

# RAPPORT D'EXPERTISE

---

## CONSIDÉRATIONS STATISTIQUES RELATIVEMENT À L'ÉVALUATION DES CRITÈRES DE L'ARTICLE 89 DU RNI

Présenté à :

**TEMBEC INC.**

Carl Vézina, ing.f.  
Jean-Denis Grenier, ing.f.

Par :

**CERFO**

Centre Collégial de Transfert  
de Technologie en foresterie

Donald Blouin, ing.f., M.Sc.

---

**Mars 2006**

## TABLE DES MATIÈRES

---

1. Justification de l'utilisation d'une variable comme élément de stratification, stratification par chantier ou stratification par production prioritaire .....	2
2. Façon de faire un plan de sondage avec des placettes linéaires .....	4
a) Méthode de prise de mesure.....	4
b) Plan de sondage .....	5
3. Justification de l'utilisation des chiffres significatifs.....	6
4. Qualité des informations relativement aux événements rares .....	10
a) Précision concernant les événements rares .....	10
b) Évaluation statistique du seuil minimal de régénération à suivre (statistique).....	10
c) Évaluation sylvicole du seuil minimal de régénération à suivre.....	11
5. Interprétation des résultats - possibilité de tirer des conclusions à partir d'une observation et à partir d'un échantillon tirés d'une population .....	12

<b>RÉFÉRENCES</b> .....	15
-------------------------	----

### LISTE DES TABLEAUX

Tableau 1 : Chiffres significatifs nécessaires à la présentation du volume total à l'hectare.....	7
Tableau 2 : Chiffres significatifs nécessaires à la présentation de l'intervalle de confiance du volume total à l'hectare .....	8
Tableau 3 : Chiffres significatifs nécessaires à la présentation du nombre de tige total à l'hectare .....	9

### LISTE DES FIGURES

Figure 1. Exemple d'établissement du plan de sondage pour un secteur d'intervention .....	6
Figure 2. Pourcentage de précision calculé en fonction du coefficient de distribution du groupe d'essences .....	11

### ANNEXES - LISTE DES FICHIERS JOINTS À L'ENVOI ÉLECTRONIQUE

RNI\_89\_sepm\_4304\_sentier.xls  
RNI\_89\_mixte\_4304\_sentier.xls  
RNI\_89\_sepm\_4304\_régénération.xls  
RNI\_89\_sepm\_4304.xls  
RNI\_89\_mixte\_4304.xls  
epc0104pvm83\_tembec.xls

## **1. JUSTIFICATION DE L'UTILISATION D'UNE VARIABLE COMME ÉLÉMENT DE STRATIFICATION, STRATIFICATION PAR CHANTIER OU STRATIFICATION PAR PRODUCTION PRIORITAIRE**

---

Une des plus importantes raisons justifiant la stratification est la réduction de la variation de l'estimé de la moyenne de la population. Cependant, beaucoup d'inventaires ou de prises de données impliquent l'analyse de plusieurs variables en même temps et une bonne stratification pour une variable n'est pas nécessairement bonne pour une autre variable (Steel et Torrie, 1980, page 561).

La création de strates a pour but de minimiser la variabilité intrastrate et maximiser la variabilité interstrate. La création de strates permet donc de diminuer l'effet de la variabilité de la variable à estimer et conséquemment, de réduire l'erreur d'échantillonnage (Rondeux, 1999, page 412). Pour la stratification, on choisira une ou des caractéristiques liées aussi étroitement que possible à la variable étudiée (Rondeux, 1999, page 414; Cochram, 1977, page 101).

Comme cité plus haut, ce qui justifie de faire de la stratification est l'utilisation de caractéristiques pouvant avoir une influence sur la variabilité de la variable mesurée.

Exemple :

Afin de répondre à la question : Quelle est la quantité de régénération résineuse présente sur un territoire donné?

Il est possible de donner une réponse générale pour l'ensemble du territoire ou encore, dans un souci d'améliorer la description de la répartition de la régénération résineuse sur le territoire, sachant que celle-ci est influencée par le type de peuplement d'origine, il est possible de subdiviser le territoire en trois strates afin de décrire la régénération résineuse présente dans les peuplements d'origine feuillue, dans les peuplements d'origine mixte et dans les peuplements d'origine résineuse. Il y aura à l'intérieur de chacune des strates une variabilité beaucoup moins grande que sur l'ensemble du territoire et des écarts importants entre les trois strates, justifiant ainsi leurs utilisations.

Autre exemple :

Afin de répondre à la question : Quel est le coût de construction des chemins forestiers sur un territoire donné?

Ici aussi, il est possible de donner une réponse générale pour l'ensemble du territoire, par contre, dans un souci d'en faire une description plus juste et être en mesure de faire une meilleure planification des coûts pour les travaux à venir, il est très intéressant de catégoriser le territoire en zones (strates) en fonction de caractéristiques topographiques contraignantes pour la construction de chemins telles que la présence de pentes fortes, la présence de lacs et de rivières, la solidité du sol, la rugosité du sol. Ainsi, les façons de faire et les coûts de construction de chemin seront très différents d'une zone à l'autre ainsi que par rapport à une zone à faible contrainte, d'où l'importance de tenir compte de ces caractéristiques du milieu. Fait important à remarquer dans ce cas-ci, le type de peuplements forestiers n'a que très peu d'influence sur le coût de construction de chemin et n'est donc pas retenu comme élément de stratification.

Concernant l'article 89, il est dit que : « dans le cas où la superficie occupée par les sentiers d'abattage et de débardage est supérieure à 25 % sans toutefois dépasser 33 %, le titulaire du permis d'intervention doit protéger entre les sentiers de débardage, la régénération préétablie en essences recherchées comme production prioritaire, de manière à ce que le coefficient de distribution des tiges non marchandes ayant une hauteur de 5 cm et plus, après coupe, soit supérieur à 80 % du coefficient de distribution de ces tiges avant coupe ».

