



GOUVERNEMENT
DU QUÉBEC

MINISTÈRE DES TERRES
ET FORÊTS

SERVICE DE LA RECHERCHE

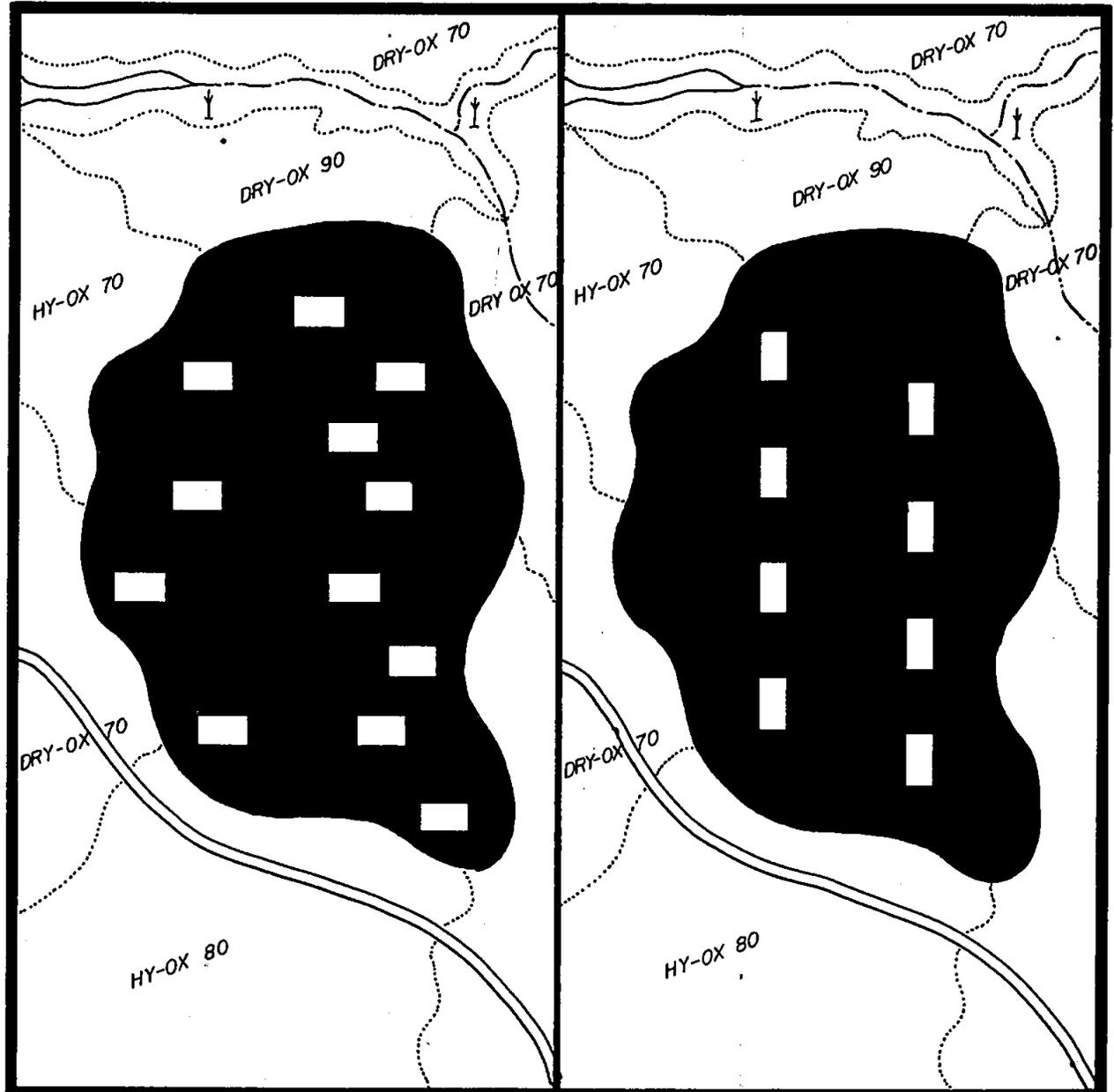
MC

O.D.C. 524.637 - 015.5(043) L.C. SD 555 V

MÉMOIRE N° 5
1971

CALCUL DE L'ERREUR D'ÉCHANTILLONNAGE

par Pierre Villeneuve



PIERRE VILLENEUVE est bachelier ès arts et bachelier ès sciences appliquées (génie forestier) de l'université Laval depuis 1962. En 1965, cette université lui décernait le diplôme de maître ès sciences forestières. Sa carrière débuta au Service de l'inventaire forestier du ministère des Terres et Forêts du Québec. En 1967, il passa au Service de la recherche à titre de chargé de recherches en dendrométrie ainsi que, par la suite, de chargé des secteurs expérimentaux. Depuis 1970, il est secrétaire du Conseil de la recherche et du développement forestiers du Québec, organisme consultatif dépendant directement du ministre des Terres et Forêts.

CALCUL DE L'ERREUR D'ECHANTILLONNAGE*

PAR

PIERRE VILLENEUVE

MEMOIRE N° 5

SERVICE DE LA RECHERCHE
DIRECTION GENERALE DE LA PLANIFICATION
MINISTERE DES TERRES ET FORETS

1971

Deuxième tirage 1975

* Résumé d'une partie de la thèse de maîtrise effectuée sous la direction de M. Gilbert Tardif et présentée par Pierre Villeneuve en 1965 à l'Ecole des gradués de l'université Laval pour l'obtention du grade de maître ès sciences (dendrométrie)

Dépôt légal

Bibliothèque nationale du Québec

RÉSUMÉ

CALCUL DE L'ERREUR D'ÉCHANTILLONNAGE

Au cours de la présente étude, nous désirons vérifier si un échantillonnage systématique peut fournir une estimation valable de la variance de la population étudiée. L'étude de quelques méthodes de sondage aléatoire a été aussi effectuée afin de vérifier la précision du calcul de la variance.

SUMMARY

DETERMINATION OF SAMPLING ERROR

The objective of this study is to determine if variances, obtained from a systematic sampling are estimates of the true variance of the population. Different random sampling methods were also applied to compare their variance with the true variance.

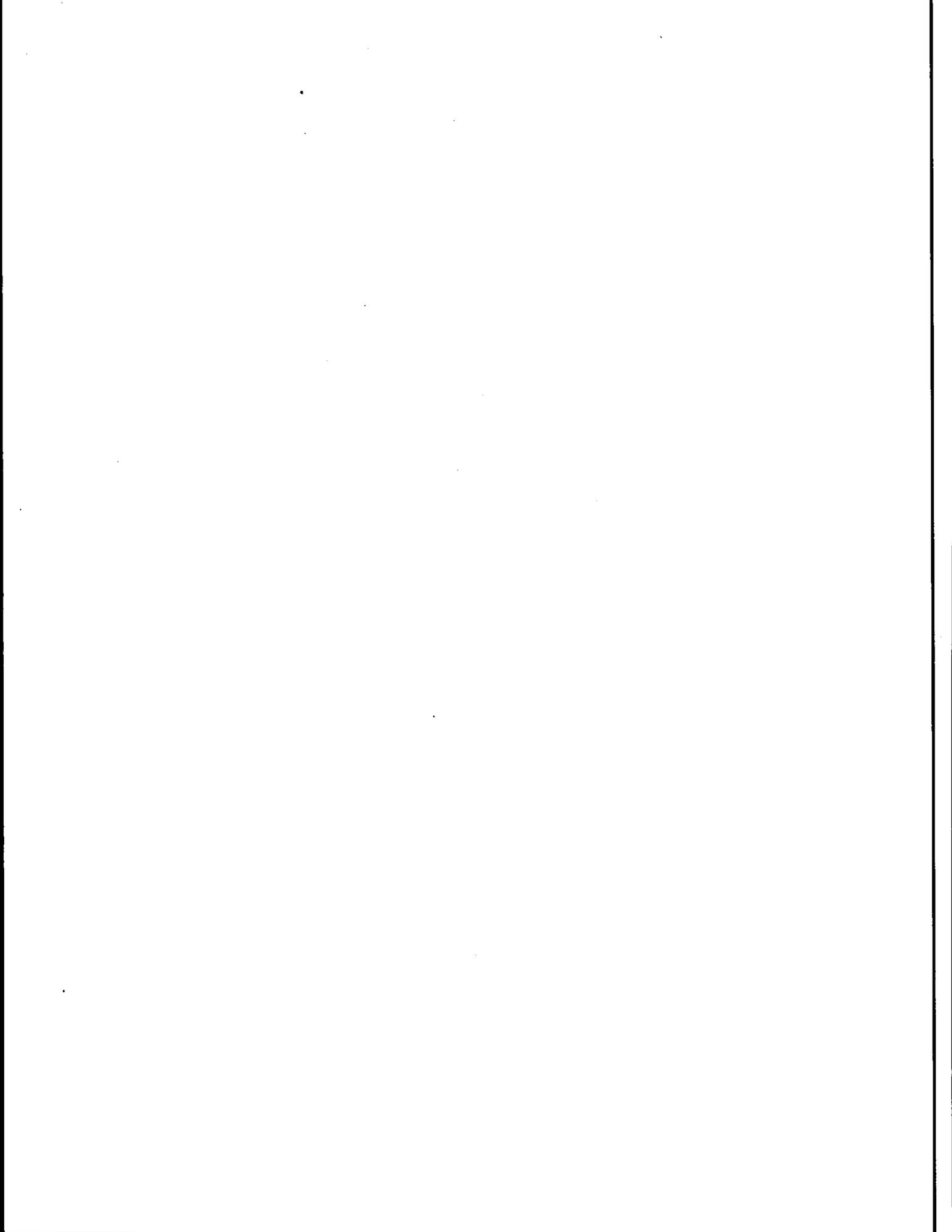


TABLE DES MATIÈRES

	page
RÉSUMÉ.....	iii
TABLE DES MATIÈRES.....	v
LISTE DES FIGURES.....	vii
LISTE DES TABLEAUX.....	vii
INTRODUCTION.....	1
CHAPITRE I TRAVAUX PRÉLIMINAIRES.....	3
CHAPITRE II REVUE DE LITTÉRATURE.....	7
CHAPITRE III DESCRIPTION DES EXPÉRIENCES EFFECTUÉES.....	11
3.1 - Sondages systématiques selon diverses fractions de sondage.....	12
3.2 - Sondages aléatoires.....	14
3.2.1 - Sondage en grappes.....	14
3.2.2 - Sondage par virées.....	15
CHAPITRE IV PRÉSENTATION ET ANALYSE DES DONNÉES EXPÉRIMENTALES.....	17
4.1 - Sondages systématiques selon diverses fractions de sondage.....	17
4.1.1 - Estimation du volume moyen par parcelle.....	17
4.1.2 - Estimation de la variance.....	23
4.1.3 - Conclusion.....	29

4.2 - Sondages aléatoires.....	29
4.2.1 - Sondage en grappes.....	29
4.2.1.1 - Définition des termes.....	30
4.2.1.2 - Somme des carrés totaux.....	31
4.2.1.3 - Somme des carrés inter-grappes.....	31
4.2.1.4 - Somme des carrés intra-grappes.....	31
4.2.1.5 - Efficacité relative.....	32
4.2.1.6 - Coefficient de corrélation.....	33
4.2.1.7 - Résultats obtenus.....	34
4.2.1.8 - Discussion des résultats.....	36
4.2.1.9 - Fonction économique.....	39
4.2.2 - Sondage par virées.....	41
CHAPITRE V CONCLUSIONS.....	47
BIBLIOGRAPHIE.....	49
APPENDICE.....	51

