

Olivier Barbary, en collaboration avec F. Bureau et A. Michel
ORSTOM (dépt. SDU)
U.R. : "Systèmes d'accumulation des hommes et des richesses"

DEFINITION ET MISE EN ŒUVRE D'UN PLAN DE SONDAGE
AREOLAIRE SUR IMAGE SPOT POUR UNE ENQUETE
DEMOGRAPHIQUE A QUITO (EQUATEUR)

Quito, Mai 1988.

O.R.S.T.O.M. Fonds Documentaire

N° :

Cote :

~~A 25747~~
ex 2

7/6/89

Plan

Pages	
6	Introduction
8	1 : Objectifs généraux, plan d'expérience, réalisation des opérations de collecte pour les enquêtes préliminaires.
8	1.1 : <u>Introduction : Objectifs généraux, plan d'expérience</u>
10	1.2 : <u>Réalisation de l'enquête "morphologie"</u>
10	1.2.1 : Sélection de l'échantillon
11	1.2.2 : Elaboration des documents d'enquête
12	1.2.3 : Collecte des données
12	1.2.4 : Codification, saisie, contrôles de cohérence et recodage
13	1.2.5 : Exploitation de l'enquête morphologie en vue du tirage d'un échantillon représentatif
20	1.3 : <u>Réalisation de l'enquête "relations bâti/population"</u>
20	1.3.1 : Sélection de l'échantillon
21	1.3.2 : Elaboration des documents d'enquête
21	1.3.3 : Collecte des données
23	1.3.4 : Codification, saisie
24	2 : Enquête préliminaire "bâti / population", exploitation pour la mise au point du premier degré du sondage : échantillonnage de pâtés de maisons
24	2.1 : <u>Introduction : le contexte et les questions posées</u>
26	2.2 : <u>Les données et les outils statistiques utilisés</u>
26	2.2.1 : Les données
26	-Données démographiques
27	-Taille des unités primaires
27	- Stratifications
31	2.2.2 : Les outils statistiques utilisés
31	-1. Phase exploratoire : corrélations, régressions linéaires, analyses de la variance.

Pages

31	-2. Phase conclusive : test des plans de sondages à l'aide du programme Varech mis au point au cours de l'expérience menée à Marseille
32	2.3 : <u>Résultats</u>
32	2.3.1 : Phase exploratoire
36	2.3.2 : Phase conclusive
36	-Probabilités d'appartenance des unités primaires à l'échantillon
39	-Stratifications
42	-Allocations optimales
45	2.4 : <u>Conclusion</u>
47	3 : Enquête préliminaire "bâti / population", exploitation pour la mise au point du second degré du sondage : échantillonnage des ménages dans les pâtés de maisons
47	3.1 : <u>Introduction : objectifs principaux</u>
48	3.2 : <u>Analyses préliminaires : les possibilités "d'explication" des caractéristiques démographiques des ménages par la morphologie du bâtiment, la morphologie et la localisation du pâté de maison.</u>
48	3.2.1 : Les questions posées, les données et les outils statistiques utilisés
48	-Les questions posées
48	-Les données
49	-Les outils statistiques utilisés
50	3.2.2 : Résultats
50	-A. Analyse de la variance à un critère : pouvoir discriminant des facteurs "explicatifs" sur les différents caractères démographiques étudiés
56	-B. Analyse de la variance multi-critères : redondance/indépendance de l'information apportée par les trois facteurs "explicatifs".
58	3.2.3 : Conclusions
59	3.3 : <u>Détermination du second degré du sondage : échantillonnage de ménages dans un pâté de maisons</u>
59	3.3.1 : Position du problème
60	3.3.2 : Formulaire, programme informatique
64	3.3.3 : Résultats
68	3.3.4 : Conclusions

Pages

70	4 : Le plan de sondage et la mise en œuvre de l'enquête
70	4.1 : <u>Introduction : le double objectif méthodologique et thématique de l'enquête, présentation du questionnaire</u>
70	4.2 : <u>Description du plan de sondage, étapes de sélection et d'inférence</u>
71	4.2.1 : Résumé des résultats obtenus lors des phases d'expérience
72	4.2.2 : Description du plan de sondage : étapes de sélection au premier et second degré
73	4.2.3 : Formulaire d'estimation pour l'étape d'inférence
73	A : Préliminaire, les problèmes posés par l'inférence dans le cas du sondage systématique
74	B : Formulaire
81	4.3 : <u>Mise en œuvre du plan de sondage : sélection de l'échantillon et organisation du travail de collecte</u>
81	4.3.1 : Définition de la base de sondage sur l'image: la limite urbaine
81	A: Les documents utilisés
81	B : Mise en œuvre et règles de décision
82	C : Validation
82	4.3.2 : Stratification de la base de sondage sur l'image : les critères retenus et la réalisation pratique, mesures de surface de chaque strate
82	A : Mise en œuvre de la classification des pixels de l'image SPOT selon la densité du bâti
82	B: Evaluation de la classification
83	C: Elaboration de la stratification
83	D: Mesures de surface de chaque strate
84	4.3.3 : Tirage de l'échantillon de pâtés de maisons
84	A : Détermination du nombre total de pâtés de maison de l'échantillon et de leur répartition dans les strates
86	B. Définition des grilles
87	C. Tirage
90	4.3.4 : Recensement des ménages dans les pâtés de maison et sélection des ménages à enquêter
90	A. Recensement des ménages
90	B. Sélection des ménages à enquêter
93	4.3.5 : Mesure de surface des unités primaires
93	A: Les méthodes de mesure employées
96	B: Evaluation de la précision des techniques de mesures
98	4.4 : <u>Conclusion</u>

Pages

99	Bibliographie sélective
99	A) Bibliographie générale
102	B) Publications et notes de travail de l'équipe ORSTOM
106	Annexes
107	Annexe 1 : documents d'enquête pour l'enquête "morphologie"
108	A) Instructions aux enquêteurs
126	B) Codification de l'enquête morphologie
141	Annexe 2 : documents d'enquête pour l'enquête "bâti / population"
142	A) Instructions aux enquêteur
170	B) Codification de l'enquête bâti/population (version espagnole) ¹
190	Annexe 3 : Listage du programme VAR2D
198	Annexe 4 : Questionnaire de l'enquête finale "démographie / migration" (version espagnole) ²

¹ : Ce manuel, en cours de traduction, est donné ici dans sa version originale en espagnol pour permettre aux personnes intéressées de retrouver les codes. La version traduite figurera en annexe de la thèse qui doit conclure ce travail.

² : Même remarque que pour le manuel de codification de l'enquête bâti/population

INTRODUCTION

Ce rapport est constitué pour l'essentiel du troisième compte rendu intermédiaire du travail que effectue O. Barbary au sein de l'équipe ORSTOM travaillant sur le projet "Utilisation de la télédétection pour l'observation des populations urbaines"¹. Il fait suite au rapport intermédiaire rédigé en Janvier 88² à l'issue de l'expérience menée à Marseille, auquel on se rapportera pour situer ce travail dans le cadre général du projet et avoir une synthèse des objectifs poursuivis, des méthodes employées et des résultats obtenus. Rappelons simplement ici le but principal poursuivi. Il s'agit de mettre au point un plan de sondage aréolaire pour des enquêtes démographiques en milieu urbain, adapté à l'utilisation d'une image satellite stratifiée comme base de sondage. A la fin du second rapport, on définissait ainsi les principales questions qui restaient à résoudre, après le travail effectué à Marseille, pour être en mesure d'appliquer la méthode.

L'expérience menée à Marseille aura donc permis de préciser un certain nombre des caractéristiques que doit avoir un sondage à but démographique en milieu urbain utilisant comme base de sondage une image satellite stratifiée. Cependant, d'une part la description de ce plan de sondage reste encore incomplète et d'autre part il reste à vérifier la stabilité des résultats acquis à Marseille lorsqu'on change de site d'application. En conséquence, trois questions principales encore en suspens peuvent être formulées :

1. L'usage de l'image satellite comme base de sondage aréolaire et support de la sélection de l'échantillon implique (...) un plan de sondage où la probabilité d'appartenance d'une unité à l'échantillon est proportionnelle à la taille de cette unité. Reste à définir quelle est la mesure de cette taille la plus efficace au sens de la précision du sondage ; (...) **a-t-on intérêt, du point de vue de la précision des estimations produites, à choisir une probabilité de sélection proportionnelle à la surface bâtie des unités plutôt qu'à leur surface totale?** Dans l'hypothèse affirmative, existe-t-il une stratégie de sélection de ces unités praticable sur l'image satellite qui garantisse une telle probabilité?

2. Le plan de sondage partiellement défini à Marseille est applicable à des unités spatiales du type districts de recensement INSEE. Ces unités correspondent en gros aux pâtes de maison physiques identifiables sur l'image satellite. Or, l'unité statistique prise en compte lors des enquêtes démographiques est soit l'individu soit, plus

¹ : Etant donné l'étroite imbrication des recherches menées par les différents membres de cette équipe, ce troisième rapport ne pouvait se concevoir sans l'intégration de certains textes présentant les travaux et résultats des autres chercheurs participant au projet. Pour plus de détails, notamment sur la conception générale du programme de recherches, les travaux en télédétection urbaine et les opérations de collecte réalisées à Quito, le lecteur peut se reporter aux articles et rapports spécifiques déjà diffusés, cités dans la bibliographie de l'équipe, pp 103 à 106 du présent rapport.

² : O. Barbary. Deuxième rapport intermédiaire à l'ORSTOM : "Estimation de caractères démographiques à Marseille, évaluation des précisions permises par l'emploi de divers estimateurs, techniques d'échantillonnage et stratifications". Quito, Janvier 1988, 184 pages.

généralement, le ménage. Le plan de sondage complètement défini doit donc comprendre plusieurs degrés, le premier degré correspond à la sélection des pâtés de maison et le dernier à celle de l'unité statistique enquêtée, le ménage.

Il reste donc à fixer le nombre de degrés intermédiaires du sondage, les taux et les techniques de sélection à pratiquer pour chaque degré. Ces choix seront faits de manière à optimiser la précision du sondage pour un coût de mise en oeuvre donné c'est à dire pour un nombre de ménages à enquêter fixé à l'avance.

3. Quel degré de stabilité offrent les résultats acquis à Marseille lorsque l'on change de site ?

En partant de ce constat, nous définissons alors les objectifs et les étapes du travail à réaliser à Quito pour parvenir à l'application de la méthode.

1. Sélection d'un échantillon de "manzanas" (pâté de maisons à Quito) et collecte d'informations morphologiques et démographiques sur cet échantillon.

2. Application des outils statistiques et informatiques mis au point à Marseille à ce corpus de données pour fournir des réponses aux questions restant posées.

3. Définition complète du plan de sondage à mettre en oeuvre lors de l'enquête démographique finale à Quito et des techniques adéquates pour sa réalisation sur document satellitaire.

4. Sélection de l'échantillon et réalisation de l'opération de collecte, codification, saisie et contrôle des données collectées.

5. Développement du programme informatique permettant l'extrapolation des données de l'échantillon aux domaines d'estimations retenus : sous-strates, strates, ensemble de l'agglomération.

6. Exploitation de l'enquête, production d'estimations de caractères démographiques et d'indicateurs de migrations sur la ville de Quito, estimation de la précision des données produites, conclusion sur l'efficacité réelle de la méthode mise au point."

Les quatre premières étapes de ce programme sont maintenant achevées et le présent rapport est d'abord le compte rendu du travail effectué pour parvenir à la définition complète du plan de sondage retenu lors de l'enquête finale "démographie / migration". En même temps nous tentons de faire le point sur l'ensemble des caractéristiques de la méthode de sondage que nous appliquons pour la première fois et de fournir les éléments pratiques nécessaires à sa mise en oeuvre.

Le premier chapitre, après avoir détaillé les objectifs et le plan d'expérience du travail à Quito, décrit les opérations de collecte correspondant aux deux enquêtes préliminaires réalisées : l'enquête "morphologie" et l'enquête "bâti/population". Dans le second chapitre, on rend compte de la première exploitation des résultats de l'enquête "bâti population", dont le but est la définition de la stratégie de sondage à appliquer aux unités primaires : les pâtés de maisons. Puis, dans le chapitre trois, la même enquête est utilisée pour aborder le second degré du sondage : l'échantillonnage des ménages au sein des pâtés de maisons. Enfin le dernier chapitre présente les objectifs et le questionnaire de l'enquête finale, définit dans son ensemble le plan de sondage et décrit les différentes étapes de sa mise en oeuvre.

Il manque, pour qu'un éventuel utilisateur de la méthode ait tous les éléments pratiques en main, un chapitre sur les coûts et délais de réalisation de l'enquête que nous n'avons pas pu rédiger à temps et qui figurera donc dans le dernier rapport intermédiaire à l'ORSTOM en même temps que les conclusions finales sur l'efficacité de la méthode ; conclusions que doit fournir la phase finale d'exploitation des résultats.

CHAPITRE 1

OBJECTIFS GENERAUX, PLAN D'EXPERIENCE, REALISATION DES OPERATIONS DE COLLECTE DES ENQUETES PRELIMINAIRES

1.1. Objectifs généraux, plan d'expérience

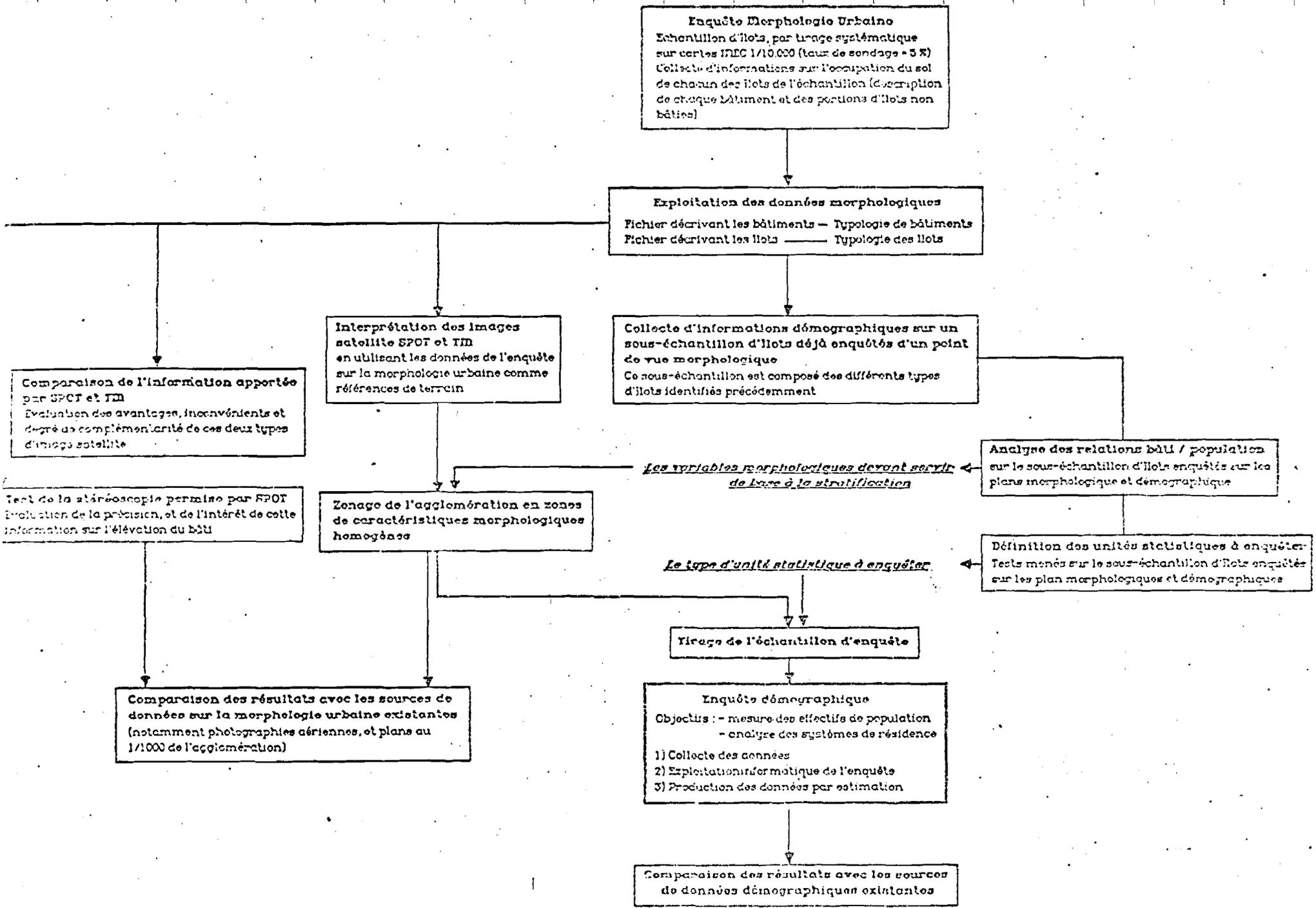
Les principales étapes du travail à Quito, décrites dans la Figure 1.1. correspondent à celles qui auraient lieu, dans une phase opérationnelle, pour toute application de la méthode :

- **Enquête sur la morphologie urbaine**, où sont collectées des informations précises sur l'occupation du sol dans un échantillon d'ilots de Quito, afin d'initialiser l'interprétation des images satellite devant servir de base de sondage à l'enquête démographique finale.
- **Analyse des relations bâti/population**, sur un sous-échantillon des ilots enquêtés du point de vue morphologique, afin de définir la stratégie de sondage à appliquer pour l'enquête finale.
- **Stratification sur image satellite et tirage de l'échantillon d'enquête**
- **Collecte démographique** : mesure des effectifs de population et analyse des systèmes de résidence
- **Exploitation de l'enquête**, production des données démographiques par estimation

Jusqu'à la production de données démographiques de l'enquête finale, nous n'utilisons que les documents de base susceptibles d'exister dans toute ville des pays en développement : image satellite, et plans de la ville, pas nécessairement actualisés. Les autres types d'information existant à Quito, mais généralement pas dans d'autres villes, ne seront consultés qu'à la fin de l'expérimentation, afin d'examiner la pertinence de nos résultats.

Cette vérification ne pourra d'ailleurs pas être très précise, étant donné le décalage de temps entre la dernière couverture aérienne (1984) et le dernier recensement (1982), et nos propres observations (1986-1987). Néanmoins, nous pourrions ainsi contrôler, certes de façon assez grossière :

- l'interprétation des images satellite, en utilisant comme référence les photographies aériennes,
- les estimations d'effectifs démographiques, par rapport à des projections sous forme de "fourchettes" des chiffres du recensement de 1982. Un deuxième passage, en Novembre 1987, sur l'échantillon de l'enquête Bâti/population nous sera d'une aide précieuse pour asseoir les projections des chiffres du recensement de 1982.



Dans la conception générale du programme de recherche, nous avons choisi de réaliser sur le site de Marseille les développements méthodologiques. Néanmoins, quelques points de méthode non abordés à Marseille font l'objet de travaux spécifiques à Quito :

- **la définition des unités statistiques à enquêter** : à Marseille, du fait que nous utilisons des données du recensement pour calculer la variance des estimations, nous sommes dans l'obligation de travailler sur les districts de recensement définis par l'INSEE, c'est à dire un découpage spatial résultant de critères administratifs, démographiques et physiques. La prise en considération de ces différents paramètres aboutit à un découpage spatial en unités de taille très variable, et ne correspondant pas systématiquement à l'îlot urbain au sens physique du terme, c'est à dire le pâté de maisons. D'autre part, les données étant agglomérées par district, elles ne permettent pas, comme on l'a déjà dit à l'issue de l'expérience marseillaise¹, d'évaluer les possibilités offertes par les sondages à plusieurs degrés. A Quito, les deux enquêtes préliminaires fournissent sur un échantillon d'îlots des informations morphologiques et démographiques disponibles, pour les premières au niveau de chaque bâtiment et pour les secondes pour chaque ménage. Ces données vont permettre de tester des procédures de sondage à plusieurs degrés qui, tout en restant praticables et en offrant de faibles risques d'erreur lors de la collecte, améliorent la précision des estimations démographiques.

- **La stéréoscopie sur image satellite en milieu urbain** : le satellite SPOT, du fait des variations possibles de l'angle de visée, permet la vision stéréoscopique. Nous testerons sur Quito l'intérêt de cette nouvelle information en télédétection spatiale, son degré de précision pour l'appréhension des hauteurs de bâti.

- **Comparaison de l'information apportée par SPOT et par TM** : disposant de ces deux types d'image à des dates très voisines, respectivement août et septembre 1986, nous pouvons évaluer les avantages et inconvénients de chacune de ces sources d'information et leur degré de complémentarité.

Dans les chapitres suivants, c'est le premier point qui sera développé ; outre cet aspect méthodologique, le travail préparatoire réalisé à Quito avant l'enquête finale a aussi pour objectif la vérification de la stabilité des principaux résultats obtenus à Marseille.

1.2 Réalisation de l'enquête morphologie

1.2.1 Sélection de l'échantillon

Pour la sélection de l'échantillon de l'enquête morphologie, comme pour toute l'expérimentation menée à Quito, nous n'utilisons que les documents de base susceptibles d'exister dans toute ville d'un pays en développement. C'est donc sur un plan de Quito non actualisé que nous avons réalisé la sélection de l'échantillon : les plans au 1/10.000 établis par l'INEC en 1982 pour le recensement général de la population.

La technique d'échantillonnage a été des plus simples ; nous avons opté pour un tirage systématique sur grille plaquée sur les plans INEC. Ayant fixé le taux de sondage à 4 %, connaissant approximativement le nombre total d'îlots de Quito et la surface totale occupée par la ville, nous en avons déduit la taille de la maille de la grille : 750 mètres de côté.

¹ : Voir à ce propos la conclusion du second rapport intermédiaire à l'ORSTOM : Mise au point de plans de sondage pour l'estimation de caractères démographiques à Marseille, (O. Barbary) Quito, Janvier 1988, pp 132 & 133.

Nous avons donc tracé sur calque un carroyage de 7,5 cms de côté que nous avons superposé au plans INEC au 1/10.000. Pour chaque intersection de la grille, on incluait dans l'échantillon l'îlot qui contenait cette intersection.

Dans le cas où une intersection de la grille tombait sur une rue, une règle systématique était appliquée pour déterminer un îlot voisin à inclure dans l'échantillon ; si nous n'avions pas procédé ainsi, le taux de sondage aurait été inférieur à 4 %.

Avec ce mode de tirage, on peut considérer que la probabilité pour un îlot d'appartenir à l'échantillon est proportionnelle à sa surface ; de ce fait, certains très grands îlots ont été tirés plusieurs fois (ils contenaient plusieurs intersections de la grille). Nous avons pratiqué un tirage "sans remise", en reportant les intersections appartenant à un même îlot sur des manzanas voisines¹.

Nous avons ainsi constitué un échantillon de 195 îlots. Les cartes de l'INEC, datant de 1982, ne couvrent pas les extensions de la ville de 1982 à 1986 ; il a donc fallu compléter notre échantillon tiré systématiquement sur document cartographique par quelques îlots de la périphérie dont le type morphologique n'était pas représenté dans l'échantillon initial. Cet ajout de 4 îlots s'est fait en fin d'enquête, une fois bien connues par l'équipe les caractéristiques morphologiques de l'échantillon enquêté et celles des extensions récentes de Quito.

Au total, notre échantillon compte donc 199 îlots.

1.2.2 Elaboration des documents d'enquête

Comme l'indique le manuel d'instructions aux enquêteurs², chaque îlot de l'échantillon est décrit par trois types de documents :

- **Carte** : sur le fond de plan INEC agrandi au 1/500, sur lequel ne figurent au départ de l'enquête que les limites de chaque îlot à enquêter, les enquêteurs établissent la cartographie de l'occupation du sol de l'îlot. Tous les bâtiments ainsi que tous les éléments non bâtis doivent figurer sur ces plans.

- **Fiche de description des bâtiments** : chacun des édifices dessinés sur la carte est décrit sur cette fiche. Il est repéré par un numéro de 1 à n.

- **Photographies** : pour chaque îlot, les chercheurs de l'Orstom font des photographies depuis un point surplombant l'îlot (étage élevé d'un immeuble ou relief), et au sol, en bordure et/ou à l'intérieur de l'îlot. Ces photographies sont destinées à préciser et contrôler l'information recueillie par les enquêteurs.

La connaissance du bâti et du non bâti acquise lors de diverses sorties sur le terrain effectuées à Quito nous a permis de modifier en l'adaptant la fiche d'enquête conçue et expérimentée à Marseille ; les conseils de collègues de l'équipe Atlas connaissant bien la ville nous ont aidé à préciser les termes à employer pour caractériser les différents matériaux de construction utilisés à Quito et à adapter les fiches en fonction des possibilités et des conditions d'enquête propres à ce site. Une relecture de ces fiches a ensuite été faite par un architecte équatorien.

¹ : De ce fait, la probabilité d'appartenance n'est plus réellement contrôlée (c'est d'ailleurs toujours le cas dans un sondage systématique), mais cela importe peu puisque l'objectif de l'opération n'est pas d'effectuer une inférence rigoureuse sur cet échantillon.

² : Voir annexe 1.

Nous avons testé avec les enquêteurs recrutés pour l'enquête le caractère opérationnel de ces fiches à l'occasion d'une première sortie commune sur le terrain. La méthode de collecte a été décrite avec précision dans le manuel d'instructions aux enquêteurs, afin qu'ils puissent s'y référer en cas d'ambiguïté.

1.2.3 Collecte des données

La procédure de collecte des données étant décrite précisément dans les points 3 et 4 du manuel d'instructions aux enquêteurs, nous ne donnerons ici que quelques précisions sur les temps et conditions de collecte.

Temps de collecte

La collecte des données (remplissage des questionnaires et établissement des schémas sur les plans au 1/500) a été réalisée par deux enquêteurs entre le 1er Octobre et le 9 Décembre 1986 ; la moyenne d'un îlot par demi-journée, estimée au début de l'enquête d'après notre expérience marseillaise, s'est donc vérifiée.

Quant aux photographies prises depuis des immeubles ou des points hauts de la ville, elles ont nécessité environ 20 matinées de travail à deux personnes ; 10/12 îlots par demi-journée semble le maximum qui puisse être fait, compte tenu des temps de déplacement et de discussion pour avoir accès aux sommets des immeubles.

Conditions de collecte

Les enquêteurs n'ont pas eu de difficulté à collecter l'information : ce type d'enquête ne réclame que peu de participation de la population, puisqu'elle ne nécessite que l'autorisation d'accéder aux terrasses des maisons ou immeubles. Pour tous les îlots, cette autorisation a pu être obtenue sans problème majeur, quelque soit le type de quartier de Quito.

1.2.4 Codification, saisie, contrôles de cohérence et recodage

Codification

La codification, réalisée par F. DUREAU sur des fiches du type de celle présentée en annexe 1, a été établie à partir de toutes les données collectées au cours de l'enquête : questionnaires, plans au 1/500, photographies. Ces photographies se sont révélées fondamentales pour préciser l'information collectée par les enquêteurs : ce sont des documents qu'il faut s'efforcer de recueillir lors de toute collecte de données-terrain en milieu urbain. Outre leur rôle de contrôle du travail des enquêteurs, ils apportent l'information qualitative complémentaire des données quantitatives des questionnaires simplifiés que nous utilisons.

Le principe général de la fiche de codification est le suivant : constituer deux fichiers, l'un décrivant les bâtiments, l'autre les "manzanas" (c'est à dire pâtés de maisons). Le fichier des bâtiments ne décrit pas les édifices un par un, mais par groupes présentant des caractéristiques morphologiques semblables ou très voisines ; ce sont ces groupes que nous appelons "genres".

Saisie

Les fiches de codification ont été saisies sous DBASEIII, sur l'IBM AT, en une dizaine d'heures à deux personnes.

A l'issue de cette phase de saisie, nous disposons de deux fichiers :

QMANZANA : fichier descriptif des 199 manzanas enquêtées (recto de la fiche de codification)
 champs 1-3 : identification manzana
 champ 4 : identification enquêteur
 champs 5-7 : mesure de surface
 champs 8-10 : typologies (sources exogènes, estimations enquêteur et codificateur)
 champs 11 - 42 : description de l'espace bâti

QBATIMEN : fichier descriptif des 2134 genres de bâtiments (verso de la fiche de codification)
 champ 1 : identification manzana
 champ 2 : identification du genre dans la manzana
 champ 3 : nombre de bâtiments du genre dans la manzana
 champs 4 à 13 : caractérisation du genre.

Contrôles de cohérence : redressement des mesures de bâtiments

L'analyse des tableaux de fréquence et histogrammes nous ont permis de corriger quelques erreurs de codification ou de saisie.

Le point principal sur lequel ont porté les corrections a été la surface des bâtiments. Il avait été demandé aux deux enquêteurs de mesurer les dimensions au sol des bâtiments en comptant leurs pas, après avoir étalonné ceux-ci en effectuant leurs premières mesures avec un décimètre. La comparaison des histogrammes des surfaces des bâtiments de chacun des enquêteurs, sur un échantillon de bâtiments censés avoir sensiblement les mêmes dimensions, nous a montré que l'un des deux enquêteurs avait sous-estimé les surfaces d'environ 20 % par rapport à l'autre enquêteur (ce dernier ayant une bonne appréciation des mesures). Nous avons donc corrigé le fichier des bâtiments en augmentant de 20 % toutes les surfaces mesurées par l'enquêteur n° 1. Les surfaces des bâtiments ont ensuite été réparties en 10 modalités.

Recodages

Au vu des histogrammes, deux autres variables ont été seuillées : les surfaces de bâtiments (7 classes) et le nombre de pans du toit (5 classes).

1.2.5. Exploitation de l'enquête morphologie en vue du tirage d'un échantillon représentatif

Afin de pouvoir sélectionner un échantillon représentatif pour l'enquête sur les relations bâti/population, il est nécessaire de définir une typologie des bâtiments identifiés et décrits lors de l'enquête morphologie, à partir de laquelle on peut caractériser l'espace bâti des manzanas et classer ces dernières.

La figure 1.2 retrace les différentes étapes de la classification des bâtiments, puis des manzanas; les figures 1.3 et 1.4 précisent le contenu des fichiers soumis aux traitements.

La classification des bâtiments en 38 types (décrits dans le tableau 1.5) est obtenue par segmentation sur 3 variables : la surface du sol, le nombre d'étages, et le matériau du toit. Les bornes des deux premières variables ont été définies d'après les résultats d'une classification automatique réalisée, avec le logiciel SPSS, sur les 4595

bâtiments regroupés en 352 "super-genres", caractérisés par leur surface au sol (10 modalités) et leur nombre d'étages (7 modalités).

La classification des manzanas est basée sur le tableau de contingence caractérisant l'espace bâti de chaque manzana, en fonction des 38 types définis précédemment : la case courante $K_{i,j}$ du tableau contient le pourcentage de surface bâtie dans l'ilot i occupée par les bâtiments de type j .

En combinant les résultats de la classification ascendante hiérarchique réalisée sur ce tableau de contingence (17 classes), la densité du bâti, et un indicateur de l'homogénéité des bâtiments au sein de chaque ilot, on obtient une classification finale en 9 classes, décrites dans le tableau 1.6.

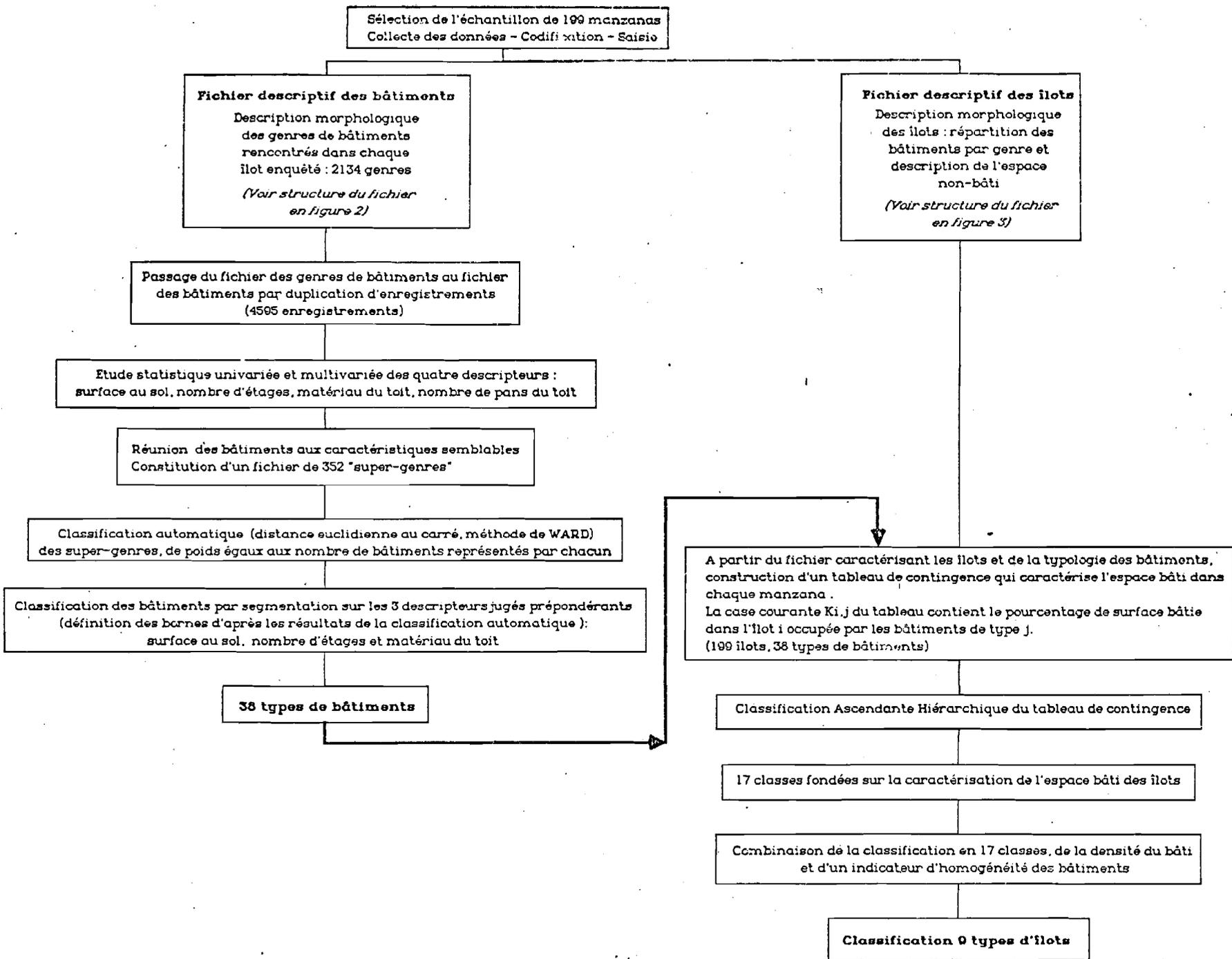


Figure 1.2 - Schéma de la démarche de classification

Figure 1.3 - Structure du fichier descriptif des bâtiments

Structure du fichier : a:qbatiment.dbf				
Nombre d'enregistrements: 2134				
Dernière mise à jour : 31/03/87				
Champ	Nom champ	Type	Dim	Dec
1	NOMANZANA	Numérique	3	
2	GENRE	Numérique	2	
3	NBATI	Numérique	2	
4	MATERIAU	Numérique	1	
5	FORMEMAT	Numérique	1	
6	COULEUR	Numérique	2	
7	NBPANS	Numérique	2	
8	NBETACES	Numérique	2	
9	SURFACEBAT	Numérique	5	
10	SURFBATC	Numérique	5	
11	SURGENRE	Numérique	7	
12	TYPEBATI	Numérique	2	
13	NOMENQ	Numérique	1	
14	SGENRE	Numérique	5	
** Total **			41	

Figure 1.4 - Structure du fichier des flots

Structure du fichier : a:qmanzana.dbf
 Nombre d'enregistrements: 199
 Dernière mise à jour : 03/04/87

Champ	Nom champ	Type	Dim	Dec
1	NOMANZANA	Caractère	3	
2	NOINEC	Caractère	7	
3	PARTINEC	Caractère	1	
4	ENQETEUR	Caractère	1	
5	SURFILOT	Numérique	6	
6	SURF1_2VOI	Numérique	6	
7	SURFTOTALE	Numérique	6	
8	TYPECAZA	Caractère	2	
9	TYPEENQUET	Caractère	2	
10	TYPEPHOTO	Caractère	2	
11	SURFBATI	Numérique	6	
12	DENSBATI	Caractère	1	
13	IMPLANTBAT	Caractère	2	
14	NBBATITOTA	Numérique	3	
15	NBBATIFINI	Numérique	3	
16	NBBATICONS	Numérique	3	

Appuyez sur une touche pour continuer...

17	NBBATI1	Numérique	2	
18	NBBATI2	Numérique	2	
19	NBBATI3	Numérique	2	
20	NBBATI4	Numérique	2	
21	NBBATI5	Numérique	2	
22	NBBATI6	Numérique	2	
23	NBBATI7	Numérique	2	
24	NBBATI8	Numérique	2	
25	NBBATI9	Numérique	2	
26	NBBATI10	Numérique	2	
27	NBBATI11	Numérique	2	
28	NBBATI12	Numérique	2	
29	NBBATI13	Numérique	2	
30	NBBATI14	Numérique	2	
31	NBBATI15	Numérique	2	
32	NBBATI16	Numérique	2	

Appuyez sur une touche pour continuer...

33	NBBATI17	Numérique	2	
34	PENTETOIT	Caractère	1	
35	NBBATIPENT	Numérique	3	
36	NBBATIFER	Numérique	3	
37	NBBATICHEM	Numérique	3	
38	VEGEBALCON	Caractère	1	
39	DIRPENTDOM	Caractère	1	
40	NIVPENTDOM	Caractère	1	
41	DIRPENTSDO	Caractère	1	
42	NIVPENTSDO	Caractère	1	
43	TYPEEAU	Caractère	1	
44	SURFEAU	Numérique	6	
45	VEGEHBTYDO	Caractère	1	
46	VEGEHBTYSD	Caractère	1	
47	VEGEHBSURF	Numérique	6	
48	VEGECBTYDO	Caractère	1	

Appuyez sur une touche pour continuer

49	VEGECBTYSD	Caractère	1	
50	VEGECBSURF	Numérique	6	
51	VOIRIETYDO	Caractère	1	
52	VOIRIETYSD	Caractère	1	
53	SURFVOIRIE	Numérique	6	
54	MURDSILOT	Caractère	1	
55	MURHRSILCT	Numérique	3	
56	TROTTOIRDO	Caractère	1	
57	TROTTOIRSD	Caractère	1	
58	CHAUSSEEDO	Caractère	1	
59	CHAUSSEESD	Caractère	1	
60	PELOUSE	Numérique	3	
61	ARBRES	Numérique	3	
62	ARBUSTES	Numérique	3	
63	VOITURESTA	Numérique	3	
64	VOITURECIR	Numérique	3	

Appuyez sur une touche pour continuer

65	TYPE17	Numérique	2	
** Total **			160	

Tableau 1.5 - Typologie des bâtiments

Surface au sol	Nombre D'étages	Béton ou Eternit	Tuiles vernies, Carrelage, Shingle	Zinc	Tuiles non vernies
1 - 29 m ²	1 ét.	1	1	1	1
30 - 79 m ²	1 ét.	3	12	21	30
80 - 224 m ²	1 ét.	4	13	22	31
225 - 14999 m ²	1, 2 ét.	5	14	23	32
1 - 29 m ²	2, 3 ét.	6	15	24	33
15 - 79 m ²	2, 3 ét.	7	16	25	34
80 - 224 m ²	2, 3 ét.	8	17	26	35
225 - 14999 m ²	2, 3 ét.	9	18	27	36
30 - 79 m ²	4, 5 ét.	10	19	28	37
80 - 224 m ²	4, 5 ét.	11	20	29	38
125 - 14999 m ²	5 à 15 ét.	2	2	2	2

(Dans chaque case du tableau, est noté le numéro du type de bâtiment, selon la surface au sol, le nombre d'étages et le matériau du toit)

Tableau 1.6 - Typologie des îlots

Type	Intitulé du	Description morphologique du type type	Nombre d'îlots
1	Non bâti	Densité du bâti nulle	6
2	Péri-urbain	Densité du bâti très faible (bâtiments petits ou moyens, tous matériaux sauf tuiles vernies)	59
3	Habitat populaire récent (1 : à étages)	Bâtiments de taille petite ou moyenne, Rdc +1 ou 2 étages , toiture en béton ou éternit, peu de tuiles vernies ou non (densité du bâti faible à très forte)	19
4	Habitat populaire récent (2 : sans étage)	Bâtiments de taille petite ou moyenne, Rdc seul très dominant, toiture en béton, éternit, zinc, peu de tuiles vernies ou non (densité du bâti faible à très forte)	20
5	Vieux centre	Bâtiments taille moyenne ou grande, Rdc + 1 ou 2 ét. , % important de toitures en tuiles non vernies (densité du bâti faible à très forte)	29
6	Habitat de standing	Bâtiments taille moyenne (80, 225m ²), Rdc, Rdc +1 ou 2 ét. , % important de toiture en carrelage ou tuiles vernies , pas de zinc (densité du bâti moyenne)	22
7	Industries et équipements	Très grands bâtiments bas , à toiture en béton, éternit ou zinc (densité du bâti faible ou moyenne)	21
8	Immeubles	Grands bâtiments élevés , à toiture surtout en béton et éternit (densité du bâti faible ou moyenne)	12
9	Viviendas	Densité du bâti forte ou très forte , très grande homogénéité des bâtiments (bas et de grande taille)	11
TOTAL			199

1.3. Réalisation de l'enquête "Relations bâti/population"

1.3.1 Sélection de l'échantillon

Comme il l'a été vu dans le chapitre précédent, les 199 îlots et les 2134 genres de bâtiments enquêtés lors de l'enquête morphologie ont fait l'objet d'une classification en, respectivement 9 types d'îlots et 38 types de bâtiments.

Etant donné les objectifs assignés à l'enquête bâti/population, trois critères ont servi à sélectionner les îlots de l'échantillon, parmi les 199 déjà enquêtés sur le plan morphologique :

- le type de bâtiments
- le type d'îlot
- la localisation au sein de la ville de Quito.

Cette sélection s'est faite par choix raisonné dans la base des 199 îlots, stratifiée par type d'îlot. Le tirage a été réalisé à partir du listing décrivant la composition par types de bâtiments des 199 îlots, eux-mêmes classés par types d'îlot, et des cartes de localisation des îlots de l'enquête morphologie. Le nombre d'îlots à tirer dans chaque type a été défini par l'application d'un taux sensiblement constant (environ 30 %) ; pour chaque type d'îlot, nous nous sommes efforcés de choisir des îlots à la fois bien représentatifs des caractéristiques de leur strate et offrant la plus grande variété de types de bâtiments et de localisations dans Quito.

Type d'îlot	Nombre d'îlots Enq. morphologie	Nombre d'îlots Enq. bâti/population	Numéros des îlots Enq. bâti/population
1 (non bâti)	6	0	-
2	59	15	158,166,256,164,216, 146,120,142,326,362, 361,319,251,148,352,
3	19	6	147,214,228,204,345 263
4	20	7	208,245,110,329,318, 342,230
5	29	8	203,121,237,249,223, 232,156,248
6	22	6	111,266,311,324,246 347
7	21	4	130,217,304,344
8	12	4	212,250,262,307
9	11	4	143,257,258,322
TOTAL	199 (196 bâtis)	54	

1.3.2 Elaboration des documents d'enquête

Pour collecter les données de l'enquête bâti/population, chaque enquêteur devait remplir deux types de documents (voir manuel d'instructions aux enquêteurs en annexe 2) :

- la fiche logement, pour décrire les personnes résidant dans les logements des bâtiments faisant partie de l'ilot ;
- la fiche récapitulative, pour rassembler les informations générales relatives aux logements appartenant à l'ilot.

Les données socio-démographiques devant être mises en relation avec les données morphologiques de l'enquête précédente, au niveau du bâtiment comme au niveau de la manzana, il importait qu'il y ait concordance entre les deux enquêtes sur l'identification des manzanas et des bâtiments au sein de chaque manzana ; c'est pourquoi chaque enquêteur recevait des copies des questionnaires et plans de l'enquête morphologie, et était tenu de conserver la même délimitation et la même numérotation de bâtiments.

Les variables démographiques figurant sur les fiches logement ont été choisies par rapport à l'objectif de l'enquête finale du dernier trimestre 1987 (estimation d'effectifs et analyse des systèmes résidentiels) : outre les caractéristiques démographiques classiques (âge, sexe ...etc), l'accent a été mis sur la saisie de la résidence, le statut d'occupation du logement et les revenus.

La formulation des questions a été soumise à l'avis de différentes personnes ayant une expérience solide de la collecte des données dans les villes équatoriennes.

1.3.3 Collecte des données

La procédure de collecte étant décrite précisément dans les points 3 et 4 du manuel d'instructions aux enquêteurs (annexe 2), nous ne donnerons ici que quelques précisions sur les temps et conditions de collecte.

Temps de collecte

La collecte a été réalisée par 5 enquêteurs, encadrés par un superviseur, entre le 11 Avril et le 8 Mai (compte tenu des jours fériés, 3 semaines effectives de travail pour 4 des enquêteurs et le superviseur ; et 2 semaines de travail pour le dernier enquêteur).

Environ 1700 logements ont pu être enquêtés durant cette période, soit une moyenne de 17 questionnaires remplis par enquêteur et par jour ; en tenant compte des interdictions d'accès à certains bâtiments qu'ont subies les enquêteurs, le temps d'enquête a donc été plus long que ce que nous avons estimé a priori (25 logements par jour et par enquêteur).

Conditions de collecte.

Cette seconde enquête s'est déroulée dans un contexte bien particulier, lié au séisme qu'a connu l'Equateur le 5 Mars 1987. Du fait de la rupture de l'oléoduc (le pétrole est la principale source de revenus de l'Etat Equatorien), l'économie nationale s'est trouvée complètement désorganisée durant la période qui a suivi le tremblement de terre : augmentation immédiate du prix du carburant (multiplié par 2), rationnements de carburant, augmentation des prix pour de nombreux produits de

base...etc. C'est donc dans un contexte économique et social très tendu que nous avons réalisé notre enquête.

Contrairement à l'enquête morphologie, l'enquête bâti/population supposait une participation de la population pour remplir le questionnaire logement. Le taux de non-réponse a fortement varié selon les îlots.

Un premier constat, bien que très classique, mérite d'être rappelé : le degré de participation à l'enquête est inversement proportionnelle aux revenus des habitants. Dans les quartiers populaires de Quito de nombreux îlots ont été enquêtés à 100 % et les taux de non-réponses se situent toujours en dessous de 10 % (nombre de logements non enquêtés/nombre de logements total de l'îlot) : ces non-réponses ne correspondent pas à des refus, mais à des logements toujours trouvés vides malgré 4 ou 5 visites de l'enquêteur, leurs occupants s'absentant longuement de leur domicile pour aller travailler souvent très loin.

Par contre, dans les quartiers de moyen et haut standing les taux de non-réponses (là, il s'agit essentiellement de refus de réponse) sont beaucoup plus élevés 15 à 30 % le plus souvent. Nous avons même dû renoncer à l'enquête dans un îlot de standing très élevé : malgré une demi-journée de discussion et avec lettres officielles à l'appui, nous nous sommes fait interdire l'accès des immeubles par les administrateurs de quasiment toutes les tours constituant l'îlot d'enquête, et, dans la seule tour où nous avons pu accéder, ce sont les résidents eux-mêmes qui ont claqué la porte au nez de l'enquêtrice....

Il faut d'autre part noter, dans chaque îlot, une certaine homogénéité dans les comportements des résidents vis à vis des enquêteurs (en rapport, bien évidemment, avec l'homogénéité de la population elle-même) d'où des taux de non-réponses très tranchés : 10 % dans un îlot "sans problème", et 70 % dans un autre îlot peu éloigné du premier, mais ayant la particularité, par exemple, d'être majoritairement occupé par des militaires et policiers à la retraite.

Ces taux de non-réponses élevés ne nous posent pas de problèmes majeurs étant donné le type d'exploitation des données que nous réalisons pour cette enquête bâti/population. Par contre, s'il en était de même pour l'enquête finale, ce serait beaucoup plus fâcheux. L'on ne peut certainement pas mettre ces taux de non-réponse au compte des enquêteurs : d'une part, des passages supplémentaires que nous avons effectués nous-même ont confirmé l'efficacité de leur travail ; d'autre part, sur des îlots réputés difficiles à enquêter, et ayant même posé les mêmes problèmes aux enquêteurs du recensement (alors que les gens sont obligés de rester chez eux durant le recensement, et que le certificat de réponse au recensement est indispensable pour nombre de démarches administratives), nos propres enquêteurs ont eu des taux de réponse satisfaisants. Il nous paraît donc certain que le climat tendu qui régnait à Quito au moment de l'enquête a joué contre les enquêteurs ; le contexte social lors de l'enquête morphologie était nettement plus favorable.

Qualité de la collecte

Les premiers contrôles des questionnaires faits dès leur remise par les enquêteurs, et ceux réalisés au cours de la codification ont montré une bonne cohérence des données collectées.

