

LE TRAITEMENT INFORMATIQUE DES DONNÉES PÉDOLOGIQUES

INTRODUCTION

En Pédologie les données deviennent de plus en plus nombreuses. Si l'on veut comparer les informations pédologiques provenant de diverses régions et saisies par différents Pédologues, il faut créer des banques de données de sol.

A cette nécessité, le Pédologue pose deux questions:

- comment entrer les données de sol en banque,
- que faire des données de sol contenues dans la banque?

On essaiera de répondre à ces deux questions.

SITUATION DE L'INFORMATIQUE EN PÉDOLOGIE

Ce n'est qu'au Congrès International de la Science du Sol, à Moscou, en 1974, qu'a été créé un groupe de travail „Informatique et Pédologie” au niveau international.

En France, il existe depuis les années 1960 des organismes qui se sont intéressés à ce genre de travaux. On peut citer une société privée: la SOGREAH*, la première à avoir travaillé dans ce sens; et puis l'ORSTOM et l'INRA**. Actuellement, l'ensemble des pédologues travaillant avec l'informatique se réunissent dans le cadre de la Société Internationale „Informatique et Biosphère”. Cette société a mis au point plusieurs glossaires:

- 1969: Description des horizons du profil,
- 1971: Description de l'environnement du profil,
- 1977: Description des horizons du profil:

* SOGREAH: Société Grenobloise d'Etude et l'Applications Hydrauliques.

** INRA: Institut National de la Recherche Agronomique. ORSTOM: Office de la Recherche Scientifique et Technique d'Outre-Mer.

du profil,
des échantillons de terre (analyses),
de l'environnement.

Elle est chargée d'envisager les bases de la création d'une banque internationale de données de sol. Pour ce, elle a produit plusieurs documents:

- 1975: les objectifs et structures juridiques et techniques d'une banque de données de sol (Agence de Coopération Culturelle et Technique),
- 1976: Etude sur l'élaboration d'une banque internationale de données de sol (glossaires et exemple de base de gestion),
- 1977: Glossaire de description de l'environnement du profil pédologique pour les sols des régions tropicales (bilingue: Français—Anglais).

La recherche dans le domaine de l'informatique pédologique se fait dans les organismes précédents, ainsi que dans les Ecoles Nationales Supérieures Agronomiques et à l'Institut National Agronomique Paris—Grignon. Des cours de 10 à 30 heures y sont donnés aux étudiants préparant un doctorat.

LA BANQUE DE DONNÉES DE SOL

On peut caractériser une banque de données par trois fonctions: les entrées, la gestion, les sorties. Ces trois fonctions permettent de répondre à la première question: comment entrer les données de sol?

LES ENTRÉES

Les informations que saisit le Pédologue proviennent soit d'analyser d'échantillons de terre, soit d'observations de terrain. Il est très important de garder toute l'information pédologique entre le moment où elle entre en ordinateur et est stockée. On distinguera trois étapes: le terrain, le glossaire, le répertoire.

Le terrain

Le pédologue est entièrement responsable de la phase terrain. Il est donc nécessaire que les outils que lui propose l'informaticien soient utilisables sur le terrain.

Ainsi la fiche de description doit pouvoir être utilisée sur le terrain (assez dure, facilement maniable...) et permettre une description convenable. Il est donc nécessaire que le Pédologue travaille en collaboration avec l'Informaticien aux diverses phases de l'élaboration de la banque des données.

Le glossaire contient tous les termes et les définitions nécessaires à la description d'un sol: horizons, profil, environnement, analyses. Il est évidemment dû au travail de Pédologue. L'informaticien n'est là que pour ordonner les diverses variables de telle sorte qu'une même logique constitue le glossaire.

Les glossaires utilisés en France reposent sur les notions de variables et de données.

La **variable** est un titre et correspond à une question: par exemple la variable **humidité** correspond à la question: quele est l'humidité de tel horizon?

