A CROP AREA ESTIMATOR BASED ON CHANGES IN THE TEMPORAL PROFILE OF A VEGETATIVE INDEX

James H. Smith and Daniel B. Ramey, Lockheed Engineering and Management Services Company, Inc.[1]

ABSTRACT

Most current crop area estimators, based on remotely sensed data, require the classification of either fields or picture elements (pixels) into crop types. This paper details current research into methods which estimate the change in crop proportion in a scene from one year to another, without requiring that individual fields or pixels be labeled as crop types. Instead, pixels are classified as vegetated or not vegetated, and the proportion of vegetated pixels in the scene is plotted as a function of time for each of two years. The plots are smoothed via polynomial regression, and the vertical distance between the curves (profiles) forms the basis of profile change methodology. Results demonstrating the feasibility of using the technique are presented.

INTRODUCTION

The United States is one of the world's major exporters of cereal grains and livestock feed grains. Thus, there is considerable interest in accurate and timely estimates of world grain production. The U.S. Department of Agriculture provides reliable estimates of U.S. grain production based on in-season statistical surveys; however, comparable estimates are not available for most foreign areas, with most foreign-provided information unavailable until well after the grain harvest. Since the economic value of production estimates declines with time, these after-harvest figures have a reduced value to U.S. consumers or producers. Thus, there has been an attempt to estimate grain production in real time using remotely sensed data, such as are obtained by the Landsat series of satellites.

One well-known approach is to obtain Landsat measurements for a sample of scenes (called segments), each measuring 5 by 6 nautical miles, and estimate the proportion of the crop of interest in each segment. These estimates are then aggregated to obtain area estimates at the regional level and, finally - given the estimated yield for the crops and varieties of interest - the total grain production estimate for the area of interest.[2]

In this paper, the authors explore an inexpensive alternative approach which estimates the annual change in acreage of the crop of interest. This approach employs simple regression models to produce either segment level or regional estimates of the annual change. The models are used to represent the profile traced out by the percent of vegetated pixels in each sampled segment. While the models used are perhaps oversimplified, the results obtained indicate that the approach is both feasible and more economical than the traditional one. This paper presents only the segment level model.

LANDSAT DATA

The Landsat system (formerly ERTS), in operation since 1972, and consisting of one or two satellites, collects data over the global land areas. Each satellite can collect data over a given site once every 18 days. (Due to overlaps in coverage between orbits, certain areas will be covered on consecutive days.) Each observation for a particular site is called an acquisition. In most agricultural applications, about half of all potential acquisitions are lost because of cloud cover. Even with two operational satellites, the minimum time between acquisitions is 8 days; often there are gaps of several weeks or months. Additional loss of acquisitions occurs in the process of registration, the spatial alignment of the segment acquisitions collected throughout the season (see refs. 4 and 5). While most current crop area estimation techniques require accurate registration, the proposed technique does not, avoiding the additional loss of scarce acquisitions.

Each Landsat is equipped with a multispectral scanner (MSS). This device contains sensors which measure the intensity of radiation in each of four wave bands:

Band 1 - 0.5 to 0.6 μm
Band 2 - 0.6 to 0.7 μm
Band 3 - 0.7 to 0.8 μm
Band 4 - 0.8 to 1.1 μm

Bands 3 and 4 are infrared wavelengths and record highly correlated responses in most agricultural regions. Therefore, we discard band 3. Band 1 corresponds to visible green wavelengths, while band 2 corresponds to visible red wavelengths. The MSS has a resolution of approximately 1.1 acres; that is, data are recorded for pixels which are about 1.1 acres in area.

In order to display the data collected in the three bands, photographic products can be produced by using a false color method (ref. 6). While the color assignments are arbitrary, the standard assignments are blue for band 1, green for band 2, and red for band 4. Thus, visible green in the scene will appear blue on the false color product, while visible red will appear green. Any objects which have high relative reflectance in the measured infrared wavelength (such as healthy vegetation) will appear red on the false color image. On the other hand, objects which have high relative reflectance in band 1 (such as bare soil) will appear blue on the imagery.