Il faut donc se poser la question pour l'évaluation de la protection de la régénération telle que visée par l'article 89. Qu'est-ce qui influence (est lié à) la protection de la régénération pouvant justifier une stratification? Est-ce le type de régénération lui-même, le type de peuplement sur pied, le type de coupe effectué (consignes de prélèvement), le type de machinerie utilisé, la présence de contraintes au déplacement de la machinerie (pente forte, rugosité, solidité du terrain, ruisseau)?

Lorsque l'on vérifie l'effet des interventions sur la protection de la régénération, le type de régénération présente et le type de peuplement sur pied n'ont pas d'influence sur la protection de cette régénération alors que le type de coupe effectué, le type de machinerie utilisé et la présence de contraintes au déplacement de la machinerie, ont une grande influence sur le succès des opérations.

Comme le type de régénération et le type de peuplements ne sont pas liés à la protection de la régénération, il n'y a aucune justification de faire une compilation de la protection de la régénération en tenant compte de ces éléments ou des éléments qui en sont déduits tels que les productions prioritaires. Puisqu'il n'y a pas de lien de cause à effet entre les productions prioritaires et la protection de la régénération, il n'y a aucune justification de créer des unités d'échantillonnage stratifiées (secteur d'intervention) en fonction de cette caractéristique.

L'utilisation de caractéristiques n'ayant pas de lien avec la variable mesurée ne fait que diminuer la grandeur de la strate, diminuer le nombre de placettes dans la strate et diminuer la précision des données, et ce, sans raison justifiée.

Pour évaluer le pourcentage d'occupation des sentiers et la protection de la régénération tel que stipulé dans l'article 89, la strate définie comme un secteur d'intervention doit donc correspondre à des aires forestières ayant fait l'objet d'un même traitement sylvicole au cours d'une année, réalisé par un type de machinerie dans des conditions de terrain semblables.

Au niveau de l'évaluation de pourcentage d'occupation des sentiers, la CMO peut être considérée comme étant le même traitement sylvicole qu'une CPRS, la CMO n'étant qu'une variante au niveau de la répartition dans l'espace des assiettes de coupe, les opérations de récolte sur le terrain étant les mêmes. Un secteur d'intervention correspond donc à l'ensemble des assiettes de coupe d'un chantier ayant subi un traitement sylvicole de CPRS ou de CMO de n'importe quelle production prioritaire.

Au niveau de l'évaluation de la protection de la régénération, le secteur d'intervention correspond à l'ensemble des assiettes de coupe d'un chantier ayant subi un traitement sylvicole de CPRS ou

de CMO de n'importe quelle production prioritaire devant faire l'objet d'une protection de la régénération.

La production prioritaire sert d'abord à déterminer s'il doit y avoir estimation du taux d'occupation des sentiers de débardage et du taux de protection de la régénération. La production prioritaire sert ensuite à identifier les essences devant faire l'objet de vérification alors que le type d'essence détermine la grandeur de la placette à utiliser. C'est la présence de la régénération totale qui fait l'objet d'une vérification de la protection de la régénération au niveau du secteur d'intervention telle que définie ci-haut.

## **2. FAÇON DE FAIRE UN PLAN DE SONDAGE AVEC DES PLACETTES LINÉAIRES**

---

### **a) Méthode de prise de mesure**

L'article 89 propose déjà de faire des sections de 60 m à partir d'une virée pour la lecture du pourcentage d'occupation des sentiers par secteur d'intervention. Afin de simplifier et de faciliter les opérations de prise de mesure sur le terrain, le principe des sections de 60 m peut être appliqué à des placettes linéaires (sections de longueur indéterminée, traversant le bloc) en respectant les principes d'échantillonnage et d'établissement d'un plan de sondage.

Les méthodes d'échantillonnage linéaire sont déjà reconnues. Van Wagner (1982) a établi différents protocoles de mesure pour l'évaluation de la quantité des résidus forestiers au sol par placette linéaire et cette méthode peut également être adaptée à l'évaluation du pourcentage d'occupation des sentiers de débardage.

Dans le compilateur du MRNF (Peng-Chea Im, 2004), la compilation des données est déjà réalisée par tronçon de « un sentier, une bande protégée » à partir des données obtenues sur un transect de 60 m. Le compilateur considère tous les tronçons, complets et incomplets, présents sur le transect de 60 m. Cette façon de faire engendre un grand nombre de tronçons incomplets faisant ainsi fortement augmenter la variabilité des données. La méthode proposée des placettes linéaires consiste à compiler seulement des tronçons complets de « un sentier, une bande protégée », ce qui aura pour effet d'être plus précis et nécessitera moins de points d'observation.

À titre d'exemple, dans le fichier « RNI\_89\_sepm\_4304\_sentier.xls », lorsque compilé selon la méthode des tronçons incomplets par transect de 60 m (Peng-Chea Im, 2004), l'échantillon constitué de 81 points d'observation « Onglet : Calcul\_Peng-Chea » permet d'obtenir une moyenne de 17,4 % de la superficie occupée par les sentiers avec une précision de 78,5 % (niveau de probabilité bilatérale de 95 %). Une compilation de la même base de données ne considérant que les tronçons complets « Onglet : Calcul\_section\_S\_B » a permis d'obtenir une moyenne de 17,7 % et une précision de 86 % (niveau de probabilité bilatérale de 95 %) avec seulement 34 points d'observation (résultats présentés à l'annexe 1).

Dans le cas du fichier « RNI\_89\_mixte\_4304\_sentier.xls », l'utilisation des tronçons incomplets permet d'obtenir une moyenne=20,0 %, avec n=146 pour une précision=88,0 % alors que

l'utilisation de seulement les tronçons complets permet d'obtenir une moyenne=22,8 % et une précision=88,6 % avec seulement n=69 points observation (résultats présentés à l'annexe 2).

