



Postface. Traitement des données historiques spatialisées. Que faire ? et comment ?

Alain Guerreau



Édition électronique

URL : <http://journals.openedition.org/cem/13840>
DOI : 10.4000/cem.13840
ISSN : 1954-3093

Éditeur

Centre d'études médiévales Saint-Germain d'Auxerre

Référence électronique

Alain Guerreau, « Postface. Traitement des données historiques spatialisées. Que faire ? et comment ? », *Bulletin du centre d'études médiévales d'Auxerre | BUCEMA* [En ligne], Hors-série n° 9 | 2016, mis en ligne le 15 décembre 2015, consulté le 04 mai 2019. URL : <http://journals.openedition.org/cem/13840> ; DOI : 10.4000/cem.13840

Ce document a été généré automatiquement le 4 mai 2019.



Les contenus du *Bulletin du centre d'études médiévales d'Auxerre (BUCEMA)* sont mis à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

Postface. Traitement des données historiques spatialisées. Que faire ? et comment ?

Alain Guerreau

- 1 Des cartes, des cartes... Des cartes ? Curieusement, il faut bien le dire, on voit se développer, lentement, des recherches qui tentent d'utiliser les coordonnées spatiales des données historiques, tandis que, dans le même temps, on s'aperçoit que les « grandes revues », depuis longtemps, ne publient plus de cartes (contrairement à ce que l'on a pu observer dans les années 50 et 60). La situation est un peu étrange. Le dossier rassemblé à l'occasion de journées d'études organisées par Marie-José Gasse-Grandjean et Laure Saligny témoigne assez bien de courants actuels qui, quoique encore marginaux, émergent progressivement. Soulignons que, lors de cette réunion dijonnaise, le public comportait non seulement des historiens, au sens restreint, mais aussi de nombreux archéologues et des conservateurs d'archives.

Évolution

- 2 Bien que la tradition française associe depuis le 19^e siècle l'histoire et la géographie, on ne peut pas dire que les historiens français, et encore moins leurs collègues européens, soient vraiment préoccupés de l'organisation spatiale des sociétés qu'ils étudient. Charles-Victor Langlois et Charles Seignobos ne se souciaient guère de géographie, et les ouvrages des années 70 censés présenter la « nouvelle histoire » n'ont pas fait mieux. Quelques travaux où la géographie jouait un rôle important (Marc Bloch, Maurice Lombard, Pierre Chaunu¹) sont restés sans postérité.
- 3 Les premiers à se préoccuper de l'organisation spatiale des sociétés furent des archéologues, et spécialement des archéologues anglo-saxons. L'ouvrage de Ian Hodder et Clive Orton, *Spatial Analysis in Archaeology* (1976) marqua un véritable tournant². Ces deux auteurs furent les premiers à considérer méthodiquement les coordonnées des objets et

des sites et à traiter cette matière à l'aide d'ordinateurs capables de mettre en œuvre des procédures statistiques relativement complexes. De nombreux préhistoriens se tournèrent alors vers ces méthodes. En France, François Djindjian fut à la fois un pionnier et un organisateur. Ses manuels constituent une référence essentielle³.

- 4 La géographie suivit, avec juste quelques années de retard. La tradition d'une géographie prétendument « humaniste » et qui n'était que littéraire (médiocrement), résista. La seule carte étudiée était la carte géologique (indiscutablement utile) et la seule analyse spatiale connue celle de la « géomorphologie », pratique assez étrange, obligatoire pour tous les étudiants de géographie, mais n'ayant à peu près aucun rapport avec les réalités sociales. Pourtant de jeunes géographes, regroupés autour du mot d'ordre d'« analyse spatiale », réussirent à recentrer leur discipline autour de l'analyse statistique des données spatio-sociales. Une nouvelle géographie avait vu le jour (années 90), non sans conflits : *'La géographie a eu beaucoup de difficulté à passer des notions aux concepts et s'est longtemps satisfaite d'un ensemble de notions imprécises (villes, paysages, etc.) qu'elle employait sans véritablement les définir. La quantification des phénomènes a provoqué une crise puisque la mesure des phénomènes obligeait à en donner des définitions précises (ex. taille des villes, biomasse, etc.)'*⁴.
- 5 Il est un peu triste d'avoir à reconnaître que rien de tel n'est encore advenu chez les historiens. La plupart continuent imperturbablement à faire comme si les sociétés planaient hors sol, comme les angelots des tableaux d'Albrecht Altdorfer. Les actes du colloque de la SHMESP tenu à Mulhouse en 2006, *Construction de l'espace au Moyen Age*⁵, me semblent assez représentatifs. Le volume se compose de 28 textes et compte 453 pages. On y trouve en tout et pour tout 29 croquis et plans, seuls trois articles comportent quelque chose qui s'apparente à une analyse spatiale élémentaire, et je ne vois employée aucune analyse statistique... J'ai moi-même présenté, dès le début des années 80, des recherches où la variable spatiale jouait un rôle-clé⁶, et j'ai tenté en 1992 de montrer schématiquement comment on pouvait essayer de reconstruire la représentation de l'espace qui avait cours dans l'Europe médiévale⁷. On le constate aisément : sur ce dernier point, j'ai été largement suivi, mais les analyses statistiques ne sont jamais citées, sinon par des géographes.
- 6 Les ressources disponibles ne cessent de s'accroître, cela seul pourrait inciter étudiants et jeunes chercheurs historiens à orienter leurs efforts dans cette direction ; notons cependant que les obstacles restent solides et que la réflexion sur les moyens de les contourner n'est pas très avancée. Voyons cela d'un peu plus près.