LISTE DES FIGURES

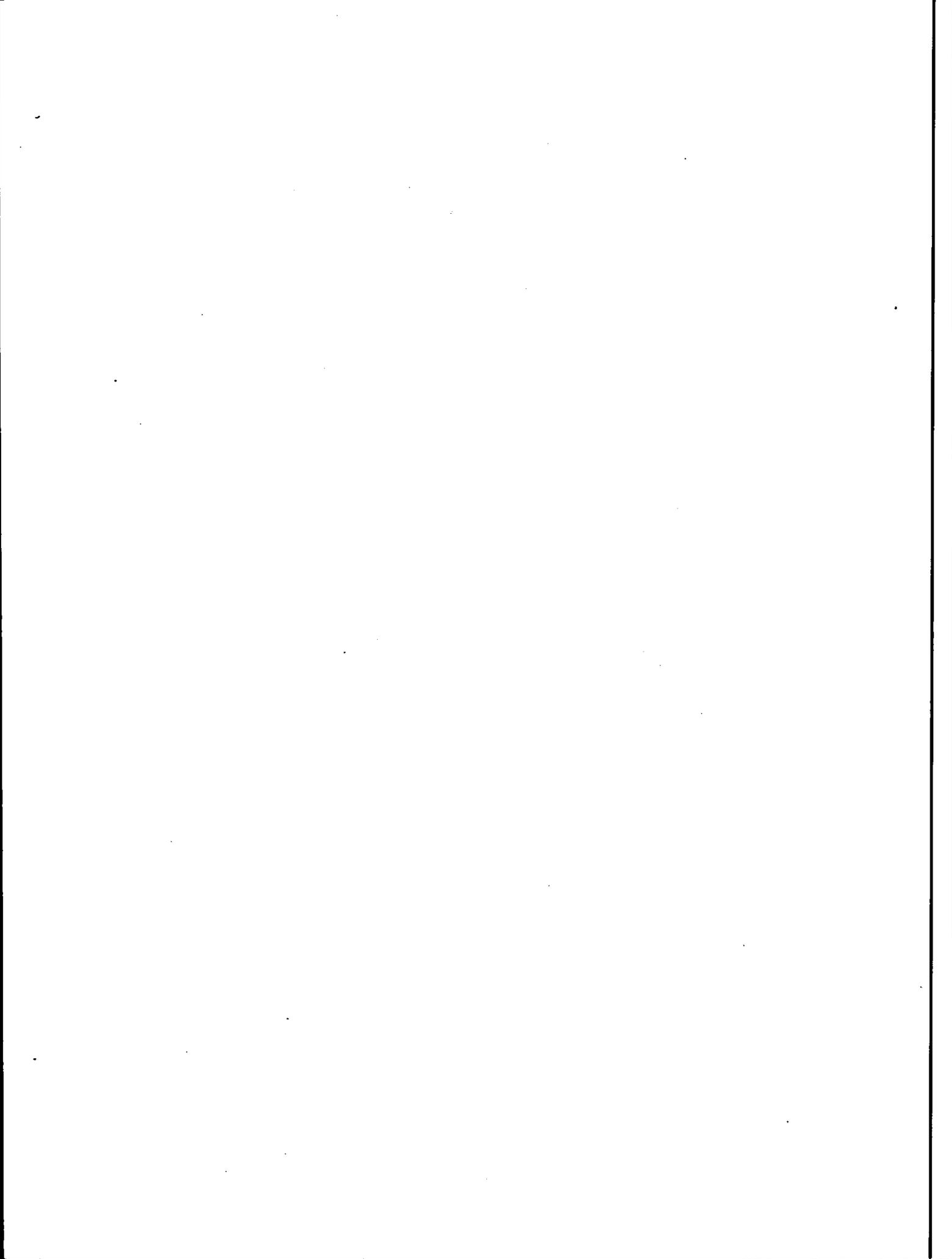
		page
FIGURE 1	Intervalles de confiance du volume moyen correspondant aux diverses fractions de sondage étudiées.....	21
FIGURE 2	Intervalles de confiance pour l'estimation de la variance totale.....	25
FIGURE 3	Intervalles de confiance pour l'estimation de la variance du <i>DRY-OX</i> 90 ans.....	27
FIGURE 4	Intervalles de confiance pour l'estimation de la variance du <i>DRY-OX</i> 70 ans.....	28
FIGURE 5	Intervalles de confiance pour l'estimation de la variance d'après des sondages par virée.....	45

LISTE DES TABLEAUX

TABLEAU I	Liste des peuplements et de leur superficie respective...	5
TABLEAU II	Calcul du volume, de la variance et du coefficient de variation par strate et pour l'ensemble de la population.	6
TABLEAU III	La proportion de chaque strate et le nombre correspondant de parcelles qui leur sont assignées.....	12

TABLEAU IV	Résultats relatifs à l'estimation du volume total et de la variance à partir des différents essais de sondage systématique avec choix aléatoire de la première parcelle.....	18
TABLEAU V	Intervalles de confiance de la moyenne pour chaque fraction de sondage.....	20
TABLEAU VI	L'erreur d'échantillonnage (E.Ec) calculée d'après la différence de chaque sondage avec le volume total de 185 877 pi ³ ainsi que le volume moyen par parcelle.....	22
TABLEAU VII	Résultats des estimations de la variance totale pour chaque fraction de sondage.....	24
TABLEAU VIII	Résultats des estimations de la variance pour chaque fraction de sondage dans le type <i>Dryopteris-Oxalis</i> 90 ans.....	24
TABLEAU IX	Résultats des estimations de la variance pour chaque fraction de sondage dans le type <i>Dryopteris-Oxalis</i> 70 ans.....	26
TABLEAU X	Variances intra-grappes et inter-grappes et la variance totale pour le sondage en grappes.....	32
TABLEAU XI	Sondage avec grappes de 3 parcelles.....	35
TABLEAU XII	Sondage avec grappes de 4 parcelles.....	35
TABLEAU XIII	Sondage avec grappes de 5 parcelles.....	36
TABLEAU XIV	Comparaison entre la variance totale des sondages en grappes de tailles différentes et la variance du <i>Dryopteris-Oxalis</i> 70 ans.....	36
TABLEAU XV	Efficacité relative des sondages en grappes.....	37

TABLEAU XVI	Changement relatif de la variance en fonction de l'accroissement de la taille des grappes.....	39
TABLEAU XVII	Distribution des parcelles lors de sondages par virées selon différentes fractions de sondages.....	41
TABLEAU XVIII	Analyse de variance pour le sondage à 5 pour 100 avec 6 parcelles par virée	42
TABLEAU XIX	Analyse de variance pour le sondage à 10 pour 100 avec 3 parcelles par virée	43
TABLEAU XX	Analyse de variance pour le sondage à 10 pour 100 avec 6 parcelles par virée	43
TABLEAU XXI	Analyse de variance pour le sondage à 20 pour 100 avec 9 parcelles par virée.....	43
TABLEAU XXII	Résultats des tests de "F" de Snedecor pour chaque fraction de sondage.....	46
CARTE montrant	la localisation de l'aire étudiée.....	4



INTRODUCTION

Le volume total ou le volume moyen par unité de surface pour l'ensemble d'un territoire boisé n'a vraiment de sens et de valeur que si l'on peut indiquer, à un seuil de probabilité donné, quelle est la marge d'erreur liée à de telles estimations. Or il arrive généralement, et cela pour fin de commodité et d'économie, que les relevés d'inventaires forestiers sont effectués à partir de parcelles-échantillons distribuées systématiquement c'est-à-dire d'une façon telle que, si "n" est le nombre total d'individus de l'échantillon, la première unité est choisie aléatoirement et le choix des (n-1) autres devient automatique. Un tel sondage ne se prête pas au calcul de l'estimation de l'erreur d'échantillonnage parce que la validité du calcul de cette erreur repose sur le caractère aléatoire des individus de l'échantillon.

Au cours de la présente étude, nous posons l'hypothèse que, pour chaque parcelle-échantillon, nous avons une juste estimation du volume et que les compilations et prises de données sont exemptes d'erreurs expérimentales. Nous ne tiendrons donc compte, ici, que de l'erreur d'échantillonnage. En effet, nous désirons déterminer si les estimations de la variance effectuées à partir d'un échantillonnage systématique sont égales à la vraie variance.

CHAPITRE I

TRAVAUX PRÉLIMINAIRES

Afin de vérifier les diverses hypothèses relatives au calcul de l'erreur d'échantillonnage, une expérience a été effectuée à l'intérieur de la forêt domaniale de Matane, sise à quelques 250 milles (400 km) au nord-est de Québec.

Une superficie totale de 50 acres (20,25 ha) de forêt a été complètement subdivisée en parcelles-échantillons de 0.05 acre ($\frac{1}{2}$ x 1 ch), soit au total 1 009 parcelles distribuées dans trois blocs de 3, 10 et 37 acres respectivement (1,2, 4,0 et 15 ha). De ce nombre 956 furent retenues. Les tiges de chaque parcelle furent dénombrées, coupées et mesurées par des mesureurs du Gouvernement.

Une étude phytosociologique de la station concernée nous a permis de la définir comme un facies du type *Dryopteris-Oxalis* tel que décrit par le Dr A. Lafond (1960). Cependant, bien que l'ensemble reflète l'aspect de ce type, l'on retrouve plusieurs flots qui appartiennent aux types *Hylocomium-Oxalis* et *Hylocomium-Cornus*. Le tableau qui suit montre la distribution de ces différents peuplements ainsi que leur superficie et leur âge respectif.

TABEAU I

Liste des peuplements et de leur superficie respective					
Groupement	Classe d'âge	Nombre de parcelles		Superficie	
		(1/20 ac)	(0,02 ha)	(acres)	(ha)
<i>Dryopteris-Oxalis</i>	90 al*	25		1.25	0,51
<i>Dryopteris-Oxalis</i>	90	112		5.60	2,27
<i>Dryopteris-Oxalis</i>	70	659		32.95	13,34
<i>Dryopteris-Oxalis</i>	50	5		0.25	0,10
<i>Dryopteris-Oxalis</i>	30-70	56		2.80	1,13
<i>Dryopteris-Oxalis</i>	30	11		0.55	0,22
<i>Hylacomium-Oxalis</i>	70	67		3.35	1,36
<i>Hylacomium-Cornus</i>	70	21		1.05	0,42
TOTAL:		956		47.80	19,35

* Indique la présence d'aulnes

En premier lieu, le volume de chaque parcelle fut compilé par types écologiques et par classes d'âge. Le volume total marchand des 956 parcelles se chiffre à $185\ 877\ \text{pi}^3$ ($5\ 260,3\ \text{m}^3$). Le volume moyen par parcelle est de $195\ \text{pi}^3$ ($5,51\ \text{m}^3$) et le volume à l'acre se chiffre à $3\ 890\ \text{pi}^3$ ($110\ \text{m}^3$).

En second lieu, les calculs de la variance du volume par strate " σ_z^2 ", et pour l'ensemble de la population, " σ^2 ", ainsi que les coefficients de variation, " C_v ", ont été effectués: tous les résultats de ces calculs sont présentés au tableau II qui suit.

TABLEAU II

Calcul du volume, de la variance et du coefficient de variation par strate et pour l'ensemble de la population.				
Strate	Volume	Nombre de parcelles	Variance	C _v %
DRY-OX 90 a1	pi ³ : 3 689 m ³ : 104,4	25	3 922.2 3,1	42
DRY-OX 90	pi ³ : 20 302 m ³ : 574,5	112	3 698.4 3,0	33
DRY-OX 70	pi ³ : 135 536 m ³ : 3 835,7	659	3 522.0 2,8	28
DRY-OX 50	pi ³ : 286 m ³³ : 8,1	5	567.0 0,48	42
DRY-OX 30-70	pi ³ : 5 041 m ³ : 142,7	56	1 095.8 0,9	37
DRY-OX 30	pi ³ : 469 m ³ : 13,3	11	1 312.5. 1,05	83
HY-OX 70	pi ³ : 15 894 m ³ : 449,8	67	2 739.4 2,2	22
HY-CO 70	pi ³ : 4 660 m ³ : 131,9	21	1 405.6 1,1	16
TOTAL:	pi ³ : 185 877 m ³ : 5 260,3	956	4 577.4 3,7	35

CHAPITRE II

REVUE DE LITTÉRATURE

Sous cette rubrique, nous avons effectué une revue des principaux travaux relatifs au calcul de l'erreur d'échantillonnage à partir de différents plans de sondage. Chez les auteurs européens, nous avons étudié les travaux de Naslund (1930, 1939), Langsaeter (1932), Matern (1949) et Pardé (1957). Ailleurs, les principaux auteurs consultés furent Schumacher and Bull. (1932), Hasel (1937a, 1937b, 1938, 1942a, 1952b), Finney (1948), Sukhatme (1954), Gautshi (1956), Yates (1946, 1948, 1949) et Cochran (1946).

L'étude de ces différents travaux nous a conduit aux conclusions suivantes:

- 1° Le sondage systématique, dans l'ensemble, donne des estimations plus précises que le sondage aléatoire;
- 2° Quand la population est distribuée conformément au schéma d'urne, toutes les méthodes de sondage donnent une estimation valable de la variance et de l'erreur d'échantillonnage, de sorte que si:

EV_s = Valeur attendue de l'erreur d'échantillonnage d'un sondage systématique.

EV_{st} = Valeur attendue de l'erreur d'échantillonnage d'un sondage stratifié.

EV_h = Valeur attendue de l'erreur d'échantillonnage d'un sondage aléatoire.

on a :

$$EV_s = EV_{st} = EV_h$$

3° Quand la variation dans la population a une tendance linéaire, nous avons :

$$EV_{st} < EV_s < EV_h$$

4° Dans les forêts composées de quelques espèces seulement à l'intérieur de peuplements bien réguliers où le gradient de fertilité est changeant et difficile à déterminer, il semble que le sondage aléatoire donnera de meilleurs résultats que le sondage systématique;

5° Si la population montre une variation périodique, l'efficacité relative du sondage systématique dépend alors du choix de l'intervalle entre les parcelles ou les unités d'échantillonnage. Si cet intervalle coïncide avec la période du cycle, alors tout l'échantillon sera composé d'individus possédant les mêmes caractéristiques et, par conséquent, il ne sera pas représentatif de la population étudiée. Dans ce cas, le sondage systématique ne peut pas être considéré comme supérieur au sondage aléatoire en ce qui concerne le calcul de l'erreur d'échantillonnage.

Bien que l'échantillonnage systématique fournisse parmi toutes les méthodes de sondage la meilleure estimation de volume et que la plupart des auteurs recommandent son emploi, il n'en demeure pas moins que sur le plan formel cette méthode ne permet pas une estimation valable de la variance et le calcul correct de l'erreur d'échantillonnage.

Néanmoins, cette erreur d'échantillonnage peut être calculée soit par des échantillonnages supplémentaires (environ 5 pour 100 de sondage additionnel), soit par la variation entre des sondages systématiques répétés, soit par un sondage systématique avec plusieurs départs aléatoires de la première parcelle (Cherng-Jiann Shiue, 1960).

