

## LE PROJET SISGEO (EMBRAPA, BRESIL):

### ETAT D'AVANCEMENT

Patrick SECHET, informaticien  
Département E, U.R. 502.

Le système SISOLOS de stockage et récupération de données de sols a été élaboré par l'Entreprise Brésilienne de Recherche Agropastorale (EMBRAPA), avec la collaboration d'un informaticien de l'ORSTOM dès 1981, pour faciliter l'accès à l'important volume des données pédologiques collectées, analysées, interprétées et publiées par le Service National d'Inventaire et Conservation des Sols (SNLCS/EMBRAPA). A l'origine, le système prévoyait seulement un fichier : les profils de sols pour lesquels sont stockées plusieurs catégories d'informations (conditions de l'environnement du profil, description morphologique de chaque horizon et analyse minéralogique de chaque fraction granulométrique). Pour exploiter le caractère géographique des informations fournies par les levés de sols, SISOLOS a été étendu et contient une deuxième collection constituée de l'information descriptive (caractéristiques physiographiques et composition taxonomique) attribuée aux unités cartographiques des cartes de sols. Avec le volume d'information que contient la banque de données (près de 9500 profils et plus de 2 200 unités cartographiques, en décembre 1984), le système est capable de supporter de nombreuses applications, essentiellement axées sur l'exploitation statistique des informations contenues. Toutefois, pour compléter la vocation géographique de SISOLOS, l'élaboration (en cours) et l'application de techniques d'infographie pour le stockage de cartes de sols est encore nécessaire. Ce dernier développement est réalisé dans un contexte plus général de Système d'Informations Géographiques à l'EMBRAPA, SISGEO, dont SISOLOS constitue la première composante.

### SISOLOS : quelques caractéristiques

SISOLOS a été conçu et développé par des informaticiens ayant une expérience d'élaboration de systèmes automatisés à vocation administrative : l'étape conceptuelle importante de structuration de l'information y a été particulièrement soignée. Plus de 500 variables ont été identifiées, répertoriées, reliées entre elles et incorporées au schéma.

En terme de stockage ce système a pour caractéristique essentielle d'utiliser les ressources traditionnelles de l'informatique, compatibles avec un système scientifique d'utilisation restreinte, sans toutefois compromettre son efficacité. L'entrée des données s'opère par transcription des originaux (publication, dans la majorité des cas), sur des formulaires spécifiques, avec codification par utilisation de tables évolutives. Cette phase est de loin la plus importante

du système et tout est fait pour améliorer le rendement (quantitatif et qualitatif) de l'équipe de transcription (10 personnes).

Après remplissage, les bordereaux sont envoyés par lots à la saisie, en vue de la critique des "mouvements" et de l'actualisation de la banque (et éventuellement des tables de codification).

Toutes les opérations qui interviennent dans la phase de stockage, y compris l'élaboration de copies de sécurité, sont dûment documentées et sont ainsi réalisées comme n'importe quelle autre tâche de routine du Centre de Traitement de Données de l'ENBRAPA.

SISSOLOS établit un équilibre entre la souplesse de stockage et les facilités de récupération, en incorporant une phase intermédiaire de réorganisation des données. Il est ainsi possible d'avoir une vision désagrégée de l'information, confortable au moment du stockage, tout en gardant une vision intégrée propice à la récupération.

Le système permet n'importe quelle extraction de données compatible avec la structure logique de l'information pédologique sur laquelle le modèle conceptuel repose. L'interrogation peut être formulée en utilisant un quelconque sous-ensemble des variables disponibles dans le dictionnaire, à l'aide d'un critère simple ou composé.

L'utilisation, comme langage d'interrogation, d'un progiciel statistique puissant (SAS, Statistical Analysis System), amplement répandu parmi les utilisateurs potentiels de l'ENBRAPA, garantit un accès facile à l'information stockée. L'exploitation des données s'opère en deux étapes : la première de sélection des variables et des observations à utiliser, de façon à produire le sous-ensemble des informations à étudier ; la seconde de traitement proprement dit de ces données par le biais des nombreuses ressources de ce logiciel (statistiques, histogrammes, graphiques, etc.).

Un manuel d'utilisation a été publié pour orienter les utilisateurs potentiels de SISSOLOS, et contient toute l'information nécessaire à l'exploitation du système : dictionnaire des données, liste et contenu des tables de codification, modalités pratiques d'accès, exemples d'utilisation, etc.

En plus du stockage et de la récupération, SISSOLOS permet une augmentation continue de la connaissance critique des données, à partir d'analyses statistiques. Cette compréhension devra mettre en évidence des relations déterministes et stochastiques qui seront stockées dans le système sous la forme de programmes de critique et en un manuel de critique.

#### Aspect géographique de SISSOLOS

On s'est préoccupé, lors du développement du système, d'incorporer l'aspect géographique, ce qui constitue une caractéristique importante de SISSOLOS, rarement rencontrée dans un système thématique d'informations.

Ainsi, en ce qui concerne les données ponctuelles contenues dans les fichiers de la banque, la géocodification est réalisée par le rattachement des profils à différents référentiels géographiques (coordonnées géographiques, "municipio", commentaire de situation au niveau local, etc.). Il est donc possible d'extraire des cartes de localisation de profils, de sélectionner des informations à partir du découpage administratif municipal, etc.

