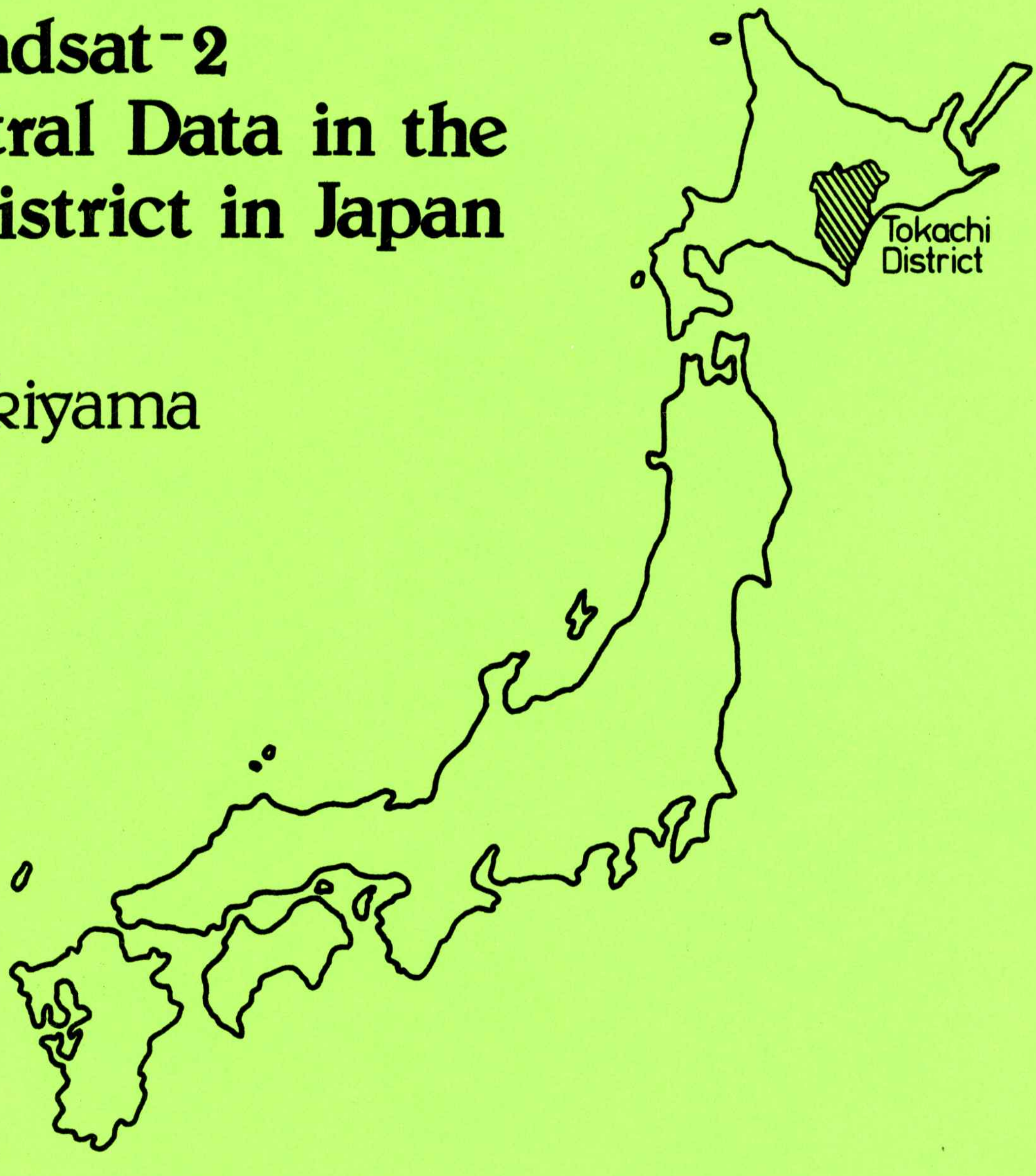
LARS Contract Report 072379

Analysis of Ground Cover Types and Grassland Condition Using Landsat-2 Multispectral Data in the Tokachi District in Japan



Tsuyoshi Akiyama

Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana 47906 USA
1979

Analysis of Ground Cover Types and Grassland Condition Using Landsat-2 Multispectral Data in the Tokachi District in Japan

Tsuyoshi Akiyama

Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana, U.S.A.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

ABSTRACT

Landsat multispectral scanner data acquired on June 11 in 1975, over the Tokachi district in Japan was analyzed using the LARSYS software to evaluate the utility of satellite data for land cover classification of large areas. Thirty-three land cover classes, including 5 grassland, 4 forest, 4 soil/vegetation and 4 urban classes were identified to exhibit spectrally separable characteristics.

Comparison of the final classification to available maps and photography proved the classification accuracy to be acceptable. Difficulties were encountered in correctly classifying classes such as GRASS1 - CLOUD4, OAT/RYE - FOREST or CLOUD2 - SNOW. Those difficulties could currently be resolved by identifying and using cloud free data, or by using Landsat data from two or more dates. Such problems should also be diminished with the availability of the spectral bands planned for Thematic Mapper.

I. INTRODUCTION

During the past decade, remarkable progress has been made in remote sensing technology, especially in the development of computer-implemented pattern recognition techniques for analysis of multispectral scanner (MSS) data.

In the field of agriculture, it has been shown that accurate identification of major crop species and the acreage estimates can be provided from the MSS data (Bauer et al., 1973 and 1978). Meanwhile, in the Corn Blight Watch Experiment conducted in 1971 over the U.S. Corn Belt region, it was shown that remote sensing can quantitatively recognize crop conditions which may be an indicator of yield (MacDonald et al., 1972.)

These and other quantitative evaluations of computer-processed Landsat data show particular promise for obtaining more accurate less costly and more timely crop information.

In Japan, after the 200 Mile International Treaty was concluded, beef became more important as a protein source instead of fish. Because of the demand for its stable supply, grassland developments were carried out in many districts in Hokkaido. However, conservation of the environment is still strongly demanded.

To meet these demands, there is a need for monitoring the grassland condition in order to manage it properly and maintain high productivity for a long time period.

For those objectives, Landsat MSS data can provide a means of obtaining this necessary information with a minimum of time and expenditure.

There are two approaches to analyze the remotely sensed data: image oriented and numerically oriented. In the present study, a numerically oriented system using pattern recognition (Swain, 1973), which is newer and in which rapid progress is being made, was employed. Numerical analysis of remote sensing data has merits such as the ability to analyze or extract information from the large quantities of data potentially available, and the potential to improve the identification or classification performance as compared to manual methods (Bauer, 1972).

The Tokachi district was selected for classification of the Landsat data using the Laboratory for Applications of Remote Sensing (LARS) system (Phillips, 1973).

II. MATERIALS AND METHODS

A. Description of the study area

The Tokachi district is a representative upland crop area within which soybeans, corn, potatoes, sugar beets and several types of pasture grass are cultivated on the largest scale in Japan.

It is located approximately in the center of the Hokkaido island and is situated between 42°30'N and 43°30'N in latitude as shown in Figure 1. This is a semi-arctic zone and the average temperature during the year and the growing season (May to October) are 5.9°C and 14.7°C, respectively. Annual precipitation in this area is 927 mm, two-thirds of which occurs during the growing season. The average amount of solar radiation during growing season is 381 cal/cm^2/day.

A computer compatible tape with multispectral digital data (Scene ID-2140-00344) acquired by Landsat-2 on June 11, 1975, was used for the present analysis. This frame covers the central part of the Hokkaido island. The multispectral scanner mounted on Landsat-2 collects data in two visible (0.5 to 0.6 μm and 0.6 to 0.7μm) and two infrared (0.7 to 0.8μm and 0.8 to 1.1μm) wavelength bands.

Analysis was conducted through the LARSYS system on approximately a quarter of the frame, covering most of the agriculture area of the Tokachi district.

In the middle of June, in this area, seedlings of upland crops other than pasture grasses possess less than 10% ground cover, and therefore appear as bare soil in the MSS data. However, the main purpose of the study was not primarily to identify the various cover types present, but rather to study evaluation of grassland condition from Landsat MSS data.

Reference data were composed of:

a) Topographic maps covering the Tokachi district at a scale of 1:50,000.

b) Aerial photographs taken on June 5, 1976 covering the Tokachi Livestock Station, and the corresponding pasture grass inventory table.

c) Land classification map produced earlier at a scale of 1:200,000.

d) A soil classification map at a large scale (ca. 1:1,000,000).

B. Preprocessing

The Landsat computer compatible tape containing digital data (Scene ID-2140-00344) was reformatted to make it compatible with the LARSYS system. In another step, geometric correction was applied (Anuta, 1973) to,

a) Rotate the frame to true north.

b) Eliminate distortion (skew) due to earth rotation during data collection.

c) Rescale the data for line printer output at a scale of 1:25,000.

d) Eliminate altitude and attitude variations.

Figure 2 shows a gray scale image of the reformatted and geometrically corrected Landsat data covering the study area. The eight insets are the candidate training areas selected for clustering.

C. General processing and analysis

The basic steps taken to perform the present analysis is summarized in the flow chart shown in Figure 3.

The analysis sequence may be divided into 5 steps as follows:

a) Select candidate training areas using gray scale maps and other reference data.

b) Calculate statistics for non-supervised classes obtained by clustering each training area.

c) Represent the clusters of each training area with spectral response curves and then merge the statistics of the clusters from each training area, calculate the separability of the clusters and make pooling decisions to obtain representative and separable training classes.

d) After obtaining satisfactory results in test classification, perform classification of the entire study area using a maximum likelihood point by point classification method.

e) Evaluate the classification results and extract information.

(1) Selection of candidate training areas

In order to become familiar with the distribution of ground features within the study area, gray tone images were generated by the processor *GDATA in the LARSYSDV environment. The image, printed by a Varian printer/plotter represents 16 gray tone levels for each channel separately in a pattern of dots at a scale of 1:125,000 (Figure 2). Color composites of the data (Figure 4) generated by the *IMAGEDISPLAY function were also useful to know the ground features.

The Tokachi district study area contains various features such as snow-capped mountains, ocean, rolling valleys covered with forests, a broad central plane for agriculture, large pasture lands and dense urban areas.

A total of eight candidate training areas listed in Table 1 were selected to train the computer to recognize specific classes of interest. They contain almost all land cover types in this district and each area consists of 5,000 to 10,000 pixels.

