Chapitre 2.2.8
Contextes et interactions spatiales
Analyser le système de peuplement

Laure NUNINGER(1), Hélène MATHIAN(2)

(1) Chrono-Environnement UMR 6249, MSHE C.N. Ledoux UAR 3124, CNRS, Université Marie et Louis Pasteur, Besançon, (2) Environnement Ville Société - EVS UMR 5600, CNRS, Université de Lyon, Lyon

Archéologues et historiens cherchent à caractériser les différentes formes d'habitat et leur organisation spatiale pour comprendre les choix d'implantation des communautés du passé. Compris dans la longue durée, l'enjeu consiste à mieux appréhender les formes d'appropriation de l'espace, les modes d'organisation socio-économique et de structuration politique et culturelle.

Les chapitres 2.2.4 et 2.2.5 ont respectivement montré, comment on pouvait répondre à la question « où » en travaillant d'une part sur la distribution spatiale* d'un semis de points, d'autre part en s'attachant à l'analyse spatiale* des critères permettant de « prédire » l'existence de cette localisation dans un environnement donné.

Ces analyses de la localisation et des configurations spatiales rendent compte d'un état du peuplement* à un moment donné, ou à plusieurs états successifs, soit de la manière dont l'homme dispose de son espace et comment il l'occupe. Dans ce cadre conceptuel, les lieux occupés ne sont pas considérés de façon autonome mais dans leurs localisations relatives que ce soit entre établissements ou par rapport à des systèmes de ressources ou de leurs accès. Néanmoins, l'analyse des configurations spatiales décrit les espacements entre les lieux, en se fondant uniquement sur les localisations (*cf.* 2.2.4.) sans les mettre en relation avec les caractéristiques* de ces lieux qui influent à l'évidence sur ces localisations. Aussi, pour faire des hypothèses sur les types de relations qui ont conduit à certaines formes de configurations spatiales, il est nécessaire d'introduire les caractéristiques des lieux et des populations qui y sont localisées et qui sont les moteurs de ces relations dans un système. Dans ce

système chaque lieu occupé peut avoir une ou plusieurs fonctions socio-économiques (lieu de pèlerinage ou marché par exemple) leur conférant un rôle plus structurant pour le territoire que d'autres (une ville par rapport à une ferme isolée par exemple) et induisant des relations entre les lieux, hiérarchiques ou non.

Considéré dans le temps long, ce que l'on appelle donc le système de peuplement peut alors se définir comme « l'ensemble des lieux occupés d'un territoire ayant entre eux des relations évolutives d'interdépendances » (Durand-Dastès *et al.* 1998). Pour appréhender ce système d'un point de vue archéologique, l'absence ou quasi-absence de sources textuelles, nous amène à composer seulement avec les vestiges matériels géolocalisés. Ces vestiges collectés soit en fouille, soit en surface grâce à des prospections pédestres nous permettent de caractériser les lieux et la nature de leurs relations potentielles.

Les questions abordées dans ce chapitre sont donc comment peut-on rendre compte de l'organisation d'un ensemble de lieux occupés en réseaux hiérarchiques ? Ce qui nécessite de s'interroger sur les relations potentielles entre les lieux occupés, soit comment les identifier de façon homogène et systématique ?

On s'appuiera sur l'expérience menée en Vaunage (Gard) dans le cadre du programme Archaeomedes[1] pour illustrer la démarche qui consiste d'une part à identifier la diversité fonctionnelle des localisations, ici des établissements* (Figure 1), pour ensuite modéliser les relations potentielles entre les établissements compte tenu de leurs fonctions et de leurs localisations relatives. La démarche qui est proposée ici s'appuie sur un modèle théorique d'évaluation de la portée de l'influence d'un lieu en fonction de son importance. Les réseaux hiérarchiques résultent de ces jeux d'influence sur le territoire qui dessinent des aires d'influence. Avant d'aller plus loin, il convient également de préciser que dans cette approche les sites sont considérés comme contemporains [2].

Dans notre cas d'étude sur la Vaunage, le choix va être motivé par les hypothèses suivantes : 1) les établissements les plus « importants », les agglomérations par exemple, jouent un rôle de pôle[3], ils ont une capacité d'influence, et se trouvent à la tête d'un réseau hiérarchisé de peuplement ; 2) la situation relative des pôles et des autres établissements est telle qu'il est aisé de se rendre à ce pôle-tête de réseau avec les moyens de déplacement de l'époque considérée.

Ainsi, dans une première étape, il s'agit d'être capable de hiérarchiser les établissements et d'identifier quels sont les établissements qui ont potentiellement joué un rôle de tête de réseau. La seconde étape permettra d'estimer les liens entre les établissements pour construire les réseaux et associer aux têtes de réseau l'ensemble des établissements, qu'ils polarisent et qui constituent son « réseau » d'influence.

[1] Archaeomedes II (1996-1999) : « Policy relevant models of the natural and anthropogenic dynamics of degradation and desertification and their spatio-temporal manifestations » (Program Environment of the European Commission DGXII), coordination Sander van der Leeuw (University of Cambridge). Pour en savoir plus : Durand-Dastès *et al.* 1998, Nuninger *et al.* 2006, Nuninger *et al.* 2012a, Favory, Nuninger et Sanders *et al.* 2012.

[2] Sur la notion de contemporanéité, *cf.* 2.2.2

[3] La notion de pôle en géographie fait référence à un lieu vers lequel converge un champ d'attraction (Brunet *et al.* 1993). Dans le cas des établissements, on fait l'hypothèse que les fonctions (économiques, administrative, religieuse…) sont un facteur d'attraction et polarisent ainsi les zones de peuplement.

