CR-157374

(E78-10184) INPE'S CROP SURVEY PROGRAM USING COMBINED LANDSAT AND AIRCRAFT DATA (Instituto de Pesquisas Espaciais, Sao Jose) 27 p HC A03/MF A01 CSCL 02C

N78-31483

Unclas

G3/43 00184

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO

INSTITUTO DE PESQUISAS ESPACIAIS

INDEX

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This paper reports the results of several remote sensing techniques used for applied crop area estimations in an agricultural region of São Paulo State. Principal data sources included agricultural statistics from the Secretariat of Agriculture, CIR imagery acquired from t ɔ INPE flight missions and multidate LANDSAT CCT's which were processed on the "Interactive Multispectral Imagery Analyser System" - INPE's Image-100. Efforts to produce area estimates concentrated on such ground features as "bare soil" (soil prepared for planting) and sugar cane. A total of 13 classes were sampled using CIR imagery and four classes - cotton, sugar cane, soybean and pastureland - were interpreted from LANDSAT CCT's through the Image-100 processor. Flight experiments conducted in 1976 proved that the use of CIR imagery at a scale of 1:10,000 instead of 1:20,000 (1975), did not improve on the interpreters' crop identification. An interpretation test to determine an optimum scale for the Image-100 classification of "bare soil" and "arboreal vegetation" showed no significant difference between scale changes but a scale of 1:150,000 was selected for training after considering field size within the study area. Areal estimates were made using a dot grid method and the Image-100 system. LANDSAT CCT's were partitioned into quadrangles and "single cell" and "clustering syntesis" algorithms were used for classification. "Bare soil" comprised 34.0% of the study area. A total of 156 Image-100 hours were spent in the study.

1 - INTRODUCTION AND APPROACH

Brazil is essentially an agricultural country. It is the world's largest coffee producer, second largest corn producer and a leading producer of sugar, cotton, rice, soybean and peanuts. The dynamics of cropping and food demand are subject to many variables which points to the necessity for timely crop production estimates.

Considering that the current methods of crop prediction in Brazil are subjective and often outdated, the need for new techniques is essential in order to obtain greater accuracy and more timely information on a regular basis. The conventional system of crop forecasting which has been used for many years by the Secretariat of Agriculture, is based on the sampling of a number of agricultural properties where questionaires are administered 5 times during the growing season. It has been observed that the main source of error is in the area estimates given by farmers, while estimates of crop productivity are more accurate. For Brazil, LANDSAT-MSS data processing offers new perspectives to meet our goals with respect to crop surveys over large regions.

The Brazilian Institute of Space Research (INPE), in agreement with the Secretariat of Agriculture of São Paulo State, has been carrying out an Agricultural Statistics Project (EAGRI) since 1975. The objective of this project is to verify the feasibility of remote sensing techniques for crop forecasting, primarily using LANDSAT data and the "Interactive Multispectral Imagery Analyser System" (IMAGE-100).

According to the project objectives INPE has had the responsibility of developing a methodology using remote sensing techniques for crop area measurement. The Secretariat of Agriculture has contributed ground data and productivity estimates.

There are many complexities inherent to a project of this type. Problems encountered during the development of a remote sensing methodology included:

1) the growing season of annual crops is from October to May, in coincidence with the rainy season which restricts the aquisition of cloud-free imagery during this period of interest;

2) major crops are planted in relatively small areas. This causes some difficulties in either spectral discrimination or in area estimates due to the low resolution of LANDSAT imagery.

Considering these problems, and the current potentialities of remote sensing techniques, efforts to produce area estimates concentrated on the following ground features:

1) BARE SOIL (area estimates of soil being prepared for planting). This information is important as it can be associated with the Secretariat of Agricultures's publication of "The Report of Planting Intentions of Farmers". Both, area estimates and the Secretariat's publications, could provide a preliminary crop prediction at the beginning of the growing season. The soil is prepared for planting from June to October. This is the best period of the year to acquire cloud-free coverage. Also, during these months bare soil is well defined on LANDSAT imagery due to the sharp contrast with neighboring vegetation.

2) SUGAR CANE - area estimates of sugar cane can also be carried out during the dry season. Another advantage is that sugar cane is cultivated over a large and continuous area, which helps in crop identification.

3) OTHER CROPS - the feasibility of using remote sensing data for coffee, citrus and pastureland surveys is also being studied.

2 - INTERPRETATION METHODOLOGY AND RESULTS

The study area (Figure 1) is the Regional Administrative ·vision of Ribeirão Preto (DIRA-RP). This region includes our test area, the Município of Jardinópolis, which contains the major representative crops of São Paulo State.