(2) Clustering and identification of cluster classes

The LARSYS processor *CLUSTER was employed for identifying and grouping spectrally similar pixels. Twelve to fifteen such groups or cluster classes were requested from each training area. A map showing the cluster to which each pixel was assigned was printed at a scale of 1:25,000 for each of the eight training areas (Figure 5). The histograms of each cluster in each channel were printed and the statistics which were used in later steps were also punched by this processor. The total of 101 cluster classes were identified with the help of reference data and the spectral response curve of each cluster.

(3) Merging statistics, separability studies, and pooling of spectral classes

An algorithm *MERGESTATISTICS in the LARSYSXP environment was employed to combine the statistics from the eight training areas into one set. Each cluster was represented on a bi-spectral plot where the average of the mean of the visible bands (channels 1 and 2) was plotted against the average of the mean of the infrared bands (channels 3 and 4).

In addition, transformed divergence was calculated by the *SEPARABILITY processor using the means and covariance matrix of each class contained in the punched statistics. Transformed divergence is a measure of the probability of successfully distinguishing between pairs of classes. Thus, the transformed divergence values for all possible class pairs were printed. Those which had transformed divergence values greater than 1500 were recognized as sufficiently separable classes. Those class pairs with a transformed divergence value less than 1500 were identified on the bispectral plot. Groups of clusters were then selected so as to maximize separability from one another and maintain representativeness.

Since no classes had been obtained for cloud, snow or shadow from the eight training areas, additional areas including those classes were selected. In some cases, statistics were punched directly from *CLUSTER as before. In other cases, the addresses of groups of pixels belonging to the same cluster were punched. The *STATISTICS processor was then employed to calculate the statistics of those classes.

The processes of combining and deleting of classes, and calculating the separability of the pools selected were repeated several times before an acceptable set of classes was obtained. Through the above mentioned processes, the total number of classes was reduced from 101 to 33. The bispectral plot of the final 33 classes is shown in Figure 6.

(4) Classification and the evaluation

As a further evaluation of the training classes selected several of the training areas were classified using the *MINIMUMDISTANCE processor in the LARSYSDV environment. This processor performs a minimum distance to means pixel by pixel classification. The classification results were stored on tape and then input into the *PRINTRESULTS processor for production of the classification map. The classification and display of 5 test areas representing almost all ground cover types was done in this way. Because of the lack of detailed reference data, the percent of correct classification could not be estimated. However, the classification maps were found to be in acceptable agreement with the reference data. Therefore, the entire study area consisting of about 1.3 million pixels was classified with the *CLASSIFYPOINTS (maximum likelihood, pixel by pixel) processor.

Three types of classification maps were generated by the *PRINTRESULTS, *GRESULTS and *PHOTO processors.

Using *PRINTRESULTS a line printer map of the entire study area was generated in which each of the 33 classes was represented by a different symbol (Table 2, Figures 12 and 13). Also, using *PRINTRESULTS, another line printer output was generated at the same scale in which 20 groups of classes were shown (Figure 14). Using *GRESULTS another classification image (Figure 15) having 10 groups (Table 2) was produced at a scale of 1:125,000 on a Varian printer/plotter. Using *PHOTO, a color coded map (Figure 16) with 10 groups represented by different colors covering the northern half of the study area was also produced.

III. RESULTS AND DISCUSSION

(1) Vegetation

Grassland: Five different classes for grassland were identified in the analysis. Class GRASS1 represents the most productive area at this season, mostly Timothy grass, which sprouts the earliest in spring among pasture grasses in this district. GRASS2 is a grassland containing Orchardgrass to some extent and also seems to be in good grassland condition. The above mentioned two classes show typical absorption by chlorophyll and high reflectance in infrared bands (Figure 7).

Meanwhile, most of the grazing grasslands were identified as GRASS3 and 4. However, we could not differentiate a kind of seashore vegetation consisting of homogeneous and native Gramineous herbage from GRASS3. Multi-temporal data should be able to aid for the further analysis in this problem.

Overgrazed or low productive grassland were identified as GRASS4. GRASS3 and 4 showed lower activity in photosynthesis compared with the GRASS1 and 2 on the basis of the response values in the 0.6-7μm band (channel 2, Figure 7).

Class GRA/FOR often appears in the marginal bushed area of grassland and it contained both components of grassland and forest. It also seems to include native grassland such as covered with Sasa which is an important natural resource for livestock in Japan.

Forest areas: Two out of four forest classes possessing lower spectral response were defined as coniferous forest (CONFOR1 and 2), and the remaining two were defined as mixed forest and deciduous forest, FOREST1 and 2, respectively. All of these four forest classes showed deep depression at channel 2 (0.6 to 0.7μm) and lower reflectance response in all four channels when compared with the above mentioned grassland classes (Figure 7).

The Tokachi district is surrounded by mountains and it was determined from the results that the northern boundary high mountains are mainly covered with coniferous trees, while in the western and southern boundaries the lower mountains are dominated by deciduous or mixed forests.

Agricultural fields: Four classes, SOI/VEG1 through SOI/VEG4, consisted of several soil types covered with several kinds of upland or paddy field crops. However, the soil cover by upland crops still remains less than 10%, hence almost all these fields look like bare soil at this time of the year (June 11) in this district. So each of the crop varieties could not be defined by spectral difference.

As for paddy field crops, which were already flooded and in which rice seedlings had just been transplanted, it was not possible to distinguish them in any wavelength band available.

In fact some of the deepest paddy fields were misclassified as class RIVER. However, as can be seen in Figure 12, we could detect some of the paddy field areas by noting the occurrence of SOI/VEG, SHADOW and RIVER together. Fukuhara et al., (1977) tried to classify several soil types in the same study area, according to the organic matter content. In this study, the soil characteristics of the four SOI/VEG classes are yet to be determined by field checking.

Besides the above mentioned four SOI/VEG classes, an OAT/RYE class was defined as a wintering upland crop. However, oat and rye pixels were sometimes classified as FOREST1, because these wintering crops were just before the harvest and were similar in spectral response with FOREST1.

Bauer et al., (1978) clarified that the middle infrared band (2.08-2.35μm) is the single most important band in obtaining crop information. This band is planned for installation in Landsat D.

Multi-temporal data would also be helpful to ease some of the above mentioned problems.

(2) Urban areas

Four classes were identified as urban. Each is spectrally different in some way (Figure 9).

URBAN1 appears in the most urbanized commercial area such as the Obihiro city area shown in Figure 12. From URBAN1 to URBAN4, the proportion of coverage by vegetation increased. Therefore, URBAN4 is supposed to be residential areas containing much vegetation.

The class SSVEG (seashore vegetation) was sometimes found to be mixed with URBAN4, so this class was grouped into class URB/VEG in Figures 14, 15 and 16.

(3) Others

Water: Classes SEA1, SEA2, SEA/SWA (sea and swamp), RIVER and LAKE were identified as aqueous areas. Classes RIVER and LAKE appeared mainly in inland area and possessed lower reflectance values than those of aalt waters (SEA1 and 2) in Figures 6 and 10.

In the display of the classification of the entire study area a 0.5% threshold* percentage was adopted. 5445 out of 1.3 million points were thresholded in this manner. Most of the pixels thresholded belonged to inland water.

*The threshold percentage defines a cut off for the probability that a pixel belongs to the class it was classified as. Pixels with probability values below the cut off are "thresholded" and identified with a blank on the classification map.

Shadow: The two classes identified as shadow have the lowest spectral response in the visual bands among the finally selected 33 earth cover features (Figure 10).

The shadows in the study area were mainly generated by clouds and mountain ridges. It was difficult to distinguish spectrally which shadows were due to clouds and which ones were due to ridges.

Clouds: Classes CLOUD1-4 showed various kinds of spectral response as shown in Figure 11. In some cases they affected seriously the identification of the ground cover.

In cases when the clouds were very thin (CLOUD4), the clouds and the shadows they cast do not mask the spectral response of the vegetation underneath perfectly. The result is that new artificial spectral responses are generated. For example, clouds were sometimes confused with snow-(CLOUD2), sand-(CLOUD3) or grasses-(CLOUD4) as can be seen. This phenomena is discussed in other reports (Virasteh, 1977).

Because of cloud cover, analysis of some part of the study area had to be abandoned.

Other cover types: Class SNOW was found in mountain areas more than 1,000 meters in altitude. Class SAND was identified on the seashore, while BRISOIL (bright soil) for a restricted small area in the northern part of the study area. These three classes have similar spectral response, even though the separability values were more than 1,500, they tended to be confused with one another.

To reduce these confusions, the area could be stratified and each ecosystem analyzed separately. Alternatively, topographic and/or other data could be registered with the MSS data to aid in the analysis.

IV. CONCLUSIONS

Computer-aided analysis techniques using Landsat MSS data showed the potential for obtaining more accurate and timely information about ground cover types and pasture condition.

As the results of classification, five grasslands, four forests and four soil/vegetation classes were identified successfully as vegetative cover types.

Because of lack of reference information, the accuracy should be examined by field checking, however the general patterns of the classification were found to be acceptable. Some difficulties occurred when thin layer of clouds spread over the study area. Several confusions such as GRASS1-CLOUD4, SOI/VEG1-4-paddy fields, GRASS3-a kind of seashore vegetation, OAT/RYE-FOREST1, CLOUD2-SNOW could not be eliminated in this analysis.

For further analysis, especially for agricultural purposes, multi-temporal data or the middle infrared bands proposed in Landsat D may be able to aid in resolving these confusions.

ACKNOWLEDGEMENT

I am thankful to the Division of Space Development, the Science and Technology Agency, Government of Japan, for granting the fellowship for the training in remote sensing technology at the Laboratory for Applications of Remote Sensing (LARS) at Purdue University, West Lafayette, Indiana.

I also thank Mr. Ronald K. Boyd, Dr. Luis Bartolucci and other senior LARS staff members for providing meticulous guidance at every step of the study. The valuable assistance and arrangement rendered by Mr. Douglas B. Morrison, during this study is also acknowledged with thanks.

I would like to express my sincere thanks to Dr. T. Okubo of Nagoya University and Mr. A. Maruoka, Dr. M. Siyomi, Mr. S. Takahashi of National Grassland Research Institute for giving the opportunity to do this study.

I also sincerely thank Mr. Y. Shimada, Mr. M. Fukuhara and Mr. G. Saito of Hokkaido National Agriculture Experiment Station for preparing useful reference data.

V. REFERENCES

1. Bauer, M.E. and J.E. Cipra. 1973. Identification of agricultural crops by computer processing of ERTS MSS data. LARS Information Note 030173. LARS, Purdue University, W. Lafayette, Indiana.

