

Recensements, bases de données locales et systèmes d'information géographique : une chance pour l'information démographique et sociale

Xavier Charoy¹

Un double constat est à l'origine de cet article : a) les recensements, tels qu'actuellement réalisés en Afrique, sont trop lourds et trop coûteux, b) les données communautaires, bien que collectées lors des recensements, ne sont pas suffisamment exploitées, voire pas du tout. Après avoir rapidement décrit le contexte actuel des statistiques socio-démographiques, l'article se concentre dans un premier temps sur le recensement de la population, pôle du domaine. Puis il propose dans ses grandes lignes une refonte du système d'information autour d'un recensement redimensionné et d'une base de données locales. Il détaille certaines caractéristiques et contraintes de la base de données locales avant de passer en revue ses utilisations techniques et institutionnelles possibles. Quelques remarques complémentaires mènent à une conclusion tournée vers la mise en œuvre des opérations évoquées.

Introduction

Le présent article est né de plusieurs idées, constations ou expériences, qui étaient au départ indépendantes les unes des autres, dont aucune peut-être n'est vraiment originale, mais dont la juxtaposition, la confrontation et la critique ont conduit l'auteur à envisager une approche renouvelée du domaine des statistiques démographiques et sociales dans les pays d'Afrique. On verra que, comme il est normal s'agissant de ce domaine, la réflexion tourne beaucoup autour du recensement de population.

Les trois idées fondamentales autour desquelles se développera cet article sont les suivantes :

- les recensements de population, tels qu'ils sont généralement conçus aujourd'hui en Afrique, sont d'une lourdeur devenue insupportable pour les pays comme pour les bailleurs, et ils ne

répondent que très imparfaitement à ce qu'on en attend ;

- les données dites « communautaires » sont encore très insuffisamment exploitées, valorisées et diffusées, alors même que la demande en la matière s'accroît considérablement ;
- le besoin d'une coordination de l'ensemble souvent disparate d'opérations statistiques couvrant le domaine socio-démographique et d'une articulation autour d'un pôle commun s'affirme un peu plus chaque jour.

Ces trois idées peuvent être développées indépendamment les unes des autres, mais en les rapprochant ici, on tentera une approche systémique du domaine socio-démographique, cherchant au sein du système statistique national un équilibre et une complémentarité entre les sous-systèmes dans les meilleures conditions globales de coût et

¹ Administrateur de l'INSEE à la retraite

d'efficacité. On se situera ainsi dans une démarche de type SNDS (stratégie nationale de développement statistique)¹, à la différence près, toutefois, qu'on s'intéressera à des opérations statistiques et à des outils plus qu'à des textes ou à des schémas organisationnels.

Les propositions qui seront faites ici constituent des pistes de réflexion qui devront être complétées et approfondies par une réflexion et par des expérimentations propres à chaque pays. Le texte qui suit ne vise en aucune façon à fournir un outil « clés en main » et il laisse de nombreux points en suspens, à l'appréciation et à l'étude des acteurs nationaux.

Le contexte actuel

L'omniprésence du problème de la pauvreté

Ce n'est pas faire preuve d'originalité que de dire que le maître mot actuel concernant les pays en voie de développement, d'Afrique sub-saharienne notamment, est la réduction de la pauvreté. Mais il n'est pas inutile de le redire ici tant la conséquence en est forte sur les besoins en informations des acteurs impliqués dans cette réduction de la pauvreté et donc sur l'organisation des producteurs d'information (les systèmes statistiques nationaux) et les outils dont ils doivent se doter. Il convient cependant de rappeler que, depuis longtemps, une des composantes – et non des moindres – de la pauvreté, la faim, avait été prise en compte, y compris sur le plan statistique, par certaines institutions au premier rang desquelles se trouvaient la FAO et l'Union européenne. Du côté de cette dernière, on se rappellera le projet « Sécurité alimentaire » (années 1980 et 90) et son importante composante statistique.

Dans le contexte de la crise générale économique et sociale à laquelle font face les pays africains, l'attente des utilisateurs vis-à-vis de la statistique a en effet beaucoup évolué. Il existe actuellement, dans le cadre du suivi des DSRP et des OMD², une demande concrète, immédiate d'informations sur l'état, le cadre de vie et les conditions de vie de la population³, c'est-à-dire pratiquement sur

¹ La démarche SNDS a été notamment évoquée par Antoine Simonpietri dans son interview publiée dans le n° 100 de STATECO.

² DSRP : Document Stratégique pour la Réduction de la Pauvreté. OMD : Objectifs du Millénaire pour le Développement

³ Il faut noter ici que la demande en ces matières porte sur des informations qui ne sont pas seulement de nature statistique, mais peuvent être très ponctuelles ou qualitatives, en tout cas peu susceptibles des traitements

l'ensemble du domaine démographique et social.

Toutefois, même si cette demande est globalement en expansion et si elle s'est au fil des années modifiée, précisée, affinée, on constate encore que celle qui provient des pays (gouvernements, autorités locales, université, centre de recherche, secteur privé, ONG) est souvent mal exprimée et celle qui provient de l'extérieur (essentiellement des organisations internationales et des partenaires au développement) est profuse et souvent peu cohérente et qu'elle exerce une pression quelque peu désordonnée sur des systèmes statistiques nationaux encore fragiles. En tout cas, claire ou non, cohérente ou non, la demande, telle qu'elle s'exprime aujourd'hui dans le contexte qu'on vient d'esquisser, est de plus en plus forte et pressante et attend une réponse meilleure et plus rapide aux besoins.

Mais surtout – et c'est là un point essentiel de notre propos – beaucoup d'utilisateurs exigent maintenant plus d'informations simples mais **finement localisées**, notamment pour dresser des **cartes de pauvreté** afin d'identifier les « poches » où elle se situe (demande émanant en partie des administrations régionales et des ONG, mais en général exprimée de façon assez floue).

Les opérations mises au point au cours des décennies précédentes (entre 1960 et 1985), généralement quantitatives, lourdes et destinées plus à décrire une situation qu'à fonder une action, ne correspondaient plus que partiellement à cette demande. Si, par la suite, les exigences de qualité et de rapidité ont été abondamment commentées⁴ et si elles ont conduit à de multiples réalisations associant les INS et les partenaires au développement, l'exigence de localisation fine est, semble-t-il, restée largement sans conséquences. Elle est une des raisons et un des fondements du présent article.

Comment les systèmes statistiques réagissent-ils ?

Force est de reconnaître, et là encore ce n'est pas faire preuve d'originalité, que les systèmes statistiques nationaux dans la plupart des pays d'Afrique subsaharienne sont en état de faiblesse. La situation de certains d'entre eux est même

statistiques habituels, notamment pour ce qui touche au cadre de vie.

⁴ Ces commentaires, généralement très critiques, voire polémiques, systématiques et outrepassés, ont volontairement occulté l'apport méthodologique et pratique considérable des enquêtes de la première génération qui, malgré leurs défauts, ont ouvert la voie aux générations suivantes.

désastreuse. On citera, probablement sans être exhaustif, les raisons suivantes :

- les ressources humaines, quoique bien formées, sont mal utilisées et mal rétribuées, donc peu motivées à rester dans la statistique publique ;
- les institutions statistiques sont faibles, souffrent d'une organisation désuète et manquent de ressources budgétaires régulières et sûres ; elles peinent à organiser un « système statistique » cohérent et à fédérer les ressources extérieures ;
- les responsabilités techniques sont diffuses sans qu'il existe de coordination forte et efficace (ceci est surtout valable pour les pays francophones d'Afrique occidentale et centrale qui ont opté pour des systèmes statistiques décentralisés où les statistiques sectorielles sont élaborées par les ministères techniques compétents) ;
- les partenaires au développement et autres institutions de coopération sont à la fois inquiets de l'insuffisance des données disponibles et réticents à s'engager dans des opérations lourdes et coûteuses, requérant une bonne coordination entre eux ; la difficulté croissante qu'éprouvent les pays à boucler le budget de leur recensement de population, avec des montages compliqués et hasardeux, en est l'illustration.

Pour être un peu brutal et sans nuance, ce constat résume cependant une situation moyenne qui rend compte de la difficulté générale qu'éprouvent les systèmes statistiques nationaux à répondre correctement à la demande. La communauté internationale a joué en cette matière un rôle paradoxal : d'une part, en imposant des mesures drastiques et brutales d'ajustement structurel, elle a privé les instituts de statistique de la faculté de recruter de nouveaux agents et même de seulement remplacer ceux qui partaient ; d'autre part, elle a conduit – et aidé financièrement – les systèmes statistiques à entreprendre des opérations plus directement liées à la demande que par le passé (et, lorsque le financement suivait, certains INS ont correctement assumé leur responsabilité technique dans le déroulement de ces opérations) ; mais la faiblesse institutionnelle demeure.

