

Information spatialisée et déformation cartographique en démographie

Daniel Delaunay

Bien des démographes affichent d'emblée une attitude sceptique envers la cartographie des statistiques censitaires, et de manière plus générale vis-à-vis de la dimension spatiale des phénomènes qu'ils étudient. Cette méfiance est-elle fondée ? La question, pour le moins, mérite d'alimenter le présent débat sur les illusions, voire les manipulations que la carte est susceptible d'entretenir.

Serait-elle inutile, et l'espace démographique dépourvu de sens, qu'elle serait effectivement source de confusions et de faux jugements. Quelques exemples viendront étayer une conviction contraire ; à savoir que la dimension spatiale propre aux phénomènes démographiques mérite l'attention, notamment au moyen de la cartographie détaillée qu'autorisent les systèmes d'information géographique (SIG). Cependant, les dangers ou les tentations d'une manipulation mensongère, ou d'une interprétation erronée des statistiques localisées, sont réels. Les altérations potentielles tiennent pour une part aux difficultés méthodologiques rencontrées par l'analyse des statistiques localisées, en géographie de la population (Pryor, 1984), mais aussi à l'usage expérimental

des systèmes d'information géographique. Le désintérêt de la théorie démographique est également à incriminer¹.

Les remarques avancées ne prétendent évidemment pas à l'exhaustivité ; elles témoignent d'une pratique plus que d'une réflexion et seront donc illustrées. Ce le sera avec le dernier recensement de population réalisé au Mexique en 1990 dont les résultats sont disponibles depuis peu avec la précision spatiale requise². Les traitements infographiques présentés sont réalisés actuellement dans le cadre d'un programme de coopération entre le Colegio de la Frontera Norte³, en Basse Californie mexicaine, et l'ORSTOM⁴, une recherche en partie financée par le CONACYT. Sont employés les logiciels *Savane* et *Rapido*⁵ développés à l'Institut. L'information nous a été gracieusement cédée par l'INEGI⁶.

Les cartes démographiques ont-elles un sens ?

L'éloquence visuelle de la bonne carte, le souvenir de nos manuels d'écolier confortent probablement cette opinion partagée que la carte excelle comme support pédagogique, facilitant la perception de la diversité régionale des populations, à l'instar de celle de leurs terroirs et de l'habitat. Le géographe entretient ce sentiment quand il conçoit la carte comme moyen d'expression plus que d'analyse ; il nous y engage en dressant le mode d'emploi des signes et des techniques qui font *parler* la carte, alors

1 • La dimension spatiale en démographie est mieux abordée par le géographe avec une prédisposition pour les études empiriques et un corpus sans doute insuffisant de travaux théoriques. On notera cependant des exceptions : la régionalisation des rapports ressources-populations par Ackerman (1959) et la théorie de la transition migratoire par Zelinsky (1971 et 1979). L'effort de modélisation est en revanche plus considérable.

2 • Sous forme d'un CD-ROM (le CODICE90) contenant l'ensemble des publications imprimées du recensement, plus les plans cartographiques urbains. La constitution des bases d'information localisée a cependant exigé un long travail d'extraction et de traitement.

3 • COLEF, 21 calle Abelardo Rodriguez, 22320 Tijuana, BC, Mexique.

4 • ORSTOM, 213 rue La Fayette, 75480 Paris cedex 10, France.

5 • Savane est un système d'information géographique convivial développé surtout par Marc Souris sur station de travail SUN. Rapido est une version simplifiée mais performante qui autorise le traitement cartographique de bases de données gérées par Savane. Ce logiciel développé par Michel Lepage fonctionne sur un ordinateur personnel de type PC.

6 • INEGI : Instituto Nacional de Estadísticas, Geografía e Informática.

média de la découverte, de la décision, du rêve (Brunet, 1987). Nous prêtons spontanément moins de rigueur à ces qualités-là, étant habitués à ce que cet objectif médiatique s'accommode de manipulations symboliques mensongères, trop souvent exclusivement soucieuses de l'effet produit. Convient-il de s'en inquiéter parce que la carte est aujourd'hui populaire dans nos journaux, sur nos écrans ? C'est-à-dire offerte à des lectures novices ou pressées, parce qu'elle se doit d'accrocher une attention usée. Ou bien parce que les instruments infographiques modernes permettent à quiconque de confectionner une carte en un temps si court qu'il en oublie les nécessaires rigueurs de l'exercice, pour le moins qu'il y soit préparé. C'est ce que semblent craindre les géographes.

La méfiance de ses collègues est autre : si l'on interroge un non-géographe, il pensera spontanément que la carte invite à la description, pas à l'explication. Il négligera ce support quand il sera persuadé que les contrastes régionaux des caractères démographiques sont dus aux variations des contextes économiques, éducationnels, culturels ; bref, de ces déterminismes familiers à la discipline. Il ne conçoit qu'avec un effort peu convaincu que les dynamiques de population puissent être également infléchies par des processus spatiaux. Reste à s'assurer qu'ils existent, les circonscrire, estimer leur impact ; une tâche difficile que la carte, outil graphique d'analyse, est peut-être en mesure d'aider.

Nul ne songe à douter de l'utilité d'une carte qui donne à voir en un coup d'oeil la localisation précise de la mortalité des enfants, par exemple. De telles qualités heuristiques guideront utilement les politiques vers les lieux de leur plus grande opportunité (Udo, 1976). L'information visuelle qu'elle apporte sera d'autant plus précieuse que le découpage géographique sera fin et les statistiques localisées détaillées. Les systèmes d'information géographique autorisent cette précision de la carte au prix d'un effort raisonnable pour peu qu'il ne soit pas nécessaire de produire les statistiques, et que celles dont on dispose soient fiables. Très peu de ces logiciels, cependant, assistent l'analyse de l'information qu'ils gèrent et représentent ; cela traduit une négligence (car les techniques statistiques existent), assez courante chez les usagers de l'outil, et peut

conduire à une altération grave du sens de la carte, du moins de celui perçu par le plus grand nombre de lecteurs. À défaut d'une analyse rigoureuse des configurations spatiales ainsi présentées, chacun sera enclin à les interpréter selon l'idée qu'il se fait de la géographie sous-jacente. Ainsi au Mexique, les populations de tradition indigène paraissent supporter les mortalités juvéniles les plus élevées. Mais une carte aussi parlante ne dit rien des discriminations de l'éducation, de l'histoire, voire du milieu où elles se sont réfugiées, des caractères évidemment pas intrinsèques à la condition indigène. En montrant plus qu'elle ne saurait expliquer, la carte risque d'abuser la perception du lecteur. Et que peut-on comprendre de la carte détaillée des 2 400 *municipios* mexicains quand la géographie du milieu physique, de l'activité économique, de la comptabilité nationale n'a que très rarement cette précision ?

Ne négligeons pas, néanmoins, la découverte toujours possible de configurations spatiales inattendues, toutes ne sont pas annoncées par la théorie et les modèles, beaucoup s'en faut. La carte peut en dévoiler quelques-unes, lors d'une phase exploratoire des données, ou quand les causalités sortent du champ statistique couvert : la part non expliquée du phénomène (les résidus d'une corrélation, par exemple) présentera alors des structures spatiales significatives. On peut même penser que, faute d'être localisée, l'analyse statistique offre des arguments fallacieux aux politiques de populations dès lors incapables de distinguer les lieux qui échappent à la causalité dégagée. Pour des raisons historiques ou même circonstanciées, telle région peut s'écarter de cette logique démographique perçue par la théorie ou confirmée par l'expérience. Nombreuses sont les études de cas qui découvrent les exceptions aux théories admises de la transition ou de la mobilité, dès que l'on considère des cultures spécifiques ou un contexte spatio-temporel singulier (Pryor, 1984 : 30).

Il me semble que la cartographie censitaire est appelée à instrumenter notre connaissance de la diversité démographique selon deux perspectives : en retrouvant les échelles intermédiaires de leur configuration, et en révélant les processus spatiaux qui agissent sur les caractères des populations.