2. Bauer, M.E., M.M. Hixson, B.J. Davis, and J.B. Etheridge. 1978. Area estimation of crop by digital analysis of Landsat data. Photogrametric Engineering and Remote Sensing. 44:1033-1043.

3. MacDonald, R.B., M.E. Bauer, R.D. Allen, J.W. Clifton, J.D. Erickson and D.A. Landgrebe. 1972. Results of the 1971 Corn Blight Watch Experiment. Proc. Eighth Int'l Symp. on Remote Sensing of Environment, pp. 157-190, Ann Arbor, Michigan, Oct. 2-6, 1972.

4. Phillips, T.L. (ed). 1973. LARSYS User's Manual, Vol 1, 2 and 3. LARS, Purdue University, W. Lafayette, Indiana.

5. Anuta, P.E. 1973. Geometric correction of ERTS-1 digital multispectral scanner data. LARS Information Note 103073. LARS, Purdue University, W. Lafayette, Indiana.

6. Fukuhara, M., N. Hayashi, J. Iisaka and Y. Yasuda. 1977. Crop inventory study by remote sensing. (In Japanese). Jap. Soc. Photogramet. Engin. Extra Issue for 1977, pp. 23-26.

7. Bauer, M.E. and M.M. Hixson. 1978. Analysis of agronomic and physical data for physical understanding. In Final Report, Vol. 1. Agricultural scene understanding, pp. 1-41. LARS Contract Report 112578. LARS, Purdue University, W. Lafayette, Indiana.

8. Virasteh, Ali. 1977. Using LANDSAT-2 for crop area estimation in Northern Iran. LARS, Purdue University, W. Lafayette, Indiana.

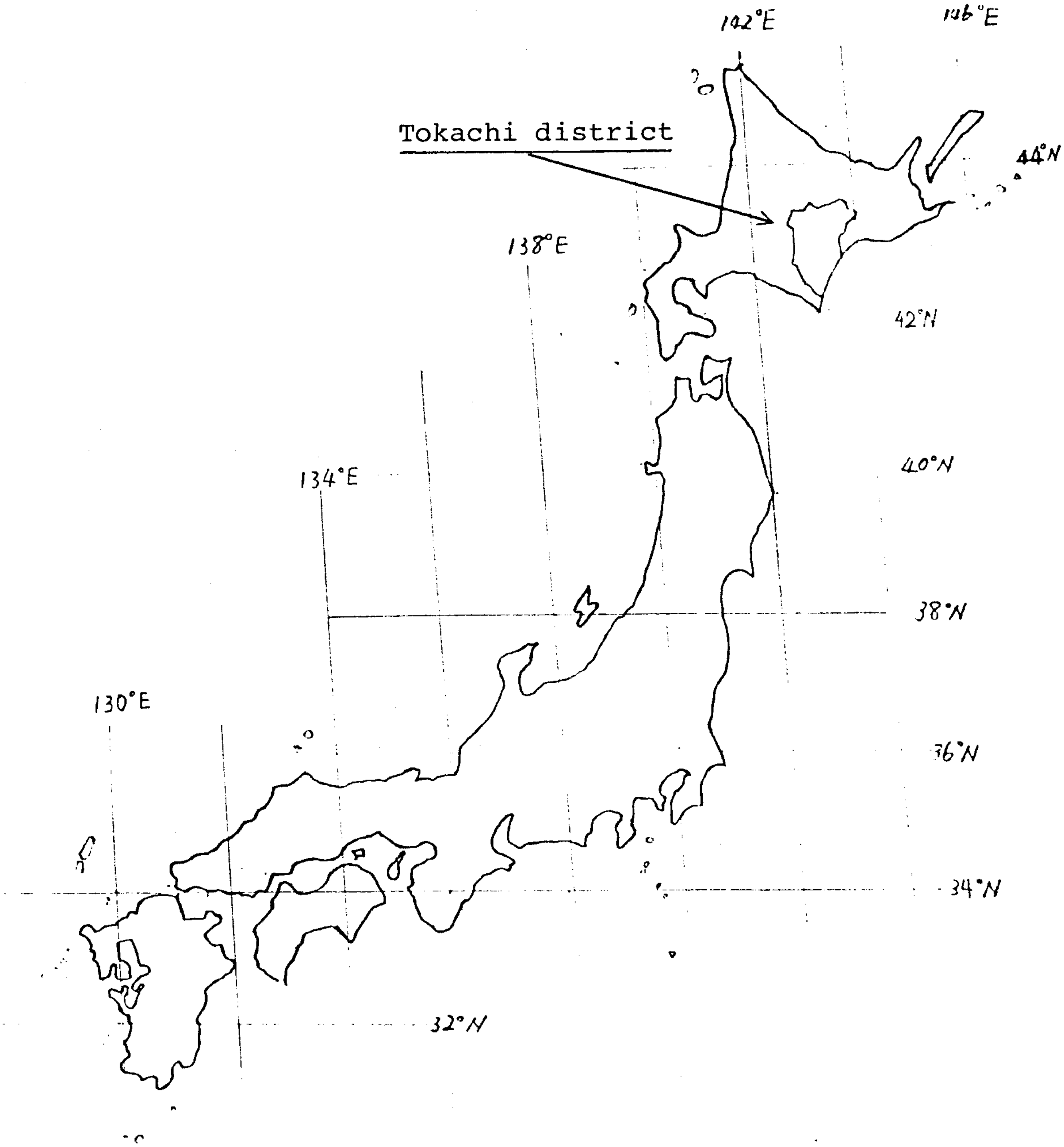


Figure 1. Location of the Tokachi district, Japan.

HALF-TONE PATTERN 'H5W4GRAY' WILL BE USED FOR THIS PLOT - THE GRAY SCALE LEVELS FOR THIS PATTERN ARE

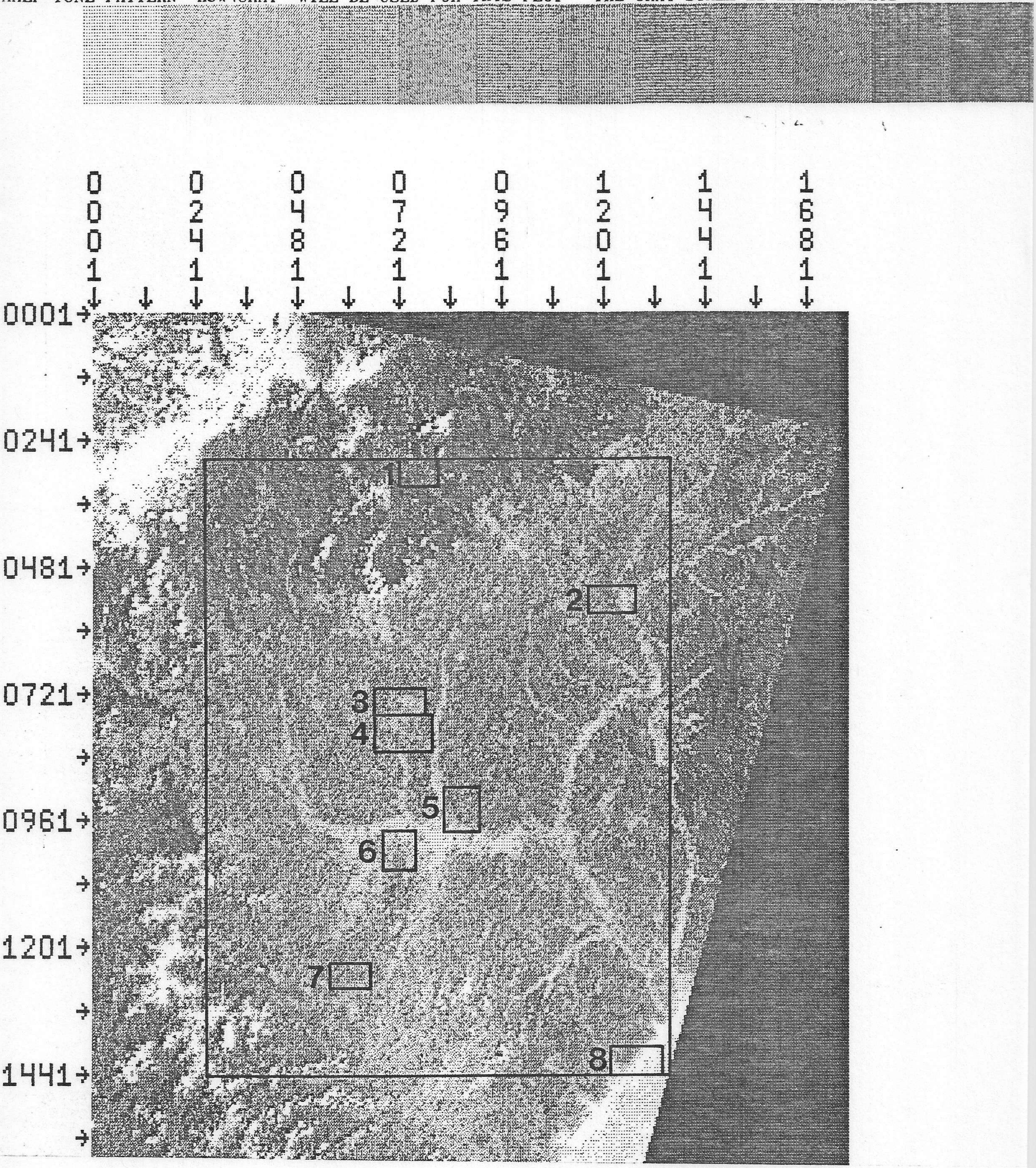


Figure 2. Geometrically corrected Landsat MSS data covering the study area. The insets are the study area and training areas selected for clustering.

