

## SISSOLOS : base de données de sols du Brésil

Patrick SÉCHET

*Ingénieur informaticien ORSTOM. EMBRAPA Divisão de Sistemas de Informação, Supercenter Venâncio 2000, 70333 Brasília, DF (Brésil)*

### RÉSUMÉ

*Le Service National d'Inventaire et de Conservation des Sols (SNLCS), de l'Entreprise Brésilienne de Recherche Agropastorale (EMBRAPA), réalise et coordonne, depuis près de trente ans, l'inventaire systématique des ressources en sols du pays. Cette action s'est traduite par l'édition de plus d'une centaine de publications (notice technique associée à la carte pédologique), où sont consignés les profils collectés sur une zone géographique dont les dimensions varient selon l'échelle et les objectifs du levé.*

*Afin de rendre les données correspondantes promptement accessibles en vue des applications les plus diverses, l'EMBRAPA, avec la participation de l'ORSTOM, a mis au point un système informatique de stockage de ces données, baptisé SISSOLOS. Les fichiers de ce système contiennent essentiellement deux collections :*

- les profils de sols, pour lesquels de nombreuses informations sont disponibles : conditions de l'environnement du profil, description morphologique et analyses physico-chimiques de chaque horizon, analyse minéralogique de chaque fraction granulométrique ;*
- les unités cartographiques de la carte des sols, généralement constituées d'associations de sols, plus ou moins complexes selon l'échelle de l'inventaire, et pour lesquelles la description des caractères physiographiques, la composition taxonomique et les paramètres relatifs à l'aptitude agricole sont stockés.*

*Toutes ces données sont logiquement regroupées par « travail », un travail correspondant, dans la plupart des cas, à une publication du SNLCS. Pour compléter le système actuel, le stockage informatique des cartes de sols est en cours de développement.*

*Les programmes nécessaires à l'exploitation des informations disponibles dans la base constituent deux sous-systèmes distincts : le premier permet la mise à disposition des données sous la forme de rapports descriptifs qui obéissent aux modèles définis par la Société Brésilienne de Science du Sol ; le second autorise le traitement des interrogations, simples ou composées, qui font intervenir n'importe quels attributs d'un profil, ou d'une unité cartographique.*

*SISSOLOS a été développé par une équipe d'informaticiens : le respect des concepts et normes adoptés par cette profession fait en sorte qu'il bénéficie d'un certain nombre d'avantages, généralement réservés aux systèmes administratifs ou para-administratifs, susceptibles d'assurer sa pérennité et son utilisation au niveau national. Ainsi, le système présente deux caractéristiques essentielles : il utilise des ressources traditionnelles en informatique, sans que son efficacité s'en trouve dégradée et il respecte l'équilibre entre les facilités de stockage et celles d'extraction.*

*Le stockage des données publiées commence par une étape de transcription : en mai 1985, près de 11 000 profils étaient disponibles, ce qui représente environ 80 % de l'information rassemblée par le propre SNLCS. Des conventions avec d'autres organismes collecteurs ont été élaborées et visent la création de nouveaux centres de transcription, ce qui permet de prévoir à court terme l'agrégation, dans les structures de SISSOLOS, de la quasi-totalité de l'information pédologique du Brésil.*

**MOTS-CLÉS :** Brésil — Système d'information — Sols — Informatique — Pédologie.

## RESUMO

## SISSOLOS : BASE DE DADOS DE SOLOS DO BRASIL

O Serviço Nacional de Levantamento e Conservação de Solos (SNLCS), da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), realiza e coordena, há quase trinta anos, o inventário sistemático dos recursos em solos do país. Este trabalho já gerou mais de cem publicações (boletins técnicos com os respectivos mapas de solos) onde se encontram os dados pontuais (perfis de solos), coletados numa área geográfica cujas dimensões variam com a escala e os objetivos do levantamento.

No intuito de tornar esses dados rapidamente acessíveis às mais diversas aplicações, a EMBRAPA, com a colaboração do ORSTOM, desenvolveu um sistema de armazenamento e recuperação dessa informação, denominado SISSOLOS. Os arquivos deste sistema contêm essencialmente duas coleções de objetos :

- os perfis de solos para os quais são disponíveis várias categorias de informações ; condições ambientais do perfil, descrição morfológica, análises físicas e químicas de cada horizonte, análises mineralógicas de cada fração granulométrica ;
- as unidades de mapeamento de solos, geralmente constituídas de associações mais ou menos complexas de solos (dependendo da escala do levantamento), para as quais a descrição das características fisiográficas, a composição taxonômica e os parâmetros de aptidão agrícola estão armazenados.

Todos esses dados estão logicamente agrupados por « trabalho », um trabalho correspondendo geralmente a uma publicação. Para completar o sistema atual, o armazenamento dos mapas de solos se encontra em fase de desenvolvimento.

Os programas necessários à recuperação dos dados armazenados constituem dois subsistemas distintos : o primeiro possibilita o acesso a informação contida no Banco sob a forma de relatórios descritivos que obedecem aos padrões definidos pela Sociedade Brasileira de Ciência do Solo ; a segundo subsistema autoriza o processamento das consultas, simples ou compostas, que envolvem quaisquer atributos de um perfil, ou de uma unidade de mapeamento.

SISSOLOS foi desenvolvido por uma equipe de técnicos de processamento de dados e, por respeitar os seus conceitos e normas de trabalho, beneficia de um certo número de vantagens, geralmente reservados aos sistemas administrativos ou para-administrativos, suscetíveis de garantir sua perenidade e uma utilização abrangente. Assim, o sistema apresenta duas características essenciais : ele usa recursos tradicionais de informática, sem prejuízo da eficácia e também mantém equilíbrio entre as facilidades de armazenamento e as de recuperação.

O armazenamento dos dados publicados tem início com uma fase de transcrição. Em maio de 1985, quase 11.000 perfis já estão disponíveis, o que representa aproximadamente 80 % da informação coletada pelo próprio SNLCS. Convênios com outras instituições, visando a criação de novos centros de transcrição, já foram firmados, o que permite esperar a agregação a curto prazo nas estruturas do SISSOLOS, da maior parte da informação pedológica disponível no Brasil.

PALAVRAS CHAVES : Brasil — Sistema de informação — Solos — Informática.

## ABSTRACT

## SISSOLOS : SOIL INFORMATION SYSTEM OF BRAZIL

The Serviço Nacional de Levantamento e Conservação de Solos (SNLCS) of Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), has been carrying out and coordinating, during the last thirty years, the systematic inventory of soil resources of the country. At present, this work has provided more than one hundred of publications (technical bulletins with the soil maps), where point data (the soil profiles) collected from a geographic zone are available.

In order to support many applications, the EMBRAPA, with the collaboration of ORSTOM, has developed a soil information system, named SISSOLOS, to make easier the access to this information. The files of this system include essentially two collections. The first one includes soil profiles for which there are many categories of data : environmental conditions of the profiles, morphological description, chemicals and physicals analysis of each horizon, mineralogical analysis of particle fraction. The second one includes the soil mapping units, generally composed of more or less complex associations of soil classes (depending on the scale of the survey), for which are stored the description of physical characteristics, taxonomic composition and some agricultural capabilities.

All these data are logically grouped by « work », corresponding to one SNLCS Technical Bulletin. In order to complete the actual system, the storage of soil maps is being developed.

The access of the information available in SISSOLOS is possible through two different subsystems : the first one permits the generation of standard reports according to models designed by Brazilian Soil Science Society, and the second one permits the interrogation processing, single or composed, encompassing any characteristics of a profile, or a mapping unit.

The system has been developed by a team of professionals of data processing, following the work concepts of this area. It presents some advantages, generally reserved to administrative systems, which will grant its continuous and universal utilization. Particularly, SISSOLOS have two essential characteristics : the utilization of traditional resources of data processing, without losing efficiency and, also, maintaining an equilibrium between the facilities of storage and those of retrieval.

The process of storing published data begins with a stage of transcription. At present (may 1985), about 11,000 soil profiles are already available, that represents approximatively 80 % of the information collected by SNLCS. In order to create new pools of transcription, some contracts were signed with order research institutes of the country, which will permit to aggregate, in a short delay, almost of soil information already available in Brazil.

KEY WORDS : Brazil — Information system — Soils — Data processing — Soil science.

## INTRODUCTION

Le Brésil, avec ses huit millions et demi de kilomètres carrés, est un pays dont l'économie s'appuie fortement sur l'agriculture et pour lequel la connaissance de la ressource « sol » est fondamentale. L'évaluation du potentiel offert par une zone déterminée à l'exploitation de telle ou telle culture, l'établissement d'une politique d'occupation adéquate ou encore la définition de normes rationnelles de travail et conservation des sols, sont fondées sur les levés pédologiques et la caractérisation de l'aptitude potentielle des terres.

Au côté de nombreuses autres institutions, pour la plupart de caractère régional, le Service National d'Inventaire et de Conservation des Sols (SNLCS), rattaché à l'Entreprise Brésilienne de Recherche Agropastorale (EMBRAPA) du Ministère de l'Agriculture, a été créé, au début des années 1950, pour étudier scientifiquement les sols du pays. Parmi les recherches menées, la cartographie des sols est l'objectif vers lequel converge une grande partie des activités du service, étant entendu qu'elle constitue un procédé traditionnel de collecte d'informations pédologiques.

La connaissance du sol, et de son comportement, dépend de sa caractérisation chimique, physique, minéralogique et micromorphologique, qui sont à la charge d'autant de secteurs spécialisés du service, et dont les déterminations fournissent les données nécessaires à l'identification des sols observés sur le terrain. Les résultats rassemblés lors du levé des sols d'une zone déterminée sont consignés dans une publication éditée par le SNLCS, constituée d'une notice technique accompagnée d'une carte de sols, susceptible d'être utilisée par d'autres études scientifiques, travaux appliqués, finali-

tés pratiques ou études détaillées sur tel ou tel aspect particulier.

La nécessité d'utiliser au mieux le volume croissant des données de sols accumulées au cours de ces inventaires, d'autant plus que ceux-ci constituent une activité assez dispendieuse, a conduit le SNLCS, dès 1979, à entreprendre le développement d'un système d'information apte à en faciliter l'accès. L'ORSTOM a collaboré à ce projet, baptisé SISSOLOS, par le biais de l'affectation d'un ingénieur informaticien, qui s'est associé à l'équipe brésilienne du projet de 1981 à 1984.

Le présent article prétend mettre en évidence les points les plus intéressants de ce système, en soulignant particulièrement les solutions techniques générales adoptées, susceptibles d'expliquer le succès rencontré par le projet au niveau national. Au cours de la rédaction de cette note, on a surtout cherché à présenter l'esprit qui a guidé l'ensemble du travail, en expliquant les motifs qui justifient chaque option choisie.

Le premier paragraphe décrit le contexte institutionnel qui a prévalu pour le développement du projet ; les données fournies par les levés de sols sont ensuite brièvement révisées, dans l'optique d'expliquer plus clairement les solutions retenues pour le traitement.

Le système archive et rend disponible toutes ces données. Les fonctions de stockage et d'extraction présentent des caractéristiques propres, et correspondent à des visions différentes de l'information : rendre compte de ces caractéristiques, en mettant l'accent sur l'interaction utilisateur-système, constitue une partie essentielle de ce document. L'intégration stockage-extraction est réalisée à l'aide d'une phase de réorganisation de la base de données, qui est présentée ensuite.

Avec le volume d'information atteint, et les outils disponibles pour la traiter, SISSOLOS autorise de nombreuses applications. Une part importante de cet exposé a été réservée à la description des applications du système, tant celles en cours de réalisation que celles sur le point d'être initiées. Il convient d'ailleurs de percevoir l'ensemble d'une façon intégrée, en considérant que le système initial complété par ses applications actuelles constitue une nouvelle génération du système. Cette dernière est elle-même en évolution vers un véritable système-pédologue, détenteur d'une irremplaçable connaissance sur les sols brésiliens.

En mai 1985, près de 11 000 profils de sols, et plus de 2 500 unités cartographiques sont accessibles par le système, ce qui représente sensiblement plus de 80 % du total de l'information disponible du SNLCS. Compte tenu du flux de données collectées ces dernières années (équivalent à 500 profils l'an), et des inventaires réalisés par d'autres entités, le système devrait se stabiliser d'ici deux ou trois ans, une fois traitée toute l'information historique disponible. Il permettra alors l'exploitation intégrée de la quasi-totalité de l'information pédologique fournie par la cartographie brésilienne des sols, pour des applications inhérentes à tous les segments de la science du sol.

