UPDATING MAPS OF SOILS THROUGH LANDSAT IMAGES
COMISSION IV

Gilberto J. Garcia
Department of Cartography - Geosciences Institute
São Paulo State University, Rio Claro - SP, Brazil

Mario Valerio Fº
Department of Remote Sensing
Space Research National Institute - S.J. Campos - SP, Brazil

Paul S. Anderson
Department of Geography - Geology
Illinois State University, Normal, IL-USA and Fulbright
Professor to São Paulo State University, Rio Claro - SP, Brazil

ABSTRACT:

Soil maps are basic in integrated studies directed toward land management. Brazil and other developing countries have few areas mapped at adequate scales and definition of soil boundaries. The objective is to analyse the usefulness of spectral characteristics of soils for the improvement of existing maps. Digital processing and statistical analysis of Landsat/TM images in one agricultural region in São Paulo state were used to produced a new soils map which was compared it to existing maps. The methodology was successfully utilized in areas without dense vegetative cover (e.g., tilled and prior to major plant growth). The authors believe that the methodology can be applied in many other agricultural regions.

Key words: Landsat, soils, spectral characteristics, maps.

INTRODUCTION:

In the integrated analysis of landscapes, for whatever purpose, information about soils is essential. Although fundamental, soil maps in Brazil (and other developing countries) at adequate scales and detail levels are rare, even in the most developed regions. The State of São Paulo, Brazil, has old soil map coverage at 1:250,000 by the Soils Commission, ((dating from the 1960s. Currently, the Agronomic Institute (IAC) of São Paulo state has conducted mapping at the scale of 1:100,000. By mid-1992, approximately fifteen percent of the state has been mapped. The maps serve well for regional studies, but are generally insufficient for local studies at scales between 1:25,000 and 1:50,000. When using traditional methods to enrich the IAC map information for use at the local level, analyses of geology, topographic relief, drainage network and field work are needed.

For a more rapid, economic and efficient enhancement of the IAC maps, we propose an alternative method. The method uses statistical analysis of the spectral character of exposed soils as seen in satellite images obtained in periods of tillage prior to major plant growth. In this paper we present our methodology and the results from a test area to produce a soils map at a scale of 1:50,000.

REVIEW

In their early (1974) study, Haralick and Shanmugan found that spectral expressions, textures and situational context are the three elements used in visual interpretation of satellite images. In terms of spectral behavior, soils exhibit spectral curves (signatures) that are much more uniform than those of rocks. Studies by Montgomery and Baungardner (1974) and Stoner and Baungardner (1981) show that reflectance of soil is a cumulative property resulting from the heterogeneous contribution of organic matter, iron oxide, moisture, granulometry, and structure. According to Stoner et alii (1980), soils can be grouped in five basic spectral curves, as shown in Figure 1.