La réponse à cette question est la **donnée**. On répond par exemple: très humide, ce qui est une donnée. La variable humidité comporte 5 données: sec — frais — humide — très humide — noyé. Pour chaque variable il ne faut donner qu'une seule réponse, qu'une seule donnée. D'où cet axiome: les données s'excluent mutuellement, les variables s'adjoignent.

Le codage

Le codage peut se faire de diverses façons. Il est nécessaire qu'il soit fait avec le Pédologue. Ce codage peut avoir des conséquences au moment du traitement statistique des données de sol. Cependant il n'est par très grave d'avoir plusieurs codages à partir d'un même glossaire. C'est ce qui a eu lieu en France, en Belgique et au Canada. Mais on a trouvé un moyen de transformer les codes et donc de pouvoir faire passer l'information entre les diverses banques.

On appelle **répertoire** le document qui donne les équivalences entre les termes du langage naturel et les codes. C'est l'élaboration de ce document qui est fondamentale, si l'on veut que l'information qui sera stockée soit fidèle à la pensée du Pédologue. Le travail demande une part importante pour l'Informaticien, mais la part du Pédologue n'est pas négligeable. Le codage peut être fait au moment de la perforation de cartes, ou bien par un programme dans l'ordinateur. Dans ce dernier cas on perfore les informations en langue naturel.

LA GESTION

Il est nécessaire que les informations contenues dans la banque des données soient fiables. Il faut donc vérifier les informations qui doivent être stockées. En effet une fois stockées, on les considère comme justes.

La vérification

La première opération effectuée lors de l'entrée des données est une vérification de ces données. Un programme vérifie si les données ou les codages sont conformes à ceux du répertoire. On sort alors les données entrées en indiquant en marge celles qui ne sont pas conformes. Comme il peut aussi bien s'agir d'erreurs de perforation que de fautes d'orthographe, de codage ou d'erreurs du Pédologue, ce premier document est adressé au Pédologue qui corrige lui-même les erreurs.

Actuellement on ne dispose pas de programme qui permette de vérifier la cohérence des données pédologiques entre elles. C'est le Pédologue qui doit faire cette vérification. Mais il faudra prévoir de tels programmes.

La modification

La vérification entraîne des modifications premier fichier. Comme il y a toujours des modifications, il est nécessaire de prévoir un programme qui permette les modifications sans tout recommencer. Ce programme existe, il suffit d'une carte programme, suivie des données **modifiées**.

Le stockage

Le stockage se fait au moyen de codes. Il peut être sur cartes perforées ou sur bandes. Dans ce dernier cas on duplique les bandes afin de pouvoir travailler sans risquer d'effacer.

Deontologie

Il est important que la déontologie soit abordée. Elle peut varier selon les pays, mais doit examiner les points suivants:

- aspect humain des échanges d'information,
- affectation des diverses responsabilités,
- aspects de la protection des informations,
- caractérisation de l'information.

Ce dernier point est le plus important pour le Pédologue. L'utilisation des données dépendra du volume de l'information, de la provenance de l'information et du type de dépôt. Pour cela toute information pédologique doit être assortie des coordonnées du dossier: de numéros de référence (du dossier, de la banque...), de la date de la prise d'information, du nom du Pédologue, etc.

Il est nécessaire que les sorties soient faciles à interpréter pour le Pédologue, ces sorties sont de plusieurs types.

Impression

Le plus souvent les sorties se font sur imprimante. Actuellement toutes les données sortent après une mise en forme. C'est ainsi que la description des horizons du profil et de l'environnement sont écrites en clair sur le listage et dans un format compatible avec ceux des dossiers, le plus souvent 21 cm × 29,7 cm. Il n'est donc pas besoin de les recopier. On les photocopie et on les insère directement dans le rapport ou l'étude. Les sorties peuvent se faire en plusieurs langues.

Il en est de même des tableaux d'analyses qui sont présentés selon un standard.

Sorties avec traitements simples

Il s'agit de pouvoir répondre à quelques questions simples. On peut sortir les données sous forme de tableaux de chiffres, d'histogrammes, de courbes du type $y = f(x)$, de points renseignés, sur cartes à échelle variable, de points dans un triangle de texture, de courbes d'isovaleurs.

