An Exercise in Soil Survey and Land Evaluation using digital processing and image interpretation

E. Nieuwenhuis
D. Shrestha
International Institute for Aerial Survey and Earth Sciences (ITC)
The Netherlands

VII - 4

Abstract

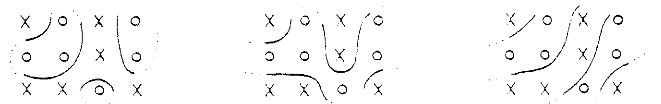
The ITC soil survey procedure is described as applied to an area in Greece. Use is made of Aerial Photography and of Landsat Imagery. Possibilities to estimate the effect of available moisture on the production of wheat and barley are investigated on the basis of digital processing of Landsat data. A methodology is indicated.

Introduction

The ITC Soils Group offers post-graduate courses in soil survey using image interpretation. In addition a programme can be followed which leads to an MSc degree. This last programme comprises research in a subject related to soil survey. This paper refers to a survey carried out in the Larisa area, Greece, as an exercise in the regular training course. Investigations are reported which are included in an MSc programme; they describe aspects of digital processing of Landsat data related to the effect of available moisture on the growth of wheat and barley.
The survey took place in 1983 and covers an area of about 120,000 ha's.

Survey Technique

Dominating in the soil survey methodology of the ITC Soils Group is the "physiographic approach". The starting point for the physiographic soil survey is the study and understanding of the processes that shaped the area. An insight in the soils distribution pattern which is of course related to these processes has to be obtained in the beginning of the survey: it will increase both the efficiency of the survey and the validity of the boundaries. On the basis of this insight it becomes possible to design an efficient layout of suitably shaped sample areas. The key-word in this issue is: "representativeness". This applies both to the location of the sample areas and to the location of individual observation points. As the soil properties are linked to the landscape, by the participation of the soil bodies in the landscape development, the soil mapping can be conceived as a more or less detailed landscape mapping. This leads to the conclusion that points of observation depend on the physiography. Moreover, the same observation points lead to another map when the formation of the area is different. Consequently, this explains why the surveyor should try and build up a concept about the physiography of the area right from the start of his survey, as it influences his choice of observation points as well as his conclusion with regard to the boundaries to be drawn, throughout the survey. The idea may be clarified with a hypothetical example. The following figure (fig.1) shows three survey situations. In each case twelve observations are done; two kinds of soils have been identified in the same distribution pattern with respect to the points of observation. Three different concepts about the formation of the area would lead to three different "maps". A proper "concept" allows for a considerable reduction of observation points.



Limestone area knobs and kettles

Alluvial area levees and basins

Aeolian area ridges and valleys

Fig 1 Soils and physiography

In the sample areas the soil distribution pattern is studied in detail. On the basis of this investigation the mapping units for the final map are established, in other words, the legend is constructed. A specially designed correlation scheme, in which soil classification and physiographic classification are matched, helps in the decision as to how the grouping of the detailed units of the sample area survey ought to be effectuated. In fig.2 the basic layout of the correlation scheme is presented.

