

La théorie des variables régionalisées, ses applications possibles dans le domaine épidémiologique aux recherches agronomiques en particulier sur le palmier à huile et le cocotier (1)

R. LECOUSTRE (2), P. de REFFYE (3)

Résumé. — La notion de variable régionalisée, ainsi que les calculs permettant l'étude de la structure d'une telle variable, sont introduits. Les applications de l'étude de l'évolution dans le temps des figures représentant les demi-variogrammes à l'épidémiologie sont évoquées. Les principes du krigeage, méthode sans biais d'estimation d'une variable spatiale de densité à partir de sondage, ses applications à la cartographie automatique, sont développés. Enfin, les possibilités d'utilisation de ces méthodes géostatistiques à la résolution de problèmes d'analyse structurale multivariable sont présentées.

INTRODUCTION

Le but du présent article est d'attirer l'attention des utilisateurs potentiels (pédologues, phytopathologistes, entomologistes et agronomes en général) sur l'existence de méthodes statistiques modernes qui permettent la description et l'analyse structurale d'une variable spatiale mieux que par ses seules moyenne et variance (nature d'un sol, aire d'extension d'une maladie ou d'une attaque de ravageur, interactions entre facteurs du sol et production). Ces techniques permettent en outre la cartographie en isovaleurs pour une variable donnée et le calcul de corrélations locales et différées quand on travaille sur plusieurs variables.

Ces méthodes développées récemment, au cours des années 60 [9, 10, 11 et 2], ont été mises au point par les géostatisticiens qui les emploient largement [13]. A l'heure actuelle, les pédologues de l'IRAT-CIRAD et de l'ORSTOM, les agroforestiers du CTFT-CIRAD, les virologues des plantes de l'ORSTOM avec les entomologistes et les phytopathologistes de l'IRHO-CIRAD les ont introduites avec profit dans le domaine des sciences agronomiques.

I. — NOTION DE VARIABLE RÉGIONALISÉE

Une variable est dite « régionalisée » lorsque les valeurs qu'elle prend dépendent de sa position dans l'espace [12] ; de plus, il est courant d'admettre une dépendance existant entre deux sondages proches déterminant des corrélations entre les valeurs mesurées dans certaines zones. Ceci indique l'existence d'une structure à l'intérieur de l'aire explorée.

Un exemple élémentaire [11] montre clairement ce que l'on entend par cette notion. Supposons que deux séries de mesures effectuées pour une variable donnée, à intervalles

réguliers, le long d'une ligne sur le terrain, aient donné les séquences suivantes :

A : 1 - 2 - 3 - 4 - 5 - 6 - 5 - 4 - 3 - 2 - 1,
B : 1 - 4 - 3 - 6 - 1 - 5 - 3 - 4 - 2 - 5 - 2.

La séquence A possède une structure symétrique très nette et dans le cas B, si la structure existe, elle est très faible et montre une forte irrégularité ; cependant ces 2 séries de 11 mesures admettent la même moyenne et la même variance. Ceci montre qu'on ne peut donc pas appréhender la distribution d'une variable spatiale uniquement à l'aide de ces notions classiques.

Il est donc nécessaire de recourir à une méthode qui analyse à la fois la localisation, la continuité, l'anisotropie et le caractère transitif d'une telle variable. Pour ce faire, on retient pour hypothèse que les valeurs prises par une variable régionalisée sont une réalisation particulière d'une fonction aléatoire stationnaire douée d'une fonction d'auto-corrélation. Cette fonction aléatoire $F(x)$ est définie par :

— son espérance mathématique (ou moyenne)
..... $m(x) = E F(x)$,
— sa variance $v(x) = V F(x)$,
— sa covariance $C(x, x') = E F(x) F(x') - m(x) m(x')$,

avec la condition dite de stationnarité : la moyenne, la variance et la covariance sont invariantes par translation :

— $m(x) = m$ la moyenne de $F(x)$ est la même en tous points,
— $v(x) = v$ la variance de $F(x)$ est la même en tous points,
— $C(x, x') = C(x-x') = C(h)$: la covariance de $F(x)$ ne dépend pas de la position des points x et x' mais seulement de la distance h qui les sépare,
— l'accroissement $F(x+h) - F(x)$ ne dépend que de h .

L'outil d'analyse est alors le demi-variogramme défini pour toute distance h comme étant la moyenne des carrés des écarts des valeurs prises par la variable en tous points distants de h l'un de l'autre.

$$G(h) = 1/2 E[F(x+h) - F(x)]^2$$

(1) Voir glossaire page 546.

(2) IRHO-CIRAD, Laboratoire de biomodélisation - Montpellier (*).

(3) CIRAD, Directeur du Laboratoire de biomodélisation (*).

(*) CIRAD, B. P. 5035 - 34032 Montpellier Cedex (France).

On montre que le demi-variogramme et la covariance se correspondent par la relation :

$$G(h) = C(0) - C(h); \text{ où } C(0) = v.$$

Aux grandes distances, les corrélations entre points ne se font plus sentir et $G(h)$ tend vers la valeur maximale $C(0)$. La valeur, notée a , de h pour laquelle ce maximum est atteint s'appelle la portée du demi-variogramme. Le comportement de la fonction $G(h)$ pour $h < a$, caractérise le degré de régularité de la variable régionalisée.

En général, pour l'exploitation simple des données, on se contente du demi-variogramme expérimental construit en tout h par la moyenne.

$$G(h) = (1/2 N_h) \sum_i [F(x_i + h) - F(x_i)]^2,$$

où N_h = nombre de couples $(x_i, x_i + h)$.

On peut observer différents types de demi-variogrammes qu'il est possible d'ajuster à certains modèles classiques.

Si les corrélations sont nulles, le demi-variogramme prend immédiatement sa valeur maximale (phénomène purement aléatoire) ce qui se traduit par un demi-variogramme plat (Fig. 1) : c'est l'effet de pépites pur.