Un problème particulier, apparu au cours de la collecte, mérite toutefois d'être signalé, étant donné l'objectif de notre enquête : il s'agit de la difficulté qu'ont eue les enquêteurs dans certains quartiers construits en continu comme le centre historique de Quito, à identifier les bâtiments repérés par leurs collègues de l'enquête morphologie. Pour la première enquête, c'est une définition physique du bâtiment, vu du dessus qui primait : dans un dédale de couloirs serpentant à l'intérieur de vieux bâtiments accolés, où rien ne traduit la limite des différents bâtiments, il est bien évidemment difficile de savoir dans quel bâtiment précis l'on se trouve, sous quelle toiture ! Dans ce cas précis, nous avons donc pris le parti

d'affecter la population enquêtée à un groupe de bâtiments, sans chercher à identifier précisément si tel logement appartenait à un bâtiment ou à son voisin, qui a, dans la quasi-totalité des cas, exactement les mêmes caractéristiques morphologiques.

Ces difficultés à identifier les bâtiments ne se sont produites que dans le cas du vieux centre de Quito ; dans tous les autres types d'implantation spatiale du bâti, il n'y a pas eu de problèmes à retrouver les bâtiments figurant sur les plans au 1/500 établis lors de l'enquête morphologie.

1.3.4 Codification et saisie

Deux types de codification ont été réalisées, générant deux fichiers :

- codification des caractéristiques individuelles notées sur les fiches logement -----> fichier où chaque enregistrement correspond à un individu enquêté (fichier QDEMINDI, comptant près de 5000 enregistrements) ;
- codification des caractéristiques des logements et des bâtiments notées sur les fiches logement et sur les fiches récapitulatives -----> fichier où chaque enregistrement correspond à un logement enquêté (fichier QDEMVIIV, comptant environ 1700 enregistrements).

Cette structuration des données en deux fichiers permet de ne pas dupliquer les informations relatives au logement pour chaque individu résidant dans celui-ci ; le numéro de logement existant dans les deux fichiers, toutes les données peuvent être mises en relation par le système de gestion de données DBASEIII, sans qu'il y ait redondance de l'information.

La codification des fiches logement a été réalisée par un étudiant, qui a entamé son travail une semaine après le début de l'enquête ; il a été aidé durant sa dernière semaine de travail par le superviseur. Ce dernier, plus qualifié, a été chargé de la codification des questions sur l'activité et les revenus. De la sorte, la codification a été terminée seulement une dizaine de jours après la fin de l'enquête.

La codification des caractéristiques individuelles a été réalisée directement sur la fiche d'enquête, où des cases avaient été prévues pour cela ; cette solution permet d'éviter le report sur une fiche de codification spécifique des données recueillies directement sous forme de codes, comme l'âge, la date de naissance, le statut de résidence...etc. Cette procédure diminue nettement les temps de codification ; elle avait été choisie parce que initialement, nous pensions effectuer la saisie nous-mêmes à Quito, sur l'IBM AT avec le logiciel DBASEIII. Pour des raisons de calendrier, nous avons finalement confié cette saisie à une société de service informatique en France, où nous étions en mission durant tout le mois de Juin 1987 : la saisie directe sur les fiches d'enquête est nettement plus complexe et longue pour les opérateurs de saisie. Lorsque les travaux de saisie sont sous-traités à des sociétés spécialisées où le travail est effectué très "mécaniquement", il nous semble bien préférable d'opter pour la codification sur bordereau spécifique.

La codification des données relatives aux logements et bâtiments, très rapide, a été réalisée sur bordereau (voir modèle en annexe 2).

CHAPITRE 2

ENQUETE PRELIMINAIRE "BATI POPULATION", EXPLOITATION POUR LA MISE AU POINT DU PREMIER DEGRE DU SONDAGE : ECHANTILLONNAGE DE PATES DE MAISONS

2.1 Introduction : Le contexte et les questions posées

La manière dont on a abordé le problème de la définition d'une stratégie de sondage à appliquer aux unités primaires à Quito résulte d'un compromis : explorer un éventail de méthodes le plus large possible tout en restant dans le cadre des contraintes fixées au départ du projet : c'est à dire d'une part l'utilisation de l'image satellite comme base de sondage et d'autre part le caractère rapide et peu coûteux des solutions recherchées. Plutôt que de dresser d'emblée la liste des conséquences qu'ont ces contraintes sur les objectifs que nous nous fixons, nous préférons les évoquer à chaque fois qu'elles interviennent dans une question posée, une réponse apportée ou une décision prise.

A propos de la sélection de l'échantillon des pâtés de maisons et de la démarche d'inférence qui doit suivre, quatre questions principales se posent.

1. Quelle probabilité doit avoir chaque unité d'appartenir à l'échantillon ?

L'éventail des possibilités est en fait assez restreint. Un sondage équiprobable est impraticable puisqu'on ne dispose d'aucune liste des unités et que les techniques d'interprétation de l'image ne permettent pas encore de dresser une telle liste du moins avec la rapidité souhaitée. On retiendra pourtant cette alternative dans la suite comme base de comparaison et également parce qu'il n'est pas exclu qu'elle puisse être praticable dans l'avenir.

L'idée qui vient ensuite "naturellement" à l'esprit, s'agissant d'un sondage aréolaire qui s'applique à des unités de taille variable, est celle d'une probabilité d'appartenance à l'échantillon proportionnelle à la surface de ces unités. C'est elle qui s'applique au cas d'une sélection aléatoire de points dans l'espace géographique de la base de sondage (on retient alors dans l'échantillon l'ensemble des pâtés de maisons "touchés" par ces points) et, par une approximation admise, au cas du tirage systématique à l'aide d'une grille de points.

Enfin une troisième possibilité est de rechercher une stratégie de sélection des pâtés de maisons qui fournisse une probabilité d'appartenance proportionnelle à une meilleure mesure de la taille des unités, c'est à dire à une mesure plus corrélée aux caractères démographiques que la surface totale. Il s'agirait par exemple d'une mesure approchée de la surface bâtie ou, mieux encore, de la surface habitable.

On ne peut pas encore obtenir par interprétation d'image satellite, une mesure, même approchée, de la surface habitable; par contre des classifications de l'image existent qui fournissent un découpage de l'agglomération en zones homogènes du point de vue de leur densité de bâti ou de leur densité de végétation. A partir de ces classifications et de l'information obtenue lors de l'enquête "morphologie", on peut, par étalonnage, obtenir une valeur approchée de la densité de bâti à l'intérieur de chaque classe et donc une approximation de la surface bâtie de chaque pâté de maisons. Mais la

réalisation d'un sondage suivant une telle probabilité de sélection pose un certain nombre de problèmes; les difficultés majeures sont les suivantes.

A: Dans l'étape de sélection de l'échantillon

Pour respecter lors d'un tirage systématique à l'aide d'une grille de points une probabilité d'appartenance approximativement proportionnelle à la surface bâtie de l'unité primaire, on doit penser à une stratégie qui n'accepte un pâté de maisons dans l'échantillon que lorsque le point de la grille qu'il contient est situé au dessus d'un bâtiment; dans le cas contraire l'unité est rejetée. Une telle stratégie complique considérablement la gestion du taux de sondage à travers la taille de la maille de la grille. **Mais surtout elle nécessite que le document servant de base au tirage permette de décider si un point donné est situé au dessus d'un bâtiment ou au contraire inclu dans une surface non bâtie**; à la limite, pour être totalement rigoureux, il faudrait être en mesure d'identifier chaque bâtiment du pâté de maisons. **La résolution spatiale actuelle des images satellite est bien évidemment incompatible avec une telle contrainte.** La probabilité d'appartenance qui résulterait de l'application d'une stratégie de ce type sur l'image satellite ne serait donc au mieux que très grossièrement proportionnelle à la surface bâtie des unités et n'améliorerait sans doute pas les estimations.

B. Dans l'étape d'inférence

Pour calculer les estimateurs sans biais associés à ce plan de sondage il faut connaître, pour chaque unité de l'échantillon et pour chaque domaine d'extrapolation, leur surface bâtie. La mesure de cette surface, qu'on devrait tirer de l'image satellite en ce qui concerne les domaines d'extrapolation (univers entier, strates et éventuellement sous-strates), ne serait bien sûr qu'approximative ce qui pourrait avoir de fâcheuses conséquences sur la précision des estimations puisqu'il s'agit des facteurs d'extrapolation. **Enfin cette même mesure au niveau des pâtés de maisons de l'échantillon, ne peut pas être effectuée sur l'image satellite et sa collecte sur le terrain, également sujette à erreurs, allongerait énormément les temps d'enquête.**

Pour l'ensemble de ces raisons, nous n'envisageons pas de retenir pour son application à Quito un tel plan de sondage, cependant il nous semble important pour les développements futurs d'en évaluer dès maintenant la précision.

Le test qui va suivre concernera donc trois plans :

1. Plan équiprobable
2. Plan à probabilité proportionnelle à la surface des unités
3. Plan à probabilité proportionnelle à la surface bâtie des unités.

Pour répondre complètement à la question posée, il est bien entendu indispensable de comparer également les trois plans dans le cadre stratifié. En effet, le problème du choix d'une probabilité d'appartenance est, on le verra, très lié à la stratification employée.

2. Quels estimateurs employer pour l'inférence de l'échantillon de pâtés de maisons au total des strates ?

Il s'agit ici de confirmer les résultats obtenus lors de l'expérience sur Marseille concernant les comparaisons entre estimation sans biais et estimation par le ratio intégrant la surface des unités, qui s'étaient conclues à l'avantage de la première¹.

3. Quel gain de précision permet la stratification de la base de sondage à partir de l'image satellite?

Au départ de notre recherche, cette question était posée différemment : il s'agissait de comparer entre elles plusieurs stratifications issues de l'image correspondant chacune à des critères ou associations de critères différents. Malheureusement, étant donné le calendrier de la phase d'application, nous ne disposions au moment du test que d'une seule classification de l'image correspondant au critère de densité du bâti. C'est donc à l'évaluation de cette stratification qu'on s'attachera et à sa validation par comparaison à une stratification sur le même critère obtenue à partir de l'enquête préliminaire "morphologie". L'évaluation d'autres stratifications reste donc à mener et c'est d'ailleurs à notre avis une des directions prometteuses à explorer dans des recherches futures.

4. Y a-t-il des possibilités d'optimisation de l'allocation aux strates?

A t-on intérêt à faire varier le taux de sondage selon les strates de densité de bâti pour améliorer la précision des estimations sur l'ensemble des strates? La réponse à cette question dépend évidemment de l'ensemble de variables qu'on considère dans l'enquête. Pour chacune d'elles on peut essayer, à partir des données de l'enquête préliminaire "bâti/population", d'obtenir une valeur approchée de la variance du caractère sur chacune des strates et, partant, de définir une "bonne" allocation. Les différentes allocations ainsi définies seront elles compatibles entre elles ou, au contraire contradictoires? Dans la première hypothèse on pourrait adopter, en ce qui concerne la stratégie d'allocation de l'échantillon aux strates, un moyen terme qui serait alors profitable à l'ensemble des estimations.

Les problèmes étant posés, les paragraphes suivants sont consacrés aux données et aux traitements statistiques auxquels elles ont été soumises, puis à l'examen des différents résultats obtenus. Enfin, en conclusion, on tentera d'apporter des réponses aux questions formulées ci dessus.

2.2 Les données et les outils statistiques utilisés

2.2.1 Les données

Puisqu'il s'agit dans cette étape d'examiner les unités primaires du sondage, c'est à dire les pâtés de maisons, il a fallu constituer en partant du fichier ménage un fichier pâté de maisons contenant les informations nécessaires aux tests envisagés. Comme à Marseille, mais cette fois ci sur l'échantillon de pâtés de maisons enquêtés, on souhaite disposer d'informations fiables concernant d'une part un certain nombre de caractéristiques démographiques moyennes des ménages du pâté de maisons et d'autre part la localisation, la taille et la morphologie de chaque îlot. La liste de ces informations est donnée dans le tableau 2.1. A propos de ces données, il y a plusieurs remarques à faire.

Données démographiques

¹ : Voir à ce sujet le deuxième rapport intermédiaire à l'Orstom : "Mise au point de plan de sondage pour l'estimation de caractères démographiques à Marseille, évaluation des précisions permises par l'emploi de divers estimateurs, techniques d'échantillonnage et stratifications". Quito, Janvier 1988, 184 Pages

On l'a vu, durant la phase de collecte, le taux de non réponses à tout ou partie du questionnaire a été très variable suivant le quartier enquêté. Des données trop partielles risquant d'ôter toute signification aux calculs de moyennes, la première étape de la constitution du fichier est la sélection, pour chaque variable, du sous ensemble de l'échantillon dans lequel l'information est suffisante pour fournir une moyenne significative ; connaissant pour tous les pâtés de maisons considérés le nombre de ménages y résidant, on a pu, lorsque, pour une variable donnée, l'information était disponible pour la grande majorité d'entre eux, étendre la donnée à l'ensemble de l'ilot. Il résulte de cette étape que, comme dans le fichier ménage, le domaine où la donnée est connue peut être différent suivant la variable considérée. Dans le tableau 2.1, on indique le nombre de pâtés de maisons pour lesquelles chaque variable est soit totalement connue, soit extrapolée avec une fiabilité acceptable.

Une fois définis ces domaines de validité, le calcul des indicateurs démographiques sur les pâtés de maisons se fait par sommation à partir du fichier ménage. Cette opération a été grandement facilitée par les possibilités de gestion relationnelle des données offerte par dbaseIII. On trouvera en annexe les programmes mis au point pour produire le fichier manzana et le listage des données correspondantes.

Signalons que si les données obtenues au niveau des ilots peuvent être entachées de certaines inexactitudes dues aux non réponses, nous avons néanmoins pu vérifier à partir des chiffres du recensement de 1982 qu'au niveau de l'ensemble de l'échantillon les valeurs obtenues pour un certain nombre d'indicateurs étaient en parfaite cohérence avec la donnée "réelle". Nous pensons donc que le degré de précision des informations qui vont servir aux tests est suffisant pour que leurs résultats soient fiables.

Enfin le lecteur remarquera qu'une partie seulement des traitements s'applique à la totalité des variables citées dans le tableau 2.1. Pour les tests menés avec le programme Varech, les variables "AGECHEF", "REVENU" et "DURESMIG" ont été supprimées d'une part pour ne pas allonger l'étape de traitement et respecter le calendrier prévu pour la réalisation de l'enquête et d'autre part parce que, lors de l'introduction des stratifications, nous aurions disposé pour ces variables d'un nombre insuffisant de données dans certaines strates.

Taille des unités primaires

On considérera dans les analyses trois mesures de taille des pâtés de maisons :

1. La surface totale de l'unité
2. La surface bâtie de l'unité (exclus la voirie interne, les jardins, les terrains vagues avec ou sans végétation, les parkings ou toute autre surface ne comportant pas de bâtiments)
3. Le complémentaire de la surface en végétation (inclus la voirie interne, les terrains vagues sans végétation, les parkings ou toute autre surface ne contenant pas de végétation). Ces données ne sont pas issues de l'interprétation de l'image satellite mais proviennent de l'enquête préliminaire "morphologie" réalisée sur l'échantillon. Le niveau de précision de la mesure de la surface totale nous paraît en accord avec celui des mesures praticables sur l'image, par contre la donnée concernant la surface bâtie est probablement plus précise que son équivalent obtenu d'après l'information satellitaire. Pour pallier à cette différence, nous avons considéré une troisième mesure : obtenue par soustraction de la surface en végétation à la surface totale, elle est sans doute plus proche de l'approximation de la surface bâtie qu'on pourrait tirer de l'indice de végétation calculé sur l'image.

Stratifications

Dans ce travail, on ne considère qu'un seul critère de stratification qui est la densité de bâti, pourtant on s'intéresse à trois stratifications différentes construites toutes les trois sur ce critère. La première est issue des observations réalisées lors de l'enquête "morphologie" et la variable lui correspondant sera appelée DENSBATI. La seconde, DENSBATIM, est la stratification de la base de sondage spatiale obtenue par

classification de l'image satellite SPOT de juin 1986¹. Quant à la troisième, DENSBATIME, elle correspond aux résultats des corrections faites sur la seconde. Nous avons en effet jugé utile de "redresser" cette seconde stratification pour deux raisons.

D'une part, pour des impératifs de calendrier, la stratification de l'image dont nous disposons (DENSBATIM) est une classification des pixels et non pas la classification d'îlots obtenue en utilisant la "valeur moyenne" de la classification des pixels de l'îlot. Les travaux de l'équipe télédétection ayant montré que cette dernière classification est nettement préférable, nous avons pris le parti de corriger la stratification dans ce sens en changeant certains îlots de strate.

D'autre part, pour alléger la collecte à réaliser dans l'enquête préliminaire démographique, lorsque qu'on pouvait identifier sans ambiguïtés plusieurs parties au sein d'un même pâté de maisons, on n'a parfois enquêté qu'une seule de ces parties. Il peut alors se produire que la densité de bâti caractérisant la partie enquêtée soit significativement différente de la densité moyenne du pâté de maisons. L'unité réellement enquêtée se voit alors affectée, si l'on s'en tient à la seconde des stratifications, à une strate qui ne correspond pas réellement à sa densité de bâti. Nous avons donc dans ce cas réaffecté ces unités aux strates présentant la même densité moyenne de bâti qu'elles.

Qu'elles aient été effectuées pour la première ou la seconde raison, ces réaffectations se font dans l'immense majorité des cas à la strate immédiatement plus dense ou moins dense que la strate d'origine (DENSBATIM).

¹ : Voir à ce sujet le chapitre 4, paragraphe 4.3.2.

Tableau 2.1 : Liste des champs du fichier "pâté de maisons".

Champ	Nom du champ	Définition	Nombre d'îlots où la donnée existe
1) Identification			
1	NOMA	Numéro de l'îlot (manzana)	53
2) Localisation, taille, morphologie			
2	SUDNORD	Latitude	53
3	ESTOUEST	Longitude	53
4	LOC	Appartenance à une des douze zones géographiques définies dans Quito.	53
5	SURFACE	Surface de la manzana	53
6	DENSBAT	Pourcentage de la surface bâtie dans la manzana, d'après l'enquête "morphologie"	53
7	DENSBATI	Classification en cinq classes de densité de bâti : résultat de l'enquête "morphologie"	53
8	DENSBATIM	Classification en cinq classes de densité de bâti : résultat du traitement de l'image SPOT	53
9	DENSBATIME	Classification en cinq classes de densité de bâti obtenue après "redressement" de la classification DENSBATIM	53
10	DENSVEG	Densité de végétation d'après l'enquête "morphologie"	53
11	SURFBAT	Surface bâtie, mesure approchée par le pourcentage de surface bâtie.	53
12	CSURFVEG	Complémentaire à la surface en végétation : mesure approchée par l'indice de végétation tiré de l'image satellite.	53

Tableau 2.1 (suite) : Liste des champs du fichier "pâté de maisons".

Champ	Nom du champ	Définition	Nombre d'îlots où la donnée existe
3) Démographie			
13	EFFECTRES	Population résidente ¹	46
14	NBMENAGE	Nombre de ménages	53
15	NBMEN1	Nombre de ménages d'une personne	46
16	NBMEN7	Nombre de ménages de sept personnes et plus	46
17	NBMENP1	Nombre de ménages de plus d'une personne (complémentaire de NBMEN1 ²)	46
18	NBMENM7	Nombre de ménages de moins de sept personnes (complémentaire de NBMEN7 ³)	46
19	EFFM5A	Population âgée de moins de cinq ans	44
20	EFFMIG	Population de migrants	44
21	EFFMIGM5	Population de migrants ayant moins de cinq ans de résidence à Quito	44
22	EFFECTMOY	Effectif moyen des ménages	46
23	AGECHEF	Age moyen des chefs de ménages (en années)	45
24	PCMOINS5	Proportion de la population de moins de cinq ans dans la population totale (en %)	44
25	PCMIG	Proportion de la population de migrants dans la population totale (en %)	44
26	PCMIGM5	Proportion de la population de migrants ayant moins de cinq ans de résidence à Quito dans la population totale (en %)	44
27	DURESMIG	Durée moyenne de résidence à Quito des migrants (en mois)	41
28	REVENU	Revenu moyen des résidents de l'îlot (en millier de sucres)	38
29	PCMENAG1	Proportion des ménages d'une personne dans la totalité des ménages (en %)	46
30	PCMEN7	Proportion des ménages de sept personnes et plus dans la totalité des ménages (en %)	46

¹ : Pour les définitions des différentes catégories de populations considérées ici, ménages, population résidente, population migrante, voir le manuel d'instruction au enquêteurs en annexe 2.

² : NBMENP1 = NBMENAGE - NBMEN1

³ : NBMENM7 = NBMENAGE - NBMEN7

2.2.2 Les outils statistiques utilisés

Pour apporter des solutions aux problèmes posés, on a procédé aux traitements sur le fichier des pâtés de maisons en deux temps.

1. Phase exploratoire : corrélations, régressions linéaires, analyses de la variance.

Dans cette étape, en identifiant et quantifiant à l'aide d'outils statistiques simples les principales liaisons existant entre caractères démographiques à estimer et mesures de la taille des unités primaires, on poursuit un double objectif. D'une part on veut écarter, pour gagner du temps dans la suite de l'expérience, l'étude de solutions à priori envisageables mais que l'absence de liaisons statistiquement significatives font apparaître illusoire. On pourra par exemple conclure dès cette étape à l'inutilité d'intégrer les informations sur la taille des unités primaires dans des estimateurs par le ratio. D'autre part, on cherche à formuler des hypothèses concernant l'intérêt d'adopter une probabilité de sélection des unités de l'échantillon proportionnelle aux diverses mesures possibles de la taille des pâtés de maisons. On passera ensuite, dans la seconde phase, à la vérification de ces hypothèses.

2. Phase conclusive : test des plans de sondage à l'aide du programme Varech mis au point au cours de l'expérience menée à Marseille¹.

Le programme Varech permet dans cette étape d'évaluer les conséquences qu'auront sur la précision des estimations les différents choix possibles quant aux probabilités d'inclusion dans l'échantillon et aux stratifications. Pour cela on considère l'échantillon des pâtés de maisons enquêtés comme une base de sondage et le programme calcule la variance des estimateurs sur cette base. Afin de pouvoir comparer entre elles l'ensemble des variances calculées, nous avons systématiquement appliqué un même taux de sondage ($2/3 = 66,6\%$) ; ce taux, totalement arbitraire, a été choisi pour limiter l'ordre de grandeur des coefficients de variation. Pour vérifier les hypothèses formulées dans la phase exploratoire, les traitements effectués à l'aide du programme se sont déroulés en trois temps. Tout d'abord on a comparé les effets, sur la précision des estimations, des différentes probabilités d'appartenance à l'échantillon qu'on peut envisager. Ces effets varient évidemment suivant le fait que la base de sondage est ou non stratifiée, l'expérience est répétée dans les deux cas. Puis on a évalué les gains de précision permis par les trois stratifications. Enfin on s'est intéressé aux allocations optimales pour chacun des caractères afin d'essayer de dégager une règle d'allocation améliorant l'ensemble des estimations.

¹ : A propos du programme VARECH, son domaine d'application, le formulaire mis en œuvre et le fonctionnement du programme, voir le chapitre 4 du second rapport intermédiaire à l'ORSTOM : Mise au point de plans de sondage pour l'estimation de caractères démographiques à Marseille, O. Barbary, Janvier 1988, Quito (184 p, voir pp 72 à 82 et 140 à 184).

2.3 : Résultats

2.3.1 Phase exploratoire

Tout d'abord on s'est intéressé aux corrélations existant entre indicateurs démographiques.

Tableau 2.2 : Matrice des corrélations entre variables démographiques. (logiciel SPSS)

COEFFICIENTS *	1	2	3	4	5	6	7	8	9	10	11
1 EFFECTRES	1	.99	.74	.76	.93	.99	.87	-.28	.23	-.10	.10
2 NBMENAGE	.99	1	.79	.69	.96	.99	.90	-.32	.24	-.08	.11
3 NBMEN1	.74	.79	1	.46	.77	.89	.74	-.48	.48	-.08	.27
4 NBMEN7	.76	.69	.43	1	.54	.72	.52	.03	.23	-.25	.10
5 EFFM5A	.93	.96	.77	.54	1	.94	.93	-.34	.19	.08	.11
6 EFFMIG	.99	.99	.80	.72	.94	1	.91	-.28	.27	-.09	.21
7 EFFMIGM5	.87	.90	.74	.52	.93	.91	1	-.15	.25	-.06	.37
8 EFFECTMOY	-.28	-.32	-.48	.03	-.34	-.28	-.15	1	-.11	-.19	.28
9 AGECHF	.23	.24	.48	.23	.19	.27	.25	-.11	1	-.48	.42
10 PCMOINS5	-.10	-.08	-.08	-.25	.08	-.09	-.06	-.19	-.48	1	-.23
11 PCMIG	.10	.11	.27	.10	.11	.21	.37	.28	.42	-.23	1
12 PCMIGM5	-.08	-.08	-.05	-.06	-.07	-.02	.26	.62	.11	-.20	.72
13 DURESMIG	.08	.07	.05	.13	.02	.02	-.22	-.29	-.03	-.01	-.65
14 REVENU	.56	.58	.45	.32	.59	.58	.62	-.22	.32	-.14	.22
15 PCMENAG1	.15	.19	.68	.07	.14	.21	.14	-.48	.50	-.18	.28
16 PCMEN7	-.20	-.23	-.26	.21	-.25	-.19	-.11	.79	.16	-.32	.32

COEFFICIENTS	12	13	14	15	16
1 EFFECTRES	-.08	.08	.56	.15	-.20
2 NBMENAGE	-.08	.07	.58	.19	-.23
3 NBMEN1	-.05	.05	.45	.68	-.26
4 NBMEN7	-.06	.12	.32	.07	.21
5 EFFM5A	-.07	.02	.60	.14	-.25
6 EFFMIG	-.02	.02	.58	.21	-.19
7 EFFMIGM5	.26	-.22	.62	.14	-.11
8 EFFECTMOY	.62	-.29	-.22	-.48	.79
9 AGECHF	.11	-.03	.32	.50	.16
10 PCMOINS5	-.20	-.01	-.14	-.18	-.32
11 PCMIG	.72	-.65	.22	.28	.32
12 PCMIGM5	1	-.65	.04	-.10	.42
13 DURESMIG	-.65	1	-.11	.07	-.21
14 REVENU	.04	-.11	1	.10	-.16
15 PCMENAG1	-.10	.07	.10	1	-.18
16 PCMEN7	.42	-.21	-.16	-.18	1

* : En gras, les coefficients supérieurs à 0,70.

De l'examen du tableau 2.2 se dégagent deux groupes de caractères.

Premièrement les effectifs qui forment un ensemble au sein duquel les valeurs des coefficients sont généralement élevées et toutes positives. En particulier les trois variables "Effectif de résidents", "Nombre de ménages", "Effectif des migrants" sont fortement corrélées entre elles et à tous les autres effectifs considérés. Un tel phénomène de bloc, qui n'est d'ailleurs pas surprenant, autorise à espérer, pour la suite des traitements, un comportement semblable de l'ensemble des effectifs.

Le second groupe est formé par les variables de type quotient pour lesquelles les corrélations sont en général médiocres, à deux exceptions près qui reflètent des liaisons évidentes : corrélation élevée entre l'effectif moyen du ménage et la proportion de ménage de plus de sept personnes et entre la proportion de migrants et la proportion de migrants ayant moins de cinq ans de résidence à Quito. Après ce qui vient d'être dit, on peut penser qu'il sera relativement facile de formuler, à l'issue des tests qui vont suivre, des conclusions valables pour l'estimation de l'ensemble des effectifs alors qu'à l'inverse il risque fort de ne pas se dégager de règles applicables à l'ensemble des quotients. Comme entre les deux groupes, il n'existe aucune corrélation linéaire significative, il faudra prendre garde qu'une stratégie de sondage améliorant l'estimation des effectifs ne détériore pas au contraire celle des quotients.

S'agissant de l'étude des différentes mesures de la taille des unités primaires qui, selon la stratégie de sélection de l'échantillon retenue, définissent la probabilité d'inclusion, il nous semble naturel d'examiner d'abord les corrélations entre ces mesures et les caractères démographiques. En effet, notre sondage à probabilités inégales sera bien sûr d'autant plus efficace que la variable définissant la probabilité sera mieux corrélée aux variables à estimer.

Le tableau 2.3 donne les coefficients de corrélation linéaire existant entre les indicateurs démographiques et les trois mesures de la taille des pâtés de maisons retenues.

L'examen des résultats conduit aux conclusions suivantes :

1. Il n'y a pas de corrélations significatives entre les caractères démographiques et la surface totale des unités. Il faut donc s'attendre à ce que l'introduction d'une probabilité proportionnelle à la surface des pâtés de maisons n'améliore pas la précision des estimations.

2. La plupart des effectifs, à l'exception du nombre de ménage de sept personnes et plus, sont corrélés positivement avec la surface bâtie. Les coefficients sont compris entre 0,55 et 0,61 ; de telles valeurs, si elles paraissent insuffisantes pour justifier l'emploi d'estimateurs par le ratio, devraient permettre en revanche un gain appréciable de précision si on retient une probabilité d'appartenance proportionnelle à la surface bâtie. Dans le cas des quotients au contraire, les coefficients de corrélation sont faibles et varient de -0,35 à 0,50, ce qui laisse peu d'espoir d'améliorer les estimations en adoptant une telle probabilité.

3. La troisième mesure, le complémentaire de la surface en végétation, apparaît bien, au regard des coefficients de corrélation avec les indicateurs démographiques, comme une approximation de la surface bâtie. Les corrélations, systématiquement inférieures à celles qu'on observe dans le cas de la surface bâtie, doivent néanmoins être suffisantes pour que l'introduction de cette mesure comme déterminant de la probabilité améliore la précision du sondage.

Tableau 2.3 : coefficients de corrélation entre variables démographiques et mesures de taille des unités primaires du sondage. (logiciel SPSS)

	(1) SURFACE	(2) SURFBAT	(3) CSURFVEG	
	Surface totale de l'ilot	Surface bâtie de l'ilot	Compiément à la surface en végétation	
1) Totaux				
1	EFFECTRES	0,068	0,572	0,442
2	NBMENAGE	0,075	0,603	0,453
3	NBMEN1	0,042	0,610	0,420
4	NBMEN7	-0,070	0,308	0,239
5	EFFM5A	0,071	0,608	0,405
6	EFFMIG	0,060	0,581	0,437
7	EFFMIGM5	0,018	0,548	0,365
2) Quotients				
8	EFFECTMOY	-0,055	-0,349	-0,236
9	AGECHEF	-0,192	0,150	0,023
10	PCMOINS5	-0,012	-0,107	-0,183
11	PCMIG	-0,034	0,147	0,095
12	PCMIGM5	-0,154	-0,155	-0,112
13	DURESMIG	0,216	0,144	0,118
14	REVENU	-0,146	0,380	0,112
15	PCMENAG1	0,173	0,504	0,396
18	PCMEN7	-0,165	-0,218	-0,196

Enfin, à propos de ces résultats, il faut garder en mémoire une des leçons de l'expérience marseillaise et considérer avec prudence les chiffres concernant les effectifs ou ratios qui décrivent les catégories de population très minoritaires. Comme à Marseille, il est probable que le passage aux quantités complémentaires améliore beaucoup les résultats. Dans la phase conclusive de l'expérience on retiendra donc également ces quantités pour évaluer avec quelles précisions on peut les estimer.

Quels modèles explicatifs des caractères démographiques peut-on maintenant construire à partir des informations relatives à la taille des unités primaires?

Il serait vain, étant donné le degré de corrélation observé ci dessus, de considérer la surface totale des pâtés de maisons comme facteur explicatif. En revanche il peut être intéressant de voir quel type de modèle rend le mieux compte des corrélations existant entre indicateurs démographiques et surface bâtie. On comparera ici trois modèles linéaires :

1. Régression entre effectifs et surface bâtie
2. Régression entre densités et densité de bâti
3. Analyse de la variance testant l'effet du classement des observations en six niveaux de densité de bâti.

Le tableau 2.4 résume les résultats obtenus lorsqu'on applique successivement ces trois modèles aux sept variables démographiques.

Tableau 2.4 : Résultats résumés des régressions et analyses de la variance entre caractères démographiques (effectifs ou densité) et surface bâtie ou densité du bâti. (Logiciel SPSS)

	Régressions entre effectifs et surface bâtie		Régression entre densités et densité de bâti		Analyse de la variance à un critère : densité de bâti en 6 cl.
	Coeff. de reg.	% de varian. expl.	Coeff. de reg.	% de varian. expl.	% de variance expliquée
1 EFFECTRES	0,52	26,7	0,67	44,6	52,6
2 NBMENAGE	0,55	30,2	0,67	45,0	53,8
3 NBMEN1	0,52	26,7	0,45	19,8	26,8
4 NBMEN7	0,21	4,2	0,42	17,3	20,9
5 EFFMOINS5A	0,53	27,8	0,66	43,4	51,0
6 EFFMIG	0,52	27,1	0,62	37,8	55,2
7 EFFMIGM5	0,47	21,8	0,37	13,7	29,3

Tout d'abord il faut dire que les parts de variance expliquées par chacun des modèles sont toutes trop faibles pour qu'on puisse "prédire" la valeur prise par le caractère démographique à partir de celle du facteur explicatif. Cela confirme qu'il y a peu d'avantage à utiliser des estimateurs par le ratio intégrant ces mesures de taille (il est même probable que la précision des estimations en souffre). Mais l'expérience ne perd pas pour autant tout intérêt. En effet on constate que les trois modèles n'ont pas des performances équivalentes : d'une part, les régressions entre densité d'une catégorie de population ou de ménages et densité de bâti sont généralement nettement meilleures que celles entre l'effectif correspondant et la surface bâtie (sauf pour les catégories très minoritaires comme les ménages d'une personne ou les migrants ayant moins de cinq années de résidence à Quito, variables pour lesquelles il faudrait étudier les quantités complémentaires), d'autre part, le modèle d'analyse de la variance s'ajuste toujours mieux aux données que les régressions. Ces constatations conduisent à avancer l'hypothèse suivante concernant la stratégie de sondage à adopter et qu'il s'agira de vérifier dans la phase conclusive de l'expérience : plutôt que de chercher une stratégie à mettre en pratique qui garantisse un tirage à probabilité approximativement proportionnelle à la surface bâtie des unités primaires, ce qui de toute façon sera difficile étant donné l'information disponible sur l'image base de sondage, n'a-t-on pas intérêt à pratiquer, à l'intérieur de strates offrant une densité de bâti homogène, un tirage avec probabilité proportionnelle à la surface totale des pâtes de maisons qui est par contre facilement réalisable?

Il nous faut maintenant vérifier, à partir du calcul sur l'échantillon de la variance des estimateurs qu'effectuera le programme Varech, le bien fondé de l'ensemble des hypothèses formulées dans la phase exploratoire.

2.3.2 Phase conclusive

Probabilités d'appartenance des unités primaires à l'échantillon

Le tableau 2.5 donne les coefficients de variation calculés par le programme dans le cas non stratifié et dans le cas d'une stratification de la base de sondage en cinq classes de densité de bâti (variable DENSBATI). Pour le cas non stratifié on a retenu quatre définitions de la probabilité d'appartenance :

1. Probabilité égale pour chaque unité (colonne 1)
2. Probabilité proportionnelle à la surface totale de l'unité (colonne 2)
3. Probabilité proportionnelle à la surface bâtie de l'unité (colonne 3)
4. Probabilité proportionnelle au complémentaire de la surface en végétation (colonne 4).

Pour le cas stratifié seules les trois premières définitions ont été appliquées, les colonnes paires donnant le coefficient obtenu avec l'allocation proportionnelle et les colonnes impaires ceux correspondants à l'allocation optimale :

1. Equiprobabilité (colonnes 5 et 6)
2. Probabilité proportionnelle à la surface totale (colonnes 7 et 8)
3. Probabilité proportionnelle à la surface bâtie (colonnes 9 et 10).

Enfin les colonnes 11 et 12 mesurent le gain dû à la stratification sous le plan à probabilités proportionnelles aux surfaces totales des unités, en pourcentage du coefficient de variation non stratifié (colonne 11) et par le facteur de division de la variance (colonne 12).

L'observation du tableau fournit quatre résultats principaux.

1. Comparé à la référence équiprobable et dans le cadre non stratifié, un sondage à probabilités proportionnelles aux surfaces totales des unités primaires (sondage aréolaire aléatoire non stratifié) ne présente pas d'intérêt ; au contraire, dans la plupart des cas il augmente l'erreur d'échantillonnage (voir colonnes 1 et 2). On trouve ici confirmation d'un des résultats importants obtenus à Marseille.

Tableau 2.5 : Coefficients de variation des estimations des variables démographiques avec différentes définitions des probabilités d'appartenance à l'échantillon, cas non stratifié et cas stratifié.

(Programme Varech)

	Non stratifié				Stratifié sur la densité du bâti (DENSBATI)						Gains dus à la stratif.	
	1 Equi.	2 P-su	3 P-sb	4 P-csv	5 Equi. alloc. prop.	6 Equi. alloc. opti.	7 P-su alloc. prop.	8 P-su alloc. opti.	9 P-sb alloc. prop.	10 P-sb alloc. opti.	11 ¹ Gain (8-2) 2 en %	12 ² Gain (2/8) ²
1) Totaux												
1 EFFECTRES	24,91	25,42	15,97	17,40	22,46	15,92	13,28	12,30	13,73	12,23	-51,6	4,3
2 NBMENAGES	26,72	26,26	15,58	17,75	24,15	16,95	13,27	12,19	13,03	11,59	-53,6	4,7
3 NBMEN1	34,10	40,58	24,03	27,99	31,40	27,39	26,25	21,79	24,49	21,19	-46,3	3,5
4 NBMENP1	26,79	26,25	15,83	17,88	24,19	16,88	13,36	12,33	13,15	11,70	-54,0	4,5
5 NBMEN7	20,55	28,87	23,92	24,90	19,61	18,29	23,10	21,91	23,41	21,94	-24,1	1,7
6 NBMENM7	28,17	27,14	15,92	18,33	25,50	17,74	13,63	12,40	13,18	11,64	-54,3	4,8
7 EFFM5A	25,53	26,54	16,11	18,42	23,24	17,37	14,50	13,63	12,66	12,25	-48,6	3,8
8 EFFP5A	25,34	26,45	16,64	18,10	22,87	16,14	13,92	12,77	14,57	12,85	-51,7	4,3
9 EFFMIG	28,31	29,62	17,80	19,79	25,05	17,23	13,84	12,79	14,53	12,96	-56,8	5,3
10 EFFNMIG	23,61	25,02	16,60	16,60	21,61	16,08	14,58	13,60	14,82	13,48	-45,6	3,4
11 EFFMIGM5	26,02	30,50	29,42	21,55	23,38	19,18	20,03	19,23	21,76	20,38	-37,0	2,5
12 CEFFMIGM5	25,38	26,42	16,61	18,12	22,95	16,29	14,06	12,92	14,53	12,83	-51,1	4,2

Equi. : sélection équiprobable des unités primaires dans l'échantillon.

P=su : probabilité d'appartenance à l'échantillon proportionnelle à la surface totale des unités.

P=sb : probabilité d'appartenance à l'échantillon proportionnelle à la surface bâtie des unités.

P=csv : probabilité d'appartenance à l'échantillon proportionnelle au complémentaire de la surface en végétation des unités.

1 : Gain en % du coefficient de variation non stratifié

2 : Facteur de division de la variance (par rapport à la variance non stratifiée)

Tableau 2.5 (suite) : Coefficients de variation des estimations des variables démographiques avec différentes définitions des probabilités d'appartenance à l'échantillon, cas non stratifié et cas stratifié. (Programme Varech)

	Non stratifié				Stratifié sur la densité du bâti (DENSBATI)						Gains dus à la stratif.	
	1 Equi.	2 P-su	3 P-sb	4 P-csv	5 Equi. alloc. prop.	6 Equi. alloc. opti.	7 P-su alloc. prop.	8 P-su alloc. opti.	9 P-sb alloc. prop.	10 P-sb alloc. opti.	11 Gain (8-2) 2 en %	12 Gain (2/8) ²
2) Quotients												
13 EFFECTRES / NB MENAGE	3,71	4,07	3,08	3,40	3,35	2,68	3,15	2,83	3,13	2,81	-30,5	2,1
14 EFFECTRES / SURFACE	29,92	25,42	23,77	23,48	15,64	13,96	13,28	12,30	13,79	12,49	-51,6	4,3
15 NB MEN1 / NB MENAGE	22,34	30,09	21,57	23,21	20,11	13,08	22,51	18,80	20,94	17,19	-37,5	2,6
16 NB MENP1 / NB MENAGE	1,14	1,54	1,10	1,19	1,03	0,93	1,15	0,96	1,07	0,92	-37,7	2,6
17 NB MEN7 / NB MENAGE	21,46	23,60	19,92	21,59	17,86	16,52	19,71	18,06	19,29	18,19	-23,5	1,6
18 NB MENM7 / NB MENAGE	2,18	2,40	2,02	2,19	1,81	1,68	2,0	1,89	1,96	1,85	-21,2	1,6
19 EFFP5A / EFFECTRES	9,88	10,91	9,57	9,66	9,69	8,02	9,09	8,29	9,37	8,57	-24,0	1,7
20 EFFP5A / EFFECTRES	1,02	1,12	0,99	0,95	1,0	0,83	0,94	0,85	0,97	0,88	-30,2	1,7
21 EFFMIG / EFFECTRES	5,55	7,58	6,33	6,23	3,99	3,49	5,10	4,73	5,06	4,75	-37,6	2,6
22 EFFNMIG / EFFECTRES	3,54	4,84	4,04	3,98	2,55	2,23	3,25	3,02	3,23	3,03	-37,6	2,6
23 EFFMIGM5 / EFFECTRES	12,95	18,07	24,31	14,88	13,17	12,17	17,19	16,17	19,12	17,58	-10,5	1,3
24 CEFFMIGM5 / EFFECTRES	1,06	1,48	1,99	1,22	1,08	0,99	1,41	1,32	1,52	1,44	-10,8	1,3

Equi. : sélection équiprobable des unités primaires dans l'échantillon.

P-su : probabilité d'appartenance à l'échantillon proportionnelle à la surface totale des unités.

P-sb : probabilité d'appartenance à l'échantillon proportionnelle à la surface bâtie des unités.

P-csv : probabilité d'appartenance à l'échantillon proportionnelle au complémentaire de la surface en végétation des unités.

2. Une autre hypothèse faite lors de l'expérience marseillaise se vérifie : s'il était possible de gérer correctement une probabilité de sélection proportionnelle à la surface bâtie, ou du moins à son approximation d'après l'image satellite, l'amélioration du sondage serait en revanche notable, du moins en ce qui concerne les effectifs : 30 à 40 % de réduction du coefficient de variation dans le cas de la mesure de la surface bâtie, 20 à 30 % s'il s'agit de l'approximation par la surface en végétation (voir colonnes 1,3 et 4 de la première page du tableau). Pour les quotients en revanche l'avantage ne va systématiquement à aucune des deux définitions de la probabilité.

3. Quand on s'en tient au tirage équiprobable et à l'allocation proportionnelle, la stratification en six classes sur le critère de densité du bâti n'apporte qu'un gain assez faible. Par contre l'allocation optimale permet une diminution sensible de la variance des estimations sur l'ensemble de la base (colonnes 1,5 et 6).

4. L'observation comparée des colonnes 1, 2 et 5 à 10 permet d'affirmer qu'un **sondage à probabilités proportionnelles à la surface totale des unités, une fois stratifié la base de sondage selon la densité du bâti, est la solution qui s'impose**. Premièrement, il est la plupart du temps meilleur que le sondage équiprobable stratifié et a fortiori non stratifié (colonnes 7 et 8 comparées aux colonnes 1,5 et 6). **Deuxièmement avec un tel plan de sondage le gain induit par la stratification, une fois optimisée l'allocation aux strates est important**, puisqu'il s'agit d'une diminution des coefficients de variation comprise entre 25 et 55% ce qui représente une **division de la variance par un facteur compris entre 2 et 5** (voir colonnes 2 et 8 ou, plus commodément, 11 et 12). Enfin la comparaison des colonnes 7 et 8 avec les colonnes 9 et 10 permet de conclure à l'inutilité de rechercher plus avant un mode de tirage qui garantisse une probabilité approximativement proportionnelle aux surfaces bâties ; cette recherche se heurte d'ailleurs comme on l'a vu à de nombreuses difficultés pratiques. On constate en effet, en examinant conjointement les colonnes 8 et 10, qu'une fois retenu le principe de la stratification sur la densité du bâti, le tirage avec probabilités proportionnelles aux surfaces bâties n'améliore que très peu, voire pas du tout, l'estimation par rapport au tirage à probabilités proportionnelles aux surfaces totales. Le résultat apparaît d'ailleurs intuitivement satisfaisant : à l'intérieur d'une même classe de densité de bâti, surface totale et surface bâtie des unités primaires sont grosso modo proportionnelles et donc présentent les mêmes corrélations avec les caractères démographiques. On peut également considérer que ce résultat valide d'une certaine manière la stratification employée : les six classes de densité du bâti ont un bon niveau d'homogénéité.

Stratifications

Le tableau 2.6 permet de comparer en termes d'amélioration de la précision du sondage, les trois stratifications construites sur l'échantillon, toutes trois fondées, comme on l'a vu, avec des degrés de précision différents, sur un critère de densité du bâti. Il indique tout d'abord, comme référence pour l'appréciation des gains, les coefficients de variation des estimations non stratifiées (colonnes 1 et 2), puis (colonnes 3 et 4) ceux obtenus à partir de la stratification de référence issue de l'exploitation de l'enquête "morphologie" (DENSBATI), sous plan de sondage à probabilité proportionnelle à la surface totale des unités (P=su). La comparaison avec les stratifications établies d'après l'image satellite, stratification "brute" ou "corrigée" (respectivement DENSBATIM et DENSBATIME), doit être faite dans le cadre de ce même plan de sondage (P=su) à l'aide des colonnes 5 et 6 ou 9 et 10. Les colonnes 7,8,11 et 12 donnent, pour les deux nouvelles stratifications, les coefficients obtenus sous plan à probabilité proportionnelle à la surface bâtie des unités (P=sb) ; elles permettent de vérifier que les conclusions de l'expérience précédente sur la probabilité d'appartenance ne sont pas affectées par ces changements de stratification.

1. Pour un même plan de sondage et une même variable de stratification on constate que de manière générale les estimations d'effectifs profitent plus de l'introduction de la stratification que les estimations de quotients. Le gain se situe autour de **50 % du coefficient de variation** pour les premiers (variance divisée environ par 4,5) alors qu'il est beaucoup plus variable pour les seconds, s'étalant entre 10 et 40 % (division de la variance par 1,3 à 2,5).

2. Une fois redressées les erreurs d'affectation de l'échantillon aux strates, la stratification issue de l'image remplit fort bien son rôle puisque les gains qu'elle permet sont tout à fait comparables à ceux apportés par la stratification de référence : une fois optimisée l'allocation aux strates, la différence entre les deux stratifications n'exède pas 1,5 % du coefficient de variation non stratifié dans le cas des estimations d'effectifs. Cette différence, tout en restant faible, est plus variable en ce qui concerne l'estimation des quotients. Elle peut même être alors à l'avantage de la stratification tirée de l'image, il est vrai dans des cas où le gain dû aux stratifications est de toute manière faible.

La stratification d'après l'image satellite non corrigée apparaît en revanche, étant donné les biais évoqués plus haut, beaucoup moins bonne que la stratification de référence du moins en ce qui concerne les effectifs. Pour les ratios les trois stratifications sont en fait globalement équivalentes, l'avantage revenant selon les cas à l'une ou l'autre. Ce phénomène, surprenant au premier abord, s'explique en réalité très simplement. Comme le suggère la première observation faite sur le tableau, dans le cas des quotients les liaisons entre variables à estimer et variables de stratification sont assez ténues. Il n'y a donc pas de raisons pour qu'une stratification seulement légèrement plus "exacte" que les autres en ce qui concerne le critère morphologique soit systématiquement plus efficace au sens du caractère démographique. La "meilleure" stratification peut donc être tour à tour chacune des trois.

3. Comme on l'a déjà remarqué, et c'est une évidence, les conséquences sur la précision des estimations du choix de la probabilité d'appartenance à l'échantillon des unités primaires ne sont pas indépendantes de la stratification retenue. On peut à ce propos formuler la règle générale suivante : **plus la redondance entre les deux variables définissant d'une part la stratification et d'autre part la probabilité d'appartenance est forte moins il y a d'intérêt à pratiquer avec cette stratification, une telle probabilité.**

Dans le cas qui nous intéresse, on trouve une bonne illustration de cette règle en comparant l'ensemble des colonnes "P-su" et "P-sb" dans les tableaux 2.5 et 2.6 : le gain sur le coefficient de variation des estimations obtenu en passant d'une probabilité d'inclusion proportionnelle à la surface totale des unités à une probabilité proportionnelle à leur surface bâtie est de l'ordre de 30 à 40% dans le cas non stratifié, il n'est plus que de 10 à 20 % lorsque la stratification de la base sur le critère de densité du bâti est "floue" (variable DENSBATIM) et il devient négligeable (on obtient même une détérioration de la précision pour certaines estimations) lorsque la stratification s'améliore (variables DENSBATI et DENSBATIME).

