

# Nombre et taille des parcelles en expérimentation agronomique sur plantes à grand développement végétatif. Application au cas du bananier.

**X. PERRIER\***

**NOMBRE ET TAILLE DES PARCELLES EN EXPERIMENTATION AGRONOMIQUE SUR PLANTES A GRAND DEVELOPPEMENT VEGETATIF.**  
**APPLICATION AU CAS DU BANANIER.**

**X. PERRIER (IRFA).**

*Fruits*, Mar. 1986, vol. 41, n° 3, p. 163-173.

**RESUME** - Dans le cas de plantes à grand développement végétatif, comme le bananier, le problème de l'efficacité des dispositifs expérimentaux est celui de l'équilibre entre le nombre de répétitions et le nombre d'individus par parcelle. Ce problème est compliqué par le risque de disparition de plants en cours d'expérimentation. Des résultats de statistique classique et de techniques de simulation sur un lot de données expérimentales permettent d'avancer certains éléments de réponse. Les possibilités d'extrapolation à d'autres conditions d'environnement sont discutées.

## INTRODUCTION

Un grand nombre de travaux ont apporté des éléments de réponse à la question souvent posée par les expérimentateurs en agronomie : «Combien doit-on faire de répétitions ?». Les statisticiens proposent des règles de décision que l'expérimentateur applique à sa culture, en fonction de la variabilité attendue, des différences qu'il souhaite mettre en évidence et des seuils de probabilité qu'il se fixe (PHILIPPEAU, 1980).

La plupart de ces travaux concernent des cultures à développement végétatif réduit (les céréales en particulier ont été beaucoup étudiées) où le nombre de plants par parcelle élémentaire est extrêmement élevé (tableau 1) ; l'unité expérimentale est en général la parcelle et l'on admet

que la valeur «moyenne» observée sur une parcelle est bien représentative du potentiel de cette parcelle. Par contre, pour des cultures arbustives et certaines cultures tropicales où le développement végétatif est beaucoup plus important, l'adoption d'un nombre comparable d'individus par parcelle conduirait à des unités extrêmement grandes où les hétérogénéités d'environnement (sol, microclimat ...) seraient considérables. On est donc contraint d'opter pour des parcelles contenant un nombre réduit d'individus (tableau 1). Le problème devient «Combien doit-on faire de répétitions et combien d'individus doivent comporter les parcelles ?». La superficie d'une expérimentation agronomique ne peut pas croître exagérément de par les hétérogénéités habituelles de l'environnement et en particulier du sol ou par de banales contraintes de place disponible ; le problème est alors celui de l'équilibre optimal (pour un seuil de probabilité fixé et pour un niveau de différence à mettre en évidence) entre le nombre de répétitions et la taille d'une parcelle élémentaire.

\* - IRFA/CIRAD - B.P. 5035 - 34032 MONTPELLIER Cedex

**TABLEAU 1 - Comparaison des superficies et des nombres de plants pour des expérimentations agronomiques sur orge (données ITCF : Institut technique des Céréales et Fourrages) ou sur bananier (données IRFA).**

	Orge (ITCF)	Bananier (IRFA)
densité	300 pl/m <sup>2</sup>	2 000 pl/ha
Parcelle élémentaire		
. surface	12 m <sup>2</sup>	200 m <sup>2</sup>
. nombre d'individus en blocs aléatoires	3 600	40
Exemple d'un essai à 6 traitements et 5 répétitions		
. surface	720 m <sup>2</sup>	6 000 m <sup>2</sup>
. nombre d'individus	108 000	1 200

Les statistiques permettent de répondre à ces questions à conditions qu'un certain nombre d'hypothèses simplificatrices soient vérifiées. Ces simplifications sont nécessaires et ont permis le développement de la panoplie de tous les outils de la statistique dite «classique». Cependant elles ne satisfont pas entièrement le statisticien :

- soit parce qu'il sait que ses hypothèses ne sont pas vérifiées, mais il ne dispose plus alors des outils nécessaires. De nombreux travaux sur la «robustesse» des méthodes montrent les limites des conclusions possibles lorsque certaines hypothèses ne sont pas entièrement vérifiées.

- soit parce qu'il ne peut vérifier la validité de ses hypothèses : tester par exemple la normalité d'une distribution nécessite un nombre important de données dont on ne dispose généralement pas en expérimentation agronomique.

Dans certains cas, le statisticien peut aujourd'hui sortir de cette impasse grâce aux moyens de calcul dont il dispose. Cette démarche, que l'on qualifie parfois de statistique informatique, est fondée en général sur l'estimation d'un paramètre statistique d'une population par le calcul de ce paramètre sur un très grand nombre de sous-populations issues aléatoirement de la population initiale. Citons pour exemple BAKER (1964) qui étudie par cette méthode l'effet de la non-normalité des distributions sur les risques de première et deuxième espèce.

Nous nous proposons ici de confronter ces deux démarches statistiques sur des résultats obtenus en culture bananière et d'en dégager certaines recommandations pour l'expérimentation agronomique.

#### DONNEES EXPERIMENTALES

Une plantation homogène de 1 400 bananiers, variété 'Grande Naine', a été faite en 1983 sur les terrains de la Station de Recherche agronomique de Nyombé, dans le sud du Cameroun. La parcelle apparemment homogène est

constituée de 20 lignes de 70 bananiers (espacement : 1,55 m sur la ligne et 3,20 m entre lignes). Un certain nombre d'observations ont été faites sur chaque bananier tout au long du cycle de végétation et de fructification.

Nous nous intéresserons à la variable production, observée sur 1 218 bananiers, les autres, à la suite d'accidents divers, n'ayant pas fleuri. La moyenne de la population est de 30,53 kg avec un écart type de 3,97 kg ; la distribution peut être admise comme normale, les paramètres de dissymétrie et d'aplatissement différent peu de leurs valeurs théoriques (figure 2).

Une cartographie de la parcelle réalisée en représentant chaque plant par sa classe de production (figure 1) ne montre pas de variations structurées (gradients, plages ..). On peut penser que la variabilité observée entre plants viendra des microhétérogénéités de sol et de la variabilité du matériel végétal. Bien qu'étant homogènes génétiquement (clones), les plants présentent une variabilité phénotypique liée en grande partie à la multiplication végétative ; le matériel de multiplication, bien que rigoureusement sélectionné, peut présenter de légères variations (quantité de réserve, état physiologique, etc.) qui pourront entraîner des différences notables de croissance ou de développement.

