

Cette fiche a été réalisée dans le cadre du rapport du CETE de Rouen sur les méthodes d'estimation de population dans des zonages thématiques

Estimation du nombre d'étage des bâtiments de la BDTopo®

La démarche générale consiste à soustraire l'altitude au sol, obtenue à partir des courbes de niveau et des points cotés de la BDTopo®, de l'altitude au toit des bâtiments dont l'information est donnée par la BDTopo®.

► Les données utilisées

Les données utilisées sont les couches LIGNE_ISO, POINT_COTE et BAT_SURF de la BDTopo® de l'IGN.

► La méthode

Étape 1 : création d'une couche de points cotés à partir des courbes de niveau

L'objectif de cette étape est d'obtenir une couche de points cotés la plus exhaustive possible sur le secteur concerné par l'estimation.

Pour ce qui concerne la BDTopo, deux couches d'information vont être utiles lors de cette phase, la table des courbes de niveau (LIGNE_ISO.tab) et la table des points cotés (POINT_COTE.tab). L'objectif est de constituer une nouvelle table de points cotés la plus exhaustive possible sur le territoire d'étude. Pour cela, on va transformer la table des courbes de niveau (polylignes) en une table de points cotés grâce à l'utilisation de l'outil p2p.mbx. Ensuite, on fusionne la table de points ainsi créée avec la table de points cotés (POINT_COTE.tab), ce qui permet d'obtenir une couverture plus complète de points cotés sur le territoire d'étude.

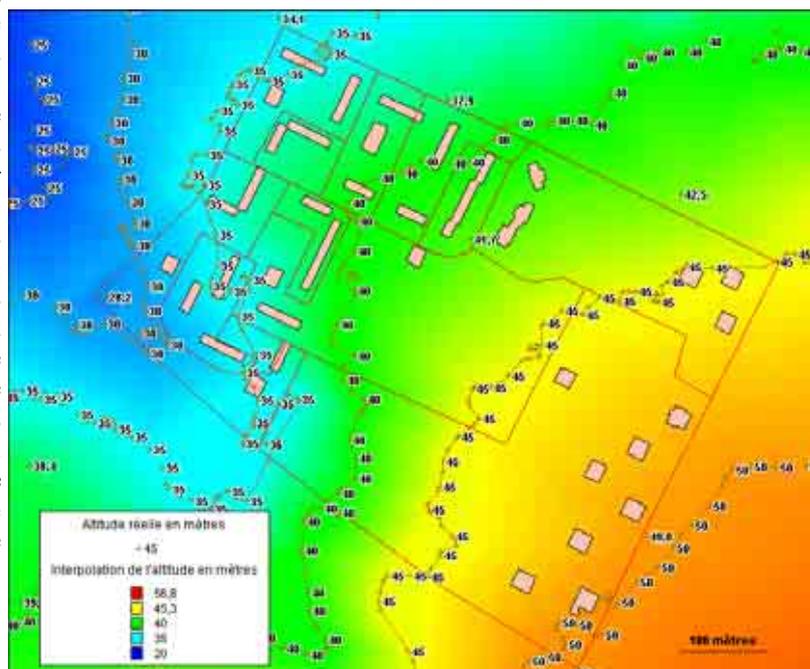
Étape 2 : génération d'une grille à partir des points cotés

Le semis de points cotés étant trop faible pour simuler une altitude relative réaliste, il est nécessaire de procéder à une interpolation de cette altitude à partir des points cotés.

La démarche consiste dans un premier temps à réaliser dans MapInfo une analyse thématique par coloration continue à partir de la couche des points cotés obtenue lors de la 1^{ère} étape. Pour obtenir une interpolation de type Modèle Numérique de Terrain, on privilégiera la méthode d'interpolation TIN en choisissant une taille de cellule la plus fine possible. 5 à 10 mètres de côté paraissent suffisants.

Logiquement, une taille de cellule de 25 mètres devrait être suffisante, c'est d'ailleurs la taille du MNT BDTopo fourni par l'IGN.

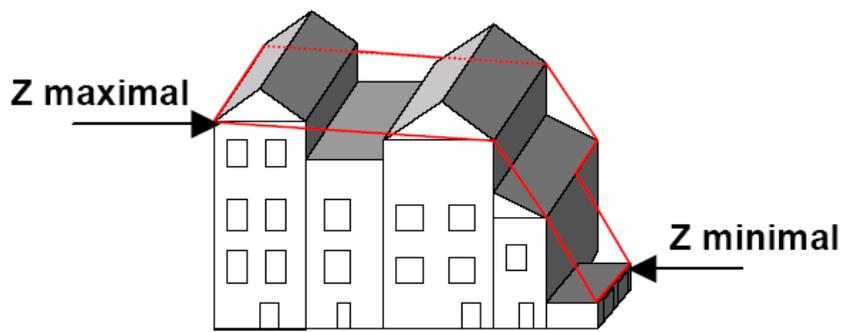
Cependant, plusieurs tests réalisés dans le cadre de cette étude ont permis de mettre en évidence une plus grande précision des résultats avec un pas de 5 mètres.



Le fichier MIG généré

Lors de l'interpolation, MapInfo génère un fichier mig correspondant à l'image rasterisée de cette interpolation. Comme il n'est pas possible d'extraire directement les résultats de cette interpolation (fichier raster), il est nécessaire de procéder à la vectorisation de ce fichier mig grâce à l'utilitaire InfoExMig.mbx.

Monde réel et modélisation



Source : La BD TOPO Pays, Version 1.2, Descriptif du contenu, janvier 1997, Édition 1.0, décembre 2002.

Cet utilitaire va permettre de générer une grille vecteur dont chaque cellule intègre un résultat partiel de l'interpolation.

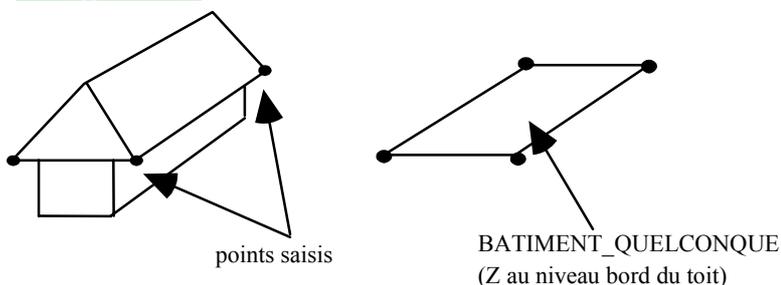
Étape 3 : estimation de la hauteur réelle des bâtiments à partir de la grille

A partir de la grille ainsi constituée, il est possible de produire une estimation de la hauteur moyenne des bâtiments grâce à la fonction « mettre à jour colonne » de MapInfo. Le calcul est réalisé à partir de la moyenne proportionnelle, en surface, de l'altitude des cellules de la grille intersectant les bâtiments (« Proportion Avg »).

Cette opération permet d'obtenir pour chaque bâtiment une estimation de l'altitude au sol. Comme la BDTopo fournit pour chaque bâtiment l'altitude du toit, il ne reste plus qu'à soustraire l'estimation de l'altitude au sol obtenue en phase 2 de l'altitude au toit des bâtiments donnée par la BDTopo afin de produire une estimation de la hauteur des bâtiments.

Il faut également préciser que la BDTopo fournit deux valeurs d'altitude, une valeur basse et une valeur haute. Les bâtiments de la BDTopo sont des objets surfaciques numérisés au bord du toit.

Les valeurs basses et hautes du z correspondent donc respectivement aux points bas et haut des bords de toit. Ces valeurs sont en règle générale assez proches, sauf en milieu urbain dense où les



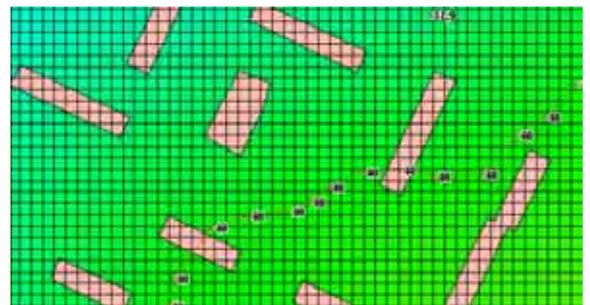
Source : La BD topographique, Spécifications détaillées, Édition 4, janvier 1997

bâtiments sont côte à côte et sont numérisés dans la BDTopo comme des entités homogènes.

Dans l'exemple ci-dessus, le bâtiment quelconque en rouge regroupe en fait plusieurs bâtiments distincts de hauteur variable.

L'objectif étant d'estimer le nombre d'étages des bâtiments par rapport à l'altitude au sol, dans le cadre de cette approche, nous utiliserons la valeur haute donnée par la BDTopo afin d'éviter des estimations d'étages négatives

4^{ème} étape : estimation du nombre d'étages à partir de la hauteur estimée des bâtiments



Superposition de la grille d'un pas de 5 mètres permettant de récupérer les valeurs de l'interpolation

L'estimation du nombre d'étages à partir de la hauteur estimée des bâtiments suppose qu'un choix soit fait en ce qui concerne la hauteur moyenne d'un étage. Dans le cadre de l'étude réalisée par le CETE Normandie-Centre en 2003, le choix s'est porté vers une hauteur moyenne de 3 mètres pour un étage. Ce choix ayant ensuite été validé par une comparaison avec la réalité terrain, il est également repris dans le cadre de cette méthode.

En divisant la hauteur estimée des bâtiments par 3, on obtient donc une estimation du nombre d'étages de chaque bâtiment.

Évaluation de la méthode

Afin de rendre compte de la pertinence de la méthode, il convient de comparer les résultats obtenus dans le cadre de l'estimation avec la « hauteur réelle » des bâtiments. Sans la possibilité pratique de vérifier la hauteur réelle des 5.704 bâtiments de la zone d'étude, on a donc utilisé la BDTopo Pays qui donne une « hauteur réelle arrondie » des bâtiments, que l'on divise

également par 3 afin d'obtenir là aussi une estimation du nombre d'étages.

Pour information, la BDTopo Pays est la base de données qui succède à la BDTopo standard, en y apportant une actualisation, une amélioration (en agglomération, équipements publics et MNT plus précis par exemple) ainsi qu'un nouveau modèle de données.

Dans le cas présent, la BDTopo Pays est plus complète que la BDTopo standard puisqu'elle fournit directement une hauteur arrondie pour tous les bâtiments.

Sur une base de 5.704 bâtiments quelconques situés sur la zone d'étude, on obtient en comparant la « hauteur réelle arrondie » provenant de la BDTopo Pays et la hauteur estimée avec la BDTopo un coefficient de corrélation de 0,77.