2.1 - VISUAL INTERPRETATION

Two CIR aerial surveys were conducted over Jardinópolis. The first was in March, 1975 (scale 1:20.000) and the second was in February, 1976 (scale 1:10.000). The objectives of both experiments were to test the feasibility of using aircraft data for area estimates and to give support to LANDSAT data interpretation.

The EAGRI Project frequently uses areal measurements taken from CIR imagery to support LANDSAT data analysis. Two different methods of area measurements were compared: the dot grid method (using $1cm^2$ and $2cm^2$ grids) and the weighing method.

Five aerial photographs from 1975 were analyzed. The maximum and minimum percentages of ground cover for interpreted crops ranged from 50% to 1%. From the results of the statistical analysis, no differences were observed between the two methods. However, the dot grid method is more time-saving (using computer analysis) than the "weighing" method and thus the 2 cm^2 grid was selected for the crop area estimates in our study.

The total area of Jardinópolis (521 km^2) was mapped by visual photointerpretation and supported by ground observations. A thematic map was constructed to show the boundaries of each crop present in the test area. Area estimates were made for each crop type. Both the thematic map and the area estimates were the first basic data compiled of the test site from IMAGE-100 interpretation.

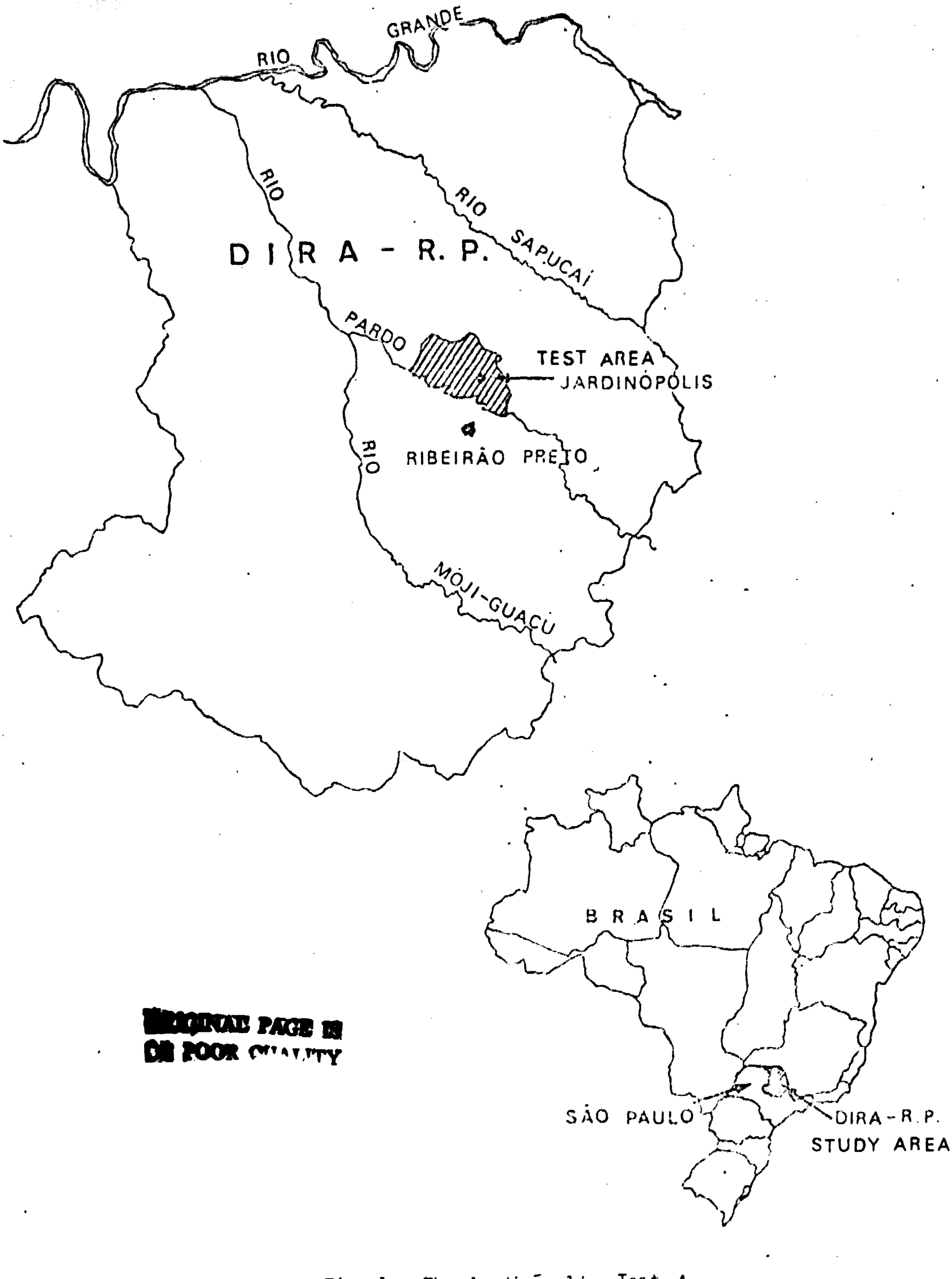

Fig. 1 - The Jardinópolis Test Area.

2.2 - AUTOMATIC INTERPRETATION

An optimum scale determination for the IMAGE-100 operation was conducted to produce an accurate and time-saving classification of "bare soil" and "arboreal vegetation". An area of approximately 1,400 km^2 was selected and analysed at four different scales: 1:125,000, 1:150,000, 1:200,000 and 1:300,000. A computer-aided interpretation (IMAGE-100) was administered by four trained technicians who were asked to classify bare soil and arboreal vegetation in four scales for one - third of the study area. A two way analysis of variance was performed. The interpreter's area estimations were run against the four different scales (tables 1 and 2).

The results show that there is a significant difference among interpreter's scores for bare soil and arboreal vegetation at $p = 0.10$ and $p = 0.05$, respectively (tables 3 and 4).