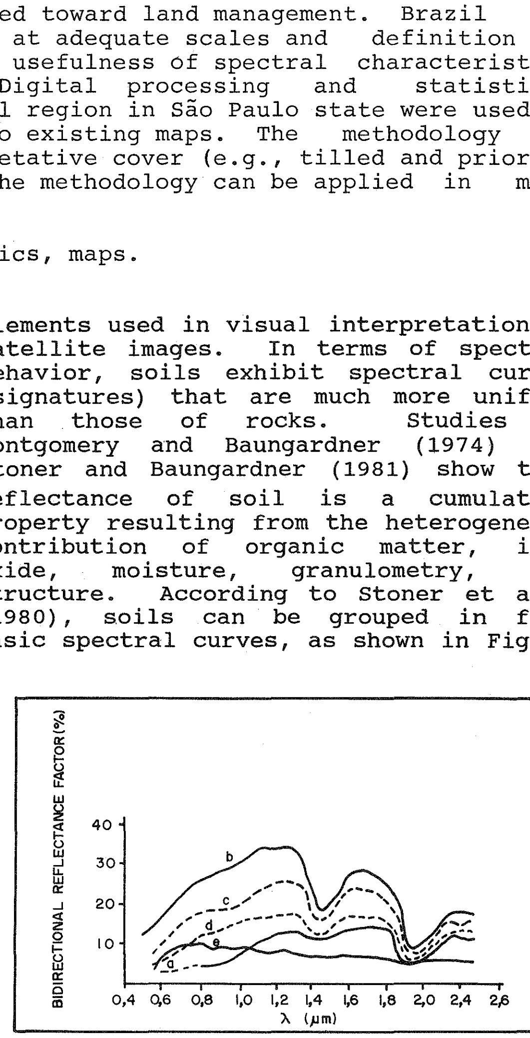

Fig.1 - Five spectral soil patterns. Dominant factors: a) organic matter (O.M.); b) low weathering with low O.M. and medium iron; c) affected by iron; d) affeted by O.M.; e) iron.

As demonstrated by Page (1974), the amount of organic matter is inversely correlated with the spectral reflectance. However, Baungardner et alli (1970) comment that levels of organic matter above two percent mask the effect of other properties. Furthermore, in the case of iron oxide, even levels above four percent can mask the effects of high levels of organic matter.

Montgomery et alii (1976) state

that the significance of iron oxides in the reflectance increases with the increase of the wavelength in the electromagnetic spectrum, especially in the regions of visible light and near-infrared.

Cipra et alii (1976) compared spectral-radiometer measurements of exposed soils with digital data from Landsat-1. Results showed that Landsat radiance and spectroradiometer reflectance values were highly correlated for all wavelength bands.

Lund et alii (1980) and Harrison and Johnson (1982) concluded that the use of spectral maps derived from Landsat data improved accuracy and/or quality of map unit delineations. More recently, Coleman and Montgomery (1987) and Everitt et alii (1989) studied the same question. The former encountered great interdependence between the moisture content and the reflectance of the respective soils. The latter studied the moisture content, organic mater and the level of iron in alfisols and vertisols, finding high correlations between reflectance and the studied variables. In both cases the soils were separated and accurately mapped on the basis of spectral, physical and chemical properties.

Agbu and Nizeyiama (1991) compared soil maps from SPOT spectral data with maps produced in the field. Although the maps based on field work were found to be better than those based on spectral analyses, the differences did not attain statistical significance at the 0.05 level, using the Kappa statistic.

Visual analysis of spectral differences are insufficient for the required studies. Although only a few (4 to 12) bands are selected from the continuum of electromagnetic energy, each band contains a continuum of extremely small variation of the intensity of reflectance in that band. Therefore, the number of combinations of bands (colors) is extremely large. Consequently, research that deals with such spectral behavior of targets in images from space satellites require digital analyses via computer processing. These capabilities exist in image analysis systems such as ERDAS or, in Brazil, SITIM (Sistema de Tratamento de Imagem) from the Space Research National Institute, INPE.

Among the methods for spectral image analysis, of particular note are those that are based on the statistical distance between probability densities that characterize the standard classes. These methods include divergence, transformed divergence, Bhattacharyia's distance and Jeffreys-Matusita (JM) distance (Swain and King, 1973, and Richards, 1986).

METHODS

1. Description of the study area

The study area is approximately 10 x 10 minutes of latitude and longitude (220 square kilometers) on the Araras topographic sheet in the state of São Paulo, Brazil (see Figure 2). The area is tropical, being 130 kilometers north of the Tropic of Capricorn. The maximum and minimum elevations are 560 and 680 meters above sea level. The prominent relief is a slightly rolling landscape. Only ten percent of the area has limitations that prevent mechanized agriculture.

According to the maps of the IGC (1982), the geology of the area includes rocks from the Tubarão Group, the Irati and Corumbataí (siltstone and shales) formations of the Passa-Dois Group, basic intrusives, sandstones from the Botucatu-Pirambóia formation, and the Cenozoic.