Le nombre d'étages estimé à partir de la BDTopo a également été comparé au nombre d'étages déterminé à l'aide de la BDTopo Pays sur la zone d'étude :

L'écart constaté est

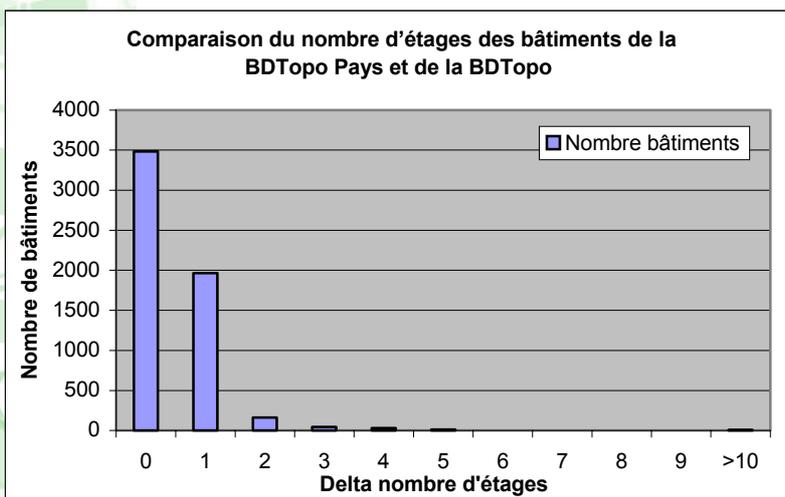
$$\text{delta} = \text{Nombre d'étages calculé BDTopo Pays} - \text{Nombre d'étages estimé BDTopo}$$

Tous types de bâti confondus

Nombre total d'îlots de l'échantillon : 552
 Nombre total de bâtiments de l'échantillon : 5.704

Delta en valeur absolue	0	1	2	3	4	5	6	7	>10	Total
Nombre de bâtiments	3482	1965	162	45	30	12	1	1	6	5704
Pourcentage	61,0	34,4	2,8	0,8	0,5	0,2	0,0	0,0	0,1	100,0

Comparaison du nombre d'étages des bâtiments de la BDTopo Pays et de la BDTopo



Les résultats sont donc dans l'ensemble assez comparables, bien que l'on constate des invraisemblances concernant le nombre d'étages de certains bâtiments de la BDTopo Pays (delta >10), comme plusieurs bâtiments d'une trentaine d'étages et même un bâtiment de 59 étages, soit 177 mètres de hauteur déclarée dans la BDTopo Pays, pour un bâtiment d'habitat collectif d'environ 10 étages.

Ces erreurs de la BDTopo Pays mériteraient sans doute une investigation plus poussée, mais qui ne relève pas du sujet de cette étude.

► Utilisation de Vertical Mapper

Pour les services disposant du logiciel Vertical Mapper, il est possible d'arriver au même résultat, en faisant un polytopoint pour obtenir le semis de points cotés et en réalisant une interpolation (triangulation) pour obtenir un fichier MNT .grd. C'est ce grd qui sera utilisé pour estimer l'altitude au sol des bâtiments.

Pour cela, on réalise un « region inspection » entre le MNT au format grd et la couche de bâtiments afin de calculer l'altitude moyenne au sol de chaque bâtiment.

► Bibliographie

La 3e dimension géographique, Utilisation des modèles numériques de terrain illustrée par la BDAlti de l'IGN, Certu, décembre 2001

La BD topographique, Spécifications détaillées, Édition 4, janvier 1997

La BD TOPO Pays, Version 1.2, Descriptif du contenu, janvier 1997, Édition 1.0, décembre 2002

Note rédigée par François-Xavier Robin et Vincent Demeules - CETE Normandie-Centre

Relue par Magali Di Salvo - CERTU

francois-xavier.robins@equipement.gouv.fr
 vincent.demeules@equipement.gouv.fr

magali.di-salvo@equipement.gouv.fr

Cette fiche a été réalisée dans le cadre du rapport du CETE de Rouen sur les méthodes d'estimation de population dans des zonages thématiques

Enquête de recensement de la population : une nouvelle méthode

L'objectif de cette annexe n'est pas de présenter une méthode d'estimation de population, mais les impacts de l'enquête de recensement de la population pour les estimations de population dans des zonages thématiques.

► Contexte

Depuis janvier 2004, le recensement de la population française est réalisé par enquête annuelle.

La commune reste le territoire de base de la collecte du recensement, mais on distingue les communes en fonction d'un seuil de population fixé à 10 000 habitants.

- ◆ Pour les communes de moins de 10 000 habitants, le recensement est exhaustif (100% de la population est recensée). Une commune sur cinq est enquêtée chaque année, soit au bout de 5 ans 100% des communes et des habitants.
- ◆ Pour les communes de 10 000 habitants ou plus, une enquête par sondage est réalisée chaque année auprès d'un échantillon de 8% de leur population. Le choix de l'échantillon se fera au niveau de l'immeuble et non plus de l'ilot, grâce au RIL (Répertoire de Immeubles Localisés).

Ainsi tous les 5 ans, on aura recensé 100% des habitants des communes de moins de 10 000 habitants et 40% de ceux des communes de 10 000 habitants ou plus.

Les données complexes (CSP, etc.) seront traitées à 100% des 40% recensés pour les communes de plus de 10 000 habitants et à 25% pour celles de moins de 10 000 habitants.

Enfin, pour « réajuster » les données et améliorer l'estimation proposée, l'INSEE se propose d'utiliser deux fichiers administratifs, le fichier de la taxe d'habitation et celui des bénéficiaires de l'assurance maladie, ce qui constitue une grande nouveauté en France.

► Le RIL (Répertoire d'Immeubles Localisés)

Le RIL est un répertoire d'adresses localisées (en coordonnées géographiques) associé à une base communale de données géographiques, Base-Ilots. Dans un premier temps il comprend les adresses des immeubles d'habitation (au sens du recensement), celles des communautés et celles des établissements du répertoire SIRENE. Il comprendra ultérieurement celles des équipements au service du public. Il ne concerne pour le moment que les communes de 10 000 habitants ou plus.

Il a été constitué pour les besoins du recensement de la population, pour lequel il constitue la base de sondage des adresses dans les communes de 10 000 habitants ou plus.

Le RIL a été constitué par l'INSEE et mis à jour à partir des fichiers de permis de construire et des fichiers de La Poste. Les communes ont également apporté une contribution sous la forme de confirmation de l'existence d'adresses et de travaux d'expertise.

On répartit ensuite les adresses en cinq groupes équilibrés à partir du RIL. Puis, chaque année, l'INSEE effectue un tirage dans le groupe de l'année d'un échantillon d'adresses représentant 40% des logements du groupe, soit 8% des logements de la commune. La collecte s'effectue alors pour l'ensemble des logements des adresses ainsi sélectionnées.

► La diffusion des données

À la fin de l'année 2008, à l'issue des cinq premières enquêtes de recensement, l'INSEE publiera pour la première fois les résultats statistiques complets sur les habitants et le logement. À partir de 2009, ces résultats seront mis à jour tous les ans.



Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE



Cependant, avant 2008, l'INSEE diffusera des résultats issus des enquêtes annuelles. Sont d'ores et déjà disponibles à partir de l'enquête de 2004 :

- ◆ Des chiffres-clés : indicateurs sur la population et les logements au niveau de la France et des régions, pour les communes de moins de 10 000 habitants ayant réalisé l'enquête de 2004 ainsi que pour une centaine de « grandes villes ».
- ◆ Une population provisoire : en janvier 2005 a été diffusé un chiffre de population pour les communes de moins de 10 000 habitants ayant réalisé l'enquête de 2004 et pour une centaine de « grandes villes ».

L'INSEE a entrepris une démarche de concertation afin de mieux faire comprendre le projet du nouveau recensement et l'adapter en fonction des attentes des utilisateurs. Il y a ainsi eu une réflexion sur la définition du dispositif de diffusion des résultats.

L'INSEE a constitué des groupes de travail au sein du Conseil National de l'Information Statistique (CNIS). De juin 2003 à l'automne 2004, un premier groupe de travail a été constitué afin de réfléchir sur la nature des résultats statistiques produits par le nouveau recensement et sur les modalités de leur diffusion. Composé d'experts (démographes, géographes, etc.) et de représentants de communes, ce groupe a remis à l'automne un rapport synthétisant ses réflexions et propositions. Celles-ci sont consacrées à la qualité et l'accessibilité des résultats du recensement ainsi qu'à leur mise à disposition.

Les propositions formulées par le groupe ont notamment influencé la présentation des résultats (tableau synthétisant les résultats à utiliser et les précautions d'emploi en fonction du niveau géographique), leur rythme de diffusion (montée en charge progressive), ainsi que leur accessibilité avec la mise en ligne gratuite sur le site Internet de l'INSEE.

La politique de diffusion de l'INSEE pour les prochaines années consiste entre autre à faciliter l'accès aux données statistiques et leur utilisation. Ainsi, outre la mise à disposition systématique sur Internet de la plus grande part possible de l'information, l'INSEE a établi une nouvelle politique tarifaire qui devrait aboutir en moyenne à une baisse des prix des produits de diffusion, et l'accès aux informations sur Internet sera gratuit. Quant aux cédéroms, leur prix sera fonction, comme pour les précédents recensements, des frais de réalisation, mais il ne dépendra plus du nombre de données qu'ils offriront.

Enfin, la rediffusion, commerciale ou non, du contenu de tout produit gratuit ou payant sera totalement libre de droits, sous réserve de la mention de la source et du respect de l'intégrité des données.

Depuis l'automne 2004, un groupe de travail, composé essentiellement de représentants de communes, est venu succéder au premier. Il a pour

mission d'approfondir la réflexion sur l'appropriation des résultats du recensement au niveau communal et infra-communal et sur leur restitution, afin de répondre au mieux aux besoins des communes.

► Questions posées par le RRP

Le recensement rénové de la population possède incontestablement des avantages majeurs : il permet de réduire la période intercensitaire et d'augmenter conséquemment la fraîcheur des informations, sans augmenter le coût de collecte, tout en simplifiant son organisation. Cependant cette nouvelle méthode soulève de nombreuses questions :

1. L'inégalité de traitement entre communes de moins de 10 000 habitants et communes de 10 000 habitants ou plus soulève des interrogations (absence de prise en compte de l'intercommunalité, pourquoi la limite de 10 000 habitants et pas un autre nombre).

L'INSEE explique que ce choix du seuil des 10 000 habitants est dû à des raisons juridiques (nécessité d'un critère objectif de distinction) et technique (à coût égal on peut assurer une collecte plus fréquente que pour le RGP). De plus, l'inégalité de traitement ne concerne que la méthode de collecte et non les résultats.

2. Le problème du possible refus des personnes de répondre à un sondage inquiète quelque peu : pour le recensement général, il était obligatoire de répondre, contrairement au RRP. Le taux de non-réponses risque donc d'augmenter et de perturber conséquemment la fiabilité des nombres d'habitants obtenus sur lesquels l'exploitation reposera.

Pour l'INSEE, utiliser le sondage permettra au contraire d'améliorer les taux de non-réponses obtenus lors des précédents recensements car il permet de concentrer l'effort de collecte d'informations sur un plus petit nombre de personnes, et donc d'assurer un meilleur suivi et par conséquent une meilleure qualité de la collecte.

3. Le découpage en IRIS 2000 a été établi sur la base du RGP90 lors de la préparation du RGP99 et n'a pas été revu depuis. Compte-tenu de l'importance prise par ce découpage dans l'enquête de recensement où il constitue la « brique de base » de diffusion, il semble pertinent de les redéfinir, ce qui est prévu. De plus, l'existence même des îlots semble compromise car la brique de diffusion élémentaire sera l'IRIS2000 en standard.

Les IRIS « atypiques », de faible taille, seront traités à part et nécessiteront sans doute un regroupement. De plus, des travaux de simulation ont montré a priori que le risque de moindre précision était minime.

4. Le RIL soulève également certaines questions. Si

on peut penser qu'il permettra de bien repérer les nouvelles entrées d'*immeubles*, on peut s'interroger sur la difficulté de déterminer tous les *logements* des immeubles d'un groupe de rotation donnée. De plus, certaines collectivités s'interrogent sur leur capacité à le mettre à jour (essentielle si on veut obtenir des résultats fiables), notamment en raison de la difficulté d'estimer le coût humain et financier d'une telle opération. Enfin, l'utilisation du RIL comme clé de répartition, permettant de ventiler approximativement des informations générales diffusées à l'IRIS, sur des sous-ensembles particuliers définis en termes d'adresses, serait très précieuse pour de nombreux utilisateurs. Or la diffusion du RIL n'est pour l'instant pas autorisée par les arrêtés de constitution du RIL soumis à l'approbation de la CNIL.