En général, cet effet de pépites, qui dépend d'une microstructure, est toujours superposé aux autres structures mais il reste possible d'ajuster le demi-variogramme observé (G_0 = effet de pépites) aux modèles théoriques suivants :

— le modèle linéaire qui est un modèle sans palier (Fig. 2)

$$G(h) = G_0 + bh;$$

— le modèle gaussien dont la portée pratique est $a' = \sqrt{3} a$ (Fig. 3)

$$G(h) = G_0 + G_1(1 - \exp(-h^2/a^2));$$

— le modèle sphérique dont la portée vaut a (Fig. 4)

$$G(h) = G_0 + G_1(3/2 h/a - 1/2(h/a)^3) \text{ pour } h < a,$$

$$G(h) = G_0 + G_1 \dots \dots \dots \text{ pour } h > a;$$

— le modèle exponentiel dont la portée pratique est $a' = 3 a$ (Fig. 5)

$$G(h) = G_0 + G_1(1 - \exp(-h/a)).$$

Des études de simulation ont montré qu'on peut suivre l'évolution d'une attaque de ravageur ou d'une maladie en observant les demi-variogrammes correspondant aux dates de contrôle simulées. S'il y a une forte propagation de proche en proche, $G_0 + G_1$ (palier) devient très grand par rapport à G_0 et la portée du demi-variogramme augmente. Les figures 6 à 8 montrent comment, d'une situation caractéristique de l'effet de pépites, on passe, par le fruit de la propagation, à une situation où le demi-variogramme peut s'ajuster au modèle sphérique. Cette remarque montre qu'une telle étude par les demi-variogrammes peut être très utile en épidémiologie.

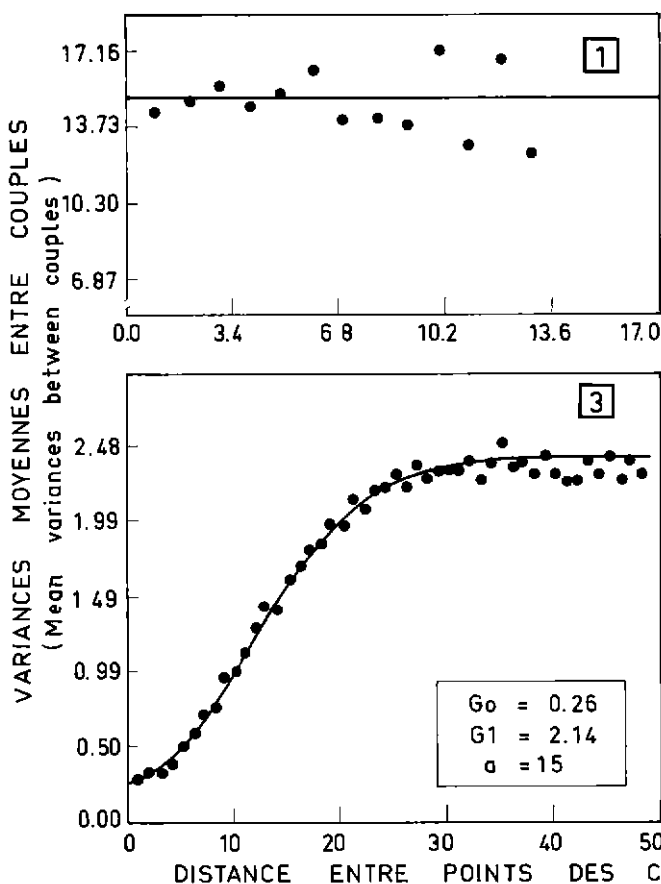


FIG. 1. — Demi-variogramme caractéristique de l'effet de pépites. Maladie à *Phytophthora*, Assinie (taux cumulé de pertes) (Semi-variogram characteristic of the nugget effect. *Phytophthora* at Assinie - cumulative rate of losses)

FIG. 3. — Demi-variogramme caractéristique du modèle gaussien (simulation) (Semi-variogram characteristic of the Gaussian model - simulation).

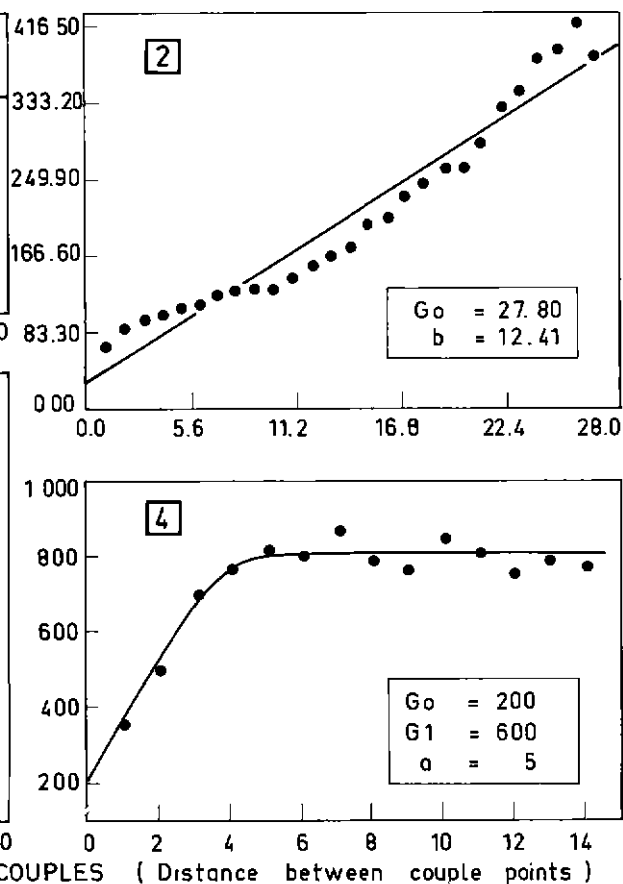


FIG. 2. — Demi-variogramme caractéristique du modèle linéaire. Mosaïque africaine du manioc (*Semi-variogram characteristic of the linear model African cassava mosaic*) ORSTOM, Adiopodoumé.

FIG. 4. — Demi-variogramme caractéristique du modèle sphérique, épaisseur du limon de couverture (*Semi-variogram characteristic of the spherical model thickness of topsoil - silt*) Hole Farm, Norfolk.