## 1. LE CONTEXTE

Le SNLCS a été créé à Rio de Janeiro en 1951, sous le nom de Commission des Sols, pour produire, adapter et normaliser les méthodologies utilisées en science du sol ; centraliser, mettre à jour et divulguer toute l'information pédologique disponible et exécuter la cartographie des sols et l'interprétation de l'aptitude agricole des terres sur tout le territoire brésilien. D'autres institutions réalisent des levés de sols ; le Projet RADAMBRASIL, en particulier, qui est le seul autre organisme à étendre ses activités au niveau national, mais aussi des institutions régionales ou d'état (l'IAC, Institut Agronomique de l'Etat de São Paulo, par exemple), des Universités, et, dans une moindre mesure, des entreprises privées.

La force de travail allouée à la cartographie des sols brésiliens peut être évaluée par le contingent de pédologues que totalisent tous ces organismes, lui-même estimé à moins de 150 personnes (EMBRAPA, 1981). Bien entendu, l'exécution de ces inventaires, à des échelles diverses, ne satisfait pas la demande, et traditionnellement les levés de sols sont effectués sous l'impulsion des institutions publiques de développement régional, ou des gouvernements de chaque état, ou encore d'autres administrations sectorielles chargées de promouvoir le développement de la production d'une culture déterminée.

Avec la création de l'EMBRAPA, en 1973, et par le biais du Programme National d'Inventaire des Sols coordonné par le SNLCS, des directives ont pu être établies pour que les efforts d'évaluation de la ressource « sol » se fassent de manière plus homogène et soient fondés sur une méthodologie uniforme. Annuellement, les projets financés sur ce programme sont harmonisés par la communauté scientifique représentant les organismes concernés. Il est ainsi possible de procéder à l'inventaire systématique des sols du pays et à la préservation de cette ressource, tout en évaluant ses potentialités pour l'implantation de nouveaux projets, l'augmentation de la productivité et la rationalisation de l'expansion de la frontière agricole.

Au fur et à mesure que les travaux de cartographie pédologique progressent dans le pays, l'ensemble des profils décrits avec les informations morphologiques et analytiques correspondantes augmente ; par ailleurs, l'analyse globale d'un grand volume de données est une tâche pratiquement impossible à réaliser sans le concours de l'ordinateur, qui favorise en particulier les évaluations statistiques.

Aussi, dans le cadre de ce même programme, le projet de traitement informatique de données de sols a-t-il été approuvé en 1979, dans l'espoir de permettre une meilleure exploitation de toute la masse de données déjà collectée, et de servir de base à de nouvelles études, que la dissémination de l'information dans de nombreuses publications rendait jusqu'alors impossible. Pour le développement de cette activité, le SNLCS a pu s'appuyer sur les compétences d'un analyste de systèmes du Département de Méthodes Quantitatives, et compter sur la collaboration d'un ingénieur informaticien de l'ORSTOM, qui est venu s'intégrer à l'équipe de statisticiens du Service.

## 2. LES DONNÉES

Le système de stockage et d'extraction de données pédologiques SISSOLOS a été développé pour faciliter l'accès aux informations accumulées lors des inventaires réalisés dans tout le pays. Une révision de la nature et de la forme de présentation de ces informations s'impose donc, dans la mesure où les structures du système respectent les relations existantes entre les données, plutôt que de s'inspirer de la vision adaptée à telle ou telle application.

### 2.1. Données fournies par les inventaires

Le levé des sols d'une zone déterminée comprend l'étude, l'identification et la cartographie des différents sols présents, y compris la représentation de leur répartition géographique, l'étude des conditions physiques de

leur environnement, de leurs caractéristiques morphologiques, physiques, chimiques et minéralogiques. L'interprétation des données afférentes à toutes ces propriétés, dans l'optique de l'évaluation de l'aptitude agricole des terres, complète ces activités.

Concrètement, l'inventaire est obtenu par la réalisation d'une succession d'étapes (OLMOS, 1981) :

- établissement d'une légende préliminaire, grâce à l'étude et l'identification de toutes les unités de sols à partir des observations et de l'examen, sur le terrain, des sols dans leurs diverses conditions d'environnement ;
- esquisse des unités cartographiques identifiées, sur les documents de base existants (cartes topographiques, géologiques, réseau hydrographique, etc.), par le biais des photographies aériennes, images de radar ou de satellite, disponibles à diverses échelles. Au cours de

cette phase, la légende préliminaire est adaptée et mise à jour, à mesure que de nouvelles unités sont identifiées ;

- collecte des profils (coupes de sols suffisamment larges et profondes pour permettre l'observation de toutes les couches — horizons — et leurs variations latérales) aux endroits les plus représentatifs, en vue de leur caractérisation analytique, selon une méthodologie rigoureusement définie (EMBRAPA, 1979) ;

- mise au point de la légende descriptive, résultat de l'actualisation continue de la partie descriptive de la légende préliminaire, en incluant les considérations sur le milieu ambiant, la description, la caractérisation et la classification des différentes unités cartographiques, et les observations sur le mode d'occupation des sols et les facteurs limitants pour l'utilisation agricole ;

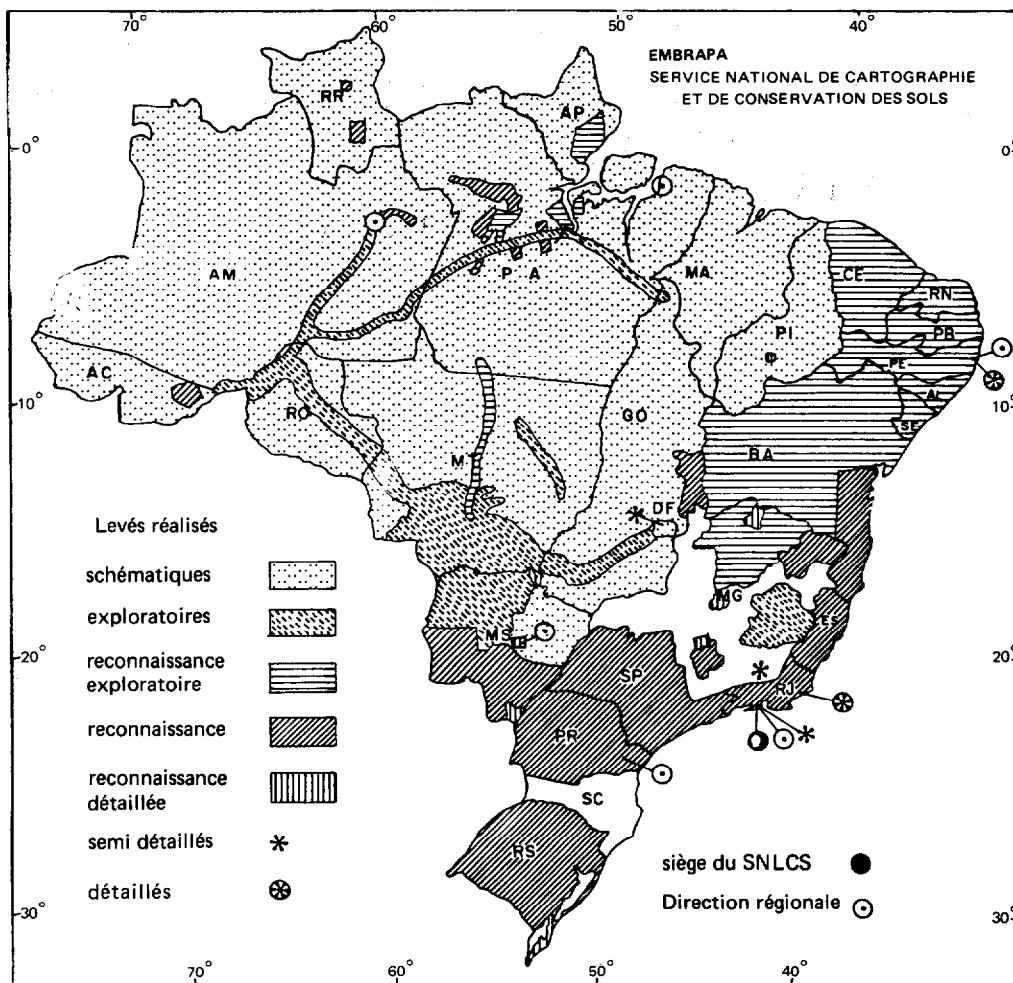


Fig. 1. — Carte de situation des levés de sols réalisés par le SNLCS au Brésil (1984).

— élaboration de la carte des sols et de la carte d'aptitude potentielle, rédaction, organisation et publication des résultats (notice technique), selon des normes bien précises et uniques pour la pédologie au Brésil (LEMOS et SANTOS, 1982).

Les levés de sols, à différents niveaux de détail (lequel dépend, en particulier, de l'objectif poursuivi et de l'extension de la zone à cartographier), réalisés depuis plus de vingt ans sur tout le territoire national (Fig.1), ont donné lieu à l'édition de plus d'une centaine de publications. Bien que les notices correspondant aux inventaires, du schématique au détaillé, forment la plus grande partie de la collection de bulletins techniques éditées, d'autres titres afférents à des travaux de thèses, études particulières, annales de réunion de classification, etc., fournissent des informations qui peuvent être utilisées par le système.

Tous ces documents contiennent les résultats des analyses morphologiques, physiques, chimiques, minéralogiques et de classification des sols réalisées dans le cadre de ces études, ce qui constitue une caractéristique commune à toutes ces publications. Toutefois, les différences de finalité qui existent entre chaque catégorie de travail font que les publications correspondantes incorporent des données complémentaires différentes, que l'on peut néanmoins rassembler en trois groupes distincts : les données générales, globalement inhérentes à l'étude considérée, les données ponctuelles qui se réfèrent aux profils d'observation du sol et les données zonales afférentes aux cartes de sols et d'aptitude potentielle, ainsi que leurs légendes.

## 2.2. Données supportées par le système

La représentation sur ordinateur de toute l'information contenue dans ces publications étant un des objectifs majeurs assignés par l'équipe du projet pour le développement du système, SISOLOS prévoit des structures de stockage pour toutes ces catégories de données, et non seulement pour les résultats des observations et analyses effectuées sur les profils, comme c'est le cas pour la quasi-totalité des systèmes d'informations pédologiques existants (MCDONALD et KLOOSTERMAN, 1983 ; LEPSCH *et al.*, 1978 ; MENK, 1984, par exemple).

Ainsi, dans la catégorie des données générales, le titre, l'année et les références de la publication, le niveau de détail du levé, l'information qui a trait aux auteurs, commanditaires et circonstances de réalisation du travail, un résumé aussi complet que possible et la relation des communes incluses dans la région inventoriée, sont autant de données stockées dans le système.

Les données ponctuelles constituent évidemment le plus gros contingent d'information disponible : outre

les profils caractéristiques, ou complets, les profils complémentaires, quelquefois baptisés d'échantillons « extras », mais aussi les échantillons de fertilité collectés par les pédologues pour compléter l'inventaire, les examens locaux réalisés lors des études préliminaires rapides et les échantillons de roches prélevés dans des objectifs de corrélation sol-environnement, prennent place dans les fichiers de la base de données.

Pour tous ces points d'échantillonnage, la description des conditions de l'environnement (y compris la localisation géographique du profil), les descriptions morphologiques des horizons et les résultats des analyses de laboratoire (physique, chimique et minéralogique) sont archivés. Dans l'état actuel du système, aucune structure adéquate à la représentation des diffractogrammes de rayons X ou à des courbes interprétatives du type profil d'argile ou de matière organique en fonction de la profondeur n'ont été prévues, ces informations n'étant pas fournies de façon routinière.

Jusqu'à présent, le stockage de cartes de sols n'a pas pu être réalisé, étant entendu que la saisie de ce type de données fait intervenir des moyens informatiques spécifiques, qu'il n'a pas encore été possible d'obtenir (ASSIS *et al.*, 1983). SISOLOS ne contient donc, dans la catégorie des données zonales, que les unités cartographiques (légende des cartes pédologiques), pour lesquelles on dispose de la composition taxonomique, de la description des conditions physiographiques et de l'information correspondant à l'interprétation d'aptitude potentielle (groupe et classe d'aptitude, nature et degré des facteurs limitants, etc.). En effet, les cartes d'aptitude potentielle fournies sont entièrement dérivées des cartes de sols correspondantes, par association des unités cartographiques qui offrent les mêmes potentialités pour l'exploitation agricole.

## 3. MÉTHODOLOGIE DE STOCKAGE

### 3.1. Processus d'entrée des données

Pour saisir l'information originale, qui se présente généralement sous la forme de publication plus ou moins ancienne, une phase de transcription sur des bordereaux aptes à la saisie mécanographique est nécessaire. L'alternative de saisie directe par utilisation de grilles d'écran adéquates sur micro-ordinateur a été écartée, non seulement pour des raisons circonstanciées (les premiers micro-ordinateurs ont fait leur apparition à l'EMBRAPA en 1981), mais aussi pour des raisons de portabilité évidente, d'une part, et parce que l'étape supplémentaire de saisie mécanographique qui résulte de ce choix ne conduit à une augmentation substantielle ni des erreurs, ni du temps total de saisie, d'autre part.

