

QUESTIONS GENERALES

CHAPITRE 6**LA QUALITE DE L'OBSERVATION :
ERREURS, CONTROLES, EVALUATION***Francis GENDREAU***1. INTRODUCTION**

Comme toute observation, l'observation en démographie vise à mesurer des grandeurs ou des phénomènes.

Comme pour toute observation, les mesures qui en résultent sont entachées d'erreurs.

Il s'agit alors d'identifier quelles sont les principales sources d'erreurs possibles de façon à y remédier au cours d'une opération de collecte :

- soit lors de la préparation de l'opération : c'est tout l'aspect traité par ailleurs de la connaissance du milieu, de l'élaboration du questionnaire, de la cartographie, ou de la formation des agents chargés de la collecte ;
- soit lors de la réalisation de l'opération : c'est le problème des contrôles à mettre en oeuvre pour déceler et corriger les erreurs qui se seront, malgré toutes les précautions, produites.

Mais l'analyse des sources d'erreurs et les remèdes que l'on y apporte ne constituent qu'un aspect du problème, qui pourrait être qualifié d'interne à l'opération : il faut prendre en compte un second volet, externe celui-là, qui consiste à comparer les résultats obtenus (éventuellement les résultats ajustés après utilisation de techniques indirectes) avec d'autres résultats disponibles par ailleurs ; c'est toute la question de l'évaluation de la qualité et de la précision des estimations, et de la cohérence de l'ensemble des données disponibles sur la population considérée.

2. LES ERREURS

Les erreurs susceptibles d'être commises au cours d'une opération de collecte sont nombreuses. On ne parlera pas ici de l'erreur de sondage traitée par ailleurs, mais des trois autres grandes catégories d'erreurs qui concernent le dénombrement, qui portent sur le contenu des réponses, ou qui apparaissent au cours de la phase d'exploitation.

2.1. Erreurs de dénombrement

Il s'agit des omissions, et, la plupart du temps dans une moindre mesure, des doubles comptes; il est rare que les deux erreurs se compensent. Ces erreurs ont une très grande importance car elles sont souvent systématiques, et touchent des catégories particulières de populations.

Les exemples les plus classiques d'erreurs de ce type sont les suivants :

- omission de villages, liée à une mauvaise cartographie et à un mauvais découpage en zones de dénombrement ;
- omission de hameaux éloignés ou isolés dans un village, ou omission des "campements de culture", ou, au contraire, inclusion à tort de hameaux soi-disant rattachés au village enquêté mais se trouvant en fait hors du champ géographique de l'opération ;
- omission de ménages du fait d'un mauvais plan du village ou du quartier, ou parce qu'il s'agit de ménages isolés (par exemple des vieillards vivant seuls, ou des jeunes immigrants urbains souvent isolés, mobiles, et peu connus de leurs voisins), ou de ménages totalement absents (dans le cas d'un dénombrement de la population résidente) ;
- omission ou double compte d'individus : il y a là de multiples occasions d'erreurs. Par exemple : omission ou double compte de migrants, d'enfants scolarisés hors de leur famille, d'enfants "confiés", d'hommes polygames,... On peut aussi signaler les omissions "préférentielles" liées à certaines catégories particulières de population : les étrangers, les réfugiés, les nomades, les pygmées... On peut enfin citer deux autres sources particulières d'erreurs de cette catégorie :
 - . les omissions et doubles comptes liés à la mauvaise compréhension (par l'enquêteur ou par l'enquêté) des concepts utilisés pour définir la population à saisir (population de droit, population de fait). On peut aussi citer ici les erreurs de dénombrement liées à la mauvaise compréhension par l'enquêteur, au cours d'une enquête par sondage, des instructions relatives au tirage des ménages-échantillons ;
 - . pour certaines sources administratives, de fréquentes omissions dues à la négligence ou à la volonté délibérée des intéressés qui ne se déclarent pas : listes électorales, recensements administratifs, registres des étrangers...
- erreurs dans les déclarations d'événements :
 - . omission de déclarations à l'état civil ;
 - . omission de déclarations d'enfants nés-vivants dans les questionnaires rétrospectifs sur la vie génésique des femmes, ou, inversement, inclusion abusive de mort-nés ;

- . erreurs dans les déclarations de naissances et de décès au cours d'une période de référence (par exemple les douze derniers mois), qu'il s'agisse d'omissions volontaires (par exemple non déclaration d'un décès) ou d'erreurs liées à la mauvaise perception de la durée de la période de référence ou à l'imprécision du souvenir de la date de l'événement.