II. — LE KRIGEAGE

Le krigage est une méthode d'estimation locale des valeurs d'une variable, en tous points d'une zone géographique, basée sur la théorie des variables régionalisées, qui tient compte du demi-variogramme. Elle a été mise au point principalement par Matheron [9, 10 et 11] et Krige [2] pour l'estimation des réserves en minerai dans l'exploitation minière, pour laquelle elle est de plus en plus utilisée.

L'estimation de la valeur $F(x_0)$, en un point x_0 quelconque, entouré de n points sondés, est obtenue par :

$$F(x_0) = \sum_{i=1}^n L_i F(x_i).$$

où $F(x_i)$ désigne la valeur de la variable au point x_i .

L_i est le coefficient de pondération du sondage x_i , son calcul est associé à la condition

$$\sum_{i=1}^n L_i = 1.$$

Soit $G(x_i, x_j)$ la valeur du demi-variogramme à une distance $h = x_i - x_j$ le long d'une ligne joignant x_i à x_j .

On calcule les L_i de telle sorte que l'espérance de la variance au point x_0 soit minimale. On aboutit au système suivant :

$$\begin{vmatrix} L \\ M \end{vmatrix} = A^{-1} B$$

où :

$$A = \begin{vmatrix} G(x_1, x_1) & G(x_2, x_1) & \dots & G(x_n, x_1) & 1 \\ G(x_1, x_2) & G(x_2, x_2) & \dots & G(x_n, x_2) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ G(x_1, x_n) & G(x_2, x_n) & \dots & G(x_n, x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{vmatrix}$$

$$B = \begin{vmatrix} G(x_1, x_0) \\ G(x_2, x_0) \\ \vdots \\ G(x_n, x_0) \\ 1 \end{vmatrix} \quad \begin{vmatrix} L \\ M \end{vmatrix} = \begin{vmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \\ M \end{vmatrix}$$

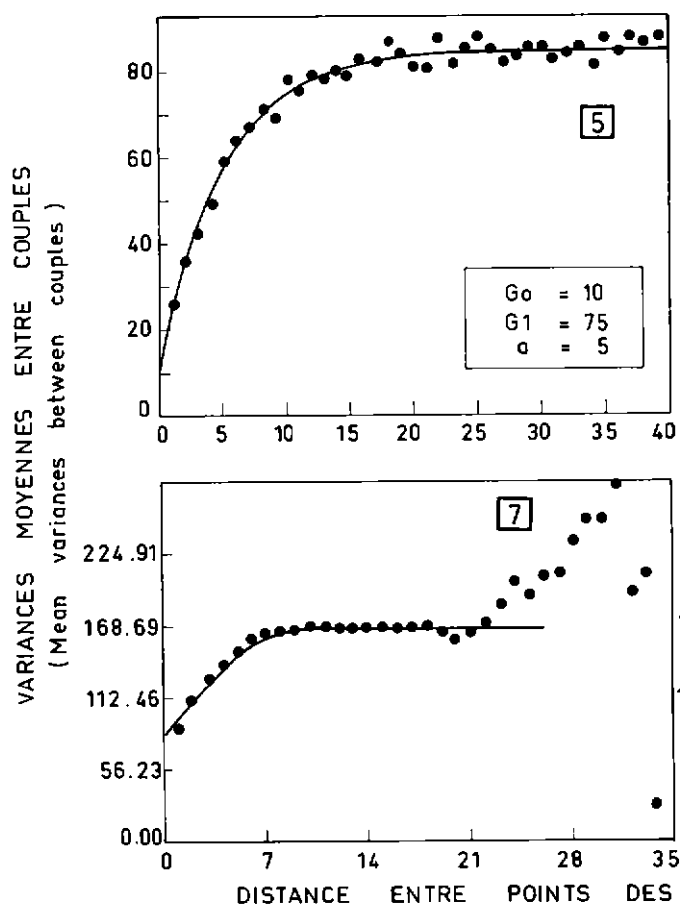


FIG. 5. — Demi-variogramme caractéristique du modèle exponentiel (simulation) (Semi-variogram characteristic of the exponential model - simulation)

FIG. 7. — Demi-variogramme de l'attaque sur le site après diffusion, pendant 4 unités de temps, ajustement au modèle sphérique (Semi-variogram of an attack on site after spread over 4 units of time, adjustment to the spherical model).

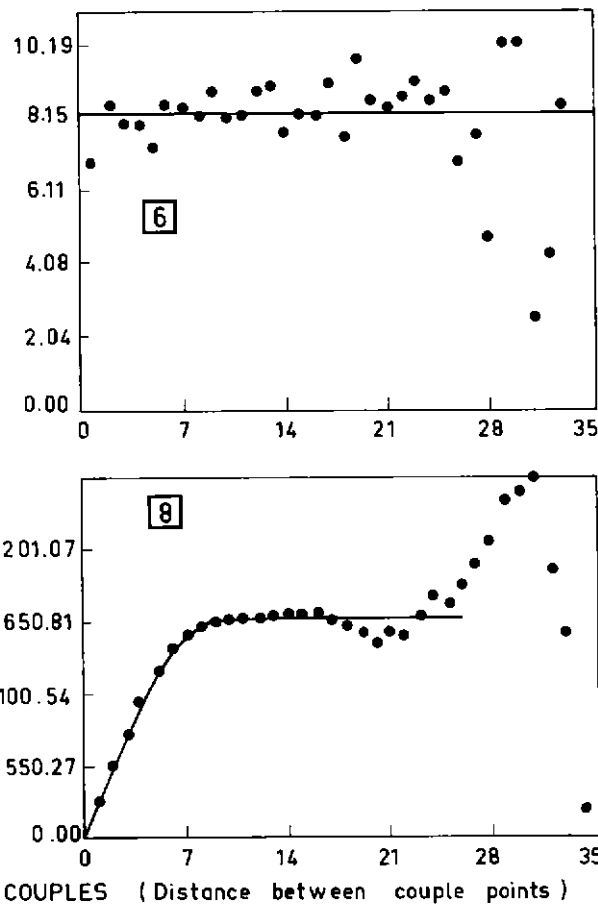


FIG. 6. — Demi-variogramme initial de l'attaque sur le site. Caractéristique d'un effet de pépites pur (Initial semi-variogram of an attack on site characteristic of a pure nugget effect)

FIG. 8. — Demi-variogramme de l'attaque sur le site après diffusion, pendant 8 unités de temps, ajustement au modèle sphérique (Semi-variogram of an attack on site after spread over 8 units of time, adjustment to the spherical model).

où M est le multiplicateur de Lagrange associé à la condition

$$\sum_{i=1}^n L_i = 1.$$

La variance d'estimation minimale en x_0 vaut :

$$s_e^2 = B' \begin{vmatrix} L \\ M \end{vmatrix}$$

où B' est la matrice transposée de B.

