

INFORMATION GEOGRAPHIQUE
ET CARTOGRAPHIE THEMATIQUE :
LE SYSTEME SAVANE

Marc SOURIS

ORSTOM
LIA, unité Infographie
70-74, route d'Aulnay, 93140 BONDY (France)

INTRODUCTION

L'ORSTOM est un organisme de recherche scientifique et technique où la multidisciplinarité est essentielle. Depuis longtemps et dans de nombreux pays, l'institut a été amené à intervenir dans le cadre d'études pluridisciplinaires, « inventaires » d'espaces régionaux ou nationaux, avec pour objectif final de fournir aux décideurs (aménageurs ou planificateurs) les données nécessaires à une bonne gestion de l'espace. Le cas de l'Equateur est exemplaire : une équipe importante a pendant plus de dix années travaillé à l'élaboration de l'inventaire national des ressources naturelles, récoltant ainsi une masse d'informations qu'il est difficile d'exploiter avec des méthodes traditionnelles.

Dans ce contexte, il est alors nécessaire de posséder un système de gestion de l'information qui puisse permettre de répondre aux questions posées par ces décideurs, comme :

- où cultive-t-on telle plante et quelle surface occupe-t-elle ?
- quelles sont les zones les plus favorables (compte tenu des facteurs physiques, économiques, sociaux...) pour développer telle culture ?
- pour quel quartier faut-il développer les transports, en fonction des activités existantes et à venir ?
- etc.

C'est en étant confronté à cette problématique, d'abord en Equateur, puis à travers les discussions avec les chercheurs de l'institut et avec nos partenaires étrangers, que le laboratoire d'infographie a été amené à développer le système que nous présentons ici, système qui s'articule autour de trois idées :

- une gestion puissante et souple des données géographiques localisées,
- une cartographie thématique, simple et rapide,
- un système autonome et permettant toute évolution de l'information.

L'application du système développé dépasse bien entendu le cadre des pays en voie de développement. Elle intéresse en fait toute communauté ayant à

gérer un espace et à déterminer des orientations sur l'affectation des différents éléments qui le composent.

Une gestion souple et puissante de données géographiques localisées

Les applications de l'informatique à la géographie et la cartographie sont diverses : saisie, stockage, restitutions graphiques. L'information géographique est complexe et variée : données multiples, hétérogènes, de sources très diverses (cartes, télédétection spatiale, relevés de terrain, données statistiques...). Ces données n'ont souvent aucun lien a priori entre elles autre que leur localisation dans l'espace, qui peut être donnée sous des formes variées. Il est néanmoins nécessaire d'avoir à sa disposition un système de gestion et d'exploitation de données qui permette de mettre en relation ces données à tout moment sur des critères définis lors de l'interrogation et sans avoir à modifier ces données : c'est le principe de la gestion relationnelle des données. Le cas des données localisées rentre dans ce cadre : on doit pouvoir comparer des objets sur leur localisation sans que ces objets soient a priori reliés entre eux par une description nominale particulière commune de la localisation (comme un nom de département, un nom de rue...) ce qui est d'ailleurs souvent impossible. Il nous a donc fallu étendre le principe de la gestion relationnelle à « l'attribut localisation » description logique de la localisation des objets que ce soient des zones, des lignes, des réseaux, des points et indépendamment de leur mode de stockage et de représentation. Grâce à de nouvelles opérations de gestion, le traitement de la localisation se banalise et rentre dans le cadre général de la gestion relationnelle. Cela donne un système puissant (toutes les données se trouvant sur un territoire sont comparables entre elles à tout moment) et souple (c'est l'utilisateur qui définit lui-même au moment de l'interrogation les cheminements dans la base de données en ayant la possibilité d'utiliser la localisation intrinsèque des données). Il est ici possible de respecter l'implantation spatiale de chaque objet : il n'est plus nécessaire de définir un maillage unique pour représenter des objets divers d'échelles de validité différentes. D'autre part l'utilisateur manipule ces objets aussi bien sur leurs contenus descriptifs que sur la localisation : il a ainsi accès à l'ensemble des opérations classiques de gestion et de traitement, à sa demande : sélection, regroupement, classification, traitements statistiques...

