



HAL
open science

Méthode de calcul du volume des peuplements sur pied dans les places d'essais de sylviculture

Pierre Ayrat, Jacques Abadie

► **To cite this version:**

Pierre Ayrat, Jacques Abadie. Méthode de calcul du volume des peuplements sur pied dans les places d'essais de sylviculture. *Annales de l'Ecole Nationale des Eaux et Forêts et de la Station de Recherches et Expériences Forestières*, 1956, 15 (1), pp.1-136. hal-03483730

HAL Id: hal-03483730

<https://hal.science/hal-03483730>

Submitted on 16 Dec 2021

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

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

I. — DÉFINITION ET PRÉSENTATION

par P. AYRAL

Conservateur des Eaux et Forêts
Chef de la 1^{re} Section de la Station de Recherches
et Expériences forestières

1^o CONSIDÉRATIONS GÉNÉRALES

Les études de production faites par la 1^{re} Section de la Station de Recherches et Expériences forestières de Nancy, dans ses places d'essais permanentes, tendent à établir quelle est, pour chacune des essences forestières, spontanées ou introduites, étudiées, et dans chaque station considérée, la méthode de traitement permettant d'obtenir soit le rendement optimum dans le temps minimum, soit le maximum de produits nécessaires à l'économie générale du Pays.

Ces études reposent essentiellement sur des « comparaisons d'inventaires » périodiques qui fournissent les « données » dont l'« interprétation » doit amener aux conclusions. L'exactitude de ces « données » est donc primordiale ; elle est d'autant plus nécessaire que les « données » doivent être réunies par plusieurs chercheurs successifs, car l'expérimentation forestière, poursuivie au rythme de la croissance des peuplements forestiers, est fatalement longue, et l'exploitation des résultats bien loin d'être immédiate. Au surplus, les différences qui peuvent être décelées dans les productions, en *quantité* et en *qualité*, résultant de traitements variés — toutes les autres conditions étant, autant que possible, identiques —, sont parfois faibles, tout au moins au début de l'expérimentation, et, pour ne pas les masquer, il est nécessaire d'employer des méthodes suffisamment précises et permettant le calcul des erreurs-types.

En raison du nombre des facteurs de variation, il nous a semblé indispensable d'utiliser des méthodes relevant du calcul statistique. Les peuplements de nos placettes d'essais ne sont pas autre chose que des « populations » dont l'étude ne peut se faire que par un « échantillonnage » correct. Le cubage direct des produits abattus est toujours possible, mais le cubage des peuplements restant sur pied — indispensable pour étudier leur évolution — ne peut se faire que d'une façon plus ou moins approchée, par l'intermédiaire d'un « tarif de cubage » calculé à partir d'un « échantillon valable » de ces peuplements. Les études de volume total, avec décomposition en « catégories de produits », peuvent alors se faire, avec une probabilité connue, sur cet échantillon.

2° MÉTHODE D'ÉCHANTILLONNAGE

Dans un rapport détaillé (3) que nous avons présenté à la Section 23 de l'Union Internationale des Instituts de Recherches forestières, lors du XI^e Congrès de cette Union — Rome, 1953 —, nous avons exposé la méthode « d'échantillonnage stratifié » que nous avons mise au point, et les vérifications expérimentales complètes auxquelles nous avons procédé dans des placettes temporaires installées dans des peuplements qui devaient être abattus à blanc-étoc dans la forêt domaniale de Saint-Avoid (parcelles N^o 134 b, 142 et 156), la forêt communale de Porcellette, toutes deux situées en Moselle, la forêt indivise d'Haguenau dans le Bas-Rhin et la forêt domaniale de Saint-Amond (parcelles 46 et 47) en Meurthe-et-Moselle. Nous avons publié tous nos résultats ainsi que les comparaisons faites, aussi bien en volume total qu'en pourcentages de catégories de produits, entre les volumes calculés et les volumes réels exactement connus par cubage direct des peuplements après abatage.

Ces comparaisons ont démontré les avantages de notre nouvelle méthode sur celle qui était employée auparavant, et qui comportait seulement le choix d' « arbres types » parmi les arbres abattus en éclaircie dans la placette et dans les zones d'isolement ; aussi n'avons-nous pas hésité à en généraliser l'emploi dans tous nos dispositifs d'essais.

Pour tous les détails, les lecteurs voudront bien se reporter à notre rapport publié dans les Comptes rendus du Congrès. Nous nous bornerons à rappeler ici, brièvement, les principes de la méthode.

« L'échantillonnage » est fait à partir d'un « inventaire détaillé », avant éclaircie, de tous les arbres de la placette d'essai. Ces arbres sont numérotés individuellement et l'on note, pour chacun d'eux, non seulement sa circonférence à 1,30 m du sol — mesurée toujours au même endroit sur une ceinture ou un trait à la peinture — mais aussi quelques caractères simples, aussi peu nombreux que possible, mais permettant d'en effectuer un « classement relatif ». Ces caractères doivent être appelés par un seul opérateur par placette pour que ce classement soit valable. L'emploi de chiffres et de lettres rend les appels et les pointages faciles et rapides. Nous avons retenu seulement trois caractères : dominance, cime et tronc, et les notations sont ainsi faites :

— Place dans les étages de végétation (*Dominance*) :

- 1 — Dominant ;
- 2 — Retardataire ;
- 3 — Dominé.

— Forme de la *cime*:

- A — Ample et complète;
- B — Normale;
- C — Etriquée.

— Qualité du *tronc*:

- D — Droit; et en outre: F — fourchu;
- T — Tordu; ta — taré

Les éclaircies sont ensuite martelées. Les numéros des arbres à exploiter sont naturellement pointés, et l'on note au surplus les « caractères » des arbres abandonnés dans la zone d'isolement, pour le cas où ces arbres seraient nécessaires à l'échantillonnage. Les « classes d'arbres » peuvent être alors constituées. Il est préférable de ne pas multiplier le nombre de ces classes; la trop grande dispersion des tiges de chaque essence entre les classes et les circonférences pourrait rendre impossible la confection de l'échantillon. Il est bien préférable de grouper les circonférences et les caractères. Les circonférences — mesurées au centimètre près — sont groupées par 10 centimètres et les caractères retenus pour la confection des classes d'arbres sont tout d'abord les caractères de dominance et de cime (1A, 1B, 1C, 2A, etc...). Les caractères de tronc n'interviennent qu'ensuite, à l'intérieur de chaque classe, pour le choix des tiges de l'échantillon.

Les inventaires étant faits dans l'ordre des numéros des arbres, il est nécessaire de procéder à plusieurs pointages successifs, en nombre puis par numéros individuels et par classe, d'abord par centimètre, puis par groupe de circonférence, pour établir le Tableau récapitulatif, par groupes de 10 centimètres de circonférences et par classes d'arbres, des tiges du peuplement avant éclaircie, des tiges de l'éclaircie, et des tiges du peuplement après l'éclaircie. L'échantillon représentatif du peuplement restant sur pied après éclaircie doit respecter, dans chaque groupe de circonférences, la proportionnalité des « classes » représentées à l'intérieur de ce groupe. Il faut donc choisir un rapport d'échantillonnage et l'appliquer successivement à chaque classe dans tous les groupes de circonférences, certains groupes pouvant être réunis si les tiges y sont insuffisamment représentées. Le rapport d'échantillonnage doit être une fraction très simple ($1/3$, $1/4$,...), et choisi de telle sorte que le nombre des tiges de l'échantillon dépasse 30, limite pratique des grands nombres en statistique.

Le Tableau récapitulatif des tiges permet de faire très rapidement le choix des numéros retenus pour l'échantillon par classe dans chaque groupe de circonférences. Ces numéros sont tirés « au hasard »,

et nous nous servons pour cela de la table de « Random Numbers » de FISHER and YATES. Pour des raisons de commodité de cubage, on prend d'abord tous les arbres de l'éclaircie qui conviennent, complétés, le cas échéant, par des arbres des classes et grosseurs déficitaires abattus dans la zone. Les catégories manquantes sont recherchées parmi les arbres qui doivent rester sur pied. Ce complément s'avère rigoureusement indispensable dans tous les peuplements en croissance, car les coupes d'amélioration respectent les plus beaux arbres et ne comprennent, en général, que peu de tiges dans les plus grosses circonférences. Les nombreux échantillonnages actuellement réalisés par cette méthode nous permettent de dire que le nombre des arbres sur pied nécessaires est, en moyenne, compris entre la moitié et le tiers du total des tiges de l'échantillon et que ces arbres sur pied représentent, en général, les plus grosses tiges de cet échantillon. La distorsion systématique que présentent des échantillons constitués uniquement par des arbres de l'éclaircie est alors manifeste, ainsi que l'imprécision, surtout dans les plus gros volumes, des tarifs de cubage construits à l'aide de tels échantillons qui, en l'absence de représentants de toutes les catégories de circonférences tirés au hasard, ne permettent pas, au surplus, le calcul de l'erreur-type.

Les cubages d'arbres sur pied rendus possibles, comme nous l'avons exposé dans notre rapport, par l'emploi des échelles en duralumin « Ecole Forestière », complétées par les « coulisses » de notre invention, que nous avons décrites et qui permettent de légers déplacements à droite ou à gauche des éléments d'échelle pour les arbres qui manquent de rectitude, sont susceptibles de combler toutes les catégories déficitaires et d'étendre l'échantillon à toutes les classes et sur tous les groupes de circonférences, ce qui est *essentiel* pour l'exactitude du tarif de cubage et le calcul de l'erreur-type (*).

Notre équipe spécialisée de cubeurs-grimpeurs se rend alors sur le terrain et procède au cubage détaillé, par billons d'un mètre et avec classement des diverses catégories de produits, de tous les arbres abattus dans la placette, des arbres abattus dans la zone choisis pour l'échantillon, et des arbres sur pied également choisis pour cet échantillon. Le volume exact de la coupe nous est donné par la sommation des volumes individuels des arbres abattus dans la placette, et seuls les numéros retenus parmi ces arbres pour l'échantillon sont ensuite repris pour l'établissement du tarif de cubage.

(*) Pour le cubage sur pied de certains gros arbres feuillus dont l'escalade est difficile, il peut être employé une méthode optique. Voir le rapport présenté à la Section 25 de l'Union Internationale des Instituts de Recherches forestières pour le Congrès d'Oxford 1956, par P. ARBONNIER, Ingénieur à la 1^{re} Section: « Mesure optique du volume bois fort tige des feuillus sur pied ».

3° LES TARIFS DE CUBAGE

Les données de l'échantillon — circonférence à 1,30 m et volume mesuré de chaque tige d'essai — nous permettent d'établir le « tarif de cubage » qui servira à calculer le volume du peuplement restant sur pied. Tous nos tarifs étaient construits suivant la méthode graphique indiquée par HUFFEL (9). Cette méthode est relativement rapide et donne une bonne approximation. Elle nécessite, cependant, de la part du dessinateur, le « sens de la continuité de la courbe », et pour l'établissement du tarif numérique, des compensations aux volumes lus sur la courbe, pour obtenir une progression régulière. Au début, on employait, pour construire la courbe, du papier millimétré ordinaire. La courbe prenait alors, en général, la forme d'une branche de parabole qui rendait difficile et assez imprécise la lecture des volumes. Pour remédier à cet inconvénient, à partir de 1938, on a adopté le papier à double échelle logarithmique, en admettant l'hypothèse que, dans cette transformation, la courbe $v = f(c)$ serait une droite de pente moyenne, facilitant ainsi sa construction et la lecture des volumes. Mais, en réalité, cette hypothèse ne s'est pas trouvée confirmée; il s'agit non pas d'une droite, mais d'une courbe à très grand rayon dont la concavité est tournée vers le bas. Nous l'avons tracée, en général, à l'aide d'un ou plusieurs arcs de cercle à grand rayon et raccordés tangentiellement entre eux, ce qui n'est qu'un palliatif insuffisant. En outre, le système des coordonnées logarithmiques entraîne d'importantes variations en valeur absolue dans les gros volumes.

Il est bien évident que, dans le tracé graphique d'une courbe dont la forme exacte n'est pas connue, le coefficient personnel du dessinateur intervient d'une façon nullement négligeable, et que la précision de nos tarifs de cubage ainsi établis s'en trouve affectée. Le souci d'exactitude que nous avons exposé, nous a amené à rechercher une méthode de calcul, non subjective, suffisamment précise tout en restant assez rapide et simple pour être appliquée à l'aide des moyens dont nous disposons. Cette méthode devait permettre aussi de se faire une idée de l'erreur probable par le calcul de l'erreur-type à l'aide d'une formule statistique valable.

Nos expériences complètes, rappelées ci-dessus, nous fournissaient des données certaines sur lesquelles il était possible d'engager des études théoriques et des discussions de méthodes. Nous voulions aussi tester statistiquement les hypothèses de notre méthode d'échantillonnage. Toutes ces nouvelles recherches nécessitaient des connaissances mathématiques et statistiques particulièrement élevées. J. ABADIE, à la fois agrégé de mathématiques et statisticien, a bien voulu nous prêter son concours éclairé. En partant de nos données, il s'est livré à de longs et savants calculs qui ont tous été exécutés, sous sa direction, par nos calculateurs.

La publication, in extenso, de ses intéressants travaux nous a semblé indispensable; on la trouvera ci-après.

Cette publication est, en effet, susceptible d'intéresser tous les chercheurs qui se penchent sur des problèmes similaires. Par ses discussions basées sur des exemples concrets pris dans nos expériences complètes, par l'élaboration de certaines formules approchées des erreurs expérimentalement vérifiées, c'est une œuvre originale qui permet d'appuyer les conclusions pratiques sur des bases solides. C'est aussi une œuvre de synthèse qui rassemble des méthodes pratiques de calcul actuellement éparses dans divers ouvrages scientifiques.

Nous en résumerons simplement ici les conclusions essentielles:

Dans chaque placette d'essai, pour une essence et une circonférence données, la distribution des volumes suit une loi normale de LAPLACE-GAUSS au sein d'une même classe (1A par exemple).

A *circonférence égale*, le caractère de dominance (1, 2 ou 3) donne des différences de volumes significatives, c'est-à-dire non imputables seulement au hasard, tandis qu'il n'en est pas de même pour le caractère de cime (A, B ou C). Les études n'ayant porté que sur le volume total bois-fort-tige, ce résultat n'a rien de surprenant, mais nous continuons à noter, lors de nos inventaires, le caractère de cime, pensant que celui-ci nous sera utile, ultérieurement pour nos études de rapidité de croissance et de qualité.

Pour ce qui est des méthodes d'évaluation de volume du peuplement restant sur pied, seul un « échantillon stratifié » comprenant des représentants, tirés au hasard, de toutes les « classes d'arbres » dans tous les groupes de circonférences, proportionnellement aux nombres correspondants de tiges dans la « population » restant sur pied, permet l'établissement d'un tarif de cubage correct et le calcul de l'erreur-type. Les tests ayant établi qu'au point de vue volume total bois-fort-tige tout au moins, une tige d'éclaircie est représentative de sa circonférence dans sa classe, il est logique, pour simplifier le travail sur le terrain, d'inclure les tiges d'éclaircie convenables dans l'échantillon; également, ce que nous venons de dire pour le caractère de cime autorise le remplacement, pour le *calcul des volumes*, en cas de nécessité, des tiges A par des tiges B de même circonférence et de caractère de dominance semblable.

La recherche d'une loi $v = f(c)$ calculable alliant une précision suffisante à une relative simplicité des calculs a permis de conclure, après des essais de plusieurs lois dont la complexité des calculs n'améliorait pas la précision de façon *rentable*, à l'adoption de la loi $V = a + b C^2$ dont les équations normales sont faciles à résoudre et qui maintient l'erreur-type entre 1 et 2 %, sauf dispersion aléatoire exceptionnelle des volumes dans l'échantillon. Sur papier millimétré, la courbe de régression du volume en fonction du carré de la circonférence (ou de la surface terrière à 4 π près) est une

droite facile à tracer. Ceci permet de se rendre compte, à priori, de l'allure des droites de régression, et de décider si un tarif unique peut convenir ou s'il est préférable de recourir à plusieurs tarifs (par classe d'arbres, avec groupement éventuel de certaines d'entre elles, ou encore par essence dans le cas des peuplements mélangés). Cette importante question a d'ailleurs été étudiée en détail par P. ARBONNIER (2), Ingénieur à la 1^{re} Section de la Station de Recherches, dans un rapport très documenté qu'il a établi pour le prochain Congrès de l'Union internationale des Instituts de Recherches forestières — Oxford, 1956 — (Section 23).

Le calcul des tarifs permet de supprimer le coefficient personnel du dessinateur et de calculer l'erreur-type. Partant des tiges de l'échantillon dont on connaît la circonférence C à 1,30 m et le volume v , l'application de la méthode des moindres carrés, pour la détermination des constantes a et b , aboutit à la résolution d'un système linéaire simple de 2 équations à 2 inconnues :

$$\begin{aligned} na + b \sum C^2 &= \sum v \\ a \sum C^2 + b \sum C^4 &= \sum v C^2 \end{aligned}$$

la sommation étant étendue à toutes les n tiges de l'échantillon. Les constantes a et b étant connues, le volume total est donné par la formule issue de la sommation pour la population totale dont le nombre d'arbres est $N = \sum n_i$

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

$\sum n_i C_i^2$ se calcule à partir de l'inventaire qui donne, pour chaque circonférence, le nombre d'arbres correspondant n_i .

La connaissance de $\sum n_i C_i^2$ conduit, en outre, automatiquement, en divisant par 4π , à celle de la surface terrière totale, puisque celle-ci, S_t , est liée à C_i^2 par la formule $C_i^2 = 4 \pi S_i$.

Le calcul des quantités $\sum C^2$, $\sum C^4$, $\sum v$, $\sum v C^2$, $\sum v^2$, et la résolution des équations normales peuvent se faire, sans difficultés, à l'aide de nos machines à calculer ordinaires. Il est cependant plus économique de faire exécuter tous ces calculs, en série, par une machine électronique. Le calcul de l'erreur-type s'en trouve d'ailleurs très abrégé, car cette machine donne en même temps la somme des moindres carrés. Cette méthode de calcul, en plus de tous les avantages qu'elle présente sur la méthode graphique, nous permet même, de cette manière, de gagner du temps sur cette dernière.

Cette méthode de la régression linéaire en C^2 , avait d'ailleurs été déjà considérée par HUFFEL (10), A. OUDIN (14) et L. SCHAEFFER (17), mais ceux-ci lui avaient préféré la méthode graphique, plus rapide à l'époque, faute de machines à calculer appropriées.

F. C. HUMMEL (11), tout en partant de formules et de raisonnements mathématiques différents, est parvenu à des conclusions concordantes avec celles de J. ABADIE. Ceci confirme bien l'intérêt de la méthode.

4° ETUDES QUALITATIVES

Les études qualitatives que nous estimons absolument indispensables pour définir la meilleure méthode de traitement nécessitent des classements en « catégories de produits » des tiges de l'échantillon et des essais au laboratoire. Les études expérimentales que nous avons faites dans nos expériences complètes portaient du principe que l'échantillon valablement constitué, représentant une image aussi fidèle que possible du peuplement resté sur pied, il était possible d'appliquer à l'ensemble de ce peuplement les pourcentages des diverses catégories de produits calculés sur l'échantillon. Cette méthode donne, sans aucun doute, des indications approximatives extrêmement précieuses sur l'évolution de la qualité du peuplement, mais les études statistiques ont démontré la complexité du problème. La méthode d'extension peut se trouver faussée par l'irrégularité de la distribution des produits, dont certains peuvent exister ou ne pas exister dans des tiges de même grosseur, tant de l'échantillon que de la population. Il y a intérêt, dans ces conditions, à ne pas multiplier les catégories, à étudier graphiquement l'évolution du volume de chaque catégorie par rapport à la circonférence et à ajuster une courbe simple par la méthode des moindres carrés.

5° CONCLUSION

La méthode de calcul du volume, en quantité et en qualité, des peuplements sur pied de nos places d'essais, que nous venons d'exposer, répond au double but que nous nous étions assigné: précision suffisante et calculable, application relativement simple, tant sur le terrain qu'au bureau.

Nous n'avons jamais perdu de vue le caractère pratique du problème qui nous préoccupait et c'est ainsi, par exemple, qu'il nous a semblé qu'il n'y avait pas lieu, pour ramener l'erreur-type de 1,6 à 1,3 % d'adopter une méthode plus précise, mais nécessitant des calculs exagérément longs, compliqués et, par suite, dispendieux.

La méthode adoptée reste pratique et d'application peu coûteuse, malgré sa valeur scientifique, attestée par les études mathématiques et statistiques qui suivent, et dont nous remercions très vivement J. ABADIE.

Nous ne saurions terminer cet exposé sans adresser nos très sincères remerciements à tous nos camarades forestiers qui nous ont facilité, sur le terrain, la réalisation de nos expériences complètes,

— MM. les Conservateurs LOPPINET à Nancy et SAUR à Metz, MM. les Ingénieurs BILLARD à Saint-Avold et LEROY à Haguenau — et à ceux de leurs Préposés qui nous ont prêté leur concours.

Nous n'aurons garde d'oublier tous nos dévoués collaborateurs qui, à tous les échelons et dans leur domaine propre, ne nous ont pas ménagé leur assistance, et notamment M. l'Ingénieur ARBONNIER qui a tout spécialement assisté J. ABADIE dans ses travaux, et a mis lui-même au point de très nombreux graphiques et plusieurs annexes.

Nancy, le 28 avril 1956.

P. AYRAL.

II. --- ÉTUDES MATHÉMATIQUES ET STATISTIQUES

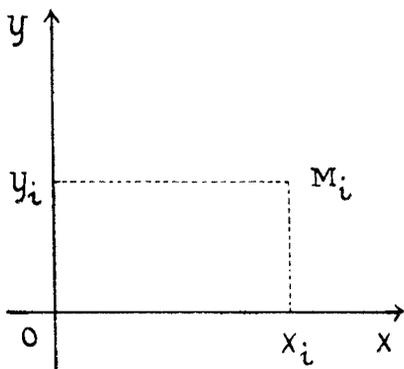
par J. ABADIE
Agrégé de Mathématiques

TITRE I

LA MÉTHODE DES MOINDRES CARRÉS

SECTION I. --- LE POINT DE VUE MATHÉMATIQUE

1° On est en présence d'un ensemble de n couples de nombres (x_i, y_i) , $1 \leq i \leq n$. La représentation graphique par « nuage de points » M_i ayant x_i, y_i pour coordonnées cartésiennes est bien familière.



On veut trouver une courbe passant « le mieux possible » à l'intérieur du nuage de points. Il est à remarquer que la locution « le mieux possible » n'a évidemment pas une signification suffisamment précise pour donner lieu à application de méthodes mathématiques.

2° La méthode dite « des moindres carrés » consiste à procéder en deux étapes de la façon suivante :

a) On se donne une famille (f) de fonctions de x dépendant d'une façon précise d'un certain nombre k de paramètres $a_1,$

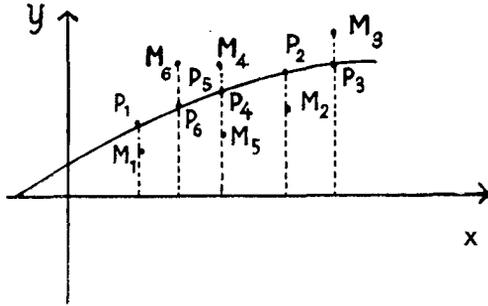
a_2, \dots, a_k ; on désignera par $f(x; a_1, a_2, \dots, a_k)$ une fonction générique de la famille (f).

b) On choisit de représenter le nuage de points par la courbe représentative de celle des fonctions précédentes qui donne à l'expression:

$$r(a_1, a_2, \dots, a_k) = \sum_{1 \leq i \leq n} [y_i - f(x_i; a_1, a_2, \dots, a_k)]^2$$

la valeur minimum; la fonction f ainsi trouvée sera désignée par $\hat{f}(x; a_1, a_2, \dots, a_k)$, ou même pour abrégé par \hat{f} . De même, $r(a_1, a_2, \dots, a_k)$ sera abrégé par r , et la valeur minimum de r par \hat{r} . De même encore, les valeurs de a_1, \dots, a_k qui assurent ce minimum seront désignées par $\hat{a}_1, \dots, \hat{a}_k$.

3° On se rendra mieux compte de la signification de la condition de minimum de r de la façon suivante: la parallèle à l'axe des y menée par chaque point M_i du nuage de points coupe la courbe



représentative d'une des fonctions $f(x; a_1, a_2, \dots, a_k)$ en un point P_i (dépendant naturellement des valeurs données aux paramètres a_1, a_2, \dots, a_k). Il s'agit de trouver celles des courbes de la famille (f) qui rend minimum la somme:

$$r = \sum_{1 \leq i \leq n} \overline{P_i M_i}^2$$

4° Il est à peu près évident que cette « courbe minimisante » occupera effectivement une position « intérieure » au nuage de points. En outre, le calcul des valeurs à donner aux paramètres a_1, a_2, \dots, a_k est théoriquement fort simple, si l'on prend une famille (f) où chaque fonction est dérivable par rapport à chacun de ces paramètres. La condition de minimum s'exprime en effet en annulant chacune des demi-dérivées partielles de la fonction $r(a_1, a_2, \dots, a_k)$. On obtient le système suivant de k équations aux k inconnues a_1, a_2, \dots, a_k :

$$\sum_{1 \leq i \leq n} \left[y_i - f(x_i; a_1, a_2, \dots, a_k) \right] \frac{\partial}{\partial a_j} f(x_i; a_1, a_2, \dots, a_k) = 0 \quad 1 \leq j \leq k$$

On posera, pour alléger les notations :

$$f(x_i; a_1, a_2, \dots, a_k) = f_i$$

Le système résolvant est donc :

$$\left\{ \begin{array}{l} \sum_{1 \leq i \leq n} f_i \frac{\partial f_i}{\partial a_1} = \sum_{1 \leq i \leq n} y_i \frac{\partial f_i}{\partial a_1} \\ \sum_{1 \leq i \leq n} f_i \frac{\partial f_i}{\partial a_2} = \sum_{1 \leq i \leq n} y_i \frac{\partial f_i}{\partial a_2} \\ \dots\dots\dots \\ \sum_{1 \leq i \leq n} f_i \frac{\partial f_i}{\partial a_k} = \sum_{1 \leq i \leq n} y_i \frac{\partial f_i}{\partial a_k} \end{array} \right.$$

On donne aux équations de ce système le nom *d'équations normales*.

5° La résolution des équations normales n'est praticable sans grosses machines spéciales que si ces équations sont « simples » et en nombre « assez petit ».

Le cas le plus simple est celui où les équations sont linéaires par rapport aux inconnues, c'est-à-dire où les fonctions de la famille (f) sont des combinaisons linéaires de k fonctions g₁, g₂, ..., g_k de x seul (c'est-à-dire de k fonctions ne dépendant pas de a₁, a₂, ..., a_k); c'est le cas où les fonctions de la famille (f) sont de la forme :

$$f(x; a_1, a_2, \dots, a_k) = a_1 g_1(x) + a_2 g_2(x) + \dots + a_k g_k(x).$$

Les équations normales sont alors :

$$\left\{ \begin{array}{l} a_1 \sum [g_1(x_i)]^2 + a_2 \sum g_1(x_i)g_2(x_i) + \dots + a_k \sum g_1(x_i)g_k(x_i) = \sum y_i g_1(x_i) \\ a_1 \sum g_2(x_i)g_1(x_i) + a_2 \sum [g_2(x_i)]^2 + \dots + a_k \sum g_2(x_i)g_k(x_i) = \sum y_i g_2(x_i) \\ \dots\dots\dots \\ a_1 \sum g_k(x_i)g_1(x_i) + a_2 \sum g_k(x_i)g_2(x_i) + \dots + a_k \sum [g_k(x_i)]^2 = \sum y_i g_k(x_i) \end{array} \right.$$

les sommes \sum étant étendues à i de 1 à n.

6° A défaut de renseignements d'ordre théorique, on simplifie davantage encore en prenant pour les fonctions de base $g_1(x)$, $g_2(x)$, ..., $g_k(x)$ des polynômes. Deux systèmes de polynômes présentent des avantages, dans certains cas, pour le calcul numérique, tout en étant strictement équivalents au point de vue théorique :

a) Système des puissances :

$$g_1(x) = 1; g_2(x) = x; g_3(x) = x^2; \dots; g_k(x) = x^{k-1}$$

b) Système factoriel :

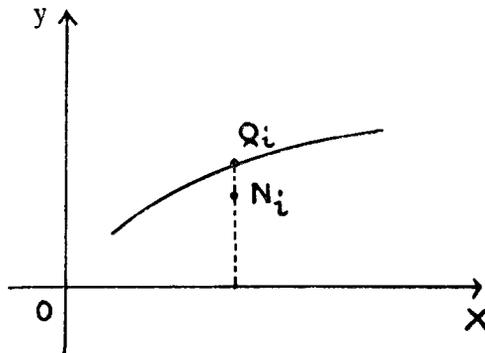
$$g_1(x) = 1; g_2(x) = x; g_3(x) = \frac{1}{2!} x(x-1)$$

$$g_4(x) = \frac{1}{3!} x(x-1)(x-2); \dots; g_k(x) = \frac{1}{(k-1)!} x(x-1)(x-2)\dots(x-k+2)$$

On verra en annexe I un exemple de calcul correspondant au système factoriel.

Les équations normales du système des puissances sont, en posant $k-1 = h$ et $a_{i-1} = b_i$ pour simplifier les notations :

$$\left\{ \begin{array}{l} nb_0 + b_1 \sum x_i + b_2 \sum x_i^2 + \dots + b_h \sum x_i^h = \sum y_i \\ b_0 \sum x_i + b_1 \sum x_i^2 + b_2 \sum x_i^3 + \dots + b_h \sum x_i^{h+1} = \sum y_i x_i \\ b_0 \sum x_i^2 + b_1 \sum x_i^3 + b_2 \sum x_i^4 + \dots + b_h \sum x_i^{h+2} = \sum y_i x_i^2 \\ \dots \dots \dots \\ b_0 \sum x_i^h + b_1 \sum x_i^{h+1} + b_2 \sum x_i^{h+2} + \dots + b_h \sum x_i^{2h} = \sum y_i x_i^h \end{array} \right.$$



7° Le nuage de points du § 1 aurait pu être porté sur « papier à double échelle logarithmique » (en abrégé: papier bilogarithmique), ce qui correspond à porter, sur papier millimétré habituel, en abscisse $X_i = \log x_i$, et en ordonnée $Y_i = \log y_i$. Le couple (x_i, y_i) de nombres positifs sera donc représenté par le point N_i , de coordonnées cartésiennes X_i, Y_i .

Soit (F) une famille de fonctions de X dépendant de k paramètres A_1, A_2, \dots, A_k : on choisit la fonction.

$$\hat{F}(X; A_1, A_2, \dots, A_k)$$

de cette famille qui rend minimum l'expression:

$$R(A_1, A_2, \dots, A_k) = \sum_{1 \leq i \leq n} [Y_i - F(X_i; A_1, A_2, \dots, A_k)]^2$$

L'interprétation graphique reste la même qu'au § 3: l'expression à minimiser est:

$$R(A_1, A_2, \dots, A_k) = \sum_{1 \leq i \leq n} \overline{Q_i N_i}^2,$$

où Q_i est l'intersection de la courbe représentative de la fonction $F(X; A_1, A_2, \dots, A_k)$ avec la parallèle à l'axe des Y menée par le point N_i .

Pour alléger les notations, on fera ici encore des conventions d'écriture analogues à celles qui ont été exposées à la fin du § 2: remplacer simplement dans ce paragraphe les lettres minuscules par des lettres majuscules. Quant aux équations normales, on les obtiendra en faisant cette même substitution dans les équations du § 4.

8° En nous plaçant encore dans le cas bilogarithmique du § 7, on simplifiera les équations normales en considérant la famille (F) des fonctions de la forme:

$$F(X_i; B_0, B_1, \dots, B_n) = B_0 + B_1 X + B_2 X^2 + \dots + B_n X^n$$

(cf. § 6).

Finalement, en revenant aux couples (x, y) primitifs, le nuage des points M_i du § 3 sera représenté par le graphique de la fonction:

$$\begin{aligned} \Theta(x) &= \text{anti } \log F(\log x; \hat{B}_0, \hat{B}_1, \dots, \hat{B}_n) \\ &= 10^{\hat{B}_0 + \hat{B}_1 \log x + \hat{B}_2 (\log x)^2 + \dots + \hat{B}_n (\log x)^n} \end{aligned}$$

c'est-à-dire, en posant $c = \text{antilog } B_0$:

$$\Theta(x) = c x^{\hat{B}_1} + \hat{B}_2 \log x + \hat{B}_3 (\log x)^2 + \dots + \hat{B}_h (\log x)^{h-1}$$

9° La méthode bilogarithmique des §§ 7 - 8, se généralise en remplaçant la fonction logarithme par n'importe quelle fonction continue et croissante dans un intervalle suffisamment étendu. L'intérêt, en recherches forestières, de la fonction logarithme est dû à ce que, sur papier bilogarithmique, le nuage des points représentant, pour chaque arbre d'un même peuplement, les variations liées du volume total et de la circonférence à 1,30 m, se laisse parfois assez bien traverser par une droite. En réalité, l'expression montre qu'une droite ne peut être toujours tenue pour une représentation satisfaisante, et la pratique conduit plutôt à chercher une fonction faiblement concave, c'est-à-dire une fonction dont la concavité de la courbe représentative est tournée vers le bas, la courbure demeurant assez petite.

La courbe la plus simple répondant à ces exigences est la parabole représentative du polynôme du second degré en X :

$$F(X; B_0, B_1, B_2) = B_0 + B_1 X + B_2 X^2$$

ce qui conduit, pour représenter y , à choisir (cf. § 8) une fonction parmi celles du type:

$$c x B_1 + B_2 \log x = c x B_1 (x \log x) B_2$$

Les calculs ont été effectivement faits pour Haguenau et Saint-Amond 47 (nous ne donnerons pas les détails des calculs de cette méthode qui paraît devoir être abandonnée définitivement), x étant la mesure en centimètres de la circonférence à 1,30 m, et y le volume total en décimètres cubes de tiges d'essais d'une place d'expériences. La fonction trouvée est:

pour Haguenau:

$$0,000021878 x^{6,34482} - 1,09556 \log x$$

c'est-à-dire:

$$0,000021878 \frac{x^{6,34482}}{(x \log x)^{1,09556}}$$

pour Saint-Amond 47:

$$0,0036905 x^{4,45424} - 0,57512 \log x$$

est représentée par une droite sur papier bilogarithmique. Les calculs ont été faits pour Haguenau et ont conduit à :

$$Y = -1,98123485 + 2,495803 X,$$

c'est-à-dire, en échelles ordinaires :

$$y = 0,010441 x^{2,4958}$$

13° Pour en revenir au cas linéaire (§ 5), il existe une méthode simple pour calculer r (§ 2) : c'est la *méthode du produit scalaire*. La régression étant :

$f(x; a_1, a_2, \dots, a_k) = a_1 g_1(x) + a_2 g_2(x) + \dots + a_k g_k(x)$,
 r est donné par :

$$\hat{r}(x; \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k) = \sum y_i^2 - \hat{a}_1 \sum y_i g_1(x_i) - \hat{a}_2 \sum y_i g_2(x_i) - \dots - \hat{a}_k \sum y_i g_k(x_i)$$

SECTION 2

LE POINT DE VUE STATISTIQUE

1° Parmi les nombreuses applications où l'arbitraire dont procède la méthode des moindres carrés est réduit par des considérations statistiques, nous nous intéressons ici à l'estimation de :

$$\sum_{1 \leq j \leq n} y_j$$

dans les conditions suivantes :

On étudie une population de N couples de nombres (x_j, y_j) au moyen d'un échantillon de n tels couples connus ; pour les N couples de la population, le premier nombre x est connu pour chacun.

Ce sont exactement là les conditions sous lesquelles s'établissent les tarifs de cubage, x_j représentant, pour la jème tige, la circonférence à 1,30 m, et y_j son volume ; N est naturellement le nombre de tiges restant sur pied dans la place, et n le nombre des tiges servant à établir le tarif ;

$$\sum_{1 \leq j \leq n} b_j$$

est alors le volume total des arbres restant sur pied.

Pour simplifier le langage, on désignera, dans ce qui suit, le couple (x, y) par le mot *arbre*, x par le mot *circonférence*, y par le mot *volume* — ce qui est loisible car la théorie demeure la même.

me quelles que soient les significations concrètes de x et y . On désignera en outre par $N(x)$ et $n(x)$ les nombres d'éléments respectifs de la population et de l'échantillon dans la circonférence x .

2° Ici, toute transformation de l'échelle des volumes (I § 9) est interdite. Il faut choisir une famille $[f]$ de fonctions et appliquer la méthode indiquée en I, §§ 2 à 6.

Une transformation de l'échelle des circonférences est parfaitement autorisée, car elle ne revient qu'à un changement éventuel de la famille de fonctions $[f]$. Une telle transformation s'utilise d'ailleurs couramment pour essayer d'avoir des points représentatifs alignés.

$$N(x)$$

Si, en outre, le coefficient d'extension $\frac{N(x)}{n(x)}$ est constant pour

chaque circonférence x , alors les coefficients (ou poids) c_i du début de I § 10 doivent être tous égaux à 1, et l'on se trouve donc dans les conditions optima.

3° Étant donné l'importance accordée par les forestiers américains à leur méthode (I § 12), il est utile de la discuter, et de discuter aussi la généralisation faite en I §§ 8 et 9.

La première critique à faire est que les méthodes statistiques pour l'estimation d'une somme, dans les conditions indiquées au § 1 ci-dessus, ne conduisent pas à des changements de l'échelle des volumes; une deuxième est que, si l'on appliquait la méthode à l'échantillon, on ne trouverait pas en additionnant les estimations des y_j , la somme, cependant connue, de ces y_j , car l'application du tarif à l'échantillon ne donne pas le volume total de celui-ci. C'est ainsi que, appliquée à l'échantillon d'Haguenau (I, § 12), l'estimation de la somme des volumes est 27 437 dm³ alors que le volume réel de l'échantillon utilisé est 27 199 dm³; l'erreur relative ainsi commise est 0,87 %, ce qui n'est pas négligeable. Cette critique subsiste pour la méthode indiquée au I, §§ 8 et 9, bien que l'erreur relative soit fortement diminuée: pour Haguenau, l'estimation du volume de l'échantillon est 27 121 dm³ (erreur relative 0,3 %) et pour Saint-Amond 47, cette estimation est 17 652 dm³ contre un volume réel de 17 711,5 dm³ (erreur relative: 0,3 %).

