General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

DC-Y1-04156
NSTL/ERL-198

E82-10285
TM-84694

AgRISTARS

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

Domestic Crops and Land Cover

December 1981

Technical Report

AN ANALYSIS OF LANDSAT MSS SCENE-TO-SCENE REGISTRATION ACCURACY

Benjamin R. Seyfarth
Paul W. Cook

(E82-10285) AN ANALYSIS OF LANDSAT MSS SCENE-TO-SCENE REGISTRATION ACCURACY (NASA) 14 p HC A02/MF A01 CSCL 02C
N82-24563
Unclas
G3/43 00285

National Aeronautics and Space Administration
National Space Technology Laboratories
Earth Resources Laboratory
NSTL Station, MS 39529

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO. DC-Y1-04156, NSTL/ERL-198	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
4. TITLE AND SUBTITLE An Analysis of Landsat MSS Scene-to-Scene Registration Accuracy		5. REPORT DATE December 1981
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) 1) Benjamin R. Seyfarth and 2) Paul W. Cook		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS 1) NASA/National Space Technology Laboratories/ Earth Resources Laboratory 2) USDA/Statistical Reporting Service/Research Division		10. WORK UNIT NO.
		11. CONTRACT OR GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS National Aeronautics and Space Administration and United States Department of Agriculture		13. TYPE OF REPORT & PERIOD COVERED Technical Report
		14. SPONSORING AGENCY CODE
15. SUPPLEMENTARY NOTES		

16. ABSTRACT

Measurements of scene-to-scene registration accuracy were made for 12 registrations done by ERL and for 8 registrations done by SRS. Comparisons were made for the eight common areas.

The results indicate that the ERL method is significantly more accurate in five of the eight comparisons. The differences between the two methods are not significant in the other three cases. There are two possible reasons for the differences. First, the ERL model is a piecewise linear model and the EDITOR model is a cubic polynomial model. Second, the ERL program resamples using bilinear interpolation while the EDITOR software uses a nearest neighbor resampling. This study did not indicate how much of the difference is attributable to each factor.

The average of all merged scene error values for ERL was 31.6 meters and the average for the eight common areas was 32.6 meters. The average of the eight merged scene error values for SRS was 40.1 meters.

17. KEY WORDS Landsat, ELAS, EDITOR, Registration, Control Points	18. DISTRIBUTION STATEMENT Unlimited

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
UNCLASSIFIED	UNCLASSIFIED	12	

NSTL FORM 13 (JAN 1978)

AN ANALYSIS OF LANDSAT MSS SCENE-TO-SCENE REGISTRATION ACCURACY

December 1981

Benjamin R. Seyfarth

Paul W. Cook

National Aeronautics and Space Administration
National Space Technology Laboratories
Earth Resources Laboratory

and

United States Department of Agriculture
Statistical Reporting Service
Research Division

TABLE OF CONTENTS

AN ANALYSIS OF LANDSAT MSS SCENE-TO-SCENE REGISTRATION ACCURACY

I. INTRODUCTION

This paper describes efforts by NASA's National Space Technology Laboratories (NSTL), Earth Resources Laboratory (ERL) and USDA's Statistical Reporting Service (SRS) to analyze the accuracy of scene-to-scene registration of Landsat Multispectral Scanner (MSS) data. Twelve areas, representing four major types of land cover categories, were selected for this study. Landsat MSS data were merged by both ERL and SRS. ERL used the Earth Resources Laboratory Applications Software (ELAS) system to perform the registration, while SRS used the EDITOR software system.

The specific objectives of this study were:

1. To determine the registration accuracies of the data sets merged by ERL and SRS.
2. To determine whether the registration techniques used were equally well suited for the different land cover categories studied.
3. To determine whether one registration technique produced significantly better registration.
4. To determine whether a significant relationship exists between the root mean square (rms) errors of control points used in scene-to-scene registration and the accuracies determined by this study.

II. LANDSAT DATA

One of the goals of this study was to determine the suitability of the ELAS and EDITOR registration techniques for a variety of land cover categories. The general surface cover categories chosen were agriculture, rangeland, forest, and mixed. The mixed category comprises areas with both forest and agriculture.

Three test sites were selected for each of the four land cover categories, for a total of 12 test sites. Two Landsat scenes were used by ERL for each site. Table 1 lists each Landsat scene and its land use category.

To minimize tape mailing between SRS and ERL, two sites from each of the four land cover categories were registered by SRS and the merged data were mailed to ERL for comparison of the two registration techniques. The two registration techniques were compared for these eight sites.

TABLE 1. LANDSAT SCENES AND LAND COVER CATEGORIES

LAND COVER CATEGORIES	LOCATION AND LANDSAT SCENE		
	MISSOURI (28/32)*	S. DAKOTA (32/29)	ARIZONA (40/37)
AGRICULTURE	21654-16100 30435-16165	21640-16313 21676-16321	21540-17174 21630-11203
	KANSAS (30/34)	IDAHO (43/30)	OKLAHOMA (30/36)
RANGELAND	21980-16270 30653-16254	22029-17405 30558-17404	22016-16283 21566-16204
	COLORADO (38/34)	S. CAROLINA (18/36)	IDAHO (46/27)
FOREST	30517-17141 21718-17095	30515-15201 30623-15181	21654-17511 30579-17561
	KENTUCKY (23/34)	N. CAROLINA (17/36)	MISSISSIPPI (23/39)
MIXED (FOREST AND AGRICULTURE)	21559-15392 30556-15474	22057-15142 21967-15125	21667-15442 21739-15461

* (XX/XX) = PATH/ROW NUMBERS.