VEGETATIVE INDICES

In order to facilitate the automatic classification of pixels as either vegetated or nonvegetated for a particular acquisition, a number of functions of the four bands have been proposed. These functions are generally referred to as vegetative indices, since the intent is that a vegetated pixel have a large value and a nonvegetated pixel a small value.

While several vegetative indices have been proposed, probably the most frequently used is Kauth greenness (ref. 7), which is approximately the second principal component of the four bands. The authors have chosen, however, to work with an alternative index which is related to the false color photographic imagery. The index is calculated as follows (ref. 8):

1. Normalize data for each band by dividing the band value for each pixel by the mean band value of all pixels in the scene.
2. Compare the normalized values for each pixel; if band 4 is greater than or equal to band 1, and band 1 is greater than or equal to band 2, call the pixel 'red.' (If false color imagery were produced from the normalized data, the pixel would appear red.)
3. If the pixel is 'red,' set the index to 1. Otherwise, set it to 0.

We will classify a pixel as vegetated if it has an index value of 1; otherwise, we will classify it as nonvegetated.[3] In the text following, we will use 'red pixel' to refer to a pixel which has an index of 1.

TEMPORAL PROFILES

In many important crop regions, the only vegetation seen in the early spring consists of evergreen trees, shrubs, some grasses, and weeds. As the crop season progresses, the winter grains planted the previous fall emerge from dormancy, and then spring grains emerge from their spring planting. Somewhat later, summer crops such as corn and soybeans emerge. These crops are generally harvested in the order in which they emerged, so that by some date in the fall, the remaining vegetation is essentially the same as in the early spring.

After plowing and before emergence of a crop, only bare soil is visible; thus, the area planted in that crop appears blue on false color imagery which was obtained before emergence. After emergence, less of the soil is visible, and the area begins to appear red in the imagery. As the crop reaches the peak of its vegetative growth, the area planted in that crop becomes bright red. As the crop matures and begins to senesce, it turns orange in the false color imagery. Finally, after harvest, the bare soil again becomes visible. Each crop has a characteristic pattern (this pattern is referred to as a profile) in the sequence of false color images. Virtually every proportion estimation technique attempts to use these profiles to classify the segment pixels as to crops.

The term temporal profile is used in this paper to denote the path traced out over time by the proportion of red pixels in a segment. Figure 1 shows plots of the temporal profile for segment 1924 for the 1978 and 1979 crop years. This segment, located in North Dakota, is predominantly spring wheat. The general shape of the temporal profile is easily seen in figure 1, particularly in the plot for 1979. Ideally, we would wish to locate the maximum of each curve, which corresponds to the maximum percentage of red pixels, giving an estimate of the percentage of vegetated pixels. We could then base a change estimate on the difference in the heights of the two plots.[4]

Because of the infrequent Landsat coverage, there are too few data points available to fit a curve to these data and estimate accurately its maximum (again, see figure 1.) Since more frequent coverage is unlikely, we must find some other approach to improve our estimate. One way would be to use data from more than one year to estimate the profile.

MULTIYEAR PROFILES

If the change in the shape of a segment's temporal profile from year to year were only a vertical shift due to an increase or decrease in the percentage of the crop of interest, then the parameters of the curve could be estimated using data from two or more years. However, planting dates, and to some extent, days from planting to harvest, vary considerably from year to year, introducing a horizontal shift in the profiles as well as a change in the horizontal scale. These effects are due primarily to differences in weather from year to year, since farmers tend to plant and harvest as early as possible, while crop maturity depends on temperature, the amount of available moisture, and solar radiation.

To put multiyear data on the same temporal scale, a simple measure of crop maturity may be constructed based on cumulative growing degree-days (ref. 9). Then the profile can be modeled as a function of this measure rather than as a function of elapsed time. This removes the effects of most of the weather conditions mentioned above.