1 – Chaque phénomène démographique varie dans l'espace selon une échelle propre : certains se confondent avec les grands traits climatiques ou ethniques, d'autres se conforment au tracé des voies de communication et du réseau urbain, la ville en son sein recèle une disposition singulière. La territorialité des comportements reproductifs ou matrimoniaux, par exemple, sera méconnue de l'analyse démographique qui marque une préférence pour les échelles extrêmes, souvent exclusives : les agrégats nationaux et l'individu ou sa famille. On imagine pourtant d'autres acteurs dont les logiques démo-économiques ne peuvent être ignorées si on veut comprendre l'extrême diversité des mouvements de population. Or, ces acteurs, et les systèmes qu'ils gèrent, se manifestent à des échelles spatiales intermédiaires. La cartographie censitaire, confrontée à celle d'autres inventaires, autorise à considérer des échelles intermédiaires susceptibles de révéler d'autres cohérences démographiques : un réseau de places centrales pour l'exode rural, une rupture climatique qui imposa des limites incontournables à un peuplement ancien, et peut-être le contour d'un trait culturel singulier, la géographie d'enclavement de l'analphabétisme...

2 – En soi, la carte n'explique rien. Mais de constater la corrélation statistique entre une géographie de la mortalité infantile et, disons, du produit régional brut et de l'éducation des mères n'explique pas, non plus, la genèse de cette configuration multivariée. Par quels processus la diversité se génère-t-elle dans l'espace, s'entretient-elle ? S'ils existent, ces processus sont probablement en mesure d'avoir une influence directe sur les caractères des populations autrement que par l'intermédiaire des facteurs reconnus. Douter *a priori* de la pertinence de certains d'entre eux en démographie serait de l'aveuglement. Haining (1992) recense les principaux. La diffusion de certaines techniques (contraceptives, par exemple), produits ou attitudes (familiales ou culturelles...), celle des épidémies, bien sûr, ont un impact reconnu. Et à leur tour, les qualités des populations, telle la densité ou l'éducation, peuvent contribuer à cette diffusion. La dispersion concerne la population elle-même et non plus ses attributs comme dans le cas de la diffusion. Les migrations

humaines justifient à elles seules l'attention à porter à l'espace démographique, ne serait-ce que pour les changements qu'elles entraînent. Les échanges et transferts de marchandises, de signes constituent le fondement des activités économiques et tissent un lien régional puissant, leur étendue est susceptible d'expliquer la prospérité et l'identité changeante des lieux. Enfin, il faut considérer les processus d'interaction quand un événement ou une attitude résultent d'une réponse à un autre événement qui survient en un site éloigné, un choix familial en fonction des migrants à qui on est lié, par exemple.

Il ne faut néanmoins pas se cacher la grande difficulté à reconnaître ces mécanismes dans les configurations spatiales dévoilées par la carte. Deux processus différents peuvent produire des géographies identiques, du fait également de la complexité des réseaux d'interdépendance qui façonnent les systèmes sociaux à cause du poids des attributs historiques. Et si la carte est conçue pour aider l'analyse, et pas seulement pour l'illustrer, une mauvaise manipulation risque de mener à l'erreur, éventuellement commise à l'insu de son auteur.

La réduction des attributs

Le pouvoir évocateur des cartes tient à leur capacité de réduction, tant de l'espace que de ses attributs : un recensement se verra ainsi réduit à quelques planches. La simplification est un outil précieux lors de l'exploration préliminaire de statistiques abondantes, afin de guider l'adoption des traitements ou des modèles, pour dégager l'essentiel. Mais du fait de ces réductions, une carte est autant le produit des choix qui ont guidé sa confection qu'une expression de la réalité.

La qualité des sources

Les progrès informatiques stimulent la diffusion des bases de données qui sont appelées à devenir notre source principale de connaissance sur l'environnement et l'activité des populations. Cependant, cette profusion de statistiques risque de contribuer à

notre ignorance si on ne sait en contrôler la validité, s'assurer de la cohérence des divers inventaires. Ceux-ci sont généralement produits par des institutions distinctes, peu disposées à révéler leurs défaillances ou à en admettre les lacunes inhérentes. Les statistiques vitales exemplifient ce danger. Pour estimer les vitalités d'une population, l'enregistrement quotidien des décès et naissances doit être associé aux dénombrements décennaux des populations de référence. On met en relation deux sources statistiques fort différentes qui subissent leurs propres déformations dans le temps et dans l'espace, pas nécessairement congruentes, et donc capables de provoquer de plus amples erreurs sur les taux calculés. La confrontation d'inventaires de nature et de délimitation territoriale différentes, du milieu naturel et de l'activité agro-pastorale par exemple, est plus délicate encore.

Avant de cartographier les statistiques censitaires, il faut s'attendre à certaines faiblesses propres à fausser le document graphique. Les sous-dénombrements sont inévitables, à moins de grossir abusivement les coûts. On les évalue *a posteriori*, mais jamais pour les divisions fines du territoire, ce qui interdit de les corriger. Qui s'intéresse aux structures, c'est-à-dire aux équilibres entre sous-ensembles de la population dénombrée, peut miser sur une répartition aléatoire des omissions en chaque groupe, sans altération systématique de la carte. Cette commode supposition n'est pas toujours remplie : dans les pays de statistiques imparfaites, il est habituel que les populations indigènes, les immigrants clandestins, les très jeunes enfants... soient plus mal recensés. Plus grave pour la carte, la couverture varie d'un lieu à l'autre, désavantageant les régions enclavées, habitées par des minorités, mal desservies. Des lacunes particulières, qui touchent la mesure de tel phénomène en certains groupes ou lieux, sont susceptibles de varier dans le temps. La cartographie des parités⁷, par exemple, doit considérer la possible défaillance des informateurs quand les enfants sont exposés à un risque de décès élevé : le sous-enregistrement touchera surtout les femmes plus âgées, appartenant aux sociétés indigènes, vivant dans les régions les plus montagneuses. Ces lacunes sont assez importantes pour

7• Descendance moyenne des femmes, ou nombre moyen d'enfants nés vivants.

fausser les cartes et leur interprétation. Il est d'usage d'éliminer les non-réponses, mais elles peuvent être importantes au point de jeter un doute définitif sur la validité des statistiques se rapportant à tel sous-ensemble : les chômeurs qui ne déclarent pas leur revenu, des femmes divorcées qui escamotent leur état civil, etc.

La cartographie d'objets spatiaux de petite taille risque d'être gênée par la règle de confidentialité qui protège les citoyens. Même si elle est souple, comme dans les recensements mexicains qui ne divulguent l'information que pour les groupes de plus de trois unités statistiques, il peut devenir impossible de cartographier certains attributs. Songeons à la concentration de quelques grandes industries pour s'en persuader. De même que toutes les localités qui réunissent moins de quatre logements sont regroupées sous une seule catégorie non référencée dans l'espace. Leur absence déforme l'image des structures spatiales.

Des erreurs de localisation viendront s'ajouter à celles des dénombrements. Il n'est pas toujours facile de faire coïncider les attributs et leurs unités spatiales sur les fonds cartographiques, même quand ces deux inventaires sont produits par la même institution, comme c'est le cas au Mexique. Plusieurs localités (de 1 à 3 % environ selon les régions) sont apparues avec des coordonnées géographiques fausses⁸. Une autre difficulté est générée par la création de nouvelles unités administratives d'un recensement à l'autre⁹, de l'imprécision des cartes de référence¹⁰, souvent de simples croquis. L'interprétation s'accommodera de ces imprécisions plus ou moins gênantes selon l'échelle de l'observation. Elles deviennent rédhibitoires pour la mise en relation de plusieurs couvertures thématiques au découpage divergent. Cette jointure géographique, qui permet d'inférer l'information d'un zonage à l'autre, suppose une distribution homogène à l'intérieur des zones ou bien un système de pondération fiable (Delaunay, 1991).

La désinformation provoquée par des statistiques douteuses est un risque difficile à contrôler dans les SIG relationnels qui

8• La règle fut de corriger les localités qui sortaient du *municipio* et, systématiquement, les plus importantes, mais il était exclu de penser contrôler l'exacte localisation de toutes.

9• Des regroupements d'unités spatiales doivent alors être opérés.

10• Pour la cartographie urbaine au Mexique qui ne dispose pas d'un cadastre homogène, on doit reconstituer le tracé des îlots sur des photos aériennes corrigées.

rassemblent et croisent une information volumineuse et de qualité disparate. L'analyse spatiale multivariée court le danger d'un nivellement par le bas, ne pouvant pas valoir mieux que la moins bonne mesure des facteurs. Essayer d'assainir les fonds cartographiques et leurs attributs constitue une tâche accaparante, autrement plus longue et fastidieuse, en amont de la confection des cartes.