Les sources : de quoi part-on ?

- 7 Traiter des données historiques spatialisées, cela signifie - au sens propre - que l'on cherche à tenir compte du fait que les faits sociaux (tous les faits historiques sont des faits sociaux) que l'on analyse sont munis de coordonnées géographiques. On peut, avec certaines restrictions, étendre à des faits dont on connaît seulement les relations topologiques (voisinage ou non, continuité ou rupture)⁸. A partir de cette définition, on peut distinguer trois catégories de sources :
 - les documents cartographiques
 - les documents incluant des indications fiables à des coordonnées (en général des toponymes)

- des objets-documents que l'on peut munir de coordonnées, e.g. bâtiments, trouvailles archéologiques, textes dont on connaît le lieu de composition, etc.
 - Le point nodal est constitué dans tous les cas par les coordonnées géographiques, sans lesquelles on ne peut que très latéralement procéder à un traitement spatial. Problème : ces documents apparaissent sans coordonnées, il faut donc les en munir, et ce n'est pas trivial.
 - pour les cartes dites « anciennes », il faut procéder à leur géolocalisation ; ce n'est pas très difficile⁹, mais long et fastidieux ; si l'on part de cartes assez simples, géométriquement exactes, on peut, avec de l'expérience et en se concentrant, géolocaliser une carte en 10 mn ; mais si l'on part d'une carte non orientée, dont on ne connaît pas exactement l'emplacement, et peu exacte géométriquement, cela peut prendre une demi-heure, voire davantage.
 - la question de la géolocalisation des toponymes est un monde en soi ! s'il s'agit juste du centroïde des communes, cela peut aller très vite, on peut en général procéder automatiquement¹⁰ ; mais s'il s'agit de la forme latine, dans une charte, d'un micro-toponyme inconnu a priori, on peut y passer la journée, sans être sûr d'aboutir à un résultat.
 - la question des objets-documents est peut-être la plus simple : si l'on connaît l'emplacement avec assez de précision, on peut géolocaliser plusieurs centaines d'objets en une journée¹¹.
- 8 La difficulté s'accroît s'il s'agit de géolocaliser des entités surfaciques, circonscriptions diverses, zones, parcelles. Là, il faut créer des polygones et leur attacher les objets considérés. La question des flux et des réseaux peut déboucher sur de réelles difficultés techniques.
- 9 On voit parfaitement que dans tous les cas se pose le problème de la saisie, et que celle-ci, dans le meilleur des cas, sera dite longue et pénible ; dans bien des circonstances, il faudra résoudre diverses difficultés pratiques, de type traditionnel (identification) ou récent (codage informatique). Avant toute recherche, il est impératif de bien évaluer le temps moyen de saisie par unité d'information, de manière à calibrer l'ampleur globale de cette saisie. La plupart des recherches subissant une contrainte de temps, il faut toujours veiller à ce que cette phase de saisie ne dépasse pas le quart de la durée totale disponible, sans quoi on finit par se trouver avec une masse de données que l'on n'a plus le temps de traiter convenablement. A cet égard, rien ne remplace une bonne pratique statistique, qui permet de se faire une idée de la masse d'information minimale pour permettre l'application des procédures adéquates à l'objectif de la recherche. Dans certaines conditions, il peut être intéressant de recourir à l'échantillonnage¹².

Des exemples très pédagogiques

- 10 Je saute à l'autre bout de la chaîne, les résultats. Le dossier présenté ici comporte plusieurs cas très parlants. Je n'en retiens que trois.
- 11 Un exemple particulièrement simple, mais frappant, est celui de la carte des dates d'arrivée du sécateur dans les villages viticoles bourguignons (Grillon-Garcia, fig. 10). Il s'agit d'une carte brute de semis de points. En utilisant des symboles différents pour les périodes successives, on s'aperçoit au premier coup d'œil que le sécateur est arrivé essentiellement du nord au sud, le Brionnais étant toutefois plus précoce que ce que l'on aurait pu imaginer¹³. Bien entendu, cette constatation, simple et massive, n'est qu'un point de départ, qui permet de construire de premières hypothèses que d'autres recherches devraient permettre d'étayer ou de rejeter.

- 12 Dans le cas des très intéressantes analyses de la forêt de Chailluz, C. Fruchart montre avec quelles méthodes on peut parvenir à reconstituer les états successifs du réseau viaire, et comment la caractérisation de ces états entraîne la construction d'hypothèses explicatives liées à l'histoire générale de la région (Fruchart, fig. 16, 17, 18). La combinaison, dans un cadre cartographique, d'un ensemble d'informations, topographiques, archéologiques, cartographiques, toponymiques et documentaires permet de faire ressurgir une structure totalement invisible en lecture directe.
- 13 Les enquêtes de N. Perreaux frappent par leur ampleur, mais le principe ici mis en œuvre est assez voisin : en projetant sur le même fond de carte européen la chronologie des rédactions de chartes d'un côté, et la répartition des églises 'romanes' de l'autre, on s'aperçoit de l'étonnante coïncidence entre la répartition de ces églises et une période précise dans la chronologie des chartes (groupes I et II, 10^e-11^e siècles) (Perreaux, fig. 8 et 16). Cette observation conforte vivement une hypothèse émise antérieurement, celle d'une relation entre les flux de donations aux institutions ecclésiastiques et les constructions (ou reconstructions) de bâtiments du culte.
- 14 On voit ce que ces trois exemples ont en commun : la simple création d'une carte fait apparaître une structure spatiale, antérieurement invisible (donc inconnue), structure étroitement liée à l'organisation sociale dont en fait elle n'est qu'un élément. Il n'est pas difficile d'imaginer que des analyses de ce type pourraient être multipliées en très grand nombre, permettant de substantiels progrès de nos connaissances.

