ERIM

ENVIRONMENTAL RESEARCH INSTITUTE OF MICHIGAN
P.O. Box 8618
Ann Arbor, Michigan 48107

Progress Report

for

NASA Contract NAS9-15476

ANALYSIS OF SCANNER DATA FOR CROP INVENTORIES

Program Manager

Quentin A. Holmes

Task Leaders

Richard J. Kauth
Richard C. Cicone
William A. Malila
John E. Colwell

Period Covered

15 June 1978 to 11 September 1978

ERIM Report Number 132400-12-P

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
132400-12-P		

4. Title and Subtitle	5. Report Date
Progress Report for June 15, 1978 - September 11, 1978 on the Analysis of Scanner Data for Crop Inventories	September 1978
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Co-Authors - R.J. Kauth, R.C. Cicone, W.A. Malila and J.E. Colwell	132400-12-P
	10. Work Unit No.

9. Performing Organization Name and Address	11. Contract or Grant No.
Environmental Research Institute of Michigan (ERIM) P.O. Box 8618 Ann Arbor, Michigan 48107	NAS9-15476
	13. Type of Report and Period Covered Progress June 15 1978 to September 11 1978

12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
National Aeronautics and Space Administration Johnson Space Center Houston, Texas 77058	

15. Supplementary Notes

16. Abstract

This report summarizes progress on six tasks of the subject contract. These tasks are:

Multisegment Training

Evaluation of Partitions for Signature Extension

Wheat vs. Small Grains

Forecasting Production of Wheat from Satellite Data

Study of Multicrop Spectral Separability

Multicrop Labeling Aids

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Crop inventory, sampling, haze correction, signature extension, Landsat, crop discrimination, yield forecasts, photointerpretation	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
UNCLASSIFIED	UNCLASSIFIED		

*For sale by the National Technical Information Service, Springfield, Virginia 22161

TASK 1
MULTISEGMENT TRAINING
(R. Kauth and W. Richardson)

1.1 OBJECTIVE

The objective of Task 1 is to develop a sampling strategy for selecting training data, applicable to proportion estimation over a wide region. The main requisites of that strategy are that it produces a representative sample and that the training sample size is small compared to the total area to be classified.

1.2 APPROACH

1. Create a conceptual basis for the problem of training in a large scale remote sensing system, incorporating the inputs from UCB, LARS, and other ERIM tasks, and consistent with LACIE operational constraints.
2. Within this framework, propose a detailed methodology for training selection.
3. Demonstrate the selection methodology in an intermediate scale exercise over a partition containing from 15 to 30 sample segments from which 5 to 10 segments are selected for training and for which a wheat proportion estimate is made.
4. Incorporate both multitemporal and across partition signature extension capability into the final procedure.
5. Incorporate the capability to work with incomplete sets of multitemporal data and to optimize selection to make estimates at several times during the growing season.

1.3 PROGRESS

A baseline version of Procedure B was previously provided to Task 2 of this contract for testing. During the last quarter such tests were carried out with necessary technical support provided by this task. (See Task 2 discussion.) Effort on the missing acquisition capability was temporarily suspended. Coding of the missing acquisition capability is about 90% complete. The problem of defining a composite procedure combining desirable aspects of both Procedure B and JSC's Procedure 1 was investigated resulting in recommendations to make key modifications in P-1. These were included in the SR&T quarterly review, Sept. 11-14, 1978.

The major effort during the quarter was in the exercising of diagnostic tools and procedures to measure the performance of certain components of Procedure B not previously examined in detail. Note that the tests on the baseline version of Procedure B under Task 2 are tests of global performance compared to other approaches. The component performance tests carried out under this task are for the purpose of identifying and isolating the sources of variance and bias in Procedure B and of establishing optimal parameter settings.

1.4 TECHNICAL DISCUSSION

The technical discussion will be limited to two topics, namely recommendations for immediate modifications to P-1 improve efficiency, and discussion of component tests of the spatial stratification portion of Procedure B. Both of these topics were presented at the quarterly review September 11, 1978.

The three performance measures for machine processing which are of greatest current interest are bias, variance and analyst support. P-1 and Procedure B are designed to reduce bias with respect to the source of labels and in this they are successful. It appears to be

worth while to trade some bias for other performance measures, such as reduced variance (i.e., greater efficiency) or analyst convenience or accuracy. Analyst accuracy and convenience may be substantially improved by asking the analyst to label only reasonable features in the scene, as indicated by blob interiors. This is discussed further in section 1.4.2., and in the Task 2 discussion.

Efficiency can be gained in part by a better stratification of the segment prior to bias correction. This is discussed in more detail in the following section 1.4.1.

1.4.1 RECOMMENDATIONS FOR P-1 MODIFICATIONS

The so called bias correction in P-1 can be thought of as a regression estimator or as a stratified sample estimate. In P-1 the spectral data is formed into two strata, a nominal wheat class and a nominal non-wheat class, using about half of the labelled samples (the Type 1 dots) to assist in forming the two strata from 40 raw unsupervised clusters. Other labelled samples (the Type II dots) which fall into these strata are then used to make a final estimate of the wheat in the scene.

$$\hat{P} = \frac{N_w}{N}\,\frac{n_{ww}}{n_w} + \frac{(N-N_w)}{N}\,\frac{n_{\bar{w}w}}{n_{\bar{w}}}$$

where

N is the total number of pixels

N_w is the number of pixels classified as nominal wheat

$n = n_w + n_{\bar{w}}$ is the total number of type II dots

n_{ww} is the number of wheat dots found in the nominal wheat class

$n_{\bar{w}w}$ is the number of wheat dots found in the nominal non-wheat class

Since the so called classes are only nominal this expression may be generalized to multiple strata, as in Procedure B. (In Procedure-B the sampling is with blobs instead of pixels, but that is not the important point in this discussion.) Then the estimate is

$$P = \sum_{i \in S} \frac{N_i}{N} \frac{n_{wi}}{n_i}$$

where the summation is over all the strata
Now the variance of this estimate is

$$\text{Var } \hat{P} = \sum_{i \in s} \left(\frac{N_i}{N}\right)^2 \text{var} \left(\frac{n_{wi}}{n_i}\right)$$

which may be shown to be.

$$\text{Var } \hat{P} = \sum \left(\frac{N_i}{N}\right)^2 \frac{P_i \left(1-P_i\right)}{n_i}$$

where

P_i is the true proportion of wheat in the i^{th} segment.

In the absence of prior information about the P_i the minimum variance estimate is found by directing the sample proportional to the size of the strata, that is,

$$N_i = N \frac{n_i}{n}$$

(This proportional allocation automatically occurs with sufficient accuracy in P-1 because there are only two strata. In Procedure-B the samples must be directed to the strata in order to approximate proportional sampling.)

Under this approximation the variance becomes

$$\text{Var } \hat{P} = \frac{1}{n} \sum \frac{N_i}{N} P_i \left(\frac{1-P_i}{n_i} \right)$$

In order to obtain a measure of efficiency we compare the variance in stratified sampling to the variance in unstratified sampling, namely $\frac{1}{n} P (1-P)$ where P is the overall true proportion of wheat in the scene.

The result is a ratio, the variance reduction factor, which is independent of the number of Type II dots drawn.

$$\text{R.V.} = \frac{\sum \frac{N_i}{N} P_i \left(\frac{1-P_i}{n_i} \right)}{P (1-P)}$$

Now it has been shown by Cochran [1] and independently by Holsztynski [2] that R.V. stays the same or increases when the number of strata is decreased by combining existing strata. Thus an indication of theory would be that the strata in P-1 should not be combined but instead should be sampled from proportionally.

In order to test this concept a group of 5 segments were examined. P-1 clustering (ISOCLASS) had been performed upon these segments and the cluster statistics were provided to us by G. Bahdwahr of JSC. The statistics included the number of pixels in each of 40 clusters, the Type I dot label assigned to the cluster and the true proportion of small grains in each. Using this information the R.V. for three different procedures can be calculated, 'P-1', 'P-Smart' and 'P-B'.

'P-1' consists of combining the clusters into two strata based on the Type I dot label. 'P-Smart' consists of combining the clusters into two strata based on the true proportion of small grains in the cluster. 'P-B' consists of leaving all the clusters separated and assuming that the samples are allocated proportional to the cluster size.

Figure 1.1 shows the results of this comparison for the 5 segments. The entries in the figure are the variance reduction factor, R.V. (Recall that a small value of R.V. is good, being a measure of the number of samples which must be labelled to achieve some stated variance of the estimate.) The table shows that 'P-smart' has a lower R.V. by about .2, which indicates that efficiency is lost in combining strata. 'P-B' averages about .1 lower R.V. than 'P-smart'.

Based on this example our recommendations, presented as the quarterly review, are as shown in Figure 1.2. Type I dots should not be used to group clusters, instead they should be used directly to make the segment proportion estimates. It appears that a factor of about 3 might be saved in labelling effort by this step.

1.4.2. DISCUSSION OF BLOB PARAMETER SETTING AND PERFORMANCE TESTS

(The following discussion is substantially the text of appropriate portions of the quarterly review presentation of Wyman Richardson.)

Figure 1.3

We have been conducting experiments on 13 Kansas segments to measure the performance of components of Procedure B.

Further testing of Procedure B on N. Dakota segments, have been carried out under Task 2 of this contract.

Figue 1.4

The components of Procedure B we have been testing are

The grouping of the data into spatial-spectral clusters (blobbing)

The grouping of blobs into B clusters

Selection of training segments

Selection of training blobs

The method of proportion estimation

Tests on the last four components were reported on last time. We now report tests on the first component. These tests have been made possible by the provision of exact ground truth.

Figure 1.5

To help in the understanding of the blob test, let me first describe the blob algorithm.

Blob is a way of grouping the pixels of a scene into clusters that are spatially and spectrally homogenous. The Blob algorithm is to add two more channels, line number and point number, to the multispectral data channels. Line and point are the spatial variables. A standard clustering algorithm is then used to add pixels to existing clusters and to creat new clusters when the pixel is not close enough to any existing cluster.

The distance function used in the clustering is this: each channel of the pixel is in turn subtracted from the corresponding channel of the candidate blob mean, this difference is squared, divided by a weight which we'll call a variance, and finally, all such terms are added up. Notice that the last 2 terms are spatial and the first ones are spectral.

The pixel is added to the existing blob with the smallest distance unless this minimum distance is less than a parameter TAU, in which case the pixel becomes the seed point of a new cluster.

In our task, we have used 6 spectral variables, brightness and greenness from 3 biophases.

Figure 1.6

In our study of Procedure B components, we are faced with the problem of specifying the parameters of the Blob algorithm, namely Var 1 through Var 6, VARL. VARP and TAU.

The spectral weights were set by referring to our previous work on finding the optimal spectral weights for grouping blobs into B clusters. A search pattern in 6 dimensional space convinced us that for B clustering, the best spectral weights, expressed as variances, were proportional to the effective ranges of the variables. Using the same proportion for blobbing weights, we have determined the spectral weights relative to each other.

We can set the proportion of the line variance VARL to the point variance VARP so that the line s.d. represents the same geographic distance as the point s.d. This proportion is not 1 to 1 because points are sampled more frequently than lines.

The next step is to determine the balance between the set of spectral variances and the pair of spatial variances. This was accomplished by holding the spectral variances constant and comparing the Blob reduction of variances for 3 sets of spatial variances, small, middle and large. (You understand that if all the parameters are increased by the same proportion, the algorithm remains exactly the same. That is why holding the spectral variances constant and varying the spatial variances and TAU is permissible.)

Figure 1.7

The blob reduction of variance factor that is the arbiter for determining the balance between the spectral and spatial weights is the same expression as before. The denominator is p times 1-p, where p is the overall wheat proportion in the segment. The numerator is an average of similar expressions, P_i times $1-P_i$, where P_i is the proportion of wheat in blob i.

When a blob is pure wheat or pure other, either P_i or $1-P_i$ is zero and the term vanishes. Thus the Blob R.V. is a measure of the purity of the blobs: the purer the blobs, the smaller the Blob R.V. It's like gold, a low score is a good score. Perfect purity produces a zero score; all blobs equally mixed produces a score of 1.

Figure 1.8

Here are the results of computing the blob reduction of variance for 8 segments and 3 settings of the spatial variances. The small setting corresponds to more emphasis on the spatial, the large setting to more emphasis on the spectral.

The general result is theat Blob R.V. is quite stable for the three settings. The vertical scale has been greatly stretched to show any trend in the curves. But since we do need to decide on a setting, we might as well use one that these gentle trends indicate is optimal.

The worst case, segment 10, has a minimum in the middle. But most of the segments indicate a minimum at a lower setting. You will notice that the average trend, represented by the dotted line goes down as you go from the upper to the middle setting. A setting of 3.46 and 6.0 midway between the 2 lower settings, was chosen.

The parameter TAU that determines whether a pixel should start a new blob, was chosen so that the number of big blobs would roughly

correspond to the number of fields. TAU was kept the same for all segments, rather than keeping the number of blobs constant, so that segments with large or small fields would not be forced into a uncomfortable mold.

TAU was determined in this way for the middle spatial setting. For the upper and lower settings, value of TAU was found to keep the number of blobs constant. You understand that when you decrease TAU, you increase the number of blobs and vice versa. And as a general rule when the number of blobs increases, the reduction of variance factor goes down. So to make the three spatial settings comparable, the number of blobs was held constant within each segment while the number of blobs was allowed to vary from segment to segment.

Figure 1.9

I've just mentioned the term big blob. By a big blob, I mean one that has interior pixels. An interior pixel is one that is surrounded on all four sides, up, down, left and right, by pixels in the blob. A pixel in a big blob that is not interior is called an edge pixel. A blob with not interior pixels is call a small blob.

Figure 1.10

Here is a gray map of the blobs in Segment 1865. The big blobs are represented by printed characters and the small blobs by blanks. Although there are many small blobs, they don't contain many pixels. They mostly represent boundaries between fields.

Figure 1.11

The parameter settings chosen produced blobs with these characteristics:

Segment 1 is good for comparing the number of blobs with the number of fields because it has very few small fields. Some of the other segments have many fields that one would not expect to correspond to big blobs.

Note that the number of big blobs in a segment is remarkably constant.

Although the small blobs are numerous, they don't represent many pixels. They average 19% of the pixels.

About half the pixels are edge pixels in big blobs.

The interior pixels average 30% of the total.

Figure 1.12

This table shows the performance of the operations of blobbing and B clustering as measured by the R.V. factor. Remember that the R.V. factor ranges from 0 to 1, 0 representing perfect stratification and 1, worthless stratification.

The table also shows the lower limit of purity imposed by the number of mixed pixels. A "mixed pixel" here refers to a mixture of wheat and non-wheat. If other crop mixtures were counted, a much higher number of mixed pixels would be observed.

The mixed pixel R.V. measuring the average purity of the pixels is analogous to the blob and B cluster R.V.

The blob R.V. is given in the "all" column and the B cluster R.V. in the right hand column.

It's not a coincidence that in every case, the pixel R.V < the blob R.V. < the B cluster R.V. A member of our staff has proved that whenever you combine groups into larger groups, the average purity, as measured by the R.V., can never get better, and only in rare

coincidences can it stay the same.

Thus the pixel R.V. is a limitation on how good the blob R.V. can be and similarly, the blob R.V. is a limitation on how good the B cluster R.V. can be.

The R.V. score for blob interiors is very good, showing that the blob operation is doing its job in the sense that although there may be some confusion at the edges of the blobs, the interiors of the blobs are quite pure.

The average R.V. for B clusters, .58, can be interpreted like this: If we were to sample the pixels of each B cluster at random in such a way that the number sampled from each B cluster were proportional to the size of the cluster, and form the stratified estimate of the per cent wheat in the segment, then that estimate would have a certain variance. We compare that variance with the variance of an estimate obtained by sampling at random from one pool of all the pixels in the segment and we find that the stratified estimate variance is only 58% of the overall estimate variance. Since variance is inversely proportioned to sample size, it follows that the stratified estimate would require only 58% of the identifications needed by the overall estimate to achieve the same variance.

This column is a misleading indication of the value of Procedure B because Procedure B samples blobs rather than pixels. This would be expected to reduce the variance because the proportion of wheat in a blob randomly chosen from a B cluster would be expected to have less variance than the proportion of wheat in a pixel randomly chosen from a B Cluster.

The purity of blob interiors offers some hope of solving a practical difficulty with sampling. Our wheat estimates are all based on samples provided by the judgments of analyst interpreters. If these

AI's are asked to identify a pixel at random, the chance is 70% that it will come from an edge or a small blob and is therefore likely to be on or close to a field boundary. What with multitemporal registration errors and mixed spectral responses, it would seem a formidable, if not impossible task to identify such a pixel. But if asked to identify a relatively pure blob interior, the AI would have a much better chance to respond accurately.

Figure 1.13

This table of different methods of estimating the wheat % gives empirical information about the accuracy of such a procedure.

The column on the left is the % of wheat in the scene computed from every pixel. The figures in this column and indeed all of the figures I have displayed except for the column on the right, are based on the ground truth data recently computed at ERIM. The column on the right is a measurement of the % wheat in the scene made by planimetry at JSC a couple of years ago. The average absolute difference between the right hand column and the left column is 8 tenths of 1%, showing that even the most careful measurements from high resolution photography are subject to an error of about 1%. I'm leaving out of the calculation the discrepancy in segment 27 which is caused by the failure of the photography to cover the top quarter of the segment.

The next four columns are the 1% wheat computed on various subsets of pixels. The big blob pixel and the edge pixel estimates are quite close to the measured truth, while the small blob pixel and the interior pixel estimates are erratic.

The estimate on small blob pixels has a bias that averages 4% and ranges from -8% to 17%. The bias in the estimate from the small blob pixels has to be balanced by a bias in the opposite direction on the

big blob pixels. The big blob bias is smaller because there are more pixels in the big blobs. The estimate on interior pixels has a bias that averages -3% and ranges from -6.5% to 2.5%.

The interior pixel estimate was computed by simply totalling the % wheat of the interior pixels and dividing by the number of the interior pixels.

Another estimate based on interiors is to assign to all pixels in a big blob the proportion of wheat found in the blob interior. This is the "extrapolated from interior" column. This estimate is less biased and less erratic than the simple-minded interior pixel estimate. In fact, its bias and average absolute error is close to the error between the 2 planimetry measurements.

This estimate, moreover, is more achievable by the AI because it is based only on interiors of blobs.

A more realistic estimate yet is obtained on the assumption that on the relatively pure interiors, the AI will identify either 100% wheat or 0%. The "extrapolation from pure interiors" is such an estimate. Either 100% or 0% is extrapolated to all the pixels of the blob and then the % wheat of all pixels is obtained.

This estimate also gives very good results and has the advantage of representing a practical sampling procedure.

1.5 RECOMMENDATIONS

An experiment to determine the ability of LACIE analysts to accurately identify the interiors of designated blobs, as opposed to their ability to identify P-1 Dots, should be conducted.

The use of a more sophisticated clustering algorithm which might produce better variance reduction with fewer clusters should be investigated.

At least two aspects of Procedure B should be implemented in a P-1 environment.

Multiple strata (No Type 1 Dots)

Dots to be labelled assigned by stratified random sample

Further ahead in time, a blob capability could be added to P-1 to assure the purity of the dots assigned to be labelled.

1.6 PLANS

During the final quarter we plan to carry out end to end evaluation of Procedure B in Kansas, similar to Task 2 evaluation of N. Dakota. This will be carried out jointly with Task 2. These tests will be carried out for local and multisegment Procedure B and for 2 biowindows as well as 3 biowindows.

REFERENCES

1. Cochran, William G., Sampling Techniques, 2nd Edition, New York, Wiley, 1963

2. Holsztynski, W., "Some Results Concerning Stratified Sampling", ERIM Internal Memo, IS-WH-2267, June 27, 1978

FIGURE 1.1. VARIANCE REDUCTION FACTOR FOR 5 SEGMENTS USING PROCEDURE-1 CLUTERING AND SEVERAL GROUPING RULES

SEGMENT #	'P-1'	'P-SMART'	'P-B'
1060	.849	.653	.600
1231	1.000	.487	.434
1033	~1.0	.892	.708
1005	.857	.731	.672
1582	.596	.463	.432
AVERAGE	.860	.645	.569

FIGURE 1.2 CONCLUSIONS

- Clusters ought not be combined in P-1, instead, samples should be drawn directly from the 40 clusters

- Type I dots are thus not used in the labelling and can be used along with the Type II dots to further reduce the variance of the areal estimate.

- A practical limitation is that the number of clusters should not approach the number of labelled samples

FIGURE 1.3

STATUS OF PROCEDURE B DEVELOPMENT

- Have Been Conducting Experiments On Kansas Segments To Measure The Performance Of Components Of Procedure B
- Test and Evaluation of Procedure B In North Dakota By Task 2 Of This Contract

Task 1

FIGURE 1.4
COMPONENTS BEING EXAMINED

- Spectral/Spatial Stratification (Blob)
- Spectral/Ancillary Data Stratification (B-Cluster)
- Training Segment Selection
- Training Blob Selection
- Proportion Estimation Algorithm

FIGURE 1.5

BLOB ALGORITHM

EACH NEW DATA POINT X

$$X = (\underbrace{X_1, X_2, \ldots, X_n}_{\text{SPECTRAL}}, \underbrace{\text{LINE}, \text{POINT}}_{\text{SPATIAL}})$$

IS TESTED FOR ADMISSION INTO EACH EXISTING BLOB BY COMPUTING

$$\text{DIST} = \frac{(X_1 - M_1)^2}{VAR_1} + \cdots + \frac{(X_n - M_n)^2}{VAR_n} + \frac{(L - M_L)^2}{VARL} + \frac{(P - M_p)^2}{VARP}$$

WHERE $(M_1, \ldots M_n, M_L, M_p)$ IS THE BLOB MEAN VECTOR

AND $VAR_1 \ldots VAR_n$, VARL, VARP ARE WEIGHTS EXPRESSED AS VARIANCES

X IS INCLUDED IN THE BLOB FOR WHICH DIST IS SMALLEST EXCEPT THAT IF DIST > TAU, X STARTS A NEW BLOB

FIGURE 1.6
SETTING WEIGHTS FOR THE BLOB VARIABLES

- Spectral Variances Proportional To The Ranges Of The Variables
- Line and Point Variances Proportioned to Equalize Geographic Distance
- Balance Between Spectral and Spatial Found By Minimizing The Blob R.V.

FIGURE 1.7

BLOB REDUCTION OF VARIANCE FACTOR

$$R.V. = \frac{\sum_{\text{all blobs}} \frac{N_i}{N} P_i(1-P_i)}{P(1-P)}$$

P_i = PROPORTION OF WHEAT IN BLOB i

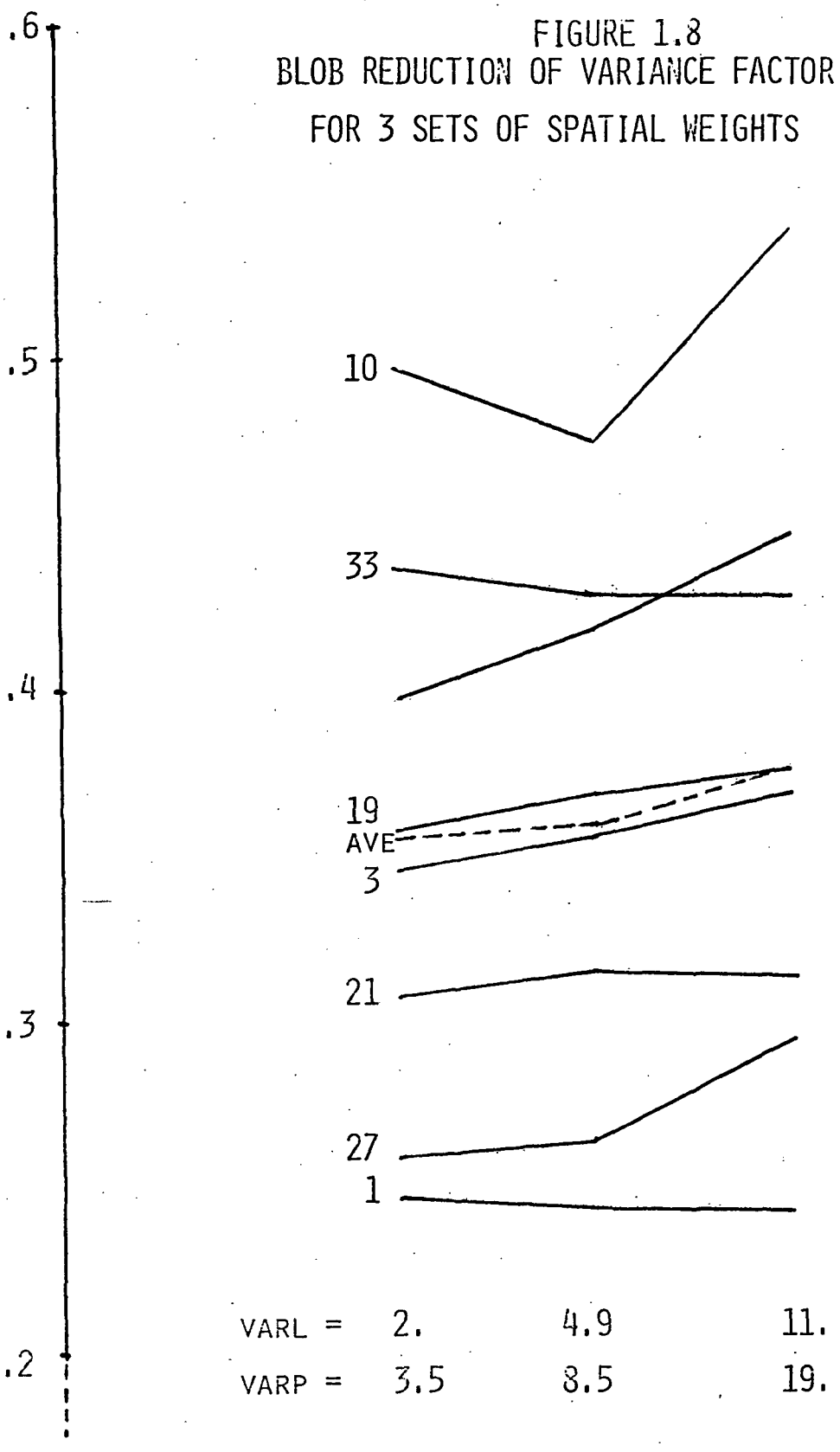
N_i = NO. OF PIXELS IN BLOB i

P = OVERALL WHEAT PROPORTION IN THE SEGMENT

N = NUMBER OF PIXELS IN THE SEGMENT

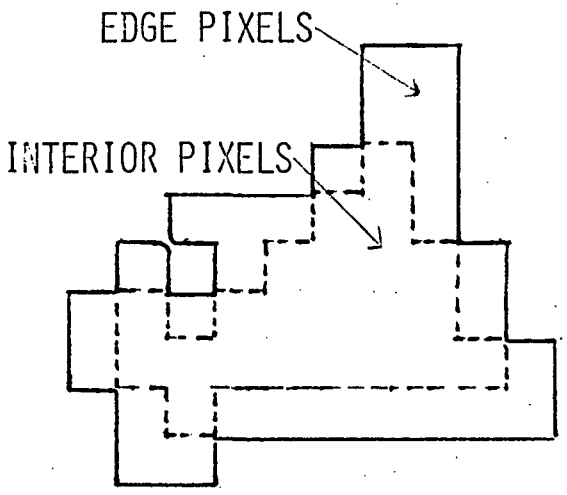
THE PURER THE BLOBS, THE SMALLER THE R.V.

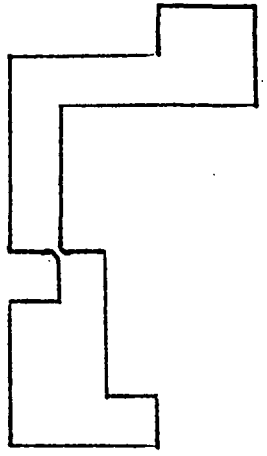
FIGURE 1.8
BLOB REDUCTION OF VARIANCE FACTOR
FOR 3 SETS OF SPATIAL WEIGHTS



VARL =	2.	4.9	11.
VARP =	3.5	8.5	19.