Dans le cas des données zonales, d'autres applications géographiques sont possibles à partir des paramètres contenus dans le fichier des unités cartographiques : par exemple, il est théoriquement possible de déterminer l'aire totale occupée par un certain type de sol, à l'intérieur d'une région entièrement inventoriée.

Par croisement de ces deux types de données (rendu possible par le fait qu'un profil est relié à l'unité cartographique qui le contient géographiquement), ces applications peuvent être étendues.

Toutefois, c'est la réalisation de la dernière phase du projet, avec l'incorporation d'un fichier représentatif des cartes de sols, qui devra permettre une exploitation complète des données collectées. Avec cette nouvelle ressource, SISOLOS autorisera la réalisation de deux types de recherche :

- recherche par zone : à l'intérieur d'un contour donné, quelles sont les caractéristiques des sols rencontrés ?
- recherche par attribut : où rencontrer les sols possédant certaines caractéristiques ?

C'est par l'utilisation en cascade des trois fichiers du système que ces deux interrogations fondamentales pourront être traitées (voir figure).

Dans le cas d'une recherche à partir d'un lieu fixé, c'est la relation "carte" qui sera initialement exploitée, pour obtenir un identifiant d'unité cartographique : ensuite, la relation "légende" fournit les paramètres régionaux et, éventuellement, un pointeur (numéro de profil) qui fournira les paramètres locaux recherchés. On peut noter ici que le pointeur n'est pas indispensable dans la mesure où l'on peut procéder à une recherche du type "point contenu dans une zone".

Dans le cas d'une recherche à partir d'un caractère fixé, c'est exactement le chemin inverse qui sera parcouru, totalement si le paramètre est local (attribut de profil), partiellement si le paramètre est zonal (attribut d'unité cartographique).

Bien entendu, tous les problèmes sont loins d'être résolus : sans aucun doute les plus complexes sont d'ordre scientifique plutôt que technique. Par exemple, une interrogation portant sur une zone chevauchant deux cartes différentes devra s'accommoder de deux légendes généralement hétérogènes !

## De SISSOLOS à SISGEO

Le dessin d'un système de stockage et manipulation, par ordinateur, de cartes de sols, constitue un travail important et suppose des moyens matériels et une méthodologie qui devront être sérieusement étudiés. Etant entendu que cette problématique est commune à plusieurs thèmes, il a été jugé plus intéressant de considérer le développement de ce dernier travail comme une activité partielle d'un projet plus ample de système d'informations géographiques plurithématique, SISGEO, dont SISSOLOS constitue le premier module.

En effet, le croisement des informations pertinentes à des thèmes distincts s'effectue à partir d'une localisation géographique commune, par superposition des cartes thématiques correspondantes : une structure unique pour le stockage des cartes, quel que soit le thème, garantirait par conséquent cette possibilité de composition d'informations.

Toutefois, il serait totalement illusoire d'espérer réaliser un véritable système d'informations géographiques sans se soucier préalablement de la structuration et du stockage de l'information thématique, thème à thème. Ceci constitue le travail le plus important et doit être mené, pour le moins parallèlement, au développement de techniques infographiques adéquates à la manipulation des objets graphiques constitués par les cartes.

Actuellement, les responsables EMBRAPA/ORSTOM de SISGEO ont donc débuté deux nouvelles opérations complémentaires à la poursuite des travaux sur SISSOLOS à Rio de Janeiro : un système de climatologie agricole (SISCLINA) à Bélem et un système de saisie et de stockage de cartes (SISMAPAS) à Brasilia. L'incorporation de la géocodification est évidemment une préoccupation constante pour les systèmes thématiques.

La principale difficulté rencontrée dans le cadre de l'EMBRAPA pour l'élaboration d'un tel système au niveau national, est constituée par la structure organisationnelle de cette entreprise, qui rend difficile l'établissement d'une coordination unique par thème, excepté dans le cas du thème sol, pour lequel l'existence d'un service national explique l'avance obtenue.

APLICACOES GEOGRAFICAS  
COM USO DOS TRES ARQUIVOS

Busca por regio geografica



| Coordenadas geograficas | Simbolo de unidade de mapeamento |
|-------------------------|----------------------------------|
| Xo, Yo                  | Zo                               |
| - -                     | -                                |
| - -                     | -                                |
| Xi, Yi                  | Zi                               |
| - -                     | -                                |
| - -                     | -                                |
| Xn Yn                   | Zn                               |

Arquivo de mapas



| Simbolo de unidade de mapeamento | Parametros regionais | Identificacao dos perfis da unidade |
|----------------------------------|----------------------|-------------------------------------|
| Zi                               | Ai, Bi, Ci           | Pi                                  |
| -                                | - - -                | -                                   |
| -                                | - - -                | -                                   |
| Zj                               | Aj, Bj, Cj           | Pj                                  |
| -                                | - - -                | -                                   |
| -                                | - - -                | -                                   |
| Zm                               | Am, Bm, Cm           | Pm                                  |

Arquivo de unidades de mapeamento



| Identificacao de perfil | Parametros                                       |
|-------------------------|--|
| Pa                      | a <sub>a</sub> , b <sub>a</sub> , c <sub>a</sub> |
| -                       | - - -  |
| -                       | - - -  |
| Pk                      | a <sub>k</sub> , b <sub>k</sub> , c <sub>k</sub> |
| -                       | - - -  |
| -                       | - - -  |
| Pl                      | a <sub>l</sub> , b <sub>l</sub> , c <sub>l</sub> |

Arquivo de perfis

Busca por atributo regional local

