

**Construction  
contrôle et possibilités d'utilisation  
de tarifs de cubage  
à double entrée**

**APPLICATION  
AU PIN LARICIO DE CORSE  
EN SOLOGNE**

PAR

**P. ARBONNIER**

Ingénieur en Chef des Eaux et Forêts  
Directeur de la Station de Biométrie du C. N. R. F.



## 1 — Généralités

### 1.0 — *Les tarifs de cubage.*

Les tarifs de cubage sont, comme chacun sait, des tables indiquant le volume moyen d'un arbre en fonction d'une ou plusieurs caractéristiques mesurées (circonférence, hauteur, décroissance, etc...). L'existence de tarifs convenables est donc une nécessité pour le praticien qui désire, pour déterminer ou prévoir la production, estimer des volumes qu'il ne peut directement mesurer.

La mesure d'un nombre très limité de caractéristiques sur un arbre ne suffit généralement pas à déterminer son volume avec exactitude: ce volume est une grandeur aléatoire caractérisée par une loi de distribution et son étude ressort ainsi de la biométrie. Les tarifs, aussi précis soient-ils, fournissent des volumes moyens: ils doivent être établis par l'analyse d'un nombre suffisant d'observations et ils s'appliquent essentiellement à l'estimation du volume de populations et non d'individus isolés.

L'aménagiste utilise fréquemment des tarifs à simple entrée (circonférence)\*. Ces derniers n'ont cependant qu'un domaine d'application limité dans le temps ou dans l'espace, car ils évoluent avec l'âge et diffèrent selon la station. Cela est si vrai qu'il existe des gammes de tarifs à simple entrée (Algan, Schaeffer).

Les tarifs à double entrée font intervenir, en plus de la circonférence\*, un autre caractère mesurable (hauteur totale, décroissance métrique). Il en existe de nombreux types (Pardé - 1961). A côté de tarifs de portée très générale comme ceux de l'Administration, on peut trouver des tarifs spécialisés applicables à une essence et à une région déterminées. Ces tarifs sont plus précis et plus commodes que les tarifs « à tout faire ». Ils demeurent cependant rares en France: leur construction nécessite la collecte et l'analyse d'une masse importante d'observations et l'entreprise a longtemps dépassé les moyens de la Station de Recherches de Nancy, notamment dans le domaine du calcul.

(\*) ou diamètre.

### 1.1 — *Les tarifs « Sologne ».*

Depuis quelques années, la Sologne a été choisie pour servir de banc d'essais à la construction de tables de production françaises et les équipes de la 1<sup>re</sup> Section — aujourd'hui Station de Sylviculture et de Production du C.N.R.F. — y ont sélectionné, inventorié, échantillonné et cubé plus de 200 placettes temporaires de Pin Sylvestre ou de Pin Laricio de Corse, essences de reboisement de cette région. Par ailleurs, la création d'une section spécialisée en Statistique — devenue la Station de Biométrie du C.N.R.F. — utilisant largement le calcul automatique, a permis d'assumer avec rapidité et sécurité le traitement de la grande masse d'observations ainsi recueillies sur le terrain. C'est donc un travail d'équipe que nous concluons aujourd'hui, auquel il convient d'associer tous les techniciens des deux stations.

Nous avons construit des tarifs de cubage à double entrée pour le Pin Laricio de Corse et le Pin Sylvestre, valables pour la Sologne. Ce n'est pas une œuvre d'avant-garde que nous présentons ici, mais cette étude particulière sur le Pin Laricio de Corse a pour but de préciser les méthodes de construction et de contrôle utilisées. Ces tarifs seront suivis d'autres, et il était nécessaire que l'utilisateur sache qu'ils ne sont pas seulement le résultat d'ajustements mathématiques, mais qu'ils ont fait l'objet d'un contrôle objectif avant leur publication. Nous nous sommes également interrogés sur les possibilités d'utilisation de ces tarifs par tous ceux qui ne disposent pas de machines à calcul coûteuses et ne peuvent recourir aux méthodes utilisées par les Stations de Recherches.

### 1.2 — *Choix de la formule de cubage.*

La circonférence à hauteur d'homme (1,30 m) et la hauteur totale de l'arbre sont les deux caractères mesurables retenus pour la confection des tarifs à double entrée. Le volume indiqué est le volume bois-fort sur écorce, découpe arrêtée à 22 cm de circonférence au fin bout (convention internationale adoptée par les Stations de Recherches).

De nombreuses formules mathématiques peuvent être adoptées pour la confection des tarifs. En dépit de l'utilisation croissante de la régression multiple — qui s'explique par les possibilités nouvelles offertes par le calcul automatique — nous avons fait confiance à la bonne vieille méthode de la régression linéaire du volume en fonction de la surface terrière, déjà citée par HUFFEL (5) et remise à l'honneur par HUMMEL (6) et ABADIE (3).

La confiance n'exclut pas l'esprit critique et il est aujourd'hui possible d'étayer sur des bases statistiques solides les constatations empiriques de HUFFEL. Lorsqu'on dispose — comme c'est le cas pour le Pin Laricio de Corse en Sologne — de données abondantes pour chaque classe de hauteur il est possible de tester la légitimité

de l'hypothèse retenue: il suffit d'appliquer le test de linéarité de la régression à toutes les classes de hauteur dont les effectifs le permettent. Le test de linéarité exige en effet, — voir par exemple VESSEREAU (10) — que la variable indépendante soit répartie en classes pour lesquelles on dispose de plusieurs valeurs de la variable dépendante. Il se résout, après une analyse de variance, à un test F de Snédécour permettant d'admettre ou de rejeter l'hypothèse de linéarité.

Le test de linéarité a été appliqué à toutes les classes de hauteur comprises entre 12 et 21 m. Les résultats obtenus figurent sur le tableau 1.2 ci-après:

L'examen de ce tableau montre que le test de linéarité a donné des résultats favorables dans 9 cas (sur 10). Seule la classe 13 m donne un test F défavorable (la probabilité correspondante est d'environ 2 %). Nous n'avons pas pensé que ce résultat suffisait à condamner l'hypothèse de linéarité qui est à la base de la construction des tarifs parce que dans le cas particulier de la classe de hauteur 13 m le nombre de classes de circonférences — 5 — est relativement faible, ce qui diminue la précision du test F.

En définitive, nous considérons qu'il est légitime d'admettre que pour un arbre donné, le volume moyen est donné par la formule:

$$(1.2) \quad V = a + b C^2$$

les constantes a et b étant des fonctions de la hauteur totale qui peuvent être déterminées par un calcul d'ajustement (méthode des moindres carrés).

### 1.3 — Matériel utilisé.

13.0 — La matière première nécessaire à l'établissement des tarifs de cubage nous a été apportée par la Station de Sylviculture et de Production (\*).

Une équipe de techniciens procède sur le terrain, pour chaque placette temporaire, aux opérations suivantes:

- a) Inventaire en circonférences, après arpentage de la placette,
- b) Détermination d'un échantillon stratifié à fraction sondée constante, l'effectif  $n_i$  de la classe de circonférence  $C_i$  étant tel que, si  $N_i$  est l'effectif de cette classe à l'inventaire:

$$\frac{n_i}{N_i} = \frac{n}{N} = C^{\alpha} \quad \text{avec } \sum_i N_i = N$$

(\*) Les placettes temporaires de Pin Laricio en Sologne sont sélectionnées dans des peuplements pleins et équiennes dont l'âge varie de 20 à 80 ans.

TABLEAU 1.2

Tests de linéarité, par classes de hauteur ; résultats globaux

Classe de hauteur (1)	Intervalle de classe (2)(cm <sup>2</sup> )	Nombre de tiges (3)	Ecartés par rapport à la régr.		Erreur (var. résiduelle)		Rapport (4):(6) (8)	Valeur de F (Seuil 5 %) (9)
			Carré moyen (4)	d. l. (5)	Carré moyen (6)	d. l. (7)		
12 m	500	94	241,7	10	199,1	82	1,21	1,95
13	500	81	256,8	4	256,8	75	3,15 *	2,49
14	500	112	284,2	7	316,2	103	0,90	2,11
15	500	119	677,1	6	476,5	142	1,42	2,16
16	500	114	962,7	9	789,7	103	1,22	1,97
17	500	122	1474,2	12	950,2	108	1,55	1,84
18	500	101	1024,1	11	1451,4	88	0,71	1,91
19	1.000	97	2397,7	6	6527,1	89	0,37	2,21
20	1.000	93	3374,3	8	3889,6	83	0,87	2,07
21	1.000	72	7948,1	7	6851,0	63	1,16	2,16

Nota: La variable indépendante étant C<sup>2</sup> l'intervalle de classe est exprimé en cm<sup>2</sup> et a dû être élargi pour les hauteurs supérieures où le « range » (en C<sup>2</sup>) est plus grand.

- c) Mesure des arbres retenus dans l'échantillon,  
 — hauteur totale,  
 — hauteur bois fort (découpe à 22 cm),  
 — circonférence à mi-hauteur (bois fort) par grimpage.

Les fiches « échantillon » de chaque placette donnent ainsi, pour un effectif moyen compris entre 20 et 30 tiges, la circonférence à 1,30 m, la hauteur totale et le volume « bois fort » (produit de la hauteur bois fort par la surface terrière à mi-hauteur).

Au total, 822 tiges ont servi à l'établissement des tarifs (tableau 1.3).

TABLEAU 1.3

*Composition des échantillons utilisés pour la construction des tarifs bruts*

Classe de hauteur (m)	Nombre de tiges	Circonférence à 1,30 m	
		minima (cm)	maxima (cm)
9	21	25	50
10	36	27	63
11	38	27	77
12	55	28	88
13	66	30	83
14	76	28	111
15	65	38	80
16	76	30	92
17	77	45	107
18	66	47	106
19	68	58	109
20	49	59	128
21	49	70	110
22	25	73	116
23	24	61	129
24	19	89	130
25	10	92	131
Total	822		

### 13.1 — Composition des classes de hauteur.

Les limites de classes sont respectivement 3,50 - 9,50, 10,50 - ... - 25,50 m. Les hauteurs totales ont été mesurées sur le terrain au dendromètre Blume-Leiss et sont arrondies à 25 cm. Il est donc

très fréquent de trouver des tiges dont la hauteur coïncide avec les limites de classe. Le classement ayant été fait au bureau par un opérateur, il a été possible de classer alternativement les tiges dans la classe inférieure ou dans la classe supérieure, la première opération ayant été randomisée pour chaque limite. Cette méthode, qui permet d'obtenir une continuité certaine dans l'évolution des volumes par classe, a été abandonnée pour les calculs plus récents en raison de plusieurs inconvénients constatés à l'usage :

— pour le *contrôle* des tarifs, il faut affecter à chaque tige un volume qui est fonction de la circonférence et de la hauteur totale. Si celle-ci coïncide avec une limite de classe, il faut à nouveau lever une incertitude — ce qui a été fait par tirage au sort mais entraîne une complication superflue.

— le classement est aujourd'hui réalisé par passage sur ordinateur des cartes « échantillon » des placettes. La manière la plus simple d'opérer consiste alors à ajouter 0,50 à la hauteur totale et à ne conserver que la partie entière de la somme, par un décalage à droite.

Ainsi :  $19,80 + 0,50 = 20,30$  appartient à la classe 20 m.  
 $19,30 + 0,50 = 19,80$  appartient à la classe 19 m.

Par ce procédé, toute hauteur coïncidant avec la limite de classe est rangée systématiquement dans la classe supérieure et il n'y a plus d'alternative.

— le mode de classement précédent est conforme aux normes de l'A.F.N.O.R.

Les tarifs établis pour le Pin Laricio de Corse en Sologne seront donc les seuls pour lesquels les arbres limites auront été alternativement répartis entre les deux classes adjacentes.

#### 1.4 — Réalisation des calculs.

La plupart des calculs ont été effectués sur l'ordinateur IBM 650 de la Faculté de Nancy :

- calcul des tarifs bruts,
- calcul des barèmes définitifs.

Provisoirement, les calculs d'harmonisation (détermination des poids, ajustement des coefficients de régression) ont été effectués sur machines de bureau. Ces opérations seront ultérieurement également automatisées.

La première tranche de calcul automatique n'a exigé que de légères modifications du programme que nous avons déjà rédigé pour les calculs courants dans les places d'expériences - cf. ARBONNIER (1). La seconde a nécessité un petit programme spécial.

## 2 — Construction des tarifs

## 2.1 — Calcul des tarifs bruts.

Pour une classe de hauteur totale H, le calcul — par la méthode des moindres carrés — de la droite représentative du volume moyen de l'arbre de circonférence C,

$$(2.1) \quad V(H) = a(H) + C^2 \cdot b(H)$$

ne présente aucune difficulté (\*) et l'utilisation d'un ordinateur ne se justifie que par la masse des données à traiter et la sécurité qu'il confère à leur traitement.

Sur deux cartes « résultats » par classe de hauteur H figurent :

- les coefficients des équations normales :  $n$ ,  $\Sigma c^2$ ,  $\Sigma c^4$ ,  $\Sigma v$ ,  $\Sigma vc^2$  et  $\Sigma v^2$ ,
- la plage d'application du tarif brut : circonférence minima et maxima de l'échantillon,
- les constantes brutes de la régression :  $a(H)$  et  $b(H)$ ,
- la somme des carrés des écarts par rapport à la droite de régression :  $Q$ ,
- la somme des carrés des écarts par rapport à la moyenne pour la variable indépendante  $c^2$  :  $\Sigma(c^2 - \bar{c}^2)^2$ .

Les deux dernières quantités interviennent dans la pondération de  $b$  pour l'ajustement.

## 2.2 — Calcul des tarifs ajustés.

22.0 — L'ajustement des constantes  $a(H)$  et  $b(H)$  a été réalisé suivant une méthode déjà éprouvée par HUMMEL (6) mais rendue ici plus rigoureuse par la pondération des valeurs brutes observées en ce qui concerne le coefficient de régression  $b$ .

## 22.1 — Ajustement du coefficient de régression.

## 22.11 — Pondération.

On sait que la méthode des moindres carrés consiste, lorsqu'on veut calculer une valeur liée  $\hat{y}_i(x)$  à partir d'un certain nombre d'observations  $y_i(x)$ , à minimiser la quantité :

$$Q = \frac{\Sigma_i (y_i - \hat{y}_i)^2}{2 \sigma_i^2}$$

$\sigma_i^2$  étant la variance liée de  $y_i$ .

(\*) Voir ABADIE (3), page 41.

Si les estimations  $s_i^2$  des variances  $\sigma_i^2$  sont très différentes, il y a intérêt à pondérer les valeurs de  $y_i$  par l'inverse de  $s_i^2$ . C'est ce que nous avons fait.

L'estimation de la variance du coefficient de régression brut  $b(H)$  a été calculée en appliquant une formule classique en statistique :

$$s_b^2 = \frac{1}{n-2} \frac{Q}{\sum_i (c_i^2 - \bar{c}^2)}$$

C'est la raison pour laquelle il est demandé à l'ordinateur de « sortir » le numérateur et le dénominateur de la seconde fraction pour chaque classe de hauteur.

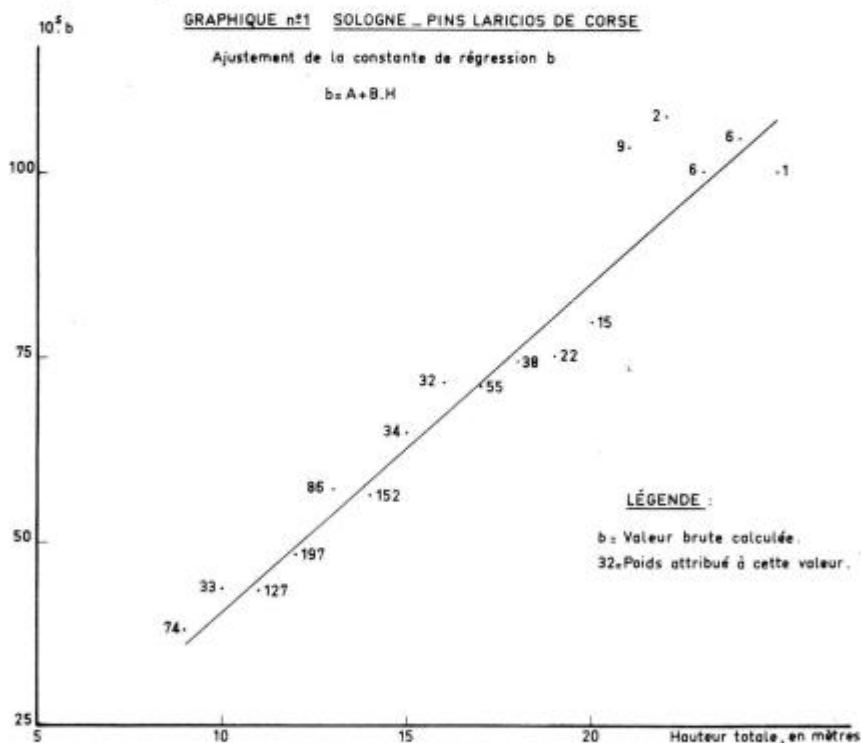


FIG. 2.1

En définitive, pour l'ajustement de  $b$ , chaque valeur brute a été affectée d'un poids  $w$

$$w = K : s_b^2$$

$K$  étant choisi de telle sorte que la plus faible valeur de  $w$  soit égale à l'unité. Les poids ont été arrondis à l'unité la plus proche. Ils

figurent sur le graphique 1 (fig. 2.1) où leur diversité peut être constatée. Il est apparu logique d'accorder une plus grande importance aux valeurs brutes qui sont connues avec une meilleure précision. C'est le cas notamment des valeurs correspondant aux classes de hauteur dont les effectifs sont les plus nombreux — toutes choses égales d'ailleurs le poids est proportionnel à  $(n - 2)$ .