In the Koppen system of climatic classification, the climate of the area is mesothermic with dry winter, type Ewa. The winter dryness extends from April to September; the rains for summer occur from October to March. June-July temperatures average 18º C (64º F), rising to 22º C (72º F) in January-February. Frosts do not occur.

The natural vegetation is classified as subtropical forest. Today the area is used for sugar cane, citrus, cotton and corn agriculture. Pastures and reforestation are found in the steeper areas. Keeping in mind the methodological considerations of the research, we selected an area predominately occupied with annual crops and obtained images from the period prior to planting. The major part (85%) of the area was free of vegetation.

The soils of the area, according to Oliveira et alii (1982), are listed below, in order of highest to lowest occurrences. Their approximate distribution, according to the pre-existing map at 1:100.000, is shown in Figure 2.

. LV - Latossolo Vermelho Amarelo
  (USA) - Quartzipsammentic Haplorthox

. LR - Latossolo Roxo - eutrófico
  (USA) - Typic Eutrorthox

. PV - Podzólico Vermelho Amarelo
  (USA) - Typic Paleudult

. TE - Terra roxa Estruturada - eutrófica e distrófica
  (USA) - Rhodic Paleudalf + Rhodic Paleudult

. AQ - Areias Quartzosas
  (USA) - Typic Quartzipasamment

. Hi - Solos Hidromórficos
  (USA) - Hydromorphic soils

. LE - Latossolo Vermelho Escuro
  (USA) - Typic Haplorthox

2. Characteristics of the images and equipment used

Analogue (1:100,000) and digital images from Landsat Thematic Mapper (TM) were obtained for six bands of visible