2. 2. Erreurs de contenu

En supposant qu'un individu ou un événement ait été correctement signalé et enregistré par l'agent chargé de la collecte, celui-ci peut commettre des erreurs dans l'enregistrement des caractéristiques qui décrivent cet individu ou cet événement :

- renseignement non fourni : l'espace réservé à la réponse sur le questionnaire est resté vide. Il est bon de rappeler ici que **toute** case doit être renseignée, y compris si la question est sans objet (par exemple en inscrivant un tiret), si la réponse est zéro (en l'indiquant) ou indéterminée (il faut alors le signaler, par exemple "N.D." ou "ind.");
- renseignement fourni mais illisible ou susceptible de donner lieu à diverses interprétations (chiffres mal formés). Ce cas peut aussi se rencontrer lorsqu'un 6 dans la colonne "âge" peut signifier 6 mois ou 6 ans, ou lorsqu'un 40 dans la colonne "âge ou date de naissance" peut signifier 40 ans ou 1940 ;
- renseignements ne correspondant pas aux instructions, par exemple :
 - . réponse indiquée ne figurant pas dans la liste exhaustive des réponses possibles (liste des ethnies, des professions...);
 - . réponse trop détaillée ou trop large : si l'on demande la sous-préfecture de naissance, il ne faut pas inscrire le village ou la commune, ou au contraire la préfecture ou la région.
- renseignements imprécis, par exemple "fonctionnaire" n'est pas une profession ;
- renseignement faux, qu'il s'agisse d'une erreur volontaire (enquêteur déclarant posséder la nationalité du pays et cachant sa qualité d'étranger) ou involontaire (question mal posée par l'enquêteur ou mal comprise par l'enquêté). Ces erreurs sont graves car fréquentes et parfois systématiques : c'est par exemple tout le problème de la déclaration de l'âge (rajeunissement ou vieillissement systématique par les enquêtés, inscription d'un âge d'1 an au lieu de 0 an...).

2. 3. Erreurs d'exploitation (1)

Après la phase de collecte de l'information sur le terrain, se situe celle de l'exploitation, au cours de laquelle de nouvelles erreurs peuvent apparaître :

- erreurs dans la gestion des dossiers (questionnaires individuels, questionnaires ménages,...) : perte d'un dossier, oubli d'un dossier ou au contraire prise en compte deux fois du même dossier...;
- erreurs de chiffrage ;
- erreurs de saisie.

(1) Pour plus de détails sur ce point, le lecteur se référera aux ouvrages traitant de l'exploitation des données, par exemple "le traitement informatique des enquêtes statistiques" cité en bibliographie.

3. LES CONTROLES

Pour éviter les erreurs, outre tout le soin à apporter à la conception de l'opération, des contrôles doivent être effectués de la manière la plus rigoureuse possible tout au long de la phase de terrain puis de celle d'exploitation. Il faut en effet être bien conscient de ce que tout point faible dans la chaîne des opérations peut compromettre l'ensemble des résultats. Il paraît d'autant plus utile de rappeler avec force cette proposition que le contrôle est une opération fastidieuse qui ne suscite généralement pas l'enthousiasme. Pourtant, c'est l'élément-clé de la réussite d'une opération.

Chronologiquement, on peut envisager les contrôles suivants :

3. 1. Le contrôle général du travail de l'enquêteur

Ce type de contrôle est essentiel : l'enquêteur doit sentir que son travail est suivi attentivement, non pas dans un but "répressif", mais en vue d'obtenir de bons résultats. Cette présence sur le terrain doit être effective de la part du contrôleur ou du superviseur dès les premières journées de terrain car il faut que très vite les éventuelles erreurs que commet l'enquêteur, sa mauvaise compréhension de certaines instructions, soient corrigées. De plus, c'est la meilleure manière pour les responsables de l'enquête de s'assurer que les instructions sont suivies de la même façon par l'ensemble des enquêteurs et que l'enquête sera homogène. Cette présence doit ensuite se poursuivre tout au long de l'opération pour corriger les "dérives" susceptibles d'apparaître graduellement.