Cette méthode est dite sans biais car elle restitue, contrairement à d'autres plus simples, le phénomène en moyenne et variance, redonne leur valeur mesurée aux points sondés et assure une variance d'estimation minimale.

Il est inutile de se servir de points de sondage situés à une distance supérieure à la portée du demi-variogramme pour estimer la valeur $F(x_0)$ en un point.

Cette méthode peut ainsi restituer des cartes d'isovaleurs, ou courbes de niveau (Fig. 9) ; d'autres méthodes graphiques permettent d'obtenir des tracés par parcelle élémentaire (Fig. 10) quelle que soit la forme ou la disposition de ces parcelles ou par plante (Fig. 11). La variable $F(x)$ doit **obligatoirement** avoir un caractère de **variable de densité** (poids, épaisseur, taux d'attaque, densité de populations, etc.).

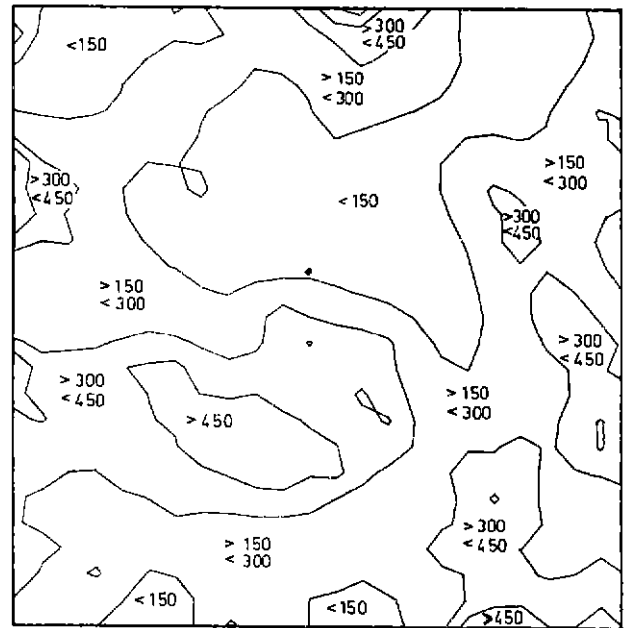


FIG. 9 — Cartographie en isovaleurs d'une variable de densité (Isovalue mapping of a density variable).

Cette technique est très utile pour délimiter les zones d'attaques d'un ravageur et étudier les caractéristiques de la propagation d'une maladie de plantes.

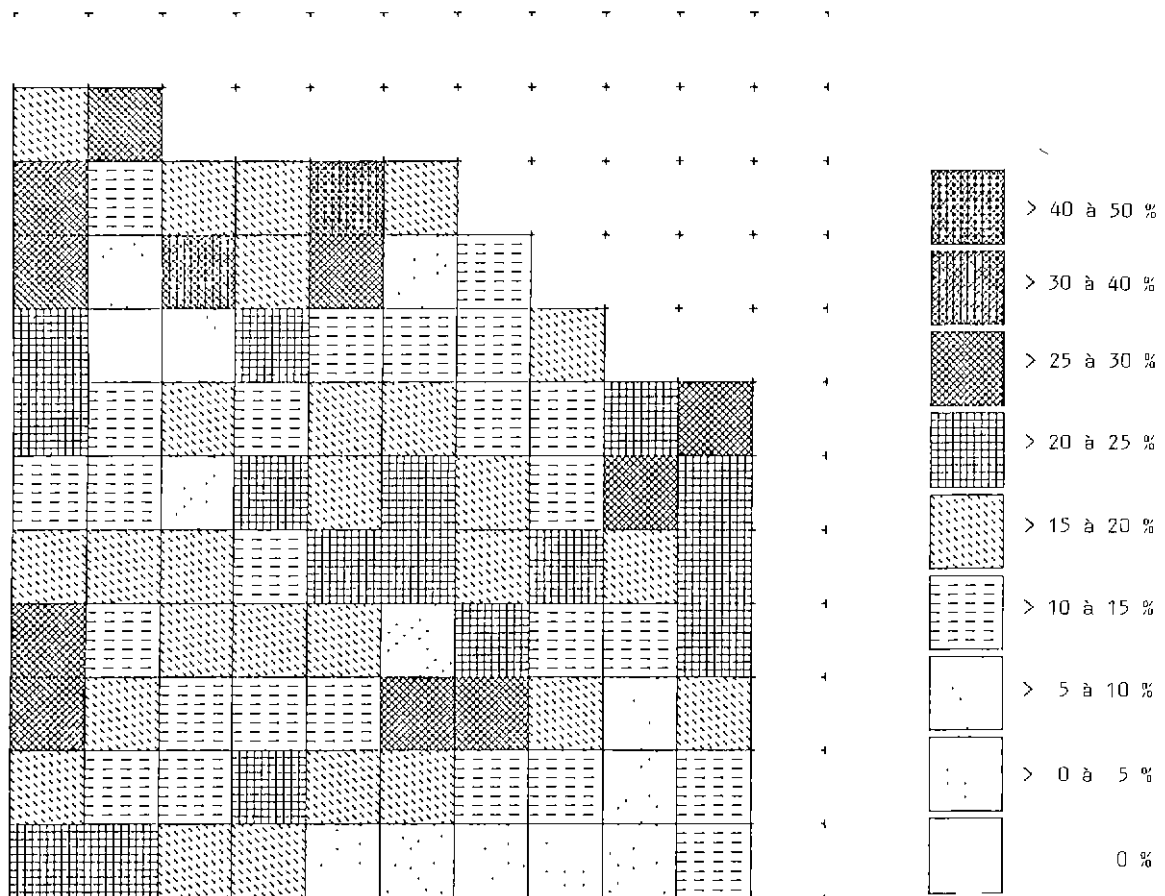


FIG. 10. — Schéma de l'extension de la maladie à *Phytophthora* sur cocotiers à Assinie. Cartographie des taux de perte cumulées par sous-parcelle (3 ans après les 1^{ers} cas) (Diagram of the spread of *Phytophthora* affecting coconut at Assinie. Mapping of the cumulative loss rate per subplot - 3 years after the first outbreak).