Tableau 2.6 : Coefficients de variation des estimations des variables démographiques avec différentes stratifications.(Programme Varech)

	Non stratifié		Stratifié DENSBATI		Stratifié DENSBATIME				Stratifié DENSBATIM			
	1 P-su	2 P-sb	3 P-su alloc. prop.	4 P-su alloc. opti.	5 P-su alloc. prop.	6 P-su alloc. opti.	7 P-sb alloc. prop.	8 P-sb alloc. opti.	9 P-su alloc. prop.	10 P-su alloc. opti.	11 P-sb alloc. prop.	12 P-sb alloc. opti.
1) Totaux												
1 EFFECTRES	25,42	15,97	13,28	12,30	14,13	12,69	14,68	13,18	17,92	16,40	15,42	14,20
2 NBMENAGES	26,26	15,58	13,27	12,19	14,08	12,57	14,16	12,68	18,13	16,44	15,18	13,72
4 NBMENPI	26,25	15,83	13,36	12,33	14,01	12,56	14,23	12,76	18,27	16,63	15,29	13,89
6 NBMENM7	27,14	15,92	13,63	12,40	14,39	12,83	14,32	12,80	18,69	16,83	15,51	13,85
8 EFFP5A	26,45	16,64	13,92	12,77	14,71	13,17	15,46	13,79	18,66	16,92	16,08	14,69
10 EFFNMIG	25,02	16,60	14,58	13,60	14,97	13,85	15,42	14,19	18,75	17,36	15,69	14,76
12 CEFFMIGM5	26,42	16,61	14,06	12,92	14,80	13,31	15,42	13,78	18,81	17,08	16,00	14,64
2) Quotients												
13 EFFECTRES / NBMENAGE	4,07	3,08	3,15	2,83	3,14	2,86	2,96	2,73	3,18	2,87	2,82	2,57
14 EFFECTRES / SURFACE	25,42	23,77	13,28	12,30	14,13	12,69			17,92	16,40	20,35	18,09
16 NBMENPI / NBMENAGE	1,54	1,10	1,15	0,96	1,14	1,0	1,06	0,94	1,10	0,95	0,96	0,83
18 NBMENM7 / NBMENAGE	2,40	2,02	2,0	1,89	1,96	1,84	1,89	1,80	2,07	1,95	1,92	1,80
20 EFFP5A / EFFECTRES	1,12	0,99	0,94	0,85	0,95	0,91	0,98	0,93	0,95	0,91	0,91	0,87
22 EFFNMIG / EFFECTRES	4,84	4,04	3,25	3,02	3,41	3,26	3,30	3,14	3,30	3,16	3,16	2,95
24 CEFFMIGM5 / EFFECTRES	1,48	1,99	1,41	1,32	1,40	1,28	1,55	1,40	1,44	1,31	2,03	1,69

P=su : probabilité d'appartenance à l'échantillon proportionnelle à la surface totale des unités.

P=sb : probabilité d'appartenance à l'échantillon proportionnelle à la surface bâtie des unités.

Plus que sa mise en évidence, qui peut paraître triviale, c'est la quantification du phénomène qui a eu une réelle importance pratique : l'équivalence, en terme de précision des estimations obtenues, des deux stratégies de sélection des unités primaires (probabilités proportionnelles aux surfaces totales ou aux surfaces bâties), justifie l'emploi de la méthode de tirage la plus simple, rapide et adaptée à la base de sondage dont nous disposons : le sondage systématique aréolaire à l'aide d'une grille de points.

Allocations optimales

D'après les résultats précédents, pour le premier degré du sondage, on peut considérer comme acquis un plan stratifié d'après la densité du bâti observée sur l'image satellite, à probabilité proportionnelle à la surface totale des unités primaires. Les colonnes 5 et 6 du tableau 2.6 montrent que dans ce cadre, l'introduction de l'allocation optimale de l'échantillon aux strates améliore de façon significative les estimations sur l'ensemble de la base. En moyenne, passer de l'allocation proportionnelle à l'allocation optimale fait gagner 20% de la variance mais l'allocation optimale, fonction de la variance au sein de chaque strate du caractère que l'on estime, doit être redéfinie pour chaque variable. Les corrélations entre indicateurs démographiques observées lors de la phase exploratoire de l'expérience peuvent faire penser que l'évolution de la variance du caractère à estimer d'une strate à l'autre est semblable pour bon nombre des indicateurs démographiques. Si tel était le cas on pourrait construire une stratégie d'allocation "intermédiaire" c'est à dire qui soit un compromis entre les différentes allocations optimales, permettant d'améliorer l'ensemble des estimations. A partir du tableau 2.7 on peut comparer, pour l'ensemble des indicateurs démographiques retenus, les répartitions de l'échantillon entre les strates correspondant aux allocations optimales.

En premier lieu, l'observation des résultats conduit, une fois encore, à nuancer les conclusions selon que l'on s'intéresse aux effectifs ou aux quotients. Pour les premiers, toutes les répartitions correspondantes aux allocations optimales sont à peu près équivalentes. Pour les seconds, on observe une plus grande variabilité dans le détail des répartitions cependant, les strates sur et sous représentées sont toujours les mêmes : respectivement les strates 4 et les strates 1 et 2.

La seconde constatation a trait à l'effet du changement de stratification. Il modifie bien entendu la répartition optimale de l'échantillon entre les strates mais, nous semble-t-il, dans des proportions raisonnables. Les allocations "moyennes" sont les suivantes (dans le cas des effectifs qui fournissent des répartitions très régulières) :

Strates :	1	2	3	4	5	Total
DENSBATIME						
Allocation proportionnelle :	6,5	23,9	28,3	23,9	17,4	100 %
Allocation optimale "moyenne" :	3	9	32	41	15	100 %
DENSBATIM						
Allocation proportionnelle :	10,9	13,1	34,8	32,6	8,7	100 %
Allocation optimale "moyenne" :	4	4	33	50	9	100 %

Tableau 2.7 : Comparaisons des allocations optimales obtenues pour l'ensemble des caractères démographiques considérés.(Programme Varech)

Répartition de l'échantillon entre les strates (%)	Stratification "DENSBATIM"					Stratification "DENSBATIME"				
	Strates :					Strates :				
	1	2	3	4	5	6	7	8	9	10
A) Valeurs connues sur 46 îlots										
NOMBRE D'ILOTS ALLOC. PROPORTION.	5	6	16	15	4	3	11	13	11	8
	10,9	13,1	34,8	32,6	8,7	6,5	23,9	28,3	23,9	17,4
1) Totaux										
1 EFFECTRES	4,6	3,7	34,2	49,7	7,8	2,8	9,3	33,8	40,2	13,9
2 NBMENAGES	3,8	3,8	32,5	50,8	9,2	2,5	8,9	31,9	41,4	15,3
4 NBMENP1	3,9	4,2	32,6	50,8	8,5	2,6	9,5	31,6	41,3	14,9
6 NBMENM7	3,0	4,4	31,4	51,9	9,3	2,5	9,4	29,6	42,5	15,9
2) Quotients										
7 EFFECTRES / NBMENAGE	2,7	4,2	40,7	48,0	4,3	0,5	12,1	33,0	37,8	16,6
8 NBMENP1 / NBMENAGE	0,1	6,9	54,0	23,3	15,8	0,1	9,8	43,7	19,2	27,2
9 NBMENM7 / NBMENAGE	7,6	7,0	42,8	41,0	1,5	0,7	18,0	37,1	31,4	12,9

Tableau 2.7 (suite) : Comparaisons des allocations optimales obtenues pour l'ensemble des caractères démographiques considérés. (Programme Varech)

Répartition de l'échantillon entre les strates (%)	Stratification "DENSBATIM"					Stratification "DENSBATIME"				
	Strates :					Strates :				
	1	2	3	4	5	6	7	8	9	10
B) Valeurs connues sur 44 îlots										
NOMBRE D'ÎLOTS	5	6	15	14	4	3	11	12	11	7
ALLOC. PROPORTION.	11,4	13,7	34,1	31,8	9,1	6,8	25,0	27,3	25,0	15,9
1) Totaux										
8 EFP5A	4,7	3,7	34,0	49,9	7,5	3,1	9,0	33,5	41,3	13,1
10 EFNMG	6,0	4,1	37,8	45,9	6,1	3,7	11,4	35,4	36,4	13,2
12 CEFFMIGM5	4,8	3,5	34,8	49,3	7,5	3,0	9,6	33,6	40,8	12,8
2) Quotients										
20 EFP5A / EFFECTRES	7,6	6,3	34,5	44,7	6,8	6,0	14,6	30,2	35,3	14,0
22 EFNMG / EFFECTRES	6,9	5,5	39,8	37,7	10,1	1,5	21,5	26,1	26,4	24,5
24 CEFFMIGM5 / EFFECTRES	4,6	7,1	54,0	29,3	5,0	1,7	42,1	20,9	23,9	11,4

Dans les deux cas les modifications apportées par rapport à l'allocation proportionnelle sont comparables : peu de changement dans les strates 3 et 5, augmentation sensible du taux dans la strate 4 au détriment des strates 1 et 2.

On peut donc conclure à une relative robustesse des stratégies d'allocation optimale vis à vis de modifications opérées soit sur la stratification soit concernant les variables à estimer. Partant de cette constatation, l'idée qui a été retenue pour construire la règle d'allocation aux strates est de moyenniser les allocations optimales au sein de l'ensemble où elle apparaissent les plus stables c'est à dire les effectifs. Par ailleurs comme on l'a vu, DENSBATIM est la stratification qu'on appliquera à la base de sondage, c'est donc à partir des allocations optimales lui correspondant que l'on effectue le calcul. On parvient ainsi à une répartition qui correspond, par rapport à l'allocation proportionnelle, à une multiplication du taux dans chaque strate par les coefficients suivants :

Strates :	1	2	3	4	5	Total
Coefficients :	0,4	0,3	1	1,5	0,9	1

Cette allocation est surtout bénéfique à l'estimation des effectifs mais elle devrait également améliorer l'estimation de la plupart des quotients (voir tableau 2.7).

2.4 Conclusions

Au début de ce travail sur le premier degré du sondage nous avons posé quatre questions :

1. Quelle probabilité doit avoir chaque unité primaire d'appartenir à l'échantillon?
2. Quels estimateurs employer pour l'inférence de l'échantillon de pâtés de maisons au total des strates?
3. Quel gain de précision permet la stratification de la base de sondage à partir de l'image satellite?
4. Y a-t-il des possibilités d'optimisation de l'allocation aux strates?

En conclusion, en faisant la synthèse des résultats obtenus, on peut formuler pour chacune d'elle les réponses suivantes.

Définition de la probabilité d'appartenance des unités primaires à l'échantillon :

L'alternative est la suivante : la probabilité doit-elle être proportionnelle à la surface totale des unités ou à une approximation de leur surface bâtie.

Dans le cadre non stratifié, par rapport à la référence équiprobable, un sondage à probabilités proportionnelles aux surfaces totales des unités primaires ne présente pas d'intérêt ; par contre, une probabilité de sélection proportionnelle à la surface bâtie, ou à son approximation d'après l'image satellite permet une amélioration notable : 30 à 40 % de réduction du coefficient de variation dans le cas de la mesure de la surface bâti, 20 à 30 % s'il s'agit de l'approximation par la surface en végétation.

Une fois stratifiée la base de sondage selon la densité du bâti, il faut par contre retenir le sondage à probabilités proportionnelles à la surface totale des unités. En effet, d'une part il est meilleur que le sondage équiprobable stratifié, d'autre part dans un tel cadre, le tirage avec probabilités proportionnelles aux surfaces bâties n'améliore plus l'estimation par rapport au tirage à probabilité proportionnelles aux surfaces totales. **Ce résultat, qui s'ajoute à la nécessité d'une méthode de tirage simple, rapide et adaptée à la base de sondage**

dont nous disposons, justifie l'emploi du sondage systématique aréolaire à l'aide d'une grille de points.

Définition de l'estimateur à employer :

Les seules informations exogènes qu'on puisse envisager d'utiliser dans l'étape d'inférence (estimation par le ratio ou par la régression), sont à nouveau les mesures de taille des unités primaires. Les coefficients de corrélation entre ces mesures et les indicateurs démographiques, compris entre 0,5 et 0,6, sont nettement insuffisants pour justifier l'emploi de ces estimateurs pour l'estimation des totaux, mais, comme toujours, l'estimateur du ratio reste utile pour l'estimation directe des quotients; pour le reste, on utilisera les estimateurs sans biais¹.

Choix des critères morphologiques de stratification :

Le gain que permet la stratification en cinq niveaux de densité de bâti choisie comme référence se situe autour de 50 % du coefficient de variation (estimation des effectifs). Au terme de l'expérience, on peut penser que la stratification issue de l'image permet une amélioration comparable, du moins une fois optimisée l'allocation de l'échantillon aux strates. C'est donc cette stratification qui sera utilisée lors de l'enquête finale ; néanmoins il serait bon, si l'on prolonge ces recherches, de tester d'autres classifications de l'image satellite.

Définition d'une règle d'allocation aux strates :

En moyenne, passer de l'allocation proportionnelle à l'allocation optimale fait gagner 20% de la variance; et la robustesse des règles d'allocation optimale vis à vis des modifications opérées soit sur la stratification soit concernant les variables à estimer autorise à "moyenner ces allocations". On parvient ainsi à une allocation qui ne modifie presque pas la représentation dans l'échantillon des strates 3 et 5 (densités moyenne et très forte), mais qui favorise nettement la strate 4 (densité forte) au détriment des strates 1 et 2 (densités faible et très faible). Cette allocation, surtout bénéfique à l'estimation des effectifs, devrait néanmoins améliorer l'estimation de la plupart des quotients.

Au terme de cette étape on est parvenu à proposer des solutions concrètes aux problèmes de sélection et d'inférence liés à l'échantillonnage des pâtés de maisons. On sait de plus quelles conséquences aura la mise en œuvre de ces solutions sur la précision des estimations. Les contraintes, liées à l'emploi de l'image satellite comme base de sondage et aux objectifs de rapidité, simplicité et faible coût que l'on s'était fixés au départ, s'avèrent compatibles avec une amélioration sensible de la précision des estimations grâce à l'emploi de la stratification, du tirage à probabilités inégales et d'une allocation de l'échantillon aux strates optimisée. Pour aboutir à la définition complète du plan de sondage, il reste à décrire les étapes de sélection et d'inférence correspondant au second degré du sondage : l'échantillonnage des ménages au sein des pâtés de maisons. C'est l'objet du chapitre suivant.

¹ : Le formulaire associé au plan de sondage que nous appliquerons pour l'enquête finale est donné en résumé au chapitre 4 du présent rapport, ou plus complètement, du moins en ce qui concerne le premier degré du sondage, dans le chapitre 4 du second rapport intermédiaire à l'ORSTOM : Mise au point de plan de sondage pour l'estimation de caractères démographiques à Marseille, O. Barbary, Janvier 1983, QUITO (184 p), pp 62 à 72.

CHAPITRE 3

ENQUETE PRELIMINAIRE "BATI POPULATION", EXPLOITATION POUR LA MISE AU POINT DU SECOND DEGRE DU SONDAGE : ECHANTILLONNAGE DES MENAGES DANS LES PATES DE MAISONS

3.1 : Introduction, les objectifs principaux et les grande lignes de la démarche

Le principal résultat de l'expérience réalisée à Marseille et de la première exploitation de l'enquête "bâti population" faite à Quito est la détermination du premier degré du sondage. Dans ce chapitre on tente de compléter la définition du plan de sondage en abordant, à partir de ces mêmes données de l'enquête "Bâti- population", le problème de l'échantillonnage des ménages au sein des pâtés de maisons. Cette tentative comporte deux volets aux objectifs bien différents.

Le premier est l'analyse, à l'aide d'outils statistiques classiques, des possibilités "d'explication" des caractéristiques démographiques des ménages par trois facteurs "explicatifs" : la morphologie du pâté de maisons, sa localisation dans l'agglomération et la morphologie du bâtiment occupé par le ménage. L'objectif assigné à cette analyse du fichier "ménages" est la mesure du degré d'hétérogénéité (ou d'homogénéité) que présentent les caractéristiques démographiques des ménages au sein des sous-ensembles définis par les facteurs explicatifs. **Plus concrètement, il s'agit en fait de statuer sur la possibilité et l'intérêt d'une stratification (ou post-stratification) du second degré du sondage.**

Dans le second volet, on aborde réellement le problème de la sélection des ménages dans chaque pâté de maisons de l'échantillon. Après avoir exposé, en termes "intuitifs", **comment un sondage à deux degrés peut permettre, en enquêtant un même nombre total de ménages, d'améliorer la précision des estimations par rapport au sondage où toutes les unités primaires sont enquêtées exhaustivement**, on formalise cette idée dans le formulaire. Il fournit les expressions des variances des estimateurs dans le cas d'un plan de sondage à deux degrés. A partir de ces expressions, on a construit le programme Var2deg qui permet de mesurer la part de variance due à chacun des deux degrés du sondage. Le principe est ensuite de considérer l'échantillon de pâtés de maisons de l'enquête "bâti population" comme une base de sondage sur laquelle le programme effectue les calculs des variances correspondant aux plans envisagés. **Les résultats permettent d'évaluer le gain de précision en fonction de la fraction de sondage au second degré et donc de déterminer la stratégie optimale de sélection des ménages dans chaque pâté de maisons.**

En conclusion de chacun des deux volets, on trouvera un résumé des principaux résultats qui débouche sur la solution pratique retenue dans le plan de sondage.

3.2 Analyses préliminaires : les possibilités "d'explication" des caractéristiques démographiques des ménages par la morphologie du bâtiment, la morphologie et la localisation de l'îlot.

3.2.1 Les questions posées, les données et les outils statistiques utilisés

Les questions posées

La solution au problème posé passe par l'évaluation et la comparaison du pouvoir discriminant des facteurs explicatifs : ceux caractérisant les unités primaires d'une part (morphologie et localisation du pâté de maisons), celui pouvant servir de base à la stratification au sein de ces unités d'autre part (morphologie du bâtiment occupé par le ménage). Il nous faut donc répondre à plusieurs questions :

1 : Quel est le pouvoir discriminant, pour chacun des caractères démographiques considérés, de chaque facteur indépendant pris isolément : morphologie du bâtiment, morphologie de la manzana, localisation ? Quelles sont les variables démographiques se prêtant réellement à une discrimination par les facteurs explicatifs ? Avec les facteurs indépendants "naturellement ordonnés" pour lesquels un codage métrique est possible, peut-on construire des modèles de régression permettant l'estimation des variables démographiques ?

2 : Lorsqu'il existe une possibilité réelle de discrimination d'une variable démographique donnée par les facteurs indépendants, les discriminations opérées par chacun des critères "explicatifs" sont-elles redondantes ou au contraire peut-on s'attendre à un effet d'addition ? Au cas où un modèle intégrant l'ensemble des critères permet en effet une discrimination plus "fine", ce modèle peut-il être limité aux effets additionnés ou doit-il intégrer les interactions entre critères (d'ordre 1 ou supérieures) ? Autrement dit, l'effet d'un critère "explicatif" donné est-il le même quelles que soient les valeurs ou modalités des autres critères ?

3 : En conclusion, la liaison existant entre caractéristiques démographiques des ménages et morphologie des bâtiments est-elle suffisamment forte pour justifier la stratification ou la post-stratification du second degré du sondage ?

Les données

Les analyses statistiques réalisées portent sur un fichier de 1381 enregistrements correspondant à l'ensemble des ménages enquêtés. Chaque ménage est décrit par quatre caractères démographiques et par les trois facteurs "explicatifs" :

Variables démographiques

1. EFFECT : Effectif du ménage (nombre de personnes résidentes).
2. REVENU : Revenu moyen des personnes du ménage (somme des revenus déclarés par chacune des personnes du ménage divisée par le nombre de personnes) ; et également son logarithme en base 10 : LOGREV.
3. AGECHIEF : Age du chef de ménage en années ou tranches d'âge quinquennales.
4. DURESMIG : Durée moyenne de résidence à Quito des migrants du ménage (somme des durées de résidence déclarées par chacun des migrants du ménage divisée par le nombre des migrants) en années ou tranches d'âge quinquennales.

Facteurs explicatifs

1. TYPÉMORF : Type morphologique du bâtiment occupé par le ménage. Typologie issue de la classification en 38 types résultant de l'exploitation de l'enquête préliminaire "morphologie" après regroupement de certaines classes morphologiquement proches pour aboutir à 14 types¹.
2. TYPÉMANZ : Type morphologique du pâté de maisons occupé par le ménage. Classification en 9 types obtenue lors de l'exploitation de l'enquête préliminaire "morphologie". Pas de type "non bâti", 8 modalités².
3. LOC : Appartenance de l'îlot à un des 12 sous-ensembles géographiques définis dans la ville, issus du croisement d'une division Nord/Sud en 4 zones avec une division Est/Ouest en 3 zones.

Cependant, du fait d'absences ou de refus de réponses, l'ensemble de ces données n'est pas disponible sur la totalité des 1381 ménages enquêtés et le nombre de ménages pour lesquels les données sont connues varie suivant les caractères considérés. Plutôt que de chercher à constituer un fichier sur lequel on dispose de la totalité de l'information, ce qui restreint beaucoup le nombre de ménages pris en compte dans la plupart des analyses, on a préféré ne conserver pour chaque analyse que l'ensemble des ménages pour lesquels les données nécessaires sont connues. Ainsi, selon les variables prises en compte dans tel ou tel traitement, l'analyse s'effectue sur un nombre plus ou moins grand de ménages. Ce nombre sera toujours indiqué pour chaque variable ou groupe de variables dans les tableaux de résultats.

Les outils statistiques utilisés

Pour aborder les questions posées on a utilisé, tour à tour et selon la nature du problème et des données, des techniques d'analyse de la variance³ à un ou plusieurs critères, avec ou sans interactions et de régression simple ou multiple⁴.

Ainsi, pour mesurer le pouvoir discriminant de chacun des facteurs explicatifs pris isolément, on a construit à l'aide du logiciel SPSS⁵ les modèles d'analyse de la variance à un critère correspondant à l'ensemble des combinaisons possibles entre variables dépendantes (chacun des quatre caractères démographiques) et variables indépendantes (morphologie du bâtiment, morphologie de la manzana, localisation de la manzana). Le pourcentage de variance expliquée par la variable indépendante (variance inter-classe) permet alors d'évaluer son pouvoir discriminant sur chacune des variables démographiques.

La construction de modèles de régression linéaire qui permettent la prévision des caractéristiques démographiques des ménages à l'aide de l'information morphologique ou de la localisation dans l'agglomération ne peut être envisagée qu'à

¹ : Pour une description de ces 14 types de bâtiments, voir annexe 1.

² : Pour une description des types de pâtés de maison, voir annexe 1.

³ : sur l'analyse de la variance, voir par exemple : L. LEBART, A. MORINEAU et J.P. FENELON : Traitement des données statistiques, méthodes et programmes, 2^e édition, DUNOD, 1982, Paris.

⁴ : sur la régression, voir E. MALINVAUD (1981, 3^e édition) : Méthodes statistique de l'économétrie. DUNOD, Paris, 846 p.

⁵ : Statistical Package for Social Sciences, programme ONEWAY.

condition que les variables exogènes puissent être codées sur des échelles métriques. Les données recueillies ne permettent un tel codage que pour les deux variables décrivant la localisation de l'îlot au sein de l'agglomération : latitude et longitude d'un point choisi comme centre du pâté de maisons. Ce sont ces deux variables qui ont été prises comme facteurs explicatifs. Malheureusement, mais cela était prévisible, les coefficients de régressions obtenus, que ce soit lors de régressions simples ou multiples¹, ne sont jamais supérieurs à 0,3. C'est évidemment très insuffisant pour espérer prévoir, avec la technique des moindres carrés, les caractéristiques des ménages. Nous n'exposerons donc pas plus avant les résultats obtenus puisqu'ils sont de toute manière dénués d'intérêt pratique.

Enfin, pour essayer de mesurer le degré de redondance existant entre les trois informations exogènes que nous considérons, cela lorsqu'au moins deux des critères explicatifs avaient un pouvoir discriminant réel sur une variable démographique, nous avons soumis les données à des analyses de la variance multi-critères². En comparant les parts de variance expliquées par les modèles intégrant plusieurs facteurs explicatifs, modèles à deux puis à trois critères, à celles expliquées par les modèles à un critère, cette technique a permis de voir dans quelle mesure les effets explicatifs des différents critères s'additionnent. Dans cette étape on a également cherché à savoir s'il existe des effets d'interaction entre les trois critères. Après avoir recodé les variables pour limiter les modalités prises par les critères obtenus par croisement et l'espace mémoire nécessaire aux calculs, on construit les modèles comportant les termes d'interactions d'ordre un (interactions entre deux critères) et les résultats permettent de conclure.

3.2.2 : Résultats

A : Analyse de la variance à un critère : pouvoir discriminant des facteurs "explicatifs" sur les différents caractères démographiques étudiés.

- Préliminaire : condition d'interprétabilité des analyses

Pour interpréter légitimement les résultats d'une analyse de variance, on doit avant tout s'assurer que les données entrent bien dans le cadre des hypothèses qui fondent la méthode. La condition principale pour cela est que la variance du caractère analysé n'apparaisse pas significativement différente d'une classe à l'autre³. Les sorties fournies par le programme SPSS contiennent, comme indicateur d'homogénéité des variances, le ratio de la variance maximale sur la variance minimale : lignes 4, 8 et 12 du tableau 3.1. Pour juger une analyse "interprétable", on a considéré que sa valeur ne devait pas être supérieure à 3. Les suppressions d'individus

¹ : On a utilisé les programme PLOT et REGRESSION de SPSS.

² : Programme ANOVA de SPSS.

³ : Ici, comme en général en analyse de la variance, on entend par classe l'ensemble des individus ayant une modalité donnée pour le caractère exogène.

trop originaux ou le recodage des variables¹ ont permis dans presque tous les cas d'atteindre ce seuil d'interprétabilité. La seule analyse pour laquelle les résultats doivent être considérés prudemment est celle "expliquant" l'âge du chef de ménage par les modalités de la variable "localisation".

- Résultats

Dans le tableau 3.1, pour chacun des trois facteurs indépendants examinés, on retient cinq indications fournies par le programme d'analyse de la variance :

1. Effectif : Nombre de ménages pris en compte dans cette analyse
2. Part de la variance expliquée : quotient (exprimé en %) de la variance inter-classes sur la variance totale.
3. Rejet de l'hypothèse d'égalité des moyennes : il s'agit du résultat de l'application du test de Fisher au seuil de confiance de 5%.
 - oui : $P < 0,05$, rejet de l'hypothèse d'égalité
 - non : $P > 0,05$, pas de rejet
4. Indicateur d'homogénéité des variances selon les classes : quotient de la variance dans la classe la plus hétérogène sur la variance dans la classe la plus homogène.
5. Nombre de classes différentes : il s'agit des classes significativement différentes au sens de la procédure de Scheffe. On donnera dans le tableau le nombre de différences constatées au seuil fixé à l'avance, 5 ou 10% selon les cas, rapporté au nombre de différences théoriquement possibles ($n.(n-1)/2$ où n est le nombre de classes de la variable indépendante).

¹ : Les caractéristiques des variables démographiques avant et après "recodage" sont les suivantes :

Variables "brutes" :

- 1. EFFECT : De 0 à 14 personnes, 1381 ménages
- 2. REVENU : De 0 à 205 000 sucres, 751 ménages
- 3. AGECHIEF : En années, 1086 ménages
- 4. DURESMIG : En années, 745 ménages

Variables "recodées" :

- 1. EFFECTR : Suppression du ménage sans résident et des ménages de plus de 9 personnes (7 ménages supprimés).
- 2. REVENUR : suppression des revenus nuls et supérieurs à 45 000 sucres (11 ménages supprimés).
- 3. LOGREV : Passage au logarithme en base 10 sur REVENUR.
- 4. AGECHIEFR : Transformation en tranches d'âge quinquennales
- 5. DURESMIGR : Transformation en tranches d'âge quinquennales

Tableau 3.1 : Résultats résumés des analyses de la variance à un critère
(Programme "ONEWAY", SPSS)

		EFFECT	EFFECTR	REVENU	REVENUR	LOGREV
L O C	Nombre de ménages	1381	1374	762	751	751
	Part de variance expliquée en % de la variance totale	4,3	3,7	13,8	23,5	27,9
	Rejet de l'hypothèse d'égalité des moyennes	oui	oui	oui	oui	oui
	Indicateur d'homogénéité des variances (VM/Vm)	3,64	2,21	135,94	35,29	2,79
	Nombre de classes différentes au sens de la procédure de Scheffe	2/66	1/66	9/66	25/66	21/66
T Y P E M O R F	Nombre de ménages	1283	1276	725	714	714
	Part de variance expliquée en % de la variance totale	3,4	3,1	15,0	21,6	23,6
	Rejet de l'hypothèse d'égalité des moyennes	oui	oui	oui	oui	oui
	Indicateur d'homogénéité des variances (VM/Vm)	4,80	3,04	88,06	13,14	2,91
	Nombre de classes différentes au sens de la procédure de Scheffe	1/91	0/91	11/91	22/91	20/91
T Y P M A N Z	Nombre de ménages	1381	1374	762	751	751
	Part de variance expliquée en % de la variance totale	2,2	2,0	13,4	21,5	24,1
	Rejet de l'hypothèse d'égalité des moyennes	oui	oui	oui	oui	oui
	Indicateur d'homogénéité des variances VM/Vm)	2,45	2,17	21,17	8,79	2,27
	Nombre de classes différentes au sens de la procédure de Scheffe	1/28	2/28	6/28	11/28	14/28

Tableau 3.1 (suite) : Résultats résumés des analyses de la variance à un critère (Programme "ONEWAY", SPSS)

		AGECHEF	AGECHEFR	DURES MIG	DURES MIGR
L O C	Nombre de ménages	1086	1086	745	745
	Part de variance expliquée en % de la variance totale	5,4	5,3	2,6	2,5
	Rejet de l'hypothèse d'égalité des moyennes	oui	oui	non	non
	Indicateur d'homogénéité des variances (VM/Vm)	4,48	4,52	2,1	2,2
	Nombre de classes différentes au sens de la procédure de Scheffe	5/66	6/66	0/66	0/66
T Y P E M O R F	Nombre de ménages	1017	1017	702	702
	Part de variance expliquée en % de la variance totale	2,9	2,9	2,3	2,4
	Rejet de l'hypothèse d'égalité des moyennes	oui	oui	non	non
	Indicateur d'homogénéité des variances (VM/Vm)	2,39	2,49	2,30	2,70
	Nombre de classes différentes au sens de la procédure de Scheffe	0/91	0/91	0/91	0/91
T Y P M A N Z	Nombre de ménages	1086	1086	745	745
	Part de variance expliquée en % de la variance totale	2,0	2,0	0,7	0,7
	Rejet de l'hypothèse d'égalité des moyennes	oui	oui	non	non
	Indicateur d'homogénéité des variances (VM/Vm)	2,32	2,23	2,12	2,18
	Nombre de classes différentes au sens de la procédure de Scheffe	2/28	2/28	0/28	0/28

- Conclusions

1 : Durée moyenne de résidence des migrants du ménage

Aucun des trois facteurs indépendants ne possède de pouvoir discriminant notable sur cette variable. L'hypothèse d'égalité des moyennes n'est jamais rejetée et aucune classe n'apparaît significativement différente d'une autre au sens de Scheffe.

2 : Effectif des résidents du ménage et âge du chef de ménage

Pour ces deux variables, certaines possibilités de discrimination existent à l'aide de la localisation et du type morphologique de la manzana, par contre le type morphologique du bâtiment n'a pas d'effet notable, aucune classe n'étant significativement différente d'une autre. Les parts de variance expliquées restent de toute manière très faibles (inférieures à 5%); les quelques différences entre moyennes à noter sont les suivantes :

Taille des ménages :

Les classes de localisation 1 (zone extrême sud-ouest, majoritairement péri-urbaine où les ménages ont des effectifs importants) et 9 (zone nord-est incluant "El Batán", quartier résidentiel de standing où les ménages sont en général de petite taille) s'opposent. De même, le type d'îlots 2 (péri-urbain), avec des ménages nombreux se différencie des types 8 (grands ensembles) et 5 (centre colonial) où les ménages sont peu nombreux.

Age du chef de ménage :

Les jeunes chefs de ménage résident surtout dans les zones 3 (extrême sud-est), 6 (sud-est) et 12 (nord-est), tandis que la zone 8 (centre nord, nouveau centre d'affaires) regroupe un nombre important de chefs de ménage âgés. Les zones 5 (centre sud colonial) et 9 (nord est résidentiel) offrent en général des valeurs intermédiaires de la variable. Quant au type morphologique de l'îlot, l'habitat de standing (classe 6) avec ses "vieux" chefs de ménage, s'oppose aux quartiers "jeunes" péri-urbain et d'habitat populaire (respectivement classes 2 et 3).

3 : Revenu moyen des personnes du ménage

Les facteurs explicatifs retenus permettent tous les trois une discrimination intéressante des valeurs prises par le logarithme en base 10 du revenu moyen des personnes du ménage. Les trois facteurs ont des pouvoirs discriminants équivalents se situant autour de 25% de la variance totale (localisation 28%, morphologie de la manzana et du bâtiment 24%). Le passage au logarithme sur cette variable améliore nettement les résultats, tant du point de vue de la variance expliquée que de celui de l'homogénéité des variances au sein des classes, cela dans la même proportion pour les trois facteurs explicatifs. Le principal résultat à noter est le suivant : le facteur "localisation" permet de discriminer quatre catégories de revenus décrites dans le tableau 3.2 ci-dessous. Les modalités des facteurs explicatifs regroupées pour former chaque catégorie y sont indiquées par leurs numéros. Pour interpréter ces regroupements en terme de localisation dans l'agglomération ou de morphologie des bâtiments ou des pâtés de maisons, il faut se reporter à la description des modalités.

Tableau 3.2 : Regroupements effectués d'après les analyses de la variance à un critère pour obtenir des catégories de revenus.

Modalités de la variable explicative et nombre de ménages concernés			Catégories de revenus et moyenne de la variable (en sucres/mois)		
L O C	1:	1, 2, 3, 4, 7, 12	236	Petits revenus	5200
	2:	6,10	87	Moyens revenus	7300
	3:	5,11	247	Hauts revenus	10700
	4:	8,9	181	Très hauts revenus	15500
	Total	751		(Moyenne : 9700)	
T Y P M O R F	1:	1,2,4,5,6	268	Petits revenus	6100
	2:	3,7,9	189	Moyens revenus	9700
	3:	8,10,11,12,13	199	Hauts revenus	12300
	4:	14	58	Très hauts revenus	17500
	Total	714		(Moyenne : 9700)	
T Y P E M A N Z	1:	2,4,7	287	Petits revenus	5850
	2:	3,5	186	Moyens revenus	9200
	3:	6,9	78	Hauts revenus	11650
	4:	8	200	Très hauts revenus	14400
	Total	751		(Moyenne : 9700)	

B. Analyse de la variance multi-critères : redondance/indépendance de l'information apportée par les trois facteurs "explicatifs".

D'après les résultats précédents, il est clair que seule la variable indiquant le revenu moyen des personnes du ménage, plus précisément son logarithme en base 10, se prête à une discrimination intéressante à l'aide des trois facteurs explicatifs considérés. **Dans la suite des traitements, où il s'agit d'intégrer dans une même analyse plusieurs facteurs, on n'a donc retenu comme variable à expliquer que cet indicateur du revenu.** Le tableau 3.3 donne la part de variance (en % de la variance totale) expliquée par les facteurs indépendants dans les trois analyses à deux critères possibles (croisements de deux facteurs) et dans l'analyse intégrant les trois facteurs. Figure également en première ligne le rappel des résultats obtenus avec l'analyse à un seul critère, enfin les colonnes "gain" expriment le gain obtenu lors de l'introduction dans l'analyse du deuxième ou troisième critère (en pourcentage de la variance expliquée par l'analyse à un critère).

L'examen du tableau 3.3 montre que la redondance d'information entre les trois facteurs "explicatifs" est importante sans être totale. On peut d'ailleurs classer les combinaisons deux à deux des facteurs de la moins redondante à la plus redondante :

- 1: localisation et morphologie du bâtiment
- 2: morphologie du bâtiment et morphologie de l'ilot.
- 3: localisation et morphologie de l'ilot.

Cependant même la dernière combinaison permet une amélioration non négligeable du pourcentage de variance expliquée si on la compare à l'analyse n'intégrant qu'un seul des deux critères.

Tableau 3.3 : Pourcentage de variance expliquée dans les analyses de variance à un, deux et trois critères.

	Localisation		Morphologie du bâtiment		Morphologie de l'ilot	
	% de var. expliquée	Gain	% de var. expliquée	Gain	% de var. expliquée	Gain
Un seul critère	28 %	0 %	24 %	0 %	24 %	0 %
D e Localisation	34,5 %	43,2 %	32,2 %	34,2 %
u x Morphologie du bâtiment	34,5 %	23,2 %	29,5 %	22,9 %
c r Morphologie de la manzana	32,2 %	15 %	29,5 %	22,9 %
Trois critères	36,5 %	30,4 %	36,5 %	52,1 %	36,5 %	52,1 %

Le passage de deux à trois critères apporte à nouveau un léger gain, mais on constate que ce dernier est très faible lorsque le troisième facteur introduit est la morphologie du bâtiment. On peut donc penser qu'une fois prise en compte dans la procédure de sondage la morphologie des manzanas (par stratification du premier degré du sondage sur critère morphologique) et la localisation (par tirage systématique assurant la bonne répartition géographique de l'échantillon), il y a relativement peu d'intérêt à stratifier le second degré du sondage (le choix de l'échantillon de ménages à enquêter dans le pâté de maison) sur un critère tel que la morphologie du bâtiment. Ce résultat important sera retenu pour l'élaboration du plan de sondage final.

Le tableau 3.4, reprenant le même principe que le tableau 3.3, fournit les chiffres permettant d'apprécier le gain qu'amène la prise en compte dans les analyses de variance, des interactions d'ordre 1. Le test a été mené après un recodage des variables indépendantes qui réduit le nombre de modalités prises : on a regroupé les modalités qui n'étaient pas significativement différentes, du point de vue de la variable à "expliquer", lors des analyses à un critère. Ce faisant, on ne modifie que très peu la part de la variance "expliquée" comme en témoigne la première ligne du tableau.

Dans une analyse de la variance à plusieurs critères, si l'on pense que l'effet du critère A peut être différent selon la modalité du critère B, on introduit dans le modèle un terme qui exprime l'interaction¹.

Tableau 3.4 : Pourcentage de variance expliquée dans les analyses de variance à un, deux et trois critères, après regroupement de certaines modalités des variables explicatives et prise en compte des interactions d'ordre 1.

	Localisation		Morphologie du bâtiment		Morphologie de l'îlot	
	% de var. expliquée	Gain	% de var. expliquée	Gain	% de var. expliquée	Gain
Un seul critère	27 %	0 %	22 %	0 %	22 %	0 %
Localisation (sans interaction)	30,5 %	38,6 %	30,3 %	37,7 %
Localisation (avec interaction)	32 %	45,5 %	31,9 %	45 %
Morphologie du bâtiment (sans interaction)	30,5 %	13 %	27,7 %	25,9 %
Morphologie du bâtiment (avec interaction)	32 %	18,5 %	28 %	27,3 %
Morphologie de la manzana (sans interaction)	30,3 %	12,2 %	27,7 %	25,9 %
Morphologie de la manzana (avec interaction)	31,9 %	18,1 %	28 %	27,3 %
Trois critères (sans interaction)	32,2 %	19,3 %	32,2 %	46,4 %	32,2 %	46,4 %

L'effet discriminant de la morphologie des manzanas est-il par exemple notablement différent selon que l'on se trouve dans telle ou telle zone de Quito?

¹ : Voir L. LEBART, A. MORINEAU et J.P. FENELON : Traitement des données statistiques, méthodes et programmes, 2^e édition, DUNOD, 1982, Paris. p 231

On peut répondre à ce type de question en examinant, à partir des résultats rapportés dans le tableau 3.4 si l'introduction des termes d'interaction est responsable d'une amélioration sensible des performances du modèle. En fait, à l'examen du tableau, il apparaît qu'aucune des interactions possibles n'a d'effet réellement important. Pour l'interaction entre morphologie de la manzana et morphologie du bâtiment par exemple, le test effectué ne permet pas de rejeter l'hypothèse selon laquelle l'effet du type de bâtiment est le même quelque soit le type de manzana. Quant au gain obtenu sur la part de variance expliquée par le modèle, il n'exède pas 5% de la variance expliquée par le modèle sans inter-action. **On pourra donc dans la suite considérer que, même si elles existent, on ne peut pas attendre de la prise en compte de ce type d'interactions (là encore par stratification croisée ou post-stratification) une amélioration sensible de la précision du sondage.**

3.2.3 Conclusions

Avant de résumer les principaux résultats acquis au cours de l'analyse du fichier "ménages", rappelons-en l'objectif principal. Il s'agissait, après évaluation et comparaison du pouvoir discriminant des facteurs explicatifs, de statuer sur la possibilité et l'intérêt d'une stratification (ou post-stratification) du second degré du sondage.

Concernant les possibilités de discrimination des caractéristiques démographiques des ménages par la morphologie du bâtiment ou du pâté de maison où ils résident, ou leur localisation dans l'agglomération, les analyses de variances à un critère fournissent les indications suivantes.

Premièrement aucun des trois facteurs indépendants ne possède de pouvoir discriminant notable sur la durée moyenne de résidence des migrants du ménage. Deuxièmement, en ce qui concerne l'effectif des résidents du ménage et l'âge du chef de ménage, certaines possibilités de discrimination existent à l'aide de la localisation et du type morphologique de l'îlot, par contre le type morphologique du bâtiment a peu d'intérêt.

Enfin les facteurs explicatifs retenus permettent tous les trois une discrimination intéressante (autour de 25% de la variance totale) du revenu moyen des personnes du ménage. Le pouvoir "explicatif" des trois facteurs apparaît donc très inégal suivant qu'on considère l'un ou l'autre des caractères démographiques. C'est en particulier le cas pour la morphologie du bâtiment occupé par le ménage qui ne permet de discriminer que des classes de revenus. **On peut donc d'ores et déjà affirmer qu'une stratification du second degré du sondage sur ce critère n'améliorerait significativement les estimations que pour une minorité des indicateurs démographiques qui nous intéressent.**

En conséquence, la suite de l'expérience n'a porté que sur l'indicateur du revenu moyen des ménages ; elle s'attache, au moyen des analyses de la variance multi-critères, à l'évaluation du degré de redondance existant entre les informations apportées par les trois facteurs "explicatifs" ainsi qu'à l'examen des interactions entre ces facteurs. Les résultats obtenus montrent que la redondance d'information entre les trois facteurs "explicatifs" est importante sans être totale. Mais, on constate que le gain lors du passage de deux à trois critères est très faible lorsque le troisième facteur introduit est la morphologie du bâtiment. **En termes de procédures de sondage, cela veut dire qu'une fois prise en compte au niveau des unités primaires, la morphologie des pâtés de maisons (par stratification sur critère morphologique) et la localisation (par tirage systématique assurant la bonne répartition géographique de l'échantillon), il n'y a plus d'intérêt à stratifier le second degré du sondage sur un critère de morphologie du bâtiment.** De même, s'agissant des interactions entre ces trois facteurs, on pourra considérer que même si elles existent, on ne peut pas attendre de leur prise en compte,

facteurs, on pourra considérer que même si elles existent, on ne peut pas attendre de leur prise en compte, là encore par stratification croisée ou post-stratification, une amélioration sensible de la précision du sondage.

Au terme de ces analyses effectuées sur les données concernant les unités secondaires du sondage, les ménages, la réponse à la question principale est donc négative. On n'a vraisemblablement pas intérêt à stratifier ou à post-stratifier le second degré du sondage en opérant une division morphologique à l'intérieur des pâtés de maisons. C'est un résultat qui, même s'il était prévu, est important pour la définition du plan de sondage final. Mais en revanche, nous allons voir maintenant qu'il existe de réelles possibilités d'amélioration de la précision des estimations si, en n'enquêtant qu'une fraction des ménages de chaque pâté de maisons, on parvient à augmenter la taille de l'échantillon d'ilots.

3.3 Détermination du second degré du sondage : échantillonnage de ménages dans un pâté de maison.

3.3.1 Position du problème

Jusqu'à présent, dans toute l'expérience réalisée à Marseille et dans les premiers traitements des données recueillies à Quito, on a toujours considéré le pâté de maisons comme l'unité observée et cette unité était connue exhaustivement, soit que les données aient déjà été agglomérées à ce niveau comme c'est le cas pour le recensement de la population à Marseille, soit qu'on ait enquêté tous les ménages des ilots comme cela a été fait à Quito lors de l'enquête bâti/population. En termes techniques, **on n'a donc pratiqué jusqu'à maintenant que des sondage en grappe** : l'unité d'échantillonnage (le pâté de maisons) regroupe un certain nombre d'unités statistiques (les ménages), lesquelles sont toutes enquêtées. Connaissant à Quito, sur l'échantillon de pâtés de maisons enquêté, les caractéristiques démographiques moyennes de la population de l'ilot et celles de chacun des ménages résidents, on peut maintenant s'intéresser à des sondages à plusieurs degrés.

Les arguments les plus souvent avancés en faveur de ce type de sondages concernent d'une part la difficulté qu'il y a à sélectionner directement les unités statistiques lorsqu'on ne dispose pas d'une liste à jour de ces dernières et d'autre part la nécessité de concentrer l'échantillon dans des secteurs géographiques déterminés afin de réduire les temps et les coûts et de transport¹. Le premier argument est déterminant lorsqu'on réalise un sondage démographique; en ce qui nous concerne le second argument l'est moins puisque les coûts et temps de transport restent raisonnables en milieu urbain. **Pour nous, c'est au contraire parce qu'il permet, par rapport au sondage en grappe, d'augmenter le nombre et la dispersion des unités primaires et cela en conservant le même nombre de ménages dans l'échantillon, qu'un sondage à deux degrés peut être profitable.**

L'expression de la variance des estimations sous un tel plan comprend deux termes. Le premier représente l'aléa survenant lors de la sélection des unités primaires et il est fonction de la variabilité de la moyenne du caractère démographique estimé d'un pâté de maisons à l'autre. Le deuxième, dû au second degré du sondage, rend compte de la

¹ : Voir par exemple L. M. ASSELIN 1984 : Techniques de sondage avec application à l'Afrique pp 105 & 106. GAETAN MORIN, Québec. L'auteur ne s'inscrit pas exactement dans la même problématique que nous. L'ouvrage est principalement axé sur la production de statistiques agricoles et les sondages à plusieurs degrés y sont considérés comme une alternative aux sondages aléatoires à un degré sur listes et non pas, comme c'est le cas pour nous, à un sondage en grappe.

variabilité du caractère dans l'ensemble des ménages d'une même unité primaire. Dans chacun des deux termes interviennent évidemment les taux au premier et second degré du sondage. Les sondages en grappes considérés jusqu'à maintenant sont donc en fait des "cas limites" de sondages à deux degrés où, le second degré étant un recensement, le deuxième terme de la variance est nul. On conçoit bien intuitivement que, si le premier terme est déterminant dans l'expression de la variance totale, on a intérêt pour réduire cette dernière tout en conservant un même nombre de ménages enquêtés, à augmenter le taux du premier degré quitte à n'enquêter qu'une fraction des ménages dans chaque ilot de l'échantillon. C'est cette idée que nous allons maintenant formaliser pour quantifier ensuite, à partir des données de l'enquête préliminaire, les gains de précision qu'on peut en attendre et déterminer les taux qu'il convient d'appliquer au premier et au second degré du sondage.

3.3.2 Formulaire, programme informatique

Formulaire

1. Préliminaire : conventions, notations

Dans ce qui va suivre, pour réduire le volume du formulaire, on se limitera aux expressions des variances concernant l'estimation des totaux, l'étude de celles ayant trait aux ratios, en tout point semblable, n'apporterait rien de plus. L'expérience va porter sur deux plans de sondages, lesquels ne diffèrent en fait qu'au premier degré. Le premier plan, qu'on a toujours pris comme référence, est équiprobable au premier comme au second degré (c'est le sondage équiprobable sur liste, classique). Le second, que l'ensemble des résultats obtenus jusqu'ici nous porte à adopter, est le plan où, au premier degré, la probabilité d'appartenance d'une unité à l'échantillon est proportionnelle à sa surface totale, et où la sélection des unités secondaires se fait de manière équiprobable. Pour simplifier les calculs dans le cas du plan à probabilités inégales, on considère que, pour les deux plans considérés, le tirage des unités primaires se fait avec remise ; au second degré en revanche, il s'effectue sans remise. Enfin il ne s'agit pas ici de démontrer l'ensemble des résultats qui conduisent aux expressions des variances dans le cas de ces sondages à deux degrés; ces résultats classiques sont, pour la plus part, repris des ouvrages de référence sur le sujet¹. L'objet des paragraphes suivants est seulement d'en faire **une synthèse qui nous conduira à l'expression de la différence qui nous intéresse c'est à dire celle entre les variances associées respectivement au sondage en grappe et au sondage où l'on n'enquête qu'une fraction des ménages de chaque ilot. C'est sur l'expression de cette différence qu'est construite le programme qui permettra les tests.**

Soient M le nombre total d'unités primaires (pâtés de maisons) contenues dans la base, m_i le nombre de ces unités incluses dans l'échantillon, N_i le nombre d'unités secondaires de l'unité primaire i et n_i le nombre des ménages enquêtés dans cette unité (on note f_2 la fraction de sondage au second degré).

