



INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE



SMR/402 - 31

COLLEGE ON SOIL PHYSICS
9 - 27 October 1989

"Méthodologie d'Analyse de la Variabilité Spatiale d'une
Parcelle Agronomique. Application à l'Echantillonnage"

M. VAUCLIN
Institut de Mécanique de Grenoble
C.N.R.S.
Saint Martin d'Hères
France

Please note: These are preliminary notes intended for internal distribution only.

Méthodologie d'analyse de la variabilité spatiale d'une parcelle agronomique. Application à l'échantillonnage

Pierre RUELLE, Dhabî BEN SALAH & Michel VAUCLIN (*)

Ecole Supérieure des Ingénieurs de l'Équipement Rural, Medjez-el-Bab, Tunisie

(*) *Institut de Mécanique de Grenoble, C.N.R.S. L.A. n° 6, B.P. 68, F 38402 Saint-Martin-d'Hères, Cedex*

RÉSUMÉ

La variabilité spatiale des rendements de trois cultures pluviales (blé tendre, vesce-avoine et blé dur) successives sur une même parcelle de 0,96 ha, (prise dans un ensemble de 14 ha), de l'humidité et de la texture moyennées sur le premier mètre de sol est étudiée par les méthodes de l'analyse statistique classique et géostatistique. Chaque variable échantillonnée en 81 points de la parcelle est caractérisée par un coefficient de variation (allant de 12 à plus de 100 p. 100) et une échelle intégrale (variant de 0 à 18,6 m) décrivant la structure spatiale. L'ensemble des données est utilisé pour estimer la précision avec laquelle les valeurs moyennes sont déterminées, compte tenu des incertitudes liées à la mesure elle-même, à l'hétérogénéité du milieu et à l'aspect nécessairement fini du domaine échantillonné. Les calculs montrent que, pour 81 observations sur 0,96 ha, les erreurs varient de 3,4 à 26 p. 100 selon les variables considérées.

Inversement, satisfaire un degré de précision souhaité conduit à rechercher une combinaison optimale : nombre d'observations — dimensions géométriques de la surface à échantillonner. Pour des conditions économiques prescrites, on montre par exemple que 45 mesures de l'humidité sur 0,85 ha, 100 observations du rendement en matière sèche-blé sur 2,5 ha seraient nécessaires pour estimer les valeurs moyennes correspondantes à 10 p. 100 près.

Mots clés additionnels : *Fonction de distribution, semi-variogramme, ~~corrélogramme~~, échelle intégrale, incertitude expérimentale, blé, vesce-avoine.*

SUMMARY

A methodology to analyze the spatial variability of an agricultural plot and its application to sampling problems.

Both classical statistical and geostatistical concepts were used to analyze the spatial variability of the yields of three rainfed crops (wheat, vetch-oat, durum wheat) successively cultivated on the same 0.96 ha-plot (taken in a 14 ha-field), and also of moisture content and textural components averaged within the first meter of soil. Each variable sampled at 81 points was characterized by a coefficient of variation (from 12 to over 100 %) and an integral scale (ranging from 0 to 18.6 m) describing the spatial structure. All the data were used to estimate the accuracy of the calculated mean values by taking into account the uncertainties in the measurements themselves, and the errors resulting from the heterogeneity of the medium and from the finiteness of the sampling area. The results show that the total relative errors were in the range 3.4 — 26 % according to the variables. Inversely, in order to satisfy a prescribed error, an optimal number of observation sampling area combination was sought. For a given set of economical conditions, it was shown, for instance, that 45 measurements of moisture content on 0.85 ha, or 100 observations of total dry matter of wheat on 2.5 ha, would be required to estimate the corresponding mean values with 10 % error.

Additional key words : *Probability density function, semi-variogram, correlogram, integral scale, experimental errors, wheat, vetch-oat.*

I. INTRODUCTION

En physique du sol, de très nombreuses études expérimentales effectuées au cours de cette dernière décennie ont montré la grande variabilité spatiale des

propriétés physico-chimiques des sols, même à l'échelle de la parcelle agronomique (VAUCLIN, 1983). Le caractère bien souvent non-aléatoire de ces variations (WEBSTER & CUANALO de la C., 1975 ; CAMPBELL, 1978 ; HAJRASULIHA *et al.*, 1980 ; GAJEM *et*

45 unités d'acide phosphorique. De plus, 33 unités d'azote ont été fournies à la fin du tallage. Un désherbage chimique a permis de maîtriser convenablement le développement des adventices. En des parcelles de 0,6 m² situées aux 81 nœuds du maillage, les mesures suivantes ont été effectuées :

- deux semaines après la levée (15 décembre 1983) : nombre de pieds par unité de surface (NP) ;
- à la récolte (25 mai 1984) : nombre d'épis par unité de surface (NE) ; rendement de matière sèche totale (MST) et épis (MSE).

Le tallage épis (TE) a été calculé par $TE = NE/NP$. De plus, la texture du sol et les humidités pondérales à 3 dates (25 avril 1981 ; 1^{er} mars 1982 et 14 mars 1983) ont été déterminées aux 81 nœuds par prélèvement à la tarière sur [0-1 m] au pas $\Delta z = 20$ cm.

B. Méthode d'analyse

Soit une parcelle de surface S de dimensions géométriques finies sur laquelle des observations d'une variable ou d'un paramètre quelconque $Z(r)$ sont effectuées en N points r_i de coordonnées (x_i, y_i) . Chaque observation est considérée comme une réalisation particulière de la variable aléatoire $Z(r)$ en tout point. La structure statistique de cette variable est alors parfaitement définie par ses fonctions de répartition et d'autocovariance. En effet, l'identification de la densité de probabilité $f(Z(r))$ permet de calculer les premiers moments définis par :

$$E\{Z(r)\} = \int_{-\infty}^{+\infty} Z(r)f(Z(r))dZ \quad (1)$$

et par conséquent la valeur moyenne :

$$\langle Z \rangle = E\{Z(r)\} \quad (2)$$

et la variance :

$$\sigma^2 = E\{Z^2(r)\} - E^2\{Z(r)\} \quad (3)$$

Sous l'hypothèse de stationnarité du 2^e ordre $E\{Z(r)f\}$ et autocovariances indépendantes de l'espace, donc de r , la connaissance de la fonction d'autocovariance :

$$C(h) = E\{Z(r) \cdot Z(r+h)\} - E^2\{Z(r)\} \quad (4)$$

permet de déterminer la fonction d'autocorrélation définie par :

$$\rho(h) = C(h)/\sigma^2 \quad (5)$$

où h est la distance entre observations.