Outils et méthodes

- 15 Le dossier montre donc bien qu'une concentration de l'attention de l'historien sur la variable spatiale des relations sociales aboutit à des résultats très novateurs. Et donc que ce domaine, jusqu'ici quasi embryonnaire, mérite d'être fortement promu. Ce point bien établi, il reste à se demander quels outils et quelles méthodes utiliser.
- 16 Il faut insister au préalable sur un point clé : la cartographie-géomatique ne remplace en aucune manière l'analyse statistique en général ni l'analyse statistique spatiale en particulier. D'une manière très générale, les méthodes graphiques, qui rendent des services considérables, ne sont pas là pour faire le travail des analyses statistiques, elles n'ont pas du tout la même puissance ; elles interviennent dans certaines phases de la démarche, mais seulement en appui et en contrepoint. La confusion est en grande partie la résultante d'une situation antérieure : avant l'arrivée des ordinateurs, les calculs étaient longs et peu efficaces, on avait recours à diverses procédures graphiques, les équations étaient présentées sous forme d'abaques. Mais tout ceci est totalement obsolète : les graphiques ont gagné en puissance grâce à tous les logiciels graphiques, mais les statistiques proprement dites ont totalement changé de nature : un outil est tout simplement né qui n'existait pas antérieurement.
- 17 La question des outils pouvait encore paraître indéfinie voici une dizaine d'années. Elle ne l'est plus, même si beaucoup de collègues ne s'en sont pas aperçus. A présent, qui dit cartographie dit Système d'Information Géographique (SIG), d'ailleurs on ne dit plus « cartographie » mais « géomatique ». De fait, pour réaliser une carte il fallait être bon dessinateur, aujourd'hui il faut des compétences informatiques¹⁴. Dans le cadre de la recherche publique, s'agissant surtout des étudiants et des jeunes chercheurs, il n'y a qu'un seul choix raisonnable, QGIS¹⁵. Ce logiciel libre et open-source est désormais

employé massivement par les professionnels (même dans des entreprises et des organismes où l'inertie fait que l'on continue aussi d'utiliser des solutions propriétaires excessivement onéreuses). QGIS dispose à lui seul d'une multitude d'outils, et des greffons (plugins) en quantité croissante (plus de 500 à la fin de 2015) facilitent l'emploi d'une multitude de procédures ; en outre QGIS peut être utilisé comme interface pour d'autres logiciels d'analyse spatiale plus spécifiques (Grass, Saga, R). La prise en mains de ce logiciel est relativement aisée, le manuel est copieux et l'on trouve en ligne de très nombreux tutoriels efficaces.

- 18 Deux autres types de logiciels sont indispensables si l'on ne veut pas se contenter de faire des cartes, mais procéder à des analyses géostatistiques. D'un côté une base de données ; on a le choix entre PostgreSQL¹⁶ et SQLite¹⁷, deux logiciels interfacés très efficacement avec QGIS. SQLite est très largement suffisant pour les données habituelles, même très copieuses, et est d'un emploi très simple. Enfin, on ne peut en aucun cas se passer de R si l'on veut faire réellement des calculs¹⁸. C'est aujourd'hui LE logiciel statistique, il est libre et open-source ; des dizaines de paquets de statistiques spatiales sont disponibles, on n'a que l'embarras du choix¹⁹. L'utilisation de R est bien plus simple que ce que prétendent ses détracteurs²⁰, il existe aujourd'hui des environnements graphiques intégrés d'une grande commodité (Rkward²¹ et RStudio²²).
- 19 Avec cette panoplie, que faire ? Dans certains cas (comme les trois mentionnés plus haut), une carte bien conçue, représentant un phénomène précis, peut être considérée comme un résultat ; à condition qu'elle fasse vraiment apparaître une relation ou un élément de structure sociale original. Mais en règle générale, des traitements sont nécessaires. Tout simplement parce que le plus souvent une traduction brute des données sur un fond de carte ne montre rien ou pas grand chose ; ou même parce qu'une traduction de ce genre peut produire une impression fallacieuse²³.
- 20 Il faut généralement commencer par des analyses statistiques classiques, en dehors des coordonnées spatiales. C'est une étape indispensable pour commencer à apercevoir les grandes répartitions, et les propriétés des distributions des caractères des objets ; sauter par-dessus cette phase, c'est presque nécessairement se condamner à l'erreur ; en passant directement à l'analyse de la variable spatiale, on a toutes les chances d'interpréter en termes spatiaux des phénomènes qui ne le sont nullement : on perd sur tous les tableaux ; on ne voit pas des structures importantes, et l'on finit par passer à côté des structures proprement spatiales. Le cas des parcelles est tout à fait éclairant : dès lors que l'on dispose d'informations (qualitatives et quantitatives) sur les parcelles correspondant à une zone donnée, il faut d'abord procéder à toutes les analyses hors-espace, c'est la seule manière de faire apparaître des gradients, des similitudes et des oppositions, qui peuvent aboutir à la construction de classes. Les analyses multivariées sont incontournables. C'est seulement ensuite que l'on peut passer à l'analyse d'une matrice de contiguïté, qui peut permettre de « régionaliser » l'espace considéré, et ainsi d'ajouter une série supplémentaire de classes. Enfin (enfin !), on peut tenter d'utiliser ce que l'on sait de la surface et des coordonnées de ces parcelles, pour essayer de reconstruire la structure de l'organisation spatiale du territoire considéré. 1. On n'y arrivera pas si l'on n'a pas identifié précédemment des classes intrinsèques 2. La cartographie devra mettre en œuvre diverses procédures, c'est peut-être alors que des graphes peuvent servir, on n'hésitera pas non plus à tester diverses méthodes d'anamorphose.
- 21 Et ce qui vient d'être dit à propos de l'instant t est encore bien plus important s'il s'agit de comparer les instants t , $t+1$, $t+n$. Sauf dans le cas de cartes de répartition ultra-simples, la