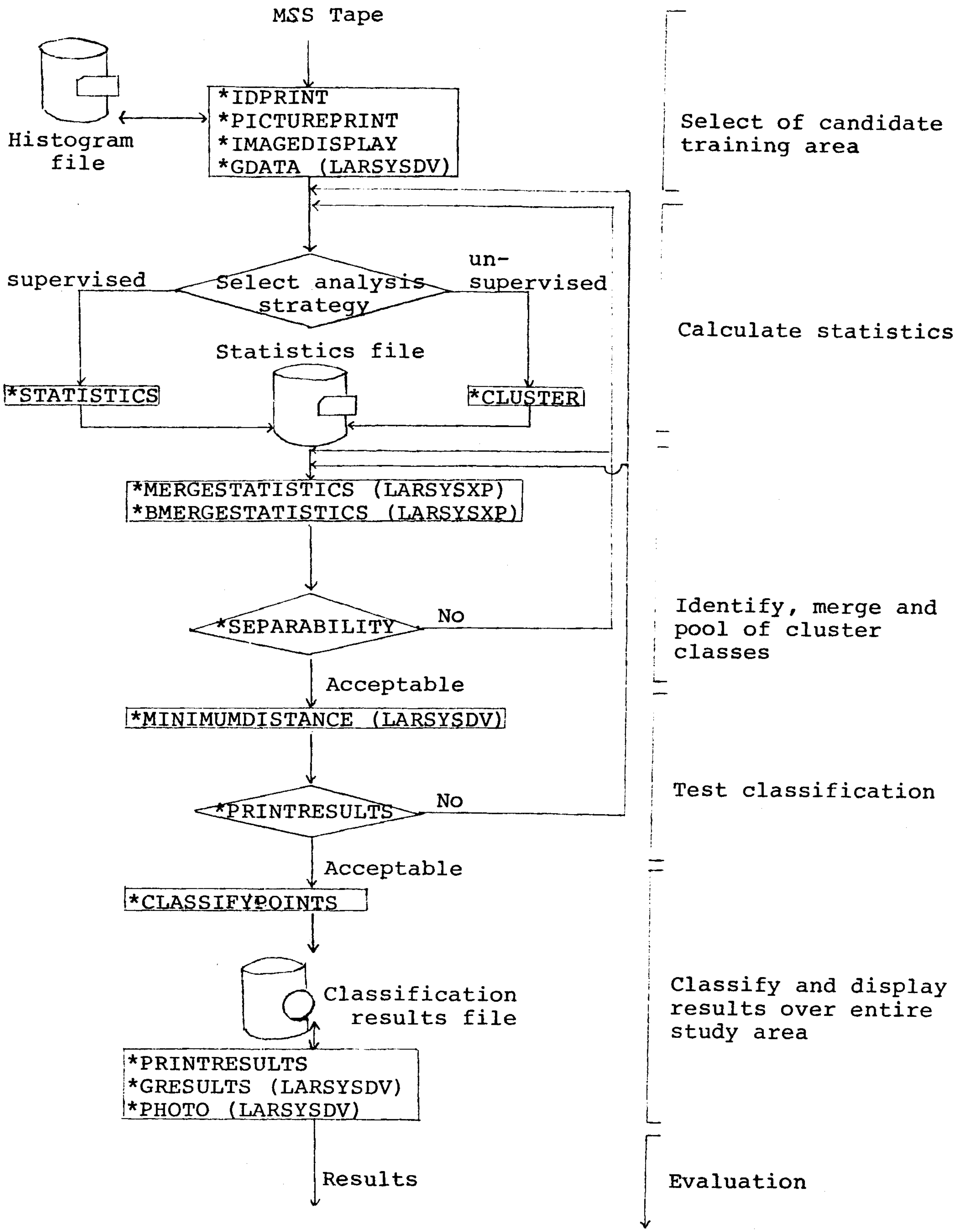


Figure 3. Flow chart of the basic steps taken to perform the analysis.

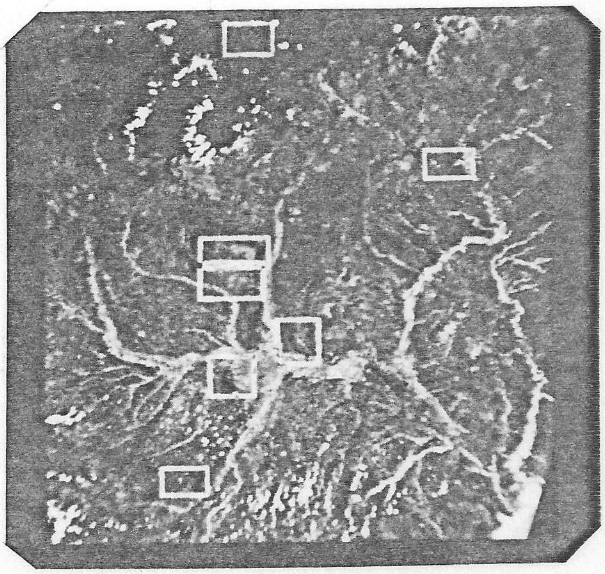


Figure 4. Color image of the study area produced on a digital display from Landsat MSS bands 4, 5 and 6. The insets show the training areas.

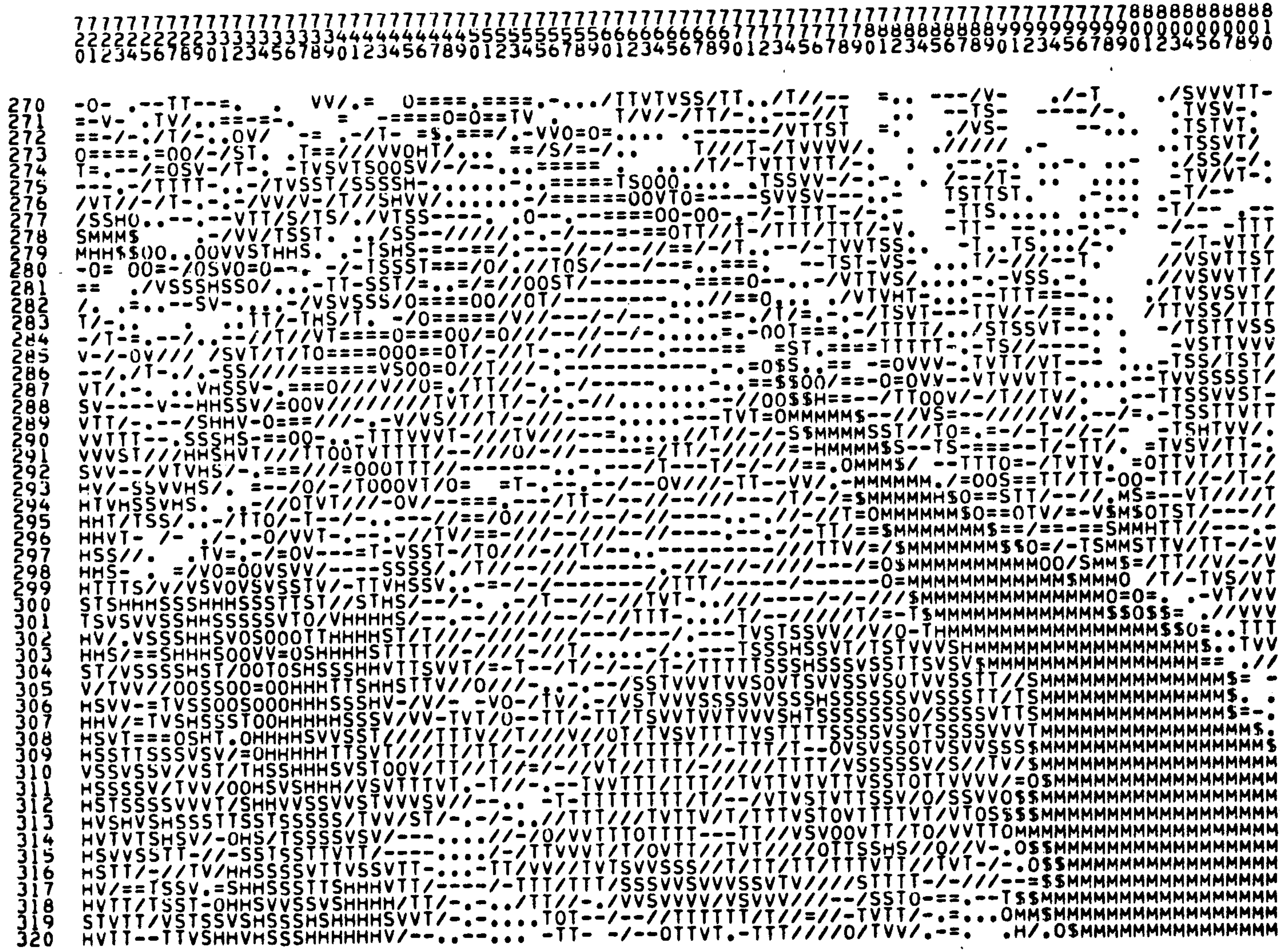


Figure 5. Cluster map of one of the training areas. (Area 1, The Lake Nukabira area) The symbols used to represent each cluster class are blank, ., -, -, =, /, O, T, V, S, H, $, M.

COINCIDENT BI-SPECTRAL PLOT (MEAN) FOR CLASS(ES)