FIGURE 1.9





BIG BLOBS

HAVE INTERIOR PIXELS

SMALL BLOBS

HAVE NO INTERIOR PIXELS

ERIM

FIGURE 1.10

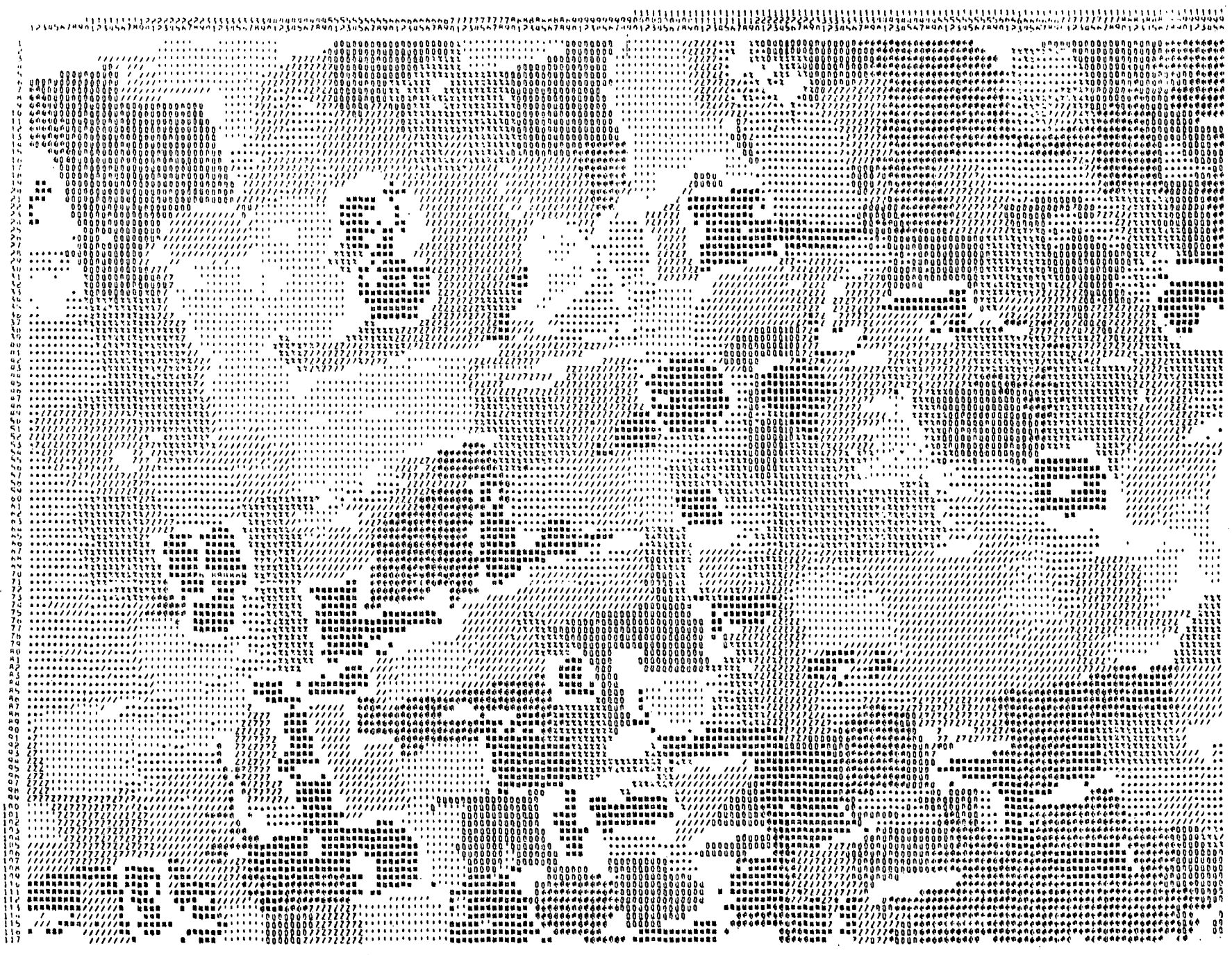


FIGURE 1.11

NUMBER AND SIZE OF BLOBS FOR KANSAS

Segment	No. Of Fields	No. Of Blobs	No. Of Big Blobs	Med. Size Of Blobs	Percent Of Pixels			
					Big	Small	Edge	Interior
1	322	659	347	18	90	11	46	43
3	764	1377	428	7	77	23	49	29
5	449	876	432	14	87	13	54	33
10	759	1782	489	6	69	31	52	17
19	518	1191	443	9	82	18	55	26
21	402	1037	434	10	86	14	52	34
26	287	590	385	29	93	7	54	39
27	399	1076	342	7	76	24	52	24
33	947	1001	468	13	85	15	56	29
8	1490	1394	449	8	76	24	49	27
14	820	1991	396	4	65	35	49	15
24	772	985	416	11	85	15	52	33
35	475	1103	344	6	84	16	47	38
				Average:	19	51	30	

FIGURE 1.12

REDUCTION OF VARIANCE FACTOR

Segment	(% Mixed Pixels)	Pixels	Blobs Or Parts of Blobs					B Clusters
			All	Big	Small	Edge	Interior	
1	11	.09	.24	.20	.56	.31	.04	.25
3	10	.11	.36	.36	.33	.42	.19	.66
5	8	.11	.41	.41	.42	.50	.21	.74
10	6	.16	.49	.49	.51	.56	.20	.86
19	12	.11	.36	.35	.42	.41	.16	.48
21	9	.08	.30	.29	.39	.37	.14	.48
26	11	.08	.26	.26	.35	.34	.09	.43
27	10	.08	.26	.26	.29	.31	.09	.56
33	18	.15	.43	.43	.47	.52	.17	.61
8	9	.18	.52	.54	.50	.60	.28	.71
14	9	.15	.39	.43	.36	.48	.18	.63
24	15	.13	.41	.40	.46	.51	.15	.53
35	7	.12	.39	.37	.45	.45	.17	.59
Average:	10	.12	.37	.37	.42	.44	.16	.58

FIGURE 1.13

VARIOUS ESTIMATES OF WHEAT PROPORTION

SEGMENT	ALL PIXELS	BIG	SMALL	EDGE	INTERIOR	EXTRAP. FROM INTERIOR	EXTRAP. FROM PURE INTERIOR	JSC WHEAT %
1	26.1	24.0	43.2	26.7	21.1	23.9	24.1	25.3
3	17.7	17.5	18.3	19.6	13.9	17.7	17.4	17.5
5	14.4	14.3	15.3	14.1	14.5	14.7	14.4	14.4
10	7.1	6.2	8.9	5.9	7.3	6.2	6.6	6.5
19	22.8	20.4	33.6	22.2	16.6	19.9	19.6	21.9
21	23.4	24.6	15.6	23.7	26.0	24.4	24.0	22.3
26	34.9	34.4	42.5	38.7	28.3	34.2	33.9	34.4
27	28.5	26.6	34.5	28.0	23.6	26.6	26.6	20.4
33	29.7	29.9	28.4	30.1	29.6	29.8	29.9	28.9
8	9.3	8.0	13.7	9.1	5.9	8.2	7.8	8.7
14	10.1	7.0	15.7	7.4	5.7	6.3	5.3	8.0
24	26.1	26.2	25.1	27.1	24.8	26.4	26.6	24.8
35	11.4	10.2	17.8	12.8	7.0	10.2	10.0	10.9
AVERAGE BIAS		- .9	3.9	.3	- 2.9	- 1.0	- 1.2	- .8
AVERAGE ABSOLUTE ERROR		1.1	5.5	1.1	3.3	1.3	1.4	.8

TASK 2
EVALUATION OF PARTITIONS FOR SIGNATURE EXTENSION
(R. C. Cicone)*

2.1 INTRODUCTION

The sampling and classification strategy of the Large Area Crop Inventory Experiment (LACIE) entails employing local signature training to determine wheat proportion estimates within 5 x 6 mile sample segments. These estimates are then aggregated over areas of interest. Multisegment signature extension is philosophically founded on the premise that representative training information can be determined using non-local procedures at an additional savings in cost and potential reduction in the variance of the estimate.

Task 2 is concerned with addressing the key issues found in Table 2.1 that pertain to non-local training techniques.

2.2 OBJECTIVE

The objective of Task 2 is to test and evaluate techniques and procedures which embody the signature extension approach to large area crop inventories using Landsat data.

2.3 APPROACH

The approach adopted to address the objective of Task 2 is outlined in Table 2.2

2.4 SUMMARY OF PROGRESS DURING THIRD QUARTER

Table 2.3 reviews progress and observations made during the First Two Quarters of this contract. Progress during the Third Quarter is outlined in Table 2.4.

*T. Wessling and O. Mykolenko contributed to work reported.

2.5 DETAILS OF PROGRESS

Efforts this quarter have concentrated on an evaluation of Procedure B conducted in a local mode. The objectives of this evaluation are detailed in Table 2.5. In addition to examining the performance of Procedure B in a local mode to provide a basis for comparison in multisegment testing, it was of great interest to examine the performance of this procedure under conditions very different from those under which it was originally developed and tested. Overall results were accurate and the Procedure behaved, in terms of bias and variance, in a predicted manner within performance expected of Procedure 1.

Table 2.6 outlines the general characteristics of Procedure B where as Table 2.7 provides information with regard to the procedures specific implementation. Procedure B differs from Procedure 1 in a variety of ways. Important differences include:

1. Modular construction [Procedure B can be easily adapted to advances in processing techniques]
2. data normalization
3. use of spatial features
4. use of multiple strata
5. use of sampling directed randomly to the strata proportional to their size
6. field interiors labelled rather than dots
7. labelled fields samples used for bias correction only (i.e., none are used for signature definition)

The specific experiment conducted is outlined on Table 2.8 using the data set illustrated in Figure 2.1. Note that the procedure is evaluated in North Dakota, a state with different crops and smaller fields than those encountered in Kansas where previous testing has been conducted.

Figures 2.2 to 2.5 are provided to illustrate field specification and labelling aspects of the procedure. Figure 2.2 is a map of field shapes in segment 1899 as determined by the BLOB program. Figure 2.3 indicates sixty fields to be labelled. The procedure considers only fields with at least one interior pixel. These interiors are illustrated on Figure 2.5. Figure 2.4 indicates the subset of sixty field interiors that would be presented to the AI for labelling.

Ground truth as provided by JSC in sub-pixel format was used for labelling. Hence, bias and variance due to analyst labelling is not a factor in this experiment. Two labelling criterion are evaluated (Figure 2.6). Proportional labelling requires that the field be labelled by the proportion of the dominant crop present, as opposed to grain/non-grain labelling in which the field is labelled by the dominant crop. The latter is considered more practicable for the AI. Its success is on how well BLOB field definition map the real fields present.

The series of tables and figures that follow indicate various performance measures used to evaluate the procedure. Much attention was placed on evaluating how well field shapes determined by BLOB correspond to actual field* shapes.

Table 2.9 provides statistics on a segment by segment basis describing the characteristics of the field shapes (blobs) determined. The definitions of 'big blobs', 'little blobs' etc. are consistent with those in Task 1. The 'interior variance reduction factor'(R) is of particular importance. This factor measures the relative purity of blob interiors. Low values indicate purer fields. If R=0, each pixel within the blob is from the same crop class, if R=1 then the field is a 50-50 mixture, and if R=.5 the field is about an 85-15 mixture. The

*A field is defined as a spatially contiguous area of like crops, as opposed to an ownership boundary.

highest average value of .352 in segments 1903 has been found to be suspect due to the misregistration of the ground truth. Figure 2.7 illustrates clearly that blob interiors are, in general, pure classes. Over 60% of big blobs are 90% or better pure non-grain, over 30% are 90% or better pure grain, and less than 10% are somewhere in between.

Since Procedure B samples only non-empty fields (big blobs), a source of bias with respect to a given crop in the procedure is the set of small blobs not sampled. The bias is estimated by the expression:

$$b = \frac{N-M}{N} \left(P_s - P_u \right)$$

where

P_s is the crop proportion in the pixels from which samples were drawn

P_u is the crop proportion in the pixels from which no samples were drawn

N is the total number of pixels

M is the total number of pixels from the sampled strata

If $P_s = P_u$ or if $M \approx N$, no significant bias is introduced. Figure 2.8 displays on a segment by segment basis (ordered by small grain content) the relationship of P_s and P_u. Note that in segments with little grain, $P_u > P_s$; whereas as grain content increases, $P_s > P_u$. This trend is being considered as a potential mechanism for bias correction at the segment level.

The next step in the procedure involves stratifying the 'big blobs' into 'b-clusters.' This step is carried out to reduce the variance of the overall estimate. The strata purity determines the extent to which the overall variance can be reduced. Figure 2.9 uses the variance reduction factor to illustrate the relative improvement that can be expected for 20, 40 and 60 strata versus the unstratified case. Note that b cluster purity is limited to within the purity of

the blobs that make up the b-clusters. There appears a general trend to purer strata as the amount of grain in a scene increases. The underlying cause is that the segments with more grain also have larger fields which tend to result in purer blob field definitions.

Turning to overall proportion estimates, Table 2.10 and Figure 2.10 are presented for reference. These represent the variance for unstratified sampling and bias on a segment basis that would be expected characteristics of Procedure B. Using proportional labelling, Figure 2.11 indicates the procedure performed very much along the lines expected in terms of bias due to not sampling small fields (Figure 2.10). Table 2.11 summarizes results for all eleven segments. Note that the measured standard deviation for the unstratified case (1 b-cluster) is comparable to that expected (Table 2.10). On the average the procedure was biased by about 1% where 31% was the measured grain ground truth proportion.

The ratio:

$$\frac{\sigma_{ij}^2}{\sigma_{1j}^2}$$

where

σ: standard deviation

i: the # of strata used

j: the # of fields sampled

is the measured variance reduction factor(R). In most cases $R \sim .5$ indicating an advantage in stratifying. Both the accuracy and R factor deteriorate if the number of samples is << the number of strata. This is a result of: (1) missing certain strata in sampling and (2) not being able to sample proportionally to the size of the strata. Other computations not shown here indicated the procedure to be unbiased with respect to the big blob ground truth (i.e., the truth ignoring the unsampled small field strata).

It was of great interest to compare the result of the second labelling criterion (grain/non-grain) to the proportional labelling approach since this labelling technique seemed more achievable using analyst interpretors. Figure 2.12 illustrates the bias measured on a per segment basis and compares favorably with both the expected bias (Figure 2.10) and that measured using proportional labelling (Figure 2.11). This is largely due to the relatively pure character of the blob interiors sampled. Table 2.12 is the counterpart to Table 2.11. We find the procedure to be largely unbiased and behaving, in terms of variance, as expected. This approach also proved to be unbiased with respect to the grain estimate based on big blobs only.

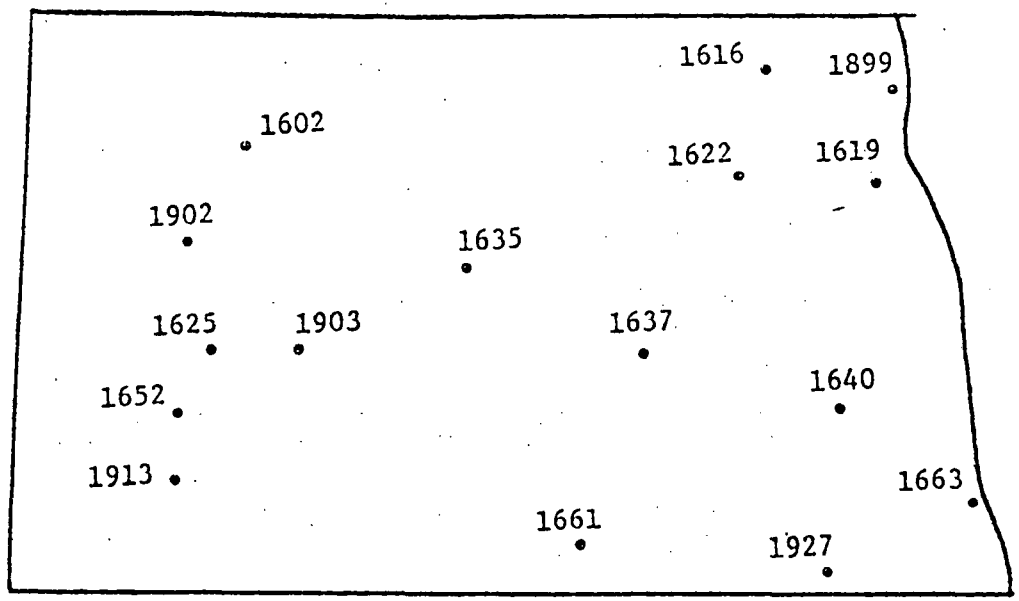
2.6 CONCLUSIONS AND RECOMMENDATIONS

Table 2.13 contains conclusions drawn based on the analysis of local Procedure B performance in North Dakota. Recommendations are shared with and presented in Task 1.

2.7 PLANS

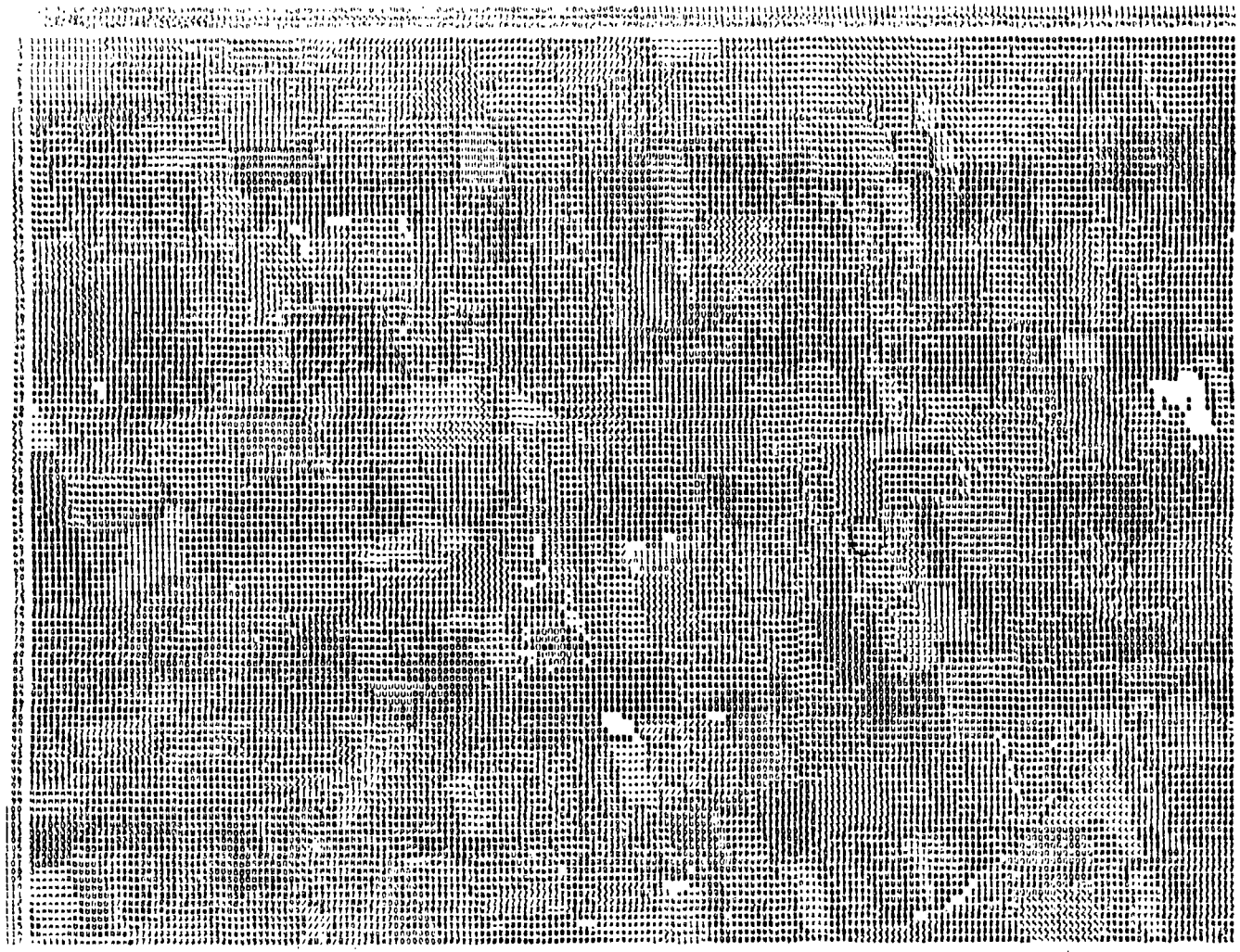
Table 2.14 lists activities recommended activities for future work. Activities for the remainder of the contract year include the completion of evaluation of multisegment techniques.

FIGURE 2.1
NORTH DAKOTA DATA SET



	BIOWINDOWS			
	1	2	3	4
1602	✓	✓	✓	✓
1616	✓	✓		
1619	✓	✓		
1622	✓	✓		
1625	✓	✓	✓	✓
1635	✓	✓	✓	✓
1637	✓	✓	✓	✓
1640	✓	✓	✓	✓
1652	✓	✓	✓	✓
1661	✓	✓		
1663	✓	✓	✓	✓
1899	✓	✓	✓	
1902	✓		✓	
1903	✓	✓	✓	✓
1913	✓	✓	✓	✓
1927	✓	✓	✓	✓

FIGURE 2.2



SEGMENT 1899 FIELD MAP

FIGURE 2.3

60 RANDOMLY SELECTED FIELDS FOR AI LABELLING

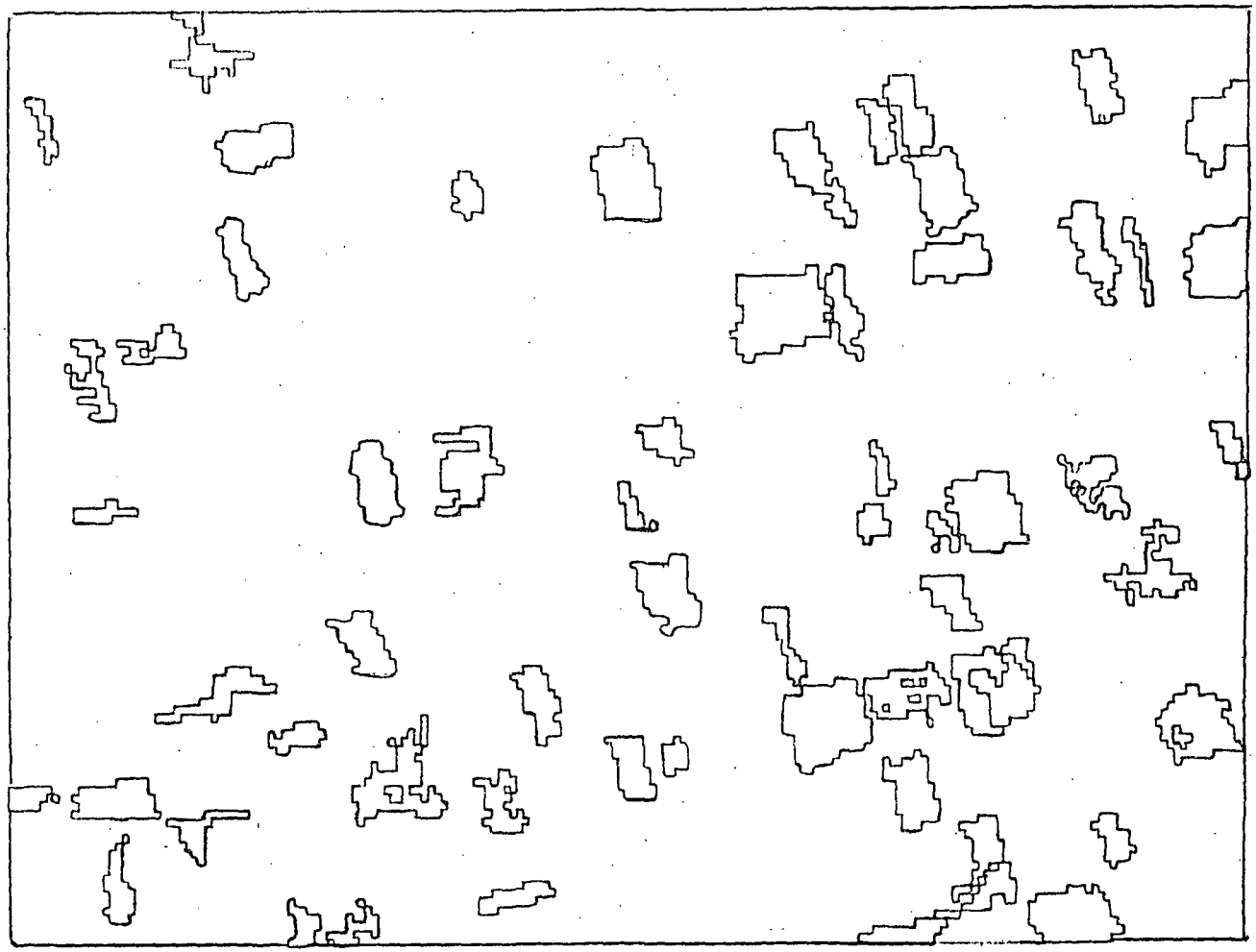


FIGURE 2.4

INTERIOR OF RANDOMLY SELECTED FIELDS

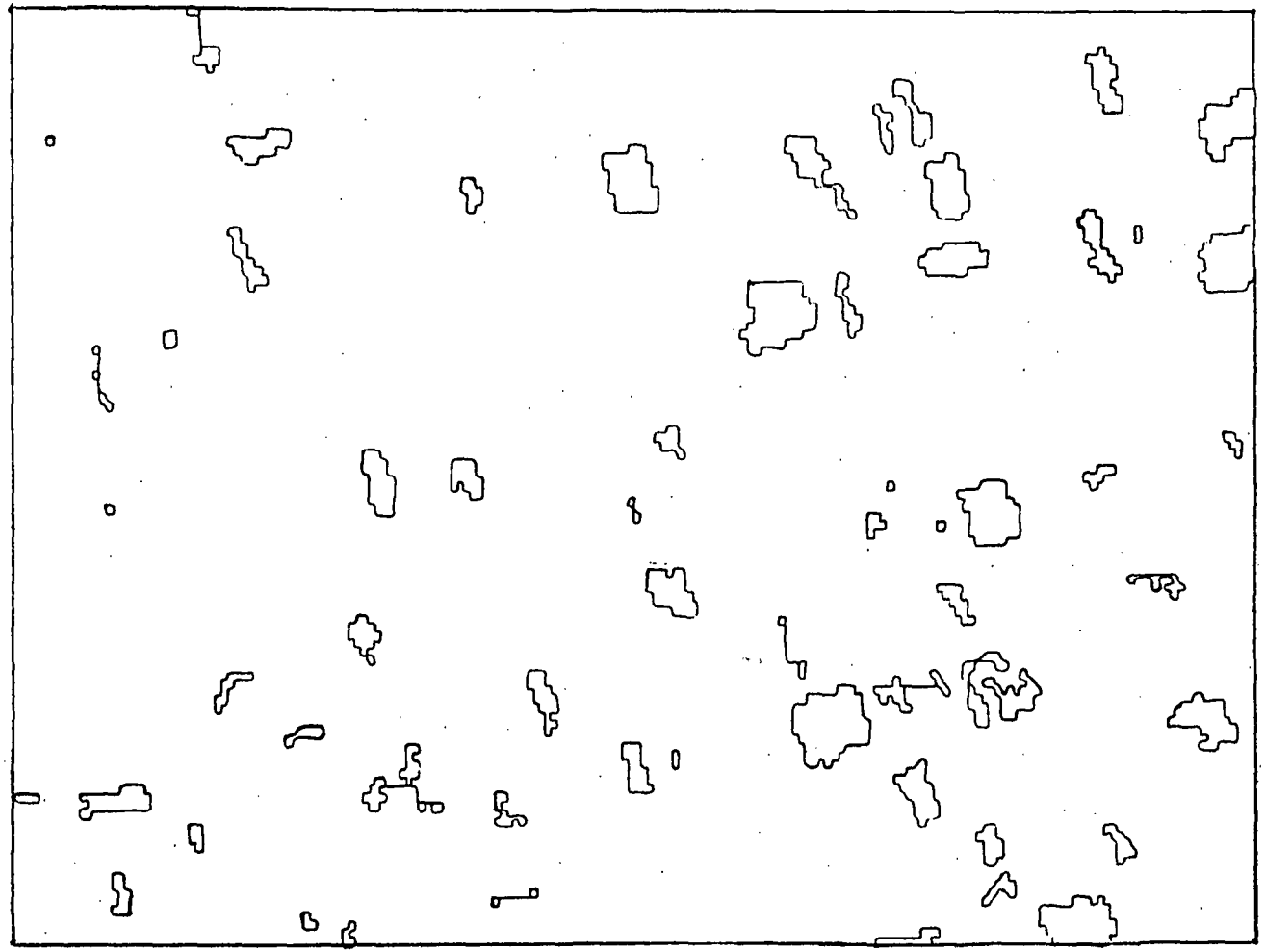
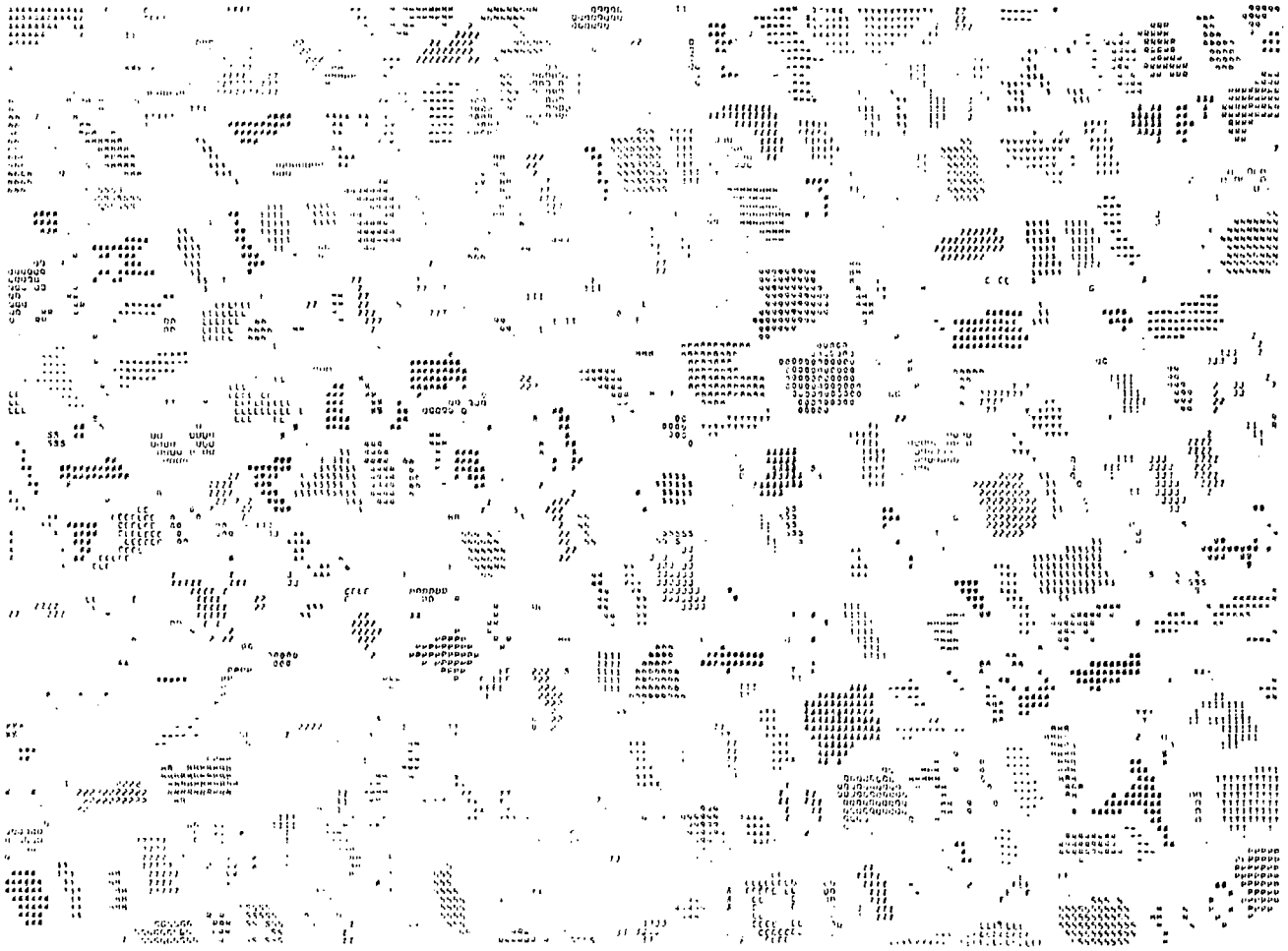


