Reprinted from

Eighth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Crop Inventory and Monitoring

July 7-9, 1982

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

COMPUTER-AIDED INVENTORY OF SUGAR CANE IN MEXICO

S. ARREDONDO G., J.A. VALDES A.,
J.I. MIRANDA V.

Direccion General de Geografia
del Territorio Nacional
Mexico, D.F.

I.- ABSTRACT

This paper presents the results of a research project on the feasibility of making sugar cane inventories through computer-aided analysis of LANDSAT MSS data. The methodology followed is described and a preliminary output product is presented. The project was conducted by the Direccion General de Geografia del Territorio Nacional (DGGTN). Which is the official cartographic agency of Mexico. It was concluded that LANDSAT data can be successfully used in a computer based method for the detection and quantification of sugar cane crops.

II.- INTRODUCTION

It is the purpose of this paper to briefly describe some early results of a research project on the applications of machine processing of LANDSAT MSS data to sugar cane crop inventories. The project was conceived as part of a broader plan of the DGGTN (National cartographic agency) for the applications of computer-aided remote sensing techniques to natural resources evaluation. Other research projects are currently being developed which seek to determine the potential of remote sensing in providing information for forest monitoring, land use-land cover mapping(1), and desertification process evaluation.

III.- BACKGROUND

Sugar cane (Saccharum officinarum L) is an important crop in Mexico. Aproximately 474,000 hectares were dedicated to its cultivation during the past year (1981) The activities related to its cultivation represent the principal income of nearly two hundred thousand families in many tropical and subtropical regions where it is cultivated.

Predictions of the yield, storage and distribution of products are made based on production forecasts and estimates of the area dedicated to sugar cane crops, collected through personal inquiries to growers, and estimates obtained in the field by supervisors of the sugar cane processing industry. Such a way of obtaining information has proven not to be the more reliable and timely source of primery data, since it depends on personal experience, and collection and summarizing of information is slow.

Remote sensing techniques have demonstrated their usefulness for obtaining information about the condition and areal extent of crops(2).

Experience in Mexico has shown that the acquisition of information through interpretation of aerial photographs, although a very precise and reliable method, is a very time consuming and expensive process.

More recently, some works on computer processing of satellite multispectral images have demonstrated the feasibility of obtaining agricultural data using multispectral images(3). Although the information provided by this process is not as complete and precise as that which aerial photographs can provide, it is faster. less expensive and easier to summarize, thus presenting a compromise between the desired information level and the timeliness of the data. This is more evident in the case in which repetitive coverage of extensive areas is needed.

IV.- OBJECTIVES

The Remote Sensing Department of the DGGTN initiated a crop inventory project using remote sensor digital data. This project has a dual purpose: first, it is intended as a first insight into the general problem of obtaining information about crops. Second, it was conceived as an attempt to fill the information requirements about areal extent of sugar cane management, for national and international commercial policy planning.

Sugar cane was selected as a logical starting point since it presents spectral, temporal and spatial characteristics well suited to the detection and quantification by multispectral scanning and pattern recognition techniques(4).

V.- RESEARCH PROCEDURES

A.- SELECTION OF TEST SITE.

The following criteria were used in selecting a test site:

* Existence of a large and homogeneous sugar cane crop ground cover condition.
* Advanced knowledge of the area.
* Availability of LANDSAT input data.
* Availability of DGGTN thematic map information.

The test site for the project is located in the subtropical southern part of the state of Jalisco (19°30' to 19°45' N; 103°10' to 103°25' W). In this site, agricultural lands are almost entirely dedicated to sugar cane cultivation, sparsely intermingled with other crops such as corn and beans. The natural vegetation of the area is characterized by deciduous and evergreen forests and shrubs, other land cover conditions are uncultivated fields, grass lands, small towns and some water reservoirs.

B.- LANDSAT INPUT DATA.

For this project, the following satellite data were used:

March 22,1981 LANDSAT ID 8222511637X0

The date selected was dictated mainly by the availability of raw data, since its acquisition is almost precluded the rest of the year due to extensive cloud cover conditions.

C.- DATA PROCESSING SYSTEM.

The LANDSAT digital data for this project were processed at the image processing facility of the DGGTN in Mexico City. DGGTN's computer system includes two PERKIN-ELMER 8/32 minicomputers and peripherical devices, two COMTAL image display devices and two off-line VERSATEC electrostatic printer/plotters. The computer software utilized is the Earth Resources Laboratory Applications Software ELAS), developed at NASA/NSTL/Earth Resources Laboratory(5).

D.- LANDSAT DATA ANALYSIS.