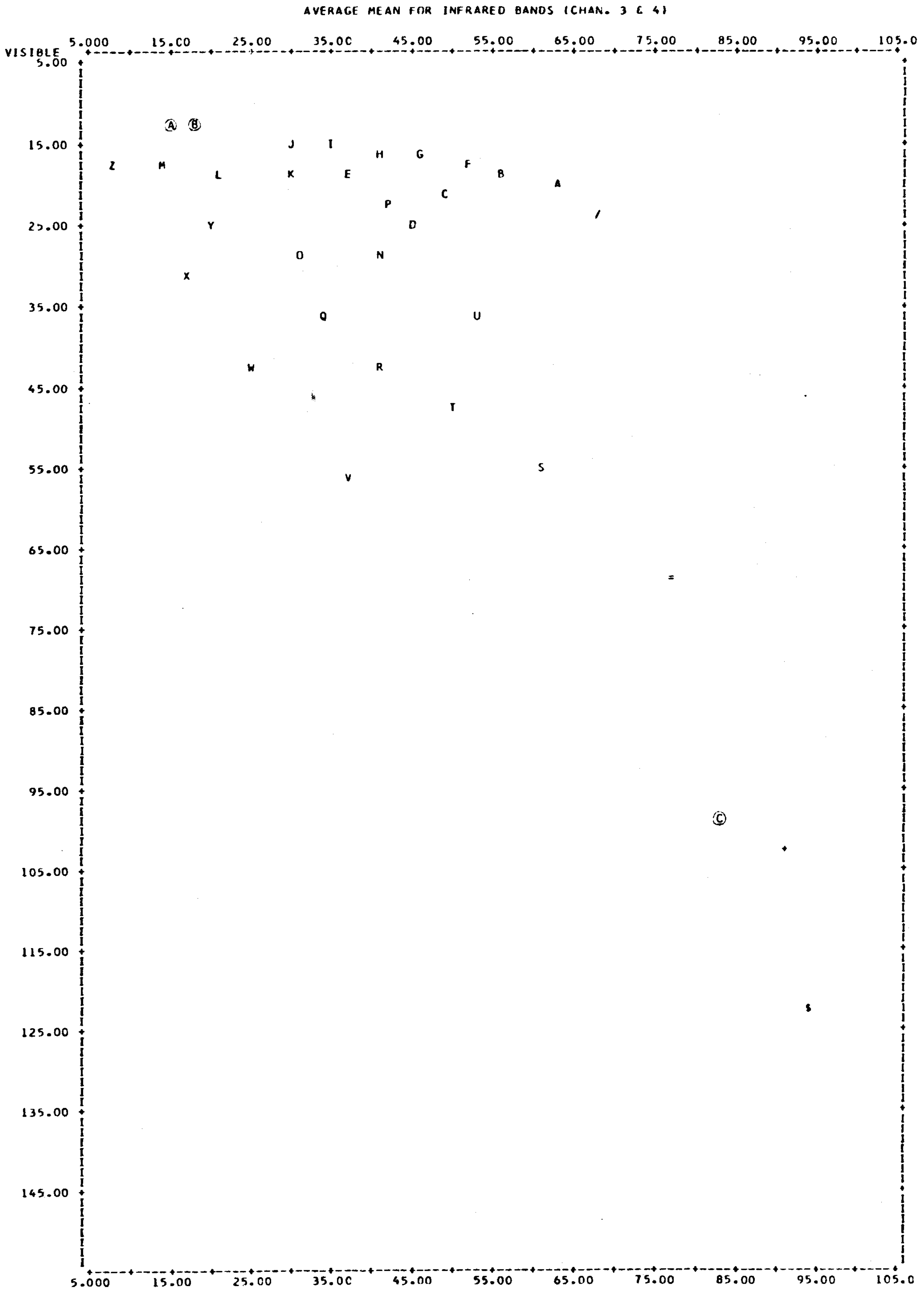


Figure 6. Coincident bispectral plot of the finally selected 33 classes. Refer Table 2 for symbols.

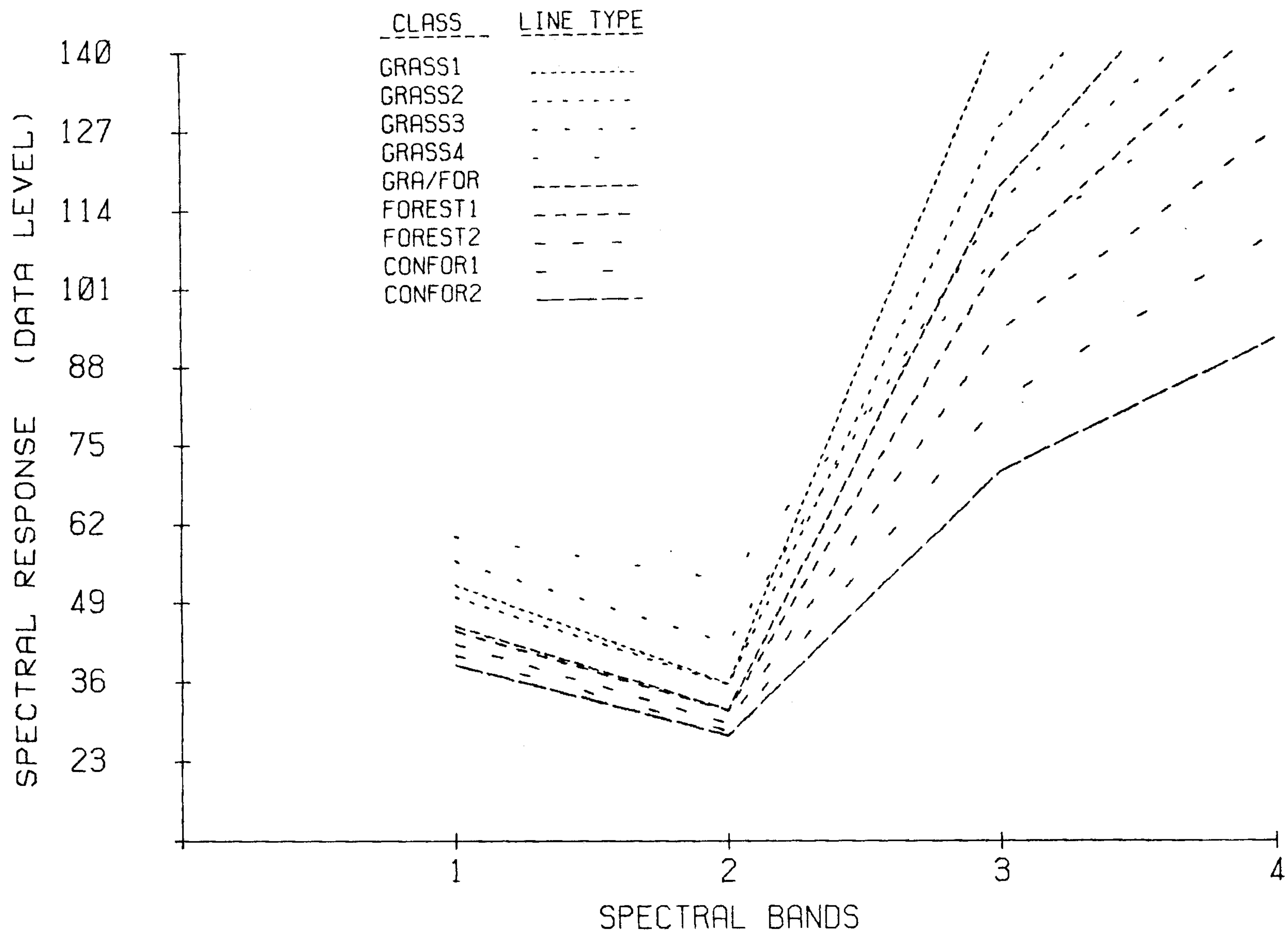


Figure 7. Calibrated spectral response curves for classes GRASS and FOREST. Refer Table 2 for symbols.

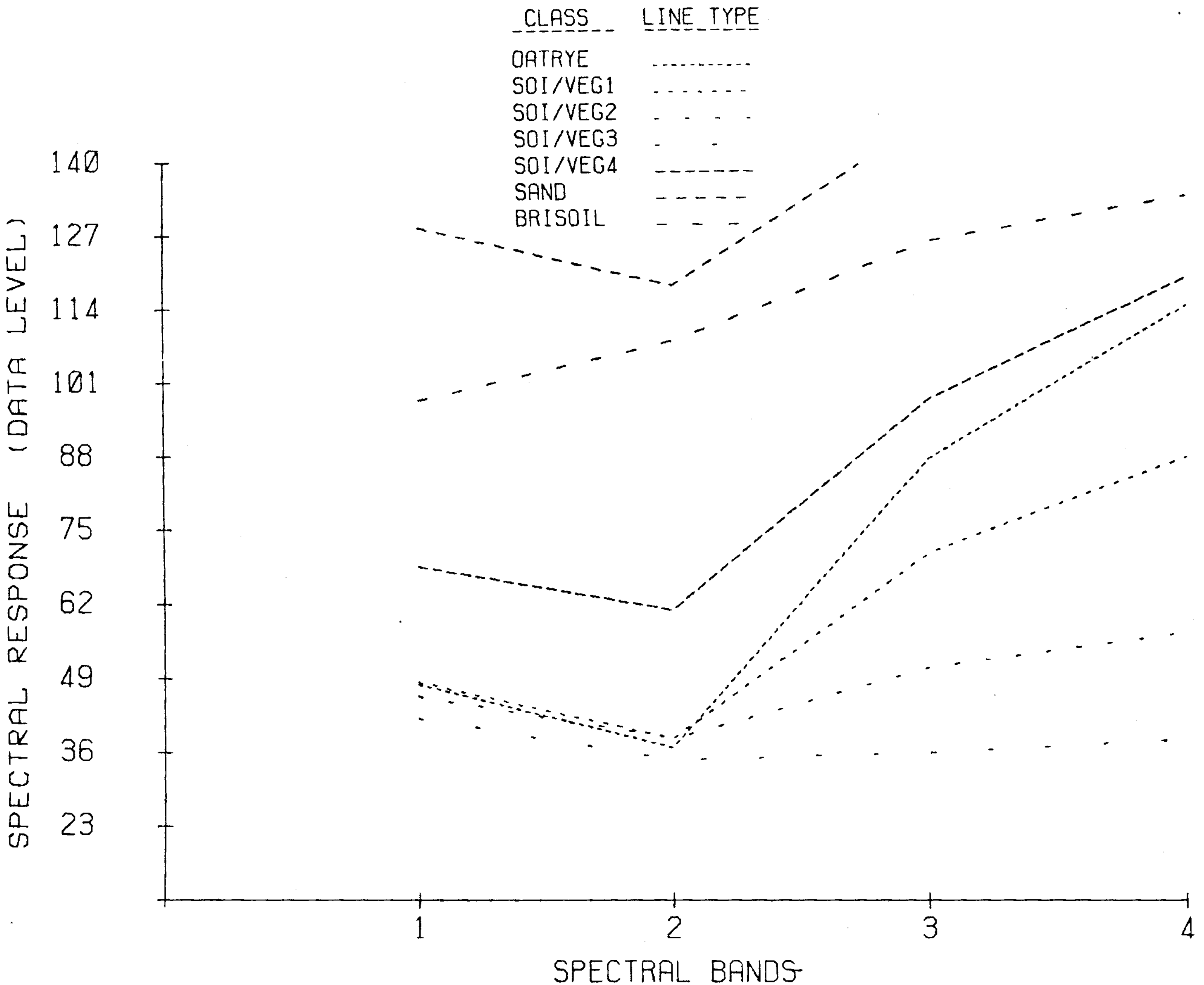


Figure 8. Calibrated spectral response curves for classes SOI/VEG, OAT/RYE, SAND and BRISOIL. Refer Table 2 for symbols.

