

DEVELOPPEMENT D'UNE METHODE DE SONDAGE DEMOGRAPHIQUE EN MILIEU URBAIN INTEGRANT L'INFORMATION SATELLITAIRE

Présentation des principaux résultats obtenus à Marseille et à Quito

1. Présentation de la recherche

En 1985, une équipe de l'ORSTOM¹ s'est constituée autour du programme "Utilisation de la télédétection pour l'observation des populations urbaines" lancé par le département "Urbanisation et socio-systèmes urbains". Nous effectuons depuis cette date des recherches visant à développer une méthode de collecte par sondage de données démographiques en ville qui intègre l'information issue de la télédétection spatiale. L'idée directrice de cette recherche est d'utiliser l'image satellite tout d'abord comme base pour la sélection d'un échantillon d'unités aréolaires à enquêter, puis comme une source d'information pour obtenir une stratification de la base de sondage sur des critères de morphologie urbaine. Mon propre travail a consisté à évaluer la précision des estimations que permettent les plans de sondages bâtis sur cette idée et à construire, sur la base de ces évaluations, un sondage, le plus précis possible, qu'on puisse facilement mettre en œuvre.

La recherche se déroule en deux étapes principales. La première, essentiellement méthodologique, s'est déroulée en France à Marseille, ville sur laquelle nous disposions d'informations exhaustives sur la démographie et la morphologie de l'agglomération qui ont permis de mesurer les variances réelles des estimations obtenues avec différents sondages.

La deuxième étape, commencée en Août 1986, se déroule en Equateur à Quito où nous sommes associés au projet franco-équatorien d'Atlas Informatisé de la ville (AIQ). L'objectif final de cette seconde étape est de vérifier, en réalisant une enquête démographique sur un échantillon sélectionné sur l'image, que la méthode mise au point possède les qualités souhaitées : rapidité et facilité de la mise en œuvre, faiblesse de coûts et précision des résultats.

Nous allons maintenant résumer les objectifs poursuivis et les principaux résultats obtenus lors des deux phases de l'expérience.

¹ : Cette équipe comprend O.BARBARY (Statistiques), F.DUREAU (Démographie), A. HAING (Photographie), B. LORTIC (Télédétection), A. MICHEL (Télédétection), I. RANOUX (Photographie) M.SOURIS (Informatique).

Attn à l'inversion des pages, suivre la
pagination rectifiée

2. Mise au point de la méthode à Marseille, approche du premier degré du sondage

Objectifs, données et outils statistiques

La phase marseillaise de l'expérience avait trois objectifs principaux :

1 : Evaluer, à partir des données du recensement général de la population à Marseille l'efficacité d'une stratification morphologique de la ville pour estimer par sondage les caractéristiques démographiques des habitants.

2 : Définir, après évaluation des différentes alternatives, les caractéristiques que doit avoir un plan de sondage utilisant comme base l'image satellite.

3 : Identifier les principales spécificités du plan de sondage défini et analyser les conséquences qu'elles ont sur la précision des estimations et la mise en œuvre pratique de l'enquête.

Les tests ont été étendus à plusieurs plans de sondage, différents soit par la probabilité de sélection des unités statistiques, soit par l'estimateur employé, ainsi qu'à plusieurs stratifications. La première est une classification morphologique synthétique de l'agglomération marseillaise issue d'une analyse typologique effectuée sur des données provenant de l'interprétation de photographies aériennes. On a ensuite reproduit l'expérience sur des classifications plus simples équivalentes à celles obtenue sur image satellite.

La base de données nécessaire à l'expérience comprend d'une part une sélection d'indicateurs démographiques issus de l'exploitation exhaustive du recensement de population effectué à Marseille en 1982, d'autre part les données morphologiques et les mesures de surface des districts de recensement qui permettront l'évaluation des procédures de sondage stratifiées et à probabilités inégales. Dans cette base, les données démographiques sont agglomérées au niveau des districts de recensement définis par l'INSEE qui correspondent généralement à des pâtés de maisons physiques. Les recherches menées à partir de ces données concernent donc exclusivement le premier degré du sondage. Ses conclusions portent sur les techniques de sélection et d'inférence à appliquer aux unités primaires du sondage : les pâtés de maisons.

L'outil statistique développé pour effectuer l'ensemble de ces tests est un programme informatique écrit sous le logiciel dbaseIII sur micro ordinateur IBM AT; il calcule, à partir des données exhaustives du recensement, la variance des estimations du caractère démographique choisi, pour un plan de sondage et une stratification donnés.