#### VARIATIONS ET DEPENDANCES ENTRE PARCELLES

Soit une population de variance  $V$  (variance entre unités expérimentales) ; si l'on répète un grand nombre de fois le calcul de la moyenne  $\bar{X}$  de  $n$  unités issues de cette population, ces moyennes auront une certaine variance  $V(\bar{X})$ .

Rappel

Si  $a$  est une constante :

$$V(aX) = a^2 v(X)$$

De même :

$$V(\bar{X}) = V\left(\frac{\sum_{i=1}^n X_i}{n}\right) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^n X_i\right)$$

Par ailleurs si  $X$  et  $Y$  sont deux variables aléatoires :

$$V(X + Y) = V(X) + V(Y) + \text{COV}(X, Y)$$

De même pour  $n$  variables aléatoires

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{COV}(X_i, X_j)$$

Ici les variances des  $X_i$  sont égales à la variance  $V$  entre unités expérimentales,

$$\text{d'où } \sum_{i=1}^n V(X_i) = nV \text{ et } \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{V}{n}$$

$$\text{donc } V(\bar{X}) = \frac{V}{n} + \frac{2}{n^2} \sum_{i=1}^n \sum_{j=i+1}^n \text{COV}(X_i, X_j)$$

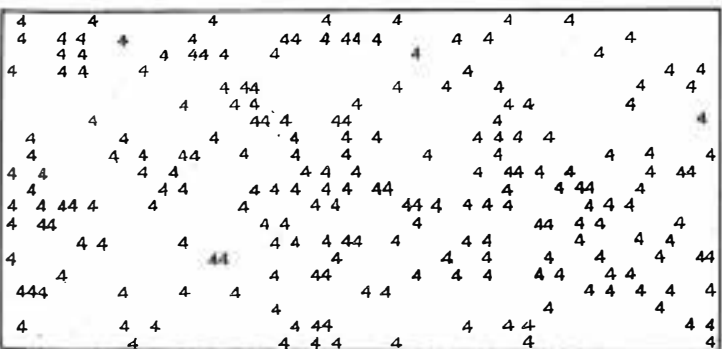
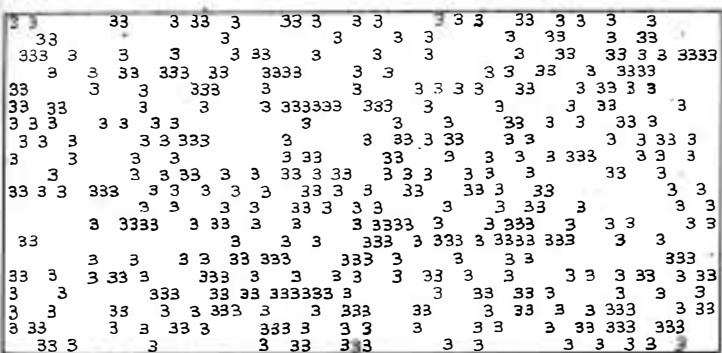
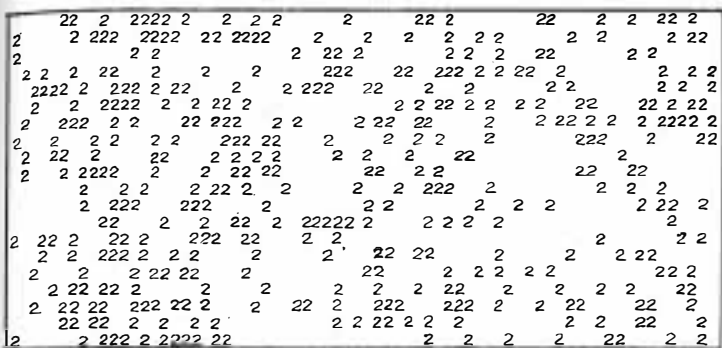
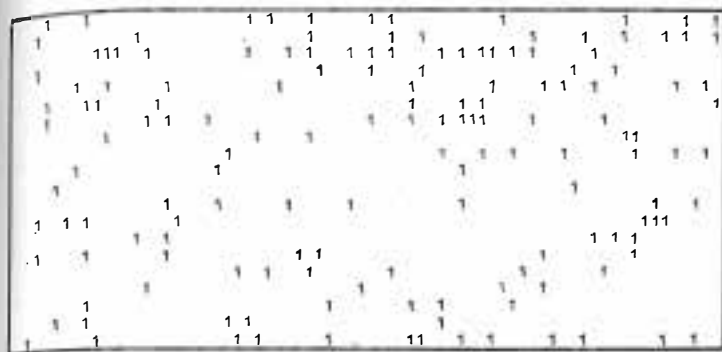


Figure 1 - CARTOGRAPHIE DE LA PARCELLE EXPERIMENTALE. Chaque plant est représenté par sa classe de production.

Si l'on fait l'hypothèse d'indépendance entre le  $X_i$ , les termes de covariance sont nuls :

$$V(\bar{X}) = \frac{V}{n}$$

La variance d'une moyenne diminue donc proportionnellement au nombre d'individus servant à calculer cette moyenne.

Connaissant la variance ( $V_1$ ) de la population des 1 218 bananiers, il est donc possible, en admettant l'indépendance entre plants (c'est-à-dire que la valeur observée pour un bananier n'est liée en aucune manière aux valeurs observées pour les autres bananiers) de calculer la variance  $V_n$  entre les moyennes de  $n$  bananiers

$$V_n = V_1/n$$

La courbe de  $V_n$  en fonction de  $n$  est donnée sur la figure 3.

#### Simulation de parcelles de différentes tailles.

Il est possible d'estimer, sur nos données, la variance entre parcelles de  $n$  individus en la calculant pour 100 parcelles de  $n$  bananiers, la position de chaque parcelle étant tirée au hasard. Les distributions des 100 valeurs calculées pour des parcelles de 5, 10, 20 ... 80 bananiers sont portées sur la figure 2. La décroissance de la variance n'est pas aussi rapide qu'elle devrait l'être si l'indépendance était vérifiée et une très nette distorsion apparaît entre la courbe théorique établie plus haut et la courbe estimée par tirage aléatoire (figure 3).

**Remarque :** ces simulations ont été faites en adoptant différentes formes de parcelles ; on n'observe pas dans notre cas d'influence de la forme des parcelles sur l'évolution de la variance. Ce résultat est lié à l'homogénéité de la parcelle ; NARAYANA (1982), dans d'autres conditions, trouve un effet très net de la forme des parcelles.