Comment sortir de cette situation, relevée depuis plus de vingt ans par tous ceux qui s'intéressent à la statistique africaine ? Comme pour tout problème complexe, il n'existe pas une réponse unique. Plusieurs ont été proposées et mises en œuvre, mais, malgré quelques succès ponctuels ou temporaires, les progrès sont souvent contestables, faute notamment d'approches globales et

cohérentes. En bref, on peut envisager deux types d'approche. L'un est de caractère plutôt **institutionnel** ; c'est celui qui s'exprime à travers les Stratégies Nationales pour le Développement Statistique (SNDS) développées dans un nombre croissant de pays sous l'égide de PARIS21. L'autre est de caractère plus **opérationnel** et **instrumental** ; il passe par la mise en œuvre de grandes opérations statistiques, telles que les recensement de population, ou encore par la confection des comptes nationaux dont l'élaboration requiert une organisation solide et une bonne coordination (cette dernière opération concernant beaucoup plus le domaine économique que le domaine démographique et social). On proposera dans cet article une approche instrumentale, touchant avant tout le domaine démographique et social mais ayant des implications plus larges. On verra en particulier que cette approche, par sa globalité et sa recherche de cohérence, peut avoir des retombées considérables sur l'aspect institutionnel.

Les trois piliers des statistiques démographiques et sociales

Il est bien connu que l'information statistique en matières démographique et sociale est basée sur trois grandes catégories d'opérations : les recensements de population, les enquêtes par sondage et les sous-produits d'opérations administratives. Bien que cette distinction soit des plus classiques, on décrira ici sommairement ces trois catégories dans le seul but d'en souligner les principales caractéristiques qui seront utiles aux développements des parties suivantes de cet article.

Les recensements de population

S'il est une opération statistique fondamentale, c'est bien, au moins dans l'esprit des statisticiens et des démographes, le recensement de population. À leurs yeux experts, il constitue le pilier incontournable de la statistique démographique et sociale. La plupart des pays du monde⁵ se sont démenés, avec des succès divers, mais plutôt bons dans l'ensemble, pour réaliser – en principe tous les 10 ans, si on suit les recommandations des Nations-Unies – un recensement de population. Pour les pays en développement, ceux d'Afrique subsaharienne en particulier, la conduite plus ou moins régulière de recensements a tenu pour beaucoup aux efforts de l'UNFP⁶ qui les a largement financés dans un premier temps, avant de devoir partager le

⁵ À l'exception notable de ceux d'Europe du Nord qui fondent leurs statistiques démographiques sur des registres permanents de population tenus par les administrations et non sur des recensements statistiques.

⁶ Plus connu sous ses anciens sigles : FNUAP en français, UNFPA en anglais.

fardeau avec d'autres bailleurs de fonds, et qui continue de contribuer fortement à leur conception et à leur mise en œuvre technique. Même si le rythme décennal n'a pas toujours pu être strictement respecté (et parfois loin de là), nombreux sont les pays d'Afrique qui possèdent maintenant des données de base sur leur population et ont acquis l'expérience nécessaire pour les collecter.

En fait, le rôle central des recensements de population dans les systèmes nationaux de statistique, proclamé par les statisticiens et les démographes, est plus réel sur le papier que dans la pratique. Peut-être en partie à cause des difficultés de financement, mais aussi d'une conception quelque peu figée des recensements par les statisticiens et démographes et un certain manque d'imagination de leur part. On reviendra sur ces deux points dans la suite de cet article.

Les enquêtes par sondage

A partir des années 60, beaucoup de jeunes pays africains qui venaient d'accéder à l'indépendance ont réalisé des enquêtes – le plus souvent assez lourdes – portant sur divers aspects de la vie des populations : démographie, migrations, conditions de vie des ménages, budgets familiaux et consommation, emploi, santé, logement, agriculture, etc. Aujourd'hui, on note une diminution du nombre d'enquêtes lourdes comportant des mesures (les enquêtes démographie-santé constituant la principale exception) et, au contraire, un fort accroissement de la quantité et de la variété d'enquêtes assez légères et plus qualitatives.

Toutes ces enquêtes étaient réalisées en utilisant la méthode des sondages et elles le sont toujours. Au début, soit schématiquement dans les années 60 et le début des années 70, c'étaient les recensements administratifs hérités de l'époque coloniale qui servaient de bases de sondage. Ceux-ci ayant été peu à peu abandonnés, des recensements démographiques ont été menés et ont pris la relève des recensements administratifs dans ce rôle technique. Cette utilisation des recensements est souvent mise en avant pour en justifier la réalisation ; sans eux, donc sans la base de sondage qu'ils seraient seuls à même de constituer, il ne pourrait y avoir de bonnes enquêtes socio-démographiques.

Toutefois, ce principe connaît de plus en plus d'exceptions, d'une part parce que les recensements sont souvent trop anciens pour constituer une base de sondage valable – et ce d'autant plus que les sociétés évoluent plus vite maintenant que naguère –, d'autre part parce que beaucoup d'enquêtes

requièrent des échantillons plus « pointus » que ceux que peuvent fournir les recensements.

Les statistiques d'origine administrative

Dans certains domaines, les statistiques de base sont essentiellement d'origine administrative, c'est-à-dire qu'elles sont établies à partir de documents, rapports, registres dont la finalité est la gestion administrative (et non la statistique). C'est en particulier le cas de ceux de l'Éducation et de la Santé où des rapports établis dans les unités de base (écoles, centres de soins) sont compilés et donnent lieu à un rapport national. Le problème est que les rapports en question ne sont pas toujours établis de façon exhaustive (nombre d'établissements ne répondent pas et le secteur privé est mal couvert voire pas du tout) et que le processus de transmission des rapports entre les différents échelons administratifs et de leur contrôle est mal maîtrisé.

D'une façon générale, en dehors des ministères de l'Éducation et de la Santé qui viennent d'être cités, la plupart des administrations détiennent des informations qui, pour n'être pas toujours de nature proprement statistique, peuvent utilement compléter un système d'information. Malheureusement, la mobilisation de ces informations est souvent très difficile dans la mesure où elles sont souvent éparées, non centralisées, partielles ou non cohérentes.

Il faut enfin souligner la totale indépendance de ces outils par rapport aux deux précédents : méthodes et circuits de collecte sont différents ; dans la plupart des pays francophones, les administrations responsables sont distinctes et les relations entre elles peu développées et très ponctuelles.

Commentaires sur les recensements de population

Il est normal qu'on concentre sur les recensements l'essentiel des commentaires et des critiques, d'une part à cause du rôle central qu'on lui attribue dans le domaine des statistiques socio-démographiques, d'autre part en raison de son contenu assez large, enfin en raison de son coût et de la charge considérable qu'il représente.

Le contenu et l'exploitation des recensements

Les recensements dans les pays africains, tels qu'ils ont été conçus au fil du temps, contiennent quatre

grandes catégories d'informations relatives aux individus et aux ménages :

- informations d'état civil : sexe, âge, statut matrimonial ;
- informations de caractère « socio-économique » : niveau d'instruction et diplômes, occupation(s) ;
- informations sur les ménages : relation entre membres, nombre d'enfants ;
- informations sur les mouvements de population : naissances, décès, nuptialité.

S'y ajoutent généralement deux catégories d'informations concernant leur cadre de vie :

- les statistiques sur les logements – caractéristiques de construction, taille, équipement –, d'où l'appellation courante : Recensement Général de la Population et de l'Habitat (RGPH) ;
- les statistiques dites « communautaires », c'est à dire relatives aux lieux de vie (villages, hameaux, campements, etc.), – infrastructures présentes, accès à certains biens ou services tels que : eau, écoles, établissements sanitaires – souvent collectées lors de la phase « cartographie » des recensements à travers un « questionnaire village »⁷.