Les sélections

Tout cela nécessite une sélection des données utiles dans l'ensemble des données de la banque. Pour cela il existe des programmes de sélection, basés sur les opérateurs logiques: ET, OU, NON, le premier, le dernier, le $n^{\text{ième}}$ et PLUS GRAND, PLUS PETIT, EGAL, INEGAL.

On peut de la sorte constituer à partir de la banque des données un fichier sur lequel on peut travailler. Ce fichier peut être sorti sur listage ou sur cartes perforées.

CONCEPTION D'UNE BANQUE INTERNATIONALE

Actuellement un groupe de travail international étudie cette question. Une proposition a été faite, qui repose sur la conception suivante. Il existerait une banque centrale liée par des contrats à des banques locales. Les banques locales auraient pour charge de stocker les informations, d'exploiter les données et de les traiter. Elles garderaient leur autonomie pour leur propres données. Elles se coordonneraient et se compléteraient entre elles par l'intermédiaire de la banque centrale. Elles partageraient les responsabilités déontologiques. La banque centrale

gérerait les informations référencées de chaque dossier de coordonnées géographiques, des noms de ceux qui ont confié le dossier à une banque locale, date de description et d'entrée des données, nom des variables décrites dans le dossier et degré de libération de l'information. Elle assurerait les interfaces entre les banques locales, proposerait des programmes de traitement, orienterait les clients sur les diverses banques locales.

LES TRAITEMENTS DES DONNÉES

Actuellement on peut distinguer trois grandes voies de traitement des données: les traitements statistiques, la cartographie automatique, les utilisations en agronomie.

TRAITEMENTS STATISTIQUES

Un grand nombre de problèmes pédologiques présentent une structure assez semblable. On peut dresser ainsi une liste ordonnée des problèmes qui se posent.

Rechercher les populations existantes

Souvent dans une première analyse on cherche à classer les sols ou leurs éléments. Il faut différencier l'ordination de la classification.

La classification a pour but de définir un nombre réduit d'individus abstraits auxquels on puisse assimiler les individus observés avec une erreur aussi faible que possible. Elle doit permettre de répartir tous les individus réels ou groupes distincts.

L'ordination représente un rangement des individus d'un ensemble en exprimant les relations entre les individus. On peut représenter une ordination par un dendrogramme. Ce dernier explicite l'axiome selon lequel la proximité est plus grande entre deux individus de deux taxons différents. D'un point de vue statistique, on calcule une distance entre chaque couple d'individus. Ces distances se basent sur plusieurs types de données: quantitatives, qualitatives rangées à deux états (présence — absence).

Ensuite on effectue des regroupements entre individus. Plusieurs méthodes existent. Ces méthodes permettent une classification automatique des sols, des individus, des unités cartographiques. On a par là une ouverture possible vers la cartographie automatique. Des applications cartographiques ont déjà été faites. On peut aussi définir des sols — types ou orthotypes.

Caractériser les populations

Il est nécessaire caractériser les populations par la dominante (mode, médiane, moyenne), la fluctuation (intervalle, écart-type, écart — médian), la forme (assymétrie, aplatissement). Il faut ensuite pouvoir comparer les populations (test de Kolmogorow — Smirnov, intervalle de confiance de la médiane, test de Student). Enfin on recherche les corrélations.

Adaptation des méthodes

Toutes les méthodes statistiques dont on a besoin relèvent soit des statistiques classiques, soit de l'analyse des données (analyses factorielles, analyses en composantes principales, analyse discriminante...).

Il est cependant nécessaire d'adapter ces méthodes aux problèmes précis que le Pédologue pose. Des études sont en cours actuellement et des résultats existent déjà dans la littérature.

CARTOGRAPHIE AUTOMATIQUE

On entend plusieurs choses sous les termes **cartographie automatique**. Pour s'y reconnaître, il faut différencier les entrées des sorties, et des programmes d'association.