La mégarde des choix

Le champ restreint couvert par les statistiques n'est évidemment qu'une parcelle de la réalité. Pourtant, les multiples décomptes des populations selon l'âge, l'état civil, le métier..., peuvent être combinés en un grand nombre d'indicateurs¹¹. Pour le Mexique en 1990, qui nous sert d'illustration, nous avons extrait deux cents ratios environ à partir de près de huit cents dénombremens. La confection des cartes censitaires procède donc de la double sélection de certains caractères et de leur expression numérique. Ces choix seront en partie guidés par la curiosité, l'expérience, ou le hasard quand les SIG offrent de pouvoir en considérer et en représenter un grand nombre à titre exploratoire. Une commodité qui ne va pas sans danger : les voir avant de les analyser laissera probablement trop d'importance aux qualités graphiques de telle variable, même si l'information thématique qu'elle recèle est médiocre. Qui avouera avoir écarté cette configuration parce qu'elle se révélait banale ou trop difficile à interpréter, ou d'avoir camouflé par une habile discrétisation, à l'aide d'une graphique judicieuse, que tel choix n'était pas le plus pertinent ?

L'habitude incitera le démographe à surcharger sa carte des indicateurs dont il est familier ; mais une pyramide des âges, une courbe de survie ou de fécondité ne supportent pas d'être répétées en autant d'unités spatiales, même grossières. Il devra considérer des cartes multiples mais simples, illustrant une

11 • Elles ne sont pas publiées selon tous les croisements que l'on aimerait trouver pour chaque unité spatiale, les recensements permettent plus que ne l'autorisent les instituts statistiques.

information dépouillée, dès lors plus difficile à interpréter. Ces choix multiplient les risques d'une déformation de la carte et de son sens. Prenons l'exemple de la fécondité qui résume bien le dilemme de la cartographie démographique entre le détail spatial et la justesse analytique. Le recensement donne à calculer plusieurs indicateurs : les parités moyennes de toutes les femmes classées par groupes d'âges, la même chose pour les mères seulement, ou plus simplement la proportion d'enfants de moins de cinq ans pour l'ensemble des femmes en âge de procréer. Si les statistiques vitales l'autorisent, on ajoutera de véritables indices de fécondité du moment. Au total, ce sont près d'une vingtaine d'indicateurs qu'il nous est permis de cartographier pour avoir une image géographique de la reproduction humaine. Mais tous ces indices présentent des limitations intrinsèques susceptibles de fausser la carte, ou d'égarer l'interprétation, indépendamment des techniques graphiques utilisées. Ou bien on retient des taux analytiquement rigoureux mais sur des bases statistiques douteuses, ou alors on privilégie des indicateurs plus solides mais qui amalgament les effets de plusieurs phénomènes : fécondité, mais aussi nuptialité, survie et migration des enfants et des mères, effet de l'âge. Dans le premier cas, le flou statistique découle du détail régional requis, lequel introduit des imprécisions quant aux populations de référence qui sont interpolées d'un recensement à l'autre. L'incertitude devient inacceptable chez les populations en transition démographique, quand les modèles d'ajustement qui reposent sur les populations stationnaires ne sont plus applicables. L'erreur potentielle devient dissuasive pour les unités administratives touchées par la migration capable de modifier rapidement les structures par âge. Le calcul des parités, soit la seconde solution, contourne ces difficultés car les femmes sont interrogées sur leur descendance au moment du recensement, le numérateur et le dénominateur du ratio proviennent de la même source. Mais les parités censitaires produisent une information plus rudimentaire faute de renseigner sur le calendrier des naissances¹². Souvent, on ne dispose que de parités moyennes qui mélangent les âges aux fécondités très changeantes, il exprimera les différences

12• A dix ans d'intervalle, on n'est pas assuré de mesurer la parité de la même génération de femmes pour une région donnée.

structurelles de ces populations régionales en même temps que leur prolificité. Quand cette information est disponible, il est impératif de comparer des parités aux mêmes classes d'âges. Mais pour les jeunes femmes, disons avant vingt ans, l'indice qui ne s'appuie que sur un nombre réduit d'événements risque d'être entaché d'une trop grande variabilité. Quand elles vieillissent, les erreurs associées à l'oubli augmentent du fait que certains enfants meurent au cours de leurs premières années de vie, biaisant tout particulièrement les événements anciens. Quand la *parité* des mères est disponible, et non plus seulement celle de toutes les femmes, on dispose d'une estimation épurée de l'influence de la nuptialité, laquelle subit de nettes variations régionales. Bref, avec cette rapide énumération, on constate que l'abondance d'attributs peut cacher une réelle indigence, celle de ne pas remplir l'exigence minimale de l'analyse démographique qui est de mesurer, autant que faire se peut, des phénomènes épurés d'influences parasites. Car considérer d'abord les qualités graphiques de la carte encouragera à garder le plus rudimentaire de ces indices. Le nombre d'enfants de moins de cinq ans par femme en âge de procréer est moins perturbé par la taille réduite de certaines populations, la déclaration d'un âge erroné, l'omission d'une naissance. Il se construit avec des effectifs larges de femmes et d'enfants, simplement décomptés lors du recensement. Sa cartographie apparaît alors souvent plus significative au prix de, ou grâce à une simplification analytique, celle d'exprimer un amalgame de phénomènes qui contribuent à la vitalité globale des populations : fécondité, mortalité des enfants, migration des femmes, nuptialité. Dans cet exemple précis, les indicateurs que l'on met en carte sous l'étiquette *fécondité* ne sont que des facettes floues du phénomène, on aura sacrifié la rigueur de l'analyse au profit de l'effet graphique, ce qui peut conduire à des interprétations abusives faute de bien distinguer les origines de cette conjonction d'effets.

Certains choix des attributs sont marqués par l'habitude, du cartographe comme celle prêtée au lecteur : ainsi les valeurs moyennes de la distribution des salaires, de la taille des ménages, de la proportion de célibataires... Quand il ne s'agit pas de la seule information disponible, la cartographie des valeurs extrêmes peut dévoiler des configurations radicalement différentes et hautement

instructives : la population en dessous d'un seuil jugé vital de revenu, par exemple. On pourra s'en persuader avec une donnée courante dans les recensements, à savoir le nombre moyen d'individus par logement. Le choroplèthe¹³ de cet indicateur se distinguera de celui de la proportion de ménages accueillant deux personnes, ou de ceux en réunissant sept et plus. Ces extrêmes traduisent des formes très différentes d'organisation de la famille, de la plus atomisée à la plus étendue. Résultat d'un processus complexe, l'indice est difficile à interpréter mais, en l'associant à d'autres mesures, il permet de cerner l'impact de la fécondité, de la migration, et surtout du poids des sociétés rurales domestiques sur la famille. Il constitue un bon point de départ pour analyser les transformations qu'induit la diffusion du travail salarié, et ses conséquences sur la reproduction des groupes domestiques.

On pourrait généraliser cette remarque en affirmant que la désinformation en matière de cartographie statistique commence avec l'élimination *a priori* des indices qui apparaissent triviaux pour l'ensemble d'un pays, mais dont la différenciation spatiale est susceptible de revêtir un sens très différent. Prenons le rapport de masculinité. Au sein de la population nationale, il traduit des écarts de la mortalité selon le genre¹⁴. Aux âges où ces différences sont minimes, les configurations régionales de l'indice peuvent dévoiler des comportements migratoires différents chez les deux sexes, la prédilection des femmes pour une destination urbaine par exemple.

La nécessité de simplifier les expressions numériques d'un même phénomène à des fins cartographiques et l'éventail de ces mesures possibles autorise à moduler le message graphique, donne une marge à l'interprétation. Or le sens graphique est immédiatement perçu et aisément modifié à l'aide des systèmes d'information géographique : l'instrument nous offre cette dangereuse facilité de pouvoir juger les cartes avant les chiffres.

13• "Etym.: quantité par aire. Une carte choroplèthe est une carte qui représente des quantités par maille, ou unité spatiale statistique, selon un figuré couvrant toute la surface de la maille." (Brunet (R.), 1992 - Les mots de la géographie - dictionnaire critique., Reclus - La Documentation française.

14• Et la migration internationale différentielle selon le genre, non négligeable au Mexique.

La transformation des attributs

Pourtant, les mises en garde des statisticiens sont nombreuses et précises. Les valeurs extrêmes expriment-elles la réalité, un artifice statistique ou une erreur ? S'est-on assuré de l'indépendance des observations ou doit-on compter avec une autocorrélation spatiale ? A-t-on la preuve statistique que la distribution de l'attribut dans l'espace ne suit pas une loi aléatoire ? Les SIG, conçus pour manipuler les cartes plus que pour les analyser, n'incitent pas à ces prudences considérées élémentaires quand on confectionnait une carte à *la main*. Une information abondante, vite consommée (la carte *produit jetable*) peut déprécier le jugement à la manière de la *junk food* le goût (et la capacité de se servir simultanément de la fourchette et du couteau). Pour le démographe ou l'économiste, la carte est un instrument graphique qui complète l'analyse des nombres mais ne doit pas faire l'économie de ses rigueurs.