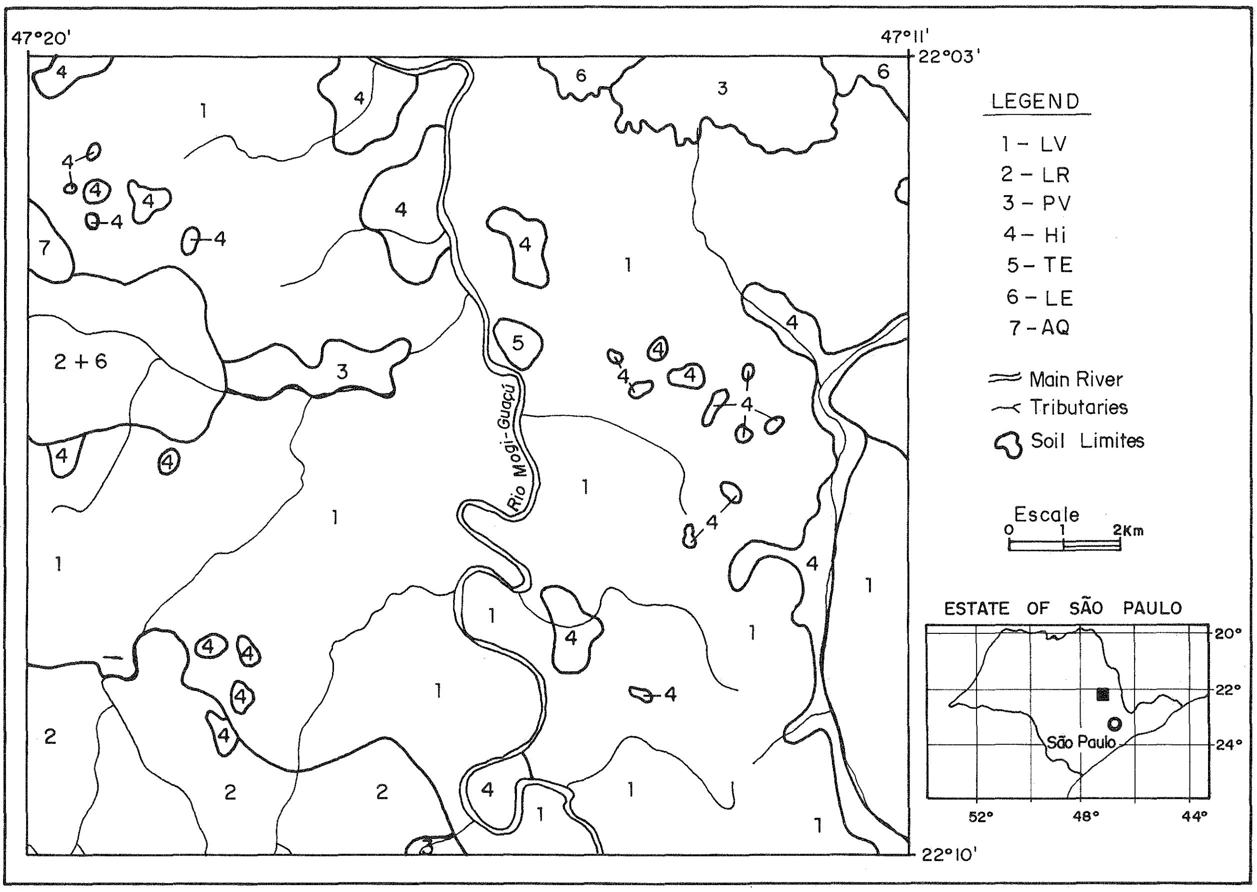

Fig. 2 - Study area as mapped in 1982 by field work of the IAC - Instituto Agronômico de Campinas, published at 1:100.000.

and reflected infrared radiation. The images were from orbit 220/75D for the month of December. The topographic quadrilateral of Araras (1:50,000) by the IBGE was used as the cartographic base for plotting information.

The digital images were processed with the SITIM-150 system on a microcomputer with a PROCON projector/enlarger. The work sequence is presented in Figure 3.

3.Analysis of the images

For selecting the subgroups of bands for generation of the color composites, the Jeffreys-Matusita distance method was used, as discussed by Swain and King (1973). The JM distance is an appropriate technique to measure the average separability between spectral classes, calculated as functions of probability density. The following researchers implemented or applied the JM method: Bendat and Piersol (1986), Andrade (1985), and Paradella (1984). This technique is a convenient alternative for selecting the best color composite images. The series of interband statistical measurements of the JM method results in a reduction of the dimensionality, processing and redundance of data.

In general, each class of interest (e.g. Typic Eutrorthox) in an image can be characterized by a function of density of probability P1 (x) that gives the values of probability densities that the pixels x belong to a class in function of x. For two classes w_i and w_j, the JM distance is defined as:

$$JM_{ij} = \int_X \left\{ \left[P_i(x)\right]^{1/2} - \left[P_j(x)\right]^{1/2} \right\}^2 dx$$

where:

JM_{ij} = JM distance between classes w_i and w_j;

$P_i(x)$ = probability density of the pixels belonging to class w_i;

$P_j(x)$ = probability density of the pixels belonging to class w_j;

X = range of interest for the X values

The software implemented on the SITIM system calculates the JM distances between classes selected by the user for all possible combinations of bands. The output includes subsets that maximize the JM average and minimum distance criteria.

As the method to classify the scene, a multi-variate analysis was applied that offers the advantage of working with both parametric and non-parametric data. Cluster analysis was adopted in order to work with a group of units characterized by diverse variables. The result is the separation of existing groups characterized by homogeneity