simple comparaison visuelle est plus un piège qu'une solution. Il est absolument nécessaire de commencer par comparer les populations hors variable spatiale ; sinon, on attribue à une évolution spatiale une évolution dans la structure qui n'a rien de spatial : contresens garanti.

- 22 Ces quelques remarques ponctuelles ne visent en aucune manière à fournir un survol des statistiques spatiales et encore moins de l'ensemble des traitements de variables historiques spatialisées : il faudrait plusieurs centaines de pages. Et, ici, on doit souligner la lacune catastrophique qui résulte de l'absence de tout manuel pour le traitement des données historiques spatialisées. Lacune qui d'ailleurs n'est qu'un élément d'une lacune plus vaste et encore plus désastreuse, celle de tout manuel de statistiques historiques digne de ce nom... On peut utiliser des manuels rédigés par des géographes, mais ce n'est qu'un pis-aller²⁴. La résistance tenace, voire hargneuse, que la plupart des historiens opposent à l'utilisation des méthodes statistiques, visible depuis longtemps, ne paraît pas faiblir. Manifestement, ces collègues, qui sentent intuitivement que le bateau se délite, préfèrent le suicide à ce qu'ils s'imaginent comme une rupture de leur pratique littéraire-subjective.
- 23 Je voudrais terminer en complétant ces considérations générales par quelques remarques relatives à trois objets précis : la statistique spatiale, la numérisation des sources des données spatiales, la toponymie.

Statistique spatiale

- 24 Il n'est pas inutile de savoir que la statistique spatiale est très tard venue. Une variable spatiale est composée de deux ensembles de nombres (appariés), ce qui en fait un objet à divers égards assez voisin d'une variable complexe. Et l'on sait que la manipulation des nombres complexes n'est pas vraiment un jeu d'enfants, les formules sont lourdes, si bien que l'utilisation pratique de ce type de variable requiert des outils de calculs vraiment puissants. Lesquels n'ont guère été disponibles avant la fin des années 60 et le début des années 70. C'est seulement à partir de là qu'il est devenu intéressant de développer des recherches dans ce domaine : antérieurement, les meilleures formules étaient condamnées à demeurer de beaux objets abstraits.
- 25 A partir de ce moment, les recherches sont allées bon train, car de nombreux domaines étaient très intéressés : météorologie, épidémiologie, bio-statistique et écologie, recherche minière, notamment. L'énorme quantité de données de tous types collectées par les satellites nécessitait des outils adaptés. Aujourd'hui, la plupart des algorithmes sont parfaitement implémentés dans des solutions libres. La seule difficulté est d'en comprendre le sens et la portée. On trouve sur internet pléthore d'information en accès libre, depuis les « grands manuels » fondateurs²⁵ jusqu'à des présentations pédagogiques rapides et synthétiques²⁶. Les recherches sont dans une très large mesure conduites dans le cadre de disciplines dont le matériel comporte une dimension spatiale importante. Pour un historien, les disciplines les plus intéressantes sont celles dont le matériau est situé à la fois dans l'espace et dans le temps. C'est tout particulièrement le cas de la bio-statistique et de l'écologie quantitative, raison pour laquelle les logiciels conçus dans ce cadre sont parmi les plus directement utilisables par l'historien²⁷.
- 26 J'ajoute une brève remarque sur une méthode relativement récente, l'anamorphose. Une carte confère aux phénomènes une valeur relative proportionnelle à la surface qu'ils

occupent. Dans bien des cas, c'est pertinent : les cartes des températures ou des précipitations représentent des phénomènes répartis sur toutes les surfaces considérées. Mais quand on passe aux structures sociales, cela devient beaucoup moins simple. On l'a noté depuis longtemps à propos de la géographie électorale. On représente couramment les résultats des élections en affectant les départements d'une couleur choisie dans une échelle. Le résultat, visuellement, est curieux : la Lozère est plus visible que la région parisienne. On a donc eu depuis assez longtemps l'idée qu'il faudrait trouver un moyen graphique qui respecte autant que possible la géographie, mais qui tienne compte de la quantité de population. Diverses solutions plus ou moins difficiles à mettre en œuvre ont été expérimentées, jusqu'à ce qu'en 2004 deux chercheurs américains, Michael Gastner et Mark Newman, publient un nouvel algorithme particulièrement bien conçu et efficace²⁸. On trouve aujourd'hui des logiciels libres qui implémentent cette procédure, un paquet R est disponible²⁹. On obtient des « cartes » fort suggestives (on dit plutôt dans ce cas « cartogrammes »)³⁰. On retient le plus souvent la population comme variable sous-jacente, mais d'autres sont tout aussi possibles. Dans le cas d'un parcellaire géolocalisé par exemple, on pourrait vouloir représenter les parcelles non pas proportionnellement à leur surface, mais à leur revenu (par exemple utiliser la valeur - ou le revenu - indiqués par l'état des sections du cadastre napoléonien, en les rapportant à la surface) : on gagnerait probablement de la sorte une tout autre vision du terroir.

