Chapitre 2.2.1.
Introduction : « Où-Localisation spatiale »

Laure NUNINGER(1), Hélène MATHIAN(2), Xavier RODIER(3)

(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, (3) CITERES-LAT UMR 7324, MSH Val de Loire UAR 3501, CNRS, Université de Tours, Université d'Orléans, Tours

1 Introduction

Les données archéologiques, outre leur caractère propre (morphologie, texture, homogénéité, hétérogénéité…) comportent également des dimensions à la fois temporelles et spatiales. C'est notamment en interrogeant ces deux dimensions que, dans les approches statistiques visant à analyser les données, on pose les questions du Quand ? et du Où ?

En archéologie, le Quand ? est lié d'une part à l'amplitude temporelle couverte par la discipline, d'autre part au processus d'élaboration de l'information lui-même contraint par la nature des sources mises en œuvre (datation, chronologie, chrono-typologie… *cf.* 1.1). S'interroger sur le Où ? pour l'archéologue est une des premières tâches qui consiste, dès le terrain, à localiser ses découvertes dans un espace géographique selon l'échelle de l'étude : des unités stratigraphiques au territoire. La seconde tâche est de situer ces découvertes au sens du rapport qu'elles entretiennent entre elles pour comprendre les modes de vie et le fonctionnement des sociétés du passé. Les sociétés se définissent aussi par leur dimension spatiale, c'est-à-dire par la manière dont elles se répartissent, circulent ou concentrent leurs activités.

Les études de cas présentées dans cette partie s'interrogent donc sur la dimension spatiale du "Où ?" *via* l'analyse de la localisation d'objets. L'approche ne porte pas tant sur le test d'hypothèses de lieux de provenance (production, circulation, diffusion) qui sont traitées dans la partie 2.1 mais sur l'étude des répartitions géographiques des artefacts ou sites archéologiques afin de mettre en évidence des formes différentes d'occupation du territoire, pouvant permettre aussi de comprendre les modalités de diffusion de ses artefacts (*cf.* 2.2.7) et de modéliser les échanges entre les lieux, soit les

interactions spatiales. Nous nous intéressons, à partir des données archéologiques et de toutes les sources mobilisables, aux rapports des sociétés du passé à l'espace dans lequel elles s'inscrivent et qu'elles produisent. Cette approche, qualifiée d'archéologie spatiale, se caractérise par un enchaînement de méthodes issues de l'analyse spatiale qui sont traitées dans les chapitres de cette partie.

2 Notion d'espace et d'analyse spatiale ?

L'argumentaire archéologique est fondé sur l'espace, en termes de proximité ou de connexion des objets archéologiques, de relation de ces objets à l'espace géographique, ou en termes de répartition spatiale. Bien que l'analyse spatiale soit depuis longtemps ancrée en archéologie (Hodder et Orton 1976), la référence à l'espace demeure souvent implicite et la conception de l'espace par les archéologues, en particulier en France, relève, le plus souvent, de la notion d'« espace support » (Pumain 1993). En d'autres termes, on peut penser l'analyse statistique des entités observées sans intégrer explicitement l'espace, simplement en localisant les données en amont ou les résultats obtenus en sortie. Dans ce cas, on aura une lecture cartographique des données ou résultats statistiques et on pourra faire des hypothèses sur la répartition de ces données mais pas analyser comment la dimension spatiale peut être un élément explicatif de la variabilité observée.

Pour l'**analyse spatiale**, malgré les nuances observées sur les définitions, on retiendra qu'il s'agit de l'« Ensemble des méthodes mathématiques et statistiques visant à préciser la nature, la qualité, la quantité attachées aux lieux et aux relations qu'ils entretiennent – l'ensemble constituant l'espace –, en étudiant simultanément attributs et localisations. […] » (Brunet *et al.* 1993). « L'analyse spatiale est une analyse formalisée de la configuration et des propriétés de l'espace géographique, tel qu'il est produit et vécu par les sociétés humaines » (Pumain, Saint-Julien 1997).

Dans cette définition, il est donc question d'analyses fondées sur des méthodes statistiques intégrant la notion de distance et/ou de relation selon leur localisation. Il y a une prise en compte de l'espace en tant que facteur/variable dans l'analyse.

3 Comment mobiliser l'analyse spatiale et pourquoi ?

Comme le lecteur pourra le constater au fil des chapitres, l'analyse spatiale peut être utilisée soit en préalable à une démarche exploratoire, soit *a posteriori* dans une approche modélisatrice.