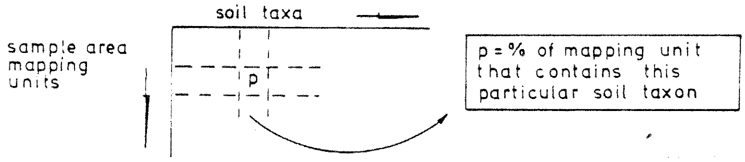


Fig.2 Correlation scheme

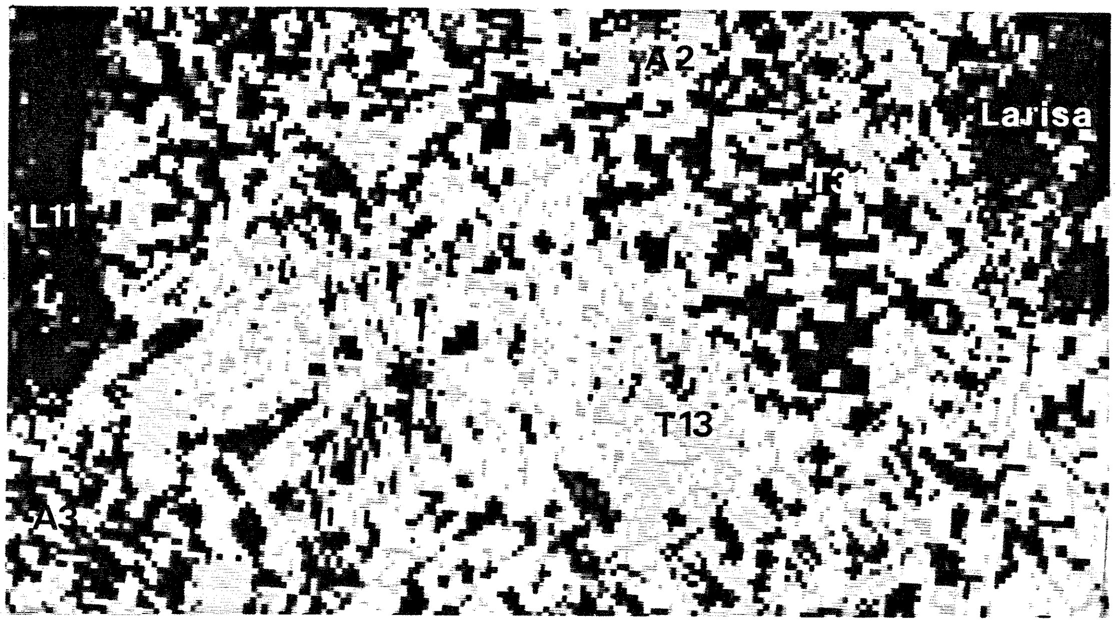
The great advantage of this procedure is the limited time it takes to add valuable information to the final soil map in giving information as to the more detailed soil distribution pattern within the mapping units. The efficiency of the use of image interpretation in selecting representative sites can hardly be overestimated. This applies to both sample areas and observation points. When properly carried out the sample area survey includes a description of the mapping units for their identification on the relevant imagery. The procedure as a whole leads to a considerable reduction in the time and consequently cost of the soil survey, the main effect of using the imagery being that of the relatively easy and fast location of the soil boundaries.

Physiography and Soils

As a preparation for the next part of this paper a reconnaissance map of the area, at the scale 1:100,000 is presented and discussed in general terms. For reference, see the map and the legend on pages 8 and 9. The surveyed area is located in a large basin, surrounded by mountain ridges. In the ridges several gaps occur through which rivers flow. The mountain ridges have been subdivided into two major units: those which consist of mainly metamorphic rock (M) and those which are composed of metamorphic rock and Limestone (L).

Only the ridges in the North and West have been included in the survey. In the East a very large unit is recognised: the actual basin (B). Two other major units occur respectively in the North-West, a major fan (F) and in the westerly central and southern part, a dissected plain, consisting of Tertiary and early Quaternary deposits (T). Entering the area in the West and leaving it in the North-East is the Pineios river. Along its course and that of its tributaries, rather narrow alluvial plains are found (A). At the foot of the mountains a landtype is distinguished consisting of coalescing fans and pediments (P).

The landtypes T and B have wheat and barley as major crops. In both areas a relation was found between the waterholding capacity of the soils and the crop growth. We will limit ourselves to a discussion of the soils of one sub-unit of landtype T, namely "T13". This unit has an undulating to rolling relief and is characterised by a contorted drainage pattern. The location is west-south-west of Larisa: wheat and barley are the only crops. For the soil distribution pattern see the first schematic cross section of the T landtype in the legend on page 9. The Calcixerollic Xerochrepts and the Pachic Calcixerolls occupy respectively the summits and the depressions. The former soils have loamy textures and are excessively drained, the latter are clayey and well-drained. In the plate below, the result of the digital processing exercise is presented; it shows the unit T13 with the distribution pattern of the two soils in light grey and white tones respectively. These tones indicate the growth of the crop in relation to the water availability; see the table on page 7.



The other units indicated in the plate are the following:

T31 : concave valleys and footslopes with many cotton fields (black);
A2 : meander plain with various crops and three different stands of barley and wheat;
A3 : river basin with mainly wheat, barley and cotton;
L11 : higher slopes and summits of the Limestone mountains with a limited growth of grasses (overgrazing).

Digital Image Processing

Landsat imagery has been used in soil survey because of its synoptic view and its multispectral characteristics. The near orthographic view of the imagery allows for its use as a base map.

