

COLLOQUE DE LILLE

La géocodification des données _____ par la connaissance de l'occupation du sol

par André BALLUT

Chargé d'études à l'IAURIF

(Institut d'Aménagement et d'Urbanisme de la Région Ile-de-France)

1 — LES BESOINS DE CONNAISSANCE DE L'OCCUPATION DU SOL

On ne peut pas concevoir la Banque de Données Urbaines sans aborder le problème de la collecte des données et de leur repérage géographique. En premier lieu on peut se demander, quels types de données sont utiles ?

Toute analyse d'un groupe humain implanté sur un territoire suppose :

- a) La connaissance de son importance, ce qui implique son dénombrement avec en corollaire celui des logements.
- b) La connaissance de ses activités qui se traduisent par des flux économiques.
- c) La connaissance du territoire sur lequel ce groupe est implanté et exerce ses activités.

Si les deux premiers points sont relativement bien connus grâce aux recensements et aux statistiques économiques, il n'en est pas de même du troisième. Et pourtant, le problème se pose de plus en plus aux responsables et décideurs de collectivités de faire des choix raisonnés sur la transformation de l'usage du territoire qu'ils gèrent.

Toute étude d'aménagement qui établit des diagnostics ou des prévisions et propose des objectifs d'action à la puissance publique doit disposer de données sur l'usage du sol (entre autre), en terme de connaissance du patrimoine et de son évolution.

Chaque fois qu'une implantation nouvelle est envisagée sur un terrain c'est en fait une modification de l'usage de ce terrain qui devrait faire l'objet d'un bilan économique préalable.

La connaissance de l'occupation du sol fournit trois types de produits :

- Une image du territoire à une date donnée.
- Le bilan chiffré des différents composants du territoire.
- Un référentiel géographique qui peut servir de réceptacle à d'autres données de caractère socio-économique.

2 — LES OUTILS DE COLLECTE DE L'INFORMATION

En matière d'occupation du sol, on s'aperçoit en fait que les sources possibles sont peu nombreuses et, soit peu fiables soit non exhaustives. On trouve deux catégories d'outils : ceux actuellement opérationnels et ceux que l'on peut prévoir à moyen terme.

Actuels :

Il y a en premier lieu ce que l'on pourrait appeler "les filtres administratifs", c'est-à-dire un observatoire placé à un point du flux administratif comme par exemple au niveau du permis de construire. L'expérience montre que cette source est à la fois difficile à exploiter géographiquement et non exhaustive.

Il y a ensuite les enquêtes, forcément localisées et difficilement extensibles à la totalité d'un territoire, et les recensements qui ne fournissent qu'une part très faible de l'information sur l'occupation du sol.

Il y a enfin la photo aérienne qui est de loin la source la plus intéressante. Cependant il faut bien noter qu'en milieu urbain elle ne résout pas tous les problèmes. Si, grâce à la photo aérienne on peut avoir une bonne appréciation de l'implantation physique de la ville, on n'a que peu d'informations sur le rôle fonctionnel des bâtiments. Un bâtiment de même apparence pourra abriter un supermarché, un entrepôt, un atelier, etc. Un problème d'identification des objets subsiste et, si la photo aérienne fournit environ 80 % de l'information nécessaire les 20 % restants sont à acquérir par d'autres moyens.

A terme :

L'outil d'analyse de l'occupation du sol le plus séduisant intellectuellement est la Télédétection spatiale. Cette technique nouvelle peut certes fournir une information, mais il faut bien en convenir insuffisamment fine actuellement sur le milieu urbain sauf, s'il s'agit au niveau d'une agglomération entière de repérer les zones d'extension périphériques.

(1) Pixel, de l'anglais Picture element : plus petite unité d'image.

Nous avons mené des travaux sur des données Landsat, la taille du pixel (1) 57 x 79 m englobe des objets urbains trop hétérogènes et la radiométrie enregistrée n'est pas, dans plus de 60 % des cas significative donc représentative d'une occupation du sol. Il y a là un bûtoir sur lequel travaillent plusieurs laboratoires de recherche faisant intervenir dans les traitements des données satellites, des notions de proximité, de fréquence, de texture et de structure.

Le nouveau satellite Spot introduit un vision nouvelle avec une taille de pixel de 20 x 20 m en multispectral et 10 x 10 m en "panchromatique". Il est encore trop tôt pour dire si l'analyse de l'occupation du sol sera suffisamment fine pour être utilisable mais d'ores et déjà l'image fournie peut être comparée à la qualité d'une vue aérienne au 1/50 000. De plus, Spot permettra, grâce à ses prises d'enregistrement décalées, une vision stéréoscopique.

