1.Classification INPE-COM.4/RPE 528.711.7:633.61(816.12)	2.Period	4.Distribution Criterion

3.Key Words (selected by the author) REMOTE SENSING AUTOMATIC CLASSIFICATION LANDSAT SUGARCANE	internal ☐ external ☒

5.Report NⓄ INPE-1929-RPE/254	6.Date October, 1980	7.Revised by Sherry Chou Chen

8. Title and Sub-title USE OF LANDSAT DATA FOR AUTOMATIC CLASSIFICATION AND AREA ESTIMATION OF SUGARCANE PLANTATION IN SÃO PAULO STATE, BRAZIL	9.Authorized by Nelson de Jesus Parada Director

10.Sector DSR/DDP Code 30.241.00	11.NⓄ of Copies 06

12.Authorship Francisco José Mendonça	14.NⓄ of Pages 24
13.Signature of the responsible	15.Price

16.Summary/Notes

Sugarcane (Saccharum spp.) plantation of São Paulo State was classified automatically using IMAGE-100 system and LANDSAT digital data. Ten segments of the size 20 x 10 km were aerially photographed and used as training areas for automatic classification. The study area was covered by four LANDSAT paths 235, 236, 237 and 238. The percentages of overall correct classification for these paths range from 79.56% for path 238 to 95.59% for path 237.

17.Remarks   This paper is submitted for publication at the International Journal of Remote Sensing.

USE OF LANDSAT DATA FOR AUTOMATIC CLASSIFICATION AND AREA ESTIMATION OF SUGARCANE PLANTATION IN SÃO PAULO STATE, BRAZIL

Francisco José Mendonça

Instituto de Pesquisas Espaciais - INPE

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

São José dos Campos, SP, Brazil

REVISED VERSION

APRIL, 1981

USE OF LANDSAT DATA FOR AUTOMATIC CLASSIFICATION AND AREA ESTIMATION OF SUGARCANE PLANTATION IN SÃO PAULO STATE, BRAZIL

ABSTRACT

Sugarcane (*Saccharum spp.*) plantation of São Paulo State was classified automatically using IMAGE-100 system and LANDSAT digital data. Ten segments of the size 10 x 20 Km were aerially photographed and used as training areas for automatic classification. The study area was covered by four LANDSAT paths, 235, 236, 237 and 238. The percentages of overall correct classification for these paths range from 79.56% for path 238 to 95.59% for path 237.

Francisco José Mendonça

Instituto de Pesquisas Espaciais - INPE
Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
São José dos Campos, SP, Brazil

Francisco José Mendonça

Instituto de Pesquisas Espaciais - INPE

Av. dos Astronautas, 1758 - C.P. 515

12.200 - São José dos Campos - SP

BRAZIL

USE OF LANDSAT DATA FOR AUTOMATIC CLASSIFICATION AND AREA ESTIMATION OF SUGARCANE PLANTATION IN SÃO PAULO STATE, BRAZIL

Francisco José Mendonça

Instituto de Pesquisas Espaciais - INPE

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

São José dos Campos - SP, Brazil

1. INTRODUCTION

Remote sensing techniques have been applied to agriculture for three basic objectives: 1) to identify different crops according to their spectral characteristics; 2) to determine the best spectral window(s) for crop differentiation, and 3) to obtain multispectral responses of different crops during their growth cycles and under different soil conditions.

Using remote sensing techniques the Large Area Crop Inventory Experiment (LACIE) project contributed the first step toward the development of a global scheme for crop forecasting purposes (MacDonald and Hall, 1978). The possibility of deriving vegetative biomass information from LANDSAT Multispectral Scanner (MSS) data was tested by Wiegand et al. (1973). In their study they concluded that the percentage of exposed soil and the crop growth stage may play a significant role to the spectral response, and through an analysis of spectral response the optimal time for crop discrimination may be selected. Automatic classification of crops using LANDSAT digital data have been experimented by various researchers

(Horton and Heilman 1973; Dietrich et al. 1975). Bauer and Cipra (1973) reported a classification accuracy of 83% for automatic classification of corn and soybean in Northeast Illinois. Higher automatic classification accuracy has been obtained in Baumgardner and Henderson's (1973) study, the percentages of correct classification was 96.1, 97.0 and 97.9 for pastureland, wheat and bare soil respectively.