#### Ajustement de FAIRFIELD SMITH.

FAIRFIELD SMITH (1938), étudiant les résultats d'un essai à blanc de 1 080 parcelles de blé, avait noté la même distorsion entre la courbe théorique et la courbe obtenue en simulant des parcelles de différentes tailles. Il proposa alors la formule

$$V_n = V_1/n^b$$

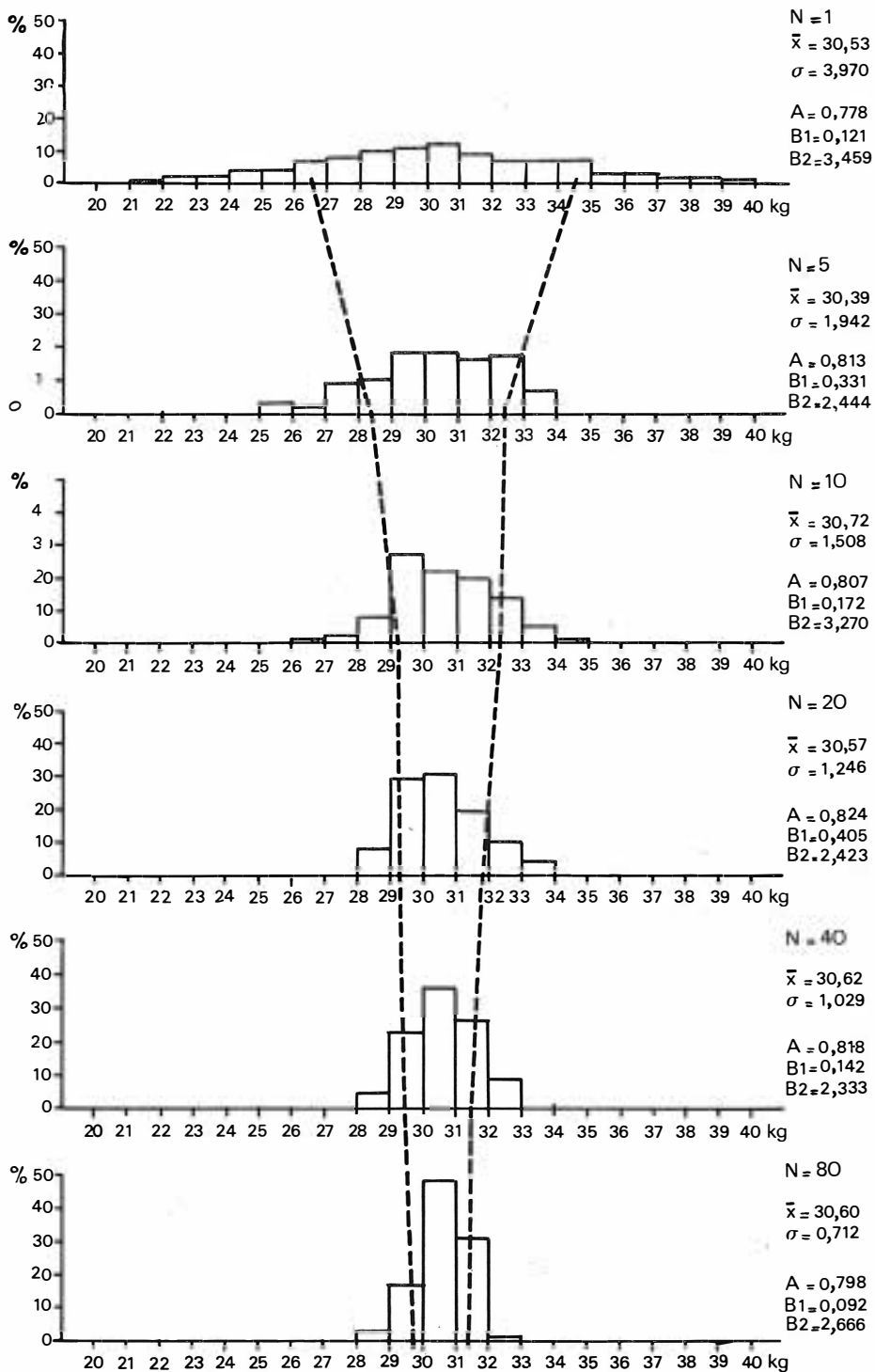


Figure 2 - DISTRIBUTIONS ET PARAMETRES STATISTIQUES DE LA VARIABLE PRODUCTION DES 1218 INDIVIDUS PUIS DES MOYENNES DE 100 TIRAGES ALEATOIRES DE PARCELLES DE 5, 10, 20, 40 ET 80 BANANIERS. Les lignes pointillées joignent les points  $\bar{x} + \sigma$  et  $\bar{x} - \sigma$ . ( $\sigma$  = écart-type; A = coefficient d'aplatissement de GEARY; B1 et B2 = coefficients d'asymétrie et d'aplatissement de PEARSON).

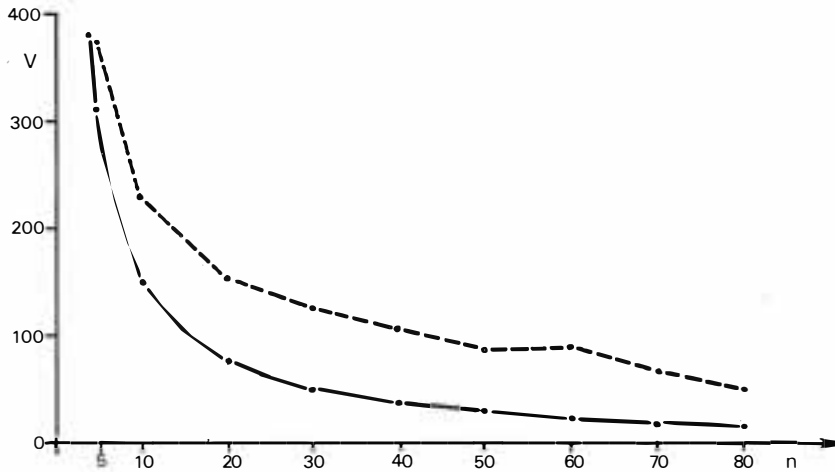


Figure 3 - EVOLUTION DE LA VARIANCE ENTRE MOYENNES DE PARCELLES DE n INDIVIDUS : —•— calculée à partir de la variance de la population  $V_1$  en admettant l'hypothèse d'indépendance ( $V = \frac{V_1}{n}$ ); - - - - estimée par tirage aléatoire de 100 parcelles de n individus.