Les transformations appliquées aux attributs à des fins thématiques (calcul d'un taux de mortalité, par exemple), ou statistiques (normaliser une distribution exponentielle) peuvent avoir des conséquences sémantiques annexes. Il est fréquent de se demander s'il est préférable de représenter le dénombrement absolu d'une sous-population ou son importance relative. La première présentera probablement une répartition exponentielle, la seconde normale, cette dernière serait donc préférable en termes statistiques. Mais montrer où se trouvent les effectifs d'adultes illettrés, par exemple, permet de guider efficacement les investissements éducatifs vers les lieux de leur meilleur impact numérique. En revanche, quiconque se soucie de justice sociale présentera la carte du pourcentage d'analphabètes pour désigner les régions les plus défavorisées. Mais ce faisant, on insistera, peut-être abusivement, sur la situation de populations moins nombreuses et isolées au détriment de ressources rares. Du choix de l'une ou de l'autre de ces deux cartes peut résulter une impression très différente chez le lecteur ou le décideur.

La présence de valeurs extrêmes justifie une prudente transformation des variables, pour autant qu'elles ne soient pas le produit d'une erreur qui serait alors à corriger. Le découpage administratif en produit un grand nombre chez les populations

modernes fortement urbanisées. Cela tient à ce que la segmentation de l'univers statistique, celui des unités élémentaires (ici les ménages), n'est pas conduite de manière aléatoire mais selon les unités territoriales. Un biais dans la distribution des attributs cartographiés est introduit avec la taille démographique des unités auxquelles ils se rapportent (Cf. *infra*). À un extrême, les capitales tendent à présenter un comportement singulier, voire atypique, auquel l'importance démographique confère une pondération déformante lors des régressions. À l'autre, les unités spatiales les moins fournies donnent à l'attribut une trop grande variabilité, que l'on doit simplement au petit nombre d'individus dont il mesure les caractères. Ces valeurs extrêmes étirent la distribution statistique, peuvent distordre l'estimation d'un modèle, d'une régression, et bien sûr les descripteurs de l'attribut (moyenne, écart-type...) ; mais surtout, elles ajoutent du *bruit* à la cartographie, lequel contribue à occulter les configurations réelles du phénomène. Il peut être convenable de filtrer les valeurs fautives, les transformer pour atténuer la nuisance statistique, agréger les unités moins peuplées, calculer des descripteurs plus robustes, une combinaison statistique des anciens, par exemple (Haining, 1990).

La maladresse de l'oeil face à de fines nuances de tonalité ou des trames oblige à transformer l'attribut en quelques valeurs discrètes, chacune représentée par une couleur ou un niveau de gris facilement discernable. Sur le document graphique, ces classes apparaîtront homogènes, une confusion virtuelle à la discrétion du *géographe* alors en mesure de forcer un intervalle pour camoufler une diversité régionale ou créer une identité géographique factice. Car, en la matière, il existe quelques techniques (Cauvin *et al*, 1986), peu de règles et pas mal de désaccords. Une classification des valeurs selon les quantiles produit un même nombre d'unités spatiales dans chaque classe. La carte y gagne un apparent équilibre, évidemment forcé pour les distributions irrégulières de la variable si elles sont déformées vers des valeurs extrêmes. Une discrétisation centrée sur la moyenne et réduite selon l'écart-type déforme le sens de la carte si la distribution n'est pas normale. Des seuils ayant une portée analytique seront camouflés si on retient des intervalles égaux, bornés par des chiffres ronds. Dans la pratique, il est rare que l'on explicite, plus encore que l'on justifie, le choix des intervalles qui donnent à la carte son allure générale,

offrant parfois des arguments à une interprétation fallacieuse. Si une méthode est systématiquement préférée, on est en droit de douter qu'elle convienne à toutes les situations. À cela, il existe parfois des excuses quand, après avoir élaboré diverses discrétisations d'un même attribut pour écarter les dangers de l'une ou de l'autre, le choix final de la carte unique ne sera qu'un compromis. Une alternative est d'opérer une interpolation spatiale des valeurs et choisir un grand nombre de classes égales (après une éventuelle transformation de la variable). L'oeil accrochera le dessin des formes simplifiées plutôt que l'ancienne délimitation officielle et tortueuse des niveaux de couleur.

La réduction de l'espace

Bien que les recensements soient supposés exhaustifs, ils ne restituent pas l'exacte dimension spatiale de leurs dénombrements. On ignore la localisation de chaque ménage. Nous est seulement donnée celle de l'unité administrative qui en regroupe un certain nombre. En fait, le nombre des lieux qui servent au repérage des données censitaires est souvent infime dans l'univers des lieux possibles, ou même habités : une trentaine de points pour les états, moins de deux mille cinq cents pour les *municipios*. Plus grave, ce choix des localisations n'est pas donné selon une distribution aléatoire ou régulière mais conformément à une géographie officielle figée par l'histoire, et probablement inadaptée à la cartographie censitaire. Cet échantillon biaisé des lieux reconnus interdit d'inférer l'espace démographique réel par une interpolation méthodique. Une telle déformation est cartographiée avec l'attribut quand sa valeur est symbolisée au centre d'une zone, ou par une même trame la remplissant : ce faisant, on suggère l'homogénéité sur toute l'étendue du polygone. La perception de l'échelle et de la structure du phénomène s'en trouve faussée, l'une et l'autre sont travesties par la forme et l'étendue du découpage. Et on aura tendance à voir dans la géométrie des limites, celle de la configuration de l'attribut¹⁵

15• Cette géométrie administrative peut, cependant, suivre la configuration historique du peuplement, ou naturelle du climat, l'un et l'autre influents. Elle pourra revêtir un sens démographique quand les politiques régionales qu'elle circonscrit ont une incidence sur le phénomène analysé.

Cependant, l'échelle propre d'un comportement démographique est rarement de perception immédiate, ni toujours possible. Avant tout, elle est conditionnée à la disposition de statistiques géographiquement détaillées. Dans les bases de données, les attributs nécessaires ne sont normalement dispensés que pour les grandes régions nationales, états ou provinces, parfois pour les communes. Impératifs de confidentialité et d'économie en sont normalement la raison invoquée¹⁶. Cette précision est pourtant nécessaire pour appliquer les méthodes susceptibles de révéler ces échelles, mesures de variation selon la distance (Matheron, 1963), ou analyse de leur variance selon différentes tailles du carroyage (Cressie, 1991). Une mauvaise appréciation de l'étendue de telle organisation spatiale résultera de la sélection inappropriée d'une région d'étude, susceptible d'être trop partielle ou arbitraire, tels les seuls *municipios* limitrophes des États-Unis, par exemple, pour étudier les singularités frontalières. Cela devrait être une règle de toute analyse régionale que de considérer l'espace plus vaste qui entoure celui que l'on étudie, au moins pour quelques variables jugées essentielles.

Au Mexique, il n'y a pas de description spatiale plus précise que celle des statistiques référencées pour toutes les localités du pays. La limitation dans ce cas est de deux ordres. Statistique, tout d'abord : le recensement de 1990, et lui seul, n'accorde que quelques attributs à ce niveau de détail géographique. Graphique ensuite quand seuls des symboles ponctuels conviennent pour représenter les dénombrements ; la surface du cercle, ou du carré, sera proportionnelle au nombre d'analphabètes. Mais la présentation de ratios devient délicate : comment faire figurer quatre-vingt mille rapports de masculinité, sinon en multipliant les cartes sélectives ?

Et à exiger plus de détail de la carte, on risque d'en biaiser gravement la perception, privilégiant le comportement d'une population minoritaire, voire atypique du point de vue de certains attributs. Regardons une carte des états mexicains, elle est proche des agrégats nationaux, elle nous apprend peu ni ne

16• Au Mexique, on dispose dans le CD-ROM CODICE90 de plus d'un millier de données chiffrées par *municipios*. Ce chiffre tombe à 71 pour les AGEB urbaines, à 37 pour les localités rurales.

nous trompe beaucoup. Avec deux mille quatre cents *municipios* de taille très inégale, ressortent les unités spatiales les plus visibles pour être les plus étendues, généralement moins peuplées¹⁷. Les villes, au contraire, n'ont pas assez de surface pour accrocher le regard, elles ne se voient pratiquement que par l'influence irradiante exercée sur les populations environnantes. Dans les cartes selon les localités, un seul symbole sera utilisé pour l'exceptionnelle population de la capitale dans sa grande variété interne, comme pour les quelques individus d'un village dans sa spécificité anecdotique. La carte détaillée dévoile plus nettement les caractères des gens dispersés en des lieux peu denses, soit des populations rurales ou isolées. L'impact visuel que laissent les populations urbaines, mieux rémunérées, moins fécondes, mieux éduquées et surtout très nombreuses, est moins perceptible. La carte biaise alors notre perception en réduisant l'essentiel, soulignant peut-être l'atypique.

