

## L'ATLAS DE QUITO : METHODES INFORMATIQUES

(Marc SOURIS)

**RESUME** - Nous présentons dans cet article un prototype de système d'information géographique pour la gestion et la planification urbaine, dans ses principes comme dans son utilisation pour le projet d'Atlas informatisé de Quito, actuellement en cours.

Le prototype est conçu, dans son système de gestion de données, comme une extension du modèle relationnel à la gestion des données localisées : les SGBD n'étant pas à même de manipuler des données de dimension deux, nous avons développé des possibilités de gestion étendant l'algèbre relationnel aux données localisées : restriction spatiale, jointure spatiale, projection spatiale. Ces nouvelles opérations nous permettent de manipuler la localisation par des opérations algébriques ou logiques : la carte considérée comme résultat final devient le résultat intermédiaire d'une série d'opérations de gestion et de manipulation de données.

Le système développe un certain nombre de fonctionnalités propres aux données localisées, dont le principe est toujours de créer des attributs en mettant les objets en relation sur un critère d'appartenance ou de distance : sélection des points à une distance donnée d'une zone, critère d'appartenance d'une zone à une autre, etc.. tout en conservant le principe de la gestion relationnelle, le système est également étendu à la manipulation d'objets particuliers, comme les images satellites, les photographies aériennes, objets qui, grâce à des opérations spécifiques développant l'aspect multi-média et la gestion orientée-objet, sont intégrés dans le processus relationnel.

Toutes les fonctionnalités développées autour de ce système trouvent leur application au niveau de la planification urbaine. Nous effectuons la mise en oeuvre concrète du prototype dans le cadre du projet d'Atlas informatisé de Quito (AIQ) mené par l'ORSTOM avec plusieurs institutions équatoriennes (Institut Géographique Militaire, Municipalité de Quito, Institut Panaméricain de Géographie et d'Histoire). L'aire métropolitaine de Quito est confrontée à des problèmes de croissance, de fonctionnement, d'approvisionnement, et un tel outil va permettre en facilitant l'accès à l'information, d'améliorer la connaissance et donc la réflexion sur les problèmes de gestion de la ville.

La base de données est constituée à partir de l'information existante - saisie du découpage en îlots, avec intégration de données sur les équipements, l'infrastructure, le foncier, les constructions, saisie du découpage en unité de recensement avec intégration des données censitaires, saisie des différents secteurs administratifs de la ville, saisie de données du milieu physique sur l'aire métropolitaine (altitude, utilisation du sol, géologie) - et également grâce à l'information récoltée par des enquêtes spécifiques, exhaustives (recensement des activités, par

exemple) comme par sondage (démographie, migration, vulnérabilité aux risques naturels, etc.). Toutes ces données sont analysées, confrontées, de manière à tirer à la fois un bilan thématique de la situation existante et une vision prospective en vue de la planification urbaine (redéfinition du POS, zones d'interventions prioritaires, aménagement des zones d'extension, etc.).

## INTRODUCTION

Nous présentons ici les méthodes et moyens informatiques mis en oeuvre pour le projet ATLAS DE QUITO, qu'il s'agisse de matériel ou de logiciel.

Informatiser un atlas exige de savoir faire au moins trois choses : *saisir des données* (dont la plupart sont localisées, donc également saisir des plans), *gérer et traiter des données* (donc posséder un système approprié), et *cartographier des données*, car les résultats attendus pour un atlas sont essentiellement visuels. Mais le projet d'Atlas de Quito n'est pas seulement la constitution d'un ouvrage, conçu ou fabriqué grâce à des machines : c'est également la constitution d'un *outil permettant de faciliter l'accès à l'information, d'améliorer la connaissance et donc la réflexion sur les problèmes de gestion de la ville*. C'est à la création d'un tel outil que nous nous sommes attachés, et c'est dans ce sens que sont conçus les développements informatiques faits autour de ce projet.