Le traitement des données satellite permet de réaliser des analyses qualitatives de l'espace humanisé. Une étude, actuellement en cours à l'IAURIF porte sur le taux d'espace vert des communes de la banlieue de Paris à partir des données Landsat. Un calcul de l'indice de végétation (1) permet d'obtenir une cartographie de l'agglomération suivant plusieurs classes de menu verdissement, faisant aussi apparaître les zones de carences en masse végétale, ou de minéralisation excessive.

Bien entendu cette démarche qualitative peut être quantifiée par des mesures de surfaces très aisées puisque toutes les données sont référencées géographiquement.

Le gros avantage de la télédétection spatiale est néanmoins sa répétitivité, c'est-à-dire la possibilité d'observer périodiquement des territoires et de pouvoir en détecter les changements. La comparaison des deux images à deux dates différentes permet de repérer les mutations de l'usage du sol. C'est cette possibilité qui présente actuellement le plus d'intérêt pour les inventaires de l'occupation du sol en offrant des moyens de mise à jour. Une expérience est actuellement en cours (2) sur des simulations Spot de la région de Paris pour définir un algorithme permettant une comparaison fructueuse de deux images à deux dates.

3 — LA CONSTITUTION D'UN OUTIL DE "GÉOCODIFICATION"

3.1. SAISIE ET STOCKAGE DE L'INFORMATION

Le problème du stockage de cette information est bien connu et en fait il n'y a que trois méthodes possibles :

- a) La grille de petits carrés.
- b) Le balayage à pas fixe.
- c) La polygonale.

a) La grille de petits carrés ne peut être suffisamment précise que si la taille des carrés est petite ce qui suppose un très grand nombre de données à stocker et traiter. Au niveau d'un territoire, passer d'une grille de carrés de 100 m de côté à une autre de carrés de 50 m de côté c'est multiplier les informations à

traiter par quatre et les temps de traitement par davantage. De plus pour couvrir un territoire homogène, de même mode d'occupation du sol, il faut le même nombre de carrés que dans un secteur hétérogène : le nombre d'informations à enregistrer est indépendant de l'occupation du sol.

b) Le balayage à pas fixe apporte une sensible amélioration à la méthode précédente :

- La précision est théoriquement absolue dans un axe.
- Le décalage d'un point de terrain ne peut être supérieur à la moitié du pas de balayage.
- Le nombre d'informations à stocker est fonction de l'homogénéité du territoire, pour chaque mode d'occupation du sol il suffit, par ligne de balayage d'un point "d'entrée" et d'un point de "sortie".

L'expérience montre que l'on diminue par cinq le nombre de données à stocker tout en améliorant sensiblement la mesure des surfaces. En revanche cette méthode permet difficilement d'individualiser des zones pour y affecter d'autres catégories de données.

c) La méthode polygonale est la seule qui fournisse à la fois la précision de mesure des surfaces et l'individualisation des zones. Elle n'a que l'inconvénient d'être plus lourde et donc plus coûteuse à mettre en œuvre mais ceci est largement compensé par tous les avantages des exploitations ultérieures possibles.

3.2. CROISEMENT DES DONNÉES, LEUR LOCALISATION FINE, LA "GÉOCODIFICATION"

La connaissance de l'occupation du sol, son relèvement et son stockage informatique fournissent une base géographique, un référentiel, dans lequel peuvent être redistribuées géographiquement d'autres données connues seulement à l'adresse ou par unités administratives.

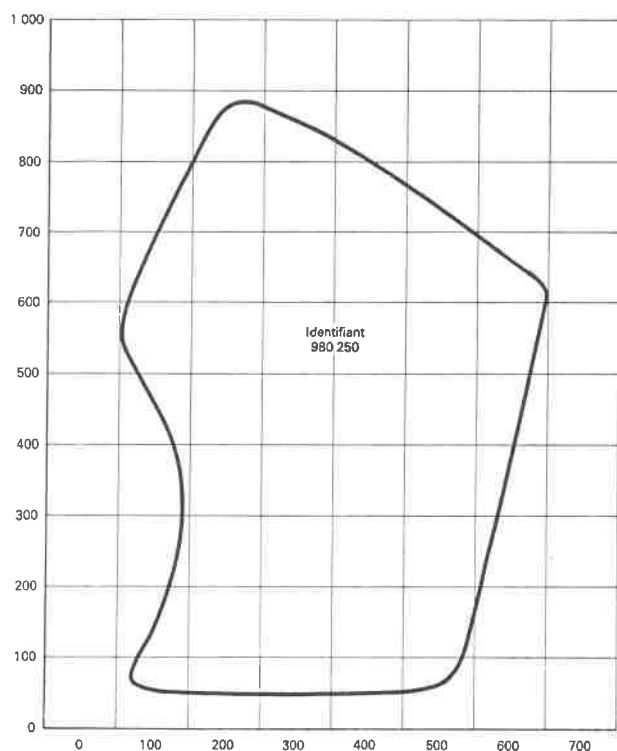
Une partie des informations de caractère socio-économique particulièrement celles issues du Recensement Général des Populations, ne sont plus exploitées que par l'intermédiaire d'un taux de sondage (au 1/5, au 1/20). On assiste, recensement après recensement à une dégradation de l'information utile. Certains pays voisins (Allemagne Fédérale, Suisse) envisagent même de ne plus faire de recensement. Il est donc nécessaire de disposer d'un outil qui permette, à partir de données connues seulement à l'adresse ou par grande unité administrative (commune, département) de faire une répartition géographique fine de l'information. Le croisement des données socio-économiques avec celles de l'occupation du sol permet cette répartition.