1. The first part of the document is a list of names and addresses.

2. The second part of the document is a list of names and addresses.

3. The third part of the document is a list of names and addresses.

4. The fourth part of the document is a list of names and addresses.

5. The fifth part of the document is a list of names and addresses.

6. The sixth part of the document is a list of names and addresses.

7. The seventh part of the document is a list of names and addresses.

8. The eighth part of the document is a list of names and addresses.

9. The ninth part of the document is a list of names and addresses.

10. The tenth part of the document is a list of names and addresses.

11. The eleventh part of the document is a list of names and addresses.

12. The twelfth part of the document is a list of names and addresses.

13. The thirteenth part of the document is a list of names and addresses.

14. The fourteenth part of the document is a list of names and addresses.

15. The fifteenth part of the document is a list of names and addresses.

16. The sixteenth part of the document is a list of names and addresses.

17. The seventeenth part of the document is a list of names and addresses.

18. The eighteenth part of the document is a list of names and addresses.

19. The nineteenth part of the document is a list of names and addresses.

20. The twentieth part of the document is a list of names and addresses.

21. The twenty-first part of the document is a list of names and addresses.

22. The twenty-second part of the document is a list of names and addresses.

23. The twenty-third part of the document is a list of names and addresses.

24. The twenty-fourth part of the document is a list of names and addresses.

25. The twenty-fifth part of the document is a list of names and addresses.

26. The twenty-sixth part of the document is a list of names and addresses.

27. The twenty-seventh part of the document is a list of names and addresses.

28. The twenty-eighth part of the document is a list of names and addresses.

29. The twenty-ninth part of the document is a list of names and addresses.

30. The thirtieth part of the document is a list of names and addresses.

CHAPITRE III

DESCRIPTION DES EXPÉRIENCES EFFECTUÉES

L'estimation de la variance d'une variable doit être effectuée à partir d'un échantillonnage aléatoire, c'est-à-dire à partir d'un sondage dont chacun des éléments a été choisi d'après un schéma d'urne. Une telle condition n'est toutefois pas essentielle si, au cours d'un relevé en forêt, le volume des parcelles est lui-même distribué d'une manière aléatoire sur le territoire. C'est là précisément l'argument invoqué par ceux qui affirment que l'estimation de la variance calculée à partir d'un échantillonnage systématique est valable.

Dans cette optique, nous essaierons d'étudier deux aspects particuliers de ce problème, à savoir:

- 1° Jusqu'à quel point les estimations de la variance calculées à partir d'un sondage systématique sont valables;
- 2° De quelle manière un échantillonnage aléatoire peut être effectué dans une forêt stratifiée.

Dans le but d'élucider ces deux questions, les expériences suivantes ont été effectuées:

3.1 SONDAGES SYSTÉMATIQUES SELON DIVERSES FRACTIONS DE SONDAGE.

Des sondages systématiques de 5, 10, 20 et 40 pour 100 ont été effectués sur tout le territoire. Pour ceux de 5 et 10 pour 100, il y eut cinq répétitions tandis que pour ceux de 20 et 40 pour 100, il n'y eut que trois et deux répétitions respectivement. Le tableau III ci-dessous résume la distribution des parcelles au cours de ces différents sondages.

TABLEAU III

Tableau montrant la proportion (W_i) de chaque strate et le nombre correspondant de parcelles (N_i) qui leur sont assignées.					
Strates	$W_i = N_i/N$	Intensité de l'échantillonnage en pourcentage			
		5	10	20	40
<i>DRY-OX</i> 90	0.1171	6	11	22	45
<i>DRY-OX</i> 90 <i>al</i>	0.0262	1	3	5	10
<i>DRY-OX</i> 70	0.6893	33	66	131	264
<i>DRY-OX</i> 50	0.0052	0	0	1	1
<i>DRY-OX</i> 30-70	0.0586	3	6	11	23
<i>DRY-OX</i> 30	0.0115	1	2	2	4
<i>HY-OX</i> 70	0.0701	3	6	13	27
<i>HY-CO</i> 70	0.0220	1	2	4	8
TOTAL:	1.0000	48	96	189	382

L'échantillon fut réparti proportionnellement au nombre total de parcelles de chaque strate et l'estimation de la variance a été effectuée pour chacune des répétitions afin de voir comment elle se rapproche et se compare à la variance de la population. A cette fin, nous avons calculé les intervalles de sécurité de ces estimations et, si elles ne sont pas biaisées, on peut espérer qu'en moyenne ces estimations sont contenues dans les limites définies par l'intervalle. Ces intervalles de sécurité sont données par l'expression:

$$\text{Probabilité} = \left(\frac{\sigma_z^2 \chi_{P_1}^2}{f} < S_i^2 < \frac{\sigma_z^2 \chi_{P_2}^2}{f} \right) = P_2 - P_1$$

dans laquelle:

χ^2 = Chi-carré avec "f" degrés de liberté aux seuils de probabilité P_1 (0.025) et P_2 (0.975)

σ_z^2 = Variance de la i^e strate

S_i^2 = Estimation de la variance de la i^e strate

Pour le calcul de $\chi_{P_1}^2$ et $\chi_{P_2}^2$, lorsque les tables de χ^2 ne présentaient pas de valeurs correspondantes à un nombre suffisamment élevé de degrés de liberté, la formule approximative suivante a été employée:

$$\chi_{P_1}^2 \approx \frac{1}{2} (\sqrt{2f-1}-U_p)^2$$

$$\chi_{P_2}^2 \approx \frac{1}{2} (\sqrt{2f-1}+U_p)^2$$

où "Up" est une variable normale réduite et égale à 1.96 au seuil de probabilité de 95 pour 100.

3.2 SONDAGES ALÉATOIRES.

Ces essais ont pour objet de vérifier si certains plans de sondage, qui satisfont aux exigences de la méthodologie statistique, peuvent s'avérer efficaces et pratiques dans l'exécution des inventaires forestiers. Les deux plans suivants ont été expérimentés:

3.2.1 SONDAGE EN GRAPPES.

Le but recherché ici est d'obtenir les dimensions de la grappe qui permettent à la fois de calculer une bonne estimation de la variance et de trouver un mode d'exécution économique des travaux sur le terrain. A cette fin, nous avons formé des grappes de 3, 4 et 5 parcelles de 0.05 acre (0,02 ha) dans la strate *Dryopteris-Oxalis* de 70 ans. Puis, pour chacun des sondages, nous avons fait l'analyse de variance, le calcul du coefficient de corrélation " ρ " et le calcul de l'efficacité relative "E" du sondage en grappes par rapport à un sondage aléatoire. Voici comment se présente la distribution de cette partie de l'expérience:

<u>Dimension de la grappe</u>	<u>Nombre de grappes</u>
3 parcelles	130
4 "	82
5 "	51

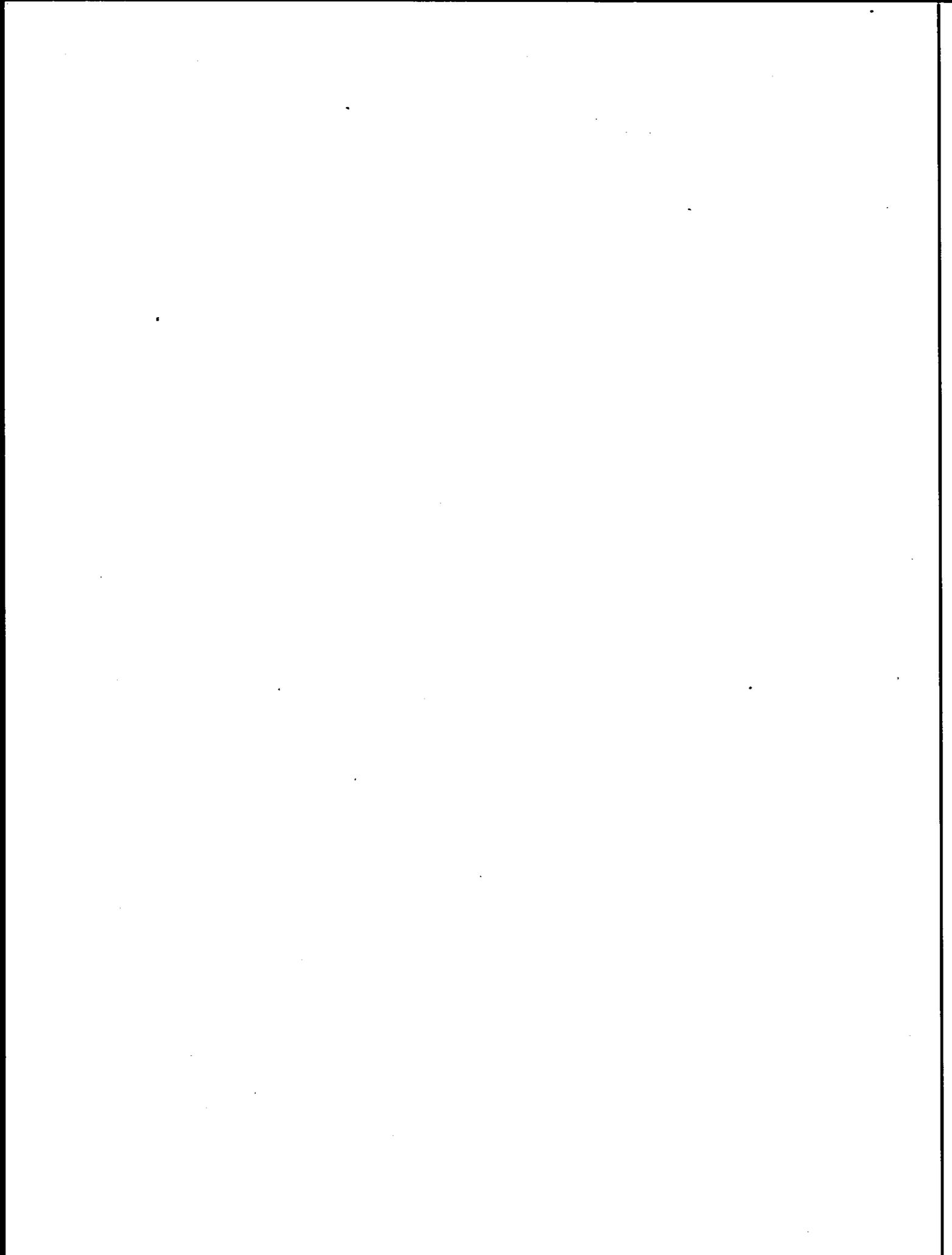
En dernier lieu, nous avons élaboré le sondage en grappes, composé de 3 parcelles par grappe, en effectuant plusieurs sondages à des intensités différentes. C'est ainsi que des sondages de 20, 30 et 50 pour 100, avec deux répétitions dans chaque cas, ont été faits. Puis le calcul de " ρ " et de "E" ainsi que l'analyse de variance furent effectués pour chacun des cas.

3.2.2 SONDAGE PAR VIRÉES.

Cette expérience consista à diviser la superficie en bandes de 33 pieds (10 m.) de largeur, de 10 chaînes (201,2 m) de longueur et orientées est-ouest. On a obtenu ainsi au total 23 virées contenant chacune 10 parcelles de 0.05 acre (0,02 ha) qui ont été choisies de manière aléatoire.

Nous avons ensuite effectué des sondages à différentes intensités, à savoir à 5 pour 100 et 10 pour 100 avec 5 répétitions et à 20 pour 100 avec 2 répétitions. Puis les variances intra-virées et inter-virées et la variance totale furent calculées pour chaque cas.

Puisque les virées et les parcelles sont choisies de manière aléatoire au sein d'une population homogène, on peut espérer que la variation intra-virées sera, en moyenne, égale à la variation inter-virées et qu'elle constituera de fait une estimation correcte de la variance totale.



CHAPITRE IV

PRÉSENTATION ET ANALYSE DES DONNÉES EXPÉRIMENTALES

4.1 SONDAGES SYSTÉMATIQUES SELON DIVERSES FRACTIONS DE SONDAGE.

Le tableau III, présenté précédemment, montre la distribution des parcelles dans chaque strate selon les diverses fractions de sondage. Pour chaque strate et l'ensemble des strates, la variance et le volume total ont été calculés. Les résultats de ces différents essais sont présentés au tableau IV qui suit. Deux aspects de ces résultats seront étudiés:

L'estimation du volume moyen par parcelle;

L'estimation de la variance.

4.1.1 ESTIMATION DU VOLUME MOYEN PAR PARCELLE.

A partir du volume total obtenu à chaque intensité de sondage, l'on a d'abord calculé l'erreur d'échantillonnage en pourcentage d'après sa différence avec le volume total de $185\ 877\ \text{pi}^3$ ($5\ 260,3\ \text{m}^3$). Ensuite, après avoir trouvé le volume moyen par parcelle pour chaque fraction de sondage, l'on a pointé les intervalles de sécurité pour la moyenne générale

TABLEAU IV

Résultats relatifs à l'estimation du volume total et de la variance à partir des différents essais de sondage systématique avec choix aléatoire de la première parcelle						
ESSAIS	5 pour 100			10 pour 100		
	Volume	Volume	Variance	Volume	Volume	Variance
A	π^3 : 9 565	199.3	5 591.0	18 305	190.6	5 351.3
	m^3 : 270,7	5,64	4,48	518,0	5,39	4,28
B	π^3 : 8 743	182.2	3 800.0	18 815	196.0	5 530.6
	m^3 : 247,4	5,16	3,04	532,5	5,55	4,43
C	π^3 : 9 079	189.2	4 828.6	17 468	182.0	5 965.2
	m^3 : 256,9	5,35	3,87	494,3	5,15	4,78
D	π^3 : 9 294	193.6	3 662.0	17 609	183.4	5 819.0
	m^3 : 263,0	5,48	2,93	498,3	5,19	4,66
E	π^3 : 9 517	198.3	4 588.0	18 546	193.2	4 356.6
	m^3 : 269,3	5,61	3,67	524,8	5,47	3,49
ESSAIS	20 pour 100			40 pour 100		
	Volume	Volume	Variance	Volume	Volume	Variance
A	π^3 : 36 127	191.2	4 673.0	74 483	195.0	4,682.6
	m^3 : 1 022,4	5,41	3,74	2 107,9	5,52	3,75
B	π^3 : 36 564	193.5	4 526.7	74 319	194.6	4 909.4
	m^3 : 1 034,8	5,48	3,62	2 103,2	5,51	3,93
C	π^3 : 36 517	193.2	5 526.9			
	m^3 : 1 033,4	5,47	4,43			

afin de se rendre compte si les moyennes obtenues se localisaient à l'intérieur des limites fixées.

