

L'ANALYSE DE LA FORME D'UN SEMIS DE POINTS

Une application à la comparaison de parcs arborés au Burkina-Faso et à l'étude de la distribution de sites archéologiques dans le Val-de-Marne

Claude GRASLAND, Catherine MERING, Université Paris VII
Marianne GUEROIS, Université Paris I

Type : TD

Niveau : Licence

Durée : 4 heures (2 heures pour chaque exemple, le deuxième pouvant faire l'objet d'un devoir sur table ou d'un devoir à la maison)

Thème : ANALYSE SPATIALE

Objectifs :

Ce TD s'inscrit dans le cadre d'un module d'introduction à l'analyse spatiale. Son objectif est de montrer l'intérêt respectif de deux méthodes d'analyse de la forme d'un semis de points, celle des « quadrats » et celle du « plus proche voisin ». On rappellera, pour situer ces méthodes, que la distribution spatiale de phénomènes ponctuels peut être résumée par un point central, ou bien caractérisée à partir de points et d'axes remarquables (distribution des densités de population en fonction de la distance au centre d'une ville, par exemple). Mais il s'avère aussi parfois utile, en géographie, d'analyser ces semis de points en en faisant ressortir la forme concentrée, aléatoire ou régulière, indépendamment de toute référence à un point ou à un axe particuliers¹.

Pour le montrer, on partira de deux exemples : le premier porte sur la comparaison de parcs arborés du Burkina-Faso, le second sur la distribution de sites archéologiques qui témoignent, dans le Val-de-Marne, de la répartition du peuplement à différentes époques. Il s'agit d'apprendre à mesurer et à comparer, dans l'espace et dans le temps, des formes de concentration ou de dispersion spatiale, et de voir en quoi l'application de ces méthodes permet de poser des hypothèses sur la genèse des formes observées.

¹ Pour plus de détails, voir en fin de « Feuille » la bibliographie sur ces aspects méthodologiques.

Premier exemple

**CONTRIBUTION DE LA METHODE DES QUADRATS
A L'ETUDE DE PARCS ARBORES EN PAYS DAGARI (BURKINA-FASO)**

Présentation :

Les parcs arborés, qui associent l'arbre au champ, sont une constante des paysages agraires des pays du Sahel. En dépit de leur apparence « naturelle », ce sont le plus souvent des créations anthropiques, produits de pratiques agro-pastorales traditionnelles. La composition des essences arborées dans ces parcs, mais aussi leur densité et la forme de leur distribution spatiale, traduisent à cet égard des besoins et des techniques différents et révèlent « la stratégie que chaque société conduit à l'égard du milieu dans lequel elle est insérée » (Pélissier, 1980). C'est dans cette perspective que l'on évaluera la contribution de la méthode des quadrats à la description des parcs arborés et à la compréhension de leur genèse récente.

Le **document 1** reproduit la carte des parcs arborés en pays dagari (Burkina-Faso), chaque espèce d'arbre étant représentée par un figuré ponctuel particulier. Une première analyse visuelle montre que toutes les espèces ne sont pas distribuées également. On peut notamment identifier deux sous-espaces (zones test n°1 et 2) qui présentent des distributions assez contrastées. Dans quelle mesure la méthode des quadrats peut-elle aider à mesurer et interpréter ce contraste ?

Documents :

- Document 1 : Cartes de la distribution des essences arborées en pays dagari
- Document 2 : Fréquence d'apparition des essences arborées dans deux zones-test
- Document 3 : Extrait de la G.U. sur les fonctions des parcs arborés
- Document 4 (corrigé) : Calcul d'un indice de concentration
- Document 5 (corrigé) : Test de la forme des distributions de nérés et d'acacias
- Document 6 : Extrait d'un texte de P. Gould sur l'intérêt du test du Chi-2

Déroulement :

Première étape : Comparer la composition des essences dans les deux zones-test

Dans un premier temps, on demande aux étudiants de commenter le tableau qui dénombre la fréquence globale des différentes essences dans chacune des deux zones-test (**Document 2**). L'une et l'autre apparaissent dominées par une seule essence ou par une association d'essences. La zone 1, qui montre une densité d'arbres plus faible que dans la zone 2 (environ 12 arbres à l'ha contre 16 arbres à l'ha dans la zone 2), se caractérise par la sur-représentation des karités (plus de la moitié des essences) et, à un moindre degré, des nérés (25% des essences) ; elle reste néanmoins plus diversifiée que la zone 2, où l'acacia prédomine largement (80% des essences présentes).

Plusieurs pistes d'interprétation peuvent être proposées pour rendre compte des différences observées : on peut notamment invoquer le rôle de facteurs écologiques locaux (on sait par exemple que le karité est une plante qui a particulièrement besoin d'eau), mais la lecture du **document 3** suggère que la fonction de chacune des essences, en relation avec les besoins particuliers des sociétés exploitant les parcs, est un des facteurs les plus déterminants. Ainsi,

la prédominance des karités, appelés aussi « arbres à beurre », témoigne, au nord, de la recherche de graisses qui peuvent notamment se substituer à la matière grasse animale, tandis que le besoin en condiments peut justifier de la présence des nérés. Au sud, la sur-représentation de l'acacia, légumineuse qui enrichit le sol, indique une volonté de fertilisation des terres et doit être mise en relation avec la présence d'un bétail important (rôle nutritif et protecteur des feuilles de l'acacia pendant la saison sèche). Sachant que la plantation d'acacias renvoie à un processus d'intensification des cultures, on peut suggérer l'existence d'un lien entre la nature des essences dominantes et l'intensité de la pression démographique s'exerçant sur chacune des deux zones. Enfin, on peut être tentés par une interprétation qui mettrait davantage en avant des facteurs culturels, liés aux différentes civilisations agraires (il arrive parfois que même à faible distance, la distinction entre les parcs souligne de véritables frontières de civilisation et n'illustre pas seulement deux modes d'exploitation), même si l'on sait que cette dichotomie entre parcs à acacias et d'autres colonies arborées se retrouve pratiquement dans toute la zone sahélienne, quelle que soit l'ethnie considérée.

Deuxième étape : Caractériser la distribution spatiale des nérés et des acacias à l'aide de la méthode des quadrats

Si la composition des parcs arborés est un indice du mode d'organisation des sociétés rurales, cette première approche gagne à être complétée par la description de la distribution spatiale des essences dominantes. La forme plus ou moins concentrée de cette distribution peut en effet témoigner des dynamiques de déboisement ou de reboisement responsables de l'évolution des parcs arborés : les essences représentées peuvent être des espèces résiduelles, ayant résisté aux déboisements et à la mise en culture de nouvelles parcelles, mais elles peuvent aussi avoir été sélectionnées et repiquées volontairement après l'abandon de certains champs cultivés ou en jachère.

Deux étapes sont nécessaires pour comparer la distribution spatiale des essences dominantes que sont les acacias au sud et les nérés au nord. Nous détaillons ici les résultats de l'étude mais nous renvoyons, pour la compréhension plus précise de la démarche générale, aux références méthodologiques données en bibliographie.

- Calcul d'un indice de concentration :

Après avoir dénombré les arbres au sein d'un maillage régulier de « quadrats » (les carreaux des deux grilles superposées à l'espace d'étude), on peut examiner la distribution des fréquences d'arbres par quadrat et déterminer si elle se rapproche d'une distribution aléatoire, concentrée ou régulière (**Document 4**). On calcule pour cela un indice de concentration qui est le rapport entre la variance de la densité d'arbres par quadrats (cette mesure de la dispersion statistique est élevée quand il y a quelques quadrats très pleins et beaucoup de quadrats vides) et la densité moyenne d'arbres par quadrat. L'indice, qui tend vers 1 lorsque la distribution est de forme aléatoire, est supérieur à 1 lorsqu'elle montre une tendance à la concentration.

La forme du semis des nérés, dans la zone 1, est proche d'une distribution aléatoire (IC=0,94), alors que celle des acacias montre une forte tendance à la concentration (IC=4,9). Cet indice donne donc une première idée de la forme du semis des arbres dans chacune des deux zones, mais n'est pas suffisant : il doit être accompagné d'un test du Chi-2 pour savoir si l'écart entre la distribution observée et la distribution théorique, de forme aléatoire, est significatif ou pas.