La transcription est réalisée par une équipe spécialisée, orientée par un coordinateur détenant une connais-

sance minimale de pédologie : une interprétation est en effet constamment nécessaire, pour cadrer les descriptions, aux nuances parfois subtiles, fournies par l'homme de terrain, dans le vocabulaire nécessairement limité imposé par la normalisation, et indispensable pour autoriser le traitement comparatif de grandes masses de données.

Les diverses équipes de transcription ont été créées par des institutions de collecte de données, utilisatrices de SISSOLOS : en plus du SNLCS de Rio de Janeiro, l'Institut Agronomique de l'état de São Paulo (IAC) à Campinas et le Centre de Recherche Agropastorale pour le « Cerrado » à Brasilia, stockent l'information correspondant à leur zone d'intervention dans le système.

Une fois remplis, les formulaires sont envoyés par lots au siège de l'EMBRAPA, qui dispose à Brasilia d'un centre informatique important, où ils sont saisis sur machines spécialisées, puis traitées par les routines de stockage : la figure 2 illustre l'enchaînement de ces opérations. Les bordereaux de codification utilisés ont été dessinés pour obtenir le meilleur rendement possible

durant l'opération de transcription, qui est sans aucun doute l'étape critique du système, au point de vue temps, coût et efforts fournis. A titre d'exemple, on peut citer qu'en trois mois sont remplis environ 5 000 formulaires, qui sont ensuite saisis en moins d'une semaine, par une équipe sensiblement égale (6 à 8 personnes), l'information correspondante étant finalement intégrée aux fichiers du SISSOLOS en l'espace de quelques minutes !

### 3.2. Caractéristiques générales de la transcription

En plus des caractéristiques ergonomiques de chaque bordereau (l'un des modèles a déjà été tiré, et rempli, à plus de 20 000 exemplaires), un certain nombre de critères ont été pris en considération :

— En premier lieu, chaque formulaire est conçu pour contenir un ensemble bien précis et complet d'informations : par exemple la description des conditions de l'environnement est complètement transcrite sur le bordereau de « PERFIL ». Ainsi, dans le cas où l'on ne dispose que de cette information pour un point d'échantillonnage déterminé (aucun horizon décrit ni collecté),

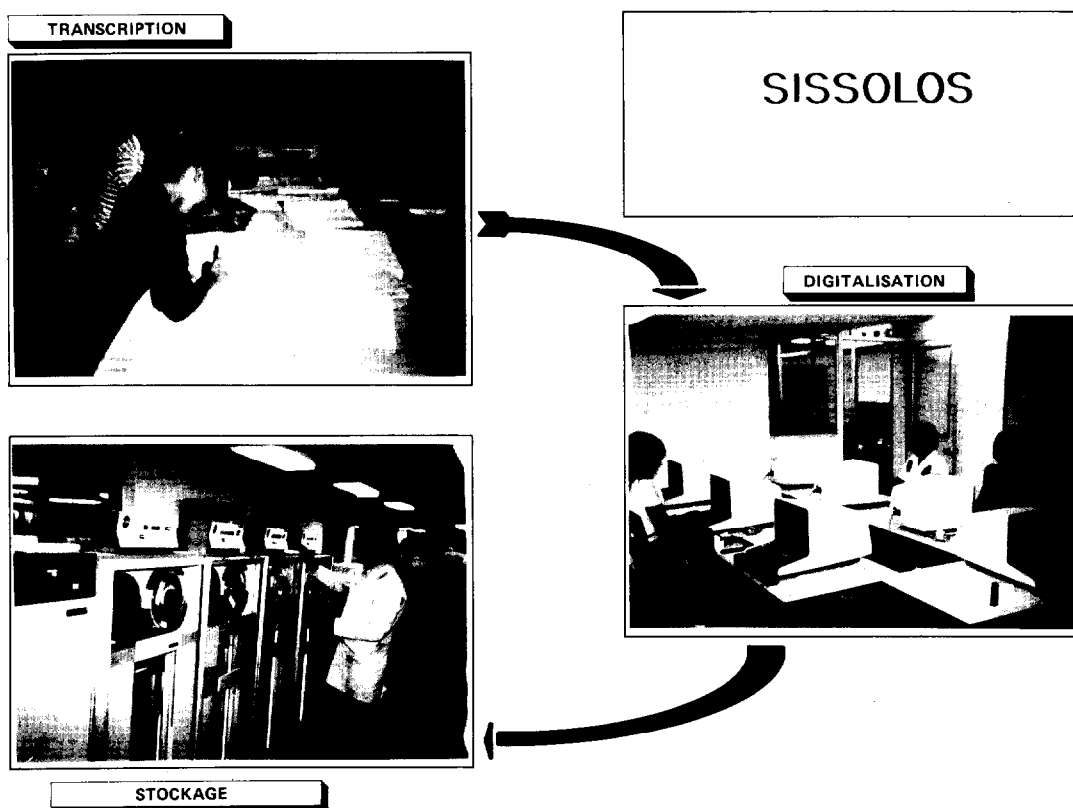


Fig. 2. — Processus de stockage des données de sols dans le système (d'après SISSOLOS — Guia de Entrada).

ce seul formulaire sera rempli, et on évite le problème des formulaires qui sont de véritables dossiers et que l'on hésite à remplir lorsqu'on ne dispose que d'une partie de l'information qu'ils prévoient.

— Lorsque c'était possible, on a fait systématiquement appel à la codification pour archiver les paramètres non numériques, de forme à réduire l'effort de transcription : il est évidemment commode de représenter par un seul chiffre la situation topographique du profil, par exemple, plutôt que d'écrire textuellement « tiers supérieur d'élévation », comme pour un système à format libre. Le risque d'erreur est toutefois plus grand : on a cherché à le minimiser en adoptant fréquemment des codes mnémomiques, PC pour PréCambrien, par exemple. La correspondance entre les codes et les descriptions est fournie par une table, pour une variable déterminée. L'ensemble des 80 tables de codification utilisées par SISOLOS joue le rôle de ce qui est parfois appelé glossaire dans d'autres systèmes (DUBROEUCQ et PIROT, 1983, par exemple).

— Les informations littérales non codifiables, par exemple la phrase « profil collecté au coucher du soleil », sont systématiquement représentées dans le système et véhiculées sur des formulaires spéciaux de textes, classés en fonction de la nature de l'information qu'ils archivent. Ainsi, la situation géographique d'un profil (dans son référentiel local) est stockée *in extenso* dans une catégorie spécifique de commentaire, différente de celle où l'on trouve les observations sur l'activité biologique le long de ce même profil, par exemple.

— Les informations redondantes ne sont normalement pas transcrites, vu qu'elles peuvent être obtenues par calcul à partir des informations de base. Une exception a toutefois été faite dans le cas des résultats d'analyses physiques et chimiques des horizons (tableaux de chiffres), de façon à se créer des conditions de vérification à posteriori. Les données redondantes correspondantes (rapport C/N, par exemple), sont transcrites sur formulaire spécifique, dans des conditions distinctes des informations de base qu'elles permettront de vérifier.

— Les valeurs obtenues lors des déterminations physico-chimiques de chaque échantillon collecté constituent deux groupes de données : un premier lot de résultats de routine, qui sanctionnent des analyses systématiquement effectuées, le second groupe étant formé de valeurs obtenues occasionnellement (sels solubles, par exemple). La codification de ces deux lots sur des formulaires séparés, l'un fréquent et l'autre éventuel, contribue à économiser les efforts de transcription.

— Tous les bordereaux ont été prévus pour recevoir un nombre maximal de 128 positions, de telle sorte que ce bloc d'information puisse être facilement manipulé dans les opérations de traitement postérieur (son image

est ainsi représentée sur une seule ligne d'impression, par exemple).

La présence de la majorité des champs qui ont été définis est optionnelle, et dépend seulement de l'existence, ou non, de l'information élémentaire que cette variable représente : seuls les paramètres qui jouent le rôle de clé, comme le numéro d'un profil ou le symbole d'un horizon, sont indispensables. L'absence d'information est donc représentée par un champ vierge, alors qu'une observation qui conclut à la non-existence d'un élément est traduite par un code non blanc.

Le glossaire constitue un fichier qui est maintenu par le système, de telle sorte que l'inclusion de nouvelles entrées puisse être réalisée à tout moment, lorsqu'un état d'une variable n'est pas représenté dans la table correspondante. Cette opération est effectuée, en routine, par le coordinateur de transcription. Celui-ci s'informe auprès des pédologues de façon à éviter le cas où cette éventuelle inclusion représenterait un synonyme d'une valeur déjà existante dans le glossaire (ARAUJO et SÉCHET, 1983).

#### 4. CORRECTION ET MISE À JOUR

##### 4.1. Mise à jour de la base

L'inclusion d'un lot d'informations nouvellement transcrites, et digitalisées, dans la base de données est réalisée à partir d'une opération de mise à jour, constituée par deux étapes consécutives : critique et reformatation des enregistrements nouveaux puis incorporation de ces enregistrements dans les fichiers du système. Les résultats fournis par SISOLOS à l'issue de la routine de mise à jour sont constitués de quatre états différents (MENEGUELLI *et al.*, 1981) :

— Une liste des enregistrements corrects, émise individuellement pour tous les transcrip-teurs. Chaque ligne de cet état constitue l'image d'un enregistrement pour lequel le système n'a pas détecté d'erreur, et permet à l'utilisateur une vérification, principalement des commentaires, mais aussi de constater d'éventuelles omissions.

— Une liste des enregistrements pour lesquels le système a détecté au moins une erreur est également émise individuellement pour chaque transcrip-teur. L'image de l'enregistrement erroné constitue une première ligne, la ligne suivante étant destinée à signaler le champ incriminé et la nature de l'erreur commise. Le programme de critique permet en effet de détecter quatre types différents d'incohérences : champ invalide (par exemple lettre lorsqu'un chiffre est attendu), valeur non représen-



tée dans la table correspondante, champ obligatoire absent et valeur extérieure à l'intervalle de définition d'un paramètre (pH supérieur à 14, par exemple). En fonction de la catégorie de l'erreur commise, l'enregistrement correspondant est soit refusé, si l'absence d'un champ clé empêche son rattachement correct, soit accepté « avec erreur », et le champ pour lequel une erreur a été détectée est alors considéré comme omis.

— Un état destiné à résumer l'opération de critique des données nouvelles est émis à l'intention du (ou des) coordinateurs de transcription. Il présente un certain nombre d'informations intéressantes pour évaluer le travail global de l'équipe et de chacun de ses membres, comme le nombre de registres transcrits, le pourcentage d'erreurs commises, etc. Pour atteindre cet objectif, chaque formulaire comporte une information signalant la responsable pour son remplissage, en l'absence de laquelle l'enregistrement correspondant est systématiquement rejeté.

— Un état destiné à résumer l'opération d'actualisation de la base de données est également fourni aux coordinateurs de transcription. Celui-ci contient la liste de toutes les opérations élémentaires anormales (tentative d'inclure un enregistrement déjà existant dans les fichiers, d'exclure ou de substituer un enregistrement inexistant, etc.), lesquelles doivent être examinées et résolues, une par une, par le coordinateur. Les statistiques fournies sur ce même listing permettent d'exercer un contrôle rigide sur le nombre d'enregistrements contenus dans la base, par comparaison avec le nombre obtenu lors de l'exécution précédente, de telle sorte qu'une erreur éventuelle de manipulation de bandes magnétiques peut être détectée et réparée avant épuisement du délai de rétention des volumes correspondants.

#### 4.2. Mise à jour du glossaire

La mise à jour du glossaire constitue une opération en tout point semblable à celle des fichiers de base : les mêmes états sont fournis et, après chaque actualisation significative, les nouvelles tables sont imprimées selon deux classifications différentes :

— Classées selon l'ordre lexicographique des descriptions pour les transcripteurs et les utilisateurs, elles facilitent la recherche du code correspondant à une valeur, ou un état (description) ;

— Classées selon l'ordre alphanumérique des codes, un exemplaire unique est fourni au coordinateur pour faciliter l'inclusion de nouvelles entrées, par choix d'un nouveau symbole.

Pour ces deux routines de mise à jour des fichiers de SISSOLOS, comme pour toutes les applications du système, des états spécifiques sont élaborés pour le per-

sonnel d'opération, indépendamment des listings émis à l'intention des utilisateurs. De cette façon, le contrôle de la production s'en trouve facilité, à partir du nombre d'enregistrements à l'entrée et à la sortie, du nombre de pages émises pour chaque état, etc.

#### 4.3. Modes de correction

Une fois saisis et traités, les formulaires sont retournés aux équipes de transcription en même temps que les états qui résultent de la phase de mise à jour de la base de données. Ceux-ci sont alors séparés par transcripteur et classés dans l'ordre adopté pour les listes de registres erronés, de telle sorte que chaque transcripteur puisse lui-même corriger les erreurs qu'il a commises. Ces corrections sont effectuées en utilisant les mêmes formulaires, de l'une des deux manières possibles :

— Correction au niveau de l'enregistrement. Dans ce cas, la correction est matérialisée sur le propre formulaire initial : si l'enregistrement correspondant a été refusé lors de la première tentative, le formulaire est envoyé de nouveau comme inclusion, tandis qu'il est proposé comme substitution si l'enregistrement avait été accepté « avec erreur ».