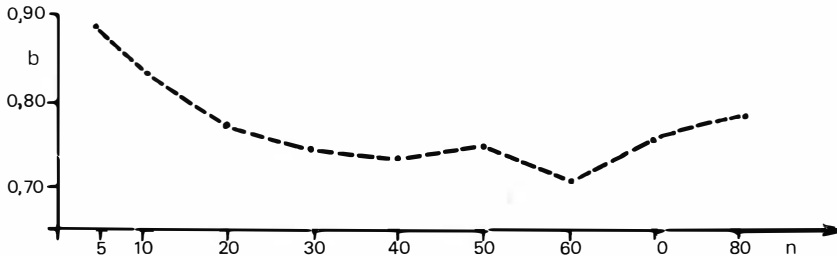


Figure 4 - EVOLUTION DU COEFFICIENT b DE L'AJUSTEMENT DE FAIRFIELD SMITH :  $V_n = V_1/n^b$ .

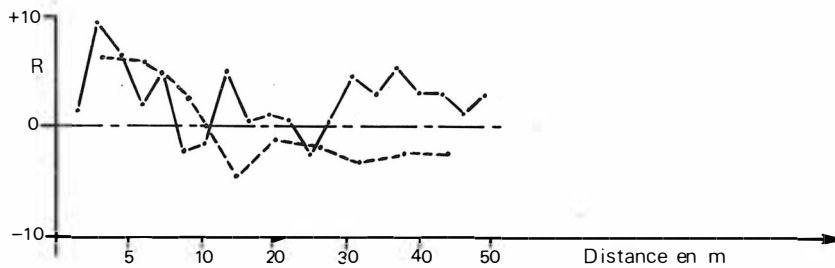


Figure 5 - EVOLUTION DU COEFFICIENT DE CORRELATION ENTRE PLANTS CONTIGUS, PUIS SEPARES PAR UN PLANT, PUIS DEUX, ETC. (— sur les lignes, plants espacés de 1,55 m; - - - sur les colonnes, plants espacés de 3,20 m).

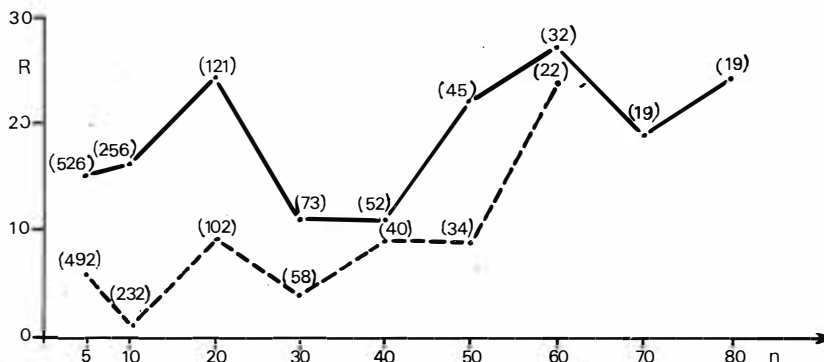


Figure 6 - EVOLUTION EN FONCTION DU NOMBRE D'INDIVIDUS PAR PARCELLE, DU COEFFICIENT DE CORRELATION ENTRE PARCELLES CONTIGUES (—) OU SEPAREES PAR UNE PARCELLE (- - -). Entre parenthèses : le nombre de couples servant au calcul.

$V_n$  étant la variance des moyennes de parcelles constituées de  $n$  unités expérimentales,  $V_1$  la variance de ces unités expérimentales et  $b$  un coefficient entre 0 et 1 mesurant la dépendance.

Si  $b = 1$ , il y a indépendance, on retrouve la formule théorique ; si  $b = 0$ , il y a dépendance totale et la variance reste constante. L'ajustement réalisé sur ses données donnait un  $b$  de 0,749.

FAIRFIELD SMITH appliquant ce calcul à différentes données de la littérature trouvait des coefficients  $b$  très variables (0,40 à 0,80) suivant les cultures. BHARGAVA (1978), sur deux vergers de 480 bananiers (plantation au carré 1,50 x 1,50), obtient les valeurs de 0,79 et 0,92.

Les estimations réalisées sur nos données pour les différentes tailles de parcelles donnent des valeurs de 0,89 à 0,70, les plus fortes valeurs correspondant aux  $n$  les plus faibles (figure 4).

Ce coefficient  $b$  apparaît être fonction de  $n$ , le modèle proposé par SMITH ne semble donc pas entièrement satisfaisant pour décrire la diminution de la variance parcellaire lorsque  $n$  augmente.

#### Variance entre parcelles et corrélations entre plants.

L'hypothèse d'indépendance ne pouvant être admise, il convient de prendre en compte dans l'expression de la variance les termes de covariance :

$$V(\bar{X}) = \frac{V_1}{n} + \frac{2}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{COV}(X_i, X_j)$$

La covariance peut s'écrire en fonction des coefficients de corrélation entre unités expérimentales (entre plants) :

$$\text{COV}(X_i, X_j) = r_{ij} \sqrt{V(X_i) \cdot V(X_j)} = r_{ij} \sqrt{V_1^2} = r_{ij} V_1$$

d'où :

$$\begin{aligned} V(\bar{X}) &= \frac{V_1}{n} + \frac{2V_1}{n^2} \sum_{i=1}^{n-1} \sum_{j=i+1}^n r_{ij} \\ &= \frac{V_1}{n} \left( 1 + \frac{2}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^n r_{ij} \right) \end{aligned}$$

On conçoit donc que la formule de FAIRFIELD SMITH, qui s'écrit

$$\frac{V_1}{n} (n^{1-b})$$

ne peut traduire la réalité que pour une structure très particulière des coefficients de corrélation.

Dans le cas particulier d'une parcelle en ligne, on trouve-

rait :

$n - 1$  corrélations d'ordre 1 :  $r_1$  (individus voisins)  
 $n - 2$  corrélations d'ordre 2 :  $r_2$  (individus séparés par un plant)  
 ...  
 1 corrélation d'ordre  $n - 1$  :  $r_{n-1}$

d'où

$$V(\bar{X}) = \frac{V_1}{n} \left( 1 + \frac{2}{n} \sum_{i=1}^{n-1} (n-i) r_i \right)$$

Il suffirait de connaître les différentes valeurs de  $r_i$  pour avoir une estimation de  $V(\bar{X})$ .

