

Apport de l'analyse statistique multivariée en archéologie : L'implantation des sites (habitats et enceintes) au Néolithique ancien et moyen dans le Bassin parisien

Michèle CHARTIER

(ArScAn – Protohistoire européenne et ArScAn – Archéologie du Bassin Parisien)
(michele.chartier@mae.u-paris10.fr)

L'objectif de cette présentation est de reprendre quelques notions de base en analyse multivariée appliquée à des problématiques archéologiques. Il s'agira de souligner les apports et les limites de la méthode tant en ce qui concerne les procédures que la nature des données faisant l'objet du traitement. Cette présentation s'appuiera sur une étude déjà publiée¹ sur l'existence ou non de différences dans les choix d'implantation des sites entre le Néolithique ancien et moyen dans le Bassin parisien.

1 - CONSTITUTION DE LA BASE DE DONNÉES

La création de la base de données, permettant de répondre à la question posée, suppose de prendre en considération une certaine exhaustivité de l'information même si son corollaire est une généralisation de cette dernière et une moindre précision des descripteurs. L'information doit donc exister pour tous les sites et décrire au mieux

des critères particulièrement difficiles à obtenir car ils ont trait à l'environnement. Des données telles que la nature des sols, la microtopographie des sites, peuvent ne pas être disponibles, ou plus simplement, la localisation précise de certains sites n'est pas toujours enregistrée pour les fouilles anciennes. De ce fait, l'information environnementale est assez restreinte et a imposé le recours à des variables qualitatives telles que : la position topographique au sens large, la proximité d'une zone humide conséquente à moins de 500 mètres, la distance à la rivière la plus proche, la nature de cette rivière : petite, moyenne, grande. Le corpus de sites compte 161 individus au total assez bien répartis sur l'espace retenu avec quelques focales privilégiées ciblées sur les vallées de l'Aisne, de l'Oise ou de l'Yonne (fig. 1).

L'effectif des sites retenus lors de certaines procédures n'a jamais été inférieur à 28. Les descripteurs archéologiques choisis sont : l'appartenance culturelle du site, la fonction du site et la partie de Bassin parisien dont relève le site.

¹ - Chartier 2005.