Il est possible d'atténuer ces déformations indésirables par le tracé d'isolignes qui joignent les valeurs identiques déduites par interpolation. C'est pratiquer une inférence quelque peu aveugle quand la localisation des données est parcimonieuse et arbitraire¹⁸. L'approximation sera meilleure si on dispose d'un grand nombre d'observations. La carte qui en résulte n'est pas nécessairement plus juste que l'*officielle*, elle est seulement soulagée du tracé indésirable des limites administratives, lesquelles ne suggèrent plus une uniformité interne factice mais un dégradé qui tient compte des zones mitoyennes. Le procédé, en tout cas, améliore le confort de la lecture en assimilant l'intensité du phénomène à un relief, la surface statistique pouvant être représentée en trois dimensions. Cependant, cette transformation, opérée sans précaution, peut fausser notre perception dans la mesure où les interpolations dépendent de la distance entre les lieux statistiques. La représentation des ratios n'en sera pas altérée, la proportion de célibataires par exemple,

17• Cela n'est pas une loi du peuplement mexicain, mais il est logique qu'un minimum de citoyens étant nécessaire pour justifier l'administration d'un *municipio*, ceux situés dans les zones semi désertiques du nord soient plus étendus.

18• Surtout pour une information produite pour quelques zones. Les centres géographiques de chaque polygone ne reflètent pas nécessairement la situation de la population, normalement plus nombreuse au chef-lieu de l'unité administrative. Le centre de gravité démographique constitue une autre alternative.

car il est facile d'admettre une répartition régulière d'un comportement. En revanche, l'effet risque d'être désastreux pour les dénombrements absolus. Une population isolée, comme celle des villes septentrionales du Mexique, verra son importance loin étalée, soit jusqu'au prochain peuplement suivant. En revanche, le poids démographique d'un plus grand centre urbain entouré de localités peu peuplées sera immédiatement contenue, et donc peu visible. De même l'effet visuel du lissage restera influencé par la taille des polygones. Le carroyage préalable de l'espace, avec attribution de valeurs nulles au vide humain, évite cette distorsion puisqu'il définit une trame régulière qui ne dépend pas de la distance entre les points documentés par le recensement.

Le sens de la carte : interprétation et modélisation spatiale

Ces quelques remarques illustrées ne sont pas des règles prétendument utiles à la confection de la carte statistique, plutôt une mise en garde envers son interprétation précipitée, ou qui se verrait égarée par les qualités graphiques recherchées. Réduire les attributs et l'espace contribue à faciliter la lecture des tendances majeures : graduation nord-sud, gravitation urbaine, contour démographique des climats... au risque d'arrêter l'analyse à ces ébauches. Car la carte simplifie ce qui est déjà une schématisation de l'espace, sur la foi de statistiques abrégées. Cette simplicité recherchée et nécessaire ne doit pas faire oublier que l'intelligence d'un tel document constitue un exercice aussi complexe que l'immense diversité dont il témoigne.

Pour réussir une carte construite avec des nombres, il convient que l'analyse statistique précède sa confection, accompagne son interprétation ; carte qui à son tour, à l'instar de toute graphique (Bertin, 1977), affinera la construction et le calibrage des modèles. Certains modèles sont normatifs, ils explorent les conséquences logiques d'une hypothèse, d'un comportement supposé ; d'autres tentent une explication, sur la foi de l'observation du réel, de données chiffrées par exemple. Confondre les uns et les autres revient à pratiquer une manipulation subtile de la connaissance : ce qui est supposé à des fins analytiques est présenté comme une expression des faits, sous le couvert d'une formalisation

sophistiquée. Sans même aborder le domaine des statistiques spatialisées, on évoquera en deux mots quelques-unes des imprudences, ou réelles difficultés, qui peuvent habiller une carte fautive d'une interprétation vraisemblable, ou le contraire.

Une première série de limitations découle de la nature des données censitaires : elles ne proviennent pas d'une expérimentation, comme en pratiquent les sciences exactes, mais de l'observation. Le modèle, qui serait défini par l'expérimentation, doit alors être emprunté à une théorie ou bien élaboré lors d'une analyse exploratoire des données (Tukey, 1977), elle-même quelque peu démunie face à la localisation des données démographiques. D'autres difficultés apparaissent lors de la validation des modèles ou hypothèses. Deux d'entre elles seront vite évoquées : la dépendance des attributs localisés et la non-normalité des distributions qui résultent du regroupement des unités statistiques primaires, ménages ou établissements économiques.

Comment comprendre les cartes censitaires quand manquent les modèles qui intègrent la dimension spatiale, mais aussi les bases théoriques pour les construire ? Les modèles, la démographie ne sait pas s'en passer, pour ajuster les données imparfaites, instruire les projections de population et les politiques démo-économiques en simulant leur dynamique ou l'impact de tel déterminant. Certains prennent en considération la division régionale pour estimer la croissance des populations ouvertes, celles dont la reproduction et la structure sont infléchies par la migration (Rogers, 1975). Est bien fourni le corpus théorique de la migration et de la distribution des lieux habités, mais on est plus démunis pour analyser les configurations spatiales de la transition vitale (Noin, 1984), de l'éducation ou du logement. Reste alors à tenter de démêler l'information observée. Les alternatives méthodologiques sont alors aussi variées que les buts poursuivis, pas nécessairement conciliables (Leamer, 1978) :

- un modèle de régression est retenu mais les facteurs pouvant être mesurés de maintes manières, l'attribut le plus approprié est recherché ;
- on aimerait réduire la complexité du modèle tout en conservant l'adéquation de l'ajustement, en particulier parce qu'un grand nombre de variables sont liées entre elles ;
- on tente de trouver des variables complémentaires pour

obtenir une meilleure explication du modèle, pour peu que l'on en dispose pour les unités spatiales considérées ;

- on transforme des variables pour réduire la part non expliquée de l'observation.

Ces cheminements complémentaires sont de règle pour aborder les données censitaires et vitales, à la fois redondantes et incomplètes. Car si l'utilité gestionnaire des SIG n'est plus à prouver, leur intérêt analytique en revanche se heurte à la difficulté de développer des enquêtes spécifiques à l'échelle nationale pour produire l'information jugée appropriée à la vérification d'une hypothèse. Un bon registre civil des événements vitaux et du profil des individus concernés est une voie vers une solution. Une autre restriction tient à la relative rareté des méthodes statistiques exploratoires intégrant la dimension spatiale, ou spatio-temporelle. L'utilité de celles développées pour les sciences du sol, ou de la biologie animale (Matheron, 1963), reste, dans la pratique, limitée. Il sera, par exemple, rarement nécessaire de s'assurer que telle configuration spatiale n'est pas due au hasard de telle loi de probabilité, tant il est souvent vérifié que celles des populations humaines sont bien structurées. Le calcul de variogrammes multivariés et multidimensionnels (Dauphine *et al*, 1988) peut s'avérer d'interprétation difficile ou inspirer des résultats triviaux dans le cas des comportements humains.

