

Analyse variographique de données satellitaires pour déterminer la taille des unités d'échantillonnage.

Exemple d'un inventaire d'occupation du sol en milieu agricole et forestier.

FOURNIER Ph.¹, GEROYANNIS H.² & GODARD V.³

Mots clés : Analyse spatiale ; Échantillons ; Inventaires ; Occupation du sol ; Sondages spatiaux ; Télédétection ; Variogrammes

Depuis le milieu des années 80, un grand nombre d'inventaires régionaux de statistiques agricoles utilisent les sondages spatiaux associés aux données satellitaires. Celles-ci sont utilisées pour améliorer les résultats d'estimation de superficie et / ou pour les cartographier. Dans le cadre des sondages spatiaux, l'unité d'échantillonnage est appelée segment. Elle a une taille non nulle, qui doit être définie en fonction de l'hétérogénéité du paysage.

En milieu agricole, l'expérimentation (FOURNIER 1988) a montré qu'un segment calé sur la taille moyenne des exploitations assurait un compromis satisfaisant entre : représentativité des différentes cultures, temps d'enquête et précision des estimations. En revanche, en milieux forestiers ou en milieux faiblement anthropisés, il n'existe pas d'information analogue permettant de calibrer le segment sur une quelconque organisation des paysages. Toutefois, certaines caractéristiques des secteurs à inventorier peuvent être perçues grâce à une analyse variographique des données télédéteectées.

Le variogramme et la dépendance spatiale

Présente depuis les années 70 dans les Sciences de la Terre, l'analyse variographique est un des outils de la géostatistique utilisé pour analyser la dépendance spatiale des échantillons entre eux. Son utilisation sur des données satellitaires est plus récente, elle date des années 80. Elle a essentiellement servi à calibrer l'espacement des unités d'échantillonnage (CURRAN 1988).

Rappelons que les données satellitaires sont spatialement corrélées. Cela signifie que les voisins immédiats d'un pixel tiré au hasard lui sont en général plus ressemblant que des voisins plus

¹Direction Régionale de l'Agriculture et de la Forêt d'Ile de France - 18, avenue Carnot, 94234 Cachan Cedex, France

²Centre d'Analyse et de Mathématique Sociale (CAMS) de l'École des Hautes Études en Sciences Sociales (EHESS) - 54, boulevard Raspail, 75270 Paris Cedex 06, France

³Université de Paris 8 & Centre de Biogéographie-Écologie URA 1514 CNRS - Le Parc, 92211 Saint Cloud Cedex, France

lointains. Cette dépendance spatiale reflète l'extension des zones homogènes qui marquent les paysages à inventorier, son étude doit donc permettre de calibrer les segments d'enquête.

Pour connaître les distances caractérisant les changements de paysage, on tire un échantillon de points et l'on compare chaque point à son voisinage immédiat puis sur des distances de plus en plus grandes et cela dans un certain nombre de directions. On en tire une variance par distance et par direction que l'on représente en général à l'aide d'une courbe, le variogramme. La formule utilisée pour calculer cette variance est la suivante :

$$\gamma (D,d) = \frac{1}{n} \sum_{i=1}^n \frac{(y_{i+d} - y_i)^2}{2}$$

où y_i est la valeur radiométrique du premier point tiré au hasard, y_{i+d} la valeur radiométrique du deuxième point du couple pris à une distance "d" de y_i dans la direction "D" (DUPLAT 1981). La recherche des points d'inflexion sur la courbe informe des ruptures dans la dépendance spatiale, et donc dans la continuité des paysages. Des segments dont la taille est calée sur ces distances doivent permettre une prise en compte satisfaisante des paysages de la zone d'étude.