(1) Cet indice est obtenu par la formule suivante : $\frac{\text{Canal 5} - \text{Canal 7}}{\text{Canal 5} + \text{Canal 7}} \times K$

(2) Cette expérience est menée par l'IAURIF et le Centre Scientifique d'IBM France, à partir de simulations Spot du GDTA.

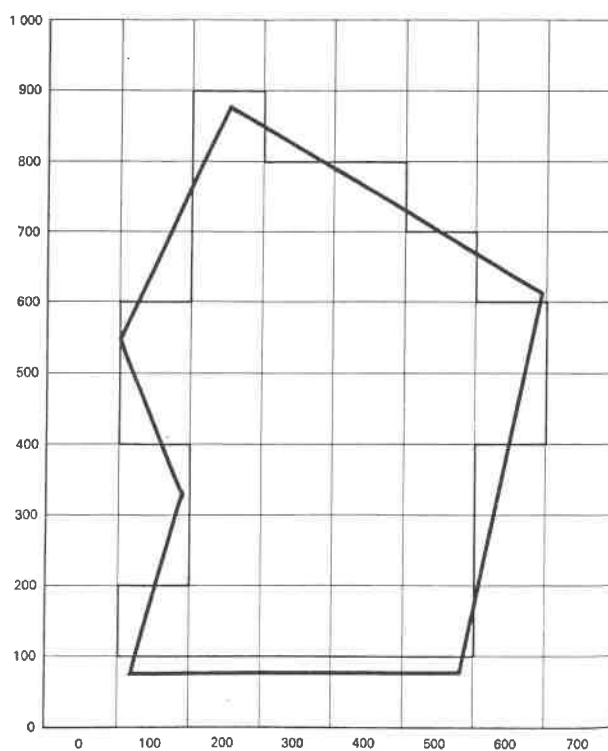
... COLLOQUE DE LILLE

ZONE A ÉTUDIER FICTIVE



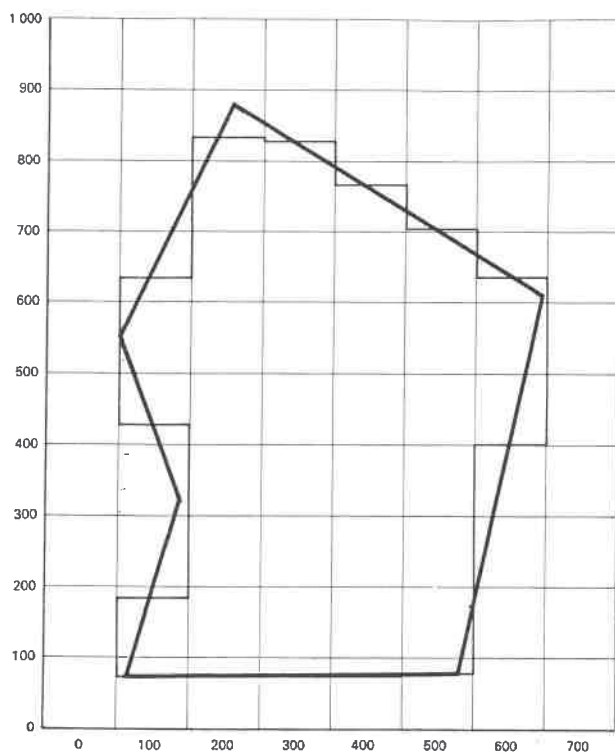
1.1) Exemple de zone à numériser.

REPRÉSENTATION DE LA ZONE EXEMPLE SOUS FORME « PETITS CARRÉS »



1.2) Méthode "a" grille de petits carrés.