Much work has been done with respect to the spectral response of various types of soil and the use of Landsat data for soil survey in general. Distinguishing soil features on the basis of spectral characteristics alone is often not feasible because of a disturbing surface cover. The surface features of the soil have then to be inferred from the characteristics of this cover. In the soil survey of the Larisa area variations were observed in the growth of barley and wheat which were related to the soil moisture availability. In the following a computer processing technique is explained; this technique aims at enhancing the imagery with respect to the differences in the vegetative cover which are due to the mentioned soil moisture availability.

A computer compatible tape of the Landsat-4 MSS data of the area was acquired, dated April 11, 1983. Since Landsat-4 lacks the RBV Camera system (like in Landsat 1, 2 and 3), the multispectral bands are referred to as MSS 1, 2, 3 and 4 for green, red and two near infrared portions of the electromagnetic spectrum, respectively. This numbering system is used in the present text.

Leaf area index (LAI) transformation is used as a basis for the processing technique. The transformation is based on two reference samples: pure vegetation and bare soil. For the pure vegetation sample an alfalfa field was selected. This field as well as a field with bare land was located during the fieldsurvey and marked on the Landsat image. The corresponding radiance values in all of the four spectral bands were obtained for alfalfa and bare soil from the haze corrected digital image of the area. These values are given below in vector notations.

$$\underset{\text{(alfalfa field)}}{\vec{V}} = [\,7, 5, 31, 32\,]$$

$$\underset{\text{(bare soil)}}{\vec{S}} = [\,18, 23, 26, 20\,] \qquad \ldots\text{(i)}$$

$$\text{MSS} \;:\; \quad 1 \quad 2 \quad 3 \quad 4$$

In fig. 4 the corresponding bar spectra are presented. Fig.5 shows the location of the alfalfa fields and bare land in the two-dimensional feature space of bands 2 and 4.

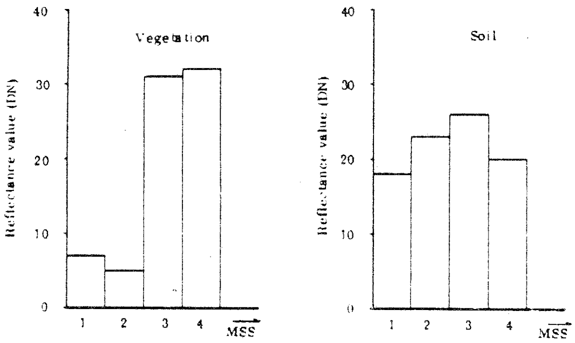


Fig. 4 Spectral response of the reference samples for the different bands

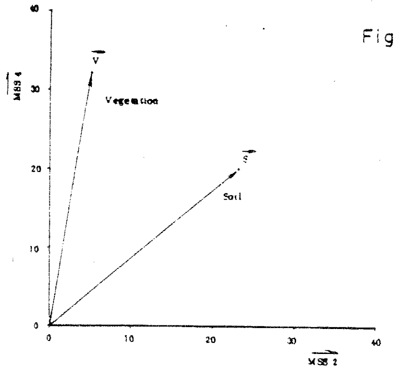


Fig.5

The location of the reference samples in the 2-dimensional feature space of bands 2 and 4

The variation of the illumination in sunlit and shadow areas, due to the topography results in inconsistent radiance values for the same object. In order to remove this effect of varied illumination, the radiance values in each pixel are normalised by the total intensity which is the sum of all the bands for that pixel.

Thus,

$$\text{I (intensity)} = \sum_{i=1}^{n} B_i \qquad \text{where B indicates the MSS band}$$

$$\text{So, the intensity for } \vec{V} \text{ is :} \quad \sum_{i=1}^{4} V_i = 7+5+31+32 = 75$$

$$\text{and for } \vec{S} \text{ the intensity is :} \quad \sum_{i=1}^{4} S_i = 18+23+26+20 = 87 \qquad \ldots\text{(ii)}$$

The reference characteristic vectors of vegetation and soil are thus normalised by the total intensity,

$$\overrightarrow{V'} = \overrightarrow{V} / \sum_{i=1}^{n} Vi \quad \text{and} \quad \overrightarrow{S'} = \overrightarrow{S} / \sum_{i=1}^{n} Si \qquad \ldots \text{(iii)}$$