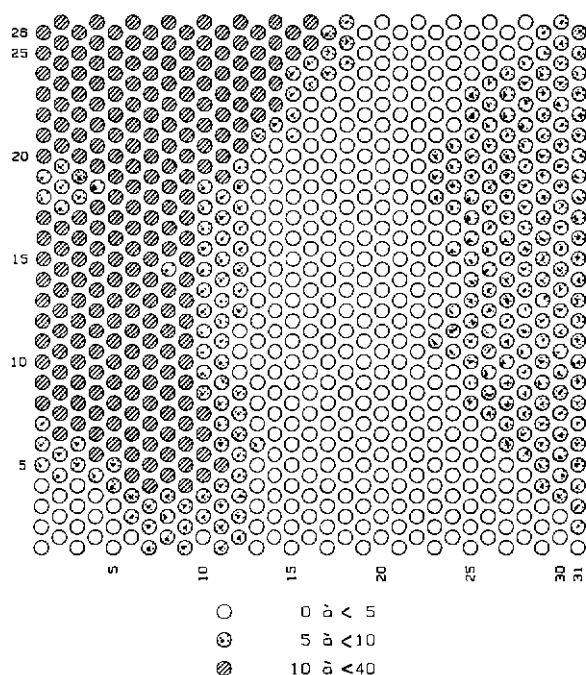


FIG. 11. — Cartographie des indices d'attaque de *Coelaenomenodera minuta* sur palmiers à huile. Estimation point par point. Méthode du krigeage) (Mapping of the index of attack by *Coelaenomenodera minuta* on oil palm. Point by point estimate. Kriging method)

III. — ANALYSE STRUCTURALE MULTIVARIABLE

Jusqu'à présent, nous n'avons pris en considération qu'une seule variable sans nous préoccuper des corrélations entre variables ; ces dernières ne sont d'ailleurs pas importantes dans un cas d'estimation globale d'isovaleurs pour une seule des variables. Par contre, dans le cas d'estimations locales de plusieurs variables, la prise en compte de ces corrélations peut en améliorer l'estimation, les rendre plus compatibles entre elles. Dans ce cas, il faut avoir recours au cokrigeage qui implique de connaître les corrélations différées entre les p variables F_i par :

— les covariances croisées :

$$C_{ij}(h) = E F_i(x) F_j(x + h) - m_i m_j,$$

$$C_{ij}(-h) = C_{ji}(h) = E F_j(x) F_i(x + h) - m_i m_j, \text{ où}$$

m_i et m_j sont les espérances de F_i et F_j ,

— ou les variogrammes croisés :

$$2 G_{ij}(h) = E (F_i(x) - F_j(x+h))(F_j(x) - F_i(x+h)),$$

qui sont les symétrisés des covariances croisées,

car :

$$2 G_{ij}(h) = 2 C_{ij}(0) - C_{ij}(h) - C_{ji}(h).$$

Seule la covariance croisée, plus riche sur le plan descriptif que le variogramme croisé, parce que non symétrique, permet de déceler les corrélations différées.

Dans la pratique, on calculera donc toujours d'abord les $C_{ij}(h)$; s'ils sont symétriques ($C_{ij}(h) = C_{ji}(h)$), on peut alors calculer les $G_{ij}(h)$ qui seront utilisés dans les estimations.

Cette méthode de calcul peut être utilisée avec profit pour dégager les effets d'une structure du milieu agropédologique sur la répartition de la production au sein d'une parcelle : effets de la toposéquence, de la hauteur de la nappe ou d'une qualité de sol déterminée.

CONCLUSION

L'agronome en général, les entomologistes, les phytopathologistes et les pédologues ont donc actuellement à leur disposition un outil mathématique permettant l'analyse structurale des variables spatiales sur lesquelles ils travaillent : la théorie des variables régionalisées et l'établissement des demi-variogrammes.

Ces méthodes permettent en outre d'établir des cartographies d'isovaleurs pour une variable donnée à partir de sondages, c'est le krigeage.

Lorsqu'il s'agit de cartographier les isovaleurs de plusieurs variables d'un site et d'en tirer leurs interactions directes ou différées, on a recours à la méthode de cokrigeage multiple.

Nous avons utilisé avec profit l'analyse des demi-variogrammes pour suivre l'épidémiologie de la Mosaïque africaine du manioc en Côte d'Ivoire (avec l'Orstom), de la maladie de la Pourriture du Cœur du palmier à huile au Pérou et en Equateur, et des maladies à *Phytophthora* du cocotier en Côte d'Ivoire.

Les techniques du krigeage permettent de dresser des cartes de l'indice d'attaque de *Coelaenomenodera minuta* palmier par palmier (Fig. 11) ; cette méthode n'est malheureusement pas utilisée pour délimiter les zones à traiter, même par voie systémique, d'une part à cause de la taille relativement importante de l'échantillon qu'elle impose de prendre (de l'ordre de 7 à 15 p. 100) et, d'autre part, du fait que, pour l'instant, c'est surtout l'aspect végétatif du palmier qui est pris en compte pour ces délimitations. Elles autorisent à alléger les contrôles au champ, en échantillonnant 7 p. 100 des sous-parcelles, destinés à cartographier la répartition des cas de Mosaïque africaine du manioc sur les parcelles expérimentales de l'Orstom en Côte d'Ivoire.

