

DE L'ENQUÊTE DÉMO-SOCIO-ÉCONOMIQUE PAR SONDAGE A L'EXPRESSION CARTOGRAPHIQUE

Ph. HAERINGER

*Géographe de l'ORSTOM
Abidjan (Côte d'Ivoire)*

Il arrive bien souvent qu'un chercheur ait à exploiter les résultats, voire les documents bruts d'une enquête effectuée par d'autres personnes que lui, ayant agi dans un cadre qui n'est pas forcément — et généralement pas — celui d'une interrogation proprement scientifique. On conçoit que dans pareil cas de nombreuses difficultés surgissent : d'abord pour la compréhension et l'interprétation des méthodes et des critères d'analyse définis par les auteurs de l'enquête (ces éléments ne sont pas toujours clairement explicités dans les rapports), ensuite et surtout pour l'ajustement des données disponibles à l'objectif poursuivi. Quant l'objectif du chercheur est de consigner sur une carte les variations dans l'espace d'un phénomène géographique continu et qu'il dispose pour ce faire d'une enquête par sondage, il y a toute chance pour que la difficulté la plus sérieuse soit, pour le chercheur, de s'accommoder d'un plan de sondage non adéquat.

1. Une information parfois considérable

Dans le domaine des sciences humaines, c'est évidemment aux enquêtes de nature socio-économique (vocabulaire qui inclut souvent une approche démographique) que l'on se trouve le plus fréquemment confronté. Menées généralement par des sociétés d'étude privées pour le compte des ministères, elles ont pris comme on sait un grand essor en Afrique, au moins dans certains pays, relayant dans une certaine mesure et dans un style bien différent les sources purement administratives (recensements, rapports annuels) de

l'époque coloniale. En raison de l'intérêt pratique et immédiat qu'on en attend, il leur est consacré des moyens dont ne disposent pas les instituts de recherche fondamentale. Les chercheurs ne peuvent donc passer à côté de la masse d'information qu'elles représentent, particulièrement quand ils se proposent de consigner dans un ouvrage général l'ensemble des connaissances acquises sur un pays donné, ce qui est le cas lorsqu'est entrepris la confection d'un atlas, qu'il soit national ou régional. C'est même, en partie, la sortie des résultats d'une série d'enquêtes régionales qui, en Côte d'Ivoire, a suscité la mise en œuvre de l'atlas auquel l'ORSTOM travaille.

Or il est bien évident qu'une enquête socio-économique n'a en général pas pour fin de restituer sur des cartes les phénomènes qu'elle étudie. Elle n'a même pas, le plus souvent, l'ambition de caractériser les diverses parties de l'espace auquel elle s'applique. En général, le but poursuivi est un bilan statistique global. Si, dans la plupart des cas, il est au moins partiellement détaillé en fonction d'un découpage spatial, l'accent est mis sur d'autres principes de stratification et la division de l'espace en secteurs est plus conçue comme une commodité ou une nécessité administrative que comme une fin en soi. Cadre de travail établi préalablement à l'enquête, elle n'est que rarement réexaminée lors de l'analyse des résultats. Lesquels ne sont donc pas mis à profit pour en améliorer la pertinence.

Les matériaux qu'offrent ces enquêtes (en réalité fort diverses) au géographe qui ne les a pas effectuées sont donc, inévitablement, assez malaisés à utiliser

lorsqu'après coup et en dehors d'elles il tente de leur donner une suite cartographique. Cette difficulté est surtout sensible lorsque l'on est contraint de se contenter de résultats publiés, donc passés au moule d'une analyse en l'occurrence non géographique, où la disparition du détail consécutive à cette analyse s'est faite en pure perte pour le géographe. Par contre, les documents bruts de l'enquête, qui conservent le détail et, surtout, permettent une localisation dans l'espace, offrent évidemment une matière beaucoup plus maléable et plus riche. On est donc conduit à tenter d'y accéder. Mais l'exploitation en reste néanmoins délicate et quelque peu décevante, essentiellement parce que le plan de sondage n'est jamais celui qu'il aurait fallu.