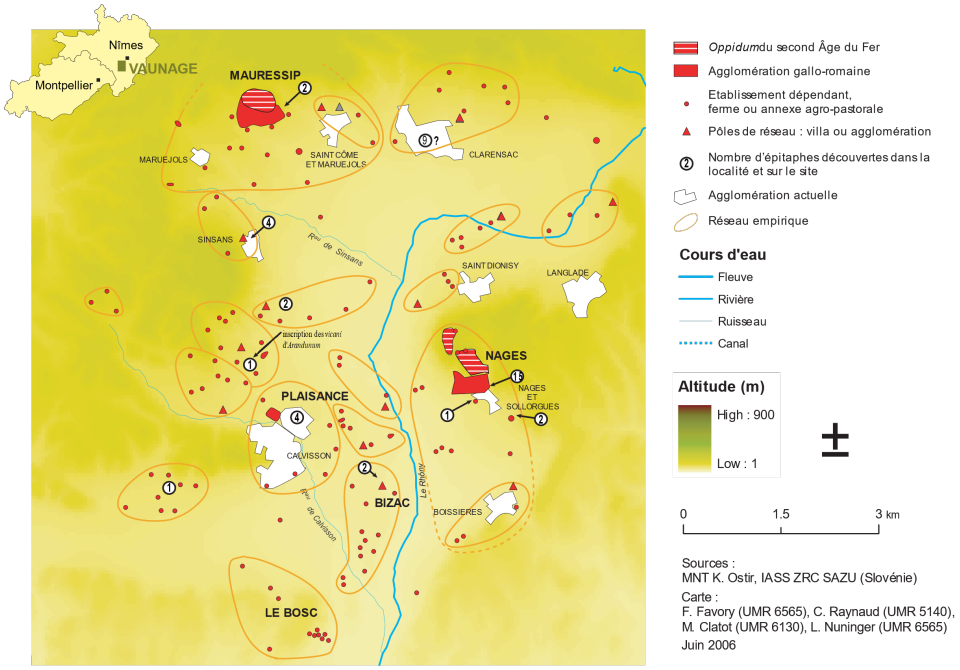


FIG. 1 : Carte de localisation de la Vaunage (Gard), sites archéologiques du 1er siècle de n.è. et réseaux d'influence définis par les archéologues (d'après C. Raynaud et F. Favory).

1 Comment différencier les établissements et pourquoi faire ?

Pour étudier les interactions entre deux lieux occupés, il faut d'abord pouvoir les comparer selon les caractéristiques qu'ils présentent. C'est un exercice courant pour l'archéologue qui établit des typologies afin d'interpréter un ensemble de données. Néanmoins, notre objectif étant d'analyser la relation d'influence qui s'exerce entre chaque établissement et les autres, la typologie doit être pensée afin d'ordonner les établissements ou les types d'établissements les uns par rapport aux autres (*cf.* 2.2.2). Avant de s'interroger sur la manière de différencier les lieux occupés ou sur les outils à utiliser pour le faire, il convient d'abord de se demander à quelle question l'on souhaite pouvoir répondre en distinguant ces lieux selon leur degré potentiel d'influence. S'agit-il d'identifier les lieux selon leur potentiel économique en fonction d'activités spécifiques ? Ou s'agit-il d'ordonner les lieux selon leur niveau hiérarchique ou leur capacité d'attraction de la population par exemple ? Selon les cas, les critères retenus pour différencier les établissements ne seront pas forcément les mêmes. Ainsi par exemple si la forme de la structure d'habitat (ronde, carrée ou ovale) peut jouer un rôle dans la détermination fonctionnelle du lieu occupé, il est peu probable que ce critère soit pertinent pour distinguer des niveaux hiérarchiques...De même, si la question est de hiérarchiser les lieux occupés les uns par rapport aux autres pour comprendre la manière dont une société structure ses relations dans un espace donné, alors le fait d'ordonner un lieu occupé par des vivants avec un lieu occupé par des morts sera

sans doute difficile à interpréter. On ne retiendra donc que les établissements (occupés par des personnes vivantes) et on utilisera les sites funéraires pour qualifier ces établissements (associés ou non à une nécropole à une distance maximale donnée, par exemple). La question des critères à considérer est donc liée au sens que l'on donne aux objets à analyser.

L'évaluation des différenciations entre les établissements va se faire sur la base d'une description multidimensionnelle des établissements, c'est-à-dire d'un ensemble de variables. C'est une construction relative au corpus étudié. On notera que le classement[4] peut se faire à partir de variables quantitatives* ou catégorielles* (qualitatives) (Figure 2).

Dans tous les cas la démarche de classification* peut se faire en plusieurs étapes :

1- Dans une phase d'exploration des données disponibles pour faire une première sélection des descripteurs* et de leurs modalités* qui devront concerner un nombre significatif d'établissements. Par exemple, si pour le descripteur superficie de l'établissement la modalité « supérieure à 10 ha » ne concerne que 2 des 100 établissements analysés (soit 2%), ce critère n'est statistiquement pas significatif. Le fait d'avoir cette modalité devient une particularité, comparée aux autres valeurs qui peuvent avoir des poids beaucoup plus importants (comme 20 ou 30 % des établissements). Elle risque de constituer la différenciation principale entre les établissements, et donc de perturber l'analyse sans apporter de connaissance supplémentaire puisque l'on connaît déjà la spécificité de ces établissements. Dans cette phase exploratoire, il s'agira donc de discuter les valeurs des descripteurs, qu'ils soient nominaux (type de matériau, par exemple) ou ordinaux provenant d'une discrétisation* (en classes de superficie, par exemple), tout en veillant à garder une cohérence thématique dans la manière dont les modalités sont définies.

