

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

[Give to AgEcon Search](https://makingagift.umn.edu/give/yourgift.html?&cart=2313)

AgEcon Search <http://ageconsearch.umn.edu> aesearch@umn.edu

Papers downloaded from AgEcon Search may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

United States Department of **Agriculture**

Economics and **Statistics** Service

Statistical Research **Division**

February 1981 ESS Staff Report No. AGESS 810223

Crop Area Estimates Using Ground-Gathered and LANDSAT Data

A Multitemporal Approach; Missouri 1979

James Weldon Mergerson

CROP AREA ESTIMATES USING GROUND-GATHERED AND LANDSAT DATA: A MULTITEMPORAL APPROACH, MISSOURI 1979. By JAMES W. MERGERSON, Statistical Research Division, Economics and Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, ESS Staff Report No. AGESS8l0223, February 1981.

ABSTRACT

This report describes a comparative study in which a unitemporal approach for obtaining crop area estimates using ground-gathered and LANDSAT data was compared to a multitemporal approach. Four channel data were used for the unitemporal approach. Eight channel data consisting of four channels from each of two dates were used for the multitemporal approach. The multitemporal data set and the two corresponding unitemporal data sets were analyzed using various procedures. Results indicated that the use of multitemporal data can significantly improve the precision of crop area estimates for corn, soybeans, and winter wheat, obtained using the unitemporal approach.

**

*
This paper was prepared for limited distribution to the research community outside the U.S. Department of Agriculture **

ACKNOWLEDGMENTS

The author would like to thank Ed Camara, Paul Cook, Tjuana Fisher, George Harrell, Richard Sigman and Sandy Stutson of the Remote Sensing Branch, Statistical Research Division, Economics and Statistics Service (ESS) for their assistance, and Kathy Reiman and Bessie Johnson for their fine word processing efforts.

CONTENTS

INTRODUCTION

The objectives of this project were to gain LANDSAT analysis experience using imagery from two different dates which cover the same land area in conjunction with ESS's conventionally gathered June Enumerative Survey (JES) ground data and to determine if this method can significantly improve the precision of crop area estimates obtained using single-date imagery in conjunction with ground data. The study area consisted of an eleven county region in Northwest and North Central Missouri.

For an explanation of general statistical methodology, ground data acquisition and processing, data processing systems hardware and software, or a genepal description of LANDSAT data, refer to the-paper-by-Hanuschak, $et.a1.¹$

This report, intended for those with some knowledge of remote sensing applications, will be useful to researchers considering the use of multi-date imagery in estimating crop areas.

LANDSAT DATA

Two scenes, with different dates, covering the same land area were selected for this project. Due to cloud cover only a portion of the scenes were analyzed. Additional information about the two scenes is provided in Table 1.

MULTITEMPORAL REGISTRATION AND DIGITAL DATA SET CREATION

Registration is the process of relating the LANDSAT row-column coordinates with map latitude-longitude coordinates by means of mathematical equations. A multi temporal LANDSAT digital data set consists of data from two different scenes taken at different dates but covering the same ground area. In creating a multitemporal digital data set, one of the original scenes is
selected as the primary scene and the other as the secondary scene. The selected as the primary scene and the other as the secondary scene. August 3 scene was selected as the primary scene and the May 14 scene was selected as the secondary scene.

Scene to map registration was performed using thirty four control points with an accuracy of 100 meters RMS (Root Mean Square Error) for the primary scene. The resulting precision calibration file was used in conjunction with the digitized segment files to predict the segment locations in the LANDSAT data. For each discrepancy between the predicted and actual location of a segment a local calibration was done. Segments were shifted as shown in Table 2.

TABLE 2: SEGMENT SHIFTS, MISSOURI 1919

 $- 4 -$

After registering the primary scene, a scene to scene overlay of the two scenes was performed with an accuracy of one-half pixel. This procedure consisted of selecting for each pixel in the primary scene the pixel in the secondary scene that most nearly represented the same area on the ground. The result was an eight channel data set; the first four channels for each pixel were from the primary scene and the second four channels were from the secondary scene.