5. Le cas de la méthode des « **extrapolations temporelles** » (en raison des enquêtes réalisées à des dates différentes) soulève des interrogations quant à leur fiabilité : la photographie statistique d'une zone sera-t-elle crédible à un instant T sachant qu'elle est issue de données collectées à cinq dates différentes ? Comment pourrait-on « extrapoler » les migrations alternantes ou avoir une vision cohérente d'un territoire qui a connu la construction d'un équipement structurant majeur sachant que certaines communes ont été sondées avant cet équipement, et d'autres après ?

L'INSEE souligne que la rénovation propose un compromis entre la fraîcheur de l'information et sa précision. Cependant, la précision apportée par le RP n'était que théorique car elle existait seulement pour les phénomènes qui intervenaient au hasard du calendrier censitaire à distance raisonnable d'un futur recensement. La nouvelle méthode permettrait, quant à elle, de pouvoir apprécier des tendances d'évolution. De plus, toutes les données publiées seront accompagnées d'indicateurs de qualité.

6. Le cas particulier des **migrations résidentielles** : si on peut considérer que ramener l'intervalle à 5 ans est un progrès par rapport au RP, il reste la question de savoir traiter des informations sur les migrations pour tout type de territoire, sachant qu'elle n'est pas collectée au même moment partout et qu'elle change de date chaque année.

Le rajout d'une question sur le lieu de résidence 5 ans auparavant devrait permettre de disposer de taux de mobilité de meilleure qualité que l'on élabore actuellement. De plus, la méthode d'estimation de l'INSEE permettra de disposer, chaque année, d'informations sur les flux entre zones.

Note rédigée par Pascale Gondeaux

pascale.gondeaux@equipement.gouv.fr

CETE Normandie-Centre

Relue par Monique Gadais et Magali Di Salvo

monique.gadais@equipement.gouv.fr

CERTU

Magali.di-salvo@equipement.gouv.fr

Cette fiche a été réalisée dans le cadre du rapport du CETE de Rouen sur les méthodes d'estimation de population dans des zonages thématiques

Utilisation cartographique de la base de données Filocom

L'objectif de cette fiche est de montrer comment exploiter des données provenant de Filocom de manière cartographique avec la BD Parcellaire de l'IGN. Les données extraites de Filocom concernent ici la population et le logement en 2003.

► Filocom

FILOCOM est la base de données issue des fichiers fiscaux sur la taxe d'habitation, l'impôt sur le revenu et la taxe foncière. Elle a été élaborée par la DAEI à partir des données de la DGI.

Les données initiales sont recueillies à l'échelle de la parcelle cadastrale. Pour des raisons de confidentialité, elles sont agrégées au niveau de la section cadastrale dans la base FILOCOM.

Il s'agit d'une base alphanumérique, le code de section cadastrale permettant de joindre les données avec le cadastre numérisé.

La base FILOCOM est très riche en informations, notamment, sur les logements (type de logement, mode d'occupation, caractéristiques du logement, statut d'occupation des résidences principales) mais aussi sur leurs occupants et sur les mutations.

Cette base est remise à jour tous les deux ans. Elle est gérée au Ministère par le CETE Nord-Picardie qui peut en faire des extractions à la demande

Avantages et inconvénients de FILOCOM

Les principaux avantages

La réactualisation : le fait que la base soit réactualisée tous les deux ans est un atout.

La richesse d'information : le grand nombre d'informations disponibles dans FILOCOM rend cette base très intéressante, surtout pour le logement (résidence principale ou résidence secondaire, nombre de pièces, nombre de personnes occupant le logement...)/

La disponibilité : la base est disponible facilement, gratuitement et sur toute la France pour le Ministère.

Les principaux inconvénients

Imprécision des données de population : la base FILOCOM n'a pas pour finalité le recensement des populations, elle est adaptée aux besoins spécifiques des services fiscaux. Ainsi, les données de population contenues dans FILOCOM intègrent des double-comptes : par exemple, un étudiant qui est rattaché au foyer de ses parents pour la déclaration d'impôt sur le revenu mais qui loue un studio et se trouve de ce fait assujéti à la taxe d'habitation est comptabilisé deux fois. C'est le cas aussi pour les gens hospitalisés pour une longue durée. Une comparaison réalisée sur Lille par le CETE Nord-Picardie fait état d'un écart de 11% sur les chiffres de population entre FILOCOM et le Recensement de la Population 1999 (il n'est que de 1.6% sur les résidences principales).

Il s'agit d'une base alphanumérique : elle nécessite que l'on dispose du cadastre numérisé pour pouvoir être intégrée dans un SIG. Dans le cas – assez fréquent – où l'on ne dispose pas de cette donnée, il est toujours possible, soit de numériser les sections cadastrales avec leur code (travail fastidieux), soit de passer par l'intermédiaire de la base îlot. En effet, les îlots initialement créés par l'INSEE étaient basés sur les sections cadastrales et le code îlot contient le code de section cadastrale. Une étude sur la compatibilité des deux bases a été réalisée par le CETE Nord-Picardie dans le cadre du programme ACTEUR du CERTU.

Pour de plus amples informations sur la base de données Filocom et ses différentes variables, on pourra s'adresser à la Direction des Affaires Économiques Internationales (DAEI) du Ministère de l'Équipement.



► La BDParcellaire

La BDParcellaire est un assemblage par département de toutes les planches du cadastre. Ce produit est fabriqué par l'IGN à partir de données cartographiques de la DGI pour la réalisation du Référentiel à Grande Échelle (RGE).

La BDParcellaire existe sous forme vecteur là où existe un Plan Cadastral Informatisé (PCI) et sous forme raster sinon.

Quelle que soit la forme, les localisants de parcelle sont fournis séparément du fond de plan, sous la forme d'un ensemble d'objets ponctuels.

La BDParcellaire fournit également de manière vectorielle les contours des sections cadastrales. C'est la couche des sections cadastrales qui a été employée dans la présente fiche.

► Le test effectué

Le secteur de test porte sur 3 communes de Seine-et-Marne (77), Saint-Fargeau-Ponthierry, Seine-Port et Nandy, communes pour lesquelles il est possible d'obtenir gratuitement un jeu test de la BD Parcellaire.

Extraction des données

L'extraction des données Filocom est réalisée sur demande par le CETE Nord-Picardie, qui produit un fichier .rtf (Rich Text Format) contenant le code de la section cadastrale ainsi que les données commandées.

Le fichier .rtf fournit donc les données pour chaque section cadastrale, mais aussi un regroupement communal des données provenant de sections dont le seuil de secret statistique ne permet pas de divulguer l'information (sections dont le nombre d'enregistrements est inférieur à 11).

Exemple : passage de « 77326 AB » à « 77326AB »

Traitement de la BD Parcellaire sous MapInfo

En parallèle, il est nécessaire de préparer la BD Parcellaire afin de pouvoir la joindre au fichier Filocom provenant d'Excel.

En effet, la BD Parcellaire contient un fichier « division parcellaire », qui ne contient pas le code de la section cadastrale dans un champ unique, mais déconcaténé au sein de trois champs distincts.

Exemple : CODDEP= « 77 », CODCOM= « 326 », CODSECT= « AB »

Là aussi, des espaces peuvent se rencontrer dans certains champs (CODSECT en particulier). Avant de concaténer les champs, il est donc nécessaire de supprimer ces espaces, par exemple en faisant un Ltrim\$ sur les champs concernés.

Il est ensuite nécessaire de re-crée un champ « section cadastrale », que l'on va renseigner par une simple mise à jour de colonne dans MapInfo : CODDEP+CODCOM+CODSECT

Également, il apparaît que la table « division parcellaire » de la BD Parcellaire est composée en fait de sous-ensembles des sections cadastrales (les feuilles cadastrales).

Avant de joindre le fichier Excel et la table MapInfo, il est donc nécessaire d'opérer un regroupement de ces feuilles cadastrales en fusionnant les objets dans MapInfo et en les regroupant par la colonne « section cadastrale » créée précédemment.

On obtient ainsi un nouveau fichier que l'on peut nommer « sections cadastrales » et qui va servir à réaliser la jointure avec les données Filocom provenant d'Excel.

Conditions de la jointure

La réunion des deux fichiers ne pose pas de problème spécifique. Une attention doit seulement être portée sur deux éléments :

les informations regroupées à l'échelle communale pour des raisons statistiques dans le fichier source n'apparaissent plus à ce stade. Il conviendra donc d'en tenir compte si des analyses à l'échelle communale sont requises ;

le fichier source Filocom ne fournit des informations que pour les sections cadastrales ayant une population et pour lesquelles le secret statistique le

permet. Lors de la jointure sous MapInfo, certaines sections ne possèdent donc pas de valeurs, sans qu'il soit possible de déterminer si cela est dû au secret statistique ou tout simplement à l'absence d'information (zone industrielle ou zone naturelle

données à la section cadastrale

Regroupement communal des sections dont la population est inférieure au seuil statistique (<11 habitants)

Section cadastrale

	nombre de personne	type logement		statut occupation			total
		logement collectif	logement individuel	autre	locataire	propriétaire occupant	
77326	13	12
77326 AB	1217	156	195	22	182	147	351
77326 AC	1513	376	187	30	296	237	563
77326 AD	1225	0	358	11	66	281	358
77326 AE	820	0	228	.	92	132	228

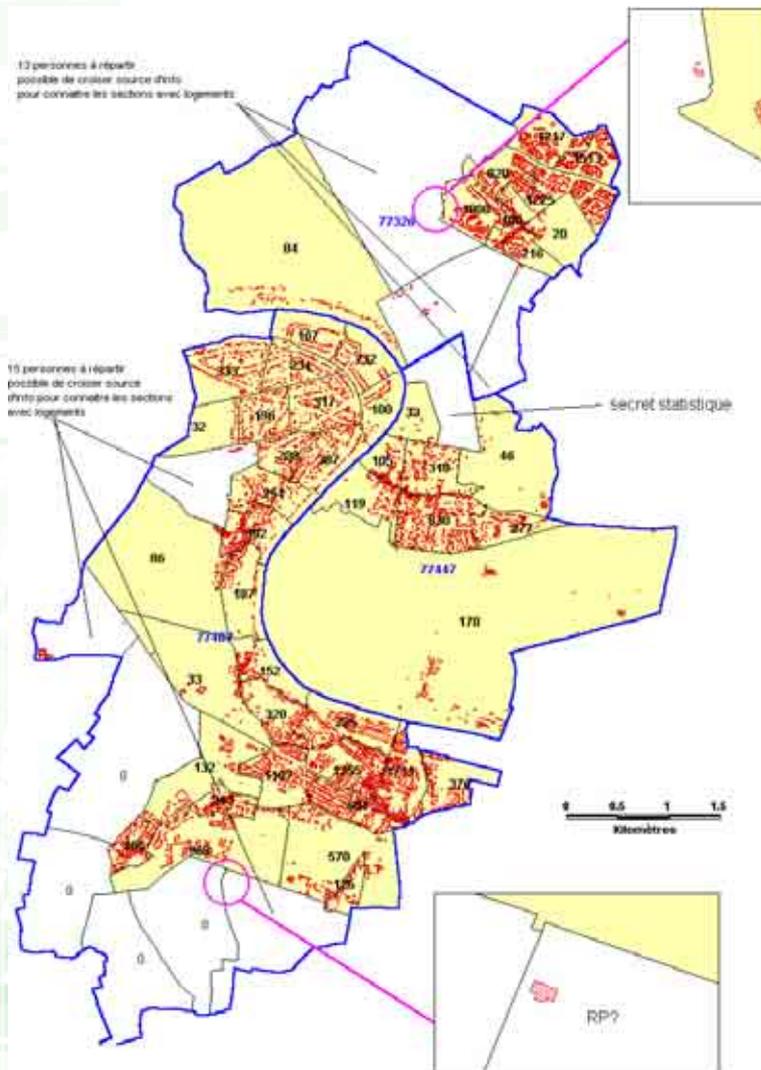
Un exemple d'un extrait Filocom

Traitement des données sous Excel

Par un simple copié-collé, les données sont importées dans Excel. La seule opération nécessaire à ce stade consiste dans la suppression des espaces contenus dans le champ de la section cadastrale.

par exemple). Seul un croisement avec d'autres sources d'information (occupation du sol, couche bâtie de la BD Topo, etc.) peut éventuellement permettre de lever cette incertitude. (Cf carte suivante).