2. L'inadéquation des plans de sondage

Il arrive qu'un sondage porte sur des entités immatérielles et donc non repérables dans l'espace : par exemple des événements démographiques, des consultations médicales ou des opérations bancaires, à condition toutefois qu'ils aient été enregistrés sur un livre ou dans un fichier. Le tirage fait et l'exploitation accomplie, il peut y avoir lieu, éventuellement, de montrer l'impact que ces événements ont pu avoir dans l'espace. Mais ce n'est pas cet impact qui pouvait fonder le plan de sondage. Que le chercheur soit géographe ou non.

En fait et tout au moins en Afrique, c'est plus souvent l'homme lui-même que les traces de ses actes que l'on interroge. Même l'exploitation d'un fichier, lequel mémorise généralement des événements ou des actes, peut n'être, par le biais de ceux-ci, qu'une façon de constituer un échantillon d'humain : élèves des écoles, malades des hôpitaux, ouvriers d'une usine. Mais ce mode d'approche d'une population est plus propre à des enquêtes très spécialisées qu'aux enquêtes socio-économiques de type courant. Dans le cadre de celles-ci, l'homme est normalement appréhendé en tant qu'habitant et donc contacté sur son lieu de résidence. La qualité du sondage importe alors beaucoup au géographe.

Ce n'est pas l'homme lui-même mais son environnement matériel qui fournit l'unité de sondage. La plus petite entité utilisable est le logement, mais l'inventaire des logements est souvent malaisé. Dans le cas d'une enquête urbaine, c'est plutôt l'unité cadastrale qui est choisie, c'est-à-dire la cour ou l'immeuble, quel que soit le nombre de logements qu'ils comportent. Dans le cas d'une enquête régionale ou natio-

nale, c'est le village ou même le canton qui sert d'unité. Au moins dans un premier temps. Car généralement il est procédé à un second tirage à l'intérieur de l'échantillon pour atteindre, cette fois, la cour ou l'exploitation familiale.

On sait que théoriquement seul un sondage parfaitement aléatoire garantit la validité des résultats. Plus exactement, c'est le seul qui permette de calculer l'erreur due à l'échantillonnage. Mais il existe bien d'autres causes d'erreurs quand on interroge des hommes sur des questions aussi complexes que leurs budgets ou leurs temps de travaux, ou quand on fait mesurer par un enquêteur de niveau modeste et de conscience professionnelle limitée les contours objectivement flous d'un champ sur brûlis forestier. Tout cela n'a rien à voir avec la vérification du poids de millions de cigarettes à la sortie d'une machine. Les auteurs d'une enquête économique sont nécessairement conscients de ces difficultés, mais puisqu'ils détiennent le moyen d'éliminer une cause d'erreur, celle de l'échantillonnage, ils ne se privent pas d'y avoir recours et de se libérer l'esprit sur ce point.

Toutefois l'approche probabiliste laisse place et même invite à des finesses d'échantillonnage qui ne sont pas toujours exploitées. Elle n'oblige pas à traiter l'ensemble d'un pays ou d'une ville comme un tout indifférencié. Ce que le géographe a le plus l'occasion de regretter, c'est qu'il n'ait pas été tenu compte, dans le plan de sondage, des variations de la densité de population. Quand tel est le cas, les régions sous-peuplées sont trop faiblement représentées pour qu'on puisse en tirer des résultats significatifs (1). La répartition des densités est pourtant, au moins en ordres de grandeur, connue. Sans du tout entacher le caractère aléatoire du sondage et tout en conservant une même taille d'échantillon, il est possible, en procédant à une stratification, de faire varier les taux de sondage en fonction de la densité et d'obtenir ainsi en toute région une égale précision de connaissance.