Les points sur lesquels doit porter ce contrôle sont multiples. Par exemple :

- déplacement de l'enquêteur dans les ménages et non pas convocation de la population (ou remplissage en chambre des questionnaires avec la seule inspiration de son imagination,...) ;
- respect par l'enquêteur des limites de sa zone;
- présentation de l'enquête auprès des enquêtés et façon de poser les questions ;
- visite systématique de tous les ménages à enquêter, y compris ceux des hameaux les plus éloignés.

3. 2. Le contrôle des questionnaires sur le terrain

Ce type de contrôle est lui aussi essentiel : il doit être systématique de la part du contrôleur, porter sur tous les questionnaires, et être réalisé régulièrement, tous les soirs. C'est ici surtout qu'apparaît le caractère rébarbatif et fastidieux du contrôle, car la relecture de quelques dizaines de questionnaires n'a rien d'enthousiasmant, surtout après une journée fatigante de terrain.

Il s'agit pour le contrôleur d'effectuer les contrôles suivants :

- complétude des documents : le nombre de documents doit être vérifié, ainsi que l'égalité entre le nombre d'unités annoncé dans un questionnaire d'un certain type et le nombre de questionnaires d'un autre type en découlant : par exemple, si dans une unité d'observation, il y a 2 femmes de 12 ans et plus à l'intérieur du questionnaire "ménage", il doit y avoir, pour une enquête sur la fécondité, 2 questionnaires "femme" ;

- respect des instructions : toutes les "cases" doivent être renseignées, l'écriture doit être lisible, les réponses doivent correspondre à celles autorisées par les instructions, elles doivent être précises...;
- vraisemblance ou cohérence des réponses : l'examen du questionnaire doit permettre de relever les erreurs les plus flagrantes ;
- exactitude des reports : un renseignement inscrit à un endroit et repris à un autre endroit doit y être reporté sans erreur; c'est le cas par exemple de la date de naissance d'un enfant de moins d'un an figurant à la fois dans le tableau des membres du ménage, dans le cartouche des naissances des 12 derniers mois, et dans le questionnaire "femme".

3. 3. Autres contrôles

Le contrôle par un nouvel entretien est très important : le contrôleur doit, accompagné par l'agent enquêteur, revenir systématiquement dans quelques ménages, pour procéder lui-même à un entretien afin de vérifier, non plus sur document, mais directement, la qualité du travail de l'enquêteur.

Ce contrôle, ainsi que les deux précédents, présente l'intérêt essentiel d'être réalisé sur le terrain et donc de corriger les erreurs à la source. En particulier, le deuxième contrôle permet de renvoyer le lendemain l'enquêteur auprès du ménage pour compléter les renseignements manquants ou rectifier les anomalies décelées.

Il y a enfin d'autres contrôles, susceptibles d'être effectués en aval des précédents :

- le contrôle de complétude des documents à chaque transfert des dossiers, pour éviter les pertes ;
- les contrôles de vraisemblance des premiers résultats, issus d'un premier dépouillement manuel rapide sur le terrain : taille moyenne des ménages et rapport de masculinité, par exemple; ces indices peuvent être calculés séparément pour chaque enquêteur ;
- les contrôles effectués au moment du chiffrage et de la saisie.

4. L'EVALUATION

Quel que soit le soin apporté à la réalisation de l'observation (conception de l'opération, formation du personnel, mise en oeuvre des contrôles), il est évident que les résultats seront entachés d'erreurs. Il est important dès lors de procéder à une analyse permettant d'évaluer la qualité de ces résultats de façon que les utilisateurs aient les éléments leur permettant d'en apprécier les biais éventuels, la précision, et donc les limites d'utilisation.

4. 1. Les objectifs de l'évaluation

L'évaluation doit fournir des indications qualitatives et, si possible, quantitatives sur la confiance que l'on peut accorder aux résultats, et doit déboucher sur la correction des données brutes et le calcul de données ajustées. Ce deuxième objectif est bien sûr plus difficile à atteindre.