4° La deuxième objection est facilement levée. Soit en effet $\theta(x)$ l'une quelconque des fonctions du I § 8. On peut alors corriger l'erreur systématique précédente en considérant la famille $[f]$ des fonctions à deux paramètres $a_1 + a_2 \theta(x)$, et en appliquant la méthode des moindres carrés à cette famille. Les équations normales sont alors:

$$\begin{cases} n a_1 + a_2 \sum \theta(x_i) = \sum y_i \\ a_1 \sum \theta(x_i) + a_2 \sum [\theta(x_i)]^2 = \sum y_i \theta(x_i) \end{cases}$$

les sommations étant étendues à l'échantillon.

On corrige ainsi, non seulement l'erreur déjà signalée au § 3 précédent, mais encore en partie celle qui consiste pour une estimation de $\sum y_i$, à traiter de la même façon une erreur de 1 % pour des y_j petits et pour des y_j grands (cette assimilation est inévitable par l'introduction des logarithmes). Le calcul a été effectivement fait pour Haguenau et pour Saint-Amond 47, avec la fonction:

$$\theta(x) = c x^{B_1 + B_2 \log x},$$

et a donné pour Haguenau:

$$y = 1,41 + 0,000021867 \frac{x^{6,34482}}{(x \log x)^{1,09556}}$$

pour Saint-Amond 47:

$$y = -10,91 + 0,0037849 \frac{x^{4,45424}}{(x \log x)^{0,57512}}$$

5° Il n'est pas du tout aussi facile de pallier la première objection signalée au § 3. Prenons pour famille [f] celle des fonctions:

$$f(x, a_1, a_2, c, B_1, B_2) = a_1 + a_2 \theta(x) = a_1 + a_2 \left[c x^{B_1 + B_2 \log x} \right]$$

que l'on écrira, en changeant la signification des paramètres:

$$f(x; a, b, c, d) = a + b x^c + d \log x$$

les logarithmes étant cette fois népériens.

Il faut, si l'on veut estimer la somme des y_j , minimiser:

$$R(a, b, c, d) = \sum_{1 \leq i \leq n} (y_i - a - b x^c + d \log x)^2$$

ce qui donne pour équations normales:

$$\left\{ \begin{array}{l} na + b \sum x_i^{c+d \log x_i} = \sum y_i \\ a \sum x_i^{c+d \log x_i} + b \sum (x_i^{c+d \log x_i})^2 = \sum y_i x_i^{c+d \log x_i} \\ a \sum x_i^{c+d \log x_i} \log x_i + b \sum (x_i^{c+d \log x_i})^2 \log x_i = \sum y_i x_i^{c+d \log x_i} \log x_i \\ a \sum x_i^{c+d \log x_i} (\log x_i)^2 + b \sum (x_i^{c+d \log x_i})^2 (\log x_i)^2 = \sum y_i x_i^{c+d \log x_i} (\log x_i)^2 \end{array} \right.$$

les sommations étant étendues à l'échantillon.

La résolution en a, b, c, d de ce système, peut être effectuée par la méthode des approximations successives; elle est très longue, et ne peut être abordée pratiquement sans machines spéciales.

6° On peut se demander à cette étape, où est l'intérêt d'une telle fonction $a + bx^c (x^{\log x})^d$. Seules en effet, des commodités de calcul et une vague invitation graphique ont fait adopter en échelles logarithmiques (I § 9) la parabole $Y = B_0 + B_1X + B_2X^2$, qui nous a conduit, après une correction, également choisie en raison de la facilité des calculs, à la fonction $a + bx^c(x^{\log x})^d$. Si l'on devait essayer des fonctions généralisant la fonction kx^α (voir I § 12 — ce qui est à considérer en raison de certains cas où une droite semble convenir sur papier logarithmique), il serait sans doute plus indiqué d'essayer des fonctions de la forme plus simple:

$$f(x; a, b, c) = a + bx + cx^\alpha$$

ou encore de l'une des formes:

$$f(x; a, b, c) = a + bx^2 + cx^\alpha$$

$$f(x; a, b, c) = a + bx^\alpha + cx^{\alpha+1}$$

$$f(x; a, b, c) = a + bx^\alpha + cx^{2\alpha}$$

ou même plus simplement $f(x; a, b, \alpha) = a + bx^\alpha$; etc... qui conduisent à des équations normales plus aisées à résoudre que celles du § 5.

7° En se plaçant toujours dans l'hypothèse formulée au § 2
 $N(x)$
 (constance du coefficient d'extension $\frac{N(x)}{n(x)}$), les méthodes statistiques conduisent à prendre, pour estimer la somme

$$\sum_{i \leq j \leq n} y_j$$

les coefficients c_i du I § 10 égaux à 1, c'est-à-dire à appliquer la méthode du I §§ 2 et 3. Seule subsiste l'indétermination due au choix de la famille [f] des fonctions. En application de la remarque faite au I § 6, on a choisi des polynômes de degrés 4, 3, 2. Le genre de machine à calculer mise à notre disposition pour ces calculs a conduit à écrire le polynôme général de degré k sous la forme :

$$u_0 + u_1 x + \frac{1}{2!} x^{(2)} + \frac{1}{3!} x^{(3)} + \dots + \frac{1}{k!} u_k x^{(k)},$$

où l'on a posé suivant l'usage :

$$x^{(n)} = x(x-1)(x-2) \dots (x-n+1).$$

Ceci permet l'application de la théorie des différences finies, si souvent utilisée dans les calculs numériques. On trouvera dans l'Annexe I les détails du calcul.

On a trouvé les formules suivantes :

HAGUENAU :

degré 4 :

$$y = 368,50 + 13,3694(x-65) + \frac{1}{2} 0,163533(x-65)(x-66) - \frac{1}{6} 0,00295010(x-65)(x-66)(x-67) + \\ \frac{1}{24} 0,0002427851(x-65)(x-66)(x-67)(x-68)$$

degré 3 :

$$y = 366,87 + 13,3216(x-65) + \frac{1}{2} 0,189793(x-65)(x-66) - \frac{1}{6} 0,00272504(x-65)(x-66)(x-67)$$

degré 2 :

$$y = 366,85 + 12,9542(x-65) + \frac{1}{2} 0,187234(x-65)(x-66)$$

SAINT-AVOLD 156 :

degré 3 :

$$y = 766,67 + 19,5171(x-86) + \frac{1}{2} 0,171342(x-86)(x-87) - \frac{1}{6} 0,00416647(x-86)(x-87)(x-88)$$

degré 2 :

$$y = 774,27 + 18,8791(x-86) + \frac{1}{2} 0,127174(x-86)(x-87)$$

SAINT-AMOND 47 :

degré 3 :

$$\hat{y} = 417,17 + 13,5100(x-65) + \frac{1}{2} 0,253472(x-65)(x-66) + \frac{1}{6} 0,01486356(x-65)(x-66)(x-67)$$

degré 2 :

$$y = 404,09 + 15,1501(x-65) + \frac{1}{2} 0,341435(x-65)(x-66)$$

8° En résumé, faute d'indications théoriques, on est conduit à adopter la régression polynômiale pour des raisons de commodités de calcul; par ailleurs, cette méthode est en accord avec les impératifs statistiques.

La méthode américaine ne semble pas à retenir, ni non plus l'amélioration signalée au I § 8 (parabole sur papier logarithmique).

9° Ayant ainsi adopté la régression polynômiale, on est conduit à en chercher une approximation rapide par groupement des valeurs de x , par groupe de 5 unités par exemple. A l'intérieur d'une classe de 5 unités, les valeurs correspondantes de x se calculent par interpolation linéaire.

C'est là une approximation de la méthode qui consiste à choisir comme famille [i] de fonctions, une certaine famille de lignes polygonales, dont les sommets sont une parabole arbitraire, et c'est aussi une approximation de la méthode polynômiale.

Une correction est nécessaire, car la méthode présente le deuxième défaut signalé au § 2. Voici une correction extrêmement facile: soit y_r la somme de $\sum y_i$ étendue à l'échantillon et soit y_e l'estimation de y_r donnée dans le précédent calcul: la correction consistera à

ajouter la constante algébrique $\frac{y_r - y_e}{n}$ à chaque valeur calculée

comme estimation de y . Cette correction ne fait par ailleurs pas sortir de la famille [i] de fonctions choisies, car elle correspond graphiquement à une translation parallèle à l'axe des y .

10° Les détails du calcul sont donnés à l'Annexe 2 pour le 2^e degré. Voici les résultats obtenus sur les 7 places temporaires de volume total réel connu au moment de cette étude (volume mesuré après abattage de tous les arbres de la place):

TABLEAU.

TABLEAU n° 1

Place	Erreur type	Erreur vraie
		$V_r - V_t$
Haguenau	1,01 %	— 0,69 %
St-Avold 142	2,19 %	2,32 %
St-Avold 156	1,82 %	0,82 %
St-Avold 134 b	1,86 %	— 1,65 %
Porcellette	1,86 %	2,32 %
St-Amond 46	1,05 %	— 1,19 %
St-Amond 47	1,23 %	— 0,13 %

Pour le calcul de l'erreur-type, consulter la section 3.

SECTION 3

LE CALCUL DE L'ERREUR

1° Nous nous plaçons dans les hypothèses déjà exposées à la Section 2, §§ 1 et 2. En particulier $\frac{n(x)}{N(x)}$ est constant lorsque x varie. Les notations et conventions sont celles qui ont déjà été constamment utilisées précédemment.

2° La première chose à faire est de calculer $r(a_1, a_2, \dots, a_k)$ (I, § 2). La méthode courante est ensuite de négliger l'erreur aléatoire sur les paramètres a_1, \dots, a_k , et d'estimer la variance de

$$\sum_{1 \leq i \leq n} y_i$$

en multipliant r par un coefficient convenable.

Hypothèse 1: l'échantillon fait partie de la population restant sur pied.

Le coefficient à prendre est:

$$\left[\frac{N^2}{n^2} - \frac{N}{n} \right] \frac{n}{n - k}$$

Hypothèse 2: l'échantillon n'a aucun élément commun avec la population (échantillon détruit).

Le coefficient est alors :

$$\left[\frac{N^2}{n^2} + \frac{N}{n} \right] \frac{n}{n - k}$$

Hypothèse 3 (englobant les deux précédentes): n_1 éléments de l'échantillon sont conservés dans la population, n_2 éléments sont détruits. On suppose en outre que :

$$\frac{n_1(x)}{N(x)} \quad \text{et} \quad \frac{n_2(x)}{N(x)}$$

sont constants lorsque x varie.

Le coefficient est $\frac{N^2}{n^2} g$, avec

$$g = \left(1 + \frac{n_2 - n_1}{N} \right) \frac{n}{n - k} = \frac{N + n_2 - n_1}{N} \cdot \frac{n}{n - k}$$

comme on le démontre à l'Annexe 3.

Ce n'est aucune de ces trois hypothèses qui est couramment réalisée, mais plutôt la suivante :

Hypothèse 4: la même que l'hypothèse 3, mais on ne postule plus la constance de $\frac{n_1(x)}{N(x)}$ et de $\frac{n_2(x)}{N(x)}$.

On peut employer une méthode approximative :

soit $r_1 = \sum_1 [y_1 - f(x_1; a_1, \dots, a_k)]^2$, la partie de r qui correspond à des arbres (x_1, y_1) de l'échantillon restant sur pied, et de même r_2 la partie qui correspond à des arbres détruits. On peut

adopter, en posant $\frac{N}{n} = u$:

$$\frac{n}{n - k} [(u^2 - u) r_1 + (u^2 + u) r_2]$$

c'est-à-dire :

$$\frac{n}{n - k} [u^2 (r_1 + r_2) - u (r_1 - r_2)]$$

Mais, dans la pratique, nous en sommes néanmoins revenus à la formule de l'hypothèse 3. (Voir à ce sujet Titre VII, § 3.)

3° Désignons par r' le résultat du calcul exposé au § 2. $\sqrt{r'}$ est une estimation de l'erreur-type relative à la méthode actuelle d'estimation du volume.

Il est d'usage d'utiliser une table de primitives de la fonction de LAPLACE-GAUSS dont voici un extrait :

TABLEAU n° 2

x	p (en %)
0,0	100 %
0,2	84,1 %
0,4	68,9 %
0,6	54,9 %
0,8	42,4 %
1,0	31,7 %
1,2	23,0 %
1,4	16,2 %
1,6	11,0 %
1,8	7,2 %
2,0	4,6 %
2,2	2,8 %
2,4	1,6 %
2,6	0,93 %
2,8	0,51 %
3,0	0,27 %
3,2	0,14 %
3,4	0,067 %
3,6	0,032 %
3,8	0,014 %
4,0	0,006 %

Cette table se lit de la façon suivante : une erreur-vraie dépassant par exemple 1,2 fois l'écart-type s'observe en moyenne dans 23 % des cas. En pratique, on utilise les résultats approchés suivants : l'écart-type se trouve dépassé par l'erreur vraie à peu près 1 fois sur 3 ; le double de l'écart-type est dépassé environ 5 fois sur 100 ; 2,5 fois l'écart-type est dépassé environ 1 fois sur 100 ; 3 fois l'écart-type est dépassé environ 4 fois sur 1 000.

4° La légitimité de l'emploi de cette table est basée sur un théorème analogue à celui de LIAPOUNOFF : la distribution de la somme de n variables aléatoires, de moyennes respectives m_1, m_2, \dots, m_n et de variances respectives $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ est « très proche ».

« en général », lorsque n est « très grand », de la distribution normale de moyenne $\sum m_1$ et de variance $\sum \sigma_1^2$.

La question critique est, non pas, semble-t-il, de savoir si les conditions d'application du théorème de LIAPOUNOFF sont réalisées, mais si, dans la pratique, n est assez grand pour que la distribution de la somme des n volumes étudiés précédemment soit suffisamment proche de la distribution normale considérée pour être pratiquement confondue avec elle.

Or, l'application du théorème, d'approximative qu'elle soit, deviendrait rigoureuse si, pour chaque x fixé, la distribution des volumes était celle de LAPLACE-GAUSS. Il a donc paru nécessaire de tester une telle hypothèse (voir Titre II), et cela d'autant plus qu'un résultat éventuellement positif justifierait un critère de l'utilité (ou de l'inutilité) des classifications en catégories 1A, 1B, 1C; 2A, 2B, 2C; 3A, 3B, 3C de M. le Conservateur AYRAL (voir Titre III).

TITRE II

ETUDE DE LA NORMALITÉ DE LA DISTRIBUTION
DU VOLUME DANS CHAQUE CIRCONFÉRENCE,
D'UNE MÊME CATÉGORIE D'UNE MÊME PLACE

1° Soit un échantillon y_1, y_2, \dots, y_n de variables indépendantes normales (m, s^2). On pose :

$$\bar{y} = \frac{1}{n} \sum_{1 \leq i \leq n} y_i,$$

et :

$$z_i = y_i - \bar{y},$$

puis :

$$r = \sum_{1 \leq i \leq n} z_i^2,$$

et enfin :

$$t_i = \sqrt{n-2} \frac{\sqrt{n} z_i}{\sqrt{(n-1)r - n z_i^2}}$$

t est distribuée dans une t — distribution à $n - 2$ degrés de liberté, qui est tabulée dans FISHER and YATES, Statistical Tables (Table III). (Cf. Annexe 4).

2° En posant :

$$\bar{y}_p = \frac{1}{p} \sum_{1 \leq j \leq p} y_j, \quad z_p = y_p - \bar{y}_p,$$

on peut mettre r sous la forme :

$$r = \sum_{2 \leq p \leq n} \frac{p}{p-1} z_p^2$$

Si l'on pose encore :

$$t_p = \sqrt{p} \cdot \frac{\sqrt{\frac{p+2}{p+1}} Z_{p+2}}{\sqrt{p_{p+1}}}$$

avec

$$p_p = \sum_{2 \leq j \leq p} \frac{j}{j-1} z_j^2$$

les variables t_1, t_2, \dots, t_{n-2} sont indépendantes, et distribuées dans des t — distributions à 1, 2, ..., $n - 2$ degrés de liberté respectivement. (Cf. Annexe 4).

3° Le § 1 fournit un critère de normalité de la façon suivante : on applique le t — test à la variable t du § 1, après avoir tiré au hasard la valeur de i parmi les nombres 1, 2, ..., n (cette méthode exige $n \geq 3$).

4° On a un critère plus fin si l'on dispose d'un grand nombre d'échantillons ayant respectivement n_1, n_2, \dots, n_i éléments ($n_i \geq 3$) : on calcule pour chaque échantillon i , et comme il a été dit au § 3, des nombres respectifs t_1, t_2, \dots, t_i .

On considère ensuite un tableau dressé de la façon suivante :

p	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001	0
---	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	-------	---

Pour plus de clarté, la façon de remplir ce tableau sera exposée sur un exemple : supposons que l'on ait trouvé $t_1 = 0.591$ pour $n_1 = 8$ (donc pour $n_1 - 2 = 6$). On voit, dans la table III citée, que, à 6 degrés de liberté, le nombre 0,591 se place entre les colonnes marquées :

0.6 (on lit 0,553) et
0.5 (on lit 0,718).

On mettra une barre dans le tableau, entre 0.6 et 0.5.

p	1	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001	0
---	---	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	-------	---

5° Disposant ensuite du tableau rempli, on compare la distribution réelle obtenue pour les nombres t à la distribution théorique, qui, en admettant la normalité de la distribution des y , est la distribution rectangulaire : celle-ci se traduit ici en distribuant le nombre

total N des barres proportionnellement aux intervalles successifs, c'est-à-dire conformément au tableau suivant :

p	1	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001	0
	$\frac{N}{10}$	$\frac{N}{20}$	$\frac{3N}{100}$	$\frac{N}{100}$	$\frac{9N}{1000}$	$\frac{N}{1000}$									

Désignons par v_i le nombre réel de barres dans le i^{me} intervalle, par v'_i le nombre théorique correspondant. On calcule :

$$\chi^2 = \sum \frac{(v_i - v'_i)^2}{v'_i}$$

La comparaison des deux distributions (réelles et théoriques) se fait en consultant la table de χ^2 (FISHER and YATES, Table IV), le nombre de degrés de liberté à prendre étant le nombre de colonnes, diminué d'une unité. On peut aussi utiliser la table plus précise de PEARSON and HARTLEY, *Biometrika Tables*, table 7.

6° Le § 2 donne un meilleur test encore, en ce sens que le nombre des barres est fortement augmenté : en effet, pour chaque échantillon (i) de n_i éléments, on possède $n_i - 2$ valeurs $t_1(i)$, $t_2(i)$, ..., $t_{n_i-2}^{(i)}$. Le principe d'application de ce nouveau test est le même que le précédent.

7° Les deux tests ont été appliqués (voir Annexe 5 pour le premier et Annexe 6 pour le deuxième), aux sept places temporaires, sur la totalité de la population restant sur pied, en tenant compte pour chaque place de la division en classes 1A, 1B, 2A, 2B, 3A, 3B. On a pris pour échantillon élémentaire par exemple la totalité des tiges restant sur pied dans la circonférence 72 de la classe 1A de la place temporaire d'Haguenau.

8° On trouve, avec la méthode exposée aux §§ 3 à 5, $\chi^2 = 7,857$ avec 9 degrés de liberté.

La table de χ^2 donne une probabilité de 55 % pour qu'une valeur de χ^2 supérieure ou égale à celle-ci soit atteinte ou dépassée, par le seul fait du hasard, si l'on suppose normale, dans chaque circonférence et pour chaque catégorie de chaque place, la distribution du volume.

La méthode du § 6 donne de même : $\chi^2 = 7,062$ avec 9 degrés de liberté, ce qui correspond à une probabilité de 63 %.

9° En conclusion, il n'y a aucune raison de douter de l'hypothèse suivant laquelle, dans chaque place, à l'intérieur d'une même classe,

la distribution du volume pour chaque circonférence est la distribution normale de Laplace-Gauss.

10° Il en résulte, en particulier :

1° la légitimation de la méthode utilisée pour évaluer l'erreur-type commise sur l'évaluation du volume total, ainsi que celle de l'utilisation de la table de la loi de LAPLACE-GAUSS (Titre I - Section 3);

2° la légitimation de la méthode utilisée pour la comparaison des classes (Titre III).

TITRE III

COMPARAISON DES CLASSES

1° Soit $(y_1, \dots, y_{n_1}), (y'_1, \dots, y'_{n_2})$ deux échantillons extraits d'une même distribution normale. La variable

$$t = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \frac{\bar{y} - \bar{y}'}{\sqrt{r_1 + r_2}},$$

les notations étant analogues à celles du titre II, est distribuée dans une t distribution à $(n_1 + n_2 - 2)$ degrés de liberté. (Voir par exemple CRAMER - *Mathematical Methods of Statistics*, 29 - 4).

2° Soit maintenant à tester une différence entre les classes 1A et 2A par exemple. Cette comparaison sera faite entre les tiges 1A et les 2A d'une même circonférence, à l'intérieur d'une même place (ce qui donne bien deux échantillons). La comparaison sera naturellement faite pour chaque circonférence de chaque place d'expérience.

Si l'on veut comparer 1 et 2, on comparera, outre 1A et 2A comme il vient d'être dit, 1B et 2B et, s'il y a lieu, 1C et 2C.

Les résultats seront représentés par des barres mises dans un tableau identique à celui qui a été utilisé au Titre II; on appliquera ensuite, comme indiqué à ce même Titre II, le χ^2 - test à la comparaison du tableau réel ainsi obtenu au tableau théorique correspondant à une distribution rectangulaire.

3° Comparaisons 1, 2. — A vrai dire, pour tenir compte du signe de t , on a quelque peu modifié le tableau des t ; d'autre part, pour avoir des groupes théoriques assez nombreux, on a été amené à dresser pour le calcul de χ^2 le tableau suivant, où sont indiqués les résultats réels des comparaisons 1, 2 et, au-dessous, les nombres théoriques:

TABLEAU n° 3

		-			1	+			
		0	0,4	0,8	0,8	0,4	0		
	3	5	14	12	23				
	11,4	11,4	11,4	11,4	11,4				

et le χ^2 correspondant est 2,55 à 4 degrés de liberté. La Table 7 de PEARSON donne, pour probabilité correspondante, 0,63.

6° On peut étudier les signes comme au § 4. Ceci aura encore l'avantage de pouvoir faire intervenir 161 comparaisons indépendantes, au lieu des 76 du § 5.

Les résultats sont les suivants :

— la catégorie A surpasse la catégorie B: 90 fois;

— la catégorie B surpasse la catégorie A: 71 fois.

Ainsi :

$$\chi^2 = \frac{(90 - 80,5)^2}{80,5} + \frac{(71 - 80,5)^2}{80,5} = 2,24$$

La Table 7 de PEARSON donne, pour probabilité, 0,13.

Le § 5, comme le § 6, n'interdit nullement d'attribuer au seul hasard les différences constatées entre les catégories A, B.

7° On peut étudier le nombre de fois où la classe A surpasse la classe B, en groupant les résultats par places. On n'a conservé que les places réunissant au moins 10 comparaisons AB :

TABLEAU n° 5

Places	A surpasse B
Haguenau	16 fois sur 37
St-Avold 134 b	25 fois sur 38
St-Amond 47	8 fois sur 17
St-Avold 156	13 fois sur 30
St-Amond 46	9 fois sur 10
Porcelette	14 fois sur 22

Le χ^2 correspondant a été trouvé (voir Annexe 7) égal à 6,544 pour 5 degrés de liberté, soit une probabilité d'environ 0,25.

Les écarts entre les nombres observés et les nombres issus de l'hypothèse de la similitude de la distribution des volumes ne sont donc pas, ici encore, significatifs.

8° On peut reprendre le § 7, en groupant les places par espèces :

TABLEAU n° 6

Espèces	A surpasse B
Epicéas	34 fois sur 74
Pins sylvestres	39 fois sur 60
Hêtres	17 fois sur 27

On trouve $\chi^2 = 3,339$ à 2 degrés de liberté, soit une probabilité d'environ 0,18, encore dénuée de significativité.

9° Conclusion:

Il est hors de doute que les volumes de tiges de classe 1 et ceux de classe 2 ne suivent pas des lois statistiques identiques, et que, pour préciser davantage, les volumes de classe 1 sont, en moyenne, et toutes choses égales d'ailleurs, supérieurs à ceux de classe 2.

Il n'a pas été possible de trouver une différence quelque peu significative entre les distributions de volume de classe A et les distributions de volume de classe B (toutes choses égales évidemment par ailleurs).

TITRE IV
COMPARAISON ECLAIRCIE —
POPULATION RESTANT SUR PIED

1° L'échantillon d'une place est composé en grande partie de tiges d'éclaircie. Pour une circonférence donnée et pour une catégorie déterminée (1 A, 1 B, 2 A, etc...), une tige d'éclaircie diffère-t-elle en moyenne significativement d'une tige restant sur pied? Dans l'affirmative, il serait impossible de s'en tenir à la méthode d'échantillonnage adoptée, qui se base implicitement sur le fait qu'une tige d'éclaircie est représentative de sa circonférence dans sa catégorie.

2° On a adopté, dans chaque place séparément, une méthode de comptage des signes + ou — des différences $V_{deb} - V_{ab}$, où V_{deb} représente la moyenne des volumes, dans une catégorie déterminée, des tiges restant sur pied dans une même circonférence et V_{ab} la moyenne analogue pour les arbres de l'éclaircie.

Ce comptage a donné les résultats suivants :

TABLEAU n° 7

	1 A		1 B		2 A		2 B		3 A		3 B	
	+	—	+	—	+	—	+	—	+	—	+	—
<i>Epicéas:</i>												
Haguenau	0	3	—	—	4	3	5	6	0,5	1,5	7	2
St-Avoid 142	1	0	0	2	—	—	1	0	—	—	—	—
St-Avoid 152	1	1	1	3	—	—	9	2	—	—	2	0
<i>Pins sylvestres:</i>												
St-Avoid 134 ^b ...	6	3	5	10	0	1	7	4	—	—	4	2
Porcelette	2	2	0	2	1	0	1	1	—	—	—	—
<i>Hêtres:</i>												
St-Amond 46	2	3	2	1	—	—	4	3	—	—	1	0
St-Amond 47	1	1	5	4	—	—	1	4	—	—	—	—
TOTAUX	13	13	13	22	5	4	28	20	0,5	1,5	14	4

On a appliqué la méthode du χ^2 au comptage global des signes +, avec groupement des classes trop faibles, en partant de l'hypothèse que la probabilité d'un signe + est égale à celle d'un signe —.

TABLEAU n° 8

	1 A	1 B	2	3
Observé	13	13	33	14,5
Théorique. ...	13	17,5	28,5	10
Δ	0	— 4,5	+ 4,5	+ 4,5
Δ^2	0	20,25	20,25	20,25
Δ^2 /théorique.	0	1,16	0,71	2,03

$\chi^2 = 3,90$, avec 3 degrés de liberté.

Il y a une probabilité de 27 % pour que cette valeur de χ^2 soit atteinte ou dépassée, ce qui n'est en aucune façon significatif.

3° Appliquons le test χ^2 , à l'épicéa seul, en groupant toutes les catégories :

	+	—	
Observé	31,5	23,5	55
Théorique	27,5	27,5	55
Δ	4	— 4	0
Δ^2	16	16	32

$$\chi^2 = \frac{32}{27,5} = 1,16 \text{ à } 1 \text{ degré de liberté.}$$

La probabilité pour que cette valeur soit atteinte ou dépassée est 55 %, ce qui n'est aucunement significatif.

Pour le Pin sylvestre, on trouve :

26 fois le signe +,
25 fois le signe —,

et il est évidemment inutile d'appliquer un test statistique pour voir que la différence observée n'est pas significative.

Pour le hêtre, on a :

16 fois le signe +,

16 fois le signe —,

ce qui aussi évidemment entraîne la même conséquence.

4° En conclusion, sur une place donnée, une tige d'éclaircie équivalent, en moyenne et au point de vue du volume total, à une tige restant sur pied de même catégorie et de même circonférence.

On a vu par ailleurs que les différences entre catégories 1, 2, 3 étaient significatives, les différences entre catégories A, B ne l'étant pas, ni pour le volume total, ni pour les volumes des différentes catégories de produit.

Tout ceci justifie donc les règles suivantes, destinées à prendre un échantillon aussi représentatif que possible de la population restant sur pied :

A) stratifier la population restant sur pied en catégories 1 A, 1 B, 1 C, 2 A, 2 B, 2 C, etc...

B) l'échantillon doit être proportionnel dans chaque catégorie et pour chaque circonférence, aux nombres correspondants de tiges dans la population restant sur pied.

C) Il est permis, en vue de simplifier le travail sur le terrain, de représenter une tige de circonférence et d'une catégorie déterminée par une tige *d'éclaircie* de même circonférence et de même catégorie.

D) Il est même permis de remplacer une tige 1 A par une tige 1 B de même circonférence dans l'échantillon (ou une tige 2 A par une tige 2 B, 3 A par 3 B), en ce qui concerne *le calcul des volumes*. Mais il se peut qu'en vue d'autres études (prévision d'accroissements par exemple), la distinction des tiges A et des tiges B ait son importance.

En ce qui concerne l'établissement d'un tarif de cubage on est en réalité obligé dans la pratique de l'échantillonnage de grouper les tiges par groupes de circonférences (de 5 cm en 5 cm, ou même de 10 en 10) et on se permettra de remplacer une tige 1 A, par exemple, par une tige 1 B du même groupe de circonférences.

TITRE V

MÉTHODES DE CUBAGE
BASÉES SUR LA CLASSIFICATION DES TIGES

1° Il a été démontré, au Titre III, qu'il existait une différence réelle, toutes choses égales par ailleurs, entre les distributions des volumes des tiges de classe 1 et de classe 2. Il est tout naturel d'étendre ceci à la classe 3.

Il n'a pu être mis en évidence aucune différence significative entre les distributions des volumes des tiges de classe A et de classe B.

En conséquence, on est logiquement conduit, dans le calcul d'un tarif de cubage, à tenir compte des classifications en classes 1, 2, 3.

2° Cette diversification introduisant une souplesse supplémentaire, on a été conduit à chercher à diminuer par ailleurs le nombre des constantes de la famille [f] de fonctions (voir Titre I). On s'est finalement arrêté à la famille des fonctions

$$a + bx^2$$

en imposant à la constante b d'être la même pour toutes les classes, la constante a prenant autant de valeurs qu'il y a de classes à distinguer. Cette méthode portera le nom de « *méthode linéaire en C²* », la circonférence mesurée en centimètres étant désignée par C dans la pratique forestière.

3° La courbe représentative d'une telle fonction $f(x) = a + bx^2$ est une droite, à condition d'adopter une échelle des carrés en abscisse. Cela permet de juger à l'œil, en bloc, de la façon dont une telle loi s'adapte à la réalité. Les graphiques confirment nettement cette adaptation.

4° Les équations normales sont, si l'on distingue trois classes:

$$\left\{ \begin{array}{l} n_1 a_1 + b \sum_1 x^2 = \sum_1 y \\ n_2 a_2 + b \sum_2 x^2 = \sum_2 y \\ n_3 a_3 + b \sum_3 x^2 = \sum_3 y \\ a_1 \sum_1 x^2 + a_2 \sum_2 x^2 + a_3 \sum_3 x^2 + b \sum x^4 = \sum yx^2 \end{array} \right.$$

où \sum_i désigne une sommation, étendue à la classe i , n_i le nombre de tiges de cette classe, et \sum une sommation sur l'échantillon total.

si l'on distingue deux classes:

$$\left\{ \begin{array}{l} n_1 a_1 + b \sum_1 x^2 = \sum_1 y \\ n_2 a_2 + b \sum_2 x^2 = \sum_2 y \\ a_1 \sum_1 x^2 + a_2 \sum_2 x^2 + b \sum x^4 = \sum y x^2 \end{array} \right.$$

avec les mêmes notations.

5° L'erreur-type sur l'évaluation du volume est $\frac{N}{n} \sqrt{g Q}$, avec :

$$Q = \sum y^2 - a_1 \sum_1 y - a_2 \sum_2 y - b \sum y x^2 \quad \text{pour 2 classes,}$$

$$Q = \sum y^2 - a_1 \sum_1 y - a_2 \sum_2 y - a_3 \sum_3 y - b \sum y x^2$$

pour 3 classes, g désignant le coefficient indiqué au Titre I, Section 3, § 2 (avec $k = 3$ pour 2 classes, $k = 4$ pour 3 classes).

6° Voici une formule d'erreur-type qui, bien que quelque peu empirique, serre la réalité de plus près. On peut chercher en effet à introduire à des étapes différentes du calcul l'effet d'une part des coefficients a_i , et d'autre part du coefficient b qui intéresse globalement l'échantillon.

La formule aura l'avantage intuitif, et d'ailleurs purement logique, d'introduire une indétermination dans la partie du calcul correspondant à la classe i lorsque cette dernière ne comporte qu'une seule tige (auquel cas naturellement on se hâterait de grouper la classe i avec une autre).

Voici la méthode de calcul pour le cas de deux classes : on calcule :

$$r_1 = \sum_1 (y - a_1 - b x^2)^2,$$

$$r_2 = \sum_2 (y - a_2 - b x^2)^2$$

par les formules :

$$r_1 = \sum_1 y^2 - a_1 \sum_1 y - b \sum_1 y x^2 + \varepsilon$$

$$r_2 = \sum_2 y^2 - a_2 \sum_2 y - b \sum_2 y x^2 - \varepsilon$$

en posant :

$$\begin{aligned} \epsilon &= b (\sum_2 yx^2 - a_2 \sum_2 x^2 - b \sum_2 x^4) \\ &= -b (\sum_1 yx^2 - a_1 \sum_1 x^2 - b \sum_1 x^4) \end{aligned}$$

On calcule ensuite :

$$\begin{aligned} r'_1 &= r_1 \frac{n_1}{n_1 - 1} \cdot \frac{N_1 + n'_1 - n''_1}{N_1} \\ r'_2 &= r_2 \frac{n_2}{n_2 - 1} \cdot \frac{N_2 + n'_2 - n''_2}{N_2} \end{aligned}$$

où n'_1 par exemple désigne le nombre d'arbres *abattus* de la classe 1 appartenant à l'échantillon, et n''_1 le nombre des arbres *restant sur pied* dans cette même classe ($n_1 = n'_1 + n''_1$).

Et est donné finalement par

$$\text{Et} = \frac{N}{n} \sqrt{\frac{n}{n-1} (r'_1 + r'_2)}$$

7° L'évaluation du volume est :

$$\begin{aligned} b\sum X^2 + N_1 a_1 + N_2 a_2 + N_3 a_3 &\text{ pour trois classes,} \\ b\sum X^2 + N_1 a_1 + N_2 a_2 &\text{ pour deux classes.} \end{aligned}$$

où les sommations sont cette fois étendues à toute la population restant sur pied, et où N_1, N_2, N_3 désignent le nombre des tiges de chaque catégorie dans cette population.

8° On peut enfin grouper toutes les tiges en une seule classe. Les formules sont alors très simples. Les équations normales sont :

$$\begin{cases} na + b\sum x^2 = \sum y \\ a\sum x^2 + b\sum x^4 = \sum yx^2 \end{cases}$$

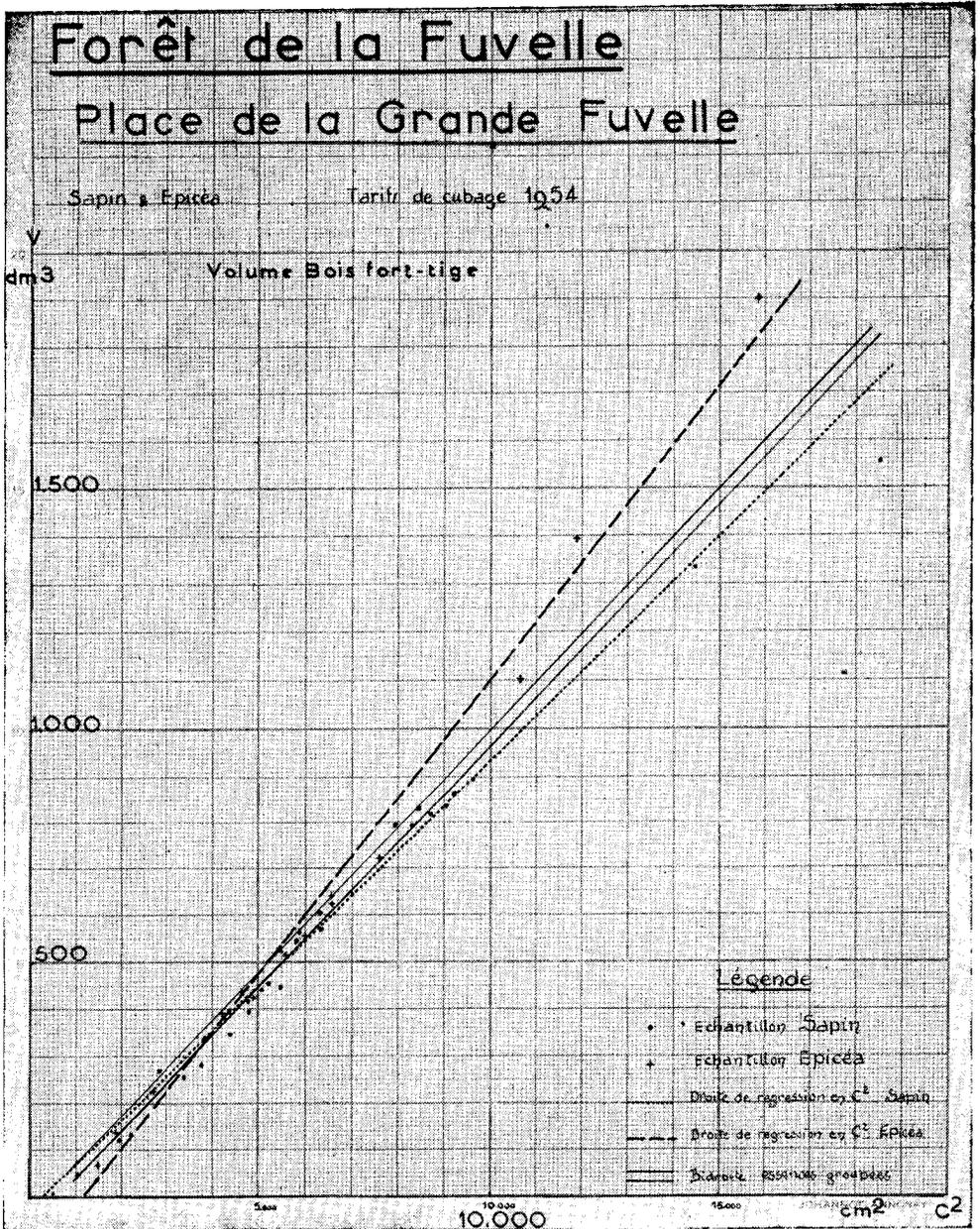
L'erreur-type est $\frac{N}{n} \sqrt{g Q}$ (prendre $k = 2$ dans g)

$$Q = \sum y^2 - a\sum y - b\sum yx^2$$

9° Pour juger de l'utilité de la distinction en classes, nous avons, pour St-Avoid 156, procédé aux calculs suivants :

— 1° un bitarif (distinction en deux classes, les classes 2 et 3 étant groupées pour n'en former qu'une seule);

— 2° un monotarif, les tiges étant toutes groupées en une seule classe.