Malgré tout, il convient de rester prudent lors des



croisements d'information, car la possibilité de connaître les sections avec logements et donc avec population repose sur la qualité de la donnée relevant de la couche bâtie.

Or, dans l'exemple traité ici, c'est le bâti provenant de la BD Parcellaire qui a été utilisé, sans qu'il soit possible de déterminer la date de sa numérisation. Cette information doit normalement faire partie des métadonnées accompagnant la livraison de la BD Parcellaire (selon le descriptif du contenu), ce qui n'a pas pu être vérifié dans le cadre du jeux-tests de l'IGN. De plus, la couche bâtie de la BD Parcellaire distingue uniquement les bâtiments en dur des constructions légères, ce qui ne permet donc pas de distinguer les bâtiments relevant de l'habitat de ceux relevant des activités économiques.

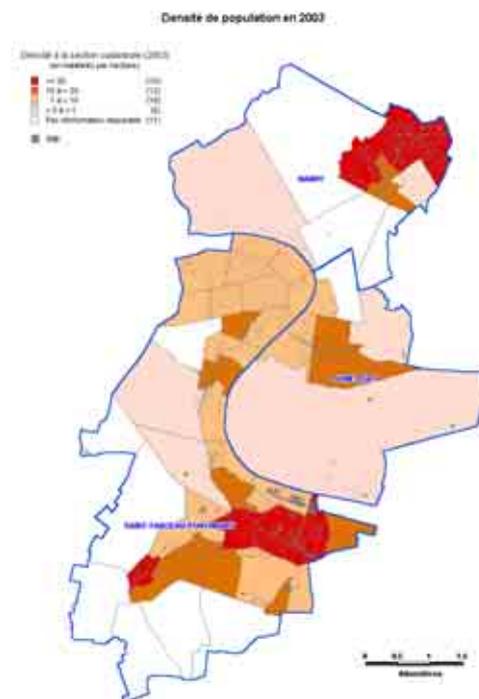
Par conséquent, mis à part certains cas précis comme la commune 77447 possédant une seule section concernée, la distinction entre sections cadastrales soumises au secret statistique et sections

n'ayant pas d'habitants reste aléatoire. Cette situation pose problème à la fois en termes d'analyse infra-communale et de représentation cartographique.

A l'échelle des trois communes étudiées, sur 57 sections cadastrales, seules 46 sections fournissent une donnée de population ; il reste donc 11 sections concernées par cette incertitude, et 10 une fois le cas de la section de la commune 77447 réglé (soit 17,5% du nombre total de sections, mais seulement 0,15% de la population totale du secteur d'étude).

Représentation cartographique

Dans le cadre de cet exemple, seule une analyse de la densité de population en 2003 a été menée. Dans la carte suivante, les sections soumises au secret statistique et celles ne possédant pas de logements (et donc d'habitants) sont regroupées sous l'expression « Pas d'information disponible ».



► Conclusion

L'analyse cartographique de données de population provenant de Filocom pose quelques problèmes liés aux limites du secret statistique. Il peut en effet parfois être difficile de distinguer sections sans population et sections soumises au secret statistique.

Lorsqu'on s'intéresse à d'autres données que la seule population, comme la répartition entre logements individuels et collectifs par exemple, le problème devient encore plus aigu.

De manière générale, toute donnée plus spécifique que la seule population pose rapidement la question du seuil statistique en dessous duquel les données manquantes peuvent devenir nombreuses.

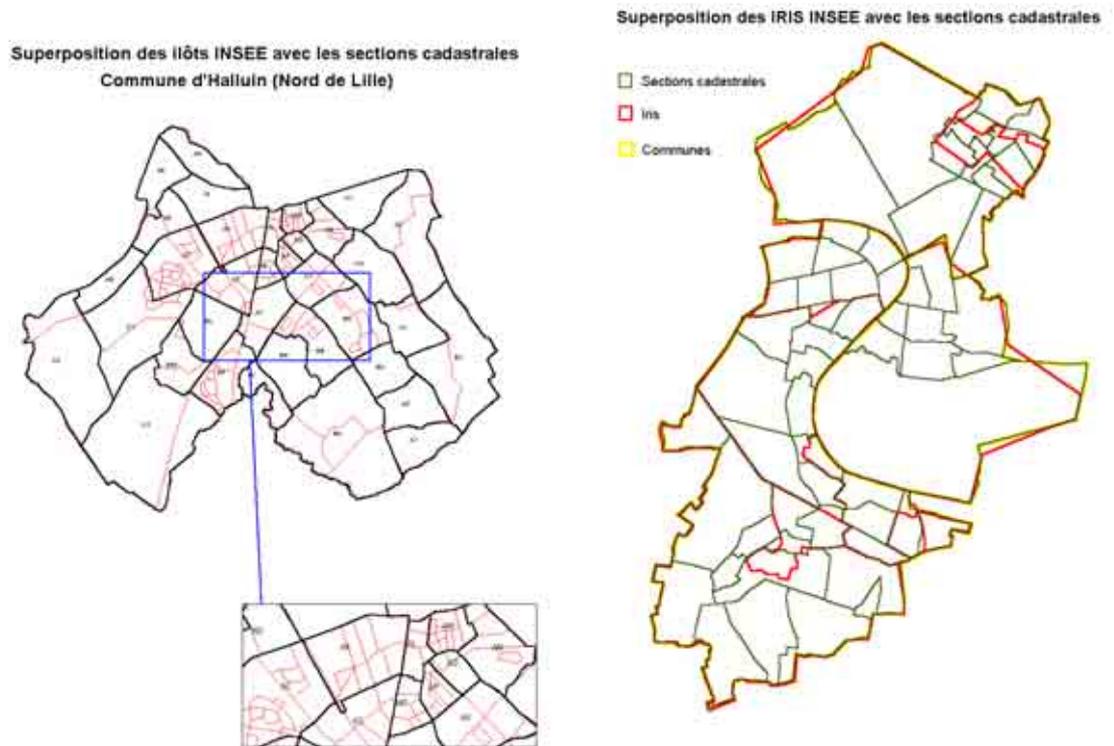
En effet, sur les 57 sections cadastrales étudiées, il apparaît que 18 sections sont concernées par le secret statistique concernant le logement individuel ou collectif (nombre de logements inférieur à 11), ce qui représente quand même plus de 30% des sections pour lesquelles il n'est pas possible de connaître la répartition exacte entre logements individuels et collectifs.

Malgré tout, il semble nécessaire de relativiser ce phénomène. Tout d'abord, nous avons vu que les poids de population concernés restent très faibles, du moins dans le cadre de l'exemple traité. Également, il est possible de considérer qu'en milieu urbain dense, les problèmes de secret statistique se posent moins que dans le cas étudié ici. Ensuite, selon le type d'analyse requis, il est toujours possible de demander, afin de limiter ce problème, une extraction de données à une échelle moins fine que celle de la section. Les regroupements de sections cadastrales permettent en effet de lever le secret statistique dans de nombreux cas, moyennant une analyse *a priori* du territoire d'étude.

Enfin, dans l'hypothèse où un îlotage des données Filocom serait réalisé afin de les utiliser comme source d'information à l'îlot sur la population, il paraît opportun de s'intéresser aux recouvrements permis entre BD Parcellaire et BD Îlots.

Dans le cadre du jeu-test de l'IGN, il n'est pas possible d'obtenir gratuitement les îlots sur la zone d'étude. Aussi, le recouvrement a été seulement observé sur une commune du nord de Lille (carte suivante). Il apparaît que dans l'ensemble ce recouvrement s'opère dans de bonnes conditions, même si certaines corrections topologiques peuvent être requises.

Par contre, le recouvrement entre base IRIS et sections cadastrales semble nettement plus aléatoire alors que les IRIS sont pourtant des regroupements d'îlots.



Note rédigée par François-Xavier Robin et Vincent Demeules - CETE Normandie-Centre
Francois-xavier.robins@equipement.gouv.fr
Vincent.demeules@equipement.gouv.fr

Relue et complétée par Magali Di Salvo -CERTU
Magali.di-salvo@equipement.gouv.fr

Cette fiche a été réalisée dans le cadre du rapport du CETE de Rouen sur les méthodes d'estimation de population dans des zonages thématiques

HERA : Décompte des Habitants et des Résidences autour des Aéroports

La méthode de l'IGN

L'objectif du Service Aménagement-Environnement de l'IGN est la mise au point de méthodologies à valeur ajoutée à partir des bases de données de l'IGN pour favoriser leur utilisation, en particulier la BDTopo®.

Cette fiche a été réalisée à partir de documents de travail de l'IGN. Les différentes illustrations en sont notamment issues.

La rédaction de cette fiche date de 2003.

► Contexte

Le STBA a passé commande auprès de l'IGN l'élaboration d'un outil permettant de quantifier le nombre d'habitants et de logements dans le cadre des 2 documents réglementaires concernant le bruit des aérodromes : le Plan de Gêne Sonore (PGS) et le Plan d'Exposition au Bruit (PEB).

Le travail réalisé par l'IGN pour cette commande a donné lieu à l'élaboration du logiciel HERA (Habitants Et Résidences autour des Aéroports), utilisable avec le SIG MapInfo®.

La méthode consiste à évaluer le nombre d'habitants et de résidences pour chaque bâtiment de type "habitation" de la base de données BD TOPO®, en utilisant les résultats du Recensement Général de la Population 1999 (RGP) et la Base-îlots® de l'INSEE.

La démarche employée par l'IGN pour élaborer le logiciel HERA s'est inspirée des résultats obtenus lors de la mise au point d'une méthodologie analogue effectuée pour l'Agence de l'Environnement et de la Maîtrise de l'Énergie (ADEME) en 2000.

Ce travail permet au final d'estimer le montant des indemnités pour insonoriser les bâtiments dont la date de construction est antérieure à celle du PEB lors de la création d'un aéroport

► Description de la méthode

La démarche suivie peut être résumée comme ceci :

La hauteur des bâtiments est estimée grâce au Modèle Numérique de Terrain (MNT) de la BD TOPO® ainsi que le nombre d'étages qui en est directement déduit, ensuite les nombres d'habitants et de résidences sont évalués au prorata de la surface développée (surface x nombre d'étages) de chaque bâtiment.

La répartition des habitants et résidences des îlots ou communes sur les bâtiments introduit automatiquement une imprécision au niveau de chaque bâtiment. Cependant la somme des nombres d'habitants et de résidences par zone PEB lisse cette erreur que des comptages terrain ont permis de vérifier

Les données

En France métropolitaine, toutes les données géographiques en entrée sont dans la projection Lambert II étendu, et au format MIF/MID (format propriétaire MapInfo®), sauf pour le MNT qui est au format ASCII X,Y,Z

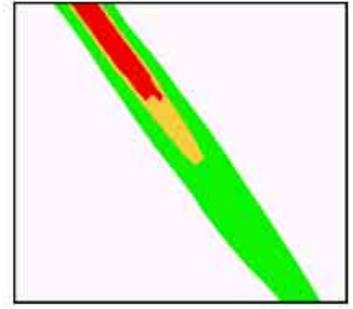
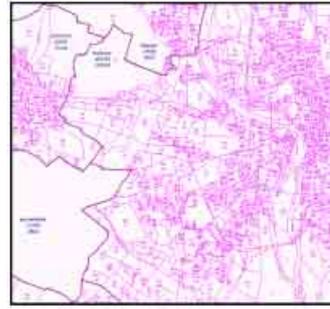
Population et logement

Les données de population et de logements prises en compte sont celles fournies par l'INSEE. La base de données utilisée est la Base-îlots® issue du dernier recensement (RGP99).

