation temporelle, ρ , et ordre de grandeur du gain potentiel en précision, sur l'écart-type d'estimation SPR pour $p = 0,5$ et $u = n$; données : 36 placettes de conifères du Rhône (source : I.F.N.) ; la stratification consiste à éliminer les 4 placettes ayant subi un accident majeur (coupe rase ou chablis).

Value of the temporal autocorrelation, ρ , and possible precision gains due to SPR estimators (% of reduction of the standard deviation of classical estimators) for $p = 0,5$ and $u = n$.

	avec stratification ⁽¹⁾			sans stratification ⁽²⁾		
	ρ	γ_2	$\gamma_{d'}$	ρ	γ_2	$\gamma_{d'}$
Volume PB	0,65	6 %	25 %	0,56	4 %	22 %
Volume MB	0,50	3 %	18 %	0,39	2 %	14 %
Volume GB	0,90	13 %	55 %	0,87	13 %	55 %
Volume total	0,73	8 %	35 %	0,56	4 %	22 %
Total volume						
Tiges/ha	0,87	13 %	55 %	0,78	9 %	38 %
Stems/ha						
Accroissement	0,62	6 %	25 %	0,55	4 %	22 %
Growth						

γ_2 = ordre de grandeur du gain pour l'estimation de $E(X_2) = 1 - G_2$ - precision gain for the estimation of $E(X_2)$.

$\gamma_{d'}$ = ordre de grandeur du gain pour l'estimation de $E(D) = 1 - G_{d'}$ - precision gain for the estimation of $E(D)$.

PB = petit bois (diamètre < 22,5 cm) - diameter < 22,5 cm.

MB = moyen bois (22,5 cm ≤ diamètre < 37,5 cm) - 22,5 cm ≤ diameter < 37,5 cm.

GB = gros bois (37,5 cm ≤ diamètre) - 37,5 cm ≤ diameter.

(1) 32 undisturbed plots

(2) 36 plots including 4 disturbed plots (clearcuts and storms).

L'accroissement est l'accroissement courant.

(d'après HOULLIER, 1983b)

Plus généralement, l'utilisation d'un plan d'échantillonnage partiellement renouvelé permet d'obtenir des informations plus fiables sur des variables auxquelles on ne peut pas appliquer les estimateurs SPR classiques, mais qui peuvent faire l'objet d'une estimation par régression. Le passage à la futaie ou la mortalité sont de bons exemples de telles variables : on peut chercher à les « expliquer » par des variables auxiliaires (âge du peuplement, densité,...), ajuster des régressions sur les placettes permanentes, puis les appliquer aux placettes temporaires, selon le principe même du SPR (cf. 2.2 et 3.1). Pour conclure sur ce chapitre du gain pratique apporté par le SPR, il faudrait, comme au 4, introduire le coût relatif des placettes permanentes et temporaires.

6. Conclusion

Les gains de précision liés au SPR sont assez sensibles ; ces estimateurs sont cependant plus compliqués que les estimateurs usuels et leur mise en œuvre doit s'intégrer dans un plan d'échantillonnage global où interviennent souvent des stratifications multiples et où on s'intéresse simultanément à un grand nombre de variables ; c'est par exemple le cas de l'I.F.N.

Il faut alors trouver un équilibre entre des stratifications de plus en plus fines — qui permettent d'améliorer la corrélation temporelle — et la diminution des effectifs d'échantillons par strate — qui conduit à des risques élevés de biais et de non optimalité des estimateurs. La stratification par les « accidents majeurs » semble s'imposer ; par contre, l'utilisation des estimateurs SPR peut n'être envisagée que pour des ensembles de strates élémentaires contenant un nombre suffisant d'échantillons permanents (de 30 à 50).

De la même manière, il faut trouver un équilibre entre le nombre de variables auxiliaires (qui augmentent elles aussi la corrélation) et la complexité des calculs qui en découle. Selon les objectifs de l'inventaire, on peut ainsi réserver les estimations SPR aux seules variables jugées essentielles pour lesquelles une grande précision est indispensable. Pour les autres variables, se limiter aux estimateurs univariés présentés ici, en fixant a priori la valeur de ρ (exemple : $\rho = 0,75$), est sans doute tout à fait suffisant.

La méthode d'optimisation du plan d'échantillonnage proposée ici n'est pas générale : la fonction de coût et le critère de précision peuvent avoir une autre forme. Elle fournit cependant des indications utiles sur l'influence des paramètres de J, de C et de la population. Un taux de remesure de 50 % semble ainsi raisonnable.