Une telle proposition se voit parfois opposer l'objection selon laquelle l'ampleur de l'échantillon n'autorise

(1) Dans l'une des enquêtes régionales qui ont été menées en Côte d'Ivoire il y a quelques années, plus de cinquante villages témoignaient de la minuscule zone dense de Korhogo, d'ailleurs très homogène, tandis que le vaste secteur de Kong n'était représenté que par un seul village, autrement dit ignoré. Malgré le très faible poids démographique de ces étendues, le géographe eut aimé savoir ce qui s'y passe, d'autant plus qu'elles correspondent à un isolat ethnique original. Or, quelques points d'enquête de moins autour de Korhogo et quelques points de plus près de Kong auraient tout changé ici sans rien changer là.

pas une analyse à l'échelon local et que cette amélioration du sondage est donc sans objet. Or on a vu que pour diverses raisons un certain découpage du domaine d'enquête intervenait presque toujours, y compris dans la présentation des résultats. Ceux-ci sont pour la plupart donnés sous forme de tableaux croisés où sont mis en corrélation deux ou trois facteurs. Dans un certain nombre de tableaux, le découpage géographique constitue l'une des variables. Si la taille de l'échantillon est conçue pour ne permettre une analyse complète qu'à l'échelon global, avec des croisements de deux, de trois, de n variables, il reste donc (en schématisant) la possibilité d'une analyse moins fine avec croisements de $n-1$ variables à l'échelon régional, ce qui suffit à donner matière à cartographie.

Au lieu d'une amélioration du sondage aléatoire simple par une stratification (dont le critère peut être aussi bien la répartition ethnique ou les conditions climatiques que la densité de population), c'est parfois une aggravation de son caractère arbitraire qui est opérée. Pour des raisons pratiques évidentes, il est souvent procédé à un sondage en grappes. Au lieu de tirer des unités isolées, on tire des groupes d'unités voisines. Les déplacements des enquêteurs s'en trouvent réduits, mais la dispersion des villages ou parcelles urbaines tirés est plus inégale encore que dans le cas précédent, même si, pour compenser le biais introduit par le critère de voisinage, la taille de l'échantillon est grossie. De vastes zones, même correctement peuplées, peuvent être complètement ignorées par le sondage (1).

3. Les limites du sondage aléatoire

Cependant, affiné ou non, le sondage aléatoire est-il vraiment la seule manière raisonnable de procéder ? Nous évoquions plus haut les multiples causes d'erreur pouvant entacher, quelles que soient les techniques d'échantillonnage, les résultats d'une enquête en milieu humain. Dans des domaines d'étude où les réponses demandées sont loin de s'exprimer par oui ou par non, ces risques d'erreur sont en général bien supérieurs aux 5 % que l'on s'accorde d'ordinaire pour l'erreur d'échantillonnage.

(1) Toujours en Côte d'Ivoire, le hasard a voulu que dans l'enquête qui couvrait le plus vaste territoire (le Nord-Ouest), toutes les grappes de village tirées se rangent sur le pourtour de la région. Aucun des cent villages-échantillon n'a donc pu rendre compte des zones intérieures, pourtant non dépourvues de population.

Dès lors l'attachement au tirage purement aléatoire, sans être vain puisqu'il écarte malgré tout un type d'erreur, ne paraît pas être une attitude indispensable. Au contraire, puisque les avantages du recours au hasard apparaissent si ténus, il semble intéressant d'y renoncer pour jouer à fond des cartes que l'approche probabiliste ne permet pas d'utiliser complètement : les connaissances acquises sur le milieu étudié ; l'initiative du chercheur, en considération de ces connaissances, dans la détermination de l'échantillon ; l'élaboration d'une problématique préalable à l'enquête, dont il soit tenu compte lors de la définition du plan de sondage (alors que d'ordinaire les questions ne sont posées, ou du moins n'interviennent, qu'au moment de l'exploitation des résultats).

Certes ces cartes sont jouables, jusqu'à un certain point, dans le cadre d'un sondage aléatoire, grâce à la technique de stratification que nous venons d'évoquer. Mais les limites en sont vite atteintes en raison de l'obligation de conserver des ensembles de grande taille. Or l'établissement d'une carte sérieuse demande une investigation très nuancée. Il faut que les accidents du milieu naturel, les particularités ethniques ou historiques, le voisinage des villes, des voies de communication puissent être pris en compte, que leur impact sur le phénomène étudié puisse être apprécié. Quel que soit celui-ci, il y a toujours plusieurs paramètres à faire intervenir et, bien entendu, ce n'est pas un découpage en grands ensembles qui peut résulter de leur croisement, mais un maillage serré et comportant inévitablement des unités petites entre toutes. Si l'on veut que l'enquête porte sur chacun des secteurs homogènes ainsi définis (homogènes relativement aux paramètres qu'on a choisi d'éprouver), et à moins de gonfler considérablement l'échantillon, on a toute chance de devoir se contenter d'un faible nombre d'observations par secteur.