The results of tables 1 and 2 show that the individual interpretations of arboreal vegetation were more consistent than for bare soil at different scales. One can conclude that the interpreters should have a better a priori knowledge of the area. In the case of bare soil, the lack of field information in areas with different soil types and the difficulty of perceiving pixel density explained the differences among the interpreters' classification performances. Arboreal vegetation is a class which can be easily identified without field information. The variation of pixel density is the prominent factor in explaining the differences among the interpreters' scores. This suggests that, besides using training samples, a proficient knowledge of the test area is required to quantitatively evaluate the capabilities of the IMAGE-100 System. The results showed no statistically significant differences among the scores between scale changes.

Considering the general field size of the region and the selection of representative training samples, the scale of 1:150,000 proved to be the most suitable for computer-aided interpretation and analysis.

TABLE 1

BARE SOIL AREA ESTIMATION (KM^2) AT DIFFERENT SCALES BY FOUR INTERPRETERS

SCALES \ INTERPRETERS	1	2	3	4	MEAN	C.V.
1:125,000	437	492	354	463	436,5	11.79
1:150,000	591	559	447	431	507,0	13.64
1:200,000	510	501	502	365	469.5	12.87
1:300,000	645	549	273	493	490,0	27.87
MEAN	545.80	538.9	391.8	443.0		
C.V.	41.49	5.54	22.24	10.85		

TABLE 2

ESTIMATED ARBOREAL VEGETATION AREA (KM^2) AT DIFFERENT SCALES BY FOUR INTERPRETERS

SCALES \ INTERPRETERS	1	2	3	4	MEAN	C.V.
1:125,000	76	108	98	79	90.3	14.71
1:150,000	80	114	80	96	92.5	15.16
1:200,000	74	123	80	79	89.0	22.20
1:300,000	69	113	86	88	89.0	17.64
MEAN	74.80	114.50	80.4	85.50		
C.V.	4.72	5.30	8.54	8.29		

TABLE 3

ANALYSIS OF VARIANCE OF BARE SOIL ESTIMATED AREA

SOURCE OF VARIATION	d.f.	S.S.	M.S.	F
Scales	3	11,037.0	3,679.00	0.57 M.S.
Interpreters	3	61,833.0	20,611.17	3.25*
Error	9	57,104.5	6,344.94	
TOTAL	15	129,975.0		

* $F_{0.10} = 2.81$ $F_{0.05} = 3.86$

TABLE 4

ANALYSIS OF VARIANCE OF ESTIMATED ARBOREAL VEGETATION AREA

SOURCE OF VARIATION	d.f.	S.S.	M.S.	F
Scales	3	32.69	10.90	0.17M.S.
Interpreters	3	3,475.69	1,158.56	18.48 **
Error	9	546.06	62.67	
TOTAL	15	4,072.44		

F 0.05 = 8.72

2.3 - AREAL EVALUATION OF BARE SOIL, ARBOREAL VEGETATION AND PASTURE/ SUGAR CANE USING LANDSAT AUTOMATIC INTERPRETATION

After an optimum scale was determined (1:150,000), a bare soil estimation was conducted for areas of the DIRA-RP region using the IMAGE-100 System. The following corresponding computer compatible tapes (CCT) were analysed:

ORBIT/POINT	DATES
178/26	09/13/75
178/27	09/13/75, 05/10/75
192/26	09/14/75
192/27	09/14/75

Support data included four LANDSAT MSS frames (channels 5 and 7) at scales of 1:1,000,000 and 1:250,000, CIR imagery at scales of 1:10,000 and 1:20,000 and ground information.