Substituting the values of (i) and (ii) in (iii), we obtain

$$\overrightarrow{V'} = \frac{1}{75} \begin{bmatrix} 7 \\ 5 \\ 31 \\ 32 \end{bmatrix} = \begin{bmatrix} .0933 \\ .0667 \\ .4133 \\ .4267 \end{bmatrix} \quad \text{and} \quad \overrightarrow{S'} = \frac{1}{87} \begin{bmatrix} 18 \\ 23 \\ 26 \\ 20 \end{bmatrix} = \begin{bmatrix} .2069 \\ .2644 \\ .2988 \\ .2299 \end{bmatrix}$$

While, $$\sum_{i=1}^{4} V'i = 1.0000 \quad \text{and} \quad \sum_{i=1}^{4} S'i = 1.0000 \qquad \ldots \text{(iv)}$$

Normalising by the total intensity has as effect, the uniform illumination throughout the image since $\Sigma B'i = 1$, where B' are the intensity normalised multispectral bands.

Now, the intensity normalised soil vector $(\overrightarrow{S'})$ is subtracted from the corresponding vegetation vector $(\overrightarrow{V'})$ and the resulting vector $\overrightarrow{VS'}$ gives the scale for the LAI and is used as the transformation vector, which serves as the weight function in the data compression from the four multispectral bands into one (LAI).

$$\overrightarrow{VS'} = \overrightarrow{V'} - \overrightarrow{S'} = \begin{bmatrix} -.1136 \\ -.1977 \\ .1145 \\ .1968 \end{bmatrix} \qquad \ldots \text{(v)}$$

The leaf area index transformation is thus performed for each pixel as follows:

$$\overrightarrow{VS'} \cdot \overrightarrow{B'} = [-.1136, -.1977, .1145, .1968] \cdot \begin{bmatrix} B'1 \\ B'2 \\ B'3 \\ B'4 \end{bmatrix} \qquad \ldots \text{(vi)}$$

The numerical LAI scale is generated by two values, namely for the highest and lowest leaf area cover, by calculating the inner product of $\overrightarrow{VS'}$ with $\overrightarrow{V'}$ and with $\overrightarrow{S'}$ respectively, through substitution of the values of (iv) and (v),

$$\overrightarrow{VS'} \cdot \overrightarrow{V'} = 0.1075$$

$$\overrightarrow{VS'} \cdot \overrightarrow{S'} = 0.0039 \qquad \ldots \text{(vii)}$$

The resulting numerical scale of the LAI is generated with seven breakpoints, at equal distances.

In order to fit the byte range of 0-255, scaling and shifting of the origin is necessary to the transformed data from the 4 MSS bands. The final leaf area index greyscale of 0-255 byte range is generated by exponential representation of the numerical LAI scale as shown in the figure 6.

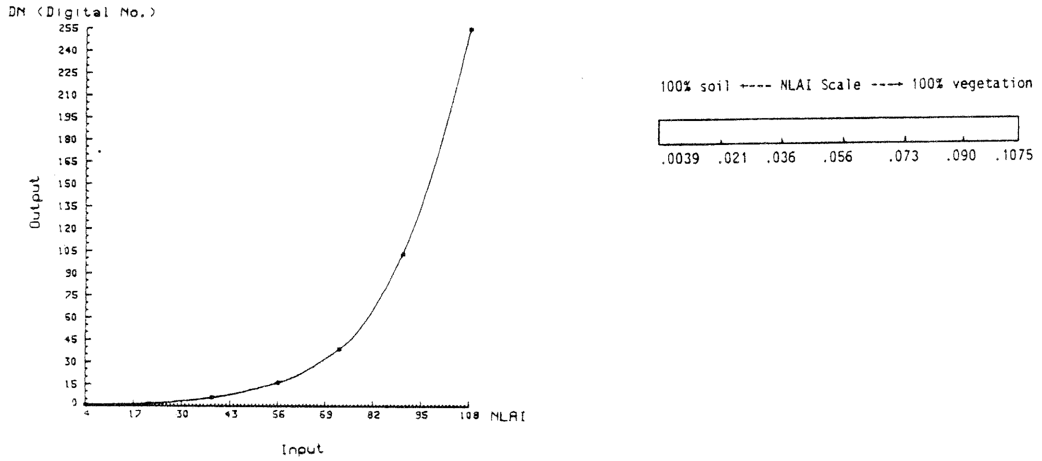


Fig.6 Exponential representation of the Numerical Leaf Area Index (NLAI)

Classification method

The observation points for respectively good and bad growth of wheat and barley were plotted on the feature space of LAI against the total intensity of the reflection. The observation points are scattered throughout the whole landscape. However the main growing areas of wheat and barley are concentrated in the southern part of the survey area, i.e. the Basin and the Tertiary Plain.