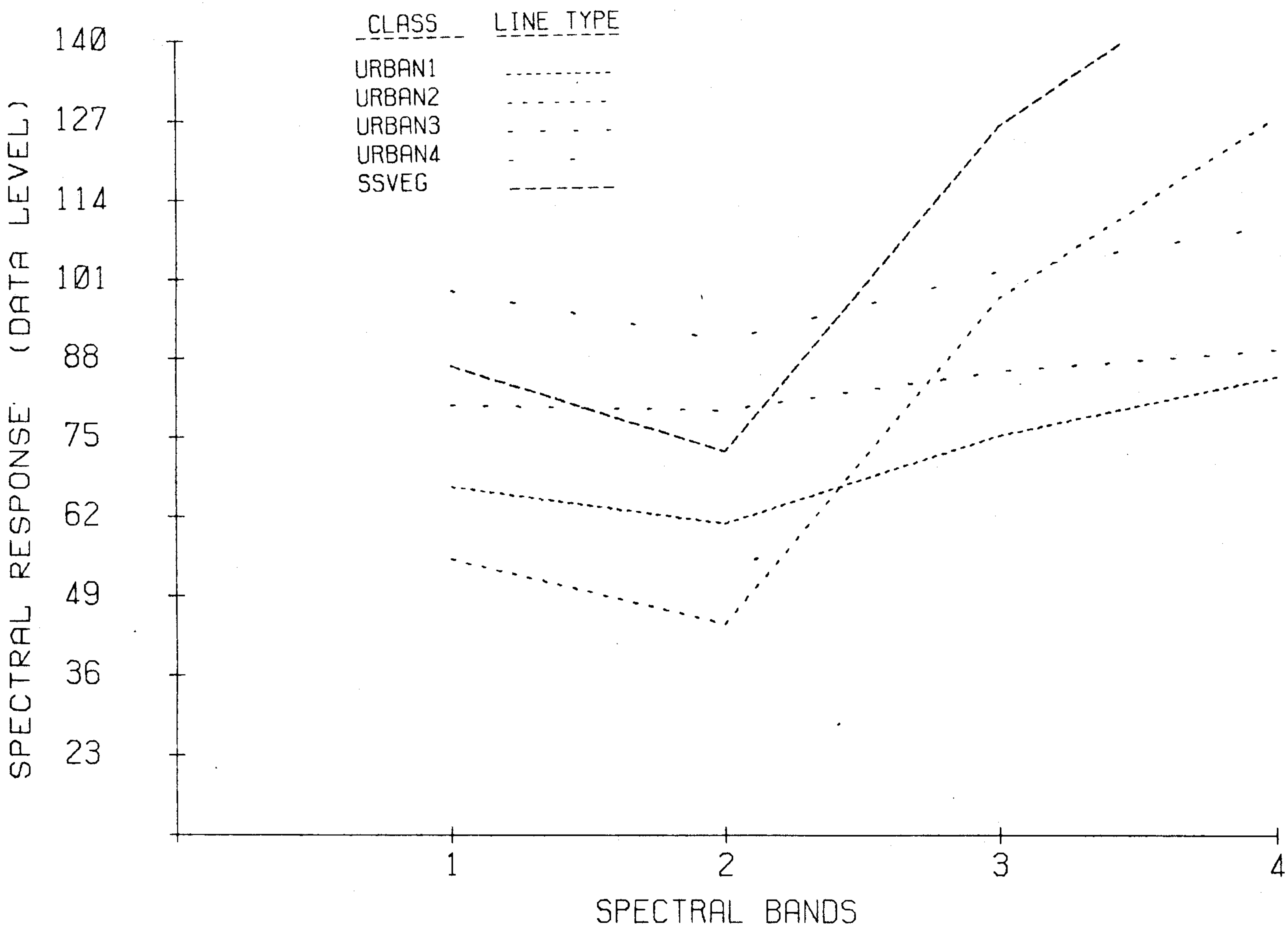


Figure 9. Calibrated spectral response curves for classes URBAN and SSVEG. Refer Table 2 for symbols.

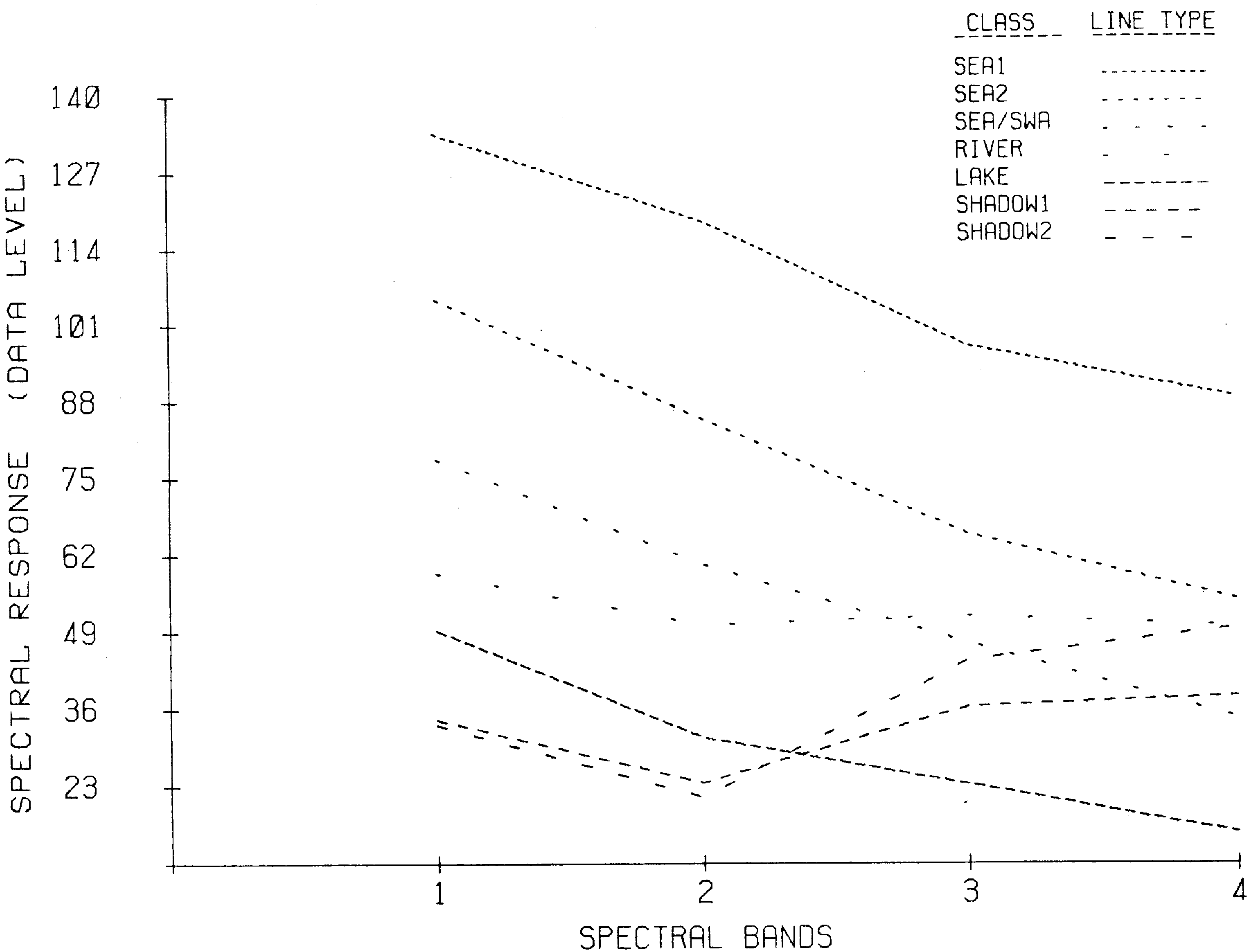


Figure 10. Calibrated spectral response curves for classes SEA, SEA/SWA, RIVER, LAKE and SHADOW. Refer Table 2 for symbols.

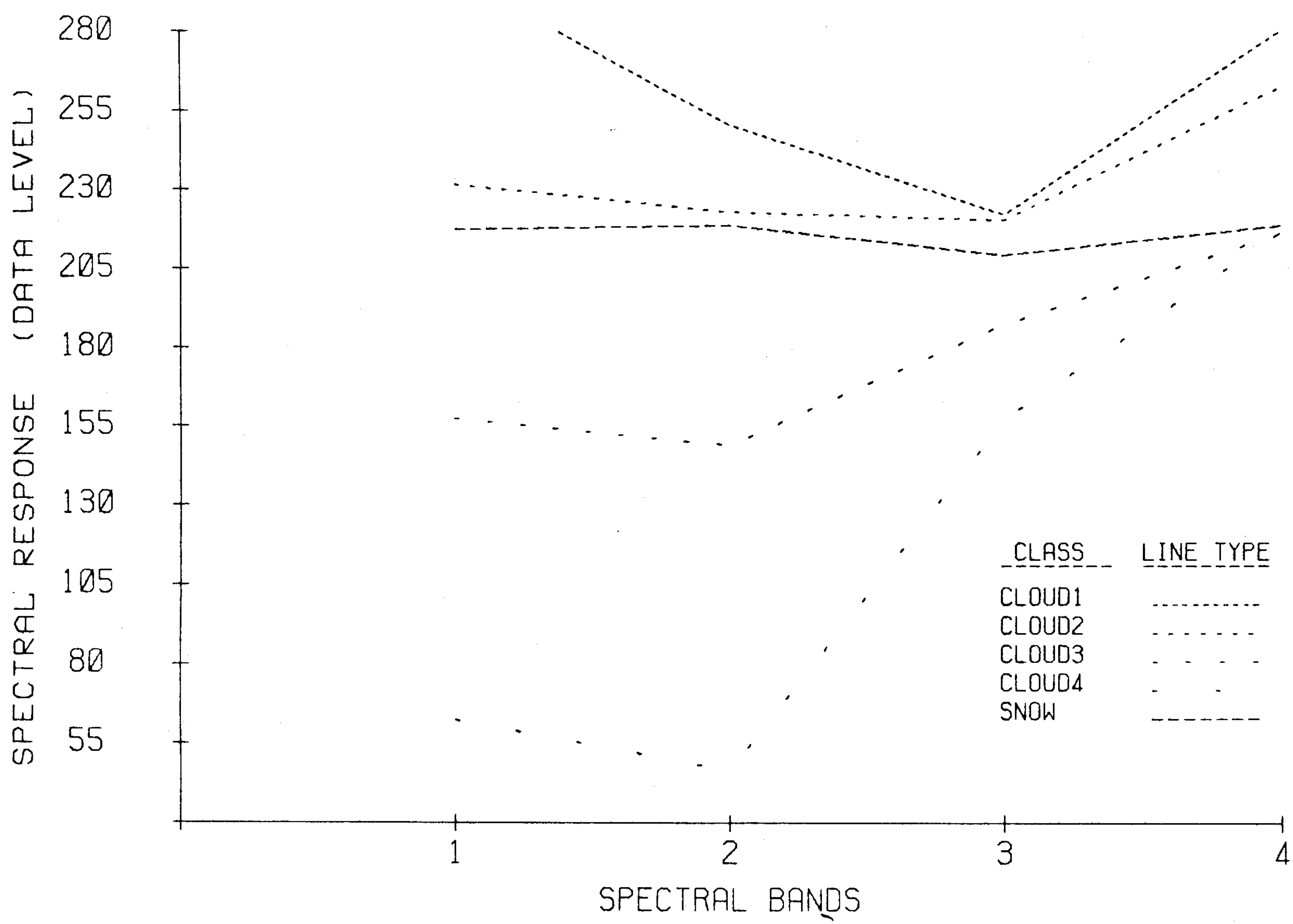


Figure 11. Calibrated spectral response curves for classes CLOUD and SNOW. Refer Table 2 for symbols.