Les corrélations entre individus voisins, puis séparés par 1 plant, 2 plants, etc., le long des lignes (1,55 m d'écartement) et des colonnes (3,20 m) sont faibles et peu structurées (figure 5). Le maximum (positif) est rencontré sur la ligne pour les plants séparés par un individu ; on peut interpréter la faiblesse de la corrélation entre voisins immédiats comme la résultante d'un effet négatif (compétition entre plants) et d'un effet positif (proximité).

Il ne paraît pas possible d'ajuster correctement les  $r_i$  à une fonction simple de  $n$  ; AASTVEIT (1982) souligne la difficulté d'une telle démarche et surtout la très grande diversité des situations possibles.

Faute d'un meilleur ajustement, l'approximation de FAIRFIELD SMITH sera donc utilisée, avec un  $b$  moyen de 0,77, cet ajustement étant satisfaisant pour les parcelles d'au moins 10 individus.

**Remarque :** de la même façon que l'on a calculé les corrélations entre plants, il est possible de calculer des corrélations entre moyennes de parcelles de  $n$  individus (figure 6) ; elles varient entre 0,10 et 0,30 pour des parcelles contiguës et entre 0 et 0,20 pour des parcelles séparées par une parcelle (la forme des parcelles a été choisie la plus carrée possible).

#### EQUILIBRE ENTRE NOMBRE DE REPETITIONS ET NOMBRE DE PLANTS PAR PARCELLE

Il est possible, en connaissant la variance des parcelles de  $n$  individus et le nombre de répétitions  $p$ , de calculer la différence (exprimée en pourcentage de la moyenne) que l'on peut mettre en évidence entre deux traitements :

$$d \% = \frac{100}{\bar{x}} (t_{\alpha} + t_{2\beta}) \sqrt{\frac{2V_n}{p}}$$

$\bar{x}$  : moyenne de la variable analysée.

$\alpha$  et  $\beta$  : les risques de 1ère et 2e espèce que nous avons choisis de 5 et 20 p. 100, valeurs habituellement retenues pour nos expérimentations.

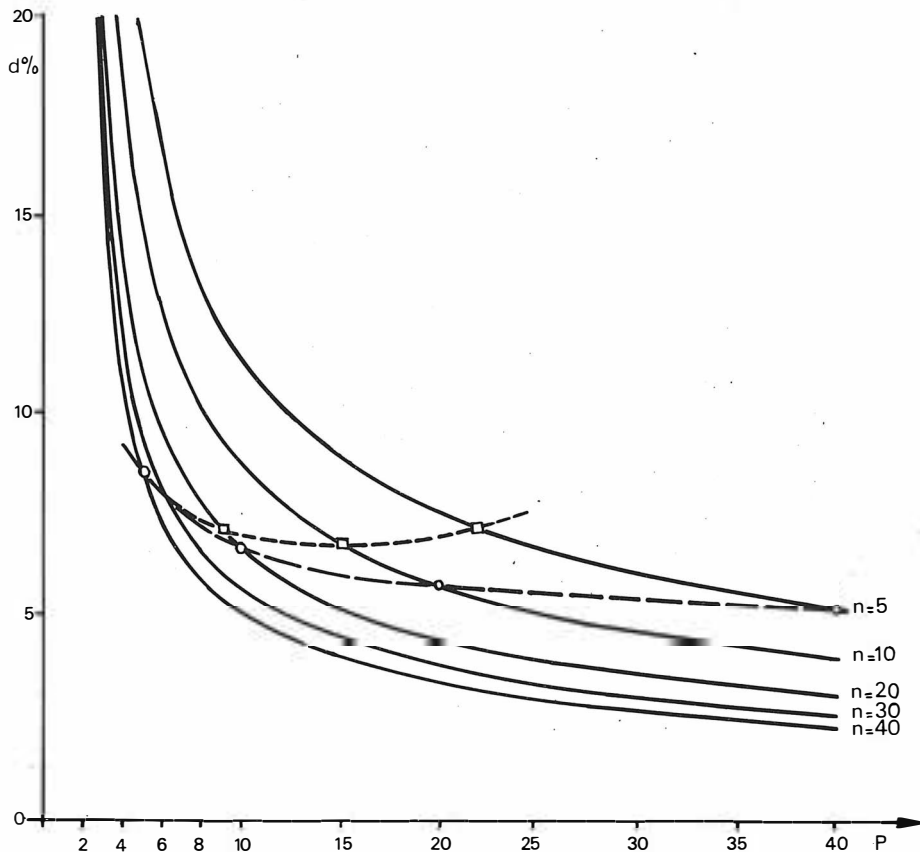


Figure 7 - EVOLUTION DE d% (différence que l'on peut mettre en évidence) EN FONCTION DU NOMBRE DE REPETITIONS (p) POUR DIFFERENTES TAILLES DE PARCELLES (n).

Différentes combinaisons de n et p pour un essai en bloc à deux traitements (○—○ 400 bananiers - pas de bordure entre parcelles ; □—□ 540 bananiers - bordure entre parcelles).

$t_{\alpha} + t_{2\beta}$  (que nous noterons  $t(p)$  est la somme des valeurs de t de STUDENT pour  $p - 1$  degrés de liberté dans le cas d'un dispositif en bloc. Pour un essai à dispositif complètement aléatoire, les valeurs seraient légèrement supérieures, les degrés de liberté étant  $2(p - 1)$ .

$V_n$  : variance résiduelle pour des parcelles de taille n que l'on peut estimer par  $V_1/n^{0,77}$ .

Nous appellerons «sensibilité» cette différence, exprimée en p. 100 de la moyenne, que l'on peut mettre en évidence

TABLEAU 2 - Différence (%) que l'on peut mettre en évidence entre deux traitements pour p répétitions de parcelles de n plants ( $\alpha = 5 \%$ ,  $\beta = 20 \%$ ).