Dans une démarche exploratoire, il s'agit le plus souvent de détecter les principaux éléments (lignes directrices, zones homogènes/hétérogènes) qui structurent la distribution d'entités spatiales, et d'identifier des spécificités ou des anomalies. À ce stade de l'analyse, il est possible d'expliquer statistiquement mais pas d'interpréter historiquement ou géographiquement les phénomènes observés. En revanche, les résultats des analyses permettent de proposer des hypothèses d'explication de ces phénomènes, qui pourront soit être confrontées à des connaissances externes, soit être mobilisées ensuite dans l'élaboration d'un modèle spatial.

Dans une approche modélisatrice, on pose un certain nombre d'hypothèses explicatives de l'organisation des entités spatiales en amont dont on évalue ensuite la validité pour un jeu de données grâce aux outils d'analyse statistique et spatiale en mobilisant des tests statistiques, par exemple. Les hypothèses peuvent être fondées sur la base d'analyses exploratoires de données ou à partir de connaissances externes ou encore de théories.

On propose de distinguer l'analyse spatiale et la modélisation spatiale, du fait que la première a plutôt une visée descriptive (*cf.* 2.2.3, 2.2.4), tandis que l'objectif de la seconde est plutôt explicatif (*cf.* 2.2.6). Néanmoins, dans la pratique plusieurs chapitres montrent que pour répondre à une problématique complexe, les deux démarches sont utilisées conjointement dans un plan analytique qui procède en plusieurs étapes (*cf.* 2.2.5, 2.2.7, 2.2.8).

D'un point de vue pratique, il est possible d'effectuer l'analyse spatiale de données archéologiques en utilisant des outils bureautiques jusqu'à des systèmes plus spécialisés. Ainsi, par exemple des analyses de distances et de localisations peuvent être réalisées avec un tableur, voire sans moyen informatique, tandis que des analyses de localisation fondées sur une approche multicritères à partir de plusieurs couches d'information géographique nécessitent le recours à un système d'information géographique (SIG) pour mettre en œuvre des méthodes comportant des calculs importants ou des optimisations trop compliquées à mettre en œuvre manuellement. Il est important de souligner ici qu'aucun environnement (SIG, SGBD…) est neutre du point de vue des calculs élaborés et, de fait, dans les résultats obtenus. Par exemple, le choix du modèle de données peut impacter directement le calcul de la distance. Dans un format vectoriel* (point, ligne, polygones), la distance calculée entre deux points sera de type euclidienne (vol d'oiseau), tandis que dans un format raster* (pixels constituant une image), elle sera de type rectilinéaire (distance de Manhattan). En termes mathématiques, pour la distance D entre A et B cela revient à avoir :

- Distance euclidienne : $\sqrt{(c^2 + d^2)} = D(A;B)$
  soit par exemple pour c=4 et d=6 => D=7.2111025
- Distance rectilinéaire : $c + d$ = D(A ;B)
  soit par exemple pour c=4 et d=6 => D=10

Quand on utilise un SIG, il convient donc de bien préciser sur quels types de données portent les calculs et surtout de rester cohérents dans les valeurs employées pour réaliser les analyses statistiques afin d'avoir des résultats fiables et comparables. Dans les chapitres qui vont suivre, les deux types de formats sont utilisés dans les analyses : vectoriel (*cf.* 2.2.6 et 2.2.8) et raster (*cf.* 2.2.5). Mais parfois, une analyse par maille (*cf.* 2.2.3) revient à transformer un espace vectoriel en espace de type raster en attribuant à chaque maille (équivalent pixel) une valeur spécifique correspondant à un choix associé à la question. Il est donc très important d'être explicite sur les données et leur transformation à chaque étape de l'analyse.

Notons enfin que les outils d'intelligence artificielle ont considérablement fait évoluer les méthodes d'analyse employées, notamment pour corriger l'échantillonnage (*cf.* 2.2.5). Ce type d'outil ne sera pas présenté dans les chapitres qui suivent car

ces méthodes sont récentes et elles n'ont pas encore trouvé toute leur place dans la pratique archéologique.

4 Se poser les bonnes questions

Si le détour par les outils présentés ci-dessus est important, savoir se poser les bonnes questions est une posture fondamentale et primordiale pour mener à bien une analyse. C'est précisément ce sur quoi les cas d'étude présentés ensuite mettent l'accent. Ainsi, chaque chapitre est proposé comme un exemple qui illustre la manière dont une problématique plus ou moins complexe peut être découpée en une série de questions plus simples que l'on peut traduire d'un point de vue analytique avec les outils statistiques. Chaque exemple s'efforce également de montrer comment ces analyses s'articulent dans un schéma analytique cohérent dont l'objectif est de contribuer à clarifier le problème initial. Il est donc important, ici, d'avertir le lecteur que les chapitres suivants ne sont pas des fiches techniques ou des modèles à suivre pas à pas. En revanche, ils ont vocation à l'accompagner dans la compréhension de démarches d'analyse spatiale développées dans plusieurs programmes de recherche et par ailleurs publiées dans des articles ou des ouvrages. Nous invitons donc le lecteur à consulter ensuite la bibliographie pour aller plus loin et prendre connaissance de l'élaboration des méthodes et des résultats obtenus par ces travaux.