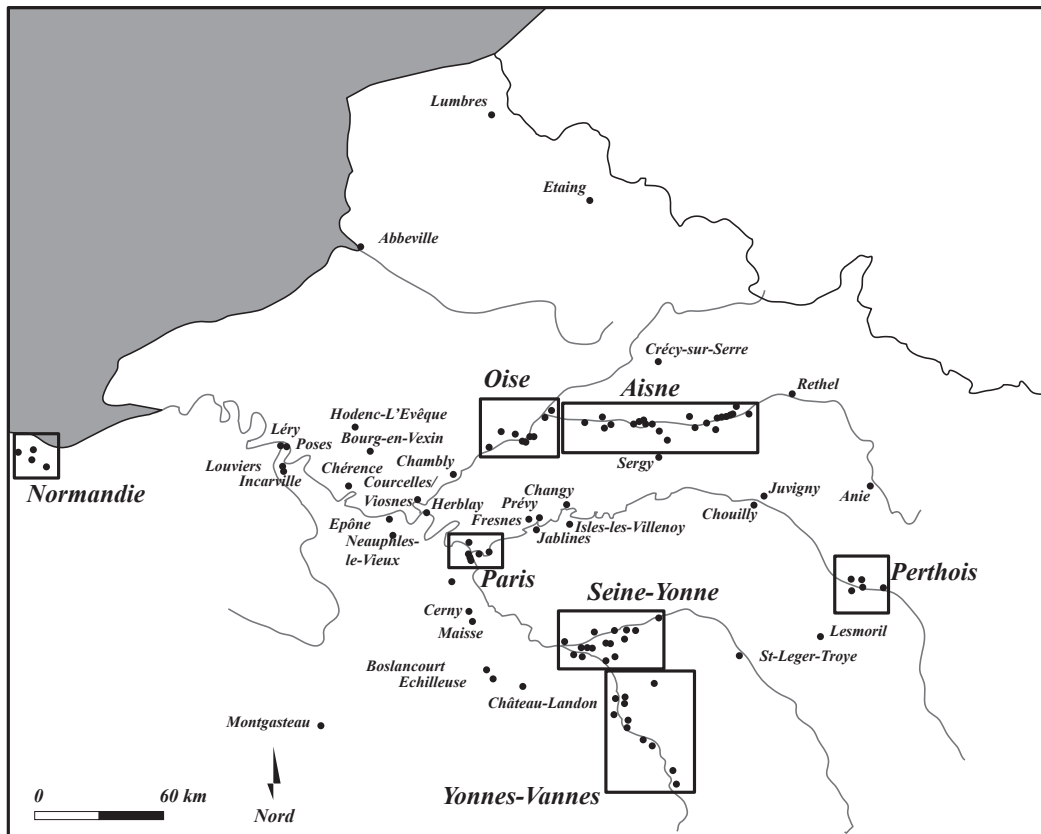


Fig.1 : localisation des sites

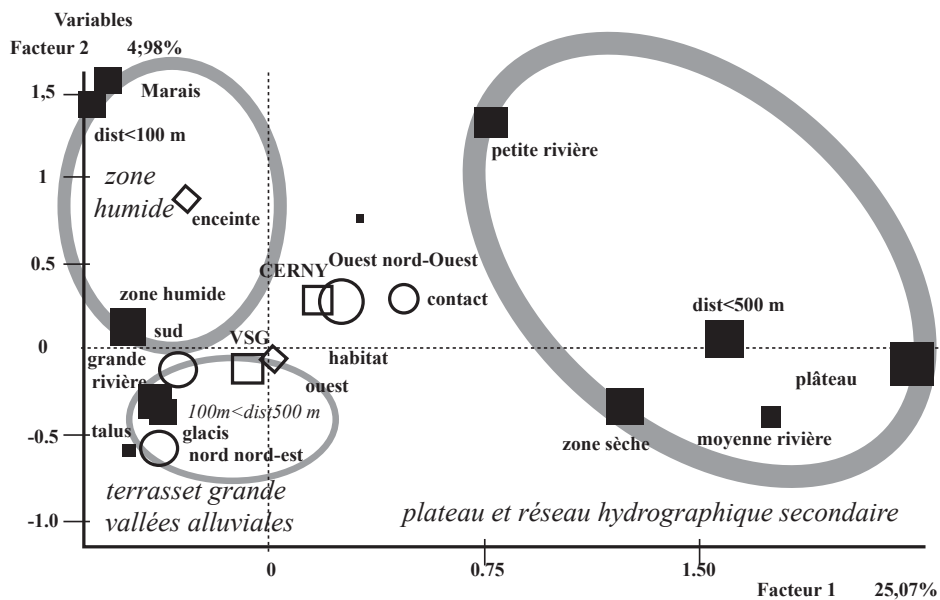


Fig.2 : Analyse du passage Néolithique ancien/moyen (490-4400 B.C.) 4 : Villeneuve-Saint-Germain - 5 : Cerny, (sites)

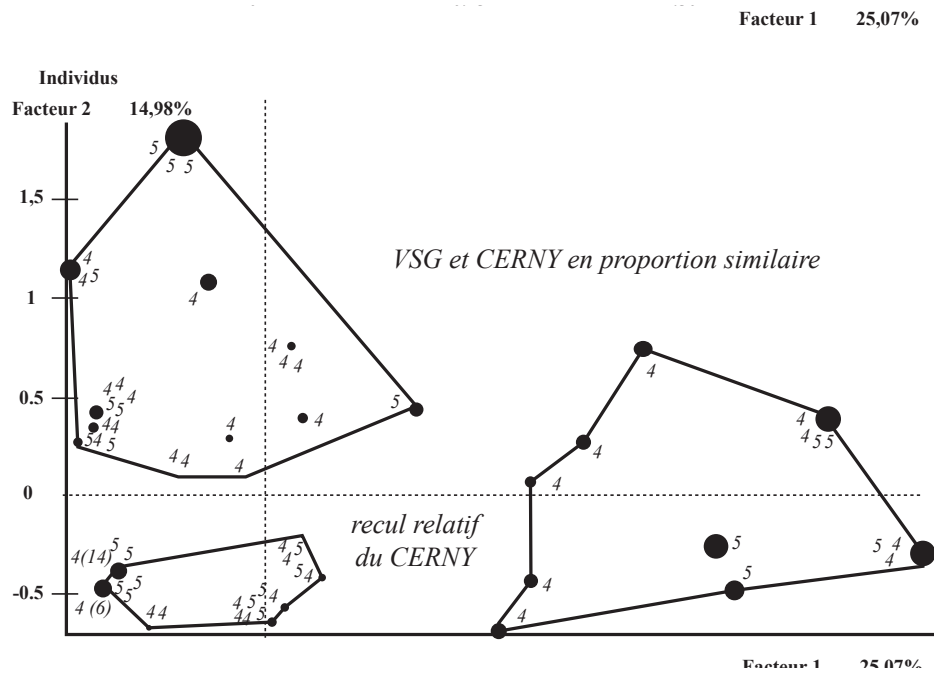


Fig. 2 Analyse du passage Néolithique ancien/moyen (400-4400 B.C.) 4 : Villeneuve-Saint-Germain - 5 : Cerny, (sites)