## **b) Plan de sondage**

Les principes statistiques à mettre en application pour l'échantillonnage systématique par placette linéaire tels qu'illustrés à la figure 1 sont :

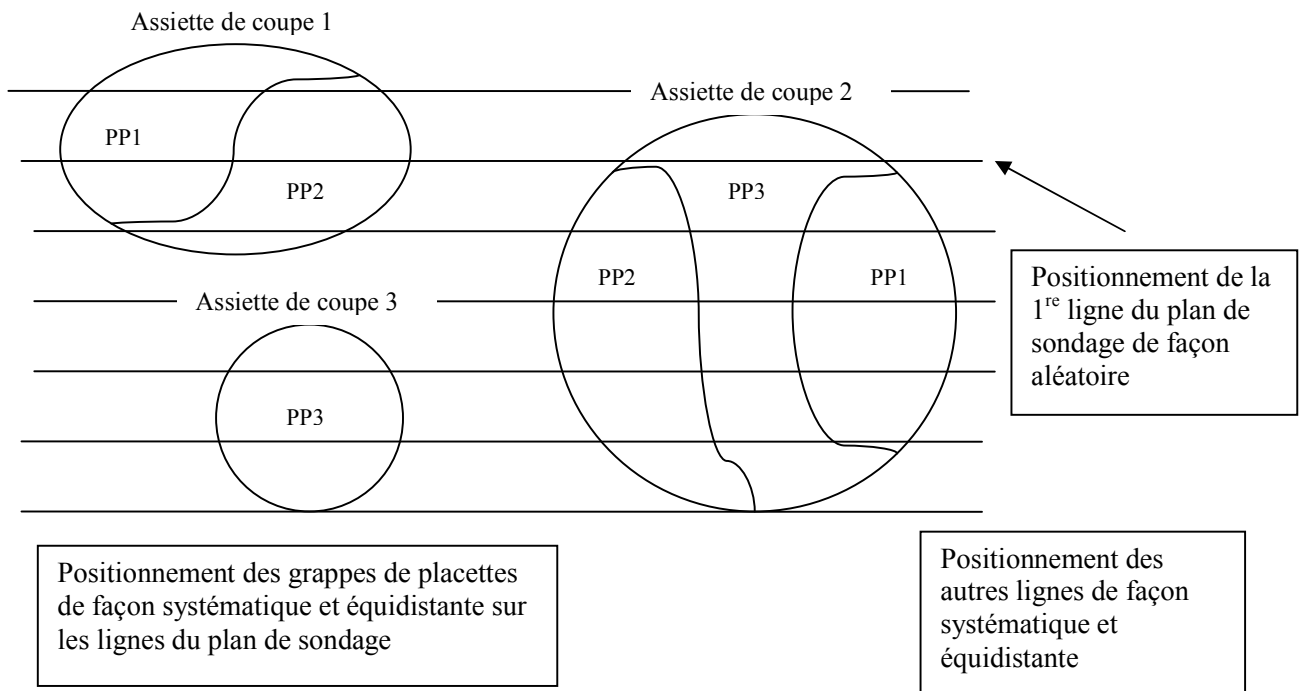
- 1- S'assurer de couvrir l'ensemble de la superficie en considérant chacune des assiettes de coupe comme faisant partie du secteur d'intervention;
- 2- Le taux d'échantillonnage est constant pour l'ensemble de la superficie considérant chacune des assiettes de coupe comme faisant partie du secteur d'intervention;
- 3- La localisation de la première placette linéaire est fixée de manière aléatoire;
- 4- Les autres placettes linéaires sont toutes équidistantes;
- 5- Les placettes linéaires doivent être perpendiculaires à l'ensemble des sentiers de débardage
- 6- Les grappes de régénération sont localisées de la même manière (équidistantes) le long des placettes linéaires.

En sachant que :

- 1- les petites assiettes de coupe ont une plus forte proportion de sentier de débardage que les grandes assiettes de coupe, et que
- 2- les bords de chemin ont une plus forte proportion de sentier de débardage que les fonds des assiettes de coupe,

la mise en application des principes d'échantillonnage permet de s'assurer de couvrir de façon objective l'ensemble des conditions retrouvées dans ce qui est défini comme un secteur d'intervention. Il sera donc normal de retrouver une grande variabilité entre chacun des points d'observation individuels. C'est toujours la moyenne de l'échantillon qui sert de référence à l'analyse des résultats.

**Figure 1. Exemple d'établissement du plan de sondage pour un secteur d'intervention**



Ainsi, le plan de sondage couvrira le secteur d'intervention qui doit contenir l'ensemble des assiettes de coupe (peu importe leur dimension et leur production prioritaire) ayant fait l'objet d'un même traitement sylvicole (CPRS ou CMO) au cours d'une année donnée avec un type de machinerie. La production prioritaire servira simplement à définir si un suivi du taux d'occupation des sentiers et de la régénération doit être effectué dans cette assiette de coupe.

Le plan de sondage permet de visualiser la répartition géographique des placettes linéaires d'échantillonnage. Le taux d'échantillonnage variera en fonction de la précision recherchée, de la variabilité de la variable mesurée au niveau du secteur d'intervention et de la méthode de calcul retenue. Le calcul du nombre de placette peut être effectué sur les bases de données déjà en possession et servir de guide afin de fixer le taux d'échantillonnage des prochaines prises de mesure (exemples de calcul dans les fichiers joints et les annexes 1 et 2).

### **3. JUSTIFICATION DE L'UTILISATION DES CHIFFRES SIGNIFICATIFS**

En utilisant comme référence l'estimation du volume d'un arbre dans le tarif de cubage de Perron (1985), on obtient le volume d'une tige en  $\text{dm}^3$  exprimé avec 3 à 5 chiffres significatifs.