Précisons également que tous les chapitres ne sont pas construits de la même façon et sont indépendants. À travers plusieurs exemples, certains se focalisent plus sur les différentes manières d'aborder un problème, tandis que d'autres se concentrent plus sur la manière d'aborder un problème à partir d'un seul exemple. On notera que chaque chapitre aborde un problème plus ou moins complexe avec les outils d'analyse spatiale et que, par ailleurs, ils font appel à des techniques statistiques relativement simples (*cf.* 2.2.3 ou 2.2.4) ou très avancées (*cf.* 2.2.6 ou 2.2.5). On insistera aussi sur l'idée que "qui peut le plus peut le moins" est une fausse route car la décomposition d'un problème est nettement préférable à l'utilisation d'une technique sophistiquée pour obtenir de bons résultats, contribuer à répondre à la question initiale et apporter, en préalable, une meilleure connaissance des données et de ce qu'elles recèlent. Dans le cas contraire le risque est grand de se perdre face à des résultats que l'on est incapable d'analyser et encore moins d'interpréter, ce qui demeure tout de même l'objectif premier d'une étude.

Enfin, à travers chacun des chapitres, on verra qu'il est possible de construire un raisonnement différent, soit dans un processus linéaire avec des données en entrée et des résultats à interpréter en sortie, soit dans un processus itératif ou par boucles avec un aller-retour régulier entre les données en entrée et les résultats obtenus au cours de l'analyse. Ces deux modes de raisonnement sont également combinés dans certaines approches exploratoires pour d'une par poser un cadre de référence en fonction des données que l'on traite (*cf.* 2.2.3), d'autre part affiner la compréhension du jeu de données que l'on analyse (*cf.* 2.2.8).

Chaque cas d'étude peut être lu indépendamment les uns des autres mais nous conseillons au lecteur de lire d'abord les plus simples (*cf.* 2.2.3 ou 2.2.4) avant de se lancer dans les exemples impliquant la combinaison de plusieurs techniques (*cf.* 2.2.5 et

suivant). En outre, avant d'aborder les cas d'étude, nous l'invitons à prendre connaissance de l'encart sur le vocabulaire employé (ci-dessous) et du chapitre 2.2.2 sur la formalisation des données qui est un préalable à toute analyse.

Petit point de vocabulaire

L'analyse spatiale utilise une terminologie qui emprunte souvent à celle de l'analyse statistique et qui peut parfois sembler ambiguë au lecteur. Sans chercher à être exhaustif, nous attirons l'attention ici sur quelques termes et nous invitons vivement le lecteur à se reporter systématiquement au glossaire.

Densité : en analyse spatiale nous parlons régulièrement de densité qu'il faut comprendre comme le rapport entre un indicateur (nombre d'individus statistiques) et une surface ou un linéaire. On parlera ainsi de densité de population (rapport entre le nombre d'habitants et la surface d'un territoire délimité). En statistique, on utilise également le terme de densité indépendamment de l'espace, comme le rapport entre la fréquence de la classe et son étendue pour comparer deux classes d'amplitudes différentes, par exemple. Tout au long de l'ouvrage il sera donc question de densité spatiale, de densité de fréquence ou de densité de probabilité.

Distribution : est aussi un terme régulièrement utilisé en analyse spatiale pour rendre compte de la façon dont les objets se répartissent dans l'espace sous forme dispersée, concentrée, régulière, irrégulière. On parle aussi de répartition spatiale. Dans ce cas, il ne s'agit pas de la distribution statistique qui renvoie aux fréquences d'apparition par classe de valeur sous la forme d'un histogramme par exemple.

Peuplement/population : en géographie ou archéologie spatiale la notion de peuplement est généralement liée à celle de système ou de dynamique et se réfère aux manières dont les sociétés humaines ont occupé et occupent l'espace. Les indicateurs statistiques utilisés peuvent mais ne sont pas systématiquement d'ordre démographique. Il peut s'agir de sites archéologiques par exemple. Les données qui constituent cette "population" sont construites pour être signifiantes d'un point de vue thématique (*cf.* 2.2.2), elles peuvent être différentes de la "population statistique" analysée, elle-même composée d'individus statistiques, c'est-à-dire d'unités statistiques quelle que soit leur nature d'un point de vue thématique.

Incertitude : en archéologie l'incertitude est le plus souvent liée à la complétude ou non des données, ce qui est différent de l'incertitude de mesure en statistique.