Il faut en effet distinguer le niveau auquel on opère :

- les données globales, fournies pour l'ensemble du champ de l'opération font généralement l'objet d'une évaluation ; il est en effet important de savoir s'il y a eu sous-estimation ou surestimation, et dans quelle proportion, de l'effectif total de la population fourni par un recensement ou du taux brut de mortalité fourni par une enquête ;
- ce type de renseignement est, de même, utile au niveau des différentes circonscriptions géographiques ou répartitions ; mais l'on sait bien, dans le cas d'une enquête par sondage, que les résultats ne peuvent être valablement fournis à un niveau géographique trop fin, ou que les taux de mortalité par sexe et par âge sont connus avec une précision moindre que le taux brut de mortalité ;
- enfin il est intéressant de connaître les caractéristiques des individus ou des événements omis ou comptés deux fois.

4. 2. Les moyens internes d'évaluation

Parmi les méthodes dont dispose le démographe pour procéder à l'évaluation des résultats d'une opération, il y a d'abord celles qui ne font appel qu'aux données recueillies lors de celle-ci :

- un premier moyen, insuffisamment utilisé, est l'étude des difficultés rencontrées sur le terrain : en effet dans certaines zones, la réalisation de l'opération a pu se heurter à des problèmes particuliers (psychologiques, climatiques, administratifs) rendant moins fiables les résultats pour cette zone, sans pour autant remettre en cause l'opération dans les autres zones ;
- le second moyen, qui concerne les recensements, est bien sûr la réalisation d'une enquête de contrôle post-censitaire : même si de telles enquêtes posent de réels problèmes de mise en oeuvre, il faut insister ici sur leur intérêt et recommander leur réalisation. Le lecteur trouvera dans la bibliographie plusieurs références sur cette question étudiée par ailleurs (voir chapitre 7) ;
- une autre méthode consiste, pour les enquêtes par sondage, à calculer l'"erreur" de sondage, c'est-à-dire la variance des divers estimateurs ;
- les méthodes de l'analyse démographique permettent de calculer des indicateurs donnant une idée de la qualité de l'observation (par exemple, pour l'âge, les indices de WHIPPLE, de MYERS ou de BACHI), ou d'ajuster sous certaines hypothèses les données brutes (structure par âge, fécondité, mortalité) ;
- dans certaines opérations, l'exploitation des résultats par enquêteur et par contrôleur ou superviseur permet de mettre en évidence les zones où la qualité de travail laisse à désirer.

4. 3. Les moyens externes d'évaluation

Une opération de collecte ne doit jamais être considérée isolément : elle s'inscrit dans une succession d'observations dont il est bon d'avoir une vue globale, car le progrès des connaissances est, dans le domaine démographique comme dans tous les autres, le résultat d'une accumulation patiente et obstinée.