- Test du Chi-2 et interprétation :

Cette étape est assez délicate car il faut faire comprendre l'intérêt d'un tel test alors que l'on dispose déjà d'une mesure de concentration. L'exemple de la centrale nucléaire de Sellafield donné par Stan Openshaw s'avère en cela particulièrement didactique (**Document 6**) : le test du Chi-2 permet dans ce cas de prouver l'existence d'un lien entre la localisation d'une centrale nucléaire et la concentration de cas de leucémie infantile dans le voisinage de cette centrale. A partir du texte, les étudiants doivent bien identifier l'origine de la controverse, à savoir qu'une « répartition causée par un processus purement aléatoire peut produire des groupements de points », et que, par conséquent, la seule mesure de la concentration n'est pas suffisamment probante.

Dans ce TD, l'enjeu du test n'apparaît pas aussi stratégique... mais il est nécessaire de l'appliquer si l'on veut pouvoir commenter les processus responsables de la distribution des essences arborées (**Document 5**).

On part de l'hypothèse selon laquelle la forme de la distribution des nérés et des acacias est le produit d'un processus aléatoire. Pour les nérés (zone 1), on ne peut pas rejeter cette hypothèse car le Chi-2 observé (soit l'écart relatif à une distribution aléatoire) reste inférieur au Chi-2 théorique, même pour un risque d'erreur de 20%, et n'est donc pas significatif. Comme le suggérait l'indice de concentration, la distribution des nérés est bien le produit d'un processus aléatoire. Pour les acacias, à l'inverse, le Chi-2 observé, largement supérieur au Chi-2 théorique, quel que soit le risque d'erreur considéré, est significatif, et permet de rejeter l'hypothèse selon laquelle cette concentration serait le produit du hasard.

Ces résultats autorisent l'introduction de quelques hypothèses sur la genèse récente des parcs arborés dans les deux zones étudiées. Au nord (zone 1), les nérés, qui montrent une distribution assez aléatoire, pourraient être considérés comme des essences « préservées » dans un contexte de déboisement et de culture extensive de champs « de brousse », à l'écart des villages. Au sud (zone 2), la concentration significative d'acacias viendrait au contraire conforter l'idée d'une intervention concertée et d'une dynamique de reboisement artificiel. L'accroissement démographique, l'intensification des cultures et la réduction des jachères autour des villages seraient à l'origine de la mise en place, par les communautés villageoises, d'espèces « fertilisantes » et propres au fourrage en saison sèche. On peut évoquer, pour renforcer l'hypothèse d'une plantation artificielle destinée à l'élevage, le fait que ces espèces apparaissent relativement éloignées de l'habitat.

Les « deux parcs emboîtés correspondent-ils à deux étapes de l'évolution agraire d'une même population » ? « Un des parcs est-il fossile et par conséquent témoin d'une civilisation différente » ? S'agit-il de « deux populations ayant des techniques, des organisations et des stratégies distinctes » (Pélissier, 1980) ? Il resterait à mener, à partir de ces hypothèses, un travail d'interprétation conséquent pour expliquer la juxtaposition à si faible distance de ces deux formes de colonies arborées.

- Limites de cette méthode

On peut, pour prolonger cette introduction à la méthode des quadrats, revenir sur le document 6 et sur les précautions prises par S. Openshaw pour faire varier la taille des mailles de la grille de manière systématique. Cet exemple peut être le support d'un examen critique de la méthode et de ses principales limites (influence de la taille et de la forme des quadrats sur les résultats), difficiles à gérer sans la puissance d'un outil informatique permettant de multiplier les simulations. De plus, cette méthode ne tient pas compte du voisinage mais uniquement de la densité à l'intérieur des mailles. Si l'on regarde la zone 1, par exemple, on distingue des structures plus ou moins linéaires dans la répartition des nérés, qui ne sont pas prises en compte par la méthode, et qui peuvent interroger le caractère aléatoire de la structure.

Deuxième exemple

APPLICATION DE LA METHODE DU PLUS PROCHE VOISIN A LA DISTRIBUTION DES SITES ARCHEOLOGIQUES DU VAL-DE-MARNE

Présentation :

Parce qu'elle s'affranchit de la grille des quadrats et des limites inhérentes au choix de cette grille, la méthode du plus proche voisin est parfois préférée pour caractériser la forme d'un semis de points. Au lieu de partir de la distribution des densités de points dans un maillage, elle s'appuie sur les caractéristiques de l'espacement entre les points. Cette méthode est ici appliquée à un exemple de géographie historique (et préhistorique) : un inventaire cartographique des sites archéologiques du Val-de-Marne (**Document 7**) permet d'étudier la distribution des découvertes archéologiques en fonction de leur localisation et de la période concernée. On doit déterminer si les sites archéologiques se distribuent au hasard dans le département du Val-de-Marne ou bien si leur localisation répond à une logique spatiale particulière.

Documents :

- Document 7 : Carte des sites archéologiques du Val-de-Marne, toutes périodes confondues
- Documents 8 à 12 : Cartes des sites archéologiques, aux périodes du Paléolithique, du Néolithique, Bronze et Fer, Gallo-Romaine et Médiévale
- Documents 13 et 14 (corrigé) : Relevé des distances au plus proche voisin pour chaque période

Déroulement :

Première étape : Détermination d'une zone-tampon autour de la zone d'étude

Avant toute analyse, on peut d'abord faire réfléchir les étudiants sur l'intérêt de la zone-tampon ajoutée autour de la zone d'étude (**Documents 8 à 12**). Le choix des limites de la zone d'étude peut en effet introduire des biais importants en tronquant certains groupements de points : au Paléolithique par exemple, si l'on reste dans les limites de la zone d'étude, le site n°13, situé à plus de 7 km de son plus proche voisin, apparaît assez isolé, alors que son plus proche voisin en dehors de cette zone d'étude n'est qu'à environ 1 km. C'est donc pour limiter ce biais, sensible pour les points les plus excentrés, que l'on a créé une zone-tampon. Les points inclus dans cette zone peuvent tenir lieu de « plus proches voisins ».

Deuxième étape : Caractérisation de la distribution des sites à chaque période

Les étudiants doivent ensuite déterminer si, à chaque période donnée, les sites archéologiques se distribuent au hasard dans le département, s'ils montrent une tendance à la concentration ou au contraire à l'espacement régulier. A chaque période, ils doivent d'abord calculer la distance qui sépare chaque point de son voisin le plus proche, et calculer la moyenne D_0 de l'ensemble de ces distances observées (**Document 13**). La distance observée moyenne doit être alors comparée à une distance théorique qui correspond, comme dans la méthode des

quadrats, à une distribution de forme aléatoire (**Document 14**). Le rapport entre les deux mesures de distance produit un indice (R) qui est égal à 1 si la distribution est aléatoire (distance observée = distance théorique) et qui est inférieur à 1 si le semis de points est plutôt concentré (distance observée < distance théorique).

Au Paléolithique, les sites archéologiques montrent une tendance à la concentration (R=0,78), qui s'estompe au Néolithique (R=0,95) et s'affirme à nouveau aux ères du Bronze et du Fer (R=0,75) ainsi qu'à l'époque Gallo-romaine (R=0,8). L'époque médiévale montre quant à elle une tendance sensible à la plus grande dispersion des sites (R=1,05). Le résumé de chacune des distributions par un indice synthétique facile à interpréter (attention : l'interprétation de cet indice est inversée par rapport à la méthode des quadrats : une valeur inférieure à 1 - distribution aléatoire - sera plutôt concentrée d'après la méthode du plus proche voisin) s'avère particulièrement utile pour cette comparaison temporelle. En principe, il faudrait procéder à un test paramétrique pour tester l'hypothèse R=1 qui correspond à l'existence d'une distribution aléatoire. Ceci n'est possible que si les étudiants maîtrisent déjà ces tests, plus compliqués que le test du Chi-2 employé dans la méthode des quadrats.

Troisième étape : Analyse de l'évolution historique du peuplement

Si l'on admet que la distribution des sites reflète la distribution du peuplement au fil des périodes étudiées, quelle tendance remarquable peut-on mettre en évidence dans l'évolution historique du peuplement, et comment l'expliquer ?