On voit que, collectant des informations assez hétérogènes par leur nature, même si elles concernent toutes la population (individus et ménages), les recensements tentent de répondre à plusieurs préoccupations à la fois. Si on considère ces informations des trois points de vue qui suivent :

- celui du degré de détail géographique avec lequel elles sont nécessaires,
- celui de leur actualité souhaitable,
- celui de la difficulté et du coût de leur collecte (une information avec un certain détail vaut-elle le coût ?),

on constate, d'une façon générale, que :

- les données nécessaires avec un grand détail géographique sont des données simples et faciles à collecter, c'est-à-dire plutôt factuelles, ne faisant pas ou que peu appel à la mémoire des enquêtés et ne mettant pas en jeu des nomenclatures complexes ;
- compte tenu de l'ampleur relative des mouvements et des changements possibles

dans les petites unités géographiques, il n'est raisonnable de publier et d'utiliser pour elles que ces données simples ;

- les résultats portant sur des variables complexes n'ont en général aucun sens pour de petites unités géographiques (a-t-on jamais calculé un taux de natalité par village ?) ;
- même les données utilisables pour de petites unités géographiques ne sont que rarement, en tout cas que partiellement, exploitées et diffusées pour ces unités (malgré de récents progrès dans ce sens) ; c'est le cas en particulier des données communautaires.

En première analyse – et de façon assez intuitive –, **un recensement ne devrait donc comporter que des données utiles et utilisables à des échelons géographiques fins** (dénombrement des logements, des ménages et des individus, état civil, niveau d'instruction en 2 ou 3 positions, activité en 2 ou 3 positions, 2 ou 3 caractéristiques des logements). **Les données de caractère socio-économique** (niveau d'instruction, activité, emploi) détaillées, **les données de mouvement démographique, la plupart des informations sur les logements**, inutiles à des niveaux fins, de toute façon inutilisables compte tenu des délais d'exploitation (elles sont caduques avant même d'être publiées), **devraient faire l'objet d'enquêtes par sondage**, telles que les enquêtes démographie-santé (EDS) ou les enquêtes conditions de vie de toutes sortes, mieux adaptées, plus riches de contenu, plus rapides à exploiter, de qualité au moins égale et plus souvent supérieure⁸ et moins coûteuses.

Un coût et une charge de travail considérables

Les recensements exigent la mobilisation d'un personnel très nombreux, qu'il s'agisse des enquêteurs et contrôleurs de terrain ou des agents de codification et de saisie, en général recrutés et formés pour l'occasion, ou de l'encadrement supérieur, prélevé temporairement sur les services statistiques du pays. Au coût que représente ce personnel (salaires et frais de déplacement) s'ajoute celui de l'équipement nécessaire (véhicules, matériel informatique, etc.). Les financements nécessaires pour faire face à ces coûts sont de plus en plus difficiles à trouver. Les pays sont maintenant contraints de faire appel à plusieurs partenaires-bailleurs dont il est difficile de coordonner les interventions (techniques et financières). Nombreux sont les pays qui, ayant pu

⁷ Les zones urbaines, recensées au même titre que les zones rurales, le sont sur la base de plans de villes, actualisés lors de la phase de cartographie si nécessaire. Il n'existe en général pas d'équivalent urbain du questionnaire village. D'ailleurs, dans la plupart des cas, les questions posées dans le questionnaire village ne sont pas pertinentes en zone urbaine, du moins pas sous la forme qu'elles revêtent en zone rurale.

⁸ On sait en effet que la qualité des recensements est entachée d'erreurs d'observation considérables allant bien au-delà des erreurs aléatoires dont sont affectées les enquêtes par sondage.

financer et réaliser la collecte, ont dû attendre longtemps avant de trouver les fonds nécessaires à l'exploitation.

L'organisation d'un recensement de population est en outre très lourde et contraignante, principalement pour la phase de collecte, mais aussi pour l'exploitation (codification et saisie).

Évolutions méthodologiques

On assiste aujourd'hui à travers le monde à une vaste réflexion sur les recensements de population et à la mise en place de méthodes plus ou moins complètement renouvelées. Leur coût, leur espacement dans le temps, les difficultés vis-à-vis de la population et d'autres considérations ont amené certains pays parmi les plus développés à repenser complètement leur approche des statistiques de population. Les pays du Nord de l'Europe travaillent depuis longtemps sur la base de fichiers de population, la France a adopté récemment un système de recensement tournant en partie par sondage⁹ et il existe d'autres pistes qui sont mises en œuvre ou explorées ici ou là.

En Afrique subsaharienne, la réflexion méthodologique est le plus souvent en panne. D'une façon générale, la conception et la réalisation des recensements n'ont guère évolué au cours des trois ou quatre dernières décennies, c'est-à-dire depuis qu'on les y a introduits et généralisés. Certes, des concepts internationaux nouveaux ont été pris en compte, mais cela n'a modifié qu'à la marge les contenus et les méthodes censitaires. Le volume des questionnaires de recensement paraît même avoir tendance à s'accroître.

Assez récemment, plusieurs documents ont été préparés dans le cadre d'un groupe de travail UNFP/PARIS21 sur les recensements¹⁰. La plupart des solutions avancées pour améliorer la situation apparaissent comme ponctuelles ou marginales et dans un seul de ces documents on ose, bien que d'une façon qu'on ne suivra que partiellement ici, remettre en cause les principes mêmes des recensements tels qu'ils se pratiquent en Afrique.

On mentionnera ici pour mémoire les dépouillements rapides de recensements à partir d'un échantillon. Plusieurs pays africains ont expérimenté depuis les années 80 ce moyen rustique, techniquement simple mais efficace d'obtenir des résultats provisoires dans un délai réduit. Mais la méthodologie générale des recensements n'était pas remise en cause pour

autant. D'autres pays s'y sont refusés, pour ne pas décourager les bailleurs potentiels de financer une exploitation exhaustive sous le prétexte que les résultats provisoires seraient bien suffisants. Ce en quoi lesdits bailleurs n'avaient pas totalement tort, s'agissant de variables complexes tout au moins.

Dans ce contexte peu évolutif, de timides voix, sensibles au coût des recensements et au délai nécessaire à leur exploitation, se sont tout de même parfois élevées pour suggérer qu'on les allège et qu'on aille plutôt vers la voie d'un **dénombrement** (comptage légèrement enrichi), comme il a été sommairement décrit ci-dessus, complété par une ou plusieurs **enquêtes socio-démographiques** par sondage. Vite étouffées par les protestations des statisticiens et démographes qui ne voulaient pas courir le risque qu'une telle proposition conduise à l'éventualité que seul le dénombrement soit financé et non les enquêtes complémentaires, ces voix n'ont eu jusqu'à présent pratiquement aucun écho et le principe d'un « recensement à tout faire » s'est maintenu. Certains pays ont préféré reculer leur recensement, allant dans quelques cas jusqu'à ne pas le faire, ou en retarder l'exploitation, plutôt que de l'alléger.

Des outils nouveaux, des réalisations et des expériences ponctuelles

Sur le plan **technique**, en revanche, les recensements africains ont largement tiré profit, au cours de la dernière décennie en particulier, des évolutions considérables qui se sont produites. Des outils nouveaux ont été développés et largement diffusés qui permettent de réaliser facilement aujourd'hui ce qui était encore inconcevable hier. On se référera ici à trois catégories d'outils :

- des **outils statistiques**, notamment ceux qui concernent l'étude des « petits domaines »¹¹ (petites zones géographiques, petits groupes sociaux, etc.) ;
- des **outils informatiques**, et on citera surtout parmi eux les systèmes de gestion de bases de données (SGBD), les outils de lecture optique, les outils de diffusion (CDRoms, DVD, Internet) et les systèmes d'information géographiques (SIG) qui sont tous venus révolutionner le monde de l'information (et pas seulement celui de l'information statistique) ;

⁹ Pour en savoir plus, on pourra consulter le site de l'INSEE :

http://www.insee.fr/fr/recensement/page_accueil_rp.htm

¹⁰ Voir bibliographie

¹¹ On parle aussi de « petits ensembles » ou de « petites aires », mais on retiendra ici le termes de « petits domaines ». Ce sujet, très prometteur mais certainement plus difficile qu'il n'y paraît, a fait l'objet de plusieurs communications lors des derniers congrès de l'Institut international de Statistique (IIS).

- des **outils de positionnement** (GPS), qui permettent d'obtenir les coordonnées géodésiques de n'importe quel point du sol.

Ces outils ont eu une influence à trois stades de la réalisation des recensements :

- au stade de la **préparation**, l'utilisation de logiciels cartographiques¹² a permis, au moins dans certains pays, des progrès spectaculaires qui n'ont pu qu'avoir des conséquences positives pour l'organisation de la collecte ;
- au stade de l'**exploitation**, les délais de disponibilité des résultats ont été généralement raccourcis ;
- au stade de la **diffusion**, le recours à des supports modernes (CDRom, Internet) a été étendu, tandis que la représentation cartographique des résultats se généralisait.