L'application de tests statistiques et de régressions linéaires multivariées suppose généralement que la distribution statistique des attributs suive une loi normale (gaussienne) et que les observations soient indépendantes. Deux conditions rarement remplies pour les données localisées, une lacune susceptible de fausser l'estimation des modèles. Ainsi, l'association statistique de deux attributs peut être un artefact dû à la dépendance spatiale de l'information (Haining, 1992 : 29). La normalité de la distribution est vérifiable à l'aide des logiciels d'analyse statistique, les techniques de transformation des variables coupables sont disponibles pour les corriger (Afifi et Clark, 1990). Mais la correction est d'autant plus fastidieuse et délicate que les attributs sont nombreux, les déformations fréquentes, sinon la règle pour les données censitaires. Cela tient à l'introduction du découpage biaisé de l'espace selon lequel sont regroupés les

ménages, unités élémentaires du recensement. La division administrative, qui introduit une segmentation de l'univers statistique, prend en compte des entités de peuplement dont la taille démographique suit une progression exponentielle, résultat de la concentration urbaine. Or cette taille influe de nombreux attributs, la plupart sans doute de ceux que les recensements enregistrent. Cela rend la régression linéaire des attributs localisés délicate à mener, et surtout à interpréter, de même que toutes les techniques statistiques qui l'emploient (analyse factorielle ou par composantes principales). En effet, que conclure honnêtement quand les corrélations changent avec le découpage effectué ? Et qu'elles n'indiquent pas une causalité, seulement une correspondance entre les géographies (Hubert, 1985). Les phénomènes de diffusion contribuent à diluer les incidences de tel facteur dans l'espace ; cela peut justifier d'introduire dans l'explication d'un phénomène des variables observées dans des sites voisins (Haining, 1990 : 29). L'analyse factorielle des taux et proportions est de surcroît difficile à conduire en toute rigueur quand chaque indice doit être pondéré par son dénominateur. Une seule pondération commune à tous les attributs, par exemple la taille démographique de l'unité spatiale observée, n'est alors qu'un pis-aller. Les difficultés augmentent quand chaque variable a subi une transformation différente.

Des corrélations de proximité entre variables localisées traduisent le fait que si elles sont liées, elles le sont plus dans un espace proche que lointain. L'estimation de ces autocorrélations spatiales est indispensable à l'analyse afin d'éviter des conclusions erronées. D'abord, lors de l'application des tests d'inférence, quand on reconstitue les structures spatiales sur un nombre réduit de points. Les méthodes supposent l'indépendance des observations, qui évidemment disparaît si elles sont associées dans l'espace. La réduction de l'intervalle de confiance doit alors être considérée avant d'avancer une interprétation. La mesure de la variabilité dans l'espace est requise également pour éviter des analyses purement visuelles sur la foi d'une cartographie dont on a signalé le flou. Mais aussi parce que certaines homogénéités ne sont pas immédiatement perceptibles ; songeons à celles que dessinent les réseaux par exemple, ou les distances *économiques* évaluées par le coût du déplacement. À supposer que la finesse

de la carte en permette la lecture, seul un oeil averti serait capable d'un tel discernement, surtout quand les unités spatiales doivent être pondérées, non par la surface qui se voit, mais par le poids démographique. Or, le calcul des variogrammes autorise plusieurs modalités de regroupement, d'association des lieux entre eux, pas seulement de proximité dans l'espace physique. Les SIG sont ici de grande utilité en automatisant ces estimations quand les critères d'association sont contenus dans la base : appartenance à un réseau de routes, algorithme de triangulation spatiale. Cette discipline s'impose quand on songe que lire une carte à des fins d'analyse reviendrait à estimer les paramètres d'une distribution, un coefficient de corrélation à... l'oeil nu, sur la foi des expressions graphiques.

Conclusion

Ces quelques remarques sur la carte comme outil de la désinformation ou/et de l'erreur ont négligé les premiers risques d'égarement, ceux associés à l'analyse statistique des données sans même considérer leur dimension spatiale. Si l'utilité de la carte ne fait aucun doute pour l'analyse exploratoire des recensements, on doit néanmoins garder à l'esprit qu'elle va d'abord refléter la disponibilité et la qualité de l'information plutôt que l'essence du phénomène. Phénomène dont l'apparence cartographique sera l'expression finale d'une série de choix sur la réduction de l'attribut, de son échelle spatiale, de sa représentation graphique. Faut-il donc rappeler cette prudence élémentaire qu'on ne peut faire l'économie des méthodes statistiques, spatialisées ou non, lesquelles doivent impérativement précéder la confection de la carte, qu'elle soit analytique ou seulement illustrative ?

Si la démographie utilise peu la cartographie censitaire, c'est faute de pouvoir conduire l'analyse pour une partition détaillée du territoire avec le même raffinement que pour l'ensemble du pays. Les difficultés surgissent dès que l'on examine des unités assez fines pour nous révéler les processus spatiaux à l'oeuvre. La qualité défailante des statistiques à ce point détaillées, celles des registres civils en particulier, interdit la construction d'indicateurs

solides. Il est ensuite fort difficile d'épurer la mesure des phénomènes de leur influence réciproque, en particulier de celle de la migration souvent méconnue pour toutes les communes ou localités d'un état. Manquent enfin les modèles théoriques, d'ajustement ou de simulation capables de guider l'interprétation de la carte démographique. Ainsi, dépourvue des outils analytiques élémentaires, l'intelligence de la carte démographique reste un exercice périlleux.

Ouvrages cités

- ACKERMAN, E. A. (1959), - *Population and natural resources*, in Hauser, P. M. et DUNCAN, O. (ed.) "The Study of Population : an Inventory and Appraisal", Univ. of Chicago Press, Chicago, pp. 621-48.
- AFIFI (A. A.) ET CLARK (V.), 1990 - *Computer-aided multivariate analysis*, Van Nostrand Reinhold Company, New York.
- BERTIN (J.), 1977 - *La graphique et le traitement graphique de l'information*, Flammarion Editeurs, Paris
- BRUNET (R.), 1987 - *La carte mode d'emploi*, Paris, Fayard/Reclus, 270 p., 269 ill.
- BRUNET (R.), 1992 - *Les mots de la géographie*, Paris, Reclus - La Documentation Française, 518 p.
- CAUVIN (C.), Reymond (H.), Serrao (A.), 1986 - *Discretisation et représentation cartographique*, GIP-Reclus, Montpellier.
- Dauphine (A.), Voiron-Canicio (C.) -1988- *Variogrammes et structures spatiales*, GIP-RECLUS, Montpellier.
- DELAUNAY (D.) -1991- *Démographie et systèmes agraires, un exemple d'inférence statistique dans un Système d'Information Géographique*, in Mullon (C.) (ed.), *Seminfor 4 : le transfert d'échelle*. Orstom, Paris, pp. 255-266.
- HAINING, ROBERT, 1990 - *Spatial data analysis in the social and environmental sciences*, Cambridge University Press, Cambridge.
- HUBERT (L.), Golledge (R.C.), Costanzo (C.M.) et Gale (N.), 1985 - Measuring association between spatially defined variables : an alternative procedure. *Geographical Analysis*, num. 17, pp. 36-46.
- LEAMER (EE.), 1978 - *Specification Searches : Ad Hoc Inference with Non Experimental Data*, Wiler, Nex York
- MATHERON (G.), 1963 - *Traité de Géostatistique appliquée*, Editions Bureau de Recherches Géologiques et Minières, Paris.
- NOIN (D.), 1984 - *La géographie de la population*, in "Les concepts de la géographie humaine, Bailly, A. et al., Masson, Paris, pp. 75-83.
- PRYOR, ROBIN (J.), 1984 - *Methodological Problems in Population Geography*, in "Geography and Population, Approaches and Applications", Clarke, J. (ed.), Pergamon Press, Oxford, pp. 25-34.
- ROGERS, 1975 - *Introduction to multiregional mathematical demography*, Wiler, New York.
- SHIVA (S. HALLI) ET VANINADHA RAO (K.), 1992 - *Advanced Techniques of Population Analysis*, Plenum Press, New York.
- TUKEY (J. W.), 1977 - *Exploratory data analysis*, Addison-Wesley Publishing Company, Massachusetts.
- UDO (R.), 1976 - *Applied Population Geography : A Survey*. I.G.U. Commission on Population Geography, Edmonton.
- ZELINSKY (W.), 1971 - The hypothesis of the mobility transition, *Geographical Review*, 61, 219-49.
- ZELINSKY (W.), 1979 - *The demographic transition : changing patterns of migration*, in *Institute of life and I.U.S.S.P., Population Science in the Service Mankind*, Conference on Science in the Service of Life, Vienna, 165-89.

Commentaires des cartes

Toutes les cartes représentent un même attribut simple : le nombre de personnes par foyer. Sa mesure est peu exposée à de grands risques d'erreur, sa variabilité n'est pas due à des effectifs insuffisants, sa distribution est proche d'une loi normale. En dépit de cette relative transparence, les cartes en donnent une interprétation floue, infléchie par l'échelle choisie, les discrétisations retenues, les zonifications changeantes.