Since 1975 the Instituto de Pesquisas Espaciais (INPE) has carried out a series of experiments using LANDSAT (ex-ERTS) data and colour infrared (CIR) aerial photographs for crop classification and area estimation studies. In a study carried out by Mendonça et al. (1978) the sugarcane visual interpretation accuracy using LANDSAT images was compared to that of the automatic classification of LANDSAT digital data and the percentages of correct classification were 83 and 89% respectively.

The world energy crisis caused Brazilian Government to initiate an economic policy giving priority to researches using biomass as energy sources. Among the potential vegetations sugarcane is the most suitable source for alcohol production in Brazil and has received governmental incentives through the National Alcohol Program (PNA).

This paper presents a methodology for sugarcane inventory study of São Paulo State, Brazil. Sugarcane was automatically classified and its area was estimated using an interative image analysis system (IMAGE-100).

2. MATERIAL AND METHODS

2.1 - STUDY AREA

São Paulo State was selected as the study area due to its high contribution (45%) to the Brazilian sugarcane production (IAA/PLANALSUCAR 1978). The coordinates of this State are $20^{\circ}00'$ ~ $25^{\circ}00'S$ and $44^{\circ}00'$ ~ $53^{\circ}00'$ W and is covered by six LANDSAT paths (234, 235, 236, 237, 238 and 239) which were taken as strata in this study. After a preliminary study of the distribution of sugarcane plantation using estimates provided by the Institute for Agricultural Economy (Instituto de Economia Agrícola - IEA) strata 234 and 239 were eliminated from the study area because of their low sugarcane occurrences. Strata 235, 236, 237 and 238 encompass approximately 162,352 Km^2 which is equivalent to 65.6% of the geographic area of São Paulo State (Figure 1).

2.2 - AIRCRAFT DATA ACQUISITION

Owing to the lack of "ground truth" data, aerial photographs were taken for intensive study areas. The information derived from aerial photographs should be used not only for training sample selections in automatic classification processing but also to evaluate the classification accuracy of the IMAGE-100 system.

Considering the different soil types, culture managements, climate and LANDSAT paths which may cause different spectral responses, ten intensive study segments of the size 10 x 20 Km^2 were distributed in the

study area to represent all the different phisiognomic patterns. The CIR aerial photographs (Kodak Aerochrome n9 2443) were taken during the period of May 27 to June 12, 1978 by a RC-10 camera which was equipped on INPE's aircraft-Bandeirante.

2.3 - LANDSAT DATA ACQUISITION

Due to the problem of cloud cover only two LANDSAT passes of the year 1978 - April and July/August - were available for these ten segments. The LANDSAT imagery (1:250,000) of channels 5 and 7 and the computer compatible tapes (CCTs) were used for analysis. The data of aircraft mission, the path/row annotation and the LANDSAT pass for each segment are shown in Table 1. Owing to the lack of LANDSAT data segments 3 and 4 were eliminated from multi-temporal comparison analysis.

2.4 - TRAINING SAMPLE SELECTION FOR AUTOMATIC CLASSIFICATION

Aerial photographs were interpreted manually and a thematic map (1:20,000) of sugarcane was derived for each segment. The sugarcane fields, which were also recognizable on LANDSAT imagery, were then delineated. For automatic classification, in order to obtain a 1:1 ratio between the resolution cell of CCT and the picture element of the image monitor of IMAGE-100 system, each segment was enlarged to an approximately scale of 1:100,000. Using auxiliary information of LANDSAT imagery, and thematic map derived from aerial photography and field data, several sugarcane training areas were selected by an electronic cursor to represent different sugarcane patterns appeared on IMAGE-100 monitor. The spectral

information of these training samples was used by the classification algoritm "MAXVER" (Velasco et al. 1978) for sugarcane identification. Other auxiliary classes such as pastureland, bare soil, water body and forest were also classified using "MAXVER" to provide a better discrimination of the study class-sugarcane. After classification an "alphanumeric theme print" of the scale 1:20,000 was produced by IMAGE-100 for each segment for classification accuracy study.