Sur la totalité de la période, la variation de la valeur de l'indice montre une tendance générale à la dispersion du semis, et à un processus très lent de déconcentration de l'habitat. On part au Paléolithique d'un semis plutôt concentré de l'habitat pour aboutir, à l'époque médiévale, à un semis plutôt dispersé, proche d'une distribution aléatoire. Néanmoins, cette tendance n'est pas tout à fait linéaire. La période Néolithique se distingue en particulier par une dynamique de diffusion du peuplement (cf. l'évolution du nombre de sites découverts). Elle est suivie d'une rétraction et d'une concentration de l'habitat jusqu'à la période médiévale.

On peut partir de ce constat pour émettre l'hypothèse selon laquelle chaque « vague » de dispersion du peuplement correspondrait à une révolution agricole (au Néolithique et au Moyen-Age), tandis que la plus faible densité et la plus forte concentration de sites pourrait témoigner d'un repli et de positions défensives en période de grande insécurité. L'interprétation de cette tendance générale et de ces fluctuations doit néanmoins être conduite avec prudence. A défaut d'un découpage chronologique plus fin, il est en effet difficile d'aller plus loin dans la formulation d'hypothèses de cette nature. Mais surtout, la localisation des sites reste dépendante de l'activité des archéologues et de leur subjectivité quant au choix des zones de fouille, et on ignore sur quelle base ont procédé les auteurs de l'Atlas du Val-de-Marne. On doit donc se garder de trop fantasmer sur la grande interprétation finale trans-historique...

Ces réserves faites, ce type d'exercice donne un aperçu de ce qui, avec davantage de précision dans les données et d'information quant à leur collecte, peut contribuer à une meilleure compréhension de l'histoire du peuplement dans une région (Archéomédès, 1997).

Remarques sur l'intérêt des étudiants pour ces TD

Si les étudiants comprennent assez bien l'intérêt de ce type d'analyse dans une perspective descriptive (résumer la distribution de chaque période à l'aide d'un seul paramètre puis étudier des tendances), il est souvent plus difficile de leur montrer l'intérêt des tests statistiques associés à la mesure des paramètres de concentration spatiale.

Notre expérience pédagogique montre qu'une certaine dramatisation (genre « Une histoire dont vous - géographe - êtes le héros ») est nécessaire pour faire accepter le test en montrant qu'il est parfois crucial de déterminer si la configuration observée est l'effet du hasard ou non. L'exemple de la centrale nucléaire de Sellafield (document 6) est de ce point de vue particulièrement efficace puisque l'on peut facilement montrer aux étudiants que le choix d'une démarche scientifique rigoureuse était ici absolument nécessaire pour contrer les affirmations d'une administration opaque et que le géographe S. Openshaw n'aurait jamais pu contrer les affirmations des experts ministériels s'il n'avait pas eu la capacité statistique de les battre sur leur propre terrain... L'exemple est d'autant plus joli que, non content de montrer les dangers de la centrale de Sellafield (« Vive la géographie citoyenne »...), l'étude d'Openshaw avait fait découvrir un autre foyer épidémiologique dont personne ne se doutait (« C'est beau la science... »).

Bref, si on a bien « chauffé la salle » à l'aide de ce texte, non seulement les étudiants trouvent tout naturel de procéder au test du chi-2 dans le cas de la méthode des quadrats, mais il peut même aussi arriver qu'ils s'étonnent qu'on ne procède pas également au test dans le cas de la méthode du plus proche voisin et reprochent à l'enseignant son manque de rigueur (!).

Quelques difficultés spécifiques à l'emploi de ces méthodes en géographie

Il est assez difficile de trouver de bons exemples d'analyse des semis de points en géographie humaine car la plupart des distributions (écoles, hôpitaux, commerces...) sont fortement corrélées avec la distribution générale de la population. Lorsque l'on teste le caractère aléatoire ou non aléatoire d'un semis d'écoles, par exemple, on teste surtout indirectement le caractère évidemment non aléatoire de la distribution de la population et le résultat peut apparaître trivial. Il faut donc soit trouver des portions d'espace où la densité de population est à peu près constante (e.g. un centre-ville), soit procéder à des corrections du semis pour éliminer le facteur banal qu'est la densité de population. On peut signaler à cet égard l'exemple très intéressant proposé par Taylor dans son manuel sur la distribution des centres commerciaux anglais vers 1970 (**Document 15**) : la répartition des commerces, qui apparaissait concentrée sur la carte ordinaire, ressort comme régulière lorsqu'elle est corrigée par la répartition de la population, et s'insère alors bien dans le modèle théorique des places centrales.

Le problème se pose également en géographie physique. Ainsi, lorsque l'on étudie la distribution d'arbres à l'intérieur d'un espace traversé par de fortes discontinuités topographiques, climatiques ou pédologiques, il n'y a pas grand sens à confronter la distribution observée à une distribution aléatoire puisque les facteurs structurels déterminent largement les répartitions observées. Si l'on veut procéder à une analyse en termes de quadrats ou de plus proches voisins, il faut choisir une zone-test où les conditions sont a priori homogènes, tout au moins par rapport aux déterminants structurels connus.

Bibliographie complémentaire :

Sur les aspects méthodologiques :

- GRASLAND C., 1999-2003, *Cours d'analyse spatiale de licence*, Université Paris 7 : <http://ibm2.cicrp.jussieu.fr/grasland/go303/index.htm>
⇒ Cours et documents de cours accessibles en ligne. Voir le Chapitre 1 consacré aux semis de points, notamment la dernière partie.
- PUMAIN D., SAINT-JULIEN T., MATHIAN H., 1997, *L'analyse spatiale. - t.1 Localisations dans l'espace* Cursus-Géographie, Armand Colin, Paris,
⇒ Manuel simple mais très didactique avec des exemples. Voir le Chapitre 2 : Lieux et distributions de lieux, pp. 51-87 et plus particulièrement pp. 75-82.
- TAYLOR P.J. , 1977, *Quantitative Methods in Geography - An Introduction to Spatial Analysis*, Waveland Press Inc., Prospect Heights, Illinois, 386 p.
⇒ Ouvrage de référence apportant toutes les informations sur la méthode des quadrats et celle du plus proche voisin, les tests, les développements possibles, la bibliographie, etc. Comporte des exercices corrigés en fin de chapitre. Voir le chapitre IV : Point pattern analysis, pp. 133-171.

Sur l'exemple des parcs arborés :

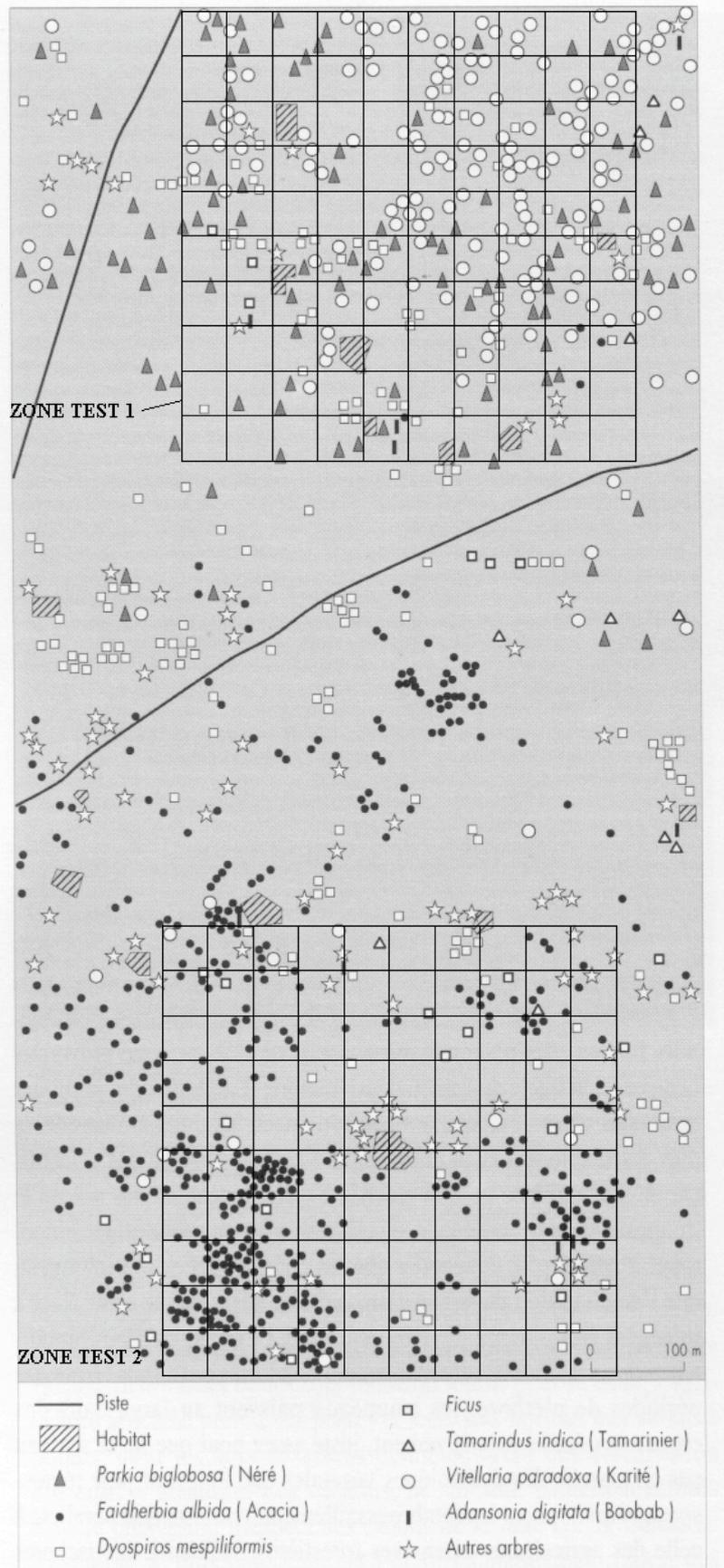
- PELISSIER P., 1980, « L'arbre dans les paysages agraires de l'Afrique noire », *Cahiers de l'Orstom*, vol. XVII, n° 3-4, 131-136.