Les données d'étude

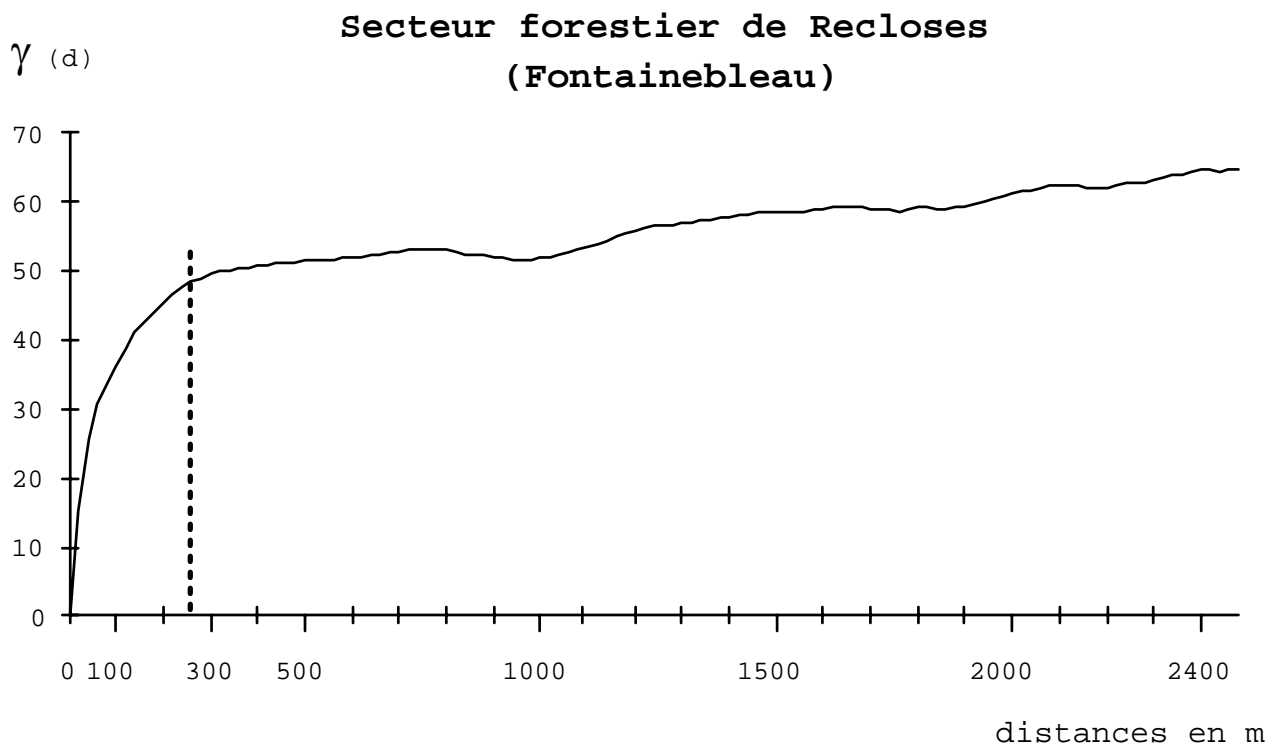
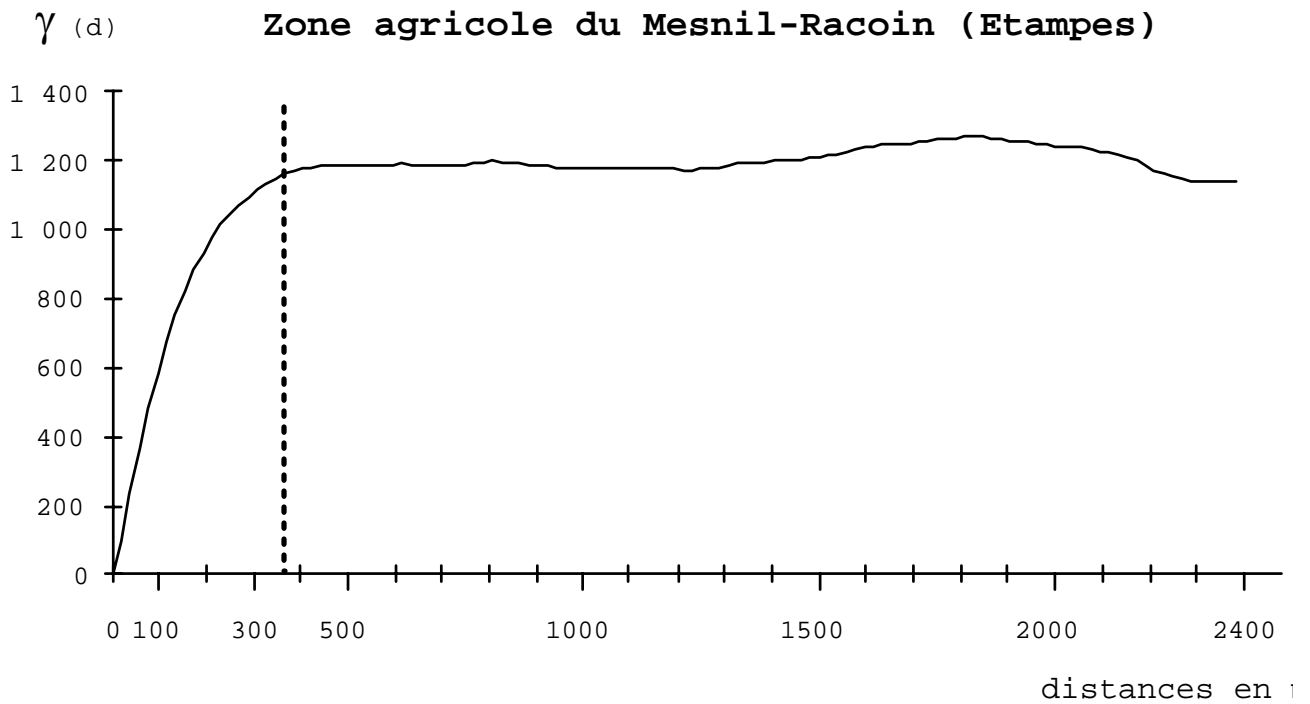
Les milieux sur lesquels ont été réalisés les variogrammes se situent en France au Sud du Bassin Parisien sur quatre sites : deux à dominante agricole et deux à dominante forestière. A cheval sur l'Essonne, le Loiret et la Seine-et-Marne, la zone d'étude fait une trentaine de kilomètres de côté et s'étend du Nord au Sud entre Brétigny et Pithiviers et de l'Ouest à l'Est entre Étampes Fontainebleau. Pour chaque site de cinq kilomètres de côté, nous avons étudié les données numériques des trois canaux bruts du capteur HRV de SPOT (KJ 40-252, 1er mai 1986) ainsi que deux néo-canaux, l'indice de végétation et l'indice de brillance.