### 22.12 — Calcul des valeurs ajustées.

Chaque valeur de  $b$  étant ainsi pondérée, la représentation graphique suggère l'existence d'une relation linéaire entre  $b$  et la hauteur totale  $H$ .

En utilisant la méthode des moindres carrés, on obtient la relation suivante :

$$(2.2) \quad 10^5 \cdot b = 445,5 H - 400,6 \quad (H \text{ en m.})$$

Si l'on calcule le coefficient de corrélation entre  $b$  et  $H$ , on trouve :

$$r^2 = 0,9614 \quad r = 0,98$$

La corrélation entre  $b$  et  $H$  est donc extrêmement marquée.

## 22.2 — Ajustement du terme constant de la régression.

### 22.21 — Correction avant ajustement.

On obtient habituellement la valeur de  $a$  à partir de la première équation normale.

$$(2.3) \quad na + b \sum_1 c_i^2 = \sum_1 v_i$$

la sommation étant étendue à l'échantillon.

Cette équation exprime que la droite de régression passe par le centre de gravité du nuage de points. On en tire :

$$a = \frac{\sum_1 v_i - b \sum_1 c_i^2}{n}$$

et l'on constate que la valeur de  $a$  dépend non seulement de  $n$ ,  $\sum_1 v_i$ ,  $\sum_1 c_i^2$ , calculés pour l'échantillon — et « sortis » par l'ordinateur — mais encore de  $b$  (\*).

(\*) Sans ajustement,  $b$  est implicitement lié à  $n$ ,  $\sum_1 v_i$  et  $\sum_1 c_i^2$  puisque ces valeurs sont des coefficients du système d'équations linéaires dont les solutions sont  $a$  et  $b$ .

Pour une classe de hauteur totale  $H$ , si l'on conserve les valeurs de  $n$ ,  $\Sigma v$  et  $\Sigma c^2$  observées sur l'échantillon, et que l'on retient pour  $b$  la valeur ajustée par la formule (2.2) on obtient les valeurs de  $a$  corrigées de l'ajustement de  $b$  — cf. HUMMEL (6).

Ces valeurs corrigées sont reportées sur le graphique 2 (fig. 2.2). Elles peuvent à leur tour faire l'objet d'un ajustement propre.

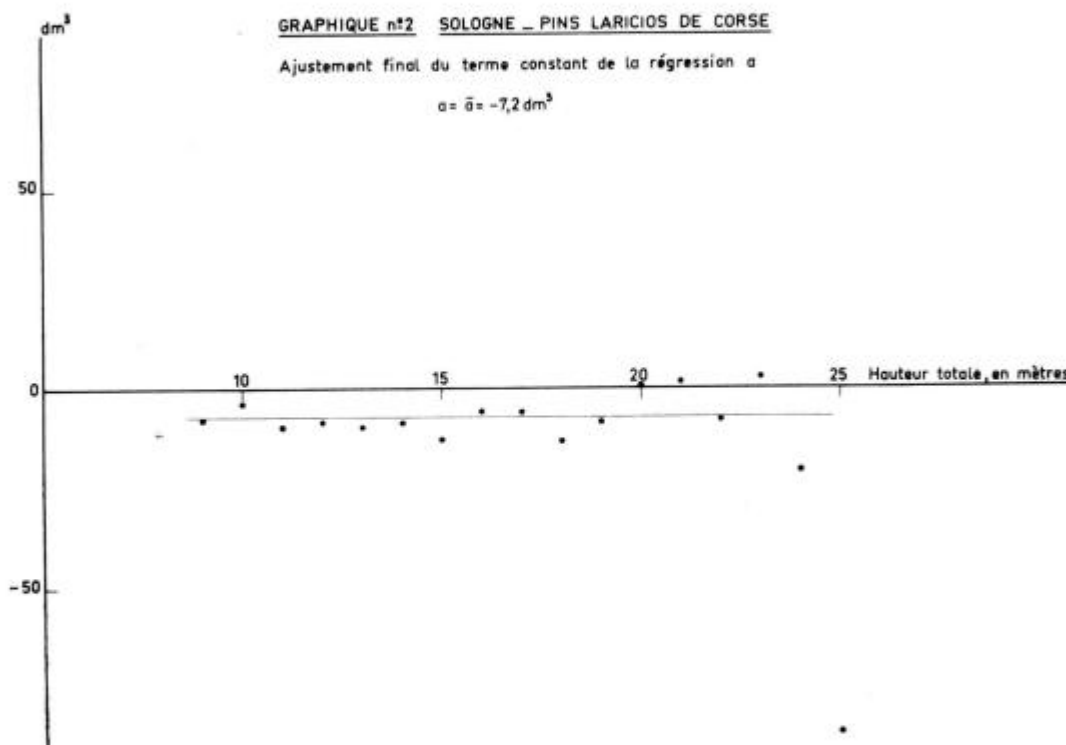


FIG. 2.2

### 22.22 — Calcul de la valeur ajustée de $a$ .

L'examen du graphique 2 suggère l'absence de liaison marquée entre  $a$  et  $H$ . Seul le point correspondant à  $H = 25$  m s'écarte de la tendance générale, mais on observera que le tarif brut pour cette classe a une précision très faible ( $w = 1$ ). Ce point n'a pas été retenu pour l'ajustement définitif.

Le calcul du coefficient de corrélation observé entre  $a$  et  $H$  donne :

$$r^2 = 0,003954 \quad r = 0,063$$

Il est aisé de vérifier que cette valeur ne diffère pas significativement de zéro. Dans ces conditions, la meilleure estimation de  $a$

est la moyenne arithmétique des valeurs observées ( $H = 25$  exceptée).

$$a = \bar{a} = -7,2 \text{ dm}^3 \approx 7 \text{ dm}^3$$

La formule condensée des tarifs de cubage devient :

$$(2.4) \quad V = 10^{-5} (445,5 H - 400,6) C^2 - 7$$

$V$  en  $\text{dm}^3$ ,  $H$  en mètres,  $C$  en centimètres.

GRAPHIQUE n°3 SOLOGNE - PINS LARICIOS DE CORSE

Tarifs de cubage par classes de hauteur

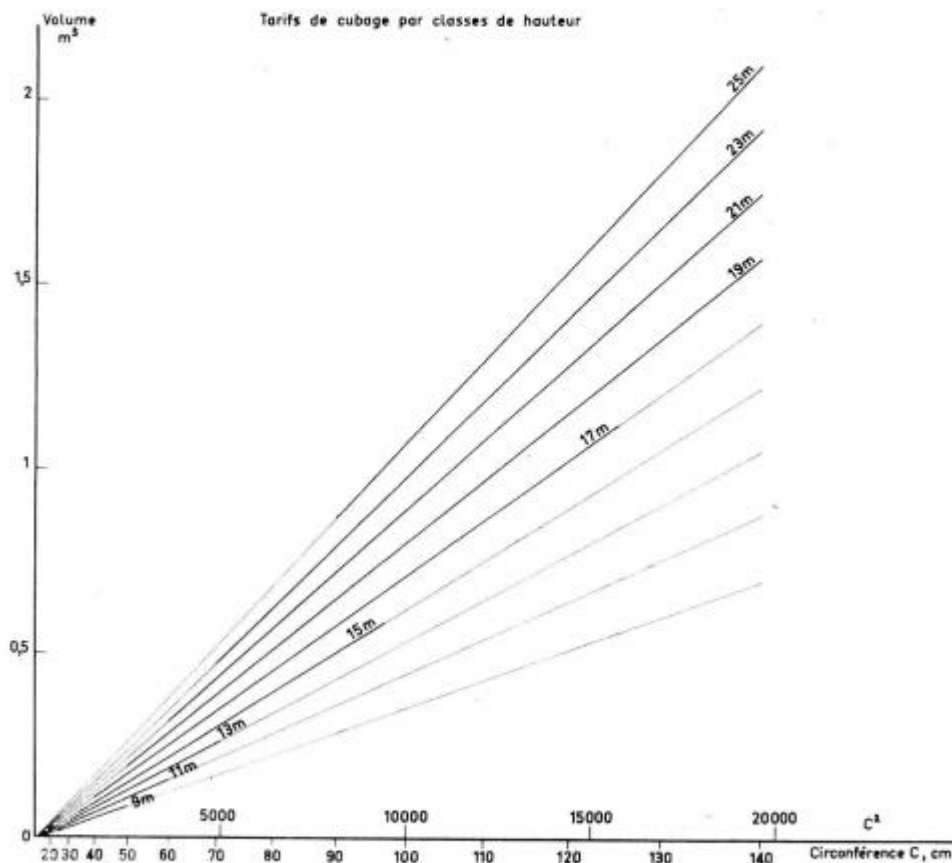


FIG. 2.3

*Remarque:* Le fait que  $a$  soit indépendant de  $H$  signifie que toutes les droites représentatives des tarifs de cubage (si l'on porte en abscisses le carré de la circonférence et en ordonnées le volume) se

coupent sur l'axe Oy au point  $C = 0 \quad V = -7 \text{ dm}^3$  (graphique 3, fig. 2.3) (\*).

### 3 — Contrôle des tarifs

#### 3.1 — Contrôle du biais.

31.0 — En dépit des garanties qu'apporte à l'établissement de ces tarifs de cubage l'emploi de méthodes objectives et d'un nombre important de données, il demeure nécessaire de contrôler que l'ajustement est satisfaisant et, en particulier, n'entraîne aucune erreur systématique — le fameux « biais » des statisticiens.

Un tel contrôle est possible sur tous les échantillons dont les tiges ont été utilisées comme données pour la construction des tarifs. Cependant, au cours de 1963, d'autres placettes ont été inventoriées et cubées en Sologne, alors que les tarifs étaient déjà calculés. Cette campagne de cubage ultérieure nous a permis de disposer d'échantillons nouveaux, *absolument indépendants de ceux qui avaient déjà été utilisés* et par conséquent, susceptibles de fournir de précieuses possibilités de contrôle. Le contrôle du biais a donc été étendu à cette seconde tranche pour apprécier s'il était opportun de réviser les tarifs.

#### 31.1 — Contrôle sur le groupe de placettes ayant fourni l'échantillon.

Ce contrôle ne peut valablement s'exercer qu'en comparant les estimations obtenues à l'aide des tarifs à des volumes réellement observés, sans que viennent interférer par exemple, des erreurs de sondage. Il doit de même s'exercer sur des groupes d'arbres et non sur des individus isolés. C'est donc sur les différents échantillons ayant servi à l'établissement des tarifs de cubage individuel de placettes que le contrôle du biais a été opéré. On a procédé aux cubages « tarif » de ces échantillons et obtenu, pour chacun d'eux, un couple de valeurs (volume réel, volume estimé). Il devenait alors possible de tester statistiquement si la moyenne des écarts observés diffère significativement de zéro. S'il n'en est rien, les tarifs ne sont entachés d'aucune erreur systématique.

(\*) On remarquera également que V étant une fonction linéaire de H on peut, sans erreur, interpoler linéairement entre deux hauteurs métriques.

Exemple:  $V(19,40) = V(19) + 0,4 [V(20) - V(19)]$

Une telle interpolation ne présente d'intérêt que si H est une hauteur moyenne connue avec une certaine précision.

On retrouvera en annexe 1 le tableau 1 mentionnant les valeurs observées (\*), les écarts qui constituent l'erreur absolue et enfin, l'erreur relative pour les 32 échantillons du premier groupe. Il se résume ainsi :

Somme des volumes estimés	:	299 043 dm <sup>3</sup>	
Somme des volumes réels	:	297 989 dm <sup>3</sup>	
		<hr/>	
erreur totale	:	1 054 dm <sup>3</sup>	soit 0,3 %
erreur moyenne	:	32,9 dm <sup>3</sup>	

Si l'on s'inquiète de la répartition des erreurs absolues, on constate que celles-ci sont normalement réparties (\*\*\*) autour de la moyenne observée (32,9) avec un écart-type estimé à 383,4 dm<sup>3</sup>. L'erreur type est de 67,8 dm<sup>3</sup>.

Le rapport de la moyenne à l'erreur type (test de Student-Fisher) :

$$t = 32,93 : 67,8 = 0,486$$

correspond à une probabilité un peu supérieure à 60 %.

On peut donc en déduire que l'utilisation des tarifs fournit une estimation absolument correcte des volumes, sans erreur systématique.

L'étude des erreurs relatives donne des conclusions identiques. Celles-ci sont normalement réparties autour d'une moyenne observée (0,364 %) qui ne diffère pas significativement de zéro (erreur type = 0,723 %).

### 31.2 — Contrôle sur le groupe de placettes cubées postérieurement à l'établissement des tarifs.

Ce contrôle concerne 21 échantillons mesurés au cours de l'été 1963. Il a donné les résultats suivants (annexe 1, tableau 2) :

Somme des volumes estimés	:	167 630 dm <sup>3</sup>	(tarif)
Somme des volumes réels	:	165 155 dm <sup>3</sup>	
		<hr/>	
erreur totale	:	2 475 dm <sup>3</sup>	soit 1,4 %
erreur moyenne	:	117,9 dm <sup>3</sup>	

(\*) Conformément à la résolution adoptée au 13<sup>e</sup> Congrès de l'I.U.F.R.O. (Vienna, 1961) par la Section 25 — rapport 25/1-§ 3.4 — nous publions les données chiffrées essentielles, mais seulement en annexe pour ne pas alourdir le texte. Les autres peuvent être obtenues de la Station de Biométrie du C.N.R.F.

(\*\*) Le contrôle de la normalité a été pratiqué suivant une méthode graphique exposée par TOMASSONE (9) en utilisant l'intervalle de confiance de la droite de Henry. Cette normalité simultanée des erreurs absolues et des erreurs relatives peut inquiéter certains... en fait, le test ne démontre pas la normalité: il indique que cette hypothèse n'a pas de raison d'être rejetée.

On remarque immédiatement que l'erreur relative est plus importante. Si cependant on examine le tableau des écarts observés, on s'aperçoit qu'une seule placette (S 126) contribue à elle seule pour 2 091 dm<sup>3</sup> (8,11 %) à l'erreur totale. En fait, cette placette aurait pu être écartée de la comparaison car son échantillon n'a pu être que partiellement cubé. Les caractéristiques du peuplement de 60 ans qui la constitue *débordent en effet les limites normales d'application des tarifs* (circonférences supérieures à 150, hauteurs totales supérieures à 25 m) et l'on serait en droit de la considérer comme aberrante — ce qui ramènerait l'erreur totale à 384 dm<sup>3</sup> (0,2 %) et l'erreur moyenne à 19,2 dm<sup>3</sup>. Cependant, cette observation isolée ne suffit pas à condamner les tarifs puisque l'hypothèse d'absence de biais peut être conservée.

La valeur observée pour le *t* de Student est en effet

$$t = \frac{117,86}{111,0} = 1,06$$

et correspond à une probabilité supérieure à 20 %.

Toutefois, l'erreur absolue n'est plus ici distribuée normalement et la validité du test de Student pourrait être contestée. Dans ces conditions, il est encore possible de faire appel à d'autres tests qui n'impliquent pas la normalité de la distribution. Tel est le cas notamment du *test de Wilcoxon* fondé sur l'étude des rangs — voir annexe 2 où ce test est exposé.

La constante *W* a pour valeur 452. Elle peut être considérée comme normalement distribuée autour d'une moyenne égale à  $21^2 = 441$  avec une variance égale à 6321 (écart-type: 79,5), on trouve alors :

$$t = 0,138$$

et la probabilité (unilatérale) est de 44 %.

Si l'on étudie, non plus les erreurs absolues, mais les erreurs relatives, on constate que celles-ci sont normalement distribuées autour d'une moyenne observée de 0,23 % avec un écart-type de 4,78 %. L'écart-type de la moyenne est alors 1,04 et la valeur du *t* de Student :

$$t = 0,23 : 1,04 = 0,221$$

autorise encore à conclure que la moyenne observée ne diffère pas significativement de zéro.

31.3 — *Contrôle sur l'ensemble des deux groupes (53 placettes).*

Comme il fallait s'y attendre, on aboutit aux conclusions suivantes :

a) Les erreurs absolues ont une distribution sensiblement normale autour d'une moyenne observée de 66,6 dm<sup>3</sup> (0,76 %) avec un écart-type estimé de 437 dm<sup>3</sup>. La moyenne ne diffère pas significativement de zéro.

b) Les erreurs relatives (en %) sont distribuées normalement autour d'une moyenne observée (0,31 %) qui ne diffère pas significativement de zéro.

c) Un test de Wilcoxon confirme que les distributions des 53 volumes réels et des 53 volumes estimés peuvent être considérées comme identiques.

d) Un autre teste de Wilcoxon montre qu'on peut admettre l'identité des distributions des 32 erreurs absolues constatées sur le premier groupe et des 21 erreurs absolues observées sur le second groupe.

Deux conclusions se dégagent de ces contrôles :

1) Les tarifs ne sont pas biaisés et fournissent une estimation absolument correcte des volumes.

2) Les tarifs calculés à l'aide des tiges de la première tranche s'appliquent correctement à la seconde tranche et ne nécessitent pas d'être révisés.

*Remarque:* Il serait cependant dangereux d'extrapoler ces tarifs, en utilisant la formule (2.4) ci-dessus. L'exemple de la placette S 126 montre que lorsqu'ils sont appliqués à des arbres dont les dimensions atteignent ou dépassent les limites que nous avons adoptées, on court le risque de surestimer les volumes. La valeur observée du terme constant corrigé pour la classe de 25 m est une autre indication en ce sens. On devra donc se montrer prudent lorsque les peuplements atteignent de telles dimensions, assez exceptionnelles du reste en Sologne.

3.2 — *Dispersion.*

Après avoir montré que les tarifs fournissent en moyenne des estimations correctes des volumes, on peut s'interroger sur la dispersion de celles-ci, c'est-à-dire sur leur précision. En fait aucun tarif, aussi précis soit-il, ne pourra jamais supprimer la dispersion inhérente à la variabilité d'un caractère (\*) (le volume) mesuré sur di-

(\*) La dispersion du volume traduit la variabilité du coefficient de forme.

vers individus d'une population, lorsque cette population est définie par une espèce ligneuse occupant une aire géographique comme la Sologne.