Il reste possible, à ce stade, qu'on ne soit pas obligé d'abandonner la rigueur du sondage aléatoire. Quand les paramètres utilisés ne sont pas trop nombreux (2 ou 3) il peut arriver qu'un même type de situation se retrouve en divers points du territoire étudié et qu'ainsi des regroupements de secteurs dispersés mais semblables puissent se faire au niveau de l'analyse. Si l'on arrive ainsi à reconstituer des ensembles assez vastes et réunissant des points de sondage suffisamment nombreux, on est ramené à un cas particulier du sondage stratifié où, simplement, les strates ne sont plus d'un seul tenant. Mais ce cas se présente assez exceptionnellement et surtout dans le cadre d'enquêtes urbaines (deux quartiers populaires ou deux quartiers résidentiels ont plus de chances de

se ressembler que deux régions rurales). De plus, que plusieurs régions ou quartiers entrent dans une même catégorie de classement n'ôte pas tout intérêt à une étude particulière de chacun d'eux, surtout lorsque l'on se propose de dresser des cartes.

4. Renoncer à la démarche probabiliste

Donc, même dans ce cas particulier, mais d'abord dans le cas général, le géographe qui aurait à concevoir une enquête par sondage de grande envergure a tout lieu d'être tenté de considérer en eux-mêmes des groupes restreints de points sondés. S'il accepte la tentation, le tirage au hasard devient une mauvaise recette. S'il doit se contenter de cinq villages dans tel massif montagneux ou dans telle région côtière, il est plus sûr qu'il en fasse lui-même le choix plutôt que de se fier au hasard. Que ces régions aient été au préalable reconnues homogènes relativement à un certain nombre de critères jugés prépondérants ne change rien à la chose : d'autres éléments peuvent se montrer localement déterminants (situation par rapport aux pentes ou plus ou moins grande proximité de la côte) ; et il ne faudrait pas non plus que le hasard désigne cinq villages établis sur le pourtour de la région, ou cinq villages-centres, ou cinq villages d'immigrés.

Faut-il considérer une telle démarche comme déraisonnable et dangereuse ? Je ne le crois pas, Pour l'avoir touché de près je pense que la cécité du hasard derrière lequel on se retranche trop souvent est autrement pernicieuse sous des dehors rassurants. Elle endort l'esprit critique et elle conduit à des pertes d'information considérables. Bien sûr, cinq villages ne peuvent suffire à fonder une statistique régionale, mais s'ils sont bien choisis et tant que l'on n'aura pas les moyens de mettre en place un appareil d'observation plus fin et plus proche de l'exhaustivité, ils demeureront des témoins précieux qui méritent mieux que de disparaître dans l'anonymat d'une statistique nationale ou macro-régionale.

Du reste, il ne faut pas s'exagérer les risques d'une analyse fondée sur un faible nombre d'observations quand elle s'insère dans une investigation de grande ampleur portant elle-même sur un univers intelligible et cohérent. Au moins en milieu rural, les transitions entre deux aspects différents d'un même phénomène sont normalement lentes ; il suffit d'en être conscient lorsque l'on tombe sur un résultat ponctuel apparaissant très en relief. De plus, il est bien évident qu'une telle démarche ne peut s'appliquer qu'à l'observation de phénomènes continus. Personne n'aurait d'ailleurs

idée d'étudier l'implantation d'équipements de nature exceptionnelle (écoles d'apprentissage, centres de vulgarisation agricole) ou la dispersion des sites touristiques à travers un sondage quel qu'il soit.