Graphique 1

L'erreur-type a été de 1,66 % pour le bitarif, et de 1,93 % pour le monotarif. Il y a donc eu là un avantage réel à calculer un bitarif. Toutefois, nous devons à la vérité de dire que l'expérience a été faite sur la place qui, à priori, semblait donner le plus grand avantage au bitarif.

En conclusion, il est recommandé d'utiliser le monotarif, très simple à calculer, sauf le cas où l'on rechercherait une très grande précision, et où une indication graphique assez nette influencerait en faveur du bitarif. Le tritarif (distinction en trois classes) ne semble pas, sauf peut-être exception qui ne s'est pas jusqu'à présent signalée, devoir être très utile.

On trouvera à l'Annexe 8 un exemple de calcul d'un monotarif, en Annexe 9 un exemple de bitarif.

10° Dans certaines places, deux espèces se trouvent mélangées intimement. Par exemple, à la place de la Grande Fuelle (1), se trouvent des Sapins et des Épicéas. Une simple inspection graphique montre que les lois sont distinctes. Voici le graphique en C², extrait d'un rapport présenté par P. ARBONNIER (2) au XII^e Congrès de l'Union Internationale des Instituts de Recherches Forestières (Graphique 1).

Il a été procédé à différentes stratifications :

1° un bitarif sapin (classe 1 d'une part, classes 2 et 3 bloquées d'autre part), et un monotarif épicéa. L'erreur-type est 1,30 %.

2° deux monotarifs, un pour les épicéas et un pour les sapins. L'erreur-type est 1,35 %.

3° un bitarif sans distinction d'essences. L'erreur-type est 2 %.

La distinction des essences apporte donc dans ce cas une nette amélioration. Il n'en est pas toujours ainsi, comme l'a montré P. ARBONNIER, et une inspection graphique préalable semble s'imposer.

(1) Place d'expériences permanente de la 1^{re} Section dans le Jura.

TITRE VI

COMPARAISON DES DIVERSES METHODES
D'ÉVALUATION DU VOLUME

1° La comparaison de deux méthodes devrait s'effectuer en faisant le rapport des variances (c'est-à-dire des carrés des erreurs-types). Si, par exemple, ce rapport est 2, la première méthode a une « efficacité » double de la seconde, ce qui se traduit par le fait qu'il faut un échantillonnage deux fois plus élevé, en général, dans la seconde méthode, pour arriver à la même approximation.

2° On peut, en première approximation, prendre comme évaluation de l'erreur-type commise dans un ajustement graphique du « nuage de points » les formules du Titre I, Section 3. Il faut bien voir que, ce faisant, on néglige des erreurs, non calculables, dues au dessinateur :

- a) erreur de courbure du dessin,
- b) erreur de « hauteur » dans la position de la courbe,
- c) erreur d'« orientation » dans cette position.

Aussi, si l'on veut comparer une méthode graphique et une méthode de calcul, la méthode du rapport de variances ne peut-elle s'appliquer, puisqu'une part de la variance de la méthode graphique est inconnue.

3° Aussi avons-nous fait cette comparaison par la valeur absolue de l'erreur relative en %. Nous avons comparé quatre méthodes, qui ont été appliquées aux sept placettes dont on connaît le volume total.

Première méthode: tarif graphique habituel sur papier bilogarithmique;

Deuxième méthode: tarif habituel précédent après correction des ordonnées (multiplier l'estimation obtenue par le rapport V_r/V_t du volume réel de l'échantillon à son estimation due à l'application du tarif).

Troisième méthode: celle du Titre I, Section 2 § 9 (régression polynomiale de degré 2, avec groupement des circonférences par classes de 5 cm).

Quatrième méthode: celle du Titre V (droites en C^2).

Voici les erreurs correspondantes, exprimées en pourcentages, la dernière ligne donnant la somme des valeurs absolues de ces erreurs :

TABLEAU n° 10

Place	1	2	3	4
Haguenau	— 1,61 %	— 0,68 %	— 0,69 %	— 0,64 %
St-Avoid 134 ^b . .	— 3,99 %	— 2,91 %	— 1,65 %	— 1,98 %
St-Amond 47 . . .	+ 0,35 %	— 0,32 %	— 0,13 %	— 0,37 %
St-Amond 142 . .	+ 4,00 %	+ 2,74 %	+ 2,32 %	+ 2,12 %
St-Avoid 156 . .	+ 0,10 %	+ 0,78 %	+ 0,82 %	+ 0,91 %
St-Amond 46 . . .	— 1,85 %	— 1,42 %	— 1,19 %	— 1,31 %
Porcelette	+ 1,92 %	+ 2,41 %	+ 2,32 %	+ 2,35 %
	13,82 %	11,26 %	9,12 %	9,66 %

On constate une nette amélioration en passant de la méthode 1 à la méthode 2, puis à la méthode 3 ou à la méthode 4, ces deux dernières étant tout à fait comparables, malgré un léger désavantage apparent de cette dernière.

4° Les méthodes de calcul indiquées au Titre I, Section 2, § 7 (polynômes), ont été essayées sur Haguenau, St-Amond, St-Avoid 156; le tableau suivant donne les erreurs-types de ces méthodes (trois dernières colonnes), ainsi que les *erreurs-types* des méthodes de calcul 3 et 4 :

TABLEAU n° 11

Place	Polynômes 2° degré Circ. groupées Méthode 3	Linéaire en C ² Méthode 4	Polynômes		
			2° degré	3° degré	4° degré
Haguenau	1,01 %	1,01 %	1,01 %	1,01 %	1,2 %
St-Avoid 134 ^b . .	1,86 %	1,86 %			
St-Amond 47 . . .	1,23 %	1,37 %	1,23 %	1,05 %	
St-Amond 142 . .	2,19 %	2,39 %	1,72 %	1,71 %	
St-Avoid 156 . .	1,82 %	1,66 %			
St-Amond 46 . . .	1,05 %	1,04 %			
Porcelette	1,86 %	1,83 %			

D'après ces tableaux, l'avantage est à la parabole du troisième degré, puis à celle du second degré, enfin les méthodes 3 (parabole du second degré avec groupement des circonférences : $y = a_0 + a_1x + a_2x(x-1)$) et 4 (méthode $y = a + bx^2$) semblent équivalentes.

Etant donné les facilités de calcul apportées par la méthode 4, il semble qu'il faille opter entre l'une des solutions suivantes :

- 1) adopter la méthode 4 ;
- 2) adopter la parabole du troisième degré, en faisant effectuer les calculs à façon, par une machine électronique spéciale.

La méthode graphique sur papier logarithmique ne doit pas être utilisée sans sa correction (méthode 2) et, même avec sa correction, semble devoir donner des résultats moins bons, sans regagner, du côté du temps utilisé, la perte de précision.

5° Antérieurement, et pendant des années, la méthode utilisée était celle décrite par P. AYRAL dans sa communication au XI^e Congrès de l'Union Internationale des Instituts de Recherches forestières, « Sur le calcul du volume, total et par catégories de produits, des peuplements restant sur pied dans les places d'essais au passage des éclaircies » (1).

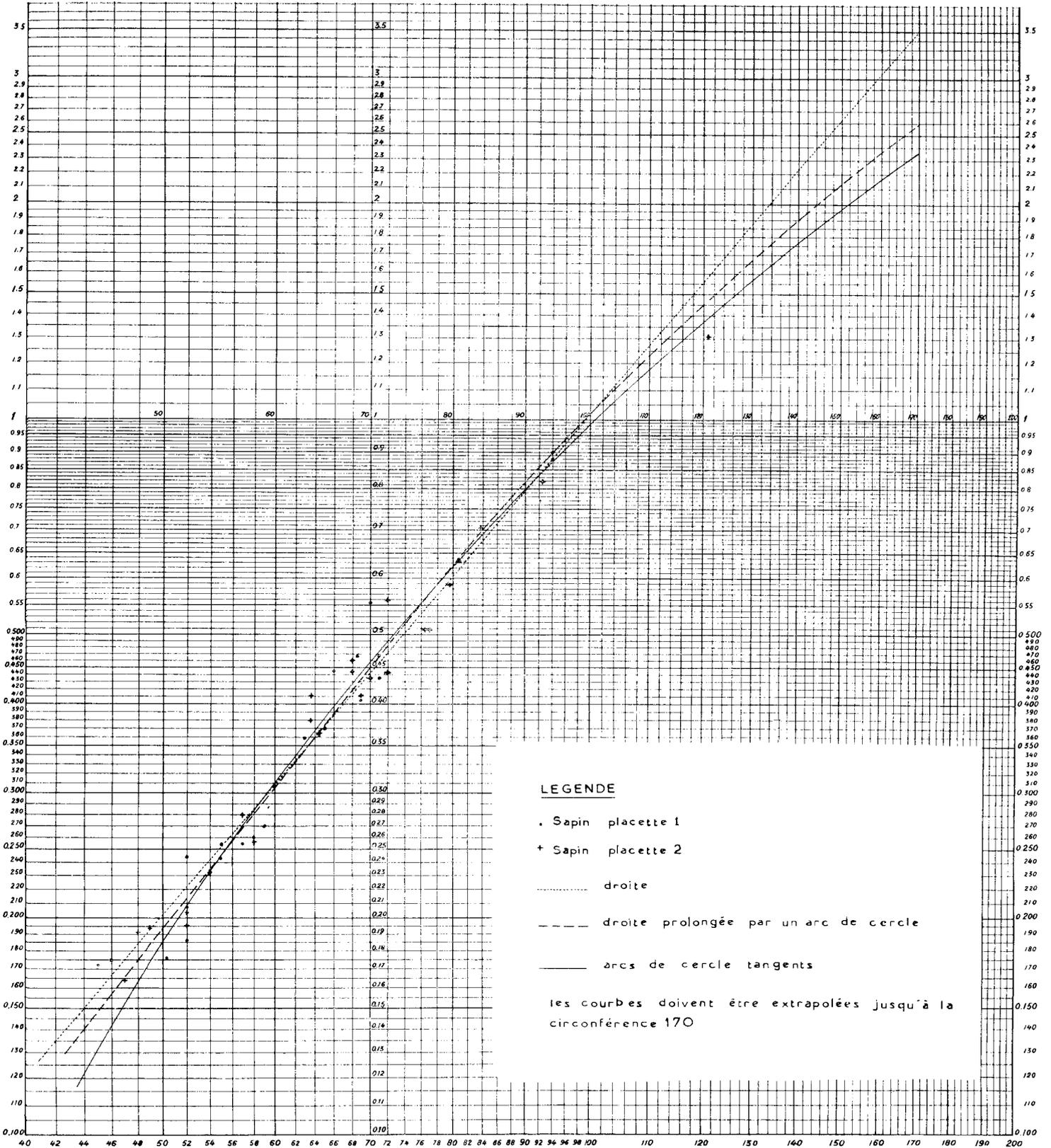
« ...Pour des raisons de commodité, résultant de l'absence d'un matériel véritablement approprié, les tiges d'essai ont été, jusqu'à présent, cubées abattues. Il fallait donc les rechercher principalement parmi les arbres à exploiter dans la place d'essai, au moment de chaque éclaircie, quitte à compléter les catégories manquantes par quelques pieds choisis dans les zones d'isolement pour les besoins de la cause et par suite, forcément, en nombre très limité. Après un inventaire des circonférences à 1,30 m de tous les arbres numérotés dans la place et le martelage de la coupe, les tiges d'essai étaient désignées à vue sur pied, dans plusieurs catégories de circonférences, parmi les arbres à exploiter représentant comme hauteur et comme forme la moyenne des arbres de même grosseur maintenus sur pied. Les arbres à cuber dans la zone étaient ensuite recherchés, le cas échéant... ».

Les inconvénients sont multiples :

— 1° la part « subjective » laissée à l'estimateur est trop grande, celui-ci pouvant avoir tendance à choisir des arbres-types « trop beaux » par exemple ;

— 2° il y a une forte distorsion de l'échantillon : $\frac{n(x)}{N(x)}$ est grand pour les faibles valeurs de la circonférence (nombreux arbres abattus

(1) Compte rendu du XI^e Congrès, Rome, 1953, p. 556.



Graphique n° 2

disparus du peuplement à cuber sur pied), très faible ou *nul* pour les grandes valeurs de la circonférence ;

— 3° en relation avec cette importante distorsion, aucun calcul d'écart-type n'est possible, et il est donc impossible de chiffrer approximativement l'erreur maximum commise ;

— 4° les arbres-types s'étendent, en général, dans les petites circonférences et il y a donc une extrapolation, souvent très forte, de la ligne de régression (appelée tarif de cubage). L'exemple suivant, en double échelle logarithmique, en donne une idée (la courbe a été extrapolée jusqu'à la plus forte circonférence de la population restant sur pied) (1).

Voir graphique n° 2.

6° Quoiqu'il en soit, il existe dans les dossiers des centaines d'échantillonnages effectués avec la méthode des arbres-types. Le problème est donc de tirer le meilleur parti possible de ces échantillons. On a pensé que la méthode linéaire en C^2 (Titre V) pouvait être préférable à la méthode graphique sur papier bilogarithmique. Le tableau suivant donne les résultats des erreurs-vraies (en pourcentage du volume réel) des deux méthodes. On a désigné par n le nombre des arbres-types, N le nombre de tiges de la population restant sur pied :

TABLEAU

(1) On se rend compte du danger de l'extrapolation : suivant la courbe adoptée, on trouve pour $C = 170$ (plus grande circonférence de la population à cuber sur pied) : 2 350 dm³, 2 600 dm³ et 3 500 dm³. Or, il y a :

— dans la placette 1, 72 arbres $C > 120$ (environ 11 m³)

— dans la placette 2, 77 arbres $C > 120$ (environ 11 m³)

L'échantillon de la placette 1 s'arrête même à $C = 94$.

P. ARBONNIER.

TABLEAU n° 12

Places d'essais et surface	Essences et âge moyen	Eclaircie f = faible F = forte	n	N	Erreur vraie méthode des arbres-types	Erreur vraie méthode linéaire en C2
St-Avoid P. 134 ^b 20 ares	P. sylv. 27 ans	f F	25 54	221 221	+ 6,02 % - 0,54 %	- 2,02 % + 2,41 %
St-Avoid P. 142 8 ares	Epicéa 65 ans	f F	20 32	65 65	+ 14,38 % - 0,88 %	+ 5,72 % + 3,71 %
St-Avoid P. 156 34 ares	Epicéa 68 ans	f F	23 42	145 145	- 4,87 % + 8,09 %	+ 7,92 % + 0,52 %
Porcellette 22 a 75	P. sylv. 60 ans	f F	15 29	107 107	- 5,60 % + 2,40 %	+ 13,36 % + 3,32 %
Haguenau P. 208 ^a 39 a 83	Epicéa 65 ans	f F	16 46	285 285	+ 21,64 % + 1,62 %	- 5,68 % + 3,56 %
St-Amond P. 46 15 ares	Hêtre 50 ans	f F	15 23	107 107	+ 1,93 % + 2,03 %	+ 6,19 % - 0,14 %
St-Amond P. 47 15 ares	Hêtre 60 ans	f F	10 18	106 106	+ 5,50 % + 2,07 %	+ 4,61 % - 1,54 %
Somme des valeurs absolues des erreurs					59,04 17,63	45,50 15,20

On constate que la somme des valeurs absolues des erreurs vraies est à peu près la même pour les deux méthodes, avec un léger avantage en faveur de la méthode linéaire en C²; par contre, la plus grande erreur est nettement en faveur de cette dernière, comme l'indique le tableau ci-après.

TABLEAU n° 13

	Méthode des arbres-types	Méthode linéaire en C ²
Eclaircie forte	8,09 %	3,71 %
Eclaircie faible	21,64 %	13,36 %

La méthode d'extrapolation linéaire en C² semble donc préférable. Elle a en outre la supériorité de pouvoir être confiée en série à une machine électronique, ce qui permet de recalculer très rapidement les nombreux tarifs déjà établis par la méthode graphique.

7° En tout état de cause, la comparaison entre les méthodes basées sur les arbres-types et celles qui s'appuient sur l'échantillonnage AYRAL donnent définitivement l'avantage, et d'une façon écrasante à ces dernières, ainsi qu'on peut le vérifier en considérant le tableau récapitulatif suivant :

	Echantillonnage AYRAL, Méthode linéaire en C ²	Arbres-types écl. forte		Arbres-types écl. faible	
		Méthode linéaire en C ²	Méthode graphique	Méthode linéaire en C ²	Méthode graphique
Somme des valeurs absolues des er- reurs vraies	0,66 %	15,20 %	17,63 %	45,50 %	59,04 %
Plus grande erreur vraie	2,35 %	3,71 %	8,09 %	13,36 %	21,64 %

Rappelons encore une fois que la méthode d'échantillonnage AYRAL permet un calcul d'erreur, ce qui est impossible avec la méthode des arbres-types.

8° Une dernière méthode d'extension consiste à appliquer la formule :

$$V = \frac{N}{n} \sum v$$

Volume vrai: 133 535 dm³
N = 367

TABLEAU n° 15
Hagueau

Comparaison des erreurs vraies
Méthode linéaire en C2 et méthode d'extension

N° échantillon	n	$\frac{n}{N}$	Méthode linéaire en C2			Méthode d'extension		
			dm ³ Volume tarif	dm ³ Erreur vraie	Erreur vraie %	dm ³ Volume calculé	dm ³ Erreur vraie	Erreur vraie %
1	25	0,068 119	137 960	4 425	3,31	147 403	13 868	10,38
2	25		132 430	1 105	0,82	138 390	4 855	3,63
3	24	0,065 395	133 822	287	0,21	129 168	4 367	3,27
4	25		132 090	1 445	1,08	135 556	2 021	1,51
5	31	0,084 468	132 980	555	0,41	133 660	125	0,001
6	31		134 148	613	0,45	136 560	3 025	2,20
7	31		134 638	1 103	0,82	125 124	8 411	6,29
8	31		127 238	6 297	4,71	131 304	2 231	1,67
9	37	0,100 817	130 016	3 519	2,63	131 287	2 248	1,68
10	37		130 475	3 060	2,29	131 684	1 851	1,38
11	37		133 201	334	0,25	133 757	222	0,002
12	36	0,098 092	131 507	2 028	1,51	127 400	6 135	4,59
13	37		132 710	825	0,61	133 509	26	0,00
14	41	0,111 716	131 035	2 500	1,87	131 198	2 337	1,75
15	40	0,108 991	134 459	924	0,69	125 368	8 167	6,11
16	41		132 471	1 064	0,79	132 774	761	0,006
17	41		136 266	2 731	2,04	137 106	3 571	2,67
18	46	0,125 340	131 809	1 726	1,29	131 322	2 213	1,65
19	46		136 492	2 957	2,21	136 604	3 069	2,29

..	IV	133 204	ZVI	0,20	132 918	• 617	0,005
22	53	133 082	— 447	— 0,33	135 506	— 1 971	— 1,47
23	52	139 049	— 5 514	— 4,12	135 790	— 2 255	— 1,68
24	52	132 587	948	0,70	128 351	5 184	3,88
25	52	132 586	949	0,71	128 337	5 198	3,89
26	61	131 630	1 905	1,42	131 157	2 378	1,78
27	61	136 795	— 3 260	— 2,44	135 279	— 1 744	— 1,30
28	61	133 666	— 131	— 0,09	132 481	1 054	0,78
29	61	133 998	— 463	— 0,34	133 004	531	0,004
30	74	136 788	— 3 253	— 2,43	136 331	— 2 796	— 2,09
31	73	134 050	— 515	— 0,38	132 206	1 239	0,92
32	73	132 750	785	0,58	131 904	2 441	1,82
33	73	130 939	2 506	1,94	129 530	4 005	2,99
34	73	135 963	— 2 428	— 1,81	133 819	— 284	— 0,002
35	92	134 362	— 827	— 0,61	135 578	— 2 043	— 1,52
36	92	132 032	603	0,45	133 260	275	0,002
37	92	132 453	1 082	0,81	132 941	594	0,004
38	91	130 818	2 717	2,03	128 567	4 968	3,72
39	111	133 533	2	0,00	135 760	— 2 225	— 1,66
40	111	132 700	835	0,62	134 897	— 1 362	— 1,01
41	111	133 164	371	0,27	135 257	— 1 722	— 1,28
42	111	134 695	— 1 160	— 0,86	136 785	— 3 250	— 2,43
43	122	133 459	76	0,06	131 816	1 719	1,28
44	122	134 266	— 731	— 0,54	132 971	564	0,004
45	122	133 782	— 247	— 0,18	132 820	715	0,005
46	123	133 289	246	0,18	134 978	— 1 443	— 1,08
47	148	133 311	224	0,16	135 051	— 1 516	— 1,13
48	148	132 758	777	0,58	135 529	— 1 994	— 1,49
49	148	132 803	732	0,54	134 942	— 1 407	— 1,05
50	148	132 133	1 402	1,04	134 009	— 474	— 0,003

où Σv est le volume de l'échantillon ; elle ne peut être évidemment appliquée qu'avec l'échantillonnage AYRAL. Ce n'est autre que la méthode qui sera appelée « horizontale » au Titre VIII Section 2. Elle possède une vertu incontestable : c'est de fournir rapidement une valeur du volume de la population. Malheureusement, il arrive parfois qu'elle soit par trop imprécise. Par exemple, dans le cas de St-Avoid 134^b, qualité charpente Titre VIII Section 2, la méthode d'extension donne une erreur vraie de — 63 %, contre — 13,8 % avec la méthode linéaire en C (avec, dans ce dernier cas, une erreur-type de 15,4 %).

On peut prendre comme autre exemple celui des 50 tarifs d'Haugenau (Titre VII). Voici la série des 50 erreurs-vraie de la méthode d'extension, comparées à celles de la méthode linéaire en C² :

Voir tableau n° 15.

On voit, il est vrai, que l'erreur vraie dans la méthode linéaire en C² dépasse l'erreur vraie de la méthode d'extension 31 fois sur 50, mais cela n'a aucun signification (le χ^2 correspond à une probabilité de 23 %). Par contre, le tableau suivant donne le nombre de fois où l'erreur vraie de la méthode d'extension tombe dans les intervalles indiqués (l'unité d'intervalle étant l'erreur-type de la méthode linéaire en C²). Pour compléter le tableau, on a indiqué aussi les résultats pour l'erreur-vraie de la méthode linéaire en C² :

TABLEAU n° 16

Nombre d'unités d'intervalle	2	3	4	5	6	7
Méthode d'extension	9	3	0	1	1	
Méthode linéaire en C ²	3	0	0	0	0	

Ainsi, si le test de pile ou face n'apporte pas la certitude que l'erreur-vraie de la méthode d'extension est en moyenne supérieure à l'erreur-vraie de la méthode linéaire en C², le tableau précédent montre que la méthode d'extension présente beaucoup plus souvent de grosses erreurs vraies et quelquefois même de très grandes. Il suffit de songer que le quintuple de l'erreur-type ne peut être dépassé, par la méthode linéaire en C², que moins d'une fois sur 100 000, et que ceci a été effectivement dépassé 2 fois sur 50 par la méthode d'extension. Enfin, le double de l'erreur-type ne doit être dépassé en moyenne qu'une fois sur 20 dans la méthode linéaire en C², et a été dépassé 14 fois sur 50 dans la méthode d'extension (contre 3 fois sur 50 dans la méthode linéaire en C²).

TITRE VII

VERIFICATION DE LA FORMULE
DONNANT L'ERREUR-TYPE

1° On a choisi, dans les places disponibles, celle où la population totale est la plus nombreuse : Haguenau, présentant un volume bois fort-tige de 133 535 dm³ pour 367 arbres.

La méthode de vérification est la suivante : 50 échantillons ont été constitués au hasard, le volume étant désigné pour chacun par V_E.

Diverses populations ont été considérées :

la population A, de volume V_A = 133 535 dm³, constituée par l'ensemble des 367 arbres ;

les populations B, de volumes V_B = V_A — V_E, chacune constituée en ôtant de la population A les arbres d'un échantillon.

Ainsi, les populations A correspondent à l'hypothèse 1 du Titre I - Section III, §2 (l'échantillon fait partie de la population à cuber) et les populations B à l'hypothèse 2 du même texte (l'échantillon ne fait pas partie de la population à cuber).

On a calculé pour chaque échantillon les estimations du volume par la régression a + bC² (voir Titre V) : ce sont les volumes-tarif V_{tA}, V_{tB}. Les erreurs-types ET % (ramenées en pourcentage du volume tarif) ont également été calculées, ainsi que les erreurs-vraies, ramenées en pourcentage du volume réel :

$$EV_A \% = \frac{(V_A - V_{tA}) 100}{V_A}$$

$$EV_B \% = \frac{(V_B - V_{tB}) 100}{V_B}$$

Le tableau suivant donne les nombres essentiels, relatifs à la population A, ainsi que les rapports

(voir tableau n° 17).

La méthode de tirage au hasard est la suivante, que l'on exposera sur deux exemples :

— pour les échantillons 1 à 4, l'échantillon comprend, à une unité près, 1/15 des arbres de la population. Celle-ci a donc été ordonnée par circonférences croissantes, puis divisée en tranches de 15 arbres. Un arbre a été pris au hasard dans chacun des tranches (utilisation des *random numbers* de FISHER and YATES) ;

TABLEAU n° 17

N = 367

V_A = 133 535

N° de l'échan- tillon	Nombre de tiges de l'échan- tillon	Volume de l'échantillon	Volume tarif A V _{TA}	Erreur type E T	Erreur vraie E V	E T A %	E V A %	
							E T A %	E T A %
1	25	10 041	137 960	2 526,3	— 4 425	1,83	—	1,81
2	25	9 427	132 430	2 053,7	1 105	1,55	—	0,53
3	24	8 447	133 822	2 129	— 287	1,59	—	0,13
4	25	9 234	132 090	3 150,8	1 445	2,38	—	0,45
5	31	11 290	132 080	1 308	535	0,98	—	0,42
6	31	11 535	134 148	2 316	— 613	1,72	—	0,26
7	31	10 569	134 638	2 387	— 1 103	1,77	—	0,46
8	31	11 091	127 238	4 123	6 297	3,24	—	1,45
9	37	13 236	130 016	1 869	3 519	1,43	—	1,84
10	37	13 276	130 475	2 858	3 060	2,19	—	1,05
11	37	13 485	133 201	2 541	334	1,90	—	0,13
12	36	12 497	131 507	2 808	2 028	2,18	—	0,69
13	37	13 460	132 710	1 813	825	1,36	—	0,45
14	41	14 657	131 035	2 507	2 500	1,91	—	0,98
15	40	13 664	134 459	1 347	— 924	1,00	—	0,69
16	41	14 833	132 471	2 114	1 064	1,59	—	0,50
17	41	15 317	136 266	1 357	— 2 731	0,99	—	2,06
18	46	16 460	131 899	2 179	1 726	1,65	—	0,78
19	46	17 122	136 492	1 658	— 2 957	1,21	—	1,83

21	40	16 660	133 264	2 201	271	1,65	0,12
22	53	19 569	133 982	2 031	— 447	1,51	— 0,22
23	52	19 240	139 049	2 905	— 5 514	2,13	— 1,93
24	52	18 186	132 587	2 242	948	1,69	0,41
25	52	18 184	132 586	2 236	949	1,68	0,42
26	61	21 800	131 630	1 990	1 905	1,51	0,94
27	61	22 485	136 795	1 385	— 3 260	1,01	— 2,42
28	61	22 020	133 606	1 553	— 131	1,16	— 0,08
29	61	22 107	133 998	1 299	— 463	0,96	— 0,35
30	74	27 489	136 788	3 534	— 3 253	2,58	— 0,94
31	73	26 315	134 050	1 668	— 515	1,26	— 0,30
32	73	26 076	132 750	1 584	785	1,19	0,49
33	73	25 765	130 939	1 388	2 596	1,06	1,83
34	73	26 618	135 963	1 435	— 2 428	1,05	— 1,72
35	92	33 987	134 362	1 120	— 827	0,83	— 0,73
36	92	33 406	132 932	1 167	603	0,87	0,52
37	92	33 326	132 453	1 075	1 082	0,81	1,00
38	91	31 879	130 818	1 253	2 717	0,95	2,14
39	111	41 061	133 533	944	2	0,70	0,002
40	111	40 800	132 700	1 046	835	0,78	0,70
41	111	40 909	133 164	1 146	371	0,86	0,31
42	111	41 371	134 695	1 085	— 1 160	0,80	— 1,08
43	122	43 819	133 459	991	76	0,74	0,08
44	122	44 203	134 266	924	— 731	0,68	— 0,79
45	122	44 153	133 782	925	— 247	0,69	— 0,26
46	123	45 238	133 280	1 085	246	0,81	0,22
47	148	54 462	133 311	910	224	0,68	0,24
48	148	54 655	132 758	970	777	0,73	0,79
49	148	54 418	132 803	798	732	0,60	0,90
50	148	54 042	132 133	1 019	1 402	0,77	1,35

— pour les échantillons 47 à 50, l'échantillon comprend 4/10 des arbres. La population a été divisée en tranches de 10 arbres et 4 ont été pris au hasard dans chaque tranche (utilisation des *random permutations of 10 numbers* de FISHER and YATES).

2° Il serait bon, et on peut l'espérer d'après ce qui a été vu au Titre II. que le rapport

$$\frac{EV_A \%}{ET_A \%}$$

suive la loi normale réduite. Pour le vérifier, on divise l'intervalle ($-\infty, +\infty$) en 10 intervalles possédant la propriété suivante: la probabilité pour qu'une variable suivant la loi normale réduite tombe dans un de ces intervalles fixés d'avance, est la même (1/10) pour chacun d'eux. Les limites de ces intervalles sont données par CRAMER, « *Mathematical Methods of Statistics* », table 2 (P. 558), et sont indiquées comme limites d'intervalle du tableau suivant:

TABLEAU n° 18

— ∞		— 1,2816		— 0,8416		— 0,5244		— 0,2533	0
	6		2		4		5		3
0	+ 0,2533		+ 0,5244		+ 0,8416		+ 1,2816		+ ∞
	6	9	5	5	5				

Dans chaque intervalle est indiqué le nombre de fois où

$$\frac{EV_A \%}{ET_A \%}$$

tombe dans celui-ci. Le nombre théorique est 5 dans chaque intervalle. On calcule χ^2 et l'on trouve $\chi^2 = 6,6$ pour 9 degrés de liberté, ce qui correspond à une probabilité de 68 %.

On peut penser que le nombre théorique 5, est un peu petit. En groupant les colonnes deux à deux par contiguïté, on trouve $\chi^2 = 2,2$ pour 4 degrés de liberté, soit une probabilité de 70 %. Enfin, on peut encore grouper les colonnes deux par deux symétriquement autour de l'axe central, ce qui revient à ne pas tenir compte du signe. On trouve $\chi^2 = 2,8$ à 4 degrés de liberté, soit une probabilité de 59 %.

En conclusion, au degré de précision atteint par cette vérification, il n'y a aucune raison de douter que la distribution de

$$\frac{EV_A \%}{ET_A \%}$$

ne soit normale.

Il en est de même de celle de

$$\frac{EV_B \%}{ET_B \%}$$

(les détails ne seront pas reproduits ici), car pour chaque échantillon, les rapports correspondant aux populations A et B sont presque identiques.

3° La formule indiquée pour le calcul de l'erreur est donc d'un emploi pratique certain dans chacun des cas extrêmes correspondant aux hypothèses 1 et 2 du Titre I, Section 3, § 2. Dans le cas intermédiaire de l'hypothèse 3, la formule est encore certainement valable. Dans le cas d'un échantillon dont la partie détruite n'est pas également répartie sur l'ensemble des circonférences (hypothèse 4), la formule n'est plus rigoureusement valable. Comme dans la plupart des cas réels la partie détruite est en majorité dans les circonférences faibles, de petites variances, le résultat est une surestimation de l'erreur-type, qui doit ainsi être légèrement inférieure à ce que donne le calcul.

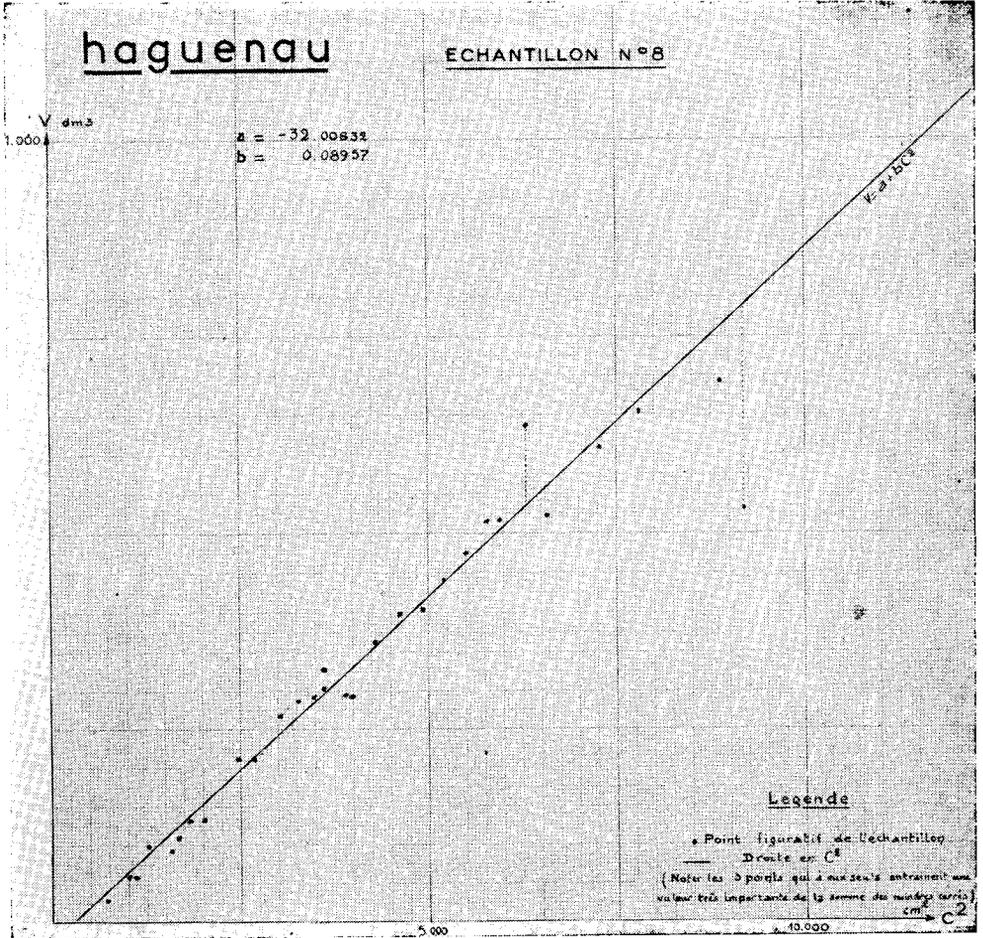
4° On remarquera qu'à une même fraction d'échantillonnage correspond des ET % fort divers. Ainsi, pour la fraction $\frac{1}{12}$, on trouve 0,98 % pour l'échantillon 5, et 3,24 % pour l'échantillon 8. Les deux graphiques suivants (V en fonction de C²) font comprendre le phénomène: voir graphiques n° 3 et 3 bis.

L'alignement des points est presque parfait pour l'échantillon 5. En conséquence, la somme des carrés des écarts est petite:

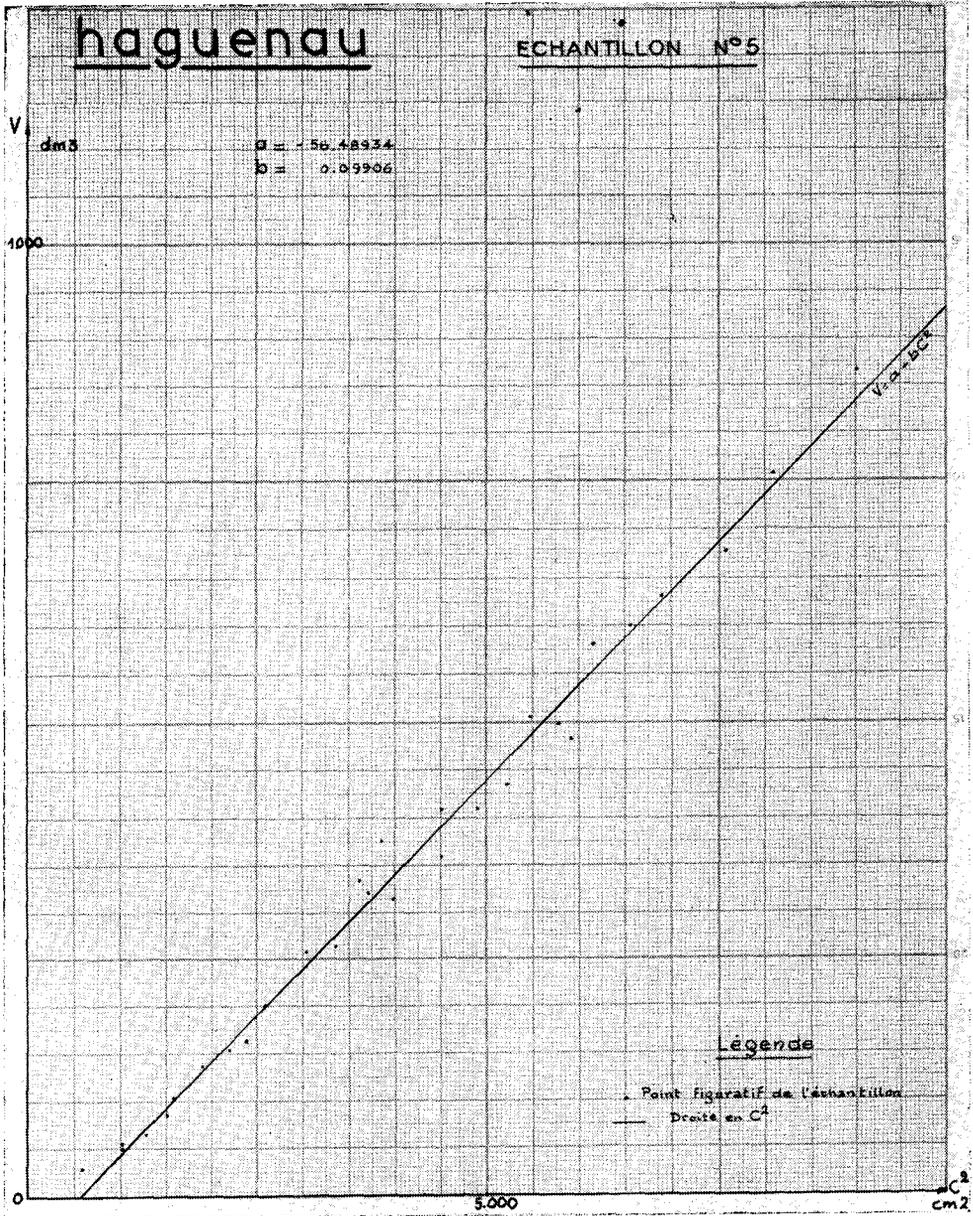
$$\sum (V - V_t)^2 = 12\,470$$

Pour l'échantillon 8, trois points s'écartent considérablement de la droite théorique. La somme des carrés des écarts est 123 916, sur lesquels les trois arbres aberrants donnent 110 061. Les 28 autres arbres n'interviennent donc que pour 13 850.