Il peut être cependant intéressant d'avoir une idée sur les erreurs possibles. A ce point de vue, l'étude qui a été faite sur la distribution des erreurs relatives — distribution que l'on peut considérer comme normale — apporte des indications assez valables puisqu'elles concernent 53 placettes.

Nous avons observé un écart-type de 4,3 % pour l'erreur relative. Compte tenu de la normalité de la distribution, on peut conclure que cette erreur ne sera dépassée en moyenne qu'une fois sur trois et qu'il y a 95 chances pour 100 que l'erreur relative ne dépasse pas 8,7 %.

Rappelons que ces résultats sont valables lorsqu'on estime individuellement chaque arbre — connaissant sa circonférence et sa hauteur totale — et pour des échantillons d'effectif comparable à ceux que nous avons cubés (généralement 22 à 24 arbres). Il n'est pas douteux que pour des effectifs plus importants la précision se trouve accrue par la compensation des erreurs commises sur l'estimation individuelle du volume de chaque tige.

#### 4 — Utilisation des tarifs

4.0 — Dans la pratique, un cubage tel que nous venons de le définir, est sans intérêt, sauf cas très particulier. L'utilisateur des tarifs procédera par sondage, par exemple en faisant un inventaire de son peuplement par classe de circonférence et un échantillonnage des hauteurs lui permettant de déterminer le ou les tarifs à appliquer. Dans ce cas, une *erreur de sondage* s'ajoute évidemment aux erreurs aléatoires dues à l'emploi des tarifs, quelle que soit encore la précision de ceux-ci. Cette erreur de sondage peut même être systématique si l'on adopte une valeur inexacte de la hauteur pour le choix du tarif.

Nous avons donc entrepris d'analyser l'influence du mode d'utilisation des tarifs et pour cela nous avons utilisé 43 placettes parmi celles qui ont été inventoriées et cubées par la Station de Sylviculture.

Lorsqu'il s'agira de peuplements (et non plus d'échantillons), nous n'en connaissons plus le volume réel, mais une estimation de celui-ci, établie suivant les méthodes définies par ABADIE (3) et assortie d'une erreur-type. A ce mode d'estimation précis mais qui exige sur le terrain la mesure effective du volume des arbres de l'échantillon et, au bureau, des calculs relativement simples mais fastidieux, on peut opposer des méthodes plus rapides substituant la mesure de hauteurs à des cubages plus ou moins détaillés. On devra cependant, ne pas oublier dans de telles comparaisons qu'en

confrontant deux estimations — même indépendantes en ce qui concerne la procédure — on combine deux erreurs de sondage, ce qui n'était pas le cas pour le contrôle du biais sur les échantillons.

Les échantillons utilisés pour le calcul du volume dans les places d'expériences sont des échantillons stratifiés par classe de circonférence, la fraction sondée étant sensiblement constante (il n'est pas possible d'obtenir une constance rigoureuse avec des échantillons limités à une trentaine de tiges). L'effectif d'une classe de circonférence dans l'échantillon est ainsi, à peu de choses près, proportionnel à celui de cette classe dans le peuplement, établi par un inventaire préalable. S'agissant de peuplements équiennes, l'échantillon est donc bien représentatif et l'on peut penser que l'utilisation correcte de tarifs valables fournira une estimation non biaisée du volume.

L'utilisation des tarifs peut se concevoir de plusieurs façons à partir des échantillons ainsi constitués :

1 — en estimant le volume  $\bar{v}$  d'un arbre moyen qu'on multipliera par le nombre d'arbres  $N$  du peuplement pour obtenir le volume de celui-ci.

$$V = N \bar{v}$$

2 — en estimant, pour chaque classe de circonférence  $C_i$  un volume moyen  $\bar{v}_i$  que l'on multipliera par l'effectif  $N_i$  de cette classe et en effectuant la sommation :

$$V = \sum_i N_i \bar{v}_i$$

3 — en calculant une régression en  $C^2$  propre au peuplement à partir des volumes de l'échantillon obtenus à l'aide des tarifs — la mesure des hauteurs au dendromètre étant substituée à celle des volumes — ce qui supprime sur le terrain l'obligation du grimpage.

Ayant ainsi calculé  $a$  et  $b$  pour l'échantillon, on cube le peuplement par application de la formule :

$$V = N a + b \sum N_i C_i^2 \quad N = \sum N_i$$

#### 4.1 — Estimation par le volume de l'arbre moyen.

4.1.0 — Si l'on admet la formule utilisée pour l'établissement des tarifs

$$v_i = a + b c_i^2$$

l'arbre de surface terrière moyenne (circonférence  $c_m$ ) est aussi l'arbre de volume moyen  $\bar{v}$ .

On a en effet :

$$V = Na + b \sum_i c_i^2 = N\bar{v}$$

et, par définition :

$$c_g^2 = \frac{1}{N} \sum_i c_i^2 \quad \text{d'où} \quad \bar{v} = a + b c_g^2$$

Le calcul de la circonférence moyenne  $c_g$  ne soulevant aucune difficulté, il reste à déterminer  $a$  et  $b$ , c'est-à-dire à choisir la classe de hauteur appropriée et le tarif correspondant.

Cette classe de hauteur peut être déterminée par le calcul. On peut aussi espérer l'observer sur un échantillon.

#### 41.1 — Hauteur moyenne calculée.

Un échantillon stratifié fournit une estimation correcte de la hauteur moyenne arithmétique et la première idée qui vient à l'esprit est d'utiliser cette *hauteur moyenne arithmétique* pour « entrer » dans le tarif correspondant. On peut cependant démontrer sans difficultés (voir annexe 3) que l'utilisation de la hauteur moyenne arithmétique conduit à une estimation erronée du volume total. Elle introduit une erreur systématique et le volume, ainsi déterminé, est *toujours inférieur*, en moyenne, à sa valeur réelle.

On peut du reste le vérifier en consultant le tableau 1 de l'annexe 3 où figurent les résultats de l'application de tarifs fondés sur l'emploi de la hauteur moyenne arithmétique au cubage des échantillons et des populations des 23 placettes de la première tranche.

Pour les échantillons, on note une erreur absolue de — 11 250 dm<sup>3</sup> soit — 5,30 % du volume réel. Les erreurs relatives sont distribuées autour d'une moyenne de — 6,95 % dont l'erreur type est estimée à 1,32 %. Le test de Student ( $t = 5,27$ ) est très fortement significatif, ce qui montre l'existence d'un biais.

Pour les populations, on observe un écart absolu de — 109 437 dm<sup>3</sup> soit — 5,30 % du volume estimé par la méthode Abadie. Les écarts relatifs sont distribués autour d'une moyenne de — 6,13 % avec une erreur type de 0,921 %. Ici encore, le test  $t$  de Student ( $t = 6,66$ ) oblige à conclure que les estimations du volume ne sont pas compatibles entre elles.

Il est cependant possible de calculer une hauteur moyenne adéquate permettant de choisir le tarif à appliquer à l'arbre moyen. C'est une moyenne pondérée par les surfaces terrières. En désignant par  $\bar{H}_i$  la hauteur moyenne de la classe de circonférence  $C_i$ , d'effectif  $N_i$  on a :

$$H_L = \frac{1}{\sum_i N_i C_i^2} \sum_i N_i C_i^2 \bar{H}_i$$

TABLEAU 4.1  
Sennely (Loiret)  
Pin Laricio de Corse - Placette S 70

	Population			Echantillon							
	$N_i$	$V_i$	$H_i$	$n_i$	Hauteurs (mesurées sur le terrain) $h_{ij}$				$S_i = \sum_j h_{ij}$	$\bar{h}_i$	$\bar{H}_i$
30	52	29	10	3	9,25	10,50	11,50		31,25	10,41	10
40	53	72	12	3	9,75	13,25	14,25		37,25	12,41	12
50	86	161	16	4	14,00	16,50	16,75	15,50	62,75	15,68	16
60	81	251	17	4	20,00	16,25	16,75	16,75	69,75	17,43	17
70	51	366	18	3	16,25	19,00	19,25		54,50	18,16	18
80	35	538	20	2	20,50	19,75			40,25	20,12	20
90	18	718	21	1	21,00				21,00	21,00	21
100	5	889	21	-					-	-	21
TOTAUX	381			20					316,75	15,84	16

1 — Volume par arbre moyen

$$v_{18} (58,6) = 249 + 5 = 254 \quad (*)$$

1 — a — Hauteur moyenne arithmétique  
La hauteur moyenne arithmétique  $\bar{H}$  est

$$v_0 = 381 \times 254 = 96.774 \text{ dm}^3$$

$$\bar{H} = \frac{S}{n} = \frac{316,75}{20} = 15,84 \approx 16 \text{ m}$$

1 — c — Hauteur moyenne « observée »  
58,6 appartient à la classe des 60 donc

$$HG = \bar{H}_{60} = 17 \text{ m}$$

L'arbre de surface terrière moyenne a pour circonférence  $C_g$ :

$$V_{17} (58,6) = 234 + 5 = 239 \quad (*)$$

$$VG = 381 \times 239 = 91.059 \text{ dm}^3$$

$$C_g^2 = \frac{\sum N_i C_i^2}{\sum N_i} = \frac{1.307.900}{381} = 3432,81$$

2. — Volume par classes de circonférence

$$\sum N_i V_i = 54.366 \text{ dm}^3 = V_{st}$$

$$C_g = 58,6 \text{ cm}$$

3. — Volume calculé 1<sup>re</sup> Section

$$V = 95.595 \pm 2004$$

$$v_{16} (58,6) = 196 + \frac{3,6}{5} (235 - 196) \\ = 196 + 28 = 224$$

4. — Erreurs observées

$$V_{tm} = 381 \times 224 = 85.344 \text{ dm}^3$$

$$\Delta_m = 85.344 - 95.595 = -10.251 \text{ dm}^3$$

$$\Delta_c = 96.774 - 95.595 = 1.179 \text{ dm}^3$$

$$\Delta_g = 91.059 - 95.595 = -4.536 \text{ dm}^3$$

1 — b — Hauteur moyenne « calculée »

$$H_L = \frac{1}{\sum n_i C_i^2} \cdot \sum C_i^2 S_i = \frac{1.190.450}{67.500} \\ = 17,64 \approx 18 \text{ m}$$

$$\Delta_{st} = 94.366 - 95.595 = -1.229 \text{ dm}^3$$

(\*) Interpolation sur tarif de cm en cm.

et pour un échantillon à fraction sondée constante :

$$H_L = \frac{1}{\sum_i n_i C_i^2} \sum_i n_i C_i^2 H_i$$

comme nous le démontrons par ailleurs (annexe 3).

Cette hauteur moyenne calculée coïncide avec la hauteur moyenne de Lorey, telle qu'elle est rappelée notamment par ASSMANN (2).

Un exemple d'application d'un tel calcul figure sur le tableau 4.1, § 1.b.

D'autre part, les résultats obtenus à l'aide de cette méthode de cubage sont reportés sur un tableau récapitulatif global des différentes méthodes d'estimation (colonnes « volume par hauteur moyenne calculée ») - tableau 1, annexe 4.

Ils seront discutés plus loin.

#### 4.1.2 — Hauteur moyenne observée.

Sans être spécialement compliqué, le calcul de la hauteur moyenne, pondérée comme il vient d'être dit, nécessite un échantillon d'une certaine importance, dont la composition reflète exactement l'inventaire du peuplement. Lorsque cet inventaire a été dressé, l'échan-

tilton s'en déduit par le choix d'une fraction de sondage  $\frac{n}{N}$ . Dans

les placettes étudiées, cette fraction était choisie de telle sorte que l'effectif total de l'échantillon soit compris entre 20 et 30 (généralement de l'ordre de 23).

Il faut donc rechercher sur le terrain — et on devrait en principe les désigner au hasard pour une classe donnée — les arbres représentant les circonférences convenables.

Il se peut que la précision exigée des cubages ne justifie pas une telle procédure et l'on peut être amené à rechercher une hauteur moyenne *observable*. On songe alors naturellement à la hauteur moyenne de la classe de circonférence qui renferme l'arbre de surface terrière moyenne  $C_g$ .

Ainsi, dans l'exemple de la placette S 70 (tableau 4.1) on a trouvé une circonférence moyenne  $C_g = 58,6$  cm. Cet arbre appartient à la classe des 60 dont la hauteur moyenne estimée sur l'échantillon est de 17,43 m. Que penser de l'estimation fondée sur le choix d'une hauteur « tarif » de 17 m?

Aussi simple et séduisante soit-elle, cette méthode n'offre pas, malheureusement, de meilleures garanties *théoriques* que celle qui

TABLEAU 4.2

*Arbre de circonférence moyenne - Hauteur observée et hauteur calculée*

Placette	Circonférence "moyenne"	Hauteur "moyenne"		Différence $H_L - H_G = x$	Placette	Circonférence "moyenne"	Hauteur "moyenne"		Différence $H_L - H_G = x$
		calculée $H_L$	observée $H_G$				calculée $H_L$	observée $H_G$	
S 1	68,4	18,53	17,56	0,97	S 73	56,0	16,90	16,33	0,57
S 7	79,7	21,52	21,35	0,17	S 91	102,1	22,46	21,96	0,50
S 8	68,2	17,59	17,50	0,09	S 92	75,3	19,38	18,95	0,43
S 9	66,1	17,15	16,75	0,40	S 93	75,7	19,19	19,05	0,14
S 12	98,6	24,32	25,06	-0,74	S 104	62,0	16,70	16,83	-0,13
S 27	54,3	15,15	14,87	0,28	S 105	70,0	19,23	19,15	0,08
S 30	107,6	25,75	25,06	0,69	S 106	48,0	14,33	15,06	-0,73
S 31	73,3	20,07	19,41	0,66	S 107	68,7	18,46	18,46	0,00
S 37	69,7	18,96	18,25	0,71	S 108	57,4	14,37	14,55	-0,18
S 41	65,2	17,70	17,35	0,35	S 109	56,4	16,24	15,58	0,66
S 42	90,3	21,14	21,37	-0,23	S 110	68,3	19,99	19,25	0,74
S 47	62,1	11,19	10,58	0,61	S 114	64,9	17,62	16,92	0,70
S 52	94,7	21,76	22,18	-0,42	S 115	71,7	17,90	17,90	0,00
S 58	78,1	20,63	19,83	0,80	S 116	56,8	15,42	15,40	0,02
S 60	73,4	16,68	16,00	0,68	S 118	74,3	20,73	19,80	0,93
S 61	75,8	18,81	18,50	0,31	S 123	46,6	12,61	12,42	0,19
S 64	49,3	14,11	13,50	0,61	S 129	72,4	16,42	16,50	-0,08
S 67	57,1	17,14	16,70	0,44	S 132	67,0	18,47	18,35	0,12
S 69	41,2	12,13	11,91	0,22	S 144	57,4	15,23	15,39	-0,16
S 70	58,6	17,64	17,43	0,21	S 152	57,2	15,17	14,42	0,75
S 71	43,2	12,95	12,17	0,78	S 153	62,1	14,78	14,94	-0,16
S 72	36,4	10,43	10,37	0,06	N=43	TOTAUX	752,95	740,91	+12,04 : 43 =0,28

 $H_L$  = hauteur moyenne de Lorey $H_G$  = hauteur moyenne (arithmétique) de la classe de circonférence  $C_G$ 

$$\Sigma x^2 = 10,7494$$

$$C = 3,3712$$

$$Q = 7,3782$$

$$s^2 = 0,175671$$

$$s^2_{\frac{x}{\bar{x}}} = 0,00408538$$

$$s = 0,42$$

$$s_{\frac{x}{\bar{x}}} = 0,0635$$

fait appel à la hauteur moyenne arithmétique. La hauteur moyenne calculée définie en 41.1 est une moyenne pondérée de la forme :

$$H_L = \frac{\sum w_i \bar{H}_i}{\sum w_i}$$

il n'y a pas a priori, de raison pour qu'elle coïncide avec l'une des hauteurs moyennes de classe  $\bar{H}_i$ .

Cela est si vrai, que si l'on compare — pour les 43 placettes étudiées — les valeurs respectives des hauteurs moyennes « calculées » et « observées » — tableau 4.2 — on peut montrer que la hauteur moyenne calculée est significativement supérieure à la hauteur moyenne observée.

La différence moyenne (0,28 m) ne peut être l'expression aléatoire d'une moyenne nulle. Le *t* de Student a pour valeur :

$$t = \frac{0,28}{0,0635} = 4,40$$

et la probabilité correspondante est infime, de l'ordre de  $10^{-5}$ .

La détermination, par l'observation, de la hauteur de l'arbre moyen — fondée sur la hauteur moyenne arithmétique de la classe contenant l'arbre de circonférence moyenne  $c_g$  — n'est donc pas absolument correcte, au sens statistique du terme.

La valeur moyenne du biais n'est pourtant pas considérable et l'on peut se demander si, compte tenu du fait que les hauteurs sont groupées par classes (de mètre en mètre), il n'est pas cependant possible d'utiliser cette hauteur moyenne « observée ». Pour tenter de répondre à cette question, nous avons utilisé cette hauteur moyenne « arrondie » pour le choix du tarif et cubé 43 placettes : les résultats se trouvent également sur le tableau 1 de l'annexe 4 où l'on relève :

$\sum V_o$  = somme des estimations par le tarif « hauteur moyenne observée »,

$\sum V_1$  = somme des estimations « Station de Recherches ».

$$\sum V_o = 3.716.552 \text{ dm}^3$$

$$\sum V_1 = 3.789.150 \text{ dm}^3$$

	- 72.607 dm <sup>3</sup>	(- 1,92 %)
--	--------------------------	------------

La comparaison avec les estimations « Station de Recherches » (tarif individuel de placette) peut concerner soit les écarts absolus, soit les écarts relatifs.