### 1. GESTION DES DONNEES

L'informatique nous sert avant tout de système informatif. Pour une application du type atlas, il est d'abord nécessaire de disposer d'un *système de gestion de données performant*. Ici, nous avons multitude de sources de données, qui sont pour la plupart géographiques, et il s'agit bien sûr de gérer également cette localisation dans l'espace. Il faut pouvoir comparer, juxtaposer, superposer, sélectionner, etc., afin de produire l'information sur laquelle sont effectués les traitements et produits les résultats. La gestion de ces données nécessite à la fois puissance et souplesse : le système que nous développons (nommé TIGRE) essaye de répondre à ces objectifs.

Le *modèle relationnel* est à la base de notre système de gestion de données : laissant à l'utilisateur le choix de la navigation dans l'ensemble des données, c'est le modèle à la fois le plus simple et le plus performant en terme de fonctionnalité. L'interrogation des données avec TIGRE est *interactive* : le résultat d'une opération de gestion sert d'entrée à l'opération suivante, ce qui permet à l'utilisateur de construire sa requête, et d'aboutir, après un cheminement qui peut être complexe, à un résultat ayant fait intervenir un grand nombre de sources de données différentes et reliées entre elles au fur et à mesure de ce cheminement, quand les opérations demandées le requièrent.

## 2. GESTION DE LA LOCALISATION

Particularité essentielle du système de gestion que nous développons pour gérer les données de l'Atlas de Quito : *devoir gérer des données géographiques*. Prenant exemple sur la gestion relationnelle des données, nous avons étendu les principes de cette gestion au traitement de la localisation. Ce qui signifie que les opérations classiques (sélection, jointure - c'est-à-dire croisement - sur un critère commun) seront également disponibles pour la localisation : sélection sur un espace géographique (par fenestrage, par masquage), croisement de deux objets sur leur localisation (on les met en relation s'ils sont au même endroit, s'ils sont à telle distance l'un de l'autre, le critère de mise en relation étant toujours ici basé sur la distance). La localisation est une donnée intrinsèque : tout objet localisé pourra, sur ce critère, être comparé à tout autre objet localisé de la base de données, ce qui ouvre le champ à une multitude d'opérations de gestion sur des données qui n'ont comme critère commun, comme moyen de liaison, que leur localisation dans l'espace. Ces nouvelles opérations entrent bien sûr dans le processus normal d'interrogation interactive et font partie du cheminement de l'utilisateur dans la base de données : l'algèbre relationnelle est ainsi étendue à la localisation.

Cette méthode de gestion de la localisation a des conséquences importantes dans l'élaboration de l'information géographique de base : chaque donnée peut en effet être conservée dans sa propre implantation spatiale sans réduire les possibilités de traitements ultérieurs, puisque c'est le système qui se charge de gérer la localisation : plus question de carroyage régulier dans lequel on ramène toutes les données pour pouvoir les comparer, en faisant fi de l'implantation spatiale d'origine. Plus question non plus de système dans lequel la localisation n'est traitée que pour cartographie, où les données localisées ne sont comparées que visuellement par la superposition de sorties graphiques. Plus question enfin d'avoir à préciser pour un objet d'un certain type son appartenance géographique à un objet d'un autre type : la localisation seule suffira à établir cette appartenance. Ainsi, le P.O.S. ne sera plus une variable de la parcelle cadastrale : saisi à part comme un découpage thématique de l'espace, c'est lors de l'interrogation des données que l'on pourra établir le P.O.S. pour chaque parcelle. Si l'on veut modifier le P.O.S., on pourra le faire indépendamment des parcelles et trouver ensuite, par une opération de gestion de données, l'impact de la modification sur les parcelles concernées. Par cette méthode de gestion de la localisation, les opérations de simulation sont rendues possibles et faciles à mettre en oeuvre.

Garder chaque objet dans sa propre implantation spatiale, avec la possibilité de comparer et de mettre en relation la spatialisations de différents objets, met bien sûr en évidence les problèmes de validité de la localisation : des objets d'échelles spatiales différentes, ou plus généralement de validité différente pour leur localisation, ne sauraient être joints sur cette localisation sans grande précaution.