D'une façon plus générale, il faut distinguer le choix d'un plan SPR de l'utilisation des estimateurs SPR. Le forestier est au moins aussi intéressé par les composantes de l'évolution des peuplements (mortalité, prélèvement, accroissement biologique, recrutement) que par la variation globale observée entre deux inventaires. Un plan SPR permet d'acquérir des informations sur ces composantes, même si les estimateurs SPR classiques ne s'y appliquent pas. Il peut aussi servir de base pour des études de modélisation, dans la perspective d'une actualisation régulière des résultats d'inventaire, ou d'études prospectives ; on rejoint alors la question (3) de l'introduction.

Remerciements

Je remercie J. BOUCHON, J. CHADOEUF, R. CHEVROU, P. DUPLAT, H. JOANNES, A. PAVE et G. PERROTTE pour leurs commentaires et critiques lors de diverses phases de ce travail. Les données utilisées proviennent de la Station de Sylviculture et de Production du C.N.R.F. (I.N.R.A.) et de l'I.F.N. (Echelons de Lyon pour le recueil, et de Nancy pour la saisie). Je remercie également M. HENRION pour la réalisation des figures.

Summary

Successive inventories in forest : theoretical properties and practical limits of sampling with partial replacement (SPR)

Sampling with partial replacement may be applied for successive forest inventories of the same population. This method is based on a sampling design, composed of both permanent and temporary plots (see figure 1), and on a statistical model : the classical linear model.

First, the theoretical basis of SPR is reviewed with emphasis on the underlying hypothesis. Estimators and their efficiencies are presented (see paragraphs 2.4, 2.5 and figure 2).

The limits of this technique are then pointed out through a discussion of the hypothesis :

- the linear underlying model may be generalized by a dynamic model of the population, showing the flexibility of the sampling design ;
- SPR does not take into account the spatial variation of forest. But it is possible to do it by generalizing SPR with the theory of the Regionalized Variables ;
- practical estimations require an estimation of the variances and temporal autocovariance of the population ; this induces a lack of optimality and a bias for SPR estimators ; this bias is major (see table 2) ; with at least 50 permanent plots it may be neglected.

Looking for an optimal SPR design, an optimization problem is defined and solved by minimizing a global precision criterium under a linear cost constraint. The main difficulty appears to be the definition of the optimization problem, but not the way to solve it. Anyway, a 50 % remeasurement ration seems to be reasonable (see figures 3 and 4).

Two examples provide some indications on the practical efficiency of SPR estimators for both regional or management inventories. They indicate that SPR would improve sensibly the precision of the estimations (see tables 1 and 3).

Finally, the problems related to the introduction of SPR in a more general sampling design are discussed.

Reçu le 13 novembre 1984.

Accepté le 4 décembre 1984.