Ces intervalles de sécurité pour le volume moyen ont été effectués à l'aide de la formule suivante:

$$x_s = \bar{X} + tS/\sqrt{n}$$

$$x_i = \bar{X} - tS/\sqrt{n}$$

où \bar{X} = moyenne générale: $195 \text{ pi}^3 (5,52 \text{ m}^3)$

t = variable de Student: en raison de la taille de "n",
on l'a assimilée ici à une variable normale réduite;

S = racine carrée de la variance de la population totale;

n = nombre de parcelles de l'échantillon;

x_s = limite supérieure de l'intervalle de confiance;

x_i = limite inférieure de l'intervalle de confiance;

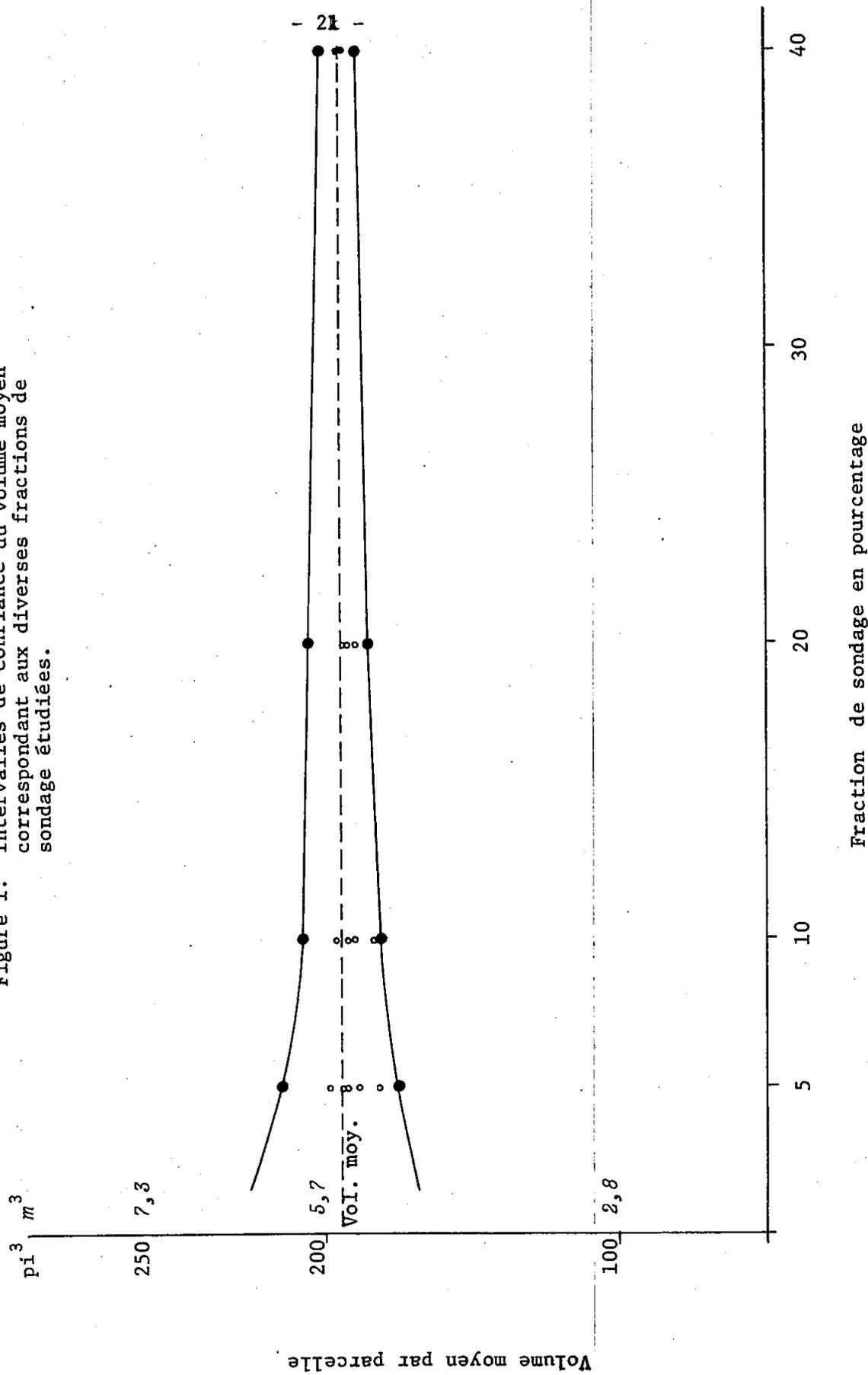
Les limites supérieures et inférieures pour chaque fraction de sondage ainsi que les intervalles de sécurité pour le volume moyen par parcelle pour l'ensemble des différents essais de chaque fraction de sondage sont présentés au tableau V qui suit.

TABLEAU V

Intervalles de confiance de la moyenne pour chaque fraction de sondage				
Intensité (pourcentage)	x_s		x_i	
	(pi ³)	(m ³)	(pi ³)	(m ³)
5%	214.3	6,06	175.7	4,97
10%	208.8	5,91	181.2	5,13
20%	204.8	5,79	185.2	5,24
40%	201.8	5,71	188.0	5,32
Limites pour le volume moyen par parcelle				
Intensité (pourcentage)	x_s		x_i	
	(pi ³)	(m ³)	(pi ³)	(m ³)
5%	203.7	5,76	186.3	5,27
10%	201.3	5,7	188.8	5,34
20%	199.4	5,64	190.6	5,39
40%	198.1	5,61	191.9	5,43

La figure 1 nous indique les intervalles de confiance pour le volume moyen par parcelle tandis que le tableau VI présente l'erreur d'échantillonnage calculée d'après la différence de chaque fraction de sondage avec le volume total.

Figure 1: Intervalles de confiance du volume moyen correspondant aux diverses fractions de sondage étudiées.



TABEAU VI

Tableau donnant l'erreur d'échantillonnage (E. Ec.) calculée d'après la différence de chaque sondage avec le volume total de 185 877 pi ³ (5 260,3 m ³) ainsi que le volume moyen par parcelle								
ESSAIS	5 pour 100		10 pour 100		20 pour 100		40 pour 100	
	Volume	E. Ec. p. 100	Volume	E. Ec. p. 100	Volume	E. Ec. p. 100	Volume	E. Ec. p. 100
A	pi ³ : 9 565 m ³ : 270,7	2.9	pi ³ : 18 305 m ³ : 518	1.5	pi ³ : 36 127 m ³ : 1 022,4	1.3	pi ³ : 74 483 m ³ : 2 107,9	0.12
B	8 743 247,4	5.9	18 815 532,5	1.2	36 564 1 034,8	0.66	74 319 2 103,2	0.04
C	9 079 256,9	2.3	17 468 494,3	6.0	36 517 1 033,4	0.78	-	-
D	9 294 263,0	0.0	17 609 498,3	5.0	-	-	-	-
E	9 517 269,3	2.4	18 546 524,8	0.2	-	-	-	-
Total:	pi ³ : 46 198 m ³ : 1 307,4		pi ³ : 90 743 m ³ : 2 568,0		pi ³ : 109 208 m ³ : 3 090,59		pi ³ : 148 802 m ³ : 4 211,1	
Vol. moy.:	9 240 261,49	0.6	18 149 513,6	2.4	36 403 1 030.2	2.1	74 401 2 105,5	0.68

Comme nous pouvons le constater au tableau IV, les résultats obtenus nous incitent à croire qu'en général les différents sondages systématiques simulés présentent une sous-estimation du volume. Cependant, il est à noter que toutes les valeurs obtenues demeurent à l'intérieur des limites fixées. Néanmoins, à l'exception de deux cas (5 pour 100, essai B et 10 pour 100, essai C), les estimations du volume ne dépassent pas l'erreur permise de ± 5 pour 100.

En résumé, il résulte de ces expériences que le sondage systématique s'avère un procédé d'échantillonnage très valable, en ce sens qu'en plus d'être plus économique que le sondage aléatoire, il présente des estimations de volume satisfaisantes et demeurant dans les limites permises.

4.1.2 ESTIMATION DE LA VARIANCE.

La variance du volume moyen de chaque strate ainsi que celle de l'ensemble des strates étant connues, nous avons déterminé, pour chaque fraction de sondage et pour chaque strate, les intervalles de sécurité. Cependant, afin de ne pas alourdir le texte, nous nous limiterons aux résultats de deux strates, à savoir le *DRY-OX* 90 ans et le *DRY-OX* 70 ans. Les tableaux VII, VIII et IX ainsi que les figures 2, 3 et 4 résument tous ces calculs.

TABLEAU VII

Résultats des estimations de la "variance totale" (π^3) pour chaque fraction de sondage					
Intensité pour 100	Répétitions				
	A Variance	B Variance	C Variance	D Variance	E Variance
5	π^3 : 5 591.0 m^3 : 4,48	π^3 : 3 800.0 m^3 : 3,04	π^3 : 4 828.6 m^3 : 3,87	π^3 : 3 662.0 m^3 : 2,93	π^3 : 4 588.0 m^3 : 3,67
10	5 351.3 4,28	5 530.6 4,43	5 965.0 4,78	5,819.0 4,66	4 356.6 3,49
20	4 673.0 3,74	4 526.7 3,62	5 527.0 4,43	- -	- -
40	4 682.6 3,75	4 909.4 3,93	- -	- -	- -

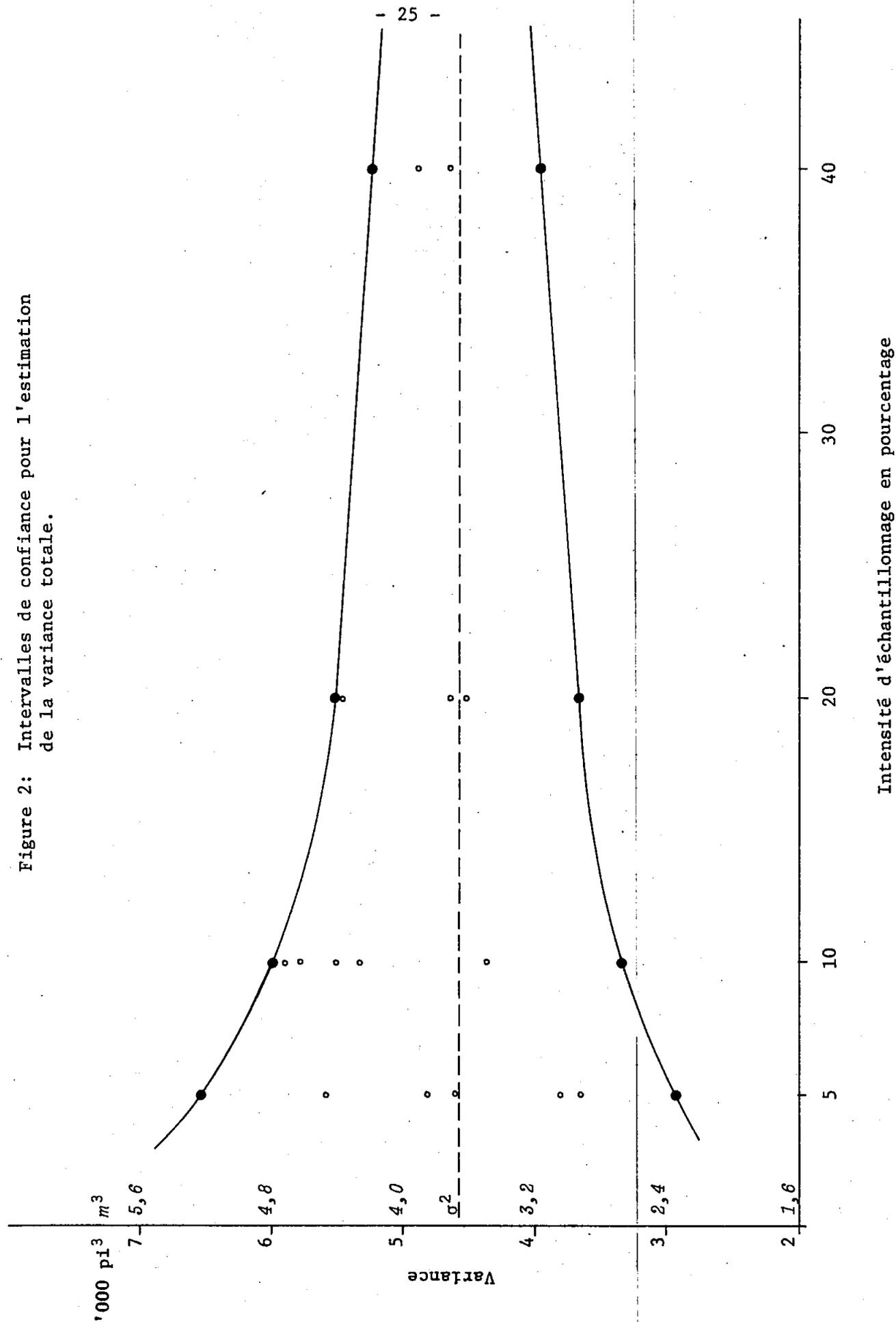
Note: La variance totale réelle est de 4 577.4 π^3 (3,66 m^3)

TABLEAU VIII

Résultats des estimations de la variance pour chaque fraction de sondage dans le type <i>Dryopteris-Oxalis</i> 90 ans.					
Intensité pour 100	Répétitions				
	A Variance	B Variance	C Variance	D Variance	E Variance
5	π^3 : 9 017.0 m^3 : 7,22	π^3 : 4 236.4 m^3 : 3,39	π^3 : 4 258.0 m^3 : 3,41	π^3 : 5 938.0 m^3 : 4,75	π^3 : 4 221.4 m^3 : 3,38
10	1 988.8 1,59	1 936.6 1,55	6 230.8 4,99	4 962.8 3,97	2 518.1 2,02
20	5 331.0 4,27	2, 288.8 1,83	5 599.5 4,48	- -	- -
40	3 645.8 2,92	3 723.0 2,98	- -	- -	- -

Note: La variance réelle de DRY-OX 90 ans est de 3 698.4 π^3 (2,96 m^3)

Figure 2: Intervalles de confiance pour l'estimation de la variance totale.