— Correction au niveau du champ. Pour les enregistrements qui ont été acceptés « avec erreur », un formulaire spécial, dit de correction partielle, est utilisé et autorise la modification de n'importe quel champ (sauf les champs-clés), repéré par sa position initiale et le type d'enregistrement où il est rencontré. Par mesure de sécurité, le contenu du champ à substituer est requis, ce qui évite des modifications intempestives.

Quelle que soit la forme adoptée pour effectuer les corrections dans la base, la même routine de correction et mise à jour est utilisée. Les bordereaux de correction partielle sont joints à ceux d'inclusion, exclusion et substitution engendrés par la correction de la dernière actualisation, eux-mêmes rassemblés avec les formulaires d'inclusion de données nouvelles, pour former un lot unique soumis au prochain traitement. Les formulaires non exploités pour la correction sont soigneusement archivés, ce qui permet ensuite une récupération facile pour des exploitations postérieures.

### 5. RELATION ENTRE L'UTILISATEUR ET LE SYSTÈME

#### 5.1. Structure de l'information

SISSOLOS met à la disposition de ses utilisateurs un ensemble de sept fichiers, chacun d'eux formé par une collection d'observations contenant les valeurs prises par les attributs qui se réfèrent à une entité déterminée. Ainsi, il existe un fichier dit de profils, constitué d'une séquence ordonnée d'enregistrements (ou observations),

chacun d'eux présentant une valeur déterminée de chaque champ (ou variable), comme par exemple le numéro du profil (clé), sa localisation, la classe de sol, etc.

Les sept fichiers qui composent la banque de données de sols sont les suivants (MENEGUELLI *et al.*, 1983) :

— fichier des inventaires (TRABALHO), pour lequel chaque observation contient l'information générale relative à une étude déterminée (notice technique d'un levé, dans la plupart des cas, mais aussi thèse, compte-rendu de réunion de classification, etc.) ;

— fichier des communes (MUNICIPIO), composé de la collection des informations relatives à chaque commune référencée dans un des inventaires stockés dans le fichier précédent ;

— fichier des profils (PERFIL) ;

— fichier des horizons (HORIZONTE), pour lequel chaque observation contient l'information relative à un horizon identifié, décrit et/ou collecté de l'un des profils stockés dans le fichier précédent ;

— fichier des fractions granulométriques (GRANULOMETRIA), formé des résultats de l'analyse minéralogi-

que opérée sur chaque classe granulométrique séparée de l'un des horizons contenus dans le fichier précédent ;

— fichier des unités cartographiques (MAPEAMENTO), contenant les unités cartographiques représentées sur la carte des sols de l'un des levés contenus dans la banque, y compris ses caractéristiques d'aptitude agricole, quand elles ont été stipulées ;

— fichier des composantes taxonomiques (COMPONENTES), destiné à recevoir les caractéristiques de chaque membre des associations de sols qui composent les unités cartographiques précédentes.

Evidemment, ces sept fichiers ne sont pas indépendants, dans la mesure où toutes ces entités obéissent à une certaine structure logique d'ensemble, représentée sur la figure 3. Cette structure se présente sous la forme d'un réseau, où les liaisons sont faites aux points de connexion (nœuds), à partir d'un champ commun (descripteur). Ainsi, un profil contient le sigle du travail dont il est tiré, le symbole de l'unité cartographique où il se situe et le code de la commune dans laquelle il a été collecté.

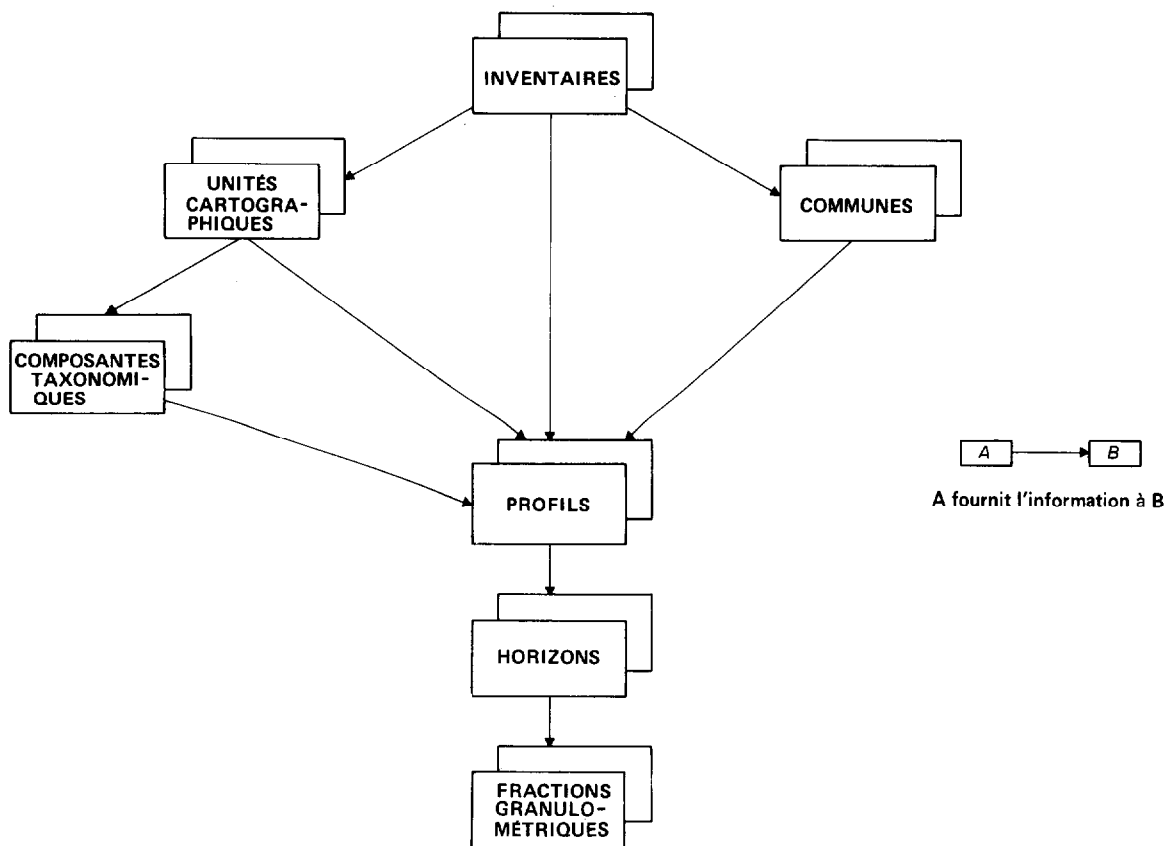


Fig.3. — Structure logique de l'information représentée dans le système (d'après SISOLOS — Manual de Uso).

## 5.2. Mode d'interrogation

Le système autorise n'importe quelle extraction d'information, pourvu qu'elle soit compatible avec les liaisons logiques existantes entre ces entités : c'est-à-dire qu'à partir d'un niveau quelconque de cette structure, toutes les informations de ce niveau ainsi que celles des niveaux supérieurs seront disponibles. Par exemple, étant donné un profil, les attributs qui se réfèrent au travail, à l'unité cartographique, à la composante et à la commune peuvent être obtenus, en plus des paramètres du propre profil, alors que les valeurs afférentes aux horizons de ce profil (résultats des déterminations physico-chimiques, par exemple) ne pourront être consultées qu'aux niveaux hiérarchiques inférieurs (HORIZONTE et GRANULOMETRIA). Une interrogation comme « combien de profils de vertisols situés dans des unités cartographiques sans facteurs limitants existent-ils dans la commune de Buritizeiro ? », est donc possible au niveau du fichier de profils, alors qu'une condition supplémentaire qui ferait intervenir le pH de l'horizon superficiel nécessiterait l'utilisation du fichier des horizons, plus volumineux.

L'existence, et la maintenance, de fichiers qui couvrent tous les niveaux de la hiérarchie, alors que théoriquement le niveau le plus bas permet la récupération d'information pour toute la structure à laquelle il appartient, se justifie par le souci d'efficacité dans le traitement des interrogations. En effet, si on devait consulter le fichier des horizons pour répondre à une sélection qui fait seulement intervenir des paramètres relatifs aux niveaux supérieurs, comme dans l'exemple choisi précédemment, plusieurs occurrences dupliquées seraient fournies (étant entendu que tous les horizons d'un même profil présentent les mêmes valeurs de ces paramètres), alors que d'autres seraient perdues (cas des profils pour lesquels aucun horizon n'a été collecté).

N'importe quelle exploitation du système est réalisée en deux étapes :

— dans un premier temps, on doit formuler en termes logiques les critères choisis pour sélectionner dans les fichiers du système la matrice contenant les seules observations et variables qui seront traitées (MENEGUELLI *et al.*, 1983) ;

— dans un second temps, il est nécessaire de programmer les opérations (émission d'états, tracés de graphiques ou d'histogrammes, calculs et analyses statistiques, etc.), qui devront être appliquées sur les données sélectionnées.

La première phase fait intervenir une succession de trois questions : quel est le fichier du système qui devra supporter la sélection ? Dans ce fichier, quelles sont les observations qui répondent au critère choisi et qui doi-

vent donc être retenues ? Et, dans ce sous-fichier, quelles sont les variables à extraire pour intervenir dans les opérations à effectuer ?

L'accès au système est permis soit directement, via terminal local ou remote en régime de traitement par lots, soit en utilisant les formulaires spécifiques pour la codification des interrogations et envoi postal. Le progiciel statistique SAS (Statistical Analysis System) a été retenu pour servir de premier langage d'interrogation de SISSOLOS, principalement en raison de son ample divulgation parmi les chercheurs et statisticiens de l'EMBRAPA. Le programme SAS nécessaire au traitement d'une interrogation obéit aux spécifications de ce langage, dans sa version 1982. La première étape de sélection correspond alors au « step DATA », qui autorise la création et le calcul de nouvelles variables ou de variables intermédiaires, en plus de toutes les variables déjà disponibles au niveau du fichier sélectionné. La seconde étape (« step PROC ») est alors considérablement facilitée par l'utilisation de ce logiciel, qui met à disposition de ses utilisateurs un grand nombre de procédures à caractère statistique, ainsi que des représentations synthétiques et des visualisations des données. La lecture de la documentation correspondante (Statistical Analysis Institute 1982) est recommandée pour évaluer les ressources considérables offertes par ce type de progiciel.


Bien entendu, le choix de ce langage, et bien que celui-ci ait répondu parfaitement à tous les besoins d'applications qui ont été détectés à ce jour, ne devra pas empêcher l'utilisation d'autres logiciels ou langages de haut niveau en guise de langages d'interrogation, en fonction des nécessités, ressources et connaissances des intéressés.

## 5.3. Documentation et dissémination du système

L'utilisation d'un système est un facteur prépondérant pour assurer sa pérennité : un effort particulier doit donc être réalisé pour élaborer une documentation claire et précise. De plus les concepteurs doivent se soucier de ce que leur travail puisse être utilisé sans leur intervention.

Dans cette optique, les responsables du SISSOLOS ont adopté, avec une certaine réussite, des mesures généralement réservées aux systèmes administratifs. Ainsi, dès que les différentes applications ont été opérationnelles, elles ont été remises aux secteurs chargés de la production et de la maintenance, moyennant l'obéissance à un certain nombre de normes imposées par le centre de traitement (programmation, documentation, utilisation de méthodes d'accès, etc.).