Les écarts absolus ont une distribution sensiblement normale autour d'une moyenne observée de  $-1\,688,5 \text{ dm}^3$  avec un écart-type estimé de  $5\,226 \text{ dm}^3$ . L'écart-type de la moyenne (erreur-type) est alors de  $797 \text{ dm}^3$ , d'où :

$$t = \frac{1\,688,5}{797} = 2,12$$

Cette valeur est significative au seuil de 5 %. ( $P = 4,2 \%$ ).

Les écarts relatifs (en %) sont distribués normalement autour d'une moyenne observée de  $-1,83 \%$  avec un écart-type estimé de  $6,03 \%$ . L'erreur-type correspondante est de  $0,92$ , ce qui donne :

$$t = 1,99$$

Valeur dont la probabilité est d'environ 5,2 %.

L'utilisation de la hauteur moyenne « observée » apparaît ainsi sujette à caution si l'on s'en tient à des considérations théoriques. Cette opinion n'est pas infirmée par la comparaison de la somme des volumes estimés par cette méthode avec celle des estimations « Station de Recherches ». Ces dernières sont, en effet, assorties d'une erreur-type et les volumes des placettes sont indépendants : l'erreur-type globale ET s'obtient alors par la formule classique :

$$ET = \sqrt{ET_1^2 + \dots + ET_{45}^2}$$

on trouve  $ET = 14\,806 \text{ dm}^3$ .

On observe d'autre part un écart :

$$\Sigma V_0 - \Sigma V_1 = -72\,607 \text{ dm}^3 = -4,9 \text{ ET}$$

Cet écart, exprimé en erreur-type est élevé, mais il faut rappeler qu'en comparant  $\Sigma V_1$  et  $\Sigma V_0$ , on cumule les erreurs d'estimation des deux méthodes.

Il n'est donc pas possible de discerner dans cette comparaison si l'importance de l'écart doit être attribuée à une erreur systématique (biais) ou à une dispersion importante des estimations fondées sur l'usage de tarifs déterminés pour la hauteur moyenne « observée ». Nous examinerons la question plus en détail au chapitre 5. Disons simplement que l'utilisateur qui n'exige pas une trop grande précision pourra se contenter de cette donnée relativement simple à mesurer : la hauteur moyenne de la classe qui contient l'arbre de surface terrière moyenne, arrondie au mètre le plus proche.

Si nous revenons à l'estimation du volume par l'utilisation d'une hauteur moyenne calculée, on a trouvé (annexe 4, tableau 1):

$\Sigma V_c$ : somme des volumes estimés par le tarif « hauteur moyenne calculée »

$\Sigma V_1$ : somme des volumes estimés « Station de Recherches »

$$\Sigma V_c = 3.798.477 \text{ dm}^3$$

$$\Sigma V_1 = 3.789.159 \text{ dm}^3$$

$$\text{Différence:} \quad + \quad 9.318 \text{ dm}^3 \quad (+ 0,25 \%)$$

L'étude de la distribution des écarts constatés aboutit aux conclusions suivantes:

— les écarts absolus sont distribués autour d'une moyenne observée de + 216,7 dm<sup>3</sup> avec un écart-type de 4 305 dm<sup>3</sup>. L'erreur-type est alors de 656,5 dm<sup>3</sup> d'où  $t = 0,33$ .

— les écarts relatifs sont distribués autour d'une moyenne observée de 0,34 % avec un écart-type de 4,80 %. L'erreur-type est ici de 0,733 % d'où  $t = 0,46$ .

Les valeurs observées dans les deux cas pour le  $t$  de Student ne sont pas significatives et par conséquent, le choix d'un tarif fondé sur la hauteur moyenne calculée (ou hauteur de Lorey) fournit une estimation du volume qui n'est pas biaisée par rapport à celle que l'on obtient par la méthode habituellement utilisée par la Station de Sylviculture et de Production. Il ne fait donc aucun doute que, lorsqu'on veut cuber un peuplement par le volume de l'arbre moyen, le tarif fondé sur la hauteur moyenne de Lorey doit être préféré — dans le cas du Pin Laricio de Corse, soulignons le.

Si l'on note avec PARDE (8) que la hauteur moyenne de Lorey est très proche de la hauteur moyenne de l'arbre de surface terrière médiane (\*) on peut trouver là une méthode simplifiée de détermination de la hauteur moyenne à adopter pour le choix du tarif. Certes, il demeure à calculer la surface terrière médiane et un tel calcul peut paraître compliqué: en réalité, il est sérieusement allégé si, au lieu de manipuler des surfaces terrières, on opère sur les carrés des classes de circonférence, qui sont à un facteur simple près les carrés des premiers nombres entiers que tout le monde connaît par cœur... L'avantage essentiel du choix de la médiane est de li-

(\*) La médiane d'une distribution observée est l'abscisse du point qui laisse de part et d'autre autant d'observations. Dans le cas d'observations groupées par classes (d'égale étendue) elle est calculée par interpolation. L'affaire est ici plus scabreuse: l'étendue des classes est variable puisqu'il s'agit d'une échelle de carrés.

miser à une seule classe les mesures de hauteur, ce qui évite d'avoir recours à un échantillon stratifié. En voici une application sur l'exemple cité au tableau 4.1.

TABLEAU 4.3

*Détermination de l'arbre de surface terrière médiane*

$c_i$ (1)	$N_i$ (2)	$c_i^2 : 100$ (3)	$N_i c_i^2 : 100$ (4)	Total cumulé (5)
30	52	9	468	468
40	53	16	848	1.316
50	66	25	2.150	3.466
60	81	36	2.916	6.382
70	51	49	2.499	8.881
80	35	64	2.240	11.121
90	18	81	1.458	12.579
Situation de la médiane $12.579 : 2 = 6.290$				

On retrouve en colonne (1) et (2) les chiffres du tableau 4.1. Les chiffres de la colonne (5) s'obtiennent en cumulant chaque fois au total précédent le chiffre figurant en colonne 4 sur la même ligne.

La médiane se situe en 6290, soit dans la classe 60, très proche de la limite 60/70 (6382).

La hauteur moyenne observée de la classe 60 (cf. tableau 4.1) est de 17,43 m, alors que la hauteur moyenne calculée a été trouvée égale à 17,64 m. Si, considérant que la médiane coïncide sensiblement avec la limite des classes 60 et 70, on avait adopté comme hauteur moyenne celle des classes regroupées (60 et 70) on aurait trouvé :

$$\frac{69,75 + 54,50}{7} = 17,75 \text{ m}$$

Valeur qui, ici, se rapproche davantage de la hauteur moyenne calculée (ou de Lorey).

Le choix d'une hauteur moyenne fondée sur la médiane de la distribution des surfaces terrières s'avère, sur cet exemple, assez encourageant. Comme ce choix est susceptible, dans la pratique, de réali-

ser un moyen terme entre le calcul de la hauteur de Lorey et la hauteur moyenne observée, il est apparu intéressant d'étendre la comparaison aux 43 placettes qui servent de support à cette étude.

Nous nous bornerons à l'énoncé des résultats de cette comparaison sans en reproduire les détails.

Somme des hauteurs calculées par la médiane	$\Sigma H_{gm}$	760,33 m
Somme des hauteurs de Lorey	$\Sigma H_L$	752,95 m

Ecart observé	7,38
Ecart moyen: $7,38 : 43 =$	0,17 m

On note tout de suite que l'écart moyen est inférieur à celui (0,28 m) qui a été trouvé lors de la comparaison concernant la hauteur moyenne « observée » (celle de la classe contenant l'arbre de surface terrière moyenne).

L'écart-type estimé de la différence ( $H_{gm} - H_L$ ) est:

$$s = 0,661$$

L'erreur-type est alors:

$$S: \sqrt{43} = 0,101$$

La valeur du t de Student de Fisher est donc:

$$t + \frac{0,17}{0,10} = 1,70$$

Valeur dont la probabilité est voisine de 10 %. Au seuil de probabilité de 5 %, la hauteur moyenne calculée à partir de l'arbre de surface terrière médiane n'est donc pas biaisée par rapport à la hauteur moyenne de Lorey. On note cependant que la dispersion est importante, l'écart-type d'une différence étant estimé à 66 cm. On avait estimé un écart-type de 42 cm seulement pour la différence entre la hauteur moyenne de la classe renfermant l'arbre de surface terrière moyenne et la hauteur moyenne de Lorey.

Pour la détermination rapide de la hauteur moyenne de Lorey (qui se confond avec ce que nous avons précédemment appelé la hauteur moyenne « calculée »), la hauteur moyenne de l'arbre de surface terrière médiane donne ainsi une estimation plus fidèle, mais plus dispersée que celle que l'on obtient en adoptant la hauteur moyenne de classe de l'arbre de surface terrière moyenne.

L'utilisation de la médiane simplifie un peu les calculs et limite les mesures de hauteur sur le terrain à celle d'une classe de cir-

conférence. C'est une méthode bâtarde dont le gain de précision ne nous apparaît pas déterminant si on la compare à celle qui fait appel à la hauteur moyenne de l'arbre de circonférence terrière moyenne. C'est pourquoi nous la jugeons d'un intérêt limité.

En définitive, le cubage d'un peuplement par le volume de l'arbre moyen se résout, après le calcul de la circonférence de l'arbre de surface terrière moyenne  $C_g$  sur les données de l'inventaire :

— le plus simplement : en déterminant sur le terrain la hauteur moyenne  $\bar{H}_g$  de la classe de circonférence qui renferme  $C_g$  et en appliquant le tarif correspondant à  $\bar{H}_g$  (arrondi), on obtient :

$$\bar{v} = v(C_g, \bar{H}_g) \quad \text{et} \quad V = N\bar{v}$$

Pour plus de précision, on peut apporter à  $V$  une correction de + 2 %.

— plus laborieusement mais plus exactement : en calculant la hauteur moyenne pondérée par les surfaces terrières  $H_L$  sur un échantillon stratifié du peuplement. On utilise alors  $H_L$  (arrondi) pour le choix du tarif et l'on calcule comme précédemment  $\bar{v}$  et  $V$ .

Sur le terrain, la méthode simplifiée suppose la mesure des hauteurs uniquement dans la classe de circonférence qui renferme l'arbre « moyen » de circonférence  $C_g$ . La méthode exacte oblige au contraire à mesurer des hauteurs dans toutes les classes de circonférence au prorata des effectifs de ces classes dans l'inventaire du peuplement.

#### 4.2 — Estimation par classes de circonférence - Stratification.

Comme nous l'avons indiqué plus haut, l'estimation du volume total du peuplement est dans ce cas obtenue en affectant un volume moyen aux arbres de chaque classe de circonférence représentée dans l'inventaire et en effectuant la sommation. Le volume moyen de chaque classe est lu sur le tarif correspondant à la hauteur moyenne arithmétique (arrondie) observée sur les arbres de l'échantillon appartenant à la classe considérée. Pour les classes extrêmes, il se peut qu'en raison de la faiblesse de l'effectif relevé à l'inventaire, il n'y ait pas d'arbres dans l'échantillon. Nous adoptons alors la hauteur retenue pour la classe contiguë.

Ce mode de cubage a été également appliqué aux 43 placettes étudiées. Les résultats se trouvent encore sur le tableau 1 de l'annexe 4. Si l'on désigne par  $V_{st}$  le volume estimé suivant cette méthode,  $V_1$  étant toujours l'estimation « Station de Recherches » correspondante, on a :

$$\begin{aligned} \Sigma V_{st} &= 3.789.658 \text{ dm}^3 \\ \Sigma V_1 &= 3.789.159 \text{ dm}^3 \end{aligned}$$

Différence +499 dm<sup>3</sup> (soit + 0,013 %)

Les deux modes d'estimation aboutissent à des totaux pratiquement identiques et cela est remarquable...

L'étude de la distribution des écarts montre que :

— les écarts absolus sont distribués autour d'une moyenne (11,6 dm<sup>3</sup>) qui ne diffère pas significativement de zéro. L'écart-type estimé est en effet 3885 dm<sup>3</sup> et l'erreur-type est alors de 592,5 dm<sup>3</sup>.

— les écarts relatifs, qui sont les plus intéressants, sont normalement distribués autour d'une moyenne (0,035 %) qui, elle non plus, n'est pas significativement différente de zéro. L'écart-type de la distribution de  $y$  est estimé à 4,6 % et l'erreur-type est ainsi de 0,701 %.

L'estimation du volume d'un peuplement par l'emploi de tarifs *correctement choisis pour chaque classe de hauteur* apparaît ainsi dépourvue d'erreur systématique si on la compare à l'estimation obtenue par la méthode de la régression en C<sup>2</sup>. Si on la rapproche de l'estimation obtenue par la hauteur moyenne calculée, on remarque que les erreurs relatives ont des distributions sensiblement comparables :

— volume par stratification : écart-type estimé de l'erreur relative : 4,6 %.

— volume par hauteur moyenne calculée : écart-type estimé de l'erreur relative : 4,8 %.

On notera que les deux méthodes utilisent l'ensemble des hauteurs mesurées sur l'échantillon. Il n'en est pas de même du cubage par la hauteur moyenne « observée », celle-ci étant calculée sur un nombre bien plus faible d'observations extraites de l'échantillon global.

#### 4.3 — Utilisation combinée des tarifs et de la régression en C<sup>2</sup>.

43.0 — Les méthodes qui viennent d'être exposées sont d'un emploi relativement facile pour le praticien puisqu'elles se ramènent après un inventaire du peuplement, à la mesure, au dendromètre, de la hauteur totale d'un certain nombre d'arbres convenablement choisis. Elles présentent cependant, une certaine dispersion par rapport à l'estimation obtenue par les méthodes plus raffinées, donc plus complexes, utilisées dans les travaux de recherche — nous évoquerons plus loin le problème de la précision des méthodes de cubage précédentes.

Pour améliorer l'estimation du volume sur pied — notamment au bénéfice de la recherche — sans pour autant consentir à mesurer effectivement sur le terrain les volumes individuels des arbres de

l'échantillon, on peut songer à utiliser les tarifs pour déterminer les volumes des arbres de l'échantillon servant à la construction d'une régression en  $C^2$  particulière à un peuplement donné.

1 — La méthode la plus rigoureuse consiste à affecter à chaque arbre de l'échantillon dont on a mesuré la circonférence (à 1,30 m)  $c_i$  et la hauteur totale  $h_i$ , un volume  $v_i$  lu sur le tarif idoine. On calcule alors, par la méthode des moindres carrés, la régression en  $C^2$  du volume qui constitue le tarif « personnalisé » du peuplement.

2 — Une approximation plus rapide est obtenue en admettant que la droite de régression en  $C^2$  du volume d'un peuplement passe, comme tous les tarifs que nous avons construits, par le point :

$$C = 0 \quad V = -7 \text{ dm}^3$$

Comme elle passe également (en vertu de la première équation normale) par le centre de gravité  $G$  du nuage des points  $(C_i^2, v_i)$  elle est alors définie sans qu'il soit nécessaire de procéder aux calculs exigés pour l'ajustement par la méthode des moindres carrés. Si cette approximation — car c'en est une — s'avérait valable, elle apporterait une simplification intéressante.

#### 43.1 — Régression en $C^2$ calculée sur les volumes « tarif » de l'échantillon.

On ne mesure plus sur le terrain le volume réel des arbres de l'échantillon, mais seulement leur circonférence à 1,30 m et leur hauteur totale. On se sert des tarifs pour leur affecter individuellement un volume et on calcule ensuite, suivant la méthode habituelle (ajustement par la méthode des moindres carrés) la droite de régression en  $C^2$  du volume d'un arbre du peuplement. On simplifie donc le travail sur le terrain ; au bureau, les calculs sont les mêmes.

Les résultats obtenus par cette méthode sont rassemblés sur le tableau 2 de l'annexe 4 (volume sur échantillon « tarif », colonne  $V_3$ ).

On obtient, pour l'ensemble des placettes, les résultats suivants :

$$\begin{aligned} \Sigma V_3 &= 3\,794\,718 \text{ dm}^3 \text{ (régression sur volumes « tarifs »)} \\ \Sigma V_1 &= 3\,789\,159 \text{ dm}^3 \text{ (régression sur volumes réels)} \end{aligned}$$

$$\text{Différence} + 5\,559 \text{ soit} + 0,15 \%$$

L'étude de la distribution des écarts observés entre les deux estimations montre que :

— les écarts absolus sont distribués autour d'une moyenne observée de  $129,3 \text{ dm}^3$ . L'écart-type est estimé à  $3\,814 \text{ dm}^3$  soit une erreur-type de  $581,7 \text{ dm}^3$  ; les deux estimations ne présentent entre elles aucun biais.

— les écarts relatifs (en %) sont distribués autour d'une moyenne observée de 0,26 avec un écart-type estimé de 4,70 soit une erreur-type de 0,716, nettement supérieure à la moyenne obtenue qui, ainsi, ne diffère pas significativement de zéro.

On peut donc dire que les deux modes d'estimation sont homogènes entre eux — ce qui confirme encore la validité des tarifs — mais que la dispersion des écarts demeure sensible, du moins pour les chercheurs. Il est du même ordre que celui qui a été observé dans certains modes d'estimation précédents — en particulier la stratification — plus commodes à mettre en œuvre.