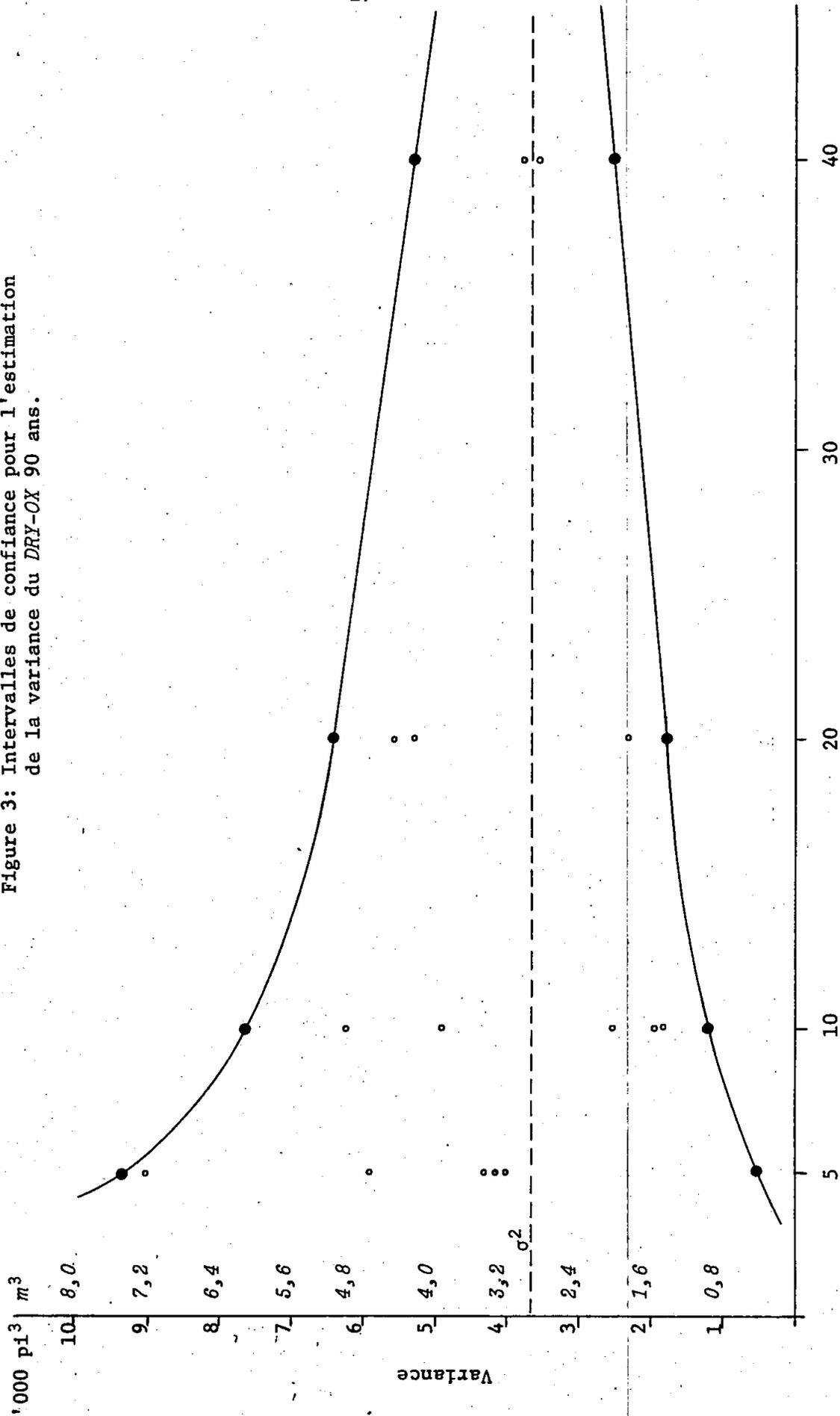
Intensité d'échantillonnage en pourcentage

TABLEAU IX

Résultats des estimations de la variance en π^3 pour chaque fraction de sondage dans le type <i>Dryopteris-Oxalis</i> 70 ans					
Intensité pour 100	Répétitions				
	A Variance	B Variance	C Variance	D Variance	E Variance
5	π^3 : 3 937.2 m^3 : 3,15	π^3 : 2 638.4 m^3 : 2,11	π^3 : 3 113.0 m^3 : 2,49	π^3 : 2 434.5 m^3 : 1,95	π^3 : 3 014.3 m^3 : 2,41
10	3 765.0 3,01	4 492.5 3,6	4 480.4 3,59	4 480.4 3,59	3 396.0 2,72
20	3 640.0 2,91	3,640.0 2,91	3,620.8 2,90	- -	- -
40	3 910.8 3,12	3 793.2 3,0	-	-	-

Note: La variance réelle du *Dryopteris-Oxalis* 70 ans est de 3 522.0 π^3 (2,82 m^3)

Figure 3: Intervalles de confiance pour l'estimation de la variance du DRY-OX 90 ans.



Intensité d'échantillonnage en pourcentage

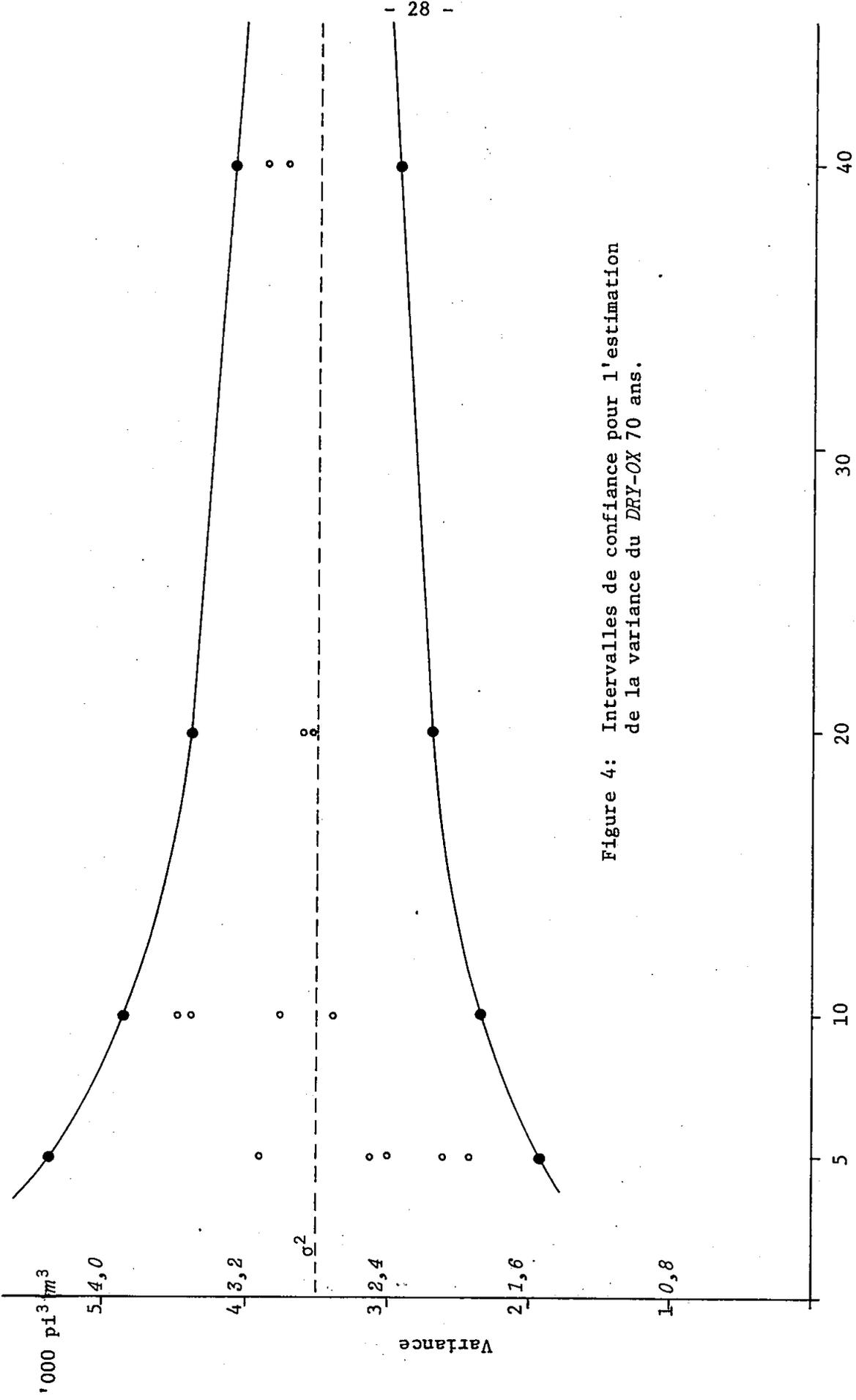


Figure 4: Intervalles de confiance pour l'estimation de la variance du DRY-OX 70 ans.

Intensité d'échantillonnage en pourcentage

4.1.3 CONCLUSION.

Spurr (1954) affirme qu'en général le sondage systématique présente une erreur d'échantillonnage maximum. Or, d'après les calculs effectués, nous remarquons la même tendance, c'est-à-dire que les sondages simulés donnent, en moyenne, une estimation de la variance supérieure à la variance de la population étudiée.

Toutefois, comme nous allons le voir par la suite, l'examen des intervalles de confiance nous incite à croire que le sondage systématique avec choix aléatoire de la première parcelle procure, en moyenne, des estimations de la variance qui sont satisfaisantes et qui se situent très bien autour de la variance des populations étudiées.

4.2 SONDAGES ALÉATOIRES.

4.2.1 SONDAGE EN GRAPPES.

Les buts précis des expériences effectuées concernant les plans de sondage en grappes sont les suivants:

- a) Trouver un moyen de calculer la précision des estimations, c'est-à-dire de calculer l'erreur d'échantillonnage, de façon correcte tout en réalisant un inventaire qui ne soit pas trop dispendieux;
- b) Chercher s'il existe une corrélation entre les volumes des grappes de taille différente;
- c) Déterminer l'efficacité relative du sondage en grappes.

Dans cette optique, nous avons effectué, pour le DRY-OX 70 ans seulement, des grappes de 3, 4 et 5 parcelles de 0.05 acre (0,02 ha).

<u>Dimension de la grappe</u>	<u>Nombre de grappes</u>
3 parcelles	130
4 parcelles	82
5 parcelles	51

Pour chaque cas, l'on a calculé les variances intra-grappes et inter-grappes ainsi que la variance totale. Pour ce faire, nous avons employé les formules statistiques suivantes:

4.2.1.1 Définition des termes.

- Soit x_{ij} = le volume de j^e parcelle de la i^e grappe;
 N = le nombre total de grappes de la population;
 m = le nombre de parcelles par grappe;
 n = le nombre de grappes de l'échantillon;
 Nm = le nombre total de parcelles de la population.

Alors,

$$\bar{X}_i = \frac{\sum_{j=1}^m \bar{X}_{ij}}{m}$$

= volume moyen par parcelle de la i^e grappe.

$$\bar{X}_N = \frac{\sum_{i=1}^N \bar{X}_i}{N}$$

= volume moyen par parcelle pour toutes les grappes de la population.

$$\bar{X}_n = \frac{\sum_{i=1}^n \bar{X}_i}{n}$$

= volume moyen par parcelle pour les grappes de l'échantillon.