En particulier, le système a fait l'objet de l'élaboration d'un manuel de production, qui regroupe l'ensemble des tâches à effectuer à la réception d'une sollicitation de l'utilisateur (saisie, mise à jour, consultation, etc.),



## SOLICITAÇÃO DE EXECUÇÃO DE SERVIÇOS

Nº DA SOLICITAÇÃO

12

---

**USUÁRIO - SERVIÇO**

CÓDIGO DO SERVIÇO

P | A | D | B | A |

SISTEMA

DATA DA SOLICITAÇÃO

RESPONSÁVEL

UNIDADE

TELEFONE

RUBRICA DO RESPONSÁVEL

---

**2 SERVIÇOS REQUERIDOS ( ASSINALAR COM X O SERVIÇO REQUERIDO, NA COLUNA MARCA )**

Nº DE ORDEM	MARCA	Nº DO SERVIÇO	DESCRIÇÃO DO SERVIÇO	NATUREZA DOS DOCUMENTOS	QUANTIDADE DE DOCUMENTOS
Ø1		Ø1	ATUALIZAÇÃO DE TABELAS	PA. D. B. A. BL Ø 12	
Ø2		Ø2	ATUALIZAÇÃO DE CADASTRO	DIVERSOS	
Ø3		Ø3	EMISSÃO DE TABELAS ( L L ) CÓPIAS	—	—
Ø4		11	GERAÇÃO DO ARQUIVO REORGANIZADO	—	—
Ø5		13	GERAÇÃO DO BLOCO SAS EM FITA	—	—
Ø6		15	PROCESSAMENTO DE CONSULTA SAS	PA. D. B. A. BL Ø Ø 3	
Ø7		2Ø	GERAÇÃO DE RELATÓRIOS - PADRÃO	PA. D. B. A. BL Ø Ø 2	
Ø8					

---

**3 MATERIAL ENVIADO PELO USUÁRIO**

Nº DE ORDEM	DESCRIÇÃO	QUANTIDADE	Nº DE ORDEM	DESCRIÇÃO	QUANTIDADE	Nº DE ORDEM	DESCRIÇÃO	QUANTIDADE
Ø1	PA. D. B. A. BL Ø 2		Ø7	PA. D. B. A. BL Ø 25		13	PA. D. B. A. BL Ø 52	
Ø2	PA. D. B. A. BL Ø Ø 3		Ø8	PA. D. B. A. BL Ø 26		14	PA. D. B. A. BL Ø 53	
Ø3	PA. D. B. A. BL Ø 12		Ø9	PA. D. B. A. BL Ø 31		15	PA. D. B. A. BL Ø 63	
Ø4	PA. D. B. A. BL Ø 21		1Ø	PA. D. B. A. BL Ø 41		16	PA. D. B. A. BL Ø 71	
Ø5	PA. D. B. A. BL Ø 23		11	PA. D. B. A. BL Ø 42		17	PA. D. B. A. BL Ø 88	
Ø6	PA. D. B. A. BL Ø 24		12	PA. D. B. A. BL Ø 51		18		

OBSERVAÇÕES

Nº EST. DE REGISTROS

---

**4 USO EXCLUSIVO DO CCE**

EXECUÇÃO DOS SERVIÇOS

TOTAL DE BYTES

DIV. SERV.

REGISTROS DIGITADOS

DENSIDADE

RELATÓRIOS DEVOLVIDOS AO USUÁRIO

O SERVIÇO FOI GRAVADO NAS FITAS ED

E	D	-					
E	D	-					
E	D	-					

Nº DE ORDEM	CÓDIGO
Ø1	
Ø2	

Nº DE ORDEM	CÓDIGO
Ø3	
Ø4	

Nº DE ORDEM	CÓDIGO
Ø5	
Ø6	

OBSERVAÇÕES

AUTENTICAÇÃO CCE

---

**5 CONTROLE**

ENTRADA NA RECEPÇÃO DO CCE

SAÍDA DA RECEPÇÃO DO CCE

RUBRICA DO USUÁRIO

Fig. 4. — Formulário de solicitação para a execução de serviços (d'après SISOLOS — Guia de Entrada).

296

Cah. ORSTOM, sér. Pédol., vol. XXI, n° 4, 1984-1985 : 285-307

un manuel du système, destiné au personnel de maintenance, qui contient toutes les informations (documentation des programmes, manuel d'analyse) nécessaires à la correction des erreurs éventuelles de traitement et, surtout, à la modification des procédures adoptées par le système en fonction des divers changements susceptibles de survenir sur la configuration des ordinateurs de l'EMBRAPA, et un manuel d'utilisation qui décrit sommairement les consignes pratiques d'utilisation de chaque service prévu pour SISSOLOS.

De cette façon, les requêtes exprimées par l'utilisateur vis-à-vis de la banque de données de sols, quand elles nécessitent l'intervention du personnel d'exécution du centre de traitement, sont perçues normalement par celui-ci, parce qu'elles sont accompagnées d'une sollicitation d'exécution de services (voir figure 4), en tout point semblable aux réquisitions utilisées par les autres systèmes en production (MENEQUELLI et SÉCHET, 1984).

La documentation scientifique du SISSOLOS se compose de deux publications, de grande diffusion, éditées par le SNLCS/EMBRAPA en 1983 et 1984. Le guide d'entrée (*Guia de Entrada*) contient toutes les informations nécessaires au stockage de données nouvelles dans le système : formes et normes de stockage de chaque observation, description détaillée de la codification des textes, processus de transferts d'informations avec le système, contrôles, mises à jour, corrections, archivage. Bien entendu, ces informations sont essentielles pour les diverses équipes de transcription : les normes de stockage sont également très utiles pour les utilisateurs.

Le manuel d'utilisation (*Manual de Uso*) contient les informations nécessaires à l'exploitation de la banque, et en particulier :

- un dictionnaire des variables qui fournit le nom codé, le format (exprimé selon les normes SAS) et la signification de chacune des 500 variables disponibles dans les fichiers) ;
- le glossaire, listing des 80 tables de codification à la date de publication de ce manuel. Cette liste étant mise à jour à la faveur de chaque actualisation des fichiers, elle est fournie périodiquement aux utilisateurs concernés ;
- l'inventaire des fichiers du système, à la date de publication de ce manuel, et les modèles de formulaires utilisés pour solliciter des extractions d'information de la banque.

Ce manuel met en évidence les modalités pratiques d'utilisation de toutes les ressources du SISSOLOS, y compris l'ouverture d'un compte-utilisateur, le stockage de données nouvelles, la correction de données erronées et l'édition d'états standards, et définit les échanges d'informations entre l'utilisateur et le système. Quelques exemples d'application, parmi lesquels un programme qui permet à n'importe quel utilisateur auto-

risé d'obtenir l'inventaire de la banque au moment opportun (étant donné la mise à jour permanente de la base), complètent ce manuel d'utilisation.

La dissémination du système a fait l'objet d'une attention spéciale de la part de ses responsables : une des alternatives qui a été largement exploitée jusqu'à maintenant consiste à présenter SISSOLOS, et ses résultats, dans les congrès et séminaires scientifiques brésiliens. A l'occasion de plus d'une dizaine de réunions de ce genre, depuis 1981, plusieurs groupes de spécialistes appartenant à la communauté scientifique, des géographes aux informaticiens, en passant par les écologistes et, bien sûr, les pédologues, ont été informés de l'existence, de la disponibilité, de l'évolution et des progrès de SISSOLOS. Le lecteur trouvera dans la liste bibliographique de la fin de cet article plusieurs références concernant les applications déjà réalisées à partir du système (ARAUJO et SÉCHET, 1983 ; ASSIS *et al.*, 1983, 1984 ; LEPRUN *et al.*, 1984 ; MENEQUELLI *et al.*, 1981, 1983 ; SANTOS et SÉCHET, 1984 ; SÉCHET, 1982).

L'édition d'un dépliant d'information de grande diffusion (« folder ») est une autre alternative envisagée pour assurer une meilleure divulgation de la banque nationale de données de sols : elle n'a pu être possible à ce jour mais devrait intervenir dans un futur proche.

## 6. INTÉGRATION STOCKAGE-EXTRACTION

### 6.1. Principe de la réorganisation

La description d'un tel système met en évidence deux perceptions fondamentalement différentes des données (Fig. 5) :

— La vision adéquate au stockage, pour laquelle les informations sont groupées par bordereaux, selon une partition propice à l'entrée des données, c'est-à-dire par rassemblement des observations qui proviennent de la même source ou suivent le même cheminement (résultats des analyses de laboratoires d'une part, des descriptions faites sur le terrain, d'autre part). Il est intéressant que ces lots soient aussi indépendants que possible les uns des autres, pour pouvoir être stockés, ou actualisés, sans avoir à effectuer de regroupement artificiel.

— La vision adéquate à la sélection des informations, pour laquelle ces mêmes données sont structurées par entités, selon un schéma approprié à l'extraction suivant des critères conformes à la liaison logique intrinsèque à ces entités. Il est cette fois important que l'on puisse récupérer les couples couleur d'horizon et taux d'argile du même horizon, par exemple, alors qu'à l'entrée ces champs sont séparés.

Bien que les bordereaux obéissent déjà naturellement à une certaine structuration par entité, une opération spécifique est nécessaire pour servir de liaison entre les

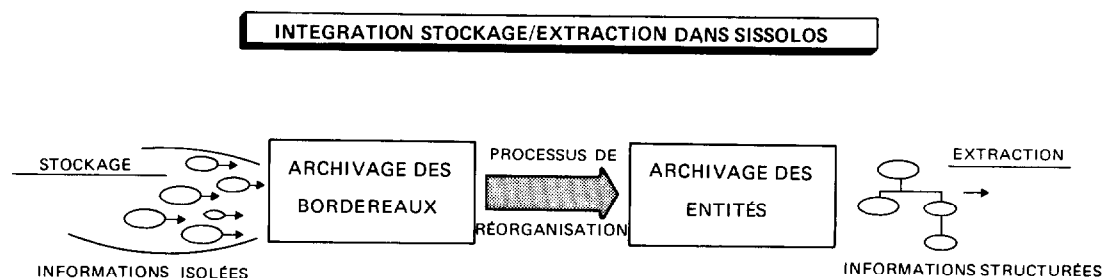


Fig. 5. — Mise en évidence de l'étape de réorganisation des données.

deux formes : SISOLOS comporte une telle étape, dite de réorganisation des données. Le rôle de cette phase consiste essentiellement à monter un enregistrement unique pour chaque observation de chaque fichier (entité).

Ainsi, chaque enregistrement du fichier des horizons est obtenu à partir de la concaténation logique du bordereau qui contient la description des caractéristiques morphologiques avec ceux qui contiennent les résultats des analyses physiques et chimiques du même horizon, en plus de ceux qui se réfèrent au profil dont cet horizon est tiré, à la commune et à l'unité cartographique où ce profil est localisé et finalement à l'inventaire dont cet ensemble est extrait. Le nouvel enregistrement ainsi formé contient donc toute l'information afférente à cet horizon, y compris la classe de sol et la localisation géographique, par exemple (tirées du profil).

Physiquement, dans un système classique de gestion de bases de données, cette concaténation logique est réalisée à l'aide de pointeurs, ce qui évite la concaténation physique et, par conséquent, la duplication des informations. Pour SISOLOS, il a été jugé préférable de traduire la concaténation logique par une concaténation physique : le problème de la duplication des informations, indésirable pour une banque des données, est évincé par la création d'un fichier réorganisé distinct de la base et uniquement destiné aux consultations.

Le principal avantage obtenu avec cette méthode réside dans le fait que tous les fichiers permanents utilisés dans le système sont organisés séquentiellement, ce qui permet que le support physique en soit la bande magnétique. La zone de travail nécessaire durant la phase de réorganisation (une heure environ) réalisée après chaque mise à jour importante, occupe moins de 10 Mégabytes de mémoire-disque et, surtout, est indépendante de la taille de la base des données. De cette façon, le coût de maintenance du système se trouve considérablement réduit, ce qui est un avantage d'autant plus appréciable qu'il s'agit d'un système scientifique dont l'utilisation est forcément réservée à un public restreint, de forme occasionnelle et sans nécessité de réponse immédiate.

## 6.2. Emission d'états standards

Les pédologues utilisent les nombreux paramètres observés sur les profils qu'ils collectent pour décider du classement de ce profil dans une catégorie prédéfinie de la classification des sols. Pour faciliter les manipulations et les comparaisons des descriptions de ces profils, les pédologues brésiliens ont défini des formats standards, progressivement améliorés au cours des années, et qui actuellement sont l'objet d'un consensus général (REUNION, 1979). Il était donc indispensable de faire en sorte que le système soit capable de restituer l'information stockée selon un modèle le plus voisin possible de la norme établie.

Un sous-système spécifique, dit de récupération descriptive ou de consultation standard, a ainsi été développé pour éditer divers états : description de profil, y compris la description morphologique des horizons, le tableau des résultats d'analyses physico-chimiques et la feuille correspondante à l'analyse minéralogique ; description générale de l'étude, avec le résumé de présentation, la liste des auteurs et la liste des communes ; caractéristiques physiographiques et composition taxonomique des unités cartographiques (légende détaillée).

Ce sous-système utilise directement le fichier des bordereaux (base de données), et est accessible à n'importe quel utilisateur par le biais de réquisitions d'édition. Pour être conforme aux normes imposées par les pédologues, la forme de présentation est sophistiquée : en contrepartie il n'y a pas de possibilité de variation de la consultation, pas plus pour le critère de sélection que pour le choix du format de sortie.

Etant entendu que cette application fournit la réponse de la banque de données sous la forme adoptée pour les publications du Service, elle est utilisée :

— par les équipes de transcription pour la vérification systématique avec les originaux. Dans cette optique, le nom du responsable pour la transcription figure sur le listing du profil (ou de l'unité cartographique), ce qui permet une récupération rapide des bordereaux correspondants lorsqu'il y a nécessité de corrections ;

— par les propres pédologues, lorsque ces informations n'ont pas encore été publiées, la sortie imprimée constituant alors un document de travail préalable extrêmement satisfaisant.