The classification is based on the location of clusters of the observation points in the feature space mentioned above. The method applied is that of the "level slicing" on the LAI axis because as expected there is not much variation in the "total intensity" axis. Four levels are used:

DN range on LAI	Class	Presentation on the image, page 3.
0 - 63	bare soil, water, settlements and inclusions of few other crops and vegetation types	black
64 - 74	poor wheat/barley growth	dark-grey
75 - 120	medium wheat/barley growth	light-grey
121 - 255	good wheat/barley growth	white

Results

Large concentrations of fields with a good stand of wheat and barley are found in the Basin and in the Tertiary Plain. Only on the higher slopes and summits, with shallow soils on hard conglomerates of the Tertiary Plain, crop growth varies from medium to poor. On the level valley bottoms the growth is generally good with few exceptions where growth is poor due to salinity, which is related to a high water table. On the Fan Terrace and on the Piedmont Plain generally the crop growth is medium. This is in agreement with the coarse texture of the soils with corresponding reduced water holding capacity.

The classes are not free from a few inclusions; it would require considerable effort to exclude these. Examples are listed below:

1. With the available data it is found difficult to separate barley from wheat due to the similar growing performance and the comparable shape of the plants; this results in similar spectral responses.

2. In some cases medium growth of wheat and barley is confused with overgrazed grasslands; in others poor growth is interpreted while other vegetation is present, affected by salts, due to the high water table. These cases occur locally in a few known areas, therefore they do not offer serious problems.

3. Occasionally alfalfa may be identified as good growth of wheat and barley. Since it is known by the fieldsurvey that the area covered by alfalfa (the reference for 100% vegetation) is insignificant, this problem is negligible.