GROWING DEGREE-DAYS

There are several methods of calculating cumulative growing degree-days; we chose the 50/86 method. We will denote this variable by G; it is computed for day n by:

Let U_j be the maximum temperature of day j (1)

$$\text{Let } t_j = \begin{cases} 0 \text{ if } U_j < 50 \\ U_j - 50 \text{ if } 50 \leqslant U_j \leqslant 86 \\ 36 \text{ if } U_j > 86 \end{cases} \qquad (2)$$

$$G = \begin{cases} 0 \text{ if } n < 60 \\ \sum_{60}^{n} t_j \text{ if } n \geqslant 60 \end{cases} \qquad (3)$$

Note that G is accumulated only from March 1 (day 60). Use of G implies these assumptions:

1. The temporal profile depends on the weather only through the growing degree-days.
2. The weather before March 1 is irrelevant.
3. No growth takes place if the maximum temperature is below 50° F; any degrees above 86 do not contribute to plant growth.

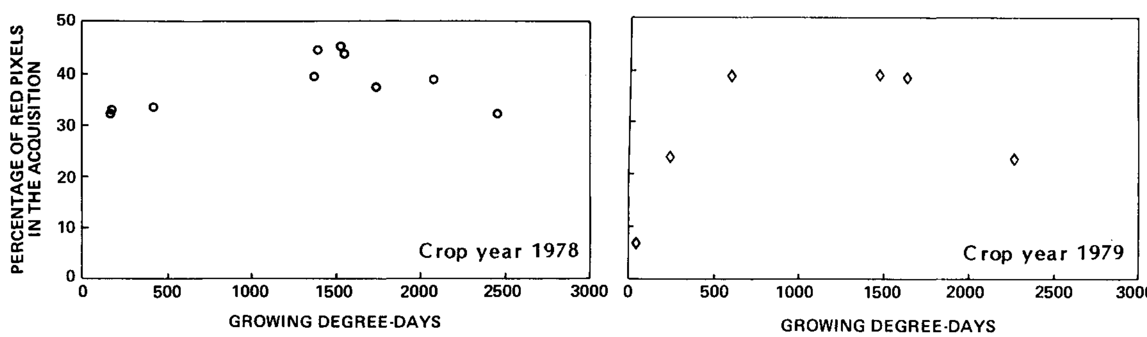

Figure 1.- Temporal profiles for segment 1924, crop years 1978 and 1979.

Thus, we assume the maximum value of the percentage of red pixels occurs at the same value of G for each of two or more years. This assumption is reasonable, provided that a major shift between crop groups (spring, winter, summer) does not occur. Fortunately, such shifts rarely occur between two consecutive years. (Generally, the major crop groups grown in an area are determined by natural conditions, such as soil and climate. Further, there is usually a large investment in infrastructure such as grain elevators, combines, and so forth, making sudden large shifts in crop percentages unlikely.) Figure 2 shows a plot of the percentage of red pixels versus G for segment 1924, for crop years 1978 and 1979.

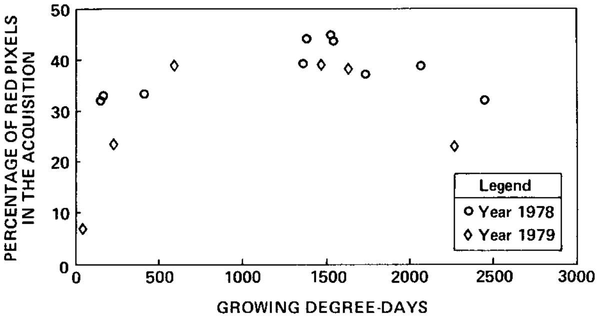

Figure 2.- Combined profiles for segment 1924, crop years 1978 and 1979.