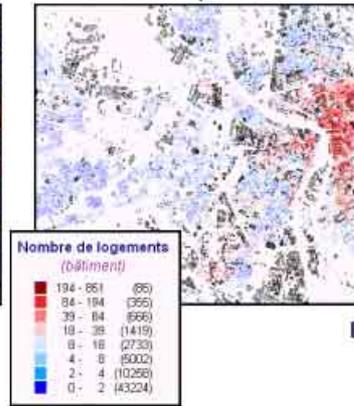
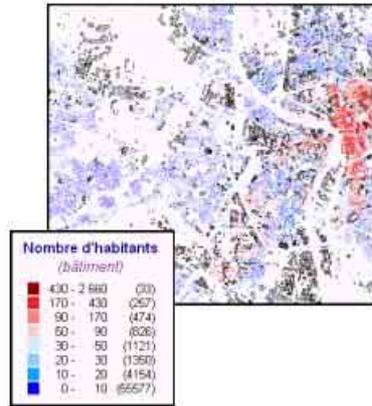


BD TOPO : bâtiments, MNT Communes et îlots INSEE

Courbes de bruit PEB



HERA



Bâtiments « d'habitation » au sein des PEB

Les données entrées dans le logiciel

Les données retenues sont :

- ♦ La population totale sans double compte de l'îlot (ou de la commune s'il n'y a pas d'îlot) pour le nombre d'habitants
- ♦ Le nombre de résidences principales de l'îlot (ou de la commune) pour le nombre de logements.

Occupation des sols

Les données d'occupation du sol et de topographie employées sont :

- ♦ la couche « limites communales » de la BD TOPO®
 - ♦ la couche bâtiments (objets surfaciques) de la BD TOPO®
 - ♦ la couche « enceintes » de la BD TOPO®
- le MNT BD TOPO®

Données sur les zones de bruit

Le STBA fournit les zones PEB autour des aéroports.

Les outils

Le logiciel HERA a été développé et fonctionne sous l'environnement MapInfo®.

Intégration des données

Cette étape consiste à intégrer les différentes

données pour qu'elles puissent être traitées sous MapInfo® par le logiciel HERA. Elle peut se décomposer en trois phases :

Import : les données sont transférées du format MIF/MID ou ASCII (pour le MNT) vers le format TAB de MapInfo®.

Fusion : les données de même nature (éléments bâtis appartenant à différentes planches de la BD TOPO®, îlots de communes différentes...) sont regroupées au sein d'une même table

Typage des bâtiments : les objets surfaciques de la classe « bâtiments quelconques » de la BD TOPO® sur lesquels sera faite la répartition de la population et des logements sont mis en évidence. Les bâtiments entièrement inclus dans un objet « enceinte » - qui peut être commerciale, industrielle, publique - sont exclus .

Correction des données géométriques

Trois types de corrections sont nécessaires afin d'améliorer les résultats (cette partie du traitement est la plus longue car les corrections doivent être faites à la main par un opérateur) :

Les limites communales : il s'agit de « fermer les surfaces » relatives aux limites communales et les mettre en cohérence lorsqu'elles ne le sont pas, en particulier lorsqu'elles sont issues de planches différentes de la BD TOPO®.

Les limites de la base îlots : l'importance des

corrections géométriques à apporter aux contours des îlots dépend de leur niveau de cohérence géométrique avec la BD TOPO®. En effet, les îlots, lors de leur saisie, sont calés soit sur GEOROUTE®, soit sur le SCAN 25®, soit sur des données cadastrales. Cette diversité des données de calage conduit à des résultats très disparates qui nécessitent de mettre en œuvre des corrections plus ou moins lourdes.

Découpe des bâtiments : afin d'optimiser les résultats pour la répartition du nombre d'habitants et de logements, il faut couper les bâtiments qui sont à cheval sur plusieurs îlots.

Affectation des données des entités administratives aux bâtiments

Après avoir affecté aux communes et aux îlots la surface totale, la surface dans l'extrait, la proportion de la surface dans l'extrait par rapport à la surface totale, les nombres d'habitants et de résidences, chaque bâtiment d'habitation est alors affecté des mêmes renseignements.

Remarque : lorsque les données à l'îlot ne sont pas disponibles, c'est la donnée à l'échelle de la commune qui est retenue pour le calcul

Calcul des données relatives aux bâtiments

la hauteur des bâtiments est déterminée par soustraction de l'altitude moyenne du bâtiment et de l'altitude moyenne au sol (calculée en faisant la moyenne des points altimétriques du MNT situés dans un buffer de 25 m autour du bâtiment concerné)

- ♦ **le nombre d'étages**, 1 niveau si la hauteur est

inférieure à 5 m, au delà, chaque étage correspond à 2,5 m

- ♦ **la surface au sol** et **la surface développée** (surface au sol x nombre d'étages).

A chaque îlot est associé la somme des surfaces développées des bâtiments « d'habitation » qu'il contient

Pour chaque bâtiment, le nombre d'habitants et de résidences est déduit du nombre total au niveau de l'îlot, au prorata de sa surface développée

Synthèse des résultats pour chaque PEB

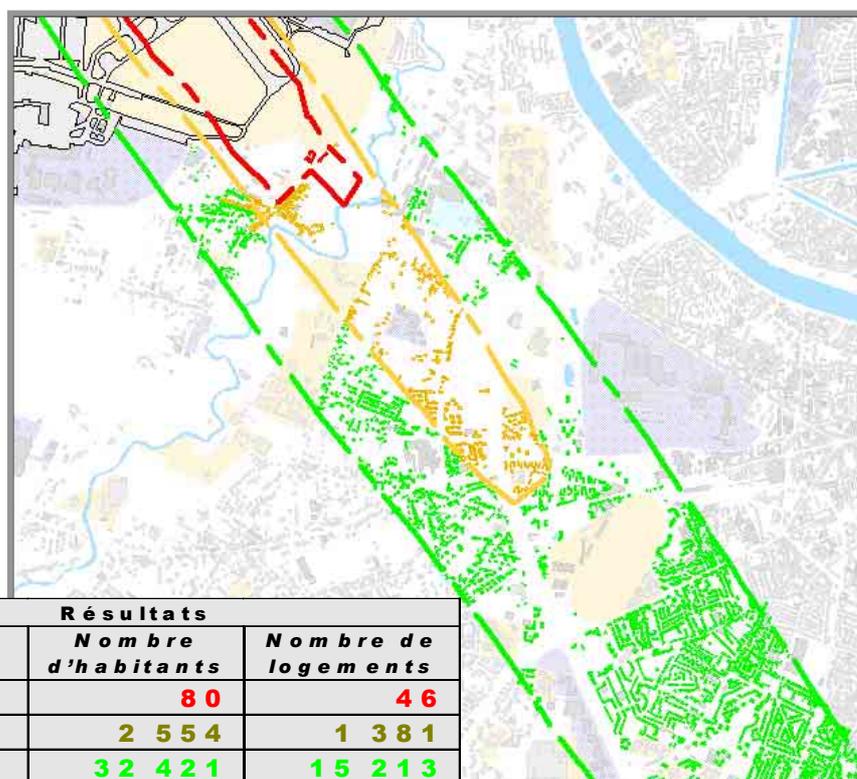
Chaque bâtiment est affecté à la zone PEB dans laquelle il se trouve. Le calcul des habitants et du nombre de résidences par PEB est alors immédiat.

► Intérêt de la méthode

La méthode présente les avantages suivants :

- ♦ Elle utilise des bases de données précises, la Base-îlots® et la BD TOPO®
- ♦ Elle est disponible sous forme d'un outil fini, le logiciel HERA
- ♦ Elle répartit la population et les logements du recensement sur chaque bâtiment au prorata de leur « volume », ce qui est plus précis que d'utiliser une base de données en 2D de type occupation du sol

La principale difficulté relative à la bonne utilisation du logiciel est mise en cohérence de la Base-îlots® et de la BD TOPO®.



Un exemple de résultat du calcul

Évaluation de la méthode

Plusieurs sources d'erreur risquent de rendre imprécise l'affectation du nombre d'habitants et de résidences à chaque bâtiment.

Les principales sont :

- ◆ la différence de date entre les prises de vue aériennes pour générer les données BD TOPO® et le dernier recensement de l'INSEE (1999). Des bâtiments nouvellement construits ou détruits durant la période donnent un excès ou un déficit d'habitants et de résidences au niveau des bâtiments de la BD TOPO®
- ◆ la résolution du MNT BD TOPO® pour le calcul des hauteurs de bâtiments, surtout ceux qui sont construits sur des terrains en pente ou ceux dont la hauteur n'est pas homogène
- ◆ la précision et l'exhaustivité de la BD TOPO®. En particulier, les objets surfaciques « bâtiment quelconque » ne correspondent pas tous à du bâti résidentiel (bureau, magasin et rez-de-chaussée, ...)
- ◆ précision et exhaustivité des données INSEE
- ◆ non prise en compte des logements et de la population présents dans les enceintes (établissements d'enseignement ou sanitaires, usines), dans les résidences secondaires, temporaires et non fixes

Toutefois, d'après les constats de l'IGN, le fait d'agréger les résultats à l'échelle d'un PEB permet de lisser ces différentes imprécisions que l'on retrouve au niveau du bâtiment. Deux comparaisons avec des données terrain ont montré un écart inférieurs à 10%. Ce chiffre n'est qu'un ordre de grandeur car il n'a pas fait l'objet d'une évaluation systématique (centre ville, zone péri-urbaine, plaine, montagne etc...).

Présentation de l'équipe

Organisme : IGN aménagement environnement

Personnes impliquées

- ◆ Philippe GERBE
- ◆ Laurent QUEYROU
- ◆ Mélanie TRAN
- ◆ Olivier BARASZ
- ◆ Sylvain TERRADE

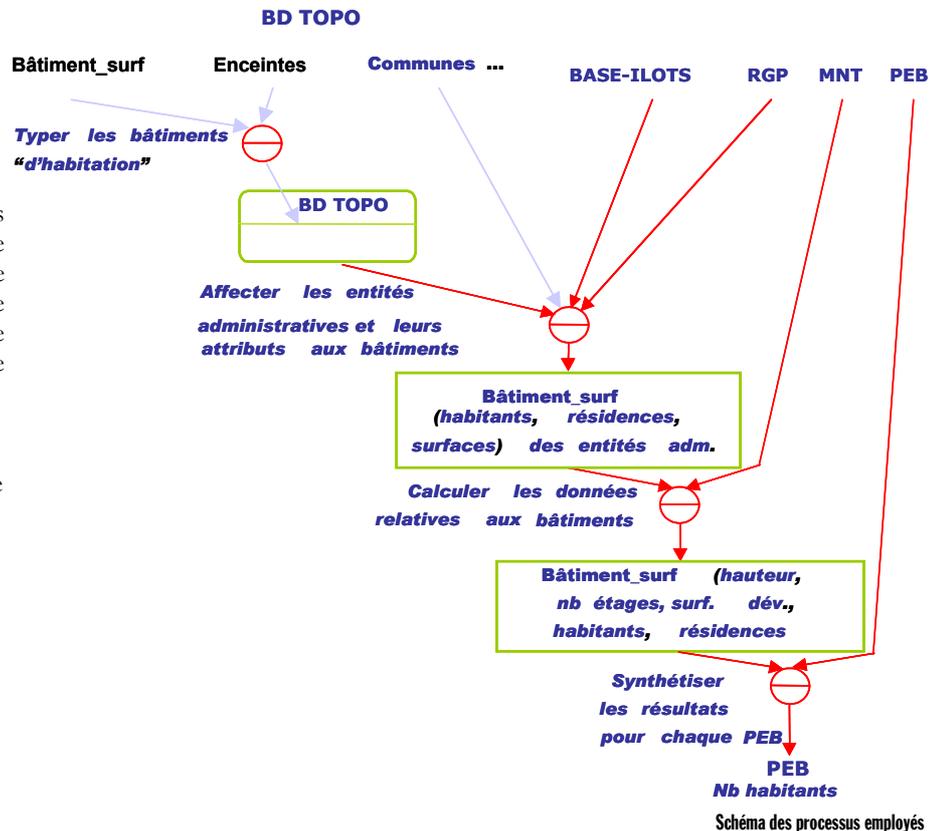


Schéma des processus employés

Note rédigée par Vincent Demeules et Vincent.demeules@equipement.gouv.fr
Sophie Martinoni - CETE Normandie-Centre

Relue par Magali Di Salvo - CERTU

Magali.di-salvo@equipement.gouv.fr

Cette fiche a été réalisée dans le cadre du rapport du CETE de Rouen sur les méthodes d'estimation de population dans des zonages thématiques

CROISEMENT DE DONNEES DEMOGRAPHIQUES AVEC UNE TACHE URBAINE AFFINEE

La méthode de la DDE31

Historiquement, la DDE 31 s'est beaucoup impliquée dans le traitement d'image d'où une utilisation de la télédétection pour l'estimation des populations.