Nous avons retenu cette image d'archives pour comparer la taille des segments estimée par analyse variographique, avec celle utilisée par le Service Centrale des Enquêtes et Études Statistiques avec ces mêmes données (FOURNIER 1988). Pour calculer les variogrammes, la dépendance spatiale est analysée sur quatre directions (N-S, E-W, NW-SE et NE-SW) et sur des distances qui n'excèdent pas la moitié des sites d'étude.

Résultats

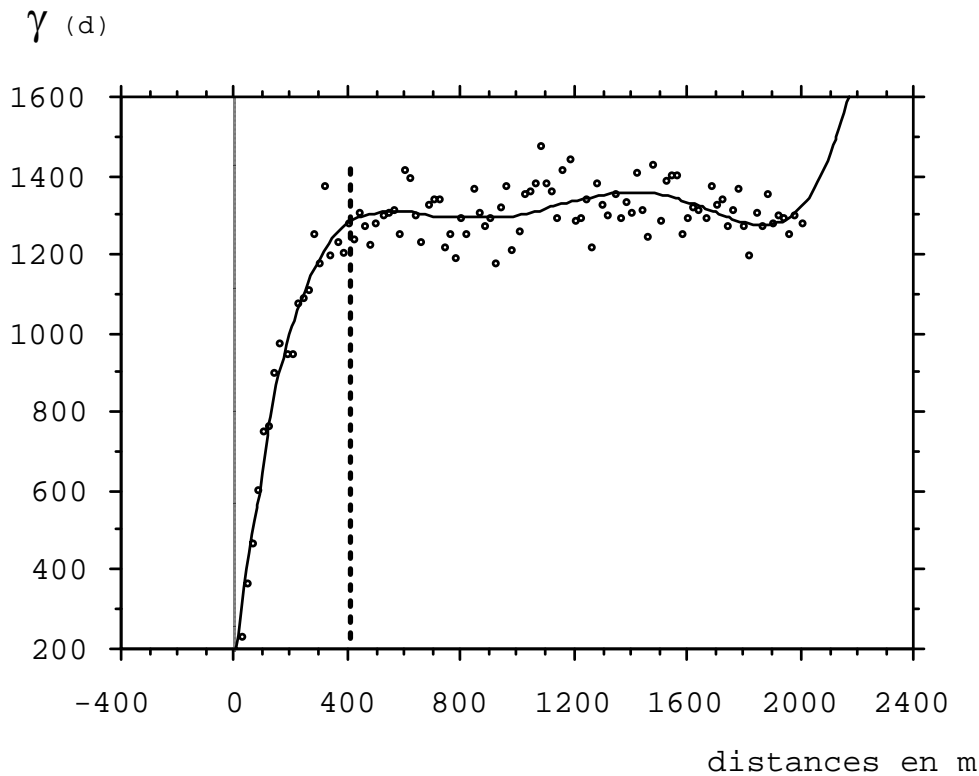
Dans un premier temps, pour chaque site et par canal nous avons réalisé le variogramme moyen des quatre directions. Celui-ci permet d'éliminer rapidement les canaux inutilisables car spatiaux ou au contraire dotés d'une variance continue et indéfiniment croissante (CURRAN 1988) à notre échelle. Une analyse plus fine par direction peut être entreprise si la dépendance spatiale se révèle nettement anisotrope.