Il semble cependant que beaucoup moins de bénéfices en aient été tirés au stade de l'analyse des résultats.

Les outils statistiques mentionnés restent insuffisamment connus en Afrique. Leur utilisation pourrait pourtant permettre de valoriser considérablement les données provenant des enquêtes en donnant des estimations relatives à de petits domaines sur des variables que personne n'a jamais songé – à juste titre – à inclure dans un recensement. Il y a un gros travail d'information et de formation à faire à leur propos. En revanche, les outils informatiques sont relativement bien maîtrisés par les statisticiens africains¹³, dans la mesure où ils peuvent avoir accès aux logiciels, et, comme on l'a déjà souligné précédemment, ils sont déjà utilisés pour une partie d'entre eux ; ces logiciels sont d'ailleurs de moins en moins coûteux. Les GPS, quant à eux, sont d'un maniement extrêmement simple, leur prix devient abordable et ils sont assez bien utilisés.

Revenons un instant sur les systèmes d'information géographiques et leur utilisation. Celle-ci s'est beaucoup développée, tant pour l'organisation de la

collecte que pour la diffusion, comme on vient de le dire. Certains pays se sont dotés, à l'occasion d'un recensement de la population, d'un équipement sophistiqué et coûteux permettant de numériser des fonds de cartes et de disposer de tout ce qui est nécessaire pour la production de résultats sous forme cartographique. Ils ont joué là le rôle que devrait tenir le service géographique national. Dans la foulée, ils ont avancé, avec le concours de l'administration territoriale dont ce devrait être la responsabilité, dans la voie d'une fixation et d'un codage des unités de base (villages, etc.), activité nécessaire au géoréférencement des données et aux exploitations fines des opérations statistiques.

Il convient ici de distinguer deux types d'exploitations fines :

- l'exploitation des données relatives aux ménages et aux individus (domaine de la statistique démographique *stricto sensu*), qui se fait du niveau géographique national au niveau du district (sous-préfecture ou équivalent) et pratiquement jamais (sauf pour de simples comptages) à des niveaux inférieurs (village par exemple) ;
- l'exploitation, ou plutôt la simple présentation, des données relatives aux logements et surtout des données communautaires, qu'on peut considérer plus comme des données « d'inventaire »¹⁴ que comme des données statistiques et qui de toute façon ne sont pas de caractère démographique.

Si les premières font maintenant couramment l'objet de présentations cartographiques, ce n'est pas encore le cas des secondes. Néanmoins, dans le cadre des recensements de population ou dans d'autres cadres, plusieurs expériences intéressantes ont été menées dans certains pays de la région (Bénin, Mali, Sénégal et certainement d'autres). Elles ont fourni une preuve concrète et tangible de l'intérêt de données localisées et font espérer que des avancées sensibles sont possibles dans ce domaine. Celle qui a inspiré en partie cet article est l'opération dite « pré-recensement » qui a été réalisée en 1995-96 dans le cadre du recensement de l'agriculture du Togo ; on en trouvera une brève description dans l'encadré n°1 (page suivante).

¹² Il en existe beaucoup. Parmi les plus connus, citons Mapinfo, Arcview, Geoconcept.

¹³ On rappellera ici le stage organisé par l'INSEE en juillet 2003 à Libourne sur la cartographie et les systèmes d'information géographique.

¹⁴ Un inventaire est une simple liste ou des unités sont repérées et localisées, sans caractéristiques descriptives.

Encadré n° 1**Le pré-recensement du Togo (1995)**

Au début des années 90, la République du Togo a souhaité entreprendre son troisième recensement de l'agriculture. Comme ses homonymes en Afrique sub-saharienne, cette opération devait être en réalité une enquête par sondage portant sur un gros échantillon (7200 ménages/exploitations). Le premier problème qui apparut alors était l'absence d'une base de sondage valable : le dernier recensement général de la population remontait à 1981 et le Togo avait connu depuis d'importants mouvements de population qui le rendaient caduc.

Une approche originale

Lorsque, en 1995, le financement de l'opération par l'Union européenne a pu être mis en place, la FAO, qui lui apportait son assistance technique, a proposé une solution originale pour constituer une base de sondage à jour : la réalisation d'un « pré-recensement », considéré comme première phase du recensement de l'agriculture.

1. inventaire exhaustif des localités rurales (villages, hameaux, etc.) ;
2. localisation des localités sur des cartes sommaires par canton ;
3. remplissage pour chaque localité d'un « questionnaire village » portant sur les infrastructures présentes et sur l'accès des habitants à certains services ;
4. inventaire des concessions et des ménages dans les concessions ;
5. remplissage pour chaque ménage d'un bref questionnaire portant sur son effectif et son activité dans l'agriculture, l'élevage ou la pêche.

Le pré-recensement comportait donc les données nécessaires à la constitution d'un échantillon stratifié pour le recensement de l'agriculture. Mais il était beaucoup plus riche qu'une simple base de sondage ; cette richesse a été exploitée et a permis la production de nombreux tableaux statistiques : 12 sur les villages et 25 sur les ménages.

Compte tenu de sa nature d'ordre assez général, le pré-recensement a été réalisé conjointement par la Direction des Enquêtes et Statistiques Agricoles du Ministère de l'agriculture – qui, avec le recensement de l'agriculture, en était le premier utilisateur – et la Direction de la Statistique du Ministère du Plan – qui y a bien vu un moyen d'obtenir des données sur la population en l'absence d'un recensement qui ne trouvait pas de financement.

Les limites

1. le champ du pré-recensement était restreint aux zones rurales (ce qui était naturel, s'agissant d'un recensement de l'agriculture) ;
2. la qualité du dénombrement des membres du ménage n'a pas été excellente, quoique suffisante pour constituer une base de sondage valable. Une sous-estimation systématique des effectifs a été notée, ce qui ne doit pas surprendre au vu de la façon dont le sujet était abordé : une seule question portant sur l'effectif de chaque ménage, sans dénombrement individuel ;
3. les données relatives aux villages n'ont pas fait l'objet d'un traitement utilisant les systèmes d'information géographiques (SIG) ; ces puissants outils n'étaient pas aussi diffusés à l'époque qu'ils le sont maintenant.

Les rubriques du questionnaire village togolais

- Localisation
- Voies d'accès : type, fréquentation
- Infrastructures d'approvisionnement en eau et électricité
- Infrastructures scolaires et centres de formation (si absence d'établissement, distance aux plus proches)
- Infrastructures sanitaires (si absence d'établissement, distance aux plus proches)
- Infrastructures socio-économiques ; magasins, marché, matériel de transformation de produits agricoles, station d'essence, cultes, etc.
- Langues parlées
- Activités économiques : type, existence de groupements de producteurs ou d'exploitation moderne
- Production fruitière
- Inventaire des îlots, hameaux, fermes

Pour une approche intégrée des statistiques démographiques et sociales

A nouvelle donne, nouvelle approche

La nouvelle donne, telle qu'on l'a décrite dans la première partie, peut se résumer en trois points :

- le renouvellement profond de la demande, notamment dans le cadre des études sur la pauvreté et du suivi de la situation alimentaire, avec un accent mis sur les données locales ;
- la difficulté croissante de financer le recensement, premier pilier des statistiques démographiques et sociales ;
- l'apparition de nouveaux outils statistiques, informatiques ou autres qui ouvrent de nouveaux horizons.

On a dit aussi que les évolutions qui ont pu être notées dans le domaine des statistiques démographiques et sociales ont concerné le contenu des opérations (des enquêtes beaucoup plus que des recensements) et les techniques d'exploitation et de diffusion des données, mais pratiquement pas l'approche globale et systémique. Autrement dit, la nouvelle donne n'a pas été prise en compte sous tous ses aspects par les systèmes statistiques nationaux, notamment pour ce qui est des recensements.

La voie dans laquelle il est ici suggéré de s'engager est celle d'une **approche intégrée** du domaine des statistiques démographiques et sociales. Certes, l'idée n'est pas neuve. Elle est déjà, avec plus ou moins de succès selon les pays, au centre du domaine des statistiques économiques avec la comptabilité nationale comme pivot. Mais, dans le domaine socio-démographique, elle n'a pas induit de progrès fulgurants. Probablement est-elle restée trop abstraite, trop intellectuelle et s'est-elle située d'une façon trop exclusive dans une logique conceptuelle et institutionnelle. Certainement s'est-elle aussi heurtée à la dispersion et au manque de coordination des intervenants nationaux comme extérieurs. Elle manquait enfin d'un « outil d'intégration ».