The LANDSAT data analysis involved the software system previously mentioned, which, among other capabilities, uses the maximum likelihood ratio pattern recognition technique for analyzing remote sensing data. The procedure involved: 1) Preprocessing the data; 2) Selecting a subset of spectral bands; 3) Developing statistical training classes for the maximum likelihood algorithm; 4) Classifying each data pixel within the test site; 5) Labeling and grouping spectral classes; 6) Geo-referencing the categorized information and 7) Estimating areas and assessing accuracy.

D.1.- Preprocessing the data.

The processing of the LANDSAT digital data began by reformatting the data into a disk file compatible with the ELAS software. Since the CCT's obtained for the project were EROS PM's geometrically and radiometrically corrected prior to data dissemination, this was the only preprocessing step required for this project.

D.2.- Selecting a subset of spectral bands.

Bearing in mind possible redundancies on the data, and the need to optimize computer costs, it was decided to select a subset of spectral bands of the LANDSAT original set of four to be used in subsequent processing steps. It has been demonstrated that the marginal increase in the information generated using a full set of spectral bands may not warrant the increase in computer time(6).

The selection process consisted in an iterative non-supervised collection of class statistics from a representative subscene of the test site, varying each time the subset of bands used. The selection criteria were the number of classes obtained and the percentage of pixels used to generate them, assuming that the more spectral classes obtained and the more pixels involved by using a particular subset of bands, the better this subset spectrally characterizes the scene. With this rationale, the subset of spectral bands selected was that of bands 4, 5 and 7, which was found to best represent the spectral variability within the sugar cane crop. Table 1 shows the results of these analyses.

Table 1.- Results of classifications using different spectral bands subsets.

Bands Subset	Number of Classes	Sampling Percentage
1, 2, 3, 4	27	2.78
1, 2, 3	26	11.90
1, 2, 4	34	13.44
2, 4	27	8.25

D.3.- Developing training statistics.

As an attempt to aproximate this research project to an operational one, it was decided to use a non-supervised approach for developing training statistics, in order to rely less heavily on ground truth data in this processing step.

Training samples were selected automatically using SEARCH, an ELAS program that uses homogeneity of spectral reflectance within a defined neighborhood of LANDSAT pixels as a selection criterion.

As with any project, an attempt was made to achieve a classification accuracy as high as possible. A high accuracy can be attained if a given classification provides enough statistical classes to account for all the cover types relevant to the analysis(7). If not attained program input parameters may be changed and a new classification attempted, untill a "good" set of classes is obtained.

The non-supervised classification produced 42 statistically diferent classes, which were considered to best represent the cover types of interest and their spectral variation within the test site.

D.4.- Classifying each data pixel.

After the statistical collection process the raw data were classified into each one of the 42 classes produced by SEARCH utilizing the principle of maximum likelihood. The ELAS program for maximum likelihood classification (MAXL) has two options: 1) it can be used as a Gaussian classifier; and 2) it can be used as a Bayesian classifier. In this project the second option was used since the apriori values calculated by SEARCH for each class, based on the relative abundance of all classes, was considered to be representative inasmuch as the wholw data file was used in the statistics calculation process.

D.5.- Labeling and grouping spectral classes.

After execution of the MAXL classification program, the spectral clsses so produced were labeled and grouped into information categories. These processes were accomplished by displaying the classified data on the color video display image processing system, which shows the spatial distribution of the spectral classes by color coding, and ascertaining the identity of each class by visual comparison with aerial photographs, and land use/cover maps. Statistical output for each class derived by SEARCH, such as mean reflectance for each band, and a two-space plot of bands seven and five also aided in this process. Once all of the spectral classes were labeled, the classes heaving the same identity were grouped into nine information categories. The information categories considered were: 1) Forest; 2) Shrub; 3) Grassland; 4) Sugar cane crop; 5) Non-irrigated soils; 6) Irrigated soils; 7) Urban area; 8) Water and 9) Barren. Figure 1 shows an electrostatic plot of a categorized data file. Although it was possible to separate several spectral classes which were found to represent diferent growth stages of the sugar cane crop in this paper they are presented as a unique category for the sake of simplicity in presentation.

D.6.- Geo-referencing the categorized information.

After categorizing the data, they were geo-referenced into the Universal Transverse Mercator (UTM) coordinate system to reduce the geometric distortion inherent in the LANDSAT data, and to refer them to a standard geographic system. This was accomplished by locating several points within the data through use of the image processing system, and locating the corresponding points on DGGTN topographic maps. The element and scan line coordinates, so that an equation could be derived to convert all of the digital data to a UTM coordinate system.

D.7.- Estmating areas and assessing accuracy.

Area estimates for each of the nine categories were determined by counting the number of pixels from the entire 72,800 ha test site that were classified into each category. The area measurement in hectares (ha) was calculated as the total pixels classified into each category multiplied by 0.250 ha/pixel. Proportion estimates for each category were also obtained. Table 2 shows the results of these estimates.