Résumé des conclusions

Objectif 1 : La stratification de la base de sondage à l'aide de la classification morphologique synthétique entraîne une diminution de 40 % en moyenne de la variance des estimations de totaux et de 50 % pour les estimations de quotients. Le critère de densité du bâti joue un rôle déterminant dans ce gain de précision : il est responsable de 50 à 75% de l'amélioration apportée par la stratification synthétique. D'autre part l'équipe de télédétection avec laquelle je travaille a montré que l'on pouvait produire des classifications de l'image

satellite qui rendent compte de manière satisfaisante de la densité du bâti en milieu urbain.

Objectif 2 : L'utilisation de l'image satellite comme base de sondage implique quasi nécessairement un plan de sondage à probabilités inégales. La technique de sélection des unités spatiales à enquêter la plus simple à mettre en œuvre est celle du tirage systématique à l'aide d'une grille de points plaquée sur l'image. La probabilité d'inclusion des unités dans l'échantillon est alors proportionnelle à leur surface. Les conclusions de l'expérience conduisent au plan de sondage suivant :

1 : 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.

2 : Cet espace est stratifié selon un ensemble de critères morphologiques fournis par diverses classifications de l'image satellite. Parmi ces critères, le premier et certainement le plus important est la densité du bâti.

3 : La sélection de l'échantillon se fait grâce à un tirage systématique où la probabilité d'appartenance à l'échantillon d'une unité spatiale donnée est proportionnelle à une mesure de sa taille. La plus "naturelle" de ces mesures est évidemment la surface totale de l'unité, c'est aussi la plus commode à mettre en œuvre lors du tirage, mais ce n'est pas l'unique possible.

4 : Les estimateurs sont sans biais lorsqu'il s'agit d'estimer des totaux mais les estimateurs par le ratio restent utiles pour estimer directement des quotients.

Objectif 3 : 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 de chaque unité primaire (pâté de maisons) enquêtée, il s'agira très probablement de la surface totale et il faut donc disposer d'un moyen fiable de la mesurer.

- La taille de chaque strate, s'il s'agit de la surface totale sa mesure s'effectue sans problème 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).

3. Application de la méthode à Quito, exploitation des enquêtes préliminaires pour la définition du plan de sondage

Les enquêtes préliminaires, données collectées et objectifs

La phase d'application à Quito a débuté par la réalisation de deux enquêtes préliminaires légères ayant pour but de collecter sur échantillons des données morphologiques et démographiques. Cette collecte répond à une double nécessité .

Premièrement il reste en suspens des questions auxquelles les données disponibles à Marseille n'ont pas permis de répondre. c'est principalement le cas en ce qui concerne la définition du second degré du sondage : la sélection d'un échantillon de ménages au sein d'un pâté de maisons. Pour conclure sur ce point, il faut disposer d'informations démographiques pour chacun des ménages de l'îlot, information que ne fournit pas le recensement général de la population à Marseille.

Deuxièmement, certains des résultats obtenus à Marseille, peuvent ne pas être directement transposables à Quito et nécessiter une adaptation aux spécificités locales. quant aux résultats fondamentaux il faut vérifier qu'ils restent valables.

Les objectifs assignés à l'exploitation des données collectées sont les suivants:

1. Définition de l'unité spatiale à enquêter, nombre de degrés du sondage, évaluation de la précision des procédures à plusieurs degrés.
2. Choix des critères morphologiques de stratification.
3. Définition de la probabilité d'appartenance des unités primaires à l'échantillon : proportionnalité à la surface totale des unités ou proportionnalité à une approximation de leur surface bâtie.
4. Définition de l'estimateur à employer: estimateur sans biais ou estimateur par le ratio
5. Définition d'une règle d'allocation aux strates.

Résumé des conclusions

Objectif 1 : 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" permet de conclure 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 au minimum cinq ménages enquêtés dans chaque pâté de maisons.

Objectif 2 : 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.

Objectif 3 : 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, la mesure de la surface totale suffit à obtenir des précisions équivalentes.

Objectif 4 : 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 de ratio pour les quotients.

Objectif 5 : Grâce à une bonne stabilité des des règles d'allocation optimale, quelque soient les caractères estimés, on peut définir l'allocation aux strates suivante pour le premier degré du sondage :

strates	taux approximatifs au premier degré (en % du nombre de manzanas de la strate)
1 : densité très faible	2.5
2 : densité faible	2.2
3 : densité moyenne	6.7
4 : densité forte	8.9
5 : densité très forte	4.4
Ensemble de la base	5.8

L'ensemble de ces conclusions conduisent au plan de sondage suivant :

1. La base de sondage est l'image satellite 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 areolaire, 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.

3. Tirage au deuxième degré : Les unités secondaires sont les ménages, la sélection est la 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 ménage dans la manzana : enquête exhaustive
- plus de dix ménages dans la manzana : dix ménages sont enquêtés

Le taux final obtenu est d'environ 1,5 % 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.