figure n°1 - Variogrammes moyens d'une zone agricole et d'un secteur forestier



Moyenne des 4 directions calculée sur l'Indice de Végétation Normalisé

figure n°2 - Variogramme moyen global de la zone d'étude



Moyenne des 4 directions calculée sur l'Indice de Végétation Norm
Ajustement par régression polynomiale d'ordre 5, avec :

$$y = 155,42 + 132,412 * x - 5,66 * x^2 + 0,112 * x^3 - 0,001 * x^4 + 3,535 E-6 * x^5$$

$$R^2 = 0,926$$

Il ressort de l'analyse du variogramme moyen de chaque canal, et pour cette date d'enregistrement, que pour les quatre sites, l'indice de végétation est le canal le plus systématiquement exploitable, car après une croissance rapide à l'origine, la courbe s'infléchit et décrit un plateau (cf. fig. 1). Cela signifie que la diversité radiométrique et donc thématique croît rapidement sur les premières dizaines voire centaines de mètres et qu'ensuite cette diversité s'estompe. Continuer de s'éloigner du point origine n'apporte pas un gain d'informations supplémentaires significatif. La rupture de pente sur le variogramme marque ainsi la dimension moyenne du parcellaire agricole ou des peuplements homogènes forestiers. C'est donc une taille pertinente pour caler le segment.

La détermination de cette rupture de pente (portée du phénomène) est effectuée ici de manière visuelle (cf. fig. 1). Comme on pouvait s'y attendre, les points d'inflexion caractérisant les milieux forestiers et les milieux agricoles sont légèrement décalés. La distance significative est

d'environ 350 m sur les secteurs agricoles, et 250 m sur les secteurs forestiers. Si l'on se réfère à un variogramme moyen global, établi sur l'ensemble de la zone d'étude (30 km de côté), il convient alors de caler le segment sur une distance d'environ 400 m (cf. fig. 2). Cet accroissement de la portée est dû à la prédominance d'un parcellaire agricole de grande taille au niveau de la zone d'enquête.

A la vue de ce dernier résultat, on peut penser que l'inventaire agricole réalisé par le SCEES en 1986 s'est appuyé sur une taille d'unités d'échantillonnage (720 m de côté) nettement surdimensionné (de + de 300m) par rapport à la diversité des paysages. Un "surdimensionnement" des unités d'échantillonnage a également été constaté sur un inventaire d'occupation du sol en milieu naturel tropical (GODARD 1994).

Discussion

Une analyse variographique de secteurs agricoles et forestiers du sud du Bassin Parisien a montré qu'il était possible de caler la taille des unités d'échantillonnage d'un inventaire sur la complexité des paysages étudiés à l'aide des données satellitaires. En même temps, par rapport à d'autres méthodes de détermination de la taille des segments, le variogramme semble minorer systématiquement la taille de l'unité d'échantillonnage.

Actuellement, il ne nous est pas possible de dire si un inventaire dont les segments ont été déterminés par analyse variographique est plus efficace qu'un inventaire dont les segments ont été déterminés par une autre méthode. Les outils d'évaluation, encore en développement, ne sont pas satisfaisants. De plus, il conviendrait de voir si les résultats obtenus sont stables en fonction de différentes dates de prises de vues et de différentes résolutions spatiales.

Bibliographie

- CURRAN (PJ) - 1988 - The Semivariogram in Remote Sensing: An Introduction. *Remote Sensing of Environment*. 24 : 493-507.
- DUPLAT (P.), PERROTTE (G.) - 1981 - *Inventaire et estimation de l'accroissement des peuplements forestiers*. Paris, ONF, 432 p.
- FOURNIER (Ph), GEROYANNIS (H), GILG (JP) - 1988 - Évaluation des surfaces des cultures annuelles et inventaire du territoire dans le Bassin Parisien par combinaison des enquêtes statistiques classiques et des données SPOT (PEPS 127). *SPOT-1. Utilisation des images, Bilan, Résultats*. Paris 23-27 novembre 1987. Toulouse, Cépadués éd. : 251-259.
- GODARD V. - 1994 - Apport de l'analyse variographique pour déterminer la taille et l'espacement des unités d'échantillonnage lors d'un inventaire d'occupation du sol en milieu naturel. *Bulletin de la SFPT*. 94-4 (136) : 33-44.