4.2.1.2 Somme des carrés totaux.

$$\begin{aligned} &= \sum_i \sum_j (\bar{X}_{ij} - \bar{X}_N)^2 \\ &= \sum_i \sum_j X_{ij}^2 - \left(\sum_i \sum_j X_{ij} \right)^2 \end{aligned}$$

d'où la variance totale

$$\sigma^2 = \frac{\sum_i \sum_j X_{ij}^2 - \frac{(\sum_i \sum_j X_{ij})^2}{Nm}}{Nm}$$

4.2.1.3 Somme des carrés inter-grappes.

$$m\sigma^2_b = m \sum (\bar{X}_i - \bar{X}_N)^2$$

d'où la variance inter-grappes =

$$\sigma^2_b = \frac{\sum (\bar{X}_i - \bar{X}_N)^2}{N - 1}$$

4.2.1.4 Somme des carrés intra-grappes.

$$= \sum_i \sum_j (X_{ij} - \bar{X}_i)^2$$

d'où la variance intra-grappes =

$$\sigma^2_w = \frac{\sum_i \sum_j (X_{ij} - \bar{X}_i)^2}{N(m - 1)}$$

Puis à l'aide des données mentionnées ci-dessus, on forme,
pour chaque sondage, le tableau suivant:

TABEAU X

Variances intra-grappes et inter-grappes et variance totale pour le sondage en grappes			
Source de variation	N.D.L.	Somme des Carrés	Carrés moyens
Variation inter-grappes	N-1	$m \sum (\bar{X}_i - \bar{X}_N)^2$	σ_b^2
Variation intra-grappes	N(m-1)	$\sum_{ij} (\bar{X}_{ij} - \bar{X}_i)^2$	σ_n^2
Variation totale	Nm-1	$\sum_{ij} (\bar{X}_{ij} - \bar{X}_N)^2$	σ^2

4.2.1.5 Efficacité relative.

La variance d'une moyenne par parcelle, établie d'après un échantillon de "n" grappes, est égale à:

$$V(\bar{X}_n) = \frac{N-n}{N} \cdot \frac{\sigma_b^2}{n}$$

Cependant, si au lieu de "N" grappes de "m" parcelles, l'on tirait "Nm" parcelles d'une manière aléatoire, alors la variance de la moyenne deviendrait:

$$V(\bar{X}_{Nm}) = \frac{Nm - mn}{Nm} \cdot \frac{\sigma_b^2}{nm}$$

et l'efficacité relative du sondage en grappes par rapport au simple sondage aléatoire est donnée par le rapport suivant:

$$E = \frac{V(\bar{X}_{Nm})}{V(\bar{X}_n)} = \frac{S^2}{mS_b^2} \dots \dots \dots (12)$$

Or, si mS_b^2 est égale à S^2 , l'efficacité de ces deux plans de sondage est identique: on aurait donc le cas où les volumes des parcelles d'une grappe ne sont nullement corrélés entre eux.

4.2.1.6 Coefficient de corrélation.

L'efficacité relative peut être envisagée sous une dimension nouvelle si on l'établit en fonction du coefficient de corrélation, "ρ" des volumes des grappes. Sukhatme (1954) exprimait le coefficient de corrélation entre les éléments d'une même grappe en fonction de la variance inter-grappes, mS_b^2 , la variance intra-grappes, S_w^2 , et la variance totale:

$$\rho = \frac{\left(\frac{N-1}{N}\right) \left(\frac{S_b^2 - \frac{S_w^2}{m}}{m}\right)}{\left(\frac{Nm-1}{Nm}\right) S^2} \dots\dots\dots (13)$$

qui, pour un "N" suffisamment grand (>30), devient:

$$\rho = \frac{S_b^2 - \frac{S_w^2}{m}}{S^2} \dots\dots\dots (14)$$

De plus, quand les grappes sont formées de manière aléatoire, S_b^2 et S_w^2 sont égales à S^2 et alors:

$$\rho = -\frac{1}{Nm-1}$$

Pour un sondage en grappes ordinaire, "ρ" a une valeur positive et il décroît à mesure que "m" croît. Cependant, pour un sondage en grappes systématique, c'est-à-dire que chaque grappe est composée d'un groupe de parcelles systématiquement espacées et distribuées sur tout le territoire, la variance intra-grappes, S_w^2 , tend à être plus grande que la variance inter-grappes, S_b^2 . Alors on aura un coefficient de corrélation négatif.

Toutefois, il est à noter que ce fait dépend de la distribution des parcelles de la grappe, comme celle, par exemple, que Cherng-Jiann (1960) a adoptée. C'est pourquoi les résultats seront différents dans nos expériences, car la distribution des parcelles à l'intérieur de la grappe n'était pas la même que celle de Cherng-Jiann Shiue.

A l'aide des équations ci-dessus, on détermine que:

$$V(\bar{X}_n) = \frac{N-n}{N} \cdot \frac{Nm-1}{m(N-1)} \cdot \frac{S^2}{nm} \left[1 + (m-1)\rho \right]$$

et

$$E = \left[\frac{m(N-1)}{Nm-1} \right] \left[\frac{1}{1 + (m-1)\rho} \right]$$

qui pour "N" grand deviennent:

$$V(\bar{X}_n) = \frac{N-n}{N} \cdot \frac{S^2}{nm} \left[1 + (m-1)\rho \right] \dots\dots\dots (15)$$

et

$$E = \frac{1}{1 + (m-1)\rho}$$

et si $\rho \rightarrow 0$ on a: $E \rightarrow 1$

Enfin tous les calculs furent effectués à l'aide de l'expression suivante:

$$E = \frac{\text{est. } S^2}{mS_b^2} \dots\dots\dots (16)$$

4.2.1.7 Résultats obtenus.

Ces différents sondages en grappes ont donné les résultats présentés aux tableaux XI, XII et XIII qui suivent:

TABLEAU XI

Sondage avec grappes de 3 parcelles			
Source de variation	N.D.L.	Somme des carrés	Carrés moyens
Variation inter-grappes	N-1=129	π^3 : 597 693 m^3 : 478,7	4 633.3 3,71
Variation intra-grappes	N(m-1)=260	π^3 : 708 615 m^3 : 567,5	2 725.4 2,18
Variation totale	Nm-1=389	π^3 : 1 306 308 m^3 : 1 046.2	3 358.2 2,69

TABLEAU XII

Sondage avec grappes de 4 parcelles			
Source de variation	N.D.L.	Somme des carrés	Carrés moyens
Variation inter-grappes	81	π^3 : 385 128 m^3 : 308,4	4 754.6 3,81
Variation intra-grappes	246	π^3 : 735 998 m^3 : 589,4	2 991.8 2,40
Variation totale	254	π^3 : 1 121 126 m^3 : 897,9	3 428.5 2,74

TABEAU XIII

Sondage avec grappes de 5 parcelles			
Source de variation	N.D.L.	Somme des carrés	Carrés moyens
Variation inter-grappes	50	π^3 : 219 541	4 390.8
		m^3 : 175,8	3,52
Variation intra-grappes	204	π^3 : 519 488	2 546.5
		m^3 : 416,0	2,04
Variation totale	254	π^3 : 739 029	2 910.0
		m^3 : 591,9	2,33

4.2.1.8 Discussion des résultats.

Si l'on se réfère au tableau II on a pour le *Dryopteris-Oxalis* de 70 ans, une variance de 3 522 π^3 (2,82 m^3). Alors, si l'on compare la variance totale obtenue à partir des sondages par grappes de différentes tailles, on obtient les résultats résumés dans le tableau XIV ci-dessous:

TABEAU XIV

Comparaison entre la variance totale des sondages en grappes de tailles différentes et la variance du <i>Dryopteris-Oxalis</i> 70 ans			
	Variance		
	Taille des grappes		
	3	4	5
Estimation	π^3 : 3 358.2 m^3 : 2,69	3 428.5 2,74	2 910.0 2,33
Variance de la population	π^3 : 3 522.0 m^3 : 2,82	3 522.0 2,82	3 522.0 2,82
Différence	- 4.6 p. 100	- 2.6 p. 100	- 17.4 p. 100

Tout d'abord, l'étude de ces chiffres révèle que les sondages en grappes présentent une variance inférieure à celle de la population totale: ce qui, à première vue, nous paraît inusité car, théoriquement, l'on devrait obtenir des résultats quasi semblables puisque nous avons employé les mêmes données et les mêmes formules de calcul. Néanmoins, cette situation peut s'expliquer par le fait que les grappes de ces sondages n'ont pas englobé toutes les parcelles de la population et que, par conséquent, l'on n'a pas tenu compte de toute la variation de la population.

En second lieu, à l'aide de la formule (16), l'on a calculé l'efficacité relative de ces sondages différents, dont les résultats apparaissent au tableau XV ci-dessous.

TABLEAU XV

Efficacité relative des sondages en grappes			
Variance	Variance		
	Taille des grappes		
	3	4	5
Inter-grappes	π^3 : 4 644.4	4 754.6	4 390.8
	m^3 : 3,71	3,81	3,52
Totale	π^3 : 3 358.2	3 428.5	2 910.0
	m^3 : 2,69	2,74	2,33
Efficacité relative	0.73	0.72	0.66

Ces quelques résultats indiquent que l'efficacité relative est inversement proportionnelle à la taille de la grappe, *i.e.* qu'à mesure que la taille de la grappe augmente, l'efficacité diminue. En d'autres termes, le sondage avec des grappes de 5 parcelles devra être plus considérable que le sondage avec grappes de 3 parcelles pour présenter une estimation de la variance avec la même précision.

De plus, les éléments d'une même grappe, en pratique, seront plus semblables entre eux que ceux qui appartiennent à des grappes différentes. Ainsi, la variation du volume entre différentes superficies d'un même territoire, sera généralement inférieure à celle qui existe entre les superficies de territoires différents. En conséquence, la variance d'une estimation basée sur un sondage en grappes sera ordinairement supérieure à celle basée sur un échantillonnage équivalent d'éléments choisis indépendamment (Sukhatme, 1954). Or, la manière selon laquelle la variance d'une estimation augmente en fonction de la taille de la grappe peut être élucidée à l'aide du concept de la corrélation entre les éléments d'une même grappe, à savoir le coefficient de corrélation " ρ " tel que défini par la formule (14) ci-dessus.

D'un autre côté, afin de mieux saisir la relation qui existe entre le " ρ " et le changement de variance en fonction de la taille de la grappe, l'on explicitera la formule (15), mentionnée à la page 34. Cette formule, exprimant la variance de la moyenne de " n " grappes, a été développée pour la première fois par Hansen et Hurwitz (1942). Elle se compose essentiellement de 3 facteurs. Le premier est le facteur pour une population limitée $N-n/(N-1)$. Le second représente l'estimation de la variance de la moyenne basée sur " M " éléments choisis aléatoirement. Le troisième facteur mesure la contribution à l'estimation de la variance du sondage en grappes.

Si $M = 1$, ce troisième facteur égale l'unité; mais si M est > 1 , le facteur $(M-1)$ mesurera alors le changement relatif de la variance, causé par un sondage en grappes au lieu du sondage ordinaire. En pratique, " ρ " est positif et décroît à mesure que " M " croît, mais son taux de décroissance est relativement faible, comparé au taux d'accroissement de " M " de sorte qu'ordinairement une augmentation de la taille des grappes conduit à une augmentation substantielle de la variance de l'estimation. Cette hypothèse est assez bien illustrée par les données du tableau XVI ci-dessous.

TABLEAU XVI

Changement relatif de la variance en fonction de l'accroissement de la taille des grappes			
	Taille des grappes		
	3	4	5
ρ	0.19	0.13	0.126
$(M-1)$	0.38	0.39	0.504

4.2.1.9 Fonction économique.

Nous ne possédons pas évidemment assez de données pour déterminer de façon satisfaisante tous les paramètres et la structure de la fonction économique d'un sondage en grappes. C'est pourquoi nous nous contenterons, pour le moment, de suggérer une méthode théorique ou une approche.

Comme nous l'avons déjà dit, la variance inter-grappes aura tendance à diminuer avec la taille de la grappe, tandis que la variance intra-grappes aura tendance à augmenter.

Nous pouvons donc poser:

$$\text{Variance inter-grappes} = S_b^2 = \frac{a'}{M^g} \dots \dots \dots (1)$$

où a' et g' sont des paramètres et "M" la taille relative des grappes, ou encore:

$$\text{Variance intra-grappes} = S_w^2 = a''M^{g''} \dots \dots \dots (2)$$

D'autre part, le coût de sondage, pour une superficie donnée, comme pour le cas étudié auparavant, peut s'exprimer comme suit:

$$C_o = C_1 Mn + C_2 b \sqrt{n} \dots \dots \dots (3)$$

où C₁ = coût de sondage d'une parcelle

M = taille de la grappe

C₂ = coût de déplacement unitaire entre les grappes

n = nombre de grappes

b = coefficient de proportionnalité

Enfin, la variance de la moyenne:

$$V(\bar{x}) = \frac{N-n}{N} \cdot \frac{S_b^2}{n} \rightarrow \frac{S_b^2}{n} \dots \dots \dots (4)$$

qui, après substitution, devient:

$$V(\bar{x}) = \frac{S^2}{n} - \frac{(M-1) a''M^{g-1}}{n} \dots \dots \dots (5)$$

en notant que:

$$S_b^2 = MS^2 - (M-1) S_w^2$$

Alors la taille de la grappe qui minimise les coûts pour une variance donnée de la moyenne (ou une erreur admissible) est déterminée en formant la fonction:

$$P = C_1 Mn + C_2 b \sqrt{n} - \lambda \left[\frac{S^2}{n} - \frac{(M-1)}{n} a''M^{g-1} - V(\bar{x}) \right]$$

En résolvant les dérivées partielles $\frac{\delta F}{\delta M}$, $\frac{\delta F}{\delta n}$ et $\frac{\delta F}{\delta \lambda}$ et égalant à zéro, il n'y a pas de solution explicite immédiate en "M". Cependant, l'expression suivante donne l'optimum de "M":

$$\frac{aM^{g-1} \left[gM - (g-1) \right]}{s^2 - (M-1)aM^{g-1}} = 1 - \left[\frac{1 + 4 C_0 C_1 M}{b^2 C^2} \right]^{-\frac{1}{2}}$$

Le terme de gauche n'est fonction que des seuls caractéristiques de la variance alors que celui de droite est lié à la fonction économique.