Each frame of the whole scene was partitioned into quadrangles equivalent to $1/_{16}$ of the original LANDSAT frame. This provided a scale of approximately 1:150,000 on the Image-100 display.

Initially, several algorithms of the Image-100 System were used. Then, spectral responses of the targets (i.e. bare soil, arboreal vegetation and pasture/sugar cane in the four MSS channels were acquired.) This preliminary analysis led to the selection of "single cell aquisition" and "clustering synthesis" algorithms for classification purposes.

Quadrangles were classified according to these algorithms. The spectral resolution used was 128 bits because a 256 resolution did not improve the classification and a 64 resolution generated too much overlapping of themes.

The results show that 12,678 km^2 were classified as bare soil for September 13, 1975. This corresponds to 34% of the DIRA-RP region which has an area of 36,801 km^2, as assessed by the Image-100. Besides bare soil evaluation, pasture/sugar cane (as a single class) and arboreal vegetation were also classified. These three classes, together, gave an area of 30,637 km^2 which corresponds to 83% of the DIRA-RP area. A pixel by pixel analysis will be carried out in the future, to determine classification accuracy by comparing the automatic classification output with aerial photographs. A total of 156 Image-100 hours were spent in this study. Only 75 hours were effectively spent for bare soil classification, and the rest was spent for radiometric correction and recording.

Table 5 shows the spectral responses in the four channels of the four LANDSAT frames analysed for bare soil, arboreal vegetation and pasture/sugar cane. These results are the average of the responses (upper and lower bounds) of each class, in all quadrangles of each frame mentioned previously.

2.4 - CROP IDENTIFICATIONS AND AREA ESTIMATES - 1975

The preliminary results of the 1975 crop inventory work for the Jardinópolis test area (Fig. 1) are discussed below:

Figures 2, 3 and 4 show area delimitations visually interpreted from LANDSAT channels 5 and 7 (May 10, 1975) at a scale of 1:250,000. This particular scene was concomitant with the harvest of the major crops. The results of areal crop estimates, based on visual interpretation of CIR imagery and automatic interpretation of LANDSAT CCT's, are presented in Table 6.

TABLE 5

LANDSAT MSS SPECTRAL RESPONSES OF BARE SOIL, ARBOREAL VEGETATION, AND PASTURE/SUGAR CANE FOR THE ANALYSED FRAMES

ORBIT/POINT	CLASSES	BARE SOIL		PASTURE-SUGAR CANE		ARBOREAL VEGETATION	
	CHANNELS \ BOUNDS	UPPER	LOWER	UPPER	LOWER	UPPER	LOWER
178/26	4	28	37	34	48	27	36
	5	37	47	40	60	25	38
	6	37	52	46	75	43	60
	7	13	23	23	46	22	40
178/27	4	32	48	38	54	31	40
	5	39	59	37	64	26	39
	6	39	69	53	71	45	60
	7	15	39	27	44	24	38
192/26	4	30	38	33	47	29	39
	5	37	50	42	62	27	25
	6	34	51	56	74	40	61
	7	11	21	25	46	21	39
192/27	4	28	38	34	44	27	39
	5	37	49	42	58	28	39
	6	35	53	48	70	38	59
	7	12	22	25	45	22	36