### 3. LES TRAITEMENTS

Le système de gestion nous offre donc, grâce à ces opérations de base, la possibilité de développer, sur un ensemble de données, de nombreux traitements spécifiques. C'est le second aspect du système développé, qui doit offrir des traitements sortant du cadre strict du système de gestion. Ces possibilités sont, sans en faire l'énumération complète :

- l'étude des données par des *moyens statistiques simples* (moyenne, fréquence, écart-type, histogrammes, corrélations, régressions, histogrammes bi-dimensionnels) ;

- la *création de nouvelles variables* (par classification, par calcul numérique, par calcul logique, par agrégation) ;

- en plus, un certain nombre d'opérations propres aux données localisées. Ce sont le calcul de *surface*, de *périmètre*, de *valeur d'appartenance* ou *agrégation* sur un critère géographique (moyenne de points dans une zone, de zones à zones, etc.), la création d'un *masque de sélection* géographique par rapport à certains objets (par exemple, l'espace défini par la condition de se situer à moins de 10 mètres des portions d'un réseau ayant une caractéristique donnée), l'*interpolation géométrique* et la création de bloc-diagrammes.

Ces traitements s'inscrivent encore dans le processus d'interrogation interactive de la base de données : on fait un histogramme pour rechercher les seuils de classification d'une variable, on classe la variable, on fait une sélection sur certaines classes, on calcule un effectif sur le résultat, etc.

D'autres opérations plus spécifiques seront traitées comme des macro-commandes, pour répondre à des problèmes bien précis. Ces modules seront alors développés au-dessus du système de gestion, en se servant de ses possibilités mais sans s'inscrire dans le processus d'interrogation interactive.

Autre possibilité importante : le croisement avec les images numériques provenant des *satellites d'observation de la terre*. En effet, la gestion de la localisation permet, en faisant coïncider information géographique et données issues de la télédétection, à la fois d'*introduire de nouvelles données* pour des objets existants dans la base de données (comme, sur chaque îlot, la valeur étant calculée uniquement sur les pixels appartenant à l'îlot, une moyenne de radiométrie, un indice de densité de végétation, un écart-type local, ou plus généralement la valeur d'une classification de l'image), et également de développer de *nouvelles techniques de traitement en télédétection* (par zone plutôt que par pixel : l'élément étudié n'est plus seulement le pixel ou le groupe de pixels, mais l'ensemble des pixels correspondant au sol à une réalité physique ou thématique connue et corrélée, si possible, à la donnée recherchée sur l'image).

Pour de telles opérations, le recalage de l'imagerie satellitaire sur une projection géographique connue du système est nécessaire. Les opérations sont ensuite rendues simples par la coexistence, dans le système TIGRE, des structures vectorielle et matricielle pour la localisation.

#### 4. LES RESTITUTIONS

Le troisième aspect du système que nous développons concerne bien sûr l'aspect *cartographie* et représentation graphique.

Notons tout d'abord que nous avons choisi de conserver la localisation en coordonnées géographiques globales (longitude, latitude), ce qui nous donne un référentiel commun, quelle que soit l'origine de la donnée géographique (échelle, projection). C'est ce référentiel commun qui permet de comparer la localisation de toutes les données géographiques d'une même base, et ce faisant de *choisir la projection géographique* servant également à la cartographie des résultats obtenus.

La représentation graphique se fait d'abord sur *écran couleur* de haute définition (1100\*900 pixels). La capacité d'affichage (256 couleurs sur 16 millions de possibilités) permet à l'utilisateur de composer son image comme il le souhaite. Le système lui offre la possibilité de visualiser les résultats de ses requêtes : images thématiques, données ponctuelles associées à des symboles, contours, réseaux, noms et valeurs. Il a le choix des couleurs et des trames : il peut créer, modifier des palettes, construire sa légende, rajouter des noms et des symboles. Toutes ces opérations sont interactives et rendues très simples par l'utilisation d'une souris.

Toujours directement sur l'écran, le système offre à l'utilisateur des possibilités de *calculs métriques* : distance entre deux points choisis sur l'écran, le long d'une ligne brisée, surface d'un polygone défini point par point. Ainsi par exemple, ayant défini une fenêtre d'étude et fait tracer le contour des îlots dans cette fenêtre, un utilisateur peut savoir immédiatement la largeur d'une voie, la distance séparant tel îlot de tel autre, en passant par la voirie, le long d'un réseau, à vol d'oiseau, etc.. Le système, en fonction de la taille de la fenêtre et de la projection utilisée, donne tous ces résultats en mètres.