FIGURE 2.5

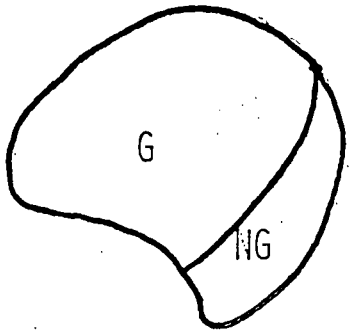


SEGMENT 1899 FIELD INTERIOR MAP

ERIM

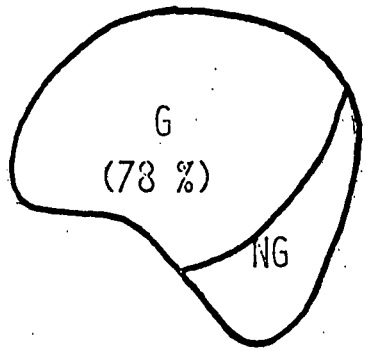
FIGURE 2.6
THE LABELLING CRITERION

GRAIN/NON-GRAIN LABELLING



100% GRAIN

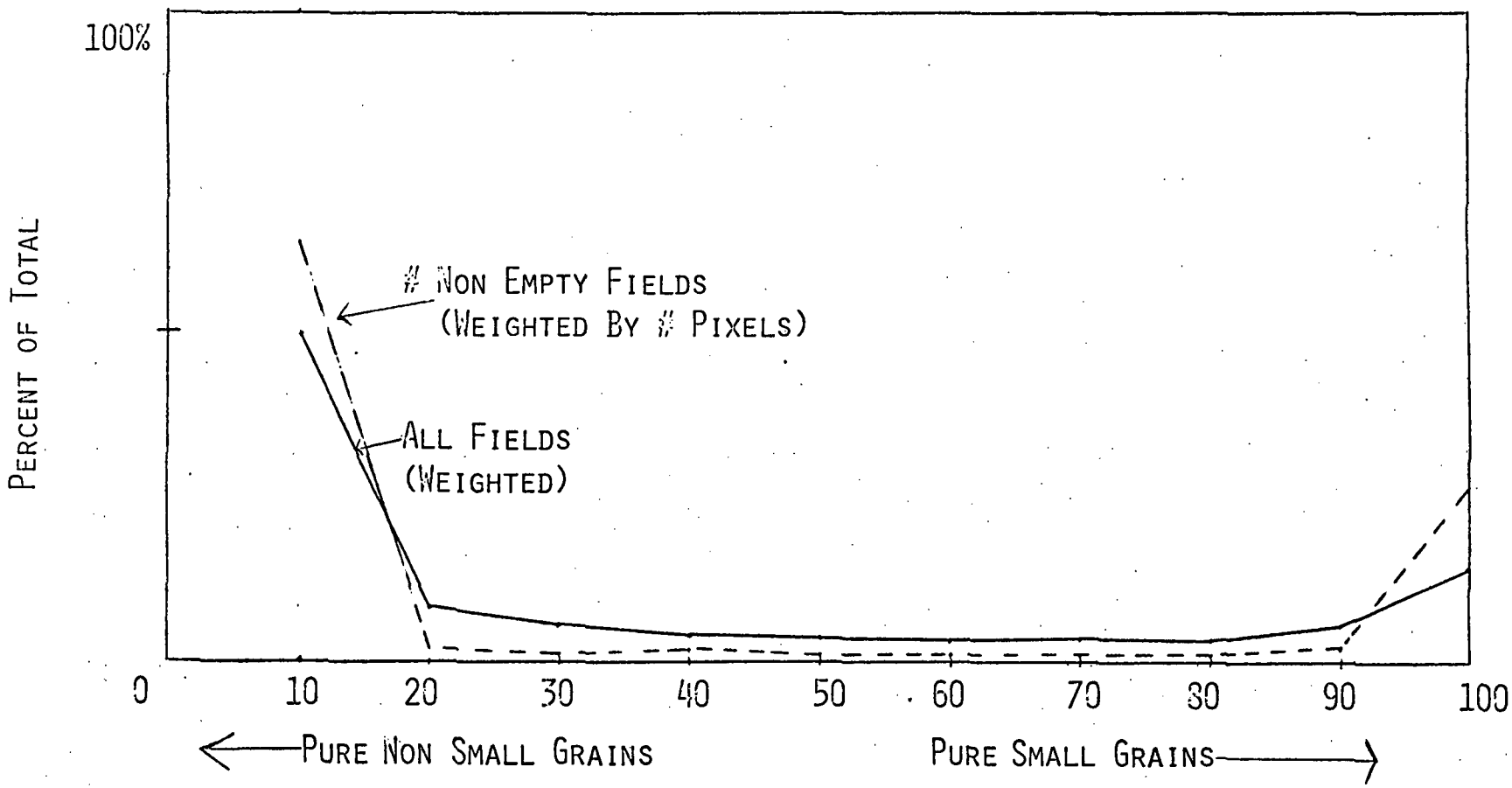
PROPORTIONAL LABELLING



78%
GRAIN PROPORTION

ERIM

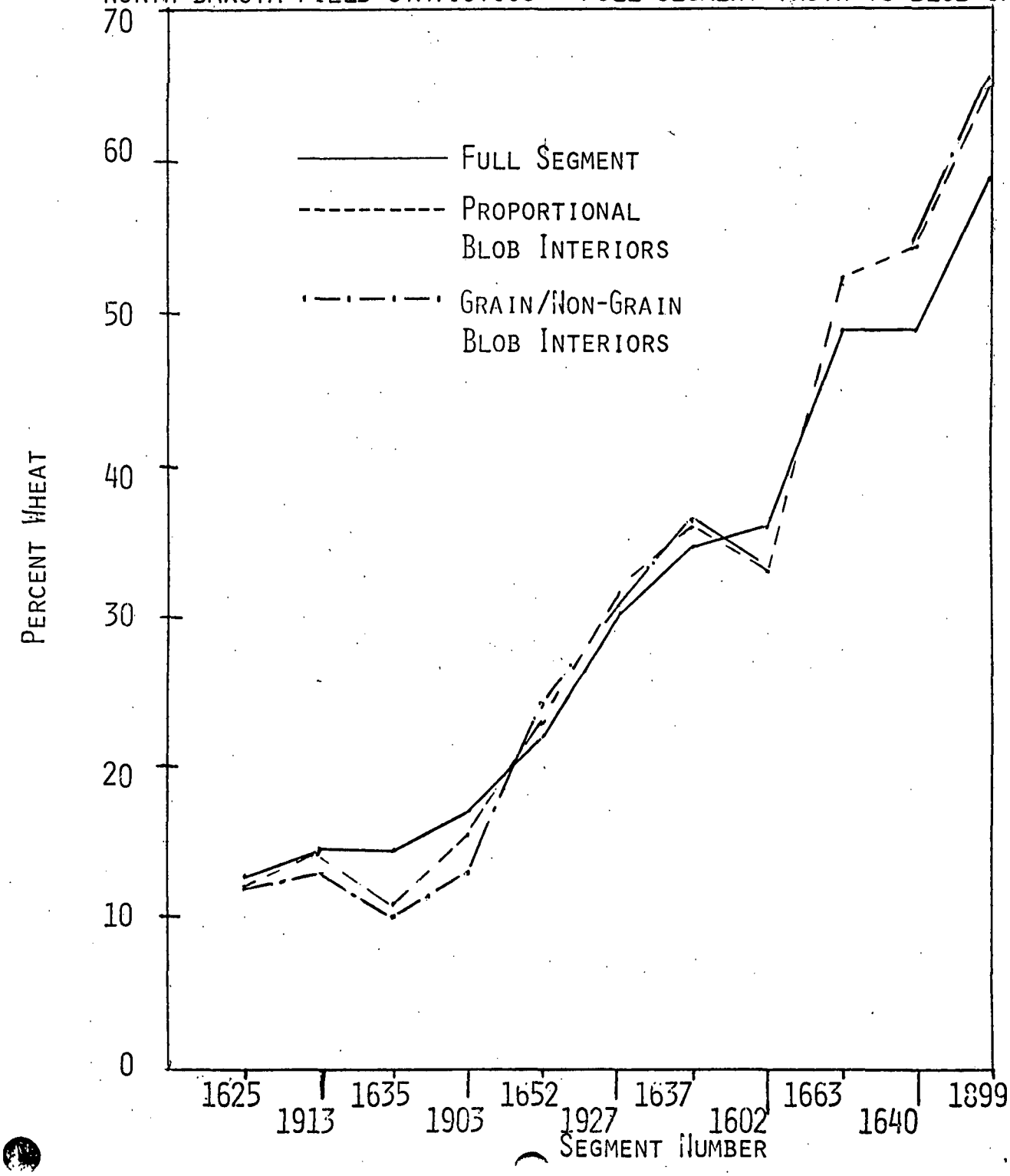
FIGURE 2.7
FIELD SAMPLE PURITY



ERIM

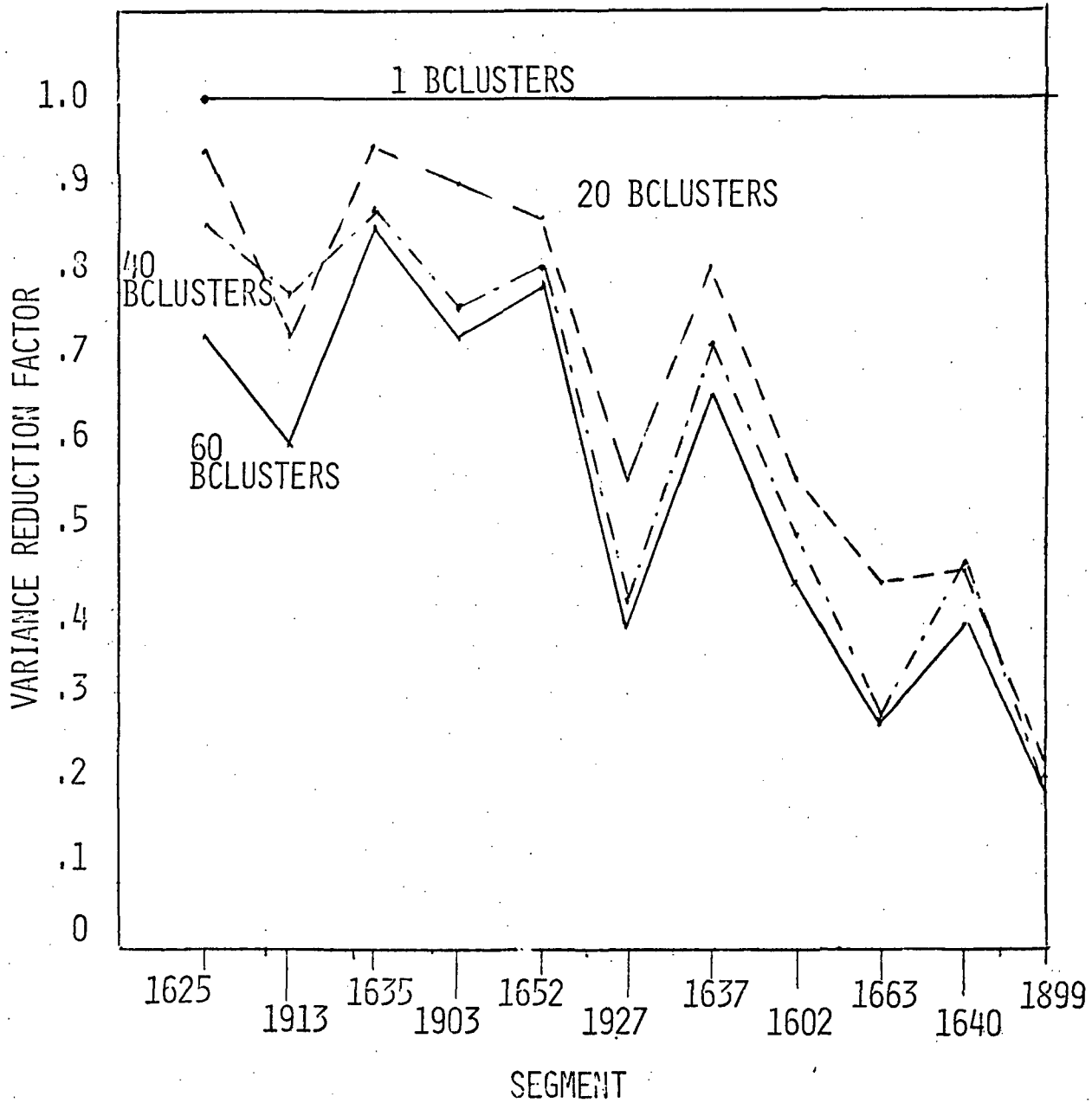
FIGURE 2.8

NORTH DAKOTA FIELD STATISTICS FULL SEGMENT TRUTH VS BLOB INTERIOR



ERIM

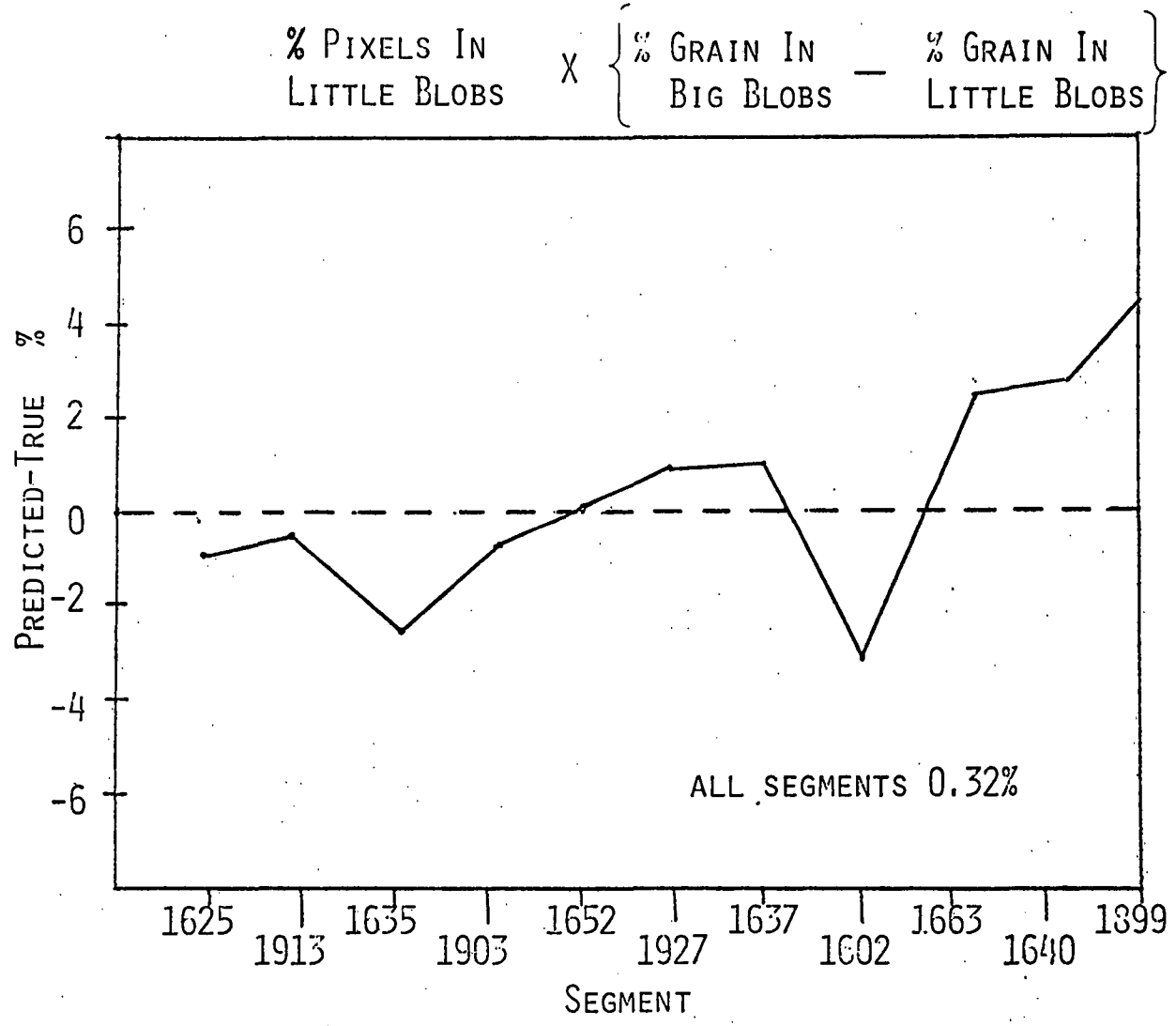
FIGURE 2.9
STRATA (BCLUSTER) PURITY



ERIM

FIGURE 2.10

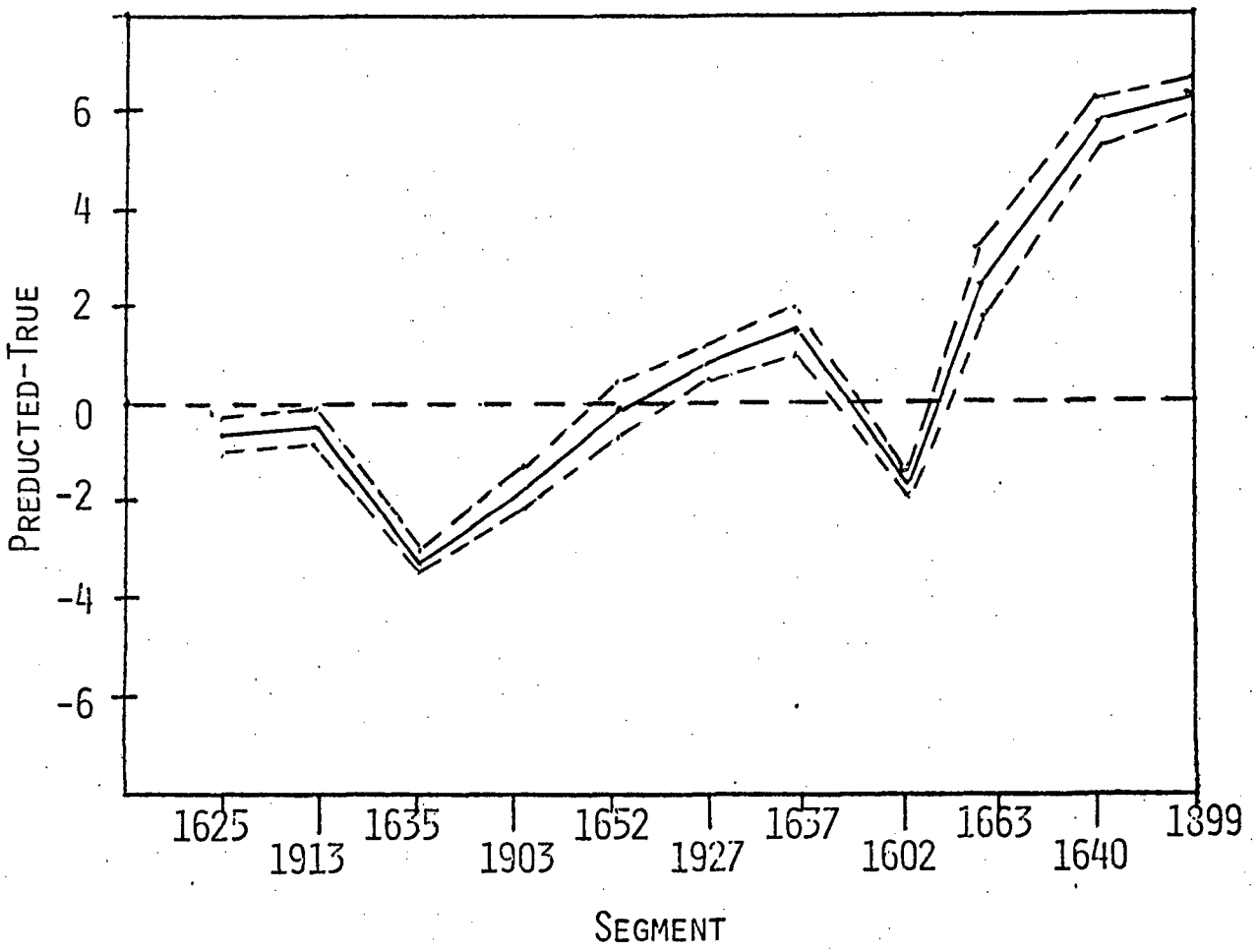
EXPECTED BIAS DUE TO SMALL FIELDS



ERIM

FIGURE 2.11

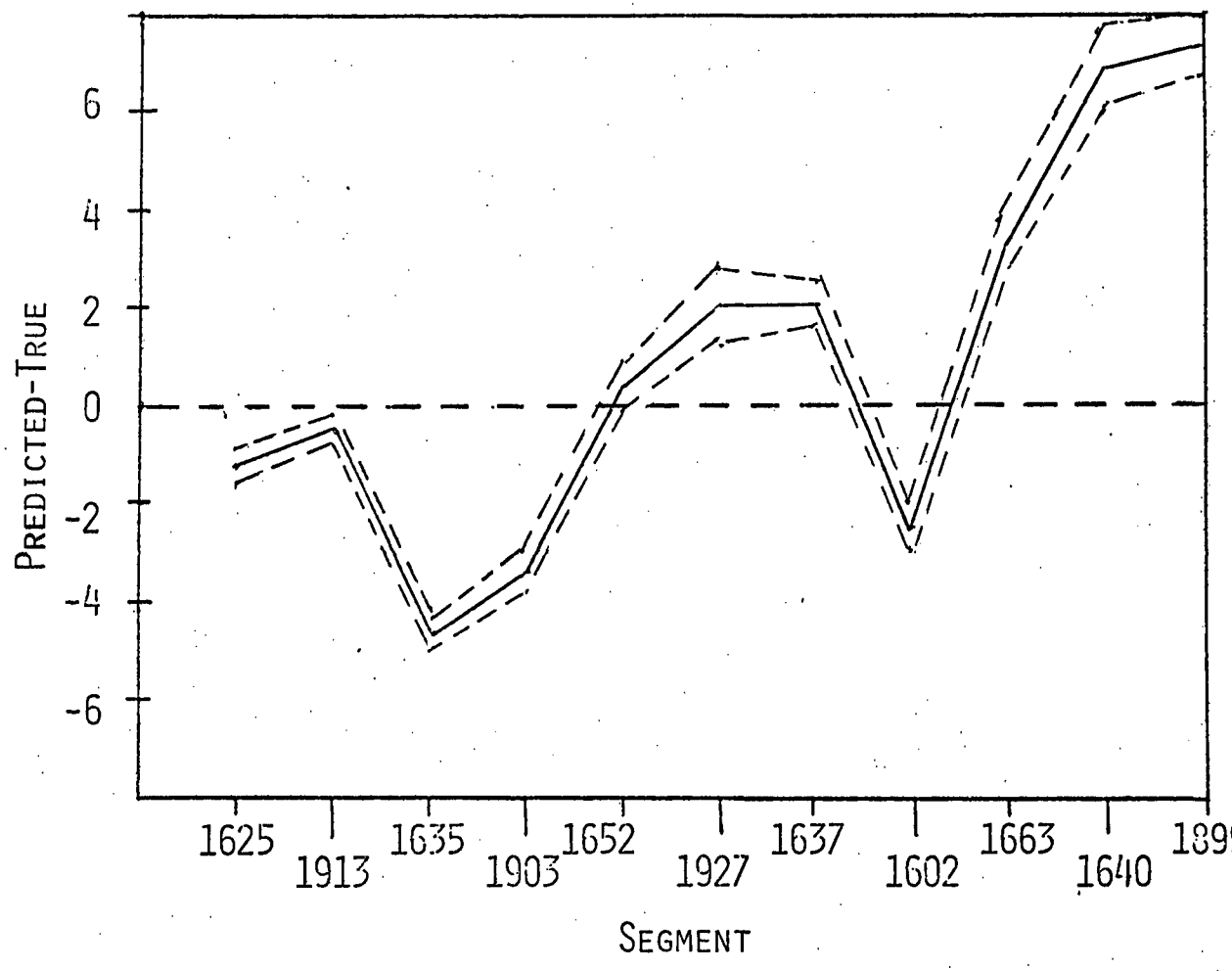
ESTIMATED BIAS
(PROPORTIONAL LABELLING)



ERIM

FIGURE 2.12

ESTIMATED BIAS
(GRAIN/NON-GRAIN LABELLING)



ERIM

TABLE 2.1
KEY ISSUES

- Multisegment Area Estimation Procedures
- Signature Extension
- Static Partition Boundaries
- Data Normalization and Preprocessing
- Procedure B

TABLE 2.2
APPROACH

- Test Design and Implementation Considerations
  - Experiment Design
  - Program Development
  - Data Preparation
- Parametric Evaluation Using Preclassification
  - Training Gain
  - Haze Correction
  - Static Partition Boundaries
  - Training Segment Selection
- Procedure Evaluation
  - Local Procedure B
  - Procedure B
  - Procedure 1

TABLE 2.3

REVIEW OF PREVIOUS QUARTERS' PROGRESS

- Evaluated Multisegment Signature Extension
  - — Using Preclassification Technique
  - — 1975-1976 Kansas Field Means Data Base
  - — 1977 North Dakota Blind Sites

- XSTAR Haze and Sun Angle Correction Resulted in Substantial Improvement in Both Field Center Classification Accuracy and Estimated Wheat Proportions

- Employment of a Viable Distance Measure to Partition or Relate Segments Appears to be a Critical Factor in Multisegment Signature Extension

TABLE 2.4

SUMMARY OF PROGRESS

- Continued Evaluation of Partitions Using The Preclassification Technique

- Initiated Procedure Evaluations
  - Conducted Extensive Evaluation Of Local Procedure B
  - Initiated Multisegment Evaluation*

*Evaluation of partitions will be conducted, but # of segments available is small.