Alors, quand les divers paramètres de cette fonction ont été déterminés, la valeur optimale de "M" peut être explicitée très facilement.

4.2.2 SONDAGE PAR VIRÉES.

Au cours de cette expérience, l'on a divisé la strate du *DRY-OX* 70 ans en bandes d'une demi-chaîne (10,06 m) de largeur par 10 chaînes (201,1 m) de longueur. On obtint ainsi un nombre total de 23 virées contenant chacune 10 parcelles-échantillons de 0.05 acre (0,02 ha). Puis on fit un choix aléatoire d'un certain nombre de virées, d'après les fractions de sondage de 5 p. cent, 10 p. cent et 20 p. cent ainsi que des parcelles sur chaque virée. Le tableau ci-dessous résume la distribution de ces différents sondages:

TABLEAU XVII

Distribution des parcelles lors de sondages par virées selon différentes fractions de sondage.			
Fraction de sondage	Nombre de parcelles	Nombre de virées	Nombre de parcelles par virées
1 : 5	12	2	6 (5 répétitions)
1 : 10	24	8	3 (5 répétitions)
1 : 10	24	4	6 (5 répétitions)
1 : 20	45	5	9 (2 répétitions)

Jusqu'à un certain point, cette méthode de sondage peut être considérée, ici, comme un sondage en grappe. Toutefois, elle diffère essentiellement du sondage en grappe en ce que les parcelles de chaque virée sont choisies aléatoirement alors que, dans le cas de grappes, le choix des parcelles est systématique.

Le but de la présente expérience est de rechercher si ce mode de sondage par virées permet d'évaluer de façon appropriée la variance de la population. Théoriquement parlant, l'on a posé l'hypothèse suivante: puisque les virées et les parcelles ont été choisies d'une manière aléatoire au sein d'une population homogène ou strate, on peut espérer que la variance inter-virées sera, en moyenne, égale à la variance intra-virées et que, par conséquent, ce procédé de sondage constitue une estimation correcte de la variance totale. Alors, pour vérifier cette hypothèse, nous avons effectué des tests de "F" de Snedecor pour les variances "inter" et "intra" et nous avons ensuite comparé l'estimation de la variance totale obtenue pour chaque sondage à la variance totale de la population étudiée. Les tableaux XVIII à XXI nous présentent les résultats de cette étude.

TABLEAU XVIII

Analyse de variance pour le sondage à 5 p. 100 avec 6 parcelles par virée						
Source de variation	N.D.L.	Variances				
		A	B	C	D	E
Var. inter-virées	5	π^3 : 1 558.6	2 193.4	4 695.8	8 235.0	5 920
		m^3 : 1,25	1,76	3,76	6,59	4,74
Var. intra-virées	6	π^3 : 1 698.8	3 568.0	4 759.3	1 249.2	1 070
		m^3 : 1,36	2,86	3,87	1,00	0,86
Var. totale	11	π^3 : 1 630.2	2 943.2	4 730.5	4 424.7	3 274
		m^3 : 1,30	2,36	3,79	3,54	2,62

TABLEAU XIX

Analyse de variance pour le sondage à 10 p. 100 avec 3 parcelles par virée						
Source de variation	N.D.L.	Variances				
		A	B	C	D	E
Var. inter-virées	2	π^3 : 3 196.0	669.0	6 955.5	7 548.0	5 501
		m^3 : 2,56	0,53	5,57	6,04	4,40
Var. intra-virées	21	π^3 : 794.2	3 786.2	2 162.2	3 103.2	4 788
		m^3 : 0,64	3,03	1,73	2,48	3,83
Var. totale	23	π^3 : 1 066.0	3 515.0	2 579.0	3 489.7	4 850
		m^3 : 0,85	2,81	2,06	2,79	3,88

TABLEAU XX

Analyse de variance pour le sondage à 10 p. 100 avec 6 parcelles par virée						
Source de variation	N.D.L.	Variances				
		A	B	C	D	E
Var. inter-virées	5	π^3 : 10 703.0	4 322.4	9 433.0	4 432.6	1 809
		m^3 : 8,57	3,46	7,55	3,55	1,45
Var. intra-virées	18	π^3 : 2 620.0	2 569.5	1 172.0	2 691.0	1 493
		m^3 : 2,10	2,06	0,94	2,15	1,19
Var. totale	23	π^3 : 4 377.2	2 950.6	2 967.8	2 200.0	1 518
		m^3 : 3,50	2,36	2,38	1,76	1,21

TABLEAU XXI

Analyse de variance pour le sondage à 20 p. 100 avec 9 parcelles par virée			
Source de variation	N.D.L.	Variances	
		A	B
Var. inter-virées	8	π^3 : 3 766.0	5 360.0
		m^3 : 3,02	4,29
Var. intra-virées	36	π^3 : 3 623.0	2 415.0
		m^3 : 2,90	1,93
Var. totale	44	π^3 : 3 649.4	2 950.6
		m^3 : 2,92	2,36

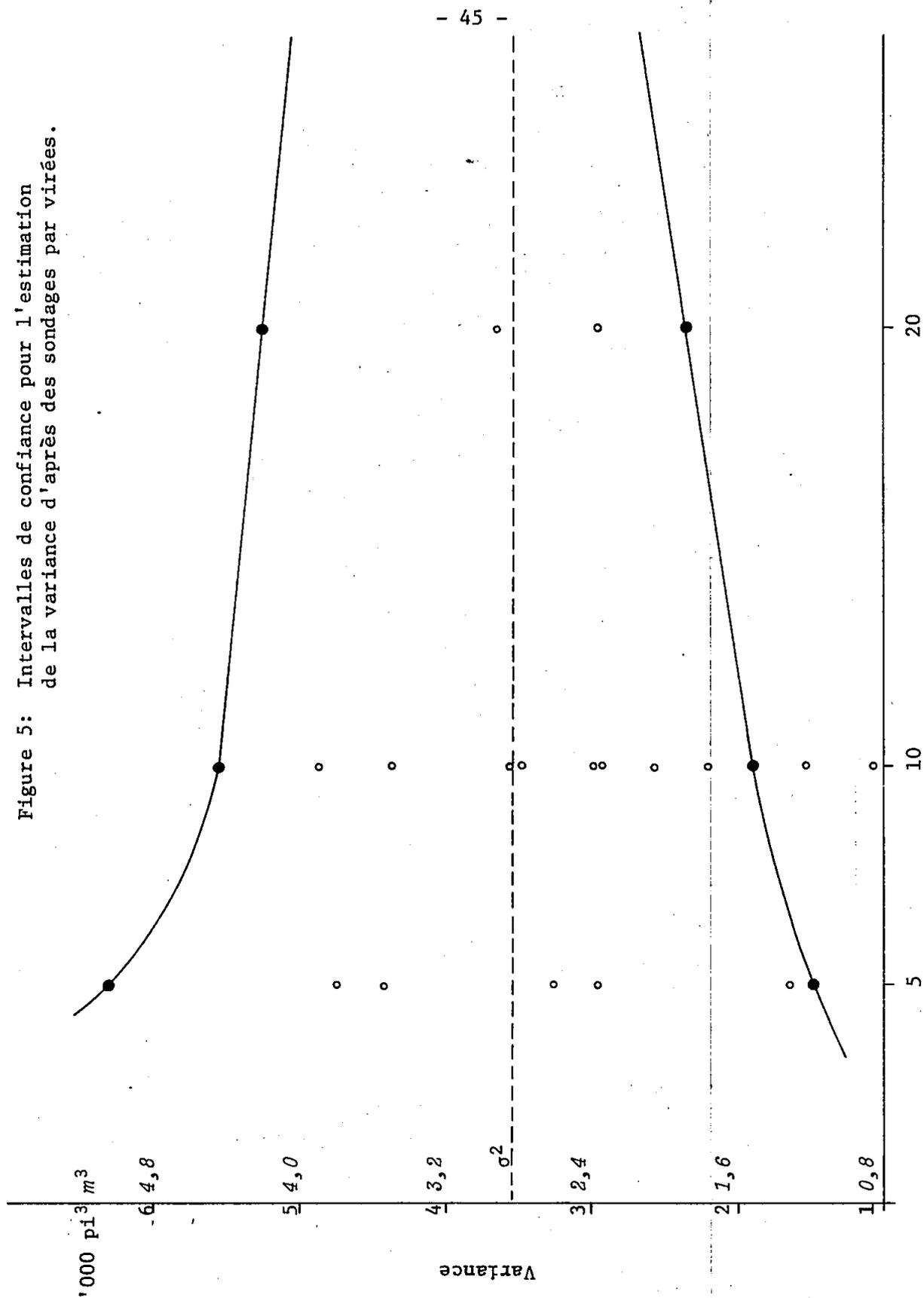
A l'aide des résultats ci-dessus, nous avons calculé les intervalles de confiance pour ces différents essais de sondages par virées, afin de comparer l'exactitude des estimations de la variance obtenues avec la variance du *DRY-OX* 70 ans qui est de $3\,522\text{ pi}^3$ ($2,82\text{ m}^3$). La figure 5 nous montre ces intervalles de confiance pour chaque fraction de sondage.

Dans l'ensemble, toutes les estimations obtenues se situent très bien autour de la variance totale " σ^2 ". Ce qui implique que le sondage aléatoire par virées présente, en général, une bonne estimation de la variance. Cependant, nous remarquons que dans le cas du sondage à 10 p. 100 d'intensité, deux points ou deux estimations se localisent à l'extérieur des limites de confiance fixées, à savoir les estimations des essais A et E.

Ce fait peut s'expliquer comme suit: au cours de l'essai "A", nous remarquons que la variance intra-virées est négligeable si on la compare avec la variance inter-virées. Alors, si nous considérons les éléments de ce sondage, nous nous apercevons que l'une des 3 virées de 8 parcelles chacune qui constituent ce sondage, contient des parcelles dont le volume est de beaucoup inférieur aux volumes des parcelles des 2 autres virées: ainsi, par exemple, le volume total de cette première virée est de $1\,417\text{ pi}^3$ ($40,1\text{ m}^3$) tandis que les deux autres virées contiennent respectivement $1\,737$ et $1\,696\text{ pi}^3$ ($49,1$ et 48 m^3). C'est ce qui explique en grande partie cette estimation de la variance inférieure aux limites fixées. Enfin, l'on peut affirmer qu'une explication analogue s'applique pour l'essai E.

Comme vérification de l'hypothèse posée au départ, nous avons effectué, pour chaque fraction de sondage, des tests de "F" de Snedecor au niveau de probabilité de 95 p. 100 dont vous trouverez les résultats au tableau XXII.

Figure 5: Intervalles de confiance pour l'estimation de la variance d'après des sondages par virées.



Intensité d'échantillonnage en pourcentage.

TABLEAU XXII

Résultats des tests de "F" de Snedecor pour chaque fraction de sondage					
Fraction de sondage	"F"				
	Répétitions				
	A	B	C	D	E
5 p. 100	0.92	0.61	0.99	<u>6.59</u>	<u>5.53</u>
10 p. 100	<u>4.93</u>	0.18	<u>3.22</u>	2.43	1.15
	<u>4.08</u>	1.68	<u>8.05</u>	0.17	1.21
20 p. 100	1.04	<u>2.22</u>			

Tous les chiffres soulignés d'un trait indiquent qu'il y a une différence significative entre les variances inter et intra-virées. Donc, pour ces cas particuliers, l'hypothèse posée ne semble pas se vérifier. Cependant, même s'il existe une différence significative entre les virées, cela n'implique pas nécessairement que l'estimation de la variance totale d'après cette méthode n'est pas bonne. Car, ce fait peut dépendre de différents facteurs, comme, par exemple, une distribution irrégulière du volume par parcelle.

CHAPITRE V

CONCLUSIONS

Posant l'hypothèse que les erreurs expérimentales sont négligeables comparativement aux erreurs d'échantillonnage, nous sommes en mesure de conclure:

- 1° Que le sondage systématique avec choix aléatoire de la première parcelle s'avère sous certaines conditions plus précis que le simple sondage aléatoire pour l'estimation de la variance.
- 2° Que le sondage systématique avec choix aléatoire de la première parcelle présente des estimations de la variance dont les écarts de leur espérance mathématique satisfont à la loi de χ^2 et que, par conséquent, il est permis de croire que le calcul de l'erreur d'échantillonnage à partir du sondage systématique est valable.

- 3° Que le sondage systématique, en plus d'être plus économique que le sondage aléatoire, présente des estimations de volume satisfaisantes et qui demeurent entre les limites fixées ou permises.
- 4° Que pour le sondage en grappes, des parcelles plus petites sont à conseiller car elles minimisent le coût des opérations et facilitent le travail sur le terrain.
- 5° Que le sondage en grappes s'est avéré une méthode appropriée pour l'estimation de la variance totale d'une population.