Entrée des données

On peut entrer les données de trois façon différentes, en vue leur utilisation cartographique. On entre les points correspondants aux observations de terrain ou d'analyse. Ces points sont repérés en x, y et z, et on y adjoint l'information pédologique. La plupart du temps ces points sont géographiquement éparpillés. Il sera difficile de faire des interpolations.

On peut entrer des points renseignés d'une façon systématique. Leur répartition géographique est régulière. Il sera facile d'intrapoler si on édicte une loi pouvant se traduire par un algorithme. Enfin on peut digitaliser des contours déjà existants. Il faut alors entrer l'information pédologique contenue dans chaque surface fermée.

Les programmes

Plusieur types de programmes existent. Ils doivent tenir compte des entrées et des sorties. Sur le plan des algorithmes, on peut en citer plusieurs:

1) Algorithme permettant de changer d'échelle cartographique. Il est bien entendu que la quantité d'information reste la même et donc que **l'échelle de précision** ou la **précision** de la carte ne change pas.

2) Algorithme permettant de choisir quelques variables et de les ressortir en **points renseignés** sur un document de type cartographique.

3) Algorithme permettant de sortir des courbes **isophènes**, c'est à dire délimitant une zone où un ensemble de variables est identique à une précision donnée.

Le problème est de savoir quelles sont les approximations les intrapolations et les types de courbes choisis pour faire passer une ligne entre deux points. On sait que l'intrapolation barycentrique n'est pas la meilleure, de même le plus souvent une courbe de premier ou deuxième degré est suffisante.

4) Algorithme permettant de passer d'un ensemble de variables pédologique à une répartition en classes d'équivalence. Ces algorithmes sont utilisés pour passer d'une carte pédologique à une carte thématique, comme par exemple une carte d'aptitude au drainage.

5) Algorithme permettant de réunir ensemble des zones ou des points estimés semblables à un niveau de précision donné.

Ces algorithmes relèvent d'études statistiques de rangement des données dans les diverses classes, ces dernières étant pondérées, de classement selon des seuils. On peut par ces algorithmes obtenir une classification automatique, qui peut se traduire cartographiquement.

Les sorties

Les sorties les plus simples se font sous une forme de listage, par une imprimante. On sort des points visualisés par des signes alpha numérique. A chaque signe correspond un phénomène particulier. On peut ensuite réunir les signes identiques et obtenir un document de type cartographique. Les points sont repartis géographiquement de manière systématique. On peut aussi sortir des points irrégulièrement répartis, mais renseignés. C'est à dire qu'à côté de chaque point sont indiquées une à quatre valeurs correspondant à une à quatre variables.

Une autre sortie consiste à affecter à chaque point une surface équivalente. On peut alors sortir, sur table traçante, un ensemble de carrés chaque carré étant renseigné par un figuré. Le document cartographique ainsi élaboré présente des surface isophènes, délimitées par des angles droits.

Enfin, la sortie la plus sophistiquée consiste à tracer des courbes ou des polygones du premier ou second degré. Cela nécessite une table traçante et des programmes de sortie élaborés.

Les résultats

On peut distinguer plusieurs types de résultats.

Cartographie assistée

Il s'agit de donner au pédologue des documents à l'échelle désirée provenant de diverses informations déjà acquises: végétation, relief, topographie, géologie etc... On obtient des documents qui peuvent être sur listage. C'est souvent suffisant.

Cartographie isophène

Il s'agit de donner à des utilisateurs, une carte dont l'information provient d'une simplification des données pédologiques. On les appelle souvent cartes thématiques, par exemple: carte de besoin en chaux etc... Les traitements de l'information sont simples, mais il est bon de prévoir une sortie en courbes.

Cartographie automatique

Il s'agit d'entrer des informations pédologiques telles qu'elles sont prélevées sur le terrain, c'est à dire en discontinu. Les algorithmes doivent permettre de classer les informations en divers taxons, ceux-ci devant être représentés cartographiquement en courbes, et à l'échelle. C'est sûrement l'aspect le plus intéressant, mais le plus compliqué. Il n'y a pas encore eu de réussite dans des cas complexes. Les premiers résultats ont été obtenus aux Pays-Bas.