La valeur λ pour laquelle $\rho(h) = 0$ est la longueur d'autocorrélation. Elle correspond à la distance au-delà de laquelle les observations sont indépendantes les unes des autres. La fonction $\rho(h)$ d'un processus stationnaire peut être caractérisée par différents échelles de longueurs. Ainsi, pour un processus bidimensionnel isotrope, l'échelle intégrale (LUMLEY & PANOFKY, 1964) définie par :

$$J = \left\{ 2 \int_0^\infty \rho(h) h dh \right\}^{1/2} \quad (6)$$

et al., 1981 ; RUSSO & BRESLER, 1981 ; VIEIRA *et al.*, 1981 ; etc...) nécessite alors la prise en compte de ces structures spatiales dans le traitement statistique des observations (SISSON & WIERENGA, 1981 ; VAUCLIN *et al.*, 1982).

En agronomie, l'analyse des rendements, généralement mesurés sur de petites surfaces élémentaires de quelques m², suppose classiquement l'indépendance statistique des observations (SNEDECOR & COCHRAN, 1980 ; CARTER *et al.*, 1983). Cependant, compte tenu des fortes interactions entre rendements et propriétés du sol, la validité de cette hypothèse de base semble très problématique.

Le but de cette étude est de présenter une analyse statistique complète de la variabilité des rendements de 3 cultures pluviales successives sur une même surface de 0,96 ha, de l'humidité et de la texture du sol et d'en tirer quelques conclusions aux plans tant de l'agronomie que de la statistique d'échantillonnage.

II. MATÉRIEL ET MÉTHODES

A. Dispositif expérimental

Une expérimentation a été conduite pendant 3 ans (1981-1984) sur une parcelle de 0,96 ha (prise dans un ensemble de 14 ha) située sur le domaine agricole de l'Ecole Supérieure des Ingénieurs de l'Équipement Rural de Medjéz-el-Bab (Tunisie) dans le but d'étudier les relations eau-sol-plantes au cours d'une rotation triennale traditionnelle vesce-avoine, blé, orge non irrigués.

Le climat est de type semi-aride : pluviométrie moyenne annuelle de 420 mm (sur 80 ans) avec de fortes variations interannuelles.

Le sol est de texture limono-argileuse à argilo-limoneuse sur une profondeur de 1,50 m.

L'ensemble du dispositif expérimental mis en œuvre est décrit en détail par ailleurs (RUELLÉ *et al.*, 1983). Alors que les campagnes de mesures effectuées sous blé (1980-1981) et orge (1981-1982) ont essentiellement porté sur l'étude de l'hydrodynamique de la zone non saturée (RUELLÉ *et al.*, 1983), les campagnes 1982-1983 et 1983-1984 ont été orientées vers les aspects agronomiques et notamment la variabilité spatiale des rendements de vesce-avoine et de blé respectivement. Toutes les opérations culturales (préparation du sol, semis, épandage d'engrais, désherbage éventuel) ont été conduites de façon mécanisée.

La vesce-avoine a été semée en lignes écartées de 0,25 m, le 10 décembre 1982, à la dose de 60 kg/ha pour l'avoine et de 30 kg/ha pour la vesce. La fertilisation a apporté 45 unités d'acide phosphorique et 33 unités d'azote. Aucun désherbage n'a été réalisé durant tout le cycle cultural. A la récolte, le 24 avril 1983, les différentes composantes du rendement en matière sèche totale (MST) de la vesce (MSV), de l'avoine (MSA), des adventices (MSAD) séparées en graminées et dicotylédones ont été déterminées sur des surfaces élémentaires de 0,5 m² situées aux 81 nœuds d'un maillage 10 × 15 m. (fig. 1).

Le blé dur a été semé (100 kg/ha) le 9 novembre 1983 en lignes distantes de 0,20 m, avec un apport de

représente physiquement la plus grande distance moyenne pour laquelle les observations sont corrélées entre elles.

Une autre possibilité d'analyser la dépendance spatiale des observations est l'utilisation du variogramme :

$$2\gamma(h) = E\{[Z(r) - Z(r+h)]^2\} \quad (7)$$

Sous l'hypothèse de stationnarité du 2^e ordre, $\rho(h)$ et $\gamma(h)$ sont 2 outils équivalents pour caractériser la dépendance entre les observations et ils sont liés par :

$$\rho(h) = 1 - \frac{\gamma(h)}{\sigma^2} \quad (8)$$

Contrairement à la fonction d'autocorrélation, l'utilisation du semi-variogramme n'impose pas à $Z(r)$ d'avoir une variance finie. L'hypothèse intrinsèque (valeur moyenne de $Z(r)$ et variance des incréments $Z(r) - Z(r+h)$ indépendantes de l'espace) permet alors le calcul de $\gamma(h)$ par l'expression :

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(r_i) - Z(r_i+h)]^2 \quad (9)$$

où $N(h)$ est le nombre de paires d'observations séparées de la distance h .

Ainsi, la détermination du semi-variogramme $\gamma(h)$ ou du corrélogramme $\rho(h)$ permet de définir la structure spatiale de la variable aléatoire $Z(r)$.

Dans le cas de 2 variables aléatoires stationnaires $Z_1(r)$ et $Z_2(r)$ échantillonnées sur la même surface S (pas nécessairement avec la même densité d'observations), aux fonctions de répartition et d'autocovariance pour chaque variable, il convient d'adjoindre la fonction de covariance définie par :

$$C_{12}(h) = E\{Z_1(r) \cdot Z_2(r+h)\} - E\{Z_1(r)\} \cdot E\{Z_2(r)\} \quad (10)$$

afin de décrire complètement la structure statistique de ces 2 variables et leur éventuelle liaison.