Sur l'exemple des sites archéologiques :

- ARCHAEOMEDES, 1998, *Des oppida aux métropoles*, Anthropos, coll. Villes, 280 p.

Document 1

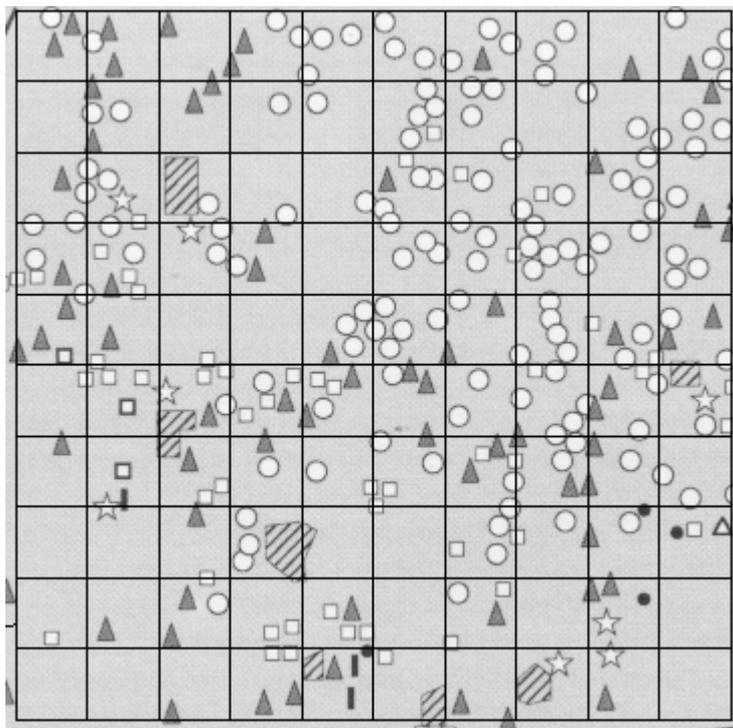
UN EXEMPLE DE « PARC » AFRICAIN : LE PAYS DAGARI DU BURKINA-FASO



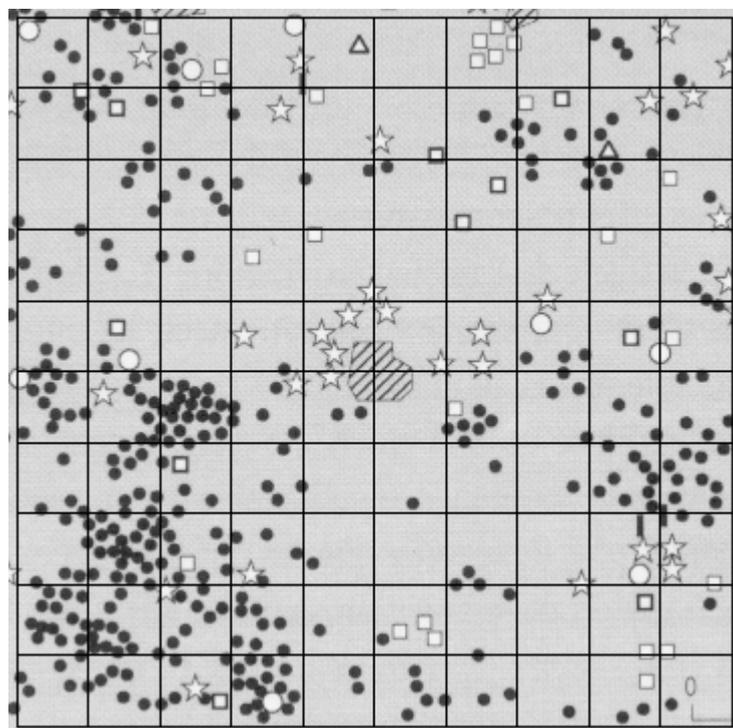
Source : Géographie Universelle, 1994, *Les Afriques au sud du Sahara*, Belin-Reclus, p. 70.

Document 1 (suite)

Zoom sur la zone d'étude n°1



Zoom sur la zone d'étude n°2



Document 2

FREQUENCE D'APPARITION DES ESSENCES ARBOREES DANS LES DEUX ZONES-TEST

| | Zone 1 | | Zone 2 | |
|----------------------|------------|--------------|------------|--------------|
| | Effectifs | Pourcentages | Effectifs | Pourcentages |
| Nérés | 68 | 25,19 | 0 | 0,00 |
| Acacia | 4 | 1,48 | 299 | 79,52 |
| Dyospiros | 48 | 17,78 | 24 | 6,38 |
| Ficus | 3 | 1,11 | 11 | 2,93 |
| Tamariniers | 1 | 0,37 | 1 | 0,27 |
| Karités | 135 | 50,00 | 8 | 2,13 |
| Baobabs | 3 | 1,11 | 2 | 0,53 |
| Autres arbres | 8 | 2,96 | 31 | 8,24 |
| TOTAL | 270 | 100 | 376 | 100 |

Document 3

LES FONCTIONS DES PARCS ARBORES

« Les « parcs » d'arbres préservés lors des défrichements, sinon sélectionnés ou repiqués, sont une constante des civilisations agraires dans les régions de savane et de forêt claire d'Afrique occidentale et centrale (...). Ces parcs sont complexes, parce qu'ils doivent satisfaire une gamme variée de besoins. On en tire du bois pour la construction, les outils, le chauffage, mais aussi des fibres ; plus encore peut-être, on en attend des graisses, que fournissent entre autres le karité, l'éleis, le cacédrat. Les parcs fournissent encore des fruits, des condiments (baobab, néré), des excitants (cola), des boissons (éleis, palmier rônier). Leur rôle alimentaire, notamment grâce aux fruits des ficus et aux pousses des rôniers, peut être essentiel en temps de disette. Enfin, à des degrés divers, les arbres contribuent à la fertilisation, par l'action de leurs racines, pièges à minéraux ; certains d'entre eux jouent un rôle particulier, et plus qu'aucun autre *Faidherbia albida*. Fournisseur, avec un travail minimal, de produits qui manquent dans la gamme des plantes cultivées ou de substituts de productions animales, le parc varie donc beaucoup selon les civilisations agraires et au gré de leur histoire. Ces formations d'apparence naturelle sont des créations culturelles, des indicateurs de l'état d'une société. La nature des arbres conservés, le traitement qui leur est réservé dépendent des productions agricoles, de l'importance du cheptel et les pratiques culturelles sont liées aux dynamiques de cette végétation « naturelle », aux nécessités de sa consommation. »

Source : Géographie Universelle, 1994, *Les Afriques au sud du Sahara*, Belin-Reclus, pp. 67-70.