En conséquence, il est clair que, s'il est normal d'espérer augmenter la précision en augmentant la fraction d'échantillonnage, cet espoir ne sera pas fatalement exaucé: on peut en effet, en augmentant l'échantillon, tomber sur quelques arbres aberrants, avec pour résultat une augmentation de l'erreur-type.



Graphique 3



Graphique n° 3 bis.

TITRE VIII

ESTIMATION DES VOLUMES
PAR QUALITES DE PRODUITS

SECTION I

ÉTUDE DES CLASSIFICATIONS EN 1, 2, 3; A, B, C,
POUR CHAQUE CATÉGORIE DE PRODUITS

1° L'intérêt de ces classifications s'étudie, pour chaque catégorie de produits, comme l'on a étudié ce même intérêt pour le volume total (voir Titre III). On a vu à ce dernier propos, que la simple étude des signes \pm des différences des volumes totaux moyens d'une même circonférence entre deux classes (1A — 1B par exemple) d'une même place conduit à des résultats comparables à ceux de méthodes plus fines. On se limitera donc, ici, à l'étude de ces signes en séparant chaque catégorie de produits.

2° Voici les résultats du pointage :

TABLEAU n° 19

Qualité	Comparaison A - B		Comparaison 1 - 2	
	A l'emporte	B l'emporte	1 l'emporte	2 l'emporte
Choix III et Choix IV.	17 fois	33 fois	31 fois	10 fois
Mine	17	10	7	16
Râperie	43	49	36	20
Quartier	23	23	7	2
Traverse	7	5	4	0
Poteau	0	2	0	1
Défibrage	7	8	4	0

3° La méthode du χ^2 a été employée pour les catégories de produits présentant assez de circonférences communes; pour certaines autres, on a utilisé la théorie de la distribution binominale; enfin la qualité « Poteau », qui ne présente pas assez de comparaisons a

été laissée de côté: voici le tableau des probabilités correspondantes (expression en %):

TABLEAU n° 20

Qualité	A - B	I. - 2
Choix III et choix IV	36 %	0,12 %
Râperie	53 %	3,2 %
Quartier	100 %	10 %
Traverse	58 %	6,25 %
Défilage	100 %	6,25 %

4° Il résulte clairement de là :

a) qu'il existe une différence significative entre les classes 1 et 2;

b) qu'aucune différence n'a pu être mise en évidence entre les classes A et B: ceci n'implique pas qu'il n'y ait pas de différence, mais simplement que, si différence il y a, celle-ci n'est pas assez importante pour être mise en évidence avec le nombre relativement petit de comparaisons dont on dispose en ce moment (1).

SECTION 2

CALCUL DES VOLUMES PAR QUALITÉS DE PRODUITS

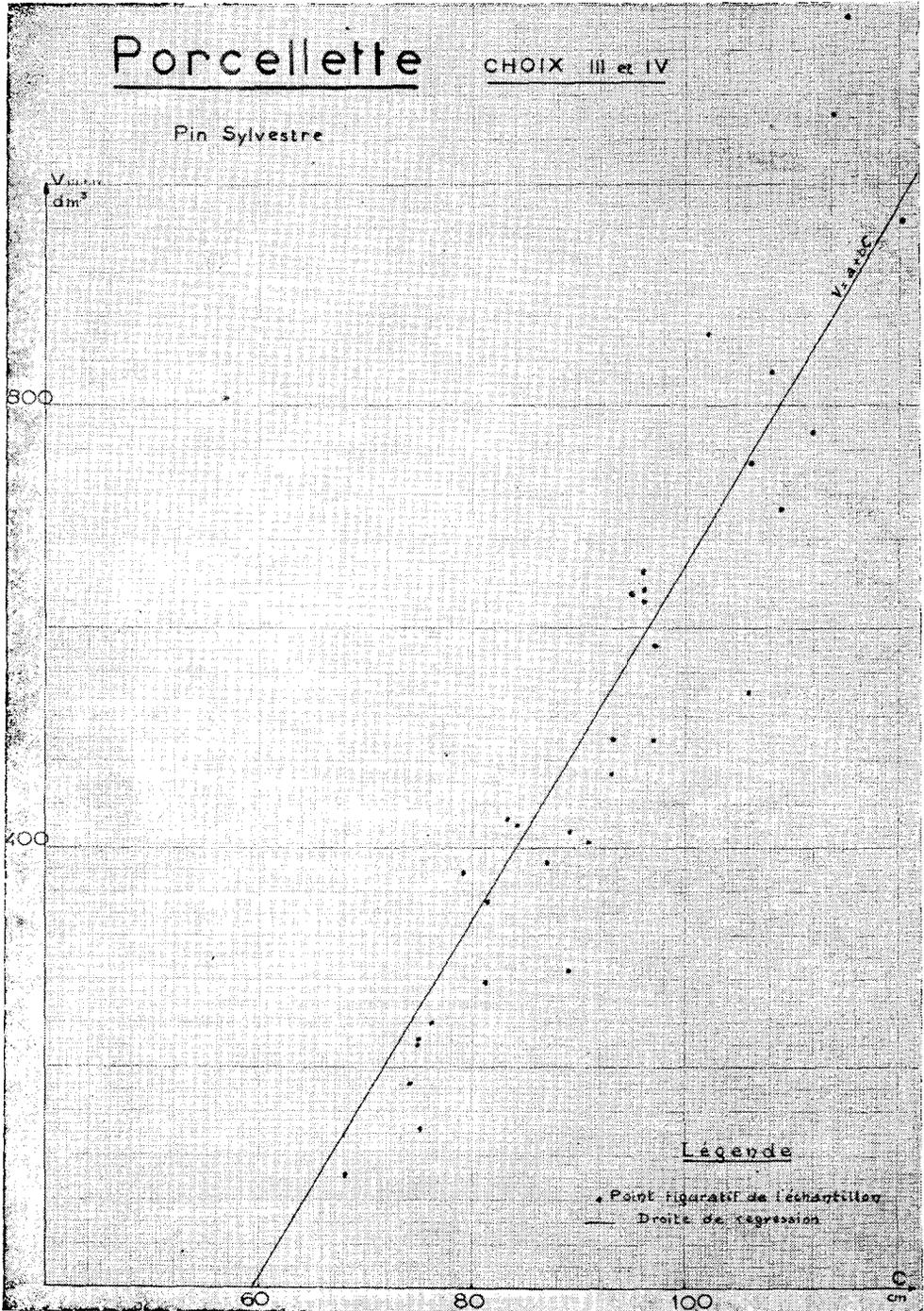
1° Le problème est ici de trouver le volume total que contient une place dans une qualité de produits (charpente par exemple). Les méthodes de régression adoptées dans le cas du volume total peuvent encore l'être ici. Toutefois, il est à remarquer que, pour une circonférence donnée, la distribution du volume n'est pas gaussienne pour une qualité de produit déterminée. On appliquera cependant les méthodes usuelles pour le calcul de l'écart-type, et l'on vérifiera expérimentalement si l'on peut adopter la règle suivante: *il est « rare » que le volume réel diffère du volume estimé de plus du double de l'écart-type*. Cependant, il restera difficile d'estimer cette « rareté ».

(1) On notera que la comparaison n'a pas pu être réalisée dans les mêmes conditions pour les meilleures qualités (Choix I et II), faute de données suffisantes.

On doit le regretter car, pour les feuillus par exemple, on peut penser qu'un arbre à cime développée sera souvent plus bas branchu qu'un arbre à cime normale ou étriquée. Dans ce cas, à circonférences égales, il est vraisemblable que la découpe au premier nœud, sinon au premier défaut, sera plus basse et le volume de bois de qualité plus faible.

Il sera intéressant d'effectuer des comparaisons lorsque l'on aura pu recueillir des données suffisantes (dernières coupes de régénération).

P. ARBONNIER.



Graphique 4

2° On a commencé par étudier le volume-charpente (Choix III + Choix IV) avec les méthodes suivantes (C désigne la circonférence, mesurée à 1,30 m, en cm, et V le volume correspondant, en dm^3):

$$\text{a) } V = \alpha + \beta C + \gamma C^2,$$

avec la restriction $V = 0$ pour $C = 59$ (1)

$$\text{b) } V = \alpha + \beta C + \gamma C^2,$$

sans restriction.

Dans chaque cas, il n'a été tenu compte que des tiges de circonférence ≥ 60 . Les détails de calcul (équations normales, etc...) ne seront pas donnés, car de meilleurs procédés ont été trouvés ultérieurement.

Voici les erreurs-types et erreurs-vraies sur Haguenau:

TABLEAU n° 20

Méthode	Erreur-type	Erreur-vraie
a	2,47 %	— 4,43 %
b	2,29 %	— 4,63 %

Les deux méthodes donnent donc des résultats identiques.

3° On a finalement adopté pour le volume charpente, la loi simple:

$$V = a + b C$$

qui semble graphiquement assez bien confirmée (voir graphique n° 4 ci-contre, Porcelette). Il n'a été tenu compte que des tiges de circonférences ≥ 60 . Les équations normales sont:

$$\begin{cases} na + b \sum C = \sum V \\ a \sum C + b \sum C^2 = \sum VC, \end{cases}$$

les sommations étant étendues aux tiges de l'échantillon dont la circonférence est ≥ 60 , et n étant le nombre de ces tiges. L'erreur-type est donnée par

$$\frac{N}{n} \sqrt{Q}$$

avec
$$Q = \frac{n}{n-2} \frac{N + n_2 - n_1}{N} (\sum v^2 - a \sum v - b \sum vc),$$

(1) En effet, la qualité « charpente » n'englobe que des grumes ayant plus de 60 cm de circonférence au fin bout pour les résineux.

les sommations étant les mêmes que ci-dessus, N désignant le nombre de tiges de circonférences ≥ 60 dans la population restant sur pied, n_2 et n_1 les nombres de tiges abattues ou restant sur pied, de circonférence ≥ 60 , dans l'échantillon. L'estimation du volume est :

$$Na + b\Sigma C,$$

la sommation étant étendue cette fois aux tiges de circonférence ≥ 60 de la population restant sur pied.

Les résultats sont consignés dans le tableau suivant ; la dernière colonne est relative à la méthode consistant à appliquer à la population restant sur pied le pourcentage de bois de charpente trouvé dans l'échantillon.

TABLEAU n° 21

Espèce	Place	Erreur-type	Erreur-vraie	Erreur-vraie (méthode de pourcentage)
Epicéa	Haguenau	2,30 %	— 4,70 %	— 6,40 %
	St-Avoid 142	7,56 %	8,80 %	— 1,79 %
	St-Avoid 156	4,49 %	6,21 %	7,74 %
Pin sylvestre.	St-Avoid 134 b	15,42 %	— 13,78 %	— 62,95 %
	Porcelette	3,15 %	4,03 %	2,15 %

Une étude spéciale est faite, pour St-Avoid 156, au § suivant.

La « méthode de pourcentage » fournit rapidement un résultat : mais, même avec un échantillonnage soigné, elle peut donner de grosses erreurs (voir St-Avoid 134 b) ; en outre, elle ne permet pas le calcul de l'erreur-type.

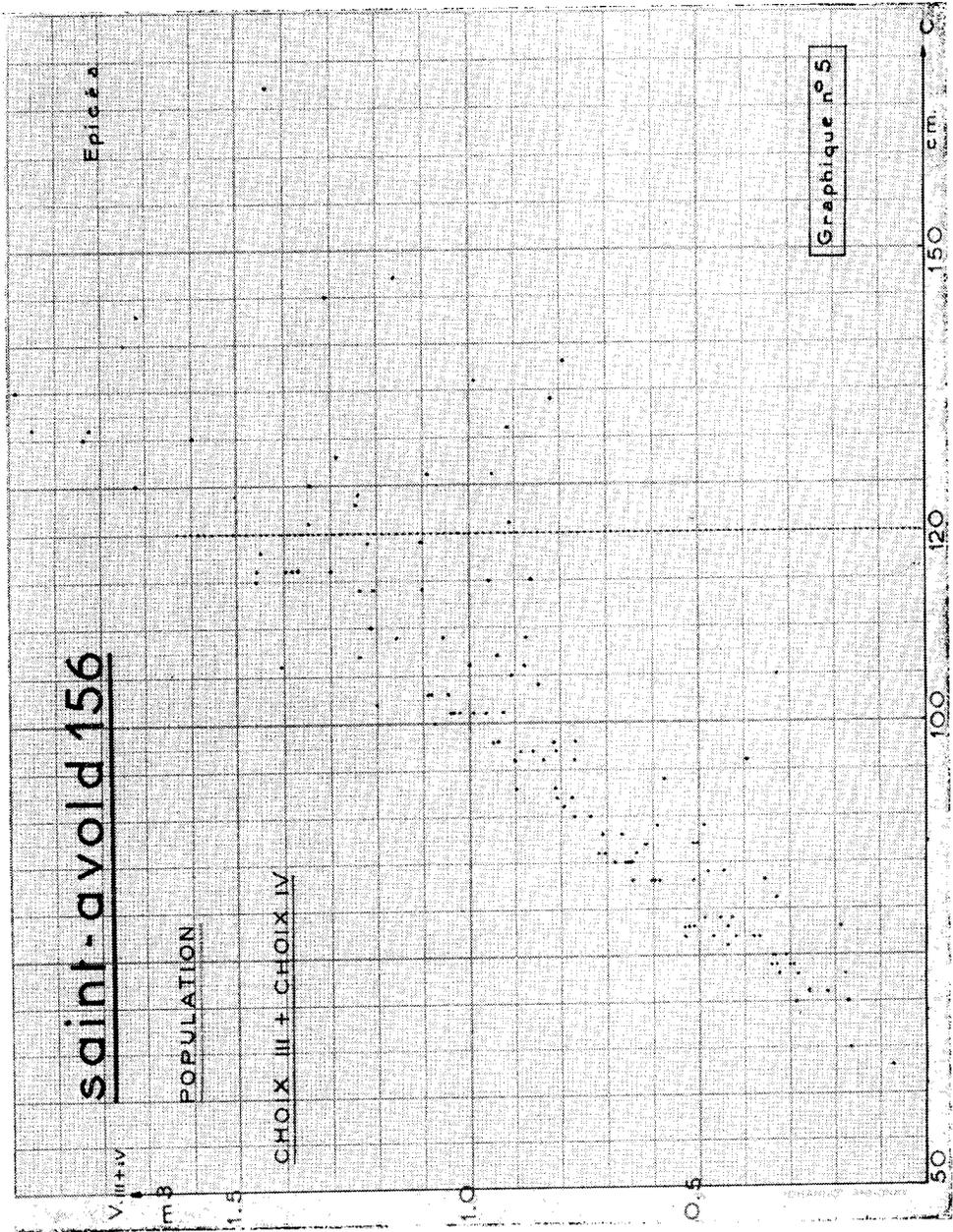
La même étude a été appliquée à la qualité « traverse » avec $C \geq 70$ (il s'agit ici de feuillus) :

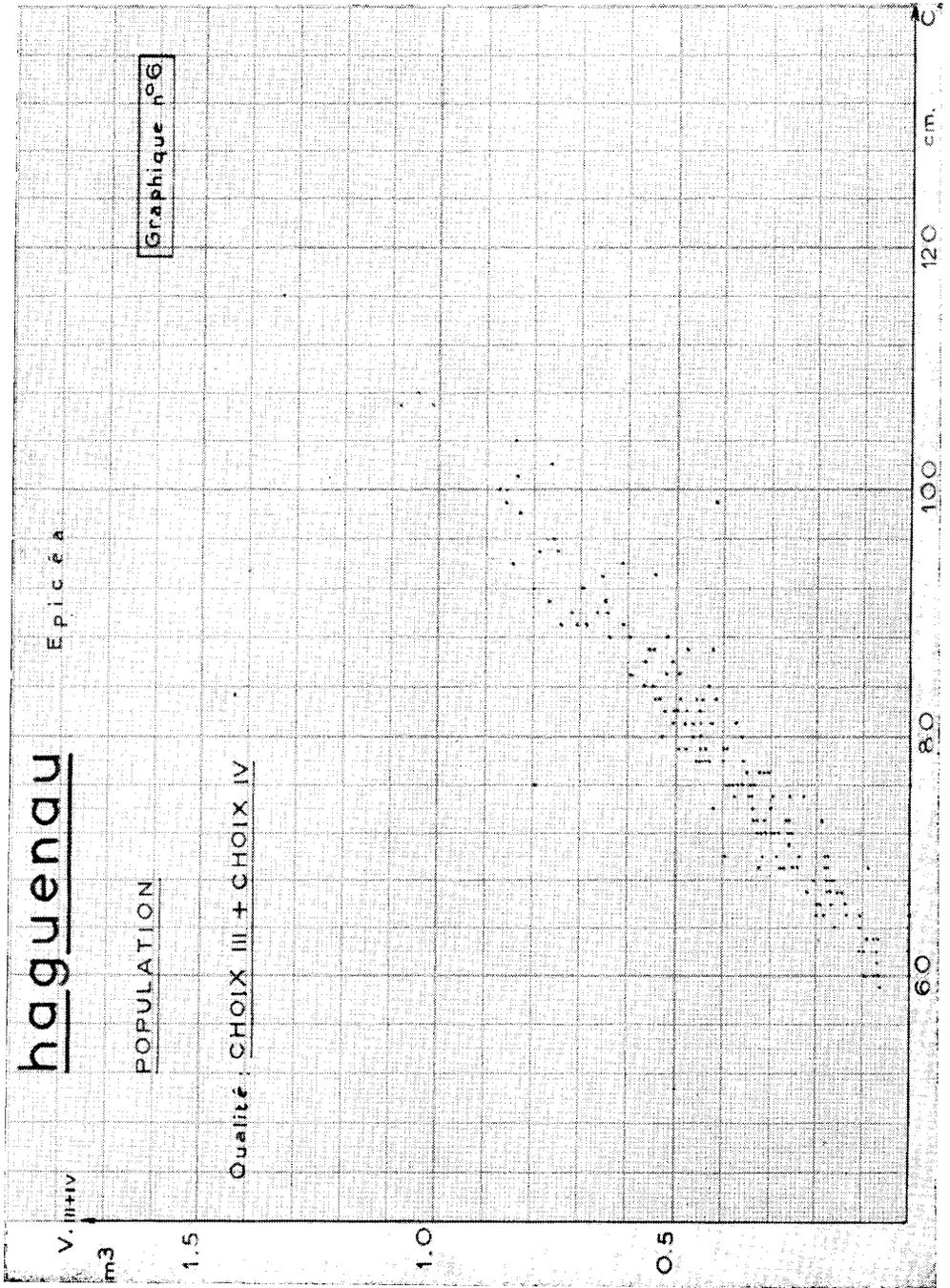
TABLEAU n° 22

Espèce	Place	Erreur-type	Erreur-vraie	Erreur-vraie méthode de pourcentage)
Hêtre	St-Amond 46	8,09 %	— 7,41 %	— 5,44 %
Hêtre	St-Amond 47	3,87 %	— 3,57 %	— 2,72 %

4° Pour St-Avoid 156, Charpente, le graphique 5 représentant V en fonction de C est le suivant (graphique sur *population*) ; le graphique 6 est relatif à Haguenau, à titre de comparaison ; il s'agit chaque fois d'épicéas.

(Voir graphiques n° 5 et 6).





On constate sur le graphique (5) un changement de loi à la circonférence 120 avec très forte dispersion pour $C \geq 120$. Or, dans cette place, il existe pour $C \geq 120$, une qualité supérieure : le bois de menuiserie. C'est à l'apparition de cette qualité supérieure qu'est lié le changement de loi observé.

On a donc été conduit à prendre :

$$V = \alpha' + \beta' C \text{ pour } 60 \leq C < 120$$

$$V = \alpha'' + \beta'' C \text{ pour } C \geq 120,$$

avec la condition que, pour $C = 120$, les deux lois donnent le même volume.

Les équations normales s'obtiennent alors en minimisant :

$$\sum_{60 \leq C < 120} (v - \alpha' - \beta' c)^2 + \sum_{C \geq 120} (v - \alpha'' - \beta'' c)^2$$

sous la condition :

$$\alpha' + 120 \beta' = \alpha'' + 120 \beta''$$

La méthode des multiplicateurs de LAGRANGE conduit aux équations :

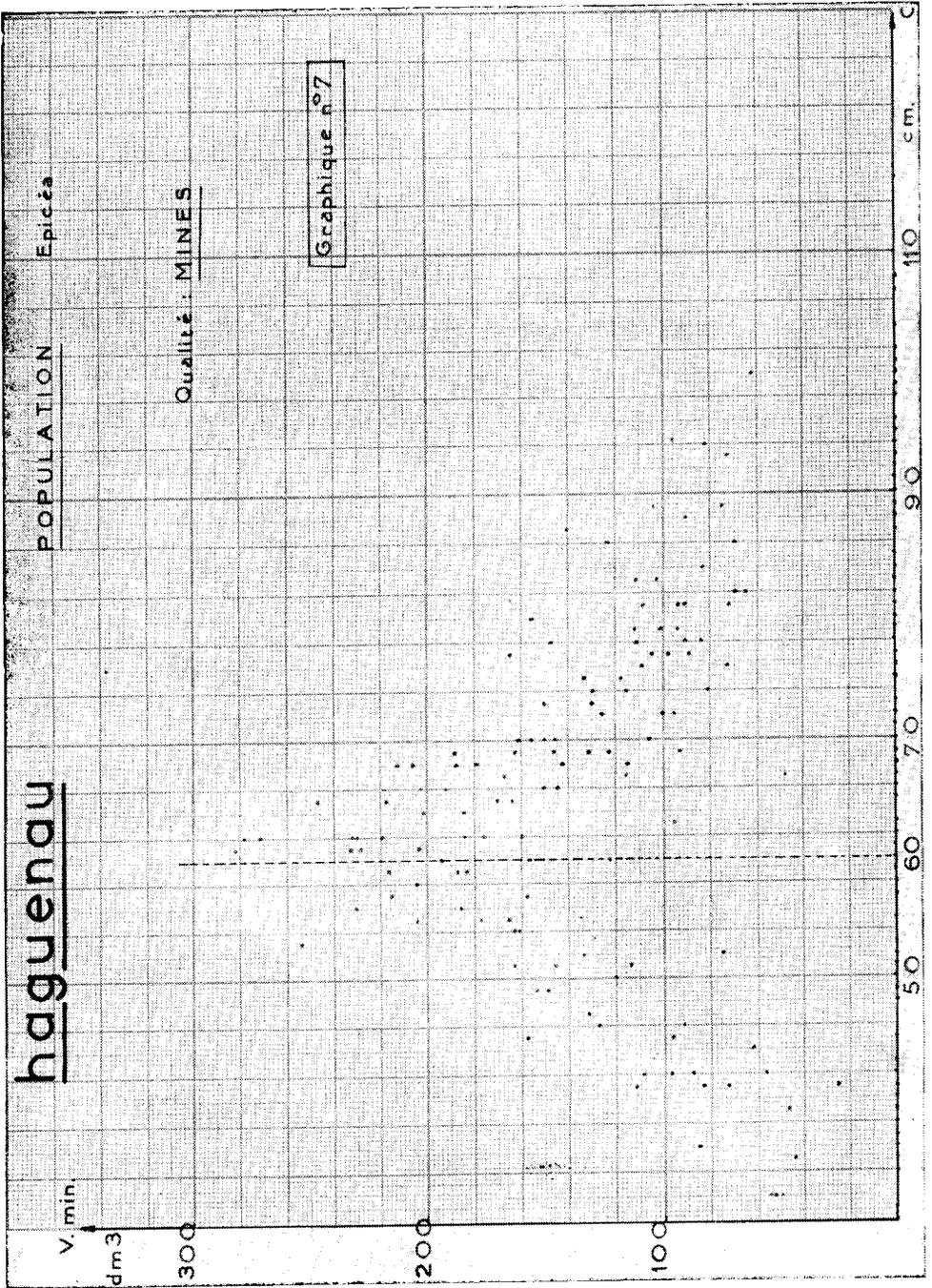
$$\left\{ \begin{array}{l} n' \alpha' + \beta' \sum' c + \lambda = \sum' v \\ \alpha' \sum' c + \beta' \sum' c^2 + 120 \lambda = \sum' v c \\ n'' \alpha'' + \beta'' \sum'' c - \lambda = \sum'' v \\ \alpha'' \sum'' c + \beta'' \sum'' c^2 - 120 \lambda = \sum'' v c \\ \alpha' + 120 \beta' = \alpha'' + 120 \beta'' \end{array} \right.$$

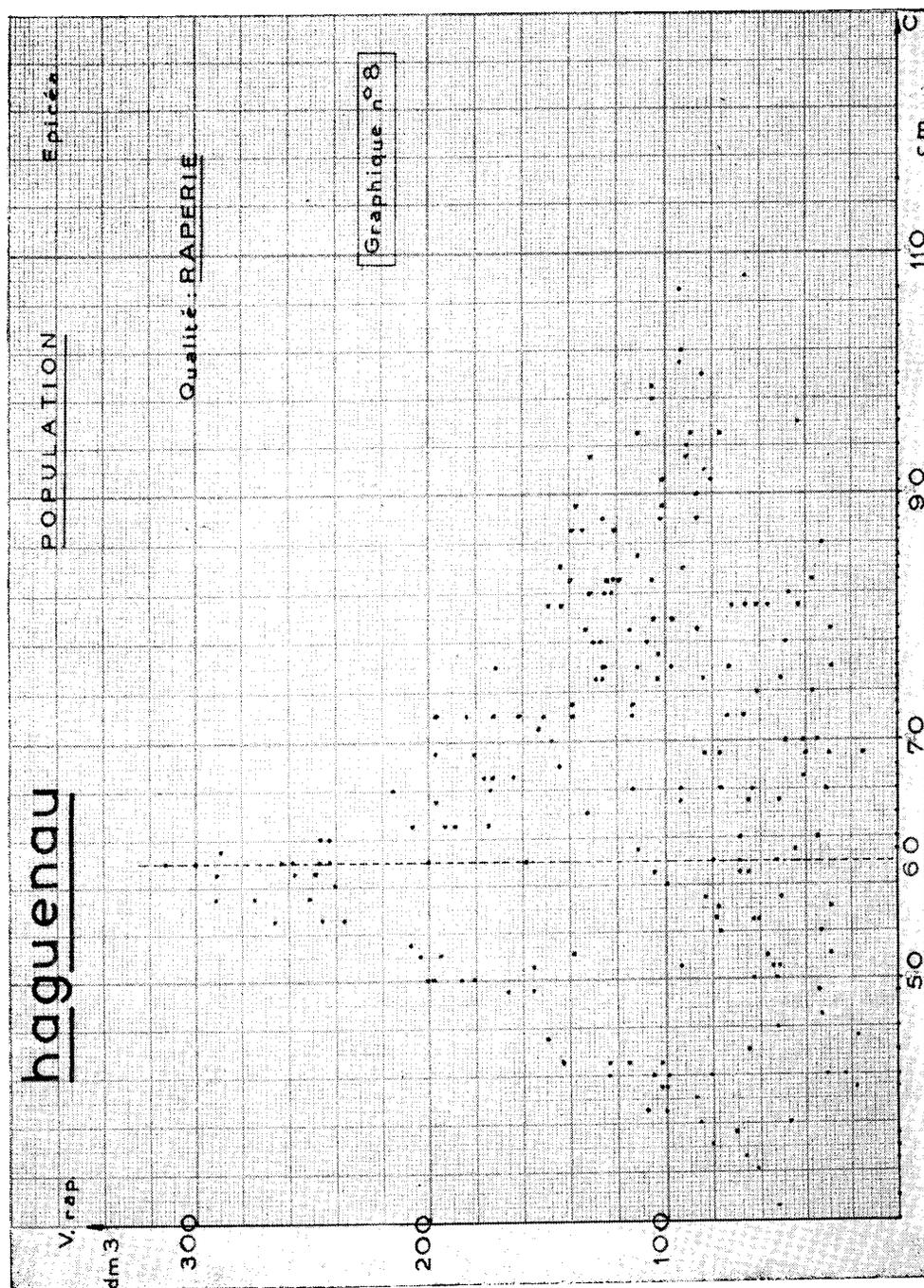
Les sommations sont étendues à l'échantillon, \sum' indique que la sommation ne se fait que pour $60 \leq C < 120$, \sum'' qu'elle ne se fait que pour $C \geq 120$; n' et n'' sont les nombres de tiges de l'échantillon dans les deux intervalles précités; λ est une inconnue auxiliaire introduite par la méthode des multiplicateurs, dont la valeur importe peu, et que l'on peut en conséquence, éliminer au cours des calculs.

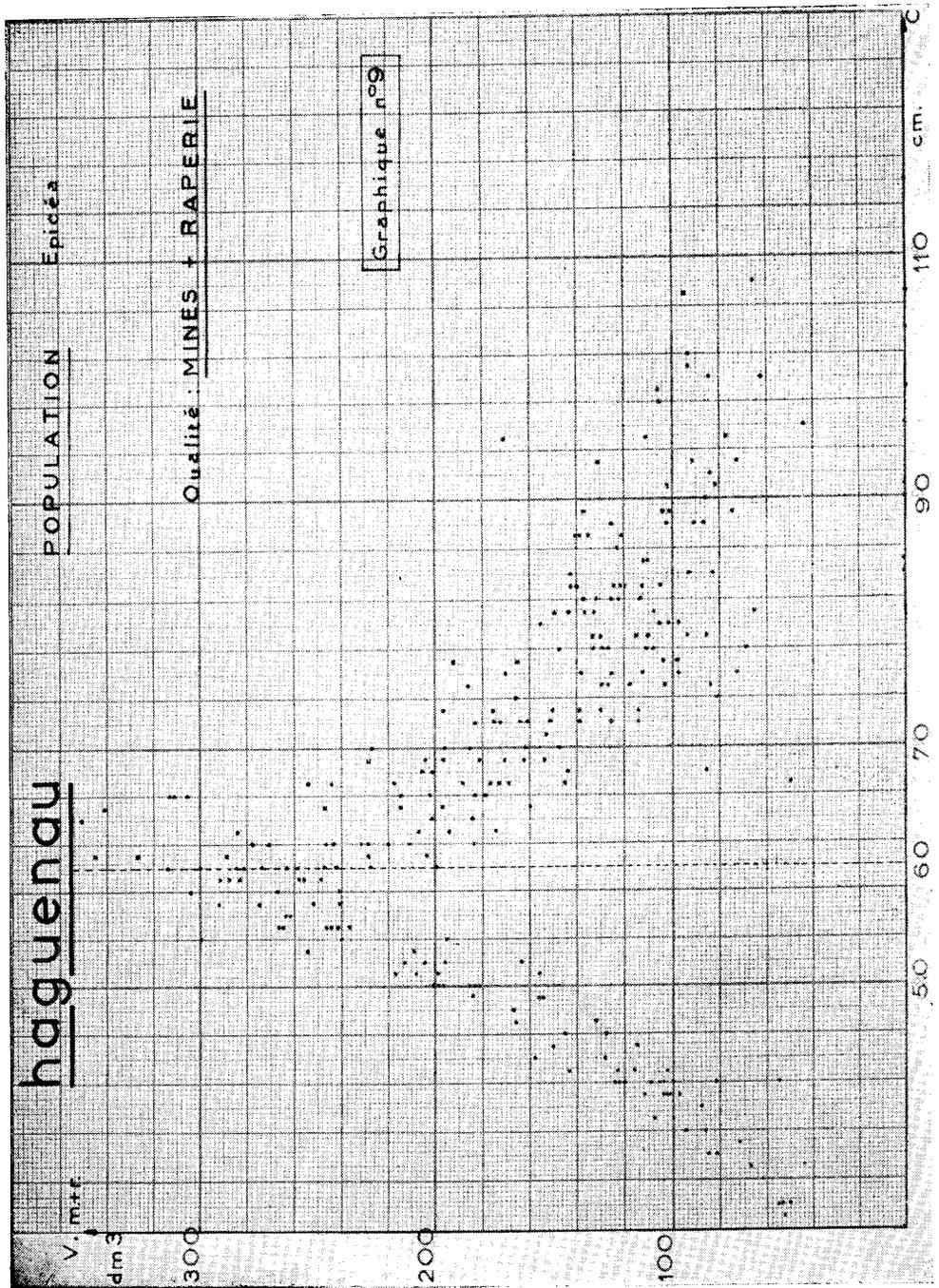
La résolution de ce système est loin d'être aussi rebutante qu'il ne le paraît, et se ramène à celle de deux systèmes linéaires, chacun à deux inconnues, assortie de quelques calculs simples (voir annexe 10 un exemple de calcul de bidroite).

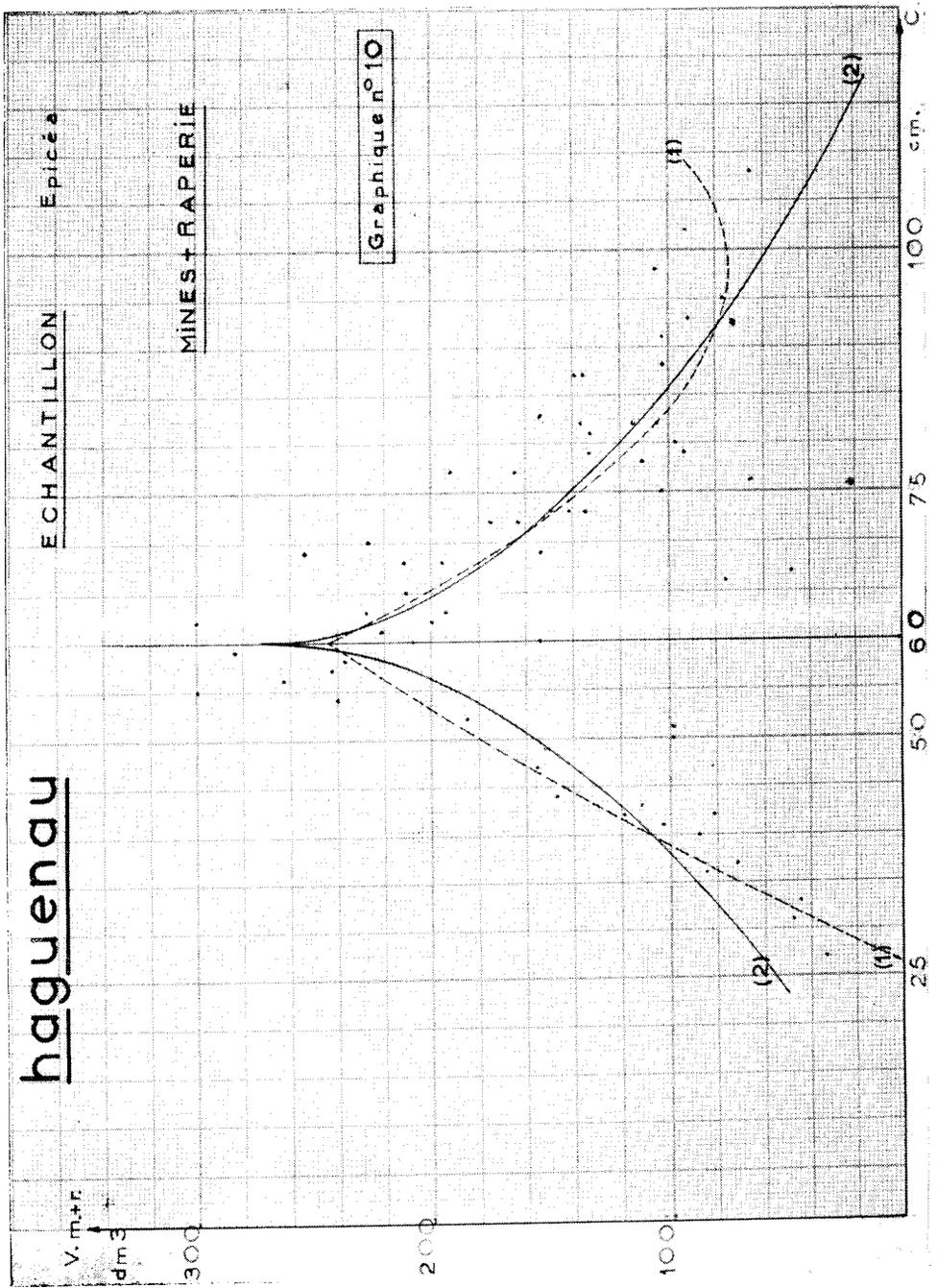
L'erreur-type est donnée par :

$$\frac{N}{n} \sqrt{\sigma}$$









$$Q = \frac{n}{n-3} \frac{N+n_2-n_1}{N} (\sum' v^2 + \sum'' v^2 - \alpha' \sum' v - \beta' \sum' vc - \alpha'' \sum'' v - \beta'' \sum'' vc)$$

les sommations étant les mêmes que ci-dessus, et n , n_2 , n_1 , N ayant les mêmes spécifications qu'au § 3.