Mais c'est la part accrue de l'initiative du maître d'enquête, non seulement au niveau de l'analyse mais surtout dans le choix des points sondés qui attire le plus de prévention. Or on est bien obligé, de toute façon, de lui faire confiance sur bien d'autres points aussi déterminants : mise au point des questionnaires, contrôle des équipes, appréhension et interprétation de faits d'observation souvent complexes, traitement mathématique des résultats. Les manques à ces niveaux ne sont pas tous vérifiables, tandis que le choix raisonné des villages peut à tout moment être jugé par quiconque, pour peu que la localisation en soit fournie dans le rapport d'enquête (ce que les auteurs d'un sondage aléatoire ne jugent pas toujours utile de faire).

Une question demeure : l'intérêt ainsi défini de la cartographie et, d'une façon plus générale, de la géographie entre-t-il en conflit avec celui de l'économiste (ou du démographe ou du sociologue) attaché à des résultats d'ensemble ? A lui (ou à eux) de répondre, mais il me semble qu'un éclairage attentif des disparités à l'échelon local ne peut qu'améliorer l'interprétation des phénomènes aux échelons supérieurs et que cela vaut bien la satisfaction de calculer des intervalles de confiance généralement illusoire et qui ne sont d'ailleurs jamais publiés. En tout état de cause et quels que soient les points de vue, il est incontestablement dommage qu'une masse considérable de travaux de terrain n'aboutisse qu'à préciser l'image globale d'une production ou d'une démographie, par exemple, sans faire notablement progresser la connaissance du pays dont elles sont issues.

5. Tirer parti malgré tout des documents disponibles

Voyons maintenant ce qui peut être fait malgré tout en matière de cartographie lorsque l'on a à sa disposition les résultats d'une enquête socio-économique par sondage d'où tout souci d'expression cartographique a été exclu. Il n'est pas douteux que les problèmes pratiques et théoriques se présentent assez différemment selon les cas d'enquête et selon le type de fait que l'on se propose d'étudier. Je me bornerai donc à énumérer les questions qui me paraissent se poser au fil d'une démarche que, par la force des choses, je conçois surtout en fonction de mes propres expériences. Loin de lui attribuer une valeur absolue, je préfère ne pas m'y attarder dans le cadre de cet

exposé et renvoie le lecteur au titre suivant, qui n'est autre que la notice d'une des planches de l'Atlas de Côte-d'Ivoire dont l'établissement m'échut.

Réexploitation des documents de terrain (fiches d'enquête)

Les difficultés sont nombreuses :

- L'obtention de ces documents n'est pas toujours facile.
- Après leur exploitation, la société d'étude qui les détient n'y prête plus guère attention. On ne les retrouve pas toujours au complet, ni en bon état de classement.
- On est malgré tout tenu, généralement, de les rendre rapidement, ce qui oblige à un énorme travail de transcription.
- Leur réexploitation, travail inévitablement considérable, est compliquée par l'absence des responsables de l'enquête : certaines clés de compréhension des écritures échappent, qu'il faut péniblement retrouver (1).

Mise en place de points d'observation supplémentaires et utilisation des autres sources existantes

Ces autres sources étant par hypothèse peu abondantes, puisque l'enquête exploitée eut lieu, on est tenté de garnir les blancs laissés par le sondage aléatoire de nouveaux points d'observation. Mais si le plan de sondage est très défavorable, les points à ajouter risquent d'être presque aussi nombreux que les points existants, l'opération devenant ainsi inabordable.

(1) Au cours de la réexploitation que j'eus l'occasion de faire des enquêtes régionales de Côte-d'Ivoire déjà évoquées, je me suis heurté, par exemple, au problème suivant. Dans la région du Centre, les parcelles agricoles avaient été mesurées en « pas » carrés et non en mètres carrés. Le pas de chaque enquêteur (ils étaient 32) avait été étalonné. Mais où se trouvait, deux ans après l'enquête, la liste des taux de conversion en système métrique ? Il me fallut beaucoup de chance pour que le responsable de l'enquête, qui n'avait pas quitté le pays, la retrouvât au fond d'un tiroir. Or la longueur des pas oscillait entre 40 cm et 1,20 m, ce qui, sur les superficies, se traduisait par des disparités allant de 1 à 9, qu'il eut été regrettable d'ignorer.