Il convient de noter que ce mode d'extraction d'informations est très complémentaire au sous-système de traitement des interrogations (récupération sélective), dans la mesure où il permet de faire figurer tous les commentaires, d'une part, et de montrer l'information totalement décodifiée, d'autre part, tandis que ni les textes ni le glossaire n'interviennent dans les ressources d'interrogation usant SAS, qui ne manipulent par conséquent que les données sous forme codifiée.

Pour obtenir le listing correspondant à un profil déterminé, ou à une unité cartographique choisie, l'utilisateur doit informer la clé sous laquelle ce profil, ou cette unité, est connu dans les fichiers de SISSOLOS, contrairement au système d'interrogations normal. Bien entendu, il pourra être intéressant d'utiliser le sous-système descriptif avec les clés obtenues par le sous-système sélectif, une fois réalisée la sélection d'un certain nombre de profils selon le critère choisi.

Pour optimiser l'exploitation, compte tenu du mode d'accès séquentiel des fichiers, le développement de cette dernière application intègre de nombreuses ressources (SÉCHET et ARAUJO, 1984). Ainsi, les utilisateurs peuvent présenter leurs sollicitations sans ordre prédéterminé, sous forme isolée (profil n° X de tel levé) ou systématique (tous les profils complémentaires de telle étude, toutes les unités cartographiques de telle autre). Le système permet de traiter en un lot unique toutes les requêtes de tous les utilisateurs, en fournissant les listings séparés pour chacun d'eux, ainsi qu'un état récapitulatif du volume d'informations transmis pour chaque réquisition, destiné au SNLCS qui administre le système.

### 6.3. Contrôle de qualité

Entre les observations faites par les pédologues sur le terrain et les résultats obtenus par les déterminations effectuées dans les laboratoires, d'une part, et les informations correspondantes finalement stockées dans la base, d'autre part, il existe de nombreuses possibilités d'erreurs, qui introduisent fatalement un nombre relativement important d'incorrections dans les fichiers. Les étapes transcription-saisie en particulier, sont passibles d'erreurs, surtout en ce qui concerne les champs alphanumériques. Par exemple, une vérification automatique réalisée en utilisant le rapport C/N (comparaison entre le rapport calculé par programme et la valeur stockée) a mis en évidence 217 occurrences d'erreurs (les deux valeurs diffèrent de plus d'une unité), sur un total de 11 103 comparaisons (MENEGUELLI et DRÜCK, 1985).

Un premier contrôle de qualité, à priori, est constitué par les routines de validation intégrées dans le processus de mise à jour de la base à partir de données nouvelles, et garantit la cohérence des informations considérées isolément, au point de vue type, format états possibles et/ou intervalle de valeurs (le cas échéant). Toutefois, à ce niveau, aucune critique réelle qui fasse intervenir plusieurs valeurs d'une même observation n'est effectuée, de telle sorte que l'on pourra rencontrer dans les fichiers un taux d'argile total inférieur au taux d'argile libre, et par conséquent un degré de floculation négatif, par exemple.

Pour SISSOLOS, le concept de critique de l'information se réfère à toute consultation réalisée dans l'optique de vérifier le comportement des données selon un critère déterminé. Pour constituer des routines de critique, à posteriori, on dispose de quelques outils de base :

— ordonner les enregistrements en fonction d'une variable déterminée afin de détecter facilement les observations pour lesquelles cette variable a été omise, qui se présenteront en première position après le tri ;

— utiliser toutes les facilités permises par le sous-système d'interrogation, afin de produire des histogrammes et des graphiques qui permettent une visualisation immédiate des données suspectes ;

— exploiter les redondances contrôlées d'information qui ont été introduites dans le système dans l'optique de la vérification. Par exemple, parmi les résultats des analyses chimiques des horizons, la somme des bases est stockée en même temps que chacun des termes de la somme (Ca, Mg, K et Na), de telle façon qu'une vérification soit aisée.

Jusqu'à présent, l'exploitation des possibilités de SISSOLOS en terme de critique des données a été limitée à la tentative d'ajuster les données stockées à l'image, la plus fidèle possible, des originaux publiés. A l'avenir, la connaissance critique acquise devra être utilisée pour favoriser une étape d'homogénéisation contrôlée des données, de nature à amenuiser les différences entre les résultats obtenus à des époques distinctes, et qui proviennent de l'évolution des méthodes utilisées pour réaliser et interpréter les levés.

## 7. APPLICATIONS DU SYSTÈME

### 7.1. Information disponible et applications spécifiques

L'accumulation des données déduites de tous les inventaires pédologiques réalisés sur le territoire brésilien constitue l'objectif principal du système, étant entendu que l'intérêt de la base de données augmente doublement avec le volume d'information disponible. En effet, du fait de son utilisation essentiellement statistique, les résultats obtenus bénéficient d'une confia-

bilité qui croît avec la masse d'information qui leur a donné naissance, d'une part, alors que la tendance à l'exhaustivité constitue un argument important pour les utilisateurs potentiels.

Depuis 1981, lorsque les routines de stockage du SIS-SOLOS sont devenues opérationnelles, le SNLCS dispose d'une équipe de transcription suffisante pour traiter tant le flux actuel de ses publications que ses données historiques propres. C'est ainsi que le volume total d'informations disponible actuellement par le système avoisine 170 000 enregistrements du fichier « des bordereaux », ou 60 000 entités (environ 150 Mbytes), avec un taux de croissance d'approximativement 30 000 enregistrements par an, comme le montre l'évolution présentée sur la figure 6.

À côté des informations tirées de plus d'une centaine de publications du Service, SISOLOS contient égale-

ment des données collectées par d'autres institutions, avec lesquelles des accords ont été pris, soit pour former une équipe de transcription locale, soit encore pour transcrire les publications correspondantes à Rio de Janeiro. En conséquence, l'utilisation comparative et intégrée des informations fournies par cinq organismes brésiliens différents est dorénavant possible.

L'utilisation d'un système unique n'empêche pas de rapprocher le système de ses utilisateurs potentiels : en l'absence actuelle d'un réseau public de télétraitement opérationnel au Brésil, la création et la maintenance, dans certaines de ces institutions, de banques de données régionales obtenues par sélection géographique dans la base centrale, joue un rôle prépondérant. Bien entendu, la cohérence des informations contenues dans les divers fichiers est garantie par deux principes : impossibilité de mettre à jour les fichiers satellites, sinon

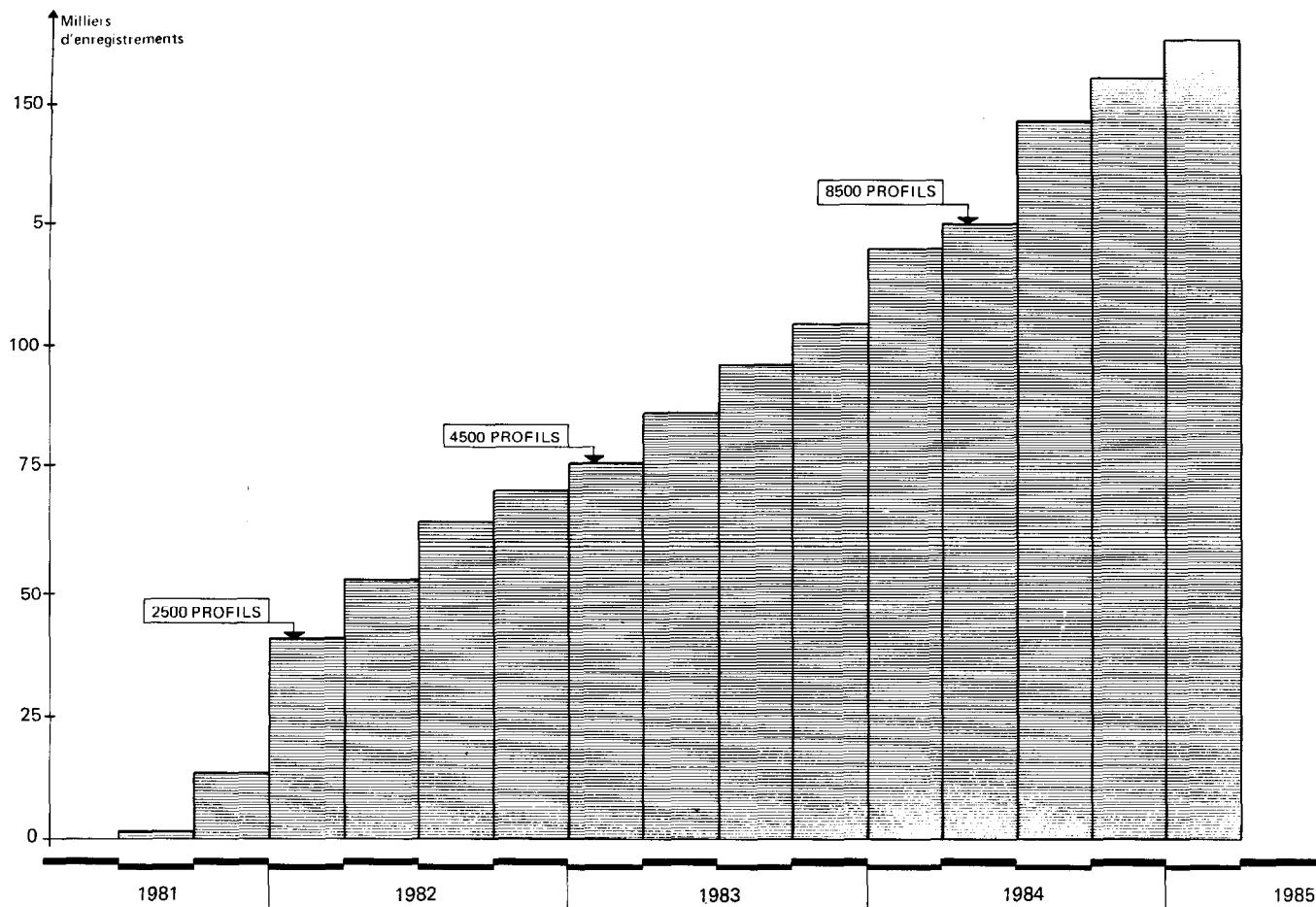


Fig. 6. — Evolution du volume d'information disponible dans la banque.



par le biais du fichier principal, et responsabilité finale de la qualité de l'information stockée dévolue à celui qui la stocke.

Les applications actuelles du système, réalisées au coup par coup, à l'initiative des pédologues le plus souvent, peuvent être classées en deux groupes : les applications taxonomiques, pour lesquelles la distribution géographique des observations ne constitue pas un paramètre essentiel, et les applications géographiques, qui reposent sur l'exploitation de cette dernière caractéristique.

Les applications taxonomiques utilisent fondamentalement des méthodes statistiques, parmi lesquels l'analyse factorielle multivariée joue un rôle important, tandis que les applications géographiques s'appuient le plus souvent sur des visualisations graphiques (ASSIS *et al.*, 1984). Pour ces dernières, une difficulté importante réside dans l'hétérogénéité de la distribution géographique des informations ponctuelles disponibles, l'exploitation des données zonales étant par ailleurs impossible sans moyens graphiques spécifiques.

Dans les deux cas, certains travaux ont déjà été effectués, quelques uns d'entre eux ayant même donné lieu à publication (LEPRUN *et al.*, 1984 ; SANTOS et SÉCHET, 1984 ; DRÜCK et MENEGUELLI, 1985 ; DRÜCK *et al.*, 1985) : tous utilisent les ressources statistiques du logiciel SAS qui, moyennant une codification relativement aisée, fournit des résultats très élaborés. La figure 7 présente un échantillon très partiel de ses potentialités lorsqu'on l'utilise pour exploiter une base de données :

- la liste des profils de « latossolos » localisés dans la commune de Santarém (état du Pará) a été choisie pour servir d'exemple de génération d'états : elle est obtenue par codification très simple (« PROC PRINT ») ;

- l'étude de la distribution statistique du degré de flocculation dans l'horizon superficiel des profils de « podzólicos vermelho-amarelos » est réalisée par utilisation d'une des nombreuses procédures statistiques disponibles (« PROC UNIVARIATE ») ;

- le comptage (*scoring*) fait partie des possibilités élémentaires de tout logiciel statistique. Il n'en rend pas moins de précieux services pour l'exploitation des grandes masses de données. Ici, on a choisi de montrer une table donnant le nombre de profils contenus dans les fichiers du système pour chaque classe de sol et chaque catégorie d'observation (profil complet, complémentaire, « échantillon extra », etc.). La procédure « FREQ » a été utilisée pour obtenir ce résultat ;

- l'élaboration d'histogrammes constitue une ressource particulièrement intéressante de SAS : l'histogramme présenté sur la figure 7 est un exemple d'utilisation de la procédure « CHART » et renseigne rapidement sur la répartition des profils collectés en fonction de la classe

de drainage : trois-quart d'entre eux possèdent un drainage modéré à accentué ;

- les valeurs caractéristiques obtenues pour un certain nombre de paramètres de l'analyse chimique de l'horizon « B » des « latossolos vermelho-escuros », sont un exemple d'utilisation de la procédure « MEANS ». On remarquera en passant le nombre élevé d'éléments contenus dans l'échantillon disponible dans la base (plus de 500) ;

- la procédure « PLOT » a été choisie comme dernier exemple : elle a été utilisée pour donner une vision synoptique de la répartition géographique des « cambissolos » sur le pays, en exploitant les coordonnées des profils stockés dans SISSOLOS.