L'estimation du tarif est :

$$N' \alpha' + \beta' \sum' c + N'' \alpha'' + \beta'' \sum'' c,$$

les sommations étant cette fois étendues à la population restant sur pied, avec N' tiges pour $60 \leq C < 120$, et N'' tiges pour $C \geq 120$.

5° Les §§ 5, 6, 7 sont consacrés aux qualités de produits « mine » et « râperie ». On commencera par une introduction graphique, faite sur Haguenau (population totale avant éclaircie). Les graphiques (7) et (8) sont relatifs aux qualités « mine » et « râperie » respectivement; pour le graphique (9), il a été effectué, sur chaque tige, la somme des volumes en mine et en râperie, pour obtenir la qualité « mine + râperie ».

(Voir graphiques 7, 8 et 9).

a) On remarquera la très forte dispersion des points en (7) et (8), contre une dispersion nettement moindre en (9); il y a en outre, en (7) et (8), un « phénomène de zéro » beaucoup plus marqué qu'en (9): bien des tiges ont un volume nul en mine, ou un volume nul en râperie, avec cependant un volume non nul en « mine + râperie », l'une des qualités excluant en quelque sorte l'autre, au moins en très grande partie.

b) On remarquera également le changement de loi vers $C = 60$. Ce phénomène est analogue à celui qui a été signalé au § 4: vers $C = 60$ commence à apparaître la qualité supérieure « charpente ». On a ainsi été conduit à considérer des lignes de régression d'équations différentes pour $C < 60$ et pour $C \geq 60$.

On a d'abord étudié le volume global « mine + râperie ». L'examen de la fig. 9 a inspiré tout d'abord la méthode suivante:

une parabole $v = \alpha' + \beta' c + \gamma' c^2$ pour $C < 60$;

une parabole $v = \alpha'' + \beta'' c + \gamma'' c^2$ pour $C \geq 60$,

avec la restriction que les volumes, pour $C = 60$, soient les mêmes. Il s'agit là de deux paraboles à axes verticaux, sécantes en A (courbe 1 du graphique 10, relatif à l'échantillon).

On a encore essayé (courbe 2 du graphique 10) un système de deux paraboles de même axe horizontal tangentes en B (pour $C = 60$).

Nous n'entrons pas dans le détail de ces méthodes, qui ne se sont guère révélées plus précises que celles que nous allons exposer maintenant et qui sont d'exécution beaucoup plus aisée. Voici toutefois les résultats obtenus, (1) désignant la méthode des paraboles sécantes à axes verticaux, et (2) la méthode des paraboles tangentes coaxiales :

TABLEAU n° 23

Essence	Place	Méthode	Erreur-type	Erreur-vraie	Erreur-vraie (méthode de pourcentage)
Épicéa	Haguenau	(1)	3,72 %	— 5,4 %	6,7 %
	Haguenau	(2)	4,11 %	5,8 %	
Pin sylvestre	St-Avold 134 b	(1)	5,19 %	— 3,9 %	0,6 %
	St-Avold 134 b	(2)	5,17 %	— 2,9 %	

7° Il a été essayé ensuite de représenter le volume par deux droites se coupant pour $C = 60$. Les calculs sont ceux qui ont été indiqués au § 4 (en remplaçant toutefois naturellement 120 par 60 dans les équations normales).

On est ainsi conduit à représenter la loi de variation de V en fonction de C par un système de deux droites en général (en abrégé une bidroite), et cela pour toutes les qualités autres que la charpente. Toutefois, l'inspection graphique de l'échantillon montre parfois une telle absence de loi qu'une droite horizontale semble suffire ; les calculs sont alors particulièrement simples : l'ordonnée de la droite est :

$$\bar{v} = \frac{\sum v}{n}$$

l'estimation du volume de la population étant

$$N\bar{v} = \frac{N}{n} \sum v,$$

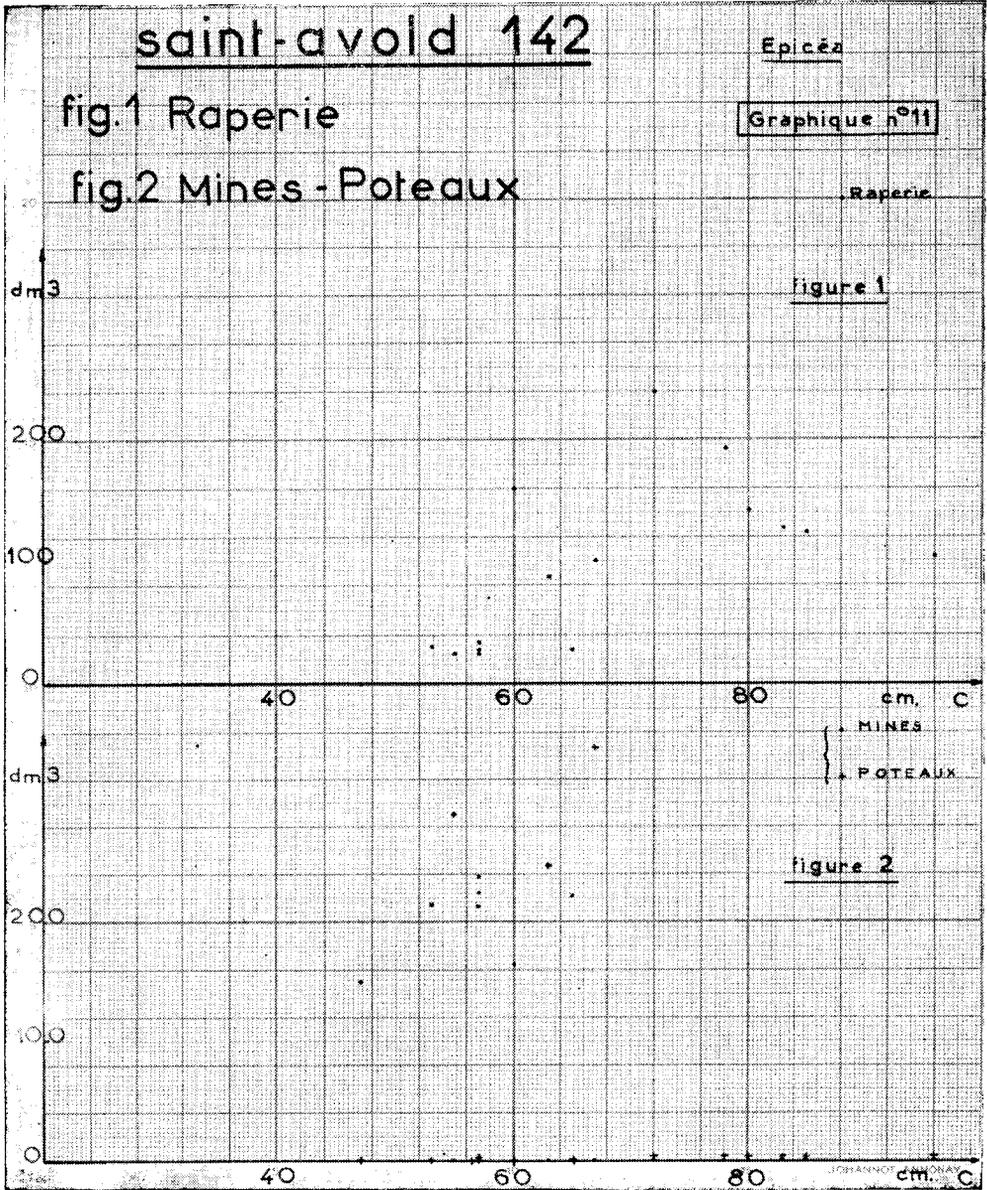
l'erreur-type est donnée par

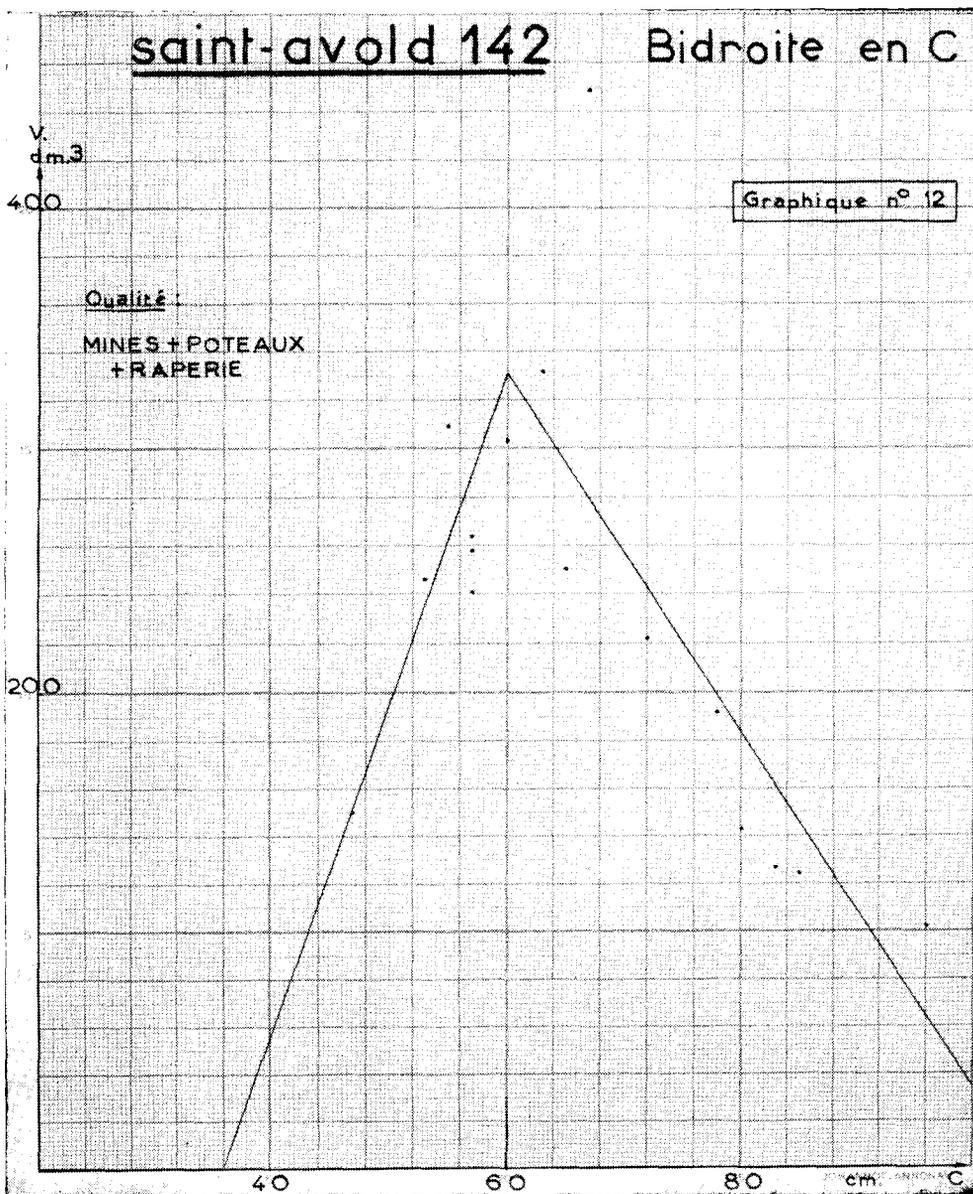
$$\frac{N}{n} \sqrt{Q}$$

avec

$$Q = \frac{n}{n-1} \cdot \frac{N + n_2 - n_1}{N} (\sum v^2 - n\bar{v}^2)$$

Les notations sont ici les mêmes qu'aux paragraphes précédents, et toutes les sommations sont étendues à l'échantillon. (Voir annexe 12.)





8° Un exemple particulièrement typique est celui de St-Avold 142 (épicéa). Les figures (1) et (2) (graphique n° 11) sont relatives aux qualités poteau, mine, râperie, respectivement; le graphique n° 12 correspond au groupement en une seule qualité appelée « poteau + mine + râperie ». Ce sont cette fois les *échantillons* (et non plus les populations comme au § 5) qui sont représentés.

(Voir graphiques n° 11 et 12).

Une loi est nettement visible sur le graphique 12. Sur les figures 1 et 2, la dispersion des points est très grande, le « phénomène de zéro » étant nettement en évidence sur la figure (2).

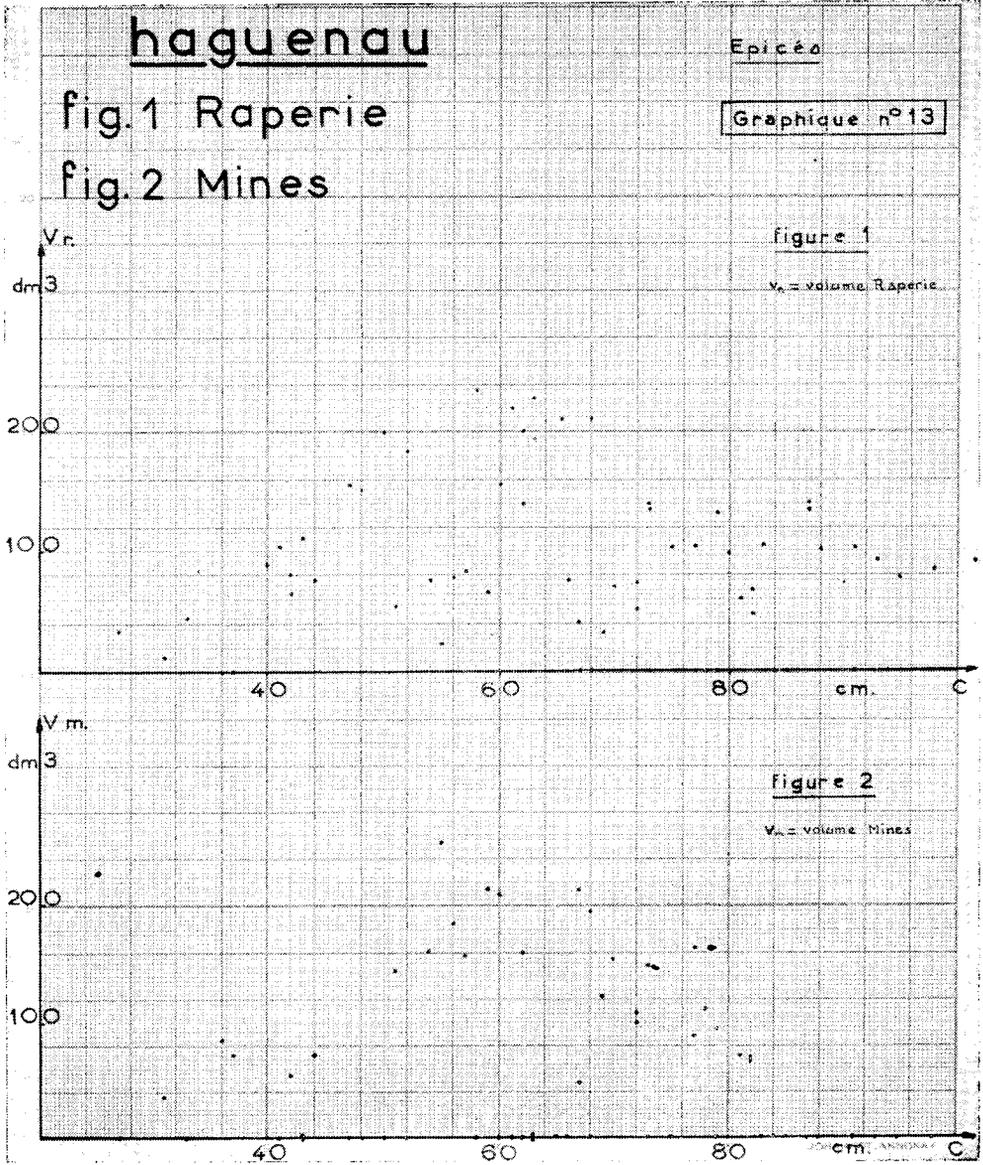
Les résultats des calculs sont consignés dans le tableau suivant :

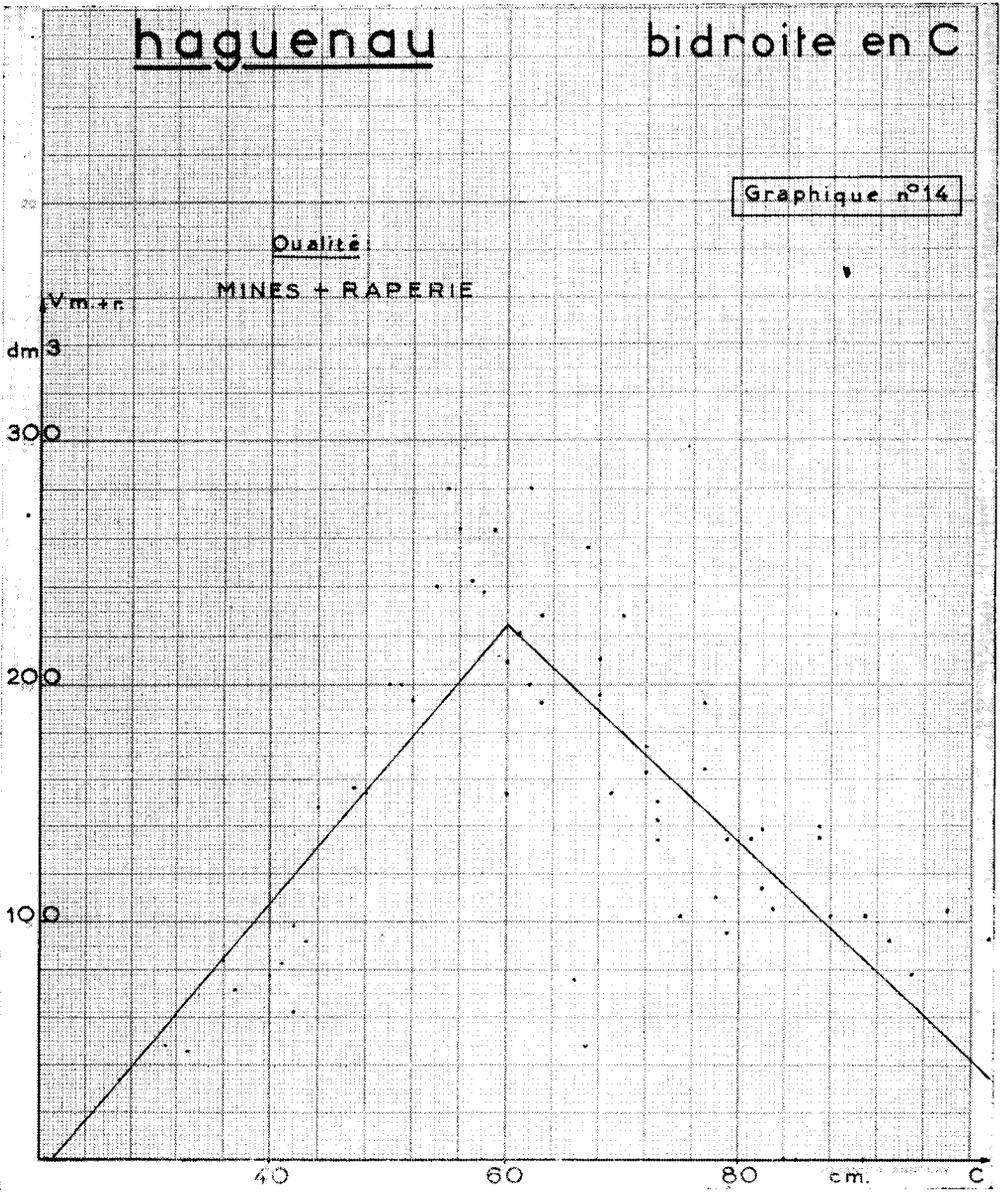
TABLEAU n° 24

Place	Qualité	Méthode	Erreur-type	Erreur-vraie
St-Avold 142	mine	bidroite	22,5 %	1 %
		horizontale	28,9 %	4 %
	râperie	bidroite	13,8 %	5,1 %
		horizontale	18,3 %	4,5 %
	poteau	bidroite	51,2 %	— 24,5 %
		horizontale	52,7 %	— 19,1 %
mine + râperie + poteau	bidroite	5,9 %	— 2,4 %	
horizontale	9,6 %	— 0,4 %		
Haguenau	mine	bidroite	11,5 %	14,8 %
		horizontale	16,5 %	16,3 %
	râperie	bidroite	8,7 %	— 1,4 %
		horizontale	9,2 %	— 0,5 %
	mine + râperie	bidroite	4,0 %	5,5 %
		horizontale	6,1 %	6,7 %
St-Amond 46	défibrage	bidroite	24,5 %	— 0,2 %
		horizontale	26,4 %	0,5 %

Il apparaît que, pour des qualités restreintes, la méthode de la bidroite ne présente pas (sauf pour St-Avold 142 - râperie, et peut-être mine, et aussi Haguenau - mine) d'avantage bien net sur la méthode de la droite horizontale; celle-ci, infiniment plus simple, doit donc, sauf contre-indication graphique très nette, être préférée.

Dans le cas des qualités « larges » (qualités restreintes groupées), les deux exemples où on a pratiqué les deux méthodes montrent





que la bidroite donne une erreur-type nettement inférieure à celle de la méthode « horizontale ».

La grandeur de l'erreur-type observée est liée à la dispersion graphique des points représentatifs, et cette dispersion elle-même dépend en partie de « l'effet de zéro ». Pour *St-Avoid 1/2* (graphique 11, fig. 1 et 2), l'échantillon comprend 16 tiges :

-- en *râperie* (fig. 1), un seul zéro, mais grosse dispersion ;

— en *mincs* (fig. 2), 9 zéros, aucune loi bien définie ; toutefois 6 sur 7 des non zéros sont groupés dans les circonférences ≤ 60 , et 8 sur 9 des zéros s'étendent des circonférences 63 à 96, ce qui rend l'écart-type nettement plus faible pour la bidroite que pour l'horizontale ;

-- en *potcau*, 13 zéros sur 16 tiges, aucune loi bien définie ;

— en *minc + potcau + râperie*, aucun zéro, loi en évidence (graphique 12).

Les mêmes phénomènes se reproduisent pour Haguenau, ainsi qu'on peut le constater par les graphiques 13 et 14 et sur les autres places.

9° Ce § montre comment la dispersion des points empêche la diminution de l'erreur-type, même lorsqu'on emploie des lois de plus en plus souples. Il s'agit de *St-Avoid 156*, *râperie*, dont le nuage de points représentatif est le suivant :

(Voir graphique n° 15).

a) un changement de loi semble se présenter à $C = 100$, qui n'est le début d'aucune qualité supérieure. La présence de charpente a conduit à considérer en outre un changement de loi pour $C = 60$.

On a pris :

une droite pour $C \leq 60$ ($V = \alpha' + \beta'C$)

une droite pour $60 < C \leq 100$ ($V = \alpha'' + \beta''C$)

une droite horizontale pour $C > 100$ ($V = \alpha'''$)

avec la condition que ces droites se coupent respectivement aux abscisses $C = 60$ et $C = 100$.

La méthode utilisée a été encore celle des moindres carrés à multiplicateurs de LAGRANGE, dont voici les équations normales l'accent ' est relatif à $C \leq 60$, le double accent '' à $60 < C \leq 100$, le triple accent ''' à $C > 100$, et les sommations sont étendues à l'échantillon ; λ et μ sont des inconnues auxiliaires introduites par la méthode de LAGRANGE, et que l'on peut, si l'on veut, éliminer) :

$$\left\{ \begin{array}{l} n'\alpha' + \beta'\Sigma'c + \lambda = \Sigma'v \\ \alpha'\Sigma'c + \beta'\Sigma'c^2 + 60\lambda = \Sigma'vc \\ n''\alpha'' + \beta''\Sigma''c - \lambda + \mu = \Sigma''v \\ \alpha''\Sigma''c + \beta''\Sigma''c^2 - 60\lambda + 100\mu = \Sigma''vc \\ n'''\alpha''' - \mu = \Sigma'''v \\ \alpha' + 60\beta' = \alpha'' + 60\beta'' \\ \alpha'' + 100\beta'' = \alpha''' \end{array} \right.$$

La résolution de ce système se ramène à celle de quatre systèmes linéaires successifs de deux équations à deux inconnues.

Le volume est donné par :

$$v = N'\alpha' + \beta'\Sigma'c + N''\alpha'' + \beta''\Sigma''c + N'''\alpha''',$$

les sommations étant étendues à la population, et l'erreur-type par :

$$\frac{N}{n} \sqrt{Q}$$

$$Q = \frac{n}{n-3} \frac{N+n_2-n_1}{N} (\Sigma v^2 - \alpha'\Sigma'v - \beta'\Sigma'vc - \alpha''\Sigma''v - \beta''\Sigma''vc - \alpha'''\Sigma'''v)$$

les sommations étant étendues à l'échantillon.

b) On peut, en négligeant les considérations théoriques, s'en tenir à deux droites :

$$V = \alpha' + \beta'C \text{ pour } C \leq 100$$

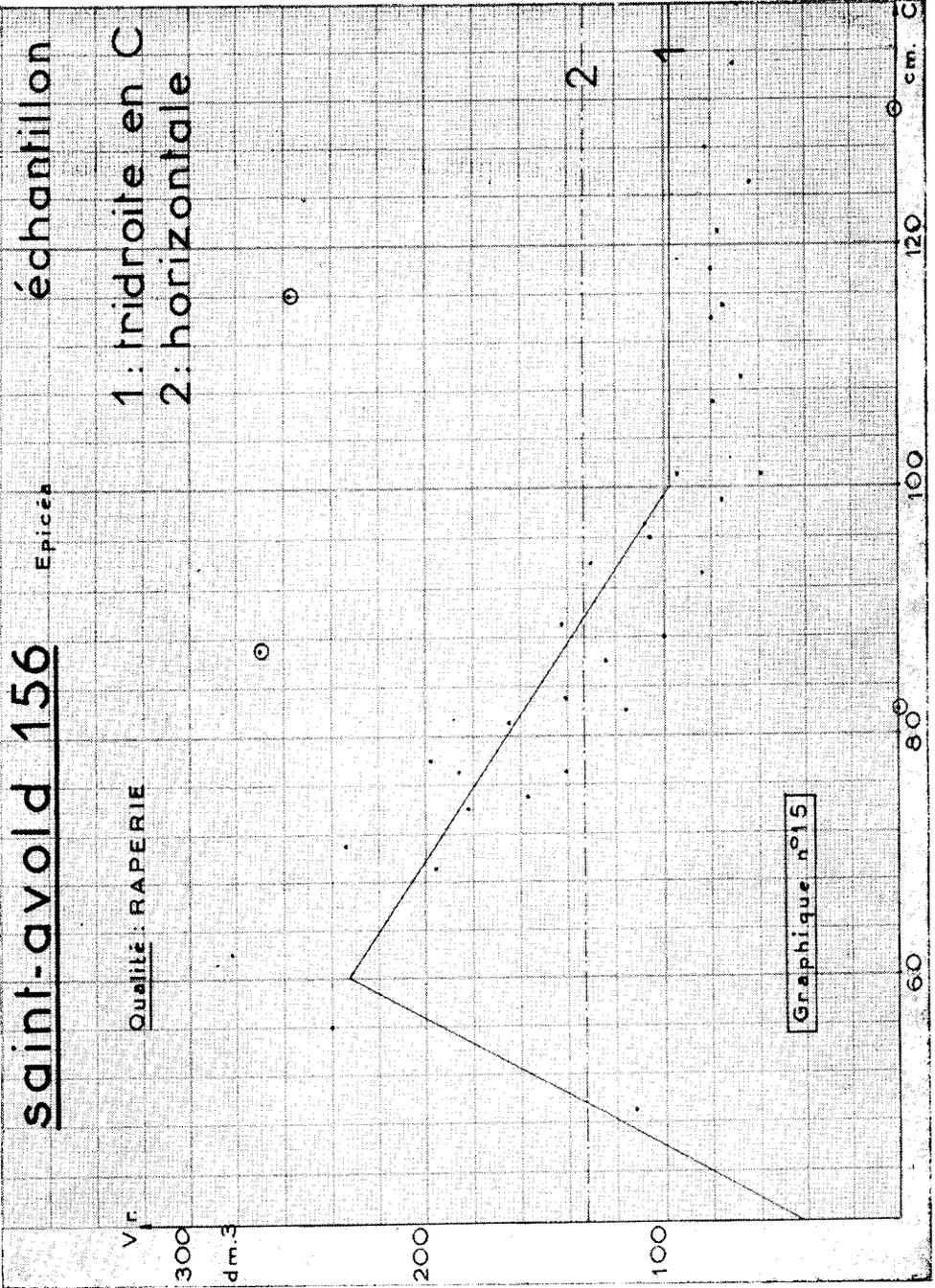
$$V = \alpha'' \text{ (horizontale) pour } C > 100,$$

avec la condition de coïncidence des volumes pour $C = 100$. Les équations normales sont :

$$\left\{ \begin{array}{l} n'\alpha' + \beta'\Sigma'c + \mu = \Sigma'v \\ \alpha'\Sigma'c + \beta'\Sigma'c^2 + 100\mu = \Sigma'vc \\ n''\alpha'' - \mu = \Sigma''v \end{array} \right.$$

$$\alpha' + 100\beta' = \alpha''$$

(l'accent est relatif à $C \leq 100$, le double accent à $C > 100$, et les sommations sont étendues à l'échantillon ; μ est le multiplicateur de LAGRANGE).



L'estimation du volume total est :

$$V = N' \alpha' + \beta' \sum' C + N'' \alpha''$$

(sommation étendue à la population)
et l'erreur-type

$$\frac{N}{n} \sqrt{Q}$$

avec

$$Q = \frac{n}{n-2} \frac{N + n_2 - n_1}{N} (\sum v^2 - \alpha' \sum' v - \beta' \sum' vc - \alpha'' \sum'' v)$$

(sommations étendues à l'échantillon).

c) On peut encore ne s'en tenir qu'aux considérations théoriques et prendre une bidroite dont le sommet est à $C = 60$.

d) On peut choisir également la méthode linéaire en C ($V = \alpha + \beta C$ sur toute l'étendue des circonférences) (1).

e) Enfin, on peut s'en tenir à la méthode horizontale.

Les résultats sont consignés dans le tableau suivant :

TABLEAU n° 25

Méthode	Erreur-type	Erreur-vraie
a	9,3 %	4,9 %
b	9,5 %	5,0 %
c	9,2 %	4,4 %
d	9,2 %	4,9 %
e	10,0 %	3,3 %

On constate que l'erreur-type est presque constante. Cela vient de ce que quatre tiges de l'échantillon (notés 0 sur le graphique échappent à toute loi, et introduisent à eux seuls 175 000 dans la somme des carrés des écarts, contre 51 000 pour les 35 autres tiges.

Dans un cas tel que celui-là, la méthode horizontale infiniment plus courte, doit être préférée.

10° En conclusion, pour chaque qualité de produits, il est conseillé de faire tout d'abord une inspection graphique rapide. S'il apparaît une droite, ou une bidroite, d'une façon assez nette, la suite du calcul est tout indiquée (se méfier cependant du phénomène signalé au § 9). Sinon, la méthode horizontale, très rapide doit être appliquée.

(1) Voir annexe II un exemple de calcul d'une telle droite.

L'inspection graphique doit être d'ailleurs limitée aux qualités les plus importantes. Sinon, on peut aboutir à des conclusions d'un intérêt fort incertain. Dans le cas de St-Avold 142 — poteau (qualité présente dans 3 tiges seulement, voir § 8), la règle du « double de l'écart-type » conduit à dire que le volume-poteau est 3 575 dm³, à 103 % près!) Etant d'ailleurs donné les faibles volumes en jeu dans cette place pour les qualités mine, râperie, poteau, le mieux est de les grouper en qualité unique mine + râperie + poteau (voir § 8). Cet exemple doit inspirer en général la marche à suivre.

BIBLIOGRAPHIE

-
1. AITKEN. — Statistical Mathematical (Edinburgh, Oliver and Boyd).
 2. ARBONNIER (P.). — Le calcul du volume bois fort-tige d'un peuplement sur pied (Rapport présenté au XII^e Congrès de l'Union Internationale des Instituts de Recherches forestières, Oxford, 1956).
 3. AYRAL (P.). — Sur le calcul du volume total et par catégories de produits de peuplements restant sur pied dans les places d'essais au passage des éclaircies (Comptes rendus du XI^e Congrès de l'Union Internationale des Instituts de Recherches forestières, Rome, 1953).
 4. BRUCE and SCHUMACHER. — Forest Mensuration (New-York, Mc Graw Hill-Book Company, 1935).
 5. CRAMER. — Mathematical Methods of Statistics (Princeton University Press, 1954).
 6. DARMOIS (G.). — Cours de Statistique de la Faculté des Sciences de Paris.
 7. EZECHIEL. — Methods of Correlation Analysis (New-York, John Wiley, 2^e édition, 1953).
 8. FISHER. — Méthodes statistiques adaptées à la recherche scientifique. — traduit par Dr. Bertrand — (Paris. Presses Universitaires de France).
 9. FISHER and YATES. — Statistical Table for Biological, Agricultural and Medical Research (Edinburgh. Oliver and Boyd, 1953).
 10. HÜFFEL. — Economie forestière, tome II (Paris, Lucien Laveur, 1905).
 11. HUMMEL. — The Volume — Basal Area Line (London, Her Majesty's Stationery Office, 1955).
 12. KENDALL. — The Advanced Theory of Statistics (London, Charles Griffin and Company, 1952).
 13. MILNE. — Numerical Calculus (Princeton University Press).
 14. OUDIN (A.). — Vues d'ensemble sur l'organisation en France des Recherches de Sylviculture et d'Economie forestière. Les Méthodes (Annales de l'Ecole Nationale des Eaux et Forêts et de la Station de Recherches et Expériences forestières, tome III, fascicule 2, 1930).
 15. OUDIN, FABRICIUS et GUILLEBARD. — Directives pour la conduite des recherches sur la production forestière (Nancy. Société d'Impressions Typographiques, 1936).
 16. PEARSON and HARTLEY. — Biometrika Tables for Statisticians (Cambridge University Press, 1954).
 17. SCHAEFFER (L.). — La méthode de corrélation appliquée à la construction des tarifs de cubage (Revue des Eaux et Forêts, juin 1932).
 18. SNEDECOR. — Statistical Methods (Iowa State College Press, 1950).
 19. VESSEREAU. — Méthodes statistiques en Biologie et en Agronomie. (Paris, Baillière, 1948).
 20. YATES. — Méthodes de sondage pour recensement et enquête. (Paris, Dunod, 1951).
-

ANNEXE I

1° Pour calculer la meilleure parabole (meilleure au sens de la méthode des moindres carrés), on peut soit calculer les différents coefficients des équations normales usuelles, soit utiliser une méthode de sommation réduisant le plus possible le nombre des multiplications à effectuer. Exposons cette méthode sur l'exemple de St-Amond 47. Les circonférences s'étendent, dans l'échantillon, de 37 à 109. On a choisi, arbitrairement, 65 comme circonférence de base pour effectuer les calculs.

On considère alors l'ensemble de toutes les paraboles d'axes verticaux, dont l'équation générale peut s'écrire :

$$V = b_0 + b_1 (C-65) + \frac{1}{2} b_2 (C-65) (C-66).$$

Dans le *tableau 1*, la *colonne (c)* indique les circonférences, en face de chacune desquelles on peut lire (*colonne (n)*) le nombre de tiges correspondant dans l'échantillon. Si, à une circonférence donnée (39 par exemple), ne correspond aucune tige, le nombre correspondant est zéro.

La *colonne (o)* se déduit, jusqu'à la circonférence 64, de la *colonne n* par *sommation vers le bas*: chaque nombre de la colonne (o), de circonférence ≤ 64 , est la somme des nombres de la colonne n de hauteur égale ou supérieure à lui (un nombre a du tableau 1 est dit de hauteur supérieure à un nombre b du même tableau s'il est à une distance supérieure du bord inférieur de la feuille: ce langage se comprend de lui-même); on vérifie aisément que chaque nombre de la colonne (o) est la somme du nombre de la même colonne situé immédiatement au-dessus, et du nombre de la colonne (n) situé immédiatement à sa gauche. C'est ainsi que :

$$6 = 1 + 1 + 0 + 1 + 0 + 0 + 0 + 2 + 1,$$

et qu'en même temps : $6 = 5 + 1$.

Nous appellerons *sommation vers le bas de la colonne (n)* depuis 37 jusqu'à 64, l'opération qui consiste à passer de la colonne (n) à la colonne (o) (jusqu'à la circonférence 64 incluse).

La *colonne (o)*, de la circonférence 65 jusqu'à la circonférence 109, se déduit de la colonne (n) par *sommation vers le haut*: chaque nombre de la colonne (o) correspondant à une circonférence ≥ 65 est la somme des nombres de la colonne (n) de hauteur inférieure ou égale à lui, il est aussi la somme du nombre de la colonne (n) situé immédiatement à sa gauche et du nombre de la colonne (o) situé au-dessous de lui.

La *colonne (1)* s'obtient à partir de la colonne (o) comme cette dernière s'était obtenue à partir de la colonne (n); on passe de même des colonnes (1) à (2), (2) à (3), (3) à (4). Toutefois, alors que les sommations vers le bas s'arrêtent à la circonférence 64, les sommations vers le haut s'arrêtent:

à la circonférence 65 pour la colonne (o),

—	66	—	(1),
—	67	—	(2),
—	68	—	(3),
—	69	—	(4).

Enfin, les nombres intermédiaires de la colonne (4) qui ne pourraient servir qu'à une sommation supplémentaire n'ont pas été écrits.

On obtient ainsi les nombres A, B, C, D, E, F, G, H, I, J indiqués dans le tableau 1, et l'on déduit les nombres S_0, S_1, S_2, S_3, S_4 qui donneront les premiers membres des équations normales.

Le *tableau 2* est destiné au calcul des seconds membres de ces mêmes équations. La *colonne (c)* indique encore les circonférences et la *colonne (v)* la somme des volumes des tiges d'une même circonférence. Lorsqu'il n'existe aucune tige d'une certaine circonférence, le volume correspondant est zéro.

Cela étant, la colonne (1) se déduit par sommations (vers le bas et vers le haut) de la colonne (o), la colonne (2) par sommations de la colonne (1). Par calcul analogue à celui du tableau 1, on obtient les nombres R_0, R_1, R_2 .