Nous croyons en effet très fortement, sans pour autant entrer à corps perdu dans une logique d'outil, qu'aucun progrès d'ordre conceptuel et institutionnel ne peut être réalisé s'il ne s'appuie sur une batterie d'outils (cette remarque est d'autant plus valable pour les pays où la ressource humaine qualifiée est rare et où peu de temps et d'énergie peuvent être déployés à des réflexions théoriques). Ceux-ci ne doivent pas dicter la voie à suivre, mais permettre de donner vie à celle qui a été choisie. De

même que, dans le cas des statistiques économiques, l'intégration, quand elle existe, se fait autour du modèle macro-économique qu'est la comptabilité nationale, de même, dans le domaine socio-démographique, il faut associer un outil à la démarche d'intégration.

Plus de détails sur une approche intégrée

On peut définir un système intégré comme un ensemble d'opérations conçues en complémentarité les unes par rapport aux autres – complémentarité de contenu (traiter tous les sujets sans dupliquer les efforts) et complémentarité technique (concepts et définition, nomenclatures, base de sondage, analyse des résultats, etc.) – chacune de ces opérations trouvant sa place dans le système intégré selon sa propre logique et sa propre cohérence interne. En réalité, l'approche proposée ici pour les statistiques démographiques et sociales ne bouleverse nullement les grandes lignes de la situation actuelle. Comme on le verra plus loin, les trois piliers des statistiques démographiques et sociales, qui ont été sommairement décrits précédemment, demeurent, avec quelques modifications de contenu. Il s'agit de proposer des pistes de réflexion en vue d'une meilleure utilisation de chacun d'eux, d'une meilleure complémentarité entre eux et d'un enrichissement des uns par les autres grâce à l'introduction d'un outil d'intégration.

Examinons d'un peu plus près les modifications qu'impliquerait l'approche d'intégration dans les différentes opérations constituant le système de statistiques démographiques et sociales.

1. Le recensement deviendrait « **recensement général de la population, des habitations et des localités** (RGPHL) ». Sa méthodologie générale resterait très voisine de celle des recensements traditionnels, mais il se concentrerait exclusivement sur les caractéristiques de base de la population et des logements d'une part, la description des lieux de vie (qu'on appellera « localités » et qui comprennent les villages ruraux et les quartiers urbains) d'autre part. L'extension de son appellation, avec l'ajout du terme « localités », met l'accent sur cette dernière orientation. Le « questionnaire localité », généralisation de l'actuel « questionnaire village », deviendrait une pièce maîtresse du dispositif. Ceci signifie, par rapport aux recensements traditionnels, une simplification drastique des questions individuelles (notamment sur les questions de formation et de profession) et des questions relatives aux habitations (focalisation sur quelques caractéristiques simples en corrélation forte avec le niveau de vie ou

permettant l'étude de petits domaines) et la suppression des questions relatives aux mouvements (naissances, décès) observés dans les ménages.

Par rapport au pré-recensement du Togo (voir encadré n° 1), le RGPLH ici esquissé pourrait avoir un questionnaire localité un peu simplifié quant aux questions relatives à l'agriculture et un questionnaire ménages nettement développé puisqu'il y aurait un dénombrement individuel des individus, comme dans tout recensement qui se respecte.

Le recensement comporterait donc les étapes suivantes :

- identification et délimitation des localités (villages, hameaux, campements, quartiers urbains, etc.) ;
 - administration pour chaque localité d'un questionnaire descriptif sur les infrastructures, les équipements, l'accès à divers services (questionnaire localité) ;
 - identification des concessions et, à l'intérieur de celles-ci, des habitations, dans chaque localité ;
 - administration d'un questionnaire descriptif des habitations ;
 - identification des ménages (les notions de logement et de ménage coïncident le plus souvent) et, au sein de ceux-ci, des individus ;
 - administration d'un bref questionnaire portant, de façon très simplifiée, sur les membres du ménage, leur état civil, leur occupation¹⁵ et les liens familiaux entre eux.
2. La série classique, modulable en fonction des besoins et des moyens, d'**enquêtes par sondage** portant sur tous les aspects imaginables de la vie des ménages et des individus, reste le deuxième pilier des statistiques démographiques et sociales. Certaines de ces enquêtes sont couramment réalisées aujourd'hui par les instituts nationaux de statistique (enquêtes démographie-santé, enquêtes conditions de vie, etc.). La nouvelle approche n'y introduirait guère que des modifications marginales de contenu par rapport aux enquêtes actuellement menées, de façon qu'elles prennent en compte les questions laissées de côté dans le recensement général qui ont été mentionnées ci-dessus et qu'elles s'intègrent de façon cohérente et sans doubles emplois dans le système. Probablement faudra-t-il aussi augmenter la

taille des échantillons de ces enquêtes pour améliorer leur représentativité régionale.

Le point commun à toutes ces enquêtes, outre le fait qu'elles reposent toutes sur une approche par les ménages, serait qu'elles utiliseraient toutes la même base de sondage, l'outil d'intégration dont il a déjà été question. On reviendra plus loin sur cette question fondamentale pour tout système statistique.

Avec l'approche proposée, l'utilisation des techniques statistiques relatives à l'étude des « petits domaines » permettrait de déconcentrer les données d'enquêtes par sondage et ainsi de valoriser leurs résultats et ne justifierait plus le recours à l'exhaustif que pour des variables fondamentales.

3. Les systèmes de collecte par **voie administrative** dans des domaines tels que l'Éducation et la Santé sont maintenus et si possible améliorés. On verra plus loin comment ils peuvent, dans le cadre proposé ici, s'intégrer beaucoup mieux qu'actuellement dans un système plus vaste d'information socio-démographique.

Une base de données locales comme outil d'intégration

Sur la base du constat résumé ci-dessus et de la nécessité qui y a été soulignée de disposer d'un **outil d'intégration des statistiques démographiques et sociales**, on proposera la création d'une **base de données locales** qui jouera ce rôle. Sous l'appellation de « fichier de villages », ce concept a été développé dès la fin des années 70, sans connaître d'ailleurs beaucoup d'applications. On pourra se reporter à l'encadré n°2 (page suivante) pour quelques détails sur les fichiers de village.

La base de données locales serait constituée à partir des **données descriptives des localités** et des **résultats résumés des dénombrements** qui y ont été effectués lors des recensements (caractéristiques fondamentales des logements, nombre de ménages et d'individus avec un minimum de détails). Associée à un **système d'information géographique (SIG)**, qui en est un complément indispensable, elle permettrait de produire des **cartes thématiques** beaucoup plus accessibles pour les utilisateurs que les tableaux statistiques classiques. On pourra se reporter, à titre d'exemple, aux représentations cartographiques des inégalités et de la pauvreté faites dans divers pays d'Afrique.

La constitution, l'enrichissement et la tenue à jour de cette base de données, d'une part, son utilisation d'autre part font l'objet des deux parties suivantes de cet article.

¹⁵ Dans la plupart des pays d'Afrique, une attention particulière devra être apportée aux occupations agricoles.

Encadré n° 2**Le « questionnaire village » : une vieille histoire**

*Dans le monde entier, ou presque, lorsque les statisticiens et démographes ont à entreprendre un recensement de population, ils sont d'une façon ou d'une autre inéluctablement conduits à identifier de prime abord les **lieux d'habitation** (villages, hameaux ou campements en zone rurale, quartiers ou îlots en zone urbaine). C'est seulement une fois en possession de cette base géographique qu'ils pourront ensuite procéder au dénombrement et à la rapide description des ménages et des individus qui peuplent ces lieux d'habitation. Ils font un large usage, dans cette opération préliminaire, des cartes et des plans existants, mais sont souvent obligés de les compléter ou de les mettre à jour pour les besoins de leur cause.*

Dans beaucoup de pays africains, les cartes et plans sont succincts ou anciens, en tout cas tout à fait insuffisants pour l'objectif poursuivi, de sorte qu'une identification spécifique des lieux d'habitation, souvent assez lourde, doit être entreprise. C'est l'opération de « cartographie ». Et puisqu'ils leur fallait pour cela arpenter tout le pays, nos statisticiens et démographes en ont souvent profité – et ils ont eu cent fois raison – pour recueillir des informations sur ces lieux d'habitation. Plus précisément, ils se sont intéressés aux infrastructures et équipements dont ils disposent ; et si un lieu d'habitation ne dispose pas de telle infrastructure ou de tel équipement, ils ont cherché à connaître la distance que doivent parcourir leurs habitants pour les trouver. C'est ce qu'on appelle souvent des « statistiques communautaires ». Il s'agit plutôt là d'un « inventaire » que d'une « statistique » (c'est d'ailleurs pourquoi l'opération de ce type qui est réalisée en France s'appelle « Inventaire Communal »). C'était une initiative potentiellement très riche, mais malheureusement bien peu mise en valeur, tant ces informations paraissaient vulgaires à côté des nobles données socio-démographiques recueillies pendant le recensement proprement dit. Et puis, le temps et les crédits manquaient trop souvent pour aller au delà de la collecte. Il faut convenir qu'un problème de présentation se posait aussi pour des informations peu susceptibles de traitements et analyses statistiques.