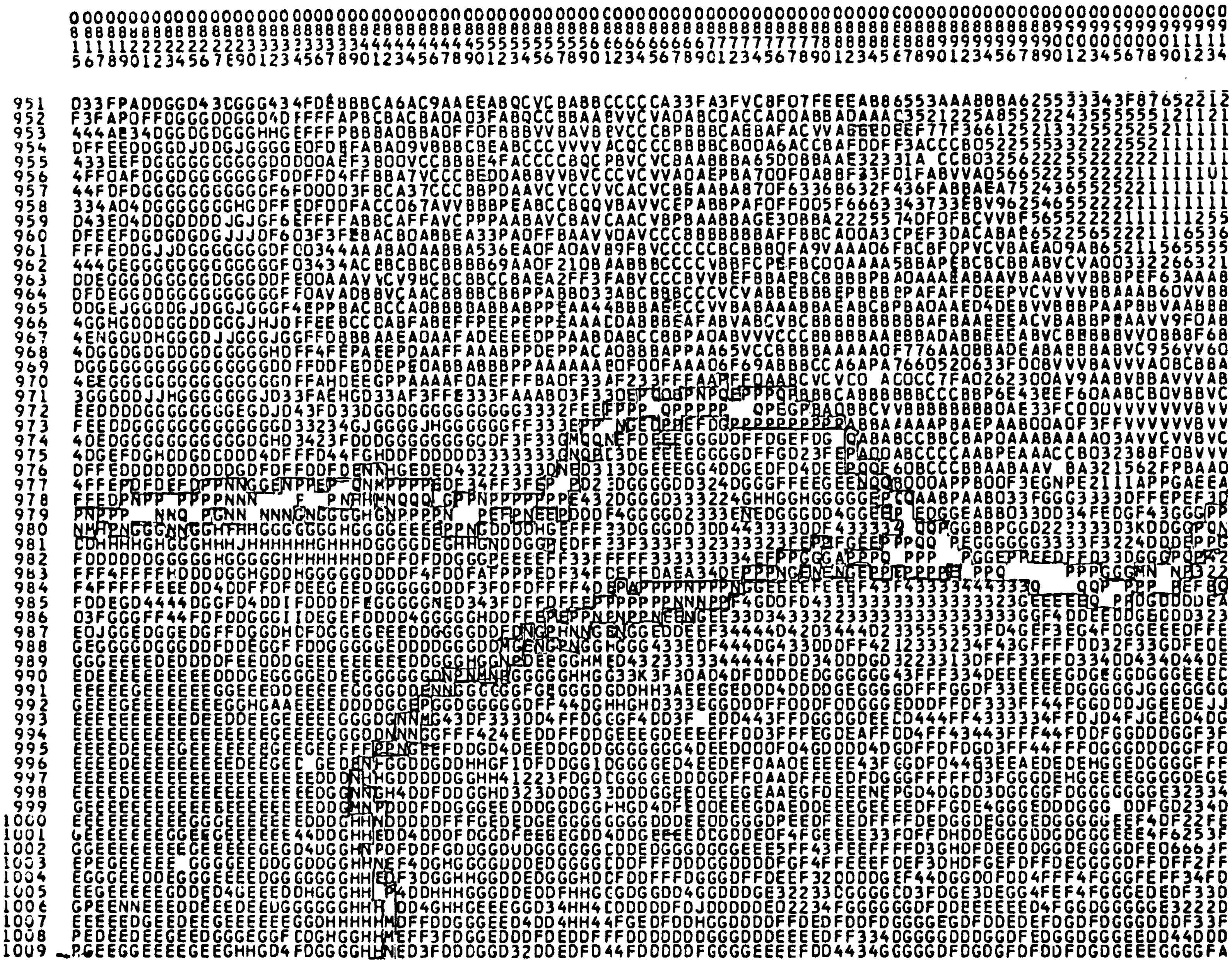


Figure 12. A part of an urban area showing the classification results on line printer output with 33 different symbols. Refer Table 2 for symbols.

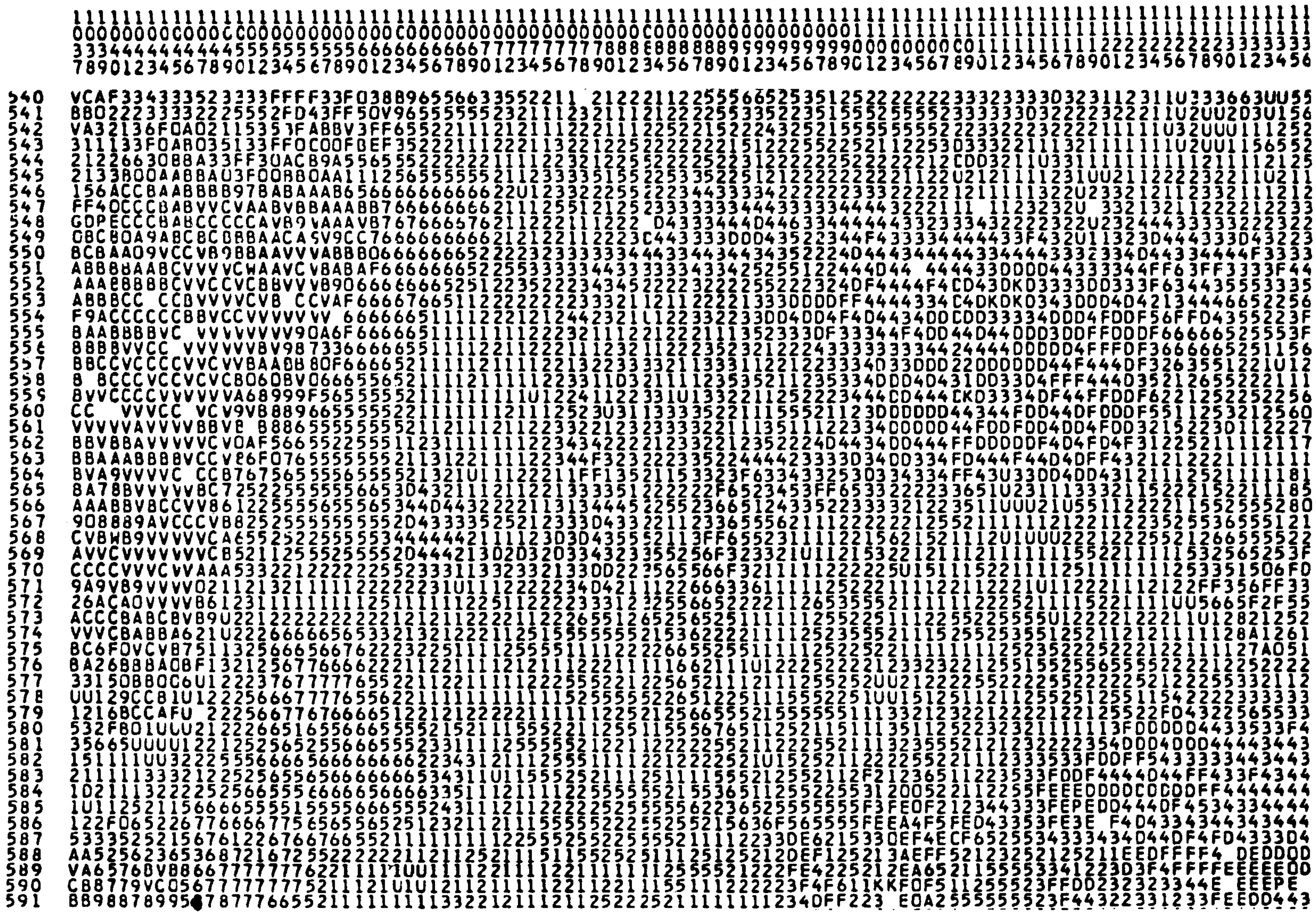


Figure 13. A part of a rural area showing the classification results on line printer output with 33 different symbols. Refer Table 2 for symbols.