2- Dans une phase d'analyse statistique, où l'on va tester le niveau de corrélation des descripteurs sélectionnés, ce qui permet de comprendre quelles relations statistiques existent entre les variables. De manière générale, il est préférable de travailler avec un ensemble de variables les moins corrélées possibles. Par exemple, si nous étudions une population d'arbres, les critères âge et hauteur ont, par hypothèse, des chances d'être corrélés tandis que la forme et la couleur des feuilles le sont sans doute moins. Si l'on introduit deux variables très corrélées, on donne du poids à la dimension qu'elles représentent. Toutefois, on peut s'autoriser cette redondance pour mettre en évidence une hiérarchie. On donne ainsi un sens hiérarchique aux classes ou mesures synthétiques qui seront issues de l'analyse multivariée*. Il faudra dans tous les cas que cette structure de relation entre les variables soit parfaitement connue pour lui donner un sens. Pour déterminer statistiquement la relation de dépendance ou d'indépendance*,

[4] Le classement vise à identifier des groupes d'établissements, groupes qui se distinguent le plus possible les uns des autres. Le classement n'est pas forcément hiérarchique, il peut s'agir de simples catégories d'établissements.

entre les variables deux à deux, on pourra utiliser différents traitements, tels que l'analyse de corrélations* (*cf.* 5.) ou l'analyse du Chi2* (*cf.* 7.) selon qu'il s'agit d'un caractère quantitatif ou qualitatif des données en entrée ou encore la nature de la relation que l'on veut tester.

3- Enfin, dans une phase d'analyses multivariées exploratoires qui va permettre de sélectionner les variables actives et illustratives (*cf.* 6.). L'analyse finale sera ensuite utilisée pour réaliser la classification hiérarchique, au sens statistique, des établissements (*cf.* 9.).

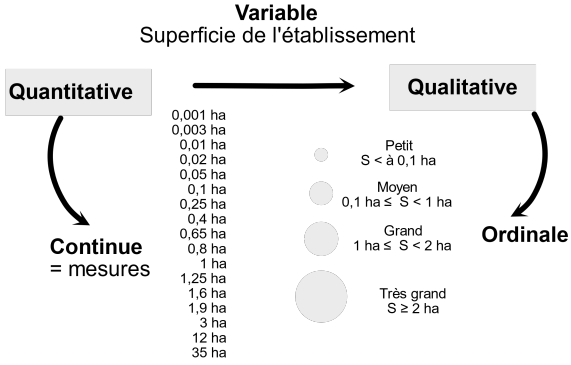


FIG. 2 : Exemple de la variable Superficie de l'établissement sous sa forme quantitative ou qualitative (catégorielle)

Dans la première expérience menée sur la Vaunage, deux types d'analyses multivariées ont été mobilisées et confrontées à partir d'un jeu de données archéologiques couvrant l'Antiquité et le Moyen-Âge (Durand-Dastès *et al.* 1998 ; Nuninger *et al.* 2006). Chacune des deux analyses aboutit à une hiérarchisation des établissements, mais de deux manières différentes.

Dans le premier cas, la structure hiérarchique est recherchée et posée en amont de l'analyse. On mobilise une analyse en composante principale* (ACP, *cf.* 6.) appliquée à un ensemble de descripteurs qui par effet cumulatif produit un indicateur quantitatif mesurant cette hiérarchie.

Dans le deuxième cas, une classification ascendante hiérarchique* (CAH, *cf.* 9.), réalisée sur les résultats d'une analyse des correspondances multiples* (ACM, *cf.* 6.), produit des classes ou catégories (modalités d'une variable qualitative) qui sont ordonnées *a posteriori* une fois les classes interprétées : la structure hiérarchique émerge de l'analyse. Ces analyses ont été menées sur un jeu de données, comprenant 679 individus statistiques (établissements archéologiques) dont la description est structurée par 16 descripteurs regroupés en 4 grandes catégories : technique et fonctionnelle, chronologique, sitologique et situationnelle)[5] (Figure 3).

[5]Pour une description précise *cf.* Durand-Dastès *et al.* 1998, p. 158 et p. 255-258.

Catégorie de descripteurs	Descripteurs	Modalités
Techniques et fonctionnels	Matériaux de construction	absence de matériaux (architecture légère)
		terre crue et tuiles
		moellons, moellons et tuiles
		plus mortier, tuileau, brique, enduit peint
		plus mosaïque, éléments de décor (corniche, moulure)
	Mobilier	3 modalités
	Activité	3 modalités
	Fonction	4 modalités
	Superficie	7 modalités
Chronologiques	Occupation antérieure et legs	non renseignée
		aucune
		dans le siècle précédant la création de l'établissement
		antérieur au siècle précédant la création de l'établissement
	Durée d'occupation	8 modalités
	Date d'implantation	13 modalités
	Statut au XVIIIe (Carte de Cassini)	8 modalités
	Statut actuel	7 modalités
Sitologiques	Topographie	Plaine
		Littoral
		Coteau
		Vallée alluviale
		Sommet
	Sol	9 modalités
	Altitude	
Situationnels	Distance à la voirie	traversé par une voie
		au bord d'une voie
		à moins de 100 m d'une voie
		à plus de 100 m d'une voie
	Nombre de chemins menant au site	5 modalités
	Nombre de liaisons avec les établissements contemporains	4 modalités