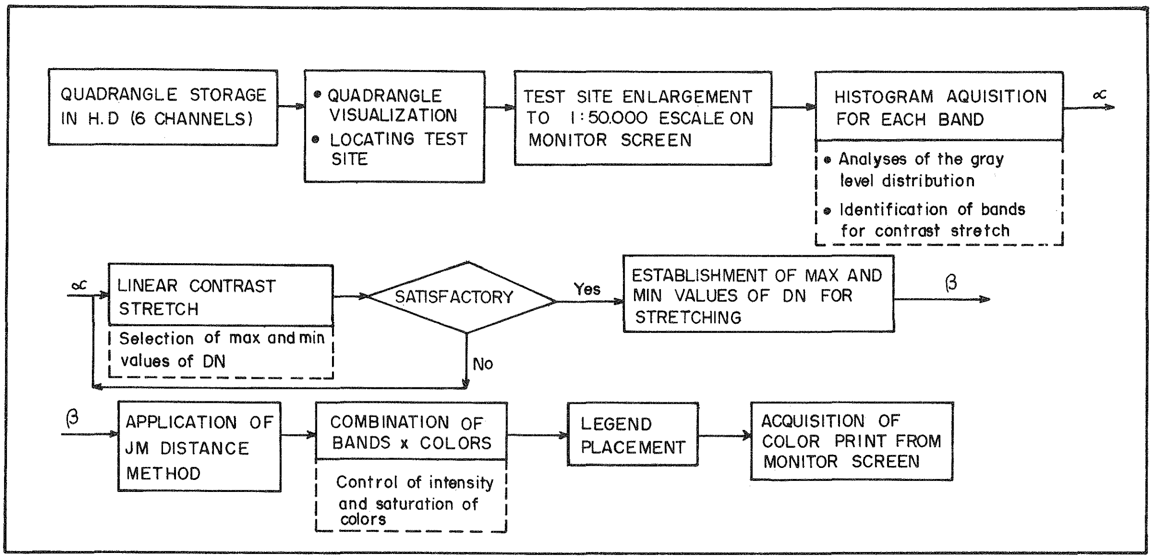

Fig. 3 - Work sequence for the production of color composite images.

between the elements within the group and by the heterogeneity between elements in different groups (Curi, 1982). The specific analysis called ISOMIX (Prelat, 1981) was used because of its great discrimination power, with the advantage of being interactive.

The degree of differentiation between the groups is measured by the similarity of the "centers of groups". If the distance between two groups is less than a specified limit, the two centers of groups are joined and have the average value of the two original groups. The process is cyclically repeated until the standard deviation of each group center is less than the specified level, or until the maximum number of permitted groups is attained. The control parameters specified by the analyst are: a) standard deviation - controls the number of classes; b) minimum group centers and maximum group centers - controls the category level to be attained; c) minimum number of pixels per group - below which it is not possible to form a new group; and d) separation threshold - to verify if the classes are similar or different.

RESULTS AND DISCUSSION

The application of the above described method of JM distance indicated, in decreasing order of efficiency, the following combinations: 254 (2B 5G 4R), 354, and 574, in the colors blue, green and red, respectively. Through visual examinations, we also considered the following additional combinations to have informative content: 274, 157, 174, 247, and 154. Tables 1 and 2 show the results of the JM distance method.

The classification of spectral information from the cluster analysis (ISOMIX) permit the union of the soil units, whether homogeneous or not, as represented in Figure 4. Table 3 presents the control parameters for the test application.

Table 1 - Rank order of best original TM bands originais

Bands	Results
4	1.4325
5	1.2604
2	1.1288
3	1.0936
7	1.0850
5	1.0682

Table 2 - Rank of the six best subsets of TM bands

Subsets	Results
4	1.4325
45	1.5821
$\boxed{245}$	1.6942
$\boxed{2345}$	1.7624
23457	1.7820
123457	1.8343

N.B.: Rectangles indicate the best subsets of three or four TM bands.