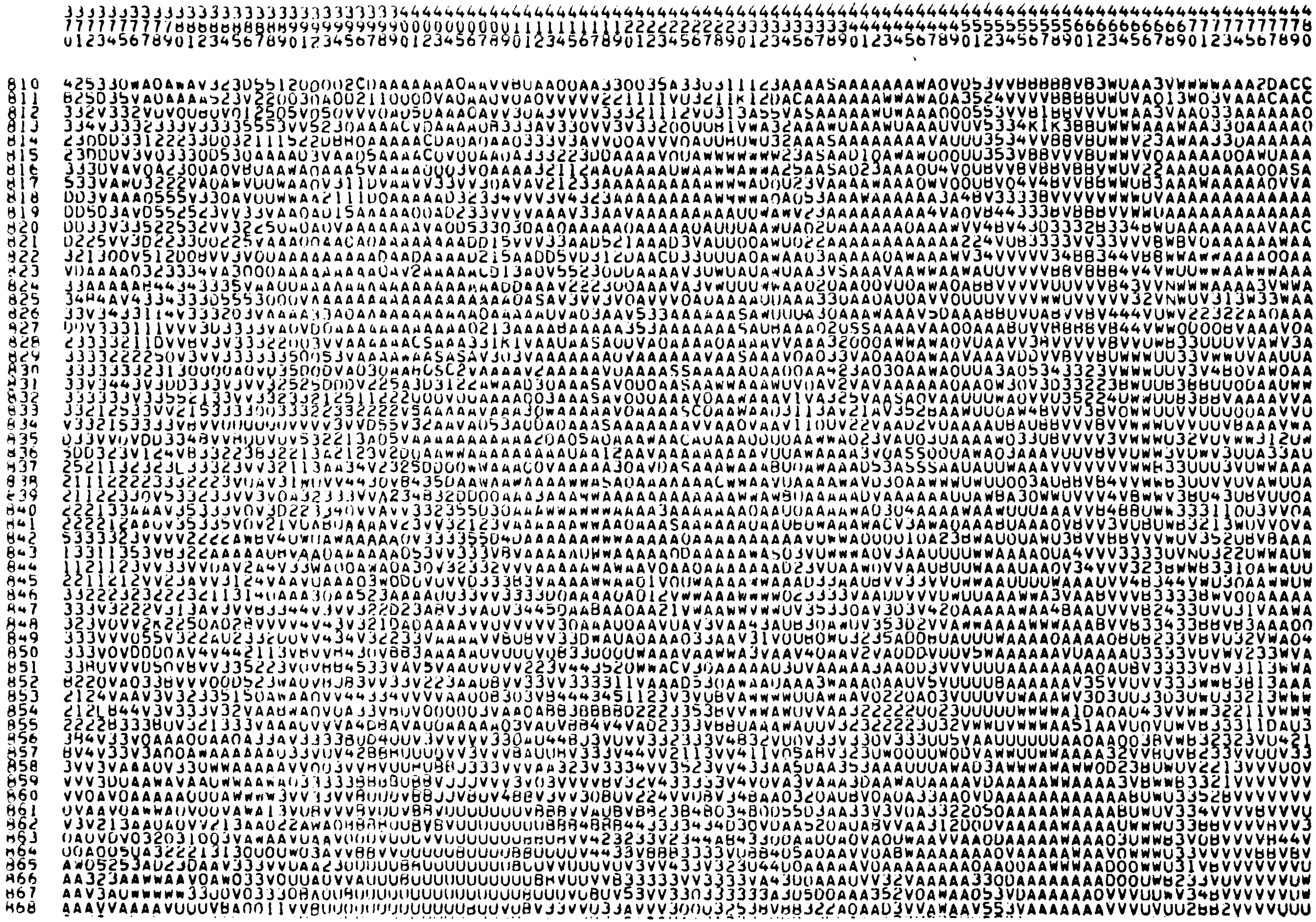


Figure 14. A part of the study area showing the classification results on line printer output with 20 symbols. Refer Table 2 for symbols.

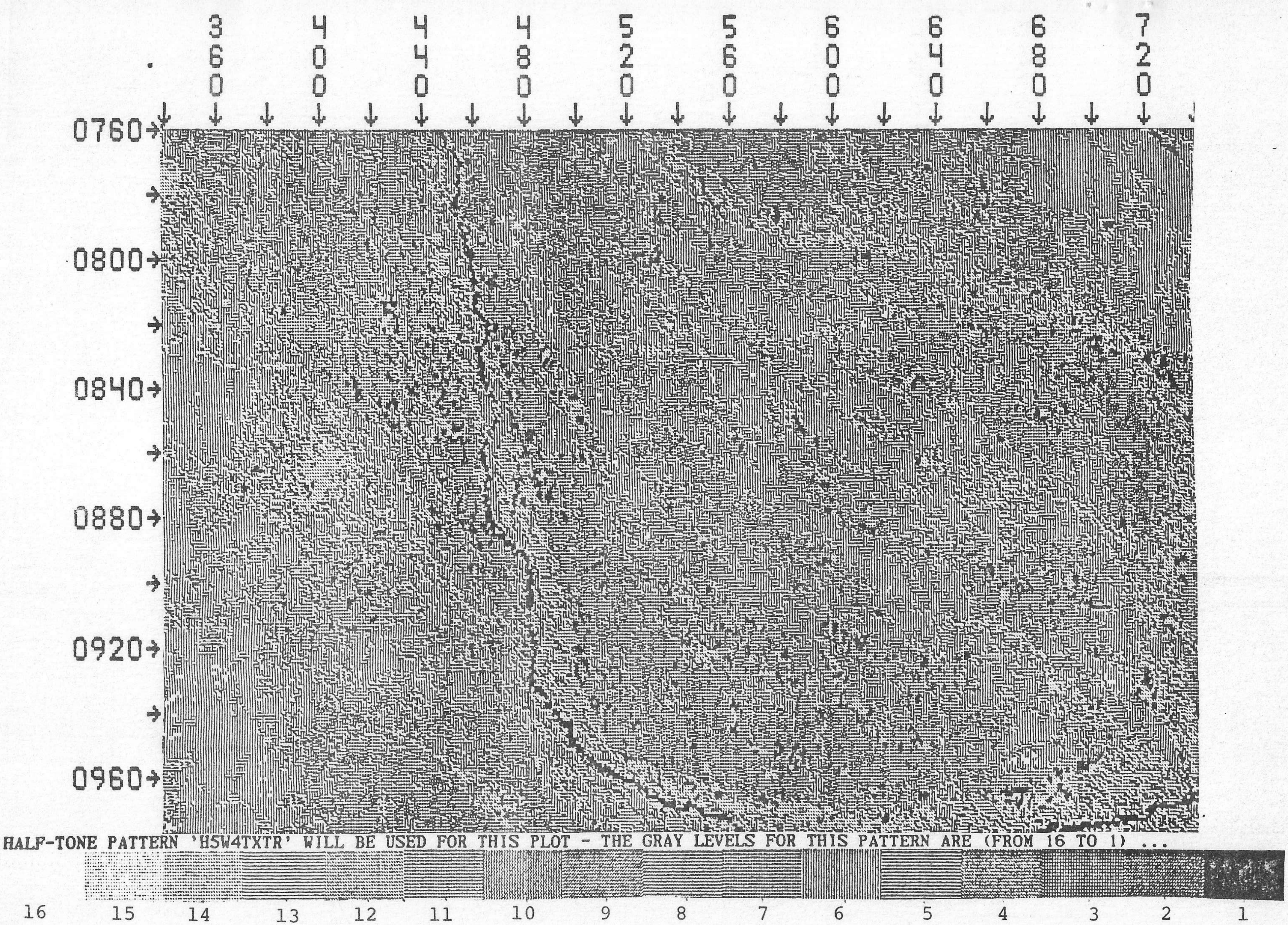
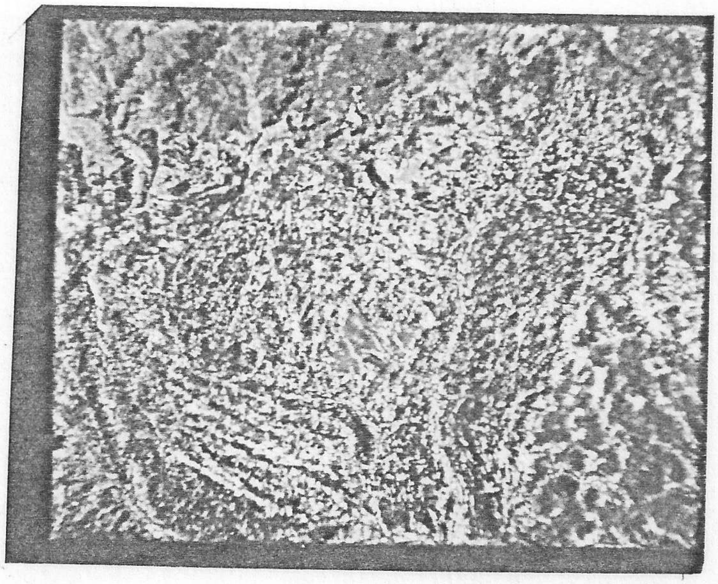


Figure 15. A part of the study area showing the classification results on Varian plotter output with 10 gray levels. Refer Table 2 for gray levels.



Color	Class
	GRASS1,2
	GRASS3,4, GRA/FOR
	SOI/VEG1,2,3,4, OAT/RYE
	FOREST1,2
	CONFOR1,2
	URBAN1
	URBAN2,3,4, SSVEG
	SAND, BRISOIL, CLOUD1,2,3,4, SNOW
	SEA1,2, SEA/SWA, RIVER, LAKE
	SHADOW1,2

Figure 16. Color coded classification results for the northern half of the study area showing 10 groups displayed with different colors.

Table 1 Selected training areas and the ground features in each.

Area#	Area Name	Line	Colum	Class	Expected ground features
1	Lake Nukabira area	270-320	720-810	12	Coniferous forest, Lake
2	Ashoro area	530-580	1160-1260	15	Forest, Grassland
3	Livestock Station(N)	715-770	670-800	12	Grazing grassland, Dry field
4	Livestock Station(S)	770-840	670-790	12	Grazing grassland, Dry field
5	Shimoshihoro area	900-990	830-910	14	Paddy field, Dry field, River
6	Obihiro city	980-1060	690-770	12	Urban area, Dry field
7	Kawanishi area	1230-1280	590-670	12	Dry field
8	Wakudo swamp area	1400-1450	1220-1340	12	Ocean,Swamp,Shrubland

Table 2. Class names and the symbols given to the class/class groups in Figures 6, 12, 13, 14 and 15.

#	Class	Name	Fig.6	Figs.12&13	Fig.14	Fig.15*
1	GRASS1	Grassland 1	A	1	1	6
2	GRASS2	Grassland 2	B	2	2	6
3	GRASS3	Grassland 3	C	3	3	6
4	GRASS4	Grassland 4	D	4	4	6
5	OAT/RYE	Oat & rye	E	0	0	7
6	GRA/FOR	Grass & forest	F	5	5	6
7	FOREST1	Decid. forest	G	6	D	5
8	FOREST2	Mixed forest	H	7	D	5
9	CONFOR1	Conif. forest 1	I	8	C	5
10	CONFOR2	Conif. forest 2	J	9	C	5
11	SOI/VEG1	Soil & vegetation 1	K	A	A	7
12	SOI/VEG2	Soil & vegetation 2	L	B	A	7
13	SOI/VEG3	Soil & vegetation 3	M	C	A	7
14	SOI/VEG4	Soil & vegetation 4	N	D	B	7
15	URBAN1	Urban area 1	O	E	U	12
16	URBAN2	Urban area 2	P	F	V	11
17	URBAN3	Urban area 3	Q	G	V	11
18	URBAN4	Urban area 4	R	H	V	11
19	SAND	Sand	S	I	I	14
20	BRISOIL	Bright soil	T	J	J	14
21	SSVEG	Seashore vegetation	U	K	L	11
22	SEA1	Sea 1	V	L	N	1
23	SEA2	Sea 2	W	M	N	1
24	SEA/SWA	Sea & swamp	X	N	W	1
25	RIVER	River	Y	P	W	1
26	LAKE	Lake	Z	Q	W	1
27	CLOUD1	Cloud 1	$	R	K	16
28	CLOUD2	Cloud 2	+	S	K	16
29	CLOUD3	Cloud 3	=	T	K	16
30	CLOUD4	Cloud 4	/	U	K	16
31	SHADOW1	Shadow 1	Ⓐ	V	S	3
32	SHADOW2	Shadow 2	Ⓑ	W	S	3
33	SNOW	Snow	Ⓒ	X	Y	15

* Numbers indicate gray level shown in Figure 15.