LES UTILISATIONS EN AGRONOMIE

Une des grandes utilisations de la pédologie est la prévision des rendements, ou la recherche des zones favorables à l'implantation de telle ou telle culture dans telle ou telle condition. Le plus souvent l'agronome n'a besoin que de quelques dizaines de variables relatives aux sols. Mais il doit les utiliser en même temps que des données provenant d'autres domaines ou expériences. Les données pédologiques servent alors à expliquer en partie un autre phénomène (variables statistiques et corrélations) ou à prendre des décisions (variables décisionnelles). Les méthodes statistiques sont différentes. Mise à part l'analyse de variance, on utilise souvent les analyses factorielles, l'analyse canonique, l'analyse discriminante à une, plusieurs populations, les régressions multiples pas à pas. Ces méthodes permettent, après une phase d'analyse et une phase de synthèse, une modélisation. On peut schématiser ces phases comme suit.

On cherche à dégager des hypothèses sur les variables ou les regroupements possibles, puis une hiérarchisation dans les variables. Ensuite on peut établir des correspondances entre deux ensembles de données parallèles, puis on vérifie la validité des critères de mesure. Enfin, on peut modéliser „l'aptitude culturale des sols” et mettre en équation les „prévisions de rendements”.

DONNEES PROVENANT DE DIVERSES BANQUES.

LA TELEDETECTION

RELATIONS ENTRE BANQUES

Les informations saisies par le Pédologue intéressent diverses autres branches de la science ou de la technique. De même pour accroître sa connaissance des sols, le Pédologue utilise des informations provenant d'autres spécialités.

On peut citer: la botanique, l'agronomie, la biologie, l'hydrologie, la géomorphologie, la géologie, la chimie, la physique, la contrainte des matériaux (génie), l'étude de la pollution, de l'environnement...

Une banque de données de sol ne peut donc être qu'un maillon dans un ensemble beaucoup plus vaste de banque de données de la Biosphère. Il est nécessaire qu'elle reçoive l'information de certaines banques: climatologie, botanique, géologie... et qu'elle en fournisse à d'autres: géologie, botanique, écologie... Il y a donc lieu d'étudier les phénomènes de transfert informatique des données entre les diverses banques. Ceci est réalisé sur des prototypes.

UN EXEMPLE: LA TÉLÉDÉTECTION

Depuis l'apparition des satellites des ressources terrestres, il existe potentiellement une grande quantité d'information. On peut se demander si elles sont utiles pour le pédologue, en quoi et comment.

La matrice des données

Les données stockées sur bandes dans les centres de réception sont directement utilisables sur ordinateur, dans la mesure où il existe des compatibilités entre systèmes.

La caractéristique de ces données est la suivante: en un point il y a peu de données. Sur LANDSAT, il y a 4 canaux, donc 4 données. Pour l'étude d'une zone, on dispose de beaucoup de points. La matrice des données a beaucoup de lignes (Individus) et peu de colonnes (caractéristiques ou variables).

On a vu que c'était l'opposé pour les matrices d'information des données sol: peu d'individus (de lignes), et beaucoup de colonnes (variables). Les méthodes de traitement des données de télédétection ne seront pas les mêmes que celles de traitement des données d'observation et analyse des sols. En principe il est possible d'utiliser conjointement les deux sources d'information, mais il se pose un certain nombre de problèmes scientifiques et techniques. En effet il n'est pas facile de savoir à quel emplacement du terrain correspond exactement un point de l'image de télédétection, et il est encore plus difficile de savoir à quels éléments

du sol sont dûes les réponses électromagnétique enregistrées par le capteur.

La télédétection supervisée ou non

L'étude à distance des objets suppose que ceux-ci puissent être entièrement définis par des propriétés physiques, telles que la réflectance ou l'émission dans certaines bandes spectrales. Actuellement les scientifiques recherchent les caractéristiques spectrales permettant de définir, et par conséquent d'identifier à coup sûr, les différents objets terrestres: sols, types de cultures...