2.5 - EVALUATION OF CLASSIFICATION ACCURACY

The thematic map derived from aerial photographs was used as "ground" information to evaluate the classification accuracy of IMAGE-100. For area estimation, an 1 x 1 cm grid was employed. The segment area (AT) and the sugarcane area within each segment (AR) were then obtained from the thematic map . Afterwards, the transparency of sugarcane thematic map was over laid on the alphanumeric theme print and a point by point comparison was carried out. Each point on alphanumeric print-out corresponds to 0.3 ha. The points classified as sugarcane by IMAGE-100, was then transferred to hectares and designated as AC . Comparing these two products the area of sugarcane on thematic map which was also classified as such on print out was called correct-classification of IMAGE-100 (ACC). In order to indicate the classification performance of IMAGE-100 the following indexes were used:

$$\text{Error of Commision (EC\%)} = \frac{AC - ACC}{AT - AR} \times 100$$

$$\text{Correct Classification (CC\%)} = \frac{ACC}{AR} \times 100$$

$$\text{Error of Omission (EO\%)} = \frac{AR - ACC}{AR} \times 100 = 100 - CC$$

To calculate the overall correct classification for each LANDSAT pass the weighing factor (Wi) using sugarcane density in each segment was used:

$$W_i = \frac{AR_i/AT_i}{\sum\limits_{i=1}^{N} AR_i/AT_i}$$

where "i" was the number of each segment.

2.6 - SUGARCANE AREA ESTIMATION FOR THE SÃO PAULO STATE

The training samples used to classify sugarcane of each segment in one path were also applied to classify the whole stratum. Estimates of these four strata were then summed to give sugarcane estimation of the whole study area. This estimate from IMAGE-100 was compared to the estimate provided by the Institute for Agricultural Economy (Instituto de Economia Agrícola - IEA).

3 - RESULTS AND DISCUSSION

Several segments could not be handled due to adverse

weather and high noise of some CCTs.

The areal extension of each segment, the results of sugarcane area interpreted from aerial photographs and estimated using IMAGE-100 system are presented in Table 2 and 3 for the pass in April and July/August respectively. Stratum 236 has the highest sugarcane density ($\simeq$52%), stratum 238 the lowest ($\simeq$13%) and strata 235 and 237 are the intermediates. The percentages of correct classification, errors of commission and omission for April are shown in Table 4 and for July/August in Table 5.

Generally speaking the percentages of correct classification for July/August are significantly higher and the errors of commission are lower than that for April. These indicate that July/August is a better season for sugarcane classification using a single LANDSAT pass. The explanation for this improvement is that the target which caused classification confusion using IMAGE-100 system was partureland, specially the artificial pasturelands which are carefully managed and gave a similar spectral response as of sugarcane.

Stratum 238 is located in the transition zone of the intensive sugarcane plantation region and rangeland of the State. In 1978 the dry season of the study area started from April and ended in the beginning of June which caused the stress condition of all the vegetation presented in the study area. These stressed vegetation gave similar spectral responses and were difficult for differentiation using IMAGE-100 system. In addition, in the transition region the sugarcane

rotation in the field is not as dynamic as the strata which are located
in the central part of the State where sugarcane is intensively planted
and can be found with different stages of development in the whole year.
This agronomic culture practice of sugarcane cultivation contributes
significantly to the similarity of the spectral patterns of sugarcane
to that of the pastureland in April. A significant reduction in error
of omission of this same stratum, using the data of the pass July/August
can be observed in Table 5.

The classification results of the segments of stratum 236
in April show the variation of commission errors from 31.64% for segment
6 to 54.22% for segment 7 while for stratum 238 vary from 8.31% for
segment 1 to 14.91% for segment 2. The soil of stratum 236 - "Terra Rossa" -
retains more soil moisture which is available for plant growth. Thus, even
in dry season the pasture of this stratum grows vigorously and presents
confusion for discrimination from sugarcane in April. Using the data of
July/August the commission error of sugarcane is reduced significantly.

The situations of the segments in strata 235 and 237 are
intermediate comparing with strata 236 and 238. In April the commission
and omission errors of these segments were almost equal, indicating the
same situation as that commented for the segments in strata 236 and 238.

The errors of commission are approximately equal for both
passes of segments 1 and 10. This may be explained by the undulated
relief with smaller sugarcane fields of segments 1 and 10. The uncontinuity
and the spatial distribution of the small sugarcane fields could then

introduce significant error of commission even in July/August - which is the optimal month for sugarcane classification due to its distinguish spectral characteristics from surrounding vegetations.