ERIM

TABLE 2.5

OBJECTIVES OF THE EVALUATION

- Evaluate Performance in a Local Mode
- Carry out Evaluation In North Dakota To Determine Procedure's Behavior In A Region Dominated By Smaller Fields And Spring Wheat
- Evaluate Spatial Features Utilized
- Evaluate Statistical Characteristics of Strata
- Evaluate Statistical Characteristics of the Proportion Estimator
- Evaluate The Contribution of Static Partitions In A Multisegment Mode

TABLE 2.6

EVALUATION OF PROCEDURE B

- General Characteristics Of The Procedure
  - — Data Normalization And Spectral Feature Selection
  - — Definition Of Spatial Features
  - — Grouping Of Field Shapes Into Multiple Strata
  - — Derivation Of Wheat Estimate For Each Strata
    - • Systematic Selection Of Segments For Sampling
    - • Random Sampling Of Field Shapes
    - • Labelling Of Fields Drawn
    - • Estimate Wheat Proportion Using Field Estimate
  - — Estimate Wheat Acreage Over Desired Region of Interest

TABLE 2.7

PROCEDURE B TEST BENCH

- Specific Implementation of Procedure B

  — Data Normalization

    SCREEN For Clouds etc.
    XSTAR Haze Correction
    COSINE Sun Angle Correction
    TASCAP Linear Transform Into Tasselled Cap Space

  — Definition Of Spatial Features

    Blob

  — Grouping Field Shapes Into Multiple Strata

    BCLUSTER

  — Derivation Of Wheat Estimate For Each Strata

    Segment Selection: LOCAL Classification
    Random Sampling: In Proportion To Strata Size
    Labelling Of Fields Draws: Proportional Or Grain/Non-Grain
    Estimate Wheat Proportion: PROP (Extrapolated Blob Interiors)

  — Estimate Wheat Acreage Over Region of Interest

    Aggregate Result of North Dakota Segments

TABLE 2.8
EVALUATION OF LOCAL PROCEDURE B

- 11 North Dakota Phase III Blind Sites
- Small Grains vs. Other
- 2 Biowindows (2 and 3)
- 4 Spectral Variables (B,G,B,G)
- 4 Strata (b clusters - 1,20,40,60)
- 5 Field Samples (Blobs - 20, 40, 60, 80, 100)
- 2 Labelling Criterion (Proportional, wheat/non wheat)
- 10 Estimates Per Case (Random Field Sample Replicates)
- 4400 Proportion Estimates Total

TABLE 2.9

NORTH DAKOTA FIELD STATISTICS

BLOB SIZE AND PURITY

	# BIG BLOBS	% PIXELS	# LITTLE BLOBS	INTERIOR VARIANCE REDUCTION FACTOR	% GRAIN IN BIG BLOBS	% GRAIN IN LITTLE BLOBS	% GRAIN IN ALL BLOBS
1602	388	63%	1175	.134	33.3	41.0	36.2
1625	351	79%	525	.312	12.0	16.2	12.9
1635	449	72%	1033	.344	12.5	20.8	14.8
1637	522	78%	809	.128	36.1	30.1	34.8
1640	421	74%	776	.133	52.2	40.3	49.2
1652	398	77%	626	.164	22.6	22.6	22.6
1663	463	77%	738	.078	51.9	39.7	49.1
1899	395	73%	945	.050	64.0	47.3	59.5
1903	369	85%	406	.352	16.4	20.8	17.0
1913	360	74%	566	.201	14.2	16.5	14.8
1927	428	83%	507	.098	31.6	24.1	30.4
							31.0 AVE

TABLE 2.10
EXPECTED VARIANCE FOR UNSTRATIFIED SAMPLING

$$\text{VAR } \hat{P} = \frac{P\,(1-P)}{N}$$

$\sqrt{\text{VAR } \hat{P}}$	N
10.342	20
7.313	40
5.971	60
5.171	80
4.625	100

TABLE 2.11

PERFORMANCE OF LOCAL PROCEDURE B
(PROPORTIONAL LABELLING OF FIELDS)

# STRATA	# FIELDS	DIFFERENCE FROM TRUE (%)	STD. DEV.
1	20	.95	10.49
	40	1.80	7.37
	60	1.19	5.39
	80	1.15	4.00
	100	1.07	4.12
20	20	.67	7.32
	40	1.04	5.58
	60	.69	3.94
	80	.49	3.43
	100	1.08	2.78
40	20	2.87	6.39
	40	.94	4.75
	60	1.41	3.50
	80	.29	2.34
	100	.80	2.66
60	20	4.35	7.19
	40	2.25	4.70
	60	1.49	3.92
	80	.93	2.63
	100	.51	2.39

DIFFERENCE DETERMINED BY ESTIMATE-TRUE WHERE TRUE IS THE PERCENT OF ALL SMALL GRAINS IN LEC GROUND TRUTH

TABLE 2.12. PERFORMANCE OF LOCAL PROCEDURE B
(GRAIN/NON GRAIN LABELLING OF FIELDS)

# STRATA	# FIELDS	DIFFERENCE* FROM TRUE	STD. DEV.
1	20	1.30	10.47
	40	2.13	7.43
	60	.74	5.22
	80	.89	5.19
	100	.85	5.19
20	20	2.25	8.54
	40	.72	5.01
	60	.90	4.05
	80	.60	3.22
	100	.94	3.07
40	20	2.24	8.82
	40	.94	4.46
	60	.32	3.86
	80	.56	3.39
	100	.92	3.06
60	20	3.61	6.96
	40	1.34	5.49
	60	1.07	3.57
	80	.49	3.29
	100	.77	2.46

*DIFFERENCE BY ESTIMATE-TRUE WHERE TRUE IS THE PERCENT OF ALL SUBPIXELS LABELLED SMALL GRAINS IN LEC GROUND TRUTH

TABLE 2.13

CONCLUSIONS

- The Application Of Local Procedure B In Phase III North Dakota Blind Sites Resulted in Accurate Estimates Of The Small Grain Acreage Throughout Eleven Sites Tested

- Field Shapes Determined Using the 'BLOB' Program Were Representative Of Contiguous Pixels Of The Same Class

- Use Of Multiple Strata Results In A Cost Reduction With Regard To Required Number of Labelled Fields For Equivalent Accuracy

- Proportional Labelling Of Fields Is Not Necessary To Achieve Accurate Estimates, Resulting In Gain In Terms Of AI Efficiency

- Segment By Segment Results Are Bounded By How Well Non-Empty Fields Represent The Scene, Over All The Segments Tested, No Significant Overall Bias Was Introduced

TABLE 2.14
PLANS

- Complete Evaluation of Partitions Using Preclassification
- Complete Procedural Evaluations
  - —Substitute Amoeba* For Blob In Procedure B
  - —Carry Out Procedure Without Definition of Spatial Features
  - —Employ UCB & Universal Strata In A Multisegment Mode
  - —Evaluate The Characteristics of P1 Dot Sampling as a Spatial Stratification Of A Scene And Compare Reduction Of Variance Characteristics to Amoeba and Blob
- Final Report

*The Portion That Determines 'Fields' or 'Patches'

TASK 4

SPECTRAL SEPARABILITY OF SPRING WHEAT FROM OTHER SMALL GRAINS

(W.A. MALILA and E.P. CRIST)

4.1 INTRODUCTION

The problem of distinguishing between spring wheat and other small grains is of interest to LACIE and similar large-area crop inventory activities. Prior studies at ERIM (See Table 4.1) gave indications of separability of spring wheat and barley under certain conditions in Landsat data from several Phase 2 Blind Sites in North Dakota. Last quarter, we expanded our analysis to include Phase 3 Blind Sites, with comparable results -- unitemporal correct classification given complete training in the 80% range on individual segments and 76% for a seven-segment case. Spring wheat and barley were most separable in the growth stage when they are turning color and ripening. Thus, accurate crop calendar estimates appear to be important for discrimination. Work presented at the last quarterly S.R.&T. review by Dr. Gautam Badhwar of NASA/JSC also indicated the importance of crop calendar estimates for individual fields within a segment.

The objective of this task is to develop a spectral classification procedure for discriminating among the spring small grains using Landsat data. Oats were found to be much more difficult than barley to spectrally distinguish from spring wheat, so major emphasis has been placed on spring wheat vs. barley separation.

4.2 SUMMARY OF PROGRESS DURING THE THIRD QUARTER

During the reporting period, a general procedure for discrimination was outlined and specific steps were selected in some cases and several alternatives were explored in others. We explored the use of the time profile of green development as a basis for estimating the shifts in crop calendar between individual fields or pixels. A mathematical

model form was selected to serve as a reference profile in the estimation procedure, together with a cross-correlation calculation to determine time shifts. Three alternative decision procedures for using shift calculations were defined and evaluation was begun using data from Segment 1663, primarily, and from two other segments.

4.3 DETAILED DISCUSSION OF PROGRESS

A general procedure for discrimination between spring wheat and barley is outlined in Table 4.2. The preprocessing steps of data screening and haze correction would incorporate the XSTAR procedures developed by P. Lambeck of ERIM (See Task 8 for a description of the most recent development, a spatially varying haze correction algorithm).

Several potential measures of crop development are listed in Table 4.3. We decided to model the green development profile using selected data from Segment 1663 (See Table 4.4) and the mathematical model form shown in Figure 4.1. After manual alignment of several sets of observations, regression fits of model parameters were made to Tasseled-Cap Greenness values (offset by 25 units), resulting in the profiles shown in Figure 4.2 for spring wheat and barley. Note that barley tended to be greener at the peak and to decline faster in Greenness late in the cycle.

To estimate crop calendar shifts of individual fields or pixels, a cross-correlation calculation was made between offset Greenness values and samples from the reference profile for various shifts. The form chosen for the cross-correlation function (See Table 4.5) has values between 0 and 1 and is independent of any scale factor that may exist between the reference and the observations. In addition, one may estimate other characteristics of the observed data, such as peak Greenness value.

Figures 4.3 and 4.4 illustrate the effects of crop calendar shifts on spring wheat and barley pixels, respectively. A summary of our

analysis of these results is presented in Table 4.6. For the remaining work reported hereafter, only shifts relative to the spring wheat profile were utilized.

Three alternatives for the use of crop calendar shift calculations in small grains discrimination are presented in Table 4.7. We began evaluation of these alternatives, with initial emphasis being placed on the first two alternatives. The data set employed is described in Table 4.8.

Since the procedure is presumed to operate on pixels that have been classified as small grains, we devised a simple test to check for acceptable small-grain trajectories (Figure 4.5) in the data set. This check tended to eliminate pixels that had apparent boundary effects.

Applying a discrete version of the procedure, data were divided into bins according to their computed crop calendar shifts from the spring wheat reference date (Table 4.9). Then, a series of stepwise linear discriminant analyses was conducted to determine which variables and acquisition dates were most valuable in separating spring wheat from barley. As noted in Table 4.10, distance from the green arm, a linear combination of Greenness and Brightness (and well correlated with Landsat MSS5), was most frequently preferred, with Greenness and Brightness at Acquisitions 3 and 5 (Days 157 and 193) also being frequent high choices for additional discrimination variables.

The discrete procedure was applied to data using Brightness and Greenness variables at the two key times and optimized decision rules for each category of shift. Classification results are presented in Table 4.11. Little difference is seen between results obtained with the shift data and without it, except for Segment 1663 which was used to determine the reference profile. Further analysis was conducted to determine which shift values were associated with the best and worst classification performances. An even finer division of shift values was made and

classifications performed. The results in Table 4.12 show that highest performance was obtained for those pixels with the least amount of shift and tended to decline as the amount of shift increased. This we believe is due to the fact that the Landsat observations were timed optimally, relative to the average crop calendar for the segment, and large shifts corresponded to suboptimal sampling times (See Figure 4.6).

To explore the possible advantages of using shift as a continuous variable, single-date classification was performed twice for Segment 1663--once using only the four channels for Day 193 and once adding computed crop calendar shift as a fifth variable in the decision rule. Results are presented in Table 4.13 for three different aggregations of the decisions. On the left, all small grain pixels that were single-class according to the JSC ground truth tape are included. In the center, those passing the small-grain trajectory test are included. The right column which shows much improved performance includes only pixels which passed a field-center test based on spatial-spectral "blobbing" of the data. In all cases, the addition of shift improved performance, with the greatest improvement of edited pixels.

Additionally, we initiated an analysis of techniques being developed and applied by G. Badhwar of JSC. We determined that more detailed information about them was required before analyses could proceed.

4.4 CONCLUSIONS AND RECOMMENDATIONS

Conclusions and recommendations are presented in Table 4.14.

4.5 PLANS

Plans for the next quarter are presented in Table 4.15.

TABLE 4.1

BACKGROUND

- Results of Prior Analyses:
  - Spectral Differences do Exist Between Spring Wheat and Barley, although they have many Similarities
    - Historical USDA Statistics Reveal Differences in Growth Patterns
  - Timing of Acquisitions is Crucial
    - Linked to Crop Calendar
    - Mid-July (Day 193-201) most important in 1976 and 1977
  - Good Estimate of Crop Calendar is Needed
    - For Segment
    - For Individual Observations
- Data Base for Analysis has been Enhanced by Recent Availability of the Phase 3 LACIE Blind Site Data

TABLE 4.2

SYNOPSIS OF PROCEDURE

- Select Segments and Acquisitions
- Digitally Preprocess Data
  - Screen
  - Haze correction
  - Check for acceptable small-grain trajectory (or accept P1 classification)
- Estimate Crop Calendar For Each Observation Considered
- Choose Features and Critical Acquisitions for Decision
- Apply Decision Rule

TABLE 4.3

POTENTIAL MEASURES OF CROP DEVELOPMENT

- GREEN DEVELOPMENT PROFILE
- ANGLE BASED ON $\left(\frac{MSS7}{MSS5}\right)$ AND $\left(\frac{MSS4}{MSS5}\right)$ RATIOS
- MULTI-DIMENSIONAL DEVELOPMENT PROFILE
- LACIE CROP CALENDAR ESTIMATES
- ANALYST INTERPRETATION OF IMAGERY

TABLE 4.4

MODELING GREEN DEVELOPMENT PROFILE

- DATA (SEGMENT 1663, PHASE 3):
  - SELECTED LACIE DOTS (JSC GROUND TRUTH)
  - CLUSTERS OF POINTS IN 5X5 SAMPLE
  - POINTS FROM THE 15 PERIODIC-OBSERVATION FIELDS
- MANUAL ALIGNMENT OF CROP CALENDARS (EARLY, MEDIAN, LATE)
- SELECTION OF MODEL FORM (FUNCTION OF TIME)
- REGRESSION FIT OF MODEL FORM TO DATA (AVERAGES OF SELECTED LACIE DOTS)
- WHEAT VS. BARLEY COMPARISON

TABLE 4.5

ESTIMATION OF CROP CALENDAR SHIFTS

- CROSS-CORRELATION CALCULATION
  - OFFSET GREENNESS VALUES, $\{G(T_i)\}$
  - SHIFTED REFERENCE PROFILE, $\{F(T_i + \tau)\}$
- FUNCTION:

$$\max_{\tau} R = \frac{2}{1 + \dfrac{\sum_{i=1}^{N} F_{i+\tau}^{2} \sum_{i=1}^{N} G_i^{2}}{\left[\sum_{i=1}^{N} F_{i+\tau}\, G_i\right]^{2}}}$$

- ADDITIONAL ESTIMATES POSSIBLE:
  - SCALE FACTOR
  - PEAK GREENNESS VALUE

TABLE 4.6

RESULTS OF CROP CALENDAR SHIFT CALCULATIONS

- Substantially Reduced Greenness Variation on any Given Day
- Average Difference in Shifts Calculated For Wheat and Barley Profiles was One Day ± 1.2 Days for Segment 1663
- Initial Discrimination Studies Showed no Major Benefit From the Two Calculations of Shift
- Therefore, It was Decided to Use a Single Shift Calculation for Interim

TABLE 4.7

ALTERNATIVE DECISION PROCEDURES USING RESULTS OF SHIFT CALCULATIONS

- DISCRETE
  - DIVIDE DATA INTO A FEW CROP CALENDAR CATEGORIES
  - DETERMINE OPTIMUM DECISION PARAMETERS FOR EACH CATEGORY
  - MAKE DECISIONS ACCORDINGLY
- CONTINUOUS
  - LET SHIFT ESTIMATE BE A VARIABLE IN THE DECISION RULE
  - ALLOW POSSIBILITY TO INCLUDE ESTIMATED PEAK GREENNESS AS A VARIABLE
  - DETERMINE A SINGLE DECISION RULE FOR ALL SHIFTS
- MAXIMUM LIKELIHOOD ESTIMATE AS AN INTEGRAL PART AND CONSEQUENCE OF THE SHIFT CALCULATION -- SEPARATE REFERENCES FOR SPRING WHEAT AND BARLEY

TABLE 4.8

DATA UTILIZED IN STUDY

Segment	Acquisitions	Number of Pixels: Sp. Wheat	Barley
1663	77138		
	77157	5282	1269
	77175	(6103)*	(1365)*
	77193		
1899	77140		
	77157	4970	5210
	77175	(5513)	(5784)
	77193		
1515	77157	5095	3249
	77175	(7228)	(3651)
	77193		

*Values in Parentheses are without Trajectory Test.

TABLE 4.9

DISCRETE PROCEDURE

- Data Separated into bins based on Shift Relative to Reference Profile

Bin	Shift (Days)	% Total Small Grain Pixels
A	-15 to -6	5%
B	-5 to +5	50%
C	6 to 15	40%
D	16 to 25	5%

- Only Pixels Passing Trajectory and Screen Test Used

TABLE 4.10

DISCRETE PROCEDURE (CONTINUED)

- STEPWISE SELECTION OF DISCRIMINATION VARIABLES REVEALED CONSISTENT PREFERENCE (SEGMENT-TO-SEGMENT) FOR CERTAIN VARIABLES AND ACQUISITIONS
  - ACQUISITION: 3 (DAY 157) AND 5 (DAY 193)
  - VARIABLES:
    - •• DISTANCE FROM GREEN ARM (A LINEAR COMBINATION OF GREENNESS AND BRIGHTNESS AND WELL CORRELATED WITH MSS5)
    - •• BRIGHTNESS
    - •• GREENNESS
- SEPARATE MULTISEGMENT DECISION RULES DEFINED FOR EACH BIN USING BRIGHTNESS AND GREENNESS FROM ACQUISITIONS 3 AND 5

TABLE 4.11

DISCRETE PROCEDURE - RESULTS

DECISION RULE BASED ON ALL THREE (3) SEGMENTS
BRIGHTNESS AND GREENNESS - ACQUISITIONS 3 AND 5

SEGMENT	RESULTS WITHOUT SHIFT	RESULTS WITH SHIFT
1663	85.6	88.0
1899	81.2	82.3
1515	86.1	84.7
ALL THREE SEGMENTS	83.9	84.4

TABLE 4.12

DISCRETE PROCEDURE - RESULTS
BY SHIFT CLASS
ALL 3 SEGMENTS

SHIFT	ACQUISITION 5 DAY RANGE	NUMBER OF PIXELS (SP. WHT./BARLEY)	% CORRECT SP. WHT.	% CORRECT BARLEY
-14 TO -10	203 TO 207	253/219	70.4	72.1
-9 TO -5	198 TO 202	299/409	79.9	74.6
-4 TO 0	193 TO 197	1078/1066	90.9	84.5
1 TO 5	188 TO 192	3729/2391	93.8	79.9
6 TO 10	183 TO 187	4113/3465	89.2	78.1
11 TO 15	178 TO 182	1996/854	89.2	71.0
16 TO 20	173 TO 177	521/242	78.3	57.9
21 TO 25	168 TO 172	141/65	82.3	67.7

TABLE 4.13

CLASSIFICATION COMPARISONS - SEGMENT 1663 ALL SPRING WHEAT AND BARLEY PIXELS MID-JULY DATA

	All Pixels	Pixels Passing SMG Trajectory Check	Pixels Passing Field Center (Blob), SMG And Screen Tests
• Number of Pixels			
- Spring Wheat	6103	5282	2482
- Barley	1365	1269	609
• % Correct Classification Local Training 4 Channels			
- Spring Wheat	82.8	84.7	90.6
- Barley	66.4	66.7	66.5
- Average	74.6	75.7	78.6
• %Correct Classification Local Training 5 Channels (Crop Calendar Shift Added)			
- Spring Wheat	84.2	86.3	92.3
- Barley	72.7	75.8	81.8
- Average	78.5	81.1	87.1

TABLE 4.14

CONCLUSIONS AND RECOMMENDATIONS

- Estimation of Crop Calendar for Individual Observations can Improve Spectal Separability of Spring Wheat from Barley

- A General Procedure for Discriminating Spring Wheat from Barley was Outlined, with Alternatives at Various Stages. Additional Testing is Recommended before Final Choices are Made

- Items that should be Further Evaluated Include: Number of Acquisitions Required, Effects of Missing Acquisitions, Effects of Boundary Pixels, and Operation in Conjunction with Procedure B

TABLE 4.15

PLANS

- CONTINUE PROCEDURE EVALUATION
- APPLY TO BLOB DATA GENERATED AS PART OF THE TEST AND EVALUATION OF PROCEDURE B
- PRODUCE FINAL REPORT