¹ : On peut trouver des exposés complets dans tous les ouvrages classiques cités en bibliographie ; quant à nous, pour cette partie, nous nous sommes servi principalement de l'ouvrage de L. M. ASSELIN, déjà cité, auquel sont en particulier empruntées les notations adoptées.

$V_1()$ sera la variance de l'estimateur employé au premier degré et $V_2()$ celle de l'estimateur employé au second degré.

Enfin on notera z_i la surface de l'unité primaire i et Z la surface totale de la base de sondage. La probabilité π_i que l'unité primaire i soit sélectionnée vaut donc :

$$\pi_i = m_1 z_i / Z.$$

2. Tirage au premier degré équiprobable

Les sélections étant équiprobables au premier comme au second degré, la variance de l'estimateur sans biais du total s'écrit : $V(\hat{Y}) = V_1(\hat{Y}_1) + \frac{M}{m_1} \sum_{i=1}^M V_2(\hat{Y}_i)$, avec

$$V_1(\hat{Y}_1) = \frac{M^2 \sigma^2}{m_1} \quad \text{et} \quad V_2(\hat{Y}_i) = N_i \left(\frac{N_i - n_i}{n_i} \right) S_i^2.$$

Les variances du caractère y , sur l'ensemble de la base de sondage (σ^2) et au sein de chaque unité primaires (S_i^2) ont respectivement les expressions suivantes :

$$\sigma^2 = \frac{\sum_{i=1}^M (Y_i - \bar{Y})^2}{M} \quad S_i^2 = \frac{\sum_{j=1}^{N_i} (y_{ij} - \bar{Y}_i)^2}{N_i - 1}$$

On obtient donc :
$$V(\hat{Y}) = \frac{M^2 \sigma^2}{m_1} + \frac{M}{m_1} \sum_{i=1}^M N_i \left(\frac{N_i - n_i}{n_i} \right) S_i^2.$$

Or, si l'on suppose f_2 constante, on a quelque soit i , $f_2 = n_i / N_i$; donc :

$$\forall i, \quad \frac{N_i - n_i}{m_1 n_i} = \frac{N_i - f_2 N_i}{m_1 f_2 N_i} = \frac{1 - f_2}{m_1 f_2}$$

Donc, en posant :

$$A_1 = M^2 \sigma^2 \quad \text{et} \quad B_1 = M \sum_{i=1}^M N_i S_i^2,$$

on obtient :

$$V(\hat{Y}) = \frac{A_1}{m_1} + \left(\frac{1 - f_2}{m_1 f_2} \right) B_1.$$

3. Tirage au premier degré à probabilités proportionnelles aux surfaces

Le tirage au premier degré s'effectuant maintenant avec des probabilités inégales, la variance de l'estimateur sans biais du total s'écrit :

$$V(\hat{Y}) = V_1(\hat{Y}_1) + \sum_{i=1}^M \frac{V_2(\hat{Y}_i)}{\pi_i} = V_1(\hat{Y}_1) + \frac{Z}{m_1} \sum_{i=1}^M \frac{V_2(\hat{Y}_i)}{z_i}$$

$$\text{avec } V_1(\hat{Y}_1) = \frac{1}{m_1} \times \left(Z \times \left(\sum_{i=1}^M \frac{y_i^2}{z_i} \right) - Y^2 \right) \text{ et toujours,}$$

$$V_2(\hat{Y}_i) = N_i \left(\frac{N_i - n_i}{n_i} \right) S_i^2 \quad \text{et} \quad S_i^2 = \frac{\sum_{j=1}^{N_i} (y_{ij} - \bar{Y}_i)^2}{N_i - 1}$$

$$\text{On a donc : } V(\hat{Y}) = \frac{1}{m_1} \times \left(Z \times \left(\sum_{i=1}^M \frac{y_i^2}{z_i} \right) - Y^2 \right) + Z \times \sum_{i=1}^M \frac{N_i}{z_i m_1} \left(\frac{N_i - n_i}{n_i} \right) S_i^2$$

A nouveau grâce au fait que f_2 est constante, on obtient :

$$V(\hat{Y}) = \frac{1}{m_1} \times \left(Z \times \left(\sum_{i=1}^M \frac{y_i^2}{z_i} \right) - Y^2 \right) + Z \times \sum_{i=1}^M \frac{N_i}{z_i m_1} \left(\frac{N_i - n_i}{n_i} \right) S_i^2$$

Donc, en posant :

$$A_2 = Z \times \left(\sum_{i=1}^M \frac{y_i^2}{z_i} \right) - Y^2 \quad \text{et} \quad B_2 = Z \sum_{i=1}^M \frac{N_i S_i^2}{z_i}$$

on obtient :

$$V(\hat{Y}) = \frac{A_2}{m_1} + \left(\frac{1-f_2}{m_1 f_2} \right) B_2$$

4. Discussion

Dans les deux cas envisagés pour le plan de sondage au premier degré, la variance de l'estimateur sans biais est donc de la forme : $V(\hat{Y}) = \frac{A}{m_1} + \left(\frac{1-f_2}{m_1 f_2} \right) B$

A et B variant suivant la probabilité de sélection retenue. A représente la variance entre unités primaires, due au premier degré du sondage. B représente la variance au sein des unités primaires, due au second degré.

Dans les deux cas également, lorsque $f_2 = 1$, le second terme s'annule et on obtient la variance du sondage en grappe de référence : $V_0(\hat{Y}) = \frac{A}{m_0}$, où m_0 est le nombre de pâtés de maisons enquêtés. On a donc, pour la différence entre les variance des deux sondages à comparer, l'expression suivante :

$$V_0(\hat{Y}) - V(\hat{Y}) = \left(\frac{1}{m_0} - \frac{1}{m_1} \right) A + \left(\frac{1-f_2}{m_1 f_2} \right) B.$$

Pour conserver, lorsque la fraction de sondage au second degré n'est plus égale à 1, le même nombre total de ménages enquêtés, il faut que f_2 vérifie : $m_1 f_2 = m_0$ (1). On a alors : $V_0(\hat{Y}) - V(\hat{Y}) = (1 - f_2) \frac{A}{m_0} - (1 - f_2) \frac{B}{m_0}$, d'où l'expression du gain sur la variance qu'il faudra maximiser :

$$V_0(\hat{Y}) - V(\hat{Y}) = \frac{1 - f_2}{m_0} (A - B)$$

Tout dépend donc du signe et de la valeur de la quantité $(A - B)$. Si B est plus grand que A , alors $V(\hat{Y})$ est supérieur à $V_0(\hat{Y})$ et le sondage en grappe où chaque unité primaire est enquêtée exhaustivement est préférable. En revanche si B est plus petit que A , alors $V(\hat{Y})$ est inférieur à $V_0(\hat{Y})$ et cela d'autant plus que f_2 est petit : c'est, dans ce cas, le sondage à deux degrés qui est le plus efficace et il faut choisir la fraction de sondage au second degré la plus faible possible.

Le principe du test à partir des données collectées lors de l'enquête "Bâti / population" va donc être de **calculer, pour un ensemble de caractères démographiques, les quantités A et B** afin tout d'abord de vérifier que la répartition de la variance entre ses deux composantes est bien semblable pour tous les caractères considérés. Si c'est le cas, **on fixera la fraction de sondage à appliquer au second degré : f_2 et on évaluera, à partir des valeurs de A et B, le gain de précision obtenu par rapport au sondage en grappes.**

Programme informatique

Le programme VAR2DEG (VARIance 2 DEGrés) effectue ces calculs à partir des fichiers de données constitués sous dbase III. Il est écrit dans le langage propre à ce logiciel² et réalise la stricte application du formulaire décrit plus haut : calcul des deux termes (A et B) de la variance d'estimation d'un total sous un plan de sondage à deux degrés. Le programme utilise deux fichiers de données correspondant aux deux degrés du sondage : le fichier des unités primaires (pâtés de maisons) et celui des unités secondaires (ménages). Le plan au premier degré, toujours considéré avec remise, peut être soit équiprobable soit à probabilité proportionnelle à une variable du choix de l'utilisateur dont la donnée doit figurer dans le fichier des unités primaires. Le plan au second degré est toujours équiprobable sans remise.

Pour lancer les calculs, on doit entrer successivement les noms des deux fichiers de données (unités primaires puis secondaires), le nom de la variable à estimer puis celui de la variable définissant la probabilité d'appartenance des unités primaires à

1) : En réalité, même une fois cette condition vérifiée, on n'enquêtera qu'approximativement le même nombre de ménage dans les deux cas, l'égalité n'étant rigoureusement vérifiée qu'en espérance. Cependant lorsque m_0 est grand on ne risque pas de trop mauvaises surprises.

2) : On trouvera le listage complet du programme dans l'annexe 3.

l'échantillon, enfin le nom d'une variable identifiant les unités primaires dans les deux fichiers (le nom de cette variable devant être identique dans les deux fichiers).

Comme pour Varech, la version actuelle de Var2deg est à l'état de maquette. Les deux programmes peuvent évidemment être appliqués aussi bien à des données exhaustives issues de recensement qu'à des données provenant d'enquêtes sur échantillon ; dans ce cas il faut bien sur tenir compte de la représentativité de ces données lorsqu'on formule les conclusions. Ils sont en fait complémentaires, Varech permettant d'effectuer des choix concernant la sélection des unités primaires et l'inférence au premier degré tandis que Var2deg aide à déterminer le taux au second degré du sondage. Dans leurs versions opérationnelles, ces deux programmes devraient donc constituer un outil informatique efficace pour évaluer les possibilités d'utilisation des techniques de sondage sur des sites où existent des données permettant une approche rigoureuse des choix à opérer.

3.3.3 Résultats

Le tableau 3.5 résume les résultats obtenus sur un échantillon de 1329 ménages répartis dans 44 pâtés de maisons enquêtés. Pour chacun des deux plans de sondage au premier degré, il donne la valeur des deux termes de la variance et la part de la variance totale due au premier degré.

Tableau 3.5 : Comparaison des variances dues au premier et second degré du sondage (Programme Var2deg)

Caractères	Equiprobable			Probabilités inégales(p=s)		
	A ₁ (x 10 ⁻⁶)	B ₁ (x 10 ⁻⁶)	$\frac{A_1}{A_1+B_1}$	A ₂ (x 10 ⁻⁶)	B ₂ (x 10 ⁻⁶)	$\frac{A_2}{A_2+B_2}$
EFFECTRES	29,02	0,13	0,99	40,75	0,25	0,99
EFFM5A	0,26	0,02	0,93	0,37	0,04	0,90
EFFPL5A	19,86	0,12	0,99	31,47	0,23	0,99
EFFMIG	4,97	0,11	0,98	7,59	0,24	0,97
EFFNMIG	8,02	0,16	0,98	13,16	0,32	0,98
EFFMIGM5	0,19	0,04	0,83	0,32	0,08	0,80
CEFFMIGM5	21,24	0,15	0,99	33,28	0,31	0,99

On constate que pour l'ensemble des caractères démographiques considérés, la variabilité des caractères entre unités primaires est responsable de plus de 90 % de la variance totale et donc que le sondage à deux degrés permettant d'enquêter un maximum de pâtés de maisons est de beaucoup préférable à un sondage en grappe. De plus, en ce qui concerne les variables

que nous estimerons réellement (en caractères gras dans le tableau), le rapport $A/(A+B)$ est si proche de 1 qu'on peut considérer que la variance totale est entièrement déterminée par le premier terme. Donc, si l'on reprend les notations du formulaire, on a, entre la variance du sondage en grappe de référence et la variance du sondage à deux degrés, la relation linéaire suivante :

$$V(\bar{Y}) \approx f_2 V_0(\bar{Y})$$

Il est donc acquis qu'il faut pratiquer le taux de sondage le plus faible possible au sein des pâtés de maisons, afin de multiplier le nombre d'unités primaires enquêtées. Mais ici intervient une contrainte supplémentaire : il faut pouvoir estimer la variance des estimateurs de façon fiable d'après les données recueillies sur l'échantillon. On sait que dans l'estimateur de cette

variance, pour chaque unité primaire, les quantités Y_i et \bar{Y}_i (total et moyenne du caractère y sur l'unité primaire i) sont remplacées par leurs estimations obtenues à partir de l'échantillon de ménages enquêté dans ce pâté de maisons. Pour que l'estimation de la variance soit correcte il faut donc que le nombre de ménages enquêtés dans chaque unité primaire soit suffisant pour permettre une estimation fiable du total et de la moyenne de y dans l'unité. De ce point de vue, une fraction de sondage au second degré constante, que l'on veut la plus faible possible, risque de poser des problèmes dans les pâtés de maisons peu peuplés. **Donc on enquêtera plutôt un nombre constant de ménages dans toutes les unités primaires**, ce qui correspond à une fraction f_2 variable suivant la taille du pâté de maisons (par taille du pâté de maisons, on entend ici le nombre de ménages qui l'occupent). Ce principe étant admis, reste donc à déterminer le nombre de ménages qu'il faut enquêter dans chaque unité pour obtenir le gain maximum sur la précision du sondage tout en conservant une estimation correcte de la variance des estimateurs. Dans ce but, on va constituer dans l'échantillon des groupes de pâtés de maisons ayant des tailles comparables. Ainsi, pour chaque groupe on pourra évaluer la part de la variance totale qui lui revient¹ et donc, quelle conséquence aura le choix des fractions de sondage au second degré sur la variance des estimateurs.

La distribution des pâtés de maisons et des ménages de l'échantillon selon la taille de l'îlot (en nombre de ménages) est donnée dans le tableau suivant :

Taille (nb de ménages dans l'îlot)	$1 \leq n \leq 10$	$10 < n \leq 20$	$20 < n \leq 30$	$30 < n \leq 40$	$40 < n \leq 50$	$50 < n \leq 60$	$60 < n \leq 80$	$n > 80$
Nb d'îlots	10	16	9	1	3	1	2	2
Nb de ménages	59	241	227	35	137	52	139	439

Pour conserver suffisamment de pâtés de maisons dans chaque groupe on considérera les quatre groupes suivants :

$1 \leq n \leq 10$:	10 Manzanas	59 Ménages
$10 < n \leq 20$:	16 Manzanas	241 Ménages
$20 < n \leq 40$:	10 Manzanas	262 Ménages
$n > 40$:	8 Manzanas	767 Ménages

¹ : voir ci dessous comment on applique pour cela aux différents groupes le principe de sommation des variance dans le cas stratifié.

Total 44 Manzanas 1329 Ménages

Dans le tableau 3.6 figurent les résultats de l'application successive du programme au fichier correspondant à chaque groupe. On n'a retenu que les quatres caractères démographiques que nous devons réellement estimer, pour lesquels on indique le rapport $A/(A+B)$ dans les deux cas de plans de sondages possibles.

Tableau 3.6 : Comparaison des variances dues au premier et second degré du sondage selon la taille des pâtés de maisons (Programme Var2deg)

Taille des îlots	EFFCTRES		EFFPLUS1		EFFNMIG		CEFFMIGM5	
	$\frac{A_1}{A_1+B_1}$ Equip.	$\frac{A_2}{A_2+B_2}$ (P-s)	$\frac{A_1}{A_1+B_1}$ Equip.	$\frac{A_2}{A_2+B_2}$ (P-s)	$\frac{A_1}{A_1+B_1}$ Equip.	$\frac{A_2}{A_2+B_2}$ (P-s)	$\frac{A_1}{A_1+B_1}$ Equip.	$\frac{A_2}{A_2+B_2}$ (P-s)
$1 \leq n \leq 10$	0,85	0,96	0,83	0,95	0,83	0,92	0,88	0,96
$10 < n \leq 20$	0,88	0,98	0,87	0,98	0,82	0,94	0,84	0,97
$20 < n \leq 40$	0,86	0,99	0,88	0,98	0,75	0,96	0,85	0,98
$n > 40$	0,995	0,997	0,994	0,996	0,993	0,995	0,981	0,986

Comme en témoigne le tableau, le terme de la variance dû au premier degré du sondage est bien déterminant quelque soit la taille du pâté de maisons considéré (particulièrement dans le cas du sondage à probabilités inégales). Il reste donc à quantifier la part de chaque groupe de pâtés de maisons dans la variance totale. Pour cela on va considérer la division opérée dans l'échantillon suivant la taille des pâtés de maisons comme une stratification. Dans ce cas la variance totale apparait comme la variance d'une somme de variables aléatoires indépendantes et on a :

$$V(\hat{Y}) = \sum_{k=1}^L V(\hat{Y}_k).$$

Si l'on considère, d'après les résultats précédents, que seul compte le premier terme, tant pour la variance totale qu'à l'intérieur de chaque "strates", on peut écrire :

$$A/n \approx \sum_{k=1}^L A_k/n_k.$$

Comme $n=1/fN$ et, pour tout k , $n_k=1/fN_k$ on obtient : $A/fN \approx \sum_{k=1}^L A_k/fN_k$.

C'est à dire :

$A/N \approx \sum_{k=1}^L A_k/N_k$

Le tableau 3.7, basé sur cette formule appliquée aux données recueillies sur les 44 pâtés de maison de l'échantillon, fournit une évaluation des variances obtenues avec les deux stratégies en concurrence :

1. le sondage en grappe où f_2 vaut 1 pour tous les pâtés de maisons sélectionnés
2. le sondage à deux degrés où l'on enquête exhaustivement les pâtés de maisons contenant jusqu'à 10 ménages et où l'on n'enquête qu'un échantillon de dix ménages dans tous les autres pâtés de maisons.

Tableau 3.7 : Variances des estimateurs, décomposées selon la taille des pâtés de maisons, pour deux stratégies de sondage au second degré (Programme Var2deg)

Taille (nb de ménages)	$1 \leq n \leq 10$	$10 < n \leq 20$	$20 < n \leq 40$	$n > 80$	Total
1) Caractéristiques de l'échantillon					
Nb d'îlots	10	16	10	8	44
Nb de ménages	59	241	262	767	1329
Contrib. à la variance : A_i/N_i	18232,3	66086,4	74786,1	757434,49	916539,4
Contrib. à la variance (%)	2,0	7,2	8,2	82,6	100
2) Sondage en grappes					
Fraction de sondage	1	1	1	1	1
Nb de ménages à enquêter	178	725	788	2309	4000
Nb d'îlots prévues	30	48	30	24	132
Contrib. à la variance	$0,02 V_0$	$0,07 V_0$	$0,08 V_0$	$0,83 V_0$	V_0
3) Sondage à deux degrés (f_2 variable)					
Fraction de sondage moyenne	1	10/15,1	10/26,2	10/95,9	
Nb de ménages à enquêter	177	730	790	2310	4007
Nb d'îlots prévues	30	73	79	231	413
Gain sur la variance (%)	0	34	62	90	80
Contrib. à la variance	$0,02 V_0$	$0,05 V_0$	$0,03 V_0$	$0,09 V_0$	$0,19 V_0$

Les conclusions à tirer de l'expérience apparaissent clairement à l'examen du tableau 3.7 :

- 1) Les pâtés de maisons les plus peuplés (contenant plus de 40 ménages) ont une contribution déterminante à la variance : à eux seuls ils fournissent plus de 80% du total. C'est donc sur cette fraction de l'échantillon qu'il importe d'avoir le taux au second degré le plus faible.

2) **A l'inverse la contribution du groupe des pâtés de maisons peuplés de moins de dix ménages est presque négligeable : 2% de la variance totale. On ne peut donc pas espérer de gain appréciable de précision en n'enquêtant que partiellement ces îlots. D'autre part on a vu que pour obtenir une bonne estimation de la variance il convenait d'avoir un nombre suffisant de ménages enquêtés dans chaque unité primaire. Nous enquêterons donc tous les ménages des pâtés de maisons de ce groupe.**

3) Pour les pâtés de maisons contenant plus de dix ménages, un effectif enquêté constant de dix ménages est un bon choix. En premier lieu, il est suffisant pour éviter toute mauvaise surprise lors de l'estimation de la variance des estimateurs. D'autre part il permet de faire décroître la fraction de sondage au fur et à mesure qu'augmente le nombre de ménages vivant dans le pâté de maisons pour arriver à une fraction d'un dixième en moyenne dans le groupe des pâtés de maisons les plus peuplés. **Ainsi l'application de cette stratégie au second degré produit, pour un même nombre total de ménages enquêtés, une division par cinq de la variance par rapport au sondage en grappe** (le coefficient de variation lui, sera divisé par 2,25).

3.3.4 Conclusions

Au début de cette expérience visant à la détermination du second degré du sondage, nous avons posé le problème suivant :

Dans un sondage à deux degrés, l'expression de la variance comporte deux termes correspondant aux aléas dus à chacun des degrés ; si le terme correspondant à la sélection des unités primaires est déterminant dans l'expression de la variance totale, on a intérêt, pour réduire cette dernière tout en conservant un même nombre de ménages enquêtés, à augmenter le taux du premier degré quitte à n'enquêter qu'une fraction des ménages dans chaque îlot de l'échantillon.

Le programme Var2deg a permis le calcul de ces deux termes de la variance. On a alors constaté que la variabilité des caractères entre unités primaires est responsable de plus de 90 % de la variance totale et donc que le sondage à deux degrés permettant d'enquêter plus de pâtés de maisons est bien préférable à un sondage en grappe où les pâtés de maisons sont enquêtés exhaustivement. On doit donc d'une part choisir un taux au second degré le plus faible possible pour optimiser la précision du sondage, et d'autre part conserver un nombre suffisant de ménages dans chaque unité primaire pour pouvoir estimer précisément la variance des estimateurs. Ce double objectif est satisfait lorsqu'on enquête un nombre constant de ménages dans chaque pâté de maisons. En effet, lorsqu'on décompose le calcul de la variance pour évaluer la contribution de groupes de pâtés de maisons également peuplés, on constate que les pâtés de maisons les plus peuplés (contenant plus de 40 ménages) fournissent à eux seuls plus de 80% de la variance totale. En revanche, la contribution du groupe des pâtés de maisons peuplés de moins de dix ménages est négligeable. C'est donc sur la fraction de l'échantillon comprenant les pâtés de maisons les plus peuplés qu'il importe d'avoir le taux au second degré le plus faible, tandis qu'on peut enquêter exhaustivement les pâtés de maisons les moins peuplés.

Le résultat de l'expérience est donc la définition de la stratégie de sélection des ménages au sein des pâtés de maisons choisis au premier degré : on enquêtera tous les ménages des pâtés de maisons contenant jusqu'à 10 ménages et seulement un échantillon de dix ménages dans tous les autres pâtés de maisons. Enfin le programme mis au point, appliqué aux données recueillies sur l'échantillon de l'enquête préliminaire "Bâti-population", permet de comparer la précision du sondage ainsi défini avec celle du sondage en grappe de référence : l'application de cette stratégie

au second degré produit, pour un même nombre total de ménages enquêtés, une division par cinq de la variance.

A l'issue des expériences menées d'une part sur les données du recensement de 1982 à Marseille et d'autre part sur les données provenant des enquêtes préliminaires réalisées à Quito, nous possédons maintenant tous les éléments pour élaborer un plan de sondage complet. Le chapitre suivant décrit d'abord ce plan dans son ensemble : étape de sélection des unités primaires et secondaires et étape d'inférence avec les formules d'estimation correspondant au premier et deuxième degré du sondage. Puis il indique comment ce plan a été mis en œuvre pour réaliser l'enquête démographique finale.

CHAPITRE 4

LE PLAN DE SONDAGE ET LA MISE EN ŒUVRE DE L'ENQUÊTE

4.1 : Introduction; le double objectif méthodologique et thématique de l'enquête, présentation du questionnaire

L'enquête finale "Démographie Migration" que nous avons réalisée à Quito possède un double objectif. Tout d'abord, sur un plan méthodologique, elle doit permettre de vérifier que la méthode mise au point possède bien les caractéristiques de rapidité, faible coût et précision que nous lui attribuons au terme de la phase d'expérience. Toujours sur ce même plan méthodologique il importe également de savoir si la technique de sondage employée demeure efficace en dehors du domaine thématique restreint des indicateurs démographiques les plus classiques. C'est une des raisons pour lesquelles nous avons choisi d'aborder dans le questionnaire de l'enquête finale plusieurs thèmes.

Ensuite l'enquête vise à produire une information fiable sur chacun des deux grands thèmes que permet de traiter le questionnaire (présenté en annexe 4) : caractéristiques démographiques générales de la population quiténienne et systèmes de résidence. La première partie du questionnaire (fiche 1) permet de saisir les caractéristiques démographiques et socio-économiques de chacune des personnes appartenant au ménage enquêté ; en dehors de modifications mineures, elle correspond à la fiche d'enquête utilisée en Avril 1987 pour l'enquête sur les relations bâti/population. La fiche 2 permet d'analyser les statuts d'occupation du logement et les modes d'accès à la propriété. Les fiches 3 et 4 sont centrées sur la question des systèmes résidentiels. Si la fiche 4 correspond à un recueil classique de biographies migratoires et professionnelles avec un seuil de durée de résidence fixé à 6 mois, la fiche 3 vise à cerner la pratique des chefs de ménage en matière de systèmes résidentiels et d'activités : cette fiche 3 correspond au recueil *intégral* de tous les changements de résidence et d'activité sur une période de 2 ans, sans notion de durée minimum. Enfin, par la fiche 5 relative aux caractéristiques d'activité et de résidence des membres de la famille ne vivant pas dans le ménage enquêté, nous tentons de replacer les pratiques résidentielles et professionnelles des chefs de ménage enquêtés dans le groupe social, éventuellement spatialement éclaté, dans lequel elles s'inscrivent.

4.2 : Description du plan de sondage, étapes de sélection et d'inférence

Au cours des expériences menées durant toute cette recherche à Marseille et à Quito, on a formulé un ensemble de conclusions qui vont maintenant trouver leur application pratique dans la définition du plan de sondage retenue pour l'enquête "démographie migration". Ces conclusions ont été exposées, au fur et à mesure de leur obtention, dans les différents rapports intermédiaires rédigés, et jusque dans les chapitres précédents du présent rapport. Pouriant, il nous semble opportun, avant de décrire dans le détail le plan de sondage, de faire une synthèse rapide des principaux résultats, en les distinguant suivant qu'ils s'appliquent à l'une ou l'autre des étapes de la définition du plan.

4.2.1 : Résumé des résultats obtenus lors des phases d'expériences

Base de sondage et stratification : parmi les stratifications actuellement disponibles sur l'image satellite, la plus efficace est une stratification en six niveaux de densité dont le premier niveau permet d'isoler les espaces non construits, exclus de la base. Dans la base de sondage, on aura donc cinq strates.

Probabilité d'appartenance des unités primaires à l'échantillon : la surface bâtie ou, plus exactement, l'approximation que l'on peut en obtenir à partir de l'information satellitaire, est une mesure de la taille des unités primaires bien meilleure que leur surface totale au sens de la précision du sondage à but démographique; mais, **une fois la base de sondage stratifiée sur le critère de densité du bâti, une probabilité de sélection proportionnelle à la surface totale suffit à obtenir des précisions équivalentes.**

Inférence au premier degré : les deux mesures de taille évoquées ci-dessus, et qui sont pour le moment les seules disponibles, ne sont ni l'une ni l'autre suffisamment corrélées aux variables démographiques pour qu'il y ait intérêt à les intégrer comme information exogène dans des estimateurs par le ratio. **On retient donc des estimateurs sans biais pour les totaux et l'estimation par le ratio pour les quotients.**

Règle d'allocation de l'échantillon aux strates : Grâce à une bonne stabilité des des règles d'allocation optimale, quelque soient les caractères estimés, **on peut définir une allocation aux strates qui améliore l'ensemble des estimations** (voir, pour les taux dans chaque strate, la description de l'étape de sélection des unités primaires)

Second degré du sondage : l'alternative retenue est un sondage à deux degrés avec comme unités primaires les pâtés de maisons et comme unités secondaires les ménages. Le test effectué sur les données provenant de l'enquête préliminaire "bâti/population" réalisé à Quito conclue à l'intérêt de pratiquer le taux le plus fort possible au premier degré, dans la limite du budget disponible et sous la contrainte d'avoir un nombre suffisant de ménages dans chaque pâté de maisons.

Spécificités du plan envisagé : Par rapport aux sondages les plus couramment pratiqués, par exemple les sondages aléatoires équiprobables sur listes, le sondage que nous proposons présente deux particularités principales :

1 : En tant que sondage à probabilités inégales, il nécessite, pour mener correctement l'inférence à partir de l'échantillon, la connaissance d'**informations exogènes** :
 - la taille (surface) de chaque unité primaire (pâté de maisons) enquêtée; il s'agit de la surface totale et il faut donc disposer d'un moyen fiable de la mesurer.
 - La taille (surface) de chaque strate; sa mesure sera faite sur l'image satellite.

2 : Le caractère systématique du sondage, en assurant une bonne répartition géographique de l'échantillon, **améliore la précision des estimations**. On évalue ce gain à 20 ou 30 % de la variance du sondage strictement aléatoire au même taux. Si l'on considère l'effet cumulé du tirage systématique et de la stratification, l'allocation aux strates étant optimale, le gain total se situe, suivant les caractères démographiques estimés, entre 45 et 95 % de la variance de départ (tirage aléatoire non stratifié).

L'ensemble de ces conclusions conduisent donc au plan de sondage suivant.

4.2.2 : Description du plan de sondage : étapes de sélection au premier et second degré

1: Base de sondage

La base de sondage est l'ensemble de l'aire métropolitaine matérialisée sur l'image satellite par la limite de la zone considérée comme urbaine; à l'intérieur de cet ensemble, les zones non bâties sont exclues.

2 : Stratification

Cet espace est stratifié en cinq niveaux de densité du bâti (stratification obtenue par classification de l'image satellite).

3 : Tirage au premier degré

Les unités primaires sont les pâté de maisons. Au premier degré nous pratiquons un sondage aréolaire systématique, la sélection de l'échantillon se faisant à l'aide d'une grille de points. Lorsque le placement initial de la grille sur l'image est aléatoire, on peut considérer que la probabilité d'appartenance à l'échantillon d'une unité spatiale donnée est proportionnelle à sa surface. La maille de la grille sera de taille variable pour respecter les taux dans chaque strate fixes par la règle d'allocation; ces taux sont donnés dans le tableau suivant.

strates	taux approximatifs au premier degré (en % du nombre de manzanas de la strate)	Coef. appliqué à la strate
1 : densité très faible	2,5	0,4
2 : densité faible	2,2	0,3
3 : densité moyenne	6,7	1
4 : densité forte	8,9	1,5
5 : densité très forte	4,4	0,9
Ensemble de la base	5,8	1

4. Tirage au deuxième degré

Les unités secondaires sont les ménages, la sélection est là encore systématique; elle s'effectue sur liste après recensement de tous les ménages vivants dans l'ilot. La fraction de sondage f_2 au second degré est variable suivant le nombre de ménages vivants dans l'ilot :

- moins de dix ménages dans le pâté de maisons : enquête exhaustive ($f_2 = 1$)
- plus de dix ménages dans le pâté de maisons : dix ménages sont enquêtés (f_2 décroît lorsque la taille du pâté de maison augmente)

Le taux final obtenu sera, nous le verrons, inférieur à 1,5 % des ménages.

4.2.3 : Formulaire d'estimation pour l'étape d'inférence

A : Préliminaire, les problèmes posés par l'inférence dans le cas du sondage systématique

Comme on l'a déjà dit dans le précédent rapport intermédiaire¹, les sondages systématiques posent un double problème théorique s'agissant de l'inférence statistique.

Premièrement, les probabilités d'appartenance à l'échantillon sont-elles suffisamment contrôlées pour que le sondage soit justiciable du calcul des probabilités.

Deuxièmement, sachant que pour ce type de sondage, il n'existe pas, dans le stricte cadre de la théorie, d'expression rigoureuse de la variance des estimations, quelle solution retenir pour l'estimation de la variance.

En ce qui concerne le premier problème, J. DESABIE le pose en ces termes : " Le sondage systématique est effectivement un sondage aléatoire, à condition que la base de la progression arithmétique soit effectivement déterminée par tirage au sort"². L'auteur formule cette règle à propos du sondage systématique dans une liste et, lorsque la condition est satisfaite, le sondage est assimilé à un sondage équiprobable sans remise. Il n'y a aucun obstacle à la transposition de cette règle au cas du sondage systématique aréolaire à l'aide d'une grille de points. Si, et c'est bien le cas ici, le premier point de la grille a été placé aléatoirement, le sondage est équivalent, en ce qui concerne les formules d'estimation des totaux, moyennes, quotients etc ..., à un sondage à probabilités inégales, proportionnelles à la surface des unités de la base.

Le second problème est généralement formulé ainsi³ : le sondage systématique est un cas particulier de sondage en grappe, sa variance dépend donc de la variance inter-grappes; comme on ne sélectionne qu'une seule grappe dans l'échantillon, il est impossible de calculer un estimateur non biaisé de la précision. Pour lever cet obstacle on trouve, dans la littérature, trois types de solutions⁴:

1. Négliger l'influence sur la précision du caractère systématique du sondage en l'assimilant au sondage aléatoire qui lui correspond.

J.M. GROBRAS⁵ par exemple dit, à propos du sondage systématique sur liste : "Si le rangement des unités dans le fichier peut être considéré comme aléatoire, ou tout au moins sans corrélation avec la variable étudiée, le tirage systématique est rigoureusement assimilable à un sondage à probabilité égales sans remise et donc de même efficacité (même espérance, même variance)". Cette assimilation peut donc être plus ou moins justifiée.

¹ : Voir second rapport intermédiaire à l'ORSTOM d'O. Barbary : Mise au point de plan de sondage pour l'estimation de caractères démographique à Marseille, Janvier 88, Quito, 184 pages. p 10 et p 131

² : J. DESABIE, 1966 : Théorie et pratique des sondages, DONOD, Paris. p 98 (

³ : notamment par L.M. ASSELIN (Techniques de sondages avec application à l'Afrique, p 49), J. DESABIE (Théorie et pratique des sondages, p 98 et p 171) ou J.M. GROBRAS (Méthodes statistiques des sondages p29 et 30).

⁴ : On ne retient pas dans cet inventaire les solutions reposant sur l'application de la théorie des processus stochastiques ou fonction aléatoire développées récemment (voir par exemple l'article de X.GUIDICELLI, J.P. LANLY, J.B. OUAKAM et M. PIETRI, 1972 dans Annales des Sciences Forestières, INRA). Pour une justification de ce choix, voir l'introduction au premier rapport intermédiaire à l'ORSTOM (Quito, Janvier 1988, p 9).

⁵ : J.M. GROBRAS : Méthodes statistiques des sondages, 1987, ECONOMICA, PARIS; p 29.

Elle l'est parfaitement dans le cas du second degré du sondage que nous utiliserons : la sélection systématique des ménages à enquêter dans le pâté de maisons peut être considérée comme aléatoire sans que l'on risque de sur ou sous estimer la variance intra unité primaire. C'est donc toujours la solution que nous retiendrons pour l'étape d'estimation de totaux (moyennes ou quotients) dans les unités primaires et pour estimer la variance de ces estimations : le sondage sera alors considéré strictement équivalent à un sondage équiprobable sans remise.

Le problème est sensiblement différent pour le premier degré. Le tirage de l'échantillon de pâtés de maisons devra être assimilé à un tirage à probabilités inégales, proportionnelles aux surfaces des unités. Les expériences faites à Marseille et à Quito ont démontré que la bonne répartition géographique de l'échantillon permise par le tirage systématique à l'aide d'une grille de points améliore nettement l'efficacité du sondage. **Il est donc certain que si l'on néglige cette caractéristique du sondage dans le formulaire d'estimation des variances, on surestimera les variances des estimations.** Les deux autres solutions visent à réduire cette surestimation.

2. Envisager le sondage systématique du premier degré comme un sondage stratifié où l'on sélectionne deux unités par strate.

Cette méthode, proposée par J. DESABIE¹, et que nous avons déjà utilisé pour évaluer l'apport du tirage systématique², pourra fournir une bonne approximation des variances des estimations mais risque de conduire à des calculs assez lourds.

3. Utiliser des méthodes robustes d'estimation de variances.

Lorsque, comme c'est le cas ici, les méthodes classiques d'estimation de variance sont mises en défaut, on peut utiliser les méthodes dites "robustes" qui permettent de calculer des intervalles de confiance même si l'on ignore les lois de probabilité des estimateurs. On peut donc, en utilisant de telles méthodes, faire abstraction de la structure probabiliste du sondage et en particulier la technique de sélection des unités d'échantillonnage importe peu. Parmi ces méthodes on retiendra la plus couramment employée, la méthode du Jackknife, dont on trouvera un exposé pratique dans le livre de J.M. GROBRAS : *Méthodes statistiques des sondages*³.

Pour ne pas trop alourdir l'exposé dans le paragraphe suivant, le formulaire n'est donné que pour la première solution envisagée; c'est d'ailleurs celle que l'on appliquera systématiquement aux variables démographiques lors de la phase d'exploitation des résultats de l'enquête. Pour les autres méthodes on se reportera aux ouvrages cités en référence.

B : formulaire

Conformément aux conclusions du paragraphe précédent et aux résultats obtenus durant toute cette recherche, le formulaire d'estimation correspond à la structure de sondage suivante :

Premier degré : sondage aléatoire stratifié avec probabilités de sélection proportionnelles aux surfaces des unités primaires.

¹ : J. DESABIE, 1966 : *Théorie et pratique des sondages*, DONOD, Paris. Chap. 7.7, p 171 : Le tirage systématique envisagé comme un sondage stratifié.

² : Voir second rapport intermédiaire à l'ORSTOM d'O. Barbary : *Mise au point de plan de sondage pour l'estimation de caractères démographiques à Marseille*, Janvier 88, Quito, 184 pages. Chap. 7, P 117 : Tentative d'évaluation de l'efficacité des tirages aréolaires assurant une bonne répartition géographique des unités sondées.

³ : J.M. GROBRAS : *Méthodes statistiques des sondages*, 1987, ECONOMICA, PARIS; P 329.

Second degré : sondage aléatoire équiprobable; recensement dans les petites unités primaires (jusqu'à dix ménages), dix ménages enquêtés dans toutes les autres unités.

Estimateurs employés : Aux deux degrés du sondage on emploie des estimateurs sans biais pour l'estimation des totaux ou des moyennes et des estimateurs par le ratio pour l'estimation directe des quotients.

Bien sûr, les résultats que fourniera ce formulaire sont des approximations, d'ailleurs en ce qui concerne les variances et leurs estimations. Comme on l'a déjà dit, l'assimilation de la procédure de sélection systématique sans remise à un plan de sondage aléatoire avec remise au premier degré doit logiquement avoir pour conséquence la surestimation des erreurs d'échantillonnage et donc une évaluation pessimiste de la précision du sondage. C'est néanmoins la seule solution compatible avec l'équipement logiciel dont nous disposons parcequ'elle ne réclame pas un travail de programmation trop important; c'est donc celle qui sera retenue, au moins dans la première étape d'exploitation des résultats de l'enquête. Par la suite il serait intéressant, si les moyens en temps et programmes informatiques le permettent, d'appliquer aux résultats des méthodes de calcul de variance mieux adaptées au sondage mis au point, en particulier pour tenter d'évaluer l'apport de la sélection systématique. Sur ce point, les travaux préparatoires réalisés à Marseille et à Quito pourront servir de base à une première approche et cela même si l'on ne dispose pas de programmes informatiques supplémentaires qui permette d'utiliser les méthode "robustes" type JACKKNIFE ou BOOT STRAP par exemple.

Plutôt qu'une présentation "littéraire", qui alourdirait beaucoup un exposé qui se veut avant tout pratique, nous avons choisi de présenter le formulaire sous forme synthétique dans les trois tableaux ci dessous. Cette présentation vise également à permettre par la suite une référence facile à la bonne formule, pour chaque cas de figure. Le tableau 4.1 fixe les notations retenues ici, qui sont celles classiquement employées en sondage; pour notre part nous les avons empruntées à l'ouvrage de L.M. ASSELIN¹. Les tableaux 4.2 et 4.3 donnent ensuite les formules à employer aux différentes étapes de l'inférence, pour les trois types de caractères envisagés.

¹ : L.M. ASSELIN 1984 : Techniques de sondages avec application à l'Afrique, Gaetan Morin éditeur, QUEBEC. Les notations sont exposées dans les chapitres 2 et 3.

Tableau 4.1 : Notations adoptées dans le formulaire

Objets	Notations	Significations
1) Base de sondage et échantillon	N	Nombre d'unités secondaires dans l'unité primaire j
	n	Nombre d'US de l'UP j contenus dans l'échantillon
	M	Nombre d'UP dans la strate k
	m	Nombre d'UP de la strate k contenus dans l'échantillon
2) Variables dont on doit estimer le total ou la moyenne (y) ou intervenant au dénominateur du quotient à estimer (x)	y_i, x_i	Valeurs des caractères pour l'US i
	y_j, x_j	Totaux des caractères dans l'UP j
	\bar{y}_j, \bar{x}_j	Moyennes des caractères dans l'UP j
	\hat{y}_j, \hat{x}_j	Estimateurs des totaux dans l'UP j
	$\hat{\bar{y}}_j, \hat{\bar{x}}_j$	Estimateurs des moyennes dans l'UP j
	Y_k, X_k	Totaux des caractères dans la strate k
	\bar{Y}_k, \bar{X}_k	Moyennes des caractères dans la strate k
	\hat{Y}_k, \hat{X}_k	Estimateurs des totaux dans la strate k
	$\hat{\bar{Y}}_k, \hat{\bar{X}}_k$	Estimateurs des moyennes dans la strate k

Tableau 4.1 (suite) : Notations adoptées dans le formulaire

Objets	Notations	Significations
Variable définissant la probabilité d'inclusion dans l'échantillon des unités primaires (z)	z_j Z_k	Surface totale de l'UP j Surface totale de la strate k
Quotients à estimer	r_j \hat{r}_j R_k \hat{R}_k	Valeurs du quotient pour l'UP j ($r_j = y_j/x_j$) Estimateurs du quotient dans l'UP j ($\hat{r}_j = \hat{y}_j/\hat{x}_j$) Valeurs du quotient pour la strate k ($R_k = Y_k/X_k$) Estimateurs du quotient dans la strate k ($\hat{R}_k = \hat{Y}_k/\hat{X}_k$)
Somme	$\sum_{i=1}^N$ $\sum_{i=1}^n$ $\sum_{j=1}^M$ $\sum_{j=1}^m$	Somme sur toutes les US de l'UP j Somme sur les US de l'UP j appartenant à l'échantillon Somme sur toutes les UP de la strate k Somme sur les UP de la strate k appartenant à l'échantillon

Tableau 4.1 (fin) : Notations adoptées dans le formulaire

Objets	Notations	Significations
Erreurs d'échantillonnage et estimation des erreurs	$V(\hat{y}_j)$ et $V(\hat{\bar{y}}_j)$	Variances des estimateurs de totaux et de moyennes dans l'UP j (second degré du sondage)
	$\hat{V}(\hat{y}_j)$ et $\hat{V}(\hat{\bar{y}}_j)$	Estimations des variances des Estimateurs de totaux et de moyennes dans l'UP j (second degré du sondage)
	$E(\hat{r}_j - r_j)^2$	Ecart quadratique moyen de l'estimateur de quotient dans l'UP j (second degré du sondage)
	$\hat{E}(\hat{r}_j - r_j)^2$	Estimation de l'écart quadratique moyen de l'estimateur de quotient dans l'UP j (second degré du sondage)
	$V(\hat{Y}_k)$ et $V(\hat{\bar{Y}}_k)$	Variances des estimateurs de totaux et de moyennes dans la strate k (premier degré du sondage)
	$\hat{V}(\hat{Y}_k)$ et $\hat{V}(\hat{\bar{Y}}_k)$	Estimations des variances des Estimateurs de totaux et de moyennes dans la strate k (premier degré du sondage)
	$E(\hat{R}_k - R_k)^2$	Ecart quadratique moyen de l'estimateur de quotient dans la strate k (premier degré du sondage)
	$\hat{E}(\hat{R}_k - R_k)^2$	Estimation de l'écart quadratique moyen de l'estimateur de quotient dans la strate k (premier degré du sondage)

Tableau 4.2 : Formulaire d'estimation de totaux, moyennes et quotients, pour une unité primaire j (deuxième degré).

	Estimateur	Variance	Estimateur de la variance
Second degré : dans l'unité primaire j, sondage systématique équiprobable assimilé à un sondage aléatoire équiprobable sans remise.			
Total	$\hat{y}_j - N \sum_{i=1}^n \frac{y_i}{n}$	$v(\hat{y}_j) = \frac{N(N-n)}{n} \sum_{i=1}^n \frac{(y_i - \bar{y}_j)^2}{(N-1)}$	$\hat{v}(\hat{y}_j) = \frac{N(N-n)}{n} \sum_{i=1}^n \frac{(y_i - \bar{y}_j)^2}{(n-1)}$
Moyenne	$\hat{\bar{y}}_j = \sum_{i=1}^n \frac{y_i}{n}$	$v(\hat{\bar{y}}_j) = \frac{N-n}{nN} \sum_{i=1}^n \frac{(y_i - \bar{y}_j)^2}{(N-1)}$	$\hat{v}(\hat{\bar{y}}_j) = \frac{N-n}{nN} \sum_{i=1}^n \frac{(y_i - \bar{y}_j)^2}{(n-1)}$
Ratio : $r_j = y_j/x_j$	<p>Estimateur</p> $\hat{r}_j = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}$ <p>Ecart quadratique</p> $E(\hat{r}_j - r_j)^2 = \frac{N(N-n)}{n(N-1)x_j^2} \sum_{i=1}^n (y_i - r_j x_i)^2$ <p>Estimateur de l'écart quadratique</p> $\hat{E}(\hat{r}_j - r_j)^2 = \frac{N(N-n)}{n(n-1)\hat{x}_j^2} \sum_{i=1}^n (y_i - \hat{r}_j x_i)^2$		

Tableau 4.3 : Formulaire d'estimation de totaux, moyennes et quotients, pour une strate donnée k, négligeant la variance intra unités primaires. (premier degré)

	Estimateur	Variance	Estimateur de la variance
Premier degré : sélection des unités primaires systématique, avec probabilité proportionnelle à leur surface, assimilée à un sondage aléatoire à probabilité inégale avec remise			
Total	$\hat{Y}_k = \frac{Z_k}{m} \sum_{j=1}^m \frac{\hat{y}_j}{z_j}$	$V(\hat{Y}_k) = \frac{Z_k}{m} \left(\left(\sum_{j=1}^M \frac{y_j^2}{z_j} \right) - Y_k^2 \right)$	$\hat{V}(\hat{Y}_k) = \frac{1}{m(m-1)} \sum_{j=1}^m \left(\frac{Z_k y_j}{z_j} - \hat{Y}_k \right)^2$
Moyenne	$\hat{\bar{Y}}_k = \frac{Z_k}{Mm} \sum_{j=1}^m \frac{\hat{y}_j}{z_j}$	$V(\hat{\bar{Y}}_k) = \frac{Z_k}{M^2 m} \left(\left(\sum_{j=1}^M \frac{y_j^2}{z_j} \right) - Y_k^2 \right)$	$\hat{V}(\hat{\bar{Y}}_k) = \frac{1}{M^2(m^2-m)} \sum_{j=1}^m \left(\frac{Z_k y_j}{z_j} - \hat{\bar{Y}}_k \right)^2$
Ratio : $R_k = \frac{Y_k}{X_k}$	Estimateur $\hat{R}_k = \frac{\sum_{j=1}^m \frac{\hat{y}_j}{z_j}}{\sum_{j=1}^m \frac{\hat{x}_j}{z_j}}$	Ecart quadratique $E(\hat{R}_k - R_k)^2 = \frac{Z_k}{mX_k^2} \left(\sum_{j=1}^M \left(\frac{y_j^2}{z_j} + R_k^2 \frac{x_j^2}{z_j} - 2R_k \frac{x_j}{z_j} \right) - Y_k^2 - R_k^2 X_k^2 + 2R_k X_k Y_k \right)$	Estimateur de l'écart quadratique $\hat{E}(\hat{R}_k - R_k)^2 = \frac{1}{m(m-1)} \sum_{j=1}^m \left(\left(\frac{Z_k y_j}{z_j} - \frac{Z_k R_k x_j}{z_j} \right) - \hat{Y}_k + \hat{R}_k \hat{X}_k \right)^2$

4.3 : Mise en œuvre du plan de sondage : sélection de l'échantillon et organisation du travail de collecte

Nous abordons maintenant, dans l'ordre chronologique de leur réalisation, les différentes étapes de la mise en œuvre pratique de la méthode : définition de la base de sondage, sélection de l'échantillon au premier et second degré et réalisation des opérations de collecte sur le terrain. Les techniques employées à chaque étape respectent le plus rigoureusement possible le plan de sondage défini ci dessus.

On fera très souvent appel, par la suite, aux résultats obtenus par les autres membres de l'équipe de recherche qui ont entièrement réalisé les deux premières étapes : définition et stratification de la base de sondage. Pour avoir des informations supplémentaires sur les méthodes et techniques employées le lecteur se reportera aux rapports et articles rédigés sur ces questions¹.