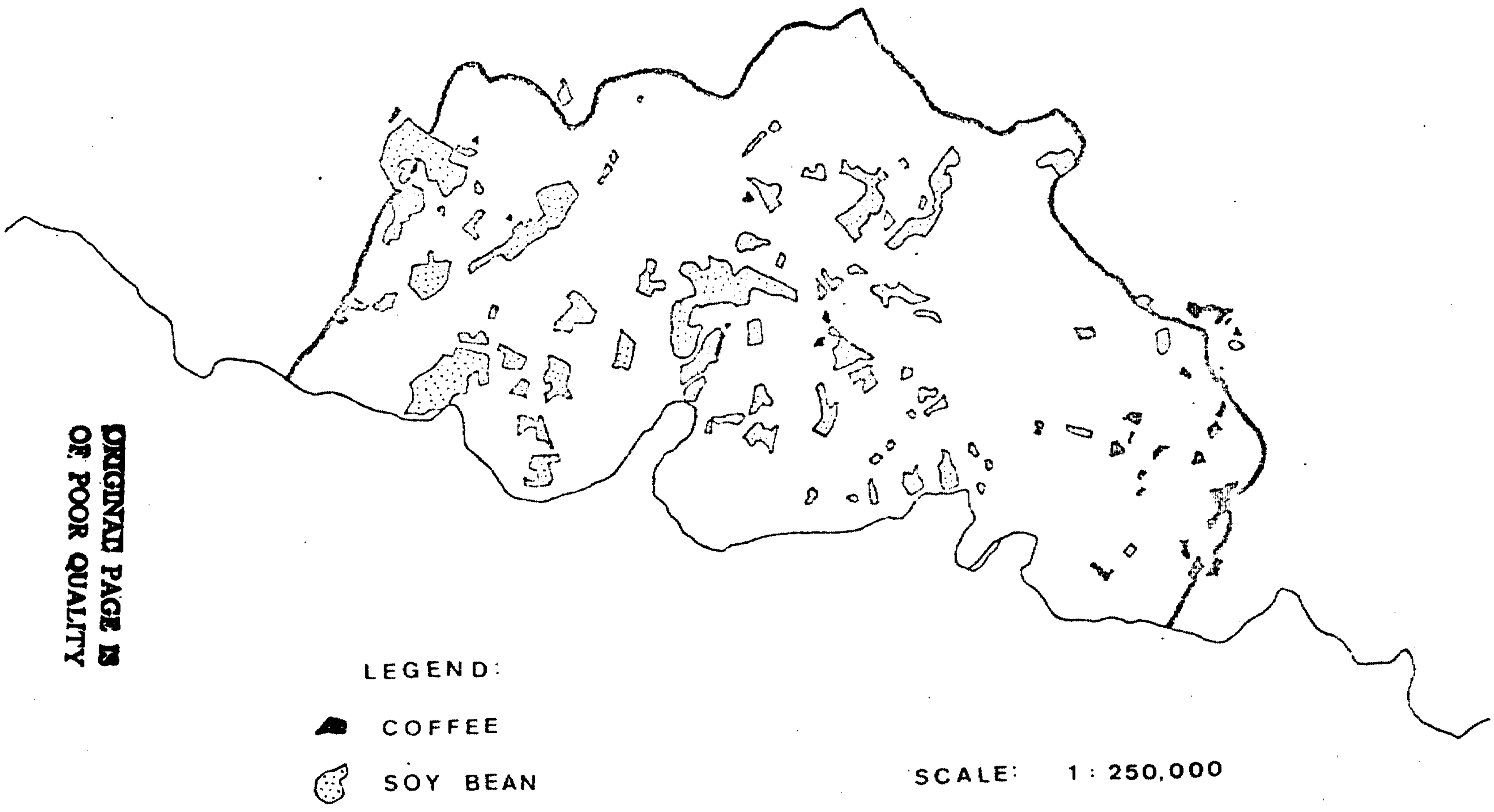

Fig. 2 - Visual interpretation of LANDSAT MSS imagery taken on May 10, 1975 (Jardinōpolis County).