FIGURE 4.1

SELECTED MODEL FORM

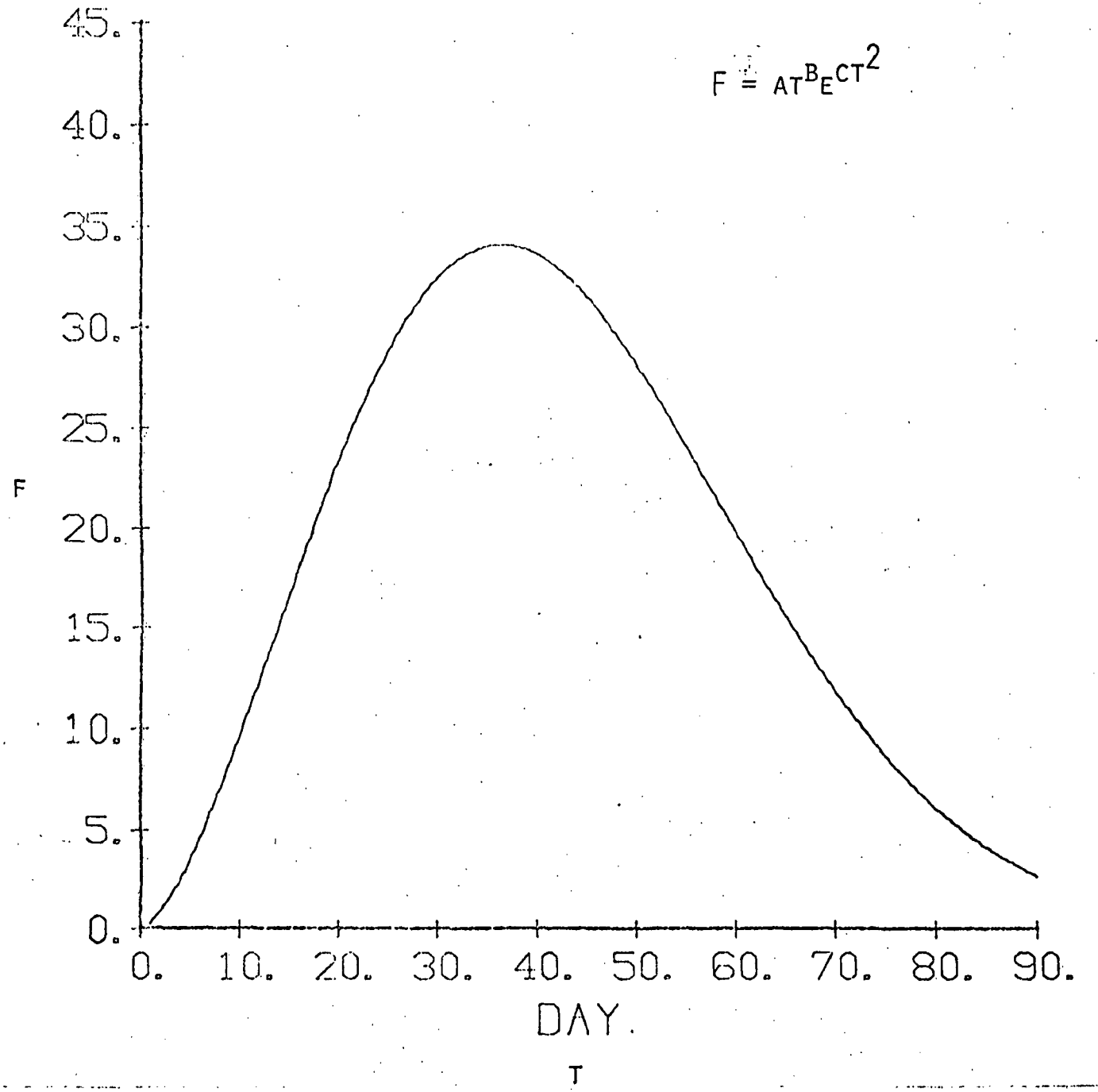


FIGURE 4.2

GREENNESS PROFILE MODELS FOR SEGMENT 1663 DATA

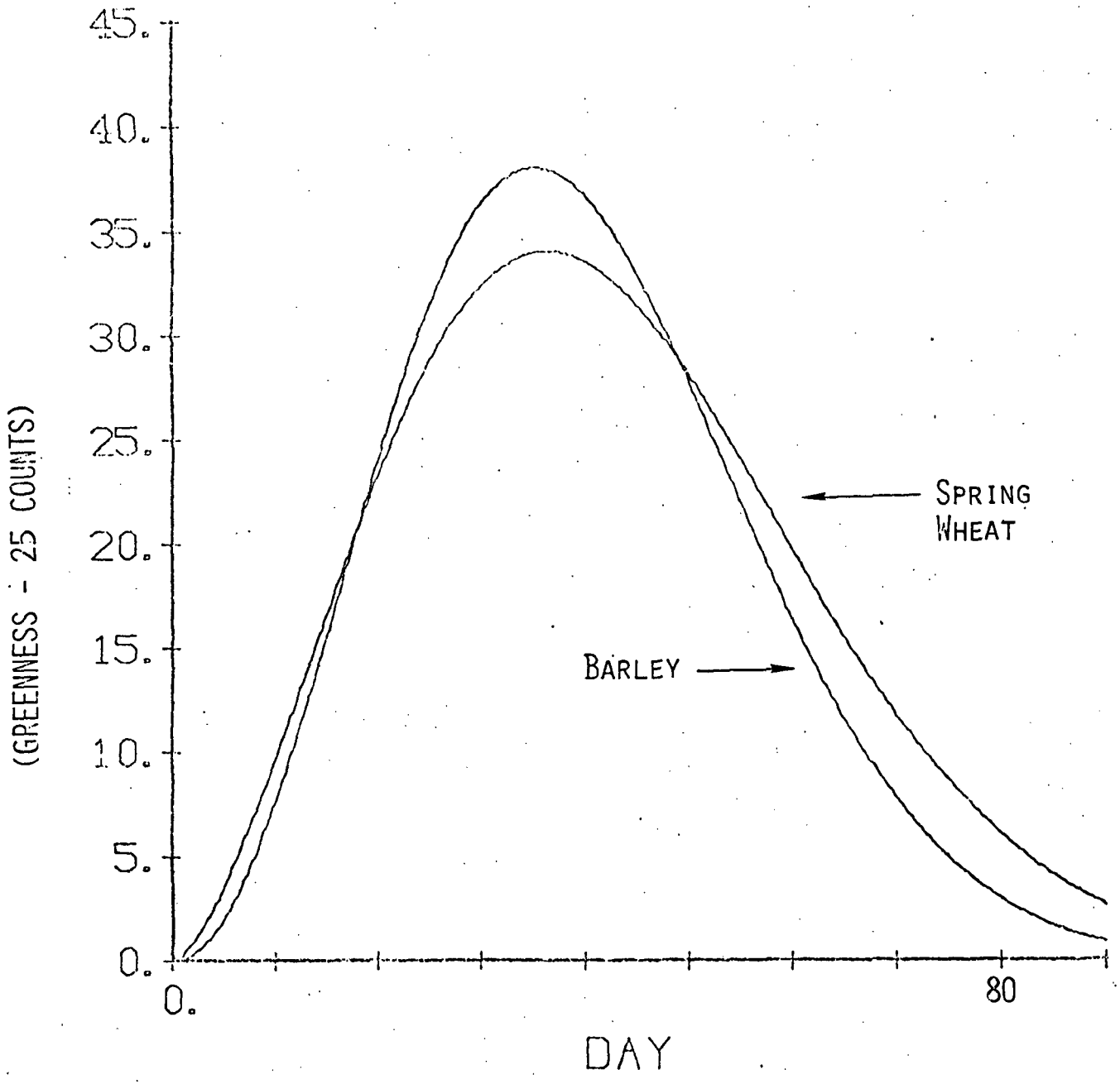


FIGURE 4.3

EFFECTS OF CROP CALENDAR SHIFTS ON 1663 GREENNESS PROFILES FOR SPRING WHEAT

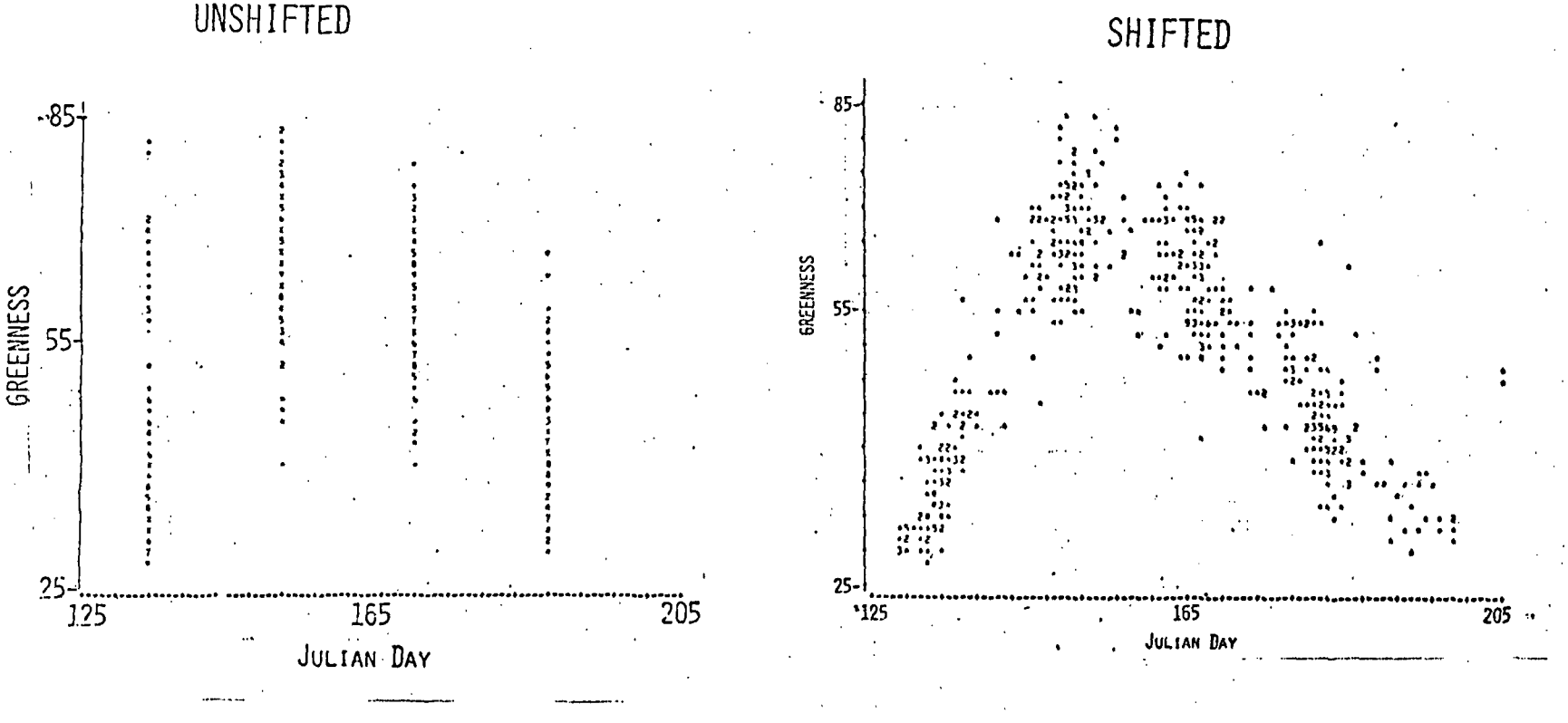


FIGURE 4.4

EFFECTS OF CROP CALENDAR SHIFTS ON 1663 GREENNESS PROFILES FOR BARLEY

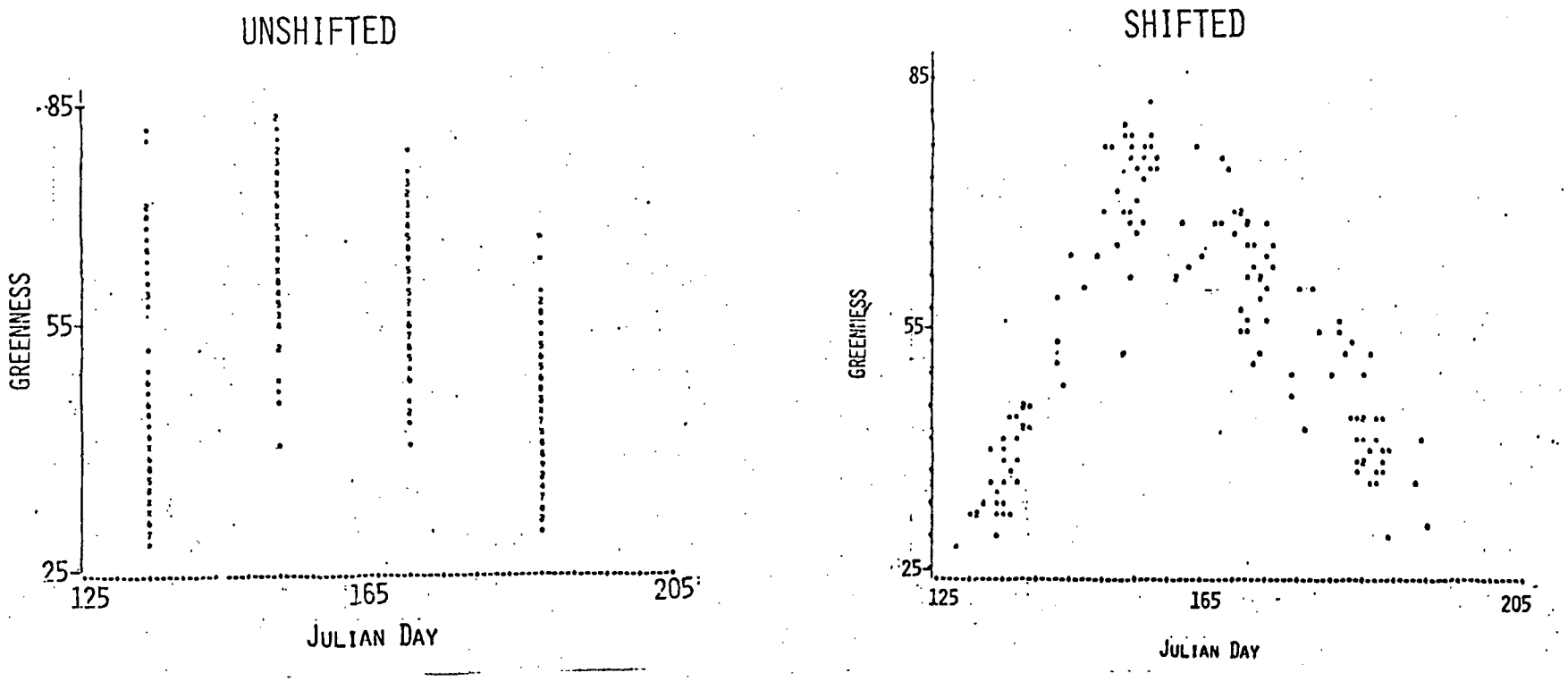
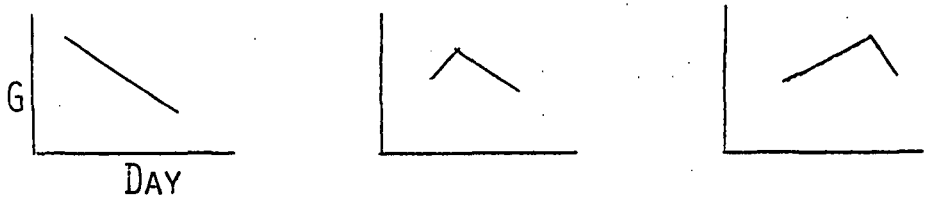


FIGURE 4.5

PROCEDURE TO CHECK FOR ACCEPTABLE SMALL-GRAIN TRAJECTORY

(MAY NOT BE REQUIRED OPERATIONALLY, IF PRECEDING CLASSIFICATION IS ACCEPTABLE)

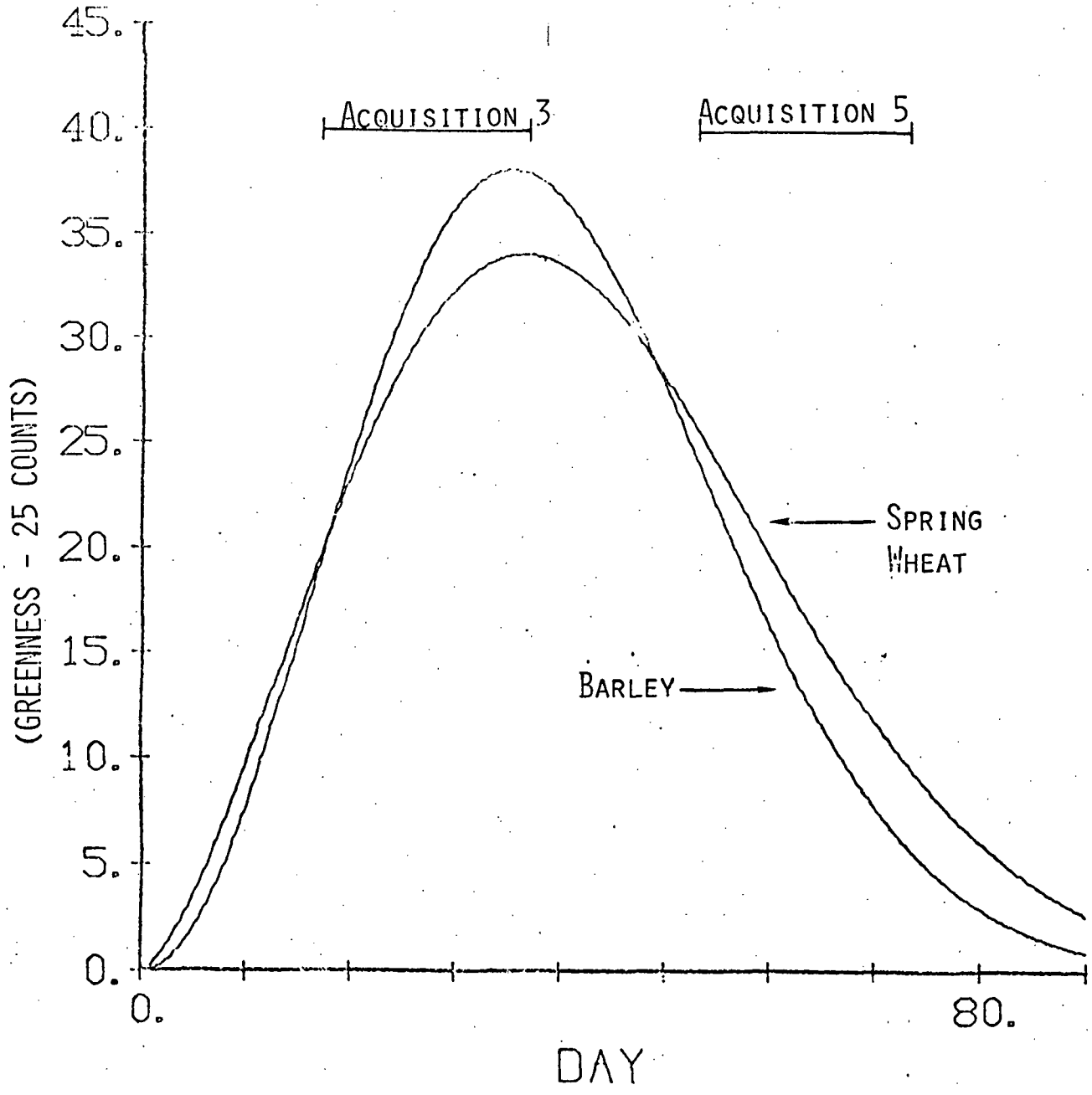
- SLOPE CHANGE IN TIME IS CONSISTENT WITH ACQUISITIONS AND LACIE CROP CALENDAR ESTIMATES FOR SEGMENT, E.G.,



- RANGE OF GREENNESS VALUES IS LARGE ENOUGH

FIGURE 4.6

DAY RANGE SELECTED FOR DISCRIMINATION - 90% OF SMALL GRAIN PIXELS



TASK 6
FORECASTING PRODUCTION OF WHEAT FROM SATELLITE DATA
(J. Colwell)

6.1 OBJECTIVE

The general objective of this task is to investigate the utility of Landsat, meteorological, and ancillary data for forecasting winter wheat yield.

6.2 APPROACH

During this report period we have concentrated on an analysis designed to make recommendations concerning future directions for crop yield forecasting. Initial activity was devoted to discussions between ERIM and JSC with respect to the focus and limits of this activity. Subsequent activity has involved review of pertinent literature and discussions with persons knowledgeable in the area of wheat yield forecasting.

6.3 TECHNICAL DISCUSSION

The analysis is proceeding, and documentation of that analysis has begun.

6.4 FUTURE PLANS

It is expected that a written report summarizing recommendations concerning future directions for crop yield forecasting will be completed during the next quarter.

TASK 7
STUDY OF MULTICROP SPECTRAL SEPARABILITY
(W. A. Malila, Task Leader)*

7.1 INTRODUCTION

This task addresses the extension of large area crop inventory technology to other important crops in addition to wheat. The objective is to conduct signature studies using currently available data to provide insights and identify potential problem areas for investigation when LACIE Transition-Year data sets become available.

7.2 APPROACH

The approach being followed is outlined in Table 7.1. Major emphasis is to be placed on corn and soybeans data available in Landsat data sets and from other sources.

7.3 SUMMARY OF PROGRESS

During the reporting period, we continued our analysis of Landsat data acquired during the CITARS project. In re-examining these data in light of understandings and new techniques resulting from LACIE, we carried out XSTAR preprocessing and Tasselled-Cap transformations of data from the Fayette and Livingston data sets. Landsat signatures of corn and soybeans were analyzed as a function of time, and clustering and classification studies continued. A simple discrimination test between corn and soybeans was devised and yielded high (90-95%) correct classification results on two CITARS data sets.

Limited amounts of corn and soybeans data exist in selected LACIE Phase 3 Blind Sites on the fringes of the U.S. Corn Belt region. Analysis of these data was initiated.

---

*J. Hemdal, J. More, and E. Crist contributed to the work reported herein.

Prior field measurements of corn and soybeans spectral reflectance were received from Purdue/LARS, but we were unable to begin analysis of them during the quarter.

7.4 DETAILED DISCUSSION OF PROGRESS

Last quarter, results of spectral clustering of field-center pixels of corn and soybeans were described. Multitemporal displays of two clusters from each crop are presented in Figure 7.1 (MSS 6 vs. MSS 5). The two corn clusters represent 93% of all corn pixels analyzed. Soil color differences for the two clusters are evident on the first acquisition date, June 10th; however, the next three dates reveal the fact that the vegetation obscures and/or shadows most of the soil and consequently dominates the signature. Similar effects are present for the two soybeans clusters shown. As noted last quarter, soybeans clusters were more numerous and varied than corn clusters -- these two largest accounting for only 37% of the total pixels. A final observation made from this figure is that the corn clusters drop in MSS 6 values from 17 July to 21 August, while soybean values continue to increase; similar trends were present in Tasseled-Cap Greenness values. The decrease for corn apparently is linked to tasseling.

Based on observations of distinctive features in the time profiles of Greenness for corn and soybeans (see Table 7.2), a simple test was formulated to distinguish between these crops. As defined in Table 7.3, a ratio of Greenness on two dates is computed and compared to a threshold value. The threshold level for ground-visited* fields in the Fayette data was established by histogramming the ratio for the collection of field means, noting a bimodality in the data, and selecting a threshold value (1.2) which separated the modes (see Figure 7.2). Results are presented in Tables 7.3 and 7.4 -- an average conditional correct classification

*CITARS ground "truth" had been obtained in two ways, one was by ground visitation and the other by photointerpretation of multidate aerial color photography.

of 93-95%. Comparable results were obtained for Livingston field means (Table 7.5), and the same threshold value applied (Figure 7.2).

Since field means are less noisy than individual pixel values, we decided to apply the test to Fayette pixel data as well. The ratio histogram in Figure 7.3 exhibits the same type of bimodality that was observed for field means. Results of the test on pixels (Tables 7.6 and 7.7) again show a high level of performance -- 93% combined correct classification.

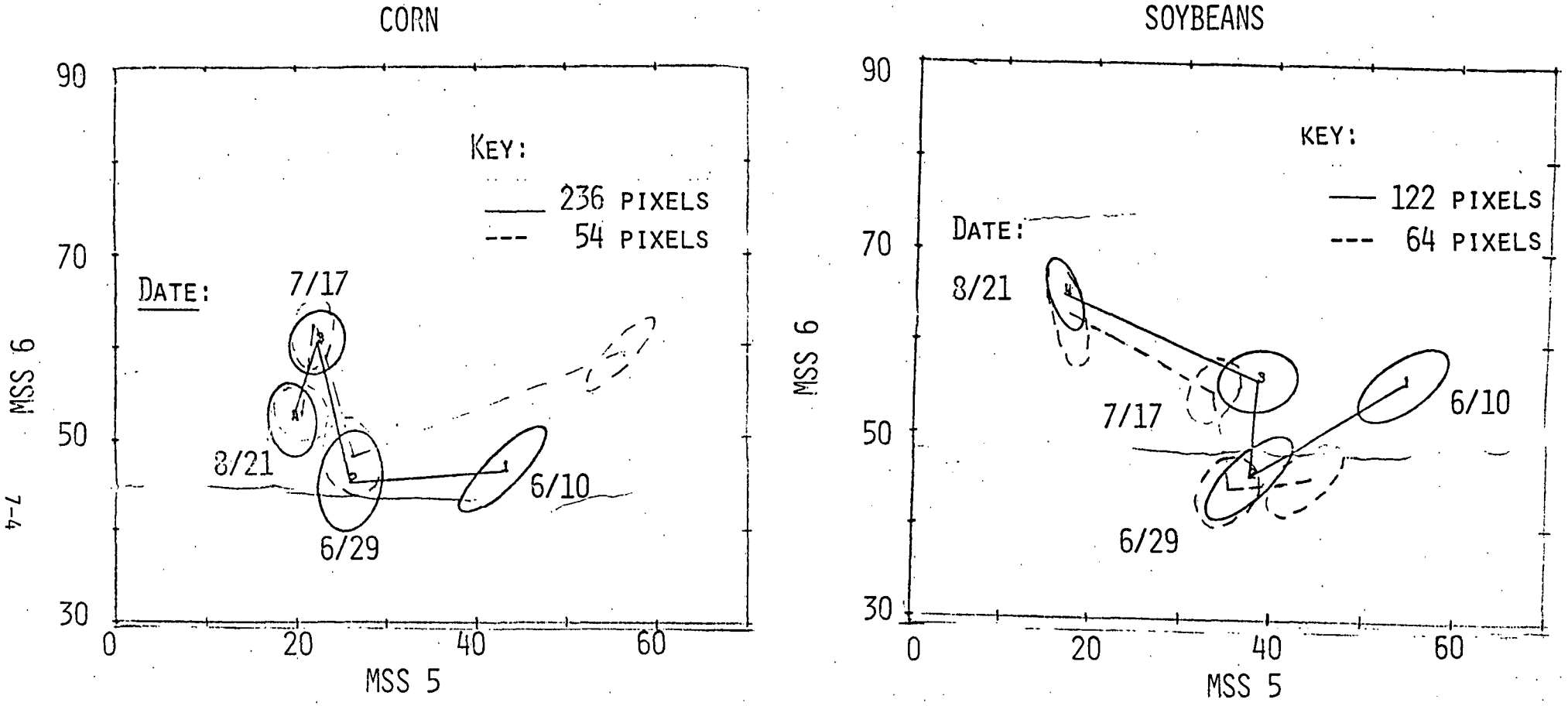
Our observations and conclusions are summarized in Table 7.8. The test developed is not intended to be a final procedure. Rather, it demonstrates that spectral differences are present and consistent in at least this limited data set and should serve as a basis for discrimination procedures in the future.

7.5 PLANS

Plans for the next quarter are presented in Table 7.9.

FIGURE 7.1

MULTITEMPORAL TRAJECTORIES OF MAJOR CORN AND SOYBEAN CLUSTERS
FAYETTE FIELD-CENTER PIXELS CITARS DATA



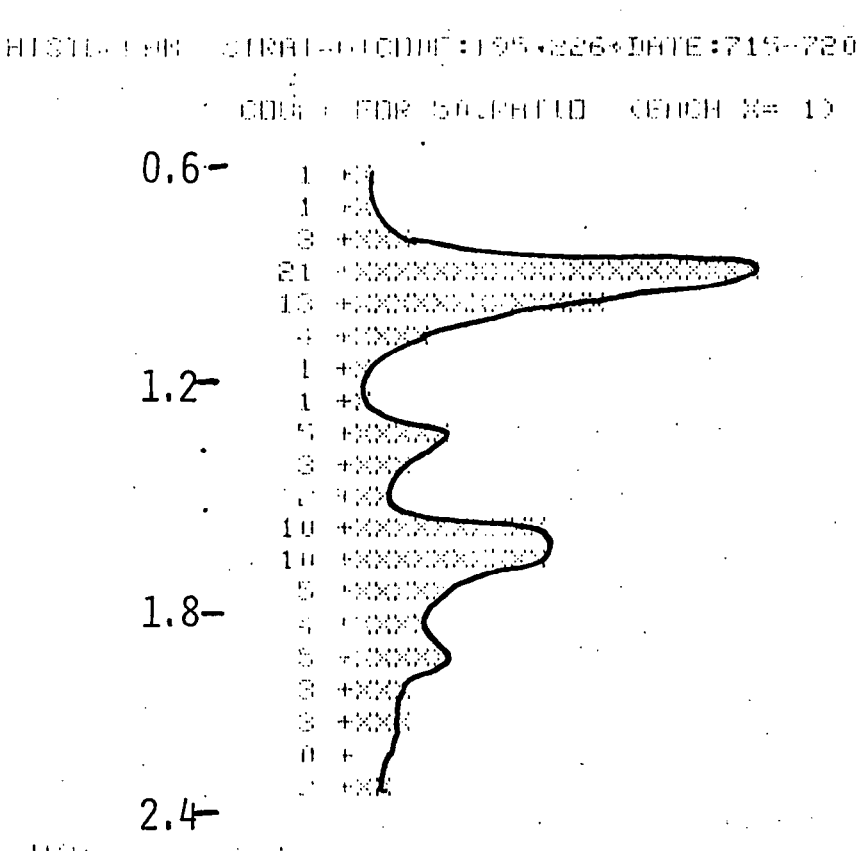
NOTES: 93% OF CORN PIXELS WERE IN THE TWO CLUSTERS SHOWN.

SOYBEAN CLUSTERS WERE MORE DIVERSE; THE ABOVE TWO REPRESENT 37% OF TOTAL.