A PROTOTYPE MODEL

To determine the feasibility of a temporal-profile-based change estimation procedure, the following model was applied successively to the data sets of available acquisitions for 10 predominantly spring small grains segments in North Dakota and South Dakota for 1978 and 1979:

$$P = \alpha Y_i + \beta_o + \beta_1 G + \beta_2 G^2 + \beta_3 G^3 \quad ; \quad i = 1,2 \qquad (4)$$

where P represents the percentage of red pixels in an acquisition; Y_i is zero for the known year (i = 1) and $Y_i = 1$ for the target year (i = 2); and G is the growing degree-day measurement for the acquisition. The coefficients $\beta_o, \cdots, \beta_3$ are themselves of no interest, but the estimate of the coefficient α , $\hat{\alpha}$, is treated as an estimate of the year-to-year change in the percentage of spring small grains. Figure 3 depicts graphically equation (4), as fit to the data of segment 1924 for crop years 1978 and 1979. Some of the shortcomings of this model are obvious at a glance:

1. The 'true' underlying profiles are probably not separated by a constant amount through the growing season (in fact, they must converge as G moves in either direction away from the peak.)
2. The general shape of the model-derived curve is clearly inadequate to describe the complexities of the 'true' profiles.
3. The number of degrees of freedom is still quite small for fitting a polynominal to the profile (note, however, that we are primarily concerned with the estimate of alpha, the year effect, and not with the betas themselves)

Table 1 is a summary of the results obtained when the technique was applied to 10 segments. The model used was actually a variant of (4) and is given by:

$$P = \alpha Y_i + \beta_o + \beta_1(2000 - G)G + \beta_2(3{,}000{,}000 - G)G^2 \quad ; \quad i = 1,2 \qquad (5)$$

Model (5) is constrained to have zero derivative at 1000 growing degree-days, which is the expected time of peak greenness for spring small grains. Very similar results were obtained using the model (4), with slightly higher sample variance and estimated bias (mean error). Table 2 is a compilation of summary statistics.

TABLE 1
Model (5) applied to 10 spring small grains segments in North Dakota and South Dakota

Segment number	State	County	Observed change, %*	Estimated change, %*	Error, %
1924	ND	La Moore	-4.53	-5.32	-0.79
1920	ND	Sioux	-1.04	-2.44	-1.40
1918	ND	Grant	-1.82	-5.82	-4.00
1676	SD	Brule	+.74	-2.85	-3.59
1658	ND	Dickey	-12.65	-9.86	+2.79
1485	SD	Dewey	+2.82	-4.68	-7.50
1461	ND	Pierce	+3.74	+3.60	-.14
1457	ND	Ward	-2.34	+.14	+2.48
1755	SD	Jerauld	+.72	-4.05	-4.77
1653	ND	Burleigh	-3.26	-.66	+2.60

*Changes represent the percentage relative to the total scene (100%). Thus, a change from 50% to 40% is recorded as a 10% decrease, not as 20%.

TABLE 2
Summary statistics

Mean observed change	-1.76%
Standard deviation of observed change	4.63%
Mean estimated change	-3.19%
Standard deviation of estimated change	3.67%
Mean error	-1.43%
Standard deviation of mean error	3.51%
Mean absolute error	3.01%
Root mean squared error	3.79%
Correlation of observed and estimated change	0.729

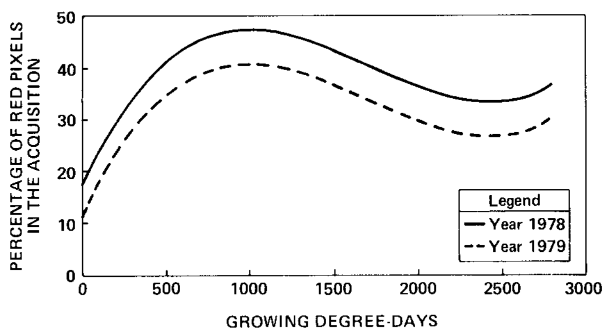

Figure 3.- Model fit to combined profiles for segment 1924, crop years 1978 and 1979.

CONCLUSIONS AND REMARKS

While in this study the technique apparently produced unbiased estimates of change, the important result is the high (0.729) correlation between the $\hat{\alpha}$'s and the observed changes. In practice, it may prove necessary to use some transformation of $\hat{\alpha}$ to produce the final estimate.