The relative differences of correct classification for each stratum comparing the results for April and July/August are shown in Table 6. Strata 235 and 238 show the highest improvement (171.81% and 111.54%). As we mentioned above these were due to the similarities of spectral characteristics of sugarcane and pastureland under stress condition. The stratum 237 shows an increment of 58.45%, and the lowest increment is in the stratum 236 (12.85%). This is because the stratum 236 covers an intensive cultivated region with the best soil type, where the highest sugarcane concentration of the whole State is found. In April and July/August or any other time of the year sugarcane is the predominant crop, thus explaining the smallest increment in correct classification from the data of April to July/August.

The Institute for Agricultural Economy (Instituto de Economia Agrícola - IEA) in São Paulo State provides information about crop production five times annually. The sampling error of the final estimate at the State level is believed around 10% (Campos and Piva 1974). Extrapolating the information of training samples of the pass July/August to classify sugarcane of the four strata gives a total estimate of 13,349.84 Km^2 which shows an overestimate of 12.57% when comparing to the data provided by the Institute for Agricultural Economy (São Paulo, 1978). No classification accuracy was obtained for the study area, however, an overall correct classification of 87.96% was achieved for the eight

segments where comparisons to aerial photographs were possible.

4. CONCLUSIONS

LANDSAT CCTs were used for area estimate of sugarcane plantation of the São Paulo State using IMAGE-100 system. The following significant results were obtained from this study:

a) crop thematic map derived from aerial photographs is efficient for training sample selection and provides precise information to evaluate the results of automatic classification using IMAGE-100 system;

b) crop density, field size and its spatial distribution, are all important factors which influence the spectral characteristics of the study crop;

c) the large areas of cultivated pastureland of the study area which have similar spectral response as sugarcane account for the commission errors encountered;

d) in dry season (April) the classification of sugarcane is more difficult due to the spectral similarities of sugarcane and that of pasture. July/August is the optimal time for sugarcane classification using IMAGE-100 system; and

e) comparing sugarcane area estimated using LANDSAT data and IEA

data a difference of 12.57% was encounted. This result is every encouraging considering the 10% sampling error at the State level using conventional survey method.

REFERENCES

Bauer, M.V. and Cipra, J. E. (1973) Identification of agricultural
crops by computer proceeding of ERTS MSS data. Proceedings of the
third symposium on significant results obtained from the earth
resources technology satellite-1, Washington, D.C., V.1, sec.A,
205-212.

Baumgardner, M.F. and Henderson Jr., A. (1973) Mapping soils, crops,
and rangelands by machine analysis of multitemporal ERTS-1 data.
LARS information note 121173, Purdue University, West Lafayette,
Indiana, 19 p.

Campos, H. de and Piva, L.H. de O. (1974) Dimensionamento de amostra
para estimativa e previsão de safra no Estado de São Paulo *(The
determination of the sample size for crop forecasting in the State
of São Paulo)*. Agricultura em São Paulo, Boletim Técnico do Instituto
de Economia Agrícola, 3 (21):65-68.

Dietrich, D.L.; Fries, R.F.; Egbert, D.D. (1975) Agricultural inventory
capabilities of machine processed LANDSAT digital data. Proceedings of
the earth resources survey symposium, Houston, Texas, V.1, sec.A,
221-232.

Horton, M.L. and Heilman, J.L. (1973) Crop identification using ERTS
imagery. Proceedings of the third symposium on significant results
obtained from the earth resources technology satellite-1, Washington,

D.C., V.1, sec.A., 27-33.

Instituto do Açúcar e do Álcool/Planalsucar. (1978) Estações Experimen tais - Relatório anual 1978 *(Experiment stations-annual report 1978)*. Piracicaba, São Paulo, Brazil.

McDonald, R.B. and Hall, F.G. (1978) LACIE: an experiment in global crop forecasting. The LACIE symposium, Houston, Texas, oct., 1-32.

Mendonça, F.J.; Lee, D.C.L.; Tardin, A.T. (1978) Estudo comparativo de interpretação automática e visual em imagem LANDSAT visando a identificação da cultura da cana-de-açúcar. *(A comparative study of sugarcane identification results using visual and automatic interpretation on LANDSAT imagery)*. Instituto de Pesquisas Espaciais (INPE-1265-NTI/104). São José dos Campos, São Paulo, Brazil.

São Paulo. **Secretaria da Agricultura** (1978). Previsões e estimativas das safras agrícolas do Estado de São Paulo 1977/78: 5º levantamento. *(Crop production estimates and prediction of São Paulo State 1977/78: 5th crop survey)*. São Paulo, Brazil, p.5.