Il est possible de faire une première vérification en considérant la précision de l'estimation de la hauteur d'un arbre. En supposant une précision de 0,5 m sur l'estimation de la hauteur, cela a pour effet de faire varier le volume d'un arbre de 5 à 30  $\text{dm}^3$  (Perron, 1985). Le même constat peut être fait entre 2 classes de diamètre de 2 cm. Ainsi, pour un sapin baumier d'un dhp de 22 cm d'une hauteur de 17 m, dire que cet arbre a un volume de 269  $\text{dm}^3$  est bien suffisamment

précis (et non 268,6 dm<sup>3</sup>) sachant qu'un arbre semblable mais de 24 cm de dhp a un volume de 315 dm<sup>3</sup> et un autre de 18 m de hauteur et de 22 cm de dhp a un volume de 286 dm<sup>3</sup>. Il semble suffisamment conservateur d'affirmer que l'utilisation de 3 chiffres significatifs est une limite raisonnable pour l'estimation du volume d'un arbre sur pied connaissant la précision (variabilité de 5 à 30 dm<sup>3</sup>) entourant cette évaluation.

Il est également possible de vérifier l'impact du nombre de chiffres significatifs utilisé dans le tarif de cubage de Perron (1985) pour l'estimation du volume d'un arbre en le transposant en volume à l'hectare (tableau 1). On peut ainsi vérifier que l'utilisation de plus de 3 chiffres significatifs ne change en rien le résultat obtenu concernant le volume total à l'hectare.

**Tableau 1 : Chiffres significatifs nécessaires à la présentation du volume total à l'hectare**

1	2	3	4	5	Nombre de chiffres significatifs
200	190	194	194,4	194,44	Volume d'un arbre en décimètre cube
50	50	50	50	50	Nombre d'arbre dans la placette
10 000	9 500	9 700	9 720	9 722	Volume dans la placette en décimètre cube
25	25	25	25	25	Nombre de placette dans un hectare
250 000	237 500	242 500	243 000	243 050	Volume dans un hectare en décimètre cube
250,00	237,50	242,50	243,00	243,05	Volume dans un hectare en mètre cube
250	238	243	243	243	Volume dans un hectare en mètre cube arrondi à 3 chiffres significatifs



On peut faire une vérification semblable au niveau de l'intervalle de confiance pour l'estimation du volume à l'hectare d'un peuplement forestier (tableau 2). On peut constater que pour un petit intervalle de confiance (erreur relative de 7 %), l'utilisation de 2 chiffres significatifs est suffisante alors que pour un intervalle de confiance plus grand (erreur relative de 14 %), 3 chiffres significatifs sont nécessaires, confirmant les déductions faites plus haut.



**Tableau 2 : Chiffres significatifs nécessaires à la présentation de l'intervalle de confiance du volume total à l'hectare**

250,00	237,50	242,50	243,00	243,05	Volume dans un hectare en mètre cube
0,07	0,07	0,07	0,07	0,07	Demi-largeur de l'intervalle de confiance en pourcentage
17,500	16,625	16,975	17,010	17,014	Demi-largeur de l'intervalle de confiance en mètre cube
1	2	3	4	5	Nombre de chiffres significatifs
20	17	17,0	17,01	17,014	Demi-largeur de l'intervalle de confiance en mètre cube en fonction du nombre de chiffres significatifs
20	17	17	17	17	Demi-largeur de l'intervalle de confiance en mètre cube arrondi à 2 chiffres significatifs
↑					
250,00	237,50	242,50	243,00	243,05	Volume dans un hectare en mètre cube
0,14	0,14	0,14	0,14	0,14	Demi-largeur de l'intervalle de confiance en pourcentage
35,000	33,250	33,950	34,020	34,027	Demi-largeur de l'intervalle de confiance en mètre cube
1	2	3	4	5	Nombre de chiffres significatifs
40	33	34,0	34,02	34,027	Demi-largeur de l'intervalle de confiance en mètre cube en fonction du nombre de chiffres significatifs
40,0	33,3	34,0	34,0	34,0	Demi-largeur de l'intervalle de confiance en mètre cube arrondi à 3 chiffres significatifs
↑					

L'utilisation de 3 chiffres significatifs semble donc être une bonne règle pouvant nous servir de référence pour nos données provenant d'inventaire forestier.

Il faut donc ajuster les unités utilisées au nombre de chiffres significatifs qu'il est correct d'utiliser dans la présentation des résultats. Ainsi, à titre d'exemple, le volume total d'un territoire estimé à  $91\,628\,861\text{ m}^3 \pm 3\,207\,010\text{ m}^3$  devrait se lire 91,6 millions de  $\text{m}^3 \pm 3,2$  millions de  $\text{m}^3$ .

Pour l'évaluation du nombre de tiges en régénération, une règle semblable peut s'appliquer. En fonction de la grandeur de la placette et du nombre de tiges qu'il est possible d'y retrouver (tableau 3), l'usage de 3 chiffres significatifs ne change pas les résultats et est suffisant pour la présentation de ceux-ci.