Si l'on pouvait définir ainsi chaque objet, il serait ensuite extrêmement simple de l'identifier sur photographies aériennes ou images de télédétection. Une analyse automatique de ces documents selon les critères de reconnaissance précédemment définis, permettrait une identification fiable. C'est ce vers quoi tend la télédétection. On suppose qu'il existe suffisamment de propriétés physiques pour pouvoir séparer entre eux tous les objets de la surface terrestre, même s'ils sont de même nature, tels que des sols entre eux.

Des équipes de physiciens, aux Etats-Unis, entre autres et aux Pays-Bas, ont proposé certains modèles permettant de retrouver automatiquement sur des enregistrements multispectraux diverses cultures. C'est ce que l'on appelle la **télédétection non supervisée**. Pendant un certain temps, on a pensé aux Etats-Unis que la réflectance des objets dans le visible et le proche infra-rouge devrait seule suffire pour les définir tous et établir des modélisations. D'où ces catalogues de réflectance établis il y a environ 10 ans sur différents cultures et plusieurs espèces arborées.

Malheureusement, la réflectance ne suffit pas à définir avec suffisamment de fiabilité tous les objets terrestres, ne serait-ce que parce que deux Bouleaux (*Betula verrucosa*) de même âge, mais poussant dans des conditions écologiques différentes, risquent d'avoir des réflectances spectrales différentes. Les causes de variation de la réflectances, dûes aux conditions atmosphériques et à la position du soleil, lors des enregistrements, peuvent être corrigées et n'introduisent plus maintenant de sources d'erreurs dans la reconnaissance automatique.

Il subsiste néanmoins beaucoup d'incertitudes, et, si l'on peut souhaiter pour son complet développement, que la télédétection puisse être non supervisée, on n'en est malheureusement pas encore là. Toute étude de télédétection nécessite encore l'établissement de zones-test où il est nécessaire de faire des mesures au sol: radiométrie, humidité, température... pour pouvoir interpréter efficacement les résultats. C'est pourquoi tous les programmes d'étude des images LANDSAT prévoient des zones-test où les mesures au sol sont nombreuses et variées.

En ce qui concerne les sols, les difficultés de modélisation sont plus grandes que pour d'autres objets tels que les arbres ou les cultures. En effet la notion d'individu est difficile à établir pour les sols que pour les végétaux. D'autre part la télédétection dans le visible et l'infra-rouge ne peut donner des indications que sur la surface des sols. Pour les sols battants, seule la croûte de battance interviendra dans la réponse, ne laissant pas présumer de la nature du sol situé dessous.

Cela explique le développement plus lent de la télédétection en pédologie, vis à vis d'autres domaines comme l'agronomie et la forêtierie. Seules les techniques de Télédétection donnant une information sur un volume et non plus sur une surface pourront permettre des études pédologiques plus poussées.

Les hyperfréquences, dans la mesure où elles ont une pénétration dans les sols, apportent de nouveaux éléments. La détection par radar en est à ses débuts, mais on peut espérer beaucoup sur son utilisation en pédologie.

La méthodologie de la télédétection en pédologie

Sur les images, comme sur les photographies aériennes, on peut définir la **texture** qui correspond aux valeurs (le plus souvent intensités de grisés) des différents points de ces images, et la **structure**, qui correspond aux relations existant entre ces différents points. Sur les photographies aériennes, ces notions de texture et de structure sont facilement reliées à des objets ou groupes d'objets bien précis à la surface terrestre.

L'interprétation est relativement simple pour les émulsions panchromatique et couleur tout au moins, la portion du spectre enregistré correspond à ce qui est perçu par l'oeil humain. Un sol sera plutôt jaune ou rouge ou noir, l'ensemble des sols d'une région, tels qu'ils sont vus dans les champs non occupés par la végétation, seront de teinte très claire. L'interprétation des émulsions infra-rouge noir et blanc et fausse-couleur est déjà plus délicate. En effet si les sols sont facilement perceptibles leur grisé ou leur couleur sur le cliché ne peut être comprise que par un raisonnement faisant référence à leurs caractéristiques spectrales.