REPRÉSENTATION DE LA ZONE EXEMPLE SOUS FORME « PETITS RECTANGLES »



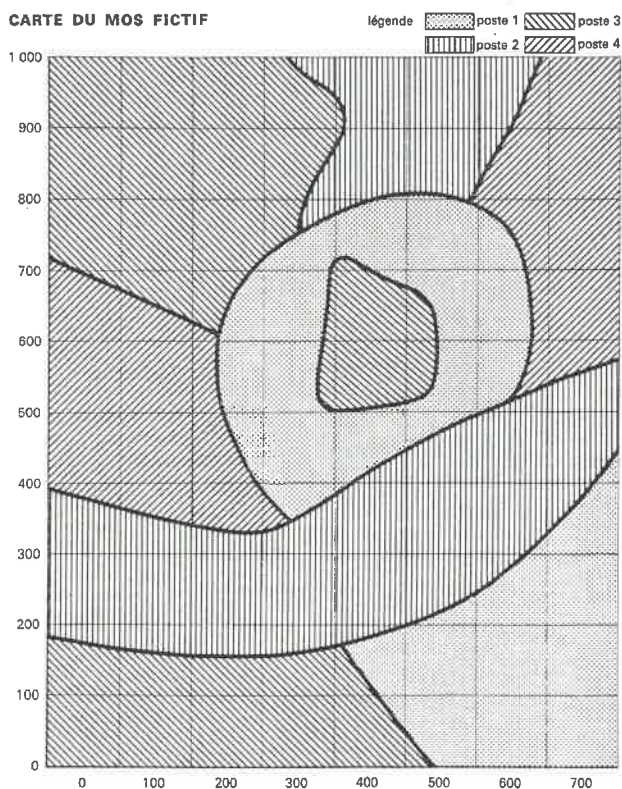
1.3) Méthode "b" balayage à pas fixe.

REPRÉSENTATION POLYGONALE DE LA ZONE FICTIVE



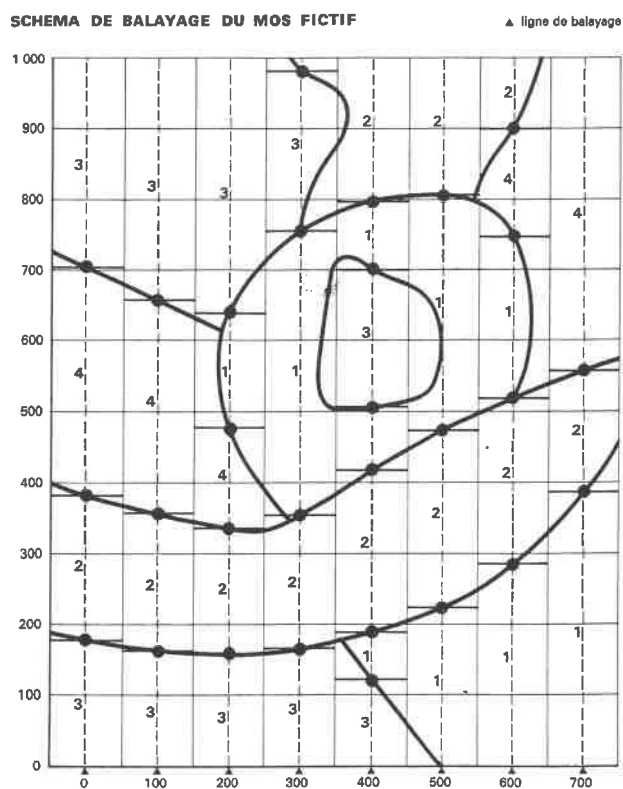
1.4) Méthode "c" polygonale.

CARTE DU MOS FICTIF



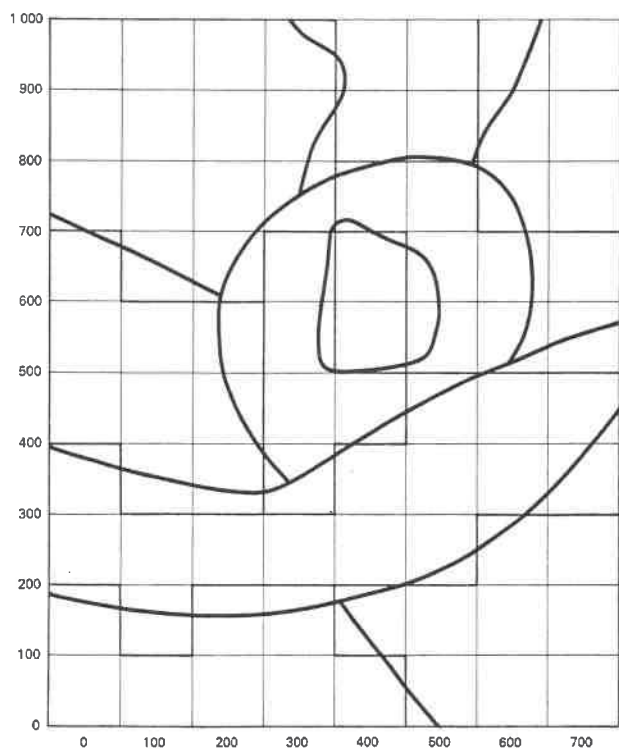
1.5) Exemple de type de MOS à numériser.