- 27 D'une manière générale, l'espace habité et vécu n'est pas du tout isomorphe à l'espace physique, et d'ailleurs plusieurs des études présentées dans ce dossier le montrent bien, en particulier celle de N. Poirier. La question des clochers de village a joué un rôle considérable dans toute l'Europe médiévale : ce n'est pas un hasard si la coutume de Lorris, qui a eu l'importance que l'on sait, définit des espaces par rapport aux clochers. En examinant en grandeur réelle de quels endroits d'une zone on voit, ou ne voit pas les clochers, en cartographiant d'autre part, les zones où l'on entend, ou pas, les cloches, N. Poirier montre bien l'écart entre le schéma abstrait et la réalité empirique. Il montre ainsi de manière très suggestive pourquoi et dans quelles conditions l'espace conçu et imaginé l'emportait sur l'espace réellement vécu (que l'espace conçu organisait).

Les sources numérisées en accès libre

- 28 Les sources historiques de tout type arrivent en masses croissantes sur internet, en mode texte ou en mode image. Jusqu'ici, le mouvement ne paraît pas se ralentir. Dans cet immense ensemble, les cartes anciennes tiennent une place importante. En France, les deux acteurs principaux sont l'IGN et les Archives Départementales. Dans la grande majorité des départements, les feuilles de plan du cadastre napoléonien ont été numérisées et sont accessibles en ligne. Et dans la quasi-totalité des départements qui manquent encore à l'appel, le travail est entrepris. Il s'agit là d'un fonds documentaire gigantesque. L'IGN offre au téléchargement libre trois groupes de documents : les feuilles de la « carte de Cassini », les feuilles de travail originales de la carte d'État-Major au 1/80000, et un stock extraordinaire de photographies aériennes de la France dans la seconde moitié du 20^e siècle (quelques photographies légèrement antérieures à 1939, la grande densité commence vers 1950), le tout en très haute définition³¹.
- 29 C'est peu de dire que l'on croûle sous l'abondance, surtout si l'on note qu'aucune de ces images n'est géolocalisée. On manque encore de recul et d'expérience.
- 30 Un autre aspect en plein essor est l'apparition à un rythme soutenu de serveurs WMS.

- 31 Web Map Service ou WMS est un protocole de communication standard qui permet d'obtenir des cartes de données géoréférencées à partir de différents serveurs de données. Cela permet de mettre en place un réseau de serveurs cartographiques à partir desquels des clients peuvent construire des cartes interactives. Le WMS est décrit dans des spécifications maintenues par l'Open Geospatial Consortium. Un service WMS sert à retourner une image visualisable sur un écran d'ordinateur. Le serveur produit des cartes aux formats image comme le JPEG, le PNG ou le GIF, ou sous forme d'éléments vecteurs comme le SVG.³²
- 32 En pratique, cela signifie qu'à partir d'un SIG fonctionnant sur un ordinateur personnel relié à internet, on peut afficher des cartes (images géolocalisées) de la zone que l'on choisit³³. Il existe à présent des sites spécialement dédiés au recensement des serveurs WMS disponibles³⁴. Naturellement, tout le monde connaît les cartes affichées par OSM (OpenStreetMap), GoogleMaps, voire Bing. On les examine ordinairement dans un navigateur, mais il est bien plus efficace de les afficher dans un SIG, au moins pour la raison que l'on peut en superposer plusieurs, et jouer éventuellement sur les transparences³⁵. Au surplus, ces cartes constituent un support rêvé comme éléments de référence lors des opérations de géolocalisation. Parmi les sites qui méritent d'être signalés³⁶, celui du BRGM³⁷ (avec la carte géologique au 1/50000) et surtout celui du cadastre³⁸, qui permet d'afficher à toute échelle le plan cadastral vectorisé de la commune choisie. C'est un outil sans équivalent pour les géolocalisation à grande échelle. On peut obtenir un flux WMS de l'IGN, à condition de s'inscrire (gratuitement), et pour une durée de six mois³⁹. Il semble difficile de surestimer l'importance que ces flux WMS pourront prendre dans les temps qui viennent.

La toponymie

- 33 Dans la plupart des pays européens, antérieurement à la confection du premier cadastre méthodique et complet, les références topographiques relevaient quasi exclusivement de la toponymie. La plupart des historiens voient cela plutôt comme un problème technique (pour eux) que comme une structure fondamentale des sociétés anciennes. C'est pourtant de cela qu'il s'agit. Jusqu'ici, la toponymie comme discipline scientifique était un domaine de la philologie, pas de l'histoire. Marc Bloch et surtout son principal disciple, André Déléage, ont consacré d'importants efforts à l'analyse et à l'utilisation des toponymes, mais ils n'ont pas eu de postérité⁴⁰. Reconnaissons cependant que, jusqu'à une date récente, la mise en œuvre historique du matériau toponymique en tant que tel était à peu près impossible : l'ampleur du matériau dépassait largement les possibilités matérielles d'un chercheur, même organisé et efficace⁴¹. Il en va désormais tout autrement : on dispose aujourd'hui à la fois d'un matériau surabondant et d'outils logiciels adaptés et puissants.
- 34 A l'heure actuelle, pour la France, on dispose de trois types de matériaux : * la base de toponymie cadastrale dite FANTOIR, * la base toponymique de l'IGN, la BDNyme, enfin les * dictionnaires topographiques publiés depuis le milieu du 19^e siècle, progressivement numérisés et mis en ligne.
- 35 La base FANTOIR est téléchargeable librement sur le site du Ministère des Finances⁴². Elle comporte tous les noms de lieu connus des services du Cadastre dans les années 70 : noms de communes, noms de hameaux et de lieux-dits, noms de rues. Moyennant une mise en

forme, un tri des données utiles et un lien avec un fichier des coordonnées des centroïdes des communes, on obtient une base toponymique de près de six millions de toponymes géolocalisés (par rattachement à une commune).