#### 43.2 — Méthode rapide - Droite de régression simplifiée.

La droite représentative du volume d'un arbre du peuplement est encore :

$$V = a + b C^2$$

mais on admet, par hypothèse, qu'elle occupe l'axe 0 V au point I

$$C = 0 \quad V = -7 \text{ dm}^3$$

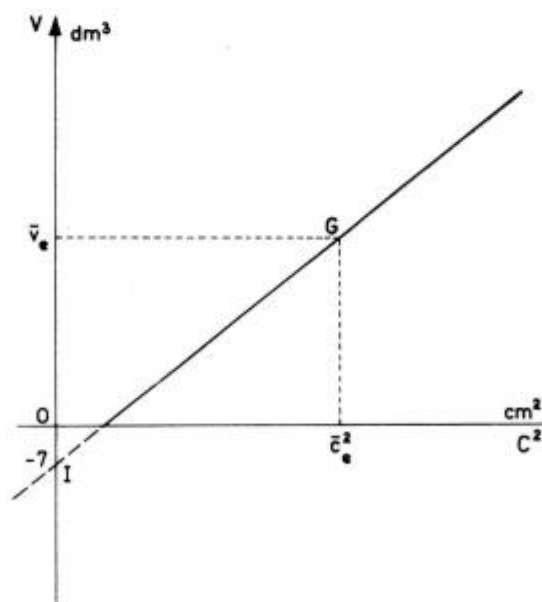


Fig. 4.3.

Elle passe aussi (première équation normale) par le centre de gravité G ( $\bar{C}^2, \bar{v}$ ) avec

$$\bar{C}^2 = \frac{1}{n} \sum_e c_e^2 \quad \bar{v} = \frac{1}{n} \sum_e v_e$$

on en déduit (fig. 4.3):

$$a = -7 \quad b = \frac{\bar{v} + 7}{\bar{C}^2} = \frac{\sum v + 7 n}{\sum C^2}$$

la sommation étant étendue aux arbres de l'échantillon.

Le volume du peuplement, d'effectif  $N_p, C_p$ , est alors:

$$V = -7 N + b \sum_p N_p C_p^2 \quad N = \sum_p N_p$$

On peut de la même manière opérer sur les surfaces terrières. Si  $g$  est la surface terrière de l'échantillon et  $G$  celle du peuplement d'effectif  $N$ :

$$a' = -7 \quad b' = \frac{\bar{v} + 7}{g} = \frac{v + 7 n}{g}$$

et

$$V = -7 N + b' G$$

On trouvera encore condensés sur le tableau 2 de l'annexe 4, les résultats obtenus par cette méthode (droite simplifiée, colonne  $V_2$ ).

On obtient pour les 43 placettes les résultats suivants:

$$\Sigma V_2 = 3\,803\,595 \text{ dm}^3 \text{ (droite simplifiée)}$$

$$\Sigma V_1 = 3\,789\,159 \text{ dm}^3 \text{ (régression sur volumes réels)}$$

$$\text{Différence} + 14\,436 \text{ dm}^3 \text{ soit } + 0,38 \%$$

L'étude de la distribution amène à constater:

— que les écarts absolus sont distribués autour d'une moyenne de  $335,7 \text{ dm}^3$  avec un écart-type estimé à  $3\,870 \text{ dm}^3$ , ce qui donne une erreur-type de  $593,2 \text{ dm}^3$ . La moyenne observée ne diffère pas significativement de zéro.

— que les écarts relatifs (en %) sont distribués normalement autour d'une moyenne de  $0,517$ ; l'erreur-type étant ici de  $0,708$ , la conclusion est identique.

Ce mode d'estimation est donc compatible avec l'estimation « Station de Recherches ».

43.3 — *Comparaison des deux méthodes précédentes.*

Comparées à la méthode rigoureuse qui fonde le calcul du tarif de cubage d'un peuplement sur la méthode des moindres carrés appliquée aux volumes vrais des arbres de l'échantillon, les deux méthodes donnent, en première analyse, des résultats comparables. Toutes deux fournissent une estimation qui est cohérente avec la première et des dispersions très voisines, les valeurs des écarts-types observés pour l'écart relatif étant respectivement 4,70 et 4,64 %.

La comparaison directe des volumes estimés le confirme. L'écart absolu  $V_3 - V_2$  de deux estimations d'un même volume est en moyenne de 206 dm<sup>3</sup>. L'écart relatif est distribué autour d'une moyenne observée de - 0,26 % avec un écart-type estimé de 0,98 %. Compte tenu de la faible valeur de l'erreur-type ( $0,967 \sqrt{43} = 0,149$ ), cette moyenne de - 0,26 % est cependant assez voisine du seuil de signification de 5 % car la valeur du *t* de Student (1,733) correspond à une probabilité d'environ 9 %. Ceci tient évidemment au fait que la méthode simplifiée constitue une approximation, comme un calcul algébrique rigoureux permet du reste de le démontrer.

Quoi qu'il en soit, les résultats obtenus par les deux méthodes diffèrent peu au regard des écarts que l'on observe par rapport aux estimations obtenues par la méthode rigoureuse. On peut donc conclure que la méthode simplifiée est à tout prendre bien plus simple et au demeurant d'une précision comparable à celle qui applique la méthode des moindres carrés aux volumes des arbres de l'échantillon déterminés à l'aide des tarifs. Cette dernière est une méthode bâtarde dont l'intérêt s'avère finalement d'autant plus limité qu'elle fournit une estimation qui n'est pas plus précise que le cubage par stratification tel qu'il a été exposé ci-dessus (§ 4.2) au prix de calculs beaucoup plus longs.

On notera enfin que la méthode simplifiée n'est applicable qu'au cas particulier du Pin Laricio de Corse, le terme constant de la régression

$$V = a(H) + b(H) \cdot C^2$$

étant indépendante de *H* seulement pour cette essence.

En définitive, aucune des méthodes étudiées au § 4.3 ne paraît applicable aux estimations nécessitées par les travaux de recherche sur la production car elles n'apportent pas un gain de précision qui justifie l'exécution de calculs plus ou moins compliqués.

5 — **Précision**

5.0 — Le choix d'une méthode d'estimation dépend toujours des objectifs que l'on se propose d'atteindre et des moyens que l'on peut

consentir à cet effet. Parmi les éléments d'appréciation qui dictent finalement la décision, la précision des méthodes occupera toujours une place importante. C'est pourquoi il est nécessaire de s'interroger sur le degré de précision des méthodes précédemment exposées pour utiliser les tarifs.

Le problème, comme nous allons le voir, est complexe et il ne sera pas possible dans le cadre de cet article d'apporter une réponse définitive à toutes les questions que soulève l'étude de la précision comparée des diverses méthodes.

La principale source de difficultés provient de l'utilisation simultanée d'un sondage et de l'application de tarifs qui, par leur définition même, traduisent des lois de variation « moyennes » du volume.

L'estimation du volume du peuplement par la mesure des caractères relevés sur un échantillon implique automatiquement une erreur de sondage. Cette erreur de sondage, que l'on peut estimer pour les techniques courantes (4) et (11), est liée aux caractéristiques de la population et à la définition du sondage. Elle est attachée au choix, dans le sein du peuplement, des arbres qui constituent l'échantillon, choix qui entraîne la désignation de l'une des multiples combinaisons qui, à priori, s'offraient au départ avec une égale probabilité.

Si l'on désigne par  $V$  le volume réel d'un peuplement  $P$  et par  $V_1$  le volume estimé à partir de la mesure effective des volumes sur les arbres de l'échantillon, on a, pour un type de sondage donné :

$$V_1 = V + e_1$$

$e_1$  est une grandeur aléatoire que l'on considère comme normalement répartie autour d'une moyenne (nulle si le sondage n'est pas biaisé) avec un écart-type  $\sigma_1$ . L'erreur-type du sondage est une estimation de  $\sigma_1$ .

Dans le cas de l'estimation « recherches » par la régression  $v = a + bc^2$ , l'estimation de  $\sigma_1$  est ce que nous avons appelé ET. Les fluctuations aléatoires de la variable  $e_1$  sont relatives à la population particulière  $P$  dont le volume est estimé par  $V_1$ .

Si l'on considère un autre peuplement  $Q$  on estimera un autre volume total et une autre erreur-type. Pour les sondages qui étaient à notre disposition (43 placettes) et qui se caractérisent, en gros, par un nombre d'arbres variant relativement peu (en moyenne 23 par échantillon), il existe une corrélation marquée entre le volume total et l'estimation de son erreur-type. Ainsi les volumes estimés par la méthode Abadie  $V_1$  et leur erreur-type ET (figure 5.1) sont en moyenne liés par la relation :

$$ET = 317 + 0,0197 V_1 \quad (ET, V_1 \text{ en dm}^3)$$

Si l'on veut entreprendre une étude valable de la dispersion, il importe au préalable de faire disparaître l'obstacle d'une variance liée à l'estimation elle-même.

Des considérations théoriques — voir KENDALL (7) vol. 2, p. 206 — montrent que la transformation « stabilisatrice » est

$$y_1 = \log (317 + 0,0197 V_1)$$

On observera toutefois que cette transformation est liée à la régression observée expérimentalement entre ET et  $V_1$ , dont le terme

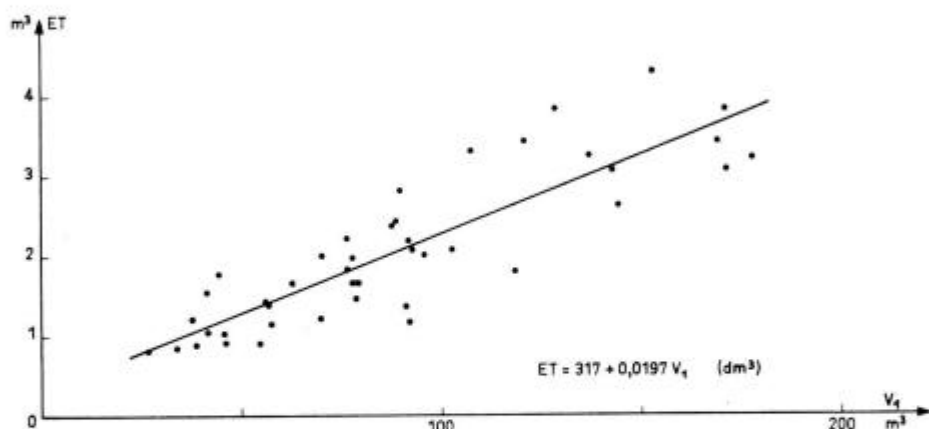


Fig. 5.1

Variation de l'erreur-type ET en fonction de  $V_1$ , estimation du volume sur pied par la méthode Abadie

$$V_1 = a N + b \sum_i N_i C_i^2$$

constant  $a$  a une valeur assez faible lorsque  $V_1$  est grand — pour la valeur moyenne de  $V_1$  dans les 43 placettes ( $V_1 = 88\,120 \text{ dm}^3$ ) on a :

$$\frac{317}{88120} = 0,0036 \approx 0,18 \times 0,0197$$

Dans ces conditions — et puisqu'il s'agit toujours d'approximation — il est beaucoup plus simple d'adopter la transformation  $y_1 = \log V_1$  qui présente l'avantage d'être indépendante des coefficients observés. Cette transformation revient du reste à étudier

les erreurs relatives au lieu des erreurs absolues (\*) et l'on peut remarquer que, dans ce qui précède, on a presque toujours pu ne pas rejeter l'hypothèse de la normalité de la répartition de ces erreurs.

L'intérêt majeur de la transformation, malgré son caractère approché, est de permettre la comparaison entre les placettes, les erreurs d'estimation appartenant alors à une même distribution normale de moyenne nulle (estimation non biaisée en vertu du mode de sondage adopté) et de variance constante en première approximation. Il n'y avait pas d'autre moyen d'exploiter les données dont nous disposions, celles-ci étant limitées à un seul échantillon par placette.

Si l'on fait appel à une autre méthode d'estimation — par exemple l'emploi de l'estimation par le volume moyen au lieu de l'estimation par la droite de régression  $v = a + b e^2$  — on obtient d'autres erreurs de sondage  $e_2$  ou  $e'_2$  définies par :

$$V_2 = V + e_2 \quad \text{ou} \quad Y_2 = \log V_2 ; \quad Y_2 = Y + e'_2$$

Même si les estimations ne sont pas biaisées, la distribution de  $e_2$  ou de  $e'_2$  sera caractérisée par des écarts-types  $\sigma_2$  et  $\sigma'_2$  puisque les précisions des estimations  $V_1$  et  $V_2$  (ou  $Y_1$  et  $Y_2$ ) sont a priori différentes. Par ailleurs, les variables  $e_1$  et  $e_2$  d'une part et  $e'_1$  et  $e'_2$  d'autre part, sont corrélées puisque les mesures ont été réalisées sur les mêmes individus constituant l'échantillon. Enfin, on peut espérer que la transformation  $Y = \log V$  stabilisera la variance de  $e'_2$ . (\*\*)

(\*) En effet :

$$e'_1 = \log V_1 - \log V = \log \frac{V_1}{V} = \log \left( 1 + \frac{e_1}{V} \right) \approx \frac{e_1}{V} \approx \frac{e_1}{V_1}$$

(\*\*) En fait, si l'on considère les volumes moyens réels  $\bar{v}_1$  des échantillons et l'estimation de leur écart-type  $s_{\bar{v}_1}$  par la formule classique — voir par exemple Cochran (4)

$$s_{\bar{v}_1} = \frac{N_1 - n_1}{N_1 \cdot n_1} s_1^2 \quad \text{avec} \quad \bar{v}_1 = \frac{1}{n_1} \sum_j v_{1j}$$

$$\text{et} \quad s_1^2 = \sum_j \frac{(v_{1j} - \bar{v}_1)^2}{n_1 - 1}$$

on obtient, par le calcul, la régression :

$$s_{\bar{v}_1} = 12,9 + 0,0858 \bar{v}_1 \quad (s_{\bar{v}_1} \text{ et } \bar{v}_1 \text{ en dm}^3)$$

dont le terme constant est faible en valeur relative. En première approximation  $s_{\bar{v}_1}$  est encore proportionnel à  $\bar{v}_1$  et la transformation logarithmique s'impose de même pour stabiliser la variance.

Supposons maintenant que, tout en gardant le même mode de sondage et la même formule d'estimation que celle qui a servi à déterminer  $V_2$  ou  $Y_2 = \log V_2$  nous utilisons dans les calculs, non plus les valeurs réelles du volume (mesurées sur le terrain), mais des valeurs estimées à l'aide des tarifs. On obtient des estimations  $V_3$  du volume (et  $Y_3 = \log V_3$ ). L'utilisation des tarifs entraîne une erreur supplémentaire que nous définirons par  $e_3$  ( $e'_3$  pour la variable  $Y_3$ ) tels que :

$$\begin{aligned} V_3 &= V_2 + e_3 & Y_3 &= \log V_3 = Y_2 + e'_3 \\ &= V + e_2 + e_3 & &= Y + e'_2 + e'_3 \end{aligned}$$

La valeur moyenne des variables aléatoires  $e_3$  ou  $e'_3$  traduit le biais éventuel de la méthode fondée sur l'emploi des tarifs. L'estimation de leur écart-type renseigne sur la précision de la méthode. Les critères de validité d'utilisation des tarifs sont donc fondés sur l'étude des différences (ou de celles des transformées) observées entre deux estimations utilisant le même mode de sondage et la même définition des grandeurs estimées à partir des données individuelles de l'échantillon. Pour une placette donnée, nous pourrions calculer sans difficulté  $V_2$ , à l'aide de volumes réels  $v_j$  aussi bien dans le cas de l'estimation par le volume moyen que dans l'application des sondages stratifiés. Nous aurons, suivant le cas :

$$V_2 = \frac{N}{n} \cdot \sum_j v_j \qquad V_2 = \sum_k \frac{N_k}{n_k} \sum_j v_{kj}$$

Pour chaque placette avec les éléments dont nous pourrions disposer, nous n'avons encore qu'une seule valeur de  $e_3$  pour étudier telle méthode d'application du tarif.

Nous admettrons, sans pouvoir actuellement le démontrer, que la transformation  $Y = \log V$  stabilise la variance de  $e'_3$  comme elle stabilise la variance ET de  $V_1$  et la variance  $s_v^2$  de  $\bar{v}$ .

Il y a là un manque de rigueur certain, mais nous n'avons pas d'autre possibilité avec le matériel dont nous disposons (\*) et l'utilisateur des tarifs exige bien moins une estimation précise de l'erreur que des indications *comparatives* sur la valeur approchée de celle qu'entraîne le choix des méthodes. Nous considérerons aussi que les résultats qui suivent sont beaucoup plus des éléments de comparaison que des chiffres absolus.

(\*) Des recherches ultérieures devront préciser ce point.

## 5.1 — Application des tarifs au cubage de l'arbre moyen.

Nous avons

$$V_2 = N \cdot \bar{v}_r = \frac{N}{n} \sum_j v_{j,r}$$

en désignant par  $v_{j,r}$  le volume réel de l'arbre  $j$  de l'échantillon,  $j = 1, \dots, n$ .

Le volume « tarif » est différent suivant que l'on adopte l'estimation  $V_0$  fondée sur l'emploi du tarif correspondant à la hauteur moyenne  $H_g$  (celle de la classe de l'arbre de circonférence  $c_g$ ) ou celle,  $V_r$  fondée sur l'emploi du tarif correspondant à la hauteur de Lorey, Ht.

51.1 — Tarif fondé sur la hauteur moyenne  $H_g$ .

Par définition :

$$e'_3 = \log V_0 - \log V_2 = \log \frac{V_0}{V_2}$$

En première approximation  $e'_3$  est distribuée autour d'une moyenne  $\mu_0$  avec un écart-type  $\sigma_0$ . L'estimation  $m_0$  de  $\mu_0$  apprécie le biais, la valeur observée de  $s_0$  de  $\sigma_0$  apprécie la précision.

Nous nous limiterons à énoncer le résultat des calculs réalisés sur les 43 placettes, en désignant par  $\bar{E}_0$  la valeur de  $e'_3$  dans la comparaison de ces méthodes.

Moyenne observée :  $\bar{E}_0 = - 0,009195$

Estimation de l'écart-type :  $s(E_0) = 0,03006$

erreur-type :  $s(\bar{E}_0) = 0,004585$

les logarithmes étant des logarithmes de base 10.

