SAMPLED LANDSAT DATA FOR CROP AREA ESTIMATIONS

Sherry Chou Chen
Instituto de Pesquisas Espaciais - INPE
C. P. 515, 12201, São José dos Campos, SP
Brazil
WG4

ABSTRACT

LANDSAT Thematic Mapper provides data with improved spectral, spatial and radiometric resolutions than that of the MSS sensor. However, these improvements lead to increased data volume and consequently involve more computer time in data analysis. This problem becomes even more acute when the analysis has to be carried out for a large geographic area within a rigid time table such as the cases of crop forecasting studies. In this paper an experiment was carried out to test if crop classification results were not significantly different between the pixel-by-pixel and the 2x2 systematic sampling schemes. Five test areas of 512x512 pixels which represent most summer crops of Paraná state were classified pixel-by-pixel and every other line and column using LANDSAT-TM data of Jan 19, 1985. The RMS difference for 25 informational classes of the five test areas was 0.13%. Comparison was also made for soybeans, corn and sugarcane of 15 municipalities. The cultivated soybean area was 23.59% and 23.36% for the pixel-by-pixel and the systematic sampling approach respectively. The official statistics provided by IBGE is 25.53% or 128,503 ha. Although results of unitemporal classifications of corn and sugarcane were not satisfactory, estimates obtained from the pixel-by-pixel and the sampling approaches were similar. The highly correlated classification results indicate that crop classification for Paraná State can be carried out by classifying the 2x2 systematically sampled LANDSAT data, in this manner, accurate area estimates can be obtained with 1/4 computer time comparing to the pixel-by-pixel classification scheme.

1. INTRODUCTION

The better spectral, spatial and radiometric resolutions of LANDSAT-TM data in comparison to that of the MSS improve the potential of remotely sensed data for natural resources studies. However, the much larger data volume of TM sensor (i. e. 343 mb vs 30 mb of MSS) causes problems for analysts in finding an efficient way to handle them; in other words to extract maximum information within a minimum computer time.This

problem becomes even more pronounced when results have to be obtained within a very short time frame for a large geographic area such as the cases of crop forecasting studies.

To give an idea about the TM data volume that has to be analyzed in an operational crop forecasting system, the Paraná state is taken as an example. Paraná is the most important Brazilian agricultural state and encompasses $\tilde{=}$ 200,000 km^2. Forty-three TM quadrant scenes cover the whole state. In order to obtain maximum information about principal summer crops of this state, ground information of sampled segments have to be collected during the second fortnight of January when most crops have been planted. The produced agricultural statistics should be available for decision-making no later than the end of April. Consequently, there is a three-month interval for digital data analysis which is a very tight time table even considering that 3 TM channels of unitemporal data are used. This is because that activities in this three-month interval include field data screening, editing, segment digitalization, classifier training and ultimately the proper classification. Viewing on the problems mentioned above, there are two possible alternatives to alleviate the data load of a digital analysis:

1) using a subset of the available TM channels for classifications,

2) substituting the classification of a systematically sampled LANDSAT data for the pixel-by-pixel classification.

For the first alternative a feature selection study was carried out and it was concluded that there was a three-dimensional spectral space for crop types of Paraná state (Chen et al. 1986). The objective of this study was to test the second alternative - if results obtained by classifying sistematically sampled LANDSAT data were not significantly different from that obtained by the pixel-by-pixel classification.

2. MATERIAL AND METHODS

LANDSAT TM data with path/row annotation of 223/76D were acquired on Jan. 19, 1985. Five subscenes of 512x512 pixels where ground information were available were used as small scale study sites. Cover types in these areas included soybeans, corn, sugarcane, coffee, pasture, forest and plowed soil. The maximum likelihood decision rule was used for classification with training statistics obtained from the feature selection study mentioned previously. After the pixel--by-pixel classification, each study site was sampled every

other line and column (i.e. 2x2 systematic sampling or 1/4 classification) and classified again using the same training statistics. Comparisons of classification results of these subscenes showed the sampling effect on crop area estimates at small scale areas of 15x15 km^2. In order to evaluate better the systematic sampling effect on a larger geographic area, 15 municipalities which are physically agglomerated were also classified. These 15 municipalities included Floresta, Florai, Maringá, São Carlos do Ivaí, São Jorge do Ivaí, Doutor Camargo, Paiçando, Ourizona, Ivantuba, Jussara, Japura, Cianorte, São Tomé, Terra Boa and Engenheiro Beltrão. Crop proportions of soybeans, corn and sugarcane of these municipalities obtained from the pixel-by-pixel and the 1/4 classification approaches were then compared. Official statistics published by the Brazilian Institute for Geography and Statistics (IBGE) for the same municipalities were also presented as references.