4.3.1 : Définition de la base de sondage sur l'image: la limite urbaine

A : Les documents utilisés

Plusieurs documents photographiques réalisés à l'Atelier de télédétection de Bondy ont servi pour parvenir à la délimitation finale de la zone urbaine :

- deux images panchromatiques SPOT :
 - Juin échelle 1/30 000 (2 tirages noir et blanc différemment contrastés afin de tenir compte de la dynamique entre les différents espaces végétaux et urbanisés) ;
 - Août échelle 1/50 000 (2 tirages noir et blanc différemment contrastés) ;
- un indice de végétation échelle 1/50 000 (2 tirages noir et blanc différemment contrastés) ;
- un écart-type local réalisé sur l'image panchromatique, copie couleur d'écran, échelle approximative 1/60 000 ;
- 1 indice de brillance , copie couleur d'écran, échelle approximative 1/60000
- 1 coefficient de densité du bâti en 6 modalités, copie couleur d'écran, échelle approximative 1/60 000.

B: Mise en œuvre et règles de décision

Quelques règles simples ont été discutées, puis retenues afin de définir, de façon la plus précise possible, le concept de "limite urbaine" applicable dans le cas d'une stratification à objectif démographique. Nous avons décidé que la limite serait tracée là où se terminait **une zone continue d'espace bâti**, en suivant si possible une limite (naturelle ou artificielle) visible sur l'ensemble des documents et sur le terrain. Une telle définition, volontairement très large, doit permettre d'inclure suffisamment d'espace pour pouvoir prendre en considération l'extension de la ville entre la date d'enregistrement de l'image (Juin 1986) et la date de réalisation de l'enquête démographique (Novembre 1987). Les lotissements prévus, repérables uniquement grâce à l'ébauche de leur réseau de voirie, ont logiquement été inclus à l'intérieur de notre limite. **Un espace au bâti très peu dense et éparse devrait être délimité en se basant sur des limites visibles (routes, rangées d'arbres, ravines, etc.) et non pas le long de la dernière maison afin de ne pas induire une densité artificiellement élevée dans les flots périphériques.** Pour la même raison, les incursions profondes du contour à l'intérieur de la ville

¹ : Par exemple, pour une première approche et une bibliographie sur le thème, on peut se reporter à la communication faite au colloque PIXIM par A. MICHEL, F.DUREAU, B. LORTIC, O. BARBARY, M.SOURIS : Segmentation et classification d'une image satellite en milieu urbain, 17 pages, Paris 1988.

devaient être limitées et n'être prise en considération que lorsque cela permettait de détourner une zone importante d'espace non bâti.

Les différentes limites obtenues indépendamment par deux personnes (F. DUREAU et A. MICHEL) ont été confrontées. Leurs tracés respectifs n'offrent pas de différences fondamentales. Une limite de "synthèse" a donc été retenue. La discrimination de la limite externe s'est faite pour l'essentiel sur l'image panchromatique brute, mais aussi à partir de l'image de l'indice de végétation, de l'écart-type local et d'un coefficient de densité du bâti. Enfin, nous sommes allés vérifier la limite ainsi obtenue par une enquête sur le terrain, qui nous a de plus permis de lever certaines ambiguïtés concernant un espace recouvert de nuages sur l'image satellite.

C: Validation

Nous avons eu à notre disposition une couverture aérienne photographique noir et blanc de Novembre 1987 à l'échelle approximative 1/40 000. Cette opportunité a été mise à profit pour estimer la fiabilité de la limite déterminée sur les documents satellitaires. La même démarche que précédemment, respectant les mêmes règles de décision a donc été appliquée, et une "limite urbaine" a été dessinée sur ces photographies par deux interpréteurs. **La comparaison de la limite issue de l'image SPOT avec celle issue de la couverture aérienne met en évidence la fiabilité du travail effectué sur les images satellite.** La seule véritable erreur d'interprétation est imputable aux carrières qui présentent sur le panchromatique et dans une moindre mesure sur la composition colorée une apparence d'espace bâti.

4.3.2 : Stratification de la base de sondage sur l'image : les critères retenus et la réalisation pratique, mesures de surface de chaque strate

A : Mise en oeuvre de la classification des pixels de l'image SPOT selon la densité du bâti

Le zonage à l'intérieur de l'espace urbain a été réalisé à partir d'une copie d'écran représentant une classification de la densité du bâti. Cette classification de type barycentrique en 6 postes a été supervisée en fonction des statistiques descriptives calculées sur l'échantillon de 164 îlots précisément décrites et mesurées lors de l'enquête Morphologie. Les canaux (XS1) et neo-canaux (Ecart-type local sur panchromatique, Indice de végétation, et Indice de brillance) utilisés pour opérer cette classification ont été choisis en fonction de leur aptitude à rendre compte de la densité du bâti. Les délais d'élaboration de cette classification ont été très courts, l'enquête devant impérativement commencer avant la fin de l'année 1987 : on s'est donc contenté de reproduire le type de classification expérimenté sur le site de Marseille.

B: Evaluation de la classification

En raison des délais très courts, la classification a été initialisée et évaluée sur un même échantillon. Cela est contraire à la rigueur scientifique, mais l'absence d'étude précise sur la détermination des canaux et neo-canaux les plus aptes à discriminer la densité de l'espace bâti nous a légitimement conduit à initialiser la classification sur le plus grand nombre possible d'individus afin de réduire autant que possible les classements de pixels ambigus. Ce choix s'est donc, dès le départ effectué au détriment de la validation de la classification. Cependant, il est nécessaire d'évaluer cette classification, en particulier pour connaître les confusions entre les classes les plus

importantes. La matrice de confusion qui suit a été établie (en effectif), en croisant la modalité de référence avec la classe d'affectation.

Matrice de confusion						
Confusion	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6
Référence 1	2	2	2	0	0	0
Référence 2	8	29	16	5	2	0
Référence 3	1	3	10	20	15	0
Référence 4	1	1	2	5	14	0
Référence 5	1	1	2	3	17	2
Référence 6	0	1	0	0	1	3

La fiabilité de cette classification est :

39% en 6 classes

60% en 3 classes

On remarque qu'un décalage important subsiste entre classe et modalité de référence. La classification a tendance à surestimer la densité. Etant donné l'objectif de la classification, stratifier l'espace quiténien selon la densité du bâti, ce décalage n'est pas gênant : cette classification respecte bien une hiérarchie de densité, même si celle-ci est décalé par rapport à une référence.

C: Elaboration de la stratification

L'interprétation de cette classification, de pixels, a nécessité un important travail de zonage, non seulement en fonction de la valeur des pixels, mais surtout en fonction des **associations de valeurs** (de la texture) de l'image classée. En effet, il est aberrant de concevoir un indicateur de densité par pixel, et l'on doit tenir compte du voisinage à fortiori s'il s'agit d'effectuer un zonage dans l'optique d'une stratification de l'espace. Les plages homogènes de petite surface n'ont donc pas été détournées. Ce zonage, pour lequel nous avons conservé les 6 postes initiaux a été effectué indépendamment par F. DUREAU et A. MICHEL, puis reporté, après confrontation, sur le tirage panchromatique au 1/30.000 préalablement délimité.

D: Mesures de surface de chaque strate

Le zonage établi sur un tirage du canal panchromatique SPOT au 1/30.000 a été reporté sur un agrandissement du même canal au 1/15.000. C'est à cette échelle qu'ont été mesurées chacune des 232 zones composant la base de sondage à l'aide d'un planimètre électronique.

Le tableau ci dessous présente les surfaces occupées par chacune des 5 strates construites, ainsi que la superficie occupée par les zones non construites incluses à l'intérieur du périmètre urbain.

Strates	non bâti	1 densité très faible	2 densité faible	3 densité moyenne	4 densité forte	5 densité très forte	Total de l'aire métropol- -itaine
Surfaces S_k (ha)	2065,8	3485,4	3258,2	3287,2	3207,1	1036,9	16340,6

4.3.3 : Tirage de l'échantillon de pâtés de maisons

Une fois définie et stratifiée la base de sondage le tirage de l'échantillon de pâtés de maison s'effectue indépendamment dans chaque strate, comme toujours dès lors qu'il s'agit d'un sondage stratifié. Pour respecter la règle d'allocation de l'échantillon aux strates, on doit construire une grille de sélection pour chaque strate, dont la taille de la maille sera déterminée pour respecter le taux prévu.

A. Détermination du nombre total de pâtés de maisons de l'échantillon et de leur répartition dans les strates

Le budget alloué à l'enquête et la période de temps dont nous disposons pour sa réalisation permettent d'enquêter entre 3000 et 3500 ménages. Combien de pâtés de maisons doit on sélectionner pour obtenir cette taille d'échantillon? La réponse à la question est immédiate si le nombre de ménages enquêtés dans chaque pâté de maisons est constant. Dans notre cas, tel qu'a été défini le plan de sondage au second degré, le nombre total de ménages enquêtés dépend du nombre de ménages que nous devons enquêter dans les pâtés de maisons de moins de dix ménages inclus dans l'échantillon. Examinons plus précisément ce point.

Appelons m le nombre de ménages total à enquêter, n le nombre de pâtés de maison à sélectionner dans l'échantillon, p_1 la proportion dans l'échantillon des pâtés de maisons contenant moins de dix ménages et m_1 le nombre moyen de ménages que contiennent ces pâtés de maison. Puisqu'on enquête exhaustivement les pâtés de maisons de moins de dix ménage et que l'on choisit dix ménages dans les autres, on a :

$$m = m_1 n p_1 + 10n(1-p_1), \text{ donc}$$

$$n = \frac{m}{p_1 \cdot m_1 + (1-p_1) \cdot 10} \quad (4.3.3.1)$$

Pour calculer n il faut donc prévoir les valeurs qu'auront p_1 et m_1 dans l'échantillon. Nous baserons cette prévision sur les données recueillies lors de l'enquête préliminaire "bâti population", mais il faut faire une remarque à ce sujet.

On se souvient que l'échantillon correspondant à cette enquête (53 pâtés de maisons) a été choisi de manière raisonnée au sein des 199 pâtés de maisons de l'enquête préliminaire "morphologie", eux même sélectionnés à l'aide d'une grille de points plaquée sur les plans réalisés par l'INEC pour le recensement de 1982. Ce mode de sélection a deux conséquences :

1. Les probabilités d'appartenance des unités à l'échantillon ne sont pas contrôlées et en particulier, rien ne permet d'affirmer qu'elles aient été égales entre elles ou proportionnelles à la surface des pâtés de maisons. On considèrera la proportion des pâtés de maisons de moins de dix ménages dans cet échantillon et la "taille" moyenne de ces pâtés de maisons comme des estimateurs de p_1 et de m_1 mais il est clair qu'ils peuvent être biaisés et avoir des variances élevées; donc les prévisions seront très approximatives.

2. La base de sondage correspondant à cet échantillon, définie par la limite urbaine adoptée par l'INEC en 1982, est beaucoup plus restreinte que celle que nous retenons pour cette enquête. La plupart des zones supplémentaires que nous considérons comme appartenant maintenant à la ville sont des zones périphériques relativement peu denses. Il est certain que la proportion d'îlots de moins de dix ménages observée sur l'échantillon des 53 îlots de l'enquête préliminaire est nettement inférieure à ce qu'elle est dans la base de sondage définie. La formule 4.3.3.1, lorsqu'on remplace p_1 et m_1 par les valeurs observées sur l'échantillon de l'enquête préliminaire fournira donc un nombre de pâtés de maisons trop faible. C'est la raison pour laquelle pour calculer n , nous avons fixé à 4000 le nombre total de ménages m .

Dans l'échantillon des 53 pâtés de maisons, on constate la répartition suivante :

- Pâtés de maisons contenant jusqu'à dix ménages : 22,7 %, en moyenne 6,9 ménages par pâté de maisons
- Pâtés de maisons de plus de dix ménages : 77,3 % .

$$\text{On a donc } n = \frac{4000}{0,227 \times 6,9 + 0,773 \times 10} \approx 430 .$$

Reste à répartir ces 430 pâtés de maisons dans les cinq strates. Le principe est simple; disposant pour chaque strate, de la surface totale mesurée sur l'image satellite (S_k) et de la surface moyenne des pâtés de maisons de la strate estimée d'après l'échantillon de 199 pâtés de maisons de l'enquête préliminaire "morphologie" (\bar{s}_k), on calcule tout d'abord une approximation du nombre de pâtés de maisons que contient chaque strate (N_k).

Le quotient $\frac{430}{\sum_{k=1}^5 N_k}$ fournit ensuite le taux global sur l'ensemble de la base qui, appliqué

à chacune des strates donne la répartition dans le cas de l'allocation proportionnelle. Enfin, par application des coefficients qui définissent l'allocation aux strates (C_k : voir tableau du paragraphe 4.2.2), on obtient le nombre de pâtés de maisons à sélectionner dans chaque strate (n_k). Cependant, la répartition entre les strates de l'échantillon n'est pas exactement la même selon que l'on considère les 53 pâtés de maisons de l'enquête "bâti/population" ou les 199 pâtés de maison de l'enquête "morphologie". Comme l'allocation du paragraphe 4.2.2 a été défini à partir de l'échantillon de 53 îlots (de plus cette allocation est le résultat d'une moyenne d'allocations optimales), si l'on applique tels quels les coefficients définis, l'échantillon obtenu compte trop peu de

pâtés de maisons. Il est donc nécessaire d' "ajuster" les coefficients à la répartition entre strates observée sur l'échantillon des 199 pâtés de maisons qui sert de base au calcul. Le tableau 4.4 ci dessous résume les calculs effectués.

Tableau 4.4 : résumé des calcul effectués pour la détermination du nombre de pâtés de maisons à enquêter dans chaque strate.

Strates	1 densité très faible	2 densité faible	3 densité moyenne	4 densité forte	5 densité très forte	Total de la base
S_k (m ²)	34.853.537	32.582.087	32.872.461	32.070.924	10.369.293	142.748.300
\bar{s}_k (m ²)	31.473	25.875	20.374	12.443	12.106	19.256
N_k	1.107	1.259	1.613	2.577	857	7.413
n_k (Alloc. prop.)	64	73	94	149	50	430
C_k (Originaux)	0,4	0,3	1	1,5	0,9	0,95
C_k (Ajustés)	0,44	0,38	1,15	1,54	0,76	1
n_k (Alloc.opt.)	28	28	108	230	38	432

E. Définition des grilles

Pour obtenir, pour chaque strate, la surface de la maille de la grille de tirage, il suffit de diviser la surface totale de la strate mesurée sur l'image par le nombre d'ilots à sélectionner. On calcule ensuite la longueur du côté de la maille. Le tableaux 4.5 ci-dessous fournit les dimensions des grilles obtenues pour chaque strate.

Tableau 4.5 : calcul des mailles de grilles dans chaque strate.

Strates	1 densité très faible	2 densité faible	3 densité moyenne	4 densité forte	5 densité très forte
S_k (en mm^2 sur l'image)	161.963	151.408	152.757	149.033	48.186
N_k	28	28	108	230	38
Surface de la maille (en mm^2 sur l'image)	5784,4	5407,4	1414,4	648,0	1268,1
Côté de la maille(en mm sur l'image)	76,1	73,5	37,6	25,5	35,6

C. Tirages

Une fois construites les cinq grilles sur papier calque millimétré, on a procédé au tirage en cinq étapes en commençant par la strate la plus dense (strate 5) pour finir par la strate la moins dense (strate 1). Une précaution importante doit être prise lors des placements initiaux des grilles sur l'image : il faut éviter d'orienter les axes principaux de la grille (horizontal, vertical ou diagonal) parallèlement à un axe principal du réseau de voirie. En effet, une telle orientation pourrait aboutir à une sur-représentation dans l'échantillon des îlots situés en bordure (ou à distance fixe) d'un grand axe, lesquels ont des caractéristiques morphologiques et démographiques spécifiques. On risquerait alors par exemple de surestimer les activités de commerce et service ou même de sous-estimer le total de population résidente. Remarquons que ce type d'erreur d'échantillonnage est en fait le pendant, dans le cas du tirage systématique aréolaire, des erreurs qui surviennent dans un sondage systématique sur liste lorsque la variable à estimer présente des variations périodiques dans la liste et que le pas de tirage est approximativement un multiple de la période¹.

Les pas des grilles sont calculés, comme on l'a vu plus haut, pour fournir dans chaque strate le nombre n_k d'îlots souhaité. Il se peut cependant que le placement initial de la grille sur l'image conduise en définitive à un nombre d'îlots légèrement différent. Cette situation ne s'est pas produite lors du processus de sélection et, par chance, le premier placement de la grille a toujours donné le nombre exact de pâtés de maisons. Dans un cas moins favorable il aurait fallu répéter l'opération en défessant aléatoirement une nouvelle position de la grille jusqu'à obtenir le bon nombre de points inclus dans la strate.

¹ : Voir sur ce point par exemple J. DESABIE, "Théorie et pratique des sondages" P. 99 ou J.M. GROBRAS, "Méthodes statistiques des sondages" P.29.

Une fois que l'on s'est assuré d'avoir dans la strate le nombre de points voulu, l'étape suivante consiste à définir sur l'image, partout où c'est possible (c'est à dire partout où le réseau de voirie apparait sans ambiguïtés), l'ilot à enquêter correspondant à chaque point. Lorsque le point est situé clairement à l'intérieur de l'ilot c'est évidemment l'ilot en question qui est sélectionné. Si le point tombe dans la voirie séparant deux (ou plusieurs) pâtés de maisons, on applique une règle systématique : nous avons pour notre part retenu les ilots situés immédiatement au nord-ouest des points. Pour faciliter l'identification des pâtés de maisons par les enquêteurs et leur repérage sur le terrain, l'ensemble des ilots ainsi repérés est reporté sur les plans utilisés par l'INEC pour le recensement (voir figure 4.6). Enfin, lorsque l'identification précise de l'ilot est impossible sur l'image ou que l'ilot est situé en dehors de la zone cartographiée par l'INEC, on effectue alors le repérage et/ou l'identification des limites directement sur le terrain avec le superviseur chargé de la zone (parfois même avec l'enquêteur en charge de l'ilot). La localisation du point sur le terrain au vu de sa situation sur l'image n'a jamais posé de problème ; en revanche l'accès au point en question a parfois été difficile.

Signalons pour finir que le respect du tirage sans remise a été total et qu'il n'a posé aucun problème de remplacement ou de répétition de la sélection puisque, lors du tirage, aucun pâté de maison n'a été sélectionné deux fois. Il n'y a d'ailleurs rien d'étonnant à cela, en effet, le tirage systématique à l'aide de la grille de points limite le risque de sélections répétées d'une même unité au cas où la taille de l'unité est suffisante pour contenir deux points de la grille. Avec le taux de sondage pratiqué, la maille de la grille dans chaque strate est suffisamment grande pour annuler ce risque.

Figure 4.6 : illustration du processus de sélection dans les différentes strates sur une partie de la base de sondage (I)

Première étape : La limite urbaine, telle qu'elle a été desssinée sur l'image support de la sélection. Ici, une partie du nord moderne de la ville, comprise entre le parc "la carolina" et l'aéroport.

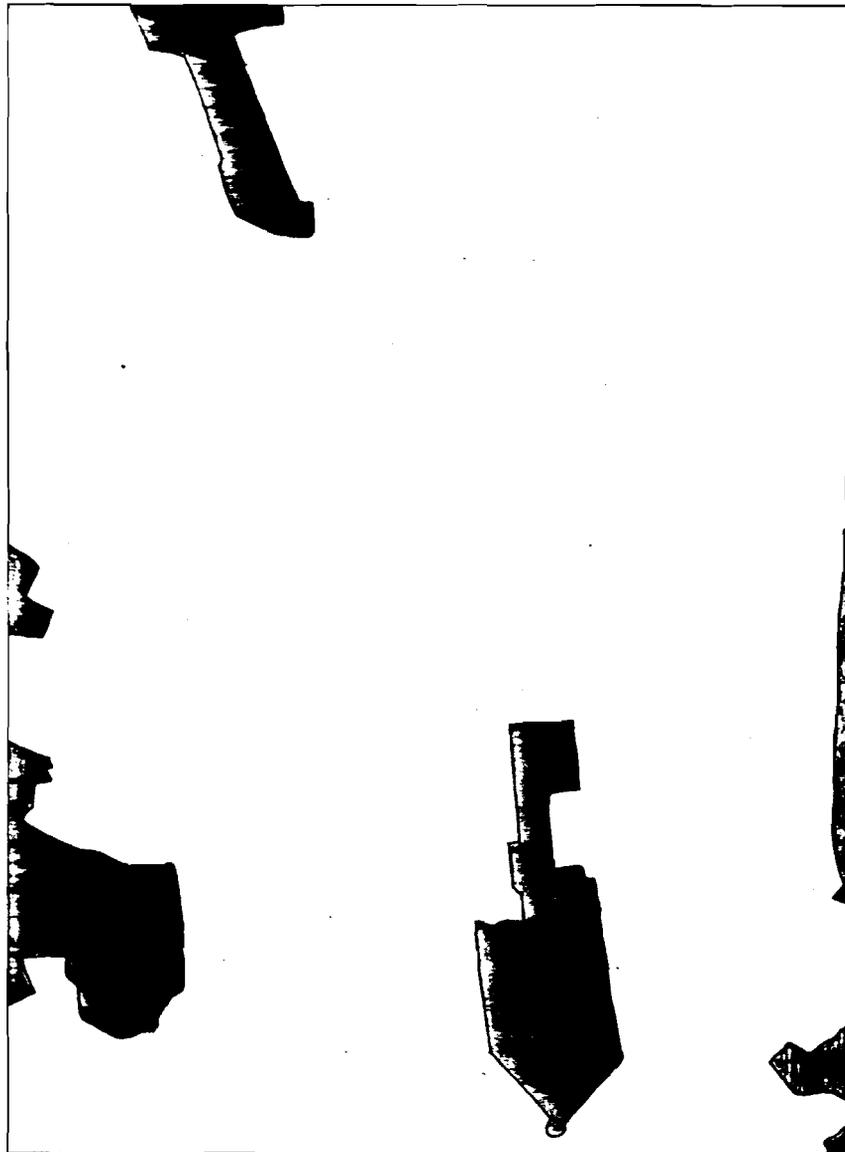


Figure 4.6 : illustration du processus de sélection dans les différentes strates sur une partie de la base de sondage (II)

Deuxième étape : Stratification de la base de sondage sur le critère de la densité du bâti :

- 1. Très peu dense*
- 2. Peu dense*
- 3. Densité moyenne*
- 4. Densité forte*
- 5. Densité très forte.*

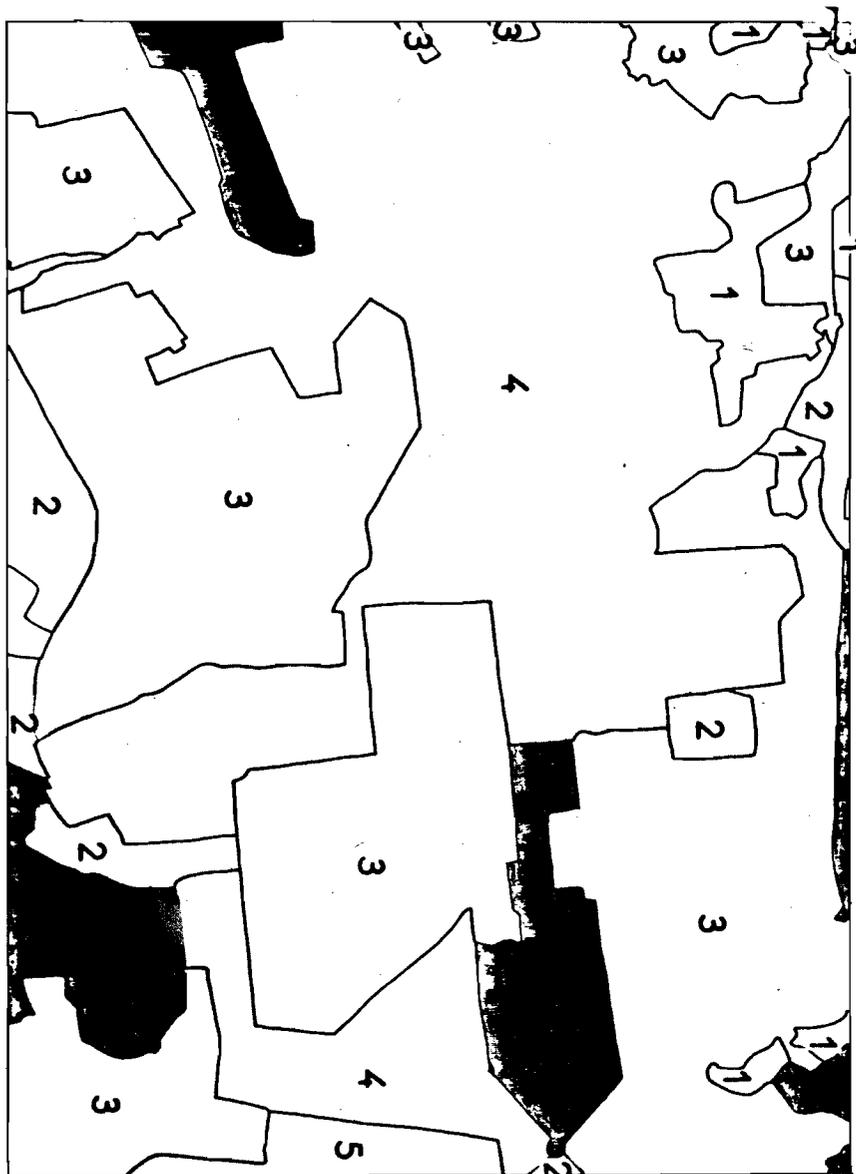


Figure 4.6 : illustration du processus de sélection dans les différentes strates sur une partie de la base de sondage (III)

Troisième étape : Dans chaque strate, la grille superposée à l'image permet la sélection des unités primaires de l'échantillon. L'allocation aux différentes strates est gérée en faisant varier la taille de la maille de la grille pour sélectionner, dans chaque strate, le nombre d'îlots voulu.

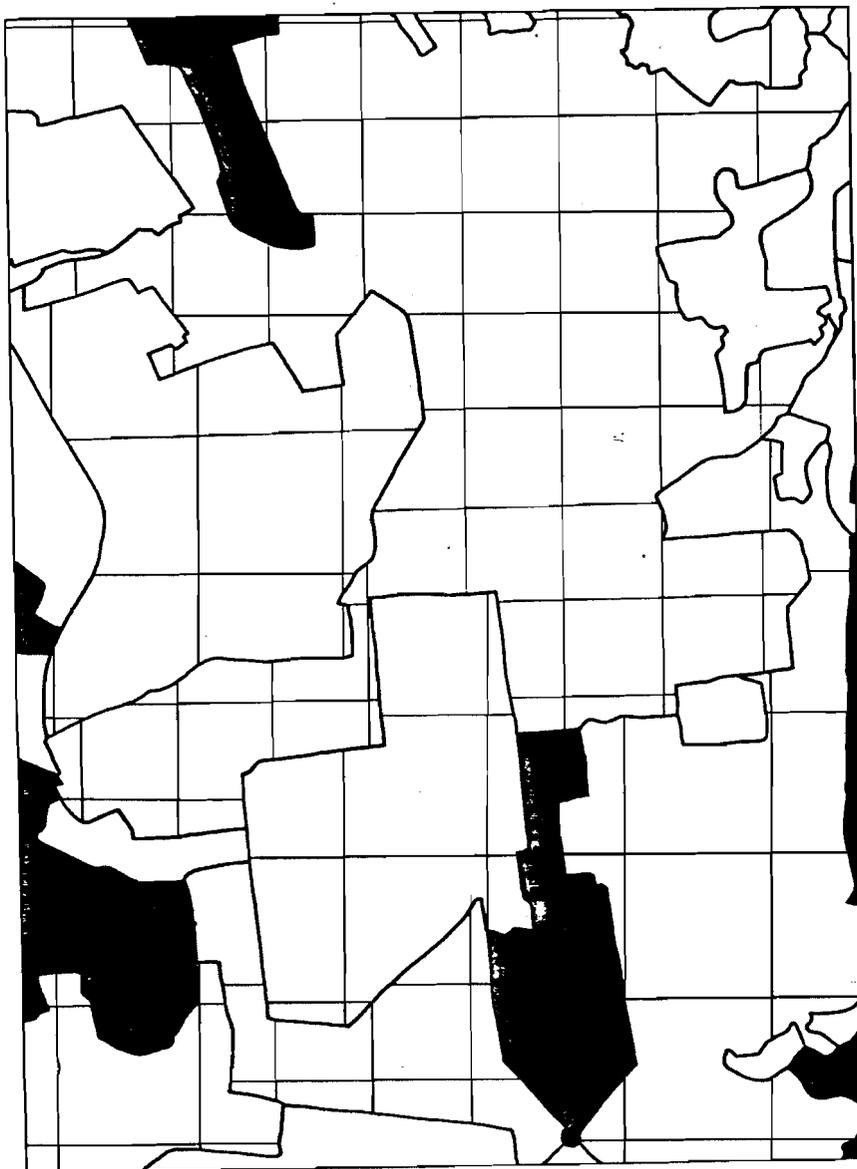


Figure 4.6 : illustration du processus de sélection dans les différentes strates sur une partie de la base de sondage (IV)

Quatrième étape : Les ilots à enquêter sont identifiés sous chaque points de la grille.



4.3.4 : Recensement des ménages dans les pâtés de maisons et sélection des ménages à enquêter

Une fois réalisée l'opération de sélection de l'échantillon d'unités primaires, le travail de collecte de l'information relative aux pâtés de maisons peut commencer ; il a un double objectif :

1. effectuer le recensement des ménages qui occupent le pâté de maisons pour pouvoir sélectionner l'échantillon de ménages à enquêter et, par la suite, pratiquer l'inférence au second degré (estimation des totaux, moyennes et quotients au sein des unités primaires)
2. mesurer la surface de chaque unité primaire pour pouvoir pratiquer l'inférence au premier degré (estimation des totaux, moyennes et quotient de la base entière, des strates ou des éventuels domaines d'étude).

Nous indiquons ici rapidement comment ont été atteints ces objectifs.

A. Recensement des ménages

Le travail de recensement des ménages dans chaque unité primaire est confié aux superviseurs qui réalisent sur le terrain un schéma du pâté de maisons permettant son identification sans ambiguïté par l'enquêteur et sur lequel figure, outre les indications que le superviseur juge utiles pour le travail de l'enquêteur, l'ensemble des bâtiments constituant l'îlot, ainsi que le nombre de ménages qui occupent chacun des bâtiments (figure 4.7). C'est ce document qui sert de base au comptage des ménages, puis à la numérotation des différents bâtiments occupés. Il est également utilisé par l'enquêteur chargé des mesures de surfaces lorsque ces dernières doivent être effectuées sur le terrain.

Signalons enfin que dès cette première étape du travail de collecte et durant toute la période d'enquête, un quartier populaire situé à l'extrême nord-ouest de la ville a été le théâtre d'affrontements armés qui ont fait plusieurs morts parmi la population. Il était donc exclu d'y mener l'enquête et l'échantillon réellement enquêté ne comprend de ce fait que 426 îlots au lieu des 432 initialement sélectionnés.

B. Sélection des ménages à enquêter

Sur le schéma ainsi réalisé, le superviseur indique à l'enquêteur la liste des ménages constituant l'échantillon. Hormis pour les pâtés de maisons comptant jusqu'à dix ménages, pour lesquels l'enquête est exhaustive, la règle de sélection systématique des ménages est la suivante : en divisant le nombre total de bâtiments par dix, le superviseur obtient le pas k du tirage ; il sélectionne ensuite "au hasard" un bâtiment dont le numéro est compris entre 1 et k puis, en partant de ce numéro il retient un bâtiment tous les k . L'enquêteur devra choisir un ménage par bâtiment figurant dans la liste. En cas de refus de réponse ou d'absence prolongée pour un ménage donné, l'enquêteur le remplace par un ménage du même bâtiment ou si c'est impossible par un ménage choisi dans un bâtiment voisin. Ces règles s'appliquent sans problèmes à l'immense majorité des cas à Quito, ville qui compte très peu de grands ensembles. Néanmoins, pour assurer une représentation équitable, au sein de l'échantillon, des ménages vivant dans de l'habitat collectif, les bâtiments contenant un grand nombre de ménages sont traités à part lors du processus de sélection des ménages. La règle est alors de sélectionner d'office un ménage dans chaque bâtiment contenant plus d'un dixième des ménages du pâté de maisons, deux ménages dans les bâtiments en contenant plus de deux dixième, etc...

En toute rigueur, la stratégie de sélection des ménages au sein des unités primaires correspond donc à un sondage à deux degrés dont le premier degré consiste à choisir un échantillon de bâtiments avec probabilité d'inclusion approximativement proportionnelle au nombre de ménages occupant le bâtiment. Mais, comme cette

probabilité n'est que très grossièrement respectée, il nous semble illusoire voir dangereux, d'intégrer ce degré intermédiaire de sondage dans le formulaire d'inférence.

A l'issu du processus de sélection et de l'opération de collecte, l'échantillon des ménages comprend 3161 ménages répartis dans les 426 flots enquêtés, ce qui correspond à un taux final d'environ 1,3 % des ménages.

4.3.5 : Mesure de surface des unités primaires

A: Les méthodes de mesure employées

Pour la mesure des surfaces des unités primaires, deux techniques ont été employées et validées :

- mesure sur le tirage photographique du canal panchromatique SPOT au 1/15000 employé pour le tirage de l'échantillon d'îlots,
- mesure sur le terrain.

256 des 426 îlots enquêtés ont pu être repérés sans ambiguïté sur le tirage photographique au 1/15 000 ; nous avons alors mesuré la surface par superposition d'un plastique transparent millimétré.

Pour les 170 îlots non repérables précisément sur image satellite, un enquêteur a été envoyé sur le terrain. Il a établi avec le maximum de précision possible un croquis mentionnant les angles et les longueurs de chaque segment droit délimitant l'îlot ainsi que leur pente (**Figure 4.8**).

Les longueurs sont mesurées au décimètre sur les terrains en pente et au pas sur les terrains plats, la longueur du pas ayant été préalablement étalonnée par comparaison avec les mesures effectuées au décimètre ; les angles sont mesurés à la boussole à chaque extrémité d'un segment, ce qui autorise une vérification. La pente de chaque segment est mesurée au clinomètre.

La surface est ensuite calculée par programme informatique à l'aide du logiciel MacDraft™ sur Apple Macintosh. Il suffit de dessiner à l'échelle choisie les contours de l'îlot en respectant les longueurs préalablement corrigées en fonction de la pente, et les angles prescrits (**Figure 4.9**).

Figure 4.8 : Croquis d'ilot pour les mesures de surfaces sur le terrain

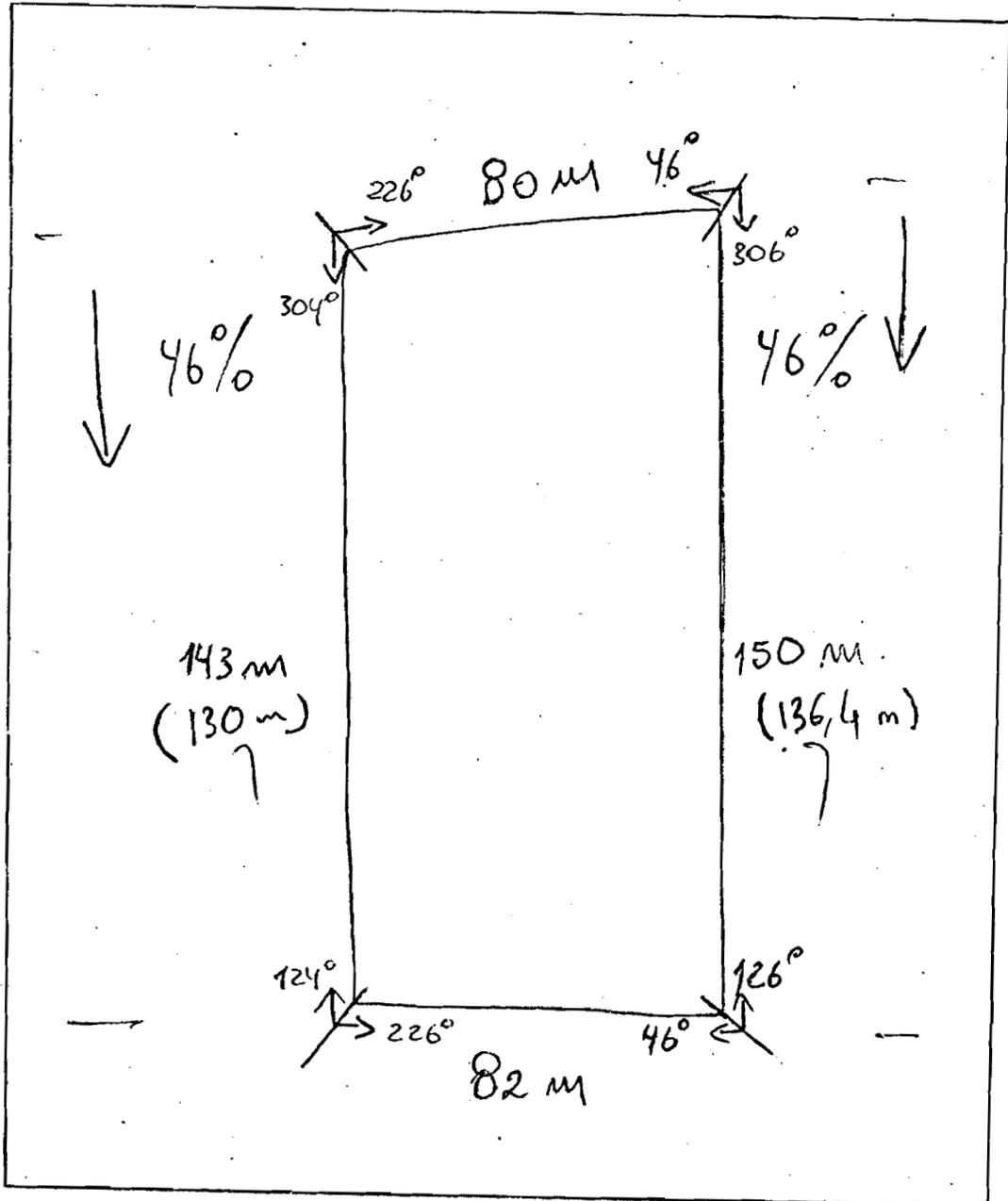
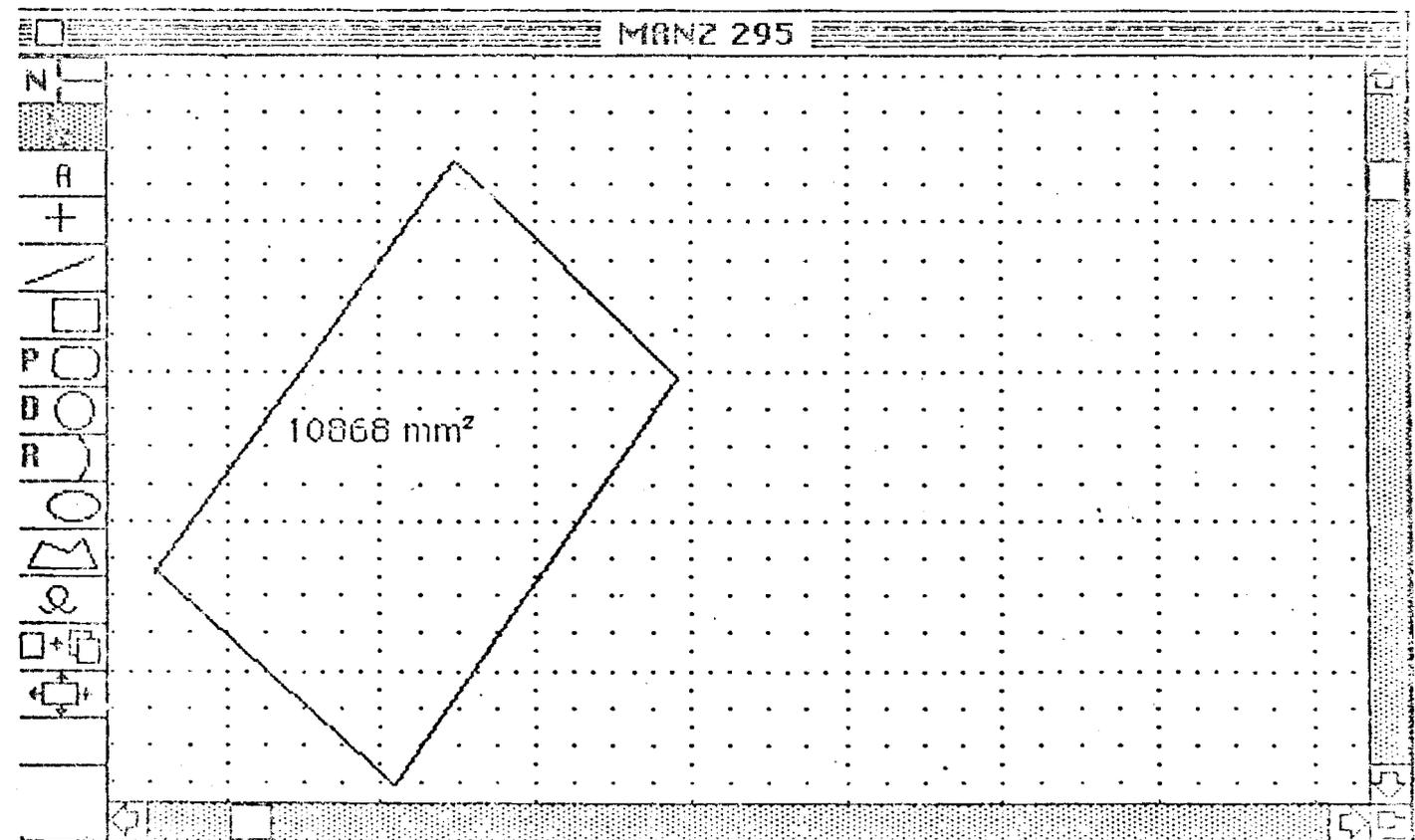
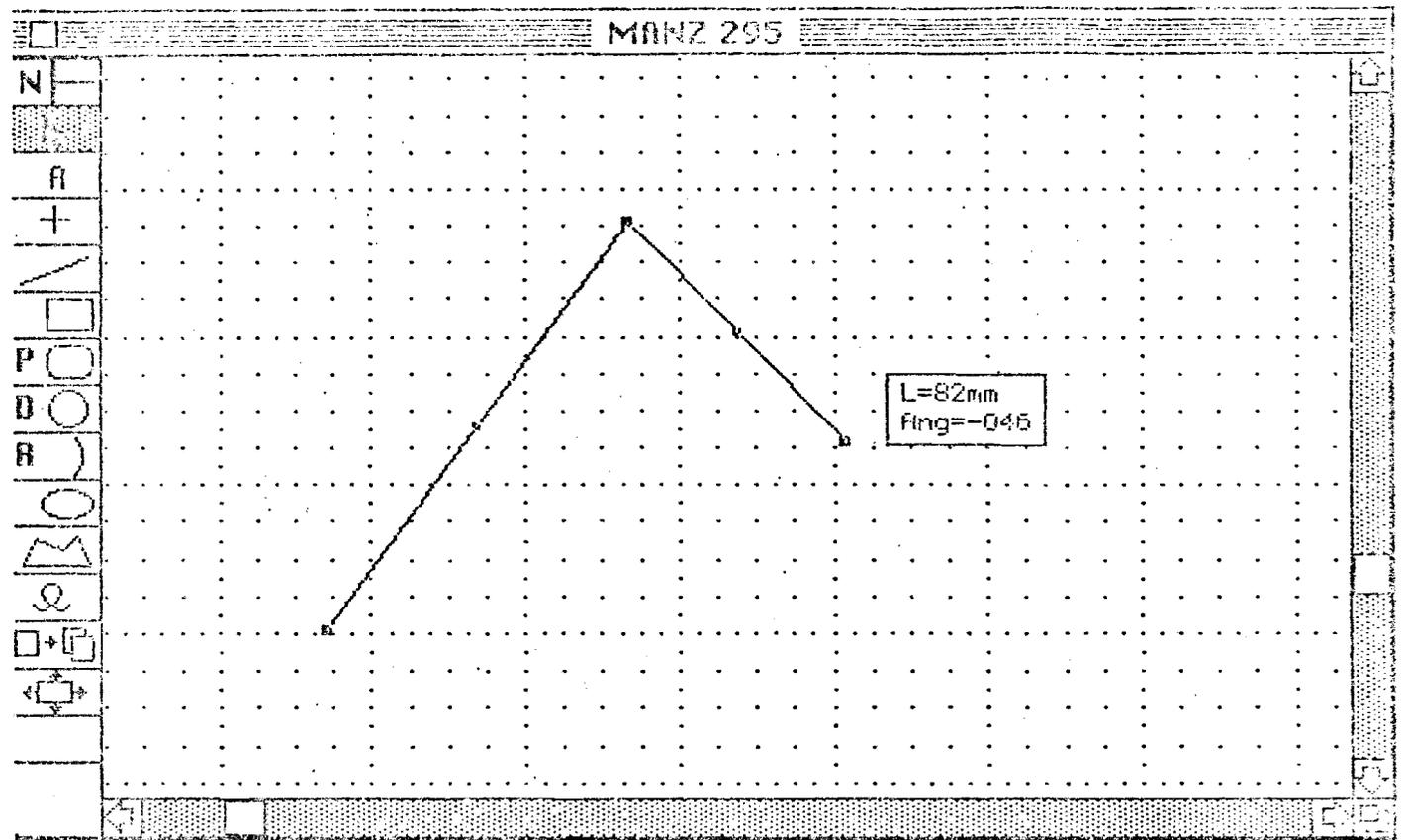


Figure 4.9 : Exemple de calcul de surface effectué à l'aide du logiciel MacDraft™ sur Macintosh



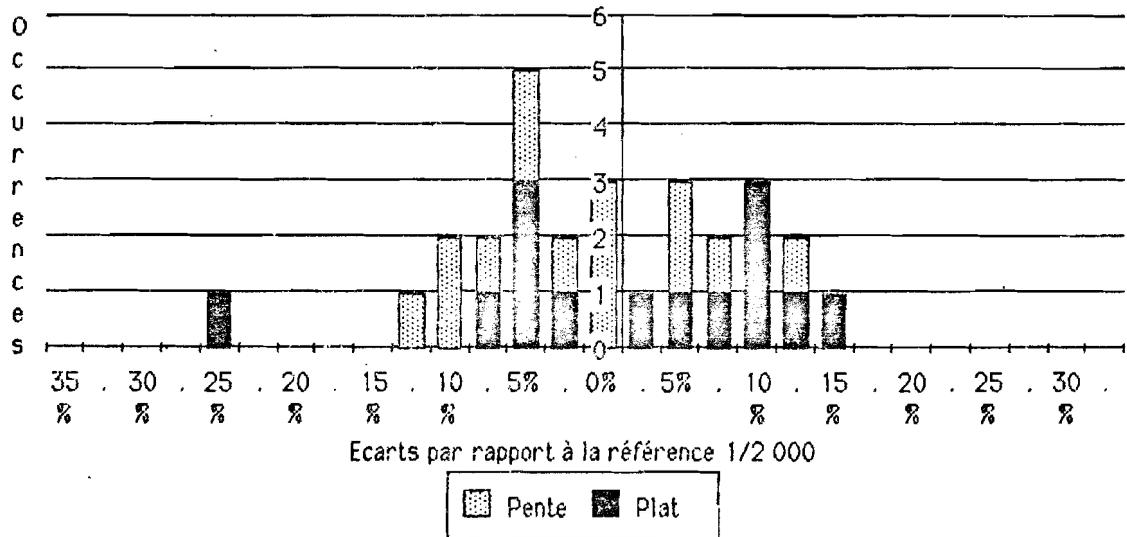
B: Evaluation de la précision des techniques de mesures

Pour deux échantillons de 30 ilots correspondant chacun à une technique de mesure, nous avons comparé les mesures ainsi obtenues avec celles, considérées comme référence, issues de la cartographie au 1/2 000 de l'IGM.

Les graphiques qui suivent correspondent aux fréquences des différences relatives entre la mesure sur image ou sur le terrain et la surface de référence. Les fréquences correspondant aux différentes modalités de pente sont superposées les unes sur les autres. Ces écarts (en pourcentage) sont classiquement négatifs à gauche de l'axe des ordonnées (**effectif**) et positifs à droite. L'axe des abscisses (**différences relatives**) est gradué de 2,5 en 2,5%.