**Tableau 3 : Chiffres significatifs nécessaires à la présentation du nombre de tige total à l'hectare**

Placette de 4 m <sup>2</sup>									
10	9	8	7	6	5	4	3	2	Nombre de tige dans la placette
2500	2500	2500	2500	2500	2500	2500	2500	2500	Nombre de placette à l'hectare
25000	22500	20000	17500	15000	12500	10000	7500	5000	Nombre de tige à l'hectare
250	225	200	175	150	125	100	75	50	Nombre de tige arrondi à 3 chiffres significatifs
Placette de 9 m <sup>2</sup>									
20	19	18	17	16	15	14	13	12	Nombre de tige dans la placette
1111	1111	1111	1111	1111	1111	1111	1111	1111	Nombre de placette à l'hectare
22220	21109	19998	18887	17776	16665	15554	14443	13332	Nombre de tige à l'hectare
222	211	200	189	178	167	156	144	133	Nombre de tige arrondi à 3 chiffres significatifs
Placette de 25 m <sup>2</sup>									
30	29	28	27	26	25	24	23	22	Nombre de tige dans la placette
400	400	400	400	400	400	400	400	400	Nombre de placette à l'hectare
12000	11600	11200	10800	10400	10000	9600	9200	8800	Nombre de tige à l'hectare
120	116	112	108	104	100	96	92	88	Nombre de tige arrondi à 3 chiffres significatifs
Placette de 100 m <sup>2</sup>									
25	24	23	22	21	20	19	18	17	Nombre de tige dans la placette
100	100	100	100	100	100	100	100	100	Nombre de placette à l'hectare
2500	2400	2300	2200	2100	2000	1900	1800	1700	Nombre de tige à l'hectare
250	240	230	220	210	200	190	180	170	Nombre de tige arrondi à 3 chiffres significatifs

Pour l'évaluation du coefficient de distribution d'une essence, en considérant un échantillon constitué de 50 grappes de 10 micro-placettes, soit 500 microplacettes, la présence d'une tige dans 261 microplacettes donne un coefficient de distribution de 52,2 %, alors que la présence d'une tige dans 262 microplacettes donne un coefficient de distribution de 52,4 %. Ici aussi la règle confirme qu'on ne peut pas être plus précis que notre unité de mesure (Steel et Torrie, 1980, page 37) qui, dans ce cas, est la présence d'une tige dans une placette et que l'utilisation de 3 chiffres significatifs est le nombre maximal à utiliser pour la présentation des résultats, soit 1 chiffre après le point au niveau des pourcentages du coefficient de distribution.

Pour un échantillon constitué de 100 microplacettes (10 grappes), seulement 2 chiffres significatifs devraient être utilisés.

Sur les sorties du compilateur du MRNF (Peng-Chea Im, 2004), fichier « RNI\_89\_sepm\_4304.xls », on devrait lire que le pourcentage de la superficie occupée par les sentiers est de 17,4 % ± 3,1 % avec une erreur relative de 17,8 % et non 17,44 % ± 3,10 % avec une erreur relative de 17,79 %, ce qui donne une fausse illusion de grande précision.

Au niveau de la protection de la régénération des tiges résineuses de 5 cm et plus, on devrait lire :

Tiges intactes : 34,6 %

Tiges affectées : 13,3 %

Tiges intactes et affectées : 47,9 %  $\pm$  11,9 %

Erreur relative : 24,8 %

#### **4. QUALITÉ DES INFORMATIONS RELATIVEMENT AUX ÉVÉNEMENTS RARES**

---

##### **a) Précision concernant les événements rares**

Selon Cochram (1977), pour une proportion entre 30 et 70 %, une estimation précise de la proportion n'est pas problématique. Ce sont les événements plus rares auxquels il faut porter une attention.

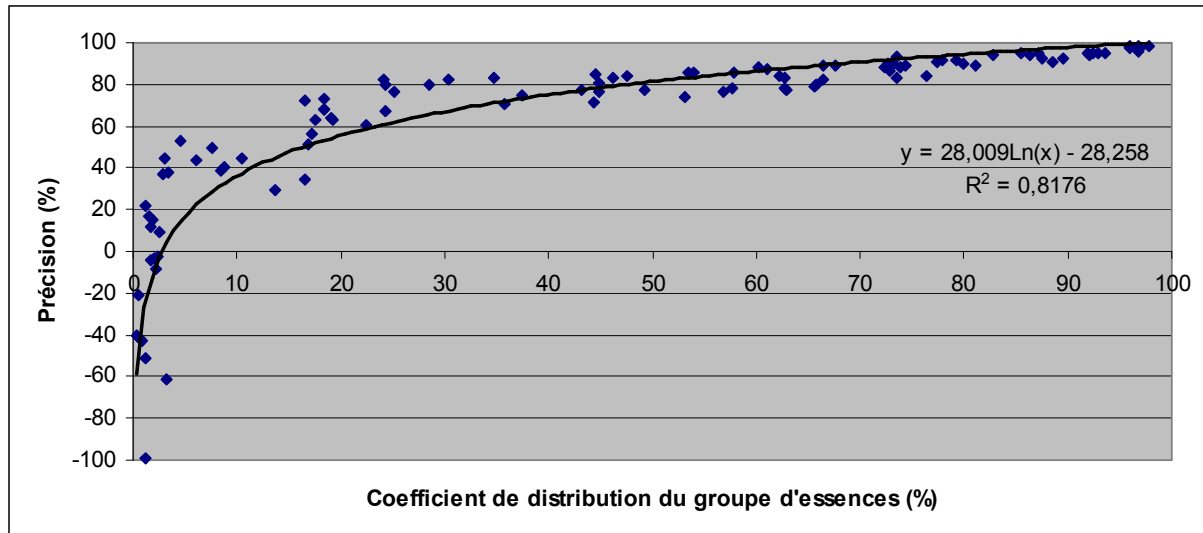
Il est reconnu que les événements rares sont toujours évalués avec une faible précision et que dans ce cas, un très grand nombre de points d'observation sont nécessaires (pouvant aller jusqu'à 10 fois plus) pour atteindre une précision donnée (Cochram, 1977, page 76).

Dans son travail sur la validation statistique de l'exigence du contrôle de pourcentage de prélèvement par essence du jardinage, Blouin (2002) faisait la démonstration, à partir d'une importante base de données de l'Outaouais, qu'il est impossible d'obtenir une précision moyenne supérieure à 80 % lorsque le taux d'occupation d'une essence ou d'un groupe d'essences est inférieur à 30 %.