TABLEAU 1

Paramètres des fonctions de distribution des rendements. L'hypothèse nulle de loi normale est rejetée (1) au seuil $\alpha' = 0,05$.
Parameters of the yield probability density functions. Null hypothesis of normal distribution is rejected (1) at the 0.05 level of significance.

Variables	m	σ^2	CV	H ₀ : loi normale	δ_1	δ_2
Blé tendre (1980-81)						
• Nbre épis/m ² (NE)	269,4	2 709,5	0,19	0,13		2,00
Vesce-avoine (1982-83)						
• Vesce : MSV (g/m ²)	132,1	3 102,5	0,42	0,34		2,78
• Avoine : MSA (g/m ²)	195,1	9 654,9	0,50	0,43		2,69
• Adventices : MSAD (g/m ²)						
— graminées	58,4	3 508,5	1,01	0,95 (1)		2,84
— dicotylédones	42,9	2 734,9	1,22	1,72 (1)		5,67 (1)
MST (g/m ²)	428,5	7 532,0	0,20	0,03		3,03
Blé dur (1983-84)						
• Nbre épis/m ² (NP)	171,0	563,4	0,15	0,52 (1)		3,18
• Nbre épis/m ² (NE)	179,5	697	0,15	- 0,19		2,29
• Tallage épis (TE)	1,13	0,039 5	0,17	0,18		2,53
• MS épis (g/m ²) (MSE)	153,3	2 036,2	0,29	0,04		2,74
• MST (g/m ²)	519,4	8 986,5	0,18	0,26		3,22

L'équation (10) conduit à :

$$\rho_{12}(h) = C_{12}(h) \sigma_1 \sigma_2 \quad (11)$$

où $\rho_{12}(h)$ est la fonction de corrélation entre les 2 variables. Contrairement à la fonction d'autocorrélation $\rho(h)$ qui prend la valeur +1 ou -1 pour $h = 0$, $\rho_{12}(0)$ qui est le coefficient de corrélation linéaire classique peut prendre évidemment toutes valeurs entre +1 et -1. Comme dans le cas d'une seule variable, l'hypothèse intrinsèque, plus faible que la stationnarité au 2^e ordre, conduit à l'utilisation du covariogramme défini par :

$$2\gamma_{12}(h) = E\{[Z_1(r) - Z_1(r+h)][Z_2(r) - Z_2(r+h)]\} \quad (12)$$

pour étudier la dépendance statistique des 2 variables $Z_1(r)$ et $Z_2(r)$.

La méthodologie utilisée ici pour étudier la variabilité spatiale de la parcelle consiste à assimiler les N observations de la variable considérée à une réalisation d'un processus stochastique bidimensionnel et à y appliquer les principes d'analyse statistique classique (recherche des valeurs moyennes et des variances) et géostatistique (recherche des structures spatiales) brièvement rappelés ci-dessus.

III. RÉSULTATS

A. Fonctions de distribution

Le tableau 1 donne, pour les différentes variables agronomiques observées, la moyenne (m) et la variance (σ^2) de l'échantillon constitué de 81 observations sur la surface $S = 9 600$ m². La variabilité spatiale est exprimée ici simplement par le coefficient de variation (CV).

On notera que la très mauvaise répartition temporelle des pluies lors des campagnes agricoles 1982-1983 et 1983-1984 explique en grande partie les faibles rendements obtenus. L'analyse détaillée des bilans hydriques est donnée dans RUELLE *et al.* (1984, 1985).

Le calcul des coefficients d'asymétrie (g_1) et d'aplatissement (g_2) et les tests de significativité à 0 et 3 respectivement montrent que l'hypothèse nulle de loi normale peut être acceptée au seuil $\alpha' = 5$ p. 100 pour l'ensemble des variables, à l'exception de la matière sèche des adventices (MSAD) et du nombre de pieds de blé par unité de surface (NP). Alors que la variable NP suit une loi lognormale de distribution, la transformation $\sqrt{x}$ conduit à des lois normales pour les rendements en matière sèche des graminées et des dicotylédones. Bien qu'il s'agisse ici de variables dimensionnelles, ce comportement s'apparente à celui d'événements rares (SNEDECOR & COCHRAN, 1980, p. 288).

Les coefficients de variation des composantes du rendement en blé (1983-1984) apparaissent comparables à d'autres résultats publiés dans la littérature. Ainsi, pour une arachide irriguée par aspersion ($S = 0,8$ ha, 20 observations effectuées sur des parcelles de 116 m²), BRESLER *et al.* (1981) donnent les valeurs suivantes : CV = 0,18, 0,22 et 0,18 pour les rendements en matière sèche totale, en fanes et en gousses respectivement. RUSSO (1984) rapporte un coefficient de variation de 0,10 pour le rendement en fruits d'une culture de piment irriguée en goutte à goutte sur une surface de 2,1 ha (25 observations sur des surfaces élémentaires de 7,5 m²).

La plus forte valeur du coefficient de variation relatif à MSE provient vraisemblablement de l'échauffage induit par des conditions climatiques sévères : déficit hydrique important, température élevée. Ce fait a du reste été confirmé par la grande variabilité observée lors du suivi de la maturation du grain, en différents points de la parcelle.

La forte variabilité de MSV et MSA (CV > 0,40) peut être expliquée, d'une part, par le semis (mélange de graines non nécessairement homogène en raison de leurs dimensions différentes) et, d'autre part, par la concurrence entre les 2 espèces et vis-à-vis des adventices. Cette concurrence se traduit d'ailleurs par une forte réduction de la variabilité du rendement en matière sèche totale (CV = 0,20).