P	2	3	4	5	6	7	8	9	10	15	20	25	30	35	40
n 5	> 100	35,4	23,9	19,1	16,4	14,5	13,2	12,2	11,4	8,9	7,6	6,7	6,1	5,6	5,2
10	84,7	27,1	18,3	14,6	12,5	11,1	10,1	9,4	8,7	6,8	5,8	5,1	4,6	4,3	4,0
15	72,4	23,2	15,6	12,5	10,7	9,5	8,7	8,0	7,5	5,8	5,0	4,4	4,0	3,7	3,4
20	64,8	20,7	14,0	11,2	9,6	8,5	7,8	7,2	6,7	5,2	4,4	3,9	3,6	3,3	3,1
25	59,5	19,0	12,8	10,3	8,8	7,8	7,1	6,6	6,1	4,8	4,1	3,6	3,3	3,0	2,8
30	55,5	17,7	12,0	9,6	8,2	7,3	6,6	6,1	5,7	4,5	3,8	3,4	3,0	2,8	2,6
35	52,3	14,7	11,3	9,0	7,7	6,9	6,3	5,8	5,4	4,2	3,6	3,2	2,9	2,6	2,5
40	49,6	15,9	10,7	8,6	7,3	6,5	5,9	5,5	5,1	4,0	3,4	3,0	2,7	2,5	2,3
45	47,4	15,2	10,2	8,2	7,0	6,2	5,7	5,2	4,9	3,8	3,3	2,9	2,6	2,4	2,2
50	45,6	14,4	9,8	7,9	6,7	6,0	5,5	5,0	4,7	3,7	3,1	2,8	2,5	2,3	2,1

entre traitements ; une bonne sensibilité correspondra à de faibles valeurs de  $d$  %.

#### Constitution et utilisation des abaques.

$$d \% = \frac{100}{\bar{x}} t(p) \sqrt{\frac{2V_1}{p n^{0,77}}}$$

Les calculs ont été faits pour des parcelles de 5 à 50 plants par parcelle et pour 2 à 40 répétitions (tableau 2 et figure 7).

L'objectif de l'expérimentateur est d'obtenir une bonne sensibilité, donc un  $d$  % le plus petit possible, compte tenu des contraintes imposées.

Pour un nombre de répétitions donné, il est évident que la taille des parcelles doit être la plus grande possible, en particulier pour les faibles valeurs de  $p$ . Par exemple pour 6 répétitions,  $d$  % est d'environ 7,4 p. 100 avec 40 bananiers par parcelle ; il n'est plus que 16,4 pour 5 bananiers.

A l'inverse, pour obtenir un  $d$  % fixé a priori, on a le choix entre différentes combinaisons de  $n$  et  $p$ . Par exemple  $p = 6$  et  $n = 40$  ou  $p = 18$  et  $n = 5$  pour une sensibilité de 8 p. 100.

Dans le cas très fréquent où la surface disponible est limitée, on peut rechercher la combinaison qui donne la meilleure sensibilité.

Par exemple (figure 7), si la superficie totale correspond à 400 bananiers, on peut choisir :

5 répétitions	40 bananiers par parcelle
10 répétitions	20 bananiers par parcelle
20 répétitions	10 bananiers par parcelle
40 répétitions	5 bananiers par parcelle

La valeur la plus faible de  $d$  % est obtenue pour le plus grand nombre de répétitions. Il faut cependant noter qu'à partir d'un certain nombre de répétitions, la décroissance devient très lente, le gain n'est que de 0,6 p. 100 en passant de 20 à 40 répétitions. Nous verrons plus loin qu'il est souvent préférable de sacrifier très légèrement  $d$  % pour conserver des parcelles de taille plus grande.

Lorsque les traitements nécessitent la présence de bordures entre les parcelles, le pourcentage de plants non observés augmente très vite lorsque la taille des parcelles diminue, d'où une diminution du nombre de répétitions possible.

Par exemple (figure 7), pour une parcelle totale de 540 bananiers, avec un seul rang de bordure entre parcelles, en choisissant les parcelles les plus carrées possible, on peut réaliser :

5 répétitions	40 bananiers observés par parcelle
9 répétitions	20 bananiers observés par parcelle
15 répétitions	10 bananiers observés par parcelle
22 répétitions	5 bananiers observés par parcelle

L'optimum n'est plus obtenu pour le plus grand nombre de répétitions mais pour environ une douzaine de répétitions de parcelles de 15 bananiers observés. La courbe est assez plate dans cette zone et autorise donc des fluctuations de part et d'autre de ces valeurs.

#### Simulation de dispositifs en blocs complets.

A l'intérieur de la parcelle de 1 400 plants, il est possible de tirer au hasard des dispositifs en blocs complets à 2 traitements ; les parcelles ont été choisies les plus carrées possible, sans bordure ; à l'intérieur de chaque bloc, la position des traitements est aléatoire. 40 tirages ont été faits pour les combinaisons :

5 blocs	40 bananiers par parcelle
10 blocs	20 bananiers par parcelle
20 blocs	10 bananiers par parcelle
40 blocs	5 bananiers par parcelle

Les histogrammes des 40 valeurs de sensibilité (figure 8) montrent des allures très différentes.

Avec de petites parcelles répétées un grand nombre de fois, le résultat est très constant et proche de la valeur attendue. Par contre, dans le cas de grandes parcelles, les résultats sont très dispersés, certains essais obtenant une sensibilité jamais atteinte avec des petites parcelles (moins de 3 %) ou inversement de très fortes valeurs (> 10 %).

La valeur estimée à partir des abaques constitue dans tous les cas un majorant de la moyenne des 40 tirages aléatoires. Il est cependant possible, pour un tirage particulier, de ne pas obtenir la sensibilité escomptée ; c'est en terme de probabilité que doit être interprétée cette notion de sensibilité.

Les écarts entre sensibilités observées et espérées sont liés à la part, dans la variance totale, de la variabilité entre blocs. Il semble, dans notre application, que l'efficacité du dispositif en blocs augmente avec la taille des parcelles.

Rappelons qu'en l'absence d'effet bloc, la puissance d'un dispositif en blocs complets est inférieure à celle d'un dispositif entièrement randomisé. Il serait logique de préférer pour les petites parcelles ce deuxième dispositif et de conserver pour les grandes parcelles le dispositif en blocs.

Ceci cependant doit être raisonné dans chaque cas, en fonction des hétérogénéités possibles du terrain.