3. RESULTS AND CONCLUSIONS

Comparisons of results obtained by using the pixel-by-pixel and 1/4 classifications for the five study sites are presented in Table 1. Twenty-five informational classes were found in these areas and the relative differences of estimated cover types obtained from these two classification approaches varied from 0 to 2.82%. The root-mean-square (RMS) difference of these 25 informational classes was insignificantly small (0.13%). For the large scale evaluation, the classified soybeans, corn and sugarcane of the 15 municipalities are presented in Table 2. It can be observed that no matter which classification approach was used, estimated crop proportions were almost equal. Comparing LANDSAT estimates to IBGE data it is concluded that an accurate soybean estimate can be obtained in a cost-effective manner by using unitemporal LANDSAT data. The imprecise estimates of corn and sugarcane are understandable. In Paraná state corn has a wide range of planting date, it may start from the end of August to the end of December. Thus, it is very difficult in using one satellite pass to successfully classify all the corn plantations. This is also the case for sugarcane. Nevertheless, these less precise estimates of corn and sugarcane caused no problem in drawing valid conclusions in this study. The similarities of classification results obtained from the pixel-by-pixel and the 2x2 sampling approaches demonstrated that it is recommended to classify each other line and colume of TM data to reduce time and cost of the full frame pixel-by-pixel classification in large geographic areas.

TABLE 1

COMPARISONS OF CLASSIFICATION RESULTS OF COVER TYPES (%) OBTAINED BY USING THE PIXEL-BY-PIXEL AND THE SYSTHEMATIC SAMPLING APPROACH

CLASS	CLASSIFICATION APPROACH		DIFF.
	PIXEL-BY-PIXEL	1/4 SAMPLING	
SITE 1 Soybeans	33.83	33.72	0.11 (0.33)*
Corn	3.17	3.16	0.01 (0.32)
Sugarcane	5.80	5.88	-0.08 (-1.38)
Coffee	2.60	2.60	0 (0)
Pasture	9.50	9.50	0 (0)
Natural forest	15.10	14.92	0.18 (1.19)
Reforestation	3.39	3.48	-0.09 (-2.65)
Plowed soil	0.75	0.76	-0.01 (-1.33)
SITE 2 Soybeans	14.95	14.96	-0.01 (-0.07)
Corn	8.90	8.68	0.22 (2.47)
Sugarcane	9.50	9.40	0.10 (1.05)
Pasture	3.82	3.80	0.02 (0.52)
Natural forest	12.48	12.54	-0.06 (-0.48)
Reforestation	5.94	5.88	0.06 (1.01)
SITE 3 Soybeans	48.08	47.76	0.32 (0.67)
Corn	0.63	0.64	-0.01 (-1.59)
Pasture	11.12	11.08	0.04 (0.36)
Plowed soil	1.60	1.56	0.04 (2.50)
Reforestation	4.15	4.08	0.07 (1.69)
SITE 4 Soybeans	51.24	50.88	0.36 (0.70)
Corn	30.52	30.48	0.04 (0.13)
Plowed soil	4.96	4.82	0.14 (2.82)
SITE 5 Soybeans	32.29	32.24	0.05 (0.15)
Sugarcane	5.79	5.64	0.15 (2.59)
Water	41.10	41.08	0.02 (0.05)

* Relative difference (%) = $\frac{\text{Diff.}}{\text{Pixel-by-Pixel estimate}}$ x 100

TABLE 2

RESULTS OF THE PIXEL-BY-PIXEL AND THE 1/4 CLASSIFICATION APPROACHES FOR 15 MUNICIPALITIES OF PARANA STATE

CROP TYPE	CLASSIFICATION APPROACH		IBGE
	PIXEL-BY-PIXEL	1/4 CLASSIFICATION	
Soybeans	23.59%	23.36%	25.53%
Corn	4.99%	4.81%	9.65%
Sugarcane	7.48%	7.36%	10.13%

4. REFERENCES

CHEN, C.S.; BATISTA, G.T.; TARDIN, A.T.; 1986. LANDSAT TM BAND COMBINATIONS FOR CROP DISCRIMINATION. Proc. 7th International Symp. on Remote Sensing for Forest Resources Development and Environmental Management. ISPRS Comm. VII. Enschede, The Netherlands.