SCHEMA DE BALAYAGE DU MOS FICTIF



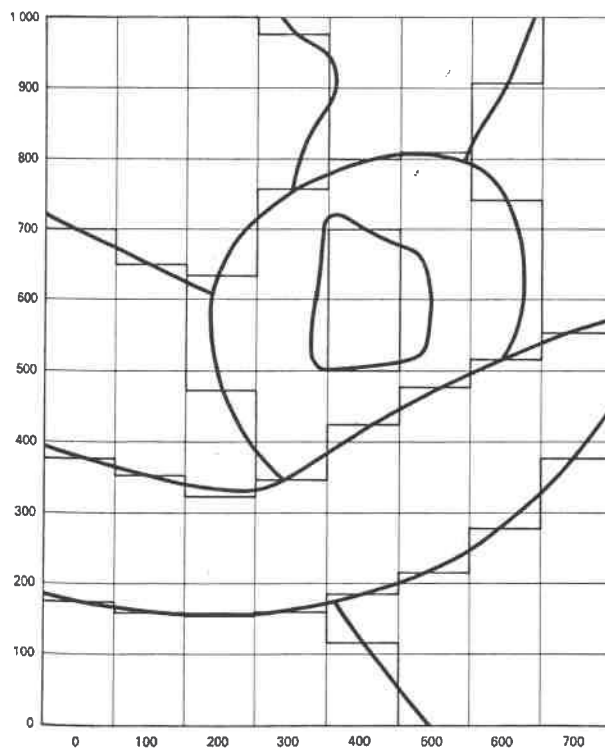
1.7) Méthode "b" balayage à pas fixe.

**REPRÉSENTATION DE L'EXEMPLE MOS
SOUS FORME « PETITS CARRÉS »**



1.6) Méthode "a" par petits carrés.

**REPRÉSENTATION DE L'EXEMPLE MOS
SOUS FORME « PETITS RECTANGLES »**



1.8) Méthode "b" balayage à pas fixe, exemple de contour de zones.

... COLLOQUE DE LILLE

4 — UN EXEMPLE SUR UNE GRANDE MÉTROPOLE : LE "MOS" DE L'ILE-DE-FRANCE

Le cas présenté ici, celui de la Région Ile-de-France, ne doit pas être considéré ici comme un modèle. La conception de l'ensemble de l'étude et les choix techniques sont liés à la spécificité de cette région.

Un premier inventaire du Mode d'Occupation du Sol (MOS) a été réalisé entre 1972 et 1974 sur 7 000 km² des 12 000 de la région Ile-de-France. Ce premier travail a permis de tirer un certain nombre d'enseignements, mais surtout, cela a fait prendre conscience aux responsables régionaux de la nécessité d'une information permanente sur l'évolution du territoire. Le principe a été retenu d'un inventaire "lourd" tous les dix ans (coordonné avec les recensements de populations) et de mises à jour intermédiaires tous les quatre ou cinq ans.

4.1. COLLECTE DE L'INFORMATION

Le choix des sources d'informations relatives à l'occupation du sol étant limité, nous avons fait appel à la photo-interprétation dont la réalisation a été confiée à l'IGN. Les renseignements sont reportés par coloriage sur des plans au 1/5 000 suivant une légende en 54 postes. A cette première étape s'ajoute une seconde qui consiste à renseigner les plans coloriés en y portant :

- les limites administratives,
- la voirie,
- les informations complémentaires.

La succession des tâches concrètes pour réaliser le MOS (le processus s'étend sur environ 3 ans)

1. Interprétation, réalisée à l'IGN, des photos aériennes de la région (1981-1982) et report des informations sur plans au 1/5 000.
2. Collecte et report par l'I.A.U.R.I.F. des informations complémentaires (notamment concernant les équipements).
3. Dessin des limites administratives, de la voirie, des contours MOS.
4. Préparation du plan de base à informatiser.
5. Informatisation des données par une méthode de saisie de la polygonale des zones grâce au système "REPERE" logiciel infographique développé par l'I.A.U.R.I.F.
6. Réalisation informatique des plans de contrôle.
7. Vérification par les collectivités locales, DDE et DDA.
8. Informatisation à l'I.A.U.R.I.F. des corrections et compléments.
9. Réalisation des fichiers informatiques des données.
10. Sélection infographique des couleurs pour les éditions cartographiques au 1/25 000 et 1/100 000.
11. Edition des bilans chiffrés d'occupation du sol pour toutes communes de la région.
12. Réalisation d'un système d'informations urbaines (croisement de l'occupation du sol et des données socio-économiques).
13. Système d'interrogation "à la demande".