III. ACCURACY ASSESSMENT METHODOLOGY

For each of the test sites analyzed, the Landsat scenes were merged by ERL using the ELAS software. SRS used the EDITOR software to merge eight of the sites. The accuracy of the merged data sets was then assessed.

The accuracy assessment procedure consisted of generating gray-scale electrostatic plots of the data and selecting test points using a digitizer. In this procedure two gray-scale plots were made of two bands of the base Landsat scene for approximately 16 different subscene areas. A third plot was made of one band of the merged Landsat scene. Then, for each subscene area, all three plots were mounted on a digitizer and matching test points were selected from each plot. In the cases of Missouri, Kansas, Colorado, and Kentucky, a fourth plot was made for each subscene area of one band of the data merged by SRS. Then all four plots were mounted on the digitizer at one time. Thus, these four comparisons were made on identical points. The four remaining comparisons are based on independently selected test points.

The points selected from the two plots of two bands of the base Landsat scene were used to estimate the random effect of a human selecting the same test points in this manner. The points selected from the plots of the merged Landsat data, when compared to the test points of the base Landsat scene, measured the random human error in addition to the registration error. A statistical procedure was used to separate the two sources of error and estimate the portion attributable to misregistration by the registration software.

IV. STATISTICAL MODEL

In the model, $[(x_{b1i}, y_{b1i}) | 1 \leq i \leq n]$ represents the set of x and y coordinates of the test points (i) selected from the first band of the base scene (b1), $[(x_{b2i}, y_{b2i}) | 1 \leq i \leq n]$ represents the set of coordinates of the test points selected from the second band of the base scene (b2), and $[(x_{mi}, y_{mi}) | 1 \leq i \leq n]$ represents the coordinates of the test points selected from the merged scene (m). These constitute all the measured variables. Several formulas will be derived using the x coordinates. Similar formulas exist for the y coordinates but will not be explicitly derived.

For the purposes of this model several assumptions are made. First it is assumed that, for each i,

$$x_{b1i} = x_i + \varepsilon_{b1i}$$

where x_i is the true location of the test point and ε_{b1i} is the error of the human selecting the points. Likewise, assume that

$$x_{b2i} = x_i + \varepsilon_{b2i}$$

Further, assume that

$$x_{mi} = x_i + \varepsilon_{mi} + \alpha_i$$

where ε_{mi} is the human error in selecting the point and α_i represents the misregistration at that point.

In addition, assume that ε_{b1}, ε_{b2}, ε_m, and α are all independent random variables, normally distributed, with mean 0 and variances σ_{xb1}^2, σ_{xb2}^2, σ_{xm}^2, and σ_α^2. It is assumed that σ_{xb1}, σ_{xb2}, and σ_{xm} are all equal to σ_{xh}, which will represent the human error in selecting points.

Then consider the differences between x_{b1i} and x_{b2i}. In particular,

$$\sum_{i=1}^{n} \frac{(x_{b1i}-x_{b2i})^2}{n}$$

which can be simplified as follows:

$$\sum_{i=1}^{n} \frac{(x_i + \varepsilon_{b1i} - x_i - \varepsilon_{b2i})^2}{n}$$

The assumption of independence yields

$$\sum_{i=1}^{n} \frac{\varepsilon_{b1i}^2}{n} + \sum_{i=1}^{n} \frac{\varepsilon_{b2i}^2}{n}$$

$$= \sigma_{xb1}^2 + \sigma_{xb2}^2 = 2\sigma_{xh}^2$$

Thus we get a simple formula for estimating the human error component:

$$\sigma_{xh}^2 = 1/2 \sum_{i=1}^{n} \frac{(x_{b1i} - x_{b2i})^2}{n} \qquad (1)$$

Now consider the difference between x_{mi} and the average of x_{b1i} and x_{b2i}.

$$\sum_{i=1}^{n} \frac{\left(x_{mi} - \frac{x_{b1i} + x_{b2i}}{2}\right)^2}{n}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left(x_i + \varepsilon_{mi} + \alpha_i - \frac{(x_i + \varepsilon_{b1i} + x_i + \varepsilon_{b2i})}{2} \right)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left(\varepsilon_{mi} + \alpha_i - \frac{\varepsilon_{b1i}}{2} - \frac{\varepsilon_{b2i}}{2}\right)^2$$