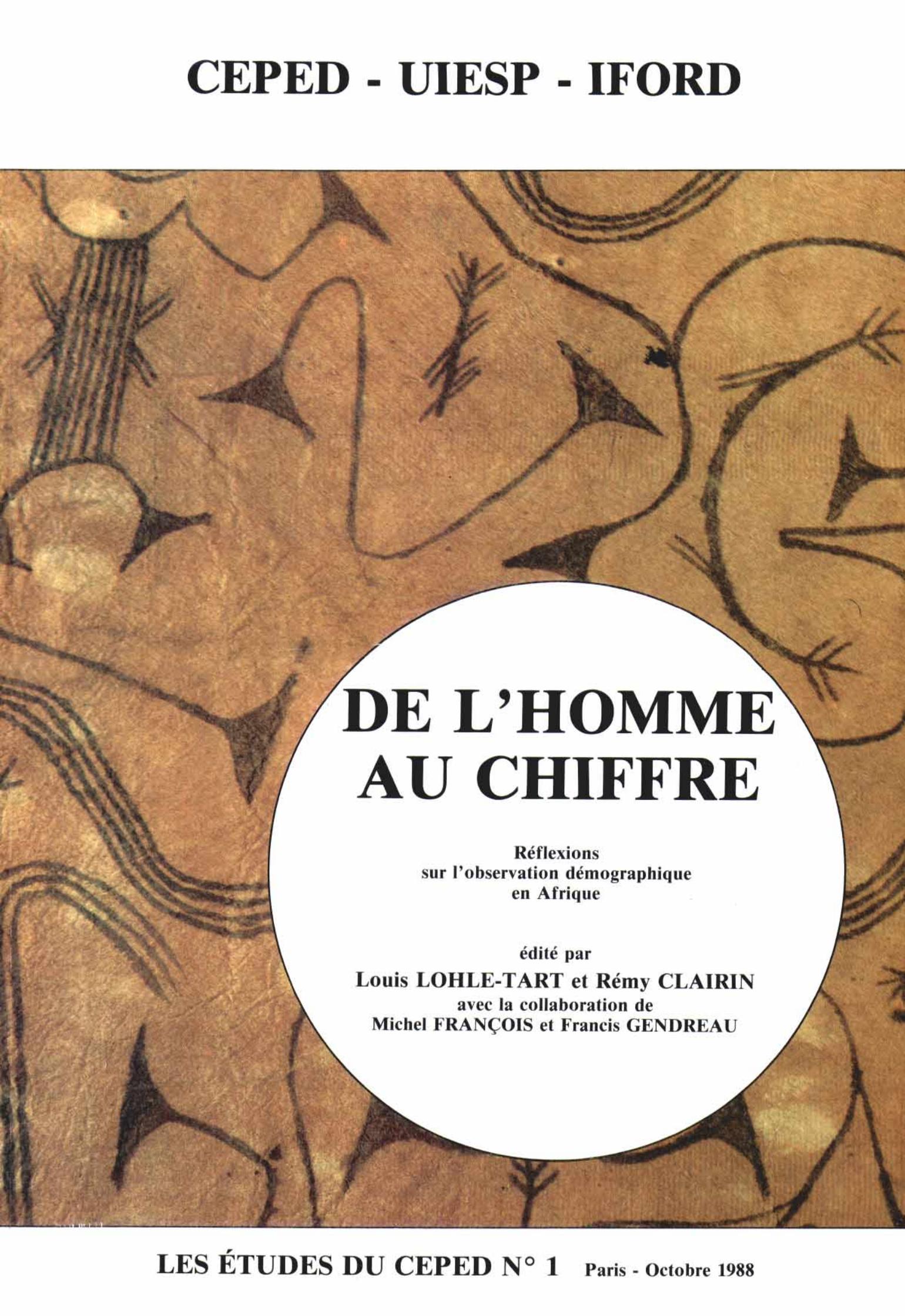
Des résultats émanant d'une nouvelle opération de collecte doivent donc toujours être replacés dans le cadre de l'ensemble des estimations antérieures. Ceci permet, grâce à un jeu complexe et délicat d'allers et retours successifs, à la fois d'évaluer la qualité de l'opération qui vient d'être réalisée, et d'affiner l'évaluation des opérations antérieures : c'est là tout le problème de la mise en cohérence des différentes estimations et de la mise en évidence de l'évolution des différents paramètres démographiques.

Il est difficile à ce sujet de fournir des méthodes précises, tant les situations sont diverses. On se contentera donc d'indiquer quelques exemples :

- comparaison des résultats d'un recensement ou d'une enquête avec des sources administratives indépendantes qui peuvent être, selon la nature des phénomènes observés : des listes électorales, des fichiers des services de santé, des statistiques hospitalières, des recensements administratifs, l'état civil...;
- analyse de cohérence entre plusieurs opérations successives, ce qui est possible dès lors que l'on dispose d'une série d'au moins 2 ou 3 enquêtes ou recensements; dans ce cas, comme dans le cas précédent, la comparaison à un niveau géographique fin (le village) est souvent instructive car des ordres de grandeur très divergents dans des effectifs de population d'un village méritent explication : soit phénomène réel lié à un événement particulier, soit erreur de l'une et/ou l'autre source ;
- un dernier exemple doit être mentionné, plus particulier car il s'agit d'une méthode d'observation intégrant dans sa conception même un caractère externe : c'est la double collecte (voir chapitre 7).

Il faut bien souligner que cette évaluation externe ne se limite pas à une confrontation point par point des diverses sources (comme la mise en rapport des effectifs relevés successivement par deux ou plusieurs opérations). Il peut s'agir d'un véritable travail de synthèse visant à mettre en lumière la cohérence ou l'incohérence des données, par exemple en vérifiant si deux effectifs sont compatibles, compte tenu des fourchettes d'estimation de la mortalité, de la fécondité et des migrations.

CEPED - UIESP - IFORD



**DE L'HOMME
AU CHIFFRE**

Réflexions
sur l'observation démographique
en Afrique

édité par
Louis LOHLE-TART et Rémy CLAIRIN
avec la collaboration de
Michel FRANÇOIS et Francis GENDREAU