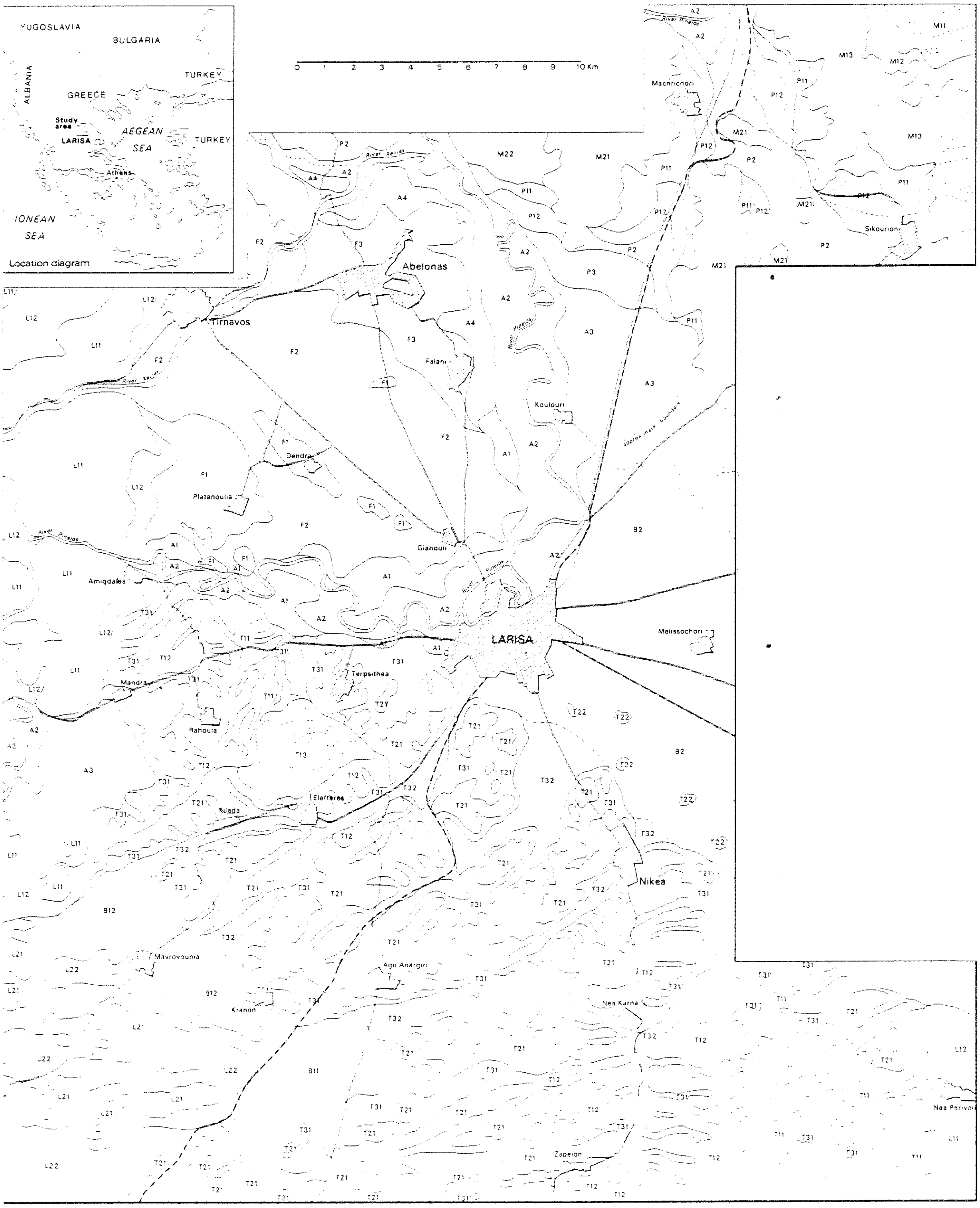
Conclusion

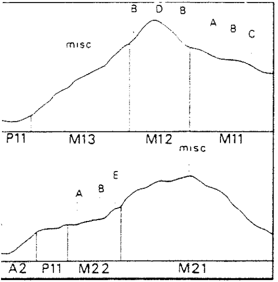
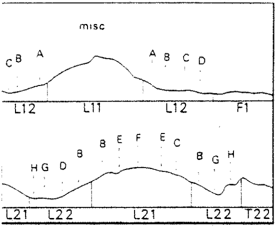
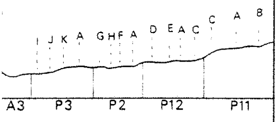
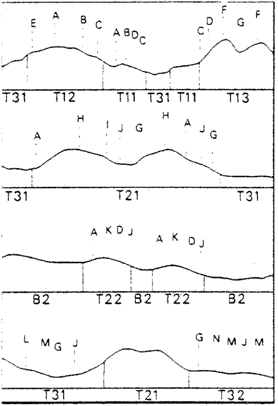
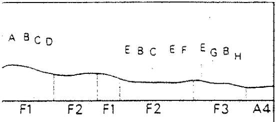
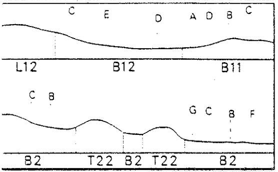
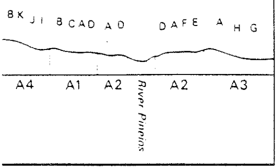
For monitoring crop growth related to soil moisture deficiency, this study has shown good results of the use of the Leaf Area Index transformation of the multispectral data. It appears to be efficient to define a range for the transformation by selecting characteristic vectors (alfalfa and bare soil in this case). This is an improvement over the so-called Hadamard transformation (also performed to obtain the Leaf Area Index) which is the difference between the intensity normalised sum of two near infrared bands and the sum of two visible bands.
However, the classification can still be improved by using multitemporal data, supported by sufficient field checking.

Acknowledgement

The soil survey was carried out by students of the 1983 ITC Soil Survey courses. Special investigations were done by C.v.d.Meeren, D. Shrestha and M. Staljanssens. The whole operation was guided by staffmembers of the ITC Soils Group, Dr G.W. Elbersen, Ir E. Nieuwenhuis, Dr W. Siderius. The digital processing was supervised by Ir N. Mulder of the ITC Image Processing Laboratory. Last but not least, the Greek Soil Survey Institute contributed considerably by making available laboratory facilities and by actively participating in the survey. All these contributions are gratefully acknowledged.