Telle parcelle est bleu foncé parce que un labour récent fait ressortir son humidité, par contre, telle tache foncée dans un champ correspond bien à une certaine hydromorphie. Mais si l'ensemble des parcelles a été labouré récemment, des taches d'hydromorphie passeront inaperçues, masquées par l'influence des travaux du sol. En ce qui concerne les images, l'interprétation devient très délicate: en effet elles correspondent à l'enregistrement de très petites portions du spectre qui n'ont aucun rapport avec ce qui est perçu par l'oeil humain.

Toute interprétation des textures doit être faire en référence avec les caractéristiques spectrales des objets dans cette bande spectrale.

Ainsi dans la bande 450 — 500 nm, un sol argileux sec peut avoir une réflectance plus forte qu'un sol sableux humide et donc apparaître avec un grisé plus clair. D'autre part, sur des enregistrements à haute altitude, l'élément le plus petit de l'image, correspond, au sol, à une grande surface. La valeur de cet élément de l'image, correspond à la moyenne des réponses des objets contenus dans cette surface unitaire.

L'interprétation de la valeur et sa mise en rapport avec des objets sur le terrain sera difficile, voire impossible. Si dans cette surface se trouve un objet, petit, mais de caractéristique spectrale très contrastée avec celles des objets voisins, sa réponse sera prédominante vis à vis des autres, et il pourra même être vu sur l'image, bien que ses dimensions soient inférieures à la résolution au sol. Ainsi une tranchée routière récemment creusée, sera parfaitement visible sur les images LANDSAT.

Sur les images à haute altitude, les structures visibles correspondent donc à des différences entre des surfaces, les contrastes qui existent entre elles sont plus ou moins forts. Ainsi une zone karstique au sein d'un ensemble cultivé aura un très fort contraste. Par contre différentes terrasses d'une plaine alluviale auront un contraste très faible et pourront, à la limite, passer inaperçues.

L'interprétation des images est donc une chose délicate. Elle suppose tout d'abord une excellente connaissance des phénomènes physiques, enregistrés dans la bande spectrale considérée. Elle suppose aussi une bonne connaissance du terrain permettant d'apprécier les phénomènes apparaissant préférentiellement à d'autres. La télédétection peut apporter une aide précieuse en pédologie pour la préparation du travail de terrain, en déterminant des zones homogènes et en permettant donc de mieux choisir les emplacements des profils.

Elle permet aussi de réaliser la cartographie en donnant une vue synthétique des phénomènes, c'est à dire des relations existant entre différents ensembles. C'est cette notion qui apparaît bien sur les images satellite à propos des contrastes.

BIBLIOGRAPHIE

- AUXENFANTS G., 1976: *Méthodes de l'analyse des données et sciences du sol. Du problème pédologique au protocole, du protocole aux résultats*, „Bulletin de Informatique et Biosphère”, n° 6 [Paris].
- GIRARD C. M., GIRARD M. C., 1975: *Applications de la télédétection à l'étude de la biosphère*, Masson, Paris.
- GIRARD M. C., 1976: *Recherche d'une méthodologie en matière de traitement statistique des données de sol. Application à la taxonomie et à la cartographie*, „Bulletin Science du Sol” n° 3 [Versailles].