Références bibliographiques

- BALLEYDIER (R.), BERTRAND (J.), 1982. La ressource forestière et sa disponibilité, une approche par l'inventaire forestier national. *Forêt méditerranéenne*, **IV** (1), 26-32.
- BENGSSON (G.), 1978. The swedish national forest survey. In *IUFRO Meeting* (S4.02 et S4.04), Bucarest, 10 pages.
- BOUCHON (J.), 1979. Structure des peuplements forestiers. *Ann. Sci. for.*, **36** (3), 175-209.
- CASSEL (C.M.), SARNDALL (C.E.), WRETMAN (J.H.), 1977. *Foundations of inference in survey sampling*. John Wiley & sons, 192 pages.
- CHEVROU (R.B.), 1982. *SPR, application aux surfaces*. Communication personnelle de la note I.F.N., IFM 722 (82), 4 pages.
- CUNIA (T.), 1965. Continuous forest inventory : partial replacement of samples and multiple regression. *Forest Science*, **11** (4), 480-502.
- CUNIA (T.), CHEVROU (R.B.), 1969. Sampling with partial replacement on 3 or more occasions. *Forest Science*, **15** (2), 204-224.
- DESABIE (J.), 1966. *Théorie et pratique des sondages*. Dunod (statistique et programmes économiques), 468 pages.
- DUPLAT (P.), PERROTTE (G.), 1983. *Inventaire et estimation de l'accroissement des peuplements forestiers*. Office National des Forêts, Section technique, 432 pages.
- FRAYER (W.E.), 1978. Sampling with partial replacement in national inventories. In *IUFRO Meeting* (S4.02 et S4.04), Bucarest, 46-51.
- FRAYER (W.E.), 1966. Weighted regression in successive forest inventories. *Forest Science*, **12** (4), 464-472.
- FRONTIER (S.), 1983. *Stratégies d'échantillonnage en écologie*. Masson, Paris et Pul, Québec, 494 pages.
- HAZARD (J.W.), 1977. Estimating area in sampling forest population on two successive occasions. *Forest Science*, **23** (2), 253-267.
- HAZARD (J.W.), PROMWITZ (L.C.), 1974. Design of successive forest inventories : optimization by convex mathematical programming. *Forest Science*, **20** (2), 117-127.
- HOULLIER (F.), 1983a. Echantillonnage en forêt : structures spatio-temporelles, objectifs de l'utilisateur et méthodes ; deux exemples : les Variables Régionalisées, le SPR. Rapport bibliographique de D.E.A., Université Lyon I, 26 pages.
- HOULLIER (F.), 1983b. Etude théorique et expérimentale de l'échantillonnage à remplacement partiel (SPR), éléments pour sa mise en œuvre. Rapport technique de D.E.A., Université Lyon I, 30 pages.
- I.F.N., 1985. *But et méthodes de l'Inventaire Forestier National*. Ministère de l'Agriculture, Paris, 65 pages.
- KENDALL (M.G.), STUART (A.), 1969. *The advanced theory of statistics*, vol. I, Charles Griffin & Co, London, 3^e Ed., 433 pages.
- KILPATRICK (D.J.), 1981. Optimum allocation in stratified sampling of forest populations on successive occasions. *Forest Science*, **27** (4), 730-738.
- LOETSCH (F.), HALLER (K.E.), ZOHRE (F.), 1973. *Forest inventory*, vol I et II, BLV Munchen, Bern, Wien, 2^e Ed., 436 + 469 pages.
- MARBEAU (J.-P.), 1976. Géostatistique forestière. Thèse de Docteur-Ingénieur, Ecole des Mines, 180 pages.
- MATERN (B.), RANNEY (B.), 1983. Variational structure in forests, implications for sampling. In *IUFRO Meeting* (S4.02), Helsinki, 1-9.
- MATHERON (G.), 1965. *Les Variables Régionalisées et leur estimation*. Masson, 305 pages.
- MATHERON (G.), 1970. La théorie des Variables Régionalisées et ses applications. Cahiers du Centre de Morphologie Mathématique de Fontainebleau, fascicule 5, 212 pages.
- NETTO (S.P.), 1983. Evaluation of different sampling designs for forest inventory in successive occasions applied to a stratified population of Eucalyptus sp. in Brazil. In *IUFRO Meeting* (S4.02), Helsinki, 67-68.

- NEWTON (C.M.), BICKFORD (C.A.), CUNIA (T.), 1974. Multivariate estimators for sampling with partial replacement on two occasions. *Forest Science*, **20** (2), 106-116.
- OMULE (S.A.Y.E.), 1981. Successive forest inventories using multistage sampling with partial replacement of units. PhD Thesis, Univ. of British Columbia, Vancouver, Canada. Résumé dans *Forestry Abstracts*, 1983, **44** (5), 216-217.
- OMULE (S.A.Y.), WILLIAMS (D.G.), 1982. Optimum allocation by dynamic programming for sampling on successive occasions with partial replacement of units. *Can. J. For. Res.*, **12** (2), 264-269.
- RANNEBY (B.), 1982. Stochastic models of variation in time and space. In *Statistics and practice, essays in honour of Bertil Matern*, Ranneby Ed., 289-297.
- SCOTT (C.T.), 1984. A new look at sampling with partial replacement. *Forest Science*, **30** (1), 157-166.
- SCHERRER (B.), 1983. Techniques de sondage en écologie. In *Frontier, Stratégies d'échantillonnage en écologie*, Masson, Paris et Pul. Québec, 63-162.
- TOMASSONE (R.), LESQUOY (E.), MILLIER (C.), 1983. *La régression, nouveaux regards sur une ancienne méthode statistique*. Masson, Paris, 177 pages.
- TRANCHEFORT (J.), 1974. *La régression, application à l'agronomie*. Institut technique des Céréales et des Fourrages, 177 pages.
- WARE (K.D.), CUNIA (T.), 1962. Continuous forest inventory with partial replacement of samples. *Forest Science*, Monograph 3, 40 pages.