Une cartographie thématique simple et rapide

Si la localisation est utilisée dans le processus de gestion de l'information, elle permet également la représentation cartographique des résultats d'une interrogation, au même titre qu'une sortie de résultats sur une imprimante. C'est « l'attribut localisation » qui est alors représenté comme résultat du processus de gestion. En permettant la production rapide de cartes simples, le système d'information géographique ouvre le champ de la cartographie à une nouvelle catégorie d'utilisateurs : plutôt qu'une carte de grande précision rassemblant dans sa légende complexe de nombreuses informations, mais d'une réalisation en tout état de cause longue et coûteuse, l'utilisateur du système demande des séries de cartes répondant chacune à une question précise et produites immédiatement. D'autre part, l'interactivité graphique permise par les matériels informatiques

donne à l'utilisateur la possibilité de construire la représentation d'un phénomène en fonction du but recherché : la réalisation d'une carte repose en effet sur une série de choix difficiles sinon impossibles à systématiser, tel le choix des trames, telle la mise en évidence des contrastes... L'utilisateur a ici l'entière maîtrise de ces choix. En face de l'écran couleur, il peut essayer successivement une série de palettes, de trames, de symboles, avant de choisir celle qui lui paraît la plus pertinente pour bien représenter la réponse à la question qu'il avait posée.

Un système autonome et permettant toute évolution de l'information

L'utilisateur devient propriétaire de ses données et capable de les mettre en œuvre et d'exploiter le système sans le concours de cartographes ou d'informaticiens.

Enfin, l'information disponible évolue sans cesse. Les études s'accumulent sur une région déterminée, apportant de nouvelles données, affinant la connaissance du milieu et de son fonctionnement. La télédétection spatiale fournit également une importante et régulière quantité d'information. Ces données nouvelles peuvent être intégrées de façon simple dans le système, de manière à compléter ou remplacer les anciennes.

PRESENTATION DU SYSTEME SAVANE

Le système comprend deux parties distinctes : la constitution de la base de données (MYGALE), l'exploitation de la base de données (TIGRE).

Avant toute informatisation, le schéma des données doit cependant être bien défini : quelles sont les familles-objets en présence (nous appellerons famille une relation), pour chaque famille quels sont les attributs de ces objets (par exemple, des parcelles forment une relation, avec des attributs comme le numéro de la parcelle, le nom du propriétaire, la catégorie du P.O.S... et les constructions forment une autre relation, avec des attributs comme numéro du permis de construire, nombre d'étages, etc.).

Constitution de la base de données

L'information géographique localisée peut être séparée en information graphique (contours des zones, lignes des réseaux, coordonnées des points) et information descriptive (valeurs des attributs de chaque objet).

Si la saisie de l'information descriptive se fait d'une manière classique sur grille d'écran sur micro-ordinateur, l'information graphique est saisie à l'aide d'une table à digitaliser reliée elle aussi à un micro-ordinateur. Un contrôle interactif de cohérence topologique, ainsi que de multiples possibilités de correction permettent d'obtenir une numérisation graphique exempte d'erreurs, associée à une facilité de manipulation pour l'opérateur de saisie.

La saisie se fait par coupure de carte, indépendamment les unes des autres. Des points de calages, puis un recalage en coordonnées géographiques dans la base de données permettent d'assurer la cohérence et la bonne juxtaposition des coupures dans la base. Ce découpage en coupures n'intervient plus dans

la phase d'interrogation, où l'utilisateur définit lui-même sa fenêtre d'étude indépendamment du processus de saisie.

Si information descriptive et information graphique sont saisies séparément, le processus d'intégration dans la base assure la liaison graphique-descriptif : une fois l'information intégrée dans la base, l'utilisateur n'a que la vision d'objets où graphique et descriptif forment un tout.

Exploitation de la base de données

L'interrogation se fait de façon interactive sur un terminal. L'exploitation de la base se fait sous forme d'une requête, en appelant des modules successifs : le résultat d'un module sert d'entrée au module suivant (par exemple on sélectionne les départements de plus 100 000 habitants, puis pour ces départements on classe le nombre de sans-emploi, etc.).

Si les données graphiques sont structurées et stockées dans la base sous forme vectorielle (descriptifs de points ou de contours), le système va utiliser pour la réalisation de certaines opérations une structure matricielle sous forme d'une image numérique de pixels associée à un fichier descriptif. Chaque pixel est relié au fichier descriptif par sa valeur numérique qui permet de retrouver les valeurs descriptives qui lui correspondent.

La taille de la matrice de pixels est paramétrable jusqu'à 1 200 x 1 200 ; la résolution du pixel dépend alors de la fenêtre géographique utilisée. Le passage de vectoriel à matriciel est un des éléments importants du système ; néanmoins, il reste au niveau interne et s'effectue d'une manière interactive lors de l'opération qui le requiert. Cette double structure (stockage vectoriel et manipulation matricielle), associée au paramétrage de la résolution matricielle, confère au système une souplesse importante dans la réalisation des différentes opérations de gestion et de manipulation de la localisation.