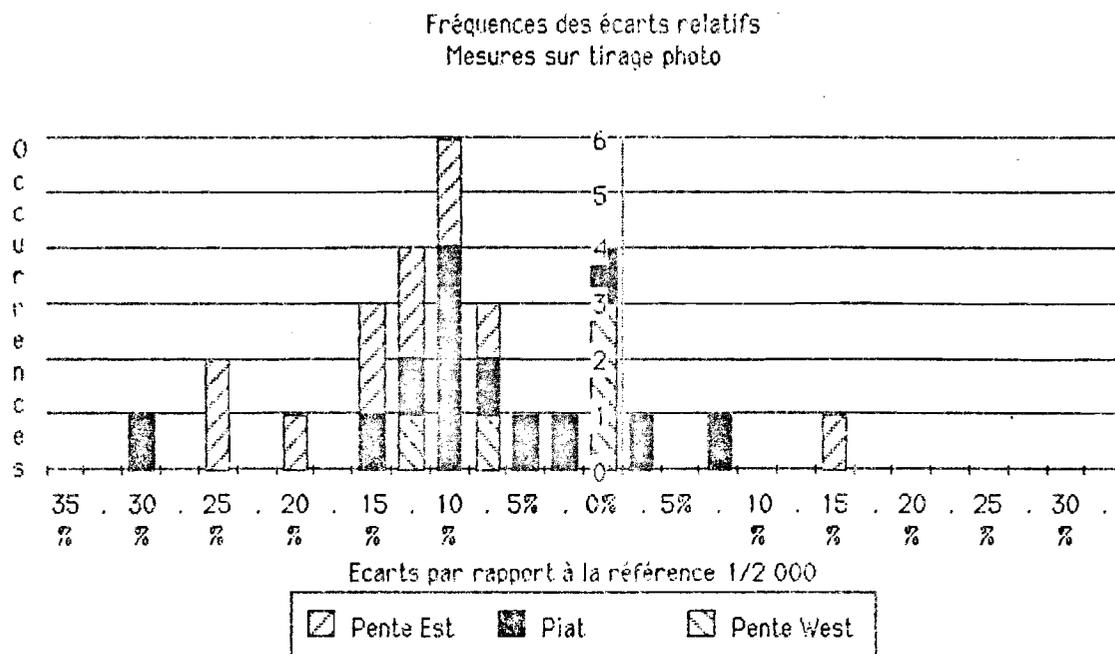
Graphique 4.10- Mesures terrains -

Fréquences des écarts relatifs
Mesures sur le terrain



Les surfaces mesurées sur le terrain apparaissent comme remarquablement centrées autour de 0%, variant de -15% à 15%. Les mesures sur pente sont légèrement plus nombreuses à avoir été sous-estimées.

Graphique 4.10 (suite) - Mesures terrains -



Le centrage des données autour de -10% montre très nettement que les surfaces, toutes modalités de pente confondues ont été sous-estimées par cette méthode de calcul. Il s'agit ici d'un biais **systematique** ; il est alors légitime de multiplier la surface ainsi obtenue par un coefficient de valeur 1,09 (1/0,92) pour tenir compte de cette sous-estimation.

La séparation entre îlots exposés Est (en majorité à gauche de -10%) et îlots exposés Ouest (en majorité à droite de -10%) apparaît clairement sur ce graphique.

L'analyse des écarts relatifs entre les surfaces mesurées sur image ou sur le terrain et les surfaces de référence conduit aux conclusions suivantes :

-On notera avec intérêt, que **c'est la mesure la plus facilement praticable dans les pays en développement (mesure sur le terrain) qui apparaît comme la plus fiable.**

- Influence de la taille de l'îlot:

Les surfaces sont globalement sous-estimées pour les îlots de forte taille quelque soit la technique de mesure employée. Sur image satellite, les mesures de surface sont plus imprécises sur des îlots de faible taille ; pour les mesures terrain, ce sont pour les îlots de taille moyenne qu'elles sont les moins précises.

- Influence de la pente:

La pente affecte peu les mesures de surface effectuées sur le terrain ; par contre sur le document satellite, très logiquement, les surfaces situées sur un flanc Ouest sont **sous-estimées** (conjugaison de la pente et de l'angle de visée du satellite SPOT), les surfaces situées sur un flanc Est "**redressées**" (influence opposée de la pente et de l'angle de visée) et les surfaces situées sur un espace plan **sur-estimées**.

4.4 : Conclusion

En conclusion, rappelons rapidement les principales caractéristiques du plan de sondage appliqué à l'enquête finale "démographie/migrations".

1. La base de sondage est l'image satellite SPOT stratifiée en cinq niveaux de densité du bâti, les zones non bâties sont exclues du tirage.

2. Tirage au premier degré : les unités primaires sont les pâtés de maisons. la sélection est aréolaire, systématique à l'aide d'une grille de points à maille variable respectant la règle d'allocation fixée. Le placement initial de la grille sur l'image est aléatoire, sous la contrainte de respect des taux. Le tirage fournit un échantillon de 432 ilots dont 426 pourront être réellement enquêtés

3. Tirage au deuxième degré : les unités secondaires sont les ménages, la sélection est là encore systématique; elle s'effectue sur liste après recensement de tous les ménages vivants dans l'ilot. Le taux au second degré est variable suivant le nombre de ménages vivant dans l'ilot :

- moins de dix ménages dans l'ilot : enquête exhaustive
- plus de dix ménages dans l'ilot : dix ménages sont enquêtés

L'échantillon de ménages enquêté comprend finalement 3161 ménages, ce qui correspond à un taux d'environ 1,3 % des ménages.

4. Les estimateurs employés :

- estimateur sans biais sous plan à probabilités proportionnelles aux surfaces des unités primaires pour l'estimation des totaux.
- estimateur du ratio sous plan à probabilités proportionnelles aux surfaces des unités primaires pour l'estimation des quotients.

Les opérations de collecte, commencées le lundi 23 Novembre 87 se sont poursuivies jusqu'à la fin Décembre 87. La phase de codification et de saisie des résultats s'est achevée à la fin du mois d'Avril 88. L'exploitation des données recueillies peut donc maintenant commencer et les résultats nécessaires pour évaluer définitivement la précision de la méthode devront être connus courant Juillet.

BIBLIOGRAPHIE SELECTIVE

A) Bibliographie générale

- Utilisation des données de télédétection aérienne et spatiale pour des estimations démographiques

- ADENIYI P. O., 1983. An aerial photographic method for estimating urban population. *Photogrammetric engineering and remote sensing*, vol. 49, n° 4, pp 545-560.
- BALLUT A., GAUTHIER M., 1983. Une méthode pour la production de données socio-économiques spatialisées dans le cadre d'un schéma directeur d'aménagement et d'urbanisme. IAURIF. Paris.
- GREEN N.E., 1956. Aerial photographic analysis of residential neighbourhoods : an evaluation of data accuracy. *Social forces*, n° 35, pp. 142-247.
- HOLZ R.K., HUFF D.L., MAYFIELD R.C., 1969. Urban spatial structure based on remote sensing imagery. *Sixth International Symposium Remote Sensing Environnement*, vol. 2, pp. 819 - 830.
- HSU S.Y., 1971. Population estimation. *Photogrammetric engineering and remote sensing*, vol. 37, n° 5, pp. 819 - 830.
- IAURIF, ROC, SIDAU, SEE, 1983. Les données urbaines. Système permanent de production rapide et économique. Recherche et synthèse. Paris.
- KRAUS S.P., SENGER L.W. RYERSON J.M. 1974. Estimating population from photographically determined residential land use types. *Remote sensing of environment*, n° 3, pp. 35 - 42.
- LO C.P., WELCH R., 1977. Chinese urban population estimates. *Annals of the Association of American Geographers*, vol. 67, n° 2, pp 246 - 250.
- PORTER P.W., 1956. Population distribution, and land use in Liberia. London School of Economics and Political Sciences. Thèse de Ph. D.
- VERNIERE M., 1978. Méthode d'analyse quantitative de la croissance urbaine dans l'espace et dans le temps. Exemple d'une banlieue de Dakar (Sénégal). *Photo interprétation*, n° 1, pp. 34 - 55.
- WATKINS J.F., 1985. Small area production estimates using aerial photography. *Photogrammetric engineering and remote sensing*, vol. 51, n° 2, pp. 1933 - 1935.

- Télédétection spatiale en milieu urbain

BALLUT A., 1980. L'étude de l'occupation du sol par satellite. Cahiers de l'IAURIF, vol. 61. Paris.

BOQUET E., 1980. Réactualisation du mode d'occupation du sol par l'utilisation des simulations SPOT (81 et 83). Mémoire de maîtrise de géographie. Université Paris VII. Paris.

CNRS, 1982. Débats des journées télédétection en milieu urbain (6 - 7 mai 1982). CERG, Paris.

IGN, 1984. Recherches méthodologiques sur l'utilisation de la télédétection spatiale en milieu urbain. IGN, Paris.

THIBAUT C. BIANCALE M. DELAVIGNE R., 1983. La télédétection par satellite appliquée au milieu urbain : test de validité de l'indice de végétation dans la proche couronne. IAURIF, Paris.

WILMET J., SOYER J., 1982. Lumbumbashi et le sud-est du haut Shaba : interprétation de données LANDSAT. Bulletin de la Société Belge d'Etude Géographiques, pp. 87 - 100.

- Analyse des données

BENZECRI J.P. et coll., 1973. L'analyse des données. Tome 1 : La taxinomie, Tome 2 : L'analyse des correspondances. DUNOD, Paris.

BENZECRI J.P. et coll., 1980. La pratique de l'analyse des données. vol. 2 : abrégé théorique, étude de cas modèles. DUNOD, Paris.

CHANDON J.L., PINSON S., 1981. Analyse typologique, théories et applications. MASSON, Paris.

DIDAY E. et coll., 1980 : Optimisation en classification automatique (2 vol.). INRIA, Paris.

JAMBU M., 1978. Classification automatique pour l'analyse des données, vol. 1 : Méthodes et algorithmes, vol. 2 : Logiciels. DUNOD, Paris.

LEBART L., MORINEAU A., FENELON J.P., 1982. Traitement des données statistiques, méthodes et programmes. DUNOD, Paris.

VOLLE M., 1981. L'analyse des données. ECONOMICA, Paris.

- Théorie des sondages et variables régionalisées

ASSELIN L.M., 1984. Techniques de sondage avec application à l'Afrique. GAETAN MORIN, Québec.

CÔCHRAN W.G., 1977. Sampling techniques (3^e édition). J. WILLEY, New York.

DELINCE J., 1983. Estimations robustes de densité dans le plan, étude comparée de techniques d'échantillonnage, application en forêt. Université Catholique de Louvain, Faculté des Sciences Agronomiques. Louvain.

- DEMING W.E., 1950. Some theory of sampling. J.WILLEY, New York.
- DEROO M., DUSSAIX A.M., 1980. Pratique et analyse des enquêtes par sondage. PUF, Paris.
- DESSABIE J., 1966. Théorie et pratique des sondages. DUNOD, Paris.
- DEVILLE J.C., ROY G., 1984. L'échantillon maître fait peau neuve. Courriers des statistiques n° 29, Paris
- GOURIEROUX Ch., 1981. Théorie des sondages. ECONOMICA, Paris.
- GROSBRAS J.M. 1987. Méthodes statistiques des sondages, ECONOMICA, Paris
- GUIDICELLI X., LANLY J.P., OUKAM J.B., PIETRI M., 1972(ISUP). Application de la théorie des processus aléatoires à l'estimation de la précision d'un inventaire forestier par échantillonnage systématique. Annales des Sciences Forestières, Institut National de la Recherche Agronomique. Paris.
- MATHERON G., 1965. Les variables régionalisées et leur estimation, une application de la théorie des fonctions aléatoires aux sciences de la nature. MASSON, Paris.
- RAJ D., 1968. Sampling Theory. MC GRAW HILL, New York.
- RAJ D., 1972. The design of sample surveys. MC GRAW HILL, New York.
- RAO J.N.K., HARTLEY H.O., COCHRAN W.G., 1962. On a simple procedure of unequal probability sampling without replacement. Journal of Royal Statistical Society B24, pp. 482 - 491.
- SUKHATME P.V., SUKHATME B.V., 1970. Sampling theory of surveys with applications. FAO, Rome.
- THIONET P., 1958. La théorie des sondages. INSEE, Paris.
- YATES F., 1981. Sampling methods for censuses and surveys (4^e édition). GRIFFIN, Londres.

- Informatique, programmation

- COHEN R., 1986. Programmer en dbase III et dbase III plus. éditions du PSI, Paris.

- Logiciels utilisés

DBASE III PLUS (ASTHON TATE) : Gestion de base de données.

SPSS PC + (SPSS inc.) : Traitements statistiques.

STAT ITCF (Institut Technique des Céréales et Fourrages) : Traitements statistiques.

ADAD MICRO (ADAD) : Analyse des données.

WORD (MICROSOFT) : Traitement de texte.

B) Publications et notes de travail de l'équipe ORSTOM
(peuvent être consultées au Département SDU de l'ORSTOM)

Articles

DUREAU F. et GUILLAUME A., 1984. La démographie depuis l'espace : un nouveau système d'observation ? Paris, STATECO n° 38, pp 5-46.

DUREAU F., 1986. A propos du traitement informatique de données localisées. Une expérience en cours : télédétection et observation des populations urbaines. Paris, ORSTOM, collection colloques et séminaires, numéro sur le traitement des données localisées, pp.263-286.

DUREAU F. et BARBARY O., 1987. Pour une méthode de collecte démographique en ville intégrant la télédétection spatiale. A paraître dans la revue informatique et sciences humaines, EHESS, 42 p.

MICHEL A., DUREAU F., LORTIC B., SOURIS M., 1987. Mise au point des méthodes d'analyse des images satellites à haute résolution et évaluation des informations fournies par ces images. 1 : Etude statistique du descripteur "densité du bâti". Présentation des classifications. 2 : Mise en évidence des réseaux routiers sur une image SPOT panchromatique. Etude de faisabilité sur Quito (Equateur). Paris, service technique de l'urbanisme, Bilan sur la télédétection urbaine en France. 10 p. et photos.

SOURIS M., 1986. Système d'information géographique et bases de données. Paris, ORSTOM, collection colloques et séminaires. Numéro sur le traitement des données localisées, pp.29-87.

Communications à des colloques

BARBARY O., 1988. Desarrollo de un método de muestreo demográfico en medio urbano que integre la información de los satélites. Presentación de los principales resultados obtenidos en Marsella y en Quito. Communication au colloque "muestreo de agrícola de áreas en los países andinos", Quito (Equateur), 2-7 Mai 1988, 8 p.

- DUREAU F. et GUILLAUME A., 1985. La population dans l'espace. Télédétection et observation démographique des villes des pays en développement. Communication au XX^e congrès général de l'Union Internationale pour l'Etude Scientifique de la Population, Florence (Italie), 5-12 Juin 1985.
- DUREAU F., LORTIC B., MICHEL A., SOURIS M., 1987. Télédétection et système d'information géographique. Communication au Forum International de l'Instrumentation et de l'Information Géographique, Lyon, (France), 10-13 Juin 1987, 15 p.
- DUREAU F., LORTIC B., MICHEL A., SOURIS M., 1987. Informatique, télédétection et observation des populations urbaines. Une recherche en cours à Quito (EQUATEUR). Communication à la "Primera conferencia Latinoamericana sobre informática en geografía", San Jose, Costa Rica, 5-9 Octobre 1987, 29 p.
- MICHEL A., EBERHARD J.M., LORTIC B., DUREAU F., 1987. El uso de la teledetección para la observación de las poblaciones urbanas. Investigación en curso en Quito (ECUADOR). Communication au deuxième "Simposio Latinoamericano sobre sensores remotos", Bogota, Colombie, 16-20 Novembre 1987, 34 p.
- MICHEL A., EBERHARD J.M., LORTIC B., DUREAU F., 1987. Utilisation de la télédétection pour l'observation des populations urbaines. Une recherche en cours à Quito (EQUATEUR) Communication au colloque SPOT, Paris, 23-27 Novembre 1987, 9 p.
- MICHEL A., DUREAU F., LORTIC B., O. BARBARY, SOURIS M., 1988. Segmentation et classification d'une image satellite en milieu urbain, communication faite au colloque PIXIM, Paris 1988, 17 p.
- SOURIS M., 1987. A geographic information system with relational architecture. Principles and exemple of use of the TIGER system. Communication à la "primera conferencia Latinoamericana sobre informática en geografía", San Jose, Costa Rica, 5-9 Octobre 1987.

Notes de travail et rapports non publiés

- Equipe ORSTOM (UR 406, ATOB, Unité d'Infographie), 1986. Intégration des données de télédétection dans un système d'information géographique : suivi de la morphologie et de la démographie d'une ville. Avril 1986, Paris, 190 p.
- Equipe ORSTOM (UR 406, ATOB, Unité d'Infographie), 1986. Intégration des données de télédétection dans un système d'information géographique : suivi de la morphologie et de la démographie d'une ville. Rapport final, Août 1987, Paris, tome 1 : 53 p., tome 2 (annexes) 538p.
- BARBARY O., 1985. Projet de thèse de mathématiques appliquées aux sciences de l'homme, sous la direction de M. BARBUT, EHESS. Sujet : "Elaboration de plans de sondage pour l'estimation de données spatio-temporelles prenant en compte les informations fournies par la télédétection. Application aux données démographiques en milieu urbain des pays en développement". Paris, 20 p.

- BARBARY O., 1986. Apport d'une stratification morphologique des districts INSEE de recensement dans l'estimation par sondage des effectifs de population à Marseille. Premier rapport intermédiaire à l'ORSTOM, Quito, 47 p.
- BARBARY O., 1988. Mise au point de plans de sondage pour l'estimation de caractères démographiques à Marseille. Evaluation des précisions permises par l'emploi de diverses techniques de sélection, de stratification et d'estimation. Deuxième rapport intermédiaire à l'ORSTOM, Quito, 185 p.
- DUREAU F., 1986. Marseille : tirage de l'échantillon de travail. Quito, 9 p.
- DUREAU F., MICHEL A., 1986. Quito : enquête morphologie urbaine, instructions aux enquêteurs. (2 versions, Français et espagnol), Quito, 19 p.
- DUREAU F., 1987. Quito : enquête bâti/population, instructions aux enquêteurs. (2 versions, Français et espagnol), Quito, 27 p.
- DUREAU F., 1987. Quito : encuesta demografía/migraciones, instrucciones para los encuestadores, Quito, 25 p.
- EBERHARD J.M., 1987. Télédétection urbaine : éléments de caractérisation géographique des quartiers de Quito à l'aide d'une image SPOT. Mémoire de DEA, EHESS, 98 p.
- MICHEL A., 1986. Projet de thèse de géographie, sous la direction de J.P. RAISON, EHESS. Sujet : "Télédétection et observation suivie de la morphologie et de la démographie des villes des pays en développement. Mise au point des méthodes d'analyse des images satellite à haute résolution et évaluation des informations fournies par ces images.", Paris, 9 p.
- MICHEL A., 1986 (et version actualisée en 1987). Le point sur la télédétection urbaine en 1986 à travers la bibliographie française et anglo-saxonne sur le sujet. Recueil des fiches de dépouillement des articles et ouvrages. Paris, ORSTOM, 49 p.
- MICHEL A., 1987. Mise au point des méthodes d'analyse des images satellite à haute résolution : étude statistique du descripteur "densité du bâti". présentation des classifications (image TM Marseille, 1983), Paris, 99 p.
- MICHEL A., 1987. Influence des facteurs météorologiques sur la probabilité d'obtention d'une image satellite SPOT ou Thematic Mapper. Etude sur le site de Quito, Equateur. Quito, 14 p.

Rapports de mission

- BARBARY O., DUREAU F., 1986. Rapport de mission Montpellier 18-19 Février 1986, CNUCS et Maison de la Géographie, Paris, 6 p.
- BARBARY O., DUREAU F., LORTIC B., MICHEL A., 1985. Rapport de mission à Marseille, Paris, 3p et annexes.
- DUREAU F., 1987. Rapport de mission à Bogota (Colombie), 13-20 Novembre 1987, Quito, 4p et annexes.
- LORTIC B., 1987. Rapport de mission à Quito, 31 Décembre 86 - 18 Janvier 87, Paris, 5 p.

- MICHEL A., 1986. Rapport de mission à Quito, 16 Septembre - 17 Novembre 1986, Quito, 7p et annexes.
- MICHEL A., 1988. Rapport de mission à Quito, 1 Septembre - 10 Octobre 1987, Quito, 9p et annexes.
- MICHEL A., 1987. Rapport de mission à Quito, 11 Mars - 2 Mai 1988, Quito, 8p.
- SOURIS M., 1986. Rapport de mission à Quito, 8-22 Décembre 1986, Quito, 5 p.
- SOURIS M., DUREAU F., 1985. Rapport de mission en Equateur, 15-29 Mai 1985 (M. SOURIS) et 19-26 Mai 1985 (F. DUREAU), Paris, ORSTOM, 6 p.

ANNEXES

Quito, Mai 1988.

ANNEXE 1

Documents d'enquête pour l'enquête préliminaire "Morphologie"

QUITO - Enquête morphologie urbaine

Octobre / Décembre 1986

A) INSTRUCTIONS AUX ENQUETEURS

SOMMAIRE

	Page
1. OBJECTIFS SCIENTIFIQUES	110
1. 1. Le programme de recherche	110
1. 2. L'enquête morphologie urbaine	111
2. TECHNIQUE D'ENQUETE	112
3. ORGANISATION DU TRAVAIL	113
4. COLLECTE DE L'INFORMATION	114
4. 1. Principes de base	114
4. 2. Description des batiments	115
4. 3. Description des éléments non bâtis	120
5. EXEMPLE D'ILOT A ENQUETER	123

I. OBJECTIFS SCIENTIFIQUES

I. I. Le programme de recherche

L'objectif du programme de recherche est de mettre au point un système d'observation permanent des populations urbaines, en intégrant l'information satellitaire sur la morphologie urbaine : suivi des effectifs et caractéristiques générales des populations, et mise en oeuvre de sondages spatiaux stratifiés sur des images satellites SPOT ou THEMATIC MAPPER, pour des investigations approfondies sur le comportement des citadins.

Deux idées sont à la base de ce programme :

- Seul un système basé sur la technique des sondages peut remplir les conditions de souplesse et rapidité nécessaires dans les villes à croissance rapide
- Une meilleure connaissance de l'espace urbain constitue un facteur important d'amélioration d'un système d'enquête démographique par sondage : pour l'établissement d'une base de sondage correcte et à jour, composée d'ilôts clairement définis, pour l'optimisation du sondage par stratification sur la morphologie urbaine, et pour l'obtention de résultats localisés, selon un découpage significatif de la ville enquêtée.

Les satellites renseignent de manière continue sur l'occupation du sol en milieu urbain : il s'agit donc, dans la méthode que nous développons, d'utiliser cette information exhaustive sur la morphologie urbaine pour recueillir rapidement, par sondage, des données relatives aux populations urbaines. Les travaux réalisés (par VERNIERE notamment) sur photographies aériennes ont montré que l'application de ce principe permet un allègement de l'échantillon d'enquête (donc, pour un même coût global et un même niveau de précision, une possibilité d'approfondir l'observation) et une spatialisation des résultats.

La réalisation de l'objectif général du programme implique des travaux, en parallèle, dans quatre directions :

- L'interprétation des images satellite à haute résolution en milieu urbain
- Les relations bâti/population et les sondages spatiaux en démographie
- L'intégration des données de télédétection dans des bases de données

localisées

-La télédétection comme système autonome d'observation démographique

Nous avons débuté nos travaux en 1985 sur la ville de Marseille (FRANCE), qui possédait dès cette date toutes les données nécessaires à une première phase de développement méthodologique.

Sur la ville de Quito, trois types d'enquête devront être réalisées :

- Enquête morphologie urbaine
- Enquête morphologie/démographie
- Enquête démographie : effectifs de population et systèmes résidentiels

1. 2. L'enquête morphologie urbaine

La première enquête que nous réalisons à Quito dans le cadre de ce programme de recherche a pour objectif de collecter des informations précises sur l'occupation du sol, c'est-à-dire la morphologie urbaine, de la ville de Quito.

L'observation au sol de la morphologie urbaine sur un échantillon d'îlots de Quito est indispensable pour interpréter les images satellite : ce n'est qu'en confrontant les données de terrain relevées au cours de l'enquête morphologie, et les données radiométriques enregistrées par les satellites, que l'on pourra interpréter correctement les images satellites de Quito. La connaissance de la morphologie urbaine sur l'échantillon d'îlots sert à établir les "clefs d'interprétation" qui seront appliquées aux images satellites sur l'ensemble de la zone urbanisée de Quito.

L'échantillon à enquêter du point de vue morphologie urbaine est constitué de 189 îlots, tirés sur les cartes INEC de 1982 au 1/10000, à l'aide d'une grille de 750 mètres de côté. L'information sur l'occupation du sol doit être recueillie précisément et exhaustivement sur ces 189 îlots, durant les mois d'Octobre, Novembre et Décembre 1986.

2. TECHNIQUE D'ENQUETE

Chaque îlot de l'échantillon est décrit par trois types de documents :

- **Carte** : sur le fond de plan INEC agrandi au 1/1000, sur lequel ne figurent au départ de l'enquête que les limites de chaque îlot à enquêter, les enquêteurs établiront la cartographie de l'occupation du sol de l'îlot. Tous les bâtiments ainsi que tous les éléments non bâtis devront figurer sur ces plans.
- **Fiche de description des bâtiments** : chacun des édifices dessinés sur la carte est décrit sur cette fiche. Il est repéré par un numéro de 1 à n.
- **Photographies** : pour chaque îlot, les chercheurs de l'Orstom feront des photographies depuis un point surplombant l'îlot (étage élevé d'un immeuble ou relief), et au sol, en bordure et/ou à l'intérieur de l'îlot.

Pour chaque îlot de l'échantillon, il doit y avoir concordance complète de l'information contenue dans ces trois documents, à partir desquels seront codés les descripteurs de la morphologie urbaine, sur une fiche de synthèse qui servira aux traitements informatiques.

3. ORGANISATION DU TRAVAIL

Tous les lundi matin, chaque enquêteur se verra attribuer la liste d'îlots à enquêter dans la semaine : il s'agit d'une liste bien précise d'îlots appartenant à l'échantillon. La liste sera établie en fonction des principes suivants :

- un îlot, en moyenne, nécessite une demi-journée (4 heures) de travail pour être enquêté correctement.
- les îlots seront regroupés géographiquement, afin de minimiser les temps de déplacement.

Pour chacun des îlots de la liste hebdomadaire, l'enquêteur recevra une carte au 1/1000, ainsi que des fiches de description.

Des modifications du tissu urbain ayant pu se produire depuis 1982 (date d'établissement de la cartographie INEC), un chercheur accompagnera l'enquêteur pour vérifier la délimitation de tous les îlots que ce dernier doit enquêter durant la semaine.

Afin de rechercher collectivement des solutions aux problèmes rencontrés au cours de la collecte d'informations et de vérifier la cohérence des cartes et fiches de description par rapport aux photographies, deux rendez-vous sont fixés chaque semaine : le mercredi et le vendredi en fin d'après midi.

En cas d'incomplétude ou d'incohérence des cartes et fiches de description, il sera demandé aux enquêteurs de retourner sur le terrain pour corriger ou compléter l'information.

4. COLLECTE DE L'INFORMATION

4. 1. Principes de base

La première opération consiste à rechercher une vision d'ensemble de l'îlot, depuis un point surplombant l'îlot : étage élevé d'un immeuble voisin de l'îlot, ou relief. Ainsi, l'enquêteur pourra procéder plus aisément à :

- la cartographie des éléments bâtis et non bâtis composant l'îlot
- la numérotation de ces éléments
- au remplissage des questions relatives aux toits des bâtiments

D'autre part, certains éléments de l'occupation du sol ne sont visibles qu'en surplombant l'îlot, et non depuis les rues jouxtant l'îlot : cette première phase d'observation depuis un point élevé est donc particulièrement importante.

Dans un second temps, l'enquêteur procédera au remplissage des fiches de description, et complétera le plan d'après ce qui est visible depuis les rues bordant l'îlot, et les cours ou passages intérieurs, s'ils sont accessibles.

Enfin, si certaines réponses demeurent impossibles à obtenir par une vision directe, l'enquêteur questionnera les résidents de l'îlot pour obtenir l'information.

4.2. Description des bâtiments

Pour chaque manzana, l'enquêteur remplit une ou plusieurs fiches de description des bâtiments, selon le nombre de ceux-ci ; chacun des bâtiments est dessiné sur le plan de l'îlot, et décrit dans une colonne de la fiche descriptive.

Définition du bâtiment

On appelle bâtiment toute construction en élévation, en cours ou achevée, à usage d'habitation ou non, qui comporte (ou comportera quand elle sera totalement achevée) au moins un toit supporté par des poteaux.

On inclue donc dans cette catégorie les hangars ou les garages couverts non fermés par des murs, les passages couverts entre bâtiments, les grandes citernes ou silos.

Les bâtiments en construction sont inclus à partir du moment où les fondations sont recouvertes d'un matériau de construction.

Limites du bâtiment

Les limites de chaque bâtiment doivent figurer sur le plan. Dans le cas où il s'agit de constructions jointives, il peut être délicat de déterminer les limites. On adoptera les conventions suivantes :

- a) Tissu urbain continu (type centre historique de Quito) : chaque immeuble est individualisé. L'observation des toitures peut permettre de lever les ambiguïtés sur les limites d'immeubles.
- b) Immeuble élevé avec garage attenant (1 seul niveau) : on distingue deux bâtiments, l'immeuble et le garage, qui ont des caractéristiques architecturales bien différentes.
- c) Villa composée de blocs de constructions de différents niveaux : on ne compte qu'un seul bâtiment pour l'ensemble (y compris le garage), si ces blocs appartiennent visiblement à une même unité d'habitation.
- d) Immeuble élevé composé de différents blocs de constructions de différents niveaux présentant une unité architecturale : on ne compte qu'un seul bâtiment.

Numéro du bâtiment

Tout bâtiment situé dans l'îlot est identifié par un numéro de 1 à n. A chaque îlot, la numérotation redémarre à 1 ; au sein d'un même îlot, deux bâtiments ne doivent pas porter le même numéro. Sur la carte, ce numéro est indiqué au centre du bâtiment ; sur la fiche, il est noté en haut de chacune des colonnes.

Stade de construction

Il s'agit de distinguer :

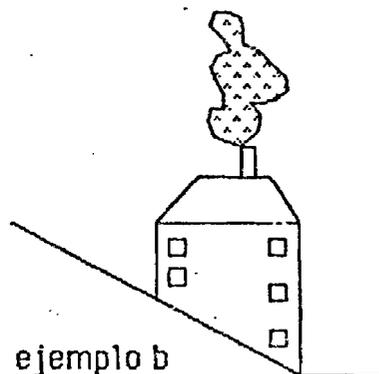
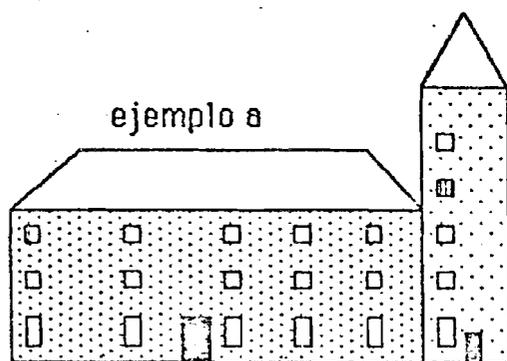
- En construction : bâtiment en chantier (à partir du moment où les fondations sont recouvertes d'un matériau de construction) non encore habitable (par exemple, toiture non terminée, vitres non posées...etc)
- Terminé : bâtiment habitable, dont l'étanchéité est assurée (toit, portes et fenêtres), même si les finitions intérieures (peinture, carrelage...etc) ne sont pas faites.

Un bâtiment habitable dans les étages inférieurs, dont le dernier niveau est constitué d'une dalle de béton avec des poteaux en béton, ou des fers à béton qui dépassent, est considéré comme "terminé".

Nombre d'étages

Nombre d'étages que comporte le bâtiment, y compris le rez de chaussée, et non compris les penthouses.

Si, en terrain plat, un même bâtiment complexe comporte différents nombres d'étages selon les parties du bâtiment, indiquer le nombre minimal et le nombre maximal (exemple a : 3/5, pour décrire un bâtiment comptant 3 étages dans une partie, et 5 dans l'autre).

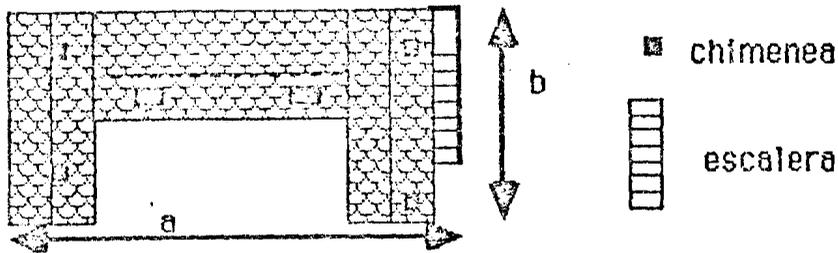


Si en raison d'une forte pente, un bâtiment compte 2 étages sur un des côtés et 3 étages sur un autre, en contrebas, indiquer le nombre d'étages minimal et maximal (2/3 dans l'exemple b), et ne pas oublier de noter la pente du sol sur le plan d'ilot.

Dimensions au sol

Les côtés a et b désignent les côtés du plus petit rectangle incluant le bâtiment (y compris les escaliers extérieurs menant aux étages du bâtiment).

Ces côtés a et b sont indiqués sur le plan de la façon suivante :

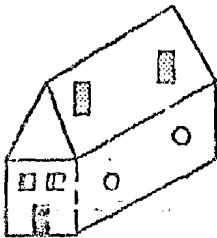


Sur la fiche de description des bâtiments, les mesures des côtés sont exprimées en mètres.

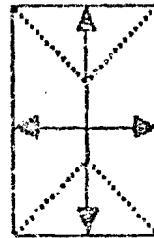
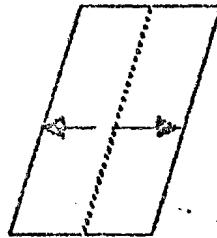
Nombre de pans du toit

Toit plat : Nombre de pans = 0. Ne rien indiquer sur le plan.

Toit avec pente(s) : Indiquer le nombre de pans sur la fiche, et le sens de la pente pour chacun des pans sur le croquis de l'ilet par une flèche dirigée vers le bas de la pente.



Toit à 2 pans



Toit à 4 pans

Matériau prédominant du toit

Il s'agit du matériau qui occupe la plus grande surface sur le toit.

Dans le cas où le toit est composé de plusieurs matériaux occupant des surfaces sensiblement égales, indiquer ces différents matériaux sur la fiche, et les délimiter sur le plan.

En dehors des matériaux les plus courants cités sur la fiche, pour lesquels il suffit de mettre une croix dans la case correspondante, on peut rencontrer à Quito d'autres types de matériaux pour les toitures : shingle, verre, carrelage ...etc. Dans ce cas, il faut indiquer en clair la dénomination du matériau, si elle est connue, ou, sinon, décrire précisément ce matériau sur la fiche "observations complémentaires".

Couleur prédominante du toit

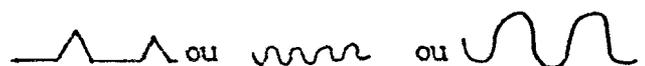
Indiquer la couleur occupant la plus grande surface sur le toit du bâtiment, en faisant attention d'être cohérent avec le matériau prédominant : c'est la couleur du(des) matériau(x) prédominant(s) que l'on doit indiquer.

Forme du matériau du toit

Sur la fiche, dans la colonne relative au bâtiment, faire un croquis, en coupe, du matériau prédominant du toit.

Exemples de formes possibles selon les matériaux :

Dalle de béton, carrelage, paille : _____

Eternit :  ou _____ ou _____

Tuiles :  ou _____

Nombre de cheminées

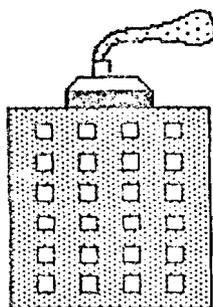
Indiquer le nombre de cheminées présentes sur le toit du bâtiment.

Présence de poteaux

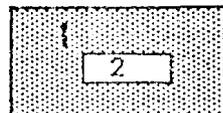
Nous retrouvons ici le cas des bâtiments habitables dans les étages inférieurs et dont le dernier étage est constitué d'une dalle en béton, d'où émergent des fers à béton ou des poteaux en bétons.

Présence de penthouse

S'il y a une penthouse (c'est à dire un petit bâtiment sur la terrasse supérieure d'un immeuble) sur un bâtiment, la dessiner sur le plan, lui attribuer un numéro et reporter celui-ci à la fois sur le plan et dans la colonne du bâtiment sur lequel elle est située (en face de la ligne "présence de penthouse") et la décrire dans une autre colonne de la fiche, comme si c'était un bâtiment indépendant.



Edificio de 6 pisos
con un penthouse



Croquis sobre el plano

Nombre de citernes

Indiquer le nombre de citernes couvertes présentes sur le toit du bâtiment : on désigne sous le terme de "citernes" les réservoirs d'eau qui servent à alimenter en eau le bâtiment en cas de coupure ou insuffisance du réseau public de distribution d'eau potable.

Présence de végétation

Indiquer s'il y a "aucune", "peu", "beaucoup" de végétation :

- sur les balcons saillants : balcons dépassant du mur extérieur du bâtiment
- sur la terrasse du toit

Autres observations sur les bâtiments

Dans cette rubrique, préciser rapidement une information incomplète dans les rubriques prévues par le questionnaire, ou indiquer la présence éventuelle d'éléments non répertoriés (capteurs solaires sur le toit par exemple). Si nécessaire, utiliser la fiche "observations complémentaires", en spécifiant bien le numéro du bâtiment dont il s'agit.

4.3. Description du non bâti

Définition

On appelle "non-bâti" toute la surface de l'îlot où il n'y a pas de bâtiment.

Chaque zone homogène (du point de vue de l'occupation du sol) constitue un "élément non bâti".

La somme des surfaces occupées par des bâtiments, et des surfaces occupées par des éléments non-bâtis, doit correspondre à la surface totale de l'îlot (ce dernier incluant les rues qui bordent l'îlot jusqu'à la moitié de la chaussée).

Pente du sol

Indiquer sur le plan de l'îlot la pente du sol par une flèche :

- Pente faible ou nulle —————>
- Pente moyenne —————>>
- Pente forte —————>>>

Identification

Chaque élément non bâti et homogène de l'îlot doit être dessiné sur le plan : l'enquêteur lui attribue une lettre de A à Z (lettres majuscules), qui indique le type de non bâti. Sur le plan de l'îlot, cette lettre doit être inscrite au centre de la zone occupée par cet élément non bâti.

Sur la nomenclature (page 14), sont répertoriés 20 types de non bâti : si un élément correspond à un de ces types, l'enquêteur utilisera toujours ce codage. Pour toute autre catégorie de non bâti, il utilisera une autre lettre (U à Z), et indiquera sa signification sur la fiche "Observations complémentaires".

Si plusieurs éléments non bâtis d'un îlot sont du même type, la lettre qui correspond à ce type sera répétée dans chacune des zones occupées par ces éléments.

Végétation

Délimiter et identifier la végétation à l'intérieur de l'îlot, ainsi que sur la voirie qui le borde. Ne pas oublier la végétation sur les trottoirs (pelouse, arbres..)

Ne délimiter les arbres un par un que s'ils sont réellement isolés.

Si il y a une rangée d'arbres ou un groupe d'arbres, ne considérer la rangée ou le groupe que comme un seul élément non bâti (lettre G)

Si l'espèce est connue, l'indiquer sur la fiche "Observations complémentaire".

Voirie

On appelle voirie : les rues y compris les trottoirs et le terre-plein, les parkings, les cours, les passages intérieurs (sans toit) de l'îlot, les allées, les escaliers au sol.

Orstom - IPOH

QUITO - Morphologie- Oct./Dec. 1986

Description du non bâti

Indiquer sur le plan la pente du sol et la présence de quebradaE
A
U

- A - rivière
- B - canal
- C - réservoir d'eau, bassin (sans toit)
- D - piscine

V
E
G
E
T
A
T
I
O
N

- E - pelouse
- F - fleurs
- G - arbres (noter l'espèce)
- H - arbustes sans fleurs
- I - arbustes avec fleurs
- J - potager
- K - pâturage
- L - champ cultivé

V
O
I
R
I
E

- M - bitume
- N - pavés ou dalles en béton
- O - pierres
- P - terre ou sable

V
O
I
T
U
R
E
S

- Q - avec des voitures en circulation
- R - avec des voitures en stationnement
- S - sans voitures

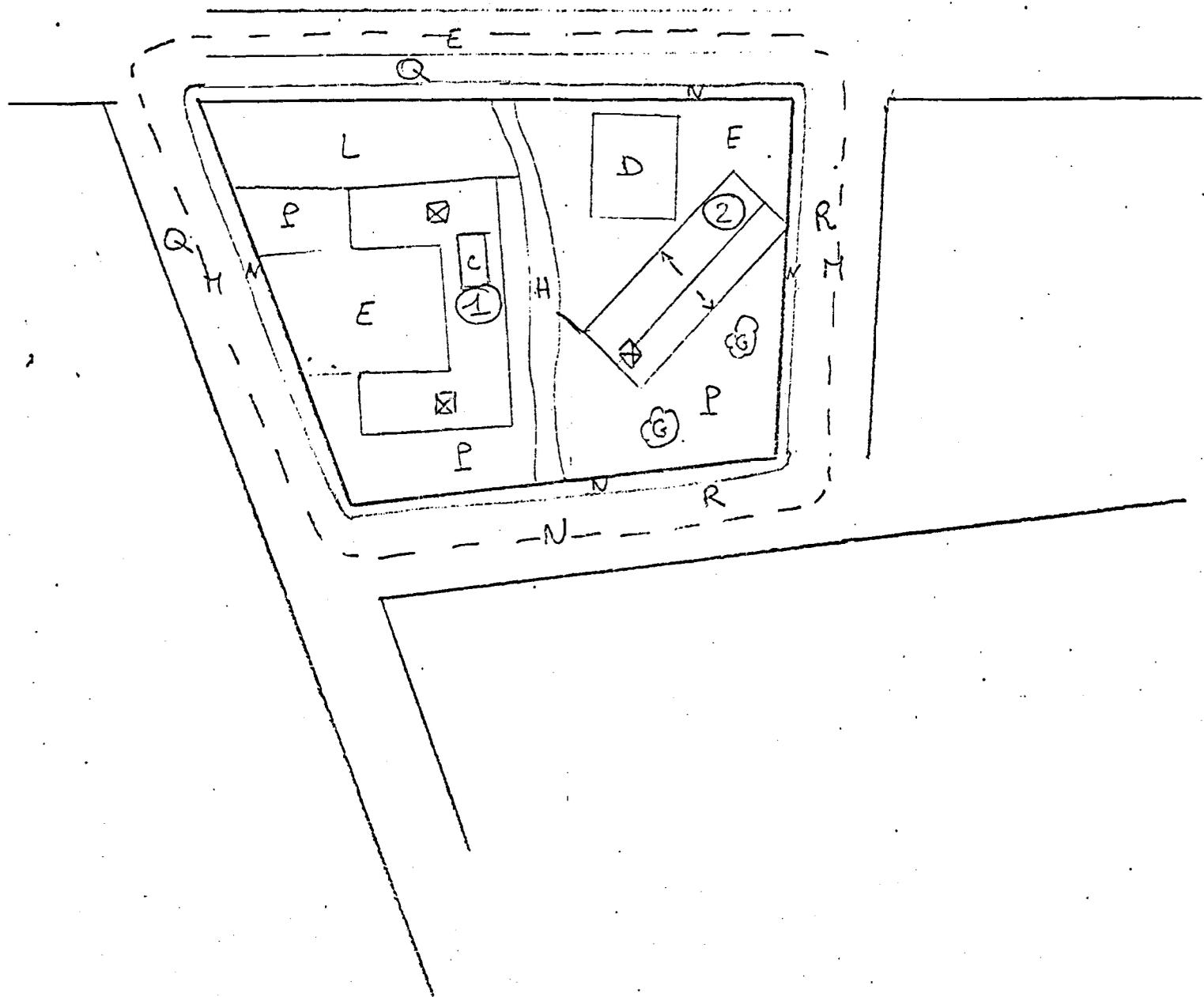
AUTRES

T - mur de cloture

Pour toute autre catégorie de non bâti (par ex., carrelage..etc.) utiliser une lettre de U à Z, et noter sa signification sur la fiche "autres observations".

Si il y a beaucoup de matériaux entreposés sur le sol, distinguer la zone qui est couverte par ces matériaux de la zone qui n'est pas couverte, chacune de ces zones correspondant à un élément non bâti.

Utiliser une lettre de U à Z encore non utilisée pour les matériaux, et indiquer la signification de cette lettre sur la fiche "Observations complémentaires."



Orstom - IPGH

QUITO-Morphologie-Oct./Dec 1986

Numero d'îlot :		DESCRIPTION DES BATIMENTS				
Date: . Heure:						
Numero du batiment correspondant au repère porté sur le plan						
D I M E N S I O N S	Bâtiment: Terminé	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	En construction	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Nombre d'étages (y compris RDC)	-----	-----	-----	-----	-----
	Dimensions au sol : Côté a	-----	-----	-----	-----	-----
	Côté b	-----	-----	-----	-----	-----
	(Indiquer ces côtés sur le plan)					
	Nombre de pans du toit (0 si plat)	-----	-----	-----	-----	-----
T O I T	Matériau prédominant du toit					
	Dalle de béton	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Eternit ou similaires	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Zinc ou similaires	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Tuiles vernies	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Tuiles sans vernis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Paille ou similaires	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Autres (shingle, carrelage..)	-----	-----	-----	-----	-----
Couleur prédominante du toit						
(Blanc, Marron foncé, Marron clair, Gris, Brique, Noir, Rouge ou Vert)	-----	-----	-----	-----	-----	
Forme du matériau du toit						
(Faire un croquis en coupe)						
Autres éléments sur le toit						
Nombre de cheminées	-----	-----	-----	-----	-----	
Présence de poteaux en béton	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Présence de penthouse (à décrire)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Nombre de tanques	-----	-----	-----	-----	-----	
Présence de végétation						
Sur les balcons saillants	-----	-----	-----	-----	-----	
Sur le toit	-----	-----	-----	-----	-----	
Autres observations sur le bâtiment		-----	-----	-----	-----	-----
A U T R E S						

Orstom - IPGH

QUITO-Morphologie-Oct./Dec 1986

Numero d'îlot: 48	OBSERVATIONS COMPLEMENTAIRES
Date: 1/10/86 Heure: 15h	
Description du type d'îlot (notamment le standing de l'habitat): Résidentiel haut standing	
Détails sur la description des bâtiments ou du non bâti de l'îlot :	

B) CODIFICATION DE L'ENQUETE MORPHOLOGIE

GENERALITES SUR LA CODIFICATIONNOMBRE D'ÎLOTS SAISIS ET CODES

* 199 îlots ont été codés et saisis, c'est à dire :

- îlots 101 à 166 inclus, sauf l'îlot 155 (n'a fait l'objet que de photos)
- îlots 201 à 276 inclus, sauf les îlots 243 (parc à côté de la Patria) et 265 (parc de la Caroline)
- îlots 302 à 362 inclus, sauf l'îlot 351 (usine qui n'a fait l'objet que de photos).

Le descriptif de ces 199 îlots (recto de la fiche de codification) correspond au fichier QMANZANS.

* A ces 199 îlots, correspondent 2133 genres de bâtiments, décrits dans le fichier QBATIMENT (saisie des données du verso de la fiche de codification)

I. RECTO DE LA FICHE DE CODIFICATIONIDENTIFIANTS**N° Manzana**

Numéro à 3 chiffres de 101 à 165, de 201 à 276 et de 302 à 362 affectés avant l'enquête morphologie aux îlots composant l'échantillon.

N° INEC

* Identifiant INEC en 7 chiffres (Zone en 3 chiffres, Secteur en 2 chiffres et Manzana en 2 chiffres) tel qu'il peut être lu sur les cartes au 1/10.000, suivi d'un code à 1 chiffre précisant si :

- 1 = la manzana INEC a été conservée entièrement pour l'enquête morphologique
- 2 = la manzana n'a été enquêtée qu'en partie
- 3 = la manzana enquêtée est plus grande que la manzana INEC

* *Dans le cas où la manzana enquêtée correspond à 2 manzanas INEC entières, deux fiches de codification ont été établies ; sur chacune, a été décrite morphologiquement la portion de "manzana 1986" correspondant à la manzana INEC. C'est le cas de la manzana 237, qui correspond en fait à 2 manzanas INEC ; il y a donc deux fiches, l'une notée 237, l'autre notée 275.*

* Si la manzana enquêtée morphologiquement en octobre-novembre 1988 *n'existe pas sur les plans INEC* de 1982 au 1/10000 (hors limites INEC de Quito, ou non existante en 1982), elle reçoit le code 999.99.99.9.

N° enquêteur

Codes :

- 1 = Ligia
- 2 = Xavier
- 3 = Olivier
- 9 = îlot ayant uniquement fait l'objet de photos.

SURFACES

Une fois connues les surfaces des îlots de l'échantillon par TIGRE, les surfaces mesurées sur carte devraient être corrigées ; il ne faut utiliser les chiffres de surface notés sur la fiche de codification qu'avec prudence. Ils ne sont destinés qu'à permettre le calcul des ratios (ex : % non bâti) n'ayant qu'une valeur *relative*, % à rassembler en classes.

Surface îlot

Surface de l'îlot en m^2 , non compris la voirie bordant l'îlot. Quand l'îlot est représenté sur la carte au 1/15000 de l'IGM, il est mesuré sur cette carte ; dans le cas contraire, la mesure est faite sur les plans INEC au 1/10.000 (Voir note dans le coin haut droite des fiches de codification des manzanas). Dans tous les cas, la mesure est faite avec une grille millimétré sur support stable.