Le savoir-faire et la détention d'outils adaptés à la télédétection (logiciel ER-MAPPER) a implicitement guidé l'équipe vers une méthode utilisant le traitement d'image.

La rédaction de cette fiche date de 2003.

► Contexte

L'équipe a été conduite à s'interroger sur le décompte de la population en 1999, avant la publication des résultats du Recensement Général de la Population (RGP), suite à une demande du préfet relative à une étude du STBA. Il s'agissait (comme pour l'IGN) d'estimer les personnes concernées par un plan d'exposition au bruit, à proximité de l'aéroport de Toulouse .

La méthode utilisée à l'époque consistait en une utilisation directe des données INSEE : la population a été directement estimée à partir des données à l'îlot datant de 1990 (la population concernée par le Plan d'Exposition au Bruit est évaluée au prorata de la surface de l'îlot incluse dans l'emprunte sonore du PEB).

Pour prendre en compte les évolutions de la population entre 1990 et 1999, les zones construites entre ces deux dates ont été repérées par zone de bruit. En faisant l'hypothèse que les constructions nouvelles avaient les mêmes caractéristiques que celles de 1990, il était possible de proposer une estimation de la variation de population entre 1990 et 1999

Après la publication des résultats de l'INSEE, toujours pour répondre à des demandes du même ordre (population exposée au bruit), l'équipe a cherché à affiner sa méthode en tenant compte de la localisation plus fine des zones bâties au sein des nouvelles entités INSEE (îlots, IRIS voire commune) ; l'objectif étant une meilleure connaissance de la répartition géographique des populations

L'idée principale qui guide la démarche est l'obtention d'une tâche urbaine à vocation résidentielle la plus précise possible

Dans un premier temps, l'équipe a utilisé la couche de bâti issue de SPOT THEMA comme tâche urbaine. Cela permettait de délimiter les espaces peuplés à l'intérieur des îlots (ou entités INSEE selon la disponibilité de la donnée) la population étant supposée répartie de manière homogène dans ces espaces. SPOT THEMA présentait également l'avantage de distinguer les grandes zones d'activités des autres zones bâties.

En revanche, SPOT THEMA n'était pas tout à fait adaptée à la démarche dans la mesure où le bâti est mal pris en compte dans les zones peu denses (bâtiment isolé occulté et contours mal définis pour un petit groupe d'habitations).

Cette mauvaise prise en compte du bâti diffus a conduit l'équipe à déterminer une nouvelle tâche urbaine à partir d'images satellitaires de résolution plus fine

► Description de la méthode

La méthode consiste en un croisement d'une tâche urbaine à vocation résidentielle avec des données démographiques à différentes échelles (îlots, IRIS, communes) ; la population résidente étant supposée homogène sur la tâche urbaine dans chaque entité INSEE.



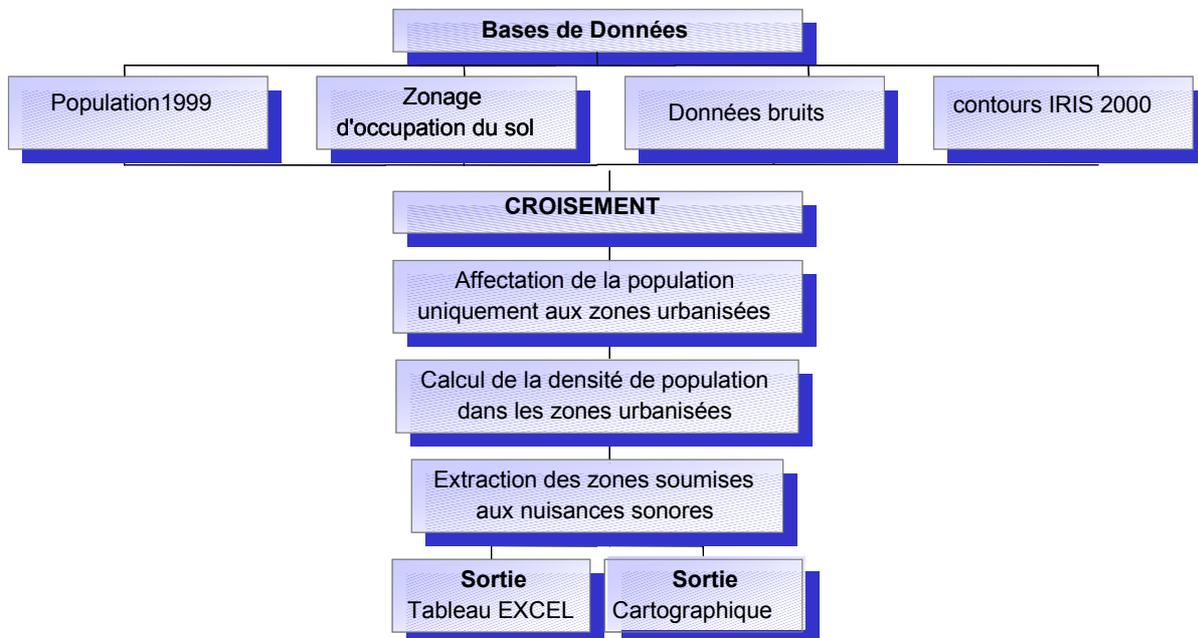


Schéma de la méthode

Les données

Données de population

- ◆ base îlots
- ◆ base IRIS
- ◆ données à la commune

Ces bases de données sont produites par l'INSEE à partir du dernier recensement de la population.

Le découpage à l'îlot est le plus précis mais il n'est disponible que pour les communes de plus de 10000 habitants.

Les trois bases sont donc utilisées pour couvrir tout le département de la Haute-Garonne en fonction de leur disponibilité.

Données d'occupation du sol

- ◆ images satellitaires IRS (résolution 5 m au sol, couverture départementale). Ce ne sont pas les images multi-bandes qui sont utilisées mais des images traitées pour se rapprocher d'une représentation en couleurs réelles. La méthode reste valable avec des photographies aériennes orthorectifiées
- ◆ atlas numérisé des zones d'activités
- ◆ couverture aérienne à différentes dates orthorectifiée par l'équipe
- ◆ SITADEL (pour les permis de construire, géoréférencés par l'équipe sur certaines zones particulières)
- ◆ Plans d'Occupation du Sol numérisés et géoréférencés

Les outils

- ◆ logiciel ER MAPPER
- ◆ logiciel MAPINFO

CETE Normandie-Centre - CERTU
Méthode d'estimation de population

► La mise en oeuvre

Constitution de la tâche urbaine à vocation résidentielle .

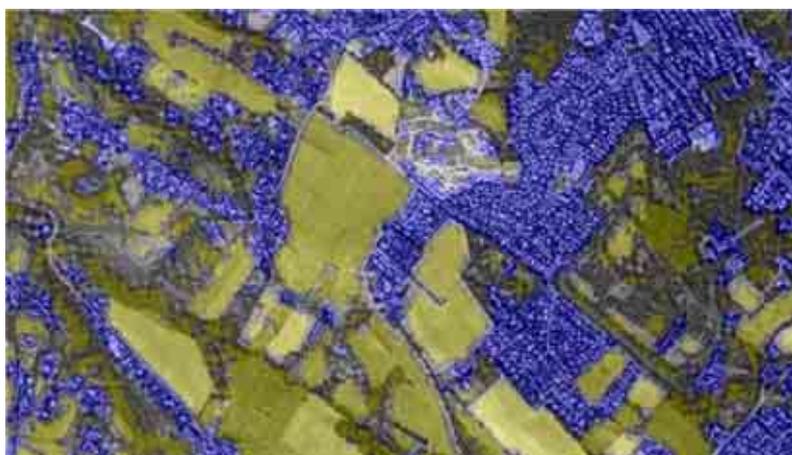
Sous ER MAPPER, un filtre de texture est appliqué qui permet de repérer les variations de textures importantes sur l'image IRS.

Ce filtre fait l'objet de réglages adaptés à l'image, à la zone d'étude et à l'échelle considérée. Il est déterminé de proche en proche par l'utilisateur.

L'application du filtre permet d'isoler le bâti mais isole aussi des éléments non bâtis comme des haies ou des parkings occupés.

Ces éléments sont exclus de la tâche urbaine à vocation résidentielle par un traitement complémentaire.

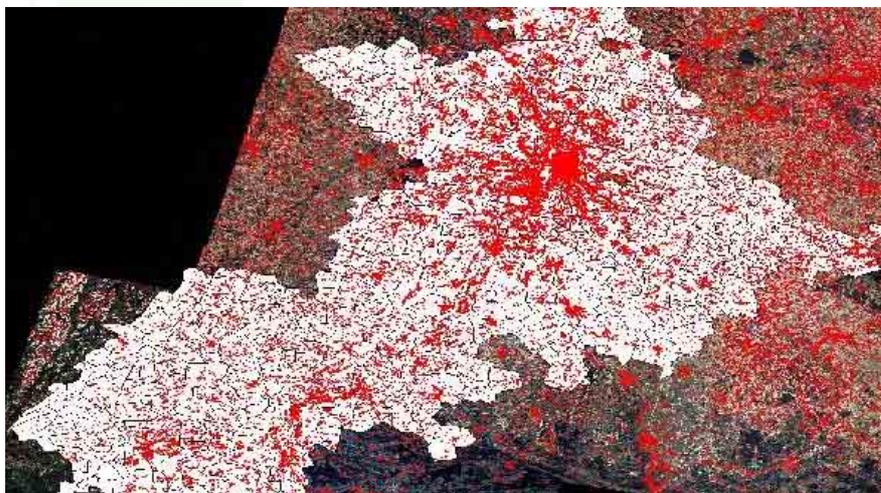
Le travail sur la texture permet également d'éliminer les grandes surfaces de toit qui correspondent généralement à un bâti commercial ou industriel (usines, supermarchés,...).



Exemple de la tâche urbaine obtenue

Les zones d'activités peuvent être exclues comme cela a priori, ou a posteriori, en utilisant un fichier géoréférencé des zones d'activité (Atlas local des zones d'activités ou POS/PLU numérisés).