Fig. 3 : Exemples de catégories, descripteurs (variables) et modalités (d'après Durand-Dastès *et al.* 1998)

Pour la première analyse, afin de produire un indicateur synthétique de hiérarchie, les descripteurs techniques et fonctionnels (superficie, la qualité des matériaux, la qualité du mobilier, et le statut présumé en fonction des traces de sépulture trouvés) ont été mobilisés[6], la combinaison de cet ensemble de variables instituant une relation d'ordre hiérarchique entre les établissements. Chacune de ces variables a été enregistrée sous une forme ordinale (variable ordinale*), c'est-à-dire que les modalités prennent des valeurs de 1 (niveau le plus faible) à 4 (niveau le plus élevé). Si les critères sont bien choisis et traduisent chacun un aspect hiérarchique, la première composante associée à l'analyse du tableau permet de traduire la corrélation entre les différents critères. Les coordonnées des établissements sur cette composante dite principale, définissent un indicateur synthétique du niveau hiérarchique des établissements (lié à ce que l'on appelle en analyse des données « l'effet de masse »). Une note surdimensionnée relativement à un des critères, est ainsi le plus souvent compensée par des valeurs faibles sur les autres s'il ne s'agit pas véritablement d'un établissement de niveau hiérarchique important. Cette première analyse permettra d'approcher dans un deuxième temps la structure spatiale. Mais ce que l'on obtient avec l'ACP ne fait que décrire quantitativement une structure hiérarchique, sans aborder les spécificités fonctionnelles des établissements d'un point de vue qualitatif.

Le but de la seconde analyse est donc de fournir une typologie des établissements sur la base d'un ensemble plus complet de critères de façon à catégoriser les établissements. Seize descripteurs à la fois techniques et fonctionnels, chronologiques, sitologiques et situationnels ont été retenus pris sous leur forme qualitative nominale cette fois. Ainsi, chacune des modalités correspond à une catégorie* sans qu'il y ait de relation d'ordre *a priori*, d'un point de vue statistique. Par exemple pour la variable superficie,

[6] Ce choix résulte notamment d'une première analyse de l'ensemble de tous les descripteurs par rapport à leur capacité à durée, l'hypothèse étant que la capacité d'un établissement à durée sur le temps long est d'autant plus grande que l'établissement était important dans le passé. Par cette analyse, la sélection des critères les plus discriminants a pu être opérée (Durand-Dastès *et al.* 1998).

chaque modalité correspond à une fourchette de valeur de superficie (< 0,1 ha, 0,1 et < 0,3 ha,...< 2 ha). D'un point de vue sémantique, ces modalités sont ordonnées, mais bien qu'elles aient des amplitudes différentes, elles auront le même poids dans l'analyse, car il s'agit de catégories. L'ensemble des modalités de tous les descripteurs constitue un tableau de critères qui est soumis à une Analyse des Correspondance Multiple (ACM) suivie d'une classification ascendante hiérarchique (CAH) sur les trois premiers facteurs*. Le résultat est un ensemble de classes ou types. Chaque classe va être décrite par un profil donnant les caractéristiques des établissements la composant, et qui vont permettre de les caractériser du point de vue thématique en définissant des types d'établissements. Dans notre exemple, six types d'établissements ont pu être identifiés, allant de l'annexe agro-pastorale ou aire technique à la capitale et/ou chefs-lieux majeurs. Les classes s'ordonnent donc *a posteriori* relativement au sens donné aux classes pour rendre compte de la hiérarchie des établissements (Figure 4).

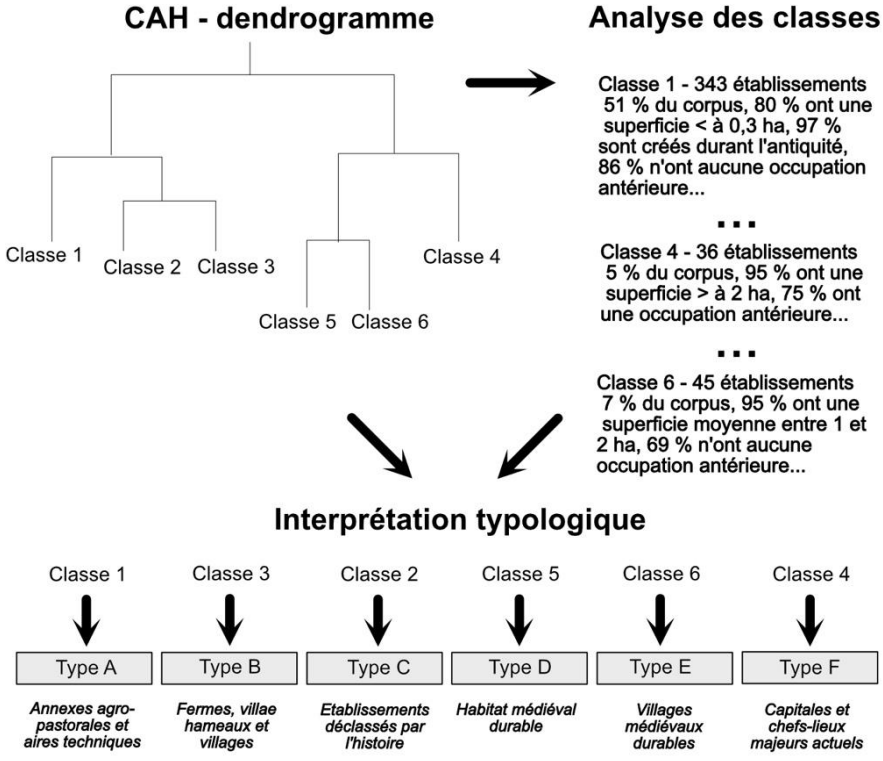


Fig. 4 : Exemple de chaine de traitement des résultats de la classification (CAH) du dendrogramme à l'interprétation.

