

OFFICE DE LA RECHERCHE SCIENTIFIQUE ET TECHNIQUE OUTRE-MER

Section de Pédologie

Travaux en informatique appliquée

VAN DEN DRIESCHE -000-

Il est désormais urgent de concerter une action dans le domaine du stockage sur ordinateur, puis de l'utilisation des données d'observation et d'analyse des sols.

La France a, dès à présent, une certaine avance dans ce domaine, grâce aux travaux effectués par la Section de Pédologie sur l'ordinateur CDC 3600, avec les programmes écrits sur demande, de 1962 à 1967, par Mme MASBOU de l'Institut Blaise Pascal du C.N.R.S.

Cet exposé n'est qu'un bref résumé des méthodes. Le texte complet, programmes compris, sera remis sous peu aux Pédologues présents aujourd'hui. L'Office accepte de traiter, par ses méthodes, les données pédologiques qui lui seront confiées ; il souhaite une adhésion très active des Pédologues étrangers à cette entreprise. Il pourrait se constituer, ainsi, une "banque de données pédologiques agissant sur un plan international pour le compte de pays collaborant entre eux".

Le traitement par correspondance, sur le CDC 3600 du C.N.R.S., est possible, après la mise sur bande magnétique des programmes.

I) - Données pédologiques obtenues au laboratoire

Une "fiche analytique" de papier calque ou de papier blanc est utilisée pour chaque profil. La première est dactylographiée et reproduite par tirage héliographique ou autre ; la seconde est manuscrite et tirée sur Xérox.

O. R. S. T. O. M.

Collection de Référence

n° 7-3580 B

Les données sont numériques ; leur inscription est faite en regard des variables choisies parmi les 52 variables répertoriées sur l'imprimé. Les variables choisies ne doivent pas être les mêmes dans les différents horizons. Les unités de mesure sont imposées et les nombres compris entre 0.01 et 9999. Aucune inscription pour "donnée manquante" n'est demandée, mais une donnée "nulle" s'inscrit .001 tandis que le mot "trace" est remplacé par un nombre compris entre .002 et .009 selon la variable. Le point décimal est toujours inscrit après les nombres entiers de 1, 2 ou 3 chiffres.- Ce format d'inscription est vérifié, après perforation, avec le programme ASV.

Outre les variables de laboratoire, la fiche porte également des inscriptions d'identification, de localisation et de classification. L'examen des variables dérivées (C/N, S/T, etc...) qui figurent, à leur place, dans les bordereaux complets imprimés par l'ordinateur lors de la mise sur bande magnétique, avec le programme BSV, peut conduire le pédologue à remplacer ou compléter certaines données de laboratoire ou de classification. Il n'en résultera qu'un léger retard. En effet, les délais de perforation et d'exécution par les programmes ASV et BSV sont réduits pour les séries ne dépassant pas 30 profils.

Après la mise sur bande, il est fait appel au programme DSV pour la sélection des enregistrements par une ou plusieurs de leurs 78 variables ; la transformation logarithmique décimale des variables demandées ; l'établissement des nuages de points demandés ; le calcul des matrices de corrélations ; l'estimation des limites de confiance des moyennes retransformées ; la perforation des données sélectionnées.

La phase la plus intéressante du dépouillement porte sur ces petits fichiers sélectionnés qui correspondent à autant d'études locales ou internationales de groupes, sous-groupes, horizons. Les programmes KMV et MMV leurs sont appliqués, pour le calcul des distances D^2 , lorsque le nombre de variables communes ne dépasse pas 6 ou lorsque

les effectifs intra-groupe ne dépassent pas la moitié du nombre de variables. Pour des effectifs plus importants ($n_1 + n_2 > v + 1$) et un nombre de variables qui s'y prête ($7 \leq v \leq 20$), on utilise les programmes KMV et RMV pour le calcul des D^2 , la recherche des constellations et la recherche des variables les plus utiles à la mesure des D^2 .

2) - Données pédologiques obtenues sur le terrain

Des fichiers de cartes sont perforés à partir de documents codés, à titre expérimental, à partir des fiches de calque intitulées "description du profil".

Les variables suivantes sont retenues : Munsell sec ; pourcentage, contraste et couleur dominante des taches ; quantité, décomposition et degré d'association de la matière organique ; charge et forme du calcaire ; texture dominante, secondaire, ses éléments grossiers ; forme de la structure particulaire, massive, fragmentaire, ses dimensions et son grade ; la porosité ; l'adhésivité, la plasticité, la friabilité ou la fragilité ; la cohésion ; quelques caractéristiques particulières ; l'enracinement ; la netteté et la régularité des limites inférieures de l'horizon. Ces variables sont de type différent, et utilisées simultanément pour deux types d'horizons, ce qui nous conduit à utiliser le programme GDL pour les indices de similitude, mis au point dans un autre domaine et offert par le Professeur GOODALL. Pour des séries ne dépassant pas 30 profils, les compléments des indices de similitude servent, ensuite, à la recherche des constellations avec le programme VMV.

Les cartes perforées, ou des bandes magnétiques, seront obtenues plus tard, si les essais sont satisfaisants, à partir de feuillets marqués ou manuscrits, de type 1232, 1231 ou 1287. L'utilisation directe, sur bande, du vocabulaire de la pédologie n'est cependant pas écartée et les travaux effectués, dans ce domaine, par l'équipe du Professeur SWINDALE sont suivis très attentivement.

3) - Données respirométriques et enzymatiques

De petits fichiers perforés, dont les enregistrements renferment 1 à 20 variables par échantillon composite de surface, sont établis à partir de feuillets quadrillés intitulés "frequency functions".

Le programme HMV pour le calcul des distances Δg (delta g) et la recherche des constellations donnent satisfaction.

4) - Données de la micromorphologie

Des bordereaux ont été établis avec l'E.N.S.A. de Grignon pour l'enregistrement des variables observées dans chaque lame-mince.

Les programmes GDL et VMV peuvent être utilisés.

5) - Données géochimiques en général

Les travaux du Comité international pour le stockage, la recherche documentaire et le traitement automatique des données géochimiques, que préside le Professeur ROUBAULT sont suivis avec le plus grand intérêt. Ils concernent, actuellement, les spécifications à joindre aux données des roches, des eaux et des sols.

Raymond Van den DRIESSCHE

Bondy, le 27.09.68