De la même manière, le tableau 2 donne les principaux paramètres de la fonction de distribution du taux moyen sur [0-1,0 m] d'éléments fins (A + L)

TABLEAU 2

Paramètres des fonctions de distribution des humidités pondérées (W) et du taux d'éléments fins (A + L).
Parameters of the probability density functions of gravimetric water contents (W) and silt + clay contents (A + L).

Variables	m	$\sigma^2 \cdot 10^4$	CV	g_1	g_2
W (g/g)					
• 1981	0,134	6,48	0,19	-0,48	3,11
• 1982	0,195	5,61	0,12	-0,35	2,88
• 1983	0,165	4,87	0,13	-0,48	2,57
A + L (g/g)	0,495	64,1	0,16	-0,47	2,91

inférieurs à 20 µm et des humidités pondérales moyennes sur [0-1,0 m], déterminées aux 3 dates de prélèvement et sur les 81 sites de mesure. Les coefficients g_1 et g_2 montrent que ces variables sont normalement distribuées. De plus, les résultats relatifs aux humidités appellent les remarques suivantes :

— Par comparaison avec d'autres résultats publiés, la variabilité est assez faible. Ainsi, BRESLER *et al.* (1981) rapportent des coefficients de variations de 0,34 et 0,27 avant et après irrigation respectivement. — Le coefficient de variation diminue quand l'humidité moyenne augmente. Ce résultat est en bon accord avec les expérimentations intensives menées par BELL *et al.* (1980).

— La répartition spatiale des humidités présente une grande stabilité temporelle au sein d'un même VAUCLIN *et al.* (1984). En effet, le calcul des coefficients de corrélation des rangs de Spearman (SA-

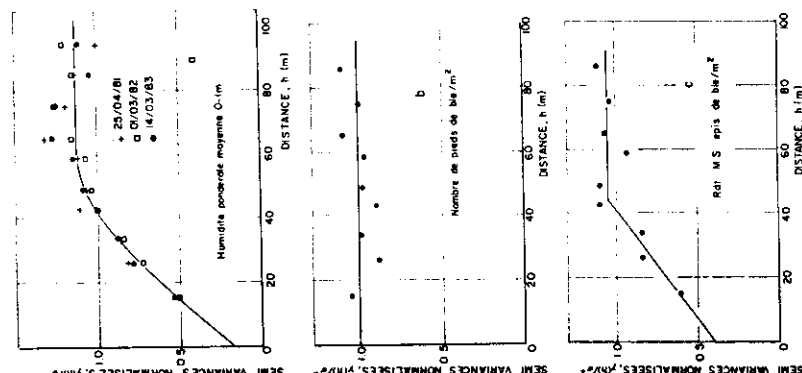


Figure 1
Semi-variogrammes normalisés. Les symboles représentent les valeurs expérimentales et les lignes continues correspondent aux modèles théoriques (voir tabl. 4).
Normalized semi-variograms. Symbols represent experimental values and continuous lines correspond to theoretical models (see table 4).

PORTA, 1978) pour les 3 séries de données ($RHO_{81}^{81} = 0,814$; $RHO_{82}^{82} = 0,910$; $RHO_{83}^{83} = 0,850$) montre que les classements des observations ne sont pas statistiquement différents : les sites les plus secs et/ou les plus humides à un instant donné restent constamment les plus secs et/ou les plus humides.

Nous reviendrons ultérieurement sur cet aspect.

B. Structure spatiale

1. Utilisation des semi-variogrammes

Afin d'identifier les structures spatiales des observations et donc de définir complètement la nature statistique des variables correspondantes, les semi-variogrammes expérimentaux moyens (toutes directions confondues) ont été calculés par l'équation (9) en respectant les lois de distribution. Toutes les valeurs de $\gamma(h)$ ont été normalisées par les variances correspondantes. A titre d'exemple, la figure 1 présente les résultats pour l'humidité pondérale aux 3 dates de prélèvement (1a), le nombre de pieds de blé par unité de surface (1b) et le rendement en matière sèche épis (1c) en 1984. Le tableau 3 synthétise pour toutes les variables considérées les coefficients des modèles théoriques ajustés empiriquement aux valeurs brutes $\gamma(h)/\sigma^2$, selon la procédure proposée par JOURNEL & HUIJBREGTS (1978).

TABLEAU 3

Coefficients des semi-variogrammes théoriques normalisés : effet de pépite (C_0); plateau (C_2); portée (a).

modèle linéaire : $\gamma(h)/\sigma^2 = C_0 + C_1 h$ pour $h \leq a$
 $\gamma(h)/\sigma^2 = C_0 + C_1 a + C_2$ pour $h > a$
) modèle sphérique : $\gamma(h)/\sigma^2 = C_0 + C_1 \left\{ \frac{3}{2} h - \frac{h^3}{2a^3} \right\}$ pour $h \leq a$
 $\gamma(h)/\sigma^2 = C_0 + C_1$ pour $h > a$
 3) pur effet de pépite.

Coefficients of the theoretical normalized semi-variograms. Nugget effect (C_0); sill (C_2) and range (a). (1) corresponds to a linear model, (2) corresponds to a spherical model and (3) is a pure nugget effect.

Variables	C_0	C_1	C_2	a(m)	modèle
1980-81					
NE	0,346	0,016	1,07	45,6	1
1982-83					
MSV	0,700	0,320	1,02	66,0	2
MSA	0,670	0,380	1,05	48,0	2
MSAD	1,06	0	1,06	< 15	3
• graminées	1,06	0	1,06	< 15	3
• dicotylédones	1,01	0	1,01	< 15	3
MST	1,01	0	1,01	< 15	3
1983-84					
NP	1,01	0	1,01	< 15	3
NE	0,628	0,388	1,02	48,5	2
TE	0,670	0,008	1,05	46,5	1
MSE	0,404	0,015	1,05	44,0	1
MST	0,552	0,508	1,06	48,5	2
W	0,180	0,95	1,13	61,5	2
A + L	0,559	0,011	1,11	52,3	1

On notera que les semi-variogrammes directionnels n'ont pas permis de mettre en évidence une anisotropie géométrique.