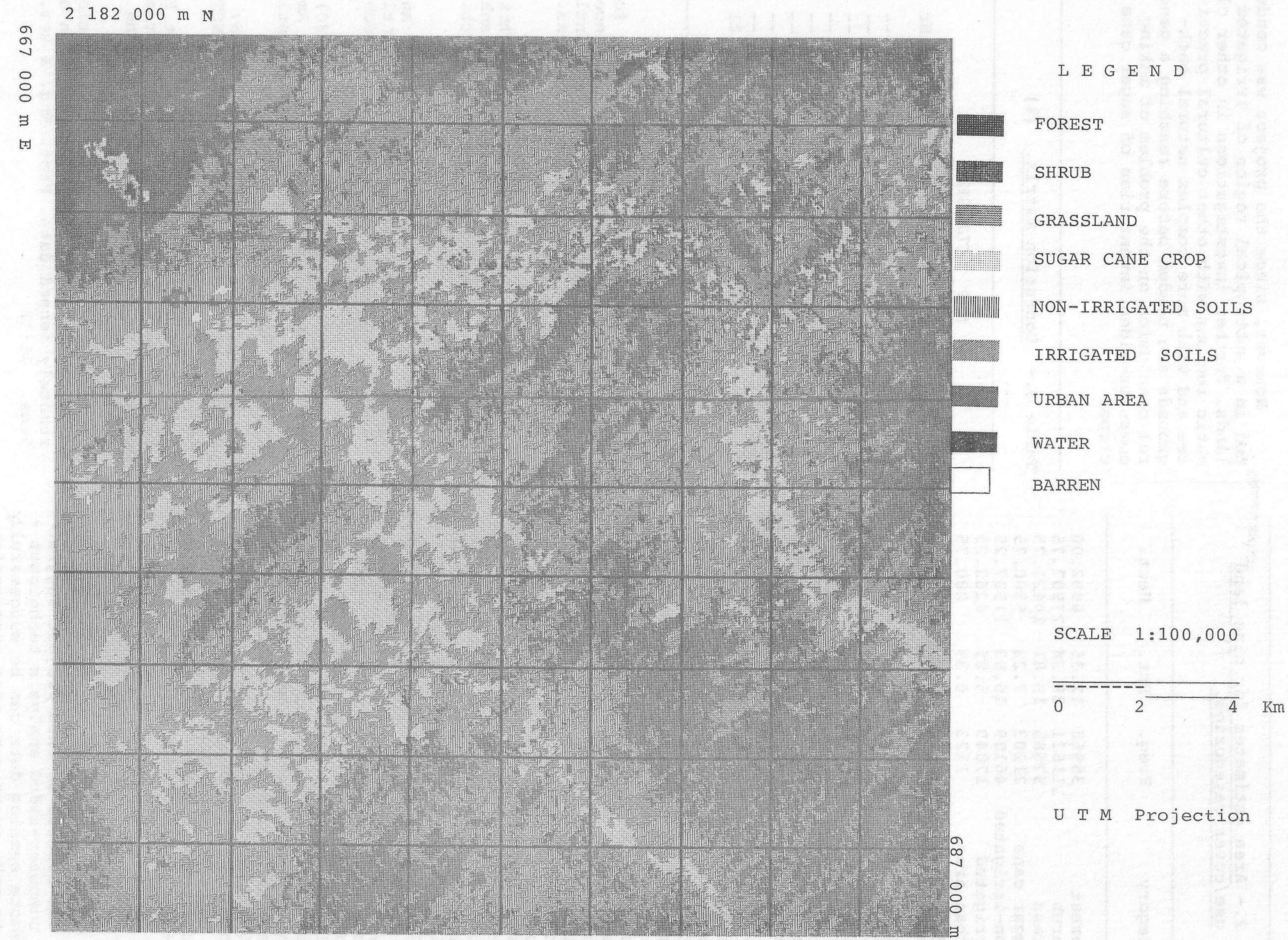

Figure 1.- A reduced portion of a preliminary cartographic product of the distri bution of sugar cane crops.

Table 2.- Area estimates for nine land use/cover categories.