La voirie interne à l'îlot est incluse dans ce calcul de surface.

Surface 1/2 voirie

Moitié de la surface en m^2 de la voirie encadrant l'îlot.

Mêmes principes de mesure que pour la surface de l'îlot.

Surface totale

Surface totale = surface îlot + surface 1/2 voirie. Exprimée en m^2 .

TYPE MORPHOLOGIQUE**Carte Caza/Godard**

Il s'agit des modalités de la légende de la carte "Esbozo de tipología del habitat", établie par P. CAZAMAJORE et H. GODARD en 1984, soit :

- 1 = Zone industrielle
- 2 = Zone de végétation
- 3 = Centre des affaires et de commerces de luxe
- 4 = Zone non urbanisée en 1984 (le code 16 des fiches de codification a été recodé 4 dans le fichier, car même signification)
- 5 = Zone urbaine "diverse" jusqu'en 1983
- 6 = Habitat classe haute
- 7 = Habitat mixte classes haute et moyenne
- 8 = Habitat classe moyenne
- 9 = Programmes de vivienda de l'Etat classe moyenne
- 10 = Programmes de vivienda de l'Etat classe populaire
- 11 = Habitat mixte classes haute et populaire
- 12 = Habitat mixte classes moyenne et populaire
- 13 = Habitat populaire consolidé ou en voie de consolidation
- 14 = Habitat populaire non consolidé
- 15 = Taudification verticale (centre) et horizontale (quartiers périphériques)

Ont été ajoutées les modalités :

- 17 = Commerce
- 99 = Type non défini

Enquêteur et Photo

Les mêmes codes sont utilisés pour le type défini par l'enquêteur et pour celui noté au moment de la prise de photo et au vu des diapos.

Après examen des modalités retenues par chacun des enquêteurs, ont été retenus les types suivants (les classes 1 à 6 se rapportant aux catégories d'habitat) :

- 1 = Alta ou Burgesia alta
- 2 = Media alta
- 3 = Media media
- 4 = Media baja
- 5 = Media y Baja
- 6 = Baja
- 7 = Industriel
- 8 = Scolaire
- 9 = Commercial
- 10 = Equipements sociaux
- 11 = Végétation (hors terrains agricoles)
- 12 = Agricole
- 13 = Bureaux

Un îlot dont seulement une partie est consacrée à l'habitation est classé avec les codes 1 à 6 selon le standing de l'habitat ; les codes 7 à 13 sont utilisés exclusivement pour les îlots ne comptant aucun bâtiment d'habitation.

DATI**Surface bâtie**

Il s'agit de la surface correspondant à la somme des surfaces de tous les bâtiments de la manzana, qu'ils soient terminés ou non.

Densité du bâti

Cinq niveaux de densité, établis d'après le plan de l'enquêteur et les diapos :

- 1 = Non bâti (pas un seul bâtiment sur la manzana)
- 2 = Densité très faible
- 3 = Densité faible
- 4 = Densité moyenne
- 5 = Densité forte
- 6 = Densité très forte

Implantation du bâti dans l'îlot

Codée d'après le plan établi par l'enquêteur et les diapos :

- 1 = Pas de type d'implantation particulier. Bâti homogène sur tout l'îlot, sans directions particulières.
- 2 = Bâtiments concentrés au centre de l'îlot, le bord des rues étant non bâti
- 3 = Bâtiments non alignés, concentrés sur le pourtour de l'îlot, près des rues, le centre de l'îlot étant vide
- 4 = Bâtiments concentrés dans un coin de l'îlot, le reste étant à peu près non bâti
- 5 = Bâtiments concentrés dans deux angles opposés de l'îlot, le reste étant à peu près non bâti
- 6 = Bâtiments alignés régulièrement près des rues, le centre de l'îlot étant vide
- 7 = Répartition des bâtiments homogène, selon un plan régulier sur tout l'îlot (en semis, et non en barre)
- 9 = Bâtiments concentrés le long de la rue principale
- 10 = Maisons accolées formant des barres continues (dans ce cas, la barre constitue *un seul* bâtiment ; c'est cette barre qui est décrite dans le tableau descriptif des bâtiments terminés, au dos de la fiche de codification)
- 11 = A rassembler avec code 10. Pas de différence significative entre ces 2 classes
- 12 = Idem 10, mais les maisons n'étant pas strictement accolées, la "barre" est divisée en n bâtiments semblables, qui sont tous décrits dans le tableau du dos de la fiche de codification.

Nombre de bâtiments

Le nombre de bâtiments total (= terminés + en construction), le nombre de bâtiments terminés et le nombre de bâtiments en construction correspondent aux bâtiments décrits par les enquêteurs sur les questionnaires.

Les pent-houses sont comptabilisés et décrits comme tout autre bâtiment de plein pied.

Répartition des bâtiments terminés par genre

Dans la colonne correspondant à chaque genre, est porté le nombre de bâtiments correspondant à ce genre ; ne sont comptabilisés et donc décrits au verso de la fiche de codification que les bâtiments terminés.

Pente du toit

S'il n'existe pas de direction majoritaire des faîtes, le code est 5 (cette variable ayant été recodée après saisie, considérer les modalités inscrites dans le fichier MANZANS et non celles notées dans les fiches de codification).

S'il existe une direction majoritaire, le code est :

- 1 = Nord / sud
- 2 = Est / ouest
- 3 = Nord est / sud ouest
- 4 = Nord ouest / sud est

S'il n'existe pas de bâtiments terminés dans la manzana, le code est 9.

Penthouse, Poteaux fer/béton et Cheminées

Pour chacune de ces trois variables, est noté le nombre de bâtiments ^{terminés} présentant l'un de ces éléments sur son toit.

Végétation sur balcons saillants

- 1 = Aucune végétation sur des balcons saillants
- 2 = Végétation sur balcons saillants peu importante
- 3 = Végétation sur balcons saillants très importante

* Dans le cas d'un *ilot ne comptant aucun bâtiment*, toutes les variables de la section "Bâti" sont codées 0 pour les nombres de bâtiments, et 9 ou 99 pour les autres variables de la section.

NON BATI

Toute la description du NON BATI ne concerne que l'intérieur de l'îlot, à l'exclusion de la voirie entourant l'îlot.

Pente du sol

Direction (mêmes codes pour dominante et sous-dominante) :

- 1 = Nord
- 2 = Nord est
- 3 = Est
- 4 = Sud est
- 5 = Sud
- 6 = Sud ouest
- 7 = Ouest
- 8 = Nord ouest

Niveau : 1 = aucune pente (sol quasiment plat)

- 2 = pente faible
- 3 = pente forte
- 4 = pente très forte

Dans le cas où il n'existe pas de pente du tout, les codes sont :

- Dominante : Direction = 9, Niveau = 1
- Sous-dominante : Direction = 9, Niveau = 9

Eau

Le type de surface en eau reprend les rubriques prévues dans le manuel d'enquête, et il est codé :

- 1 = Rivière
- 2 = Canal
- 3 = Réservoir d'eau, bassin (sans toit)
- 4 = Piscine

Si pas de surface en eau dans la manzana, le code est 9.

La surface en m² est estimée d'après le plan établi par l'enquêteur et les diapos de la manzana ; elle correspond au total des surfaces en eau, tous types 1 à 4 confondus, de la manzana.

Végétation dans l'îlot

Le type de surface occupée par de la végétation reprend les rubriques prévues dans le manuel d'enquête, et il est codé :

- 1 = Pelouse (en fait, toute surface en herbe non consacrée spécifiquement à du pâturage)
- 2 = Fleurs
- 3 = Arbres
- 4 = Arbustes sans fleurs

- 5 = Arbustes avec fleurs
- 6 = Potager
- 7 = Pâturage
- 8 = Champ cultivé

Si pas de surface occupée par de la végétation dans la manzana, le code est 9.

La surface en m^2 est estimée d'après le plan établi par l'enquêteur et les diapos de la manzana ; elle correspond au total des surfaces occupées par de la végétation, tous types 1 à 8 confondus, de la manzana.

Dans le cas où il n'existe pas de végétation sous-dominante (pas du tout de végétation dans l'îlot ou végétation d'un seul type, noté "dominant"), le code du type de végétation sous-dominante est 9.

Voirie dans l'îlot

On appelle "voirie" toute surface non bâtie et non occupée par de la végétation.

Le type de surface occupée par de la voirie reprend les rubriques prévues dans le manuel d'enquête, et il est codé :

- 1 = Bitume
- 2 = Pavés ou dalles en béton
- 3 = Pierres
- 4 = Terre ou sable
- 5 = Carrelage
- 6 = Matériaux de bâtiment entreposés sur le sol
- 8 = Terre battue (tennis)

Si pas de surface occupée par de la voirie dans la manzana, le code est 9.

La surface en m^2 est estimée d'après le plan établi par l'enquêteur et les diapos de la manzana ; elle correspond au total des surfaces occupées par de la voirie, tous types 1 à 8 confondus, de la manzana.

Dans le cas où il n'existe pas de voirie sous-dominante (pas du tout de voirie dans l'îlot ou voirie d'un seul type, noté "dominant"), le code du type de voirie sous-dominant est 9.

Murs de clôture

Il s'agit de murs de séparation à l'exclusion des murs constitutifs des bâtiments (soutenant leur toiture).

A l'intérieur de l'îlot, les codes sont :

- 1 = Aucun mur de clôture
- 2 = Murs de clôture peu nombreux
- 3 = Murs de clôture très nombreux

Sur le périmètre de l'îlot, le code correspond au % du périmètre de l'îlot occupé par des murs.

1/2 VOIRIE BORDANT L'ÎLOT**Trottoir**

Dans le cas où la chaussée de circulation est bordée par un trottoir pour le cheminement des piétons, on utilise les codes suivants pour décrire les matériaux dominants et sous-dominants des trottoirs :

- 1 = Bitume
- 2 = Pavés ou dalles en béton
- 3 = Pierres
- 4 = Terre ou sable
- 5 = Carrelage
- 6 = Matériaux de bâtiment entreposés sur le sol
- 7 = Pelouse
- 8 = Arbres

Dans le cas où tous les trottoirs bordant l'îlot sont construits dans le même matériau, le matériau dominant est codé d'après les codes cités ci-dessus, et le matériau sous-dominant est codé 9.

Dans le cas où il n'existe pas de trottoir, le code des matériaux dominants et sous-dominants est 9.

Chaussée

Dans le cas où il y a une chaussée de circulation, on utilise les codes suivants pour décrire les matériaux dominants et sous-dominants des chaussées :

- 1 = Bitume
- 2 = Pavés ou dalles en béton
- 3 = Pierres
- 4 = Terre ou sable
- 5 = Carrelage

Dans le cas où toutes les chaussées bordant l'îlot sont construites dans le même matériau, le matériau dominant est codé d'après les codes cités ci-dessus, et le matériau sous-dominant est codé 9.

Dans le cas où il n'existe pas de chaussée (limite d'îlot = quebrada, mur ou végétation) le code des matériaux dominants et sous-dominants est 9.

Le terre-plein central n'est pas considéré comme faisant partie de la chaussée ; c'est pourquoi il n'est pas prévu de code désignant de la végétation pour les chaussées.

Pelouse, Arbres et Arbustes

Ces trois variables décrivent le % de la longueur de voirie bordant l'îlot occupé par chacun de ces types de végétation, que ce soit sur les trottoirs ou sur les terre-plein centraux.

Dans la quasi-totalité des cas, cette végétation est intermittante, c'est à dire que la pelouse alterne avec le béton sur les trottoirs, et que les arbres ne forment pas une ligne continue de végétation mais sont espacés entre eux de quelques mètres ; c'est néanmoins la longueur totale occupée par cet agencement (pelouse/béton, par exemple) qui est prise en compte pour le calcul du %.

Voitures

En stationnement : % de longueur de bordure d'îlot occupé par des voitures en stationnement (au maximum)

En circulation : % de longueur de bordure d'îlot occupé par des voitures en circulation. Ce qui correspond, en fait, au % du périmètre de l'îlot disposant d'une chaussée où peuvent circuler des voitures.

2. VERSO DE LA FICHE DE CODIFICATION

* N° Manzana, N° INEC et N° Enquêteur : idem recto de la fiche de codification.

* Ne sont décrits que les bâtiments considérés comme terminés par l'enquêteur.

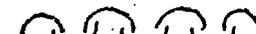
* Le nombre de bâtiments de chaque genre est noté sur le recto de la fiche de codification ; il peut y avoir au maximum 17 genres pour une seule manzana. Sont regroupés dans chaque genre les bâtiments de la manzana qui ont des caractéristiques du bâti semblables ou très voisines.

* Les codes utilisés pour les descripteurs des bâtiments sont ceux notés sur le verso de la fiche de codification, soit :

Matériau du toit

Dalle de béton
Eternit ou similaires
Tôle ou similaires
Tuiles vernies
Tuiles non vernies
Paille ou similaires
Carrelage
Autres
Non observé

Forme du matériau du toit

1	
2	
3	
4	
5	
6	Autre
9	Non observé

Couleur du toit

1 Blanc	9 Bleu clair
2 Orange	10 Bleu foncé
3 Marron clair	11 Gris
4 Marron foncé	12 Noir
5 Brique	13 Autre
6 Rouge	99 Non observé
7 Vert clair	
8 Vert foncé	

* Les codes 9 ou 99 n'ont en fait pas été utilisés pour les descripteurs des bâtiments, puisqu'il n'est pas apparu de cas "non observé".

* Couleur du toit

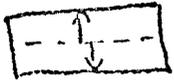
A l'issue de la codification, il s'est avéré nécessaire de regrouper des modalités pour le descripteur "couleur du toit", les qualificatifs "clair" et "foncé" pour les couleurs bleu et vert n'ayant pas la même signification selon les enquêteurs. On a donc finalement :

- Couleur code 7 : vert clair ou vert foncé (code 8 supprimé)
- Couleur code 9 : bleu clair ou bleu foncé (code 10 supprimé)

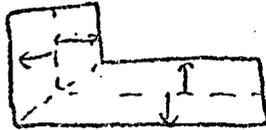
* Le code 14 = jaune a été ajouté, afin de rendre compte de quelques cas d'Etenit jaune.

* **Nombre de pans du toit**

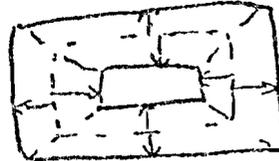
Dans le cas de bâtiments en angle, la codification a été fait comme suit :



2 pans



2 pans



2 pans

* **Forme du matériau du toit**

A été ajoutée la modalité 6 =

(ex : marché national, dans le sud de Quito)

* **Surface au sol**

Il s'agit de la surface moyenne des bâtiments composant le genre de bâtiment dans la manzana.

L'analyse des surfaces de bâtiments selon les enquêteurs (histogrammes, tableaux de fréquences et indicateurs statistiques) a conduit à augmenter les surfaces mesurées par l'enquêtrice de 20 %, afin d'homogénéiser ses mesures avec celles de l'enquêteur ; au vu des tests effectués, il était net que l'enquêtrice avait systématiquement sous-estimé les dimensions des bâtiments, alors que celles de l'enquêteur étaient correctes.

CODIFICATION QUITO - FICHE MANZANA

N° Manzana N° INEC N° enquêteur (Ligia=1, Xavier=2)
(Entière=1, en partie=2)

Surface îlot Surface 1/2 voirie Surface totale
(En m² - Mesure sur JGM 1/15000)

TYPE Carte Caza/Godard Enquêteur Photo

BATI Surface (m²) (D'après quest./plan)

Densité du bâti (D'après plan/quest./photo) Implantation du bâti dans l'îlot

Nombre de bâtiments : Total Terminés En construction

Répartition des bâtiments terminés par genre :

Genre	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Nbre																	

Pente du toit (direction majoritaire des faîtes) (Sans pente 0, N/S1, E/W2, NE/SW3, NW/SE4)

Nbre de bâtiments terminés avec penthouse

Nbre de bâtiments terminés avec poteaux fer/béton

Nbre de bâtiments terminés avec cheminées

Présence de végétation sur balcons saillants (Aucun 1, Peu nomb. 2, Très nomb. 3)

NON BATI

Pente du sol : Dom. : Direction (Ss pente=0, N=1, NE=2, E=3, Niveau (Aucun 1, Faible 2,

Ss-dom. : Direction (SE=4, S=5, SW=6, W=7, NW=8) Niveau Forte 3, Très forte 4)

Eau : Type Surface (m²)

Végétation dans l'îlot : Hors bâti : Type dom. Type ss-dom. Surface (m²)

Couvrant le bâti : Type dom. Type ss-dom. Surface (m²)

Voirie dans l'îlot : Matériau dom. Matériau ss-dom. Surface (m²)

Présence de murs de clôture : dans l'îlot (Aucun 1, Peu nomb. 2, Très nomb. 3)

sur le périmètre de l'îlot (En % du périmètre de l'îlot)

1/2 VOIRIE BORDANT L'ÎLOT Trottoir : Mat. dom. Mat. ss-dom. Chaussée : Mat. Dom Mat. Ss-dom.

Pelouse Arbres Arbustes (en % longueur de bordure)

Voitures : En stationnement En circulation

CODIFICATION QUITO - FICHE MANZANA

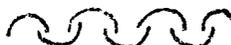
N° Manzana N° INEC : N° enquêteur (Ligia=1.
Xavier=2)

N° du genre de bâtiment dans la manzana	Descripteurs des bâtiments terminés					
	Matériau du toit	Forme mat. du toit	Couleur mat. du toit	Nbre pans du toit	Nbre étages (Rdc inclus)	Surface au sol (m ²)
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

Matériau du toit

- 1 Dalle de béton
- 2 Eternit ou similaires
- 3 Zinc ou similaires
- 4 Tuiles vernies
- 5 Tuiles non vernies
- 6 Paille ou similaires
- 7 Carrelage
- 8 Autres
- 9 Non observé

Forme du matériau du toit

- 1 
- 2 
- 3 
- 4 
- 5 
- 6 Autre
- 9 Non observé

Couleur du toit

- | | |
|----------------|----------------|
| 1 Blanc | 9 Bleu clair |
| 2 Orange | 10 Bleu foncé |
| 3 Marron clair | 11 Gris |
| 4 Marron foncé | 12 Noir |
| 5 Erique | 13 Autre |
| 6 Rouge | 99 Non observé |
| 7 Vert clair | |
| 8 Vert foncé | |

ANNEXE 2

Document d'enquête pour l'enquête préliminaire "Bâti / population"

QUITO - Enquête bâti/population

Avril 1987

A) INSTRUCTIONS AUX ENQUETEURS

SOMMAIRE

	Page
1. OBJECTIFS SCIENTIFIQUES	144
1. 1. Le programme de recherche	144
1. 2. L'enquête bâti/population	145
2. TECHNIQUE D'ENQUETE	146
3. ORGANISATION DU TRAVAIL	147
4. COLLECTE DE L'INFORMATION	148
4. 1. Principes de base	148
4. 2. Recensement, délimitation et numérotation des logements	149
4. 3. Définition de la population à enquêter. Statut de résidence	151
4. 4. Renseignements relatifs aux personnes	152
5. LA FICHE RECAPITULATIVE	163
6. EXEMPLE DE MANZANA ET DE LOGEMENT	165

I. OBJECTIFS SCIENTIFIQUES

I. I. Le programme de recherche

L'objectif du programme de recherche est de mettre au point un système d'observation permanent des populations urbaines, en intégrant l'information apportée par les images satellite sur la morphologie urbaine : suivi des effectifs et caractéristiques générales des populations, et mise en oeuvre de sondages spatiaux stratifiés sur des images satellites SPOT ou THEMATIC MAPPER, pour des investigations approfondies sur le comportement des citadins.

Deux idées sont à la base de ce programme :

- Seul un système basé sur la technique des sondages peut remplir les conditions de souplesse et rapidité nécessaires dans les villes à croissance rapide
- Une meilleure connaissance de l'espace urbain constitue un facteur important d'amélioration d'un système d'enquête démographique par sondage.

Les satellites renseignent de manière continue sur l'occupation du sol en milieu urbain : il s'agit donc, dans la méthode que nous développons, d'utiliser cette information exhaustive sur la morphologie urbaine pour recueillir rapidement, par sondage, des données relatives aux populations urbaines. Les travaux réalisés (par VERNIERE notamment) sur photographies aériennes ont montré que l'application de ce principe permet un allègement de l'échantillon d'enquête (donc, pour un même coût global et un même niveau de précision, une possibilité d'approfondir l'observation) et une spatialisation des résultats.

La réalisation de l'objectif général du programme implique des travaux, en parallèle, dans quatre directions :

- L'interprétation des images satellite à haute résolution en milieu urbain
- Les relations bâti/population et les sondages spatiaux en démographie
- L'intégration des données de télédétection dans des bases de données localisées
- La télédétection comme système autonome d'observation démographique

Nous avons débuté nos travaux en 1985 sur la ville de Marseille (FRANCE), qui possédait dès cette date toutes les données nécessaires à une première phase de développement méthodologique.

Sur la ville de Quito, trois types d'enquête sont réalisées :

- Enquête morphologie urbaine (Octobre-Décembre 1986)
- Enquête bâti/population (Avril 1987)
- Enquête démographie : effectifs de population et systèmes résidentiels (Octobre 1987)

L'enquête morphologie urbaine à Quito a été réalisée sur le terrain du 1er octobre au 9 décembre 1986 : cette première enquête a permis de collecter des informations précises sur l'occupation du sol, c'est à dire la morphologie urbaine, de 199 îlots de la ville de Quito, tirés sur les cartes INEC de 1962 au 1/10000, à l'aide d'une grille de 750 mètres de côté.

1. 2. L'enquête bâti/population

Par l'enquête bâti/population, il s'agit de recueillir des données démographiques et socio-économiques sur un sous-échantillon des 199 îlots ayant fait l'objet de l'enquête morphologie urbaine fin 1986 ; ainsi, nous disposerons à la fois des données sur la morphologie urbaine et sur les habitants de ce sous-échantillon d'îlots.

Nous pourrions donc tenter de répondre aux deux questions suivantes :

- Quels sont les descripteurs de la morphologie urbaine les plus pertinents pour l'observation démographique à Quito ?
- Quelles sont l'unité spatiale de collecte et la stratégie de sondage pertinentes pour une enquête démographique à Quito ?

Ce sont ces descripteurs et stratégie de sondage qui seront retenus pour l'enquête démographique générale d'octobre 1987.

L'échantillon à enquêter en avril 1987 du point de vue socio-démographique est constitué de 55 îlots, choisis de façon raisonnée parmi les 199 îlots de l'enquête morphologie urbaine ; les îlots ont été sélectionnés de manière à rassembler les différents types de bâtiments, les différents types d'îlots, et les différentes localisations au sein de la ville de Quito. La définition des types de bâtiments et types d'îlots est issue de l'exploitation des données recueillies au cours de l'enquête morphologie urbaine de fin 1986.

2. TECHNIQUE D'ENQUETE

Chaque îlot de l'échantillon de l'enquête bâti/population a déjà été enquêté en octobre-décembre 1988, lors de l'enquête morphologie urbaine ; dès le début de l'enquête bâti/population, l'enquêteur reçoit donc pour chaque îlot :

- un plan au localisant tous les bâtiments de l'îlot : sur ce plan, chaque bâtiment est identifié par un numéro de 1 à n. Ce numéro est indiqué au centre du bâtiment qu'il concerne. (Voir exemple de plan page 23)
- un jeu de fiches décrivant tous les bâtiments de l'îlot: chaque colonne de la fiche correspond à un bâtiment, dont le numéro (cohérent avec celui noté sur le plan) est indiqué en haut de la colonne. (Voir exemple de fiche page 24)

Ce plan et ces fiches permettent à l'enquêteur de repérer et identifier sans ambiguïté les bâtiments de l'îlot : l'enquêteur doit **conserver la même délimitation et la même numérotation de bâtiments** que celle apparaissant sur les documents de l'enquête morphologie urbaine (plan et fiches).

Pour collecter les données de l'enquête bâti/population, l'enquêteur reçoit deux types de fiches :

- **La fiche récapitulative** : il y a une fiche récapitulative par îlot. Sur cette fiche, l'enquêteur récapitule des informations générales relatives aux logements appartenant à l'îlot.
- **Les fiches logement** : ces fiches sont utilisées pour décrire les personnes résidant dans les logements des bâtiments faisant partie de l'îlot.

Il doit y avoir concordance complète de l'information entre ces quatre types de documents (Plan et fiches de l'enquête morphologie, fiches récapitulatives et fiches logement de l'enquête bâti/population).

3. ORGANISATION DU TRAVAIL

L'enquête débute par une **réunion d'information** le Jeudi 9 Avril, au cours de laquelle sont exposés les objectifs de l'enquête et la technique de collecte de l'information.

Le lendemain, chaque enquêteur **visitera** avec le superviseur et les chercheurs de l'Orstom **tous les îlots** qu'il aura à enquêter pendant les deux semaines, afin qu'il n'y ait pas d'ambiguïté sur la localisation et la délimitation des îlots et de leurs bâtiments ; il **recevra les documents d'enquête** relatifs à ces îlots (cartes et fiches descriptives des bâtiments et questionnaires vierges en nombre suffisant), afin de pouvoir démarrer le travail de collecte dès le Samedi 11 Avril. Enfin, on procédera avec lui à une **expérience de collecte sur quelques logements**, pour que les concepts et définitions utilisés pour l'enquête soient bien assimilés, et de façon homogène, par tous les enquêteurs.

La liste d'îlots de chaque enquêteur est établie en fonction des principes suivants :

- les îlots de chaque enquêteur sont, dans la mesure du possible, regroupés géographiquement, afin de minimiser les temps de déplacement.
- un enquêteur peut enquêter environ 400 logements en deux semaines de travail à temps complet

Chaque enquêteur organise lui-même ses horaires de travail pour réaliser la collecte des informations sur le terrain, en fonction des contraintes d'horaires (de présence à leur domicile) propres à la population à enquêter.

Tous les deux jours, chaque enquêteur aura un **rendez-vous avec le superviseur** de l'enquête. Cette rencontre permettra à l'enquêteur d'exposer les problèmes rencontrés au cours de la collecte, afin de rechercher ensemble une solution à ces problèmes de collecte de l'information ; c'est également à cette occasion que chaque enquêteur remettra au superviseur les questionnaires déjà remplis.

Entre deux rendez-vous d'un enquêteur avec le superviseur, celui-ci contrôlera l'avancée du travail, ainsi que la complétude et la cohérence des questionnaires ; en cas d'incomplétude ou d'incohérence des questionnaires, il sera demandé à l'enquêteur, lors du rendez-vous suivant, de retourner sur le terrain pour corriger ou compléter l'information.

L'enquête se terminera par un **rendez-vous final** dans les bureaux de l'IGM le Lundi 27 Avril ; tous les enquêteurs devront remettre lors de cette réunion les derniers questionnaires remplis.

4. COLLECTE DE L'INFORMATION

4. 1. Principes de base

* La première opération consiste à identifier, sur le terrain, les bâtiments dessinés sur le plan au 1/500 et décrits dans les fiches de l'enquête morphologie : il s'agit de retrouver, sur le terrain, à quels bâtiments correspondent les bâtiments identifiés lors de l'enquête morphologie urbaine, et quelles sont leurs limites. Cette opération a pour but de réutiliser les mêmes numéros de bâtiments, afin de pouvoir mettre en correspondance les résultats de l'enquête morphologie et ceux de l'enquête bâti/population.

* Démarrer l'enquête par le bâtiment N°1, et poursuivre, une fois ce premier bâtiment enquêté, par le bâtiment N°2...etc, jusqu'au dernier bâtiment de l'îlot. Cette façon de procéder, en suivant l'ordre des numéros de bâtiments facilitera les travaux de contrôle et de saisie de l'information collectée.

* Pour les bâtiments ne comptant pas de logement, aucune fiche logement n'est remplie ; l'enquêteur note uniquement l'usage du bâtiment (garage individuel, tienda, industria...etc) dans la colonne "usage principal" de la fiche récapitulative, dans la ligne relative à ce bâtiment.

* Tous les logements faisant partie des bâtiments situés dans l'îlot doivent être enquêtés. Pour chaque logement, il est rempli autant de fiches que nécessite le nombre de personnes vivant dans le logement.

* Pour remplir les fiches logement, il n'est pas nécessaire que l'enquêteur rencontre chaque personne vivant dans le logement : les questions doivent être posées soit au chef de ménage, s'il est présent, soit à sa femme ou à toute autre personne du ménage pouvant répondre aux questions se rapportant aux différents membres du ménage.

* Après avoir enquêté un logement, et donc rempli les fiches logement correspondantes, l'enquêteur reporte quelques informations sur la fiche récapitulative relative à la manzana (Voir mode de remplissage de cette fiche page 21).

4. 2. Recensement, délimitation et numérotation des logements

Définition du logement

On appelle "logement" tout local utilisé pour l'habitation : un local non destiné originellement à l'habitation, mais qui est habité au moment de l'enquête est assimilé à un logement.

Délimitation des logements

Plusieurs situations peuvent se rencontrer :

- Bâtiment à étages, divisé en appartements : chaque appartement (c'est à dire ensemble de pièces ayant une entrée commune, indépendante, qui occupe une partie d'un bâtiment) constitue un logement.

S'il existe des chambres de service avec des entrées indépendantes, directes, sur les parties communes de l'immeuble (couloirs), ces chambres sont considérées comme des logements indépendants.

Si une chambre de service (ou une pièce louée) n'a pas d'entrée indépendante de celle desservant l'appartement dont elles font partie, on considère qu'elle est incluse dans cet appartement, et ne constitue pas un logement indépendant.

- Villa ou maison : un bâtiment non divisé en appartements (unités indépendantes d'habitation) constitue un seul logement.

S'il existe des chambres de service avec des entrées indépendantes, permettant de rejoindre directement la rue sans passer par la maison, ces chambres sont considérées comme des logements indépendants.

Si une chambre de service (ou une pièce louée) n'a pas d'entrée indépendante de celle desservant la maison dont elle fait partie, on considère qu'elle est incluse dans cette maison, et ne constitue pas un logement indépendant.

Numérotation des logements

Chaque logement est repéré par trois numéros :

- Un numéro de manzana
- Un numéro de bâtiment, qui va de 1 à Π dans chaque manzana. C'est le numéro attribué au cours de l'enquête morphologie urbaine.
- Un numéro de logement. Pour chaque bâtiment de l'îlot, la numérotation des logements va de 1 à Π , Π étant le nombre de logements dans le bâtiment.

Ces trois numéros sont notés dans la partie supérieure de la fiche logement, et doivent être reportés sur chaque ligne de la fiche récapitulative.

Type de logement

Noter, en clair, le type du logement **individuel**, en utilisant les catégories suivantes:

Maison ou villa, appartement, pièce louée... etc.

S'il s'agit d'un logement **collectif**, indiquer en clair de quel type de logement collectif il s'agit, en utilisant les catégories suivantes :

Caserne, hôtel, hôpital, internat, couvent ...etc.

Condition d'occupation du logement

Quatre situations peuvent se produire :

- Logement en réparation ou en construction, donc inhabitable au moment de l'enquête
- Logement habitable, mais innocupé
- Logement habitable, occupé, dont on a pu enquêter les habitants
- Logement habitable, occupé, dont on n'a pas pu enquêter les habitants

Nombre de pièces utilisées pour dormir

Noter, dans les deux cases prévues à cet effet, le nombre de pièces généralement utilisées pour dormir, qu'elles soient ou non effectivement occupées au moment de l'enquête. Une pièce dont ce n'est pas la vocation première, mais qui est également utilisée pour dormir (exemple : un salon) doit être comptée dans le calcul du nombre total de pièces utilisées pour dormir au même titre qu'une chambre à coucher.

4. 3. Définition de la population à enquêter. Statut de résidence.

La population décrite dans les colonnes de la fiche logement comprend deux catégories, suivant le statut de résidence des personnes :

- **présents** : toutes les personnes ayant dormi dans le logement pendant au moins huit nuits, consécutives ou non, pendant le mois précédant le passage de l'enquêteur (qu'elles soient ou non présentes au moment du passage de l'enquêteur ou la nuit précédant son passage);

Si ces personnes résident la majeure partie de l'année dans le logement enquêté, on les appelle "**Résidents présents**".

Si ces personnes résident la majeure partie de l'année dans un autre logement que celui enquêté, on les appelle "**Non résidents présents**".

- et
- **Résidents absents** : toutes les personnes qui résident habituellement (c'est à dire la majeure partie de l'année) dans le logement, mais en sont absentes depuis plus d'un mois (tourisme, voyage d'affaires, travail temporaire ou saisonnier à l'extérieur de Quito), pour quelle que raison que ce soit (sauf un départ définitif du logement ou les cas particuliers cités ci-dessous) et qui doivent revenir vivre dans le logement dans un délai inférieur à 6 mois.

Cas particuliers

Ne font pas partie de la population à enquêter dans les logements individuels (car sont comptabilisés dans les logements collectifs) :

- Militaires, élèves internes,
- Emprisonnés ou hospitalisés pour une longue durée (supérieure à un mois)

Font partie de la population à enquêter :

- Emprisonnés ou hospitalisés pour une courte durée (inférieure à 1 mois)
- Travailleurs exerçant leur activité en dehors de Quito, vivant donc en dehors de la capitale une partie de l'année, mais revenant régulièrement habiter dans leur logement de Quito (chaque fin de semaine, chaque mois)
- Personnes venant d'emménager dans le logement pour y résider de façon définitive après un déménagement (transfert du mobilier...etc. dans le nouveau logement).

4.4. Renseignements relatifs aux personnes

Le plus simple est de réaliser l'interview en deux temps :

1) d'abord, demander de donner une liste des personnes habitant dans le logement, et noter leur nom et leur statut de résidence.

Demander à la personne interrogée de commencer par le chef de ménage, et poursuivre par les autres personnes vivant avec lui dans le logement, en respectant l'ordre général suivant :

Chef de ménage, épouse, enfants célibataires, enfants mariés avec les membres de leur famille, autres parents, personnes non apparentées au chef de ménage.

Numéroter ces personnes de 1 à N dans chaque logement.

2) puis, questionner pour avoir les différents renseignements sur chacune de ces personnes, en redonnant un par un les noms des personnes de cette liste.

Les renseignements relatifs à chaque personne occupent une colonne de la fiche logement.

Numéro de la fiche logement

Quatre cases sont prévues dans la partie supérieure de la fiche logement pour indiquer son numéro : dans les deux cases de gauche, l'enquêteur indique le numéro d'ordre de la fiche pour le logement considéré ; dans les deux cases de droite, il indique le nombre total de fiches que comporte le logement.

Exemples

- 1ère fiche d'un logement ne comptant qu'une seule fiche :

0	1	/	0	1
---	---	---	---	---

- 2ème fiche d'un logement comptant trois fiches :

0	2	/	0	3
---	---	---	---	---

Nom et prénom

Ces renseignements sur l'identité des personnes enquêtées ne feront pas partie du fichier exploité sur ordinateur ; il faut donc insister auprès des enquêtés sur le fait que l'exploitation de l'enquête est faite de façon **anonyme**. Ces renseignements ne sont demandés que pour faciliter la collecte de l'information, afin d'éviter que des personnes du logement soient omises ou au contraire comptées deux fois, ainsi que pour faciliter la collecte de l'information sur les liens de parenté au sein du logement.

Quel lien de parenté (ou autre relation) avez-vous avec le chef de ménage?

S'il existe un lien de parenté avec le chef de ménage, noter, en clair, le lien de parenté existant entre la personne et le chef de ménage.

Exemples :

- Frère de 1, épouse de 1, neveu de 1... etc. (Rappel : le chef de ménage est toujours noté 1)
- Nièce de 1, employée domestique : pour rendre compte du cas où une personne de la famille du chef de ménage a également une relation de "service" avec le chef de ménage.

S'il n'existe pas de lien de parenté avec aucune personne du logement, noter en clair la nature de cet autre lien.

Exemple : employée domestique, nurse, gardien.

S'il n'existe pas de lien de parenté avec le chef de ménage, mais qu'il existe un lien de parenté avec une autre personne du logement (elle même non apparentée au chef de ménage), noter en clair ce lien, de la façon suivante :

Exemple : la personne N° 12 est la fille de l'employée domestique, qui porte le N° 11.

On note donc pour la personne 12 : fille de 11.

Si l'on ne connaît pas le lien (de parenté ou autre), noter "INCONNU".

Sexe de la personne

Indiquer le sexe de la personne, en notant dans la case prévue à cet effet :

- 1, pour les hommes
- 2, pour les femmes
- 0 = inconnu

Quelle est votre date de naissance ou quel âge avez-vous?

Si la personne connaît précisément sa date de naissance, la noter dans les six cases prévues à cet effet (ne rien noter dans les deux cases situées en dessous, pour l'âge):

Exemple : 6 Août 1938

1	0	6	0	8	3	8
---	---	---	---	---	---	---

Si la personne ne connaît pas précisément sa date de naissance, noter son âge dans les deux cases prévues à cet effet (ne rien noter dans les six cases situées au dessus, pour la date de naissance). Il s'agit de l'âge atteint au dernier anniversaire, et non de l'âge qu'aura la personne à son prochain anniversaire. Ainsi, pour un enfant de moins d'un an, on note 00 :

Exemple : 23 ans

2	3
---	---

Dans le cas où l'âge de la personne est inconnu, noter, dans les deux cases prévues pour l'âge :

Exemple : date de naissance et âge inconnus

9	9
---	---

Où êtes-vous né?

Si la personne est née à Quito (stricto-sensu, non compris les villes de la vallée de Los Chillos, San Antonio, Cumbaya, Tumbaco), noter "Ici" sur la ligne inférieure. Ne rien indiquer sur les autres lignes.

Si la personne est née en Equateur en dehors de Quito, indiquer en clair, sur chacune des trois lignes prévues (ne rien indiquer sur la ligne "pays" :

- le nom de la province
- le nom de la paroisse
- le caractère rural ou urbain du lieu de naissance : urbain s'il s'agit du chef-lieu de la paroisse, rural s'il s'agit d'une autre localité que le chef-lieu de la paroisse.

Si la personne est née à l'étranger, indiquer seulement le pays de naissance.

Si le lieu de naissance est inconnu, noter "inconnu" sur la ligne supérieure, et mettre un trait pour les autres lignes.

Si le lieu de naissance n'est connu que de façon imprécise (par exemple, on connaît la province, mais pas la paroisse), indiquer l'élément de réponse qui est connu, et noter "inconnu" pour les autres éléments, non connus.

Exemples :

- Personne née à l'étranger :

Pays CHILE
 Province /
 Paroisse /
 Rural/Urbain /

- Personne née à Quito :

Pays /
 Province /
 Paroisse /
 Rural/Urbain AQUI

- Personne née à Cuenca, dans la ville même :

Pays /
 Province AZUAY
 Paroisse CUENCA
 Rural/Urbain Urbano

- Personne née dans la paroisse de Cuenca,
 en dehors de la ville de Cuenca :

Pays /
 Province AZUAY
 Paroisse CUENCA
 Rural/Urbain Rural

- Personne dont on ignore totalement le lieu de naissance :

Pays DESC
 Province /
 Paroisse /
 Rural/Urbain /

- Personne dont on ne connaît que le pays de naissance :

Pays /
 Province PICHINCHA
 Paroisse DESC
 Rural/Urbain DESC

Dépuis quelle date ou depuis combien d'années vivez-vous dans ce logement?

Avec cette question, il faut saisir la date d'arrivée ou la durée d'occupation dans le logement actuel :

- pour les **résidents présents ou absents** : il s'agit de la date à laquelle **la personne en a fait sa résidence principale**. Par résidence principale, on entend, si la personne habite dans plusieurs logements au cours de l'année, la résidence où elle habite le plus de temps au cours de l'année.
- pour les **présents** : il s'agit de la dernière date d'arrivée de ces personnes dans le logement.

Si la personne a **toujours vécu dans ce logement**, noter "TOUJOURS".

Si la personne a **habité autrefois dans ce logement**, en est partie pour résider ailleurs, et y est revenue s'installer, noter la date du début de la dernière installation.

Si la personne **connait sa date d'arrivée** dans le logement, la noter dans les 6 cases prévues à cet effet (ne rien noter dans les deux cases situées en dessous, pour la durée).

Si la personne **connait la durée écoulée depuis son arrivée**, noter la durée en années dans les deux cases prévues à cet effet (ne rien noter dans les six cases situées au dessus, pour la date d'arrivée). Il s'agit de la durée en années entières ; ainsi, si la personne occupe ce logement depuis moins d'un an, noter 00.

Dans le cas où la **date d'arrivée est totalement inconnue**, tout comme la durée, noter 99 dans les cases prévues pour noter la durée d'occupation du logement.

Exemples :

- Personne occupant le logement depuis le 5 Septembre 1983 :

1	0	5	0	9	8	3
---	---	---	---	---	---	---
- Personne occupant le logement depuis 1952 (jour et mois inconnus) :

9	9	9	3	5	2
---	---	---	---	---	---
- Personne occupant le logement depuis 38 ans :

3	8
---	---
- Personne dont on ignore depuis quelle date ou depuis combien de temps elle occupe le logement :

9	9
---	---

Depuis quelle date ou depuis combien d'années vivez-vous à Quito?

Le mode de remplissage de la fiche pour cette question est similaire à celui concernant la date d'arrivée dans le logement.

Avec cette question, il faut saisir la date d'arrivée ou la durée de séjour à Quito (stricto-sensu, non compris les villes de la vallée de Los Chillos, San Antonio, Cumbaya, Tumbaco) :

- pour les résidents présents ou absents et pour les présents ayant leur résidence principale à Quito, mais dans un autre logement : il s'agit de la date à laquelle la personne a eu sa résidence principale à Quito, qu'il s'agisse de ce logement ou non. Par résidence principale, on entend, si la personne habite alternativement à Quito et en dehors de Quito au cours de l'année, la résidence où elle habite le plus de temps au cours de l'année.

- pour les présents n'ayant pas leur résidence principale à Quito : il s'agit de la dernière date d'arrivée de ces personnes à Quito.

Si la personne a toujours vécu à Quito, noter "TOUJOURS".

Si la personne a habité autrefois à Quito, en est partie pour résider ailleurs, et y est revenue s'installer, noter la date du début de la dernière installation.

Si la personne connaît sa date d'arrivée à Quito, la noter dans les 6 cases prévues à cet effet (ne rien noter dans les deux cases situées en dessous, pour la durée).

Si la personne connaît la durée écoulée depuis son arrivée, noter la durée en années dans les deux cases prévues à cet effet (ne rien noter dans les deux cases situées au dessus, pour la date d'arrivée). Il s'agit de la durée en années entières ; ainsi, si la personne occupe ce logement depuis moins d'un an, noter 00.

Dans le cas où la date d'arrivée est totalement inconnue, tout comme la durée, noter 99 dans les cases prévues pour noter la durée d'occupation du logement.

Où habitiez-vous avant de résider à Quito?

Pour les résidents et les présents ayant leur lieu de résidence principal actuel à Quito, il s'agit du lieu de résidence principal de la personne juste avant qu'elle n'ait son lieu de résidence principal à Quito.

Pour les présents ayant leur lieu de résidence principal actuel en dehors de Quito, il s'agit du lieu de résidence principal actuel.

Si la personne a toujours eu sa résidence principale à Quito, noter "Ici" sur la ligne inférieure. Ne rien indiquer sur les autres lignes.

Si la personne résidait auparavant en Equateur en dehors de Quito, indiquer en clair, sur chacune des trois lignes prévues (ne rien indiquer sur la ligne "pays") :

- le nom de la province
- le nom de la paroisse
- le caractère rural ou urbain du lieu de naissance : urbain s'il s'agit du chef-lieu de la paroisse, rural s'il s'agit d'une autre localité que le chef-lieu de la paroisse.

Si la personne résidait auparavant à l'étranger, indiquer seulement le pays de naissance.

Si le lieu de résidence antérieur est inconnu, noter "inconnu" sur la ligne supérieure, et mettre un trait pour les autres lignes.

Si le lieu de résidence antérieur n'est connu que de façon imprécise (par exemple, on connaît la province, mais pas la paroisse), indiquer l'élément de réponse qui est connu, et noter "inconnu" pour les autres éléments, non connus.

Exemples :

Voir question sur le lieu de naissance

De quel type de propriétaire s'agit-il?

Indiquer, en clair, dans la colonne du chef de ménage, si le propriétaire est :

- Une personne physique (naturelle)
- Une société privée
- Un organisme semi-public
- Un organisme public
- Un organisme municipal
- Un organisme religieux
- Un organisme diplomatique
- Un organisme de bienfaisance
- Une organisation syndicale
- Une coopérative ...etc.

Si la réponse est inconnue, noter "INCONNU".

Quel lien de parenté ou autre relation avez-vous avec le propriétaire?

Cette question ne concerne que les chefs de ménage, et seulement dans les cas où le propriétaire est une personne physique et ne réside pas dans le logement (que ce soit le chef de ménage ou une autre personne qui soit propriétaire).

Si il existe un lien de parenté entre le chef de ménage et le propriétaire noter, en clair, ce lien de parenté, dans la colonne Chef de ménage :

Exemples :

- Frère de P, petit-fils de P, neveu de P...etc. (Le propriétaire est toujours noté P)

Si le chef de ménage n'a pas de lien de parenté avec le propriétaire du logement, noter sans lien familial (abrégé SLF)

Si le chef de ménage n'a pas de lien de parenté avec le propriétaire, mais qu'il existe un lien de parenté entre une autre personne du logement et le propriétaire (elle même non apparentée au chef de ménage), noter en clair ce lien, dans la colonne relative à cette personne, de la façon suivante :

Exemple : la personne N° 12 est la fille du propriétaire.

On note donc pour la personne 12 : fille de P.

Activité

Cette question concerne chaque personne vivant dans le logement, sans condition d'âge minimum.

Il s'agit de la situation au mois d'Avril 1987.

Profession

Plusieurs cas peuvent se produire.

Si la personne ne travaille pas, pour décrire les différentes situations, utiliser les termes suivants :

- | | |
|---|---|
| - Etudiant | - Travaux ménagers |
| - Retraité | |
| - Ne travaille pas, recherche un emploi | - Ne travaille pas, ne recherche pas d'emploi |

Si la personne travaille, indiquer sa profession en détail :

Exemples

- | | |
|-------------------------|--------------------------------|
| - Professeur de musique | - Professeur de l'Université |
| - Menuisier | - Vendeur d'assurances |
| - Chauffeur de camions | - Mécanicien automobile...etc. |

Une personne peut avoir plusieurs professions pendant l'année. Noter toutes ses professions, et souligner la profession principale (c'est à dire celle qui lui rapporte le plus de revenus de l'année).

Donner un numéro à chacune de ces professions. Ces numéros seront utilisés pour indiquer la branche d'activité et le statut professionnel pour chaque profession de la personne.

Branche d'activité

Pour obtenir une réponse correcte, demander :

- "Quel est l'activité principale de l'établissement où vous travaillez?"

L'enquêteur doit indiquer exactement le type d'activité de l'établissement. Ne pas accepter une réponse vague, et insister pour avoir plus de détails.

Si une personne a plusieurs professions, utiliser les numéros des professions, pour décrire la branche d'activité de chaque profession.

Exemples

- | | |
|----------------------------|---------------------------|
| - Education cycle primaire | - Consultations médicales |
|----------------------------|---------------------------|

- Usine de textile
- Boulangerie
- Menuiserie
- Ministère de l'Agriculture...etc.

Statut professionnel

Indiquer la situation dans la profession (définitions du recensement de l'INEC) :

- **Patron** : personne qui dirige sa propre entreprise, ou qui exerce une profession pour son propre compte, et qui a un ou plusieurs employés rémunérés (ex : directeur d'une usine).
- **Travailleur indépendant** : personne qui dirige sa propre entreprise, ou qui exerce pour son propre compte un travail, mais qui n'a aucun employé rémunéré (ex : vendeur ambulante).
- **Employé ou salarié** : employé qui travaille pour l'Etat ou dans une entreprise privée, et qui reçoit une rémunération, une solde, un salaire, une commission ou une somme forfaitaire.
- **Travailleur familial sans rémunération** : personne exerçant une activité non rémunérée dans une entreprise ou un commerce d'un membre de sa famille, la majeure partie de la semaine.
- **Apprenti** : personne travaillant dans une entreprise pour apprendre un métier, ne recevant pas de rémunération, ou une rémunération minimale.
- **Autres** (préciser)

Si une personne a plusieurs emplois, utiliser les numéros de ses professions, pour décrire le statut professionnel dans chacun de ces emplois.

Dans la mesure du possible, noter le montant des revenus annuels ou mensuels (préciser la période de référence), procurés par chaque emploi.

Autres sources de revenus

Noter si la personne a une autre source de revenus que les emplois décrits plus haut : indiquer la source de revenus et son montant annuel approximatif.

Exemples : mise en location d'un appartement (250.000 sucres/mois), propriétaire d'un taxi (X sucres/mois)...etc

5. FICHE RECAPITULATIVE

L'enquêteur remplit une fiche récapitulative par manzana (ou plusieurs fiches récapitulatives si le nombre de logements dans la manzana est supérieur à 21). Sur ces fiches, sont inscrites des informations globales concernant la manzana.

Chaque ligne de la fiche correspond à un logement.