Dans le cas où les critères s'ordonnent naturellement et structurent le premier axe factoriel, il est également possible d'utiliser les coordonnées de la projection des établissements sur l'axe 1 comme indicateur quantitatif du poids respectif des établissements les uns par rapport aux autres, à l'instar de la composante principale de l'ACP. Cette méthode est très réductrice et sans doute moins robuste sur le plan statistique mais elle permet d'estimer une valeur à partir d'une information initialement qualitative. C'est le choix qui a été retenu pour l'étude des réseaux de l'Age du Fer sur la même région (Nuninger 2002 ; Nuninger *et al.* 2006). Dans cette dernière analyse, parmi les descripteurs techniques et fonctionnels, seuls la superficie, les matériaux de construction et la fonction (ou statut symbolique et idéologique) ont été intégrés avec l'ajout d'un quatrième critère : la durée d'occupation. Considéré comme un descripteur chronologique dans la classification de la première expérience sur l'Antiquité et le Moyen-Âge, il a été utilisé ici, comme une variable fonctionnelle et intemporelle, avec l'attribution d'une quantité de temps à chaque établissement comme la superficie indique une quantité d'espace occupé. Sachant que ces deux critères sont indépendants, on fait l'hypothèse ensuite que plus la quantité de temps et/ou d'espace est élevée, plus l'établissement aura des chances d'être hiérarchiquement important.

À ce stade, il est important de noter que les classifications obtenues sont des modèles construits à partir de démarches différentes. Ainsi, si globalement les résultats sont cohérents, cela permet de valider le modèle hiérarchique. Toutefois, il est également possible qu'un établissement hiérarchiquement faible dans la première analyse (ACP) se retrouve dans une classe de niveau hiérarchique élevé dans la seconde analyse (ACM/CAH). Ces cas particuliers révélés par l'analyse sont très intéressants car ils invitent l'archéologue à réexaminer le profil de ces établissements atypiques, soit pour confirmer leur position atypique (liée par exemple à une fonction très particulière – un relais routier par exemple), soit pour revoir les critères descriptifs qui lui sont associés (erreur ou incertitude* sur les données en entrée).

2 Comment estimer les liens entre établissements ?

Une fois établie la valeur hiérarchique des établissements les uns par rapport aux autres, il s'agit d'estimer les relations hiérarchiques potentielles qu'ils entretiennent. La valeur et la nature d'un lien entre deux lieux, est rarement une donnée statistique directement mesurable sur le terrain. Il existe des cas où l'on peut quantifier par exemple des pièces dont on présume par ailleurs l'origine. On peut alors assimiler le lien entre ce lieu et le lieu d'origine à partir de la quantité de pièces observées en ce premier lieu qui peut être vu comme un « flux » entre les deux lieux. Mais ces observations sont rares et ne couvrent pas l'ensemble des relations. Si cette mesure ne peut être produite à partir des observations issues de la collecte de terrain, on peut utiliser un modèle, c'est-à-dire une fonction générique, pour les produire, il s'agira alors de mesures théoriques ou simulées. Ici il s'agit d'un « modèle d'interaction* spatiale » qui à partir des caractéristiques de chacun des deux lieux, et de leur localisation permet si ce n'est de calculer un « échange » théorique, d'évaluer un lien potentiel entre les deux lieux. On obtient ainsi des valeurs calculées « théoriques » et non « observées ». Le calcul est réalisé à partir d'une fonction mathématique (*cf.* ci-dessous) qui comprend des paramètres, qu'il s'agit d'adapter à l'étude, c'est-à-dire qui permettent de calibrer la

fonction, soit sur la base d'observations empiriques, soit à partir de considérations théoriques. La fonction comprend également des variables qui correspondent aux valeurs des caractéristiques mobilisées dans la fonction pour chaque entité considérée (par exemple leur poids hiérarchique relatif ou la distance qui les sépare des autres lieux).

Afin d'illustrer notre propos, nous présenterons le modèle utilisé pour construire des réseaux locaux d'habitat sur la micro-région de la Vaunage. Il est inspiré d'un modèle générique appelé « modèle de potentiel » (Pumain, Saint-Julien 2001) qui permet d'évaluer l'intensité de l'attraction/influence d'un lieu sur un autre en fonction du niveau hiérarchique de ce lieu et de l'espacement entre les deux lieux. L'étape précédente a permis de construire un indicateur synthétique permettant de différencier les établissements de deux points de vue : celui quantifiant l'importance (ACP sur variables ordinales), celui qualifiant puis hiérarchisant le caractère fonctionnel (ACM sur variables qualitatives puis interprétation). Nous présenterons ici uniquement l'articulation entre le questionnement formulé à chaque étape du modèle et les outils statistiques utilisés, l'analyse proprement dite et les résultats ayant déjà été publiés (Durand-Dastès *et al.* 1998 ; Nuninger *et al.* 2006).