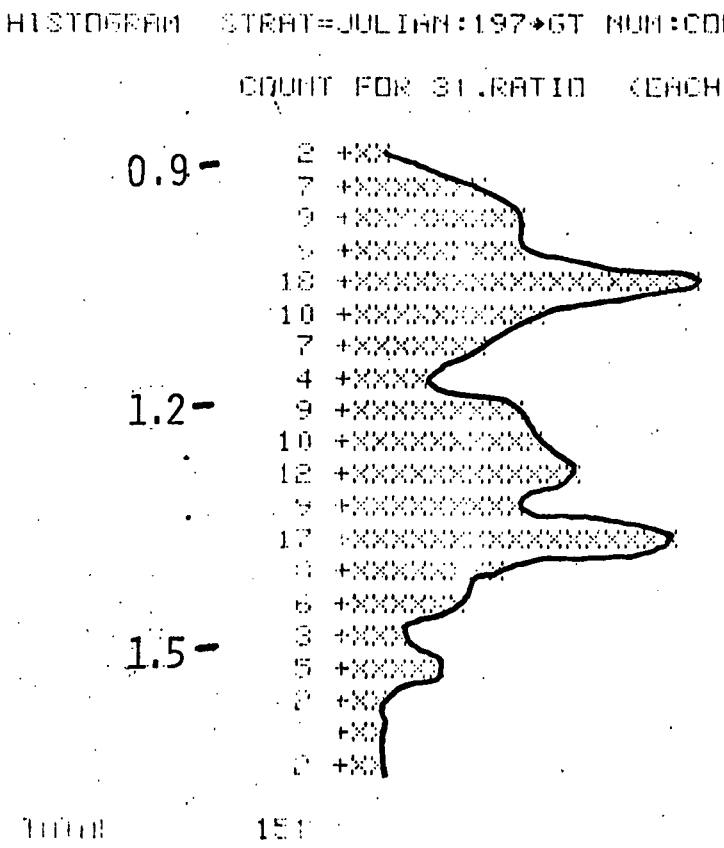
ERIM

FIGURE 7.2

HISTOGRAMS OF R RATIO SHOWING EVIDENCE OF BIMODALITY



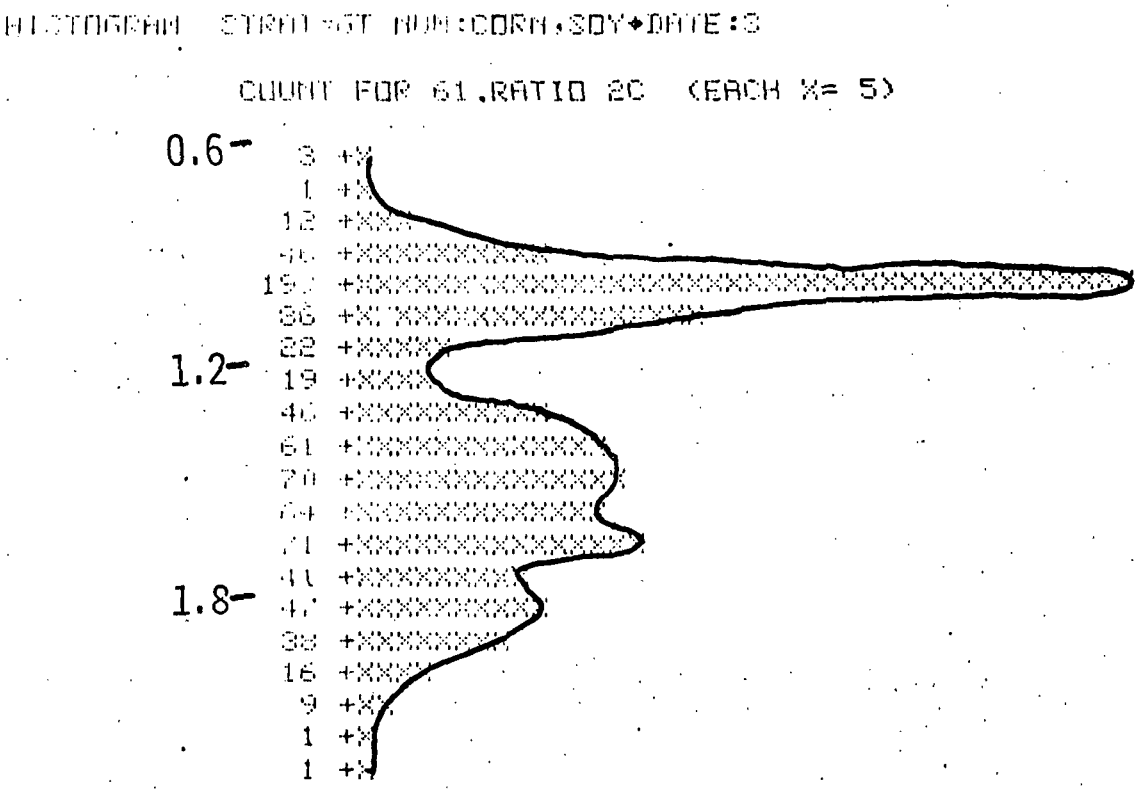
FAYETTE PI FIELDS



LIVINGSTON PI FIELDS

FIGURE 7.3

HISTOGRAM OF R RATIO SHOWING BIMODALITY



FAYETTE PIXELS

ERIM

TABLE 7.1

APPROACH

- Analyze Available Landsat Data on Corn and Soybeans
  - 1973 CITARS Data (Illinois and Indiana)
    Re-examine in light of understanding and new techniques gained through LACIE:
    - •• Structure of Landsat data (Tasselled Cap Transformation)
    - •• Spectral/spatial and Multitemporal Clustering
    - •• XSTAR Preprocessing (Correction for Sun Angle and Haze)
    - •• Quantitative use of Periodic Ground Observations
  - 1977 LACIE Blind Site Data (Selected Sites on Fringes of a U.S. Corn Belt)
- Analyze Available Data from Other Sources
  - Field Measurements of Spectral Reflectance
  - Simulation Model Data

ERIM

TABLE 7.2

DISTINCTIVENESS OF TIME PROFILES OF GREENNESS FOR CORN AND SOYBEANS

- CITARS FAYETTE DATA SHOWED THAT:
  - TIME TRENDS OF GREENNESS CONTAIN POTENTIALLY VALUABLE INFORMATION
  - SOIL COLOR EFFECTS WERE LESS EVIDENT IN GREENNESS THAN OTHER BANDS
  - SOYBEANS GREENNESS USUALLY PEAKED ON LAST (4TH) ACQUISITION (8/21)
  - CORN GREENNESS USUALLY PEAKED ON 3RD ACQUISITION (7/16)
  - TIMING OF CORN DECREASE APPARENTLY LINKED TO TASSELING
- SIMILAR TRENDS WERE OBSERVED IN LACIE SEGMENT DATA THAT WE EXAMINED
- THEREFORE, A SIMPLE TEST WAS FORMULATED TO DISTINGUISH BETWEEN CORN AND SOYBEANS (BASED ON RATIO OF GREENNESS VALUES ON TWO DATES)

ERIM

TABLE 7.3

SIMPLE TEST FOR CORN VS. SOYBEANS

- Compute Greenness Ratio:

$$R = \frac{\text{Greenness (Day 214 - 232)}}{\text{Greenness (Day 178 - 197)}}$$

- If $R \leq K$, Label as Corn

  $R > K$, Label as Soybeans

- Results on Fayette Ground-Visited Fields

Ground Truth

Assigned Class	C	S
C	7	2
S	0	12
Total Fields:	7	14

Ground Truth

Assigned Class	C	S
C	100%	14%
S	0%	86%

93% Average Conditional Percent Correct Classification

ERIM

TABLE 7.4

RESULTS FOR FAYETTE PI DATA

PI Field Means

		Ground Truth C	Ground Truth S
Assigned Class	C	39	0
	S	4	54
Total Fields:		43	54

		Ground Truth C	Ground Truth S
Assigned Class	C	91%	0%
	S	9%	100%

95.5% Average Conditional Percent Correct Classification

ERIM

TABLE 7.5

RESULTS FOR LIVINGSTON DATA

FIELD MEANS (GROUND VISITED)

	GROUND TRUTH C	GROUND TRUTH S
ASSIGNED C	8	0
CLASS S	1	18
TOTAL FIELDS	9	18

	GROUND TRUTH C	GROUND TRUTH S
ASSIGNED C	89%	0%
CLASS S	11%	100%
		94.5%

FIELD MEANS (PI)

	GROUND TRUTH C	GROUND TRUTH S
ASSIGNED C	58	5
CLASS S	6	82

	GROUND TRUTH C	GROUND TRUTH S
ASSIGNED C	90%	6%
CLASS S	10%	94%
		92%

TABLE 7.6

COMBINED RESULTS ON FAYETTE PIXELS (GROUND-VISITED AND PI)

	Ground Truth C	Ground Truth S
Assigned C	317	20
Class S	39	477
Total Pixels:	356	497

	Ground Truth C	Ground Truth S	
Assigned C	89%	4%	
Class S	11%	96%	
Combined:			93%
Ground Visited:			89%
PI Pixels:			94%

ERIM

TABLE 7.7

BREAKDOWN OF RESULTS FOR PIXELS-FAYETTE

GROUND-VISITED PIXEL SET

		Ground Truth C	Ground Truth S
Assigned Class	C	61	13
	S	9	126
Total Pixels		70	139

	C	S	
C	87%	9%	
S	13%	91%	
			89%

PI PIXEL SET

		Ground Truth C	Ground Truth S
Assigned Class	C	256	7
	S	30	351
		286	358

	C	S	
C	90%	2%	
S	10%	98%	
			94%

TABLE 7.8

SUMMARY

- Distinctive Characteristics of Multitemporal Signatures of Corn and Soybeans Have Been Noted
- A Simple Test on Greenness Profiles Performed Very Well in Discriminating Between these Crops
- More Testing and Development are Needed
- Refinements and Generalizations are Needed to
  - Account for Crop Calendar Differences
  - Discriminate Against Other Crops

ERIM

TABLE 7.9

PLANS

- Continue Analysis of Signature Characteristics, Trajectory Models, and Classifiers
- Apply Preprocessing Transformations To Other CITARS Segments And Analyze
- Analyze Correlation Between Periodic Agronomic Observations And Spectral Data

TASK 8
MULTICROP LABELING AIDS
(R. C. Cicone and R. Balon)

8.1 INTRODUCTION

The accuracy of operational large area crop inventory systems modeled after LACIE depends critically on the correctness of crop labels generated by Analyst Interpreters. Task 8 is being conducted as a partial response to a request from the multicrop inventory planning committee for support from the SR&T community in adapting LACIE technology to a multicrop environment. The critical issue addressed is that of analyst labeling.

8.2 OBJECTIVES

Our purpose through Task 8 is to analyze the methods of presentation of Landsat data for purposes of human interpretation to assess how well they convey the information relevant to crop discrimination. We endeavor to develop new data presentation techniques in the form of false color image products as well as graphic displays of spectral information, which stand to expedite correct labeling of crops.

8.3 APPROACH

Points of approach for this task are presented in Table 8.1. We have been progressing along a path with the following three major stages:

1. Implement a model of the color production characteristics of the system used to produce the film products, taking account of the physical considerations of film production and the psychological/perceptual considerations of the fact that the film is interpreted through human visual processes.
2. Analytic evaluation of LACIE film products with regard to information content and between scene stability of color meaning.

3. Development of a frame work for new image products which take advantage of the knowledge which has been gained about color production.

In another branch of approach we have been working toward the preparation of color keys for corn and soybean crop cycles in order to assist the Analyst Interpreters in the move to multicrop inventory.

8.4 SUMMARY OF PROGRESS IN PREVIOUS QUARTERS

Table 8.2 describes our progress to date. A model of the color production of the Production Film Converter after Juday of NASA/JSC has been implemented. The model features representation of colors in the coordinates of a space designed to be uniform with respect to human perception of color differences. Figure 8.1 shows the volume of color space which the PFC can represent. Table 8.3 describes this space and Table 8.4 diagrams the mapping from Landsat channel signals to color space.

Analysis of LACIE image products with regard to sensitivity and information content has been completed. It has been found that non-linear characteristics of the PFC and of color vision yield, under existing color mappings, a distorted view of Landsat data to the eye. That is to say the amount of color expansion afforded to the data is not uniform throughout the range of the data. The following experiment serves to illustrate this point. A regular sample of points is taken from the plane of data concentration for an acquisition (the Kauth Brightness/Greenness Plane) and mapped to color space coordinates. A spherical distribution is constructed about each of these color space points. Figure 2 shows this schematically. The distribution is then transformed by the inverse mapping from color to data space. A display of these distributions in data space show directly how the data is broken up into classes by color in the image. Figure 3 presents a particular example of this display technique. Each ellipse boundary represents one distinguishable color or class on an image. Where the ellipses overlap, the data is not being

"resolved" to the fineness of the sample grid. The spacing of the sample grid is based on liberal estimate of 3 counts of noise in Landsat data. Where ellipses in the sensitivity plot (Figure 3) do not overlap, the image is sensitive to less than a 3 count difference between data points. Figure 4 overlays Figure 3 to show the distribution of data to which the sensitivity ellipse plot applied. Note that in the area of data concentration there is undersirable overlapping of ellipses (image classes). From examining acquisitions of a number of sample segments in this manner we conclude that resolution of the data into color classes is generally less than desireable whenever the data fills out the full range of an agricultural scene. Elognated ellipses indicate that distance relationships in the data are not preserved by the image, i.e., the image presents a distorted view of the data to the eye. The pattern of distortion as it appears in Figure 3 is similar from one acquisition to another. Further explanation of the sensitivity study and the color production model will be found in an interim technical memorandum to be released next quarter.

An analysis of an image in terms of classification accuracy was performed to test retention of information from the original 4 channel of Landsat data to the false color image. Table 8.5 shows classification accuracy of wheat vs. non-wheat at steps along the data-to-perception conversion process. Significant loss of information occurs in dropping a channel of information and because of the less than desirable level of data to color expansion (resolution). Overall wheat recognition drops from 97.2% to 81.8% in the complete process. These experimental results support the conclusion drawn from sensitivity analysis, namely that currently used image products do not display the full information content of the data.

Work last quarter centered on assessment of image color consistency between scenes for LACIE image products and on establishing a framework

for definition of Uniform Color Space image products. These items are discussed in the next section.

8.5 DESCRIPTION OF PROGRESS IN THIRD QUARTER

LACIE Film products are generated from Landsat channels which are prepocessed (scaled and biased) in order to improve the contrast present in the image produced. The parameters of the transformation for each channel are calculated anew for each Landsat scene based on the mean and standard deviation of data points for each channel. This initial, inconsistent manipulation of the data has inherent dangers to it. The upshot is that a point of data space will not be mapped to one color consistently. Instead, the color of a data point varies depending on the statistics of the scene. This creates problems in interpretation of imagery. Color inconsistency has in some instances led to unacceptables discrepancies between interpretation and ground truth. Alternative film products have been proposed with the idea of reducing color inconsistency between scenes. One of these alternatives, the "Kraus Product", has been adopted as a supplement to the original product, "Product One". However, the alternative product is essentially very similar to product one in that it also applies a preprocessing to the data which depends on scene statistics. An experiment was performed to assess the between scene color consistency of Kraus Product imagery. The Kraus Product calculates a preprocessing scale factor for each channel based on a parameter μ which measures the average brightness of the scene across three channels. To evaluate the consistency of these product, we allow the scene mean to vary throughout the normal range of agricultural data, which is represented by the Tasselled Cap of the Kauth Brightness/Greenness plane. For each of these hypothetical scene means we calculate the color of a fixed reference point (Figure 8.5). This simulates the appearance of the reference point in wide range of scenes. The color will change and the degree of change is expressed by the distance in Uniform

Color Space units between the colors. Five UCS units is a just noticeable difference, fifteen units is an easily distinguishable difference and fifty units is a highly contrasting difference between colors. As our reference color we use the color of our reference data point when the scene mean coincides with it. Figure 8.6 portrays the color difference strata which result from this experiment. As one would expect, the stratification occurs primarily in the Brightness direction. This is because the Kraus Product uses an average scene brightness as the sole parameter in preprocessing the data. The important point to be taken from this figure is that large color inconsistency occurs in Kraus imagery when average Brightness is different from one scene to another. For example, as the scene mean moves away from the reference point by dropping a few counts in Brightness (moving to the left in Figure 6) the color of the reference point changes by 12 UCS units, a visually discernible difference. If the scene mean is moved downward in Brightness about 15 counts the resulting color change in the reference point comes to 53 UCS units. The original color and the new color may be characterized as highly contrasting. The color on a Kraus Product image is insensitive to the scene average Greenness coordinate for a constant Brightness. However, the color representation of a particular data point varies dramatically with the scene average Brightness. The color attributes examined included hue, saturation and lightness. A similar pattern, as displayed in Figure 8.6, was observed in restricting the analysis to chromatic attributes only, i.e., hue and saturation. Considering the fact that external effects such as light haze level and sun angle are responsible for significant variation in average scene Brightness, one may commonly find inconsistent color representation in the Kraus Product.

During this quarter, preliminary work has been done on design of new image products. The justification for this work is summarized in Table 8.6. The new products take advantage of the known structure of

Landsat data and the known structure of the PFC's color space to improve upon the shortcomings of information distortion and color inconsistency inherent in LACIE film products. The XSTAR spatially varying haze correction algorithm will be used to standardize representation of Landsat data in Brightness, Greenness and Yellowness coordinates. The standardized Brightness/Greenness plane can then be mapped by linear transformation onto a chosen plane of Uniform color space. Choice of the color space plane to use will be arrived at by experimentation and interaction with Analyst-Interpreters. Figure 8.7 schematizes one possibility.

8.6 PLANS

Our plan for this quarter is to experiment with Uniform Color Space image product design and produce examples of imagery for inspection. We also plan to produce color keys for currently used products to aid Analyst-Interpreters in the move to multicrop inventory. Reports to be released include:

1. Uniform Color Space Analysis of LACIE Image Products
2. AI Color Key for Multicrop Inventory
3. Final Report

TABLE 8.1

APPROACH

- USE PRINCIPLES OF COLOR SCIENCE TO MODEL THE COLOR PRODUCTION CHARACTERISTICS OF THE PFC AND THE PSYCHO-PHYSICAL PROCESS OF NORMAL COLOR VISION
- ANALYZE TECHNIQUES FOR PRODUCING FALSE COLOR IMAGERY FROM LANDSAT DATA IN TERM OF
  - WITHIN IMAGE CHARACTERISTICS
    - SPECTRAL INFORMATION CONTENT
    - VARIATION OF COLOR SENSITIVITY IN IMAGE DATA SPACE
  - BETWEEN IMAGE CHARACTERISTICS
    - EVALUATE IMAGES IN TERMS OF COLOR DISTORTION
    - EVALUATE IMAGES IN TERMS OF COLOR CONSISTENCY
- EVALUATE TRENDS OF COLOR VARIATION FOR CORN AND SOYBEANS AS A FUNCTION OF DATA
- EXAMINE ALTERNATIVE SINGLE AND MULTI-DATE IMAGE PRODUCTS

TABLE 8.2

PROGRESS TO DATE

- IMPLEMENTED THE JUDAY PFC MODEL AND L*, A*, B* UNIFORM COLOR SPACE MODEL
- CONDUCTED ANALYSIS OF CHARACTERISTICS OF
  - PRODUCTION FILM CONVERTOR
  - PRODUCT 1 IMAGE PRODUCT
  - KRAUS PRODUCT AND ALTERNATIVE IMAGE PRODUCTS
- ANALYZED BETWEEN-IMAGE COLOR CONSISTENCY
  - PRODUCT 1 COLOR TRAJECTORIES
  - KRAUS PRODUCT
- DEVELOPED FRAMEWORK FOR ALTERNATIVE IMAGES
- INITIATED EVALUATION OF COLOR CHARACTERISTICS OF CORN AND SOYBEANS AS FUNCTION OF DATE

TABLE 8.3

CHARACTERISTICS OF L^*, A^*, B^* SPACE

- L^*, A^*, B^* IS A CURVILINEAR TRANSFORMATION, OF THE CIE 1931 XYZ SPACE, DESIGNED FOR PERCEPTUAL UNIFORMITY
- LIGHTNESS AXIS L^*: COLORS OF CONSTANT L^* VALUE ARE PERCEIVED AS BEING EQUALLY LIGHT
  - AS L^* INCREASES FROM 0 → 50 SPECTRUM LOCUS EXPANDS CHROMATICITY DIFFERENCES ARE MORE PERCEPTIBLE IN LIGHT COLORS
  - AS L^* INCREASES FROM 50 → 100, THE BOUNDARY OF ATTAINABLE OBJECT-COLORS SHRINKS, APPROACHING THE COLOR OF THE ILLUMINANT
- CHROMATICITY AXES (A^*, B^*) FORM A CHROMATICITY DIAGRAM DIFFERENT AT EACH L^* VALUE
- DEFINING EQUATIONS

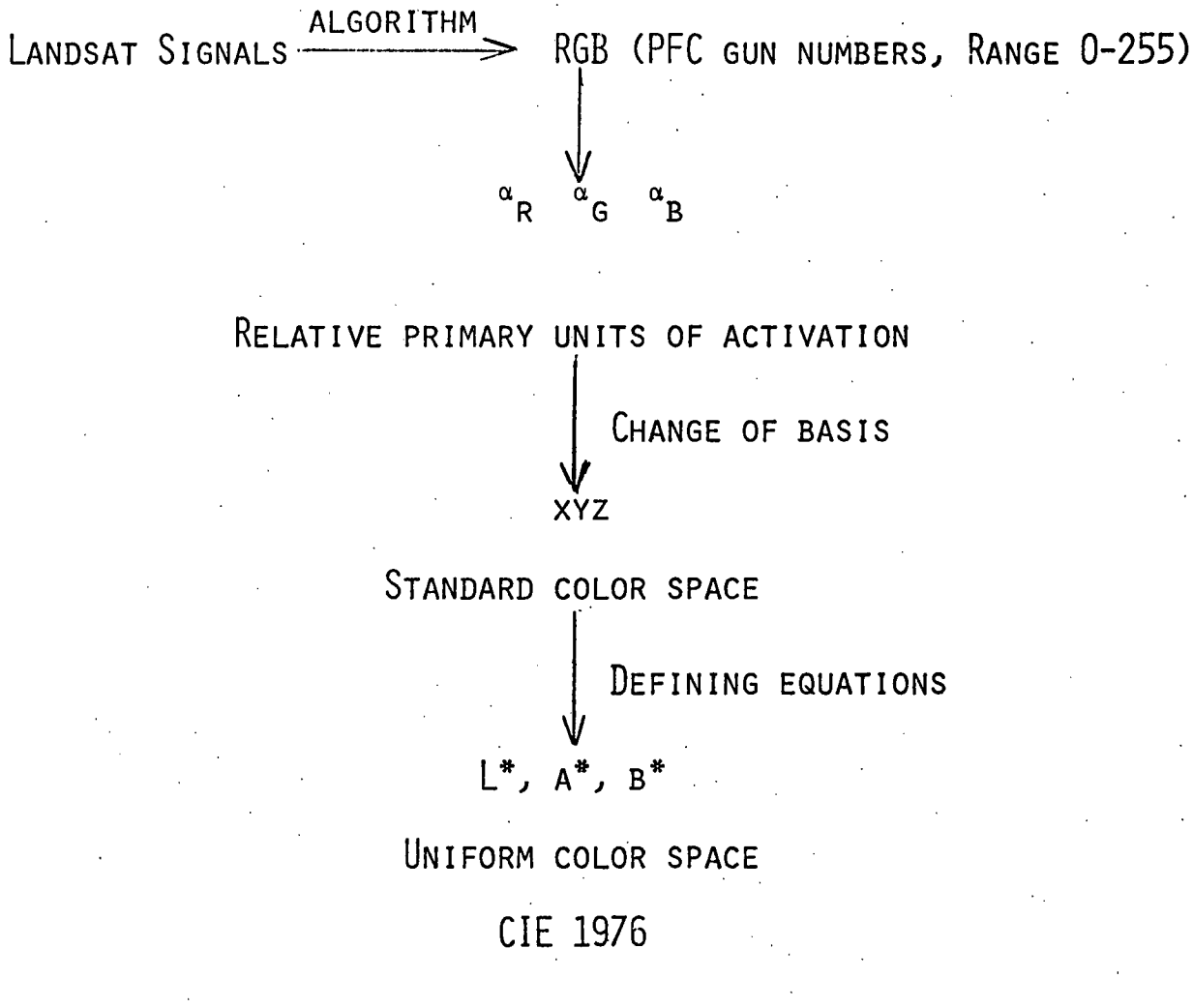
$$L^* = 25(100\ Y/Y_0)^{1/3} - 16$$

$$A^* = 500\left[(X/X_0)^{1/3} - (Y - Y_0)^{1/3}\right]$$

$$B^* = 200\left[(Y/Y_0)^{1/3} - (Z - Z_0)^{1/3}\right]$$

TABLE 8.4

MAPPING LANDSAT SIGNALS TO L*, A*, B* SPACE



ERIM

TABLE 8.5

IMAGE INFORMATION CONTENT

Coordinate Space	Percent Correct Classification	
	Wheat	Non-Wheat
Landsat 4-Channels	97.2	94.6
Landsat 3-Channels (Bands 4, 5, 7)	87.3	95.2
Color Space With 0 Count Resolution	84.8	95.0
Color Space With 5 Count Resolution	81.8	91.6

TABLE 8.6

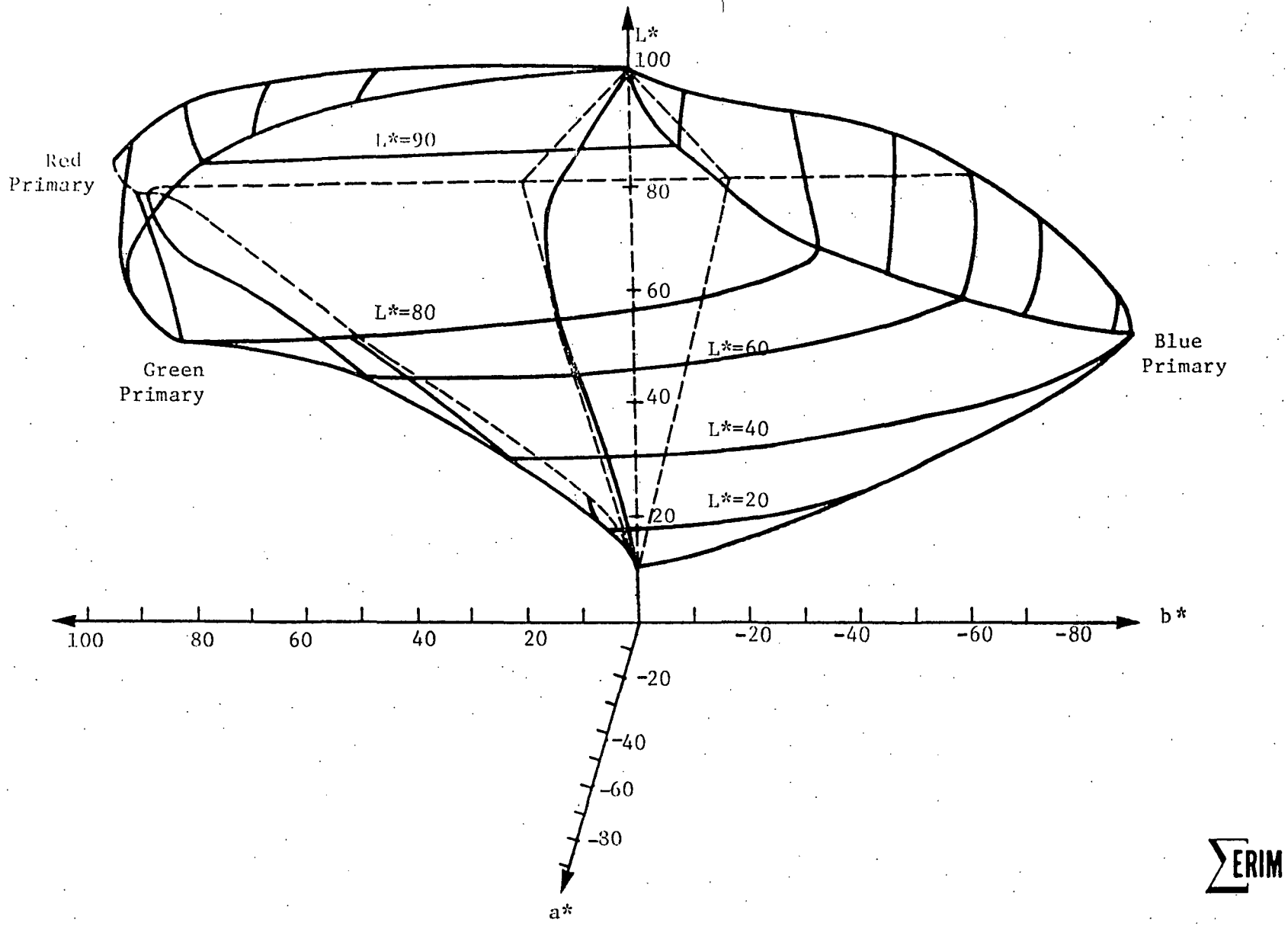
JUSTIFICATION FOR UNIFORM COLOR MAPPING

- A Scene Dependent Linear Transformation Of Landsat Data Into PFC Colorgun Counts Cannot Preserve Information Content And Simultaneously Preserve Color Consistency

- Images That Are Color Consistent, Maximize Content Of Information Displayed, And Conform To Currents AI Interpretation Techniques Can Be Produced Using Uniform Color Mapping Techniques.