Les *équations normales* sont, comme on le démontre au § 3 ci-après:

$$\begin{aligned} S_0 b_0 + S_1 b_1 + S_2 b_2 &= R_0 \\ S_1 b_0 + (2S_2 + S_1) b_1 + (3S_3 + 2S_2) b_2 &= R_1 \\ S_2 b_0 + (3S_3 + 2S_2) b_1 + (6S_4 + 6S_3 + S_2) b_2 &= R_2 \end{aligned}$$

et sont résolues dans le *tableau 3*. On y retrouve aux trois premières lignes les équations normales sous la forme:

S_0	S_1	S_2	R_0
S_1	$2S_2 + S_1$	$3S_3 + 2S_2$	R_1
S_2	$3S_3 + 2S_2$	$6S_4 + 6S_3 + S_2$	R_2

La méthode de résolution adoptée est la suivante : on divise les deux membres de chaque équation par le coefficient correspondant de b_0 , ce qui donne les équations (4), (5), (6). On retranche ainsi membre à membre l'équation (4) des équations (5) et (6), pour obtenir (7) et (8), qui forment un système aux deux seules inconnues b_1 , b_2 .

On recommence le cycle précédent : division de chaque équation par le coefficient des inconnues « les plus à gauche », et soustraction membre à membre de l'équation (9) de l'équation (10), pour obtenir l'équation (11) à la seule inconnue b_2 . Une dernière division donne : $b_2 = 0,341435210$.

Il ne reste plus ensuite qu'à opérer par substitutions successives dans (9) et dans (4) pour obtenir b_1 et b_0 .

Le tableau 4 donne l'établissement du tarif par *sommations inverses*. Les colonnes (c) et (n) donnent, pour chaque circonférence de la *population*, le nombre correspondant de tiges. Les nombres imprimés en gras dans les colonnes (0), (1), (2) sont respectivement b_0 , b_1 , b_2 . Les *sommations inverses vers le haut* se font de la façon suivante : on retranche b_2 à b_1 (pour obtenir 14,8086), puis encore b_2 au résultat (pour obtenir 14,4672), etc... On aboutit ainsi à la partie supérieure de la colonne (1).

Pour obtenir le nombre de la colonne (0) immédiatement supérieur à b_0 , on retranche de celui-ci le *nombre intercalaire* 14,8086 de la colonne (1) (pour obtenir 389,28). De même, pour obtenir le nombre suivant, on retranche à 389,28 le nombre intercalaire 14,4672 (pour obtenir 374,81), et ainsi de suite.

Les *sommations vers le bas* procèdent du même principe, en échangeant toutefois les opérations d'addition et de soustraction.

La colonne (0) donne le tarif, et il suffit d'effectuer $\sum nV$ pour avoir l'estimation du volume total de la place.

Le tableau (5) donne le calcul de l'erreur-type par la formule :

$$\frac{N}{n} \sqrt{\frac{N * n_2 - n_1}{N} * \frac{n}{n-3} (\sum v^2 - b_0 R_0 - b_1 R_1 - b_2 R_2)}$$

2° Les calculs ont été faits avec une précision bien supérieure à celle qui était indispensable.

On trouvera un autre exposé de la méthode de sommation précédente dans :

FISHER : « Les méthodes statistiques adaptées à la recherche scientifique » (traduit en français, Presses Universitaires de France).

FISHER étend la méthode aux paraboles d'ordre 3, dont voici, schématiquement, les équations normales :

b_0	b_1	b_2	b_3	
S_0	S_1	S_2	S_3	R_0
S_1	$2S_2 + S_1$	$3S_3 + 2S_2$	$4S_4 + 3S_3$	R_1
S_2	$3S_3 + 2S_2$	$6S_4 + 6S_3 + S_2$	$10S_5 + 12S_4 + 3S_3$	R_2
S_3	$4S_4 + 3S_3$	$10S_5 + 12S_4 + 3S_3$	$20S_6 + 30S_5 + 12S_4 + S_3$	R_3

correspondant à la famille de fonctions :

$$v = b_0 + \frac{b_1}{1} (c-65) + \frac{b_2}{1 \cdot 2} (c-65)(c-66) + \frac{b_3}{1 \cdot 2 \cdot 3} (c-65)(c-66)(c-67)$$

(rappelons que la base, prise ici à $C = 65$, est arbitraire).

3° Les formules sont relativement aisées à démontrer.

Commençons par poser : $C-65 = x$,

puis : $x^{(n)} = x(x-1) \dots (x-n+1)$, pour n entier > 0 .

On aura successivement :

$$x^{(2)} = x(x-1) = x^2 - x$$

$$x^{(3)} = x(x-1)(x-2) = x^3 - 3x^2 + 2x$$

$$x^{(4)} = x(x-1)(x-2)(x-3) = x^4 - 6x^3 + 11x^2 - 6x,$$

qui donne de proche en proche :

$$x^2 = x^{(2)} + x$$

$$x^3 = x^{(3)} + 3x^{(2)} + x$$

$$x^4 = x^{(4)} + 6x^{(3)} + 7x^{(2)} + x$$

Les coefficients de ces derniers polynômes ne sont d'ailleurs autres que les « différences de zéro » (table XXII de FISHER and YATES). C'est ainsi que, par exemple :

$$x^8 = x^{(8)} + 28x^{(7)} + 266x^{(6)} + 1050x^{(5)} + 1701x^{(4)} + 966x^{(3)} + 127x^2 + x$$

Cela étant, pour la parabole du second degré :

$$V = b_0 + b_1 x + \frac{1}{2} b_2 x^{(2)},$$

les équations normales sont :

$$nb_0 + b_1 \sum x + \frac{1}{2} b_2 \sum x^{(2)} = \sum V$$

$$b_0 \sum x + b_1 \sum x^2 + \frac{1}{2} b_2 \sum x \cdot x^{(2)} = \sum Vx$$

$$\frac{1}{2} b_0 \sum x^2 + \frac{1}{2} b_1 \sum x \cdot x^{(2)} + \frac{1}{4} b_2 \sum x^{(2)} \cdot x^{(2)} = \frac{1}{2} \sum V \cdot x^{(2)}$$

On commence par transformer les expressions en x à l'aide des formules trouvées précédemment :

$$x \cdot x^{(2)} = x^2 (x-1) = x^3 - x^2 = x^{(3)} + 2x^{(2)}$$

$$x^{(2)} \cdot x^{(2)} = x^4 - 2x^3 + x^2 = x^{(4)} + 4x^{(3)} + 2x^{(2)}$$

En posant :
$$s_n = \frac{1}{n!} \sum x^{(n)} \quad (n! = n^{(n)}),$$

on trouve :

$$\sum x^2 = 2s_2 + s_1$$

$$\sum x \cdot x^{(2)} = 6s_3 + 4s_2$$

$$\sum x^{(2)} \cdot x^{(2)} = 24s_4 + 24s_3 + 4s_2$$

valeurs que l'on substitue dans les équations normales précédentes.

4° Il reste maintenant à justifier que S_0, S_1, S_2, S_3 , et

$$R_0 = \sum V, R_1 = \sum Vx, R_2 = \frac{1}{2} \sum Vx^{(2)}$$

s'obtiennent effectivement comme il est indiqué respectivement aux tableaux 1 et 2.

Cette démonstration se trouve dans KENDALL, « The Advanced Theory of Statistics », 3.10, et on ne la reproduira pas ici. KENDALL, il est vrai, ne procède que par sommations vers le haut, mais il est facile de modifier sa démonstration pour retrouver la méthode de FISHER; cette dernière a eu pour nous l'avantage de n'introduire que des nombres acceptés par notre additionneuse électrique.

Les sommations inverses du tableau 4 procèdent du même principe.

5° La méthode s'étend naturellement à des paraboles de degrés quelconques. Pour la parabole du quatrième degré par exemple, les calculs donnent les équations normales dont voici le schéma :

b_0	b_1	b_2	b_3	b_4	
S_0	S_1	S_2	S_3	S_4	R_0
S_1	$2S_2 + S_1$	$3S_3 + 2S_2$	$4S_4 + 3S_3$	$5S_5 + 4S_4$	R_1
S_2	$3S_3 + 2S_2$	$6S_4 + 6S_3 + S_2$	$10S_5 + 12S_4 + 3S_3$	$15S_6 + 20S_5 + 6S_4$	R_2
S_3	$4S_4 + 3S_3$	$10S_5 + 12S_4 + 3S_3$	$20S_6 + 30S_5 + 12S_4 + S_3$	$35S_7 + 60S_6 + 30S_5 + 4S_4$	R_3
S_4	$5S_5 + 4S_4$	$15S_6 + 20S_5 + 5S_4$	$35S_7 + 60S_6 + 30S_5 + 4S_4$	$70S_8 + 140S_7 + 90S_6 + 20S_5 + S_4$	R_4

correspondant à la famille de fonctions :

$$V = b_0 + \frac{b_1}{1} (c-65) + \frac{b_2}{1.2} (c-65)(c-66) + \frac{b_3}{1.2.3} (c-65)(c-66)(c-67) + \frac{b_4}{1.2.3.4} (c-65)(c-66)(c-67)(c-68)$$

Remarquons pour finir que, si l'on utilise une machine imprimant sur bande les résultats des sommations successives, il est naturellement inutile de les recopier dans les différentes colonnes des tableaux (1), (2), (4). Cela sera mis à profit à l'annexe 2.

ANNEXE I — TABLEAU I

St-Amond 47

(a)

c	n	o	1	2	3	4
37	I	I	1	I	I	I
38	I	2	3	4	5	
39		2	5	9	14	
40	I	3	8	17	31	
41		3	11	28	59	
42		3	14	42	101	
43		3	17	59	160	
44	2	5	22	81	241	
45	I	6	28	109	350	
46	I	7	35	144	494	
47	I	8	43	187	681	
48		8	51	238	919	
49		8	59	297	1 216	
50		8	67	364	1 580	
51		8	75	439	2 019	
52	I	9	84	523	2 542	
53	I	10	94	617	3 159	
54	I	11	105	722	3 881	
55		11	116	838	4 719	
56		11	127	965	5 684	
57	I	12	139	1 104	6 788	
58		12	151	1 255	8 043	
59		12	163	1 418	9 461	
60		12	175	1 593	11 054	
61	I	13	188	1 781	12 835	
62	I	14	202	1 983	14 818	
63	2	16	218	2 201	17 019	
64		16	234	2 435	19 454	127 328
		A	C	E	G	I

(b)

c	n	o	1	2	3	4
		B	D	F	H	J
65	I	20				
66		19	311			
67	I	19	292	3 416		
68		18	273	3 124	30 882	
69	I	18	255	2 851	27 758	239 927
70		17	237	2 596	24 907	
71	I	17	220	2 359	22 311	
72		16	203	2 139	19 952	
73	I	16	187	1 936	17 813	
74	2	15	171	1 749	15 877	
75	I	13	156	1 578	14 128	
76		12	143	1 422	12 550	
77		12	131	1 279	11 128	
78	I	12	119	1 148	9 849	
79	2	11	107	1 029	8 701	
80	I	9	96	922	7 672	
81	I	8	87	826	6 750	
82		7	79	739	5 924	
83		7	72	660	5 185	
84	I	7	65	588	4 525	
85	I	6	58	523	3 937	
86		5	52	465	3 414	
87	I	5	47	413	2 949	
88	I	4	42	366	2 536	
89		3	38	324	2 170	
90		3	35	286	1 846	
91		3	32	251	1 560	
92		3	29	219	1 309	
93	I	3	26	190	1 090	
94		2	23	164	900	
95		2	21	141	736	
96		2	19	120	595	
97		2	17	101	475	
98		2	15	84	374	
99		2	13	69	290	
100	I	2	11	56	221	
101		I	9	45	165	
102		I	8	36	120	
103		I	7	28	84	
104		I	6	21	56	
105		I	5	15	35	
106		I	4	10	20	
107		I	3	6	10	
108		I	2	3	4	
109	I	I	I	I	I	

$$S_0 = A + B = 36$$

$$S_1 = D - C = 77$$

$$S_2 = F + E = 5 851$$

$$S_3 = H - G = 11 428$$

$$S_4 = I + J = 367 255$$

$$2 S_2 + S_1 = 11 779$$

$$3 S_3 + 2 S_2 = 45 986$$

$$4 S_4 + 3 S_3 = 1 503 304$$

$$6 S_4 + 6 S_3 + S_2 = 2 277 949$$

NOTA. — Le tableau (b) de droite devrait normalement se trouver sous le tableau (a) dont il constitue le prolongement. Cette séparation a été motivée par des facilités de mise en page.

ANNEXE I — TABLEAU 2

St-Amond 47

(a)

(c)	v	o	1	2
37	88	88	88	88
38	117	205	293	
39		205	498	
40	126,5	331,5	829,5	
41		331,5	1 161	
42		331,5	1 492,5	
43		331,5	1 824	
44	322,5	654	2 478	
45	162	816	3 294	
46	175	991	4 285	
47	172,5	1 163,5	5 448,5	
48		1 163,5	6 612	
49		1 163,5	7 775,5	
50		1 163,5	8 939	
51		1 163,5	10 102,5	
52	298,5	1 462	11 564,5	
53	257,5	1 719,5	13 284	
54	284	2 003,5	15 287,5	
55		2 003,5	17 291	
56		2 003,5	19294,5	
57	304	2 307,5	21 602	
58		2 307,5	23 909,5	
59		2 307,5	26 217	
60		2 307,5	28 524,5	
61	347,5	2 655	31 179,5	
62	346	3 001	34 180,5	
63	781,5	3 782,5	37 963	
64		3 782,5	41 745,5	377 163,5
		A	C	E

(b)

(c)	v	o	1	2
		B	D	F
65	387,5	13 929		
66		13 541,5	267 014,5	
67	443	13 541,5	253 473	3 461 625,5
68		13 098,5	239 931,5	
69	489,5	13 098,5	226 833	
70		12 609	213 734,5	
71	571	12 609	201 125,5	
72		12 038	188 516,5	
73	554,5	12 038	176 478,5	
74	1 085	11 483,5	164 440,5	
75	570,5	10 398,5	152 957	
76		9 828	142 558,5	
77		9 828	132 730,5	
78	695,5	9 828	122 902,5	
79	1 318	9 132,5	113 074,5	
80	618,5	7 814,5	103 942	
81	662	7 196	96 127,5	
82		6 534	88 931,5	
83		6 534	82 397,5	
84	700	6 534	75 863,5	
85	737,5	5 834	69,329,5	
86		5 096,5	63 495,5	
87	737	5 096,5	58 399	
88	846,5	4 359,5	53 302,5	
89		3 513	48 943	
90		3 513	45 430	
91		3 513	41 917	
92		3 513	38 404	
93	923	3 513	34 891	
94		2 590	31 378	
95		2 590	28 788	
96		2 590	26 198	
97		2 590	23 608	
98		2 590	21 018	
99		2 590	18 428	
100	1 118	2 590	15 838	
101		1 472	13 248	
102		1 472	11 776	
103		1 472	10 304	
104		1 472	8 832	
105		1 472	7 360	
106		1 472	5 888	
107		1 472	4 416	
108		1 472	2 944	
109		1 472	1 472	
100	1 472	1 472	1 472	1 472

$$R_0 = A + B = 17 711,5$$

$$R_1 = D - C = 225 269$$

$$R_2 = E + F = 3 838 789$$

NOTA. — Le tableau (b) de droite devrait normalement se trouver sous le tableau (a) dont il constitue le prolongement. Cette séparation a été motivée par des facilités de mise en page.

ANNEXE I --- TABLEAU 3

St-Amoud 47

	b_0	b_1	b_2	
(1)	36	77	5 851	17 711,5
(2)	77	11 779	45 986	225 269
(3)	5851	45 986	2 277 949	3 838 789
(4)	I	2,138888880	162,5277779	491,9861110
(5)	I	152,9740260	597,220770	2 925,571430
(6)	I	7,859511100	389,3264400	656,0910970
(7)		150,8351372	434,6929921	2 433,585319
(8)		5,720622220	226,7986621	164,1049860
(9)		I	2,881908032	16,13407443
(10)		I	39,64580330	28,68656260
(11)			36,76389527	12,55248817
(12)			I	0,341435210

$b_0 = 404,0890506$
 $b_1 = 15,15008957$
 $b_2 = 0,34143521$

$b_0 + 2,138888880 \quad b_1 + 162,5277779 \quad b_2 = 491,9861110$
 $b_0 = 404,0890506$

$b_1 + 2,881908032 \quad b_2 = 16,13407443$
 $b_1 = 15,15008957$

$b_2 = 0,341435210$

(c)	(n)	0	1	2
34	1	103,80		
			4,5658	
35		108,36		
			4,9072	
36		113,26		
			5,2486	
37		118,50		
			5,5900	
38		124,09		
			5,9314	
39	1	130,03		
			6,2729	
40	1	136,30		
			6,6143	
41		142,91		
			6,9557	
42		149,87		
			7,2971	
43	3	157,17		
			7,6386	
44	4	164,80		
			7,9800	
45	3	172,78		
			8,3214	
46	3	181,11		
			8,6629	
47	1	189,77		
			9,0043	
48	1	198,77		
			9,3457	
49	1	208,12		
			9,6872	
50		217,81		
			10,0286	
51	3	227,83		
			10,3700	
52	3	238,20		
			10,7114	
53	1	248,92		
			11,0529	
54	1	259,97		
			11,3943	
55	1	271,36		
			11,7357	
56	2	283,10		
			12,0772	
57	2	295,18		
			12,4186	
58	1	307,59		
			12,7600	
59	1	320,35		
			13,1015	
60	4	333,46		
			13,4429	
61	2	346,90		
			13,7843	
62	2	360,68		
			14,1257	
63	4	374,81		
			14,4672	
64	2	389,28		
			14,8086	
65	1	404,0890506		
			15,15008957	
66	2	419,23		0,34143521
			15,4915	
67	3	434,73		
			15,8329	
68	1	450,56		
			16,1743	
69	3	466,73		
			16,5158	
70	2	483,25		
			16,8572	
71	2	500,11		
			17,1986	
72	3	517,30		
			17,5400	
73	3	534,84		

(c)	(n)	0	1	2
			17,8815	
74	3	552,73		
			18,2229	
75	4	570,95		
			18,5643	
76	1	589,51		
			18,9058	
77		608,42		
			19,2472	
78		627,67		
			19,5886	
79	3	647,25		
			19,9301	
80	3	667,18		
			20,2715	
81	1	687,46		
			20,6129	
82	1	708,07		
			20,9543	
83	2	729,02		
			21,2958	
84	1	750,32		
			21,6372	
85	2	771,96		
			21,9786	
86	3	793,94		
			22,3201	
87	2	816,26		
			22,6615	
88	1	838,92		
			23,0029	
89	2	861,92		
			23,3444	
90	1	885,26		
			23,6858	
91		908,95		
			24,0272	
92		932,98		
			24,3686	
93	2	957,35		
			24,7101	
94		982,06		
			25,0515	
95		1 007,11		
			25,3929	
96		1 032,50		
			25,7344	
97		1 058,23		
			26,0758	
98		1 084,31		
			26,4172	
99	1	1 110,73		
			26,7587	
100	2	1 137,49		
			27,1001	
101		1 164,59		
			27,4415	
102		1 192,03		
			27,7829	
103		1 219,81		
			28,1244	
104		1 247,94		
			28,4658	
105		1 276,40		
			28,8072	
106	1	1 305,21		
			29,1487	
107		1 334,36		
			29,4901	
108		1 363,85		
			29,8315	
109	1	1 393,68		
			30,1730	
110		1 423,85		
			30,5144	
111		1 454,37		
			30,8558	
112	1	1 485,22		
			31,1972	

$$S_n V = 52 678,38$$

ANNEXE I — TABLEAU 5

St-Amoud 47

ΣV^2	=	11 926 089	$n_1 = 16$	$n_2 = 20$	$n = 36$	$N = 106$
$- b_0 R_0$	=	7 157 023				
$- b_1 R_1$	=	3 412 845	$n = 3$		N	
$- b_2 R_2$	=	1 310 697	n	$= 0,916666$	$\frac{N}{N + n_2 - n_1} = 0,963636$	
r	=	45 524				
		↓				
			$\frac{N}{n} = 2,944444$			
			$\frac{N}{n - 3} \rightarrow$		N	
			$\frac{45 524}{n} \rightarrow$	$= 49 662$	$\frac{N}{N + n_2 - n_1} = 51 532$	
			diviser par		prendre la	↓
					racine	↓
					$\sqrt{51 532}$	
					↓	
					227	
					↓	
					N	
					multiplié par	$\frac{N}{n} = 668 = ET$

ANNEXE 2

La méthode employée est la même que celle de l'annexe 1, mais :

1° Les arbres ont été groupés par classes de circonférences, de 5 en 5. La classe 35, par exemple, comprend les tiges de circonférence 35, 36, 37, 38, 39.

2° Les nombres intermédiaires, dans les sommations, n'ont pas été recopiés, sauf pour les colonnes (0) et (1) du tableau 1, à titre indicatif. Une bande imprimée par une de nos additionneuses automatiques montre bien l'inutilité d'un tel report de nombres.

3° La résolution du système d'équations, entièrement analogue à celui de l'annexe 1, n'est pas donnée (d'où l'absence du tableau 3).

4° Le tableau 4 a, qui remplace le tableau 4 est beaucoup plus court. Il ne donne le tarif que pour le début de chaque groupe de circonférences. Le tableau se remplit de droite à gauche, sauf évidemment pour la colonne indiquant le début de chaque groupe. La colonne à gauche, calculée après coup, sera utilisée pour l'interpolation linéaire de chaque groupe.

5° Le tableau 4 b est l'extension, par interpolation linéaire, du tableau 4 a, et donne le tarif provisoire destiné à être corrigé après achèvement du tableau.

6° Le tableau 5 indique le calcul de l'erreur-type ainsi que la correction, — 5,23, à apporter au tarif du tableau 4 b. On en déduit une estimation corrigée 52 256,12 du volume de la placette.

Remarque. — On constatera que la circonférence prise ici pour origine est $C = 35$.

ANNEXE 2 — TABLEAU 1

St-Amund 47

(c)	(n)	(o)	(1)	(2)	(3)	(4)
35	2	36 = S ₀				
40	3	34	215 = S ₁			
45	3	31	181	766 = S ₂		
50	3	28	150	585	1 898 = S ₃	
55	1	25	122	435		3 540 = S ₄
60	4	24	97	313		
65	3	20	73	216		
70	4	17	53	143		
75	4	13	36	90		
80	3	9	23	54		
85	3	6	14	31		
90	1	3	8	17		
95	—	2	5	9		
100	1	2	3	4		
105	1	1	1	1	1	1
110						

2 S ₂ + S ₁	=	1 747
3 S ₂ + 2 S ₁	=	7 226
4 S ₄ + 3 S ₃	=	19 854
6 S ₄ + 6 S ₃ + S ₂	=	33 394
10 S ₅ + 12 S ₄ + 3 S ₃	=	100 324
20 S ₆ + 30 S ₅ + 12 S ₄ + S ₃	=	324 948

Nota: 35 signifie: 35, 36, 37, 38, 39.

40 signifie: 40, 41, 42, 43, 44,
etc., etc...

Les bandes n° 1 et 2 indiquent comment, à l'aide d'une machine à additionner imprimante on peut se dispenser de remplir toutes les lignes du tableau 1.

La machine imprime en bleu les données. Les résultats partiels sont imprimés en rouge et suivis d'un losange. Le résultat définitif est suivi d'un astérisque.

La bande n° 1 correspond à la sommation de la colonne (1) du tableau 1. Les nombres suivis de losange sont donc ceux de la colonne 2 dont le total est 766. Ils sont repris dans la bande n° 2 où ils donnent ceux de la colonne (3) dont le total est 1 898 et ainsi de suite.

1	1
1 ◊	1 ◊
3	4
4 ◊	5 ◊
5	9
9 ◊	14 ◊
8	17
17 ◊	31 ◊
14	31
31 ◊	62 ◊
23	54
54 ◊	116 ◊
36	90
90 ◊	206 ◊
53	143
143 ◊	349 ◊
73	216
216 ◊	565 ◊
97	313
313 ◊	878 ◊
122	435
435 ◊	1 313 ◊
150	585
585 ◊	1 898 ◊
181	1 898
766 ◊	1 898 *
766 *	

ANNEXE 2 — TABLEAU 2

St-Amond 47

(c)	(v)	(o)	(i)	(2)
35	205	17 711,5 = R ₀	142 974 = R ₁	605 840 = R ₂
40	449			
45	509,5			
50	840			
55	304			
60	1 475			
65	1 320			
70	2 210,5			
75	2 584			
80	1 980,5			
85	2 321			
90	923			
95	—			
100	1 118			
105	1 472	1 472	1 472	1 472

$$R_0 = 17\,711,5$$

$$R_1 = 142\,974$$

$$R_2 = 605\,840$$

ANNEXÉ 2 — TABLEAU 4 a

St-Amond 47

Résultat de la division des nombres de la colonne (1) par 5	Début de la classe	Milieu de la classe	(0)	(1)	(2)
4,83	30	32	95,85	24,1868	
6,50	35	37	120,0417465	32,53528919	
8,17	40	42	152,57	40,8837	8,3484664
9,84	45	47	193,46	49,2321	
11,51	50	52	242,69	57,5806	
13,18	55	57	300,27	65,9290	
14,85	60	62	366,20	74,2774	
16,52	65	67	440,47	82,6259	
18,19	70	72	523,10	90,9743	
19,86	75	77	614,07	99,3228	
21,53	80	82	713,40	107,6712	
23,20	85	87	821,07	116,0196	
24,87	90	92	937,09	124,3681	
26,54	95	97	1 061,46	132,7165	
28,21	100	102	1 194,17	141,0650	
29,88	105	107	1 335,24	149,4134	
	110	112	1 484,65		

St-Amond 47

(c)	(n)	V	(c)	(n)	V
34	1	105,52	73	3	541,30
35		110,36	74	3	559,49
36		115,20	75	4	577,69
37		120,03	76	1	595,88
38		126,54	77		614,07
39	1	133,05	78		633,94
40	1	139,56	79	3	653,80
41		146,07	80	3	673,67
42		152,58	81	1	693,53
43	3	160,76	82	1	713,39
44	4	168,93	83	2	734,93
45	3	177,11	84	1	756,46
46	3	185,28	85	2	778
47	1	193,46	86	3	799,53
48	1	203,31	87	2	821,06
49	1	213,15	88	1	844,27
50		223	89	2	867,47
51	3	232,84	90	1	890,68
52	3	242,69	91		913,88
53	1	254,21	92		937,09
54	1	265,72	93	2	961,96
55	1	277,24	94		986,84
56	2	288,75	95		1 011,71
57	2	300,27	96		1 036,58
58	1	313,45	97		1 061,45
59	1	326,64	98		1 088
60	4	339,82	99	1	1 114,54
61	2	353,01	100	2	1 141,09
62	2	366,19	101		1 167,63
63	4	381,05	102		1 194,18
64	2	395,91	103		1 222,39
65	1	410,77	104		1 250,61
66	2	425,62	105		1 278,82
67	3	440,48	106	1	1 307,03
68	1	457	107		1 335,24
69	3	473,53	108		1 365,13
70	2	490,05	109	1	1 395,01
71	2	506,58	110		1 424,89
72	3	523,10	111		1 454,77
			112	1	1 484,65
A reporter.: 65			Total . . . 106 $\Sigma nv = 52811$ —		

ANNEXE 2 — TABLEAU 5

St-Amond 47

(c)	Volume réel V_r	Volume tarif V_t	$V_r - V_t + 5,23$	$(V_r - V_t + 5,23)^2$
<i>Abattus</i>				
38	117	126,54	— 4,31	18,58
40	126	139,56	— 8,33	69,38
44	172,5	168,93	8,80	77,44
44	150	168,93	— 13,70	187,60
45	162	177,11	— 9,88	97,61
46	175	185,28	— 5,05	25,50
47	172,5	193,46	— 15,73	247,43
52	298,5	242,69	61,04	3 725,88
53	257,5	254,21	8,52	72,59
54	284	265,72	23,51	552,72
57	304	300,27	8,06	80,28
61	347,5	353,01	— 0,28	0,07
62	346	366,19	— 14,96	223,80
63	396	381,05	20,18	407,23
63	385,5	381,05	9,68	93,70
67	443	440,48	7,75	60,06
71	571	506,58	69,65	4 851,12
73	554,5	541,30	18,43	339,66
74	486	559,49	— 68,26	4 659,42
78	695,5	633,94	66,79	4 460,90
/				
20				
<i>Debout</i>				
37	88	120,03	— 26,80	718,24
65	387	410,77	— 18,04	325,44
69	489,5	473,53	21,20	449,44
74	599	559,49	44,74	2 001,66
75	570,5	577,69	— 1,96	3,84
79	653,5	653,80	4,93	24,40
79	664,5	653,80	15,93	253,76
80	618,5	673,67	— 49,94	2 494
81	662	693,53	— 26,30	691,60
84	700	756,46	— 51,23	2 624,51
85	737,5	778	— 35,77	1 279,49
87	737	821,06	— 78,33	6 135,58
88	846,5	844,27	7,46	55,65
93	923	961,96	— 33,73	1 137,71
100	1 118	1 141,09	— 17,86	318,97
109	1 472	1 395,01	82,22	6 760,12
/				
16				
Total				
36 arbres	17 711,5	17 899,95	$r = 45,525,56$	
		— 188,45		
		— 188,45		
		— 5,23		
	36			

Correction sur le volume total

Volume estimé avant correction: 52 811
 Correction: — $5,23 \times 106 =$ — 554,38

Volume estimé après correction: 52 256,62

Le volume réel cubé après abattage
 de tous les arbres est 52 187

Calcul de l'erreur-type

$n_1 = 16$ $n_2 = 20$ $n = 36$ $N = 106$

$$\frac{n - 3}{n} = 0,916666$$

$$\frac{N}{n} = 2,944444$$

$$\frac{N}{N + n_2 - n_1} = 0,963636$$

diviser par \rightarrow = 49 663 diviser par \rightarrow = 51 534

$$\frac{n - 3}{n}$$

$$\frac{N}{N + n_2 - n_1}$$

racine carrée de 51 534

↓
227 multiplié

$$\text{par } \frac{N}{n} = \text{ET} = 668$$

ANNEXE III

Le calcul de l'erreur

1° Nous nous plaçons dans le cas linéaire, où l'on veut représenter un nuage de n points (x_i, y_i) , formant échantillon, par une fonction de la famille :

$$f(x_i, a_1, a_2, \dots, a_k) = a_1 g_1(x) + a_2 g_2(x) + \dots + a_k g_k(x).$$

On reprend les notations du Titre I, en désignant par \hat{f} , \hat{a}_1 , \hat{a}_2 ,, \hat{a}_k la fonction et les valeurs des paramètres qui assurent le minimum de

$$\sum_{i=1}^n (y_i - f_i)^2$$

Il est classique d'estimer la variance de $y - \hat{f}$ par :

$$\frac{1}{n-k} \sum_{i=1}^n (y_i - \hat{f}_i)^2$$

(voir par exemple EZECHIEL « Methods of Correlation Analysis », page 133), donc la variance de :

$$\sum_{i=1}^n (y_i - \hat{f}_i) \text{ par } \frac{n}{n-k} \sum_{i=1}^n (y_i - \hat{f}_i)^2.$$

Cela suppose essentiellement que la variance de $y - \hat{f}(x)$ ne dépend pas de x et que les variances des coefficients a_i peuvent être négligées. Malheureusement s'il est possible de négliger la variance des a_i , la variance du volume croît très vite avec la circonférence. On appliquera néanmoins les formules où la variance de y ne dépend pas de x , quitte à vérifier expérimentalement l'applicabilité des formules : c'est l'objet du Titre VII, et l'on verra que l'accord avec l'expérience est aussi bon qu'on peut le désirer, et même surprenant.

On peut voir intuitivement la nécessité d'un facteur qui, au dénominateur de la variance de

$$\sum (y_i - \hat{f}_i)^2,$$

s'annule pour $n = k$: en effet, dans ce dernier cas :

$$\sum (y_i - \hat{f}_i)^2$$

est nul. Et cependant, cette erreur nulle ne peut être d'aucune utilité, car il était évident a priori que la courbe dépendant de k paramètres allait passer par les k points (x_i, y_i) , et il n'y a plus d'erreur à calculer ; c'est ce que traduit la forme indéterminée $\frac{0}{0}$ prise par :

$$\frac{\sum (y_i - \hat{f}_i)}{n - k}.$$

Cette remarque intuitive est assez pratique lorsque l'on veut avoir rapidement une formule empirique, si l'on a les données suffisantes pour effectuer une vérification.

2° On adoptera comme estimation de la variance de $y - f$, d'après ce qui précède, l'expression :

$$\text{var}(y - \hat{f}) = \frac{\sum_{1 \leq i \leq n} (y_i - \hat{f}_i)^2}{n - k}$$

Le calcul suivant supposera encore que $\text{var.}(y - f)$ est indépendant de x .

Reprenons les notations du Titre I, Section 3. Admettons par la pensée que la population de taille N soit extraite au hasard d'une surpopulation homogène infinie. Supposons en outre que l'échantillon de taille n soit extrait, partie de la population de taille N pour n_1 éléments restant sur pied), et partie du restant de la surpopulation (pour les n_2 éléments détruits).

Désignons par \sum_1 toute sommation relative à la partie non détruite de l'échantillon, par \sum_2 toute sommation relative à la partie détruite de celui-ci, par \sum_0 et \sum_p respectivement des sommations relatives à l'échantillon et à la population.

On prend $\sum_p f_1$ comme estimation de $\sum_0 y_1$, et l'erreur commise est :

$$\epsilon = \sum_1 y_i - \sum_1 \hat{f}_i$$

Il est facile de voir que, si la fonction 1 figure parmi les fonctions g_k (la fonction 1 est définie par $g(x) = 1$ pour toute valeur de x), alors

$$\sum_0 \hat{f}_j = \sum_0 y_j$$

(c'est une des équations normales).

Supposons maintenant que :

$$\sum_p \hat{f}_i = \frac{N}{n} \sum_e \hat{f}_j,$$

hypothèse qui est vérifiée lorsque $\frac{N(x)}{n(x)}$ est une constante indépendante de x (Titre I, Section 3, § 2).

On aura :

$$\varepsilon = \sum_p y_i - \frac{N}{n} \sum_e \hat{f}_j,$$

puis :

$$\varepsilon = \sum_p y_i - \frac{N}{n} \sum_e y_j,$$

expression dont l'espérance est évidemment nulle, et que l'on peut écrire :

$$\varepsilon = \sum_p (y_i - \hat{f}_i) - \frac{N}{n} \sum_e (y_j - \hat{f}_j).$$

La variance de ε se calcule facilement :

$$\begin{aligned} \text{var}(\varepsilon) &= E \left\{ \left[\sum_p (y_i - \hat{f}_i) - \frac{N}{n} \sum_e (y_j - \hat{f}_j) \right]^2 \right\} \\ &= E \left\{ \left[\sum_p (y_i - \hat{f}_i) \right]^2 \right\} + \frac{N^2}{n^2} E \left\{ \left[\sum_e (y_j - \hat{f}_j) \right]^2 \right\} - 2 \frac{N}{n} E \left\{ \left[\sum_p (y_i - \hat{f}_i) \right] \left[\sum_e (y_j - \hat{f}_j) \right] \right\} \\ E \left\{ (y_i - \hat{f}_i)^2 \right\} &= \text{var} \cdot (y - \hat{f}) \\ E \left\{ (y_i - \hat{f}_i)(y_j - \hat{f}_j) \right\} &= 0 \text{ si } i \neq j \end{aligned}$$

Donc, en se souvenant que n_1 termes et n_1 seulement sont communs à \sum_p et \sum_e :

$$\text{var}(\varepsilon) = \left(n + \frac{N^2}{n^2} \cdot n - 2 \frac{N}{n} \cdot n_1 \right) \text{var} \cdot (y - \hat{f})$$

$$\text{var}(\varepsilon) = \frac{N^2}{n^2} \left(\frac{n}{N} + 1 - 2 \frac{n_1}{N} \right) \text{var} \cdot (y - \hat{f}),$$

et finalement :

$$\text{var}(\varepsilon) = \frac{N^2}{n^2} \left(1 + \frac{n_2 - n_1}{N} \right) \frac{n}{n-k} \sum_e (y_j - \hat{f}_j)^2$$

ANNEXE 4

1° Nous donnons ici les démonstrations, par des procédés élémentaires, des résultats énoncés au titre II, §§ 1 et 2. On se référera à la définition de la *t*-distribution de STUDENT, telle qu'elle est donnée par exemple dans CRAMER « *Mathematical Methods of Statistics* » 18.2. Nous adopterons les mêmes notations que cet auteur $E(y)$ pour désigner l'espérance mathématique de y et $D^2(z)$ pour la variance de z .

Soit un échantillon y_1, y_2, \dots, y_n de variables indépendantes normales (m, σ^2) . On pose :

$$\bar{y} = \sum_{i=1}^n \frac{1}{n} y_i$$

et l'on étudie la variable $z_i = y_i - \bar{y}$. L'expression $y_i - \bar{y}$ a sa valeur indépendante de l'origine choisie pour mesurer les y_i : on peut donc supposer sans inconvénient que l'origine est telle que $m = 0$.

Après ce changement d'origine, on peut écrire :

$$\begin{aligned} E(y_i) &= 0 & E(y_i^2) &= \sigma^2 \\ E(y_i y_j) &= E(y_i) E(y_j) = 0 & \text{pour } i \neq j \end{aligned}$$

d'où l'on déduit, en remarquant que :

$$z_i = \left(1 - \frac{1}{n}\right) y_i - \frac{1}{n} \sum_{j \neq i} y_j$$

$$E(z_i) = 0$$

$$D^2(z_i) = E(z_i^2) - [E(z_i)]^2 = \left(1 - \frac{1}{n}\right)^2 \sigma^2 + \frac{1}{n^2} \cdot (n-1) \sigma^2 = \frac{n-1}{n} \cdot \sigma^2$$

z_i est donc normale $\left(0, \frac{n-1}{n} \cdot \sigma^2\right)$ et $\sqrt{\frac{n}{n-1}} z_i$ est normale $(0, \sigma^2)$.

Dans un espace à n dimensions, l'échantillon est représenté par le vecteur \vec{y} dont les coordonnées sont (y_1, y_2, \dots, y_n) . On considère encore le vecteur \vec{Y} dont les coordonnées sont toutes égales à \bar{y} et le vecteur \vec{z} (z_1, z_2, \dots, z_n) . Rappelons que l'origine est telle que : $E(y_i) = 0$.