##### **b) Évaluation statistique du seuil minimal de régénération à suivre (statistique)**

À partir d'une série de relevés d'inventaire d'éclaircie précommerciale de Tembec en Mauricie « fichier : epc0104pvm83\_tembec.xls », le coefficient de distribution par essence et par groupe d'essences a été calculé ainsi que les statistiques associées. Il est possible de remarquer, à la figure 2, une augmentation de la précision avec l'augmentation du pourcentage du coefficient de distribution des espèces. En fonction de l'équation de régression présentée, il est possible de constater qu'une précision de 90 % est atteignable en moyenne pour des essences ou groupe d'essences occupant plus de 68 % alors qu'une précision de 80 % et plus est atteignable seulement pour les essences ou groupe d'essences occupant plus de 48 %. En considérant un minimum acceptable de 70 % de précision, l'essence ou le groupe d'essences doit occuper plus de 33 % de l'espace. Ainsi, cette démonstration vient confirmer les appréhensions de Cochram (1977) à savoir qu'il faut porter une attention particulière aux événements rares. En fait, ceci vient confirmer que selon les méthodes d'échantillonnage présentement utilisées, il est impossible d'obtenir une précision satisfaisante pour les essences ou groupe d'essences de faible importance (coefficient de distribution < 33 %) et que de tels résultats ne peuvent pas être considérés comme fiables pour la prise de décision.

**Figure 2. Pourcentage de précision calculé en fonction du coefficient de distribution du groupe d'essences**



Il est possible de confirmer ces observations au moyen des résultats du fichier « RNI\_89\_sep\_m\_4304\_régénération.xls »; le coefficient de distribution des tiges résineuses intactes et affectées de 5 cm et plus est de 47,9 % avec une précision de 75,2 %, celui de dhs de 2 cm et plus, 16,7 % avec une précision de 62,1 %, et celui de dhs de 6 cm et plus, 3,3 % avec une précision de 28,4 %. Ces trois points sont conformes au modèle établi à la figure 2. Ceci permet de confirmer qu'en moyenne un coefficient de distribution inférieur à 33 % ne peut pas être évalué avec une précision supérieure à 70 %, et donc, ces résultats ne devraient jamais être utilisés dans la prise de décision d'évaluation de la qualité d'une intervention. Les coefficients de distribution situés entre 33 et 48 % sont évalués avec une précision entre 70 et 80 % laissant eux aussi place à une marge d'erreur très grande (entre  $\pm 20$  % et  $\pm 30$  %).

Les résultats de la figure 2 servent à inféoder les résultats de coefficient de distribution de la régénération seulement et ne peuvent pas être appliqués à l'interprétation de résultats concernant le pourcentage d'occupation des sentiers.

### c) Évaluation sylvicole du seuil minimal de régénération à suivre

Pominville et Doucet (1993) font des recommandations quant aux coefficients de distribution de la régénération nécessaire au maintien de la production résineuse en pin gris, épinette noire et sapin baumier. À l'aide de placette de 4 m<sup>2</sup>, ils donnent comme exemple que le coefficient de distribution de la régénération résineuse nécessaire à une production équivalente à 750 tiges à l'hectare à maturité est de 32 % et 37 % pour le pin gris et l'épinette noire respectivement. Pour produire 1000 tiges à l'hectare à maturité, les coefficients de distribution de la régénération nécessaires sont de 42 % et 44 % respectivement, de 52 % pour la production de 1250 tiges à l'hectare et de 60 % pour la production de 1500 tiges à l'hectare.

En terme de volume marchand, Pominville et Doucet (1993) donnent comme exemple qu'un peuplement de pin gris de 124 m<sup>3</sup>/ha a un coefficient de distribution des tiges marchandes de 30 % (basé sur des placettes de 4 m<sup>2</sup>) et qu'un coefficient de distribution de la régénération de 37 % est nécessaire. Pour l'épinette noire, un peuplement de 94 m<sup>3</sup>/ha avec un coefficient de distribution des tiges marchandes de 30 % a besoin d'un coefficient de distribution de la régénération de 38 % pour assurer la reconstitution d'un peuplement équivalent.

À l'analyse des résultats de Pominville et Doucet (1993), il est possible de déduire pour le coefficient de distribution de la régénération qu'un seuil situé entre 30 % et 40 % est le niveau minimum à observer pour assurer le maintien d'une production résineuse équivalente.

## **5. INTERPRÉTATION DES RÉSULTATS - POSSIBILITÉ DE TIRER DES CONCLUSIONS À PARTIR D'UNE OBSERVATION ET D'UN ÉCHANTILLON TIRÉS D'UNE POPULATION**

---

La population consiste en toutes les observations possibles d'une variable. La population représente l'ensemble des individus, des éléments ou unités de même nature auxquels s'adresse l'échantillonnage (Rondeux, 1999). Dans le cas d'un peuplement forestier, la population est représentée par l'ensemble des arbres, alors que dans le cas d'un inventaire forestier, il peut être plus aisé de considérer la forêt comme une population de placettes.

L'échantillonnage (ou communément appelé l'inventaire forestier) vise à estimer la valeur d'une (ou plusieurs) caractéristique(s) d'une population à partir d'un échantillon (Rondeux, 1999). L'échantillon réfère donc à une partie de la population (un sous-ensemble) et ce sont les informations obtenues à partir de l'échantillon qui sont utilisées pour estimer les caractéristiques de la population (Blouin *et al.*, 2005). C'est le respect de la mise en application des principes de la théorie de l'échantillonnage qui permet de s'assurer que les estimations obtenues au moyen d'un échantillon sont représentatives de la population (tel que présenté au point 2).