ORIGINAL PAGE IS
OF POOR QUALITY

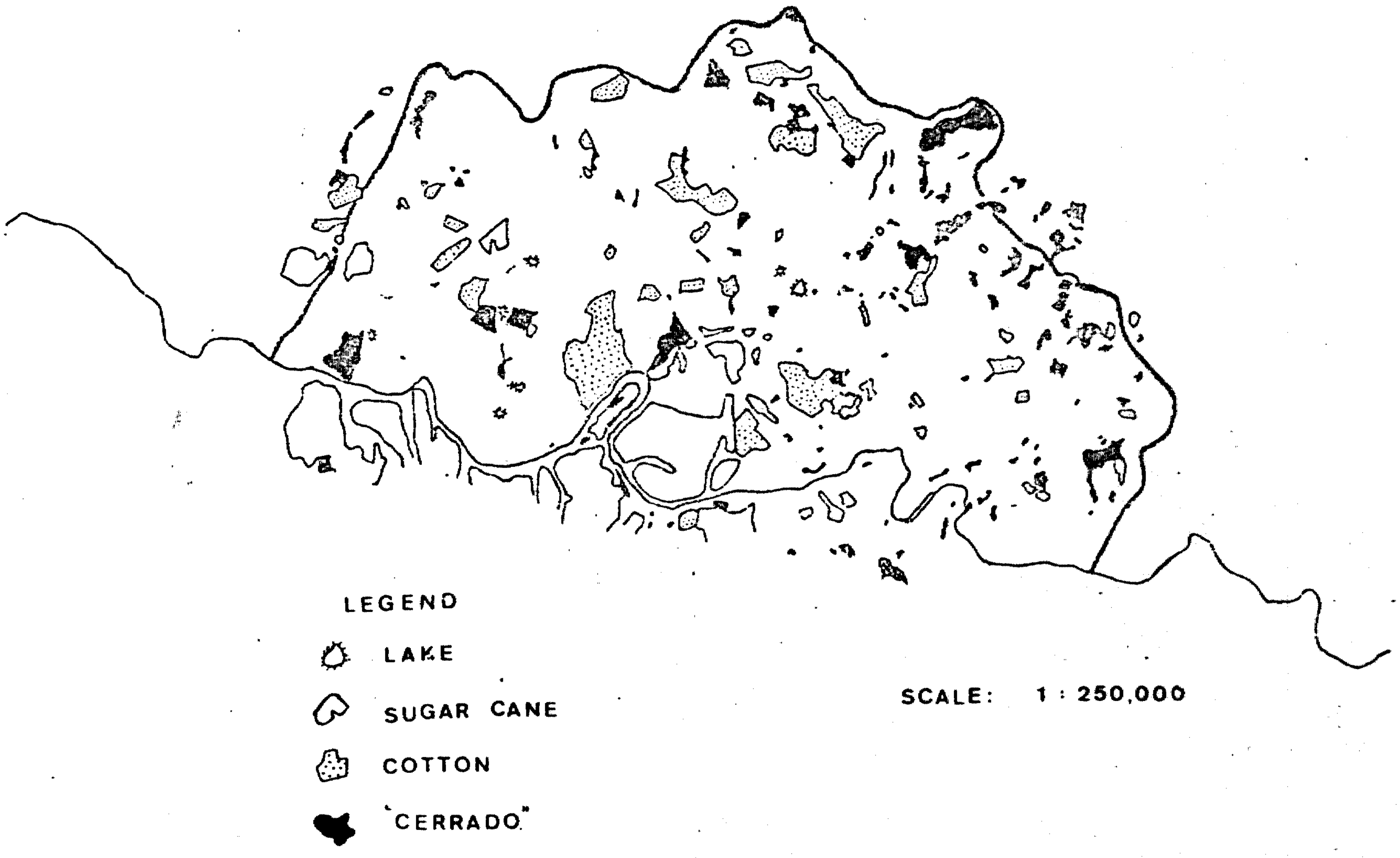

Fig. 3 - Visual interpretation of LANDSAT MSS imagery taken on May 10, 1975 (Jardinópolis County).