Désignons par $\|\mathbf{y}\|$ la longueur du vecteur \vec{y}

$$\|\mathbf{y}\|^2 = \sum y_i^2$$

Il est immédiat de vérifier que \vec{Y} est orthogonal à

$$\vec{z} = \vec{u} - \vec{v}$$

puisque le produit scalaire

$$\vec{v} \cdot (\vec{u} - \vec{v}) = \sum_{1 \leq i \leq n} \bar{y} (y_i - \bar{y}) = \bar{y} \sum_{1 \leq i \leq n} (y_i - \bar{y}) = 0.$$

On en déduit (théorème de Pythagore), en remarquant que

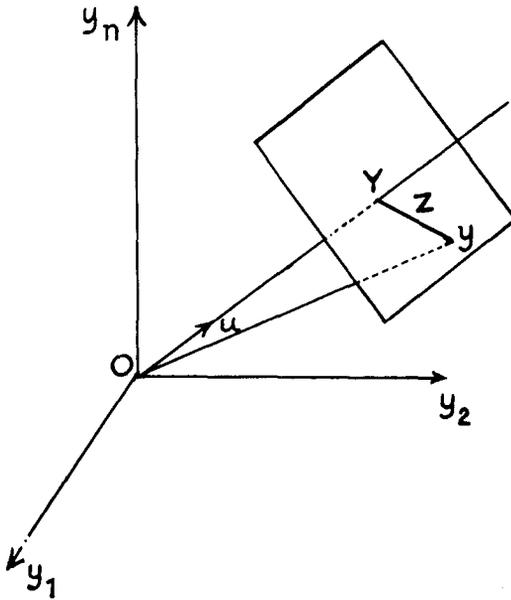
$$\vec{u} - \vec{v} = \vec{z} :$$

$$\|\mathbf{y}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{z}\|^2$$

d'où

$$r = \|\mathbf{z}\|^2 = \|\mathbf{y}\|^2 - n\bar{y}^2$$

avec les notations de (II, 1).



Z est un vecteur arbitraire d'un espace $(n - 1)$ dimensions (orthogonal au vecteur unitaire \vec{u} parallèle à \vec{Y} et dont les composan-

tes sont égales à $\frac{1}{n}$. Comme $\sqrt{n} \cdot \bar{y}$ est normale $(0, \sigma^2)$ en prenant \vec{u} comme n^{me} vecteur de coordonnées on trouve :

$$r = \|\mathbf{y}\|^2 - (\sqrt{n} \bar{y})^2$$

comme somme de $(n - 1)$ carrés de variables indépendantes normales $(0, \sigma^2)$. Ces variables sont les $(n - 1)$ composantes de \vec{Z} dans l'espace à $n - 1$ dimensions orthogonal à \vec{u} .

On a vu d'autre part que la variable

$$\sqrt{\frac{n}{n-1}} z_1$$

est normale $(0, \sigma^2)$ on peut la choisir comme $(n - 1)^{\text{me}}$ coordonnée pour mesurer r dans cet espace à $(n - 1)$ dimensions.

Alors
$$r = \frac{n}{n-1} z_1^2$$

est une somme de $(n - 2)$ carrés de variables normales $(0, \sigma^2)$ indépendantes entre elles et indépendantes de z_1 .

Il en résulte que la variable :

$$t_1 = \frac{\sqrt{\frac{n}{n-1}} z_1}{\sqrt{\frac{1}{n-2} (r - \frac{n}{n-1} z_1^2)}}$$

est une variable t de STUDENT à $(n - 2)$ degré de liberté. C'est le résultat du Titre II, § 1 (avec l'indice général i au lieu de l'indice particulier 1).

2° On va vérifier que r se met sous la forme (voir Titre II, § 2)

$$r = \sum_{2 \leq p \leq n} \frac{p}{p-1} (y_p - \bar{y}_p)^2$$

avec
$$\bar{y}_p = \frac{1}{p} \sum_{i \in J_p} y_i$$

et
$$r = \sum_{i \in I} n (y_i - \bar{y})^2$$

il suffit de montrer l'identité des coefficients des termes tels que y_a^2 et $y_a y_b$.

a) coefficient de y_q^2 .

Le terme en y_q^2 n'existe dans

$$(y_p - \bar{y}_p)^2 = \left[\left(1 - \frac{1}{p}\right) y_p - \frac{1}{p} \cdot \sum_{1 \leq j \leq p-1} y_j \right]^2$$

que pour $p \geq q$ et sa valeur est

$$\begin{cases} \left(1 - \frac{1}{q}\right)^2 & \text{si } p = q \\ \frac{1}{p^2} & \text{si } p > q \end{cases}$$

le coefficient de y_q^2 dans

$$\sum_{2 \leq p \leq n} \frac{p}{p-1} (y_p - \bar{y}_p)^2$$

est donc :

$$\begin{aligned} & \frac{q}{q-1} \left(1 - \frac{1}{q}\right)^2 + \sum_{q+1 \leq p \leq n} \frac{p}{p-1} \cdot \frac{1}{p^2} \\ & = 1 - \frac{1}{q} + \sum_{q+1 \leq p \leq n} \left(\frac{1}{p-1} - \frac{1}{p}\right) = 1 - \frac{1}{q} + \left(\frac{1}{q} - \frac{1}{n}\right) = 1 - \frac{1}{n} \end{aligned}$$

or dans r le coefficient de y_q^2 est précisément $1 - \frac{1}{n}$ comme il est

aisé de le voir

b) coefficient de $q_q y_s$ ($q > s$).

Le coefficient de $q_q y_s$ n'existe dans

$$(y_p - \bar{y}_p)^2$$

que si $p \geq q$ et sa valeur est :

$$\begin{cases} -2 \left(1 - \frac{1}{q}\right) \frac{1}{q} & \text{si } p = q \\ \frac{2}{p^2} & \text{si } p > q \end{cases}$$

le coefficient de $y_q y_s$ dans

$$\sum_{2 \leq p \leq n} \frac{p}{p-1} (y_p - \bar{y}_p)^2$$

est donc

$$\begin{aligned}
 &= \frac{2q}{q-1} \left(1 - \frac{1}{q}\right) \frac{1}{q} + 2 \sum_{q+1 \leq p \leq n} \frac{p}{p-1} \cdot \frac{1}{p^2} \\
 &= \frac{2}{q} + 2 \sum_{q+1 \leq p \leq n} \frac{1}{p(p-1)} = \frac{2}{n} \\
 &\text{ou } \frac{2}{n}
 \end{aligned}$$

est précisément le coefficient de $y_q y_s$ dans r .

La formule est donc établie.

3° La variable

$$Y_p = \sqrt{\frac{p}{p-1}} Z_p$$

est normale $(0, \sigma^2)$ car la démonstration est identique à celle qui a été donnée pour

$$\sqrt{\frac{n}{n-1}} Z_1,$$

au § I. On considère les variables

$$t_p = \sqrt{p} \frac{Y_{p+2}}{\sqrt{Y_2^2 + Y_3^2 + \dots + Y_{p+1}^2}} \quad 1 \leq p \leq n-2$$

que l'on se propose de démontrer être distribuées dans des t -distributions indépendantes dont le degré de liberté est égal à l'indice p .

a) les variables Y_p sont normales $(0, \sigma^2)$ comme on vient de le voir.
 b) les variables normales Y_p sont indépendantes. Il suffit de montrer l'orthogonalité, c'est-à-dire:

$$E \{ Y_q Y_s \} = 0$$

quels que soient q et s avec $2 \leq s < q \leq n$
 c'est-à-dire encore:

$$E \{ Z_q Z_s \} = 0.$$

En appliquant la propriété:

$$E \{ y_i \cdot y_j \} = \begin{cases} 0 & \text{si } i \neq j \\ \sigma^2 & \text{si } i = j \end{cases}$$

on trouve successivement

$$\begin{aligned}
 E \{ \gamma_q \gamma_s \} &= E \{ (y_q - \bar{y}_q) (y_s - \bar{y}_s) \} \\
 &= E \left\{ \left[\left(1 - \frac{1}{q}\right) y_q - \frac{1}{q} \sum_{1 \leq i < q} y_i \right] \left[\left(1 - \frac{1}{s}\right) y_s - \frac{1}{s} \sum_{1 \leq j < s} y_j \right] \right\} \\
 &= -\frac{1}{q} \left(1 - \frac{1}{s}\right) \sigma^2 + \frac{1}{sq} (s-1) \sigma^2 = 0
 \end{aligned}$$

$$E \{ \gamma_q \gamma_s \} = 0$$

démontre l'orthogonalité énoncée.

c) t_p est distribuée dans une t-distribution à p degrés de liberté ainsi qu'il résulte de a) et de b) (1).

d) les variables t_p ($1 \leq p \leq n - 2$) sont indépendantes. on posera :

$$u_p = \frac{t_p}{\sqrt{p}}$$

d'où successivement :

$$\begin{aligned}
 \gamma_3 &= u_1 \sqrt{\gamma_2^2} \\
 \gamma_4 &= u_2 \sqrt{\gamma_2^2 + \gamma_3^2} \\
 &\vdots \\
 \gamma_p &= u_{p-2} \sqrt{\gamma_2^2 + \gamma_3^2 + \dots + \gamma_{p-1}^2} \\
 &\vdots \\
 \gamma_n &= u_{n-2} \sqrt{\gamma_2^2 + \gamma_3^2 + \dots + \gamma_{n-1}^2}
 \end{aligned}$$

En posant

$$\sqrt{\gamma_2^2} = \lambda$$

on démontre par récurrence la formule

$$\gamma_p = \lambda u_{p-2} \sqrt{1 + u_1^2} \sqrt{1 + u_2^2} \dots \sqrt{1 + u_{p-3}^2}$$

(1) Il suffit de se rappeler que la variable t de STUDENT n'est autre que le rapport d'une variable normale réduite et de la racine d'un χ^2 divisé par son nombre de degrés de liberté et indépendant de la variable du numérateur. Ici σ que l'on aurait du reste pu, par un simple changement d'unité, prendre égal à 1 disparaît au numérateur et au dénominateur de la fraction.

Tout d'abord, la formule est évidemment vraie pour $p = 3$.

Ensuite, démontrons que si elle est vraie jusqu'au rang p elle l'est aussi jusqu'au rang $p + 1$.

$$\begin{aligned} \gamma_{p+1} &= u_{p-1} \sqrt{(\gamma_2^2 + \gamma_3^2 + \dots + \gamma_{p-1}^2) + \gamma_p^2} \\ &= u_{p-1} \sqrt{(\gamma_2^2 + \gamma_3^2 + \dots + \gamma_{p-1}^2) + u_{p-2}^2 (\gamma_2^2 + \gamma_3^2 + \dots + \gamma_{p-1}^2)} \\ &= u_{p-1} \sqrt{\gamma_2^2 + \gamma_3^2 + \dots + \gamma_{p-1}^2} \sqrt{1 + u_{p-2}^2} \end{aligned}$$

d'où finalement

$$\gamma_{p+1} = \lambda u_{p-1} \sqrt{1 + u_1^2} \sqrt{1 + u_2^2} \dots \sqrt{1 + u_{p-3}^2} \sqrt{1 + u_{p-2}^2}$$

Nous avons donc, en gardant provisoirement Y_2 le tableau de changement de variables suivant :

$$\left\{ \begin{aligned} \gamma_3 &= \lambda u_1 \\ \gamma_4 &= \lambda \sqrt{1 + u_1^2} \cdot u_2 \\ \gamma_5 &= \lambda \sqrt{1 + u_1^2} \sqrt{1 + u_2^2} \cdot u_3 \\ &\vdots \\ &\vdots \\ \gamma_p &= \lambda \sqrt{1 + u_1^2} \sqrt{1 + u_2^2} \dots \sqrt{1 + u_{p-3}^2} \cdot u_{p-2} \\ &\vdots \\ &\vdots \\ \gamma_n &= \lambda \sqrt{1 + u_1^2} \sqrt{1 + u_2^2} \dots \sqrt{1 + u_{n-3}^2} \cdot u_{n-2} \end{aligned} \right.$$

Un calcul simple de différentielles extérieures (ou encore la formule de changement de variables par la méthode du jacobien) donne

$$\begin{aligned} d\gamma_2 d\gamma_3 &= \lambda dY_2 du_1 \\ dY_2 d\gamma_3 d\gamma_4 &= \lambda^2 dY_2 du_1 (1 + u_1^2)^{1/2} du_2 = \lambda^2 (1 + u_1^2)^{1/2} dY_2 du_1 du_2 \\ dY_2 d\gamma_3 d\gamma_4 &= \lambda^3 (1 + u_1^2)^{1/2} dY_2 du_1 du_2 (1 + u_1^2)^{1/2} (1 + u_2^2)^{1/2} du_3 \\ &= \lambda^3 (1 + u_1^2)(1 + u_2^2)^{1/2} dY_2 du_1 du_2 du_3 \end{aligned}$$

On va montrer par récurrence la formule :

$$\begin{aligned} dY_2 d\gamma_3 \dots d\gamma_p &= \lambda^{p-2} (1 + u_1^2)^{\frac{p-3}{2}} (1 + u_2^2)^{\frac{p-4}{2}} \dots (1 + u_{p-2}^2)^{\frac{p-9}{2}} (1 + u_{p-3}^2)^{1/2} \dots \\ &\dots dY_2 du_1 du_2 \dots du_{p-3} \end{aligned}$$

Tout d'abord la formule est évidemment vraie pour $p = 3$.

Ensuite supposons la formule vraie jusqu'au rang p. On a :

$$\begin{aligned} dY_2 dY_3 \dots dY_p dY_{p+1} &= \left[\lambda^{p-2} (1+u_1^2)^{\frac{p-3}{2}} \cdot (1+u_2^2)^{\frac{p-4}{2}} \dots (1+u_{p-3}^2)^{1/2} dY_2 du_1 \dots du_{p-2} \right] \\ &\quad \cdot \left[\lambda (1+u_1^2)^{1/2} (1+u_2^2)^{1/2} \dots (1+u_{p-3}^2)^{1/2} (1+u_{p-2}^2)^{1/2} du_{p-1} \right] \\ &= \lambda^{p-1} (1+u_1^2)^{\frac{p-2}{2}} (1+u_2^2)^{\frac{p-3}{2}} \dots (1+u_{p-3}^2) (1+u_{p-2}^2)^{1/2} dY_2 du_1 du_2 \dots du_{p-1} \end{aligned}$$

On a donc :

$$dY_2 dY_3 \dots dY_n = \lambda^{n-2} (1+u_1^2)^{\frac{n-3}{2}} (1+u_2^2)^{\frac{n-4}{2}} \dots (1+u_{n-3}^2)^{1/2} dY_2 du_1 du_2 \dots du_{n-2}$$

On va calculer la loi de probabilité des variables conjointes u_1, u_2, \dots, u_{n-2} .

En se référant encore à CRAMER, on est amené à calculer :

$$e^{-\frac{1}{2}(Y_2^2 + Y_3^2 + \dots + Y_n^2)} dY_2 dY_3 \dots dY_n$$

Or, il résulte des formules déjà démontrées que :

$$\frac{Y_n^2}{u_{n-2}^2} = Y_2^2 + Y_3^2 + \dots + Y_{n-1}^2 = \lambda^2 (1+u_1^2)(1+u_2^2) \dots (1+u_{n-3}^2)$$

ce qui donne : $Y_2^2 + Y_3^2 + \dots + Y_{n-1}^2 + Y_n^2 = \lambda^2 (1+u_1^2)(1+u_2^2) \dots (1+u_{n-2}^2)$

ainsi :

$$\begin{aligned} &e^{-\frac{1}{2}(Y_2^2 + Y_3^2 + \dots + Y_n^2)} dY_2 dY_3 \dots dY_n = \\ &= e^{-\frac{Y_2^2}{2} (1+u_1^2)(1+u_2^2) \dots (1+u_{n-2}^2)} \cdot Y_2^{n-2} (1+u_1^2)^{\frac{n-3}{2}} (1+u_2^2)^{\frac{n-4}{2}} \dots \\ &\quad \dots (1+u_{n-3}^2)^{1/2} dY_2 du_1 du_2 \dots du_{n-2} \end{aligned}$$

On va effectuer une intégration partielle relativement à la variable Y_2 pour les bornes $-\infty$ et $+\infty$. On a, en posant $v = hY_2^2$.

$$\int_{-\infty}^{+\infty} e^{-hy_2^2} y_2^{n-2} dy_2 = \int_{-\infty}^{+\infty} e^{-v} \left(\frac{v}{h}\right)^{\frac{n-2}{2}} \left(\frac{v}{h}\right)^{-\frac{1}{2}} \frac{dv}{h}$$

$$= \frac{1}{h} \int_0^{+\infty} e^{-v} v^{\frac{n-1}{2} - 1} dv$$

C'est-à-dire: $\frac{1}{h} \frac{\Gamma(\frac{n-1}{2})}{h^2}$

à un facteur numérique près ne dépendant que de n (ce facteur est

$$\Gamma\left(\frac{n-1}{2}\right)$$

où Γ désigne une des fonctions eulériennes, mais sa valeur ne nous intéresse pas).

Ainsi, à un facteur constant près la densité de probabilité des variables conjointes u_1, u_2, \dots, u_{n-2} est:

$$\frac{(1+u_1^2)^{\frac{n-3}{2}} (1+u_2^2)^{\frac{n-4}{2}} \dots (1+u_{n-3}^2)^{\frac{1}{2}}}{[(1+u_1^2)(1+u_2^2) \dots (1+u_{n-2}^2)]^{\frac{n-1}{2}}}$$

$$= \frac{1}{(1+u_1^2)^{2/2}} \cdot \frac{1}{(1+u_2^2)^{3/2}} \dots \frac{1}{(1+u_{n-2}^2)^{\frac{n-1}{2}}}$$

L'indépendance est maintenant évidente et il serait aisé de vérifier à nouveau que $t_p = \sqrt{p} u_p$ est une variable t de STUDENT à p degrés de liberté.

ANNEXE 5 (voir Titre II)

L'extrait suivant du tableau 1 indique le calcul de t pour Haguénau, classe I A, circonférence 72 ($n = 8$). L'arbre tiré au hasard est indiqué par une croix : c'est le n° 3 ($i = 3$).

TABLEAU I (Extrait)

C	V	$V - \frac{\Sigma V}{n}$	$V - \bar{V}$	$(V - \bar{V})^2$	$r = \Sigma V (V - \bar{V})^2$	$(n-1)r$	nz^2	$(n-1)r - nz^2_3$	$\frac{nz^2_1}{(n-1)r - nz^2_3}$	$\frac{(n-2)nz^2_3}{(n-1)r - nz^2_3} = t^2_3$	t_3
72	435		26,3	691,69							
	455,5		5,8	33,64							
	+ 475		13,7	187,69							
	470,5		18,2	331,24							
	502,5		41,2	1697,44							
	437,5		23,8	566,44							
	441,5		19,8	392,04							
	464		2,7	7,29							
	461,3				3 807,47	27 282,20	1 501,52	25 780,77	0,05824	0,34044	0,591

Ce nombre 0,591 donne une barre dans le tableau 2, ainsi qu'il a été indiqué en II § 4 entre 0,6 et 0,5. En réalité, le tableau 2 suivant donne le nombre total de barres ainsi trouvées pour l'ensemble des 7 places.

TABLEAU 2

P	1	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0	Total
y_i		12	15	5	8	11	14	11	9	12	8	105
y'_i		10,5	10,5	10,5	10,5	10,5	10,5	10,5	10,5	10,5	10,5	105
$y_i - y'_i$		1,5	4,5	- 5,5	- 2,5	0,5	3,5	0,5	- 1,5	1,5	- 2,5	0
$(y_i - y'_i)^2$		2,25	20,25	30,25	6,25	0,25	12,25	0,25	2,25	2,25	6,25	82,5

L'application du χ^2 (II, § 5) donne lieu au calcul suivant : $\chi^2 = \sum \frac{(y_i - y'_i)^2}{y_i} = \frac{82,50}{10,5} = 7,857$ à 9 degrés de liberté.

La table 7 de PEARSON and HARTLEY (*Biometrika Tables*) montre que la probabilité correspondante est approximativement 55 % (55,442 % pour $\chi^2 = 7,8$).

ANNEXE 6 (voir titre II)

L'extrait du tableau 1 indique le calcul de t_1, t_2, \dots, t_6 pour Ha guenau, classe 1 A, circonférence 72. La succession des tiges est assurée par le hasard, en utilisant les tables XXXIII et XXXIII-1 de FISHER and YATES. Pour les notations, se référer au II, § 2, en remplaçant x par C et y par V.

TABLEAU 1 (extrait)

C	V	Sommation de V	p	\bar{V}_p	$Z_p = V - \bar{V}_p$	Z_p^2	$\frac{p}{p-1} Z_p^2$	r_{p-1}	$t_{p-2}^2 = (p-2) \frac{\frac{p}{p-1} Z_p^2}{r_{p-1}}$	t_{p-2}	P
72	437,5	437,5	1	—	—	—	—	—	—	—	—
	502,5	940	2	470	32,5	1 056,25	2 112,50	—	—	—	—
	441,5	1 381,5	3	460,5	— 19	361	541,50	2 112,50	0,256331	0,5063	8-7
	475	1 856,5	4	464,1	10,9	118,81	158,41	2 654	0,119374	0,3455	8-7
	464	2 320,5	5	464,1	— 0,1	0,01	0,01	2 812,41	0,000018	0,0042	10-9
	455,5	2 776	6	462,6	— 7,1	50,41	60,49	2 812,42	0,086032	0,2933	8-7
	479,5	3 255,5	7	465	14,5	210,25	245,29	2 872,91	0,426900	0,6534	6-5
	435	3 690,5	8	461,3	— 26,3	691,69	700,50	3 118,20	1,521072	1,2330	3-2

On lit dans l'avant dernière colonne :

$t_1 = 0,5063$ à 1 degré de liberté, soit P entre 0,8 et 0,7 d'après la table III de FISHER and YATES.

$t_2 = 0,3455$ à 2 degrés de liberté, soit P entre 0,8 et 0,7, etc...

Chaque valeur de P permet de mettre une barre dans une colonne du tableau 2, ainsi qu'il a été indiqué en (IV, § 4). Le tableau 2 suivant contient en fait le nombre total de barres ainsi trouvées dans les différentes colonnes, pour l'ensemble des 7 places, et se continue par le calcul du χ^2 .

TABLEAU 2

P	1	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0	Total
v_i	18	24	25	21	19	19	22	16	15	13		192
v'_i	19,2	19,2	19,2	19,2	19,2	19,2	19,2	19,2	19,2	19,2		192
$v_i - v'_i$	— 1,2	4,8	5,8	1,8	— 0,2	— 0,2	2,8	— 3,2	— 4,2	— 6,2		0
$(v_i - v'_i)^2$	1,44	23,04	33,64	3,24	0,04	0,04	7,84	10,24	17,64	38,44		135,60

On en déduit :
$$\chi^2 = \sum \frac{(v_i - v'_i)^2}{v'_i} = \frac{135,60}{19,2} = 7,0625 \text{ à } 9 \text{ degrés de liberté.}$$

La table 7 de PEARSON and HARTLEY (Biometrika Tables) donne pour la probabilité correspondante, approximativement 63 % (P = 63,7 % pour $\chi^2 = 7$).

ANNEXE 7 (voir Titre III)

Prenons pour exemple la comparaison classe 1 - classe 2. Considérons plus particulièrement Haguenau, classe 1A - classe 2A, cir conférence 76, dont voici les volumes des tiges correspondantes :

1 A	2 A
500,5	478
509,5	517
466,5	
556,5	

Les quantités \bar{V}_1 , \bar{V}_2 , r_1 , r_2 ont déjà été calculées : voir Annexe 5, l'indice 1 étant relatif à la classe 1 A, l'indice 2 à la classe 2 A, les notations restant à part cela les mêmes. L'extrait suivant du tableau 1 donne les calculs préliminaires, qui ont été effectués pour toutes les comparaisons à l'intérieur d'une même place :

TABLEAU I (extrait)

C	n_1	n_2	$[n_1, n_2] = \frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}$	\bar{V}_1	\bar{V}_2	$\bar{V}_1 - \bar{V}_2$	r_1	r_2	$r_1 + r_2$	$\frac{r_1 + n_0}{[n_1, n_2]}$	$\sqrt{\frac{r_1 + r_2}{[n_1, n_2]}}$	$t = \frac{\bar{V}_1 - \bar{V}_2}{\sqrt{\frac{r_1 + r_2}{[n_1, n_2]}}}$
76	4	2	5,333333	508,2	497,5	+ 10,7	4 132,76	760,50	4893,26	917,48	30,29	+ 0,353

A la valeur de 0,353 de t, à $n_1 + n_2 - 2 = 4$ degrés de liberté, correspond par la table III de FISHER and YATES une probabilité P comprise entre 0,8 et 0,7, ce qui permet de mettre une barre, comme on l'a déjà vu, aux Annexes 5 et 6, dans le tableau 2. Le tableau suivant indique le nombre total de barres ainsi que le calcul de χ^2 (on a placé à gauche de 1 les t négatifs, à droite les t positifs).

TABLEAU II

Comparaison Classe 1 - Classe 2

P	-			+			Total
	0	0,4	0,8	1	0,8	0,4	
v_i	3	5	14	12	23		57
v'_i	11,4	11,4	11,4	11,4	11,4		57
$v_i - v'_i$	-8,4	-6,4	2,6	0,6	11,6		0
$(v_i - v'_i)^2$	70,56	40,96	6,76	0,36	134,56		272,3

$$\chi^2 = \sum \frac{(v_i - v'_i)^2}{v'_i} = \frac{272,3}{11,4} = 24 \text{ à } 4 \text{ degrés de liberté.}$$

La table 7 de PEARSON and HARTLEY (*Biometrika Tables*) donne $P = 0,008 \%$.

Les calculs pour la comparaison classe A - classe B sont analogues aux précédents, et nous ne les reproduisons pas.

Voici le calcul du χ^2 du Titre III § 7.

Places	v_i Nombre de signes +	Nombre de comparaisons	v'_i Nombre théorique	$v_i - v'_i$	$(v_i - v'_i)^2$	$\frac{(v_i - v'_i)^2}{v'_i}$
Hagenau	16	37	18,5	- 2,5	6,25	0,337
St-Avoid 134 b	25	38	19	6	36	1,894
St-Amond 47	8	17	8,5	- 0,5	0,25	0,029
St-Avoid 156	13	30	15	- 2	4	0,266
St-Amond 46	9	10	5	4	16	3,200
Porcellette	14	22	11	3	9	0,818
						6,544

Les comparaisons par espèces (III; § 8) donnent lieu à des calculs analogues, que nous ne reproduisons pas.

ANNEXE 8

Les calculs de la régression $V = a + bC^2$ sont présentés en 4 tableaux qui se comprennent d'eux-mêmes. La méthode exige :

- 1° une table de carrés et de puissance quatrième ;
- 2° une machine automatique effectuant aisément les multiplications.

Dans le tableau 1, les valeurs de C^4 ont été écrites en négligeant les trois derniers chiffres. Une correction additive égale à $500n$ a été effectuée sur la valeur obtenue de $\sum C^4$.

La colonne (5) n'est pas remplie, car notre machine additionne automatiquement, si on le désire, les résultats des multiplications successives. Il en est de même pour la colonne (6).

Le tableau 2 indique la résolution des équations normales, et le tableau 3 est le calcul de l'erreur.

Le tableau 4 donne le volume total de la placette par application de la formule :

$$V = Na + b \sum_{n_1} C_1^2$$

la sommation étant étendue à la population restant sur pied.

Le volume estimé, qui est $624,559 \text{ m}^3$, n'a pu être comparé au volume réel qui n'est pas connu pour la place étudiée ici.

Nota. — La place de la Bénague est une place d'expériences permanente de la 1^{re} Section dans les Pyrénées (Sapin).

ANNEXE n° 8 - PLACETTE 3 - TABLEAU 1

La Bénague

Sapins unitarii

Colonne (3) C¹ Correction de 500 × n (58 tiges) = 29 000.

C (1)	C ² (2)	C ¹ (3)	Volume (4)	(2) × (4) (5)	V ² Carré de (4) (6)
<i>Abattus</i>					
89	7 921	62 742	757,5		
95	7 921	81 450	847		
95	9 025	81 450	835,5		
95,5	9 120	83 174	832,5		
102	10 404	108 243	1 134,5		
104	10 816	116 985	1 022,5		
106	11 236	126 247	1 168		
107	11 449	131 079	1 156		
110	12 100	146 410	1 305,5		
111	12 321	151 807	1 233,75		
111	12 321	151 807	1 327,5		
111	12 321	151 807	1 186		
118	13 924	193 877	1 469		
120	14 400	207 360	1 576		
136	18 496	342 102	2 024		
137	18 769	352 275	2 050,5		
/					
16					
<i>Debout</i>					
79	6 241	38 950	580		
84	7 056	49 787	680,5		
105	11 025	121 550	1 217,5		
105	11 025	121 550	1 395		
107	11 449	131 079	1 295		
108	11 664	136 048	1 300		
115	13 225	174 900	1 210		
115	13 225	174 900	1 497,5		
117	13 689	187 388	1 545,5		
119	14 161	200 533	1 451		
120	14 400	207 360	1 683		
120	14 400	207 360	1 564		
121	14 641	214 358	1 494,5		
122	14 884	221 533	1 693		
125	15 625	244 140	1 687,5		
129	16 641	276 922	1 757,5		
129	16 641	276 922	1 475		
130	16 900	285 610	1 790		
131	17 161	294 499	1 736,5		
132	17 424	303 595	1 914		
132	17 424	303 595	1 872,5		
132	17 424	303 595	1 731,5		
134	17 956	322 417	1 987		
135	18 225	332 150	2 079		
135	18 225	332 150	2 062		
136	18 496	342 102	2 001		
138	19 044	362 673	2 069		
138	19 044	362 673	2 109		
139	19 321	373 301	1 897		
141	19 881	395 254	2 040		
141	19 881	395 254	2 315		
143	20 449	418 161	2 243		
143	20 449	418 161	2 166,5		
146	21 316	454 371	2 536,5		
146	21 316	454 371	2 479		
148	21 904	479 785	2 569		
148	21 904	479 785	2 469,5		
154	23 716	562 448	2 278		
157	24 649	607 573	2 664		
162	26 244	688 747	2 869		
164	26 896	723 394	2 605		
180	32 400	1 049 760	3 749		
/					
42					
58	931 289	16 519 518 000	99 603,25	1 777 937 998	192 230 457
↓ n	Σ C ²	Σ C ¹ après correction	Σ V	Σ V C ²	Σ V ²

ANNEXE n° 8 — TABLEAU 2

I.a Bénague — Placette 3

Sapins unitarif

a	b	c	
58 (n)	931 289 (ΣC^2)	99 603,25 (ΣV)	
931 289 (ΣC^2)	16 519 548 000 (ΣC^4)	1 777 937 998 ($\Sigma V C^2$)	
I	16 056,70680	1 717,297414	a + 16 056,70680 b = 1 717,297414
I	17 738,36913		a = - 114,200976
	1 681,66233	191,817796	
	b	0,114064395	

ANNEXE 8 — TABLEAU 3

La Bénague — Placette 3

Sapins unitarif

Calcul de ET

$b = 0,114064395$
 $a = -114,200976$

$n_2 = 16$	$n_1 = 42$	$n = 58$	$N = 355$
$\frac{n-2}{n} = \frac{56}{58} = 0,965517$	$\frac{N}{N+n_2-n_1} = \frac{355}{329} = 1,079027$		$\frac{N}{n} = \frac{355}{58} = 6,120$
$\Sigma V^2 \dots\dots\dots = 192\ 230\ 457$			
$\Sigma V \times (-a) \dots = 11\ 374\ 759$			
$\Sigma V C^2 \times (-b) \dots = -202\ 799\ 397$			

r

805 919

↓ diviser par
 $\frac{n-2}{n}$

834 702

↓ diviser par g
 N

$\frac{834\ 702}{N + n_2 - n_1}$

773 569

↓ √

880

↓

multiplier par $\frac{N}{n}$

ET =

5386

ANNEXE n° 8 -- TABLEAU 4

La Bénague -- Placette 3

Sapins unitarif

Cir- conférence C_i	C_i^2	Nombre d'arbres n_i	Cir- conférence C_i	C_i^2	Nombre d'arbres n_i	Cir- conférence C_i	C_i^2	Nombre d'arbres n_i
63	3 969	1	114	12 996	4	147	21 609	3
68	4 624	1	115	13 225	5	148	21 904	5
70	4 900	1	116	13 456	3	149	22 201	1
76	5 776	1	117	13 689	8	150	22 500	1
77	5 929	1	118	13 924	6	151	22 801	7
80	6 400	1	119	14 161	6	152	23 104	3
81	6 561	1	120	14 400	6	153	23 409	2
82	6 724	1	121	14 641	6	154	23 716	2
83	6 889	3	122	14 884	5	155	24 025	4
84	7 056	2	123	15 129	6	156	24 336	3
89	7 921	1	124	15 376	6	157	24 649	2
90	8 100	1	125	15 625	5	159	25 281	1
91	8 281	3	126	15 876	5	161	25 921	1
93	8 649	2	127	16 129	4	163	26 569	2
94	8 836	2	128	16 384	8	164	26 896	2
95	9 025	1	129	16 641	8	165	27 225	1
97	9 409	2	130	16 900	7	166	27 556	3
98	9 604	2	131	17 161	10	170	28 900	1
99	9 801	2	132	17 424	5	171	29 241	1
100	10 000	3	133	17 689	10	173	29 929	1
101	10 201	2	134	17 956	6	175	30 625	1
102	10 404	5	135	18 225	8	178	31 684	2
103	10 609	9	136	18 496	5	179	32 041	1
104	10 816	4	137	18 769	9	181	32 761	1
105	11 025	8	138	19 044	6	183	33 489	1
106	11 236	3	139	19 321	4	193	37 249	1
107	11 449	6	140	19 600	8	198	39 204	1
108	11 664	1	141	19 881	3			
109	11 881	6	142	20 164	8		5 830 928	355
110	12 100	7	143	20 449	6			
111	12 321	5	144	20 736	3		$\Sigma n_i C_i^2$	$\Sigma n_i = N$
112	12 544	5	145	21 025	7			
113	12 769	10	146	21 316	1			

$a = \dots 114,200976$; $b = 0,114064395$
 $N = 355$
 $n_i =$ nombre d'arbres par circonférence C_i
 $V_{\text{unitarif}} = b \Sigma n_i C_i^2 + N a$
 $V_{\text{unitarif}} = 0,114064395 \times 5 830 928 + 114,200976$
 $= 624 559 \text{ dm}^3$

ANNEXE 9

Estimation d'un volume à l'aide de 2 classes

On a suivi ici, sur St-Avoid 157, la méthode donnant comme régression :

$$V = a_1 + b C^2 \text{ pour les tiges de classe 1}$$

$$V = a_2 + b C^2 \text{ pour les tiges de classe 2 ou 3.}$$

Le calcul de l'erreur-type a été fait par les deux méthodes indiquées respectivement au Titre V, §§ 5 et 6 (tableaux 3 A et 3 B).

La résolution du système des équations normales se fait par une méthode d'élimination aisée à retrouver sur le tableau 2

L'erreur-type trouvée est 1,68 %. Avec le même échantillon, mais en groupant comme dans l'Annexe 8, tous les arbres en une seule classe, l'erreur-type est 1,93 % : l'amélioration apportée par la distinction en deux classes qui est donc ici sensible, ne s'est pas révélée aussi importante dans les autres places.

Ce problème a du reste été étudié par P. ARBONNIER dans son rapport déjà cité au XII^e Congrès de l'Union Internationale des Instituts de Recherches.

ANNEXE n° 9 — TABLEAU 1

St-Avoid 156

Epicéas (2 classes)

C	C ²	C ⁴ 3 chiffres négligés remplacés par 000	V	(2) × (4)	V ²
(1)	(2)	(3)	(4)	(5)	(6)
CLASSE 1					
<i>Abattus</i>					
78	6 084	37 015	710		
82	6 724	35 212	812,5		
89	7 921	62 742	830,5		
94	8 836	78 074	986		
99	9 801	96 059	1 155		
101	10 201	104 060	1 114		
109	11 881	141 158	1 196,5		
116	13 456	181 063	1 450,5		
/					
8					
n ab					
(1)					
<i>Debout</i>					
77	5 929	35 153	656		
77	5 929	35 153	693		
83	6 889	47 458	807		
86	7 396	54 700	843,5		
93	8 649	74 805	1 015		
97	9 409	88 529	996,5		
101	10 201	104 060	1 050		
105	11 025	121 550	783,5		
107	11 449	131 079	1 336		
114	12 996	168 896	1 346		
115	13 225	174 900	1 370		
118	13 924	193 877	1 533		
121	14 641	214 358	1 449		
125	15 625	244 140	1 439		
128	16 384	268 435	1 878		
131	17 161	294 499	1 885		
135	18 225	332 150	2 073,5		
140	19 600	384 160	1 846,5		
147	21 609	466 948	2 018,5		
/					
19					
27	315 170	4 180 246 500	32 280	435 539 784	45 284 482,50
n deb.	ΣC^2	ΣC^4	ΣV	$\Sigma V C^2$	ΣV^2
(1)	1	1	1	1	1
CLASSE 2					
<i>Abattus</i>					
40	2 401	5 764	188,5		
62	3 844	14 776	371,5		
69	4 761	22 667	378		
71	5 041	25 411	502,5		
75	5 625	31 640	560,5		
81	6 561	43 046	782,5		
88	7 744	59 969	633		
/					
7					
n ab					
(2)					
<i>Debout</i>					
56	3 136	9 834	254		
74	5 476	29 986	515,5		
78	6 084	37 015	530		
87	7 569	57 289	652,5		
96	9 216	84 934	688,5		
/					
5					
12	67 458	422 337 000	6 057	37 530 072,5	3 407 035
n2 déb.	ΣC^2	ΣC^4	ΣV	$\Sigma V C^2$	ΣV^2
2	2	2	2	2	2
		4 602 583 500		473 069 856,5	49 231 517,5
		ΣC^4 corrigé		$\Sigma V C^2$	ΣV^2

NOTA. — La classe 2 comprend les classes 2 et 3 bloquées.