L'échantillon est constitué d'un certain nombre d'observations. L'observation est le relevé d'une variable. En foresterie, l'observation peut prendre différentes formes : un arbre, une placette, une micro-placette, un point de prisme, etc. Les statistiques calculées tels que la moyenne et l'écart-type servent à décrire l'échantillon alors que les statistiques tels que l'erreur standard et l'intervalle de confiance servent à inférer les résultats de l'échantillon à la population. Le calcul de l'intervalle de confiance permet d'évaluer la précision relative de l'inférence effectuée à la population à partir de l'échantillon.

En aucun temps l'analyse des données au niveau des points d'observation individuels (placettes individuelles) est utilisée dans l'interprétation des résultats. Ce sont les statistiques descriptives (moyenne et écart-type) qui servent à qualifier l'échantillon (représentant l'ensemble des points d'observation) alors que ce sont les statistiques inférentielles (erreur standard et intervalle de confiance) qui permettent la généralisation des résultats à toute la population.

En se référant à la figure 1, il n'est pas permis de porter un jugement seulement sur une portion du secteur d'intervention, la production prioritaire 2 dans l'assiette de coupe 2 par exemple. L'interprétation des résultats portera sur la moyenne de l'échantillon, soit l'ensemble des placettes effectuées dans le secteur d'intervention. Il est tout à fait normal d'observer une grande variabilité sur le terrain, c'est ce que l'on peut comprendre à partir de l'interprétation des

résultats. Ainsi, à partir des résultats du fichier « RNI\_89\_sepm\_4304\_sentier.xls », le pourcentage d'occupation des sentiers de 17,4 % a un intervalle de confiance de  $\pm 3,1$  %, ce qui signifie qu'il y a 95 % de probabilité que l'intervalle de confiance obtenu, soit entre 14,3 % et 20,5 % d'occupation de sentiers, contienne la vraie moyenne du secteur d'intervention (la population). En observant l'écart-type qui est égal à 17,0 %, sachant que 95 % des valeurs des tronçons sont situées à plus ou moins 2 écarts-types autour de la moyenne, il y a donc approximativement 95 % des valeurs de tronçon qui ont une valeur entre 0,0 % et 51,4 % d'occupation de sentiers. Il n'est donc pas difficile d'imaginer que certaines portions du secteur d'intervention sont occupées par une forte proportion de sentiers étant donné la très grande variabilité des données.

Le même constat peut être fait en observant les résultats sur la régénération résineuse, fichier « RNI\_89\_sepm\_4304\_régénération.xls ». Le coefficient de distribution de la régénération résineuse totale de 5 cm et plus de 47,9 % a un intervalle de confiance de  $\pm 11,9$  %, ce qui signifie qu'il y a 95 % de probabilité que l'intervalle de confiance obtenu, soit entre 36,0 % et 59,8 %, contienne la vraie moyenne de la population. En observant l'écart-type qui est égal à 28,1 %, sachant que 95 % des valeurs des grappes sont situées à plus ou moins 2 écarts-types autour de la moyenne, il y a donc approximativement 95 % des valeurs de grappes qui ont une valeur entre -8,3 % et 104,1 % de coefficient de distribution. Ceci peut sembler absurde mais permet de mieux comprendre ce que représente une faible précision des résultats (75,2 %). Dans ce cas, au niveau des données des grappes individuelles sur le terrain, les statistiques nous démontrent que toutes les réponses sont possibles.

Au niveau de l'observation des seuils de protection de la régénération des tiges résineuses et feuillues, basée sur les résultats du compilateur du MRNF (Peng-Chea Im, 2004) « fichier : RNI\_89\_sepm\_4304.xls », il est interprété que la protection de la régénération des tiges de classe 1 et 2 est insuffisante. En effectuant un test de comparaison de moyenne (test de t), il est possible de constater dans le fichier « RNI\_89\_sepm\_4304\_régénération.xls » qu'il n'existe aucune différence significative entre le seuil à respecter (protection de 80 %, 55 % et 35 %) et la moyenne obtenue de tiges intactes (résultats présentés à l'annexe 3). Les données sont beaucoup trop imprécises pour arriver à déceler quelque différence significative que ce soit. On peut donc se questionner sur l'interprétation faite des résultats présentés dans les compilateurs du MRNF (article 89) dans le cas de données de faible précision (< 80 %). Les résultats ayant une faible précision ne peuvent être utilisés qu'à titre indicatif seulement et ne doivent pas servir à la prise de décision d'acceptation ou de refus du critère évalué.

Enfin, si les principes énoncés dans le présent rapport sont respectés, il n'y a pas de problème à fournir le compilateur complet au MRNF. Si l'on désire tout de même présenter qu'un bilan des résultats, on doit présenter au minimum : le nombre de points d'observation contenu dans l'échantillon, la moyenne, l'intervalle de confiance, la précision et l'interprétation du résultat pour chacun des critères faisant l'objet d'une évaluation. Dans le cas des résultats n'étant pas suffisamment précis (< 80 %), tels que démontrés dans les sections précédentes, indiquer qu'aucune interprétation des résultats n'est possible.

---

Donald Blouin, ing.f., M.Sc.