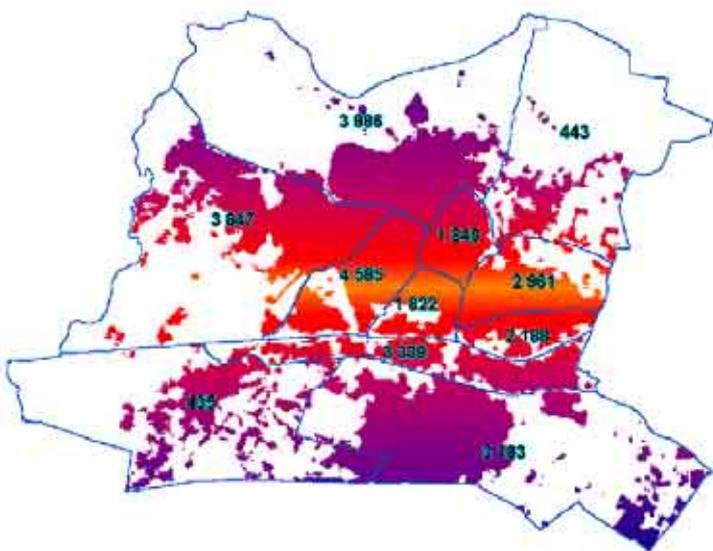
En utilisant une classification automatique, on obtient une couche raster du bâti.



Tache urbaine obtenue à partir de l'image satellite représentée à l'échelle du département



Représentation de la tache bâtie (hachuré jaune) sur les îlots INSEE (contour rouge)



Estimation de la population touchée par le bruit

L'étape suivante consiste à vectoriser la couche bâtie ainsi obtenue, sous ER MAPPER.

La généralisation (ou lissage des contours) s'effectue sous MAPINFO. Le fichier généré sous ER MAPPER étant très important (nombreux points générés lors de la vectorisation), la généralisation nécessite souvent un morcellement du fichier de base.

Au final, la tache urbaine à vocation résidentielle est produite sous la forme d'un fichier .TAB de MAPINFO

Croisement avec les données de population

Une fois la tâche urbaine créée, la méthode consiste à effectuer un croisement sous MAPINFO entre celle-ci et les données de population : la population de l'entité INSEE (îlot, IRIS ou commune) est affectée à la surface de l'entité comprise dans la tâche urbaine.

La localisation de la population est ainsi affinée au sein des entités INSEE.

Estimation de la population

Pour estimer une population dans une zone donnée, par exemple une zone de risque ou une zone exposée à un certain niveau de pollution, il suffit alors de croiser la donnée de population localisée plus précisément au sein des entités INSEE (IRIS dans l'illustration ci-dessous) avec la zone considérée (des zones de bruit dans l'illustration ci-dessous)

Les partenaires associés à la démarche

Pour la mise au point de la tâche urbaine à vocation résidentielle et la détermination des filtres d'analyse, la DDE 31 a travaillé en partenariat avec SCOT (Services et Conception de systèmes en Observation de la Terre) qui est une filiale de recherche du CNES.

► Autre(s) application(s)

La détermination d'une tâche urbaine à vocation résidentielle peut servir d'autres applications

La représentation cartographique

La localisation affinée du bâti résidentiel au sein des entités INSEE peut améliorer la représentation cartographique de certaines données relatives à la population et au logement : elle permet de recentrer les centroïdes des entités INSEE sur le bâti résidentiel. (Exemple : Atlas du logement social)

Le suivi de l'urbanisation

La détermination des tâches urbaine à vocation résidentielle à différents horizons offre des possibilités de suivi de l'urbanisation, de la consommation de l'espace, ou de contrôle du droit à construction des communes.

MAPBRUIT

La DDE 31 a déterminé les zones de bruit critiques (ZBC) en s'aidant de la tâche urbaine à vocation résidentielle : 905 ZBC ont ainsi été définies le long des axes routiers concernés par le bruit.

► Intérêts de la méthode – Difficultés rencontrées

La méthode présente deux principaux avantages :

- ♦ elle assure une prise en compte du bâti dispersé ou isolé en zone rural, ce que ne permettent pas les couches bâties issues de SPOT THEMA ou de la BD CARTO
- ♦ la tâche urbaine à vocation résidentielle peut être actualisée très facilement puisqu'elle repose sur l'exploitation de photographies aériennes ou d'images satellites. A ce propos il est intéressant de souligner que la DDE 31 s'est aussi orientée vers cette méthode considérant que la BD TOPO était trop ancienne sur son département.

Les difficultés qu'elle soulève sont les suivantes :

- ♦ la mise au point de la méthode est longue. La détermination des filtres sous ER-MAPPER se fait de proche en proche ; elle est très dépendante de l'expérience de l'utilisateur. Aujourd'hui, avec la pratique, la détermination du filtre nécessite environ une demi-heure à l'échelle d'une commune
- ♦ La principale faiblesse de la méthode est inhérente aux données de population : il s'agit de l'insuffisance de réactualisation du recensement INSEE. Le pas de temps qui sépare deux recensements est trop important pour garantir la validité des données de population après que quelques années se soient écoulées entre le dernier recensement et la date d'exploitation des données

Pour l'équipe de la DDE 31 une façon de pallier l'absence de réactualisation plus fréquente est d'utiliser d'autres sources d'information comme la base de données SITADEL dans laquelle sont recensés et localisés les permis de construire. Cela permet de compléter les données de population en intégrant les habitants des nouveaux logements (la donnée de population est affectée en supposant que les nouvelles constructions possèdent les mêmes caractéristiques que les autres logements au sein d'une même entité INSEE).

Ce palliatif demeure toutefois insuffisant puisqu'il n'intègre pas certaines dynamiques urbaines comme

l'attraction de tel ou tel quartier ou la vacance des logements dans tel autre : l'augmentation de la tâche urbaine à vocation résidentielle n'est pas forcément un bon indicateur de l'augmentation de la population.

► Évaluation de la méthode

La méthode présentée n'a pas fait l'objet d'une évaluation de terrain, c'est-à-dire que la population estimée n'a pas été comparée à la population existante sur un secteur donné.

En revanche, la détermination de la tâche urbaine a été évaluée à partir de comparaisons avec la couche bâtie de SPOT THEMA. Cette comparaison a mis en avant la plus grande précision de la tâche urbaine déterminée par la DDE 31.

La DDE 31 va participer à un groupe de travail animé par le CNES pour tester SPOT 5. Dans ce cadre, la DDE 31 évaluera l'utilisation de SPOT 5 pour la détermination de la tâche urbaine à vocation résidentielle

► Présentation de l'équipe

Organisme : DDE 31 - Service Urbanisme Aménagement / SIG

Personnes impliquées

- ♦ Mireille MULLER
- ♦ Pascal ASTRUC
- ♦ Michel SEGONDY

Note rédigée par Vincent Demeules et Vincent.demeules@equipement.gouv.fr
Sophie Martinoni - CETE Normandie-Centre

Relu par Magali Di Salvo - CERTU

Magali.di-salvo@equipement.gouv.fr

Cette fiche a été réalisée dans le cadre du rapport du CETE de Rouen sur les méthodes d'estimation de population dans des zonages thématiques

DENSITE DE POPULATION ET MORPHOLOGIE DU BATI

La méthode du CETE Normandie-Centre

L'équipe a travaillé pour le compte du CERTU sur la relation pouvant exister entre densité de population et morphologie du bâti. L'objectif étant d'en déduire une méthode pour décompter des populations. Ce travail n'a pas encore abouti à l'heure actuelle. La rédaction de la fiche date de 2003.

► Contexte

L'équipe s'est intéressée au problème du décompte de population par le biais des études de pollution atmosphérique dans le cadre de projets routiers. Il s'agissait d'estimer des populations exposées à une pollution d'origine routière.

Dans ce cadre, l'étude commandée par le CERTU consistait à mettre en œuvre et tester une méthode qui permettrait d'évaluer simplement - mais précisément - une population dans une zone urbaine donnée par l'intermédiaire de la morphologie du bâti. Le travail a été mené sous forme d'une étude de cas sur les agglomérations rouennaise et lyonnaise (le CETE de LYON a été associé au travail).

Cette étude a fait l'objet d'un rapport d'études CERTU « Densité de population et morphologie du bâti, des perspectives pour l'estimation des populations ? » (Août 2001).

► Description de la méthode

Le but recherché était de mettre en évidence un rapport simple entre densité de population et morphologie de bâti qui aurait constitué une méthode d'estimation des populations relativement aisée. En principe, si un tel lien était avéré, l'estimation des populations reposerait sur l'application de ratios forfaitaires (à déterminer) correspondant aux principaux types de bâtis rencontrés en milieu urbain. Dans cette hypothèse, il suffirait, pour quantifier le nombre de personnes résidant dans un secteur géographique particulier, de découper celui-ci en zones de bâti homogène, de calculer leur surface et de leur affecter le ratio de densité correspondant.

Les données

Données de population

Les données de population et de logements prises en compte sont celles fournies par l'INSEE. La base de données utilisée est la BASE-ILOTS issue du dernier recensement (RGP99).

La donnée retenue est la population totale sans double-compte

Données d'occupation du sol

- ◆ La couche bâtiments (objets surfaciques) de la BD TOPO
- ◆ Le SCAN25 de l'IGN
- ◆ Une couverture aérienne de l'agglomération

Des reconnaissances sur le terrain ont été nécessaires pour compléter ces données

Les outils

Les différents traitements ont été réalisés sous les logiciels Arcview et MapInfo.

► La démarche

Découpage de l'aire d'étude en zones de bâti homogène

L'aire d'étude comprend sept communes (dont Rouen). Elle couvre une zone intégrant des tissus urbains très variés : centre-ville récent et ancien, habitat résidentiel récent et ancien, habitat collectif de différentes tailles, zones d'activité économique, espaces de loisirs

Au total, sept types de bâtis - directement inspirés de la typologie CLUSTERS (Classification for Land Use Statistics : Eurostat Remote Sensing programme) mise en place par Eurostat - ont été définis. Ces morphologies sont les suivantes :

- ◆ type 1 : habitat de centre-ville
- ◆ type 2 : petit collectif (moins de 5 étages)
- ◆ type 3 : grand collectif (5 étages ou plus)
- ◆ type 4 : habitat semi collectif
- ◆ type 5 : habitat individuel
- ◆ type 6 : les zones d'activité industrielles et commerciales ou de services
- ◆ type 7 : les établissements scolaires et de santé qui abritent des populations particulièrement sensibles à la pollution atmosphérique

Le découpage de l'aire d'étude suivant ces sept typologies a été réalisé dans un premier temps à partir de clichés aériens au 1/30 000 réalisés par l'IGN. Ce travail a ensuite été affiné par un repérage sur le terrain pour les zones où les photographies aériennes ne permettaient pas de trancher entre deux types de bâti. Ainsi, une visite sur le terrain s'est souvent avérée nécessaire pour différencier le « petit collectif » du « grand collectif », l'évaluation du nombre d'étages d'un immeuble à partir de photographies aériennes n'étant pas aisée à cette échelle.

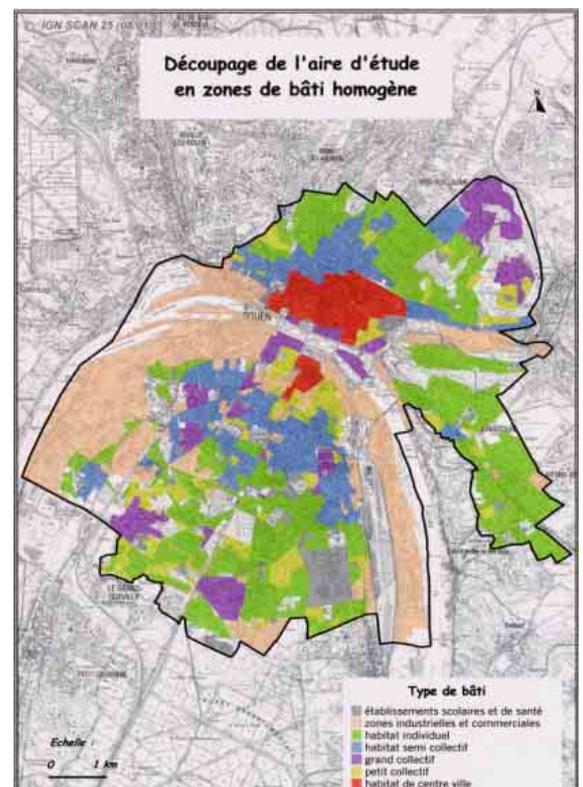
Croisement entre densité d'habitants et types de bâti.