Les commentaires suivants peuvent être faits :

a) Tous les semi-variogrammes normalisés présentent un palier (C_2) légèrement supérieur à 1 et qui est atteint plus ou moins rapidement. Cela met en évidence l'existence de variances finies. Les dimensions géométriques de la parcelle d'essai semblent donc satisfaisantes pour décrire toute la variabilité spatiale des variables considérées. Si tel n'était pas le cas, cela se traduirait par des fonctions $\gamma(h)/\sigma^2$ continuellement croissantes sans palier, mettant en évidence un effet de dérive caractéristique d'un processus non stationnaire au 2^e ordre.

b) Seules les observations MSAD, MST (1982-1983) et NP ne présentent pas de structure spatiale pour le pas d'échantillonnage adopté.

c) Dans le cas des variables spatialement structurées, l'effet de pépite (C_0) généralement élevé traduit une assez forte variabilité à une échelle inférieure à celle de l'échantillonnage. On notera que les dimensions du maillage ne permettent pas une étude fine du comportement des fonctions $\gamma(h)/\sigma^2$ au voisinage de l'origine. De plus, les valeurs de la portée (a) sont sensiblement les mêmes pour les différentes variables agronomiques structurées, mais inférieures à celles relatives à l'humidité et à la texture du sol. Il doit donc exister un certain déterminisme qui tend à imposer sa propre structure. Ainsi, pour le blé (1983-1984) alors qu'à la levée le nombre de pieds par unité de surface est totalement aléatoire, (semis non déterministe), à la récolte, les différentes composantes du rendement présentent une structure spatiale dont la portée est de 44 à 48 m environ.

En revanche, pour la culture de vesce-avoine (1982-1983), les rendements MSV et MSA admettent une structure avec toutefois un fort bruit mais le rendement en matière sèche totale a un caractère aléatoire qui semble induit par les adventices (graminées et dicotylédones). De plus, le taux important d'impuretés des semences et le mode de semis expliquent probablement la forte variabilité à courte distance ($C_0 = 0,670$ et 0,700).

d) Il convient également de noter (fig. 1a) que les humidités mesurées aux 3 dates présentent très sensiblement la même structure bien que les valeurs moyennes soient très significativement différentes (voir tabl. 2). Ce résultat couplé au test des coefficients de corrélation de rang, traduit clairement la stabilité temporelle de la répartition spatiale de l'humidité liée pour une large part au déterminisme imposé par la granulométrie. En effet, les co-variogrammes (Eq. 12) $\gamma_{W,A+L}(h)/\sigma_{W,A+L}^2$ normalisés par les écarts-types de l'humidité σ_W et du taux d'éléments fins σ_{A+L} mettent en évidence (fig. 2) que l'humidité est fortement corrélée au taux (A + L) jusqu'à une distance de 56 m environ. Cette valeur est d'ailleurs comparable aux portées données tableau 3. Les valeurs positives de $\gamma_{W,A+L}(h)$ indiquent qu'en tout point le sol sera d'autant plus humide que (A + L) sera plus élevé. On notera également que le palier $C_2 = 0,816$ est voisin des valeurs des coefficients de corrélation linéaire entre W et (A + L) : $r = 0,593, 0,712, 0,785$

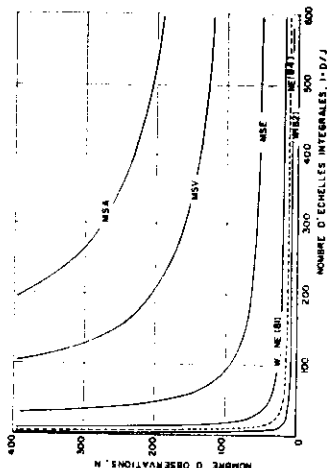


Figure 4
Relations entre le nombre d'observations et le nombre d'échelles intégrales nécessaires pour estimer les valeurs moyennes des différentes variables à $\epsilon_T = 10$ p. 100 près.
Relationships between the number of observations and the number of integral scales required to estimate the mean values of the different variables with 10 % error.

10 p. 100 près et les dimensions du domaine à échantillonner exprimées par le nombre d'échelles intégrales. Les calculs ont été effectués en utilisant dans l'équation (1-9) $\epsilon_T = 0,10$, $(\alpha') = 1,96$, $CV_m = 0,05$ et les coefficients de variation CV obtenus lors de l'expérimentation et donnés tableaux 1 et 2.

Pour une variable quelconque, caractérisée par un coefficient de variation et une échelle intégrale, le nombre d'observations décroît de façon non linéaire à mesure que la surface à échantillonner augmente. On notera que ces courbes présentent une asymptote horizontale ($N_{\min} = t^2(\alpha') / (CV^2 + CV^2)/\epsilon_T^2$ d'après l'équation 14) et une asymptote verticale ($I_{\min} = 2 CV^2/\epsilon_T^2$ d'après l'équation 15).

L'allure générale de ces courbes laisse présager l'existence d'un choix optimal (N_{opt} , I_{opt}) dont le principe de détermination est donné dans l'annexe II. On présente dans le tableau 4 les valeurs N_{opt} , I_{opt} et donc $S_{opt} = \pi/4(I_{opt})^2$ obtenues par la résolution du système (II-4) avec $\epsilon_T = 0,10$ et les hypothèses économiques suivantes :

— Coût d'acquisition d'une observation de rendement (mise en place des placettes, prélèvements de végétation, pesée et détermination des rendements) : $C_N = 60$ FF correspondant à 1 h de travail d'un technicien agricole.

— Charges financières de mise en culture d'un ha de blé (charges de mécanisation, engrais, semences, désherbants et fongicides) : 4 000 FF. Cela correspond au coût par unité d'échelle intégrale $C_I = 4 000 J/D$.

Ces valeurs correspondent approximativement au département de l'Isère (France). Pour l'association vesce-avoine, ces coûts sont supposés réduits d'un tiers. Les valeurs C_I sont données tableau 4. En ce qui concerne les variables sol, on a supposé arbitrairement que la détermination de $(A + L)$ constituait un élément de la reconnaissance pédologique préalable effectuée en sol nu ($C_L = 0$) et que les observations de W faisaient partie du suivi de la culture ($C_L \neq 0$).