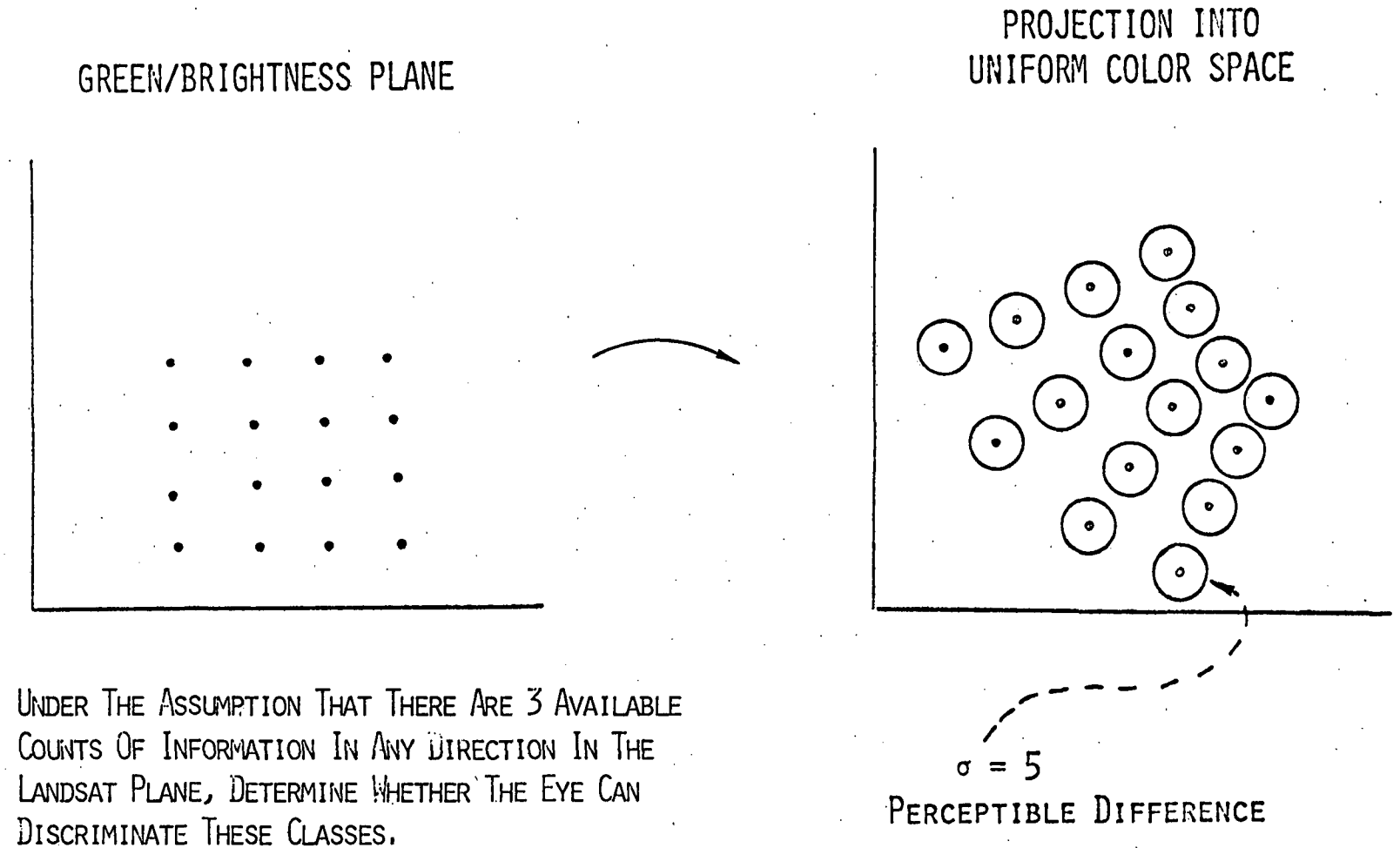
FIGURE 8.1

BOUNDARY OF COLORS ATTAINABLE BY THE PRODUCTION FILM CONVERTER IN L*,A*,B* UNIFORM COLOR SPACE



ERIM

FIGURE 8.2



ERIM

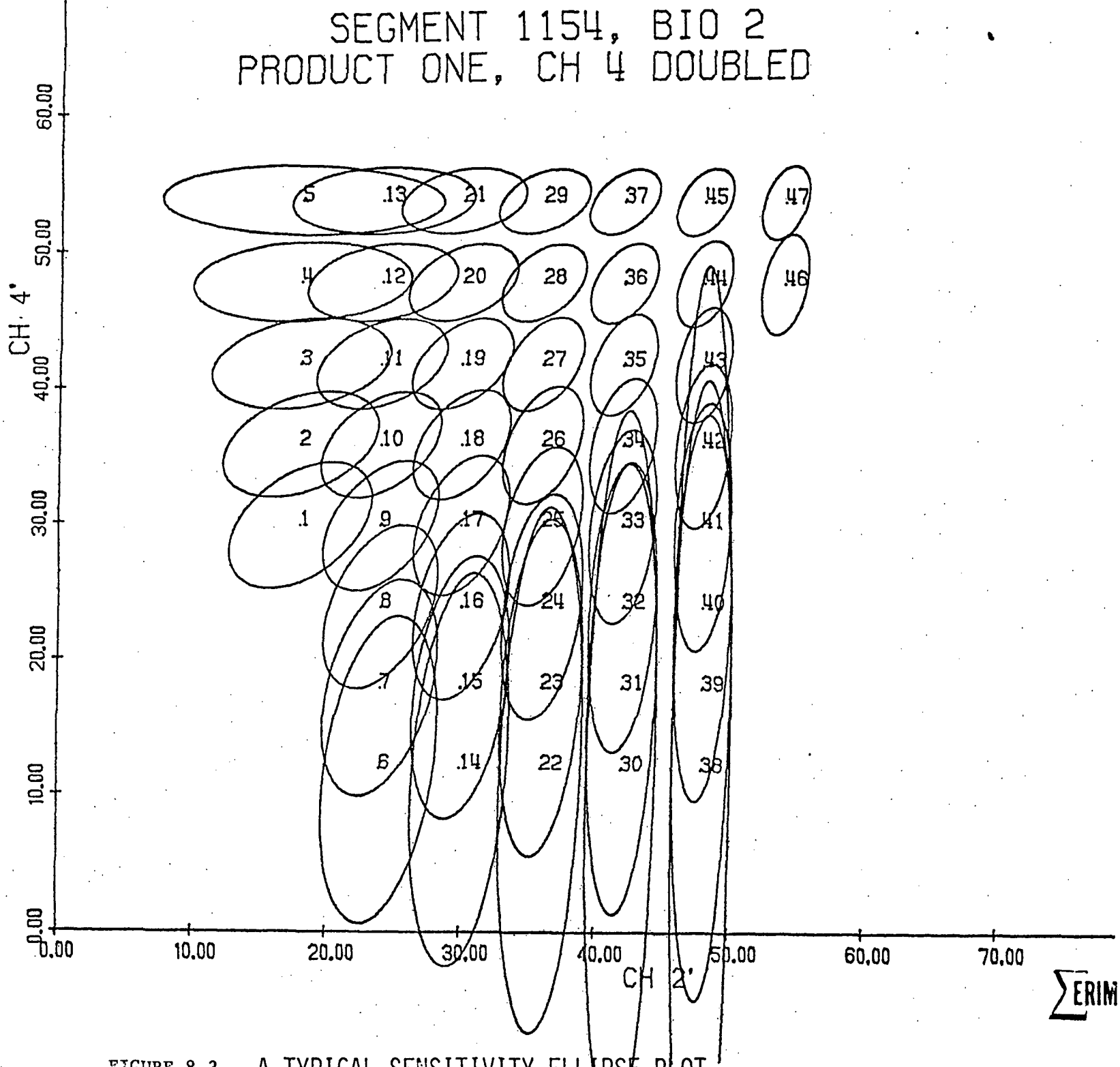
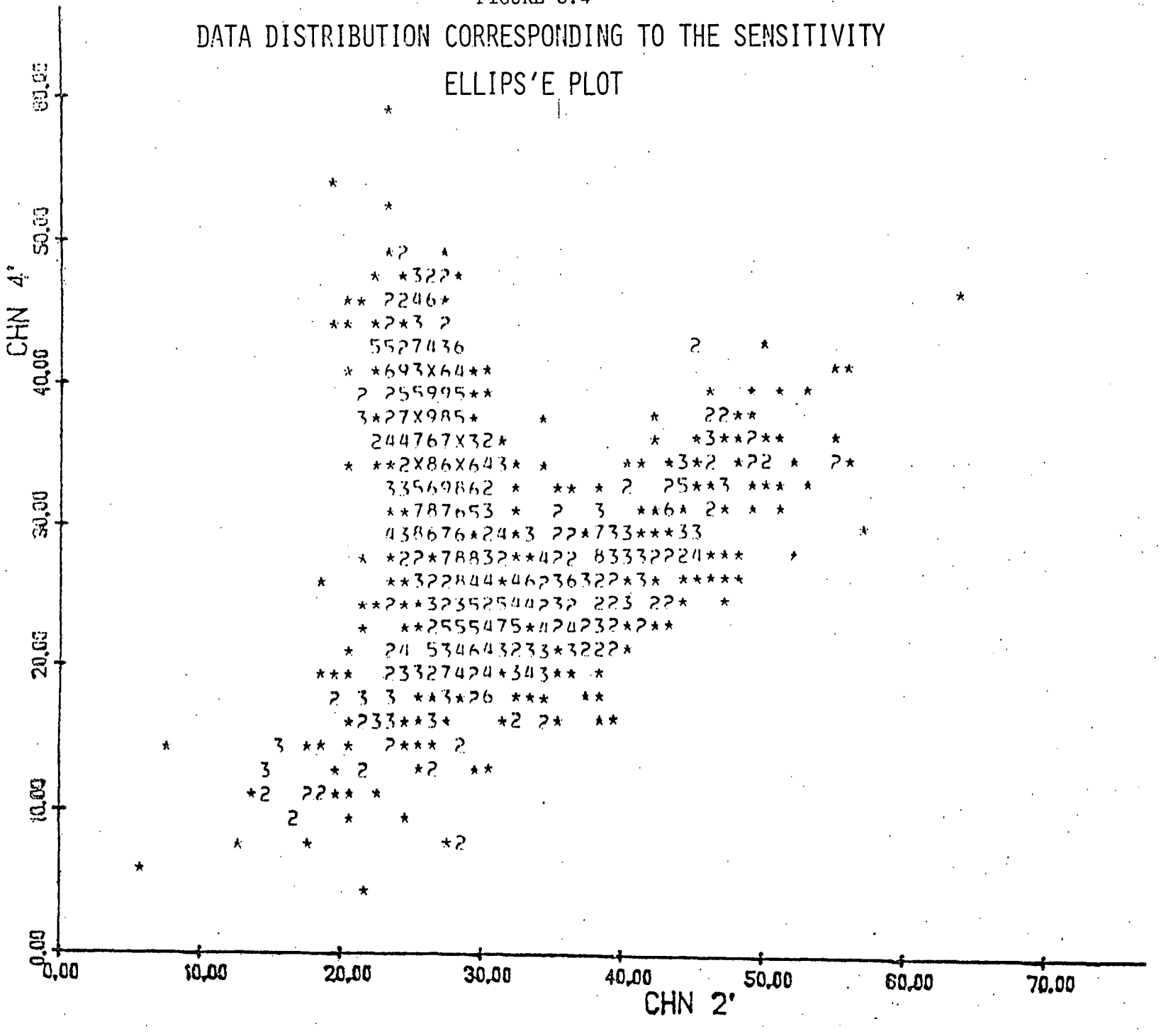


FIGURE 8.3 A TYPICAL SENSITIVITY ELLIPSE PLOT

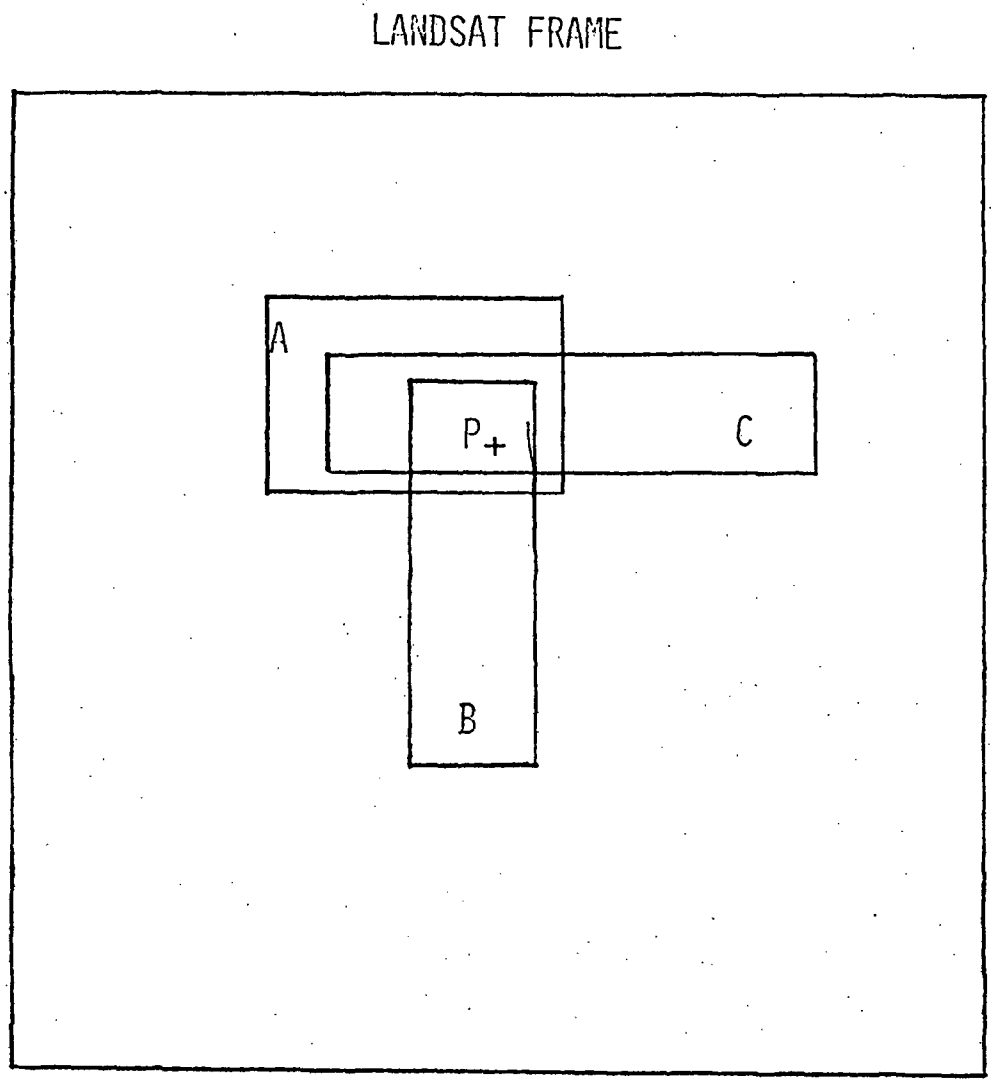
FIGURE 8.4

DATA DISTRIBUTION CORRESPONDING TO THE SENSITIVITY ELLIPS'E PLOT



ERIM

FIGURE 8.5



DETERMINE COLOR OF PIXEL 'P' IN THE CONTEXTS OF SCENES A, B AND C.

ERIM

FIGURE 8.6

COLOR CONSISTENCY IN KRAUS PRODUCT
(16 LEVEL QUANTIZATION OF PFC)

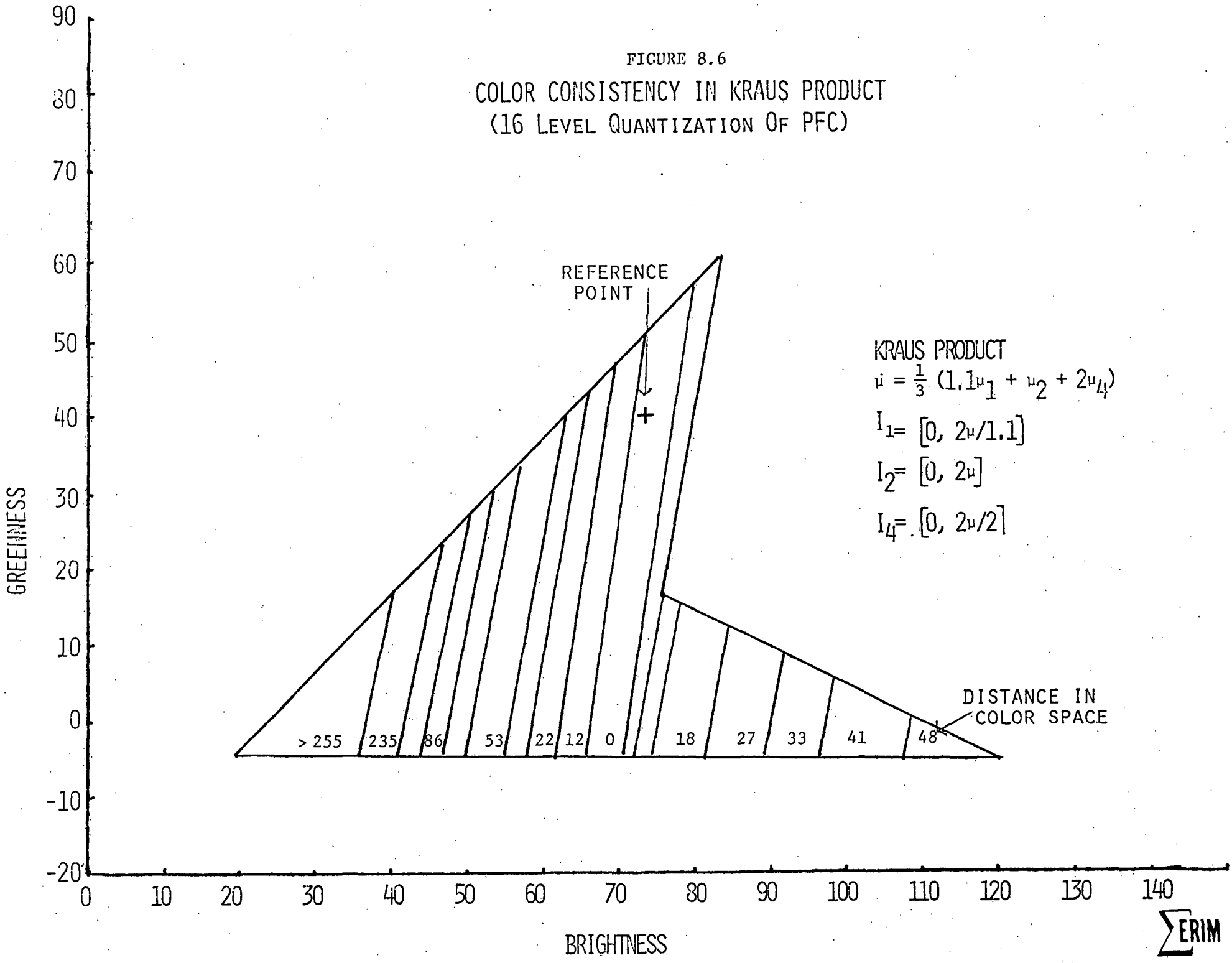
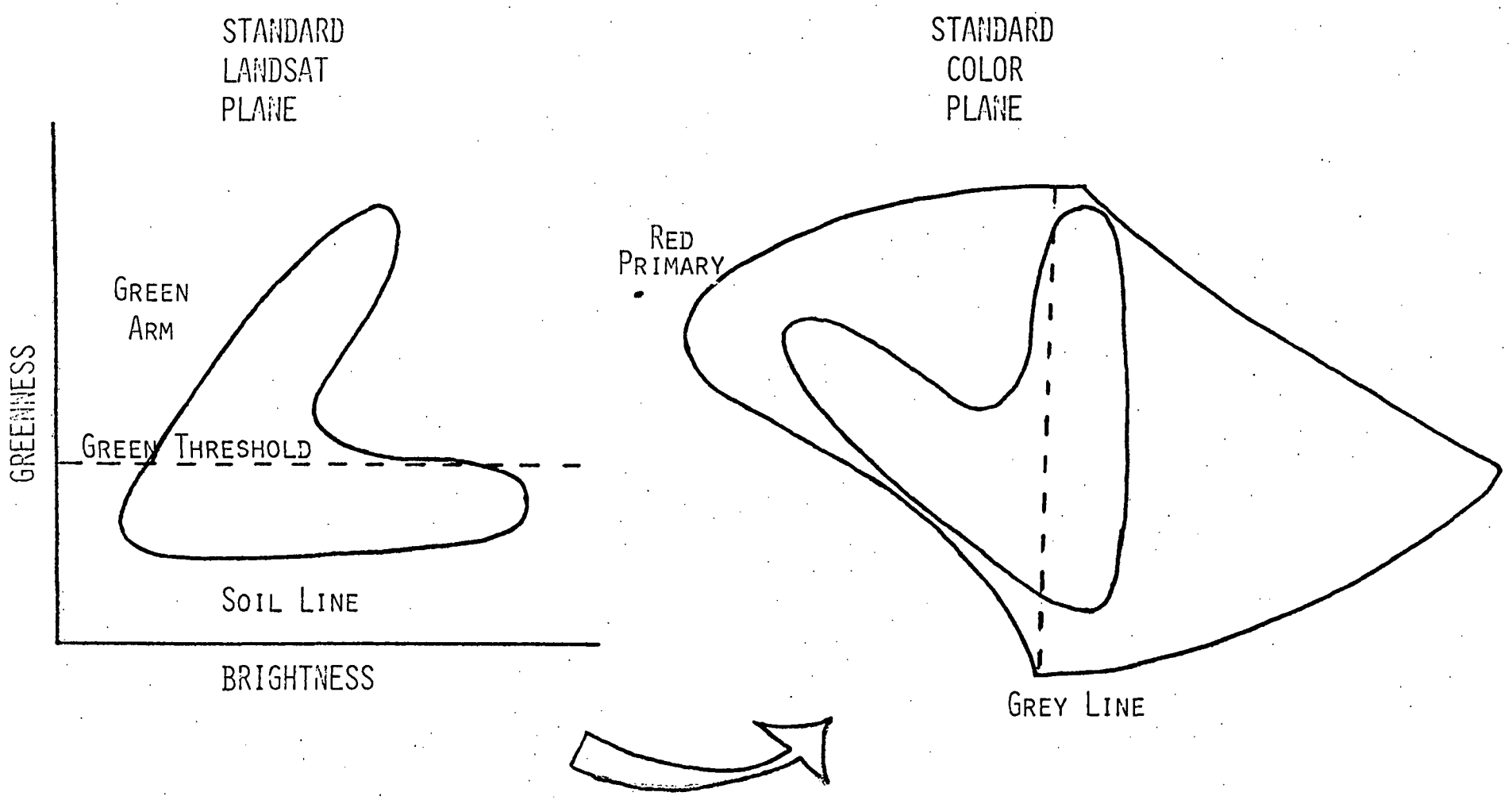


FIGURE 8.7

UNIFORM COLOR SCALE IMAGE PRODUCT



TASK 8b
MULTICROP LABELING AIDS
(P. F. Lambeck)

8b.1 INTRODUCTION

The analysts' ability to generate crop labels, the machine's ability to classify data correctly, and the researchers' ability to devise improved data analysis procedures all depend on the quality and consistency of the data which is analyzed and processed. During the second quarter we identified localized atmospheric variations within the boundaries of LACIE segments as a serious detriment to continued progress in developing labeling aids for analysts and in developing other new or improved analysis techniques. Since these localized atmospheric variations were not addressed by the XSTAR algorithm as implemented at that time, and since a viable means to remedy this shortcoming had already been proposed, a portion of this quarter's effort was focussed on this crucial problem.

8b.2 OBJECTIVE

The objective of this effort was to develop a spatially varying XSTAR haze correction procedure to compensate Landsat data for localized atmospheric distortions so that more uniform observation conditions could be simulated. The key issues addressed by this effort are listed in Table 8b.1.

8b.3 APPROACH

The approach that was taken to developing the spatially varying XSTAR haze correction is outlined in Table 8b.2. The first two items listed in the table are the obvious ones, however the last two items also had to be considered in order to arrive at a procedure which was sufficiently general and practical. The manner in which the spatially varying XSTAR correction was made to fit these latter two constraints is discussed in section 8b.4.

8b.4 PROCEDURE

The previous "global" or "spatially invariant" XSTAR procedure has been defined in Reference 8b.1. The spatially varying procedure which was developed during this quarter is outlined in Table 8b.3.

The first step of the spatially varying XSTAR correction is to recalibrate and screen the data as in Reference 8b.1 but with two of the SCREEN threshholds made less stringent by shifting them in the negative Tasselled Cap yellow direction. Specifically, the second equation under step 3b and the second equation under step 3c are modified as follows:

$$z_3 + z_1/10. < z_{Cmin} + z_{Cbias} \quad (z_{Cmin} = -7.5) \qquad (8b.1)$$

$$z_3 + z_1/7. < z_{Hmin} + z_{Cbias} \quad (z_{Hmin} = -3.25) \qquad (8b.2)$$

(For the examples shown later in this report we have used $z_{Cbias} = -1.$, however subsequent testing indicates that $z_{Cbias} = -1.7$ is perhaps the optimum value to use.)

For the global XSTAR correction, the SCREEN threshholds z_{Cmin} and z_{Hmin} were used to edit out dense haze or cloud areas which were atypical of normal scene conditions and which were thus more likely to represent localized atmospheric variations. However, some of this variation which was screened out was potentially correctable by a spatially varying XSTAR procedure. Hence, for the new procedure these threshholds have been relaxed somewhat. (The threshholds separate dense haze and clouds from bright fields.) Relaxing these threshholds also allows the spatially varying XSTAR correction to see more of the spatial variations and to track them more accurately.

The second step of the spatially varying XSTAR procedure is to divide the scene into 5 line by 5 pixel blocks, and to calculate a floating point mean value for each block, using only signal values

from pixels within the blocks which pass the SCREEN procedure (i.e., "good" pixels). Mean values for blocks with no good pixels or with fewer good pixels than half the average number of good pixels per block (truncated to integer form) are not used. For these "unknown" blocks, mean values are estimated by interpolating or extrapolating from neighboring block mean values, as described in steps 3 and 4, below.

The result of the above two steps is indicated for an example scene by the block code map on the left in Figure 8b.1. In this map, numeric values indicate blocks with "known" mean values at this stage of processing. The numbers are the last digit of the average Tasselled Cap yellow value, after rounding to an integer form such that the average yellow value for the whole scene (using only good pixels) ends with the digit "5". The digit "6" thus represents blocks with an average yellow value in the range $-11.9 \pm .5$, while the digit "4" represents blocks with an average yellow value in the range $-13.9 \pm .5$. In general, less "yellow" indicates more haze. The characters "C", "H", and "W" represent blocks with unknown mean values at this stage of processing whose "non-good" pixels are mostly "cloud", "dense haze" or "water", respectively. The character "?" is used for blocks with all good pixels, but with an insufficient number of good pixels because the blocks overflow the boundaries of the scene.

In the third step of the spatially varying XSTAR procedure, the mean values of the 5 line by 5 pixel blocks are smoothed, using a non-recursive moving window filter. The filter presently used (and recommended) has a window size of 5 blocks by 5 blocks, with a three block distance between half amplitude points in both the along-track and across-track directions. The weighting factors used in this filter are given in Table 8b.4. These weighting factors correspond to along-track and across-track cross-sections of the filter weighting function which are gaussian from the peak of the half amplitude points and which are mirror-image gaussian from the half amplitude points to zero.

In this third step, smoothed mean values are calculated for all blocks with "known" mean values, and for all blocks with at least one near neighbor (either along-track or across-track) which has a "known" mean value. The smoothed mean value is the weighted average of the available "known" mean values within the window of the filter.

The result of this third step is indicated for the example scene by the block code map on the right in Figure 8b.1. In this map zeros indicate blocks with smoothed mean Tasselled Cap yellow values within ± .5 counts of the average for the whole scene. Minus signs correspond to less yellow smoothed mean values (indicating more haze) and plus signs correspond to more yellow smoothed mean values (indicating less haze). Other characters indicate blocks which still have "unknown" mean values. These are addressed in step 4.

The combination of steps 2 and 3 above is equivalent to passing a 25 line by 25 pixel moving window filter over the scene, and then sampling every 5th line and every 5th point to obtain the smoothed block mean values. The along-track or across-track cross-section of this equivalent filter weighting function, and its frequency response are plotted in Figure 8b.2. The distance between the half amplitude points of this equivalent filter is 15 pixels. The procedure described in steps 2 and 3 is considerably less expensive than applying the equivalent filter directly to the scene. The lumps in the frequency response for this procedure, visible to either side of .2 and .4 cycles per pixel, are caused by the block averaging procedure in step 2. If we are to economize through block averaging, they cannot be avoided. Using a smaller block size would shift the lumps toward higher frequencies and would decrease their amplitude slightly. A block size of 1 line by 1 pixel would shift the lumps off the end of the frequency scale (beyond one half the sampling rate of Landsat).

Although these lumps pass frequencies which may cause artifacts in the corrected data, their amplitudes seem to be small enough so that these artifacts would not be significant.

Step 4 of the spatially varying XSTAR procedure is used to assign smoothed mean values to blocks which still have "unknown" mean values after step 3. In this step only those blocks which have "unknown" mean values, but which have at least one near neighbor (either along-track or across-track) with a smoothed mean value, are assigned smoothed mean values according to the procedure of step 3. Step 4 is iterated until all blocks have smoothed mean values.

The end results of step 4 for the example scene, and for the same scene on the following day, are indicated in Figure 8b.3. Note the dissimilarity of the block code maps for the two days shown in the figure. This indicates that there were significant spatial variations in the atmospheric condition on at least one of the two days (in this case, mostly on the first day). The spatial variations in the block average yellow values on day 140 (the second day, which was relatively clear) are not necessarily due to atmospheric variations. Such spatial variations can also be caused by the composition of the scene. In particular, soil color, moisture, canopy density, canopy yellowing, and land use (e.g., urban areas) are suspected to influence the Tasselled Cap yellow component in minor amounts. We do not know to what extent these yellow variations are detrimental to the performance of XSTAR in crop survey applications. An evaluation of this question would have to be the subject of a future research effort, however we believe that the benefits of the spatially varying XSTAR correction will significantly outweigh the detriments.

In step 5 of the spatially varying XSTAR procedure, the multiplicative and additive correction factors appropriate for each block mean value are calculated. This is done as specified in Reference 8b.1, with each block mean used as the value for $\hat{x}$ (steps 5 and 6 in the reference).