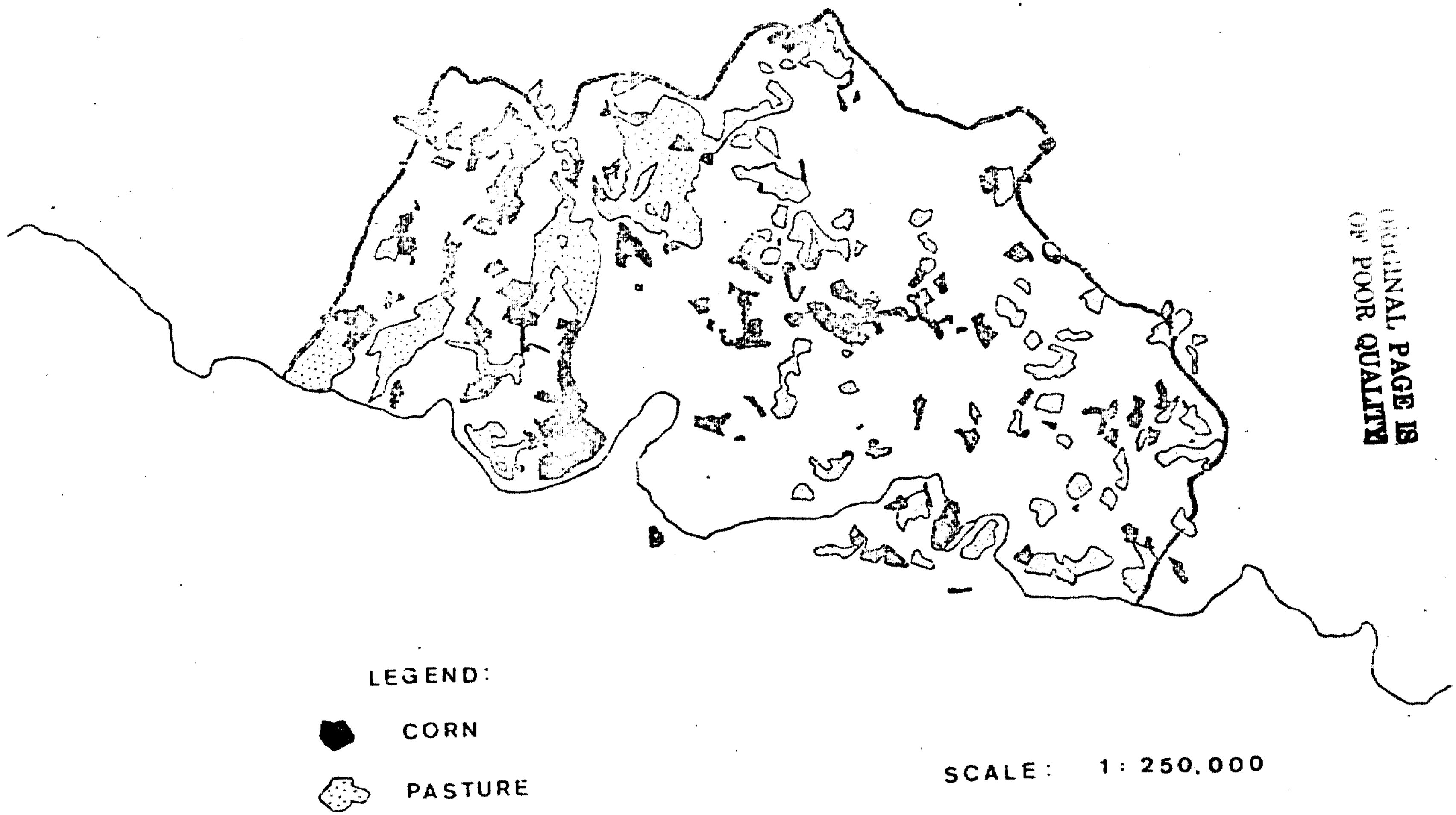

Fig. 4 - Visual interpretation of LANDSAT MSS imagery taken on May 10, 1975 (Jardinópolis County).

TABLE 6

PRELIMINARY RESULTS OF AREAL CROP ESTIMATES (KM^2) IN THE MUNICÍPIO OF JARDINÓPOLIS BY VISUAL AND AUTOMATIC INTERPRETATION

CLASSES	AREAL ESTIMATE (KM^2) IN 1975		PERCENT CHANGE
	VISUAL INTERPRETATION USING CIR FILM (A)	AUTOMATIC INTERPRETATION OF LANDSAT DATA (B)*	$\frac{B - A}{A} \times 100$
COTTON	42.31	33.40	- 21.06
SUGAR CANE	25.09	21.40	- 14.71
SOY BEAN	66.06	70.67	+ 6.97
PASTURELAND	182.86	166.13	- 9.15

* LANDSAT CCT May 10, 1975.

Comparing the percent change between the four classes, soybean was overestimated on the IMAGE-100 compared to the conventional interpretation method (+ 6.97%). A moderate information loss was scored for pastureland (- 9.15%) and increasing losses of interpretability were shown for sugar cane (- 14.71%) and cotton (-21.06%), which ranked last among the four classes. It is apparent that automatic interpretation more closely approximates visual interpretation with increasing size of the target area. Both the crop area delimitations and area estimations could be improved if CIR imagery and sequential LANDSAT data were acquired during the growing season.

2.5 - CROP IDENTIFICATIONS AND AREA ESTIMATES - 1976

In the 1976 flight mission over the Jardinópolis test area, a scale change to 1:10,000 was made with the objective of determining if greater accuracy of crop identification could be achieved. The aerial survey was conducted during the first fortnight of February, midway through the growing season, for the major crops. The visual interpretation of the imagery was supported by ground truth data, which combined to give 100% identification accuracies for: cotton, coffee, sugar cane, pastureland, soybean, orchard and cerrado. High values were obtained for corn (98%) rice (94%), peanuts (92%), and sorghum (82%). The use of CIR imagery at a scale of 1:10,000 (1976), instead of 1:20,000 (1975), did not improve on the interpreters' crop identification accuracy. It is concluded that with the high identification accuracies achieved at a scale of 1:20,000, smaller scales could be used to reduce operational and interpretation costs.