La méthode du cokrigeage sera employée sur la Station IRHO-CIRAD de La Mé pour suivre les effets du niveau de la nappe phréatique, en zone hydromorphe, sur la production de régimes des palmiers. Une vaste étude sur les composantes agropédologiques de la production, menée pour Palmindustrie recourra également à ces méthodes.

Les lecteurs intéressés par ces techniques de calcul et d'analyse se reporteront avec bénéfice aux ouvrages de référence cités dans la bibliographie.

BIBLIOGRAPHIE

- [1] BURGESS T. M., WEBSTER R., Mc BRATNEY A. M. (1981, 1982). — Optimal interpolation and isarithmic mapping of soil properties. *J. Soil Sci.*, 1981, 31, p. 505-524 ; 1982, 32, p. 643-659 (Trad. fr. : Document IRAT).
- [2] KRIGE D. G. (1966). — Two dimensional weighted moving average trend surfaces for ore-evaluation. *J. South African Inst. Mining and Metallurgy*, 66, p. 13-38.
- [3] LECOUSTRE R. et RÉFFY P. de (1983). — Théorie de la mesure et de l'estimation de l'extension des attaques de *Coelaenomenodera* sur les plantations de palmiers à huile. *Rapport interne IRHO*, 20 p.
- [4] LECOUSTRE R. (1985). — Contribution à l'étude de la Pourriture du Cœur de Shushufindi. Etudes statistiques épidémiologiques préliminaires. Etudes sur la régionalisation de la variable pourcentage de pertes cumulées. *Rapport interne IRHO*, Doc. LM/BMST N° 3.
- [5] LECOUSTRE R. (1985). — Contribution à l'étude de la Pourriture du Cœur à Palmorient. Etudes statistiques épidémiologiques préliminaires. Etudes sur la régionalisation de la variable pourcentage de pertes cumulées. *Rapport interne IRHO*, Doc. LM/BMST N° 3 bis.
- [6] LECOUSTRE R. (1986). — Contribution à l'étude de la Maladie

à *Phytophthora* à Samo et Assinie. Etudes sur la régionalisation de la variable pourcentage de pertes cumulées. Etudes statistiques épidémiologiques *Rapport interne IRHO*, Doc. LM/BMST N° 4.

- [7] LECOUSTRE R., FAUQUET C., FARGETTE D. et THOUVENEL J.-C. (1986). — Epidemiology of the African Cassava Mosaic Virus : automatic mapping of the disease spreading. *Communication au Congrès d'épidémiologie de Fort Lauderdale, Floride, U.S.A.* (6-8 août 1986)
- [8] MARBEAU J. P. (1976). — Géostatistique forestière, état actuel et développement nouveaux pour l'aménagement de la forêt tropicale *Thèse de doctorat « Géostatistique forestière » — Ecole des Mines de Paris.*

- [9] MATHERON G. (1963). — Principles of geostatistics. *Economic geology*, 58, p. 1246-1266
- [10] MATHERON G. (1965). — La théorie des variables régionalisées et ses applications. *Cah. Centre de Morphologie, Ecole des Mines de Paris*, N° 5 (trad. anglaise en 1971).
- [11] MATHERON G. (1965). — *Les variables régionalisées et leur estimation*. Masson, Paris.
- [12] NARBONI Ph. (1979). — Application de la méthode des variables régionalisées à des faits du Gabon. — *Note statistique n° 18*, CTFT, Nogent-s.-Marne.
- [13] TRANGMAR B. B., YOST R. S., UEHARA G. (1985). — Application of geostatistics to spacial studies of soil properties *Adv. Agron.*, 38, p. 45-95

SUMMARY

The regionalized variable theory : possible applications to agronomical research, in particular to oil palm and coconut, with respect to epidemiology.

R. LECOUSTRE, P. de REFFYE, *Oléagineux*, 1986, 41, N° 12, p. 541-548.

The concept of regionalized variables and calculations enabling the structure of such a variable to be studied are introduced. Applications to epidemiology of the study of the evolution in time of figures representing semi-variograms are discussed. The principles of kriging, an unbiased method which estimates spatial density variables using sample surveys and its applications to automatic mapping are developed. Finally, the possibility of calling on these geostatistical methods to solve multivariable structural analysis problems is presented.

RESUMEN

La teoría de las variables regionalizadas y sus posibles aplicaciones en el aspecto epidemiológico, en las investigaciones agronómicas, especialmente en la palma africana y el cocotero.

R. LECOUSTRE, P. de REFFYE, *Oléagineux*, 1986, 41, N° 12, p. 541-548.

Se introduce el concepto de variable regionalizada, como también los cálculos que permiten estudiar la estructura de semejante variable. Se mencionan las aplicaciones del estudio de la evolución temporal de las figuras que representan los medios variogramas, en el campo de la epidemiología. Se desarrollan los principios del « kriegeage », que es un método sin distorsiones para evaluar una variable espacial de densidad a partir de sondeos, como también las aplicaciones del citado método a la cartografía automática. Por último, se presentan las posibilidades de utilización de estos métodos geoestadísticos para solucionar problemas de análisis estructural multivariable.

GLOSSAIRE

- Isovaleurs** : Les endroits où la variable prend la même valeur (courbe isovaleur = courbe de niveau).
- Anisotropie** : Ne présente pas les mêmes propriétés dans toutes les directions.
- Krigeage** : Méthode de calcul par interpolation de la valeur prise par la variable en un point à partir des points sondés les plus proches.
- Cokrigeage** : Lorsque 2 variables sont traitées ensemble, on tient compte de la covariance.
- Portée pratique** : Distance à partir de laquelle les valeurs prises aux points sondés sont négligeables.
- Demi-variogramme** : Courbe montrant l'évolution de la covariance prise par la variable en fonction de la distance.

GLOSSARY

- Isovalues** : Points where variables take on the same value (isovalue = contour line).
- Anisotropy** : Does not have the same properties in all directions.
- Kriging** : Calculation method involving interpolation of the value taken by the variable at one point, based on the nearest points sampled.
- Co-kriging** : When two variables are considered together, the covariance is taken into account.
- Practical range** : Distance where values taken at sampled points are negligible.
- Semi-variogram** : Curve showing the evolution of the covariance taken by the variable according to distance.