For details concerning scene to map registration and scene to _nscene overlaying refer to the papers by Hanuschak, et.al. $^\omega$ and Ozga, et.al. $^\circ$

ANALYSIS

The data was analyzed using various analysis procedures. These procedures will be referred to as MA1, MA2, MA3, MA4, UA1, UA2, UA3, and UA4 and will also have the extensions EP and PUR. MA indicates multitemporal analysis and UA indicates unitemporal analysis. EP indicates the use of equal probabilities and PUR indicates the use of prior probabilities proportional to unexpanded reported acres. Eight channel data was used for the multitemporal analysis. The multitemporal data set consisted of four channels from each of the two dates. Four channel data was used for the unitemporal analysis. Clustering was not performed for MA4, UA3, and UA4. May data was used for UA3 and August data was used for UA4. For MA3, UA1, and UA2, categories were automatically grouped and 95% pattern transmission was retained. For a description of the automatic grouping technique refer to the paper by Craig, $e^{i\theta}$. May data was used for II_1 , . May data was used for UA2 and August data was used for UA1. Clustering was performed using Bolt, Beranek, and Newman, Data Processing Facility in Cambridge, Mass. (BBN) for MA1 and the ILLIAC IV Computer in Sunnyvale, Calif. for MA2. The analysis procedures are summarized in Table 3.

All analysis was performed using EDITOR. EDITOR is an interactive data analysis system for processing LANDSAT deta. For a description of EDITOR refer to the paper by Ozga, et.al. $\tilde{}$. Several modifications were made to EDITOR to enable processing of eight channel data. The basic analysis steps for multitemporal data are almost the same as for unitemporal data. One difference is in the creation of the Multi-Window file. Since one eight-channel tape is created for the left side of the multitemporal window and a second eight-channel tape is created for the right side, two Multi-Window files are created and combined into one file. This file contains data for each window. A window is a rectangular array of pixels.

Table A2 gives the number of training pixels available after the elimination of boundary pixels and questionable field pixels. Signatures were created for covers having 150 or more training pixels. Table $A\bar{A}$ gives the number of categories for each cover type by analysis procedure.

After creating statistics using various approaches all segment data was classified using the corresponding signatures. Boundary pixels were included

TABLE 3: ANALYSIS PROCEDURES

NlIMRRR OF CATEGORIES

 $\mathcal{L}^{\text{max}}_{\text{max}}$

for each classification. Tables A5 thru A20 contains categorization results for each procedure.

A regression estimator with JES data as the dependent variable and LANDSAT classified pixels as the independent variable was used. For the purpose of estimating crop areas, ESS's evaluation criterion is not the percent of pixels classified correctly but is how precisely the crop area is estimated for the area of interest. Maximization of the R-square values minimizes the variance of the regression estimates. Thus, the major criterion used to compare the various procedures was the respective R-squares. Another measure for procedures was the respective R-squares. Another measure for evaluation is relative efficiency (RE). The relative efficiency is the ratio of the variance of the direct expansion estimate to the variance of the regression estimate. Table 4 gives percent correct, R-square, and RE measures for the major crops in the analysis district by classifier.

Table 5 compares the best multitemporal results with the best unitemporal results.

SUMMARY AND CONCLUSIONS

A multitemporal data set and each of the two corresponding unitemporal data sets were analyzed using various procedures. Results indicate that the use of multitemporal data can significantly improve the precision of crop area estimates for corn, soybeans, and winter wheat. These results strongly suggest that the multitemporal approach be used when our objective is to estimate crop areas for both corn and soybeans at the analysis district level, if we are able to obtain cloud free imagery for two appropriate dates covering the same land area. A full state analysis using the multitemporal approach could entail additional complexity due to possible nonoverlap land area coverage of the multitemporal LANDSAT images due to clouds or satellite drift.

Registration and the creation of the multitemporal data set require an elapsed time of about four days. The analysis of multitemporal data requires about the same amount of time as the analysis of unitemporal data and the analysis steps are basically the same.