Tous les exemples présentés sur cette figure ont été obtenus à partir de la codification élémentaire de la procédure correspondante. L'utilisation des nombreuses options disponibles pour chaque procédure permet d'améliorer progressivement chaque résultat, tandis que d'autres procédures sont utilisables pour l'analyse de régression, l'analyse de variance, l'analyse discriminante, etc.

## 7.2. Lignes d'exploitation prioritaires

A côté de ces applications isolées, l'équipe de méthodes quantitatives du SNLCS a lancé un certain nombre d'activités prioritaires à partir de la base de données, lesquelles peuvent être rangées en activités de développement (opérations nouvelles visant à compléter les ressources offertes par SISSOLOS), et activités de consolidation (amélioration des qualités du produit fourni). Un certain équilibre est respecté entre ces deux lignes directives, de telle sorte que l'assimilation progressive des ressources offertes par SISSOLOS, qui conditionne son exploitation adéquate de la part des utilisateurs, ne soit pas contrariée par une évolution technique exagérée du système. Par ailleurs, on doit éviter que celui-ci stagne dans l'état actuel, qui conduirait fatalement à une sclérose inacceptable par les mêmes utilisateurs.

Conformément à la systématique de programmation scientifique de l'EMBRAPA, ces activités ont été formulées en termes de projets de recherche liés au projet initial (voir fig.8), lesquels ont été soumis et approuvés par la communauté pédologique nationale.

Parmi les activités de consolidation, on peut citer les projets d'implantation dans d'autres organismes, dans le but d'augmenter le flux d'entrée des données, le projet de critique, dont la finalité est d'assurer un meilleur contrôle de qualité sur les informations stockées et le système de traitement sur micro-ordinateur des données de laboratoire, susceptible de fournir au SISSOLOS des données nouvelles sur support magnétique, éliminant ainsi une partie du travail de transcription.



En ce qui concerne les données nouvelles collectées sur le terrain, deux alternatives ont été étudiées pour en accélérer la saisie. L'élaboration de fiches de terrain à remplir par le pédologue « in situ » et directement utilisable pour la saisie mécanographique, à l'image de la solution adoptée par divers systèmes existants, ou bien la mise à disposition des mêmes pédologues d'un programme conversationnel sur microordinateur, permettant la saisie des observations. Cette dernière solution a été préférée à la première et devra être implantée à court terme, parce qu'elle permet l'émission immédiate d'un état standard de description générale du profil et caractéristiques morphologiques des horizons, très utile pour le pédologue à ce niveau.

Des modifications plus substantielles, comme le stockage de toposéquences, de diffractogrammes de rayons X, de bloc-diagrammes, etc., qui pourraient être désirables, nécessiteront des altérations significatives, mais néanmoins compatibles avec la structure globale du système.

### 7.3. Perspectives d'évolution

L'évolution de SISSOLOS dans les prochaines années devrait s'orienter en fonction de deux projets importants : le stockage des cartes de sols, d'une part, le soutien au système brésilien de classification des sols, d'autre part.

Le stockage sur ordinateur des cartes de sols, dans des fichiers reliés au système à partir de la légende de

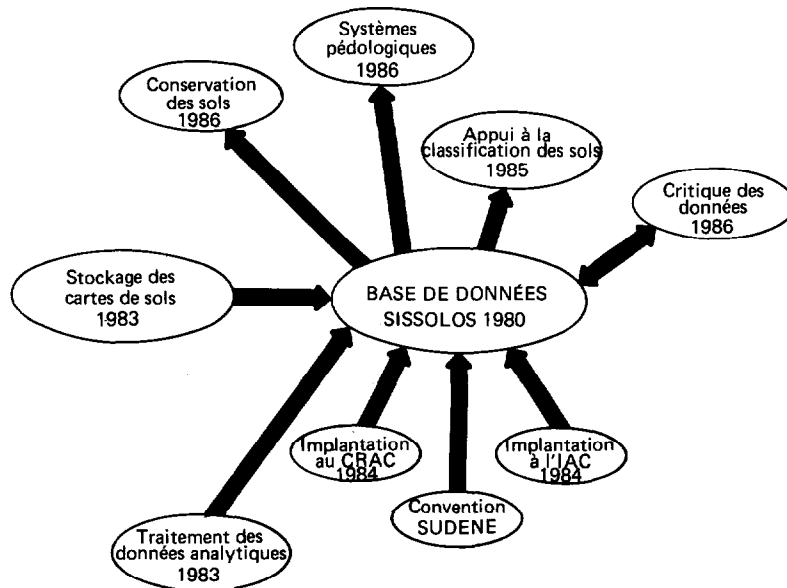


Fig. 8. — Projets de recherche en cours de réalisation liés au SISSOLOS.

Les responsables de SISSOLOS se sont préoccupés dès le début de la façon selon laquelle le système devrait accompagner les améliorations successives susceptibles d'être apportées à la méthodologie adoptée par la cartographie pédologique. C'est ainsi que l'on dispose de diverses techniques pour mettre en œuvre ces transformations, certaines d'entre elles déjà largement utilisées.

Par exemple, l'utilisation d'un glossaire, actualisable en routine par l'utilisateur, donne une grande facilité pour l'inclusion de nouvelles valeurs pour tous les paramètres tabulés. D'autres modifications sont possibles pour permettre l'incorporation d'un nouveau groupe de paramètres, avec la définition et le traitement d'un nouveau bordereau.

la carte, constitue un complément logique de la base de données dans son état actuel, dans la mesure où l'accès à la carte est indispensable pour répondre à des questions du type : « dans un endroit déterminé, quel type de sol peut-on rencontrer, et avec quelle probabilité ? », ou encore : « quelle est la distribution géographique de telle ou telle caractéristique de sol ? ».

Théoriquement, de telles opérations peuvent être effectuées manuellement, par exemple pour la première question, en localisant le lieu considéré sur la carte des sols, ce qui permet d'en déduire un symbole d'unité cartographique. La légende est alors utilisée pour donner les classes de sols susceptibles d'être rencontrées dans la région, et avec quelle probabilité (assimilée à la proportion des composantes taxonomiques de l'association).

S'il y a lieu, les profils collectés dans cette unité conduiront à des résultats de détermination analytique, par exemple.

Dans le second cas, un processus inverse doit être suivi pour, à partir des valeurs analytiques des profils représentatifs d'une unité, établir une carte thématique correspondant à la cartographie de la caractéristique étudiée.

En pratique, ces opérations sont difficilement réalisables, peu répétitives et très sujettes à erreurs, compte tenu du temps nécessaire pour manipuler de grandes quantités d'informations, d'autant plus qu'une information bidimensionnelle (la carte) intervient également.

Dans la mesure où SISOLOS prévoit une structure adéquate au stockage et à l'extraction des cartes pédologiques associées aux inventaires, ces applications géographiques seront possibles au même titre que les autres exploitations, comme en témoigne la figure 9.

Les bénéfices apportés par l'intégration des cartes de sols dans le système seront multiples (SÉCHET *et al.*, 1985) :

- stockage, reproduction et extraction de l'information cartographique ;
- réduction, ampliation et montage des cartes par juxtaposition de plusieurs feuilles ;
- calcul automatique de surface des unités cartographiques ;
- production de cartes interprétatives à partir de la carte pédologique, comme la carte d'aptitude agricole des terres, par exemple ;
- composition d'information par superposition avec d'autres cartes thématiques, pour des opérations de zonage en vue de l'aménagement du territoire, par exemple.

Le projet d'intégration des cartes de sols dans SISOLOS a vu le jour en 1983 et a déjà franchi, avec le concours de l'ORSTOM, l'obstacle initial que représente la saisie des cartes à la table à digitaliser.

Un projet d'exploitation systématique du système, au service de l'équipe de pédologues chargée d'élaborer le système brésilien de classification des sols, a été présenté en 1984. L'objectif qui lui a été fixé est de tenter

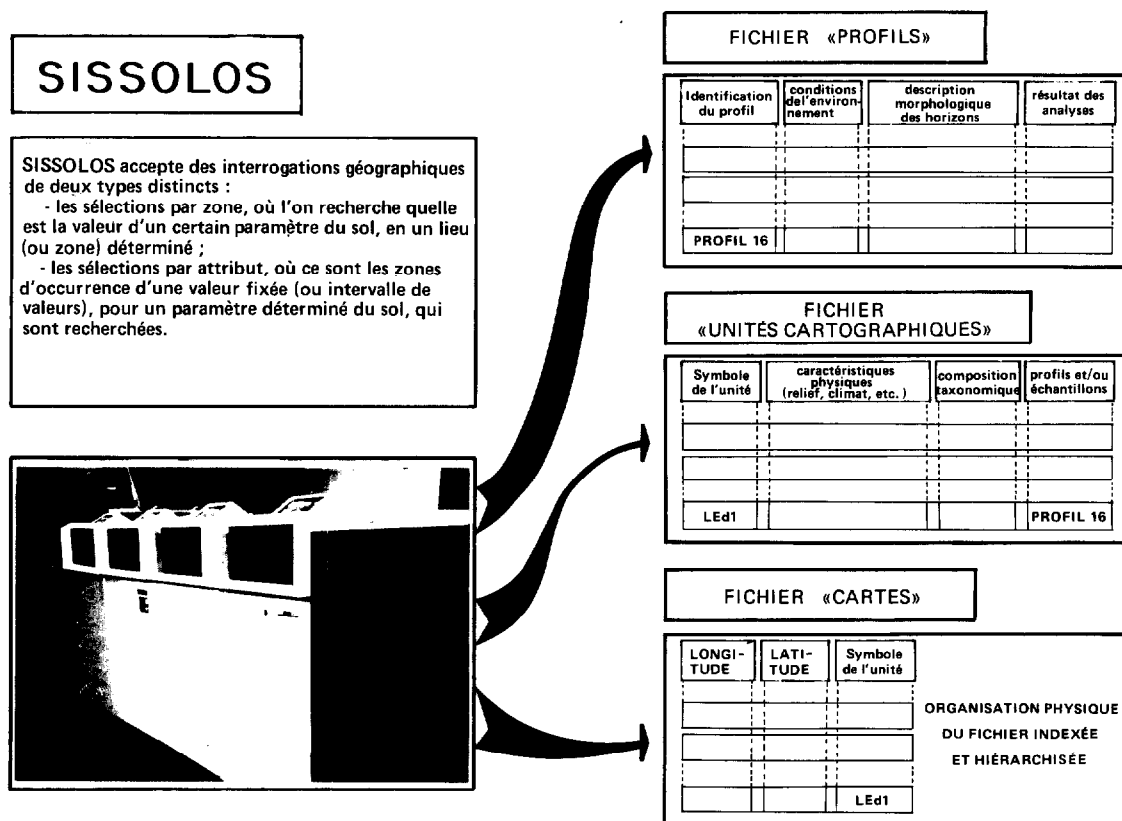


Fig. 9. — Application géographique à partir du chaînage des fichiers de SISOLOS (d'après SÉCHET *et al.*, 1984).

de quantifier, a posteriori, pour chaque classe de sol, les limites des paramètres discriminatifs utilisés dans la classification. Quelques résultats significatifs ont déjà été obtenus, à partir d'une analyse exploratoire systématique qui utilise les ressources de statistique descriptive du SISSOLOS.

Cette activité comme, dans une moindre mesure toutefois, les interrogations de la banque menées dans le cadre du travail sur le contrôle de qualité, conduit à un certain nombre de résultats du type : « pour telle classe de sol, si tel paramètre prend telle valeur, alors tel autre a une valeur comprise dans tel intervalle (éventuellement, avec telle probabilité) ».

La structuration de ces résultats en règles de production, l'organisation et le stockage de celles-ci en un système cohérent qui autorise le chaînage et l'extraction de ces inférences, devrait permettre de construire un système-expert (système-pédologique), capable de restituer non seulement de simples données d'observations, mais également les règles d'interprétation déduites de la manipulation de ces observations, obtenues d'une façon semblable au processus d'accumulation d'expérience du pédologue.