Le rapport :

$$t = \frac{\bar{E}_0}{s(\bar{E}_0)} = \frac{0,009195}{0,004585} = 2,006$$

correspond (pour 42 degrés de liberté) à une probabilité peu différente de 5 %. Sur ces résultats, on peut penser que l'utilisation des tarifs fondés sur la hauteur de l'arbre de surface terrière moyenne est quelque peu biaisée (par défaut) et qu'il serait opportun d'adopter une correction systématique. Nous en reparlerons.

La valeur de l'écart-type,  $s(E_0) = 0,03006$  correspond à

$$\log \frac{V_0}{V_2} = 0,03006 \quad \frac{V_0}{V_2} = 1,072$$

soit une erreur relative de 7,2 %. La dispersion constatée n'est donc pas négligeable. Elle reste du même ordre de grandeur que celle qui résulte de la comparaison directe des couples  $V_{10}$ ,  $V_1$ .

### 51.2 — Tarif fondé sur la hauteur de Lorey.

Soit  $E_c$  la valeur de  $e'_2$  pour cette comparaison.

$$E_c = \log V_c - \log V_2 = \log \frac{V_c}{V_2}$$

on obtient par le calcul les résultats suivants :

Moyenne observée :	$\bar{E}_c =$	—	0,00315
Estimation de l'écart-type :	$s(E_c) =$		0,02587
erreur-type :	$s(\bar{E}_c) =$		0,003944

soit pour le  $t$  de Student une valeur de

$$t = \frac{\bar{E}_c}{s(\bar{E}_c)} = \frac{0,00315}{0,003944} = 0,799 \text{ non significative}$$

La valeur de l'écart-type correspond cette fois à

$$\log \frac{V_c}{V_2} = 0,02587 \quad \frac{V_c}{V_2} = 1,061$$

soit une erreur relative de 6,1 %.

Cette erreur traduit une dispersion un peu plus faible que celle qui a été observée dans le cas précédent. Cela semble tout à fait normal. Lorsqu'on adopte la hauteur moyenne de l'arbre de surface terrière moyenne, on utilise une partie seulement de l'information qu'on pourrait recueillir sur la variation de la hauteur totale dans les différentes classes de circonférence du peuplement. La hauteur moyenne de Lorey intègre au contraire, dans une certaine mesure, ces variations et fournit une estimation dans l'ensemble plus précise du volume de l'arbre moyen.

## 5.2 — Utilisation d'un tarif pour chaque classe de circonférence.

Soit  $\bar{E}_{st}$  la valeur de  $e'_3$  dans cette comparaison.

Nous avons de même :

$$E_{st} = \log V_{st} - \log V'_2 = \log \frac{V_{st}}{V'_2}$$

$V'_2$  est cette fois l'estimation fondée sur les volumes réels en appliquant la formule classique :

$$V'_2 = \sum_i N_i \bar{v}_{i,r}$$

$\bar{v}_{i,r}$  étant le volume réel moyen des arbres de l'échantillon appartenant à la classe de circonférence  $C_i$  :

$$\bar{v}_{i,r} = \frac{1}{n_i} \sum_j v_{i,j} \quad j = 1, 2, \dots, n_i$$

Le calcul donne, pour les 43 placettes :

$$\text{Moyenne observée : } \bar{E}_{st} = 0,001686$$

$$\text{Estimation de l'écart-type : } s(E_{st}) = 0,01973$$

$$\text{erreur-type : } s(\bar{E}_{st}) = 0,003008$$

d'où une valeur du  $t$  de Student-Fisher :

$$t = \frac{0,001686}{0,003008} = 0,561 \quad \text{non significatif}$$

La valeur de l'écart-type correspond à

$$\text{colog } 0,01973 = 1,046$$

correspondant par conséquent à une variation de 4,6 %.

Ce chiffre confirme les résultats déjà observés lors de la comparaison directe de  $V_1$  et de  $V_{st}$ . Cette méthode minimise l'erreur inévitable que l'on introduit en substituant à un volume réel un volume « tarif ». Elle n'exige qu'un échantillon correctement choisi et quelques opérations arithmétiques à la portée de tous. Aucune correction de biais n'est nécessaire. L'erreur de sondage n'est évidemment pas supprimée, mais on sait que celle-ci est en principe plus faible lorsque l'estimation tient compte de la stratification comme c'est le cas ici. Rappelons encore la remarquable concordance observée entre la somme des estimations obtenues de la sorte et celle des estimations fournies par la méthode Abadie.

On peut du reste schématiser par une sorte de « coefficient de fidélité » les liens qui unissent à l'estimation « Station de Recher-

ches » celle que l'on obtient par les méthodes d'application des tarifs. Il suffit de calculer les coefficients de corrélation correspondants :

Volume moyen, hauteur $H_g$ de la classe $C_g$	$r_o = 0,9922$
Volume moyen, hauteur $H_t$ de Lorey	$r_c = 0,9948$
Volume stratifié	$r_{st} = 0,9957$

On voit que ces coefficients de corrélation sont très voisins de 1 mais que demeure le classement hiérarchique toujours observé jusqu'à présent, dans l'ordre de précision croissante : hauteur de l'arbre « moyen »  $C_g$ , hauteur de Lorey, hauteur individualisée à la classe de circonférence.

### 5.3 — Amélioration de l'estimation par l'emploi d'une régression.

Nous avons pu constater que l'utilisation des tarifs entraîne, et cela est normal, une dispersion supplémentaire de l'estimation qu'ils permettent d'obtenir ; on peut se demander si, par une correction appliquée à posteriori, on ne peut pas améliorer l'estimation obtenue en la rendant plus précise. Ceci s'applique naturellement aux travaux de recherches car, en pratique, l'utilisateur courant des tarifs pourra sans doute se contenter d'un écart-type de 5 à 6 % pour l'erreur relative par rapport à l'estimation obtenue par les méthodes des Stations de Recherches.

Il existe une possibilité relativement simple d'étalonner l'estimation « tarif »  $V$  sur l'estimation « Station de Recherches »  $V_1$ . Elle consiste à se servir des valeurs observées  $V$  et  $V_1$  pour calculer une régression linéaire.

$$V' = A + B V$$

$V'$  étant le volume « corrigé » de  $V$ ,  $A$  et  $B$  des constantes calculées par la méthode des moindres carrés appliquée aux couples  $(V, V_1)$  observés sur les 43 placettes.

53.1 — Une telle régression est surtout intéressante, à priori, pour le cas où l'estimation « tarif » est calculée à partir du volume de l'arbre moyen de circonférence  $C_g$  et de Hauteur  $H_g$ . Cette méthode est en effet la moins précise de celles que nous avons étudiées.

La régression trouvée a l'expression suivante :

$$(53.1) \quad V'_o = 1,00621 V_o + 1151,8 \quad V'_o, V_o \text{ en dm}^3$$

Quel est le résultat de cette correction? Il peut être analysé par la comparaison des nouveaux couples ( $V_1, V'_0$ ) et l'étude de la distribution des variables:

$$x_0 = V'_0 - V_1 \quad y_0 = 100 \cdot \frac{V'_0 - V_1}{V_1}$$

l'étude de l'erreur relative  $y_0$  étant du reste plus intéressante que celle de  $x_0$  puisqu'elle se rattache à la transformation logarithmique.

Tout d'abord — et en vertu de l'équation normale qui permet de calculer le terme constant de la régression, — on a:

$$\Sigma x_0 = \Sigma (V'_0 - V_1) = 0$$

Par contre, il n'y a pas compensation totale des erreurs relatives en %:

$$\Sigma y_0 = 18,37 \% \quad \bar{y}_0 = 1,193 \%$$

Pour ce qui est de la dispersion, les écarts-types estimés sont respectivement:

$$\begin{array}{ll} \text{écart-type de } x_0 & s_{x_0} = 5220 \text{ dm}^3 \\ \text{écart-type de } y_0 & s_{y_0} = 4,39 \% \quad s_{\bar{y}_0} = 0,67 \% \end{array}$$

On rapprochera ces résultats de ceux qui ont été obtenus par la comparaison directe des volumes  $V_0$  et  $V_1$  (§ 41.2)

$$\begin{array}{l} \text{écart-type de } (V_0 - V_1) = 5226 \text{ dm}^3 \\ \text{écart-type de } 100 \cdot \frac{V_0 - V_1}{V_1} = 6,03 \% \end{array}$$

L'ajustement est sans effet sur la dispersion de l'erreur absolue mais il diminue notablement celle de l'erreur relative. L'efficacité de l'ajustement peut s'exprimer par

$$R = \frac{(6,03)^2}{(4,39)^2} = 1,89$$

Ainsi la précision se trouve améliorée de 89 % par l'utilisation de la régression. Il n'y a pas lieu de s'étonner de la faible influence de l'ajustement sur la dispersion des écarts absolus. En effet, le coefficient de régression est très voisin de 1 (on peut même montrer qu'il n'en diffère pas significativement) et de ce fait, la transformation qui conduit de  $V_0$  à  $V'_0$  s'apparente étroitement à un changement d'origine, opération qui ne modifie pas la variance. En défi-

nitive, on améliorera davantage l'estimation  $V_{st}$  en appliquant une correction de biais *constante* plutôt qu'une correction proportionnelle au volume mais la correction la plus efficace — et cette efficacité est loin d'être négligeable — est celle qui est obtenue par l'équation 53.1.

53.2 — Dans l'estimation « tarif » réalisée en affectant une hauteur à chaque classe de circonférence, la précision de la méthode n'est guère améliorée par une correction fondée sur l'emploi d'une régression linéaire.

Si  $V_{st}$  est le volume estimé par cette méthode, en posant encore

$$x_{st} = V_{st} - V_1 \quad y_{st} = 100 \cdot \frac{V_{st} - V_1}{V_1}$$

le calcul conduit à l'équation

$$(53.2) \quad V'_{st} = 0,99511 \cdot V_{st} + 419,9 \quad V_{st} \text{ (} V'_{st} \text{ en dm}^3 \text{)}$$

comme dans le cas précédent :

$$\Sigma x_{st} = 0$$

L'ajustement des erreurs relatives est meilleur

$$\Sigma y_{st} = 6,21 \quad \bar{y}_{st} = 0,14 \%$$

Quant à la dispersion, elle est caractérisée par les estimations suivantes :

$$\begin{array}{lll} \text{écart-type de } x_{st} & s_{x_{st}} = 3880 & \\ \text{écart-type de } y_{st} & s_{y_{st}} = 4,58 & s_{\bar{y}_{st}} = 0,698 \end{array}$$

Pour les estimations non corrigées  $V_{st}$  on avait observé les résultats suivants (§ 4.2) :

$$\begin{array}{ll} \text{écart-type de } (V_{st} - V_1) & : 3885 \text{ dm}^3 \\ \text{écart-type de } 100 \cdot \frac{V_{st} - V_1}{V_1} & : 4,6 \% \end{array}$$

On voit que l'ajustement par la régression n'apporte dans ce cas aucune amélioration. Faut-il s'en étonner? Nous avons déjà vu que l'estimation  $V_{st}$  fournissait, *pour l'ensemble des placettes*, une estimation du volume pratiquement identique à celle que l'on avait obtenue par la méthode « Station de Recherches ». D'autre part, le coefficient de la régression est ici encore très peu différent de 1 (il n'en diffère pas significativement) et le terme constant lui-même est faible en valeur relative. La régression modifie très faiblement les estimations  $V_{st}$  par deux opérations qui se compensent

plus ou moins : réduction légère (coefficient 0,995) suivie d'une augmentation constante (+ 420 dm). Il semble bien qu'on ait atteint, avant l'ajustement, le plancher de la précision des méthodes fondées sur l'utilisation des tarifs pour l'effectif « échantillon » considéré.

On retiendra de ce qui précède qu'une méthode relativement simple, économiquement intéressante, mais moins précise qu'une méthode plus complexe, peut être améliorée à peu de frais par un étalonnage des estimations obtenu par comparaison avec des données précises recueillies en nombre beaucoup plus faible. La régression trouvée au § 53.1 peut en effet servir pour améliorer les estimations du volume des placettes qui seront cubées ultérieurement en ne mesurant que la hauteur moyenne  $\bar{H}_g$  et l'on est en droit de penser que la précision sera sensiblement la même que celle que l'on obtiendrait par un échantillonnage complet permettant la détermination des hauteurs moyennes de classe.

## 6 — Conclusion

Diverses méthodes, toutes statistiquement valables, s'offrent pour la construction des tarifs de cubage. L'avènement du calcul automatique rend aujourd'hui ces méthodes applicables dans la pratique. Tel est le cas notamment des tarifs fondés sur l'utilisation de la régression multiple mais il n'est peut-être pas toujours indispensable de recourir aux méthodes les plus compliquées.

En présence des données qui devront servir à construire des tarifs à double entrée, il paraît indiqué dans une première analyse de tester la linéarité du volume par rapport à la surface terrière (ou carré de la circonférence). Si le résultat de ces tests est favorable — comme ce fut le cas pour le Pin Laricio de Corse en Sologne — une méthode valable et économiquement intéressante consiste dans le double ajustement des coefficients des droites de régression :

$$v(H, C) = a(H) + b(H) \cdot C^2$$

Dans cette hypothèse, on peut utiliser les programmes de calcul déjà existants, c'est-à-dire traiter très rapidement les données mêmes si l'ajustement des constantes se fait à l'aide de machines de bureau. Cette rapidité de traitement constitue actuellement pour nous, un des principaux avantages de cette méthode en même temps que l'économie qu'elle entraîne quant au prix de revient des tarifs.

La famille des tarifs ainsi construits doit être soumise à un contrôle permettant notamment de déceler l'existence d'un biais. Ce contrôle peut être également accéléré par le calcul automatique. Nous avons exposé comment il s'exerçait avant de conférer aux tarifs un label de garantie préalable à leur publication.

Ces tarifs, une fois testés, peuvent être utilisés de plusieurs manières. Nous en retiendrons deux :

1 — *Calcul du volume de l'arbre moyen du peuplement*: en utilisant le tarif adapté à la hauteur moyenne  $H_g$  de la classe qui renferme l'arbre de surface moyenne, dont la circonférence  $c_g$  détermine le volume  $\bar{v}_g$ .

Le volume est alors  $V_o = N \bar{v}_g$ .

On obtient par cette formule une estimation dont l'écart-type, pour l'erreur relative en % (\*), est de l'ordre de 6 %.

On peut en améliorer la précision par un « étalonnage » fondé sur la comparaison des estimations  $V_o$  obtenues par la formule précédente et des estimations  $V_1$  fournies par les méthodes habituelles de la Station de Sylviculture et de Production pour un certain nombre de peuplements. Dans le cas du Pin Laricio de Corse en Sologne, l'étalonnage conduit à l'estimation corrigée  $V'_o$  par la formule :

$$V'_o = 1,00621 V_o + 1152$$

( $V'_o$  et  $V_o$  en  $\text{dm}^3$ )

L'écart-type de l'erreur relative en % est alors ramené à 4,4 % soit un gain de précision d'environ 90 %.

2 — *Calcul du volume de l'arbre moyen pour chaque classe de circonférence* :

Après l'inventaire, stratifié en classe de circonférence, on définit un échantillon à fraction sondée constante. Dans chaque classe de circonférence  $C_1$  d'effectif  $N_1$  relevé à l'inventaire, on choisit au hasard  $n_1$  arbres dont on mesure la hauteur totale  $H_{1,j}$  ( $j = 1 \dots n_1$ ). L'échantillon est correctement stratifié si :

$$\frac{n_1}{N_1} = \dots = \frac{n_i}{N_i} = \dots = \frac{n}{N}$$

La fraction de sondage  $\frac{n}{N}$  est définie par le nombre d'arbres  $n$  que l'on consent à mesurer. Dans les échantillons de la Station de Sylviculture et de Production qui ont servi à cette étude  $n$  est gé-

(\*) Nous employons ici le vocable erreur dans un sens très particulier et à défaut d'autre terme valable. Il s'agit en fait de la différence de deux estimations : celle obtenue à l'aide des tarifs et celle que l'on obtient par la méthode « Station de Recherches » qui est l'estimation la plus précise d'un volume réel inconnu.

néralement compris entre 20 et 25. Ce fait doit être souligné, car *c'est dans cette hypothèse que les chiffres cités relatifs à la précision demeurent valables*. Il est certain d'autre part, que la précision d'un sondage est liée à la fraction d'échantillonnage et, toutes choses égales d'ailleurs, croît avec celle-ci.

A partir des mesures  $H_{1j}$  relevées dans la classe  $C_1$ , on calcule la hauteur moyenne arithmétique de classe

$$\bar{H}_1 = \frac{1}{n_1} \sum_j H_{1,j}$$

et l'on utilise le tarif correspondant à cette hauteur moyenne pour lier le volume moyen  $\bar{v}_1$  de l'arbre de la classe  $C_1$ . Le volume total  $V$  s'obtient par la formule:

$$V = \sum N_1 \bar{v}_1$$

Une telle estimation, pour le Pin Laricio de Corse, est dépourvue de toute erreur systématique. L'erreur relative (en %) a un écart-type de 4,6 %. Elle n'est pas améliorée par une correction analogue à celle qui a été conseillée pour l'estimation par le volume de l'arbre moyen.

En définitive, si ces tarifs ne permettent pas, avec des effectifs identiques, d'obtenir une estimation aussi précise que celle qu'on obtient par la méthode habituelle de la Station de Sylviculture et de Production (erreur-type 2 %), ils sont susceptibles de satisfaire des utilisateurs — et ils sont la majorité — moins exigeants que les chercheurs.