Egalement sur l'écran, on peut y *définir directement des objets*, comme des zones thématiques, des données ponctuelles, des réseaux, et ce, soit pour créer un *masque de sélection géographique* (par exemple : on trace une ligne le long d'une portion de voie, puis on demande la zone se trouvant à moins de 10 mètres de cette ligne pour pouvoir ensuite sélectionner les activités se trouvant dans cette zone), soit pour *effectuer des simulations* (ainsi : sur une image thématique représentant par exemple le prix du mètre carré par îlot, image créée à partir des données de la base et visualisée sur l'écran, on y définit directement une zonification que l'on veut tester en vue d'améliorer le découpage des taux d'imposition. Pour chaque îlot, on crée une nouvelle variable indiquant l'appartenance de l'îlot à telle ou telle zone, ce qui permet ensuite, par une opération d'agrégation de variables par îlot sur les zones, de calculer pour chaque zone un indice à partir des données dont on dispose sur les îlots, comme la densité de population, la densité du bâti, le revenu moyen des habitants, etc.).

La sortie cartographique sur papier se fait grâce à un traceur à plume ou un système de recopie d'écran. Les possibilités graphiques d'un traceur sont très différentes de celles d'un écran : meilleure définition graphique, grande taille et choix possible de l'échelle de sortie, mais peu de choix de couleurs ou de trames

efficaces. Il est souvent long et difficile d'obtenir un document d'une bonne qualité graphique, et plusieurs essais sont généralement nécessaires. Pour les cartes de travail, un bon compromis est offert par la recopie d'écran, noir et blanc ou couleur, qui, sur une taille réduite, donne de bons résultats graphiques et une facilité d'utilisation incontestable.

## 5. CONSTITUTION DE LA BASE

Gestion, traitements, représentation des résultats, pour cela il faut bien sûr commencer par constituer la base de données sur la ville de Quito : c'est également un aspect important de notre travail. Les données doivent être recensées, évaluées (intérêt, coût, volume, fiabilité), avant d'être saisies et intégrées dans la base de données.

L'information géographique pose le problème de la *saisie de la localisation*. Nous avons donc développé un programme spécifique de saisie graphique, MYGALE/PC. Conformément au principe de la gestion relationnelle, chaque thème est saisi séparément, même si plusieurs thèmes sont présents sur une même carte.

La saisie graphique avec MYGALE/PC se fait sur une *table à digitaliser*, munie d'une loupe de grande précision (la saisie se fait à la précision de 0,1 mm) et reliée à un micro-ordinateur. Le programme offre une saisie graphique supervisée (contrôle de la fermeture des zones, de la cohérence graphique) ainsi que de nombreuses fonctions d'édition et de correction graphique. Il permet de saisir des zones, des lignes, des points, ainsi que des éléments non thématiques pour constituer des fonds cartographiques.

Si le domaine à saisir est grand, il peut être *découpé en feuilles* qui seront digitalisées séparément. C'est l'intégration des données graphiques dans la base et le recalage en coordonnées géographiques qui assurera le bon placement de chaque feuille saisie dans l'ensemble géographique.

La base de données sur Quito se compose actuellement des données suivantes :

- données d'occupation du sol par îlot. Le découpage des îlots propre aux données de la Mairie de Quito a été saisi : 6500 îlots, saisis en 110 feuilles distinctes à l'échelle 1:2 000. La saisie a été effectuée d'avril à septembre 1989. Après formation des trois opérateurs, l'ensemble des îlots a été re-dessiné sur support indéformable puis digitalisé, le temps de saisie sur la table étant légèrement inférieur au temps de dessin. Chaque feuille a ensuite été vérifiée (clés, digitalisation des arcs), puis corrigée si nécessaire (durée deux mois) ;

- données de population et logement du recensement de 1982. La saisie des 6200 îlots du recensement de 1982 de l'INEC (Institut national équatorien de statistique) a été effectuée par modification du fond municipal, dessin des modifications d'après les plans de l'INEC, digitalisation de ces modifications et changement des clés, vérification et correction (décembre-avril 1989). Disposant des données par individu, chaque variable a été sommée par îlot pour un certain nombre de ses modalités, le résultat devenant un attribut sur l'îlot ;