Tout le monde ne se désintéressait pourtant pas des fichiers de village. On en veut pour preuve le très important travail qui a été mené vers la fin des années 70 par Didier Gasse et qui a donné lieu à une publication du groupe AMIRA (voir Bibliographie). Mais peut-être leur constitution était-elle trop conçue comme une opération en soi, plus ou moins liée au recensement de la population, mais trop indépendante du « système statistique » dans son ensemble, ce qui n'a fait qu'en rendre le financement plus improbable. En outre, il n'existait pas à l'époque d'outils tels que les systèmes de gestion de bases de données ou les systèmes d'information géographique pour en tirer un bon parti.

La base de données locales

Il n'est pas question d'entrer ici dans les détails de la mise au point d'une base de données locales ni d'aborder précisément les questions de concepts, de définitions ou de nature des variables à intégrer qui y sont liées. On pourra se reporter au document de référence mentionné dans l'encadré n°2, sans oublier toutefois que, compte tenu de son ancienneté, beaucoup d'aspects doivent être repensés, notamment pour tenir compte des nombreux et capitaux développements intervenus dans les technologies de l'information depuis sa rédaction. Les remarques qui suivent n'ont donc pas pour objet de réactualiser le document en question – ce qu'il faudra faire le jour venu – mais seulement de signaler ou de rappeler quelques points qui devront être pris en considération lorsqu'on décidera de passer à l'acte.

La conception

Au contraire de la comptabilité nationale, qu'on a mentionnée comme outil de synthèse dans le domaine économique, une base de données, qui

peut aussi être considérée comme un outil de synthèse, n'est pas un modèle obéissant à des règles strictes et complexes. Sa conception est très libre et très ouverte et chaque pays peut organiser la sienne comme il le veut (en tenant compte des besoins d'information) et comme il le peut (en tenant compte pragmatiquement des données existantes) ; c'est donc une affaire de connaissance générale des besoins et des outils, de disponibilité des données, de moyens et ... de bon sens.

La conception d'une telle base de données doit être évolutive, ce qui permettra de démarrer avec les éléments existants ou facile à obtenir, puis d'étendre progressivement le champ à mesure que de nouvelles sources de données sont rendues disponibles, de nouveaux outils mis au point ou de nouvelles organisations mises sur pied.

La réalisation initiale

Pour la réalisation pratique, il faut distinguer entre la phase – qui peut être lourde, on ne doit pas se le cacher – de numérisation de fonds de cartes (frontières, limites de circonscriptions adminis-

tratives, etc.) et la phase d'introduction des données initiales. La complexité de ce travail dépendra des unités locales retenues (on en reparlera plus loin sous le titre « Définition des localités »), du nombre et de la nature des données retenues pour figurer dans la base ainsi que du logiciel choisi. Plus encore que la conception, la réalisation doit mettre en jeu toutes les compétences possibles, à l'intérieur comme à l'extérieur du système statistique, et on pense ici en particulier aux compétences cartographiques ; il se peut en effet que l'institution responsable des travaux cartographiques dans le pays ait déjà procédé à la numérisation des fonds de cartes ou que cela ait été fait dans le cadre d'un autre projet.

Dans les pays d'Afrique subsaharienne, où un questionnaire localité est généralement administré de façon uniforme à travers un pays donné lors de la phase préparatoire du recensement de population, ce questionnaire et certains résultats du recensement lui-même semblent bien être les sources les plus évidentes à partir de laquelle constituer la base. Le contenu du questionnaire localité est à la fois simple, diversifié et cohérent. Pour les recensements futurs, il devra être conçu en fonction des besoins de la base. On peut naturellement imaginer une réalisation de la base indépendante du recensement, mais elle risque de devenir vite complexe, car elle nécessite le recours à des sources nombreuses, dispersées, pas nécessairement homogènes.

Le contrôle de qualité

La plupart des données de la base initiale, en particulier celles qui concernent les équipements et les infrastructures, sont recueillies par les enquêteurs auprès de notables des villages (chefs de village, anciens, maîtres d'école, autres personnes d'influence), et peuvent naturellement comporter des erreurs. Un contrôle de ces données pourra le plus souvent être fait par recoupement avec les informations détenues par les ministères sectoriels compétents (Éducation, Santé, Agriculture, Hydraulique, Travaux publics, etc.). Il s'agit d'un travail long, méticuleux et quelque peu fastidieux, mais nécessaire pour garantir la qualité et donc asseoir la crédibilité de la base.

Une difficulté réside dans le fait que les ministères (administrations centrales) ne connaissent pas toujours les réalités du terrain et que les informations détaillées ne remontent pas jusqu'à eux. Mais elles existent, plus ou moins bien tenues, dans les circonscriptions régionales et locales, et c'est probablement là qu'on pourra le mieux les mobiliser. On verra plus loin quelle organisation on peut imaginer pour résoudre ce problème.

L'enrichissement

Les données issues des recensements et enquêtes ne sont pas les seules à posséder un intérêt au niveau local. On peut souhaiter les compléter par des données, statistiques ou descriptives, provenant d'autres sources : statistiques administratives élaborées par les ministères sectoriels compétents, comme pour le contrôle de qualité, voire certaines grosses ONG. Par exemple, à la simple identification des établissements scolaires que permet le questionnaire localité, on pourra ajouter le nombre d'élèves (par sexe, par classe, etc. si on le souhaite) en intégrant à la base des données provenant des rapports annuels à partir desquels sont élaborées les statistiques de l'Éducation. Ainsi pourra-t-on enrichir la base initiale. Encore faut-il que ces sources secondes soient d'une qualité et d'une homogénéité suffisante pour être intégrables dans la base et diffusables. On retrouve là le problème évoqué à propos du contrôle de qualité.

Ainsi pratiqué, l'enrichissement de la base de données locales, s'il est bénéfique à la base elle-même, l'est aussi aux statistiques administratives sectorielles. En effet – on l'a mentionné dans le paragraphe qui leur a été consacré – ces dernières pèchent souvent par manque d'exhaustivité. Le travail d'enrichissement qu'on suggère ici permet de repérer les manques et d'y remédier.

La mise à jour

Les données individuelles (l'individu étant ici le village ou le quartier urbain) qu'on peut envisager de mettre dans la base de données locales « vieillissent » vite : construction d'une route ou d'une adduction d'eau ici, mise en service d'une école ou d'un centre de soins là, implantation d'une usine ailleurs, forte dépopulation en un autre endroit, etc. Elles vieillissent plus vite que les données statistiques de tendances lourdes de la démographie (natalité, mortalité, etc.), celles-ci n'étant en outre pas finement localisées. Il faut donc, que ce soit pour l'utilisation directe des données de la base en vue de la diffusion ou pour son utilisation comme base de sondage, en actualiser le contenu à intervalles plus rapprochés que les dix ans qui séparent – dans le meilleur des cas – deux recensements de la population.

La mise à jour peut se faire de deux façons : par une nouvelle enquête localités ou par l'incorporation de données de source extérieure actualisées (et on est alors ramené à la problématique du contrôle et de l'enrichissement évoquée dans les paragraphes précédents). Mais la particularité d'une base de données de ce genre est qu'il n'est pas absolument nécessaire (même si c'est mieux) qu'elle soit mise à jour de façon uniforme. On peut ainsi envisager une mise à jour par roulement des données sur les

localités (y compris, le cas échéant, le dénombrement sommaire de leur population), à raison, par exemple, de 1/5 par an, ou le faire plus fréquemment dans les zones où la base vieillit le plus vite – c'est-à-dire où les changements affectant la population et son cadre de vie sont les plus rapides – que dans celles où elle reste assez stable, ou encore dans celles où la demande est la plus forte. On peut enfin envisager de ne faire porter la mise à jour de la base que sur certaines données, par exemple celles qui évoluent le plus rapidement ou qui sont le plus demandées.