L'ensemble de ces résultats appelle les commentaires suivants :

1) A l'évidence, la définition d'un protocole expérimental d'échantillonnage est conditionnée par l'existence ou non de structures spatiales. A titre indicatif, on a également reporté (tabl. 4) le nombre N_{min} d'observations nécessaires (équation 14) pour estimer les valeurs moyennes à 10 p. 100 sur la parcelle d'étude en supposant l'absence de structure ($J = 0$). On notera que le caractère non normal des lois de distribution de MSAD (graminées et dicotylédones) ne permet pas le calcul rigoureux de N_{min} . Ainsi, par exemple, il faudrait 10 observations de NE sur $S = 0,96$ ha ($J = 0$) et 62 observations de NE sur $S = 2$ ha ($J = 12$ m), bien que ces 2 variables présentent la même variabilité spatiale. D'une manière générale, l'échantillonnage d'une variable structurée nécessitera un plus grand nombre d'observations sur de plus grandes surfaces que celui d'une variable aléatoirement distribuée. Ce phénomène couplé à de fortes variabilités peut conduire à devoir effectuer plusieurs centaines d'observations sur des surfaces égales ou supérieures à 10 ha (NE en 1980-81, MSV, MSA, MSE, par exemple).

2) Pour une précision souhaitée, le choix optimum nombre d'observations-surface à échantillonner dépend de l'environnement socio-économique dans lequel l'expérimentation est réalisée. Les équations (II-4) permettent d'en étudier l'incidence. Par exemple, toutes choses égales par ailleurs, la même étude effectuée sur une culture irriguée nécessiterait un plus grand nombre d'observations sur de plus petites surfaces puisque le coût C_I est plus élevé qu'en culture pluviale, en raison des charges financières d'irrigation. On notera également que les mêmes calculs effectués avec des données tunisiennes ($C_N = 1,80$ DT ; 130 DT de charges financières pour la mise en culture d'1 ha de blé) conduisent sensiblement aux mêmes estimations N_{opt} , S_{opt} que celles données tableau 4.

3) A conditions économiques données, il est bien évident que les valeurs N_{opt} et I_{opt} seront d'autant plus grandes que la précision voulue sur la valeur moyenne sera plus sévère.

4) L'existence de l'asymptote verticale $I_{\min} = 2 CV^2/\epsilon_T^2$ (fig. 4) impose pour chaque variable structurée une surface minimum à échantillonner ($S_{\min} = \pi(I_{\min})^2/4$) dont les valeurs sont données tableau 4 afin que le nombre requis d'informations ne soit pas infini.

5) Les équations (I-4) et (I-8) montrent clairement que dans le cas d'une variable stationnaire, l'espérance $\langle Z \rangle$ (équation I-1) et la moyenne spatiale $\bar{Z}$ (équation I-2) sont équivalentes à condition que l'échelle caractéristique D du domaine étudié soit très nettement supérieure à l'échelle intégrale J du processus stochastique ($D/J \gg 1$ afin que $1 \rightarrow \infty$).

V. CONCLUSIONS

Dans cette étude, la variabilité spatiale des rendements agricoles de l'humidité et de la texture du sol, assimilés à des processus stochastiques bidimension-

nels a été analysée en termes de fonctions de distribution et d'autocovariance. Chaque variable étudiée a été caractérisée par un coefficient de variation et une échelle intégrale.

On a montré que, pour 3 cultures pluviales mises en place successivement sur la même parcelle, les rendements pouvaient ne pas être aléatoirement distribués dans l'espace. Un déterminisme lié très certainement aux propriétés texturales du sol et mis notamment en évidence par la grande stabilité spatio-temporelle de la distribution de l'humidité tend à imposer sa structure à la végétation au fur et à mesure du développement de celle-ci.

Un calcul d'erreur fait apparaître que, compte tenu des valeurs des coefficients de variations (de 12 à 20 p. 100) et des échelles intégrales (de 0 à 18 m), les valeurs moyennes sont estimées au mieux à 3 p. 100 (cas des variables aléatoires) et au pire à 26 p. 100 (cas des variables structurées) près pour 81 observations sur une surface de 0,96 ha.

Il apparaît que, dans le cas de variables structurées, la source principale d'imprécision provient du fait que la surface échantillonnée est finie. Prétendre améliorer la précision sur l'estimation des valeurs moyennes nécessite non seulement d'augmenter le nombre d'observations mais aussi la surface. Il est ainsi possible, pour des conditions économiques définies, de trouver un optimum (N_{opt} , S_{opt}) pour satisfaire un degré de précision souhaité.

A l'évidence, il nous semble que cela conduit à remettre en cause la conception des essais agronomiques classiques tant au niveau des surfaces des blocs considérés que de leurs dispositions relatives.

En effet, en assimilant la portée ou la longueur d'autocorrélation à la dimension caractéristique d'un bloc, 2 observations situées à des distances supérieures à ces valeurs seront indépendantes et situées dans 2 blocs différents. Il est alors possible d'appliquer les résultats de l'analyse statistique classique. Il n'y a pas d'avantage à concevoir une expérimentation en blocs randomisés que lorsque l'échelle intégrale J est inférieure à la dimension caractéristique équivalente D de la parcelle élémentaire. On notera que l'espérance mathématique et moyenne spatiale sont équivalentes si $J/D \ll 1$.

A l'inverse, un échantillonnage aléatoire du champ complet est plus judicieux dans le cas $J > D$. Il convient alors de mettre en place un dispositif expérimental sans répétition et de prendre en compte l'aspect fini du domaine considéré.