Table 3 - Control parameters for Isomix

Parameters	Values
STANDARD DEVIATION	2,00
MIN GROUPCNTRS	4
MAX GROUPCNTRS	6
MIN PIX/GROUP	50
SEPARATION THRESHOLD	2,50

The delimitations of the soils in the study area based on the composition 2B 5G 4R are shown in Figure 4. Comparing this map with the one in Figure 2 reveals that much of the information coincides, especially for the largest

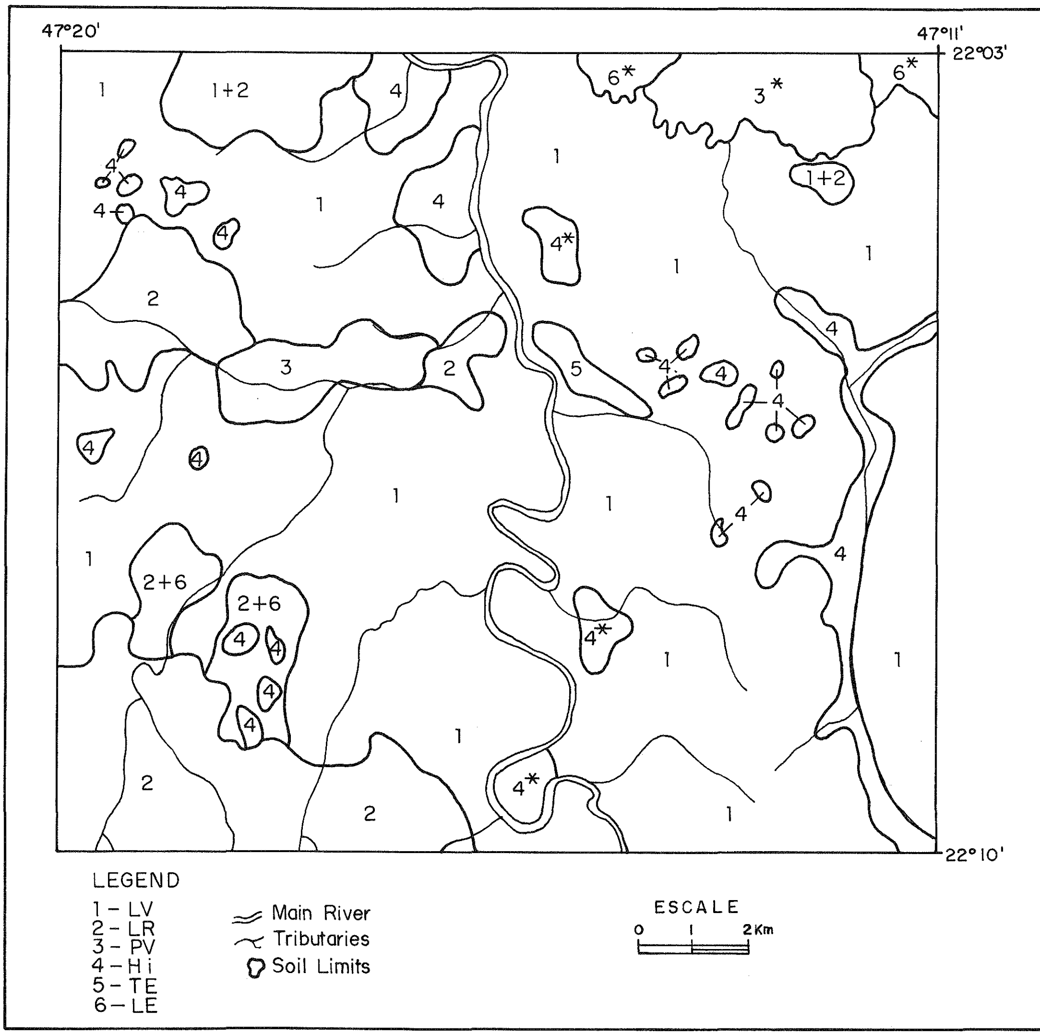

Fig. 4 - Spectral map of soils of the study area + field work map.

patches of soils. Nevertheless, the new map presents a greater richness of details, has better definition between the soil boundaries, and reveals important transitions between soils.