Des multiples usages d'une base de données locales

Conçue comme un outil statistique, la base de données locales se révèle, à la réflexion, être aussi un instrument propre à des usages institutionnels et de gestion interne.

Un outil de diffusion

C'est la première utilisation normale d'une base de données que de servir à élaborer des tableaux et autres produits destinés à l'analyse et à la diffusion. Comme de toute base de données statistique, on peut tirer d'une base de données locales des tableaux statistiques classiques. Mais le plus intéressant sera de l'associer à un système d'information géographique pour permettre la production de cartes thématiques. Ainsi pourront être valorisées les données d'inventaire des infrastructures et des services présents dans les villages et quartiers avec une incomparable facilité de lecture et de compréhension.

Un outil de suivi des politiques

Contrôler la base de données, l'enrichir et la tenir à jour sont des nécessités, on vient de le voir. Mais il faut se garder, lors de l'introduction de données actualisées, de supprimer purement et simplement les données anciennes pour les remplacer par les nouvelles. En effet, comme toujours en statistique, s'il y a une demande pour des données à jour, il y en a aussi une – peut-être encore plus importante, en tout cas plus riche – pour la comparaison de situations successives.

On retrouve ici la question du suivi des politiques qui a pris une grande importance dans le cadre des DSRP. Savoir combien il existe aujourd'hui de points d'eau dans une région est intéressant ; plus intéressant encore est de savoir comment la situation a évolué en deux ans, quatre ans, etc. Il faudra donc d'abord assurer une mise à jour régulière de la base, ensuite trouver un moyen pour la conserver telle qu'elle se présente à certaines dates (par exemple au 1^{er} janvier de chaque année,

ou après une mise à jour importante). Ce n'est ainsi qu'en permettant des comparaisons temporelles que la base prendra tout son sens et toute sa valeur, sans pour autant, et loin s'en faut, qu'elle soit le seul instrument de suivi envisageable.

Une base de sondage

Les recensements, tant qu'ils ne sont pas devenus obsolètes, constituent la base de sondage au premier degré de la plupart des enquêtes socio-démographiques. Les unités primaires sont le plus souvent les villages ou les zones de dénombrement du recensement. On les stratifie en général selon des critères géographiques ; dans le cas des enquêtes agricoles, il peut s'agir de zones agro-écologiques. On n'utilise donc en fait que très peu les données du recensement pour la stratification, un peu plus pour réaliser des tirages à probabilité inégale et il s'agit alors de la population totale ou du nombre de ménages. Les données du questionnaire localité sont généralement ignorées, de même que celles relatives aux logements.

Si une base de données locales est constituée, on peut disposer d'une base de sondage très riche, permettant de stratifier les échantillons selon les critères – pas uniquement géographiques ou démographiques comme actuellement – les plus adaptés à chaque enquête¹⁶. Encore faut-il qu'elle soit correctement tenue à jour, car on sait que la qualité d'une base de sondage s'érousse rapidement.

Les données d'inventaire fournies par le questionnaire localité peuvent aussi permettre de constituer des échantillons d'infrastructures ou d'équipements dont les caractéristiques seront étudiées par une enquête approfondie ; par exemple, un échantillon d'écoles repérées lors de l'administration du questionnaire localité pourrait faire l'objet d'une enquête détaillée sur les conditions dans lesquelles l'enseignement est dispensé, avec l'avantage que la source questionnaire localité permet d'avoir en même temps les établissements publics et les autres.

Quoi qu'il en soit, l'utilisation de la base de données locales comme base de sondage de toutes les enquêtes démographiques et sociales et des enquêtes auprès des établissements à caractère social renforcerait son rôle d'outil d'intégration.

¹⁶ On pourra par exemple utiliser des critères tels que l'existence de certaines infrastructures, les caractéristiques de l'habitat, etc.

Un outil de coordination des systèmes nationaux de statistique

On a vu que les informations de base, issues du questionnaire localité, sont très diversifiées et concernent les domaines de responsabilité de nombreux ministères : Éducation, Santé, Hydraulique, Travaux Publics, etc. Il en va de même des informations pouvant servir à contrôler la qualité de la base ou à la mettre à jour, voire à en enrichir le contenu. On tient donc, avec une base de données localisées, un outil de synthèse dans lequel des données de diverses sources sont juxtaposées ou confrontées. Pour en assurer la qualité initiale, l'enrichissement et la mise à jour, il faut que s'établisse une concertation entre les fournisseurs ou détenteurs de données, c'est-à-dire entre ministères ou institutions spécialisés. N'est-ce pas là une façon concrète de développer et de faire vivre la coordination statistique ?

Un outil au service de la décentralisation

Beaucoup de pays s'engagent sur la voie de la décentralisation administrative, par laquelle de nouvelles responsabilités sont données aux échelons régionaux du pays (régions, provinces, départements, peu important les vocables employés). Parallèlement, et pour permettre aux autorités locales de bien connaître le territoire qu'elles administrent et ainsi de mieux exercer leurs responsabilités, la statistique se décentralise aussi. Qu'elle se donne avec une base de données locales la possibilité de fournir aux autorités régionales les informations nécessaires à leur action est un puissant levier pour renforcer sa position, accroître son influence et obtenir des moyens.

Un outil d'animation des services régionaux de statistique

Les services décentralisés de la statistique, lorsqu'ils existent – et c'est de plus en plus souvent le cas avec l'extension de la décentralisation administrative qu'on a évoquée au paragraphe précédent – manquent souvent de motivation et de moyens et leur rôle n'est pas toujours clairement défini. Leur confier la responsabilité de travaux relatifs à la base de données locales (contrôle, enrichissement, mise à jour et peut-être diffusion locale) peut être une façon de leur donner une raison d'être simple et concrète : répondre à la demande locale (on a dit qu'elle était forte, même si elle est imprécise) tout en contribuant à la vie d'un outil national.

On a vu que beaucoup de données susceptibles de contribuer au contrôle, à l'enrichissement et à la mise à jour de la base étaient détenus par les ministères compétents, mais beaucoup de ces données ne remontent pas aux administrations

centrales et sont détenues par les services régionaux de ces administrations, plus opérationnels. Les services déconcentrés de la statistique auraient ainsi à se rapprocher de leurs homologues des ministères de l'Éducation, de la Santé, des Travaux Publics, etc. pour obtenir et valider les informations qu'ils détiennent avant qu'elles soient intégrées dans la base. On assisterait alors à la « décentralisation de la coordination statistique ».

Vers un programme de travail à moyen terme

On a beaucoup parlé du recensement comme pivot du système de statistiques démographiques et sociales. On a dit que, simplifié, il devait être complété par une batterie d'enquêtes par sondage. Chaque institut national de statistique devrait donc préparer périodiquement un programme pluriannuel d'enquêtes articulées autour du nouveau recensement et de la base de données localisée, le soumettre aux partenaires-bailleurs et en discuter avec eux. L'existence même d'un tel programme, solidement structuré et argumenté, montrerait que l'institut national de statistique a des objectifs et une stratégie et les partenaires-bailleurs seront à coup sûr plus enclins à s'engager sur cette base que sur des opérations isolées dont aucune n'offre de visibilité d'ensemble, chacun se concentrant sur les opérations (recensement simplifié, enquêtes par sondage, amélioration des statistiques administratives) qui sont le plus en relation avec son programme. Cela ne signifie pas que tous les partenaires-bailleurs joueront le jeu, mais ils le joueront d'autant plus volontiers qu'ils sentent qu'il y a en face d'eux une logique et une volonté.

Quelques remarques

L'agriculture et le secteur informel

Les recensements de l'agriculture et les enquêtes sur le secteur informel, qu'on range classiquement dans le domaine des statistiques économiques plutôt que dans celui des statistiques socio-démographiques, sont conceptuellement et techniquement très proches des enquêtes socio-démographiques : même base de sondage, même méthode d'interview, même organisation des opérations de collecte et de l'exploitation. De plus, ils comprennent au moins une composante de nature socio-démographique, l'identification des membres du ménage-exploitation ou du ménage-entreprise.

Les services de statistique agricole seront parmi les utilisateurs privilégiés d'une base de données locales qui pourra leur servir de base sondage pour

leurs enquêtes, et notamment pour le recensement. Ils pourront en affiner l'analyse en intégrant des données sur les infrastructures et équipements accessibles par les exploitations agricoles. Ils pourront aussi contribuer à son contrôle, à son enrichissement et à sa mise à jour. Leur association à la conception et à la gestion de la base est donc essentielle¹⁷. Ce besoin de statistiques communautaires est d'ailleurs exprimé par la FAO dans son programme mondial de recensement de l'agriculture pour les années 2010¹⁸.