L'exploitation des données recueillies a commencée fin Avril et les résultats nécessaires pour évaluer définitivement la précision de la méthode doivent être connus courant Juillet.

4. Conclusion

En conclusion, il nous semble important de situer la recherche que nous venons de présenter dans le cadre général des méthodes de collecte de données. Au départ de notre projet, il y a une idée qu'exprime bien L. M. ASSELIN dans son livre "techniques de sondage avec application à l'Afrique". Pour lui, "les techniques de sondages, outil économique et rigoureux de production d'information, doivent être au coeur de l'information quantitative requise par les stratégies de développement". C'est aussi notre conviction et plus encore s'agissant de produire des données sur les populations urbaines. En effet, face aux évolutions rapides que connaissent la plupart des villes des pays en développement, la rapidité de mise en œuvre et d'exploitation des enquêtes par sondage, ajoutée à leur caractère économique et rigoureux, fait apparaître cette méthode comme la seule réponse possible aux besoins d'information exprimés tant par les scientifiques que par les gestionnaires.

Nous pensons que le plan de sondage aréolaire présenté ici est une solution bien adaptée aux problèmes techniques et pratiques que pose la réalisation de telles enquêtes en milieu urbain, en particulier parce qu'il ne nécessite pas une liste tenue à jour des unités statistiques. De plus la méthode présentée nous semble avoir un large domaine d'application. L'utilisation de l'image satellite, la thématique démographique du questionnaire et les sites d'application retenus (Quito et Marseille) ne doivent pas être perçus comme des éléments limitants mais simplement comme le cadre dans lequel on a pu mettre au point le plan de sondage. La méthode est en fait facilement transposable sur d'autres villes, elle convient à la collecte d'information sur de nombreux thèmes (démographie, migration, économie, santé etc...) et enfin, si l'on dispose d'une autre source d'information pour construire la base de sondage spatiale (cartographique, cadastrale, photographique), on peut l'appliquer en l'absence d'image satellite. Les outils statistiques et informatiques développés pendant cette recherche permettent d'ailleurs, sur les sites où existent des données suffisantes, d'adapter le plan de sondage et d'évaluer à l'avance la précision des estimations qu'il fournit.

Département **S**ociété, **D**éveloppement, **U**rbanisation
(S. D. U.)

ATLAS INFORMATISE DE QUITO

- PRESENTATION ET PREMIERS RESULTATS -

Documents de Travail

ORSTOM

SOMMAIRE

AVANT-PROPOS

1. L'Atlas Informatisé de Quito (AIQ) : les objectifs généraux

1.1. Bref historique du programme Atlas

Informatisé de Quito (AIQ)

H.R. GODARD

1.2. Atlas Informatisé de Quito (AIQ) :

termes de référence scientifiques et
techniques

Equipe AIQ

1.3. L'Atlas Informatisé de Quito (AIQ) :

production de connaissance et dévelop-
pement de méthodologies et d'outils

Equipe AQI

1.4. L'Atlas Informatisé de Quito (AIQ) :

un outil au service de la gestion et de
la planification urbaine

H.R. GODARD

M. SOURIS

2. L'AIQ : de nouveaux outils et de nouvelles méthodologies au service de la gestion urbaine

2.1. Configuration du matériel informatique

MATRA-SUN

M. SOURIS

2.2. La constitution de la base de données

urbaines sur Quito : premières données et
organisation de la saisie graphique

M. SOURIS

- 2.3. L'utilisation de la télédétection pour
l'observation des populations urbaines :
un programme du Département SDU en
cours à Quito
- F. DUREAU
B.LORTIC
- 2.4. Développement d'une méthode de
sondage démographique etc...
- A. MICHEL
M. SOURIS
- 2.5. Influence des facteurs météorologiques
sur la probabilité d'obtention d'une image
satellite SPOT ou THEMATIC MAPPER
- A. MICHEL
3. L'AIQ : premiers axes de réflexion, premiers documents de
travail
- 3.1. Quito et ses limites : problèmes de
définition et zone d'étude AIQ
- R. VALLEJO
- 3.2. La croissance spatiale de Quito à
partir des plans historiques (1760, 1888,
1921, 1946)
- H.R. GODARD
J.VEGA
- 3.3. Micro-zonification des dommages
occasionnés à Quito lors du séisme du
5 mars 1987
- M.A de CASTRO
- 3.4. L'évolution spatiale du réseau
bancaire
- H.R.GODARD
J. VEGA
- 3.5. Structure et dynamique des centres
de Quito et de Guayaquil
- H.R. GODARD

3.6. Composition et évolution du salaire

minimum mensuel salariés du secteur

public et travailleurs en général

(1970-1987)

H.R. GODARD
CI.de MIRAS