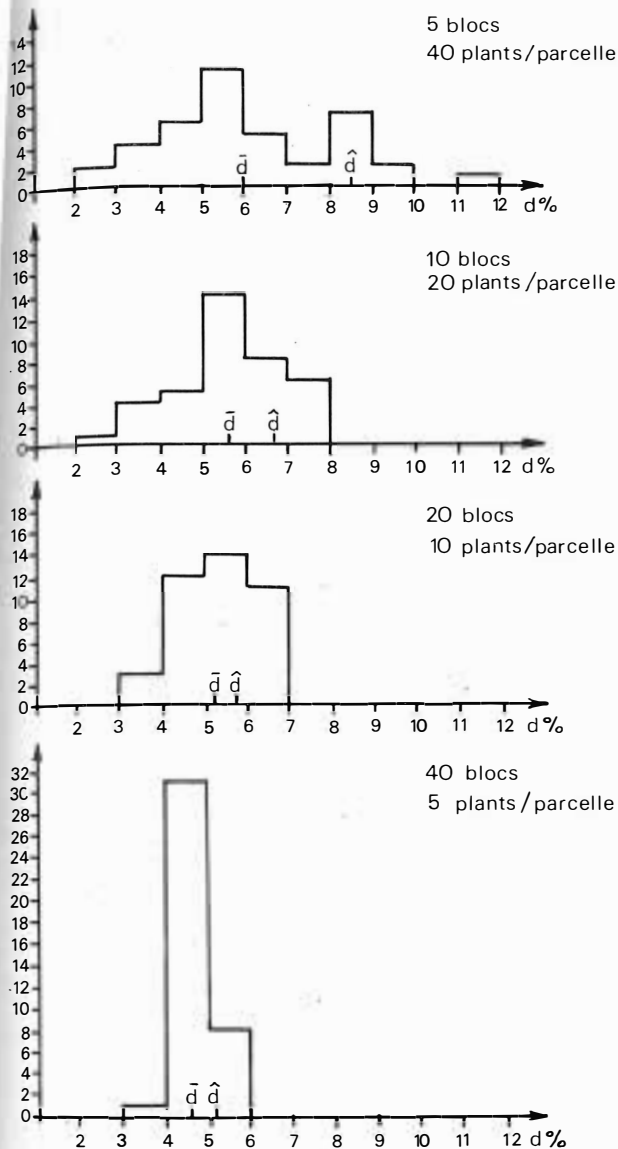


Figure 8 - SIMULATION DE DISPOSITIFS EN BLOCS COMPLETS A DEUX TRAITEMENTS. DISTRIBUTION DES 'SENSIBILITES' OBTENUES SUR 40 TIRAGES ALEATOIRES: ( $\bar{d}$  = moyenne des sensibilites obtenues sur 40 tirages;  $\hat{d}$  = sensibilite estimee a partir des abaquages).

### RISQUE DE DISPARITION DE PARCELLES

Un autre problème de l'expérimentation sur culture à grand développement végétatif, donc à petit nombre d'individus par parcelle expérimentale, est le risque de disparition de plants ; il n'est évidemment plus possible de tenir compte de la moyenne d'une parcelle lorsque le nombre d'unités servant à l'établir devient trop faible.

### Probabilité de disparition de plants.

Dans notre parcelle de 1 400 bananiers, la proportion observée de disparition est de 13,52 p. 100. En admettant une répartition aléatoire de ces disparitions, il est possible de calculer, en utilisant la loi binomiale, la probabilité de disparition de  $m$  individus dans une parcelle initiale de taille  $n$ .

$$P(X = m) = C_n^m P^m (1 - P)^{n - m} \text{ avec } P = 0,1352.$$

Ces calculs ont été faits pour des parcelles de 5, 10, 20 et 40 plants (tableau 3).

Une estimation de ces probabilités peut également être obtenue en tirant au hasard des parcelles de 5, 10, 20 et 40 bananiers et en comptant le nombre de plants disparus. Ce tirage répété un grand nombre de fois (100 fois) donne une estimation de la fréquence de disparition de  $m$  individus (tableau 3).

La concordance des probabilités calculées et des fréquences observées permet d'accepter l'hypothèse de répartition aléatoire.

### Probabilité de disparition de répétitions.

Si l'on admet que l'on doit abandonner une parcelle lorsque le nombre de plants disparus atteint un certain seuil, on peut calculer pour chaque taille de parcelle la probabilité d'abandon d'une parcelle. Ce calcul a été fait pour des parcelles de 5, 10, 20 et 40 bananiers en fixant le seuil à 20 p. 100 de disparition (tableau 4). Cette probabilité est élevée pour les petites parcelles et diminue lorsque la taille de la parcelle augmente.

De la probabilité de disparition d'une parcelle, on passe à la probabilité de disparition d'une répétition, c'est-à-dire de l'un ou des deux traitements. Elle est bien sûr supérieure à la probabilité de disparition d'une parcelle et devient très forte pour les parcelles de petite taille (tableau 4).

On est donc amené à concevoir une nouvelle notion de risque correspondant à la probabilité d'avoir, en fin d'essai, un nombre donné de répétitions. A titre d'exemple, en fixant cette probabilité à 0,8 (c'est-à-dire que l'on espère 8 fois sur 10 ce résultat), on peut calculer, dans le cas d'un essai de 400 plants, le nombre de répétitions probable en fin d'essai et la sensibilité correspondante (tableau 4).

La sensibilité est nettement diminuée particulièrement pour les plus petites parcelles : l'intérêt des petites parcelles se trouve donc diminué. On peut imaginer, en présence de bordure, la très faible sensibilité obtenue avec des parcelles de 5 plants seulement.

TABLEAU 3 - P(X) Probabilité de disparition de 0, 1, 2 ... plants pour des parcelles de différentes tailles (loi binomiale p = 0,1352).

f(X) Fréquence de disparition observée au cours de 100 tirages aléatoires.

	5 plants		10 plants		20 plants		40 plants	
	P(X)	f(X)	P(X)	f(X)	P(X)	f(X)	P(X)	f(X)
X = 0	.484	.51	.234	.19	.055	.08	.003	.02
1	.378	.37	.366	.38	.171	.22	.019	.01
2	.113	.11	.254	.22	.254	.17	.057	.05
3	.018	.01	.107	.15	.238	.24	.113	.13
4	.001	.00	.029	.05	.159	.12	.164	.19
5	.0	.00	.006	.01	.079	.09	.184	.15
6			.001	.0	.031	.05	.168	.11
7			.0	.0	.010	.02	.128	.12
8			.0	.0	.002	.01	.082	.08
9			.0	.0	.001	.0	.046	.05
10			.0	.0	.0	.0	.022	.07
11					.0	.0	.009	.02
12					.0	.0	.004	.0

TABLEAU 4 - Probabilités de disparition des parcelles et de répétitions. Conséquences sur la sensibilité des essais.