Autre fait démonstratif : la simple localisation des villages d'enquête des six études régionales me demanda une année de correspondance et d'efforts. Il fallut retrouver (à Abidjan, en brousse, en France) les six responsables et les forcer à se souvenir... Peu d'entre eux possédaient une carte de repérage. Beaucoup d'incertitudes demeurèrent.

Interprétation

Passage d'une information ponctuelle insuffisante à une cartographie par plages

Moins l'information ponctuelle est bonne dans sa répartition, plus on se trouve dans l'obligation d'aboutir à une représentation cartographique par plages. Cependant, à mesure qu'il devient plus nécessaire, ce passage s'avère aussi plus problématique.

Recherche d'un découpage

Quand le semis des points de sondage est très serré, il n'est pas impossible de fonder le découpage sur la seule observation des résultats, mais dans le cas plus général d'un semis déficient, on ne peut que faire appel à des données extérieures à l'enquête. Dans beaucoup de domaines, la solution peut être trouvée dans la recherche de corrélations entre les valeurs à traiter et les découpages connus du milieu physique et humain, qu'ils soient réels (frontières ethniques, limites des formations végétales ou géologiques) ou fictifs (lignes d'égale distance à un pôle d'attraction ou à une voie de communication).

Mise à profit de l'expérience des cadres locaux

Elle peut constituer un dernier recours lorsque l'information chiffrée est insuffisante. Ce sont les cadres de l'échelon inférieur qui sont les plus précieux, à condition qu'ils couvrent le pays d'un réseau serré (par exemple, techniciens et moniteurs d'agriculture, contrôleurs des marchés, maîtres d'école). En ce cas, l'absence de rigueur des témoignages individuels et leur caractère intuitif peuvent être palliés par les multiples recoupements qu'ils occasionnent. Une condition semble cependant indispensable : que la collecte de ces témoignages soit postérieure à l'analyse des résultats chiffrés et qu'elle s'appuie sur une pré-maquette de carte, sur une carte-hypothèse.

En somme, il n'est pas impossible de parvenir dans certains cas à une cartographie correcte. Encore faut-il pouvoir disposer d'éléments extérieurs à l'enquête qui n'existent pas obligatoirement, ou créer une information complémentaire dont le coût peut être prohibitif. Mais de toute façon les difficultés à résoudre sont telles que l'on est quelquefois amené à se demander si l'enjeu en vaut la peine. Pourtant l'expérience montre que lorsque la sortie d'une carte est annoncée, elle est toujours attendue avec beaucoup d'impatience par tout le monde. D'où l'intérêt qu'il y aurait à ce que le souci de cartographier les résultats les plus importants figure désormais dans les programmes d'enquête.

Manuscrit reçu au SCD le 24 février 1972

DISCUSSION

Plusieurs interventions revinrent sur le cas précis des enquêtes régionales ivoiriennes pour jeter divers éclairages sur l'antagonisme qui oppose, au moins en apparence, les objectifs des statisticiens et le souhaitable report de leurs résultats sur une carte. M. TISSANDIER fit remarquer de son côté que l'exploitation des enquêtes statistiques agricoles du Cameroun avait posé les mêmes problèmes.

Tout en estimant qu'il était difficile d'imposer notre loi aux statisticiens, M. SAUTTER exprima l'avis qu'il était possible de composer et qu'il fallait donc s'efforcer de participer à la conception de telles enquêtes et à l'élaboration de leur plan de sondage. Mais la mise en route de l'Atlas de Côte d'Ivoire fut postérieure à l'achèvement des enquêtes régionales. M. SAUTTER

fit valoir que même dans ce cas beaucoup de choses restaient possibles à travers une réexploitation des fiches d'enquête et une « re-stratification » du sondage, comme l'expérience de M. HAERINGER le prouve.

M. BERTIN relança ensuite le débat sur le recours au hasard dans une enquête, pour suggérer qu'un sondage aléatoire ne se trouve justifié que lorsqu'on ne peut procéder à une autre approche logique : « Le hasard constitue une limite ; une préconnaissance est supérieure au hasard ». Et M. GOUROU termina sur cette phrase : « Le géographe procède du connu vers l'inconnu, le statisticien dit qu'il ne sait rien ».

J.P. T. et Ph. H.