While results on a sample of 10 segments are certainly not conclusive evidence that this technology is reliable in all applications, they do indicate that change estimation based on temporal profiles is potentially viable. This technology is relatively insensitive to sample unit size, and it can be applied simultaneously to collections of segments (strata) to produce stratum-level change estimates.

Perhaps the most appealing feature of this technology is its efficiency; it requires little human intervention and a fraction of the computer time used by conventional proportion estimation procedures. Also, it may not require the precise registration from acquisition to acquisition that is required by more conventional proportion estimation procedures.

Current research activities at the National Aeronautics and Space Administration, Lyndon B. Johnson Space Center (NASA/JSC) in Houston, Texas, are centered around the use of multiple segments in constructing the profiles and estimating change. The great increase in degrees of freedom realized by this approach will allow investigation of more sophisticated profile models and estimation techniques.

ENDNOTES

[1]This material was developed under NASA contract NAS 9-15800 and prepared for the Earth Resources Applications Division of the National Aeronautics and Space Administration, Lyndon B. Johnson Space Center, Houston, Texas.

[2]For a more complete description of the approaches which have been used in the past, the reader is referred to references 1, 2, and 3.

[3]It should be noted that pixels appearing orange and purple in false color imagery are generally vegetated; for example, wheat which has turned (i.e., has begun to ripen, thus turning visibly yellow) appears orange.

[4]The underlying assumption is, of course, that the change in percentage of vegetated pixels occurs in the crop of interest. This is generally true in regions with one predominant crop, e.g., spring small grains in parts of North Dakota.

REFERENCES

1. Proceedings of Technical Sessions - The Large Area Crop Inventory Experiment (LACIE) Symposium, vols. 1 and 2. JSC-16105, NASA/JSC, Houston, Tex., Oct. 1978.

2. MacDonald, R. B.; and Hall, F. G.: Global Crop Forecasting. Science, vol. 208, May 1980, pp. 670-679.

3. Erickson, J.; Dragg, J.; Bizzell, R.; and Trichel, M.: Research Advances in Satellite-Aided Crop Forecasting [presented at Proceedings of the International Society for Photogrammetry and Remote Sensing, Commission VII, International Symposium, Toulouse, France, September 1982].

4. Grebowski, G. J.: LACIE Registration Processing. Proceedings of Technical Sessions, The LACIE Symposium, vol. 1, JSC-16105, NASA/JSC, Houston, Tex., Oct. 1978, pp. 87-98.

5. Landsat Data User's Handbook. 76 SDS 4258. NASA/Goddard Space Flight Center, Greenbelt, Md., Sept. 1976.

6. Juday, R. D.; Johnson, F.; Abotteen, R. A.; and Pore, M. D.: Generation of Uniform Chromaticity Scale Imagery From Landsat Data. Proceedings of Technical Sessions, The LACIE Symposium, vol. 2, JSC-16105, NASA/JSC, Houston, Tex., Oct. 1978, pp. 899-910.

7. Kauth, R. J.; Lambeck, P. F.; Richardson, W.; Thomas, G. S.; and Pentland, A. P.: Feature Extraction Applied to Agricultural Crops as Seen by Landsat. Proceedings of Technical Sessions, The LACIE Symposium, vol. 2, JSC-16105, NASA/JSC, Houston, Tex., Oct. 1978, pp. 705-721.

8. Cate, R. B.; Phinney, D. E.; Kinsler, M. C.; Sestak, M. L.; Hodges T.; and Dishler, J. J.: Interpretation of Landsat Digital Data Using a Cubic Color Model Based on Relative Energies. JSC-13776, NASA/JSC, Houston, Tex., Feb. 1980.

9. Weekly Weather and Crop Bulletin. National Oceanic and Atmospheric Administration/U.S. Department of Agriculture Joint Agricultural Weather Facility, Washington, D.C.