Document 4

CALCUL D'UN INDICE DE CONCENTRATION

- pour les nérés (zone 1)

| Nb de nérés N | Nb de quadrats contenant N nérés K | Nb total de nérés N*K | Ecart à moyenne K(N-D) ² |
|------------------|---|--------------------------|---|
| 0 | 52 | 0 | 24,04 |
| 1 | 29 | 29 | 2,97 |
| 2 | 18 | 36 | 31,36 |
| 3 | 1 | 3 | 5,38 |
| Total | 100 | 68 | 63,76 |

Densité moyenne D = nb de nérés / nb de quadrats = 0,68 nérés à l'ha

Variance V(D) = 63,76 / (100-1) = 0,64

Indice de concentration IC = V(D)/D = 0,94

- pour les acacias (zone 2)

| Nb d'acacias N | Nb de quadrats contenant N acacias K | Nb total d'acacias N*K | Ecart à la moyenne K(N-D) ² |
|-------------------|---|------------------------------|--|
| 0 | 25 | 0 | 252,81 |
| 1 | 19 | 19 | 90,30 |
| 2 | 13 | 26 | 18,10 |
| 3 | 10 | 30 | 0,32 |
| 4 | 11 | 44 | 7,40 |
| 5 | 6 | 30 | 19,87 |
| 6 | 3 | 18 | 23,86 |
| 7 | 1 | 7 | 14,59 |
| 8 | 1 | 8 | 23,23 |
| 9 | 4 | 36 | 135,49 |
| 10 | 1 | 10 | 46,51 |
| 11 | 2 | 22 | 122,30 |
| 14 | 2 | 28 | 234,14 |
| 19 | 1 | 19 | 250,27 |
| 21 | 1 | 21 | 317,55 |
| Total | 100 | 318 | 1556,76 |

Densité moyenne D = nb de points / nb de quadrats = 3,18 acacias à l'ha

Variance V(D) = 1556,76 / (100-1) = 15,57

Indice de concentration IC = V(D)/D = 4,89

NB : Le fait que, dans le calcul de la variance, la somme des écarts à la moyenne soit divisée par (K-1), peut gêner les étudiants familiers de la formule qui divise ces écarts par le nombre total d'individus (K). Sans entrer dans le détail, on peut juste signaler que c'est (K-1) qui est utilisé parce qu'on se situe dans le cadre de tests d'hypothèses (cette méthode se réfère à la loi de Poisson pour poser des hypothèses sur la genèse des formes de distribution).

Document 5

TEST DE LA FORME DES DISTRIBUTIONS DE NÉRÉS ET D'ACACIAS

- Test du Chi-2 pour la distribution des nérés (zone 1)

| Nb de nérés par quadrat N | Nb de nérés observé K | Nb de nérés théorique K* | Chi-2 observé (K-K*)/K* |
|---------------------------|-----------------------|--------------------------|-------------------------|
| 0 | 52 | 50,66 | 0,035 |
| 1 | 29 | 34,45 | 0,862 |
| 2 et + | 19 | 14,36 | 1,499 |
| Total | 100 | 100 | 2,397 |

Extrait de la table du Chi-2 :

Nb de degrés de liberté = 2

Chi2 (2, 0,01) = 9,21

Chi2 (2, 0,05) = 5,99

Chi2 (2, 0,1) = 4,6

Chi2 (2, 0,2) = 3,22

- Test du Chi-2 pour la distribution d'acacias (zone 2)

| Nb d'acacias par quadrat N | Nb d'acacias observé K | Nb d'acacias théorique K* | Chi-2 observé (K-K*)/K* |
|----------------------------|------------------------|---------------------------|-------------------------|
| 0 | 25 | 4,16 | 104,45 |
| 1 | 19 | 13,22 | 2,52 |
| 2 | 13 | 21,03 | 3,06 |
| 3 | 10 | 22,29 | 6,77 |
| 4 et + | 33 | 39,29 | 1,01 |
| Total | 100 | 99,99 | 117,82 |

Extrait de la table du Chi-2 :

Nb de degrés de liberté = 4

Chi2 (4, 0,01) = 13,28

Chi2 (4, 0,05) = 9,49

Chi2 (4, 0,1) = 7,78

Chi2 (4, 0,2) = 5,99

Document 6

INTERET DU TEST DU CHI-2 L'EXEMPLE DE LA CENTRALE DE SELLAFIELD

« Dans le Nord de l'Angleterre, certains se sont vivement émus d'une relation possible entre les fuites radioactives de la centrale nucléaire de Sellafield et l'apparition de la leucémie chez l'enfant. Lorsque l'on représentait les cas d'enfants leucémiques sur une carte, des groupements apparaissaient à la périphérie de la centrale. Après de nombreuses enquêtes (...), un verdict d'absence de preuve (...) fut prononcé. Les statisticiens témoignant pour le gouvernement et la centrale nucléaire firent remarquer qu'une répartition causée par un processus purement aléatoire pouvait produire des groupements de points. De plus, en prenant une petite zone de référence, la densité d'enfants leucémiques était très élevée et très au-dessus de la moyenne nationale. Le même nombre d'enfants dans une zone plus vaste donnerait des valeurs plus basses. Les groupements observés étaient simplement le résultat d'un processus aléatoire et du choix arbitraire des aires d'étude. Ils n'étaient, sur le plan statistique et donc (?) scientifique, pas significatifs mais seulement un artefact de la façon avec laquelle ils avaient été estimés dans l'espace géographique... estimèrent les statisticiens au service du gouvernement.

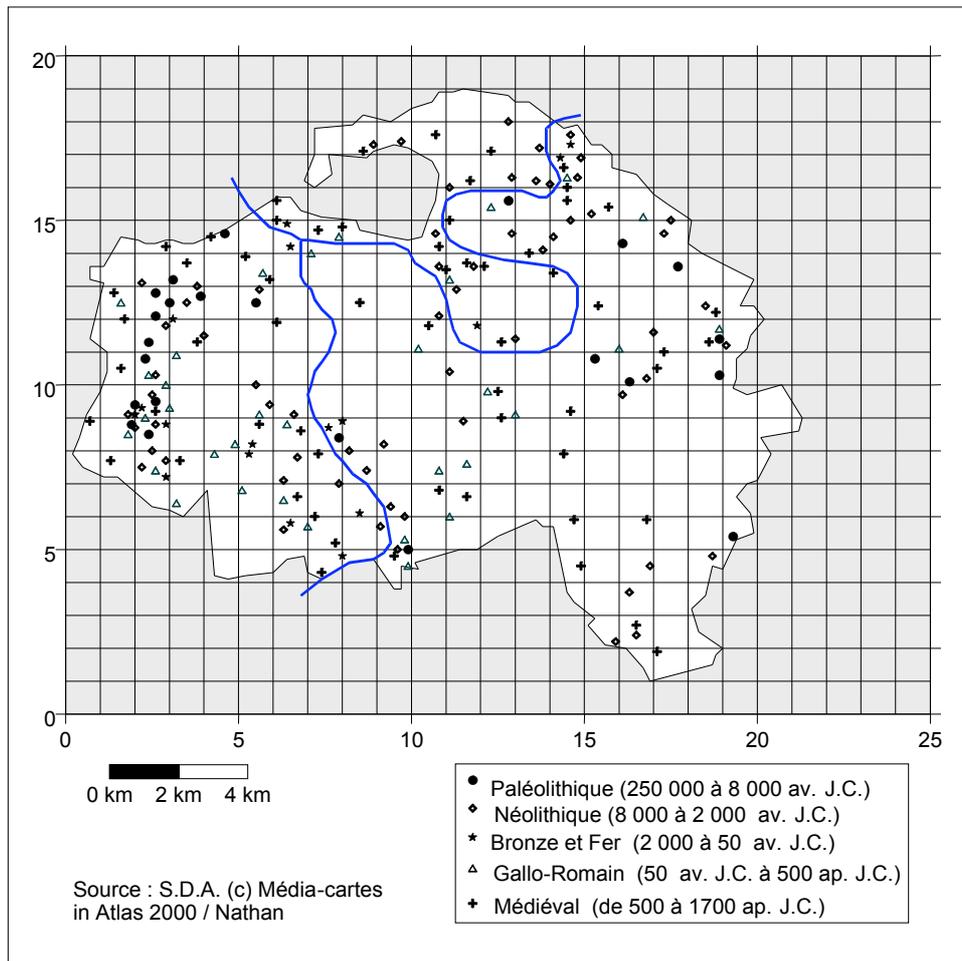
Cependant, cette conclusion ne satisfait pas le géographe Stan Openshaw (1988), l'un des chercheurs les plus créatifs pour fabriquer, grâce à la puissance des ordinateurs modernes, des béquilles à l'esprit humain. En utilisant la carte des cas de leucémie comme un système d'information géographique, il programma l'ordinateur pour parcourir la carte avec une petite grille circulaire, comme une machine à écrire se déplaçant sur une page caractère par caractère. A chaque point, l'ordinateur dénombrait les enfants dans une cellule de la grille et estimait la probabilité de trouver cette valeur par hasard d'après une distribution de Poisson. Si la probabilité était inférieure à 0,002, le traceur de l'ordinateur dessinait à cet endroit sur la carte une cellule de cette dimension et se déplaçait ensuite sur le point suivant, et ainsi de suite. A la fin d'un parcours complet, il repartait avec une grille aux mailles légèrement plus grandes, testait de nouveau chaque cellule, et ainsi de suite, inscrivant un cercle de la taille correspondante en chaque endroit de la carte où il était hautement improbable qu'un groupe de cas se constitue par hasard à cette échelle géographique. Ainsi, neuf millions « d'hypothèses spatiales » furent testées rigoureusement.