Les autres statistiques économiques

La base de données locales qu'on a proposé de constituer pour servir l'étude et l'analyse du domaine socio-démographique essentiellement à partir des questionnaires localité remplis à l'occasion des recensements de population contient des données de caractère plutôt économique, relatives à l'existence d'infrastructures et d'équipements. Il n'est pas inenvisageable d'y introduire des données économiques plus détaillées. Si, par le questionnaire localité, on a repéré par exemple l'existence d'unités industrielles, on peut envisager de compléter la base par des informations sur la production de ces unités ou sur le nombre de personnes employées. La problématique est la même que celle qu'on a citée à propos de l'enrichissement de la base, transposée au domaine économique. Mais on ne développera pas plus ce point, qui sort de l'objectif poursuivi ici.

Définition des localités

On n'entamera pas un débat, qui n'a pas sa place dans un texte de portée générale et serait d'ailleurs sans fin, sur la définition qu'il convient de retenir pour les échelons sur la base desquels sera construite la base de données locales (village, commune, etc.). On se contentera d'une remarque, mais celle-ci est fondamentale pour la conception même de la base.

Toutes les données ne sont pas pertinentes ni disponibles aux mêmes niveaux. Prenons un exemple dans le domaine agricole qu'on vient d'aborder. Si on s'intéresse à la vulgarisation et plus précisément aux effectifs de vulgarisateurs, on ne peut espérer de données que pour des circonscriptions agricoles couvrant un nombre de villages qui peut être assez élevé au sein desquelles œuvre une équipe de vulgarisateurs. Il en va de même de toutes les données qui se rattachent à des

unités administratives locales¹⁹. Faut-il pour autant se priver de celles qu'il ne serait pas possible d'obtenir par village ?

Par ailleurs, si on cherche à déconcentrer certains résultats en utilisant les méthodes d'étude des petits domaines, on doit se souvenir qu'il existe une limite à la déconcentration. Les résultats d'une enquête nationale portant sur 3 ou 4000 ménages ne pourront pas être déconcentrés pour chacun des villages d'un pays, au nombre de plusieurs milliers. Il faudra se contenter d'échelons intermédiaires (départements, sous-préfectures, districts, communes, etc.), et qui regroupent chacun plusieurs dizaines de villages et plusieurs milliers, voire dizaines de milliers, de ménages.

Il résulte de ces observations qu'il peut s'avérer utile de concevoir une base comportant des échelons géographiques emboîtés : hameau, village, commune, sous-préfecture, département. Il ne paraît cependant pas opportun, pour des raisons de commodité de constitution et de gestion de la base, de multiplier ces échelons.

Les aspects financiers

La question ici posée est de savoir à quelles économies conduirait la réduction drastique du questionnaire du recensement qui est proposée et dans quelle mesure ces économies permettraient, à **coût total égal**, de faire plus et mieux, c'est-à-dire de financer d'autres enquêtes, d'augmenter la taille des échantillons, ou encore de développer les questionnaires d'enquêtes déjà réalisées. Une telle analyse ne peut être réalisée que sur la base d'un programme à moyen terme tenant compte de la demande à satisfaire dans un pays donné et des structures de coûts dans le dit pays. Elle n'a donc pas sa place ici.

Vers un début de mise en œuvre

On pourra s'effrayer de la multiplicité des défis que suggère cet article. Défis techniques, par exemple ce qui touche à l'échantillonnage de la batterie d'enquêtes qui complètera le recensement ou à l'usage des méthodes d'étude de petits domaines. Défis institutionnels : comment faire travailler ensemble ministères et personnes dont certains intérêts peuvent diverger ? Défis politiques aussi : il faudra convaincre tous les intervenants nationaux ou extérieurs d'entrer dans une logique d'intégration alors que chacun œuvre dans une

¹⁷ Il est significatif que l'inventaire communal français – voir encadré n° 2 – ait toujours été réalisé en collaboration par l'INSEE et le service statistique du Ministère de l'Agriculture (SCEES).

¹⁸ Voir bibliographie.

¹⁹ En Afrique subsaharienne, le village n'est en générale pas une unité administrative ; le chef de village, qu'il soit choisi par la population ou nommé par l'administration, n'est qu'un correspondant de celle-ci et ne détient d'autorité que coutumière..

logique sectorielle. Défis financiers enfin : les bailleurs joueront-ils ce jeu complexe et parfois contraire à leurs normes ? On risque donc de ne pas bouger devant l'ampleur de la tâche.

Mais reprenons successivement les deux thèmes de cet article et voyons comment on peut envisager de les mettre en œuvre de façon échelonnée dans le temps. N'oublions pas pour autant le but ultime qui a été ici proposé, qui est la combinaison, au sein d'un système intégré de statistiques démographiques et sociales, d'un recensement de la population, de l'habitat et des localités), d'une base de données locales établie fondamentalement à partir du recensement mais pouvant incorporer d'autres sources, et enfin d'une batterie d'enquêtes socio-démographiques, plus corsée que celle qu'on observe d'habitude.

Le premier thème, qui faisait l'objet des deux premières parties de cet article, était celui du recensement de la population, qu'on a proposé de faire évoluer vers un « recensement de la population, de l'habitat et des localités » (notamment allégé par rapport à ce qui se fait actuellement en Afrique sur le plan des individus, à peine sur ceux de l'habitat et des localités). C'est ce thème qui est le plus difficile à traiter et qui fait l'objet principal des défis qu'on vient de mentionner. Et on conçoit bien qu'il ne peut y avoir de changement en ce domaine dans un pays donné que si : a) un recensement se profile à terme de deux ou trois ans, b) la ferme volonté existe de la part des statisticiens et démographes de faire évoluer les choses, c) une approche résolument systématique est suivie. C'est donc une affaire d'assez longue haleine.

Il en va tout autrement du deuxième thème, annoncé puis détaillé dans les parties suivantes, celui de la base de données locales. Beaucoup de pays sont peu avancés sur ce plan. À une nuance près, cependant : il existe presque partout un gisement de données locales pas ou peu exploité, provenant du questionnaire village du dernier recensement. Il n'est nul besoin de négociations avec les partenaires ni de prérequis techniques pour exploiter ce gisement, à condition toutefois qu'il ne soit pas trop ancien et que les données qu'il contient ne soient pas complètement obsolètes. L'exploitation de ce gisement sous forme cartographique permettra de tester l'intérêt que les utilisateurs peuvent porter à ce genre de données et à leur présentation ; les pays qui ont fait un tel test peuvent témoigner de l'intérêt qu'il a en effet provoqué. Ce sera aussi une sorte d'opération pilote à partir de laquelle on pourra élaborer un programme répondant au mieux aux attentes pour le prochain recensement.

Ce n'est que lorsque tous les partenaires, nationaux et extérieurs, du système statistique national auront eu la preuve par l'exemple que la base de données locales, malgré son aspect très simple, est une source d'information puissante, qu'ils pourront considérer favorablement une évolution drastique qui impliquerait l'allègement du recensement et la révision de la nature, du contenu et de la taille des enquêtes par sondage dans un esprit de système. Cela représente du travail et de l'imagination. Mais l'enjeu en vaut la peine.

Références Bibliographiques

FAO (2006), *A system of integrated agricultural censuses and surveys. Vol 1: World Programme for the Census of Agriculture 2010*. 174 pages. (traduction en français en cours d'impression)

Accessible sur le site de la FAO : www.fao.org/es/ess/census/wca2010

Gasse D. (1978), *Les fichiers de village : problématique générale et expérience gabonaise*. AMIRA-AFIRD, 96 pages.

Nations Unies (Février 2007), *Principles and Recommendations for Population and Housing Censuses, Draft 2*, 371 pages.

Accessible sur le site de la division statistique des Nations Unies :

<http://unstats.un.org/unsd/demographic/sources>

PARIS21 (2004), *Guide pour l'élaboration d'une Stratégie Nationale de Développement Statistique*, 40 pages

UNFP (2003), *Counting the people. Constraining census costs and assessing alternative approaches*, 139 pages

Accessible sur le site de Paris21 : www.paris21.org/pages/task-teams/teams/list , cliquer sur « census » puis sur « documents ».

Pour en savoir plus sur l'Inventaire communal français, on peut consulter le site de l'INSEE : www.insee.fr/fr/ico98/ico98asp