ANNEXE 9 — TABLEAU 2

St-Avoid 156

Epiciés (2 classes)

b	a (1)	a (2)	
ΣC^2 (1)	u_1	0	ΣV (1)
315 170	27		33 280
ΣC^2 (2)	0	u_2	ΣV (2)
67 458		12	6 057
ΣC^4	ΣC^2 (1)	ΣC^2 (2)	$\Sigma V C^2$
4 602 583 500	315 170	67 458	473 069 856,5
11 672 962 98	1	0	1 232,592 593
5 621,50000	0	1	504,750000
3 678 967 742 379 215 147	315 170	67 458	388 576 196 34 049 425,50
$544\ 400\ 611 = u$ (u)			$50\ 544\ 235 = v$ (v)
			$b = \frac{v}{u}$
			$a_1 = 1\ 232,592\ 593$ $a_1 = 148,830085$ $a_2 = 504,750000$ $a_2 = 17,175617$

NOTA: L'indice (1) indique la classe 1 L'indice (2) indique la classe 2

ANNEXE n° 9 — TABLEAU 3 A

St-Avoid 156

Epicéas (2 classes)

Calcul de l'erreur par la formule générale du Titre II § 5

Valeur des données

$a_1 = 148,830085$	$a_2 = - 17,175617$	$b = 0,092843825$
$n' = 15$	$n'' = 24$	$n = 39 \quad N = 145$
$\frac{n-3}{n} = \frac{36}{39} = 0,923077$	$\frac{N}{N + n' - n''} = \frac{145}{145 + 15 - 24} = \frac{145}{136} = 1,066176$	$\frac{145}{36} = 1,066176$
$\frac{N}{n} = \frac{145}{39} = 3,717948$		

$$Q = \sum V^2 - a_1 \sum_1 V - a_2 \sum_2 V + b \sum V C^2$$

$$= + 49 231 517,5$$

$$- 4 953 065,2$$

$$+ 104 032,7$$

$$- 43 921 614,7$$

$$Q = \boxed{460 870}$$

diviser par $\frac{n-3}{n}$



$$\boxed{499 276}$$

diviser par $\frac{N}{N + n' - n''}$



$$\boxed{468 286}$$

$$\sqrt{468 286}$$



$$\boxed{684,3}$$

$$\boxed{2 544}$$

$$\times \frac{N}{n}$$

$$\leftarrow \boxed{684,3}$$

$$\text{Et } \% = \frac{2 544 \times 100}{149 180} = 1,71 \%$$

ANNEXE n° 9 - TABLEAU 3 B

St-Avoid 156

Épicéas (2 classes)

Calcul de l'erreur par la méthode du titre II § 6

Valeur des données

$$\begin{aligned}
 a_1 &= 148,830085 & a_2 &= 17,175617 & b &= 0,092843825 \\
 n'_1 &= 8 & n''_1 &= 19 & n_1 &= 27 & N_1 &= 102 \\
 n'_2 &= 7 & n''_2 &= 5 & n_2 &= 12 & N_2 &= 43 \\
 n &= n_1 + n_2 = 27 + 12 = 39 & N &= N_1 + N_2 = 102 + 43 = 145
 \end{aligned}$$

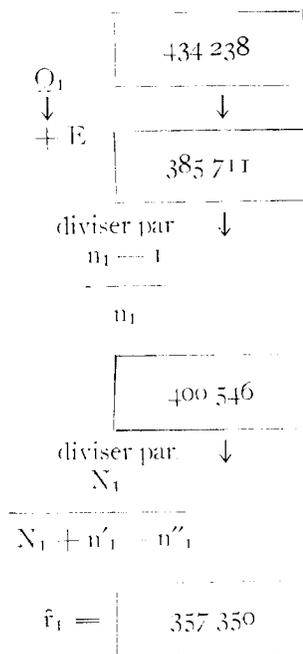
$$\frac{N}{n} = \frac{145}{39} = 3,717948$$

$\frac{n_1 - 1}{n_1} = \frac{26}{27} = 0,962962$	$\frac{n_2 - 1}{n_2} = \frac{11}{12} = 0,916666$	$\frac{n - 1}{n} = \frac{38}{39} = 0,974358$
$\frac{N_1}{N_1 + n'_1 - n''_1} = \frac{102}{102 + 8 - 9} = 1,120879$	$\frac{N_2}{N_2 + n'_2 - n''_2} = \frac{43}{43 + 7 - 5} = 0,955555$	

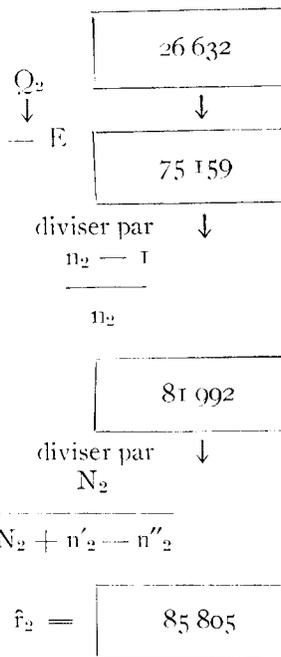
Terme correctif: $E = b [\sum_2 VC^2 - a_2 \sum_2 C^2 - b \sum_2 C^4] = b \times F$

$$\begin{aligned}
 &37\,530\,072,5 \\
 &1\,158\,632,8 \\
 &-- 39211\,382,5 \\
 \hline
 &35\,526\,722,8 \\
 &522\,677 \times 0,092843825 = 48\,527
 \end{aligned}$$

Classe (1) $\left\{ \begin{aligned} \sum_1 V^2 &= 45\,824\,482 \\ \sum_1 V &= 33\,280 \times (-a_1) = -4\,953\,065 \\ \sum_1 VC^2 &= 435\,549\,784 \times (-b) = -40\,437\,179 \end{aligned} \right.$



Classe (2) $\left\{ \begin{aligned} \sum_2 V^2 &= 3\,407\,035 \\ \sum_2 V &= 6\,057 \times (-a) = -104\,032 \\ \sum_2 VC^2 &= 37\,530\,072 \times (-b) = 3\,484\,435 \end{aligned} \right.$



$$\hat{r}_1 + \hat{r}_2 = 443\,155 \xrightarrow{\text{diviser par } \frac{n - 1}{n}} \boxed{454\,817} \downarrow \sqrt{\quad}$$

$$\text{ET} \boxed{2\,507} \xleftarrow{\text{multiplier par } \frac{N}{n}} \boxed{674,4}$$

$$\text{ET } \% \boxed{1,68} \quad \frac{2\,507 \times 100}{149\,189} = 1,68 \%$$

ANNEXE 10 (*)

Nous présentons ici les calculs qui permettent d'obtenir, à partir de l'échantillon, les équations des deux droites qui constituent ce que l'on a appelé d'une façon abrégée une « bidroite en C ». Voir par exemple les graphiques n° 12 et 14.

Nous faisons appel d'une part à la méthode des moindres carrés et d'autre part à la méthode des multiplicateurs de LAGRANGE.

La qualité de produit étudiée est la qualité « râperie » sur St-Avoid 134 B pour $C < 60$ $V = \alpha' + \beta' C$ V en dm^3
 $C \geq 60$ $V = \alpha'' + \beta'' C$ C en cm^3
 avec la condition que les volumes soient identiques pour $C = 60$, c'est-à-dire :

$$\alpha' + 60\beta' = \alpha'' + 60\beta''$$

Nous devons donc par application de la méthode des moindres carrés, avoir

$$\begin{array}{l} \text{minimum} \\ \text{lorsque} \end{array} \quad \begin{array}{l} \Sigma (v' - \alpha' - \beta'c')^2 + \Sigma (v'' - \alpha'' - \beta''c'')^2 \\ \alpha' + 60\beta' = \alpha'' + 60\beta'' \end{array}$$

les sommations intéressent l'échantillon où :

v' est le volume râperie des tiges dont la circonférence c' est inférieure à 60 cm,

v'' est le volume râperie des tiges dont la circonférence c'' est égale ou supérieure à 60 cm.

La méthode des multiplicateurs de LAGRANGE amène à minimiser l'expression :

$$(1) \Sigma (v' - \alpha' - \beta'c')^2 + \Sigma (v'' - \alpha'' - \beta''c'')^2 - 2\lambda(\alpha' + 60\beta' - \alpha'' - 60\beta'')$$

Cette méthode est classique et le lecteur pourra en trouver la démonstration dans tout ouvrage d'analyse mathématique un peu poussé. Nous ne la reproduisons pas ici.

Pour minimiser l'expression (1), il suffit d'annuler toutes les demi-dérivées partielles, ce qui donne les équations normales suivantes :

(*) Les annexes 10, 11 et 12 ont été rédigées par P. ARBONNIER, suivant les indications de J. ABADIE.

$$(2) \begin{cases} n' \alpha' + \beta' \Sigma c' + \lambda = \Sigma v' \\ \alpha' \Sigma c' + \beta' \Sigma c'^2 + 60 \lambda = \Sigma v' c' \\ n'' \alpha'' + \beta'' \Sigma c'' - \lambda = \Sigma v'' \\ \alpha'' \Sigma c'' + \beta'' \Sigma c''^2 - 60 \lambda = \Sigma v'' c'' \end{cases}$$

avec

$$\alpha' + 60 \beta' = \alpha'' + 60 \beta''$$

soit cinq équations linéaires aux cinq inconnues α' , β' , α'' , β'' et λ .
 α' , β' , α'' , β'' sont elles-mêmes des fonction linéaires de λ .

Les quatre premières équations peuvent être résolues en considérant λ comme variable et le système (2) étant linéaire, les inconnues

$$\alpha', \beta', \alpha'', \beta''$$

sont elles-mêmes des fonctions linéaires de λ .

Posons alors :

$$(3) \begin{cases} \alpha' = \alpha'_1 + \lambda \alpha'_2 & \beta' = \beta'_1 + \lambda \beta'_2 \\ \alpha'' = \alpha''_1 + \lambda \alpha''_2 & \beta'' = \beta''_1 + \lambda \beta''_2 \end{cases}$$

les α'_1 , α'_2 , β'_1 , β'_2 , etc... sont indépendantes de λ .

Les équations du système (2) se résolvent alors simplement puisque chacune d'elles donne naissance à deux nouvelles équations où ne figure plus λ .

La première donne par exemple :

$$\begin{array}{ll} \text{et la seconde} & n' \alpha'_1 + \beta'_1 \Sigma c' = \Sigma v' & n' \alpha'_2 + \beta'_2 \Sigma c' = -1 \\ & \alpha'_1 \Sigma c' + \beta'_1 \Sigma c'^2 = \Sigma v' c' & \alpha'_2 \Sigma c' + \beta'_2 \Sigma c'^2 = -60 \end{array}$$

Entre ces quatre équations associées verticalement, il est aisé de calculer α'_1 , β'_1 , α'_2 et β'_2 .

Il en est de même pour les deux autres équations du système (2) qui donnent α''_1 , β''_1 , α''_2 et β''_2 .

La condition

$$\alpha' + 60 \beta' = \alpha'' + 60 \beta''$$

$$\text{devient} \quad \alpha'_1 + 60 \beta'_1 + \lambda (\alpha'_2 + 60 \beta'_2) = \alpha''_1 + 60 \beta''_1 + \lambda (\alpha''_2 + 60 \beta''_2)$$

et donne la valeur de λ qu'il suffira de reporter dans les équations (3) pour avoir en définitive les paramètres exacts des deux droites :

$$V = \alpha' + \beta' C \qquad V = \alpha'' + \beta'' C$$

Les calculs correspondants sont présentés ci-après.

Tableau 1. — Il comporte le calcul des différents coefficients des équations (2) ainsi que de $\sum v^2$ et $\sum v''^2$ qui sont utilisées pour le calcul de l'erreur-type. Les données sont celles de l'échantillon. Pour certaines colonnes on n'a reporté que les totaux : la machine à calculer imprime sur la bande les multiplications partielles et les additionne en même temps pour donner le résultat définitif. Les calculs sont alors contrôlés par lecture de la bande.

Tableaux 2. — Le tableau 2 A comporte le calcul des valeurs de $\alpha'_1, \beta'_1, \alpha'_2, \beta'_2$ ainsi qu'il a été expliqué ci-dessus.

Le tableau 2 B comporte le calcul analogue pour $\alpha''_1, \beta''_1, \alpha''_2$ et β''_2 .

Le tableau 2 C comporte le calcul de λ par report dans l'équation :

$$\alpha' + 60\beta' = \alpha'' + 60\beta''$$

ou encore :

$$\alpha'_1 + 60\beta'_1 - \alpha'_2 - 60\beta'_2 + \lambda (\alpha''_1 + 60\beta''_1 - \alpha''_2 - 60\beta''_2) = 0$$

et le calcul des paramètres définitifs.

Tableau 3. — C'est le tableau de la composition de la population à cuber.

Le volume est calculé suivant la formule :

$$V_t = N' \alpha' + \beta' \Sigma C' + N'' \alpha'' + \beta'' \Sigma C''$$

les sommations étant étendues à la population.

Tableau 4. — Calcul de l'erreur-type suivant la méthode plusieurs fois exposée, en partant de la valeur des moindres carrés pour l'échantillon qui est ici :

$$Q = \sum v^2 + \sum v''^2 - \alpha' \sum v' - \alpha'' \sum v'' - \beta' \sum v' c' - \beta'' \sum v'' c''$$

$$E_T = \frac{N}{n} \sqrt{Q \cdot \frac{n}{n-3} \cdot \frac{N+n_2-n_1}{N}}$$

n étant le nombre de tiges de l'échantillon dont n_1 ont été cubées sur pied et n_2 abattues.

On trouve une erreur-type de 5,49 % pour une erreur-vraie de — 3,83 %.

ANNEXE n° 10 — TABLEAU I

St-Avold 134 B

Qualité étudiée: Râperie

C < 60

C ≥ 60

C'	C'^2	V'	V'^2	V' C'
(1)	(2)	(3)	(4)	(5)
<i>Abattus:</i>				
32		53		
32		46		
36		60,5		
37		70		
38		67		
40		90,5		
42		87		
44		0		
48		139		
49		98,5		
53		152		
54		182		
58		0		
59		220,5		
/				
14				
<i>Debout:</i>				
41		110		
45		148,5		
47		136		
49		152		
50		180		
50		150,5		
53		190		
54		203		
55		209,5		
56		166,5		
57		191		
59		220,5		
/				
12				
1 238	60 664	3 323,5	532 428	167 480
(Σ C')	(Σ C'^2)	(Σ V')	(Σ V'^2)	(Σ V' C')

C''	C''^2	V''	V''^2	V'' C''
(1)	(2)	(3)	(4)	(5)
<i>Abattus:</i>				
62		120		
66		268,5		
75		133		
78		120,5		
/				
4				
<i>Debout:</i>				
60		228		
61		258,5		
63		285		
65		178,5		
70		181		
72		140,5		
88		67		
/				
7				
760	53 252	1 980,5	407 584	132 223
(Σ C'')	(Σ C''^2)	(Σ V'')	(Σ V''^2)	(Σ V'' C'')

ANNEXE n° 10 TABLEAU 2 A

St-Avoid 134 B

Qualité étudiée: Râperie.

 $c' < 60$

α'_1	β'_1	abc	
$(n') = 26$ $(\Sigma C') = 1238$	$(\Sigma C') = 1238$ $(\Sigma C'^2) = 60664$	$(\Sigma V') = 3323,5$ $(\Sigma V' C') = 167480,5$	
I	47,61538480	127,8260231	$\alpha'_1 + \underbrace{47,61538480 \times \beta'_1}_{256,1114702} = 127,8260231$ $\alpha'_1 = 128,2845471$
I	49,00161557	135,2831183	
	1,38623077	7,4561952	
			$\beta'_1 = 5,378754669$

 $c' < 60$

α'_2	β'_2	abc	
$(n') = 26$ $(\Sigma C') = 1238$	$(\Sigma C') = 1238$ $(\Sigma C'^2) = 60664$	— I — 60	
I	47,61538480	— 0,03846153860	$\alpha'_2 + \underbrace{47,61538480 \times \beta'_2}_{0,3436162100} = 0,03846153860$ $\alpha'_2 = 0,3051546714$
I	49,00161557	— 0,04846526662	
		— 0,01000372802	
			$\beta'_2 = 0,00721649554$

ANNEXE n° 10 — TABLEAU 2 B

St-Avoid 134 B

Qualité étudiée: Râperie.

$c'' \geq 60$

α''_1	β''_1	abc	
$(n'') = 11$ $(\Sigma C'') = 760$	$(\Sigma C'') = 760$ $(\Sigma C''^2) = 53\ 252$	$(\Sigma V'') = 1\ 980,5$ $(\Sigma V'' C'') = 132\ 223$	
I 1	69,09090900 70,06842100	180,0454545 173,9776310	$\alpha''_1 + 69,090909 \times \beta''_1 = 180,0454545$ — 428,8760040 $\alpha''_1 = 608,9214585$
	0,97751200	— 6,0678235	
		$\beta''_1 = - 6,20741588$	

$c'' \geq 60$

α''_2	β''_2	abc	
$(n'') = 11$ $(\Sigma C'') = 760$	$(\Sigma C'') = 760$ $(\Sigma C''^2) = 53\ 252$	+ 1 + 60	
I 1	69,09090900 70,06842100	0,0909090900 0,0789473684	$\alpha''_2 + 69,090909 \times \beta''_2 = 0,0909090900$ — 0,8454588930 $\alpha''_2 = 0,9363679830$
	0,97751200	— 0,0119617216	
		$\beta''_2 = - 0,0122369052$	

ANNEXE n° 10 — TABLEAU 2 C

St-Avold 134 B

*Qualité étudiée: Râperie.*Calcul de λ et des paramètres réels

$$\begin{aligned}
 \alpha'_1 &= -128,2845471 & \alpha'_2 &= 0,3051540714 & \alpha''_1 &= 608,9214585 & \alpha''_2 &= 0,9363679830 \\
 \beta'_1 &= 5,378754669 & \beta'_2 &= 0,00721649554 & \beta''_1 &= -6,20741588 & \beta''_2 &= -0,0122369052 \\
 x_1 &= \alpha''_1 & & + 60 \beta''_1 & - \alpha'_1 & & - 60 \beta'_1 \\
 &= + 608,9214585 & & - 6,20741588 \times 60 & + 128,2845471 & & - 5,378754669 \times 60 \\
 x_1 &= 42,9357732 \\
 x_2 &= \alpha''_2 & & + 60 \beta''_2 & - \alpha'_2 & & - 60 \beta'_2 \\
 &= 0,9363679830 & & - 0,0122369052 \times 60 & 0,3051540714 & & + 0,00721649554 \times 60 \\
 x_2 &= 0,3299887320 \\
 \lambda &= -\frac{x_1}{x_2} = -127,3854813
 \end{aligned}$$

Calcul de α' et α''
 β' et β''

$$\begin{aligned}
 \alpha' &= -128,2845471 & + & -127,3854813 & \times & 0,3051540714 \\
 & \alpha'_1 & & \lambda & \times & \alpha'_2 \\
 \beta' &= 5,378754669 & + & -127,3854813 & \times & 0,00721649554 \\
 & \beta'_1 & & \lambda & \times & \beta'_2 \\
 \alpha'' &= 608,9214585 & + & -127,3854813 & \times & 0,9363679830 \\
 & \alpha''_1 & & \lambda & \times & \alpha''_2 \\
 \beta'' &= -6,207415880 & + & -127,3854813 & \times & -0,0122369052 \\
 & \beta''_1 & & \lambda & \times & \beta''_2
 \end{aligned}$$

d'où



$\alpha' = -167,1568198$	$\alpha'' = 489,6417752$
$\beta' = 6,298031404$	$\beta'' = -4,648612010$

ANNEXE n° 10 — TABLEAU 3

St-Avoid 134 B

Qualité étudiée: Râperie.

C < 60			C ≥ 60		
Cir- conférence C'	Nombre d'arbres n'	C' × n'	Cir- conférence C''	Nombre d'arbres n''	C'' × n''
—	—	—	—	—	—
26	1		60	8	
28	1		61	10	
31	3		62	6	
32	1		63	7	
34	2		64	6	
35	1		65	3	
36	2		66	3	
37	4		67	4	
38	3		68	1	
39	6		69	2	
40	5		70	4	
41	2		71	4	
42	1		72	2	
43	6		73	3	
44	5		74	1	
45	10		75	1	
46	4		76	1	
47	12		77	2	
48	3		78	2	
49	9		80	1	
50	8		88	2	
51	3				
52	3			73	4872
53	6			(N'')	(Σ n'' c'')
54	7				
55	12				
56	15				
57	3				
58	5				
59	5				
	148	7096			
	(N')	(Σ n' c')			

Calcul du volume

$$\begin{aligned}
 V_{\text{taillis}} &= N' \alpha' + \beta S_n C' + N'' \alpha'' + \beta'' S_n C'' \\
 &= 33\,047,43\,393
 \end{aligned}$$

ANNEXE n° 10 — TABLEAU 4

St-Avold 134 B

Qualité étudiée: Râperie.

Calcul de ET %

$$n_2 = 18$$

$$n_1 = 19$$

$$n = 37$$

$$N = 221$$

$$Q = \frac{(+ 532,428) (+ 407,584) (- 167,1568198 \times 3,323,5) (+ 6,298031404 \times 167,480,5) (+ 480,6417752 \times 1,980,5) (- 4,64861201 \times 132,223)}{\sum V^2 + \sum V'^2 - \alpha' \sum V' - \beta' \sum V' C' - \alpha'' \sum V'' - \beta'' \sum V'' C''}$$

$$\frac{n-3}{n} = \frac{34}{37} = 0,918918$$

$$Q =$$

85678

$$\text{diviser par } \frac{n-3}{n}$$

93237

$$g = \frac{N}{N + n_2 - n_1} = \frac{221}{221 + 18 - 19} = \frac{221}{220} = 1,004545 \quad \text{diviser par } g$$

$$Q_1$$

92815

$$\sqrt{Q_1}$$

304

$$\text{multiplier par } \frac{N}{n}$$

$$\frac{N}{n} = \frac{221}{37} = 5,972972$$

$$ET$$

1815

$$ET \%$$

5,49

$$\frac{1815 \times 100}{33\,047,43393} = 5,49 \%$$

↓ Vol. tarif

Volume réel 31 828,00000

Volume tarif 33 047,43393

— 1 219,43393

E V
(erreur vraie)

— 3,83 %

$$\frac{- 1\,219,43393 \times 100}{31\,828} = - 3,83 \%$$

↓ Vol. réel

NOTA. — Le volume réel en râperie mesuré après coupe à blanc a été trouvé égal à 31 828 dm³.

ANNEXE 11

Cette annexe présente un exemple de calcul de régression linéaire du volume en fonction de la circonférence.

La qualité étudiée est la qualité « râperie » sur la place d'expérience totale de Porcelette.

Il n'y a plus ici de distinction entre les circonférences inférieures ou supérieures à 60.

L'expression du volume est la suivante :

$$V = \alpha + \beta C$$

V volume râperie en dm³,

C circonférence de la tige, en cm, à 1,30 m.

L'application de la méthode des moindres carrés conduit à minimiser l'expression

$$\sum (v - \alpha - \beta c)^2$$

la sommation étant étendue à toutes les tiges de l'échantillon dont la circonférence est c et le volume râperie v.

Le système des équations normales est alors extrêmement simple :

$$\begin{aligned} n\alpha + \beta \sum c &= \sum v \\ \alpha \sum c + \beta \sum c^2 &= \sum vc \end{aligned}$$

et permet la résolution en α et β .

Le *tableau 1* rassemble les données de l'échantillon : il montre le calcul des coefficients des équations normales ci-dessus ainsi que le calcul de la quantité $\sum v^2$ qui servira au calcul de l'erreur-type.

Le *tableau 2* comporte la résolution des équations normales et donne les valeurs de α et β .

Le *tableau 3* rassemble les données de la population et comporte le calcul des éléments nécessaires à l'application de la formule

$$V_t = N \alpha + \beta \sum C$$

Le *tableau 4* est relatif au calcul de l'erreur-type à partir de la valeur des moindres carrés Q donnée par la formule :

$$Q = \sum v^2 - \alpha \sum v - \beta \sum vc$$

et l'erreur-type est alors

$$E_T = \frac{N}{n} \sqrt{Q \cdot \frac{N + n_2 - n_1}{N} \cdot \frac{n}{n-2}}$$

n étant le nombre de tiges de l'échantillon dont n_1 ont été cubées sur pied et n_2 abattues.

Porcelette

Qualité étudiée: Râperie.

Tableau n° 1

C (1)	C ² (2)	V (3)	V ² (4)	VC (5)
68		141		
68		322,5		
75		220		
83		100		
87		168,5		
91		168		
96		85		
96		71,5		
102		0		
/				
9				
<i>Debout:</i>				
64		314		
68		158,5		
74		170,5		
75		190		
75		185,5		
76		132		
79		114,5		
81		181		
81		139		
84		127,5		
89		210,5		
89		116		
93		106		
93		109,5		
95		69		
96		61,5		
97		119		
97		115		
106		0		
106		0		
108		0		
109		308,5		
112		0		
114		0		
115		0		
121		0		
/				
26				
3 163	293 711	4 204	785 503	346 534
(ΣC)	(ΣC^2)	(ΣV)	(ΣV^2)	(ΣVC)

La quantité ΣC^2 n'a pas été calculée puisqu'elle l'a déjà été pour l'établissement du volume bois fort tige.

NOTA. — L'utilisation des machines à calculer électriques avec bandes imprimantes donne directement les résultats des colonnes (4) et (5) par sommation des opérations partielles.

ANNEXE n° 11

Porcelette

Qualité étudiée: Râperie.

Tableau n° 3

Cir- conférence (1)	Nombre d'arbres (2)	Cir- conférence (1)	Nombre d'arbres (2)	
52	1	87	1	
53	1	89	4	$\beta = 4,24443907$
62	1	90	2	$\alpha = 503,690306$
63	1	91	1	Volume tarif = $\beta \times S_n C + N \alpha$
64	1	92	2	$= 4,24443907 \times 9\,540 + 107 \times 503,690306$
66	1	93	5	$V_{\text{tarif}} = 13\,402,91$
67	1	94	2	
68	3	95	2	
69	3	96	5	
71	1	97	4	Volume réel = 12 901,50
72	2	98	2	Volume tarif = 13 402,91
73	1	99	3	$= 501,41$
74	1	102	2	$= 501,41 \times 100$
75	6	103	1	Erreur vraie = $\frac{501,41 \times 100}{12\,901,50} = 3,88\%$
76	1	104	1	et l'erreur- type est 8,03
77	2	105	1	
78	3	106	2	
79	4	107	1	
80	1	108	3	
81	5	109	1	
83	5	112	3	
84	1	113	2	
85	3	114	1	
86	1	115	1	
		119	2	
		121	1	
		123	1	
		142	1	
		N = 107		$S_n C = 9\,540$

ANNEXE n° 11

Porcellette

Qualité étudiée: Râperie.

Tableau 4

$$\beta = -4,24413907$$

$$n_1 = 26$$

$$n_2 = 9$$

$$n = 35$$

$$\alpha = 503,600306$$

$$N = 107$$

$$\sum V^2 = \quad \quad \quad + 785\,503$$

$$\sum V = 4\,204 \times (-\alpha) = -2\,117\,514$$

$$\sum V C = 346\,534 \times (-\beta) = +1\,470\,842$$

Q =

138 831

↓

diviser par $\frac{n-2}{n}$

$$\frac{n-2}{n} = \frac{33}{35} = 0,942857$$

147 245

↓

diviser par g

$$g = \frac{N}{N + n_2 - n_1} = \frac{107}{90} = 1,188888$$

R =

123 851

↓

\sqrt{R}

351,9

↓

multiplier par $\frac{N}{n}$

$$\frac{N}{n} = \frac{107}{35} = 3,057142$$

ET

1 076

↓

ET %

8,03

$$\frac{1\,076 \times 100}{13\,402,91} = 8,03$$

↓
Vol. tarif

et
l'erreur vraie
est
= 3,88

ANNEXE 12

Cette annexe est relative au mode de calcul le plus simple qui puisse être adopté pour l'estimation de certaines qualités de produits. C'est la méthode appelée dans le texte « méthode horizontale ».

Lorsque la grande dispersion des points figuratifs de l'échantillon — souvent liée au « phénomène de zéro » — ne permet pas de déceler une loi de variation du volume étudiée en fonction de la circonférence, il est loisible d'admettre qu'il n'y a pas de corrélation définie entre V et C et de représenter en conséquence le volume par une horizontale. Comme dans une régression linéaire, la droite passe par le centre de gravité, l'ordonnée de cette horizontale est le volume moyen de l'échantillon.

$$\bar{v} = \frac{\sum v}{n}$$

La qualité étudiée ici est la qualité « râperie » sur la place d'expériences totale d'Hagnenau.

Le *tableau 1* présente les calculs nécessaires à l'obtention de $\sum v$ et $\sum v^2$, cette quantité intervenant dans le calcul de l'erreur-type.

Le *tableau 2* donne l'estimation du volume « râperie » de la population par application de la formule.

$$v_t = n\bar{v} = \frac{N}{n} \sum v$$

Le *tableau 3* est relatif au calcul de l'erreur-type. On obtient d'abord la somme des moindres carrés Q par la formule bien connue :

$$Q = \sum v^2 - n \bar{v}^2$$

et l'erreur-type est cette fois :

$$E_T = \frac{N}{n} \sqrt{Q \cdot \frac{n}{n-1} \cdot \frac{N+n_2-n_1}{N}}$$

où n est le nombre de tiges de l'échantillon dont n_2 sont cubées sur pied et n_1 abattues.

C ≥ 20

Tableau 1

Tableau 2

Calcul de h

C	V	V ²
<i>Abattus</i>		
27	34,5	
31	12,5	
33	45,5	
40	88	
41	103	
42	64,5	
43,5	112	
44	75	
47	156,5	
48	152,5	
52	185,5	
55	26,5	
58	237,5	
60	0	
60	154,5	
61	221	
62	200	
62	139	
63	228,5	
65	211,5	
66	77	
68	0	
68	211,5	
72	53,5	
73	141,5	
73	134	
/		
26		
<i>Debout:</i>		
36	0	
37	0	
42	81,5	
50	199,5	
51	54	
54	77,5	
56	78,5	
57	83,5	
59	66,5	
63	194,5	
67	41	
67	0	
69	32	
70	70,5	
72	74,5	
73	0	
75	103,5	
77	0	
77	103	
78	0	
79	0	
79	133,5	
80	98	
81	61,5	
82	48,5	
82	67	
83	105	
84	0	
85	0	
87	135	
87	139	
88	102,5	
91	0	
91	102,5	
93	92,5	
95	78,5	
98	105,5	
102	93	
108	66,5	
/		
39		
	5 754	795 688
n = 65	Σ V	Σ V ²

$$\bar{V} = \frac{\Sigma V}{n} = \frac{5 754}{65} = 88,523076$$

$$h^2 = 7 836,334975$$

$$n = 65$$

$$N = 285$$

$$V_{\text{tarif}} = \frac{N}{285} \times \bar{V} = 25 229,07663$$

$$V_{\text{réel}} = 25 102,50$$

$$V_{\text{tarif}} = 25 229,07$$

$$V_{\text{réel}} - V_{\text{tarif}} = - 126,57$$

$$\text{Erreur vraie} = \frac{- 126,57 \times 100}{25 102,50} = - 0,504 \%$$

↓
Volume réel

ANNEXE n° 12

Haguenau

Qualité étudiée: Réperic.

Tableau 3

Calcul de l'erreur-type

$$N = 285$$

$$V = 88,523076 \quad h^2 = 7\,836,334975$$

$$n_1 = 39 \quad n_2 = 26 \quad n = 65$$

$$\sum V^2 = n \bar{V}^2 = 795,688 = 65 \times 7\,836,334975$$

$$\frac{n-1}{n} = \frac{64}{65} = 0,984615385$$

$$\frac{286\,326,2273}{n-1} = 286\,326,2273$$

↓

290 800,076

↓

$$g = \frac{N}{N+26-39} = \frac{285}{272} = 1,0470411$$

diviser par g

$$Q = 277\,496,5862$$

↓

$$\sqrt{Q} = 526,77$$

↓

$$\frac{N}{n} = \frac{285}{65} = 4,38461539$$

multiplier par $\frac{N}{n}$

$$ET = 2\,309,68$$

↓

$$\frac{2\,309,68 \times 100}{25\,229,07} = 9,15 \%$$

↓
Vol. tarif

$$ET \% = 9,15$$

SUMMARY

Method of volume calculation for stands within experimental plots of silviculture

Studies about forest production and more particularly the ones about treatment methods should be based on the most accurate data. They should be qualitative and quantitative. Therefore, in order to build the tariff tables necessary to volume estimation of stands, P. AYRAL, Leader of the first Section of the Research Station, worked on a method of « stratified sample » whose application has already given excellent results. The most important characters are:

- 1) Place of the trees in stories of vegetation (1, 2, 3);
- 2) Crown development (A, B, C);
- 3) Quality of the stem (D, T, F, etc...).

The stratified sample should include sample tree of each class of such characters in every girth class, proportionally with the corresponding number of stems in the stand in order to avoid a systematic distortion. Volume measurements of standing trees, necessary to that system, are realized with the « Ecole Forestière » duralumin ladder.

One must then prepare, with this representative sample, a tariff table, suitable to the stand after its thinning. The practical basis of studies was constituted of seven clear felled stands, most accurately measured, tree by tree, after numeration.

Instead of the graphical method, improper by many ways, J. ABADIE has worked on a mathematical method, less subjective and allowing calculation of the standard-error of volume estimation.

Several mathematical laws (as $V = f(c)$) have been tested ($V =$ volume; $c =$ girth at b. h.).

The following formula has been adopted:

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

whose normal equations are easy to solve and which keeps the standard-error, between 1 and 2 %.

The constants a and b are calculated according to the method of least squares.

The constants once known, the volume of a stand of N stems is given with the formula :

$$V = Na + b \sum n_i c_i^2$$

It is practical and even cheaper to have these calculations made by an electronic machine. It will be noted that, if a scale of squares is used for abscissae, the law should be graphically represented by a straight line.

Some conclusions, though less general, but very interesting, are to remember :

1) in each considered stand, within the same class of characters, the volume distribution for a given girth is the usual distribution of LAPLACE-GAUSS.

2) for a given girth, there is a very significant difference between the volumes of trees of class 1 and trees of class 2 for instance.

Every thing equal besides, there are no differences between the volumes of a tree « A » and a tree « B ».

At last, the qualitative study of stands, considered as necessary for specifying the influence due to methods of treatment, has been undertaken, the methods of investigation, made clear.

A graphic study (diagram) will be started, which will permit to know what simple curve should be calculated, in order to take into account the distribution of the different qualities of products according to the b. h. girth.

ZUSAMMENFASSUNG

Rechnungsmethode der Massen der stehenden Bestände in den Waldbauversuchsflächen

Die Studien über den forstlichen Ertrag, ins besondere diejenigen, die der Einfluss der Behandlungsmethoden betreffen, müssen auf so genauen Angaben gegründet sein wie möglich. Sie müssen qualitativ sowie quantitativ sein. Deshalb hat P. AYRAL, der Leiter der 1. Abteilung der Versuchsanstalt, zum Aufstellen der Massentafeln, die zur Abschätzung der bleibenden Bestände dienen, eine Methode ausgearbeitet, deren Anwendung vortreffliche Ergebnisse hervorgebracht hat. Folgende Charaktere sind behalten worden:

- 1° Platz der Stämme in den Höhenstufen (1 - 2 - 3);
- 2° Form der Krone (A - B - C);
- 3° Güte des Stammes (D - T - F, u.s.w....).

Die beschränkt zufällige Probenahme muss Vertreter aus allen Baumklassen in allem Umfangsgruppen, im Verhältnis zu den entsprechenden Stammzahlen in dem stehen bleibenden Bestand begriffen, um eine systematische Verzerrung zu vermeiden. Die Kubierung der stehenden Bäume, die diese Methode fordert wird dank dem Leiter « Ecole Forestière » ausgeführt.

Nun muss noch, mit der erhaltenen repräsentativen Probenahme, eine Massentafel aufgestellt werden, die für den Bestand nach der Durchforstung gültig sei. Die praktische Grundlage der Forschungen ist von sieben kahl abgeholzten und, nach Numerierung, Baum nach Baum, sorgfältig gemessenen Beständen gebildet.

Der graphischen Methode, deren Nachteile wohlbekannt sind, hat, dank der Forschungen von J. ABADIE, eine minder subjektive mathematische Methode unterschoben werden können und die, ausserdem, das Berechnen einer Standardabweichung ermöglicht.

Mehrere Gesetze der Form $V = f(c)$ sind probiert worden ($V = \text{Masse}$; $C = \text{Umfang}$).

So ist man dazu geführt worden die Formel $V = a + bc^2$ anzunehmen, deren normalen Gleichungen leicht zu lösen sind und die die Standardabweichung zwischen 1 und 2 % festlegt.

Die Koeffizienten a und b werden mit der Methode der kleinsten

Quadrate gerechnet. Wenn jene einmal bekannt sind wird die Masse eines Bestandes von N Stämmen durch die Formel:

$$V = N a + b \sum n_i c_i^2$$

gegeben. Praktisch und dazu auch sparsam ist die Verwendung einer elektronischen Maschine.

Es kann bemerkt werden dass sich, wenn man also Abszisse eine Quadratskala annimmt, das in Frage stehende Gesetz auf dem Millimeterpapier als eine gerade Linie äussert.

Einige minder allgemeine, aber höchst interessante Schlussfolgerungen müssen auch im Sinne behalten werden:

1) in jedem studierten Bestand, innerhalb einer selben Klasse ist die Verteilung der Masse für jeden Umfang, die normale Verteilung von LAPLACE-GAUSS;

2) bei gleichem Umfang, gibt es einen bezeichnenden Unterschied zwischen den Massen, zum Beispiel der Klasse 1 und diejenigen der Klasse 2. Wenn alles gleich ist, gibt es keinen Unterschied zwischen den Massen eines Baumes « A » und eines Baumes « B ».

Schliesslich hat die qualitative Forschung der forstlichen Bestände, die als absolut unerlässlich betrachtet wird, um den Einfluss der Behandlungsmethoden zu bestimmen, in Schwung gesetzt werden können. Die zu benützendenden Forschungsmethoden sind bestimmt worden. Zuerst muss eine graphische Forschung ausgeführt werden, die dann ermöglichen wird, zu sehen welche einfache Kurve man gerechnet werden muss, um am besten über die Verteilung der verschieden Qualitäten je nach den Umfängen, Bericht zu erstatten.

(Traduction M. DÉCOURT.)