- données sur les activités en rez-de-chaussée, à partir du recensement effectué en 1987. Chaque côté d'îlot a été répertorié et un point saisi pour localiser les activités s'y trouvant (35 000 points localisés sur les plans au 1:2 000 et saisis, mai-juillet 1989) ;

- équipements de santé, par point ;
- établissements d'enseignement, par point ;
- banques, entreprises, hôtels, loisirs ;
- géologie et usage du sol sur l'aire métropolitaine (avril 1989) ;
- divers autres découpages de la ville : secteurs, quartiers (juin 1989)

permettant, après des opérations d'agrégation (par exemple la moyenne par secteur de la densité de population de chaque îlot du secteur), de représenter un phénomène dans divers découpages d'échelles différentes ;

- est en cours de saisie, l'ensemble des points cotés, donnant l'altitude dans la ville (à peu près un point par croisement de rues), ainsi que les courbes de niveaux à 10 mètres sur l'aire métropolitaine ;

- le dessin du réseau d'égout est en cours de préparation, ainsi que l'élaboration des données par tronçon. Idem pour le réseau d'adduction d'eau.

L'ensemble SAVANE (MYGALE, TIGRE, ainsi que les différents modules d'introduction et de modification de données) constitue dès lors un système d'information qui devra nous permettre de répondre aux objectifs scientifiques du programme de recherche sur Quito. Les développements informatiques à venir iront dans trois directions :

- *faciliter la gestion urbaine* en produisant information et simulation, en fonction des demandes émanant essentiellement de la Mairie de Quito ;
- *intégrer et gérer de nouvelles sources de données*, comme les photographies aériennes, en utilisant si besoin est de nouveaux moyens de stockage comme les CD-ROM ;
- *rechercher des méthodes d'actualisation* pour les données graphiques.

## 6. LA CONFIGURATION MATERIELLE

Fonctionnalités des logiciels et possibilités des matériels sont étroitement liées. Pour assurer leur cohérence, le programme nécessite une configuration informatique qui doit répondre à plusieurs critères :

- posséder suffisamment de *puissance* pour gérer un important volume de données, traiter la localisation, manipuler des images ;
- avoir un *environnement graphique* performant (écran couleur de haute résolution, périphériques graphiques classiques) ;
- décentraliser la saisie graphique sur des postes autonomes.

De plus, l'équipement doit être d'un coût total permettant son investissement rapide, aussi bien pour l'Orstom que pour nos partenaires.

Face à ces contraintes, nous avons composé une configuration informatique qui se compose des matériels suivants :

- une *station de travail* MATRA-SUN 3/110C, munie de 4 Mo de RAM, d'un disque dur de 140 Mo, d'un écran graphique 1100\*900 pouvant afficher 256 couleurs sur une palette de 16 millions, d'une console alphanumérique, d'une imprimante. La puissance du processeur (2 Mips) , le système UNIX, et le concept de station de travail en font une machine très bien adaptée aux fonctionnalités requises pour la gestion des données et leur manipulation graphique sur écran. Seule la capacité du disque est insuffisante ;

- un *micro-ordinateur compatible AT* muni d'une carte EGA, d'une imprimante, et reliée à un *digitaliseur* BENSON 6301 (surface utile 1m20\*0m80) pour la saisie graphique. Le micro est relié à la station de travail par une voie asynchrone pour le transfert des fichiers de saisie ;

- un *traceur à plume* BENSON 1333 ;

- un *système de recopie d'écran* Teta-scan à transfert thermique, relié à la sortie vidéo de la station de travail.

L'ensemble de cette configuration, hormis la recopie d'écran et la console alphanumérique, a été mis en place début mars 1988, l'installation électrique ayant été fournie par l'IGM. Jusqu'à présent, ce matériel nous a donné entière satisfaction, à l'exception d'une panne du digitaliseur, machine pourtant des plus fiables; sa remise en marche n'a pas nécessité d'intervention technique spécialisée. Il va sans dire en effet que le manque de maintenance sur place constitue le principal problème technique auquel nous pouvons être confronté, et qu'aucune réponse satisfaisante ne peut actuellement lui être apportée.