Physiographic soil map, Larisa area (Greece)



LANDTYPE	SUB LANDTYPE	SYMBOL	MAPPING UNITS	SLOPE %	TYPE OF MAPPING UNIT	MAIN AND ASSOCIATED SOILS	%	AREA (ha.)	AREA %	SCHEMATIC CROSS SECTION
MOUNTAINS mainly metamorphic (extension of Olympus massif) M	higher level highly dissected, mainly schist M1	M11	upper "martsia" catchment area	16–30	2) complex	typic xerochrepts lithic xerochrepts typic xerorthents	A 40 B 40 C 20	525	0,5	
		M12	hilly mountain plateau	20–30	complex	typic xerochrepts lithic xerochrepts lithic xerorthents	A 40 B 40 D 20	490	0,5	
		M13	scarp and steep slopes	>40	4) miscellaneous			2940	2,5	
	lower level schist and gneiss M2	M21	complex steep slopes moderately dissected	30–60	miscellaneous			3040	2,5	
		M22	undulating to rolling terrain	8–15	complex	lithic xerochrepts lithic xerorthents typic xerochrepts	B 50 E 30 A 20	515	0,5	
MOUNTAINS metamorphic and limestone L	limestone L1	L11	higher slopes and summits	4–25	miscellaneous			5870	5	
		L12	lower slopes and footslopes	4–20	complex	lithic xerorthents typic xerochrepts dystric lithic xerochrepts calcic haploxeralfs	A 60 B 25 C M* D M	2780	2,5	
	schist / limestone complex L2	L21	higher slopes and summits	3–20	complex	calcixerollic xerochrepts typic xerochrepts lithic xerochrepts dystric lithic xerochrepts	E 40 B 40 F M C M	2145	2	
		L22	footslopes and dissected erosional terraces	2–5	1) association	calcic haploxeralfs vertic haploxeralfs typic calcixerolls typic xerochrepts	D 40 G 40 H M B M	3565	3	
MAJOR PIEDMONT FORMATION P	coalescing fans P1	P11	apices	3–5	3) consociation	typic xerochrepts typic xerorthents typic haploxeralfs	A 80 B M C M	2180	2	
		P12	lower fans	2–4	consociation	typic haploxeralfs typic xerochrepts typic xerofluvents fluventic xerochrepts	C 80 A M D M E M	1805	1,5	
	piedmont plain (incl colluvial aprons) P2	P2		0–4	consociation	typic haploxeralfs calcic haploxeralfs typic palexeralfs typic xerochrepts	F 70 G M H M A M	2345	2	
	terrace (possibly lacustrine) P3	P3		0–2	consociation	entic chromoxererts vertic calcixerolls typic haploxeralfs typic xerochrepts	I 70 J M K M A M	625	0,5	
DISSECTED TERTIARY PLAIN T	undulating to rolling T1	T11	subparallel drainage pattern	0–9	association	calcixerollic xerochrepts petrocalcic xerochrepts typic xerorthents typic calcixerolls	A 50 B 20 C M D M	5830	4,5	
		T12	trellis drainage pattern	0–9	association	calcixerollic xerochrepts typic calcixerolls calcic haploxeralfs typic xerorthents	A 55 D 30 E M C M	6125	5	
		T13	contorted drainage pattern	0–9	complex	calcixerollic xerochrepts pachic calcixerolls typic xerorthents typic calcixerolls	F 60 G 20 C M D M	1030	1	
	gently undulating T2	T21	interfluves	2–6	association	calcixerollic xerochrepts tops / slopes typic calcixerolls vertic calcixerolls pachic calcixerolls	 H/A 40 I 30 J M G M	11310	9,5	
		T22	isolated hummocks	3–6	association	calcixerollic xerochrepts typic calcixerolls vertic calcixerolls petrocalcic palexerolls	A 40 D 40 J M K M	75	<0,5	
	valley system T3	T31	concave valleys and footslopes	0–4	complex	vertic calcixerolls vertic haploxerolls pachic calcixerolls typic chromoxererts	I 40 L 30 G 20 M 20	15335	12,5	
		T32	level valley bottom	0–2	consociation	typic chromoxererts vertic calcixerolls cumulic haploxerolls pachic calcixerolls	M 80 J M N M G M	3905	3	
FAN TERRACE F	higher level F1	F1		0–2	consociation	typic haploxeralfs calcixerollic xerochrepts typic xerochrepts palexerollic chromoxererts	A 85 B M C M D M	1885	1,5	
	mid level F2	F2		0–2	consociation	typic haploxeralfs calcixerollic xerochrepts typic xerochrepts fluventic xerochrepts	E 75 B M C M F M	8590	7	
	lower level F3	F3		0–3	association	calcixerollic xerochrepts typic haploxerolls typic haploxeralfs typic xerofluvents	B 50 G 20 E 20 H M	2975	2,5	
BASINS B	"mavrovounia" basin B1	B11	outer fringe	0–1	consociation	typic chromoxererts vertic calcixerolls typic calcixerolls palexerollic chromoxererts	A 75 B M C M D M	2090	1,5	
		B12	centre part	0–1	consociation	typic pelloxererts palexerollic chromoxererts typic calcixerolls	E 75 D 20 C M	1885	1,5	
	"karla" basin B2	B2		0–2	consociation	vertic calcixerolls typic calcixerolls typic chromoxererts calcixerollic xerochrepts	B 70 C 25 F M G M	17595	14,5	
MAJOR RIVER VALLEYS A	terraces A1	A1		0–4	consociation	fluventic xerochrepts typic xerochrepts calcic haploxeralfs typic xerofluvents	A 70 B M C M D M	1465	1	
	meander plain A2	A2		0–2	association	fluventic xerochrepts typic xerofluvents typic haploxeralfs typic xerorthents	A 50 D 40 E M F M	6865	5,5	
	river basins A3	A3		0–1	consociation	entic chromoxererts fluventic xerochrepts cumulic haploxerolls	G 70 A 20 H M	3780	3	
	saline fringe A4	A4		0–5	association	fluventic xerochrepts typic halaquepts typic xerochrepts typic xeropsamments	I 45 J 30 B 20 K M	1105	1	
*M = minor soils (< 15%)						TOTAL		120640	100	