## RÉFÉRENCES

---

- Blouin, D., S. Côté, F. Guillemette, H. Lapierre, C. Roy, A. Tremblay et P. Marier, 2005. Statistique appliquée à la foresterie. Notes du cours de formation en ligne développées par le Consortium de Recherche FORAC et le Centre collégial de transfert de technologie CERFO. 91 pages
- Blouin, D. et G. Lessard, 2002. Validation statistique de l'exigence du contrôle du pourcentage de prélèvement par essence du jardinage. Avis scientifique du CERFO, Rapport : CERFO 2002-10, 11 pages.
- Cochran, W.G., 1977. Sampling Techniques (third edition). John Wiley & Sons inc. US, ISBN 0-471-16240-X, 428 p.
- Im, P.-C., 2004. Programme de compilation des données pour le suivi de l'article 89 du RNI. Direction de l'assistance technique. Ministère des Ressources naturelles et de la Faune.
- Perron, J.-Y., 1985. Tarif de cubage, volume marchand brut. Service de l'inventaire forestier, Ministère de l'Énergie et des Ressources, Gouvernement du Québec. 55 pages
- Pominville, P. et R. Doucet, 1993. Coefficients de distribution de la régénération nécessaire au maintien de la production des peuplements de pin gris, d'épinette noire et de sapin baumier. Direction de la recherche, MRN. Note de recherche forestière n° 48, 15 pages.
- Rondeux, J., 1993. La mesure des arbres et des peuplements forestiers. Les Presses agronomiques de Gembloux. ISBN 2-87016-041-0, 521 p.
- Steel, R.G.D., et J.H. Torrie, 1980. Principles and procedures of statistics – A Biometrical Approach (Second edition). McGraw-Hill inc. US, ISBN 0-07-060926-8, 633 p.
- Van Wagner, C.E. 1982. Aspects pratiques de la méthode d'échantillonnage linéaire. Institut forestier national de Petawawa, Service canadien des forêts, Chalk River, Ontario. Rapport d'information PI-X-12(F)

## ANNEXE 1

---

Fichier : RNI\_89\_sepm\_4304\_sentier.xls

Onglet : Calcul\_Peng-Chea

Compilation Peng Chea, Section de 60 m incluant les tronçons incomplets

	Sentier	Total	
Somme	251,2	1440	
Nb placettes	81	81	
Moyenne			0,1744
Variance de l'échantillon			0,0288
Écart-type			0,1697
Variance de la moyenne			0,0003557
Erreur standard			0,0189
Coefficient de variation			97,3
Limite inférieure de l'intervalle de confiance			0,1369
Limite supérieure de l'intervalle de confiance			0,2120
Intervalle de confiance			0,0375
n (pour précision de 90 %)			375
n (pour précision de 85 %)			167
n (pour précision de 80 %)			94
Erreur relative			21,5
Précision			78,5

Fichier : RNI\_89\_sepm\_4304\_sentier.xls

Onglet : Calcul\_section\_S\_B

Compilation considérant que les tronçons complets

	Sentier	Total	
Somme	134,9	760,8	
Nb placettes	34	34	
Moyenne			0,1773
Variance de l'échantillon			0,0050
Écart-type			0,0710
Variance de la moyenne			0,0001
Erreur standard			0,0122
Coefficient de variation			40,0
Limite inférieure de l'intervalle de confiance			0,1526
Limite supérieure de l'intervalle de confiance			0,2021
Intervalle de confiance			0,0248
n (pour précision de 90 %)			66
n (pour précision de 85 %)			29
n (pour précision de 80 %)			17
Erreur relative			14,0
Précision			86,0



## ANNEXE 2

---

Fichier : RNI\_89\_mixte\_4304\_sentier.xls

Onglet : Calcul\_Peng-Chea

Compilation Peng Chea, Section de 60 m incluant les tronçons incomplets

	Sentier	Total	
Somme	466,7	2339	
Nb placettes	146	146	
Moyenne			0,1995
Variance de l'échantillon			0,0216
Écart-type			0,1470
Variance de la moyenne			0,0001
Erreur standard			0,0122
Coefficient de variation			73,7
Limite inférieure de l'intervalle de confiance			0,1755
Limite supérieure de l'intervalle de confiance			0,2236
Intervalle de confiance			0,0240
n (pour précision de 90 %)			212
n (pour précision de 85 %)			94
n (pour précision de 80 %)			53
Erreur relative			12,0
Précision			88,0

Fichier : RNI\_89\_mixte\_4304\_sentier.xls

Onglet : Calcul\_section\_S\_B

Compilation considérant que les tronçons complets

	Sentier	Total	
Somme	284,2	1245,8	
Nb placettes	69	69	
Moyenne			0,2281
Variance de l'échantillon			0,0118
Écart-type			0,1086
Variance de la moyenne			0,0002
Erreur standard			0,0131
Coefficient de variation			47,6
Limite inférieure de l'intervalle de confiance			0,2020
Limite supérieure de l'intervalle de confiance			0,2542
Intervalle de confiance			0,0261
n (pour précision de 90 %)			90
n (pour précision de 85 %)			40
n (pour précision de 80 %)			23
Erreur relative			11,4
Précision			88,6

## ANNEXE 3

Fichier : RNI\_89\_sepm\_4304\_régénération.xls

Onglet : Calcul et test de t

Comparaison des moyennes des coefficients de distribution de la régénération

	<b>Intact 5cm</b>	<b>Intact d&gt;2cm</b>	<b>Intact d&gt;6cm</b>
Nb placettes	24	24	24
Moyenne	3,458	0,667	0,125
Variance de l'échantillon	6,433	0,754	0,114
	<b>80%</b>	<b>55%</b>	<b>35%</b>
	<b>5cm</b>	<b>d&gt;2cm</b>	<b>d&gt;6cm</b>
Nb placettes	24	24	24
Moyenne	3,833333333	0,916666667	0,116666667
Variance de l'échantillon	5,06318841	0,67514493	0,03905797
Calcul de $t_0$ (variances égales)			
	264,412	32,862	3,523
	46	46	46
	5,748	0,714	0,077
	0,083	0,083	0,083
	0,479	0,060	0,006
	0,692	0,244	0,080
	-0,38	-0,25	0,01
$t_0 =$	-0,542	-1,025	0,104
Degrés de liberté	46	46	46
t de la table	2,013	2,013	2,013
Interprétation	ÉGALE	ÉGALE	ÉGALE