Variables et individus

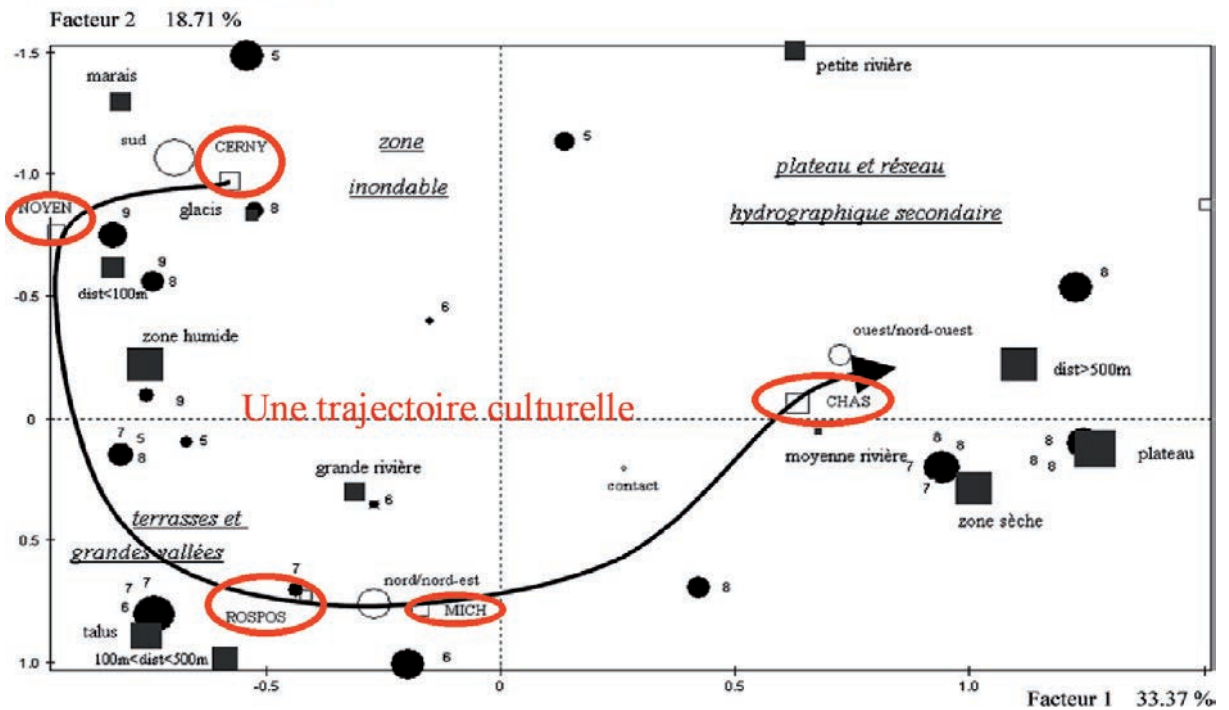


Fig. 3 Analyse des sites du Néolithique moyen (400-300 B.C.) : le cas des enceintes, 5 : Cerny, 6 : Rössen/PostRössen, 7 : Michelsberg, 8 : Chasséen, 9 : Noyen, (sites).

2 - LE CHOIX DE L'ANALYSE FACTORIELLE

L'analyse factorielle est une procédure assez commune qui permet de résumer des tableaux de données importants, de hiérarchiser une information complexe afin de déceler des éléments explicatifs. En simplifiant beaucoup, la procédure consiste à combiner les différentes variables (combinaison linéaire) pour créer de nouveaux axes résumant l'information initiale en supprimant l'information redondante. Le type d'analyse factorielle dépend de la nature des variables retenues. Ici les caractères sont de type qualitatif (les variables quantitatives ayant été transformées en caractères qualitatifs) donc l'analyse factorielle des correspondances multiples (AFCM) a été choisie. La variable « appartenance culturelle » a été traitée comme une variable « illustrative » donc n'intervenant pas dans le calcul des nouveaux axes afin de conserver, à l'interprétation, l'indépendance de la liaison entre cette variable et la caractérisation du milieu.