Celles-ci sont obtenues à partir des bulletins municipaux (1) ou à partir des fichiers administratifs (Ministère de l'Éducation Nationale pour les écoles, Ministère de la Jeunesse et des Sports pour les équipements sportifs, Observatoire Régional de Santé pour les établissements de soins, etc.). Pour chaque type d'information il n'y a pas qu'une source unique fiable, mais plusieurs qu'il faut croiser entre elles.

L'ensemble des informations étant reporté sur le plan il est procédé à la segmentation du contour de chaque zone figurant un type d'occupation du sol, ceci afin de schématiser le dessin, de le clarifier pour faciliter le travail de l'étape suivante la numérisation des données.

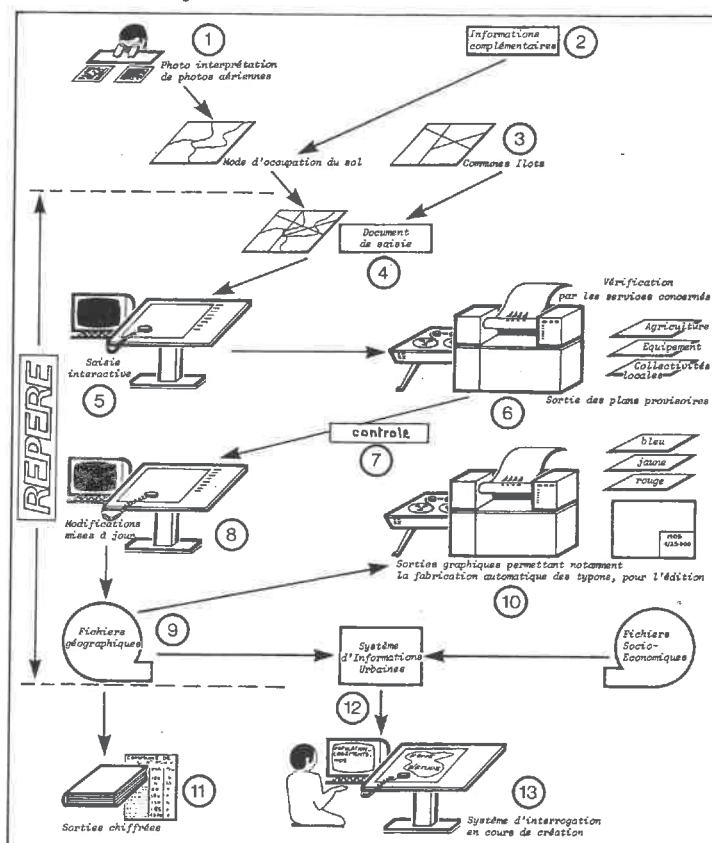
4.2. NUMÉRISATION DES DONNÉES ET CONTRÔLE