L'analyse présentée ici souffre néanmoins de quelques limitations :

a) L'ensemble des variables considérées a été assimilé à des processus stationnaires. L'absence de plusieurs réalisations ne permet évidemment pas de justifier théoriquement cette hypothèse. Néanmoins, l'étude du comportement des semi-variogrammes empiriques permet de justifier le concept d'intégrale d'échelle utilisé ici. En effet, la présence d'un palier met en évidence l'existence d'une variance finie et rend ainsi licite l'utilisation du corrélogramme et donc

la méthodologie proposée. L'absence de palier indique une dérive qu'il convient d'identifier avant de poursuivre l'analyse sur les résidus (DELHOMME, 1978). Une autre possibilité consiste à reprendre l'expérimentation sur une plus grande surface si faire se peut.

b) La conception *a priori* d'un schéma d'échantillonnage (nombre d'observations, surface) conçu pour estimer les valeurs moyennes avec un degré de précision souhaité suppose la connaissance préalable du coefficient de variation et de l'échelle intégrale. Cette nécessité constitue à l'évidence le point délicat de l'approche considérée. Néanmoins un examen de la littérature sur le sujet (malheureusement très restreint) montre (tabl. 5) des ordres de grandeur sensiblement voisins de ceux indiqués ici, montrant les conditions différentes tant de culture que d'alimentation hydrique. Il serait souhaitable d'intensifier ce type d'études avant d'en tirer des conclusions de portée plus générale.

TABEAU 5

Exemples de résultats bibliographiques.
Review of bibliographical results.

Variables	CV	J (m)	Sources
Rendement total	0,18	13	
Humidité du sol :			
• avant irrigation	0,34	21	BESSIER <i>et al.</i> (1981) :
• après irrigation	0,27	21	culture d'arachide
Perméabilité du sol à saturation	0,64	25	irriguée par aspersion $S = 0,8$ ha
Rendement fruit	0,10	34	RUSCO (1984) :
Pression de l'eau dans le sol à mi-distance des goutteurs	0,34	36	culture de piment
			irriguée par goutte à goutte $S = 2,1$ ha

c) Cette analyse ne fournit qu'une réponse partielle au problème de l'échantillonnage. En effet, elle ne permet que de déterminer la surface à considérer et le nombre d'observations mais pas leur position dans l'espace. Cet aspect qui peut être abordé par les techniques de krigeage (RUSSO & BRESLER, 1982) et de simulation conditionnelle (DELHOMME, 1978) sera abordé ultérieurement.

Le sujet est donc loin d'être épuisé. Nous espérons seulement que cette étude en suscitera beaucoup d'autres.

Reçu le 1^{er} décembre 1984.
Accepté le 17 février 1986.

REMERCIEMENTS

Cette étude a été partiellement réalisée dans le cadre de la coopération scientifique et culturelle entre la Tunisie et la France. L'aide des personnels de l'ESIER et des Services Techniques du Ministère Tunisien de l'Agriculture a été précieuse pour la conduite des expérimentations.

Vautin M., 1983. Méthodes d'étude de la variabilité spatiale des propriétés d'un sol. In : *« Variabilité spatiale des processus de transfert dans les sols »*, Coll. Soc. Hydrotechn. Fr., INRA, Gr. travail *« Dispersion en milieux poreux »*, Avignon, 24-25 juin 1982. Les Colloques de l'INRA, n° 15, INRA, Paris, France, 9-43.

Vautin M., Vieira S. R., Bernard R., Hatfield J. L., 1982. Spatial variability of surface temperature along two transects of a bare soil. *Water Resour. Res.*, 18, 1677-1686.

Vautin M., Vachaud G., Passet de Sillans A., 1984. Time stability of fields observations of soil water contents. In Udluft P., Merkel B., Friedl K. H. (ed.) : *« Recent Investigations in the case of aridation »*, Int. Symp., Bonn, Munich, 1-5 oct. 1983, C. R., Vol. 1, Techn. Univ., Munich, 231-240.

Vieira S. R., Nielsen D. R., Biggar J. W., 1981. Spatial variability of field-measured infiltration rate. *Soil Sci. Soc. Am. J.*, 45, 1040-1048.

Weber R., Cuatrecasas de la C. H. E., 1975. Soil transect correlations of North Oxtowshire and their interpretation. *Soil Sci. Soc. Am. J.*, 26, 178-194.

Ruelle P., Ben Salah D., Vachaud G., 1985. Bilan hydrique sous cultures par mesures tensiométriques, neutroniques et gravimétriques. II. — Campagne agricole 1983-84. *Rapp. Min. Agric. Tunis.*, 78 p. (à paraître).

Russo D., 1984. A geostatistical approach to the trickle irrigation design in a heterogeneous soil. 2. — A field test. *Water Resour. Res.*, 20, 543-552.

Russo D., Bresler E., 1981. Soil hydraulic properties as stochastic processes. I. — An analysis of field spatial variability. *Soil Sci. Soc. Am. J.*, 45, 682-687.

Russo D., Bresler E., 1982. Soil hydraulic properties as stochastic processes. II. — Errors of estimates in a heterogeneous field. *Soil Sci. Soc. Am. J.*, 46, 20-26.

Saporta G., 1978. *Théorie et méthodes de la statistique*. Ed. Technip, 386 p.

Sisson J. B., Wierenga P. J., 1981. Spatial variability of steady state infiltration rates as a stochastic process. *Soil Sci. Soc. Am. J.*, 45, 692-701.

Snedecor G. W., Cochran W. G., 1980. *Statistical methods*. The Iowa State Univ. Press, 7^e ed., 507 p.

D'autre part, LUMLEY & PANOFSKY (1964) donnent en 1^{re} approximation :

$$\epsilon_1^2 = \frac{2(CV)^2}{1} \quad (1-8)$$

où 1 est le nombre d'échelles intégrales, D/1 contenues dans la surface S assimilée à un cercle de diamètre D.