The regionalized variable theory : possible applications to agronomical research, in particular to oil palm and coconut, with respect to epidemiology

R. LECOUSTRE (1) and P. de REFFYE (2)

INTRODUCTION

The aim of this article is to draw to the attention of potential users (soil scientists, phytopathologists, entomologists and agronomists in general) the existence of modern statistical methods which make it possible to describe and structurally analyse spatial variables more effectively than methods involving

only means and variances (soil type, spread of disease, extent of pest attacks, interaction between soil factors and production). These techniques also enable isovalues to be used for mapping purposes with respect to a given variable and provide the basis for calculations of local and differed correlations when working with several variables.

These modern methods, developed during the 1960s [9, 10, 11 and 2], were perfected by geostatisticians who use them extensively. They have now been successfully applied to the field of agronomy by soil scientists at IRAT-CIRAD and at ORSTOM, agroforestry specialists at CTFT-CIRAD, plant virologists at ORSTOM and entomologists and phytopathologists at IRHO-CIRAD.

(1) IRHO-CIRAD, Biomodelling laboratory, Montpellier (*).
 (2) CIRAD, Director of the biomodelling laboratory, Montpellier (*).
 (*) CIRAD, B.P. 5035, 34032 Montpellier Cedex (France).

I. — THE CONCEPT OF REGIONALIZED VARIABLES

A variable is « regionalized » when its values are dependent on its position in space [12]. Moreover, it is commonly accepted that there exists a dependence between two closely linked samples which determines correlations between values measured in certain zones. This indicates that there is a structure within the area explored.

A basic example [11] clearly shows what is meant by this concept. Let us say that two series of measurements taken for a given variable at regular intervals along 1 row in the field give the following sequences :

- A : 1 - 2 - 3 - 4 - 5 - 6 - 5 - 4 - 3 - 2 - 1,
- B : 1 - 4 - 3 - 6 - 1 - 5 - 3 - 4 - 2 - 5 - 2.

Sequence A is characterized by a well defined symmetric structure whilst that of sequence B, if there is one, is difficult to determine and very irregular. Nonetheless, these two series of 11 measurements have the same mean and variance. This shows that it is impossible to describe spatial variable distribution using traditional concepts only.

A method is therefore required which simultaneously analyses localization, continuity, anisotropy and the transitive character of such a variable. To do this, we adopt the hypothesis that values determined by a regionalized variable are a particular product of a random stationary function with an autocorrelating function. This random function $F(x)$ is defined by :

- its expected value (or mean) $m(x) = E F(x)$,
- its variance $v(x) = V F(x)$,
- its covariance $C(x, x') = E F(x) F(x') - m(x) m(x')$,

where the so-called state of stationarity : the mean, variance and covariance are invariant by translation :

- $m(x) = m$: the mean of $F(x)$ is the same at all points,
- $v(x) = v$: the variance of $F(x)$ is the same at all points,
- $C(x, x') = C(x - x') = C(h)$: the covariance of $F(x)$ is not dependent on the position of points x and x' but is dependent on the distance (h) between these two points,
- the increment $F(x+h) - F(x)$ is only dependent on h .

The analysis tool is therefore the semi-variogram, defined for any distance h as the mean of the squares of deviates of values determined by the variable at all points separated from each other by h .

$$G(h) = 1/2 E [F(x+h) - F(x)]^2.$$

It can be seen that the semi-variogram and covariance correspond through the relationship :

$$G(h) = C(0) - C(h), \text{ where } C(0) = v.$$

At great distances, correlations between points can no longer be ascertained, with $G(h)$ tending towards maximum value $C(0)$. The value, a , of h corresponding to this maximum is called the semi-variogram range. The behaviour of function $G(h)$ with $h < a$ characterizes the degree of regularity of the regionalized variable.

To exploit data on a simple basis, a semi-variogram constructed for any h by the mean is generally sufficient :

$$G(h) = (1/2 N_h) \sum [F(x_i + h) - F(x_i)]^2.$$

where N_h = the number of couples $(x_i, x_i + h)$.

Different types of semi-variograms can be observed which can be adjusted to certain traditional models.

If the correlations are null, the semi-variogram immediately takes on its maximum value (this is strictly a random phenomenon), which is represented by a flat semi-variogram (Fig. 1) : we are dealing with a pure nugget effect.

This nugget effect, dependent on a microstructure, is generally always superimposed on other structures, though the semi-variogram observed can be adjusted to the following theoretical models :

- the linear model, without a plateau (Fig. 2)
 $G(h) = G_0 + bh$;
- the Gaussian model, whose practical range is $a' = \sqrt{3a}$ (Fig. 3)
 $G(h) = G_0 + G_1 (1 - \exp(-h^2/a^2))$;

— the spherical model, whose range is a (Fig. 4)

$$G(h) = G_0 + G_1 (3/2 h/a - 1/2 (h/a)^3) \text{ when } h < a, \text{ and } G(h) = G_0 + G_1 \text{ when } h \geq a ;$$

— the exponential model, whose practical range is $a' = 3a$ (Fig. 5)

$$G(h) = G_0 + G_1 (1 - \exp(-h/a)).$$

Simulation studies have shown that the evolution of pest attacks or diseases can be followed by observing the semi-variograms corresponding to the simulated control dates. If there exists considerable step by step propagation, $G_0 + G_1$ (plateau) becomes very large compared to G_0 and the semi-variogram range increases. Figures 6-8 show how, from a typical nugget effect situation, we pass through the result of propagation to a situation where the semi-variogram can be adapted to the spherical model. This remark reveals that studies based on semi-variograms could be very useful in the field of epidemiology

II. — KRIGING

Kriging is a method involving local estimates of variable values at all points of a geographical zone, based on the theory of regionalized variables, which takes semi-variograms into account. This method was primarily developed by Matheron [9, 10 and 11] and Krige [2] to estimate mineral resources during the exploitation of mines, for which it is used more and more frequently.