Il s'agit, à partir de la base Îlot, de calculer la densité d'habitants présents dans chaque îlot de bâti homogène et ensuite de comparer les résultats obtenus pour les îlots présentant un même type de bâti

Les principales étapes mises en oeuvre pour aboutir à ce résultat peuvent être résumées de la manière suivante :

- ◆ Le calcul de la densité d'habitants a été réalisé à partir des bases *Îlots...15* et *Contours...Îlots* de l'INSEE et de l'IGN. Le premier travail a été de rechercher tous les îlots compris dans l'aire d'étude : plus de 2200 îlots ont ainsi été sélectionnés sur sept communes. Parmi toutes les données attributaires affectées à chaque îlot, seuls les champs correspondant à l'identifiant, la surface, la commune et la population de l'îlot ont été conservés

- ◆ La seconde étape a consisté en un croisement de la base de données ainsi constituée avec le découpage de l'aire d'étude en zones de bâti homogène afin d'affecter à chaque îlot une catégorie de bâti. A ce stade, un second tri a dû être effectué parmi les îlots retenus. Seuls les îlots présentant un bâti homogène à 90% au moins ont été conservés (c'est-à-dire qu'au moins 90% de la surface de l'îlot doit correspondre à un type de bâti identique sur la base du découpage en zones de bâti homogène). Le principe de cette sélection est exposé dessus.
- ◆ La dernière étape consiste à calculer la densité d'habitants de ces îlots et à comparer les résultats obtenus pour chaque type de bâti. Il s'agit en particulier, suite à une analyse statistique simple d'observer la dispersion des densités calculées pour un même type de bâti .



Découpage de l'aire d'étude en sept types de bâti

Principe du tri des îlots présentant un bâti homogène à 90% au moins

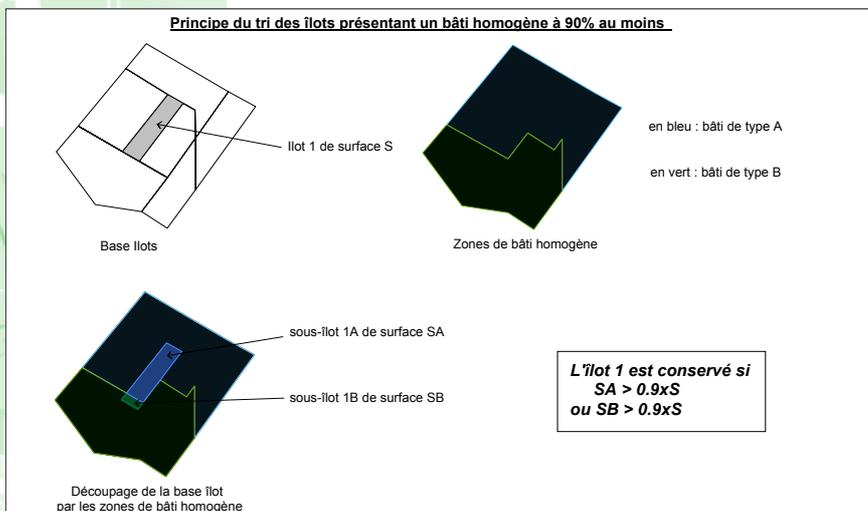


Illustration de la seconde étape

Les résultats

Résultat général

Indépendamment de la typologie de bâti, un premier calcul de densité d'habitants par îlot a été mené afin d'observer la répartition globale de cette densité sur l'aire d'étude.

Résultat par type de bâti

L'objectif est d'essayer de caractériser chacun des types de bâti retenus par sa densité d'habitants. Les îlots exploités sont donc ceux dont le bâti est homogène à 90% soit un total de 1558 îlots (260 centre-ville, 104 petit collectif, 119 grand collectif, 407 semi-collectif, 668 habitat individuel).

- ◆ tous types de bâti confondus, on observe une

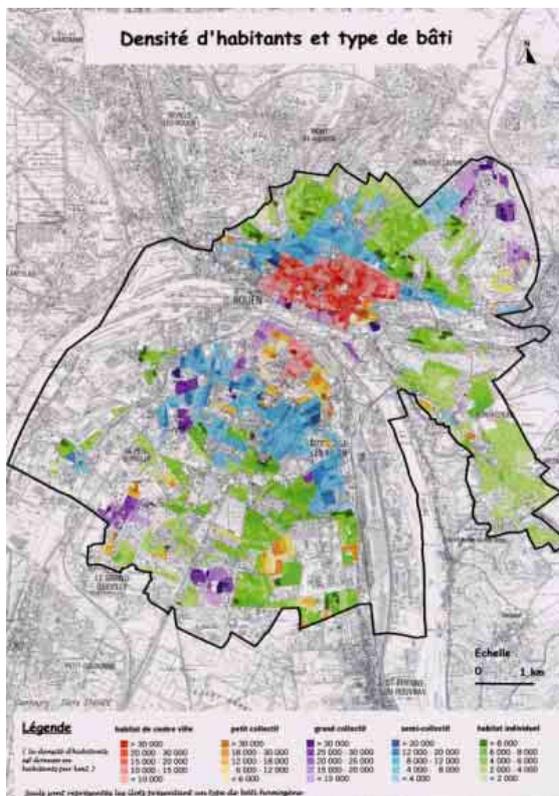
Type de bâti	Centre ville	Petit collectif	Grand collectif	Semi collectif	Individuel
Moyenne	15 325	12 035	21 686	8 087	4 667
Médiane	15 181	10 592	21 043	7 323	4 435
Écart type	6 474	6 910	8 968	3 707	1 740
Maximum	29 897	28 641	48 794	19 494	9 951
Minimum	3 078	2 262	5 836	2 206	1 089
Centile 0.9	24 412	22 644	31 797	13 263	6 904

grande dispersion des résultats (l'écart type est très élevé). Celle-ci est cependant plus importante pour l'habitat collectif (en particulier pour les catégories « petit collectif » et « grand collectif »). Il est probable que la différence du nombre d'étages des bâtiments au sein d'une même catégorie soit à

- l'origine de la divergence des valeurs observées
- malgré la forte dispersion des résultats, on constate une différence assez nette entre la densité des cinq types de bâti. En partant de la plus densément peuplée vers la moins densément peuplée, les cinq catégories de bâti se classent comme suit : grand collectif, habitat de centre ville, petit collectif, habitat semi collectif et habitat individuel

La très forte dispersion des valeurs de densité dans une même classe de bâti a montré qu'il n'est pas envisageable d'utiliser la moyenne de densité comme estimateur de la densité à l'îlot ; sauf, peut-être pour l'habitat individuel qui présente une plage de valeur plus restreinte.

Le type de bâti n'est donc pas un paramètre suffisant pour estimer la densité de population à l'échelle de l'îlot.



Représentation de la densité d'habitants pour chaque type de bâti

Densité de population et Coefficient d'Occupation du Sol

La tentative de caractérisation d'un lien direct entre densité de population et morphologie de bâti s'étant traduite par un échec, le CETE Normandie-Centre a choisi d'étudier la relation entre la densité de contenu (population) et densité de contenant (COS). L'idée de cette démarche provient de l'analyse de deux études déjà réalisées sur le sujet et ouvrant des perspectives intéressantes : Les densités urbaines et le développement durable de Vincent Fouchier (1997), Relation entre la démographie et les surfaces de plancher bâties de la DREIF (1997)

Le COS est un descripteur quantifiable de la densité bâtie. Pour cette étude, c'est le COS brut hors sol qui a été utilisé par le CETE Normandie-Centre. Il se définit comme suit

$$COS = \frac{\sum (\text{Nombre.d'étages} \times \text{Emprise.au.sol})}{\text{Surface.de.l'îlot}}$$

Ce travail a été mené sur la même aire d'étude que la comparaison entre densité d'habitants et le type de bâti.

L'emprise au sol et le nombre d'étages des bâtiments ont été calculés à partir de la BD TOPO de l'IGN. Ces calculs ainsi que le croisement avec la base Îlots...15 ont été réalisés sous SIG (ArcView GIS).

Les différentes étapes qui ont été mises en œuvre pour l'étude de la relation entre le COS et la densité sont les suivantes :

- découpage de l'aire d'étude en zones de bâti homogène
- sélection pour chacun des quatre types de bâti d'un échantillon de vingt à quarante îlots
- calcul de la densité de population pour chacun des îlots sélectionnés
- calcul du COS pour chacun des îlots sélectionnés
- exploitation des résultats (calcul du coefficient de corrélation entre COS et densité de population, discussion suivant les types de bâti)
- conclusion (mise au point éventuelle d'une relation de proportionnalité entre COS et densité de bâti).

Les résultats obtenus sont les suivants

- Le lien existant entre densité de bâti et densité de résidents est fonction de la typologie du bâti
- Les coefficients de corrélation entre densité de résidents et COS brut calculés pour les quatre types de bâti considérés sont compris entre 0.55 et 0.88. Les résultats obtenus sont plus ou moins encourageants selon les types de bâti
- L'habitat collectif est le type de bâti pour lequel la plus forte corrélation entre densité de contenant et densité de contenu est observée.

- ◆ A l'opposé, densité de bâti et densité d'habitants ne semblent, avec une corrélation de 0.55, que faiblement liés pour le tissu urbain de centre ville. Le cas de l'habitat semi-collectif et de l'individuel se situent à une position intermédiaire entre ces deux cas limites
- ◆ la vue de ces premiers résultats, il n'est pas possible d'affirmer que, de façon générale, densité de bâti et densité d'habitants sont significativement corrélées. La tendance est toutefois remarquable et mérite d'être approfondie

Il est probable que les résultats obtenus auraient été meilleurs si la densité humaine (habitants et gens travaillant dans l'îlot) avait été prise en compte plutôt que la densité des habitants seuls

► Suites à donner

Dans l'étude du lien entre morphologie de bâti et densité de population, les résultats obtenus à partir du COS sont beaucoup plus encourageants que ceux issus du seul paramétrage par type de bâti.

De nouvelles investigations apparaissent cependant nécessaires à l'équipe pour définir des coefficients de passage COS/densité adaptés à chaque type de bâti. Il s'agirait tout d'abord de réitérer la démarche sur Rouen et d'autres agglomérations métropolitaines en travaillant à partir de la densité humaine (habitants et emplois) plutôt qu'avec la seule densité d'habitants.

Enfin, dans l'hypothèse optimiste où la relation entre COS et densité humaine serait établie avec suffisamment de certitude pour estimer des populations, un travail méthodologique devra être engagé qui consistera à définir les outils les plus efficaces non seulement dans la détermination du COS mais aussi dans l'exploitation de la densité humaine dans le volet sanitaire des études d'impact

► Présentation de l'équipe

Organisme : CETE Normandie-Centre – DEIOA – Environnement

Personnes impliquées

- ◆ Sophie MARTINONI
- ◆ Vincent DEMEULES
- ◆ Jean-François BURKHART (CETE de Lyon)

Note rédigée par Vincent Demeules et Vincent.demeules@equipement.gouv.fr
Sophie Martinoni - CETE Normandie-Centre

Relue par Magali Di Salvo - CERTU

Magali.di-salvo@equipement.gouv.fr