L'équation (1-6) s'écrit donc :

$$\epsilon_1^2 = \frac{2(CV)^2}{1} \cdot \frac{1}{N} + 2(CV)^2 \left(\frac{1}{1} + 1(\alpha) \sqrt{\frac{1}{N}} \right) \quad (1-9)$$

ANNEXE II

L'optimum économique (N_{opt} , I_{opt}) est classiquement obtenu par la minimisation de la fonctionnelle :

$$M = IC_1 + NC_N - \mu(CV, I, N) - \epsilon_1^2 \quad (11-1)$$

avec : C_1 : coût financier afférent à la mise en culture d'une surface équivalente à la valeur d'une échelle intégrale.

C_N : coût d'acquisition d'une observation.

μ : paramètre de Lagrange.

L'optimum est donné par annulation des dérivées partielles de M par rapport à I, N et μ . Cela conduit à résoudre le système :

$$f(CV, I, N) = \epsilon_1^2 \quad (11-2a)$$

$$\frac{C_1}{C_N} = \frac{\partial f}{\partial N} \quad (11-2b)$$

Pour une précision ϵ_1 et un coefficient de variation CV donnés, la dérivation de l'équation (11-2a) conduit à :

$$df = \frac{\partial f}{\partial I} dI + \frac{\partial f}{\partial N} dN = 0 \quad (11-3)$$

En reportant l'équation (11-3) dans (11-2b) le système à résoudre devient :

$$f(CV, I, N) = \epsilon_1^2 \quad (11-4a)$$

$$\frac{C_1}{C_N} = \frac{dN}{dI} \quad (11-4b)$$

RÉFÉRENCES BIBLIOGRAPHIQUES

- Bell K. R., Blanchard B. J., Schugge T. J., Wilczek N. W., 1980. Analysis of surface moisture variations within large-field scales. *Water Resour. Res.*, 16, 795-810.
- Bresler E., Dasher S., Russo D., Dagan G., 1981. Spatial variability of crop yield as a stochastic soil process. *Soil Sci. Soc. Am. J.*, 45, 600-605.
- Campbell J. B., 1978. Spatial variation of sand contents and pH within contiguous delineation of two soil mapping units. *Soil Sci. Soc. Am. J.*, 42, 460-464.
- Carter T. E., Burton J. W., Cappy J. J., Israel D. W., Boerma H. R., 1983. Coefficients of variation, error variances and resource allocation in soybean growth analysis experiments. *Agron. J.*, 75, 691-696.
- Delhomme J. P., 1978. Kriging in the hydrosocieties. *Adv. Water Resour.*, 1, 251-266.
- Gajem Y. M., Warrick A. W., Myers D. E., 1981. Spatial dependence of physical properties of a typical cornfield soil. *Soil Sci. Soc. Am. J.*, 45, 709-715.
- Haan C. T., 1977. *Statistical methods in Hydrology*. The Iowa State University Press, 378 p.
- Hajrasulha S., Banjabassi N., Melley J., Nielsen D. R., 1980. Spatial variability of soil sampling for salinity studies in Southwest Iran. *Irrig. Sci.*, 1, 197-208.
- Journel A. G., Huijbregts Ch. J., 1978. *Mining Geostatistics*, Academic Press, 600 p.
- Lumley J. L., Panofsky A., 1964. *The structure of atmospheric turbulence*. John Wiley et Sons, Inc. New York.
- Ruelle P., Aouina M. S., Vautin M., Vachaud G., 1983. Bilan hydrique sous cultures par mesures tensiométriques, neutroniques et gravimétriques. I. — Campagne Agricole 1981-82 : Culture d'orge. *Rapp. Min. Agric. Tunis.*, 76 p.
- Ruelle P., Ben Salah D., Vautin M., Vachaud G., 1984. Bilan hydrique sous cultures par mesures tensiométriques, neutroniques et gravimétriques. II. — Campagne agricole 1982-83 : culture de vesse-avone. *Rapp. Min. Agric. Tunis.*, 95 p.

ANNEXE I

Par définition, la moyenne d'une population $Z(t)$ est donnée par le premier moment de sa loi de distribution :

$$\langle Z \rangle = \frac{1}{S} \int_S Z(t) dZ \quad (1-1)$$

La valeur moyenne d'une seule réalisation sur la surface finie S de dimension caractéristique D = (4 S, π)^{1/2} est :

$$\bar{Z} = \frac{2}{D^2} \int_0^D Z(t) dr \quad (1-2)$$

dont on ne peut pratiquement que l'estimation :

$$\bar{Z}_n = \frac{1}{n} \sum_{i=1}^n Z(t_i) \quad (1-3)$$

Ainsi, les incertitudes ϵ_n et ϵ_i sont définies par :

$$\epsilon_n^2 = \frac{\langle (\bar{Z} - \bar{Z}_n)^2 \rangle}{\langle \bar{Z}^2 \rangle} \quad \epsilon_i^2 = \frac{\langle (Z - \bar{Z})^2 \rangle}{\langle \bar{Z}^2 \rangle} \quad (1-4)$$

La variance relative totale est donnée par :

$$\epsilon_n^2 = \epsilon_m^2 + \epsilon_i^2 \quad \epsilon_m^2 = \frac{\langle (\bar{Z} - \bar{Z}_n)^2 \rangle}{\langle \bar{Z}^2 \rangle} \quad \epsilon_i^2 = \frac{\langle (Z - \bar{Z})^2 \rangle}{\langle \bar{Z}^2 \rangle} \quad (1-5)$$

Compte tenu des relations (1-4), les expressions (1-5) s'écrivent également :

$$\epsilon_n^2 = \epsilon_m^2 + \epsilon_i^2 + \epsilon_j^2 + 2 \epsilon_n \epsilon_j \quad (1-6)$$

Il est bien connu (théorème limite central) que, pour des observations indépendantes les unes des autres (non structurées, $J = 0$) ou situées à des distances supérieures à $J \neq 0$ et normalement distribuées, on a :

$$\epsilon_m^2 = \frac{CV^2}{N} \quad \epsilon_i^2 = \frac{CV^2}{N} \quad (1-7)$$

où CV_m et CV_i sont les coefficients de variation associés aux mesures individuelles et à la variabilité spatiale respectivement ; (α') est la variable de Student au niveau de la probabilité α' .