La carte municipale (fig. 44) localise 2 403 valeurs de l'attribut et celle des localités près de 80 000 points (fig. 51). Les cartes régionales sont composées de 192 zones (les régions de programmation) (fig. 45 à 50). Certaines font varier la méthode de discrétisation, soit les classes choisies pour représenter les valeurs contiguës de l'attribut ; d'autres sont simplement des expressions différentes du même phénomène, telle l'importance relative d'une valeur de la distribution plutôt que d'un indicateur synthétique comme la moyenne. Ces variations cartographiques illustrent les dangers d'une interprétation visuelle qui, de surcroît, privilégie les unités géographiques vastes mais peu peuplées. Enfin, d'autres cartes appliquent des méthodes de lissage supposées atténuer certaines de ces déformations mais qui se révèlent plus difficiles encore (fig. 52, 53).

Ne sont vraiment discernables que les grandes configurations nationales, la compréhension du détail nous échappe, en dehors de quelques présences urbaines. Le phénomène qui coïncide le plus avec la taille des ménages est évidemment la fécondité, c'est à peu près tout ce que nous en dit l'analyse factorielle qui signale également les facteurs normalement associés aux parités élevées : piètre éducation et faible urbanisation. Bien des questions restent en suspens. L'homogénéité du nord est intrigante (l'immigration ne suffit pas à l'expliquer), de même la concentration de bourgs peuplés de familles élargies autour de la capitale, ou encore l'atomisation des foyers dans le centre de Oaxaca, aux environs côtiers de la ville de Veracruz. Il conviendrait de simuler la variation géographique du phénomène sur la base de la nuptialité, de la reproduction des familles, de la migration avant de rechercher d'autres différenciations de type culturel ou historique.