Nombre de plants par parcelle	Probabilité de disparition d'une parcelle	Probabilité de disparition d'une répétition	Nombre initial de répétitions	Sensibilité (d %)	Nombre espéré de répétitions (P=0,8)	Sensibilité (d %)
5	0.33	0.55	40	5,2	16	8,6
10	0.27	0.47	20	5,8	10	8,8
20	0.20	0.36	10	6,7	6	9,5
40	0.12	0.23	5	8,4	4	10,4

### APPLICATIONS

Pour résumer les chapitres précédents, on peut estimer la sensibilité d'une expérimentation en blocs complets à deux traitements :

$$d\% = \frac{100}{\bar{X}} t(m') \sqrt{\frac{2V_1}{m'n^2b}}$$

où :

- .  $V_1$  est la variance entre plants ;
- . b, un coefficient traduisant le niveau de liaison entre plant ;
- . n le nombre d'individus par parcelles ;
- .  $t(m')$  la somme des t de STUDENT correspondant aux risques de 1ère et 2e espèces et pour  $m' - 1$  degrés de liberté ;
- .  $m'$  le nombre espéré de répétitions en fin d'essai qui est une fonction du nombre initial de répétitions (m), de la probabilité (P) de disparition d'une répétition et du risque de disparition que l'on accepte (T) ;

$$m' = f(m, P, T)$$

. m étant lui-même une fonction du nombre de plants par parcelle et de k le nombre total de plants de l'expérimentation (en absence de bordure  $m = k/n$ ) ;

. P, probabilité de disparition d'une répétition, est fonction de la fréquence de disparition de plants, du pourcentage de disparition que l'on accepte dans une parcelle et du nombre de plants de ces parcelles.

Il apparaît clairement que, pour des plantes à grand développement végétatif, la détermination de l'équilibre entre nombre de répétitions et nombre de plants par parcelle devient complexe.

Il importe de fixer :

- les risques de 1ère et 2e espèce ;
- la nécessité ou non de bordures et la forme des parcelles si les bordures sont nécessaires ;

- le pourcentage de disparition entraînant l'annulation d'une parcelle ;
- la probabilité d'obtention de la sensibilité souhaitée.

Doivent être connus a priori :

- la variation entre plants ;
- le niveau de liaison entre plants ;
- la fréquence de disparition de plants.

L'application numérique présentée permet de dégager certaines règles pour l'expérimentation agronomique sur bananier.

Dans les conditions de l'étude, il convient d'éviter les trop grandes parcelles qui conduisent à de trop faibles nombres de répétitions. A l'inverse, les petites parcelles présentent divers inconvénients :

- proportion trop importante de plants non utiles si des bordures sont nécessaires ;
- risque élevé de disparition de parcelle ;
- nécessité, dans la pratique, de disposer de parcelles assez grandes : application de produits, suivi de l'évolution d'une variable lorsque la mesure ne peut être répétée sur les mêmes individus (populations de nématodes par exemple).

Il est donc souhaitable d'éviter les très petites parcelles, d'autant que le gain obtenu en diminuant la taille devient vite très faible et qu'inversement l'augmentation de la taille peut laisser espérer une meilleure efficacité du dispositif en bloc.

Concrètement, lorsque l'on peut consacrer 200 à 250 plants à un traitement comme c'est souvent le cas dans les essais menés à l'IRFA, il semble souhaitable d'adopter des

parcelles de 20 à 30 bananiers répétées 8 à 12 fois.

L'extrapolation de ces résultats à d'autres conditions d'environnement paraît possible. Le niveau de variabilité entre individus, observé dans l'application présentée, est faible, mais il a été rencontré dans des expérimentations menées au Cameroun, en Côte d'Ivoire ou aux Antilles. Il est sans doute possible de l'atteindre dans la plupart des cas en apportant le maximum de soin aux différentes techniques culturales. Il est probablement comparable pour les différents clones utilisés habituellement.

Le coefficient *b*, qui traduit le niveau de liaison entre plants, ne doit pas, pour des densités comparables, s'écarter beaucoup de la valeur adoptée.

La fréquence de disparition a été retrouvée dans la grande majorité des expérimentations. Elle n'est modifiée que dans des conditions particulières : accidents climatiques, attaques non contrôlées d'un parasite ...

Il apparaît donc possible d'utiliser ces résultats comme base de réflexion pour le choix, dans d'autres conditions, de l'équilibre entre taille des parcelles et nombre de répétitions.

Par ailleurs, soulignons que ces conclusions s'appliquent à des essais menés en station expérimentale où la place est souvent limitée mais où le niveau de technicité permet une sélection rigoureuse du matériel de plantation, un suivi technique optimal, et une précision certaine des observations (PEARCE, 1953). Par contre, lorsque les expérimentations sont menées chez les producteurs où les surfaces sont moins limitées mais la technicité plus faible, il convient de choisir des parcelles élémentaires beaucoup plus grandes en conservant, si possible, le même nombre de répétitions.

## BIBLIOGRAPHIE

- AASTVEIT (A.H.). 1982.  
Random variation and correlation in agricultural field trials.  
*Biométrie Pramimétrie*, tome XXII, vol. 2, p. 91-114.
- BAKER (G.A.) and BURTON (J.H.). 1964.  
Uniformity field trials and Monte Carlo simulations.  
*Hilgardia*, vol. 35, n° 22, Nov. 1964, p. 615-625.
- BARGHAVA (P.N.), ASHA SAX SENA, BATRA (P.F.). 1978.  
Size and shape of plots in field trials with banana.  
*The Indian Journal of Horticulture*, vol. 35, n° 1, p. 68-73.
- DAGNELIE (P.). 1980.  
Théorie et méthodes statistiques.  
*Presses agronomiques, Gembloux*, vol. 2, 463 p.
- FAIRFIELD SMITH (H.). 1938.  
An empirical law describing heterogeneity in the yields of agricultural crops.  
*Journ. Agric. Sci.*, XXVIII, p. 1-23.
- NARAYANA REDDY (M.) and RAMANATHA CHETTY (C.K.). 1982.  
Effect of plot shape on variability in Smith's variance law.  
*Expl. Agric.*, vol. 18, p. 333-338.
- PEARCE (S.C.). 1953.  
Field experimentation with fruit trees and other perennial plants.  
*Commonwealth Agricultural Bureaux*, 131 p.
- PHILIPPEAU (G.). 1979.  
Puissance d'une expérience.  
*Pub. de l'Institut technique des Céréales et Fourrages*, 24 p.