Velasco, F.R:D.; Prado, L.O.C.; Souza, R.C.M. (1978) Sistema MAXVER - Manual do Usuário *(MAXVER - User's Manual)* Instituto de Pesquisas Espaciais (INPE-1315-NTI/110). São José dos Campos, São Paulo, Brazil.

Wiegand, C.L.; Gausman, H.W.; Cuellar, J.A.; Gerberman, A.H.; Richardson, A.J. (1973) Vegetation density as deduced from ERTS-1 MSS response.

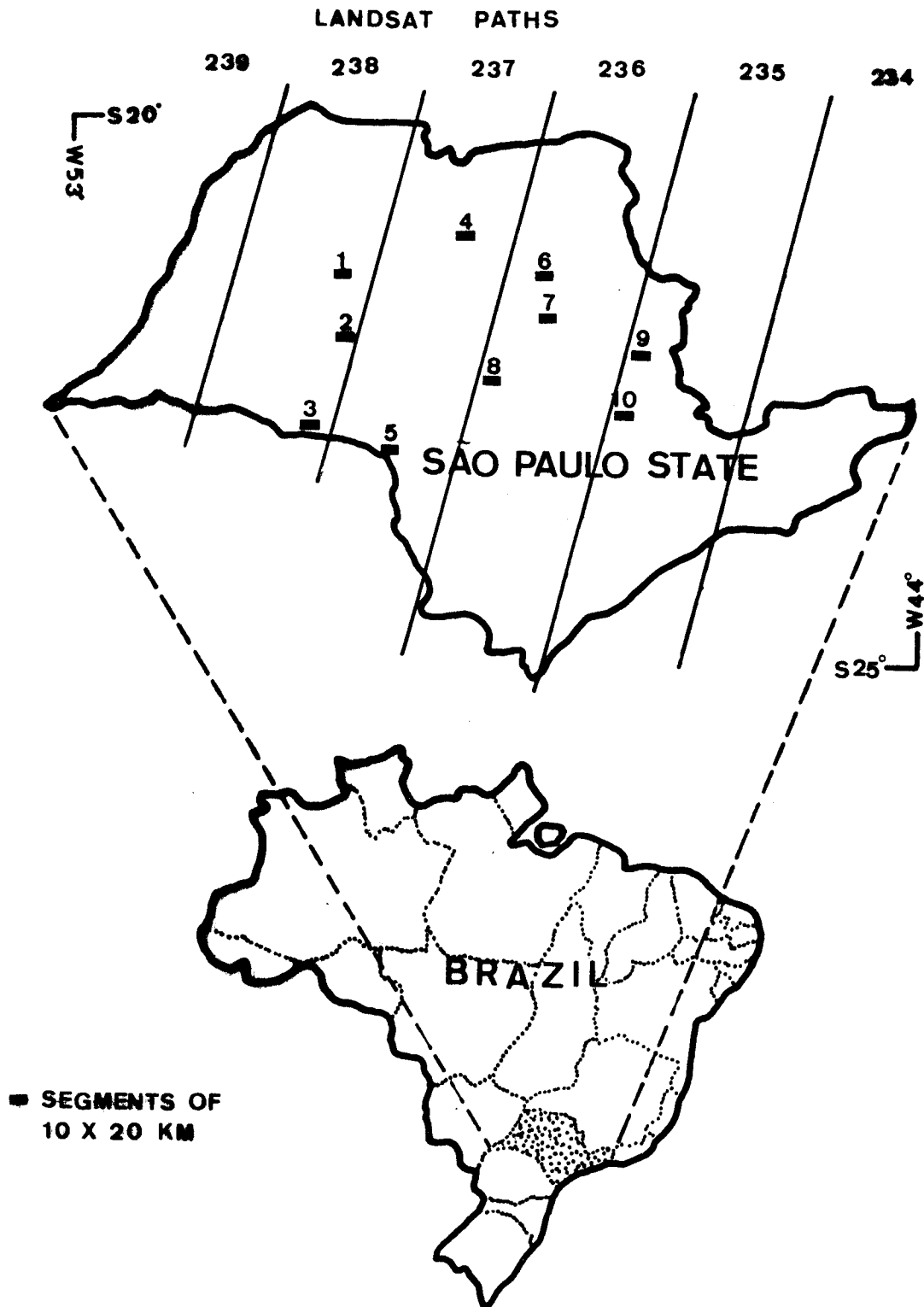Proceedings of the third earth resources technology satellite-1 symposium, Washington, D.C.,V.1, sec. A, 93-116.