- 36 La base BDNYME est la base de tous les toponymes utilisés par l'IGN pour la construction des cartes au 1/25000⁴³. Elle comporte environ 1,7 millions de toponymes bien géolocalisés (ils sont liés à leur position sur les cartes). Elle n'est pas directement en accès libre, mais toute institution de recherche ou d'enseignement obtient sans frais un code d'accès qui permet à ses membres de télécharger la plupart des fichiers de l'IGN⁴⁴.
- 37 Le Comité des Travaux Historiques et Scientifiques, institution ancienne qui a patronné la publication de la majorité des « dictionnaires topographiques », a entrepris une numérisation (en mode texte) suivie d'une mise en ligne sous deux formes : une interface d'interrogation, et une page de téléchargement des fichiers balisés. Après une première phase de financement par le CTHS lui-même, ce sont les services d'Archives publics qui ont pris le relais, le travail avance régulièrement, essentiellement de par l'activité de M.-J. Gasse-Grandjean (21 dictionnaires en octobre 2015)⁴⁵.
- 38 Il reste que, pour la France, la source toponymique la plus riche et la plus homogène, et de loin, est constituée par les états des sections du cadastre napoléonien ; dans ces registres, toutes les parcelles sont affectées d'un toponyme, nulle part la documentation n'est plus détaillée. Ces registres sont conservés essentiellement dans les Archives Départementales. Leur consultation est longue et pénible, en particulier du fait de leur grand format. Quelques départements ont d'ores et déjà entrepris de les numériser et de les mettre en ligne (les Alpes-Maritimes, les Bouches-du-Rhône, le Finistère, par exemple ⁴⁶), d'autres ont commencé les opérations ou les ont au moins programmées. On peut raisonnablement espérer que l'essentiel sera en ligne dans quelques années, fournissant ainsi un stock qui comportera probablement plus de vingt millions de microtoponymes. Cependant cette masse restera à numériser en mode texte et à géolocaliser.
- 39 Naturellement, cela ne change rien au fait, signalé ici par plusieurs auteurs, que les terriers des 14^e-16^e siècles en particulier comportent souvent des toponymes que l'on ne retrouve plus par la suite⁴⁷. Mais, même en tenant compte de ce point, on voit que la masse dès maintenant exploitable est impressionnante ; comme on l'a dit plus haut, l'outillage informatique et les procédures statistiques nécessaires sont disponibles. Il s'agit là d'un véritable domaine de recherche historique, jusqu'ici juste effleuré⁴⁸ : c'est une situation tout à fait inédite, la possibilité existe ici d'un renouvellement important de notre connaissance des sociétés européennes antérieures au 18^e siècle.
- 40 Les historiens universitaires actuels perçoivent bien, même si ce n'est qu'intuitivement, que la conjoncture générale leur est de plus en plus défavorable, alors même que leur discipline peine à trouver les voies du renouvellement. Dans leur grande majorité ils optent malheureusement pour une attitude de repli, qui se croit défensive, et qui n'est que suicidaire. On ne peut pas faire plus que d'encourager les plus jeunes et les plus courageux à se saisir des nouvelles sources et des nouveaux instruments qui sont là, à leur portée. Il ne fait aucun doute que le traitement méthodique des données historiques spatialisées est l'un de ces domaines jusqu'ici à peine entrevus, mais qui peuvent procurer d'opulentes moissons à celles et ceux qui voudront y travailler et contribuer ainsi à l'avenir de la discipline. Les journées de Dijon nous en ont donné un avant-goût.