1) Association = The soil association is a group of defined and named taxonomic soil units, regularly associated geographically in a defined proportional pattern
2) Complex = The soil complex consists of two or more taxonomic units which are so intimately associated geographically that they cannot be separated by boundaries at the scale used
3) Consociation = A mapping unit of which at least 70% of the area consists of pedons which belong to the same taxonomic category
4) Miscellaneous = Areas that have little or no natural soil or that are too nearly inaccessible for orderly examination

GLOSSARY

Data Compression/Reduction	Transformation of observed values into useful, ordered or simplified information.
Density/Level Slicing	The process of converting the continuous grey tone of an image into a series of density intervals of slices each corresponding to a specific digital range.
DN (Digital Number)	The value of reflection recorded for each pixel on Landsat CCT.
DN Range	The range of digital numbers in which the light intensity is registrated from maximum to minimum or vice verse.
Feature Space	The plot of the pixels in function of their digital numbers in different bands.
Haze Correction	One kind of radiometric correction where the atmospheric effects on the data are removed.
LAI (Leaf Area Index)	Also referred to as Biomas; plan area of leaf to a given land area.
RBV (Return Beam Vidicon)	A camera system which operates by shuttering 3 independent cameras simultaneously, each sensing a different spectral band in the range of visible to near infrared portion of the electromagnetic spectrum. Used in Landsat 1, 2 and 3.
Pixel (Picture Element)	A unit whose first member is a resolution cell and whose second member is the grey shade assigned by a digital count to that resolution cell.

SELECTED REFERENCES

Kristof S.J. & Zachary A.L. 1974 - Mapping Soil Features from Multispectral Scanner Data, Photogrammetric Engineering, 1974 pp.1427-1434

Markham B.L. & Barker J.L.- 1983 Spectral Characterisation of the Landsat-4 MSS Sensors, Photogrammetric Engineering and Remote Sensing, vol.49, No 6, June 1983 pp. 811-833.

Meeren C.v.d. 1984 - Soil Moisture Characteristics for the Growing of Wheat. MSc.thesis, ITC Enschede, May 1984.

Mulder N.J., 1980 - A view of Digital Image Processing, 14th Congress of the International Society of Photogrammetric, Hamburg 1980.

Shrestha D.,1984 - Digital Processing of Landsat MSS data for Wheat and Barley Growth-Soil Type Relationship Study in Larisa,Greece, MSc.thesis, ITC Enschede May 1984.

Staljanssens M., 1984 - Drainage Pattern Analysis of the Tertiary Plain of Larisa Area, paper in process.

Westin F.C. & Frazee C.J., 1976 - Landsat Data, its use in a Soil Survey Programme, Soil Science Society of America Journal, Vol.40 1976.