Category	Freq.	Pct.	Hect.
1.- Forest	35968	12.35	8992.00
2.- Shrub	111631	38.34	27907.75
3.- Grass	57685	19.81	14421.25
4.- Sugar cane	21203	7.28	5300.75
5.- Non-irrigated	46109	15.83	11527.25
6.- Irrigated	17049	5.85	4262.25
7.- Urban area	1123	0.39	280.75
8.- Water	85	0.03	21.25
9.- Barren	347	0.12	86.75
TOTAL	291200		72800.00

After area estimation the classified image was assessed for accuracy by comparison with photointerpreted reference data. The photography was acquired on March 24 of 1981 in support of this project. The coverage consisted of 9 x 9 - inch black and white-infrared prints at a scale of 1:20,000. These prints were photointerpreted to determine the area occupied by sugar cane crops. The interpretations were input to the computer by digitizing polygons picked from the photographs. After the digitizing was checked, a file having data values corresponding to the category numbers was generated and registered to the classified LANDSAT image. The accuracy of the classification was then determined by a program which creates contingency tables of the two input images for comparison. The confusion matrix resulting from this process is shown in table 3. Notice the relatively high accuracy in sugar cane classification, which confirms the success of this project. The confusions between sugar cane and other ground cover conditions were mainly due to sugar cane in early growth stages, since more advanced stages, mature and pre-harvest, have well-diferentiated spectral response patterns. Some confusions with forests situated on slopes oriented towards the sun were also observed, but they were removed in a post-classification refinement process using a spatial-spectral reclassifier.

VI.- CONCLUSION

The results of this project suggest that computer-aided analysis techniques of remote sensing data can be successfully used to make sugar cane crop inventories.

However, since the project was conducted in a subtropical region of irrigated lands, further investigations in other climatic regions with other cultural practices and with more complex natural backgrounds are needed before reaching a general conclusion on the problem of making computer-aided inventories of sugar cane crops.

Table 3.- Confusion matrix. (%)

	MAP CLASSES								
	Fr	Sh	Gr	Sc	Ni	Ir	Ua	Wt	Br
Fr	98	1	1	--	--	--	--	--	--
Sh	--	59	33	--	8	--	--	--	--
Gr	--	20	75	--	5	--	--	--	--
Sc	--	--	--	83	3	14	--	--	--
Ni	--	--	10	--	90	--	--	--	--
Ir	--	--	--	9	2	89	--	--	--
Ua	--	2	11	--	60	2	25	--	--
Wt	--	33	10	--	--	--	--	57	--
Br	--	--	--	--	74	5	--	--	21

Percent correct overall = 80.69 %

VII.- REFERENCES

1. Camara, R.A. (1981). A methodology for updating agricultural forest and range resource inventory in Mexico.7th. Intl. Symp. Machine Processing of Rem. Sens. Data. LARS/Purdue Univ.

2. Bauer. M.E. (1977). Crop identification and area extimation over large geographic areas using LANDSAT MSS data. LARS Tch. Rep. 012477. LARS/Purdue Univ.

3. Gusman, A.; et al.(1976). Computer analysis of LANDSAT images for crop identification in Mexico. Comunicaciones tecnicas. PR -76-17. IIMAS/UNAM. Mexico.

4. Bartolucci, L.A. & Wagner,T.W. (1980). Remote sensing applications in Latin American and the Caribean. Proc.14th. Intl. Symp. on Rem.Sens. of Environ. ERIM.

5.- Graham, M. H., et al. (1980). ELAS: A Geobased Information System. NASA/NSTL/ ERL. Report 183.

6.- Thiruvebgadachari, S. (1981). Cost Effective computer processing of airborne scanner data for regional level mapping. Phot. Eng. & Rem. Sens. 47:12.

7.- Hixson, M., et al. (1980). Evaluation of several schemes for classification of remotely sensed data. Phot. Eng. & Rem. Sens. 46:12

Sergio Arredondo G. is a staff member of the Remote Sensing Department of the Direccion General de Geografia del Territorio Nacional. He is a Biologist from the National University of Mexico, where he has participated in ecological surveying projects and coursed studies on computer science. He also has colaborated in Remote Sensing projects with the applied mathematics and sistems research institute of the National University of Mexico. He is a member of the American Society of Photogrammetry and of the Sociedad Mexicana de Fotogrametria, Fotointerpretacion y Geodesia.

José A. Valdés A. is a staff member of the Remote Sensing Department of the Direccion General de Geografia del Territorio Nacional. He is an Agronomist from the Instituto Tecnologico y de Estudios Superiores de Monterrey, campus Queretaro, where he presented a Bachelor's thesis on the applications of Remote Sensing to agronomic surveys. Presently he is planning to obtain an M.S. degree in agronomical Remote Sensing at Purdue University. He is a member of the Sociedad Mexicana de Fotogrametria, Fotointerpretación y Geodesia.

Jose I. Miranda V. Agronomist (1970), University of Michoacan, 10 years experience as photointerpreter of land capability in the General Direction of Geographic Studies of the National Territory (DGGTN) in Mexico. Presently Investigator Analyst in Remote Sensing in the same agency.