NOTES

1. M. BLOCH, *Les caractères originaux de l'histoire rurale française*, Paris, 1931. M. LOMBARD, *Espaces et réseaux du haut Moyen Age*, Paris, 1972. P. et H. CHAUNU, *Séville et l'Atlantique (1504-1650)*, Paris, 1955-1960. La liste pourrait être allongée ad libitum. Citons encore F. GOGUEL, *Géographie des élections françaises : de 1870 à 1951*, Paris, 1951. Goguel était plus sociologue qu'historien, mais cet ouvrage, constitué presque entièrement de cartes, a eu, lui, une large postérité, et son influence indirecte se fait encore sentir.
2. I. HODDER, C. ORTON, *Spatial Analysis in Archaeology*, Cambridge, 1976.
3. F. DJINDJIAN, *Méthodes pour l'archéologie*, Paris, 1991. *Manuel d'archéologie*, Paris, 2011.
4. <http://grasland.script.univ-paris-diderot.fr>. Les cours et documents de cours proposés par Claude Grasland sont une mine d'information et de réflexion pour la statistique spatiale.
5. En ligne et téléchargeable : <http://www.shmesp.fr/spip.php?article69>.
6. « Analyse factorielle et analyses statistiques classiques : le cas des ordres mendiants dans la France médiévale », in *Annales. Économie, Sociétés, Civilisations*, 36/5 (1981), p. 869-91- http://www.persee.fr/doc/ahess_0395-2649_1981_num_36_5_282792 ; « Les pèlerinages du Mâconnais : une structure d'organisation symbolique de l'espace », in *Ethnologie française*, 12 (1982), p. 7-30 - <http://www.jstor.org/stable/40988694>.
7. « Quelques caractères spécifiques de l'espace féodal européen », in N. BULST, R. DESCIMON, A. GUERREAU (éds), *L'État ou le roi : les fondations de la modernité monarchique en France (XIVe-XVIIe siècle)*, Paris, 1996, p. 85-101 (texte présenté dans un colloque en 1992) - <https://halshs.archives-ouvertes.fr/halshs-00521075>.
8. Je n'utilise pas ce terme au sens mathématique classique, mais en référence à la distinction, que l'on opère en géomatique, entre des structures topographiques (tous les points ont des coordonnées, aucun élément ne peut bouger) et des structures topologiques (les éléments sont liés les uns aux autres par des relations spatiales déterminées, indépendamment de leurs coordonnées).
9. C. AUDELAN, M. HUMBERT, C. LESCUYER, C. DE VERGNETTE DE LA MOTTE, « Géolocaliser des cartes anciennes. Procédure » (dans le dossier).
10. Si les noms des communes dont on dispose sont conformes à ceux définis par l'INSEE, il est facile d'obtenir une correspondance avec l'un ou l'autre des fichiers des communes munies des coordonnées de leur centroïde, par exemple <http://professionnels.ign.fr/geofla> ou <http://professionnels.ign.fr/rgc> (librement téléchargeables).
11. Il suffit de disposer d'un fond de carte géolocalisé (éventuellement reçu en WMS, voir plus loin), et de créer une couche vectorielle-points ; on clique sur le point choisi et l'on remplit les cases qui doivent contenir les informations liées à ce point ; et lorsque tous les points sont traités, on enregistre dans un format courant (genre shape).
12. L'échantillonnage spatial est une affaire complexe, mais sur laquelle on peut trouver beaucoup de littérature. Parmi bien d'autres, P. AUBRY, *Le traitement des variables régionalisées en écologie : apports de la géomatique et de la géostatistique* - <https://hal.archives-ouvertes.fr/tel-00003736>; ou encore A. MOREAU, S. MORICE, S. SALVADOR-BLANES, H. BOURENNANE, « Distribution spatiale des objets archéologiques et mouvements de matière à l'échelle des versants cultivés : le cas de Mougou (Indre-et-Loire, France) » - <http://archeosciences.revues.org/2987>.

13. Mais le Brionnais faisait-il partie des vignobles « bourguignons » ? Cette carte suggère plutôt le contraire, incitant à rechercher plutôt un rattachement de cette zone aux vignobles ligériens ou auvergnats (Sancerre, Saint-Pourçain...).
14. Les indications que l'on donne ici sont, par construction, condamnées à être obsolètes dans bien moins de dix ans. Elles peuvent peut-être rendre quelques services d'ici là... NB tous les logiciels cités ici sont libres et disponibles pour toutes les plates-formes. NB en 2015 il n'y a plus de doute : un enseignant qui oriente ses étudiants vers des logiciels propriétaires commet une faute professionnelle.
15. <https://fr.wikipedia.org/wiki/QGIS> ; <http://www.qgis.org/fr/site/>.
16. <https://fr.wikipedia.org/wiki/PL/pgSQL>.
17. <https://fr.wikipedia.org/wiki/SQLite> ; <https://www.sqlite.org/>. Presque tout le monde l'utilise sans s'en apercevoir : sa légèreté en fait un choix obligatoire pour tous les systèmes embarqués comme les 'smartphones', et il sert de base de données en arrière-plan de logiciels très utilisés (Firefox, Skype). Le choix entre PostgreSQL et SQLite dépend avant tout de l'environnement : sur un serveur, ou une machine multi-utilisateurs, PostgreSQL s'impose du fait de son système de contrôle des connexions et des droits ; mais c'est un outil difficile à gérer ; si l'on travaille uniquement sur une machine individuelle, SQLite s'impose sans discussion. Une base tient dans un seul fichier transportable et utilisable sur n'importe quelle plate-forme. On trouve des interfaces d'administration et d'utilisation très bien conçues comme SQLiteStudio (<http://sqlitestudio.pl/>).
18. <https://www.r-project.org/>.
19. Voir les sites miroirs du CRAN (The Comprehensive R Archive Network). Parmi les plus récents, le paquet *cartography* (français !) (<https://cran.r-project.org/web/packages/cartography/vignettes/cartography.html>).
20. On trouve autant de manuels et de tutoriels que l'on veut, dans toutes les langues, sous tous les formats. Très nombreux forums d'aide et de discussion, par exemple, en français <http://forums.cirad.fr/logiciel-R/>. On peut visiter régulièrement avec profit <http://www.r-bloggers.com/>.
21. <https://rkwart.kde.org/>.
22. <https://www.rstudio.com/>.
23. C'est tout le (gros) problème des cartes lisibles et illisibles ; on se référera toujours aux travaux de base de Jacques Bertin, en particulier sa *Sémiologie graphique. Les diagrammes, les réseaux, les cartes*, Paris, 1967. Il est dommage que Bertin ait cru que « la graphique » dispense du recours aux calculs, mais cela ne retire rien à l'extrême intérêt de ses réflexions sur les fondements de la perception visuelle.
24. Une série de trois manuels-papier de taille modeste, mais particulièrement bien conçus, très pédagogiques : M. BÉGUIN, D. PUMAIN, *La représentation des données géographiques. Statistique et cartographie*, Paris, 1994 ; D. PUMAIN, T. SAINT-JULIEN, *L'analyse spatiale. 1. Localisations dans l'espace*, Paris, 1997 ; D. PUMAIN, T. SAINT-JULIEN, *Les interactions spatiales*, Paris, 2001. On utilisera avec le plus grand profit un manuel téléchargeable, qui est à la fois une introduction aux statistiques générales, aux statistiques spatiales et à la cartographie, le tout dans le cadre de R et dans une perspective d'apprentissage de la programmation : Groupe ElementR, *R et espace*, 2014 – http://framabook.org/docs/Respace/RetEspace_final_20140901.pdf.
25. N. CRESSIE, *Statistics for Spatial Data*, New-York, rev. ed. 1993 (900 pages : une véritable encyclopédie ; à consulter si l'on comprend les notations mathématiques et si l'on veut approfondir un point particulier) – <https://rongxie.files.wordpress.com/2011/01/statistics-for-spatial-data-revised-version-1993.pdf>.
26. Mentionnons notamment une présentation très utile du cadre général de manipulation des données spatiales dans R. S. BIVAND, E. J. PEBESMA, V. GOMEZ-RUBIO, *Applied Spatial Data Analysis with R*, New-York, 2008, – <http://gis.humboldt.edu/OLM/r/Spatial%20Analysis%20With%20R.pdf>; ou