For future projects, in which the multitemporal approach will be used, scene to map registration could proceed after the acquisition of the secondary scenes. After the primary scenes are acquired scene to scene overlaying could then begin immediately.

TABLE 4: CLASSIFIER EVALUATION

 $\mathcal{L}^{\text{max}}_{\text{max}}$, where $\mathcal{L}^{\text{max}}_{\text{max}}$

 $\mathcal{L}^{\text{max}}_{\text{max}}$ and $\mathcal{L}^{\text{max}}_{\text{max}}$

• No pixels classified as winterwheat

 $\mathcal{L}(\mathcal{L})$ and $\mathcal{L}(\mathcal{L})$

 $\mathcal{A}^{\mathcal{A}}$

TABLE 5: MULTITEMPORAL vs UNITEMPORAL COMPARISONS

REFERENCES

- 1. Hanuschak, G., R. Sigman, M. Craig, M. Ozga, R. Luebbe, P. Cook, D. Kleweno, and C. Miller, Obtaining Timelv Croo Area Estimates Using Ground Gathered and LANDSAT Data. Economics, Statistics, and Cooperatives Service, U.S. Department of Agriculture, August 1979.
- 2. Ozga, M., W. Donovan, and C. Gleason, An Interactive Svstem for Acreage Estimates Using LANDSAT Data. 1977 Machine Processing of Remotely Sensed Data Symposium, June 1977.
- 3. Craig, M. Area Estimates by LANDSAT: Arizona 1Q7Q. Economics, Statistics, and Cooperatives Service, U.S. Department of Agriculture, January 1980.
- 4. Ozga M., S. Faerman, R. Sigman, Editor Multitemporal System. Economics, Statistics, and Cooperatives Service, U.S. Department of Agriculture, November 1979.
- 5. Craig, M., R. Sigman, and M. Cardenas, Area Estimates by LANDSAT: KANSAS 1Q76 Winter Wheat. Economics, Statistics, and Cooperatives Service, U.S. Department of Agriculture, August 1978.
- 6. Sterling, T., S. Pollack, Introduction to Statistical Data Processing. Prentice-Hall, Inc., 1968.
- 7. Landgrebe, D.A., Final Reoort NASA Contract NASQ-14016. Laboratory for Applications of Remote Sensing, Purdue University, 1976.

APPENDIX A

TABLE A1: TABULATION OF SEGMENT DATA BY COVER TYPE

TABLE A2: TABULATION OF USABLE TRAINING DATA

TABLE A3: ANALYSIS DISTRICT DESCRIPTION

* Counties partially contained

 $\hat{\theta}$

TABLE A4: NUMBER OF CATEGORIES BY COVER TYPE AND ANALYSIS PROCEDURES

ANALYSIS PROCEDURE

TABLE A5: CLASSIFIER PERFORMANCE - MA1, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 67.03 percent correct

TABLE A6: CLASSIFIER PERFORMANCE - MA1, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 63.20 percent correct

TABLE A7: CLASSIFIER PERFORMANCE - MA2, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 66.78 percent correct

TABLE A8: CLASSIFIER PERFORMANCE - MA2, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 63.04 percent correct

 \sim

TABLE A9: CLASSIFIER PERFORMANCE - MA3, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 68.34 percent correct

TABLE A10: CLASSIFIER PERFORMANCE - MA3, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 63.30 percent correct

TABLE A11: CLASSIFIER PERFORMANCE - UA1, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 58.43 percent correct

TABLE A12: CLASSIFIER PERFORMANCE - UA1, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 50.83 percent correct

TABLE A13: CLASSIFIER PERFORMANCE - UA2, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 54.44 percent correct

TABLE A14: CLASSIFIER PERFORMANCE - UA2, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 44.85 percent correct

 $\mathcal{A}^{\text{max}}_{\text{max}}$ and $\mathcal{A}^{\text{max}}_{\text{max}}$

PERCENT

TABLE A15: CLASSIFIER PERFORMANCE - MA4, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 65.04 percent correct