Table 7 presents estimated crop percentages and equivalent areas (km^2) from a sample of 63 CIR images. Errors of estimations were calculated for 13 classes. The mean percent error of estimation was 21.69%. A total of 100 images cover the Jardinópolis area. A calculation was made to determine the number of photos needed for crop area estimations with a 10% error. After the termination of this phase of the project it was concluded that the CIR imagery at a scale of 1:10,000 was inadequate for crop inventory work. First, there was no improvement in crop identification accuracy compared to a scale of 1:20,000. Second, the photographic sampling units at this scale produced a high variance in the crop estimations.

Table 8 shows the results derived in both the 1975 and 1976 aerial CIR surveys.

The comparison of these two tables made possible the verification of crop changes from one year to another. For example, it

TABLE 7

ESTIMATES CROP PERCENTAGE AND AREA (KM^2) BASED ON A SAMPLE OF 63*
CIR IMAGES OF JARDINÓPOLIS (1976)

CLASS	PERCENTAGE+	$\bar{A}$REA (KM^2)+	SAMPLING ERROR %	ERROR OF ESTIMATION %	NUMBER OF PHOTOS NEEDED FOR 10% ERROR OF ESTIMATE (α = 0.10)
Cotton	6.65 ± 1.69	34.64 ± 8.80	15.19	25.41	92
Peanut	0.96 ± 0.30	5.00 ± 1.55	18.04	31.25	95
Rice	2.24 ± 0.33	11.67 ± 1.71	8.93	14.73	79
Coffee	1.12 ± 0.36	5.84 ± 1.87	19.96	32.14	95
Sugar Cane	10.05 ± 2.43	52.36 ± 12.67	14.49	24.18	91
Corn	14.52 ± 1.63	75.65 ± 8.50	6.71	11.23	69
Pastureland	34.32 ± 3.03	178.80 ± 15.81	5.29	8.83	58
Soy bean	7.79 ± 1.30	40.59 ± 6.75	9.94	16.69	83
Orchard	4.20 ± 1.04	21.88 ± 5.39	14.68	24.76	92
Cerrado	5.26 ± 0.78	27.40 ± 4.06	8.92	14.83	79
Artificial Forest	0.50 ± 0.13	2.61 ± 0.69	15.71	26.00	93
Sorghum	1.93 ± 0.76	10.06 ± 3.97	23.74	39.38	97
Other	10.46 ± 1.31	54.50 ± 6.80	7.47	12.52	73

* A total of 100 images cover the entire Jardinópolis area without overlap.

+ Estimates with a 90% confidence interval.

TABLE 8

COMPARISON OF CROP AREAL ESTIMATES USING AERIAL PHOTOGRAPHS IN BOTH AIRCRAFT EXPERIMENTS (1975 AND 1976)

CLASSES	ESTIMATED AREAS (KM^2)		PERCENT DIFFERENCE BASED ON 1975 DATA
	1975	1976	
Cotton	42.32	39.65	- 18.12
Peanut	4.52	5.00	+ 10.62
Rice	4.76	11.67	+145.17
Coffee	7.94	5.84	- 35.96
Sugar Cane	25.09	52.37	+108.73
Corn	81.38	75.66	- 7.03
Soy bean	66.06	40.59	- 38.56
Orchard	31.55	21.88	- 30.65
Pastureland	182.86	178.83	- 2.20
Cerrado	35.19	27.41	- 22.11
Artificial Forest	2.58	2.61	+ 1.16
Sorghum	-	10.06	-
Other	36.82	54,50	+ 48.02
TOTAL	521.07	521,07	-

was observed that areas planted with sugar cane increased more than 100%. This induced a decrease in areas planted with other crops.

3 - GENERAL CONCLUSIONS

The following conclusions were drawn from this stage of the EAGRI research and development project:

1) To meet the objectives of crop forecasting using remote sensing techniques, it is essential to observe the spectral changes of the target crops during their growing season using sequential imagery.

2) The two flight experiments (1975 and 1976) showed that an effective scale change from 1:20,000 to 1:10,000 did not improve the accuracy of crop identification. Consequently, the increase in the quantity of imagery at 1:10,000 only increases time and operational costs.

3) Quantitative techniques and digital processing of LANDSAT data in the IMAGE-100 System permitted rapid automatic interpretation and areal estimations. These results were encouraging and will lead to the increasing use of the IMAGE-100 System for the analysis of large agricultural regions.