- Informatique et Biosphère. Glossaire de pédologie. Description des horizons en vue du traitement informatique*, 1969, Paris, p. 82.
- Informatique et Biosphère. Glossaire de pédologie. Description de l'environnement en vue du traitement informatique*, 1971, Paris, p. 173.
- Informatique et Biosphère. Rapport de synthèse sur les objectifs et les structures juridiques et techniques d'une banque de données de sol*, 1975, Directeur scientifique: M. C. Girard, Doc. ronéo., Agence de Coopération Culturelle et Technique, janvier.
- JONGEN P., 1975: *L'utilisation des méthodes informatiques dans l'étude de mise en valeur des terres*, International Symposium on „The World Food Problem”, février, Kartoum.
- LEGROS J. P., 1973: *Eléments de statistiques appliqués à la science du sol*, 105 p., SES n° 214, décembre, Montpellier.
- LEGROS J. P., 1974: *Le stockage informatique des données pédologiques: avantages et contraintes*. Actes du Colloque Informatique et Environnement, 18 p., mai, Arlon.
- LEGROS J. P., 1975: *Cartographie automatique. Exposé des méthodes et des problèmes*, Doc. ronéo., SES n° 265, octobre, Montpellier.
- NGO CHAN BANG, 1975; *Application de l'informatique et des statistiques modernes pour l'expérimentation agronomique et l'aptitude culturale des sols*, Doc. ronéo. Informatique et Biosphère, décembre, Paris.
- TROUJAN S., 1973: *Prévision des rendements agricoles à partir des variables du milieu*. 1^{er} Colloque Informatique et Biosphère. p. 51—74, Paris.
- Van den DRIESSCHE R., 1974: *La banque de données pédologiques de l'ORSTOM*. Cahier ORSTOM, Série Pédol. Vol. XII, n° I, p. 125—132, Paris.
- VIANELLO G.: *Proposte per una metodologia di base per la formazione dei piani comprensoriali. Regione: Emilia, Romagna, MB I*, Bologna.
- Soil site coding scheme for data input to the soil data file of the Canada Soil Information System (CANSIS)*, 1974, Soil Research Institute, Ottawa.

MICHEL C. GIRARD

OPRACOWANIE INFORMATYCZNE DANYCH GLEBOWYCH

Streszczenie

Autorzy proponują rozwiązanie problemu, który powstał w 1974 roku w Moskwie podczas Kongresu Gleboznawców. Dotyczy on olbrzymiej liczby danych gleboznawczych, jakie zostały zgromadzone na podstawie badań terenowych i laboratoryjnych. Jak wprowadzić te informacje do banku danych? Jak je wykorzystać? Dane terenowe i szczegółowe opisy stanowią wejście do systemu banku informacji, gdzie zostają zakodowane. Tam kontroluje się ich prawdziwość, ujednolica je i gromadzi na kartach, taśmach magnetycznych itp. oraz porządkuje. Wyjście z banku polega na prostej obróbce danych i na ich selekcji. Ten etap jest już bliski wspólnemu bankowi międzynarodowemu. Dalszym etapem informatyki glebowej jest obróbka statystyczna i kartografia automatyczna.

Ujednolicenie i uzgodnienie danych glebowych, pochodzących z różnych banków, jest możliwe dzięki teledetekcji. W pierwszym etapie jest to wynik zdalnej obserwacji nie kontrolowanej, w drugim zaś etapie kontrolowanej w różnych pasmach widma elektromagnetycznego. Satelita LANDSAT służy dobrymi obrazami syntetycznymi dzięki wykorzystaniu kontrastu między strefami glebowymi.

W ten sposób można będzie konstruować podstawę ogólnoswiatowego banku danych glebowych.

SOIL DATA PROCESSING

Summary

Suggestions are made as to how to solve the problem discussed in Moscow at the 1974 Congress of Soil Scientists. The problem itself consists in the large amount of soil data that have been achieved from field and laboratory investigations. The question arises of how to store these data in a data base and how to utilize them. Field data and detailed specifications form the inputs of the data base system where they are coded. In the base, the data are verified, unified, recorded on cards and magnetic tapes, etc., and are subject to ordering bias. The output consists in a simple processing and selection of data. Thus, we are just near a common, international data base.

The next step of soil informatics is statistical processing of data and automatic cartography.

A unification and adjustment of soil data coming from different base becomes feasible due to remote sensing. In the first stage, this is achieved by an uncontrolled remote observation, and in the second stage by an observation controlled in various bands of the electromagnetic spectrum. LANDSAT provides high-quality images owing to the utilization of contrast between soil zones.

Thus, the principles of the world base of soil data can be established.