Pour l'élaboration de ces règles, les textes d'observation stockés dans la base pourront être utilisés (ceux-ci représentant plus de 40 % du volume d'information disponible), à l'aide d'un logiciel approprié au traitement de l'information littérale, par le biais de mots-clés : une normalisation préalable de la forme de présentation de ces textes devrait toutefois y être incorporée.

## CONCLUSION

Pour une institution de recherche scientifique, les applications de l'informatique peuvent être classées en quatre grandes catégories : automatisation des processus administratifs, gestion des données documentaires, analyses des données produites par la recherche et systèmes destinés au soutien des actions de recherche.

Les systèmes qui s'encadrent dans le premier groupe (personnel, comptabilité, matériel, etc.), et dans une moindre mesure ceux du deuxième groupe comme la gestion du fond documentaire et la dissémination sélective de l'information bibliographique, ne sont pas à proprement parler spécifiques d'un organisme de recherche. Pour cette raison, les applications correspondantes peuvent être développées par des équipes d'informaticiens, rompus à ce genre d'activités.

Parmi les données produites par la recherche, on peut distinguer les résultats d'expérimentation et ceux fournis par les inventaires. L'analyse des données issues de l'expérimentation fait généralement appel à l'informatique par le biais de la statistique. L'ordinateur devient indispensable beaucoup plus à cause de la complexité

des calculs que du volume des données à traiter. Pour ce type d'activités, l'intervention de l'informaticien apparaît totalement superflue, étant donné le caractère très spécifique de ces analyses (l'efficacité de la partie traitement des données importe alors peu, pourvu que le résultat soit obtenu et l'objectif atteint). Un conseil en techniques statistiques est préférable, le chercheur pouvant rapidement garantir son autosuffisance dans l'utilisation des outils informatiques auxquels il peut accéder.

La disponibilité des informations fournies par les inventaires, sous la forme de bases de données, est fondamentale pour assurer son utilisation dans les actions de recherche menées en aval et, ainsi, évaluer la portée des résultats produits par ces recherches. Par exemple, la connaissance de la proportion, de la distribution géographique et des caractéristiques de variabilité des vertisols au Brésil, paraît essentielle pour analyser l'intérêt d'une recherche sur l'utilisation agricole de ce type de sol.

Ces bases de données stockent une connaissance historique dont le volume accumulé est important et, en raison de leur caractère de généralité, elles doivent respecter les normes de cohérence, intégrité, sécurité et possibilité d'utilisation partagée garanties par les systèmes centralisés (EMBRAPA 1980). Pour toutes ces raisons, elles doivent être réalisées avec l'intervention d'informaticiens (JACCON et SÉCHET, 1980).

Néanmoins, l'analyse de ces systèmes requiert, en plus de l'application de techniques conventionnelles de traitement de données, une interaction harmonieuse entre l'analyste et le chercheur. Ceci exige une connaissance raisonnable du domaine scientifique de la part de l'informaticien, en même temps qu'une perception suffisante des difficultés intrinsèques de l'informatique, de la part du chercheur.

Par ailleurs, les systèmes d'information scientifique font partie de l'effort de recherche dont le résultat n'est pas immédiat et, pour cette raison, ne sont généralement pas prioritaires aux yeux des administrateurs. De plus, leur utilisation est sporadique et leur dissémination est restreinte, réservée aux seuls spécialistes d'un segment de la communauté scientifique, du moins tant que l'évolution vers le système-expert reste embryonnaire.

Aussi, le coût de leur développement, et surtout leur coût opérationnel, doivent être réduits. Les ressources utilisées pour SISSOLOS, qui se contente de supports de stockage *off-line*, se limitent à une vingtaine de bandes magnétiques, environ 40 minutes de temps d'unité centrale à chaque actualisation-réorganisation (trimestrielle), et environ 2 minutes de temps d'unité centrale lors de chaque extraction d'information.

Les données recueillies par les inventaires sur les ressources naturelles présentent trois dimensions discriminatoires :

— Un aspect chronologique, quand les observations dépendent du temps : les fichiers sont alors constitués par des séries temporelles. Par exemple, dans le domaine du climat on stocke des séquences de valeurs pour quelques paramètres (températures, pluie, etc.).

— Un aspect taxonomique, quand à chaque entité correspond un certain nombre d'attributs : on traite des relations. Ainsi, un herbier se compose de collections d'espèces végétales, pour lesquelles on emmagasine des caractéristiques descriptives, morphologiques, médicinales, etc.

— Un aspect géographique, quand les entités considérées présentent une certaine dimension spatiale : on s'intéresse alors à la distribution géographique de ces observations. Par exemple, une banque de données sur l'occupation des sols archive les cartes correspondantes et

leurs légendes, avec la localisation, la forme et le contenu des unités homogènes.

Généralement, l'un ou l'autre de ces aspects prédomine, bien que les trois devraient être pris en considération : pour une banque de données de sols, le facteur temps n'est pas primordial. SISOLOS constitue une solution pour le traitement des deux autres aspects et, à ce titre, pourrait servir d'exemple pour de nouvelles réalisations dépassant le cadre de la pédologie, tant dans le domaine géographique (géologie, phytogéographie, géophysique, hydrogéologie, etc.), que dans le domaine taxonomique (sciences biologiques et humaines, par exemple).

*Manuscrit accepté par le Comité de rédaction le 17 janvier 1986*

## BIBLIOGRAPHIE

- ARAUJO (A.R. de) et SÉCHET (P.), 1983. — Dados de solos : armazenamento e recuperação. *In* : Congresso Nacional de Informática, 16., São Paulo, SP, 1983. *Anais*. São Paulo, SUCESU : 219-24.
- ARAUJO (A.R. de), COSTA (M.M.do C.), SÉCHET (P.) et DRUCK (S.), 1985. — Inteligência artificial e ciência do solo : perspectivas. *In* : Congresso Brasileiro de Ciência do solo, 20., Belém, PA, 1985. *Programa e Resumos*. Belém, SBSC, (à paraître).
- ASSIS (D.S.), MENEGUELLI (N.do A.) et SÉCHET (P.), 1983. — SISOLOS : Um sistema geográfico. *In* : Congresso Brasileiro de Geociências e tecnologias aplicadas aos estudos ambientais, 1., Rio de Janeiro, RJ (à paraître).
- ASSIS (D.S.), MENEGUELLI (N.do A.) et SÉCHET (P.), 1984. — SISOLOS : Uma fonte de informação sobre o meio ambiente. *In* : Congresso Brasileiro de Defesa do meio ambiente, 1., Rio de Janeiro, RJ. *Anais*. Rio de Janeiro, Clube de Engenharia : 381-8.
- DRUCK (S.) et MENEGUELLI (N.do A.), 1985. — Estudo da distribuição do carbono orgânico em Podzólicos Vermelho-amarelos. *In* : Congresso Brasileiro de Ciência do solo, 20., Belém, PA, 1985. *Programa e Resumos*. Belém, SBSC, (à paraître).
- DRUCK (S.), LEPRUN (J.C.) et SILVEIRA (C.O.), 1985. — Estudo estatística preliminar do grau de flocculação de diferentes latossolos brasileiros. *In* : Congresso Brasileiro de Ciência do solo, 20., Belém, PA, 1985. *Programa e Resumos*, Belém, SBSC (à paraître).
- DUBROEUCQ (D.) et PIROT (F.), 1983. — Enregistrement informatique de données de sols de l'Amazonie Vénézuélienne. Paris, *I.D.T. ORSTOM* n° 60, 103 p.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA, 1980. — Departamento de Métodos Quantitativos, Brasília, DF. *Plano Diretor de Informática 1981-1985*. Brasília, EMBRAPA-DID, 152 p.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA, 1981. — Departamento Técnico-Científico, Brasília, DF. *Programa Nacional de Pesquisa de Levantamento de Solos*. Brasília, EMBRAPA-DID, 37 p.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA, 1979. — Serviço Nacional de Levantamento e Conservação de Solos, Rio de Janeiro, RJ. *Manual de métodos de análise de solo*. Rio de Janeiro, Iv.
- JACCON (G.) et SÉCHET (P.), 1980. — Base de données hydroclimatologiques du nord-est brésilien. Paris, *ORSTOM, Sér. Hydrol.*, vol. XVII, 3-4 : 177-219.
- LEMOES (R.C.de) et SANTOS (R.D.dos), 1982. — *Manual de descrição e coleta de solo no campo*. 2. ed. Campinas, SBSC-EMBRAPA/SNLCS, 46 p.
- LEPRUN (J.C.), MENEGUELLI (N.do A.) et SÉCHET (P.), 1984. — SISOLOS : Uma alternativa para determinar condições potenciais de degradação de solos no país. *In* : Encontro nacional de pesquisa sobre conservação do solo, 5., Porto Alegre, RS, 1984. *Programa e Resumos*. Porto Alegre, SBSC, p.77.
- LEPSCH (I.F.), MENK (J.R.F.), OLIVEIRA (J.B.), SILVA (O.A.C.), MASSON (W.), BORDONI (O.), 1978. — Apoio computacional em levantamento de solos. Campinas, Instituto Agrônomo do Estado de São Paulo (IAC), *Boletim Técnico* 210, 44 p.
- MCDONALD (K.B.) et KLOOSTERMAN (B.), 1983. — The Canadian soil information system (CANSIS). *General user's manual*. Ottawa, Ontario, Land Resource Research Institute, 112 p.
- MENEGUELLI (N.do A.), ARAUJO (A.R.de), ASSIS (D.S.) et SÉCHET (P.), 1981. — SISOLOS : Um sistema de armazenamento de dados de solos. *In* : Congresso Brasileiro de Ciência do Solo, 18., Salvador, BA, 1981. *Programa e Resumos*. Salvador, SBSC, p. 69.

- MENEGUELLI (N.do A.), ASSIS (D.S.) et SÉCHET (P.), 1983. — SISSOLOS : Acesso aos dados de solos através do computador. In : Congresso Brasileiro de Ciência do Solo, 19., Curitiba, 1983. *Programa e Resumos*. Curitiba, PR, SBCS, p.98.
- MENEGUELLI (N.do A.), ASSIS (D.S.), ARAUJO (A.R.) et SÉCHET (P.), 1983. — SISSOLOS : Manual de uso. Rio de Janeiro, RJ, EMBRAPA-SNLCS, *Documentos*, 4, 245 p.
- MENEGUELLI (N.do A.) et SÉCHET (P.), 1984. — SISSOLOS : Guia de entrada. Rio de Janeiro, RJ, EMBRAPA-SNLCS, *Documentos*, 9, 91 p.
- MENEGUELLI (N.do A.) et DRUCK (S.), 1985. — SISSOLOS : controle de qualidade dos dados. In : Congresso Brasileiro de Ciência do solo, 20., Belém, PA, 1985. *Programa e Resumos*. Belém, SBCS (*à paraître*).
- MENK (J.R.F.), 1984. — GERA : Sistema computarizado de apoio a Jevantamento de solos (versão 2). I. Aspectos, estrutura e organização. Campinas, Instituto Agrônomico do Estado de São Paulo (IAC), 10 p.
- OLMOS (I.L.J.), 1981. — Bases para leitura de mapas de solos. Rio de Janeiro, EMBRAPA-SNLCS, *Série Miscelânea*, 4, 91 p.
- REUNIAO TECNICA DE LEVANTAMENTO DE SOLOS, 10., Rio de Janeiro, 1979. *Súmula*. Rio de Janeiro, EMBRAPA/SNLCS, 1979. 83 p. (EMBRAPA. SNLCS. Série Miscelânea, 1).
- SANTOS (R.D.dos) et SÉCHET (P.), 1984. — Contribuição do SNLCS para o conhecimento dos solos do trópico úmido brasileiro. In : Simposio do Trópico Umido, 1., Belém, PA, 1984. *Resumos*. Belém, EMBRAPA-CPATU, p.27-8.
- SÉCHET (P.), 1982. — Banco Nacional de dados de solos da EMBRAPA : Projeto SISSOLOS. *Inf. DMQ*, Brasília, DF, 1(2) : 8.
- SÉCHET (P.) et ARAUJO (A.R.de), 1984. — Emissão de relatórios operacionais em bancos de dados científicos. In : Congresso Nacional de Informática, 17., Rio de Janeiro, RJ, 1984. *Anais*. Rio de Janeiro, SUCESU, (2)22.
- SÉCHET (P.), ASSIS (D.S.), ARAUJO (A.R.de) et SAROLDI (M.J.), 1985. — Uso da informática no manuseio e na exploração dos mapas de solos. In : Congresso Brasileiro de Ciência do solo, 20., Belém, PA, 1985. *Programa e Resumos*. Belém, SBCS (*à paraître*).
- STATISTICAL ANALYSIS SYSTEM INSTITUTE, Raleigh, North Carolina. *SAS User's guide : 1 — Basics*. Raleigh, 1982.
- STATISTICAL ANALYSIS SYSTEM INSTITUTE, Raleigh, North Carolina. *SAS User's guide : 2 — Statistics*. Raleigh, 1982.