Nombre de bâtiments, de fiches récapitulatives, de fiches logement

Dans ces trois cas, il s'agit du nombre total dans la manzana.

Ordre des logements dans la fiche

Commencer par le bâtiment n°1, logement n°1.

Puis continuer en prenant les bâtiments par ordre de numéro, et, dans chaque bâtiment, les logements par ordre de numéro : (Bâtiment 1, Logement 1), puis (Bâtiment 1, Logement 2), puis (Bâtiment 2, Logement 2).....etc.

Nombre de personnes dans le logement

Il s'agit du nombre de total de personnes du logement, réparties selon leur statut de résidence :

- Nombre de Résidents Présents
- Nombre de Résidents Absents
- Nombre de Non Résidents Présents

Usage du bâtiment

Noter en clair quel est l'usage principal du bâtiment en utilisant de préférence les catégories suivantes :

- | | |
|---|-----------------------|
| - Habitation | - Service domestique |
| - Commerce (préciser le type de commerce) | - Bureaux privés |
| - Entrepôt | - Education |
| - Industrie | - Culture, spectacles |
| - Artisanat (préciser garage, menuiserie...etc) | - Sport |
| - Bureaux publics | - Santé |

Si un bâtiment a plusieurs usages, noter ces différents usages, en soulignant l'usage principal.

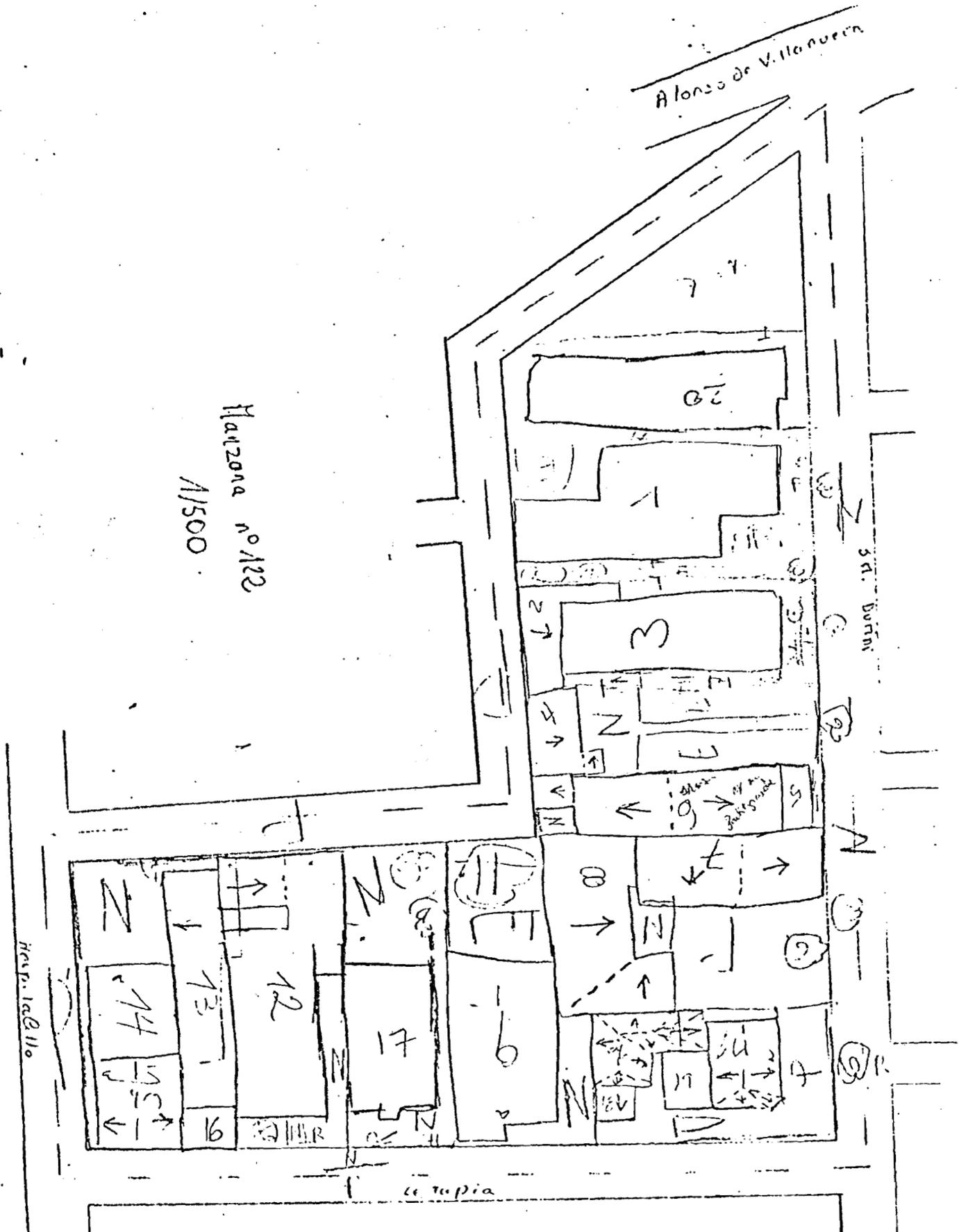
Si un bâtiment ne sert qu'à l'habitation, indiquer "Seul. Habitation".

Observations

Indiquer dans cette case :

- les problèmes éventuels rencontrés au cours de l'enquête (refus de réponse, absence prolongée des habitants...etc),
- des remarques sur l'état du bâtiment, si cet état explique la non-utilisation du bâtiment (donc, pas d'habitants, ou d'usage du bâtiment),
- toute autre observation pour décrire une situation non prévue dans les questionnaires.

6. EXEMPLE DE MAZANA ET DE LOGEMENT



N° Manzano Nom Enquêteur N° Bâtiment N° Logement N° de la fiche logement /

Type de logement..... Condition d'occup. du log. :
 Inhabitable
 Habit. innoc.
 Habit. oc. non enq.
 Habit oc. enq.

Nombre de pièces util. pour dormir

N°	RUBRIQUES	N° personne <input type="text"/> <input type="text"/>					
11	Nom et prénom						
10	Statut de résidence	Résid. présent <input type="checkbox"/> Résid. absent <input type="checkbox"/> Non res. prés. <input type="checkbox"/> <input type="checkbox"/>	Résid. présent <input type="checkbox"/> Résid. absent <input type="checkbox"/> Non res. prés. <input type="checkbox"/> <input type="checkbox"/>	Résid. présent <input type="checkbox"/> Résid. absent <input type="checkbox"/> Non res. prés. <input type="checkbox"/> <input type="checkbox"/>	Résid. présent <input type="checkbox"/> Résid. absent <input type="checkbox"/> Non res. prés. <input type="checkbox"/> <input type="checkbox"/>	Résid. présent <input type="checkbox"/> Résid. absent <input type="checkbox"/> Non res. prés. <input type="checkbox"/> <input type="checkbox"/>	Résid. présent <input type="checkbox"/> Résid. absent <input type="checkbox"/> Non res. prés. <input type="checkbox"/> <input type="checkbox"/>
11	Lien de parenté avec le chef de ménage
11	Sexe	<i>Homme =1</i> <i>Femme =2</i> <input type="checkbox"/>					
12	Date de naissance ou age	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
12	Lieu de naissance : Pays Province Paroisse Rural/Urban
14	Date d'entrée dans le logement ou durée d'occupation du logement	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
15	Date d'arrivée à Quito ou durée de séjour à Quito	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
13	Lieu de résidence antérieur : Pays Province Paroisse Rural/Urban

| N° pa. | RUBRIQUES | N° personne <input type="text"/> |
|--------|---|--|--|--|--|--|--|
| 17 | Statut d'occupation du logement
<i>(préciser, selon le cas, le montant du loyer, ou de l'antichresis)</i> |
.....
<input type="text"/> |
| 17 | Si propriétaire, coût construction
achat, héritage ou donation |
.....
<input type="text"/> |
| 18 | Type de propriétaire |
.....
<input type="text"/> |
| 18 | Lien de parenté (ou autre) avec
le propriétaire du logement |
.....
<input type="text"/> |
| 19 | Profession

Activité : Branche d'activité

Statut professionnel |
.....
.....
.....
<input type="text"/> |
| 20 | Autres sources de revenu
(agricole, foncier...etc)
<i>(indiquer la source et le montant annuel approximatif de chaque source de revenu)</i> |
.....
.....
<input type="text"/> |

m - IPGH

QUITO-Morfología -Oct./Dec 1986

Número de manzana: 122		DESCRIPCION DE LOS EDIFICIOS				
Hora:		1	2	3	4	5
Número del edificio que corresponde a la señal localizada en el plano: /		8	2	9	2	8
<u>Edificio:</u>	Acabado:	<input checked="" type="checkbox"/>				
	En construccion	<input type="checkbox"/>				
<u>Número de pisos (incluido P.B.)</u>		2	1	2	1	2
<u>Dimensiones en el suelo:</u>		15	15	10	15	12
	Lado a	20	7	14	7	20
	Lado b					
(Indicar este(os) lado(s) en el plano)						
<u>Número de pendientes del techo (0 si plano)</u>		0	1	0	1	0
<u>Material predominante en el techo</u>						
	Losa de hormigón	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
	Eternit o similares	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Zinc o similares	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Teja con barniz	<input type="checkbox"/>				
	Teja sin barniz	<input type="checkbox"/>				
	Paja o similares	<input type="checkbox"/>				
	Otros (chova, embaldosado, techaline...)					
<u>Color predominante en el techo</u>		gris	gris	gris	gris	gris
(Blanco, Café obscuro, Café claro, Gris, color ladrillo, Negro, Rojo, o Verde)						
<u>Forma del material del techo</u>		—	~	—	~	—
(Hacer un croquis de la sección)						
<u>Otros elementos sobre el techo</u>						
	Número de chimeneas					
	Presencia de pilares de hormigón	<input type="checkbox"/>				
	Presencia de penthouse (a describir)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
	Número de tanques de reserva			2		
<u>Presencia de vegetación:</u>						
	Sobre balcones sobresalientes					
	Sobre el techo					
<u>Otras observaciones sobre el edificio</u>						

B) CODIFICATION DE L'ENQUETE BATI / POPULATION

NOMBRE DEL ENCUESTADOR

1	Ramiro VASQUEZ	9	DESCONOCIDO
2	Marcelo AREVALO		
3	Henri SANCHEZ ALBA		
4	Cesar ALVAREZ MONTUFAR		
5	Katia VALENCIA		
6	Jorge ROJAS ALMEIDA		
7	O. BARBARY et F. DUREAU		

ESTATUTO DE RESIDENCIA DE LOS HABITANTES

1	Residente presente	9	DESCONOCIDO
2	Residente ausente		
3	No residente presente		

SEXO

1	Hombre	9	DESCONOCIDO
2	Mujer		

LAZO DE PARENTESCO (O OTRA RELACION) CON EL JEFE DEL HOGARPERSONAS QUE TIENEN SOLO UN LAZO DE PARENTESCO CON EL JEFE DEL HOGAR

- | | |
|----------------------------------|------------------------------------|
| 01 Jefe del hogar | 07 Primo(a) del jefe del hogar |
| 02 Conyuge del jefe del hogar | 08 Cunado(a) del jefe del hogar |
| 03 Hijo(a) del jefe del hogar | 09 Suegros del jefe del hogar |
| 04 Hermano(a) del jefe del hogar | 10 Yerno(nuera) del jefe del hogar |
| 05 Padres del jefe del hogar | 11 Nieto(a) del jefe del hogar |
| 06 Tio(a) del jefe del hogar | 12 Sobrino(a) del jefe del hogar |
| | 13 Otros parientes |

PERSONAS QUE TIENEN UN LAZO DE PARENTESCO Y UNA RELACION DE SERVICIO CON EL JEFE DEL HOGAR

- | | |
|----------------------------------|------------------------------------|
| 22 Conyuge del jefe del hogar | 27 Primo(a) del jefe del hogar |
| 23 Hijo(a) del jefe del hogar | 28 Cunado(a) del jefe del hogar |
| 24 Hermano(a) del jefe del hogar | 29 Suegros del jefe del hogar |
| 25 Padres del jefe del hogar | 30 Yerno(nuera) del jefe del hogar |
| 26 Tio(a) del jefe del hogar | 31 Nieto(a) del jefe del hogar |
| | 32 Sobrino(a) del jefe del hogar |
| | 33 Otros parientes |

PERSONAS QUE NO TIENEN UN LAZO DE PARENTESCO CON EL JEFE DEL HOGAR, PERO TIENEN UN LAZO DE PARENTESCO CON OTRA PERSONA DE LA VIVIENDA (ella misma no emparentada con el jefe del hogar)

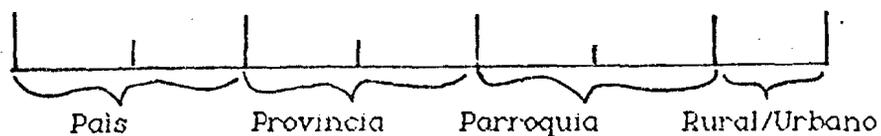
- | | |
|---------------|----------------------------------|
| 42 Conyuge | 47 Primo(a) |
| 43 Hijo(a) | 48 Cunado(a) |
| 44 Hermano(a) | 49 Suegros |
| 45 Padres | 50 Yerno(nuera) |
| 46 Tio(a) | 51 Nieto(a) |
| | 52 Sobrino(a) del jefe del hogar |
| | 53 Otros parientes |

PERSONAS QUE NO TIENEN LAZO DE PARENTESCO CON NINGUNA PERSONA DE LA VIVIENDA

- | | |
|--|---------------------------------|
| 60 Amigo(a) | 64 Empleado(a) domestico |
| 61 Colega de trabajo | 65 Guardián |
| 62 Nurse | 66 Otras relaciones de servicio |
| 63 Jardinero | 67 Otro tipo de relacion |
| 69 Tipo de relación desconocido (salvo lazo de parentesco) | |

CODIGO GEOGRAFICO

El lugar de nacimiento y el lugar de residencia anterior estan descritos de la manera siguiente :

**Caso de un país extranjero**

Anotar el código del país.

Provincia = 00

Parroquia = 00

Rural/urbano = 0

Caso de un lugar totalmente desconocido

Provincia = 99

Parroquia = 99

Rural/urbano = 9

Caso de un lugar parcialmente desconocido

Utilizar los códigos para las informaciones conocidas.

Para las informaciones desconocidas, anotar 99 o 9.

Rural/urbano

- 1 Rural (si se trata de una localidad que no sea la cabecera parroquial)
- 2 Urbano (si se trata de la cabecera parroquial)
- 3 DESCONOCIDO

CODIGOS DE PAISESAMERICA

- 01 Ecuador
- 11 Argentina
- 12 Bolivia
- 13 Brasil
- 14 Colombia
- 15 Costa Rica
- 16 Cuba
- 17 Chile
- 18 Guatemala
- 19 Haiti
- 20 Honduras
- 21 Jamaica
- 22 Malvinas
- 23 Mèxico
- 24 Nicaragua
- 25 Panamá
- 26 Paraguay
- 27 Perú
- 28 Puerto Rico
- 29 Republica Dominicana
- 30 El Salvador
- 31 Trinidad Tobago
- 32 Uruguay
- 33 Venezuela
- 34 Curazao
- 35 Otros paises de America
del Sur o Central
- 40 Estados Unidos
- 41 Canada
- 42 Otros paises de America del Norte

EUROPA

- 51 Alemania
- 52 Espana
- 53 Francia
- 54 Gran Bretana
- 55 Union Soviètica
- 56 Otros paises de Europa

ASIA

- 61 Corea
- 62 China
- 63 India
- 64 Japon
- 65 Otros paises de Asia

70 AFRICA

80 OCEANIA

99 DESCONOCIDO

CODIGOS DE PROVIDCIAS DEL ECUADOR

- 01 Azuay
- 02 Bolivar
- 03 Canar
- 04 Carchi
- 05 Chimborazo
- 06 Cotopaxi
- 07 El oro
- 08 Esmeraldas
- 09 Galapagos
- 10 Guayas
- 11 Imbabura
- 12 Loja
- 13 Los rios
- 14 Manabi
- 15 Morona Santiago
- 16 Napo
- 17 Pastaza
- 18 Pichincha
- 19 Tungurahua
- 20 Zamora Chinchipe

- 99 DESCONOCIDO (en Ecuador)

- 00 Si se trata de un pais extranjero

CODIGOS DE PARROQUIAS

Codigo INEC de las parroquias (Censo de 1982)

99 DESCONOCIDO

00 Si se trata de un país extranjero

CALIDAD DE OCUPACION DE LA VIVIENDA

(solo en la columna del jefe del hogar)

- 01 Propietario (jefe del hogar)
- 02 Propietario (otra persona que vive en la vivienda)
- 03 Inquilino
- 04 Ocupación de favor (préstamo de un amigo, pariente, colega...etc)
- 05 Por servicios (portero, guardián, jardinero...etc del edificio)
- 06 Anticresis
- 07 Ocupación ilegal de un local teóricamente abandonado
- 08 Otro tipo de ocupacion

99 DESCONOCIDO

MONTO DEL ALQUILER MENSUAL O DEL ANTICRESIS

(solo en la columna del jefe del hogar)

En el caso 03 (Inquilino), indicar el monto del alquiler mensual de la manera siguiente (la unidad es de 1000 sucres):

monto < 1000 sucres	000
1000 - 1999 sucres	001
2000 - 2999 sucres	002
10.000 - 10999 sucres	010 ...etc
DESCONOCIDO	999

En el caso 06 (anticresis), indicar el monto mensual del anticresis de la misma manera, y el monto del anticresis mismo en tres casilleros situados encima.

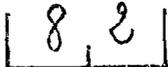
Si se trata de otra calidad de ocupacion de la vivienda (por ejemplo, propietario o por servicios), anotar 998

COSTO DE LA CONSTRUCCION O DEL COMPRA DE LA VIVIENDA


 Año de la construcción
 o del compra


 Costo de la construcción
 o del compra

Año de la construcción o del compra (o de la herencia o donación)

Ejemplo : año 1982 

Si la fecha esta desconocida : 99

Costo de la construcción o de la vivienda

Indicar el costo de la manera siguiente (la unidad es de 100.000 sucres) :

costo < 100.000 sucres	000
100.000 - 199.999 sucres	001
200.000 - 299.999 sucres	002
1.000.000 - 1.099.999 sucres	010
1.100.000 - 1.199.999 sucres	011etc
Donación	996
Herencia	997
Proprietario no vive en la vivienda	998 (Y fecha de compra = 98)
DESCONOCIDO	999

CLASE DE PROPIETARIO

- | | |
|--|------------------------------|
| 00 Persona que vive en la vivienda (jefe del hogar o otra persona) | |
| 01 Persona física (natural), que no vive en la vivienda | |
| 02 Sociedad privada | |
| 03 Organismo semi-público | 08 Organización sindical |
| 04 Organismo público | 09 Cooperativa |
| 05 Organismo municipal | 10 Organismo religioso |
| 06 Organismo diplomático | 11 Otra clase de propietario |
| 07 Organización de beneficencia | 99 DESCONOCIDO |

LAZO DE PARENTESCO CON EL PROPIETARIO

Esta pregunta concierne solamente a los jefes del hogar. Para los otras personas, utilizar el código 98 (salvo el caso si la persona no tiene parentesco con el jefe del hogar, pero tiene una relación de parentesco con el propietario).

- 00 El jefe del hogar no tiene lazo de parentesco con el propietario (que está una persona física)
- 01 El propietario es el jefe del hogar
- 02 El propietario no es una persona física

EL JEFE DEL HOGAR TIENE UN LAZO DE PARENTESCO CON EL PROPIETARIO

Indicar, en los casilleros del jefe del hogar :

- | | |
|-------------------------------|---------------------------------|
| 03 Primo(a) del propietario | |
| 04 Conyuge del propietario | 09 Cunado(a) del propietario |
| 05 Hijo(a) del propietario | 10 Suegros del propietario |
| 06 Hermano(a) del propietario | 11 Yerno(nuera) del propietario |
| 07 Padres del propietario | 12 Nieto(a) del propietario |
| 08 Tio(a) del propietario | 13 Sobrino(a) del propietario |
| | 14 Otros parientes |

UN OTRA PERSONA DE LA VIVIENDA TIENE UN LAZO DE PARENTESCO CON EL PROPIETARIO

Indicar en los casilleros del jefe del hogar : 00

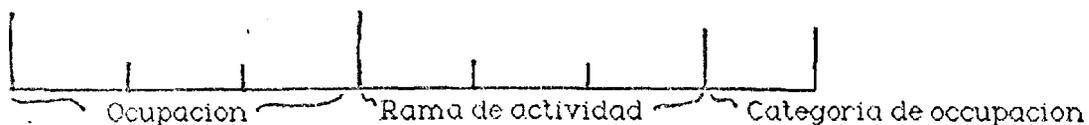
Indicar en los casilleros de la persona que vive en la vivienda y que tiene un lazo de parentesco con el propietario :

- | | |
|-------------------------------|---------------------------------|
| 15 Primo(a) del propietario | |
| 16 Conyuge del propietario | 21 Cunado(a) del propietario |
| 17 Hijo(a) del propietario | 22 Suegros del propietario |
| 18 Hermano(a) del propietario | 23 Yerno(nuera) del propietario |
| 19 Padres del propietario | 24 Nieto(a) del propietario |
| 20 Tio(a) del propietario | 25 Sobrino(a) del propietario |
| | 26 Otros parientes |

99 DESCONOCIDO

ACTIVIDAD

La actividad está descrita de la manera siguiente :



Si una persona tiene **varios ocupaciones**, anotar los códigos de la primera ocupación en los casilleros. Anotar los códigos de la segunda ocupación por encima de estos casilleros.

Si una persona **no trabaja**, indicar el código de su tipo de ocupación. Para la rama de actividad y la categoría de ocupación, escribir 000 y 0.

OCUPACION

Si la persona **trabaja**, utilizar los códigos del INEC (Censo 1982).

Si la persona **no trabaja**, utilizar los códigos siguientes :

001 Estudiante

002 Jubilado

003 Quehaceres domésticos

004 Pensionista

005 No trabaja, busca un trabajo

006 No trabaja, no busca un trabajo (los pequeños niños están en este caso)

Si la ocupación está **DESCONOCIDA** : 999

RAMA DE ACTIVIDAD

Para describir la rama de actividad, anotar dos informaciones :

**Sector de actividad**

- 1 Sector publico
- 2 Sector semi-publico
- 3 Municipio
- 4 Sector privado
- 8 Persona que no trabaja
- 9 Sector DESCONOCIDO

Rama de actividad

Utilizar los codigos siguientes :

Rama 0 : Agricultura, silvicultura, caza y pesca

- 01 Agricultura
- 02 Ganaderia
- 03 Silvicultura y extracción de madera
- 04 Caza y pesca
- 09 Otras actividades de la rama 0

Rama 1 : Explotación de mines y canteras (salvo para la construcción)

- 11 Explotación de mines de carbón
- 12 Producción de petróleo crudo y gas natural
- 13 Extracción de minerales metalicos
- 14 Extracción de minerales para la construcción (piedra, arcilla y arena)
- 15 Extracción de otros minerales
- 19 Otras actividades de la rama 1

Rama 2 : Industrias manufactureras

- 21 Productos alimenticios, bebidas y tabaco
- 22 Textiles, prendas de vestir e industrias de cuero
- 23 Industrias de la madera y productos de la madera, incluidos muebles
- 24 Fabricación de papel y productos de papel : imprentas y editoriales
- 25 Fabricación de sustancias quimicas y de productos quimicos, derivados del petróleo y del carbón de caucho y plásticos
- 26 Fabricación de productos minerales no metálicos, exceptuando los derivados del petróleo y del carbón (salvo para la construcción)

- 27 Industrias metálicas básicas (de hierro, acero, metales no ferrosos)
- 28 Fabricación de productos metálicos, maquinaria y equipo, material de transporte
- 29 Otras industrias manufactureras (joyas, instrumentos de música, artículos de deporte ...etc)

Rama 3 : Electricidad, gas y agua

- 31 Generación, transmisión y distribución de electricidad
- 32 Producción y distribución de gas
- 33 Suministro de vapor y agua caliente
- 34 Obras hidráulicas y suministro de agua (captación, purificación y distribución)
- 39 Otras actividades de la rama 3

Rama 4 : Construcción y materiales de construcción

- 41 Fabricación de materiales de construcción a base de arcilla
- 42 Fabricación de materiales de construcción a base de cemento (cemento, bloques...etc)
- 43 Fabricación de otros materiales de construcción
- 44 Obras públicas
- 45 Paredes maestras
- 46 Tubería, electricidad
- 47 Pintura, vidrios, embaldosado, decoración
- 48 Carpintería de obra
- 49 Otras actividades de la rama 4

Rama 5 : Comercio por mayor y por menor

- 51 Comercios ambulantes
- 52 Comercios generales fijos
- 53 Comercios especializados fijos (librería, ferretería, papelería, farmacia, venta de muebles...etc)
- 54 Comercios de carburante, combustible y materiales de construcción
- 55 Tienda
- 56 Comercios de alimentación
- 57 Restaurantes, bar y hoteles
- 58 Import-export
- 59 Otro comercio

Rama 6 : Transportes, almacenamiento y comunicaciones

- 60 Taxímetros
- 61 Transporte urbano colectivo de pasajeros

- 62 Transporte terrestre interurbano de pasajeros por carretera
- 63 Transporte terrestre de carga por carretera
- 64 Transporte ferroviario
- 65 Transporte aéreo
- 66 Transporte por agua (marítimo, por ríos, ...etc)
- 67 Depósito y almacenamiento (como servicio independiente)
- 68 Comunicaciones (correo, telégrafo, teléfono, radio...etc)
- 69 Otros servicios relacionados con el transporte o las comunicaciones

Rama 7 : Establecimientos financieros, seguros, bienes inmuebles y servicios prestados a las empresas

- 71 Establecimientos financieros (bancos, establecimientos de cambio...etc)
- 72 Seguros
- 73 Arrendamiento y explotación de bienes inmuebles
- 74 Servicios prestados a las empresas (jurídicos, contabilidad, publicidad, asesoría técnica...etc)
- 79 Otros establecimientos financieros y otros servicios prestados a las empresas

Rama 8 : Servicios comunales, sociales y personales

- 80 Defensa
- 81 Administración pública
- 82 Instrucción pública (escuelas primarias y secundarias)
- 83 Instrucción pública (universidades, escuelas profesionales) y investigaciones
- 84 Servicios médicos, odontológicos y otros servicios de sanidad
- 85 Servicios de diversión y esparcimiento y servicios culturales (cine, teatro, biblioteca, museo, jardín botánico o zoológico, sala de baile, estadio...etc)
- 86 Servicios de reparación (de cuero, eléctricas, de automóviles, de relojes...etc)
- 87 Servicios domésticos para el mantenimiento de los hogares (cocineras, nineras, jardineros...etc)
- 88 Servicios personales directos (peluquería, fotografía, lavandería ...etc)
- 89 Otros servicios comunales, sociales y personales

Rama 9 : Actividades no bien especificadas

- 91 Persona que no trabaja (ejemplo : estudiante, quehaceres domésticos) y no tiene otras fuentes de ingresos (alquiler, pensión...etc)
- 92 Persona que no trabaja y vive de otras fuentes de ingresos (alquiler, pensión...etc) *Para los jubilados, no utilizar este código 92 : en este caso, indicar la rama de actividad de su ex-trabajo.*
- 99 DESCONOCIDO (rama de actividad desconocida para una persona que trabaja o que está jubilada)

CATEGORIA DE OCUPACION

- 1 **Patrono** (o socio activo) : persona que dirige su propia empresa o negocio, o que ejerce una profesión u oficio y que tiene uno o más empleados remunerados (ej: el dueño de una fábrica).
- 2 **Trabajador por cuenta propia** : si la persona explota su empresa o negocio, o ejerce por cuenta propia una profesión u oficio, pero no tiene ningún empleado remunerado (ej: vendedor ambulante).
- 3 **Empleado o asalariado del sector publico** : cuando la persona es un funcionario, empleado u obrero del Estado, que recibe una remuneración, sueldo, salario o pago a destajo, bajo nombramiento, contrato o jornal.
- 4 **Empleado o asalariado del sector privado** : cuando la persona es un empleado u obrero de una empresa particular, que recibe una remuneración, sueldo, salario o pago a destajo, bajo nombramiento, contrato o jornal.
- 5 **Empleado o asalariado del sector semi-publico** : cuando la persona es un funcionario, empleado u obrero de una empresa del sector semi-publico, que recibe una remuneración, sueldo, salario o pago a destajo, bajo nombramiento, contrato o jornal.
- 6 **Trabajador familiar sin remuneración** : la persona realiza una ocupación no remunerada en una empresa o negocio de un familiar, en la mayor parte de la semana.
- 7 **Aprendiz** : se trata de una persona que trabaja sin remuneración (o minima) en una empresa para aprender una practica profesional.
- 8 **Otros**
- 9 **DESCONOCIDO**
- 0 **Si la persona no trabaja**

FUENTES Y MONTO DE LOS INGRESOS**Fuentes de ingresos**

- 01 Sin fuente de ingresos
- 02 Sólo actividad
- 11 Actividad y fuente de ingresos ligada a los bienes inmuebles (ej. : alquiler)
- 12 Actividad y fuente de ingresos ligada a un comercio (ej. : propietario de una tienda)
- 13 Actividad y fuente de ingresos ligada a los transportes (ej. : propietario de un taxi)
- 14 Actividad y pensión
- 15 Actividad y jubilación
- 16 Actividad y otro tipo de fuente de ingresos
- 21 Sólo fuente de ingresos ligada a los bienes inmuebles (ej. : alquiler)
- 22 Sólo fuente de ingresos ligada a un comercio (ej. : propietario de una tienda)
- 23 Sólo fuente de ingresos ligada a los transportes (ej. : propietario de un taxi)
- 24 Sólo pensión
- 25 Sólo jubilación
- 26 Sólo pensión y jubilación
- 27 Sólo otro tipo de fuente de ingresos
- 28 Jubilación y fuente de ingresos ligada a los bienes inmuebles

- 97 Otras fuentes de ingresos conocidas, actividad desconocida
- 98 Actividad conocida, otras fuentes de ingresos desconocidas
- 99 Actividad y otras fuentes de ingresos desconocidos

Monto de los ingresos

Se trata del monto total de los ingresos por mes : ingresos ligados a la(s) actividad(des), y ingresos de otras fuentes (ejemplo : alquiler de un departamento, propietario de un taxímetro...etc)

Indicar el monto de la manera siguiente (la unidad es de 1000 sucres por mes) :

monto < 1000 sucres	000
1.000 - 1.999 sucres	001
2.000 - 2.999 sucres	002

10.000 - 10.999 sucres	010	
11.000 - 11.999 sucres	011etc

997 Monto de las otras fuentes de ingresos conocido, ingresos ligados a la actividad desconocidos

998 Monto de los ingresos ligados a la actividad conocido, otras fuentes de ingresos desconocidas

999 DESCONOCIDO (ingresos ligados a la actividad y a las otras fuentes de ingresos desconocidos)

DATOS RECAPITULATIVOS SOBRE LAS VIVIENDAS

Véase el ejemplo de ficha de codificación en la página 19.

Numeración de los edificios

Nuevo número : se trata del número en la encuesta habitat/poblacion

Viejo número : se trata del número de la encuesta Morfología urbana

Nuevo edificio : número en la encuesta Morfología = 98

99 DESCONOCIDO

Numeración de las viviendas

Número anotado en las fichas de encuesta.

96 Edificio no encuestado. Número de viviendas en el edificio desconocido.

Tipo de vivienda

- 1 Casa o villa
- 2 Departamento
- 3 Cuarto
- 4 Vivienda colectiva
- 6 Edificio que no existe en el momento de la encuesta habitat/poblacion (pero que existaba en el momento de la encuesta morfología urbana)
- 8 Si no se trata de una vivienda
- 9 DESCONOCIDO

Condición de ocupación de la vivienda

- 1 Si no se trata de una vivienda
- 2 Vivienda inhabitable en el momento de la encuesta
- 3 Vivienda habitable, ocupada en el momento de la encuesta
- 4 Vivienda habitable, desocupada en el momento de la encuesta
- 5 Vivienda en construcción
- 6 Edificio que no existe en el momento de la encuesta habitat/poblacion (pero que existaba en el momento de la encuesta morfología urbana)
- 9 DESCONOCIDO

Numero de piezas utilizadas para dormir

Se trata del número de piezas anotado en la ficha de la encuesta.

99 DESCONOCIDO

Numero de personas en la vivienda

Se trata del número total de personas en la vivienda (residentes presentes, residentes ausentes, y no residentes presentes)

99 DESCONOCIDO

Uso general del edificio

- 1 Solo vivienda
- 2 Vivienda y comercio
- 3 Vivienda y taller
- 4 Solo comercios, oficinas, equipamientos
- 5 Solo artesanal o industrial
- 6 Bodega, banos ... etc de una vivienda
- 7 Utilizado para los animales
- 8 Garage
- 9 DESCONOCIDO

Encuesta

- 1 Vivienda ocupada encuestada
- 4 Vivienda ocupada no encuestada (no quiere contestar)
- 5 Vivienda ocupada no encuestada (en viaje)
- 8 No ocupada (sin residentes)

QUITO -- ENQUETE BATI/POPULATION - RECAPITULATIF LOGEMENT

<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° Manzana	Nouv. N° Bat.	Anc.N° Bat.princ.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° vivienda	Type de vivienda	Type d'occ. de la viv.	Nbre de pièces	Nbre de personnes	Usage géné. du bat.	Déroule. enquête		
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° Manzana	Nouv. N° Bat.	Anc.N° Bat.princ.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° vivienda	Type de vivienda	Type d'occ. de la viv.	Nbre de pièces	Nbre de personnes	Usage géné. du bat.	Déroule. enquête		
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° Manzana	Nouv. N° Bat.	Anc.N° Bat.princ.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° vivienda	Type de vivienda	Type d'occ. de la viv.	Nbre de pièces	Nbre de personnes	Usage géné. du bat.	Déroule. enquête		
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° Manzana	Nouv. N° Bat.	Anc.N° Bat.princ.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.	Anc.N° Bat.
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
N° vivienda	Type de vivienda	Type d'occ. de la viv.	Nbre de pièces	Nbre de personnes	Usage géné. du bat.	Déroule. enquête		

ANNEXE 3

Listage du programme VAR2D


```

      ENDTEXT
*
* Ouverture du fichier de procedures
*
      SET PROCEDURE TO VAR2PRO.PRG
*
* Creation des variables communes
*
NOMFICH=' '
REPVAR=' '
NB=0
A1=0
A2=0
B1=0
B2=0
*
* Affichage menu
*
      WAIT ' Appuyez sur une touche pour afficher le menu.'
      ECRMENU = .T.
      DO WHILE ECRMENU
      CLEAR
      MENU = ' '
      DATE = CDOW( DATE() ) + ' ' + STR( DAY( DATE() ), 2, 0 ) + ' ' ;
      + CMONTH( DATE() ) + ' ' + STR( YEAR( DATE() ), 4, 0 )
      @ 1,2 SAY 'Varédeg, version 1.'
      @ 3,3 SAY DATE
      @ 5,15 SAY ' ' MENU'
      @ 6,15 SAY ' ' ****'
      @ 12,15 SAY '<1> : Entré des parametres et calculs.'
      @ 14,15 SAY '<Q> : Quitter le programme'
      @ 17,1 SAY ' '
      BOUCLE1 = .T.
      DO WHILE BOUCLE1
      ACCEPT ' Choisissez une option : ' TO MENU
      MENU = UPPER( MENU )
      IF MENU = '1' .OR. MENU = 'Q'
      BOUCLE1 = .F.
      ENDIF
      ENDDO
      CLEAR
*
      DO CASE
*
      CASE MENU = 'Q'
      SET COLOR TO
      CLEAR
      CLOSE PROCEDURE
      RETURN
      CASE MENU = '1'
* Entrée des noms de fichiers
      CLEAR
      @ 2,5 SAY "FICHER DES UNITES PRIMAIRES"
      @ 3,5 SAY "*****"
      DO QUESFICH
      FICHPRIM=NOMFICH
      @ 10,5 SAY "FICHER DES UNITES SECONDAIRES"
      @ 11,5 SAY "*****"

```

```

DO QUESFICH
FICHSEC=NOMFICH
* Entrée des noms de variables
CLEAR
@ 2,5 SAY "VARIABLE A ESTIMER"
@ 3,5 SAY "*****"
DO QUESVAR
VARY = REPVAR
@ 10,5 SAY "VARIABLE EXOGENE DEFINISSANT LA PROBABILITE"
@ 11,5 SAY "*****"
DO QUESVAR
VARZ = REPVAR
@ 18,5 SAY "VARIABLE IDENTIFIANT LES UNITES PRIMAIRES"
@ 19,5 SAY "*****"
DO QUESVAR
IDPRIM = REPVAR
* Entrée du code des valeurs manquantes
CLEAR
@ 2,5 SAY "VALEUR DE CODE AFFECTEE AUX DONNEES MANQUANTES"
@ 3,5 SAY "*****"
DO QUESVAL
VALMANQ = NB
CLEAR
*
* Ouverture des zones de travail
*
SELECT 1
USE &FICHPRIM
SELECT 2
USE &FICHSEC
WAIT
*
* Calculs
*
DO CALPRIM
*
* Edition
*
SET PRINTER ON
SET DEVICE TO PRINT
*
@ 2,5 SAY "VARIABLE ESTIMEE :"
@ 2,25 SAY VARY
@ 5,5 SAY "VARIABLE EXOGENE DEFINISSANT LA PROBABILITE :"
@ 5,53 SAY VARZ
*
@ 7,5 SAY "PLAN EQUIPROBABLE"
@ 8,5 SAY "*****"
@ 10,10 SAY "1° terme (inter-grappes)"
@ 10,40 SAY "2° terme (intra-grappes)"
@ 13,10 SAY "A ="
@ 13,13 SAY A1
@ 13,40 SAY "B ="
@ 13,43 SAY B1

```

```
*
@ 15,5 SAY "PLAN A PROBABILITES INEGALES"
@ 16,5 SAY "*****"
@ 18,10 SAY "1° terme (inter-grappes)"
@ 18,40 SAY "2° terme (intra-grappes)"
@ 21,10 SAY "A ="
@ 21,13 SAY A2
@ 21,40 SAY "B ="
@ 21,43 SAY B2
```

```
*
SET PRINTER OFF
SET DEVICE TO SCREEN
ENDCASE
ENDDO
```

```

*****
*
* VAR2PRO.FRG: FICHIER DE PROCEDURES POUR LE PROGRAMME VAR2DEG*
* (Version n°1)
*
*****
*
* 1) Procédures de calcul
*
-----
* 1 : PROCEDURE CALPRIM
*
* Sommation sur les unités primaires.
*
-----
      PROCEDURE CALPRIM
*
* Initialisation
*
      A1=0
      A2=0
      B1=0
      B2=0
      N=0
      Y=0
      Z=0
      SA1=0
      SA2=0
      SB1=0
      SB2=0
      NI=0
      SYI=0
      S2YI=0
*
* Boucle sur les unités primaires
*
      SELECT 1
      DO WHILE .NOT. EOF()
      NOMANZ=&IDPRIM
      DO CALSEC WITH VARY, NOMANZ
      SELECT 1
      ZI=&VARZ
      SA1=SA1 + SYI^2
      N=N + 1
      Y=Y + SYI
      Z=Z + ZI
      SA2=SA2 + SYI^2/ZI
      SB1=SB1 + NI*S2YI
      SB2=SB2 + NI*S2YI/ZI
*
      ? RECNO()
*
      ?? NOMANZ, SA1, SB1, SA2, SB2
*
      WAIT
      SKIP
      ENDDO

```

*

* Resultats

*

```
A1=N^2*(SA1/N - (Y/N)^2)
A2=Z*SA2 - Y^2
B1=N*SB1
B2=Z*SB2
RETURN
```

*

* 2 : PROCEDURE CALSEC

*

* Sommaton sur les unités secondaires.

```
PROCEDURE CALSEC
PARAMETERS Y, NO
SELECT 2
SET FILTER TO &IDPRIM=NO .AND. &Y#VALMANQ
COUNT TO NI
SUM &Y, &Y^2 TO SY, SY2
S2YI= (SY2 - SY^2/NI)/(NI-1)
SYI=SY
?? NO, NI, SY, SY2, S2YI
WAIT
RETURN
```

*

* 2) Procédures utilitaires

*

* 1 : PROCEDURE QUESFICH

*

* Entree d'un nom de fichier dans le programme VAR2DEG.

```
PROCEDURE QUESFICH
BOUCLE = .T.
DO WHILE BOUCLE
? 'Entrez le nom du fichier, sans extention ni chemin
d"accés.'
? '(8 caractères maximum).'
? 'il doit être dans le repertoire dbaseiii du disque dur.'
? ''
ACCEPT ' nom :' TO NOMFICH
NOMFICH = TRIM(NOMFICH) + '.DBF'
NOMFICH = UPPER(NOMFICH)
IF LEN(NOMFICH) >0 .AND. LEN(NOMFICH) <13
BOUCLE = .F.
ENDIF
ENDDO
RETURN
```

```
*
*
*-----
* 2 : PROCEDURE QUESVAR
*
* Entree d'un nom de variable dans le programme VAR2DEG.
*-----
```

```
    PROCEDURE QUESVAR
    BOUCLE = .T.
    DO WHILE BOUCLE
    ? ''
    ACCEPT 'Entrez le nom de la variable.(12 caractères
maximum) :' TO REPVAR
    REPVAR = TRIM(REPVAR)
    REPVAR = UPPER(REPVAR)
    IF LEN(REPVAR) >0 .AND. LEN(REPVAR) <12
    BOUCLE = .F.
    ENDIF
    ENDDO
    RETURN
```

```
*
*
*-----
* 3 : PROCEDURE QUESVAL
*
* Entree de la valeur de code dans le programme VAR2DEG.
*-----
```

```
    PROCEDURE QUESVAL
    INPUT 'Entrez la valeur du code :' TO NB
    RETURN
```

```
*
```

ANNEXE 4

Questionnaire de l'enquête finale "Démographie / migration"

RUBRICAS		N° de persona <input type="text"/>	N° de persona <input type="text"/>	N° de persona <input type="text"/>	N° de persona <input type="text"/>	N° de persona <input type="text"/>	N° de persona <input type="text"/>
Dónde vivía usted antes de residir en Quito?		Pais Provincia Parroquia Nombre de la localidad		
Cuál es el último grado, curso o año que aprobó en la enseñanza regular?		Ninguno <input type="checkbox"/> Primaria <input type="checkbox"/> grado Secundaria <input type="checkbox"/> curso Superior <input type="checkbox"/> año		Ninguno <input type="checkbox"/> Primaria <input type="checkbox"/> grado Secundaria <input type="checkbox"/> curso Superior <input type="checkbox"/> año		Ninguno <input type="checkbox"/> Primaria <input type="checkbox"/> grado Secundaria <input type="checkbox"/> curso Superior <input type="checkbox"/> año	
Actividad principal:	Ocupación	
	Rama de actividad	
	Categoría de ocupación	
	Ingresos (suces)	<input type="text"/> por <input type="text"/> <i>Monto Duración</i>		<input type="text"/> por <input type="text"/> <i>Monto Duración</i>		<input type="text"/> por <input type="text"/> <i>Monto Duración</i>	
Secunda actividad:	Ocupación	
	Rama de actividad	
	Categoría de ocupación	
	Ingresos (suces)	<input type="text"/> por <input type="text"/> <i>Monto Duración</i>		<input type="text"/> por <input type="text"/> <i>Monto Duración</i>		<input type="text"/> por <input type="text"/> <i>Monto Duración</i>	
Tiene usted otras fuentes de ingresos? (para cada tipo de fuente, anotar "no" si no tiene ingresos de este tipo - la fuente precisa y el monto mensual, si tiene ingresos de este tipo)		Transportes <input type="text"/> Comercio <input type="text"/> Inmobiliario <input type="text"/> Agricultura <input type="text"/> Pension, beca <input type="text"/> Otra <input type="text"/> No quiere responder <input type="checkbox"/>		Transportes <input type="text"/> Comercio <input type="text"/> Inmobiliario <input type="text"/> Agricultura <input type="text"/> Pension, beca <input type="text"/> Otra <input type="text"/> No quiere responder <input type="checkbox"/>		Transportes <input type="text"/> Comercio <input type="text"/> Inmobiliario <input type="text"/> Agricultura <input type="text"/> Pension, beca <input type="text"/> Otra <input type="text"/> No quiere responder <input type="checkbox"/>	

2. VARIABLES SOBRE LAS CONDICIONES DE OCUPACION DE LA VIVIENDA

Usted ocupa la vivienda en calidad de qué ?

- El jefe del hogar es el propietario → Fecha y costo de la construcción o de la compra de la vivienda :
- Una otra persona que vive en la vivienda es propietario → Número de la persona []
- Inquilino → Monto del alquiler mensual (sucres) []
- Anticresis → Monto mensual (sucres) []
Monto del anticresis mismo (sucres) []
- Ocupacion de favor (gratis ; préstamo de un amigo, pariente, colega)
- Por servicios → Tipo de servicio :
Monto del alquiler mensual (sucres) []
- Otro tipo de ocupación → Precizar :
- Se niega a responder
- DONACION O HERENCIA [] Fecha []
- COMPRA [] Fecha [] Costo (sucres) []
- CONSTRUCCION : compra del terreno [] Fecha [] Costo (sucres) []
construc. del edificio [] Fecha [] Costo (sucres) []

Si el propietario no vive en la vivienda :
de qué clase de propietario se trata ?

- Persona física (natural) que no vive en la vivienda → Lazo de parentesco del jefe del hogar con el propietario :
- Sociedad privada
- Organismo semi-público
- Organismo público
- Organismo municipal
- Organismo diplomático
- Organización sindical
- Cooperativa
- Organismo religioso
- Otra clase de propietario
- Ninguno
- Lazo de parentesco :
- Se niega a responder
- Se niega a responder

3. LUGARES DE RESIDENCIA Y ACTIVIDADES DEL JEFE DEL HOGAR DURANTE LOS DOS ULTIMOS AÑOS

Calendario	Fecha del cambio de residencia	LUGAR DE RESIDENCIA (cualesquiera que sea la duración) <i>País, provincia, parroquia, localidad, dirección</i>	Fecha del cambio de actividad	ACTIVIDAD PRINCIPAL <i>Ocupación, rama de actividad, categoría de ocupación, ingresos mensuales</i>	RAZÓN DEL CAMBIO DE RESIDENCIA O DE ACTIVIDAD
Diciem. 87					
Noviem. 87					
Octubre 87					
Septiem. 87					
Agosto 87					
Julio 87					
Junio 87					
Mayo 87					
Abril 87					
Marzo 87					
Febrero 87					
Enero 87					
Diciem 88					
Noviem 88					
Octubre 88					
Septiem 88					
Agosto 88					
Julio 88					
Junio 88					
Mayo 88					
Abril 88					
Marzo 88					
Febrero 88					
Enero 88					

4. BIOGRAFIA MIGRATORIA, PROFESIONAL

DEL JEFE DEL HOGAR DESDE SU NACIMIENTO

N° Manzana
 N° Edificio
 N° Vivienda

Ficha suplementaria

Eddad (años)	Fecha (dia, mes, año)	LUGAR DE RESIDENCIA (+ 6 meses) Pais, provincia, parroquia, localidad, o localidad, barrio y dirección si se trata de Quito	Ocupación, rama de ocupación.	ACTIVIDAD de actividad, categoría Ingresos mensuales (suces)	RAZON DEL CAMBIO DE RESIDENCIA O DE ACTIVIDAD
Nacim. hasta	<input type="text"/> hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	
..... hasta hasta <input type="text"/>		Ocup..... Rama.....	Categ..... Ingresos <input type="text"/>	

5 - VARIABLES SOBRE LOS PARIENTES DEL JEFE DEL HOGAR Y DE SU CONYUGE QUE NO VIVEN EN LA VIVIENDA ENCUESTADA

Tipo de parentesco	Apellido y nombre	Sexo (M=Masculino F=Femenino)	Edad o año de nacimiento	Si fallecido(a), año del deceso	Lugar de residencia actual <i>País, Provincia, parroquia, localidad o localidad, barrio y dirección si se trata de Quito</i>	Ocupación principal actual	Relación económica actual con este parente (Códigos binarios)
PADRE del jefe del hogar			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
MADRE del jefe del hogar			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
CONYUGE del jefe del hogar			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
PADRE del conyuge del jefe del hogar			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
MADRE del conyuge del jefe del hogar			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
HIJOS (ASÍ) del jefe del hogar y de su conyuge (1)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(2)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(3)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(4)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(5)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(6)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(7)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(8)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(9)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(10)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(11)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(12)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>
(13)			edad o año de na			ocup. rama..... categ.....	<input type="checkbox"/> <input type="checkbox"/>

1 = El jefe del hogar no recibe nada y no manda nada a este parente
 2 = El jefe del hogar recibe una ayuda financiera de este parente
 3 = El jefe recibe una ayuda en especie (ej.: productos agric.) de este parente

4 = El jefe del hogar alberga un hijo de este parente
 5 = Este parente recibe una ayuda financiera del jefe del hogar
 6 = Este parente recibe una ayuda en especie del jefe del hogar

7 = Este parente alberga un hijo del jefe del hogar
 8 = No quiere responder