Les résultats étaient visibles : directement sur Sellafield, accumulés cellules après cellules, en une large tache noire, indiquant des groupements significatifs à presque toutes les échelles géographiques. Les regroupements n'étaient pas du tout un artefact mais étaient bien réels et dignes de susciter l'intérêt des scientifiques et des médecins (...). Mais l'analyse géographique produisit quelque chose de plus. Sur la côte Est de l'Angleterre septentrionale, il y avait sur la carte un autre pâté noir que personne n'avait imaginé. Au centre, il semblait y avoir un grand incinérateur urbain qui ne fonctionnait pas aux hautes températures pour lesquelles il avait été conçu, libérant ainsi des effluents plastiques et chimiques sur une vaste surface. »

Source : Gould P., 1992, « Epidémiologie et maladie », in Bailly, Ferras, Pumain (dir.), *Encyclopédie de Géographie*, Paris : Economica, pp. 958-959.

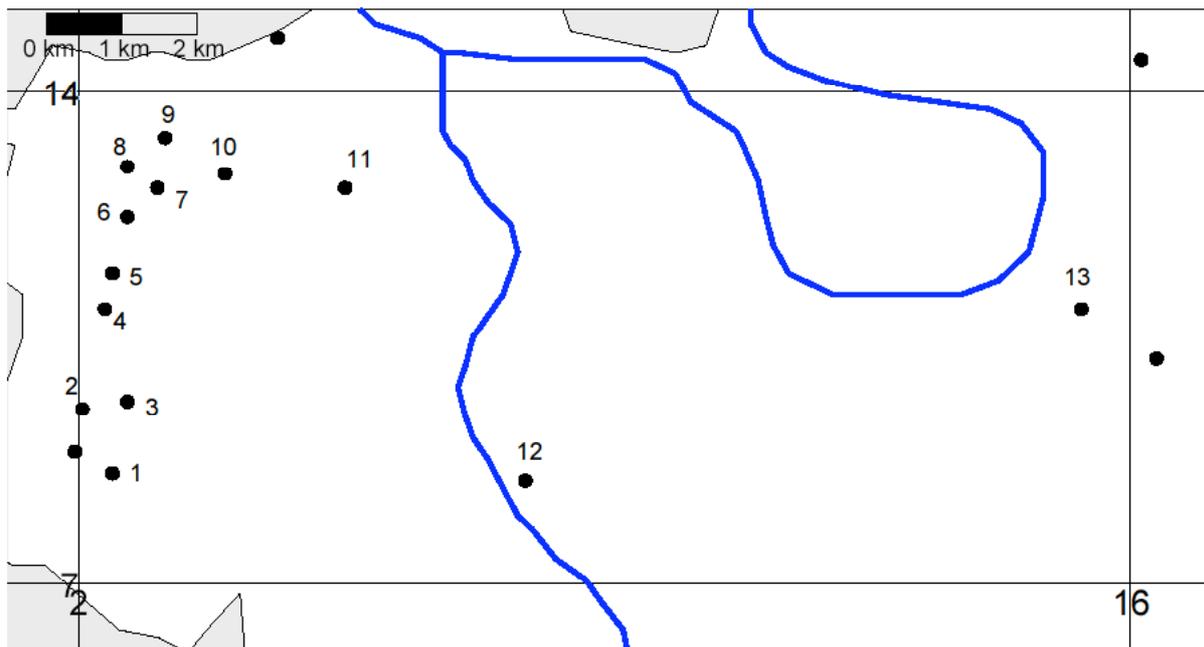
Document 7

CARTE DES SITES ARCHEOLOGIQUES DU VAL-DE-MARNE



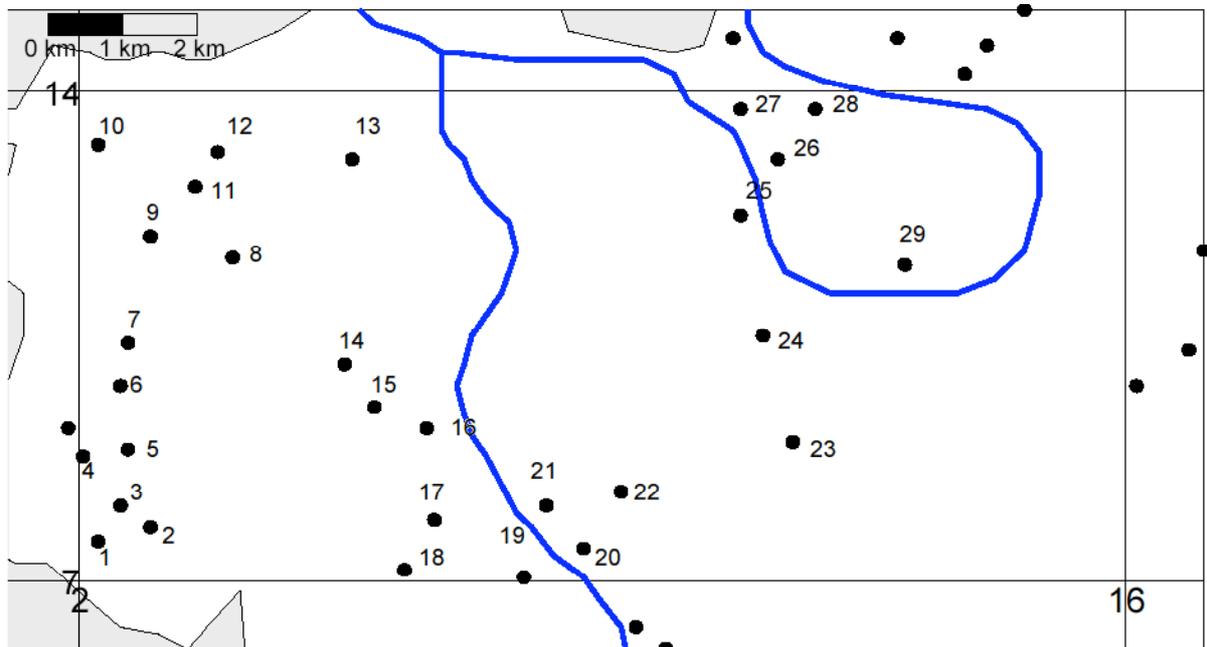
Document 8

EPOQUE PALEOLITHIQUE



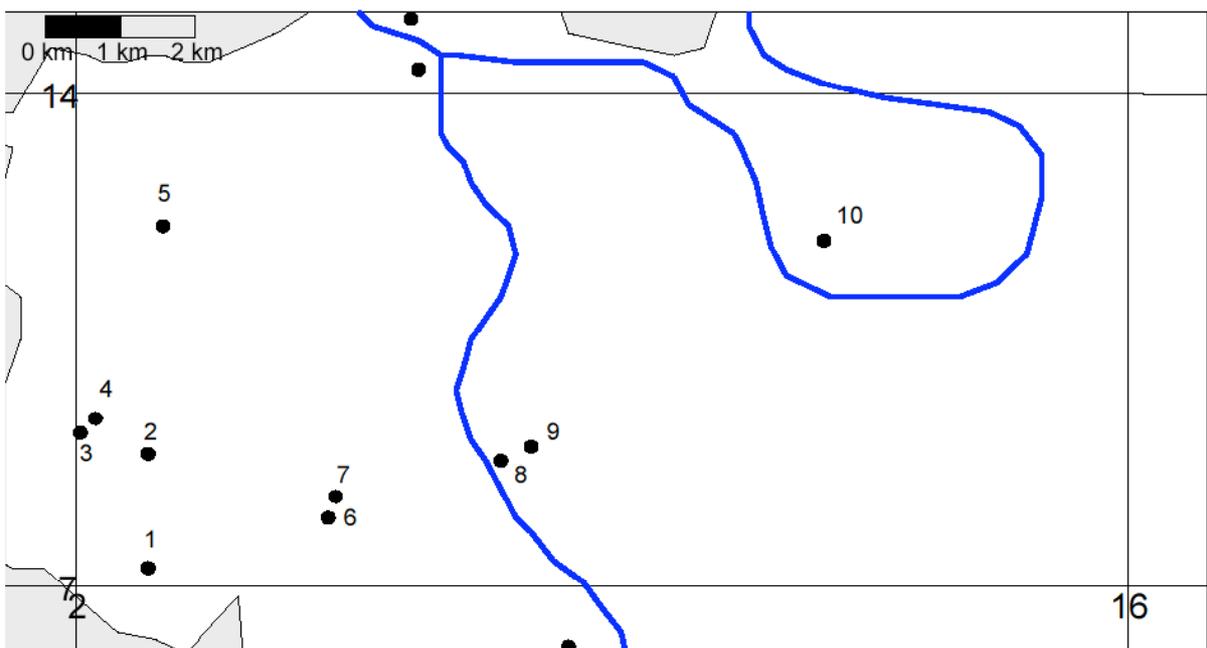
Document 9

EPOQUE NEOLITHIQUE



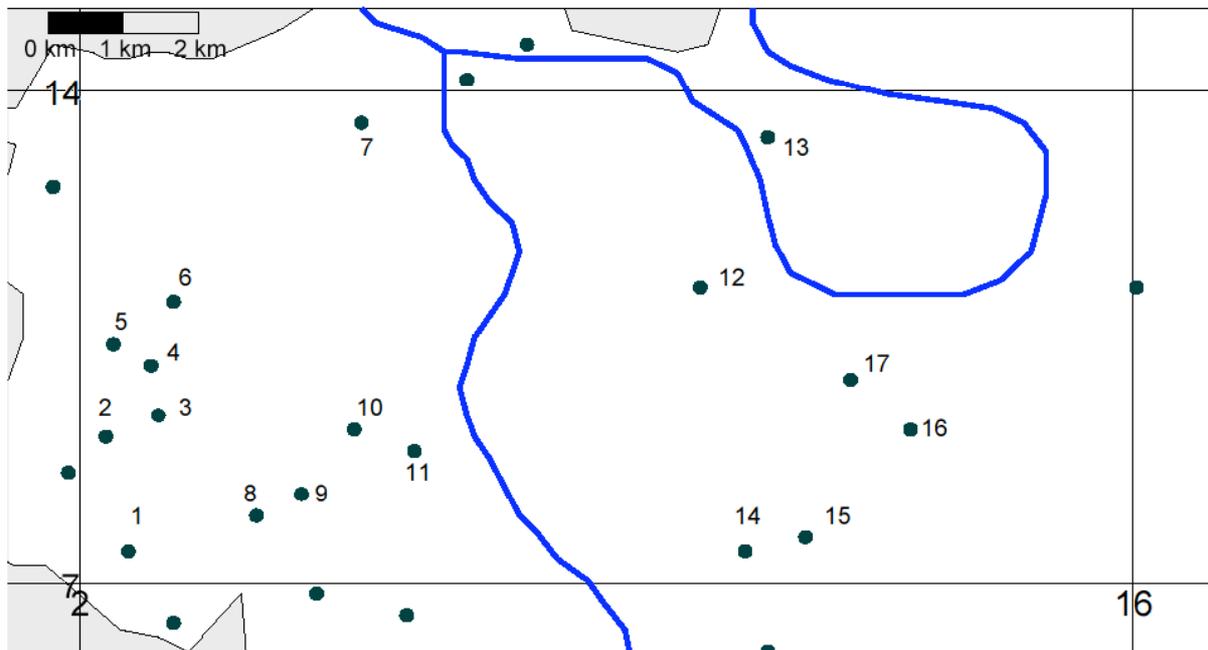
Document 10

EPOQUE « BRONZE ET FER »



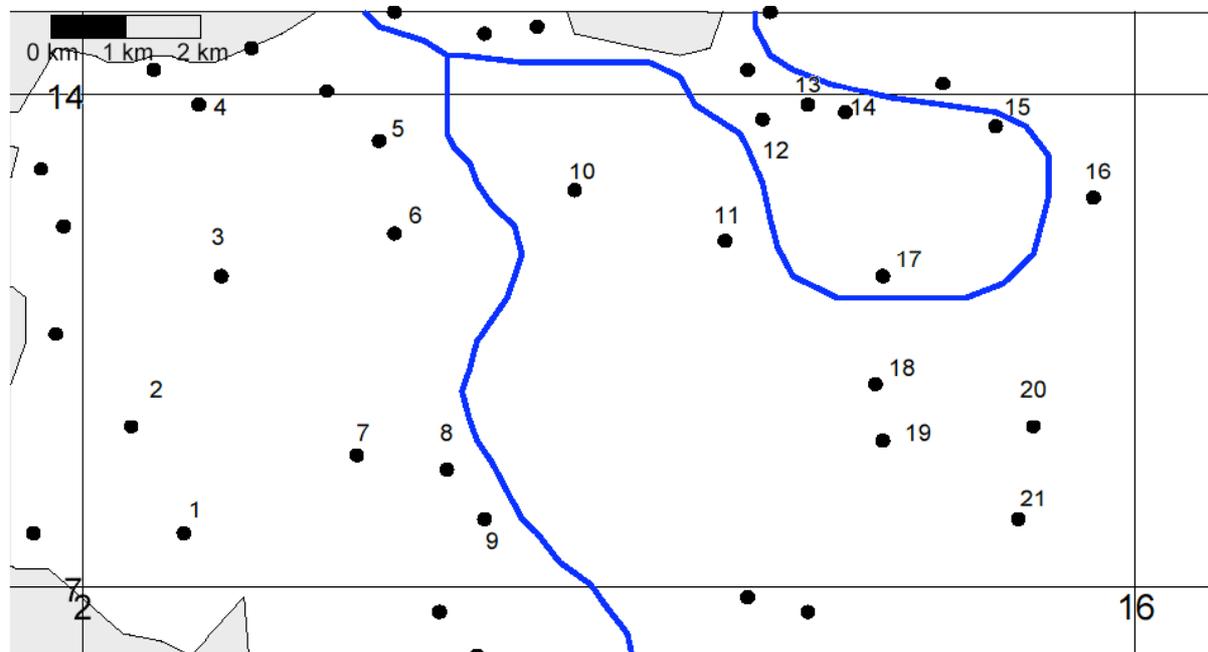
Document 11

EPOQUE GALLO-ROMAINE



Document 12

EPOQUE MEDIEVALE



Document 13

RELEVÉ DES DISTANCES AU PLUS PROCHE VOISIN POUR CHAQUE PÉRIODE (EN KM)

(a) Paléolithique

| n | Dmin(i) |
|----|---------|
| 1 | 1 |
| 2 | 0,5 |
| 3 | 1 |
| 4 | 0,5 |
| 5 | 0,5 |
| 6 | 0,5 |
| 7 | 0,5 |
| 8 | 0,5 |
| 9 | 0,6 |
| 10 | 1 |
| 11 | 1,6 |
| 12 | 4,6 |
| 13 | 1,2 |