3 - LE DÉROULEMENT DE L'ANALYSE : UNE APPROCHE CHRONOLOGIQUE DE L'ÉVOLUTION DES IMPLANTATIONS.

L'étude portait sur la séquence du Néolithique ancien et moyen du Bassin parisien.

Deux sorties de l'AFCM sont présentées ici : le passage du Néolithique ancien au Néolithique moyen (fig. 2) et le cas des enceintes du Néolithique

Période	Culture	Sigle - code (AFC)	datation (B.C.)
Néolithique ancien	Rubané moyen champenois	RMC - 1	5200-5100
	Rubané récent du Bassin parisien	RUB - 2	5100-5000
	Rubané final du Bassin parisien	RRF - 3	5000-4900
	Villeneuve-Saint-Germain	VSG - 4	4900-4700
Néolithique moyen	Cerny	Cerny - 5	4700-4400
	Rössen/post-Rössen	RÖSPOS - 6	4400-4200
	Michelsberg	MICH- 7	4200-3500
	Chasséen	CHAS- 8	
	Noyen	Noyen- 9	

Tab eau chronologique

moyen (fig. 3).

Ces analyses montrent que les étapes du peuplement sont complexes avec, en l'état de la documentation, des mouvements de flux et de reflux. La fin du Néolithique ancien (le Cerny surtout) est nettement caractérisée par une diversification des choix, tous les espaces sont investis : zones humides, terrasses, plateaux. Il s'y superpose toutefois des spécificités régionales (la Bassée) et on observe une permanence dans l'occupation des zones humides voire marécageuses ; le potentiel d'implantation à proximité du réseau hydrographique secondaire n'est jamais complètement délaissé. On remarque également la mise en évidence d'une trajectoire culturelle (fig. 3) en ce qui concerne les enceintes du Néolithique moyen. Il semble que des choix d'implantation spécifique y soient inféodés.

CONCLUSION

La pratique de l'analyse multivariée par les contraintes qu'elle impose, en termes d'effectif des corpus, du choix des descripteurs mais aussi par la prise en compte possible de l'information de nature qualitative, est particulièrement intéressante pour les problématiques archéologiques. Toutefois les corpus encore trop lacunaires ou certains biais introduits par les conditions de fouilles en interdisent (ou devraient en interdire) parfois l'usage. De telles analyses n'apportent pas, en général, de résultats révolutionnaires mais sont souvent une aide utile à une meilleure connaissance des tableaux d'information dont on dispose. En effet, ces derniers sont souvent hétérogènes avec de l'information redondante qui masque l'information essentielle. Ces analyses permettent aussi la mise en évidence de quelques « originalités » ou épiphénomènes qui ne sont pas toujours perceptibles au sein du tableau initial mais qu'il est cependant souhaitable d'exploiter

et d'interpréter à leur juste place. On touche là à une difficulté rencontrée lors de l'exploitation des résultats de telles analyses. Il s'agit de la lecture des sorties graphiques où une part non négligeable de subjectivité peut parfois entrer en ligne de compte dans la détermination des regroupements d'individus, dans l'interprétation des associations de variables ou dans la définition et la signification des nouveaux axes créés.

BIBLIOGRAPHIE

- CHARTIER M. 2005. L'analyse statistique des implantations des sites du Néolithique ancien et moyen dans le Bassin parisien : l'apport de quelques critères spatiaux. *Bulletin de la Société préhistorique française* 102, 2 : 269-280.
- DUBOULOZ J., BOSTYN F., CHARTIER M., COTTIAUX C. et LE BOLLOCH M. 2005. La recherche archéologique sur le Néolithique en Picardie. *Revue Archéologique de Picardie* 3-4 : 63-98.
- PERNAUD J.-M., CHARTIER M. et TRESSET A. (avec la collab. de Augereau A., Leroyer Ch., Sidéra I., Thiébault S.) 2004. Gestion des territoires et évolution de l'exploitation des ressources animales et végétales aux Ve et IVe millénaires en Bassin Parisien : bilan des connaissances et perspectives de travail. In : BODU P. et CONSTANTIN Cl. (ed.), *Approches fonctionnelles en Préhistoire, XXVe Congrès préhistorique de France* : 409-425. Paris : Société préhistorique française.
- SANDERS L. 1990. *L'analyse statistique des données en géographie*. Montpellier : GIP Reclus, collection Alidade.