Fig. 1 - Location map of the study area and the ten segments for aerial photographing.

TABLE 1

INFORMATION OF AIRCRAFT AND LANDSAT DATA USED FOR THE STUDY

AIRCRAFT DATA (1978)		LANDSAT DATA (1978)		
NUMBER OF SEGMENTS	MISSION DATE	PATH/ROW ANNOTATION	PASS	
			APRIL	JULY/AUGUST
1	27/05	238/27	08/04	12/08
2	23/05	238/27	08/04	12/08
3	26/05	238/28	-	-
4	12/06	237/27	07/04	-
5	24/06	237/28	07/04	11/08
6	12/06	236/27	06/04	05/07
7	12/06	236/27	06/04	05/07
8	26/05	236/27	06/04	05/07
9	28/05	235/27	05/04	04/07
10	12/06	235/27	05/04	04/07

TABLE 2

SUGARCANE AREA ESTIMATION USING LANDSAT DATA

OF APRIL 1978

SEGMENT NUMBER	STRATUM	AREAL EXTENT OF SEGMENT (ha) (AT)	SUGARCANE AREA (ha)		
			AERIAL PHOTOGRAPHS (AR)	LANDSAT DATA (AC)	DENSITY (AR/AT) x 100
1	238	20,013	3,295	2,453	16.5
2		20,000	2,072	3,181	10.4
4	237	12,023	3,720	4,745	30.9
5		20,216	4,381	5,534	21.7
6	236	11,468	6,202	6,286	54.1
7		10,612	5,796	6,739	54.6
8		19,879	10,030	12,951	50.5
9	235	20,492	5,550	4,195	27.1
10		20,000	7,432	5,529	37.2

TABLE 3

SUGARCANE AREA ESTIMATION USING LANDSAT DATA

OF JULY/AUGUST 1978

SEGMENT NUMBER	STRATUM	AREAL EXTENT OF SEGMENT (ha) (AT)	SUGARCANE AREA (ha)		
			AERIAL PHOTOGRAPHS (AR)	LANDSAT DATA (AC)	DENSITY (AR/AT) x 100
1	238	6,617	995	1,213	15.0
2		20,000	2,072	3,412	10.4
5	237	20,216	4,381	5,604	21.7
8'		19,350	9,645	10,055	49.8
6	236	11,468	6,202	6,218	54.1
7		13,480	7,108	6,574	52.7
9	235	20,492	5,550	5,533	27.1
10		14,600	6,484	7,036	44.4

TABLE 4

POINT-BY-POINT COMPARISION OF SUGARCANE CLASSIFICATION ACCURACY FOR

LANDSAT PASS IN APRIL 1978

SEGMENT NUMBER	STRATUM	CORRECT CLASSIFICATION (CC)	ERROR OF OMISSION (EO)	ERROR OF COMMISSION (EC)
1	238	32.3	67.7	8.3
2		24.5	75.5	14.9
4	237	64.6	35.4	28.2
5		54.2	45.8	20.0
6	236	74.5	25.5	31.6
7		71.2	28.8	54.2
8		79.0	21.0	51.0
9	235	49.1	51.0	9.8
10		33.0	67.0	24.5

TABLE 5

POINT-BY-POINT COMPARISON OF SUGARCANE CLASSIFICATION ACCURACY FOR

LANDSAT PASS IN JULY/AUGUST 1978

SEGMENT NUMBER	STRATUM	CORRECT CLASSIFICATION (CC)	ERROR OF OMISSION (EO)	ERROR OF COMMISSION (EC)
1	238	76.0	24.0	8.1
2		84.8	15.3	9.2
5	237	94.2	5.8	9.3
8'		96.2	3.8	8.0
6	236	88.7	11.3	13.7
7		80.1	19.9	13.8
9	235	92.0	8.0	2.8
10		79.3	20.7	23.4

TABLE 6

COMPARISONS OF CORRECT CLASSIFICATION OF SUGARCANE FOR LANDSAT PASSES

OF APRIL AND JULY/AUGUST 1978

STRATUM	LANDSAT PASS		RELATIVE DIFFERENCE (%)
	APRIL	JULY/AUGUST	
235	39.8	84.1	111.3
236	74.8	84.4	12.8
237	60.3	95.6	58.5
238	29.3	79.6	171.7