encore une présentation synthétique https://www.ciser.cornell.edu/NYCRDC/documents/Deanenon-technical_overview-1.pdf.

27. Un des meilleurs outils est constitué par le paquet `ade4` pour R, développé par le groupe de bio-statistique de Lyon. Ce groupe maintient un site qui comporte une grande quantité de matériel pédagogique, auquel il est très intéressant de se reporter : <http://pbil.univ-lyon1.fr/ade4/> et <http://pbil.univ-lyon1.fr/R/>.

28. <http://f.hypotheses.org/wp-content/blogs.dir/1381/files/2014/11/PNAS-2004-Gastner-7499-504.pdf>.

29. <http://scapetoad.choros.ch/> (excellent) ; <http://www.omegahat.org/Rcartogram/>.

30. Une foule d'exemples : <http://www.worldmapper.org/index.html>.

31. <http://www.geoportail.gouv.fr/accueil>, onglet « remonter dans le temps ».

32. https://fr.wikipedia.org/wiki/Web_Map_Service.

33. <http://www.portailsig.org/content/comment-ajouter-un-fond-de-donnees-geographiques-gratuit-dans-son-sig>.

34. par exemple <http://georezo.net/wiki/main/webmapping/webservices>.

35. Pour QGIS, deux greffons facilitent grandement les opérations : OpenLayers Plugin et QuickMapServices.

36. On trouve des sites qui fournissent des listes d'adresses, par exemple <https://www.cerege.fr/spip.php?article276> ou <http://georezo.net/wiki/main/webmapping/webservices>. La liste ne cesse de s'allonger.

37. <http://infoterre.brgm.fr/geoservices-ogc>.

38. https://www.cadastre.gouv.fr/scpc/pdf/Guide_WMS_fr.pdf. Un script perl simple et très efficace permet de récupérer les feuilles du cadastre : <http://svn.openstreetmap.org/applications/utills/cadastre-france/cadget>.

39. <http://professionnels.ign.fr/api-sig> (on peut obtenir un code permanent si l'on est enregistré, cf note 44).

40. D. MÉHU, « Les méthodes d'André Déléage dans le domaine de la toponymie et de l'archéologie des terroirs », *Annales de Bourgogne*, 83-2011, pp. 231-272. On lira avec intérêt la lettre de Marc Bloch à André Déléage du 29 juin 1929, où il discute pied à pied plusieurs interprétations de toponymes (in *Annales de Bourgogne*, 83 (2011), p. 99-100).

41. Les exceptions sont rares, comme P.-H. BILLY, *La condamne. Institution agro-seigneuriale. Étude onomastique*, Tübingen, 1997.

42. <https://www.data.gouv.fr/fr/datasets/fichier-fantoir-des-voies-et-lieux-dits/> (le fichier brut doit être soumis à un intense traitement préalable pour être utilisable, un peu de programmation est nécessaire).

43. <http://professionnels.ign.fr/bdnyne>.

44. <http://professionnels.ign.fr/enseignement-recherche>. Notons que, lorsque l'on dispose d'un code et que l'on peut télécharger des fichiers, les données toponymiques sont incluses dans la base BDTOPPO, qui est disponible par département ; lorsqu'on a tous les départements, il reste à partir à la recherche des nombreux fichiers toponymiques catégorisés et à les fondre dans un fichier unique, on ne peut guère y parvenir sans quelques lignes de programme.

45. <http://cths.fr/dico-topo/dictionnaires/cartes.php>. Une utilisation méthodique suppose le chargement des fichiers xml et leur mise en forme.

46. <http://www.basesdocumentaires-cg06.fr/archives/indexCAD.php> ; http://doris.archives13.fr/dorisuca/jsp/system/win_main.jsp ; http://mnesys-portail.archives-finistere.fr/?id=recherche_cadastre.

47. Voir l'intéressante thèse de B. VUE, *Microtoponymie et archéologie des paysages à Neuilly-l'Évêque (52) du XIII^{ème} au XX^{ème} siècle : comportement des microtoponymes au fil des siècles : le nom, l'espace, l'homme qui le nomme* - <http://www.theses.fr/1997NAN21013>.

48. A GUERREAU, « Toponymie et nouvelles techniques : une expérience de géotoponymie historique dans un cadre communal (Broye, Saône-et-Loire) », in *Nouvelle Revue d'Onomastique*, 55 (2013), p. 3-46.

AUTEUR

ALAIN GUERREAU

CNRS