Mathématiquement, ce modèle s'exprime sous la forme suivante :

$$F(_{i,j}) = P_i / aD_{ij}{}^{b},$$

Où l'influence $F(_{i,j})$ d'un lieu *i* sur un lieu *j* est fonction de l'importance ou poids mesuré par P_i et est inversement proportionnelle au carré de la distance D_{ij} qui le sépare du lieu *j,* les paramètres *a* et *b* décrivent la forme de décroissance de l'influence en fonction de la distance. En particulier *b* exprime le « frein » de la distance : une valeur faible diminue l'effet de la distance alors qu'une valeur élevée en renforce l'effet.

Dans notre exemple, le « poids » de l'établissement *i* est mesuré par l'indicateur hiérarchique issu de l'ACP présentée précédemment. Chaque établissement est relié à l'établissement *i* qui exerce sur lui l'influence la plus forte, influence calculée selon la formule précédente. Le calibrage du modèle, c'est-à-dire l'estimation des valeurs de *a* et *b*, a été effectué à partir d'un « réseau témoin » connu des archéologues (Figure 1) : c'est-à-dire qu'on a testé différentes valeurs des paramètres *a* et *b* pour trouver ceux qui permettaient de détecter ce réseau (Figure 5). Il est important de noter, qu'il existe différentes méthodes pour calibrer le modèle, aussi dans une autre expérience, sur la même région, les paramètres *a* et *b*, ont été défini en fonction de l'emprise visuelle (Figure 6), à partir de l'hypothèse qu'un pôle doit pouvoir contrôler l'ensemble de son territoire du regard[7] (Nuninger *et al.* 2006).

Dans ces exemples, le résultat est que chaque lieu est relié à un et un seul établissement. Cela permet d'identifier les pôles et leur réseau. On peut aller jusqu'à tracer les aires d'influences associées autour des réseaux constitués. Toutefois, cette dernière

[7] Chaque établissement j est "relié" à l'établissement i qui a la plus grande "influence" sur lui. L'influence de i sur j est mesurée par F(i,j) en fonction de son poids hiérarchique (Hi) et de la distance D(i,j) qui les sépare. Plus les établissements sont proches, plus le poids est fort, plus ils sont éloignés, moins i a d'influence sur j. En revanche cette influence s'annule si j n'est pas "à vue" de i.

méthode produit des frontières rigides entre les aires d'influences, or, une aire d'influence est une zone floue et un lieu peut être attiré par plusieurs zones d'influences, notamment quand il est localisé en périphérie. Pour pallier cette difficulté, on peut s'inspirer de l'approche probabiliste proposée par D. L. Huff (1964) et interpréter l'influence de i sur j $F(i,j)$ comme une probabilité que j appartienne au réseau de i. Ainsi l'établissement j peut être rattaché à plusieurs têtes de réseaux avec certaines probabilités.

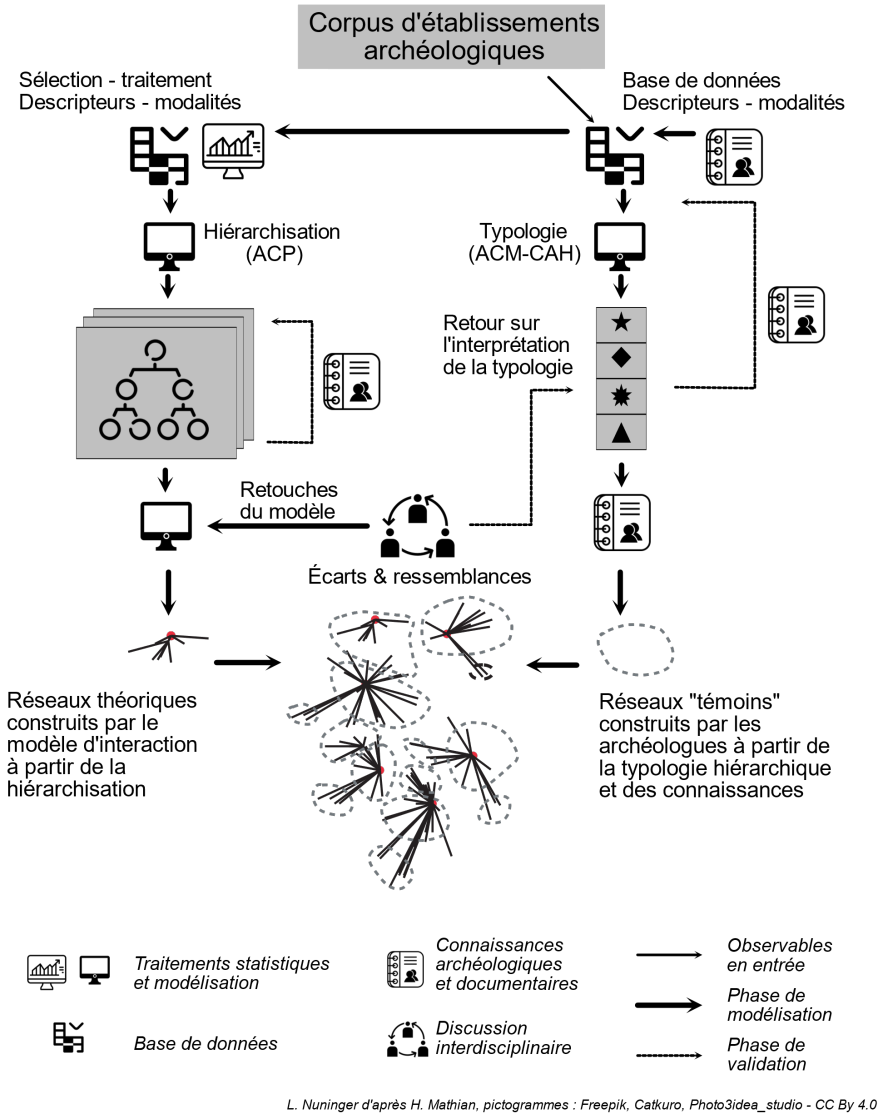
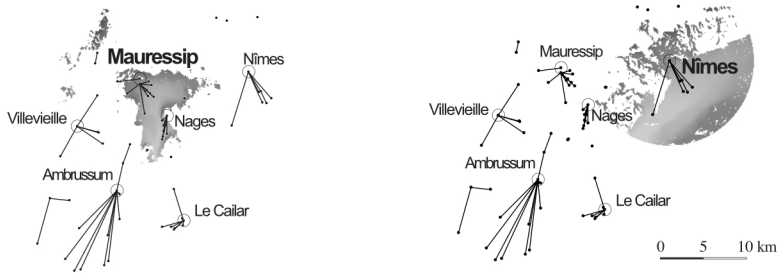