The assumption of independence yields:

$$\sum_{i=1}^{n} \frac{\varepsilon_{mi}^2}{n} + \sum_{i=1}^{n} \frac{\alpha_i^2}{n} + 1/4 \sum_{i=1}^{n} \frac{\varepsilon_{b1i}^2}{n} + 1/4 \sum_{i=1}^{n} \frac{\varepsilon_{b2i}^2}{n}$$

$$= \sigma_{XH}^2 + \sigma_\alpha^2 + 1/2 \; \sigma_{XH}^2$$

Using formula (1) to replace σ_{XH}^2 we get:

$$\sigma_\alpha^2 = \sum_{i=1}^{n} \frac{\left(x_{mi} - \frac{x_{b1i} + x_{b2i}}{2}\right)^2}{n} - 3/4 \sum_{i=1}^{n} \frac{(x_{b1i} - x_{b2i})^2}{n} \qquad (2)$$

These two formulas yield values for σ_{XH} and σ_α in terms of pixels. The rms error of the control points is expressed in meters so an expression of the misregistration in meters is desirable.

$$\sigma_X = 57\sigma_\alpha$$

Similarly, a formula could be derived for σ_y. Combining σ_X and σ_y, one gets a value for the total rms error, σ_M:

$$\sigma_M = \sqrt{\sigma_X^2 + \sigma_y^2}$$

A useful expression can be derived to determine the radius which contains 90% of the data. Assuming a normal distribution, the density function is

$$\frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right)}.$$

This can be approximated as

$$\frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)} \quad , \text{ where } \sigma = \frac{\sigma_x + \sigma_y}{2} .$$

Integrating over a circle of radius R gives

$$Q = \iint \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)} dx\, dy$$

$$\int_{\theta=0}^{2\pi} \int_{r=0}^{R} \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} r\, dr\, d\theta$$

$$= 1 - e^{-\frac{R^2}{2\sigma^2}}.$$

Solving this for R gives

$$R = \sigma\sqrt{2 \ln(\frac{1}{1-Q})}$$

or

$$R = (\frac{\sigma_x + \sigma_y}{2}) \sqrt{2 \ln(\frac{1}{1-Q})} \cdot$$

For Q = 0.9 the radius becomes

$$R90 = (\frac{\sigma_x + \sigma_y}{2})\ 2.146 \cdot$$

V. RESULTS AND CONCLUSIONS

The results of this study are listed in Table 2. For the 12 sites tested, the ELAS algorithm had an average merged scene error (σ_M) of 31.6 meters. In the eight test sites with comparable values for the EDITOR algorithm, the average of ELAS σ_M values was 32.6, while the average of EDITOR values was 40.1. Further, each σ_M value is lower for the ELAS data sets. Thus an F-test for significance was done to test whether the ELAS values were significantly smaller. The results are listed in Table 2.

An analysis was done to determine if any relationship existed between rms errors and σ_M values. However, this failed to yield any substantial results. The Spearman's rank correlation coefficient between these two variables was 0.159, a value too low for significance.

The last analysis was to see if a significant difference existed for σ_X values among the four general categories. However, in all four categories there was a significant difference between some pair of σ_M values. Thus differences among categories are less than differences within categories.

In conclusion, one may state that the ERL scene-to-scene registration is significantly better in five of the eight comparisons. There are two possible reasons for the differences. First, the ERL model is a piecewise linear model and the EDITOR model is a cubic polynomial model. Second, the ERL program resamples using bilinear interpolation while the EDITOR software uses a nearest neighbor resampling. This study did not indicate how much of the difference is attributable to each factor.

TABLE 2. TEST RESULTS

Land Cover Test Areas	ERL Control Point RMS Error ($\hat{\sigma}_{XH}$)	Number of Test Points NSTL/ERL	Number of Test Points SRS	ACCURACY ASSESSMENTS: R90 NSTL/ERL	R90 SRS	Model RMS Error ($\hat{\sigma}_M$) NSTL/ERL	Model RMS Error ($\hat{\sigma}_M$) SRS	F-Test Results* ($\hat{\sigma}_M$) ERL Vs. SRS
Agriculture								
Arizona	34.9	43	43	50.1	65.1	35.008	43.521	Significant
South Dakota	49.7	52	64	28.9	64.3	26.916	42.753	Significant
Missouri	59.8	58		61.1		40.401		
Range								
Kansas	55.4	64	64	55.9	56.7	37.603	38.813	Not Significant
Idaho	45.7	63	64	48.2	69.4	31.810	46.058	Significant
Oklahoma	49.4	70		42.0		28.320		
Forest								
Colorado	44.0	63	63	54.4	56.9	36.875	39.035	Not Significant
South Carolina	49.4	64	64	50.7	49.4	33.488		
Idaho	49.6	72		39.0		25.741	32.575	Significant
Mixed								
Kentucky	39.8	62	62	49.3	56.5	32.894	39.552	Significant
North Carolina	39.0	72	64	51.0	58.3	33.691	38.447	Not Significant
Mississippi	39.8	71		24.5		16.236		
Average	46.4			46.3	59.6	31.582	40.094	

$* H_0: \sigma^2_{SRS} \leq \sigma^2_{ERL}$, $\alpha = 0.10$

$H_A: \sigma^2_{SRS} > \sigma^2_{ERL}$