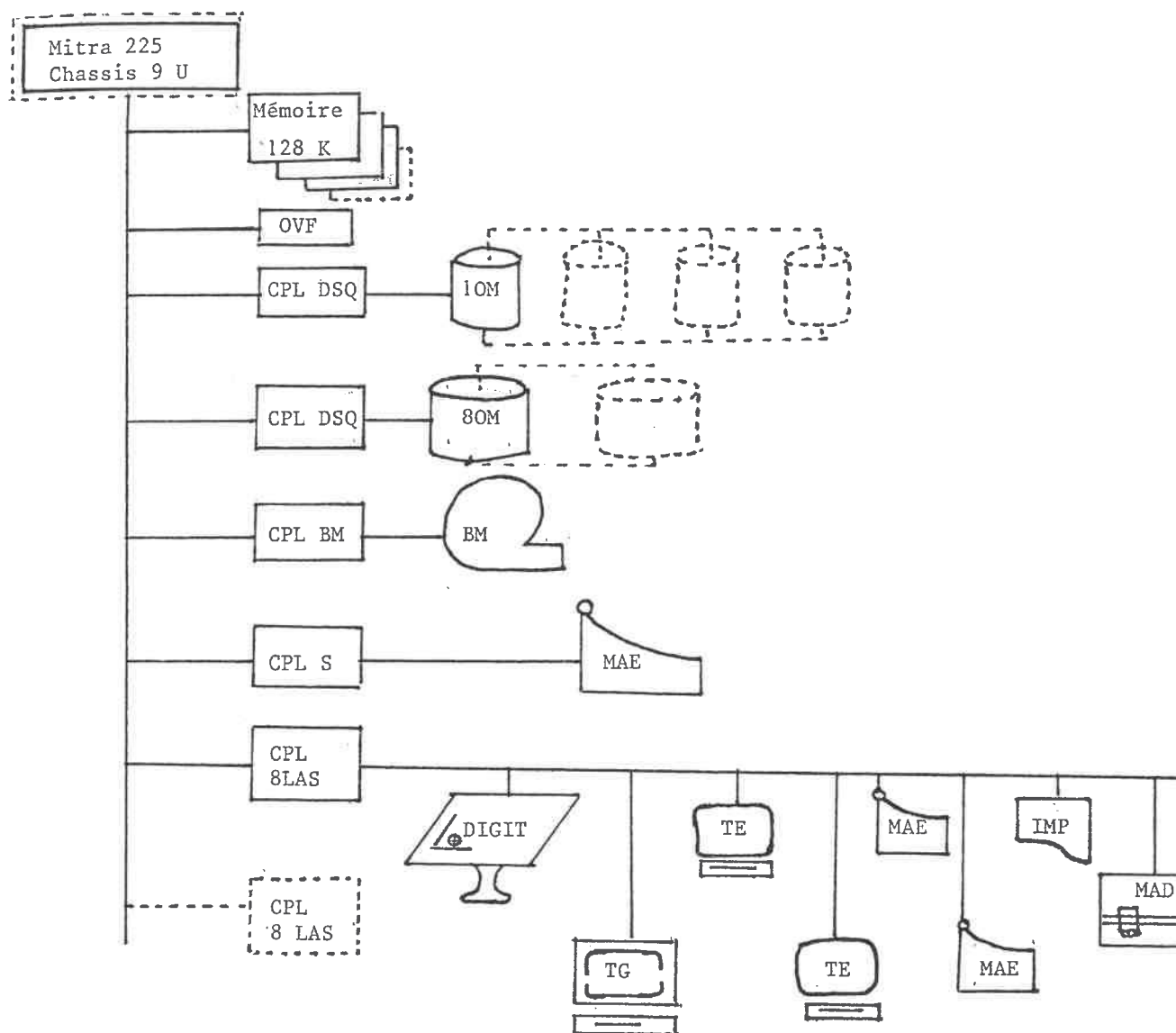
La mise sur support informatique de l'information est faite en utilisant une configuration infographique de matériel français (matériel GIXI du groupe CISI) avec langage APL. Le logiciel de saisie et de traitement de l'information, "REPERE" a été entièrement conçu à l'IAURIF. C'est un système qui répond aux besoins de stockage, d'exploitation numérique et graphique ainsi que de mise à jour des données.

L'information est saisie et stockée suivant plusieurs niveaux :

- niveau des limites administratives,
- niveau de la voirie (contour des îlots),
- niveau des modes d'occupation du sol,
- etc.

(1) Nous avons écrit à la plupart des communes urbaines de la région pour demander le bulletin municipal qui inclut souvent un plan figurant les principaux équipements de la commune.





Abreviations et explications

Mitra 225 : mini ordinateur 16 bits fabriqué par la SEMS
 Mémoire 128 K : mémoire vive (accès très rapide) de 128 x 1024 octets (1 nombre = 4 octets)
 OVF : opérateur virgule flottante câblé
 CPL : coupleur
 DSQ : disque
 10 M : 10 Mega octets 10 x 1024 x 1024 octets
 BM : Bande magnétique
 MAE : Machine à écrire
 CPL S : coupleur système
 CPL 8 LAS : coupleur de 8 lignes asynchrones
 DIGIT : table à digitaliser format A0
 TG : terminal graphique
 TE : Terminal écran non graphique
 IMP : Imprimante
 MAD : Machine à dessiner

... COLLOQUE DE LILLE

Des niveaux supplémentaires sont prévus pour le stockage d'informations spécifiques qui seront à localiser géographiquement et à croiser avec les données relatives à l'occupation du sol.

Des plans de contrôle sont produits sur traceur Benson 14 à partir d'un logiciel (partie intégrante de "REPÈRE") qui permet de reproduire en couleurs 130 postes de la légende. Ces plans, au 1/5 000, sont dessinés sous deux formes :

- plan rectangulaire dans le découpage régional,
- plan "à la commune" ne faisant apparaître que la commune concernée avec en légende le bilan chiffré des surfaces des différents modes d'occupation du sol.