A comparison between the two soils maps was made with a 2x2 mm grid sample (5425 grid cells, each representing 200 x 200 meters on the earth surface). The fieldwork map is used as the reference. Comparing the homologous squares, the percentage of coincidence (86.9%) was calculated with the following formula:

Percent = Number of correct predictions of map units x 100
Corret Total number of map units in the sample

Figures 2 and 4 indicate that the main discrepancies appear in the spectral-data map in patches of homogeneous or associated soils that are absent from the map based on field work. Subsequent field checks have verified the existence of the patches. Of the total 13.1% error in coincidence, 7.8% resulted from deficiencies in the map based on field work, whereas 5.3% were from insufficiencies of the spectral map. Of the 5.3%, 1.7% derived from patches of hydromorphic soils whose reflectances are masked by dense natural vegetation. We note that both types of errors are from omission of detail, not from excessive detail that was incorrect.

The results of the study match well the existing literature. The spectral differences between soil types are derived mainly from the different percentages of organic matter and iron oxide (Montgomery and Baungardner, 1974, and Stoner and Baungardner, 1981). In this case, the soils in the area are quite distinct between themselves, especially in relation to iron oxide.

When the best spectral combinations are analyzed, the TM4 and TM5 bands were found to be present in the composites considered to be the most informative. The spectral curves shown in Figure 1 confirm that the greatest differentiations occur precisely in the intervals that correspond to bands 4 and 5. Because the soils of the study area have quite similar concentrations of organic matter (C/N relation from 9 to 12), the spectral differences are more dependent on the quality and quantity of iron oxides present ($Fe_2 O_3$). The average percentages of iron oxides in the soil types are: LR = 34%, TE = 26%, LE = 22%, PV = 10%, LV = 5%, AQ = 2.0%, and

Hi = 0.5%. In a consistent way and as expected, the spectral response decreased from the higher to lower levels of iron oxide. The results were darker tones for LR and lighter tones for AQ.

In the case of Hi, the dominant factor was the level of moisture for the absorption of the infrared radiation. In terms of our purely visual interpretations of the images, we were able to distinguish with confidence the LR, LE, and LV soils. The AQ soils were confused with the more sandy areas of LV soils. PV soils could not be distinguished, possibly because of their low representation in the study area.

These results are comparable to those of other authors who compared soil maps from field work and spectral data: Cipra et alii (1980), Lund et alii (1980), Harrison and Johnson (1982) and Agbu and Nizeyiama (1991). Of note, however, is that this current study is conducted with completely different soils and conditions, being tropical soils in Brazil.

CONCLUSIONS

For the study area, the presented methodology permits us to make the following conclusions:

1. The employed mathematical models of JM distances and ISOMIX were sensitive and effective for the data analyses of these tropical soils.

2. For areas of exposed soils, their spectral character was of extreme usefulness for the improvement of the quality and precision of the final map.

3. The methodology shows that it is not only possible to redefine the limits of the soil units, but also the detection and delimitation of associated soil units.

4. The authors believe that the high degree of discrimination obtained permits the suggestion that the methodology is valid for application in areas with soils that are more similar to each other than are those in the study area.

BIBLIOGRAPHY

AGBU, P.A.; NIZEYAIAMA, E. 1991. Comparisons between spectral mapping units derived from SPOT image texture and field soil map units. Photogrammetric Engineering and Remote Sensing, 57:397-405.

ANDRADE, L.A. 1985 . Critérios de seleção de atributos visando a escolha dos quatro canais mais significativos do Thematic Mapper. In: 12º Congresso Brasileiro de Cartografia, Anais, Brasília - Brazil, Vol 1, pp. 205-226.

BAUNGARDNER, M.F.; KRISTOF, S.J.; JOHANSEN, A.L.; ZACHARY, A.L. 1970. Effets of organic matter on the multispectral properties of soils. Indiana Academy of Sciences, 79:413-422.

BENDAT, J.S.; PIERSOL, A.G. 1986. Randon data analysis and measurement procedures. John Wiley, N.York, 304p.

CIPRA, J.E.; FRANZMEIER, D.P.; BRAVER, M.E. and BOYD, R.K. 1980. Comparison of multispectral measurements from some nonvegetated soils using Landsat digital data and a spectro radiometer. Soil Sci. Soc. Am. J., 44:80-84.