An estimate of the value $F(x_0)$ at any one point x_0 , surrounded by n points sampled, is obtained by :

$$F(x_0) = \sum_{i=1}^n L_i F(x_i).$$

where $F(x_i)$ designates the value of the variable at point x_i and where L_i is the weight coefficient of the sample x_i , its calculation associated with the condition

$$\sum_{i=1}^n L_i = 1.$$

i.e. $G(x_i, x_j)$ the value of the semi-variogram at a given distance $h = x_i - x_j$ along the line joining x_i and x_j .

L_i is calculated in such a way that the expected variance value at point x_0 is minimum, which results in the following system :

$$\begin{vmatrix} L \\ M \end{vmatrix} = A^{-1} B$$

where

$$A = \begin{vmatrix} G(x_1, x_1) & G(x_2, x_1) & \dots & G(x_n, x_1) & 1 \\ G(x_1, x_2) & G(x_2, x_2) & \dots & G(x_n, x_2) & 1 \\ \dots & \dots & \dots & \dots & \dots \\ G(x_1, x_n) & G(x_2, x_n) & \dots & G(x_n, x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{vmatrix}$$

$$B = \begin{vmatrix} G(x_1, x_0) \\ G(x_2, x_0) \\ \dots \\ G(x_n, x_0) \\ 1 \end{vmatrix} \quad \begin{vmatrix} L \\ M \end{vmatrix} = \begin{vmatrix} L_1 \\ L_2 \\ \dots \\ L_n \\ M \end{vmatrix}$$

where M is the Lagrange multiplier associated with the condition

$$\sum_{i=1}^n L_i = 1.$$

The minimum estimate variance at x_0 is

$$s_c^2 = B' \begin{vmatrix} L \\ M \end{vmatrix}$$

where B' is the transposed matrix of B

This method is termed unbiased because unlike other more simple methods, it plots the mean and variance of the phenomenon, restores their value measured at sample points and ensures a minimum estimate variance.

It is not worthwhile using sample points located at a distance greater than the semi-variogram range to estimate the value $F(x_0)$ at any one point.

This method can also plot isovalue maps or contour lines (Fig. 9), whilst other graphic methods enable the plotting of experimental plots (Fig. 10) whatever their shape or arrangement or by plant (Fig. 11). It is absolutely **obligatory** that the variable $F(x)$ have a **variable density** character (weight, thickness, rate of attack, population density, etc.).

This technique is very useful for defining pest attack zones and for studying plant disease propagation characteristics.

III. — MULTIVARIABLE STRUCTURAL ANALYSIS

Up until now, this article has taken into account only one variable, without dealing with correlations between variables. Whilst in the case of overall isovalue estimates of only one variable, these correlations do not play an important role, in the case of local estimates of several variables, taking them into account can improve estimates, which gives rise to more compatibility between variables. In this case, co-kriging is called for which implies knowing differed correlations between p variables F_i through :

— **crossed covariances :**

$$\begin{aligned} C_{ij}(h) &= E F_i(x) F_j(x+h) - m_i m_j, \\ C_{ij}(-h) &= C_{ji}(h) = E F_j(x) F_i(x+h) - m_i m_j, \\ &\text{where } m_i \text{ and } m_j \text{ are the means of } F_i \text{ and } F_j ; \end{aligned}$$

— **or crossed semi-variograms :**

$$2 G_{ij}(h) = E(F_i(x) - F_j(x+h))(F_j(x) - F_i(x+h)),$$

which are symmetries of crossed covariances because

$$2 G_{ij}(h) = 2 C_{ij}(0) - C_{ij}(h) - C_{ji}(h).$$

Only the crossed covariance, more descriptive than the crossed semi-variogram because it is asymmetric, enables differed correlations to be detected.

In practice, the $C_{ij}(h)$ should always be calculated first, if it is symmetric ($C_{ij}(h) = C_{ji}(h)$), the $G_{ij}(h)$ can then be calculated which will be used in estimates.

This method of calculation can be successfully used to isolate

the effects of a structure in an agropedological environment on production distribution : effects of the toposequence, water table height or quality of a given soil.

CONCLUSION

Agronomists in general, entomologists, phytopathologists and soil scientists presently have at their disposal a mathematical tool enabling the structural analysis of the spatial variables with which they work : the theory of regionalized variables and the establishment of semi-variograms.

These methods also make it possible to draw up isovalue maps from samples for a given variable, this is kriging.

To map isovalues of several variables from one site and establish their direct or differed interactions, the multiple co-kriging method is used.

We successfully used semi-variogram analysis to follow the epidemiology of African cassava mosaic in Côte d'Ivoire (with ORSTOM), bud rot affecting oil palm in Peru and Ecuador and *Phytophthora* of coconut in Côte d'Ivoire.

Kriging techniques make it possible to draw up maps of the attack index by *Coelaenomenodera minuta* tree by tree (Fig. 11). Unfortunately, this method is not used to determine the zones to be treated, even systemically, because of the relatively large size of the sample which must be taken (from about 7 to 15 p. 100) and because at present, it is primarily the oil palm's vegetative aspect which is taken into account to determine these zones. These techniques make it possible to reduce field controls by sampling 7 p. 100 of the sub plots, intended to map the distribution of African cassava mosaic on experimental plots at ORSTOM in Côte d'Ivoire.

The co-kriging method will be used on the IRHO-CIRAD La Mé Station to follow the effects of water table height, in waterlogged zones, on oil palm bunch production. A vast study on agropedological production components, undertaken by Palmindustrie, will also bring these methods into play.

Readers interested in these analysis and calculation techniques can refer to the reference works given in the bibliography for additional information. ■

**POUR TOUS VOS PROBLÈMES
DE SACS AGRICOLES EN MATIÈRE PLASTIQUE
POUR**

**SEMENCES BOUTURES - ENSACHAGES
FILMS DE PAILLAGE**

Adressez-vous à

ALLIBERT 01 B.P. 1610 à ABIDJAN 01
R. C-I

PREMIER SPÉCIALISTE en AFRIQUE dans l'extrusion de la gaine et de la fabrication des sachets agricoles.

FOURNISSEUR DES PLANTATIONS : de palmiers à huile - café - cacao - ananas - horticulture...

●
DEVIS SUR DEMANDE EN 48 HEURES
●

ALLIBERT

ABIDJAN B.P. N° 1610