TABLE A16: CLASSIFIER PERFORMANCE - MA4, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 53.60 percent correct

TABLE A17: CLASSIFIER PERFORMANCE - UA3, PUR

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 55.76 percent correct

TABLE A18: CLASSIFIER PERFORMANCE - UA3, EP

NUMBER OF PIXELS CLASSIFIED INTO

 \sim

OVERALL PERFORMANCE = 33.73 percent correct

TABLE A19: CLASSIFIER PERFORMANCE - UA4, PUR

NUMBER OF PIXELS CLASSIFIED INTO

 \mathbf{r}

OVERALL PERFORMANCE = 58.54 percent correct

 \sim

TABLE A20: CLASSIFIER PERFORMANCE - UA4, EP

NUMBER OF PIXELS CLASSIFIED INTO

OVERALL PERFORMANCE = 48.03 percent correct

 $\sim 10^{-1}$

APPENDIX B

Testing Hypotheses Between Two Values of R^2

 $\label{eq:2} \frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1}{\sqrt{2}}\sum_{i=1}^n\frac{1$

In performing this test I made use of the distribution of t, where

$$
t = \frac{\sqrt{(n-3)(1+r_{um})} \cdot (r_{yu}-r_{ym})}{\sqrt{2D}}
$$

and D is the determinant of

$$
\begin{bmatrix}\n t & r_{yu} & r_{um} \\
r_{yu} & t & r_{ym} \\
r_{um} & r_{ym} & t\n\end{bmatrix}
$$

with n - 3 degrees of freedom.

- r_{yu} coefficient of correlation between reported area and number of pixels classified into a given cover type using an unitemporal procedure
- r_{ym} $-$ coefficient of correlation between reported area and number of pixels classified into a given cover type using a multitemporal prooedure
- r_{um} coefficient of correlation between number of pixels classified into a given cover type using an unitemparal prooedure and number *ot* pixels classified into a given cover type using a multitemporal prooedure

Results of MA3,PUR vs UA1,PUR are as follows:

 $\label{eq:2.1} \begin{split} \mathcal{L}_{\mathcal{A}}(t) & = \mathcal{L}_{\mathcal{A}}(t) \left[\mathcal{L}_{\mathcal{A}}(t) \right] \mathcal{L}_{\mathcal{A}}(t) \quad \text{and} \quad \mathcal{L}_{\mathcal{A}}(t) & = \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\mathcal{A}}(t) \mathcal{L}_{\$

APPENDIX C

Statistical comparison of classification results using one-factor Aaalysis of Variance and Newman-Keuls Range Test

 $\omega_{\rm{c}}$

man a tha ann an 1992.
Tagairtí an t-Saomhair an 1992 an 1992 an 1992.

n
Santa Constantino (1948), contra la componente
Santa Constantino (1948), constantino (1949)

Similar analysis procedures were performed for MA3, PUR, MA3, EP, UA1, PUR, and UA1,EP. The statistical procedures for testing the classification results are outlined below.

I. Apply arcsin transformation to overall percent correct for each classification

II. Calculate sum of square SS = $[(55.76)^{2} + (52.71)^{2} + (49.85)^{2} + (45.48)^{2}]/1$ - $[(55.76 + 52.71 + 49.85 + 45.48)^{2}/4]$

 $= 57.36$

III. Calculate mean square

 $MS = 57.36/3 = 19.12$

IV. Calculate F test and determine if significant $F = 19.12/[821/13476] = 313.84$ (significant)

 $F_{3,00} = 2.60$

(95%)

V. Arrange transformed percents in descending order

VI. Calculate standard error

$$
SE = \sqrt{\frac{821/13476}{1}} = 0.247
$$

 $- 26 -$

VII. Prepare a table of differences

VIII. Prepare a list of least significant ranges

 R_{4} = 3.633 (0.247) = 0.90 R_3 = 3.314 (0.247) = 0.82 R_2 = 2.772 (0.247) = 0.68 (95%)

IX. Compare

 \bar{x}