COLEMAN, T.L.; MONTGOMERY, O.L. 1987. Soil moisture, organic matter and iron content effect on the spectral characteristics of selected vertisols and alfisols in Alabama. Photogrammetric Engineering and Remote Sensing, 53:1659-1663.

CURI, P. 1982. Análise Multivariada. Int. Biociencias - UNESP, Botucatu - SP, 64 p.

EVERITT, J.H.; ESCOBAR, D.E.; ALANIZ, M.A. and DAVIS, M. 1989. Using multispectral video imagery for detecting soil surface conditions. Photogrammetric Engineering and Remote Sensing, 55:467-471.

HARALICK, R.M.; SHAMUGAN, K.S. 1974. Combined spectral and spatial processing of Erts imagery data. Remote Sensing of Environment, 3:3-13.

HARRISON, W.D. and JOHNSON, M.E. 1982. Improving mapping unit delineation accuracy using Landsat MSS spectral maps. In: Proc. Western Region Tech. Work Planning Conf. Nat. Coop. Soil Survey, SCS, Washington, D.C., pp. 112-128.

IGC - INSTITUTO GEOGRÁFICO E CARTOGRÁFICO 1982. Mapa Geológico do Estado de São Paulo. São Paulo, SEPLAN.

LUND, L.J.; WEISMILLER, R.A.; KRISTOF, S.J.; KIRSCHNER, F.R. and HARRISON, N.D. 1980. Development of spectral mapas for soil-vegetation mapping in the Big Desert Area, Idaho. In: Proc. Sixth Ann. Symp. on Machine Processing of Remotely Sensed Data. Lab. for Application of Remote Sensing. Purdue Univ.,West Lafayette, Indiana, pp. 84-96.

MONTGOMERY, O.L.; BAUNGARDNER, M.F. 1974. The effects of the physical and chemical properties of soils on the spectral reflectance of soils. LARS Information Note. Pardue University, West Lafayette - USA, 36 p.

MONTGOMERY, O.L.; BAUNGARDNER, M.F.; WEISSMILLER, R.A. 1976. An investigation of the relationship between spectral reflectance and chemical physical and genetic characteristic of soils. LARS Information Note. Pardue University, West Lafayette - USA, 147 p.

OLIVEIRA, J.B.; MENK, J.R.F.; BARBIERI, J.L.; ROTTA, C.L.; TREMOCOLDI, W. 1982. Levantamento pedológico semidetalhado do Estado de São Paulo - Quadrícula de Araras. IAC, Campinas - SP, 180 p.

PAGE, N.R. 1974. Estimation of organic matter in Atlantic coastal plain soils with color-diference meter. Agronomy Journal, 66:652-653.

PARADELLA, W.R. 1984. Avaliação de critérios de seleção de atributos espectrais de imagens digitais MSS-Landsat em discriminação litológica no baixo vale do rio Curuça, BA., São José dos Campos, INPE-SCT. Bol. Tec. 3248-PRE/591. São José dos Campos - SP, 28 p.

PRELAT, A.E. 1981. ISOMOIX. An interactive clustering program. Stanf. Univ., Stanford, Ca, 20 p.

RICHARDS, J.A 1986. Remote Sensing digital image analysis: An introduction. Springer-Verlag, Berlin, 284 p.

STONER, E.R.; BAUNGARDNER, M.F. 1981. Characteristic variation in refletance of surface. J. of Soil Science Soc. of America, 45:1161-1165.

STONER, E.R.; BAUGARDNER, M.F.; BIEHI, L.L.; ROBINSON, B.F. 1980. Atlas of soil reflectance properties. Research Bulletin 962. Pardue University. West Lafayette - USA, 75 p.

SWAIN, P.H.; KING, R.G. 1973. Two effective feature selection criteria for multispectral remote sensing. LARS Information Note. Pardue University, West Lafayette - USA, 24 p.