Fig. 44 : Le nombre de personnes par logement : cartographie municipale

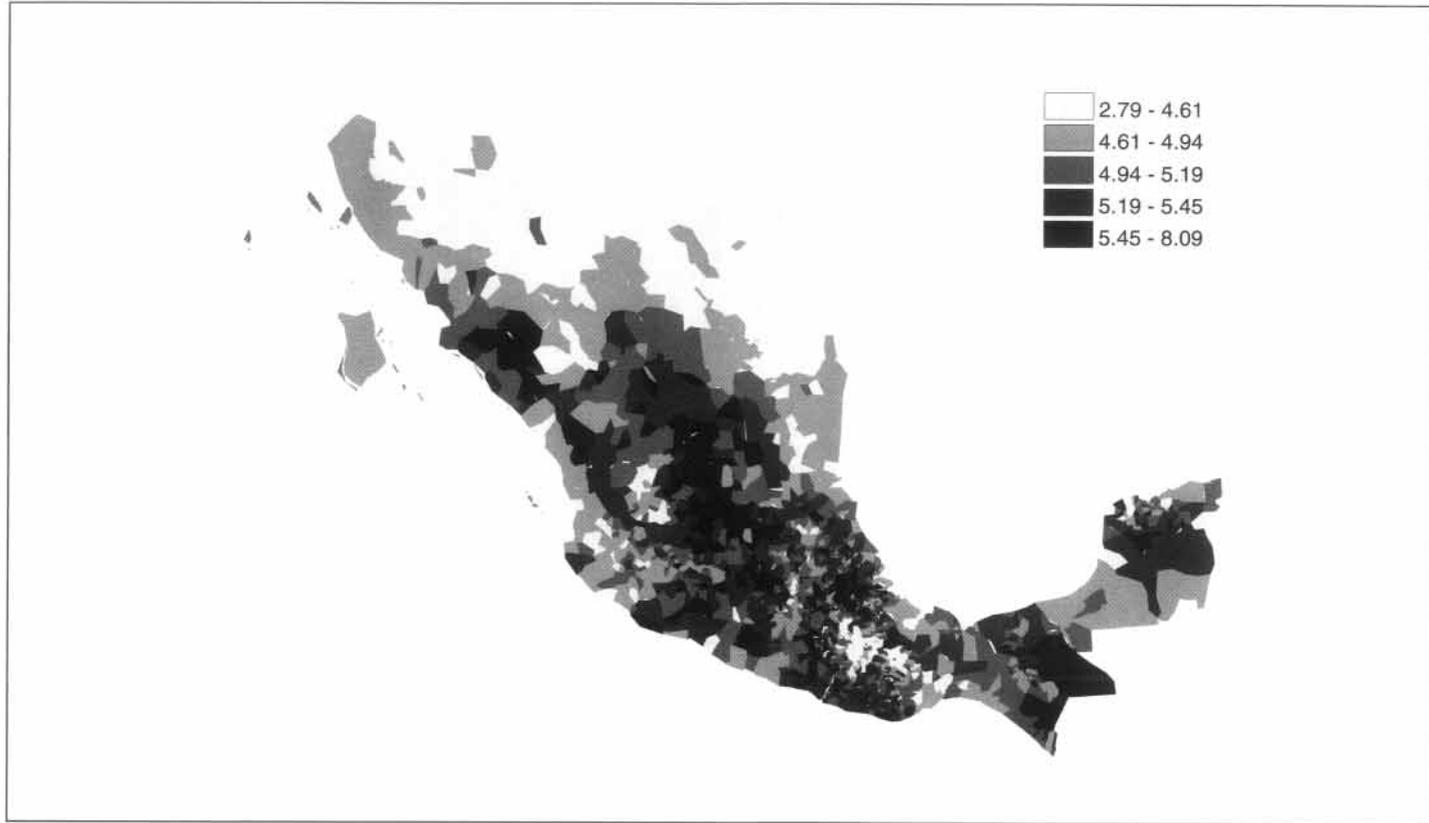


Fig. 45 : Proportion de logements abritant deux personnes

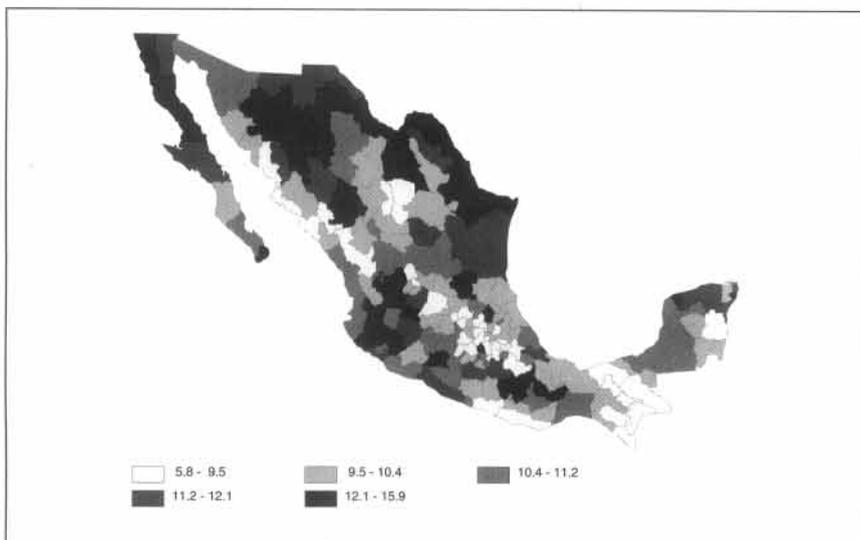


Fig. 46 : Proportion de logements abritant cinq personnes

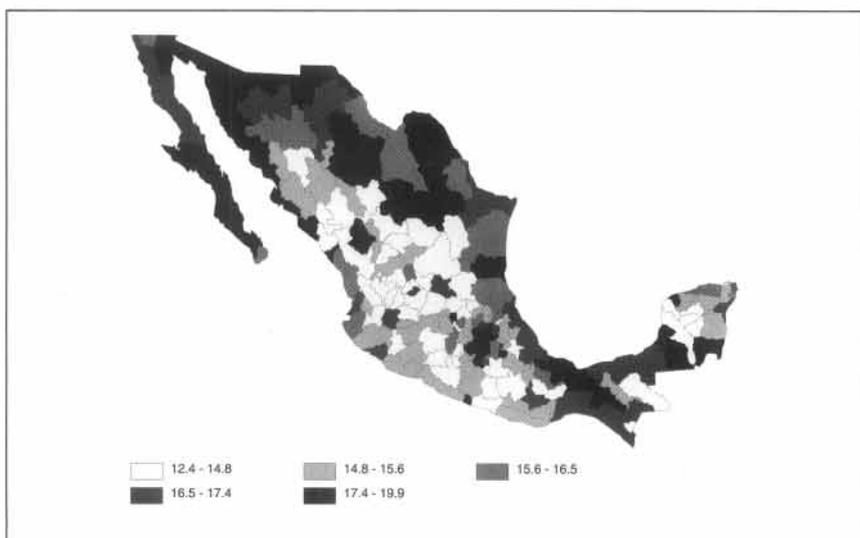


Fig. 47 : Proportion de logements abritant neuf personnes

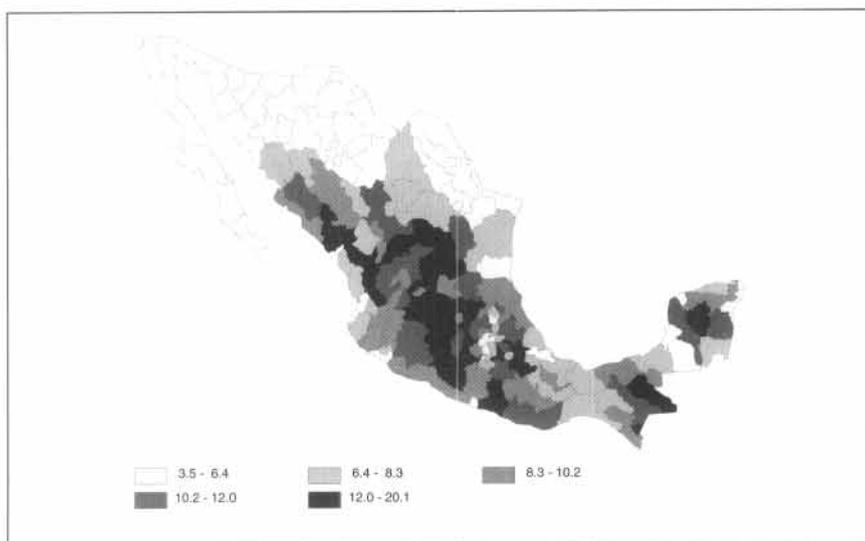


Fig. 48 : Nombre moyen de personnes par logement

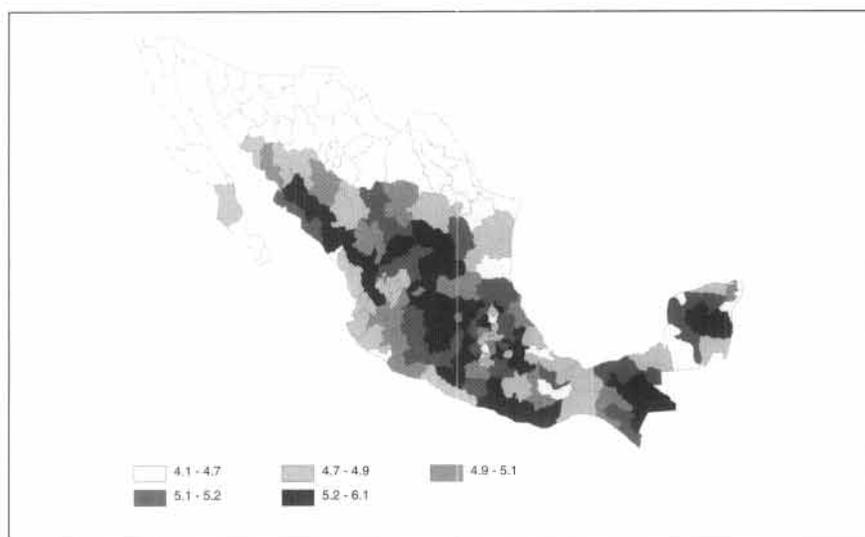


Fig. 49 : Valeurs régionales de l'attribut discrétisées selon les quantiles ; les points représentent les *municipios* qui rassemblent moins de cinq personnes par logement

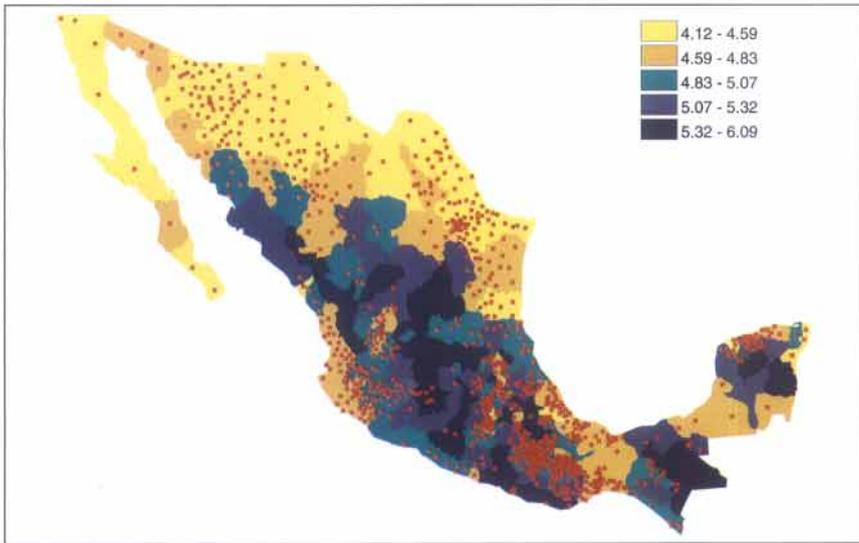


Fig. 50 : Valeurs régionales de l'attribut discrétisé selon les moyennes mobiles ; les points représentent les *municipios* rassemblant plus de cinq personnes par logement

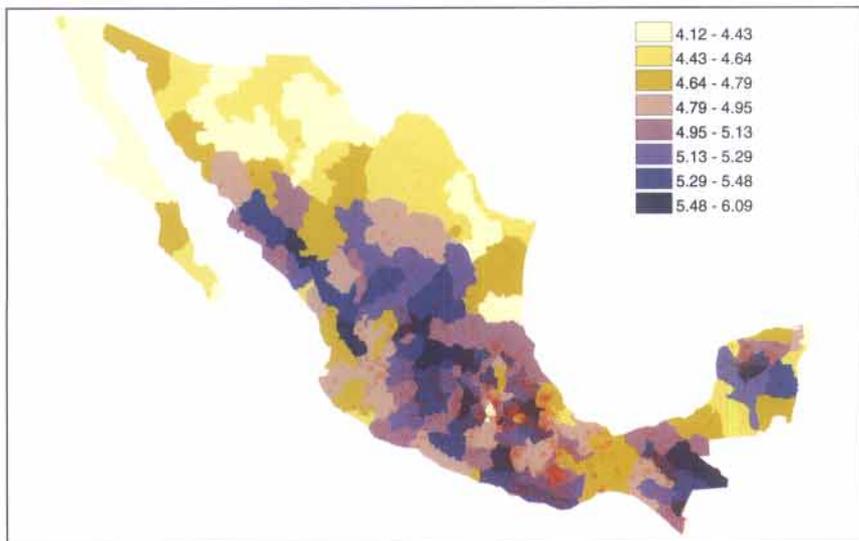


Fig. 51 : Les localités de plus de quatre logements sont classées selon la taille démographique de ces foyers, ces 80 000 points donnent la géographie la plus fine de l'attribut.



Fig. 52 : Interpolation des valeurs régionales de l'attribut

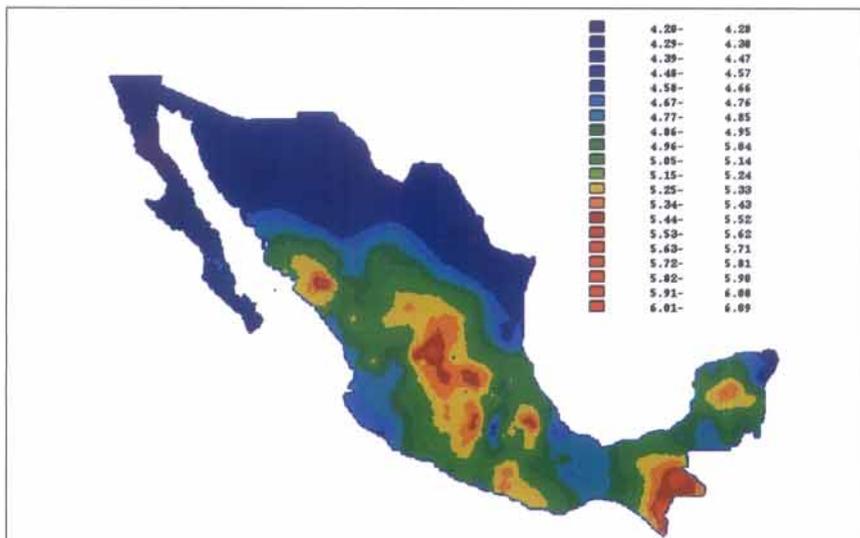


Fig. 53 : Interpolation des valeurs municipales de l'attribut