Les différents modules du système

L'utilisateur a le choix de la résolution en pixels des images balayées créées par le système lors des interrogations (*CONFIG*), le choix de la projection géographique de travail et de restitution (U.T.M., Lambert, Mercator, module *MAP*), le choix de la fenêtre géographique d'étude (*WIND*). Cette fenêtre, qui permet de définir l'espace géographique sur lequel porteront les opérations relatives à des données localisées, peut être choisie directement (longitude, latitude), par le choix d'une coupure, ou par le choix de la taille du pixel pour la projection et la résolution courante.

Le module *QUEST* regroupe les opérations algébriques d'interrogation des données relationnelles : restriction, projection algébrique, jointure, restriction spatiale, projection spatiale (transformation vecteur-laser), jointure spatiale. D'autres modules permettent de manipuler les relations (réinitialisation de la base après requête, sauvegarde temporaire). Ce module est donc à la base des opérations thématiques pour sélectionner, croiser, comparer des objets.

Le module *THEM* est la synthèse de plusieurs opérations de base : il permet la création de relations thématiques et multithématiques par combinaison d'attributs provenant d'une ou plusieurs relations zonales.

Le module *LIST* permet d'imprimer le contenu descriptif d'une relation.

Le module *STAT* permet d'effectuer des calculs statistiques simples sur les attributs (moments, histogrammes, corrélations, régressions). Ces calculs peuvent se faire à tout moment de l'interrogation et permettent ainsi de modifier, si besoin est, le cours de l'interrogation (pour le choix des seuls de classification par exemple).

Classifications, calculs, statistiques, apport d'information extérieure : les modules qui suivent permettent, à partir des données stockées dans la base, de modifier ces données ou d'en créer de nouvelles :

CLAS : classification d'un attribut. Permet de passer d'un attribut numérique à un attribut qualitatif nominal, ou de redéfinir des valeurs d'un attribut qualitatif par regroupement de valeurs.

CRIS : création d'attribut par calcul numérique, calcul logique, combinaison de valeurs, à partir des attributs existants.

STUP : création d'attribut par calcul statistique d'un attribut par rapport à un autre.

SURF : calcul de surfaces ou périmètres pour les objets d'une relation zonale.

COCA : création d'attribut pour une relation zonale par calcul résultant du croisement spatial avec une autre relation : par exemple, calcul pour chaque zone d'une première relation (zonale) de la moyenne de valeurs numériques associées aux points de la seconde relation (ponctuelle).

STAD : création d'attribut pour une relation zonale par calcul résultant du croisement spatial avec une image numérique : moyenne de radiométrie...

La cartographie des résultats se fait sur écran graphique et sur traceur. Un système de recopie d'écran couleur permet également d'avoir à tous moments, et très rapidement, une image de la carte en cours ou en fin de traitement.

PALET : création ou modification du fichier de correspondance valeur d'attribut-index de couleur ou de plume, numéro de trame.

CART : création de segments graphiques pour le tracé cartographique : tracés d'arcs, de symboles, de textes... Le module crée un segment graphique qui sera interprété par le module de tracé propre au périphérique de dessin utilisé.

SUN : tracé sur écran graphique. Permet le tracé d'images de relation, de segments graphiques, la création et l'utilisation de palettes de couleurs, la manipulation interactive sur l'écran (textes, caissons de légendes, habillage, calculs de distances et de surfaces...).

BENSON : tracé sur traceur à plume. Permet le tracé d'images et de segments graphiques, avec choix de l'échelle de restitution, habillage des documents (cadre, titre, amorces de projection...), légende associée.

Description du matériel utilisé

L'ordinateur utilisé est une station de travail SUN 3/110 (4Mo RAM 140 Mo disque, écran graphique couleurs 1150*900) et un micro-ordinateur compatible AT avec écran EGA (pour la saisie graphique). Le matériel graphique est composé d'une table à digitaliser BENSON 6301, une table BENSON 1333, un système de recopie d'écran D-SCAN. Le coût global de cette configuration est de l'ordre de 500 KF.

Une configuration plus réduite est en cours de réalisation : le traitement et la visualisation se fera entièrement sur micro-ordinateur. Si le coût ne dépassera pas les 35 KF, les possibilités de visualisation et de tracé seront limitées, et la saisie des données graphiques devra être traitée par le service de numérisation graphique de l'unité d'Infographie au centre ORSTOM de Bondy.