Fig. 5 : Modélisation des réseaux d'établissements, confrontation de deux approches dans une démarche conjointe (d'après Durand-Dastès *et al.* 1998)



Pour un établissement j donné, il s'agit de rechercher le pôle i qui maximise F(i,j) en fonction de son poids hiérarchique (Hi) et de la distance (Dij). Une fois celui-ci déterminé, la condition de la maîtrise visuelle est testée. Si l'établissement se trouve dans le champ visuel du pôle, il est affecté au réseau d'habitat de ce pôle. Dans le cas contraire il en est exclu, et l'influence du pôle est considérée comme nulle. L'influence de i sur j est donc évaluée suivant la formule suivante :

$$Fij = [Hi / (1 + a\ Dij^2)] .EV(i,j)$$

où EV(i,j) vaut 1 si l'établissement j est sous le contrôle visuel de l'établissement i et 0 sinon.

FIG. 6 : Paramètrage de la fonction d'interaction par la maitrise visuelle : exemple des pôles de Nîmes et Mauressip (Gard) pour le 2e siècle av. n.è. (d'après Nuninger *et al.* 2006).

Les modèles d'interaction peuvent être adaptés à différents phénomènes d'attraction. Dans les phénomènes étudiés en archéologie, les distributions des intensités d'un phénomène suivent rarement une décroissance linéaire[8]. Par exemple, autour d'une zone d'émission de monnaie, la densité de monnaie est en général anormalement élevée à proximité du centre émetteur, tandis que cette densité dans la périphérie est plus faiblement représentée. Dans ce cas, la courbe de décroissance de la densité n'est pas linéaire mais adopte une forme en cloche selon la fonction gaussienne* (*cf.* 4.). Si on prend l'exemple du peuplement, il est fréquent d'observer que les habitats ruraux ne s'installent pas trop près des agglomérations, dont la périphérie immédiate est généralement exploitée directement par la population de l'agglomération. Dans ce cas, le pouvoir d'attraction de l'agglomération chute donc rapidement à proximité de l'agglomération, puis plus lentement au fur et à mesure que l'on s'en éloigne, décrivant une courbe concave. Cela se traduit en théorie par une fonction de type puissance négative par exemple (Figure 7). Dans le cas des haches (*cf.* 2.2.7) les analyses de distributions distances par rapport à un lieu origine montre que les modalités de diffusion sont variables d'un type de hache à l'autre et peuvent également avoir une trajectoire plus

[8] Même si ce type de décroissance linéaire, plus simple à mettre en œuvre, a été utilisé dans les premiers modèles d'interaction spatiale utilisé en archéologie dès les années 70, notamment sous le nom de Xtent model (Renfrew, Level 1979). Observant que la forme des territoires dominés par les cités antiques étaient liées à la fois à la taille des capitales et aux distances qui les séparaient. Les auteurs ont utilisé une formule simple pour estimer l'influence politique des centres antiques (polis) sous forme de territoires hypothétiques. Avec ce modèle, il montre comment certains d'entre eux tombent sous la domination de territoires voisins plus importants. Le modèle Xtent intègre la taille des établissements et les distances qui les séparent comme facteurs déterminant la domination politique, selon la fonction suivante : $I = C^a - k\ X\ d$ où I est la force de l'influence qui dépend du poids C (taille) de chaque établissement et d la distance qui les séparent. Dans cette fonction les deux paramètres a et k déterminent l'équilibre entre la taille du centre et la distance. L'importance de la distance augmente de manière linéaire tandis que l'importance de la taille augmente de manière exponentielle. Ainsi, les grands centres seront plus compétitifs que les petits, même si la distance est plus grande.

complexe de type sinusoïdale qui implique l'utilisation d'autres modèles probabilistes plus complexes comme celui de Hägerstrand (1953, *cf.* Pumain, Saint-Julien 2001).