Do 0,78

(b) Néolithique

| n | Dmin(i) |
|----|---------|
| 1 | 0,5 |
| 2 | 0,5 |
| 3 | 0,5 |
| 4 | 0,6 |
| 5 | 0,4 |
| 6 | 0,5 |
| 7 | 0,5 |
| 8 | 1,1 |
| 9 | 1 |
| 10 | 0,6 |
| 11 | 0,6 |
| 12 | 1,4 |
| 13 | 1,8 |
| 14 | 0,7 |
| 15 | 0,7 |
| 16 | 0,7 |
| 17 | 0,8 |
| 18 | 0,8 |
| 19 | 0,9 |
| 20 | 0,8 |
| 21 | 0,8 |
| 22 | 0,9 |
| 23 | 1,4 |
| 24 | 1,4 |
| 25 | 2,1 |
| 26 | 0,9 |
| 27 | 0,8 |
| 28 | 0,8 |
| 29 | 0,8 |

Do 0,87

(c) Bronze/Fer

| n | Dmin(i) |
|----|---------|
| 1 | 1,5 |
| 2 | 0,8 |
| 3 | 0,3 |
| 4 | 0,3 |
| 5 | 2,7 |
| 6 | 0,3 |
| 7 | 0,3 |
| 8 | 0,4 |
| 9 | 0,4 |
| 10 | 4,7 |

Do 1,17

(d) Gallo-Rom.

| n | Dmin(i) |
|----|---------|
| 1 | 0,6 |
| 2 | 0,7 |
| 3 | 0,7 |
| 4 | 0,6 |
| 5 | 0,6 |
| 6 | 0,9 |
| 7 | 0,6 |
| 8 | 0,6 |
| 9 | 0,8 |
| 10 | 0,8 |
| 11 | 0,8 |
| 12 | 0,8 |
| 13 | 1 |
| 14 | 1 |
| 15 | 2,2 |
| 16 | 2,2 |
| 17 | 1,5 |

Do 0,96

(e) Médiéval

| n | Dmin(i) |
|----|---------|
| 1 | 1,6 |
| 2 | 1,5 |
| 3 | 1,2 |
| 4 | 0,6 |
| 5 | 0,6 |
| 6 | 2,4 |
| 7 | 1,2 |
| 8 | 1 |
| 9 | 0,7 |
| 10 | 2,1 |
| 11 | 1,7 |
| 12 | 0,6 |
| 13 | 0,5 |
| 14 | 0,5 |
| 15 | 1,4 |
| 16 | 0,9 |
| 17 | 1,6 |
| 18 | 0,7 |
| 19 | 1,2 |
| 20 | 0,7 |
| 21 | 1,2 |

Do 1,14

Document 14

SYNTHÈSE DES DISTRIBUTIONS DE DISTANCES AU PLUS PROCHE VOISIN

| Epoque | nb. de sites (n) | surface (s) | densité (d) | Dist. Théor. Dt | Dist. Obs. Do | Indice Do/Dt |
|---------------|------------------|-------------|-------------|-----------------|---------------|--------------|
| PALEOLITHIQUE | 13 | 98 | 0,13 | 1,37 | 1,08 | 0,78 |
| NEOLITHIQUE | 29 | 98 | 0,3 | 0,92 | 0,87 | 0,95 |
| BRONZE ET FER | 10 | 98 | 0,102 | 1,57 | 1,17 | 0,75 |
| GALLO-ROMAINE | 17 | 98 | 0,17 | 1,2 | 0,96 | 0,8 |
| MEDIEVALE | 21 | 98 | 0,21 | 1,08 | 1,14 | 1,05 |

Do = Distance observée moyenne au plus proche voisin

Dt = Distance théorique moyenne au plus proche voisin, compte tenu de la densité d du semis de points

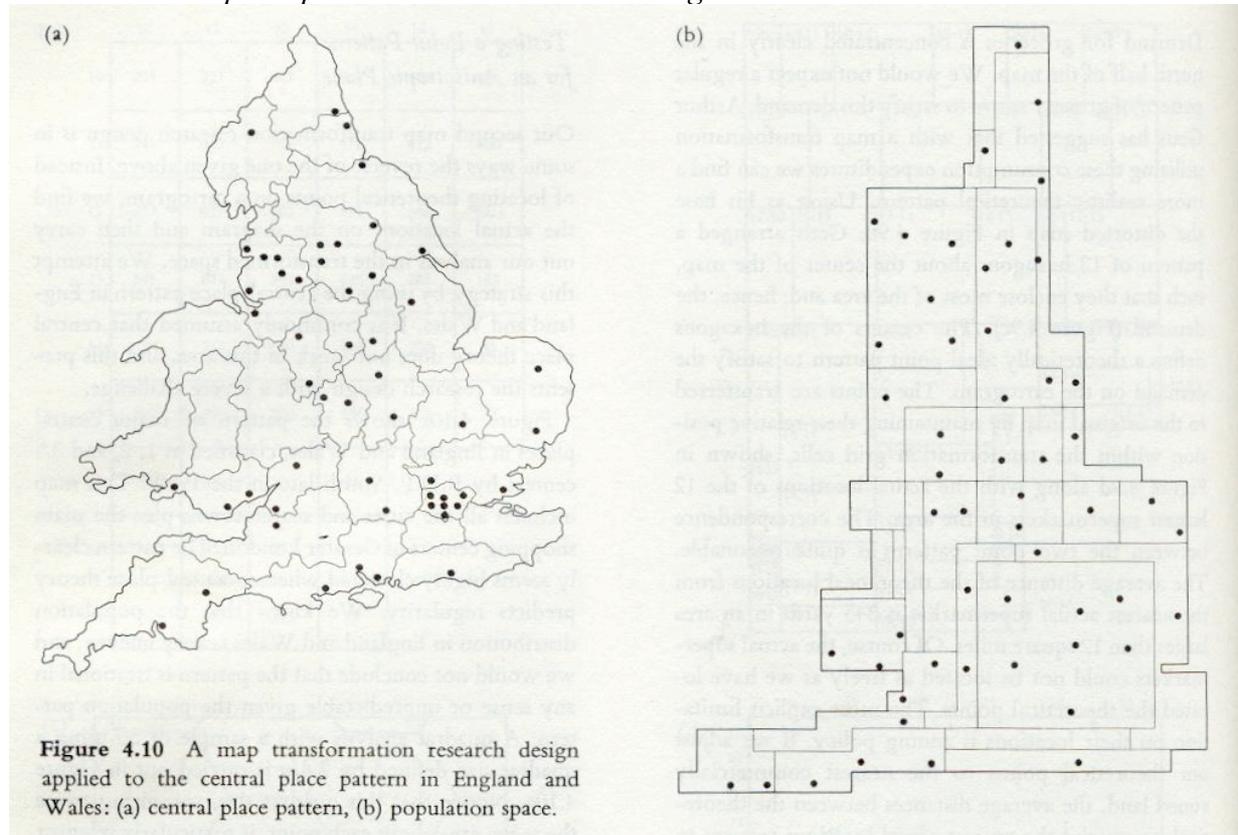
Rappel de la formule correspondant au calcul de cette distance théorique, dans le cas où la forme du semis est issue d'un processus aléatoire :

$$Dt = 0,5 / \sqrt{d}$$

Document 15

EXEMPLE D'EXERCICE POUR UN DEVOIR SUR TABLE

Distribution des principaux centres commerciaux anglais vers 1970



Source : Taylor P.J., 1977, *Quantitative methods in geography*, Waveland press, p. 154.

Le géographe anglais P.J. Taylor s'est intéressé à la distribution des principaux centres commerciaux en Angleterre au début des années 1970. Il a pu démontrer que leur distribution n'était pas aléatoire si on l'examinait sur une carte ordinaire (a) mais qu'elle ne l'était pas non plus si on l'examinait sur une carte anamorphosée où les surfaces des régions sont proportionnelles à leur population (b).

Questions :

- (1) Sachant qu'elles ne sont pas aléatoires, de quel type sont les distributions (a) et (b) ?
- (2) **Sans faire de calculs**, indiquez la démarche qu'a pu utiliser P.J. Taylor pour démontrer que les deux distributions ne sont pas aléatoires.
- (3) Quel est l'intérêt théorique de l'étude de P.J. Taylor ?