Pour ces derniers, les tarifs ne sont pas non plus dépourvus d'intérêt, car leur emploi allège considérablement les mesures sur le terrain en supprimant le cubage sur pied, donc le grimpage et le transport du matériel encombrant que cette opération nécessite. On peut alors multiplier les placettes temporaires et, avec des équipes plus mobiles, rassembler un nombre bien plus important de données pour l'établissement de tables de production. En consentant un sacrifice sur la précision de l'estimation des volumes individuels des placettes — sacrifice d'autant plus léger qu'un nombre accru d'observations améliorera la compensation des erreurs — on recueillera une masse plus importante de données sur les autres caractéristiques des peuplements qui doivent figurer dans les tables de production: nombre de tiges, circonférence moyenne, etc... Nous ne sommes pas loin de penser qu'au total, le bilan sera positif.

Dans cette optique, la collecte des données nécessaires à l'établissement des tables de production se ferait en deux étapes :

1 — Cubage précis, par mesure directe des volumes sur un nombre de placettes suffisant pour rassembler le matériel nécessaire à la construction des tarifs.

2 — Utilisation des tarifs avec des équipes légères se bornant à des inventaires et des mesures de hauteur, et par conséquent, susceptibles de prospecter, dans le même temps, un nombre accru de peuplements.

---

## BIBLIOGRAPHIE

- (1) ARBONNIER (P.). — (1962): Le forestier à l'heure de l'électronique. R.F.F., n° 3.
  - (2) ASSMANN (E.). — (1961): Waldertragskunde - B.L.V. Verlagsgesellschaft Munich.
  - (3) AYRAL (P.) et ABADIE (J.). — (1956): Méthode de calcul du volume des peuplements sur pied dans les places d'essais de Sylviculture. Annales E.N.E.F. Tome XV, fasc. I.
  - (4) COCHRAN (W.-G.). — (1963): Sampling Techniques - 2<sup>e</sup> éd. John Wiley and Sons. New-York.
  - (5) HUFFEL. — (1919): Economie Forestière, tome II, 2<sup>e</sup> édition. La Maison Rustique. Paris.
  - (6) HUMMEL (F.-C.). — (1955): The Volume-Basal Area Line. Forestry Commission Bulletin n° 24. Her Majesty's Stationery Office, Londres.
  - (7) KENDALL (M.-G.). — (1951): The advanced Theory of Statistics (2 vol). Charles GRIFFIN, Londres.
  - (8) PARDÉ (J.). — (1961): Dendrométrie. Editions de l'Ecole Nationale des Eaux et Forêts. Nancy.
  - (9) TOMASSONE (R.). — (1963): Méthode graphique de calcul de la moyenne et de l'écart-type d'une distribution normale. Test de normalité. R.F.F., n° 10.
  - (10) VESSEREAU (A.). — (1960): Méthodes statistiques en biologie et en Agronomie. 2<sup>e</sup> Ed. J.B. Baillièrre et fils. Paris.
  - (11) YATES (F.). — (1951): Méthodes de sondage pour recensements et enquêtes, traduit sous la direction de G. DARMOIS. Dunod et Masson. Paris.
  - (12) CENTRE DE FORMATION AUX APPLICATIONS INDUSTRIELLES DE LA STATISTIQUE. — (1961): Formation aux méthodes statistiques (2 fasc.). Institut de Statistique de l'Université de Paris.
-

## RÉSUMÉ

---

Pour la préparation des tables de production entreprises par la Station de Sylviculture et de Production du C.N.R.F., des tarifs de cubage ont été calculés en 1963 par la Station de Biométrie pour le Pin Laricio de Corse en Sologne, grâce au calcul automatique. Les méthodes employées sont décrites ainsi que les contrôles permettant d'assurer que les tarifs ne sont pas biaisés.

La formule de cubage est la suivante :

$$V = 10^{-5} (445,5 H - 400,6) C^2 - 7$$

V, volume bois fort - tige sur écorce (découpe à  $C = 22$  cm) en  $\text{dm}^3$ ; C, circonférence à 1,30 m; H, hauteur totale en mètres  $9 \leq H \leq 25$  m.

Le contrôle s'est exercé à la fois sur 32 échantillons ayant fourni les données du calcul et sur 21 échantillons cubés ultérieurement. Pour l'erreur relative (en %) on observe un écart-type de 4,3 %, l'effectif moyen des échantillons étant de 23 tiges.

On étudie ensuite les procédés d'application des tarifs au cubage des peuplements par mesure des hauteurs sur un échantillon stratifié par classes de circonférence, avec fraction sondée constante :

1 — par le volume de l'arbre moyen  $\bar{v}$  et  $V = N \bar{v}$

2 — par la détermination du volume moyen  $\bar{v}_i$  des arbres de la classe  $C_i$  avec :

$$V = \sum_i N_i \bar{v}_i$$

3 — par la construction d'une droite de régression donnant une estimation

$$V = a \cdot N + b \sum_i N_i C_i^2$$

Les coefficients a et b sont calculés soit par la méthode des moindres carrés appliqués aux couples  $C_i, v_i (h_i)$ , soit en admettant

l'hypothèse que la droite passe par les points  $C = 0$   $V = -7 \text{ dm}^3$  et  $G (\bar{c}^2, \bar{v})$ .

Cette étude conduit aux constatations suivantes :

1 — Le volume de l'arbre moyen est celui qui correspond à la circonférence  $C_g$  telle que :

$$C_g^2 = \frac{1}{N} \sum_i N_i C_i^2$$

mais il reste à choisir le tarif approprié.

1. a — Pour le Pin Laricio de Corse en Sologne, on montre que l'estimation correcte de  $\bar{v}$ , donc de  $V$  est obtenue en utilisant le tarif correspondant à la hauteur moyenne de Lorey,  $H_L$  définie par :

$$H_L = \frac{\sum_i N_i H_i C_i^2}{\sum_i N_i C_i^2} = \frac{\sum_i H_i G_i}{G}$$

$G = \sum N_i G_i$ , surface terrière du peuplement.

1. b — L'application à la détermination du volume de l'arbre moyen  $\bar{v}$  du tarif correspondant à la Hauteur moyenne  $H_g$  de l'arbre de circonférence  $C_g$  fournit une estimation légèrement biaisée de  $\bar{v}$ . Il est proposé d'améliorer cette estimation et de corriger le biais par l'emploi d'une régression donnant une estimation corrigée  $V'$  :

$$V' = 1,00621 V + 1151,8$$

qui a été calculée à partir des estimations  $V$  et de celles,  $V_1$  obtenues par la méthode Abadie (régression en  $C^2$ ). Cette correction ramène de 6 % à 4,4 % l'écart-type de l'erreur relative par rapport à l'estimation Abadie.

2 — En utilisant un tarif par classe de circonférence, fondé sur la hauteur moyenne arithmétique des arbres de l'échantillon appartenant à cette classe, on obtient une estimation  $V$  du volume non biaisée par rapport à l'estimation Abadie avec un écart-type estimé de 4,6 % pour l'erreur relative.

3 — La détermination d'une droite de régression à l'aide des volumes « tarif » de chaque arbre de l'échantillon ( $c_i, h_i$ ) donne également une estimation non biaisée du volume mais au prix de calculs qui n'apportent pas, en contre-partie, une amélioration substantielle de la précision.

On conclut que les tarifs construits sont valables et l'on recommande les modes d'estimation 1. b et 2 ci-dessus. La simplicité de l'estimation 1. b convient particulièrement aux utilisateurs courants qui n'exigent pas une estimation d'une grande précision et ne consentent qu'à un petit nombre de mesures sur le terrain.

Pour les travaux de recherches, en vue de la confection de tables de production, on peut procéder en deux étapes, la première permettant de recueillir les données nécessaires à la confection et au contrôle des tarifs, la seconde utilisant ceux-ci pour simplifier et accélérer le travail sur le terrain.

## SUMMARY

---

For the preparation of yield tables undertaken by the Silviculture and Yield Experiment Station of the C.N.R.F., some volume tables have been worked out through electronic computing by the Biometry Station for Corsican Pine in Sologne in 1963. Methods used as well as further tests allowing to ensure tariffs are not biased are described.

The relationship between volume and b.h. girth is as follows:

$$V = 10^{-5} (445,5 H - 400,6) C^2 - 7$$

where  $V$  represents stem volume over bark (to a 22 cm girth limit) in  $\text{dm}^3$ ,  $C$  b.h. girth, and  $H$  total height in metres  $9 \leq H \leq 25$  m. Checking has been carried out on the 32 samples which had supplied computing data as well as on 21 samples which have been measured later. As to the relative error (in %), a standard error of 4,3 % has been recorded, from an average number of samples of twenty three stems.

Then have been studied the possible methods of application of volume tables to stand volume estimates by measuring heights in a sample stratified in girth classes, with uniform sampling fraction:

1. through volume of mean tree  $\bar{v}$  and  $V = N \bar{v}$
2. through the determination of mean volume  $\bar{v}_i$  of trees in class  $C_i$  with

$$V = \sum_i N_i \bar{v}_i$$

3. by the plotting of a regression line giving the following estimate:

$$V = a \cdot N + b \sum_i N_i C_i^2$$

Coefficients  $a$  and  $b$  have been calculated either by means of least squares method applied to couples  $C_i, v_i(h_i)$ , or by assuming that regression line crosses points  $C = 0, V = -7 \text{ dm}^3$ , and  $G (c^2, \bar{v})$ .

This study enabled to make the following statements:

1 — Mean tree volume is that corresponding to girth  $C_g$  determined by:

$$C_g^2 = \frac{1}{N} \sum_1 N_1 C_1^2$$

but the suitable tariff has yet to be selected.

1. a. For Corsican Pine in Sologne, it has been shown that a correct estimate of  $\bar{v}$  can be obtained through application of a tariff based on Lorey mean height,  $H_L$ , as defined by

$$H_L = \frac{\sum_1 N_1 H_1 C_1^2}{\sum_1 N_1 C_1^2} = \frac{\sum_1 H_1 G_1}{G}$$

where stand basal area  $G = \sum N_1 G_1$ .

1. b. Application of a tariff based on mean height  $H_g$  of the tree the girth of which is  $C_g$ , to determine the volume of mean tree  $\bar{v}$ , gives a slightly biased estimate for  $\bar{v}$ . It has been suggested to improve this estimate and correct the deviation by using a regression which gives a corrected estimate  $V'$ :

$$V' = 1,00621 \quad V + 1151,8$$

which has been calculated on  $V$  estimates and other estimates  $V_1$ , computed according to Abadie's method (regression in  $C^2$ ). This correction brings down the standard deviation of the relative error in comparison with Abadie's estimate to 4,4 % instead of 6 %.

2. By using a girth class tariff, based on arithmetic mean height of trees of the sample taken from this class, an estimate  $V$  of the volume, unbiased as compared with Abadie's estimate can be obtained, with an estimated standard deviation of 4.6 % corresponding to the relative error.

3. Determination of a regression line with the help of tariff volume estimates for each sample tree ( $c_1, h_1$ ) also gives an unbiased estimate for the volume but at the cost of such computations as do not bring a significant increase in accuracy in compensation.

As a conclusion, it may be stated that the tariffs obtained are suitable and that the above-mentioned methods of estimation 1 b and 2 can be recommended. Owing to its simplicity, estimate 1 b is parti-

cularly convenient for common users who do not insist on an estimate with high accuracy and admit to make but a small number of measures in the field.

The investigations aiming at the construction of yields tables, can be dealt with in two stages, the first stage being the collection of the data necessary to the building and control of tariffs, the other one making use of these tariffs for easier and more rapid field work.

---

## ZUSAMMENFASSUNG

---

Die Station für Biometrie der forstlichen Versuchsanstalt (C.N.R.F.) in Frankreich hat im Jahre 1963 Massentafeln für stehende Bäume mit Hilfe von elektronischen Rechenmaschinen aufgestellt. Diese Tafeln wurden vorerst für die Korsische Schwarzkiefer (*Pinus Nigra Arn. Poiretiana Asch. et Graebn.* (1804) in der Sologne berechnet und sind eine Vorarbeit für Ertragstafeln die von der Station für Waldbau und Ertragskunde vorbereitet werden. Es werden die angewendeten Arbeits- und Berechnungsmethoden beschrieben, desgleichen auch Kontrollmethoden die eine Verzerrung durch systematische Fehler (bias) der Tarife mit Sicherheit ausschliessen.

Die Kubierungsformel lautet:

$$V = 10^{-5} (445,5 H - 400,6) C^2 - 7$$

V, Derbholz, Schaft mit Rinde (Zopfstärke 7 cm mit Rinde = Umfang C = 22 cm) in dm<sup>3</sup>; C, Umfang in Brusthöhe (1,30 m) in cm; H, Gesamthöhe in m. (9 m ≤ H ≤ 25 m).

Die Kontrolle der Formel wurde sowohl an 32 Stichproben die zugleich die Berechnungsgrundlagen bildeten als auch an weiteren 21 Stichproben die nachträglich kubierte wurden durchgeführt. Bei einer durchschnittlichen Stammzahl von 23 Bäumen pro Stichprobe beträgt der relative mittlere Fehler 4,3 %.

Im weiteren werden die Anwendungsmöglichkeiten der Tarife für die Massenermittlung in Beständen untersucht, wobei die Baumhöhen einer nach Stärkeklassen (Umfang in Brusthöhe) stratifizierten Stichprobe bei gleichbleibender Stichprobenintensität Eingang finden.

1 — nach dem Massenmittelstamm  $\bar{v}$ ;  $V = N \bar{v}$

2 — nach der mittleren Masse  $\bar{v}_i$  der Stämme einer Stärkeklasse  $C_i$ ;

$$V = \sum_i N_i \bar{v}_i$$

3 — durch die Konstruktion einer Regressionsgeraden die eine Schätzung erlaubt:

$$V = a \cdot N + b \sum_1 N_1 C_1^2$$

Die Koeffizienten  $a$  und  $b$  werden nach der Methode der kleinsten Quadrate für  $C_1$  und  $v_1$  ( $h_1$ ) bestimmt oder nach Annahme der Hypothese, dass die Gerade durch die Punkte  $C = 0$   $V = -7 \text{ dm}^3$  und  $G$  ( $c^{-2}$ ,  $\bar{v}$ ) geht.

Diese Untersuchung führt zu folgenden Feststellungen:

1 — der Massenmittelstamm entspricht dem Kreisflächenmittelstamm,

$$C_v^2 = \frac{1}{N} \sum_1 N_1 C_1^2$$

es muss jedoch der geeignete Tarif gewählt werden.

1. a — Für die korsische Schwarzkiefer in der Sologne konnte geseigt werden, dass bei einer Tarifwahl nach den Lorey'schen Mittelhöhen (H<sub>L</sub>) der Massenmittelstamm ( $\bar{v}$ ) und damit die Bestandesmasse ( $V$ ) mit genügender Genauigkeit bestimmt werden können.

$$H_L = \frac{\sum_1 N_1 H_1 C_1^2}{\sum_1 N_1 C_1^2} = \frac{\sum_1 H_1 G_1}{G}$$

$G = \sum N_1 G_1$ , Bestandeskreisfläche.

1. b — Bei einer Tarifwahl nach der Höhe des Kreisflächenmittelstammes ergibt sich ein geringer systematischer Fehler des Massenmittelstammes ( $\bar{v}$ ). Dieser systematische Fehler kann jedoch mit Hilfe einer Regression korrigiert werden.

$$V' = 1,00621 \quad V + 1151,8$$

Diese Korrektur wurde nach den Werten  $V$  und jenen von  $V_1$  (= nach der Methode von Abadie, Regression nach  $C^2$ ) bestimmt.

Im Vergleich zum  $V_1$ -Wert nach Abadie wird der relative mittlere Fehler von 6 % auf 4,4 % vermindert.

2 — Verwendet man einen Tarif für jede Stärkenklasse unter Berücksichtigung der arithmetischen Mittelhöhe der Stichprobe dieser Stärkenklasse so erhält man eine Bestandesmasse die im Vergleich zur Methode von Abadie keinen systematischen Fehler aufweist. Der relative mittlere Fehler beträgt 4,6 %.

3 — Die Bestimmung der Regressionsgeraden nach den « Tarifmassen » jedes Stichprobenbaumes ( $c_1$ ,  $h_1$ ) zeigt gleichfalls ein von

systematischen Fehlern freies Ergebnis. Die etwas langwierige Berechnung zeitigt jedoch keine nennenswerte Erhöhung der Genauigkeit.

Man gelangt zur Schlussfolgerung, dass die berechneten Tarife brauchbar sind und empfiehlt vor allem die oben angeführten Methoden 1. b und 2. Die Einfachheit der Methode 1. b entspricht besonders den Anwendungsmöglichkeiten in der Praxis, wenn an die Genauigkeit keine allzu hohen Anforderungen gestellt werden und die Messungen im Gelände nur in einem beschränkten Ausmass durchgeführt werden können.

Sollen die Tarife jedoch in der Forschung Anwendung finden, insbesondere für die Erstellung von Ertragstafeln, so sollte in zwei Etappen vorgegangen werden: Eine erste Etappe dient der Erhebung des Ausgangsmaterials für die Erstellung und Kontrolle der Tarife, während die zweite Etappe durch Anwendung der Tarife die Geländearbeiten vereinfacht und beschleunigt.

---

## **ANNEXES**



Annexe I — Tableau I

## CONTROLE DU BIAIS

1<sup>er</sup> groupe — Placettes ayant fourni les tiges pour le calcul des tarifs.