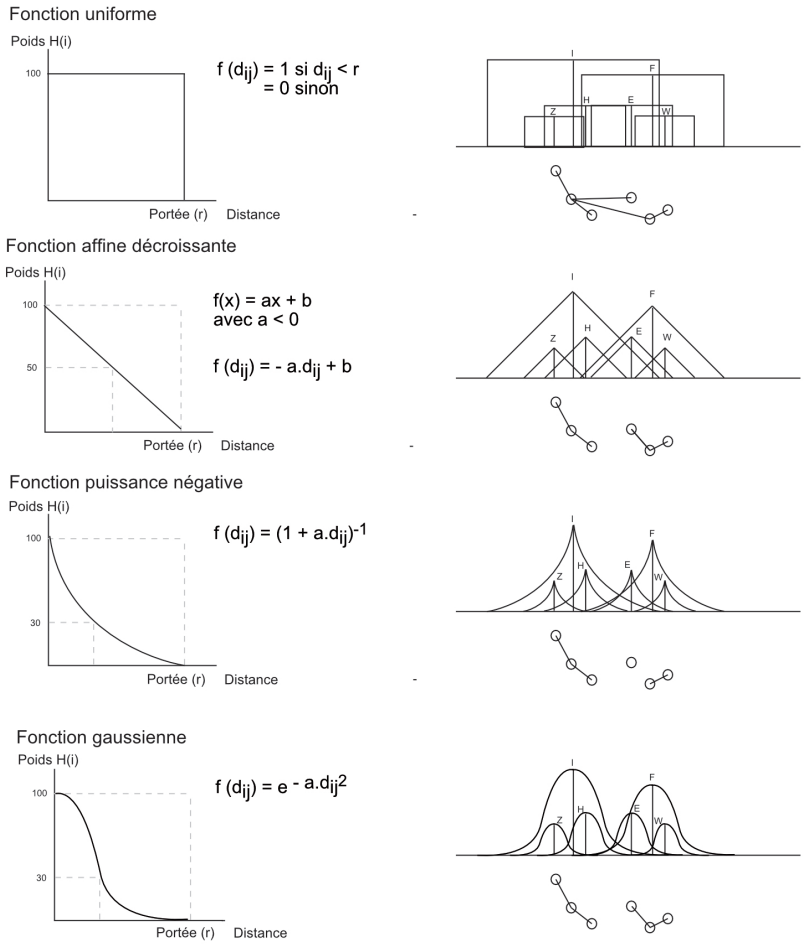


FIG. 7 : Représentation graphique de la construction des liens entre lieux par le modèle d'interaction selon différentes fonctions de lissage de la distance.

Conclusion

Dans l'exemple présenté, le choix d'un modèle d'interaction pour modéliser les relations entre les établissements n'est qu'une possibilité parmi d'autres correspondant à un type de données et surtout un type de conception socio-économique et politique qui doit être assumé par hypothèse et qui guidera ensuite le questionnement. Ce que nous avons voulu montrer dans ce chapitre, c'est la manière dont les questions se posent et dont le raisonnement se construit en mobilisant des méthodes et des techniques statistiques à chaque étape. Dans chacune d'elle, il s'agit soit d'organiser et de classer

des informations à partir de données observées, soit de construire une information manquante (le poids hiérarchique par exemple). Dans tous les cas, la construction est relative au corpus étudié et n'a pas de valeur absolue et il faut insister sur l'idée que « toute classification demeure provisoire et toute interprétation une interprétation d'attente » (Favory *et al.* 1994, p. 126). La classification et/ou l'information hiérarchique dérivée sont des outils qui ont le mérite de nous faire sortir du cadre fixiste imposé par les typologies fondées sur la morphologie des habitats, longtemps considérés par les archéologues comme autant d'entités singulières hors de tout contexte territorial et temporel.

Finalement, pour aborder la question du système de peuplement, même avec un jeu de données similaire, il n'y a pas de méthode clef en main mais une démarche à construire qui soit adaptée au questionnement initial de l'archéologue et à sa conception par hypothèse de la structure du réseau qu'il souhaite tester (type centre-périphérie, linéaire…). Il convient d'expliciter cette hypothèse initiale, les jeux de données mobilisés, puis les étapes de la démarche afin de préciser ensuite quels types d'outils statistiques pourront être mobilisés. L'article présentant trois expériences menées en Languedoc sur la modélisation des réseaux d'habitat (Nuninger *et al.* 2006) montre bien comment une même famille d'outils peut être mobilisée dans des démarches différentes.

Il est également important de noter que chaque étape d'une démarche peut ouvrir sur d'autres expériences. Ainsi, par exemple l'analyse des variables et les expériences de classification sur des corpus qui s'enrichissent dans le temps permet de vérifier la pertinence de certains descripteurs pour hiérarchiser les établissements (Nuninger *et al.* 2012a). L'information élaborée sous la forme de poids hiérarchique peut aussi être réinvestie dans une analyse de voisinage, telle que celle développée dans le chapitre 2.2.5 mais avec une approche prédictive afin de caractériser des contextes d'occupation, comme cela a été élaboré dans les programmes Archaedyn et IHAPMA[9] (Nuninger *et al.* 2012c, Nuninger *et al.* 2016) puis mis en application sur trois micro-régions d'Alsace (Nüsslein *et al.* 2020).

[9] ArchaeDyn I (2005-2007) : « Dynamique spatiale du peuplement et ressources naturelles : vers une analyse intégrée dans le long terme, de la Préhistoire au Moyen Âge » (Action concertée incitative – ACI Espaces et Territoires, contrat ET28) ; ArchaeDyn II (2009-2012) : « Dynamique spatiale des territoires de la Préhistoire au Moyen Âge » (ANR-08-BLAN-0157-01), coordonnés par François Favory et Laure Nuninger ; IHAPMA : « Introducing the human (f)actor in predictive modelling for archaeology », PHC Van Gogh, MAE – FRNL, 2010-2011, coordination Ph. Verhagen, L. Nuninger.