Les plans de contrôle sont envoyés dans les DDE (GEP) et DDA ainsi qu'aux municipalités pour vérification.

Après introduction des éventuelles corrections dans le fichier informatique, celui-ci est prêt pour exploitation.

4.3. LE SYSTÈME DE CROISEMENT DES DONNÉES SOCIO-ÉCONOMIQUES AVEC CELLES DE L'OCCUPATION DU SOL, BASE DE GÉOCODIFICATION

Le principe consiste à localiser finement, en s'appuyant sur la connaissance de l'occupation du sol, des données répertoriées seulement "à l'adresse", à "l'îlot" ou à "la commune". L'ensemble du logiciel de cette "Banque de données" n'est pas encore opérationnel mais en cours d'élaboration à partir d'un travail réalisé à l'étranger pour une Banque de Données Urbaines.

Trois types d'exploitation sont en cours de réalisation :

- 1) Une cartographie standard de l'occupation du sol sur l'ensemble de la région Ile-de-France à l'échelle du 1/25 000 suivant une légende abrégée à 54 postes.
- 2) Une présentation standard de bilans chiffrés par commune de l'occupation du sol en 150 postes.
- 3) Un système d'exploitation "à la demande" qui traite soit des données "MOS" uniquement, soit des données MOS croisées avec des données socio-économiques (population, logement, emploi,...) :
 - a) sur un multiple de zones élémentaires (exemple multiple d'îlots),
 - b) sur une zone quelconque à définir graphiquement.

C'est cette troisième exploitation qui est, de loin, la plus utile aux aménageurs.

4.4. LA MISE À JOUR

Le problème de la mise à jour d'un tel inventaire est délicat, une mise à jour complète consiste à refaire de manière exhaustive un inventaire et une mise à jour statistique ne fournit pas d'informations localisées suffisamment finement. D'autre part prévoir des ou-

tils de mise à jour plusieurs années à l'avance consisterait à anticiper dangereusement sur l'évolution des techniques. Nous avons étudié trois méthodes de mise à jour.

MÉTHODE STATISTIQUE

Testée sur 400 000 hectares avec un service de 25 000 points cette méthode est suffisamment fiable (écart-type) pour permettre de repérer que tel mode d'occupation du sol a crû de 1 % ou que tel autre a diminué de 2 % mais en revanche il n'est pas possible de localiser ces évolutions et de ce fait l'information perd tout son intérêt quand on traite de l'occupation du sol.

MÉTHODE PONCTUELLE

Les secteurs d'évolution d'une agglomération sont en général connus et la méthode consiste à faire déterminer par un groupe ad-hoc d'experts les secteurs en mutation. La mise à jour complète étant limitée à ces secteurs, le reste de l'agglomération est réputée n'ayant pas évolué. On voit immédiatement que cette méthode peut engendrer des distorsions dans la mesure où une partie importante de l'évolution est dispersée dans le tissu bâti.

MÉTHODE PERMANENTE UTILISANT LES DONNÉES SATELLITES

Cette troisième méthode est en cours d'étude et il serait prématuré d'en tirer toutes les conclusions. Cependant il apparaît que les images satellites d'un même secteur à deux dates, comparées entre elles, permettent de repérer les changements dans l'occupation du sol. Des essais ont été faits, d'autres sont en cours, à partir des données "Landsat", les mutations, d'une certaine taille sont bien repérées, en revanche il est plus difficile de savoir ce qu'il y avait avant et ce qu'il y a après. Grâce à SPOT l'information sera plus fine et un travail est en cours pour comparer deux simulations SPOT sur la banlieue de Paris, l'une de 1981 l'autre de 1983. Les premiers résultats permettent d'être optimiste et de penser qu'une observation annuelle fournira un repérage suffisant des mutations.

Entre le premier coup de pioche sur un terrain et la mise à disposition d'un logement il s'écoule entre 3 et 4 ans. Si cette mutation de terrain n'est pas repérée la première année elle le sera très probablement la deuxième ou la troisième. Une agglomération comme celle de Paris subit une évolution de 0,4 % de son territoire chaque année, si l'imagerie satellite permet de repérer toutes les petites zones ayant subi un changement, il sera ensuite très aisé, à l'aide de photos aériennes ou même quelquefois de visites sur le terrain d'identifier ces mutations et de les enregistrer dans le fichier d'occupation du sol. Le travail de mise à jour se trouvera diminué par un coefficient 100. Toutefois cette méthode n'étant pas encore totalement évaluée, il est trop tôt pour avancer des affirmations.