	Nombre de tiges	Volume (dm <sup>3</sup> )		x = erreur absolue		y = erreur relative(%)	
		réel	tarif	+	-	+	-
S 1	26	9,812	9,298		514		5,24
S 7	25	14,151	14,292	141		1,00	
S 8	26	8,962	8,978	16		0,18	
S 9	28	9,009	8,777		232		2,58
S 12	25	24,130	24,860	730		3,03	
S 27	32	6,078	5,910		168		2,76
S 30	24	31,080	31,360	280		0,90	
S 31	25	11,481	11,218		263		2,29
S 37	26	10,684	9,965		719		6,73
S 41	20	6,164	6,194	30		0,48	
S 42	19	12,728	13,429	701		5,51	
S 47	24	3,986	4,067	81		2,03	
S 52	18	14,957	14,769		188		
S 58	20	9,997	10,787	790		7,90	
S 60	22	8,224	8,531	107		1,30	
S 61	19	8,364	9,013	649		7,76	
S 64	24	3,461	3,263		198		5,72
S 67	21	5,232	5,463	231		4,41	
S 69	18	1,401	1,343		58		4,14
S 70	20	4,918	4,912		6		0,12
S 71	20	1,881	1,878		3		0,16
S 72	17	930	885		45		4,84
S 73	22	4,562	4,772	210		4,60	
S 91	24	24,415	23,762		653		2,67
S 92	24	11,262	10,841		421		3,74
S 93	25	12,319	11,750		569		4,64
S 104	22	5,711	5,733	22		0,38	
S 105	24	9,126	9,582	456		4,76	
S 106	20	2,521	2,657	136		5,12	
S 114	23	7,009	7,125	116		1,63	
S 115	24	8,923	8,964	41		0,46	
S 116	24	4,511	4,865	354		7,28	
N = 32	731	297,989	299,043	5,091	4,037	58,73	47,09
		Moyennes		$\bar{x} = 0,054 \pm 32 = 32,93 \text{ dm}^3$		$\bar{y} = 11,64 \pm 32 = 0,364 \%$	

## Calcul des variances

$$\Sigma x^2 = 4.591.646$$

$$C = 34.716$$

$$Q_x = 4.556.930$$

$$s_x^2 = 146.998 \quad s_x = 383,4$$

$$s_x^2 = 4.593,68 \quad s_x = 67,8$$

$$t = \frac{32,93}{67,8} = 0,486$$

$$\Sigma y^2 = 527,2656$$

$$C = 8,4681$$

$$Q_y = 518,7975$$

$$s_y^2 = 16,7354 \quad s_y = 4,09$$

$$s_y^2 = 0,5230 \quad s_y = 0,723$$

$$t = \frac{0,364}{0,723} = 0,503$$

Annexe 1 — Tableau 2

## CONTROLE DU BIAIS

2<sup>e</sup> groupe — Placettes cubées postérieurement à l'établissement des tarifs.

Placette	Nombre de tiges	Volume (dm <sup>3</sup> )		x = erreur absolue		y = erreur relative(%)	
		réel	tarif	+	-	+	-
S 107	23	8.133	8.421	288		3,42	
S 108	24	4.439	4.504	65		1,44	
S 109	24	4.945	5.015	70		1,40	
S 110	23	9.359	8.974		385		4,29
S 118	23	10.967	11.029	62		0,56	
S 123	20	2.277	2.100		177		8,43
S 126 (*)	24	23.680	25.771	2.091		8,11	
S 129	26	9.010	9.159	149		1,63	
S 132	24	8.034	8.273	239		2,89	
S 144	23	4.089	4.655	566		12,16	
S 152	24	5.543	5.088		455		8,94
S 153	22	5.056	5.090	34		0,67	
S 173	24	8.907	9.028	121		1,36	
S 174	23	10.116	10.343	227		2,24	
S 180	25	11.539	11.747	208		1,80	
S 181	22	16.518	16.122		396		2,40
S 182	22	6.145	6.227	82		1,33	
S 187	23	3.723	3.568		155		4,16
S 193	23	4.186	4.161		25		0,60
S 194	22	5.805	5.767		38		0,65
S 197	24	2.684	2.588		96		3,58
<b>N = 21</b>	488	165.155	167.630	4.202	1.727	39,01	33,05
			<i>Moyennes</i>	$\bar{x} = 2,475$	$z = 117,86$	$\bar{y} = 5,96$	$z = 0,284$

## Calcul des variances

$$\Sigma x^2 = 5.563.891$$

$$C = 2911.696$$

$$Q = 5.172.195$$

$$s_x^2 = 258.609,75 \quad s_x = 508,54$$

$$s_x^2 = 12.314,75 \quad s_x = 110,97$$

$$t = \frac{117,86}{110,97} = 1,062$$

$$\Sigma y^2 = 459.0724$$

$$C = 1.6915$$

$$Q = 457.3809$$

$$s_y^2 = 22.8690 \quad s_y = 4,78$$

$$s_y^2 = 1,0890 \quad s_y = 1,043$$

$$t = \frac{0,284}{1,043} = 0,272$$

(\*) Voir § 31.2 du texte.

## Annexe 2

**Test de Wilcoxon**  
**Comparaison des volumes réels et des volumes tarifs**  
**sur le 2<sup>e</sup> groupe de placettes**

Le test de Wilcoxon (\*) permet de tester l'hypothèse que deux échantillons sont tirés de la même population. Il est intéressant à plusieurs titres car :

- il n'implique aucune hypothèse sur la normalité de la distribution;
- il permet la comparaison de deux échantillons d'effectifs différents  $m$  et  $n$ ;
- tenant compte de l'importance des différences entre les valeurs observées, il est plus puissant que d'autres tests tels que le test des signes.

Si l'on désigne par :

$x_1, x_2, \dots, x_1, \dots, x_m$ , les valeurs constituant le premier échantillon et :  
 $y_1, y_2, \dots, y_1, \dots, y_n$ , les valeurs du second échantillon, on calcule une constante  $W$  qui représente le double du nombre de paires d'observations  $(x_i, y_j)$  pour lesquelles  $x_i > y_j$  augmenté du nombre de paires pour lesquelles  $x_k = y_l$ .

La quantité  $W$  obéit à une loi de distribution qui a été étudiée et tabulée pour l'hypothèse nulle. La loi est discontinue et en général dissymétrique et, si les échantillons ne comportent pas d'observations de même valeur, la moyenne  $\mu$  de la distribution est égale à  $m n$  et sa variance  $\sigma^2$  à :

$$\sigma^2 = \frac{1}{3} m n (m + n + 1)$$

Pour les valeurs faibles de  $m$  et  $n$  ( $m \leq n \leq 11$ ) il existe des abaques donnant les probabilités de dépassement unilatéral.

Pour  $m + n < 40$ ,  $n > 10$  et  $m \leq n$  les valeurs critiques  $W_\alpha$  correspondant à la zone critique unilatérale (gauche) ont été tabulées pour les seuils de confiance  $\alpha = 0,05$ ,  $\alpha = 0,025$  et  $\alpha = 0,01$ . La zone critique unilatérale droite s'en déduit facilement.

Si  $m$  et  $n$  sont grands et peu différents et si, en outre, les nombres de valeurs égales dans les classes groupant des observations égales ne sont pas trop différents on peut admettre que  $W$  suit la loi normale de moyenne  $m n$  et de variance  $\sigma^2$  (notons que la valeur de  $\sigma^2$  est différente de l'expression indiquée plus haut lorsque les classes renferment plusieurs observations égales). On peut alors appliquer le test de Student.

(\*) On trouvera des indications plus complètes sur ce test dans un ouvrage cité en bibliographie (12), fasc. I, pp. 189-195 et 203-206.

## Annexe 2 — Tableau 1

*Comparaison volume réel, volume tarif sur 2<sup>e</sup> tranche échantillon*Test de Wilcoxon

(1)	Volume réel (2)	Volume tarif (3)	Contribution à W (4)	(1)	Volume réel (2)	Volume tarif (3)	Contribution à W (4)
0		2 100	0	Report			124
1	2 277			24			
2		2 508	2	25	8 113		
3	2 684			26		8 273	26
4		3 568	4	26		8 421	26
5	3 723			27	8 907		
6				28		8 974	28
7	4 089			29	9 010		
8		4 161	8	30		9 028	30
9	4 186			30		9 159	30
10				31	9 359		
11	4 439			32			
12		4 504	12	33	10 116		
12		4 655	12	34		10 343	34
13	4 945			35	10 967		
14		5 015	14	36		11 029	36
15	5 056			37	11 539		
16		5 088	16	38		11 747	38
16		5 090	16	38		16 122	38
17	5 543			39	16 518		
18		5 767	18	40			
19	5 805			41	23 680		
20				42		25 771	42
21	6 145						
22		6 227	22	TOTAUX	m = 21	n = 21	W = 452
23	8 034						

$$N = m + n = 42 \quad W = 452 \quad m n = 441 \quad W - m n = 11$$

$$\sigma^2 = \frac{1}{3} \cdot m n (N + 1) = \frac{441.43}{3} = 6321 \quad \sigma = 79,5$$

$$\frac{W - m n}{\sigma} = \frac{11}{79,5} = 0,1384 \quad P = 89 \%$$

Probabilité unilatérale 44,5 %

La valeur de  $W$  peut être facilement calculée en procédant comme il est indiqué sur le tableau 2.1

— dans la colonne (2) on range dans l'ordre croissant les valeurs  $x_i$  — ici les volumes réels — en les séparant par des blancs.

— dans la colonne (3) on reporte, toujours dans l'ordre croissant, les valeurs  $y_j$  (volume tarif) — dans l'intervalle qui sépare deux valeurs  $x_i$  et  $x_{i+1}$ , si  $x_i < y_j \leq x_i + 1$ ; sur la ligne  $x_i$ , si  $x_i = y_j$ ;

— dans la colonne (1) un chiffre égal au double du nombre de valeurs figurant dans la colonne (2) *au-dessus* de la ligne, augmenté du nombre de valeurs figurant *sur* la ligne.

Exemple (voir tableau): sur la ligne où 4186 figure en colonne (2) on trouve en (1)  $9 = 4 \times 2 + 1$ . Il est du reste beaucoup plus facile de remplir la colonne (1) au vu d'un exemple que de l'expliquer...

— dans la colonne (4), sur chaque ligne où figure une valeur de  $y_j$  on reporte le produit du nombre lu dans la colonne (1) par le nombre de valeurs figurant sur la ligne. La somme des nombres de la colonne (4) donne  $W$ .

Suivant les valeurs de  $m$  et  $n$  on consulte les abaques ou les tables donnant directement la valeur limite (unilatérale) de  $W$ . Si  $m$  et  $n$  sont grands et dépassent les valeurs tabulées on calcule:

$$\mu = m n \quad \text{et} \quad \sigma^2 = \frac{1}{3} m n (m + n + 1)$$

et l'on utilise l'approximation normale par un test de Student.

#### PUISSANCE DU TEST.

Lorsque  $m$  et  $n$  sont petits  $m + n \leq 6$  la puissance du test de Wilcoxon est à peine inférieure à celle du test  $t$  de Student. Si  $m$  et  $n$  sont très grands il a été également montré que l'efficacité du test tend vers  $3/\pi$  (95,5 %) dans les environs de  $H_0$ .

## Annexe 3

**Calcul de la hauteur moyenne  
pour le choix d'un tarif de cubage unique**

1. — Montrons d'abord que la hauteur moyenne arithmétique fournit une estimation biaisée du volume, si l'on utilise le tarif correspondant.

Soit un peuplement dont l'inventaire a été stratifié par classes de circonférence  $C_j$ . Soit encore  $N_j$  l'effectif de la classe  $C_j$ ,  $\bar{H}_j$  la hauteur moyenne et  $V_j$  le volume de l'arbre moyen de la classe.

On a, par définition :

$$(1) \quad V = \sum_j N_j V_j = N \bar{v} \quad N = \sum N_j$$

$$(2) \quad V_j = a + b C_j^2$$

et, du fait de l'existence des classes de circonférence :

$$(3) \quad b = p + q \bar{H}_j.$$

d'où :

$$(4) \quad V = \sum_j N_j V_j = a N + p \sum_j N_j C_j^2 + q \sum_j N_j \bar{H}_j C_j^2 = N \bar{v}$$

Le volume de l'arbre de surface terrière moyenne et de hauteur moyenne (arithmétique) est :

$$\bar{v}' = a + (p + q \bar{H}) C_g^2 \quad C_g^2 = \frac{1}{N} \sum_j C_j^2$$

$$\bar{H} = \frac{1}{N} \sum_j N_j H_j$$

de (4) on tire :

$$(5) \quad \bar{v} = \frac{V}{N} = a + p C_g^2 + q \frac{1}{N} \sum_j N_j \bar{H}_j C_j^2$$

$$\bar{v} - \bar{v}' = q \frac{1}{N} [\sum_j N_j \bar{H}_j C_j^2 - N \bar{H} C_g^2]$$

$$= q \frac{1}{N} \sum_j [N_j (H_j - \bar{H}) (C_j^2 - C_g^2)] = K \text{ cov. } (H_j, C_j^2)$$

Annexe 3 — Tableau 1

*Utilisation du tarif déterminé  
par la hauteur moyenne arithmétique (de l'échantillon)*

	Estimation du volume "échantillon"				Estimation du volume "population"			
	$v_e(\bar{H})$	$v_{e,r}$	$v_e(\bar{H}) - v_{e,r}$	%	$V_p(\bar{H})$	$V_1$	$V_p(\bar{H}) - V_1$	%
S 1	8.320	9.812	- 1.492	- 15,21	69.090	80.257	- 11.167	- 13,9
S 7	14.050	14.151	- 101	- 0,71	141.624	144.313	- 2.689	- 1,9
S 8	8.502	8.962	- 460	- 5,13	75.864	79.170	- 3.306	- 4,2
S 9	8.036	9.009	- 973	- 10,80	79.786	87.648	- 7.862	- 9,0
S 12	24.850	24.130	720	2,98	174.944	171.027	- 3.917	- 2,3
S 27	5.312	6.078	- 766	- 12,60	64.408	69.688	- 5.280	- 7,6
S 30	29.688	31.080	- 1.392	- 4,48	175.654	177.413	- 1.759	- 1,0
S 31	10.075	11.481	- 1.406	- 12,25	80.197	92.598	- 12.401	- 13,4
S 37	8.866	10.684	- 1.818	- 17,02	105.710	128.628	- 22.918	- 17,8
S 41	5.580	6.164	- 584	- 9,47	67.239	76.236	- 8.997	- 11,8
S 42	13.053	12.728	325	2,55	106.485	107.440	- 955	- 0,9
S 47	4.008	3.986	22	0,55	38.744	38.429	315	0,8
S 52	14.328	14.957	- 629	- 4,21	145.668	152.791	- 7.123	- 4,7
S 58	10.260	9.997	263	2,63	142.614	136.961	5.653	4,1
S 60	7.832	8.224	- 392	- 4,77	73.336	76.269	- 2.933	- 3,8
S 61	8.189	8.364	- 175	- 2,09	81.890	79.704	2.186	2,7
S 64	2.976	3.461	- 485	- 14,01	47.864	55.914	- 8.050	- 14,4
S 67	4.452	5.232	- 780	- 14,91	40.068	41.086	- 1.018	- 2,5
S 69	1.386	1.401	- 15	- 1,07	24.948	26.599	- 1.651	- 6,2
S 70	4.480	4.918	- 438	- 8,91	85.344	95.595	- 10.251	- 10,7
S 71	1.720	1.881	- 161	- 8,56	31.390	33.842	- 2.452	- 7,2
S 72	799	930	- 131	- 14,09	39.057	44.209	- 5.152	- 11,7
S 73	4.180	4.562	- 382	- 8,37	62.320	67.864	- 5.544	- 8,2
Total	200.942	212.192	-11.250	-159,95	1.954.244	2.063.681	-109.437	-141,0
Moyenne			- 5,30 %	- 6,95			- 5,30 %	- 6,13%

*Echantillon*  $v_{e,r}$  volume réel

$v_e(\bar{H})$  volume tarif déterminé par  $\bar{H}$

*Population*  $V_1$  volume estimé « Abadie »

$V_p(\bar{H})$  Volume tarif déterminé par  $\bar{H}$

la covariance est positive et non nulle — la hauteur et la circonférence d'un arbre étant évidemment corrélées — et  $K$  étant positif

$$\bar{v} - \bar{v}' > 0$$

La hauteur moyenne arithmétique fournit donc une estimation biaisée par défaut du volume.

## 2. — Détermination de la hauteur moyenne « tarif » H<sub>L</sub>.

Par définition de H<sub>L</sub> on a :

$$\begin{aligned} \bar{v} &= a + (p + q \bar{H}_L) C_g^2 \\ &= a + p C_g^2 + q \frac{1}{N} \sum_i N_i \bar{H}_i C_i^2 \quad \text{d'après (5)} \end{aligned}$$

d'où l'on tire :

$$H_L C_g^2 = \frac{1}{N} \sum_i N_i H_i C_i^2$$

$$(6) \quad H_L = \frac{1}{N C_g^2} \sum_i N_i H_i C_i^2 = \frac{1}{\sum_i N_i C_i^2} \sum_i N_i C_i^2 \bar{H}_i$$

chaque valeur  $H_i$  doit donc être pondérée par la quantité  $N_i C_i^2$ .

Si la fraction sondée est constante, on a :

$$\frac{n_i}{N_i} = \frac{n}{N} = C^r$$

$$(7) \quad N_i = n_i \cdot \frac{N}{n}$$

De plus la hauteur moyenne de classe est estimée sur l'échantillon par :

$$(8) \quad \bar{H}_i \approx \bar{h}_i = \frac{1}{n_i} \sum_j h_{ij}$$

la formule (6) devient en utilisant (7) et (8)

$$H_L = \frac{1}{N} \cdot \sum_i \frac{N}{n} n_i C_i^2 \frac{1}{n_i} \sum_j h_{ij}$$

$$H_L = \frac{1}{\sum_i n_i C_i^2} \cdot \sum_i C_i^2 S_i(h_j) \quad \text{avec} \quad S_i(h_j) = \sum_j h_{ij}$$

Cette formule permet, à partir d'un échantillon à fraction sondée constante — cas des échantillons de la Station de Sylviculture et de Production —

de calculer une hauteur moyenne valable pour le choix d'un tarif unique. On peut alors déterminer le volume de l'arbre moyen  $\bar{v}$  pour l'application de la formule :

$$V = N \bar{v}$$

Naturellement, dans la pratique  $H_t$  peut être arrondi à la classe la plus proche.