Cependant, il est regrettable que le temps et les fonds nous aient manqué pour effectuer un tel travail sur de plus grandes superficies. Comme on a pu le constater, nous avons entrepris ces recherches sur une superficie assez restreinte et, de plus, tous les calculs ont été effectués à peu près dans une seule strate, le *DRY-OX* 70 ans. Alors, il serait très souhaitable que d'autres travaux du genre soient effectués sur des superficies plus considérables et à l'intérieur de types forestiers et de classes d'âge variés afin de pouvoir établir des comparaisons entre différents plans de sondage et, surtout, de pouvoir vérifier si le sondage systématique donne, en moyenne, des estimations exactes de la variance quel que soit le genre de forêt inventorié.

Au cours du présent travail, nous avons effleuré un point qui possède une importance énorme en inventaire forestier: le facteur coût. Le travail sur le terrain n'a comporté aucune donnée se rapportant au "coût" du sondage. Tout ce que nous avons pu faire, fut de suggérer une fonction ou une formule susceptible de servir au calcul de ce facteur; mais nous espérons qu'une étude précise et plus élaborée sera entreprise le plus tôt possible sur ce sujet qui est d'une importance capitale en inventaire forestier.

BIBLIOGRAPHIE

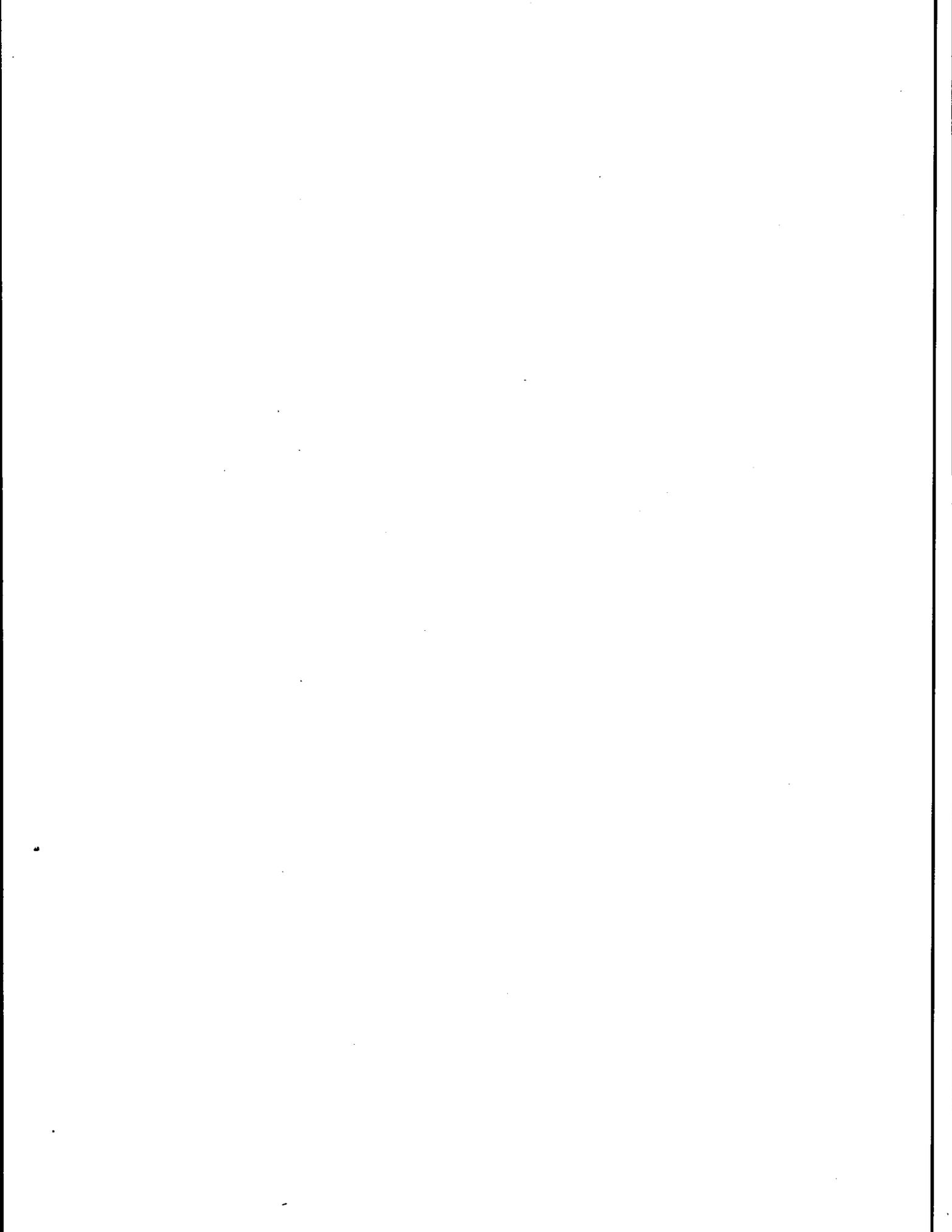
- CHERNG-JIANN Shiue, 1960. *Systematic sampling with multiple random starts*. Forest Science, vol. 6 (1): 42-50.
- COCHRAN, W.G., 1946. *Relative accuracy of systematic and stratified random sampling for a certain class of populations*. Ann. Math. Stat. 17: 164-177.
- FINNEY, D.J., 1948. *Random and systematic sampling in timber surveys*. Forestry 22: 64-99.
- GAUTSHI, W., 1956. *Some remarks on systematic sampling*. Ann. Math. Stat. 28: 385-394.
- HANSEN, M.H. et W.N. HURWITZ, 1942. *Relative efficiencies of various sampling units in population inquiries*. Jour. Amer. Statis. Assoc. 37: 89-94.
- HASEL, A.A., 1937 a. *Analyses of sampling methods for volume determination in ponderosa Pine Forest*. U.S.F.S. California Forest and Range Exp. Station 43 pp.
- HASEL, A.A., 1937 b. *Arrangement of cruise plots to permit a valid estimate of sampling error*. U.S.F.S. California Forest and Range Exp. Stat. 13 pp.
- HASEL, A.A., 1942 b. *Sampling error of cruises in the California Pine region*. Jour. For. 40: 211-217.
- LAFOND, A., 1960. *Notes pour l'identification des types forestiers sur les concessions de la "Quebec North Shore Paper Co."* Manuel de poche 56 pp.

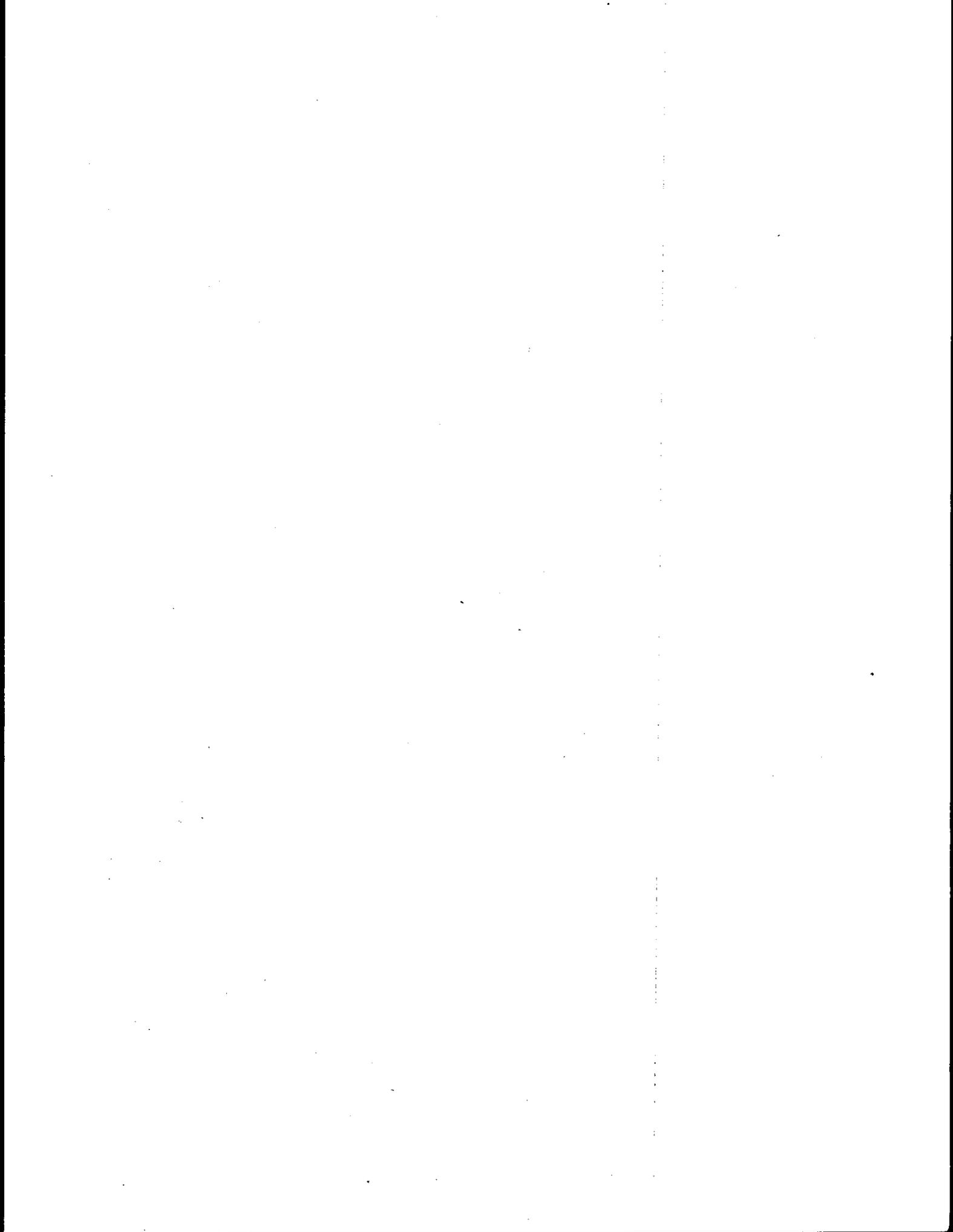
- LANGSAETER, A., 1932. *Accuracy in forest strip survey*. Medd. Norske Skog for. 4: 432-563; 5: 405-445. U.S.F.S. Trans. 354 A, 354 B, 355 (1938).
- MATERN, B., 1949. *Forest inventories sampling techniques*, U.N. Sci. Conf. Conserv. Util. Res. 7 pp.
- MORICE, E., 1957. *Méthode statistique*. Inst. Nationale de la statistique et des études économiques pour la métropole et la France d'outre-mer. Imprimerie nationale. Paris - 2 vol.
- NASLUND, M., 1930. *On computing the standard error in strip survey*. Svensk. Skogsvfören. Tidskr. 28: 309 ss.
- NASLUND, M., 1939. *On computing the standard error in strips and plot surveying*. Medd. Statens. Skogsfören, 31: 301-344.
- PARDE, J., 1957. *Recherches sur l'application aux tailles des méthodes mathématiques - statistiques d'inventaire*. Ann. Ec. Nation. des eaux et forêts et de la Stat. Rech. et Exp., Tome XV, fasc. 2.
- SCHUMAKER, F.X. et H. Bull, 1932. *Determination of the errors of the estimate of a forest survey with special reference to the bottom - land hardwood forest region*. Jour. Agr. Res., 47: 719-734.
- SPURR, S.H., 1952. *Forest inventory*. Ronald Press Co., N.Y. 476 pp.
- SUKHATME, P.V., 1954. *Sampling theory of surveys with applications*. Iowa state college press, Ames. 485 pp.
- YATES, F., 1946. *A review of recent statistical development in sampling and sampling surveys*. Jour. Roy. Stat. Soc., 109: 12-30.
- YATES, F., 1948. *Systematic sampling*, Roy, Soc. of London Philos. Trans., 241 (A-834): 345-377.
- YATES, F., 1949. *Sampling methods for censures and surveys*. Charles Griffith Co., Ltd. London, 318 pp.

APPENDICE

COEFFICIENTS UTILISÉS POUR CONVERTIR LES MESURES ANGLAISES EN MESURES MÉTRIQUES (RENNIE 1967)

Pouces (po)	x	2,54	=	Centimètres (cm)
Pieds (pi)	x	0,305	=	Mètres (m)
Chânes (ch)	x	20,12	=	Mètres (m)
Acres (ac)	x	0,405	=	Hectares (ha)
Milles (mi)	x	1,6	=	Kilomètres (km)
Milles carrés (mi ²)	x	2,59	=	Kilomètres carrés (km ²)
Pieds cubes (pi ³)	x	0,0283	=	Mètres cubes (m ³)
Pieds cubes à l'acre (pi ³ /ac)	x	0,07	=	Mètres cubes à l'hectare (m ³ /ha)
Variance en pi ³	x	0,0008	=	Variance en (m ³)





Le ministère des Terres et Forêts est responsable de l'administration et de la gestion des forêts publiques dans l'intérêt général du Québec. Le Ministère doit donc se préoccuper de connaître les ressources forestières dont il dispose. L'inventaire forestier constitue un élément indispensable à une telle connaissance. Dans les limites de sa juridiction, le Service de la recherche contribue à une meilleure connaissance des ressources forestières. En effet, une partie du budget du Service est consacrée aux recherches ayant pour but de développer ou d'améliorer les méthodes utilisées lors de l'inventaire forestier.



L'ÉDITEUR OFFICIEL DU QUÉBEC
SERVICE DE LA REPROGRAPHIE
Septembre 1975