Step 6 is the final step in the spatially varying XSTAR procedure. In this step the multiplicative and additive correction factors calculated for each block mean in step 5 are interpolated between block

centers (in two dimensions) to determine the appropriate correction factors for each pixel. These correction factors are then applied pixel by pixel. Pixels which are near the borders of the scene, so that only one or two block centers are within the interpolating range (±4 lines and ±4 pixels) of the pixel, are corrected by interpolating the correction factors calculated for those blocks which have centers within the interpolating range.

The interpolation used in step 6, above, is not a linear interpolation (although a procedure which interpolated linearly in the along-track and across-track directions could have been used). Instead, an interpolating filter is used. The weighting factors for this filter are specified in Table 8b.5. The weighting factors correspond to along-track and across-track cross-sections of the filter weighting function which are gaussian from the peak to the half amplitude points and which are mirror-image gaussian from the half amplitude points to zero (generally similar to the smoothing filter which was applied to the block mean values). The distance between the half amplitude points of this filter is 5 pixels in both the along-track and across-track directions. Note that the sum of those weighting factors which are located on a 5 line by 5 pixel grid is always unity (for proper steady-state response of the filter). When the filter is applied, the multiplicative and additive correction factors calculated for each block are associated with the pixel at the center of the block, and zeros are associated with all other pixels. Applying the filter (non-recursively) to the correction factors associated with the pixels (by the above procedure) produces the interpolated result. For pixels near the borders of the scene, correction factors which echo the pattern of the block mean correction factors along the border are dummied in for pixels outside the border as required by the filter.

The along-track or across-track cross-section of this interpolating filter weighting function, and its frequency response are plotted in

Figure 8b.4. The desired characteristics of this filter are that it have a relatively flat frequency response from zero to 1/15 cycles per pixel and have relatively little response above 2/15 cycles per pixel, and that it require no more than 9x9 weighting factors. The filter seems to suit these requirements quite well.

8b.5 COST/BENEFIT ANALYSIS

The block size for calculating the block averages and the dimensions of the moving window filter used to smooth the block mean values were determined from a cost/benefit analysis of the spatially varying XSTAR procedure. In this analysis, cost and performance of the new procedure were compared to the former spatially invaraint (global) XSTAR procedure. Figure 8b.5 indicates typical reductions in the RMS error in matching signal values (pixel by pixel) in consecutive day data which has spatial haze variations. In general we found that the smaller we made the effective size of the moving window smoothing filter, the better we could do. However, the cost of the correction increased with decreasing window size. This cost was strongly related to the block size used for calculating block averages, and varied as shown in Figure 8b.6. We finally chose to use a 5 line by 5 pixel block size and a moving window filter which, in combination with the block size, produced the equivalent of a 15 line by 15 pixel moving window (measured between half amplitude points).

The performance of the spatially varying XSTAR correction is compared to the former global XSTAR correction and to no transformation (but with sun angle correction) for three data sets in Table 8b.6. One of these data sets is segment 1640 which was also presented in Figures 8b.1 and 8b.3. All three scenes had significant spatial haze variations. The lower bound on the RMS error which is achievable by an ideal correction is not known, but could be limited by such factors as scene misregistration, view angle effects, and round-off error following the correction.

Previous test results have indicated that this lower bound may be between 3 and 6 counts, depending on the scene being processed. We believe that the improvement we have achieved is significant and that the spatially varying XSTAR correction should eventually become a part of the standard LACIE procedure.

REFERENCE

8b.1 Lambeck, P.F., "Revised Implementation of the XSTAR Haze Correction Algorithm and Associated Preprocessing Steps for Landsat Data," ERIM Memo Number IS-PFL-1916, November 1, 1977.

TABLE 8B.1

KEY ISSUES ADDRESSED

- STANDARDIZED DATA PRODUCTS
- MORE USEFUL DATA FOR DEVELOPING ANALYSIS TECHNIQUES
- MORE EFFECTIVE AI AIDS
- IMPROVED MACHINE TRAINING AND CLASSIFICATION

ERIM

TABLE 8B.2
APPROACH

- USE LOW-PASS MOVING WINDOW FILTER TO ESTIMATE HAZE DIAGNOSTIC FOR EACH PIXEL
- CALCULATE XSTAR CORRECTION FOR EACH PIXEL
- ADAPT PROCEDURE TO AREAS WITH VARYING PROPORTIONS OF "GOOD" PIXELS
- DEVELOP ECONOMICAL IMPLEMENTATION

ERIM

TABLE 8B.3
PROCEDURE

- SCREEN DATA (USING LESS STRINGENT CLOUD AND DENSE HAZE THRESHOLDS)
- CALCULATE MEAN SIGNAL VALUES FOR 5 LINE BY 5 PIXEL BLOCKS (USING ONLY "GOOD PIXELS)
- CALCULATE SPATIALLY SMOOTHED MEAN VALUES FOR BLOCKS USING A MOVING WINDOW FILTER
- USE SPATIAL INTERPOLATION/EXTRAPOLATION TO ESTIMATE MEAN VALUES FOR BLOCKS WHICH HAVE AN INSUFFICIENT NUMBER OF GOOD PIXELS
- CALCULATE XSTAR CORRECTION APPROPRIATE AT EACH BLOCK CENTER
- INTERPOLATE XSTAR CORRECTION TO APPLY TO EACH PIXEL

TABLE 8B.4

LOW-PASS MOVING WINDOW SPATIAL FILTER

WEIGHTING FACTORS FOR EACH BLOCK

DISPLACEMENT (NUMBER OF BLOCKS)						
	2	.07029	.19484	.26513	.19484	.07029
	1	.19484	.54003	.73487	.54003	.19484
	0	.26513	.73487	1.00000	.73487	.26513
	-1	.19484	.54003	.73487	.54003	.19484
	-2	.07029	.19484	.26513	.19484	.07029
		-2	-1	0	1	2

DISPLACEMENT

(NUMBER OF BLOCKS)

TABLE 8B.5

INTERPOLATING FILTER FOR CALCULATING APPROPRIATE SPATIALLY VARYING XSTAR CORRECTION FOR EACH PIXEL

WEIGHTING FACTORS FOR EACH PIXEL

DISPLACEMENT (IN PIXELS)									
4	.01102	.03761	.06736	.09396	.10497	.09396	.06736	.03761	.01102
3	.03761	.12837	.22992	.32068	.35829	.32063	.22992	.12837	.03761
2	.06736	.22992	.41180	.57435	.64171	.57435	.41180	.22992	.06736
1	.09396	.32068	.57435	.80107	.89503	.80107	.57435	.32068	.09396
0	.10497	.35829	.64171	.89503	1.00000	.89503	.64171	.35829	.10497
-1	.09396	.32068	.57435	.80107	.89503	.80107	.57435	.32063	.09396
-2	.06736	.22992	.41180	.57435	.64171	.57435	.41180	.22992	.06736
-3	.03761	.12837	.22992	.32068	.35829	.32068	.22992	.12837	.03761
-4	.01102	.03761	.06736	.09396	.10497	.09396	.06736	.03761	.01102
	-4	-3	-2	-1	0	1	2	3	4

DISPLACEMENT (IN PIXELS)

TABLE 8B.6

RMS ERROR IN REMOVING DIFFERENCES IN CONSECUTIVE DAY DATA (IN LANDSAT COUNTS)

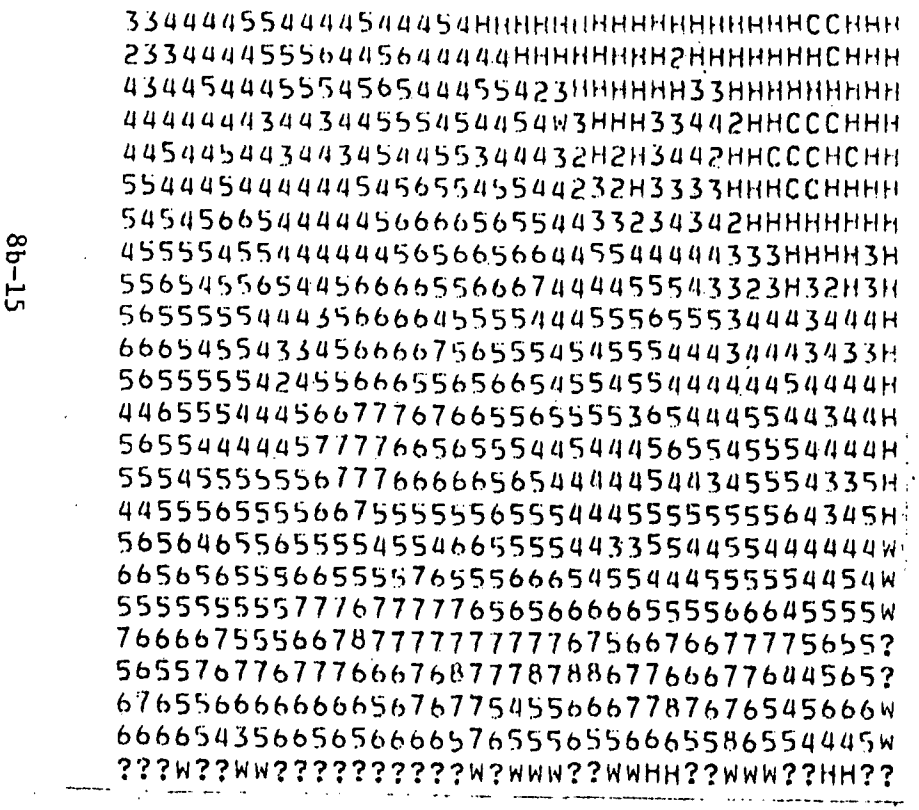
	UNTRANSFORMED	GLOBAL XSTAR		SPATIALLY VARYING XSTAR	
SCENE	RMS ERROR	RMS ERROR	IMPROVEMENT OVER UT	RMS ERROR	IMPROVEMENT OVER UT
SEG# 1619 77175-6	11.3	9.9	12.5%	9.0	20.6%
SEG# 1640 77139-40	13.0	11.3	12.9%	9.8	24.6%
SEG# 1927 77193-4	15.8	11.1	29.7%	9.3	41.5%

FIGURE 8B.1

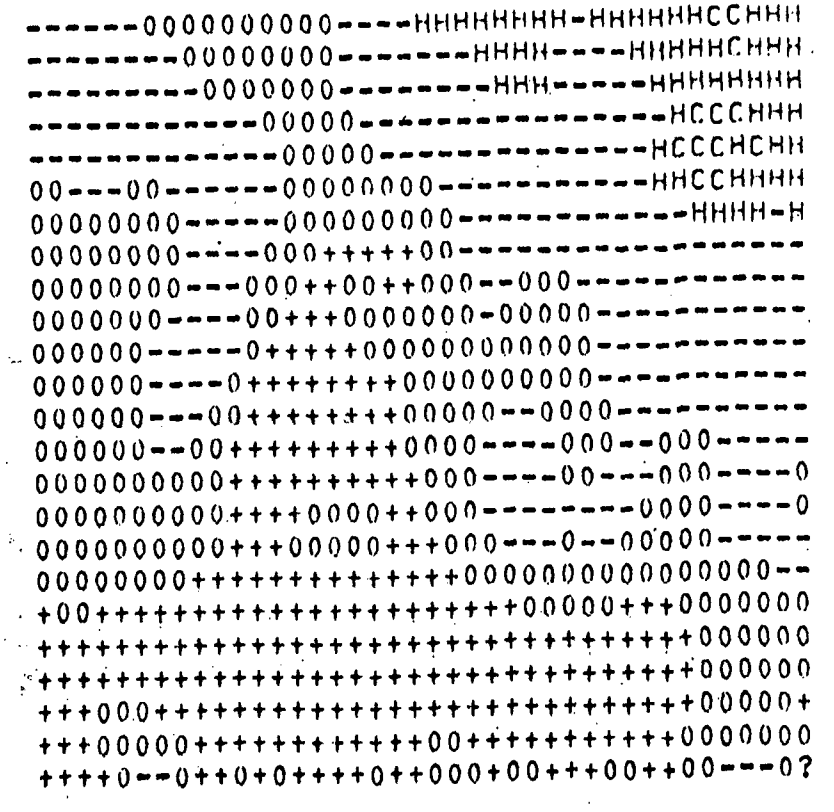
SPATIALLY VARYING XSTAR HAZE CORRECTION

(SEG # 1640 77139)

BLOCK CODE MAP
BLOCK SIZE = 5 LINES X 5 PIXELS
CODE = 5 MEANS YELLOW = -12.9 ± .5



SMOOTHED BLOCK MAP
FILTER SIZE = 3 BLOCKS X 3 BLOCKS
CODE = 0 MEANS YELLOW = -12.9 ± .5



ERIM

FIGURE 8B.2
LOW-PASS MOVING WINDOW SPATIAL FILTER

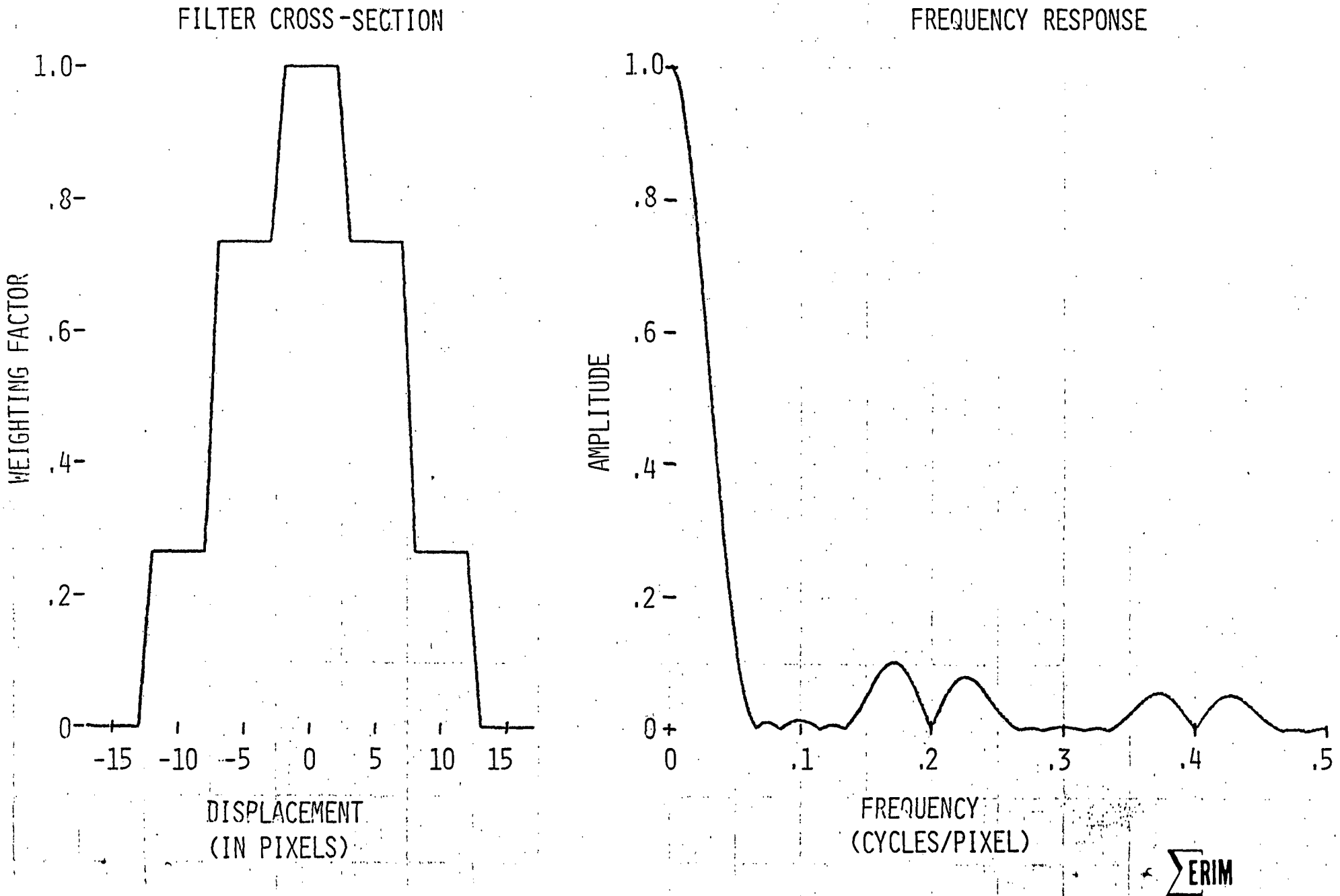
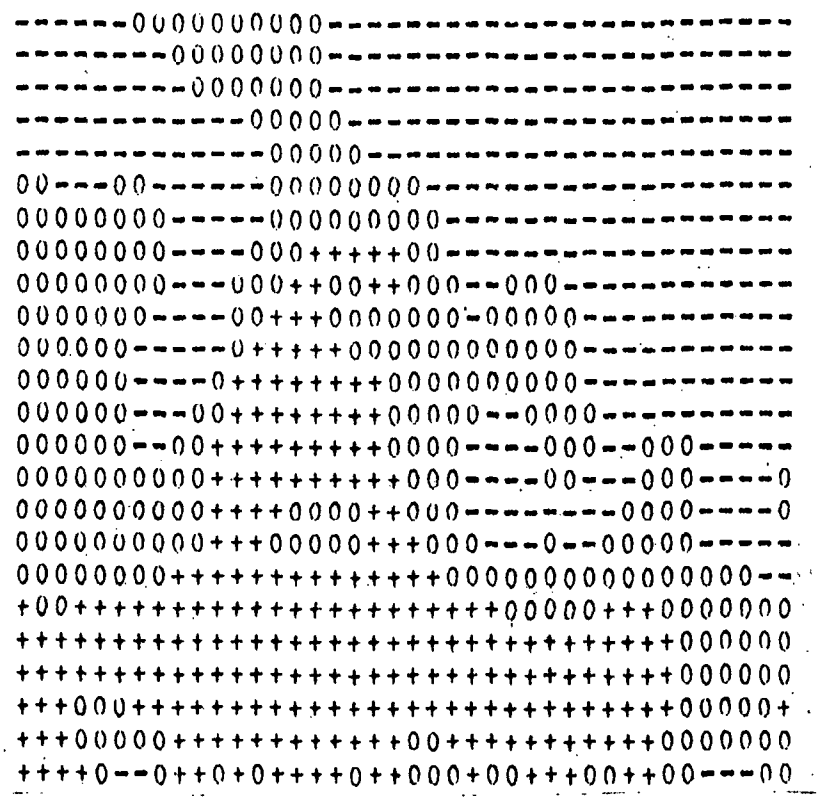
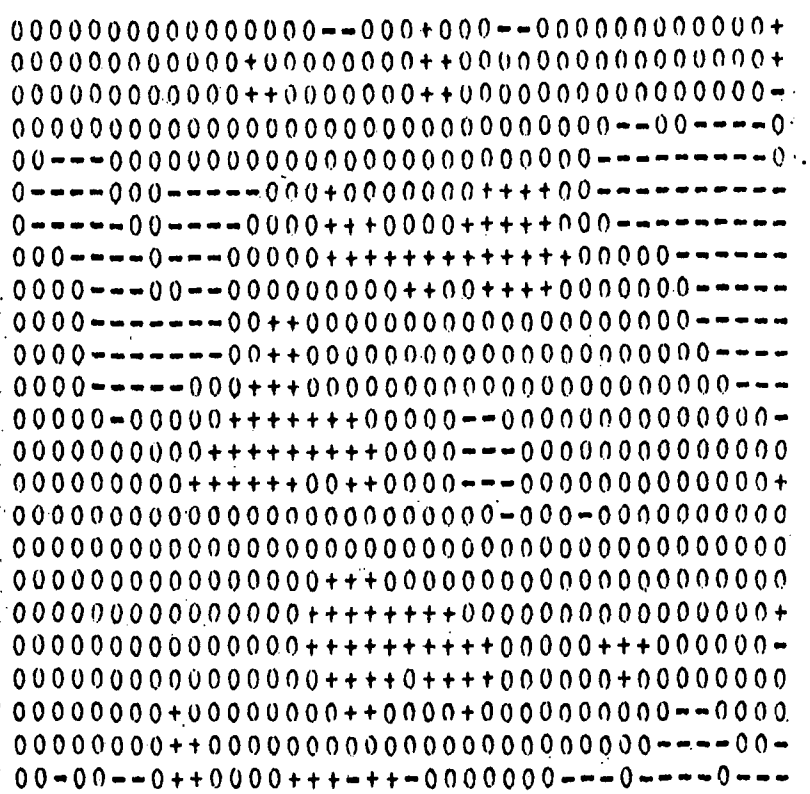


FIGURE 8B.3
SPATIALLY VARYING XSTAR HAZE CORRECTION
(SEG # 1640)
SPATIALLY SMOOTHED AND INTERPOLATED BLOCK CODE MAPS

DAY = 77139
CODE = 0 MEANS YELLOW = -12.9 ± .5



DAY = 77140
CODE = 0 MEANS YELLOW = -10.2 ± .5

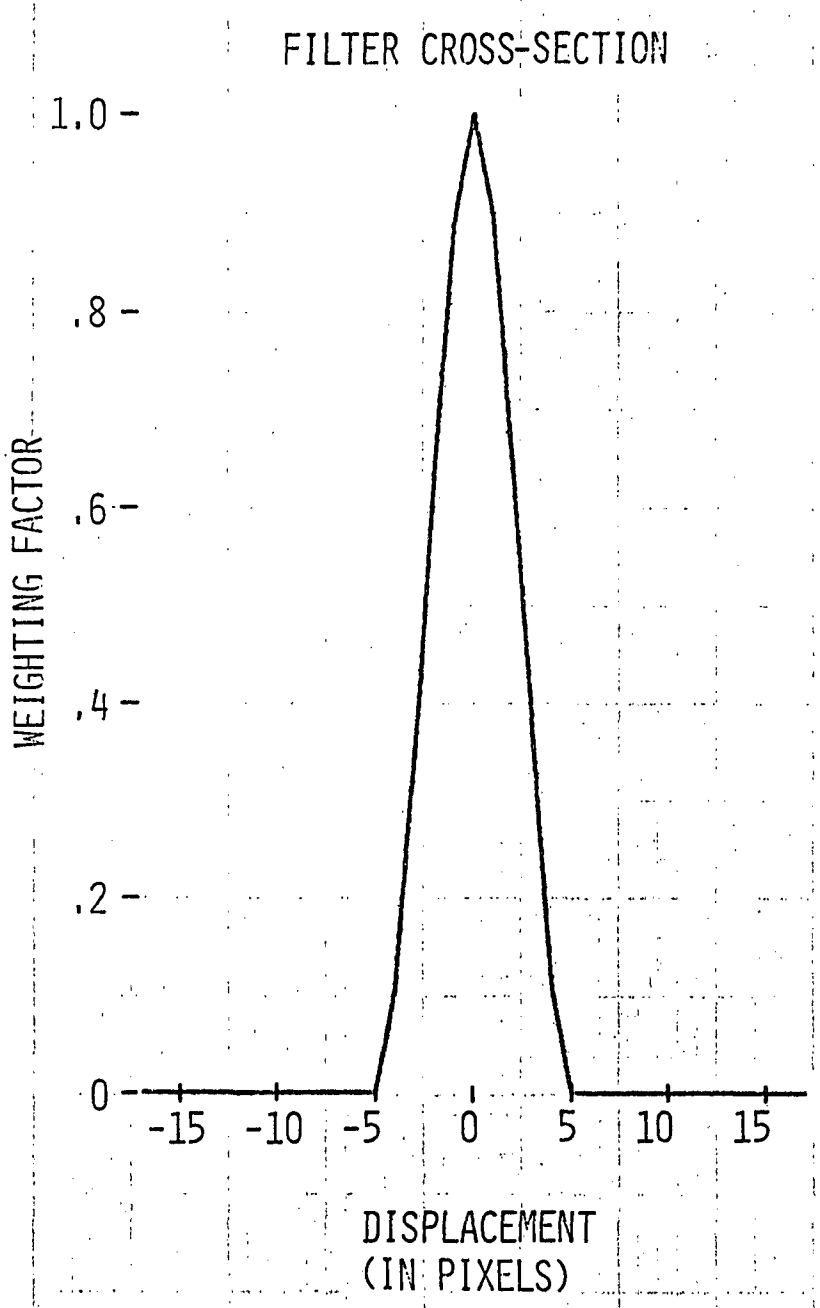


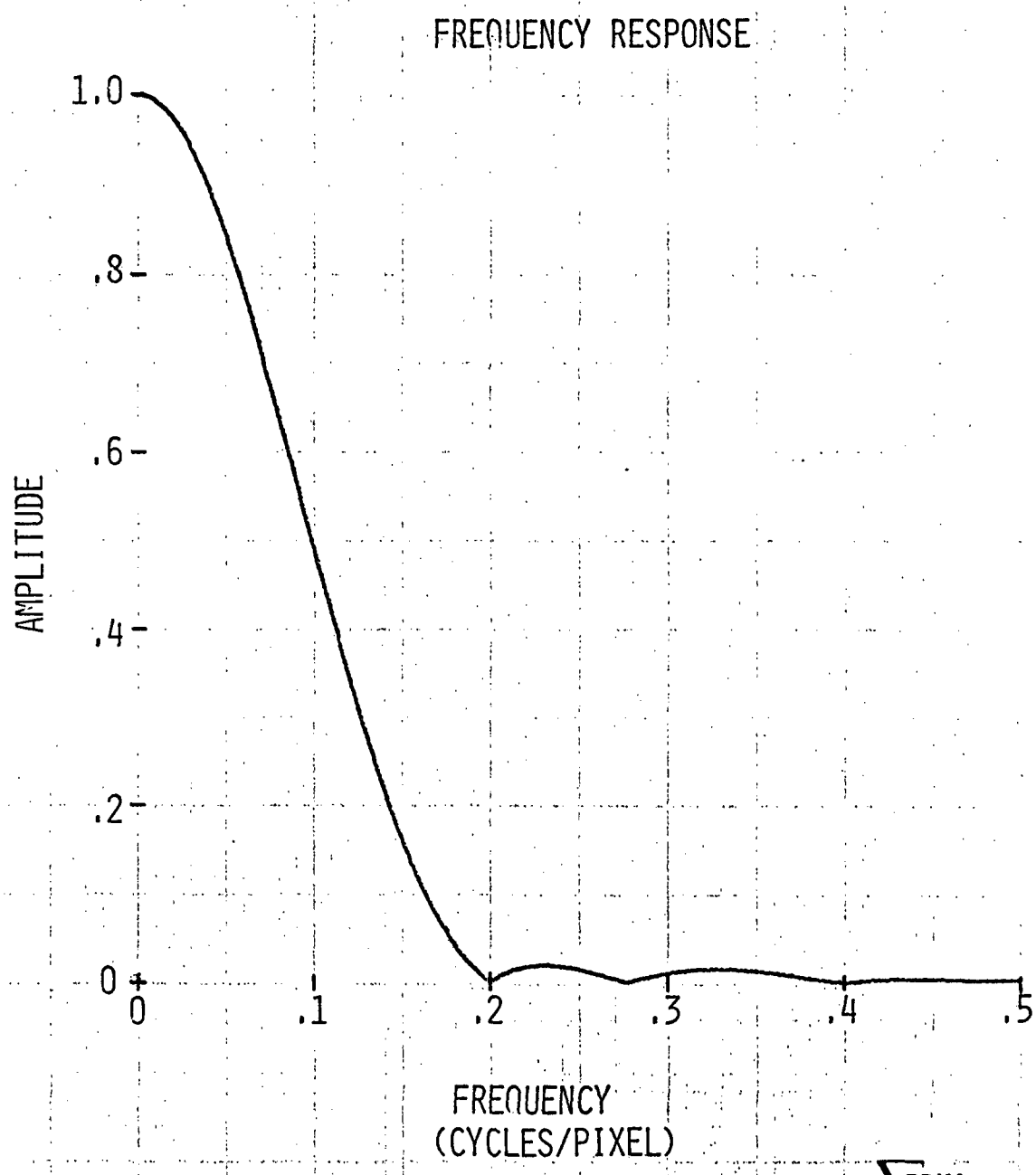
"-" REPRESENTS LESS YELLOW AREA (MORE HAZE)
"+" REPRESENTS MORE YELLOW AREA (LESS HAZE)

ERIM

FIGURE 8B.4

INTERPOLATING FILTER FOR APPLYING XSTAR CORRECTION TO EACH PIXEL





ERIM

FIGURE 8B.5

PERFORMANCE OF SPATIALLY VARYING XSTAR HAZE CORRECTION
ON FOUR CONSECUTIVE DAY DATA SETS WITH VARYING HAZE CONDITIONS

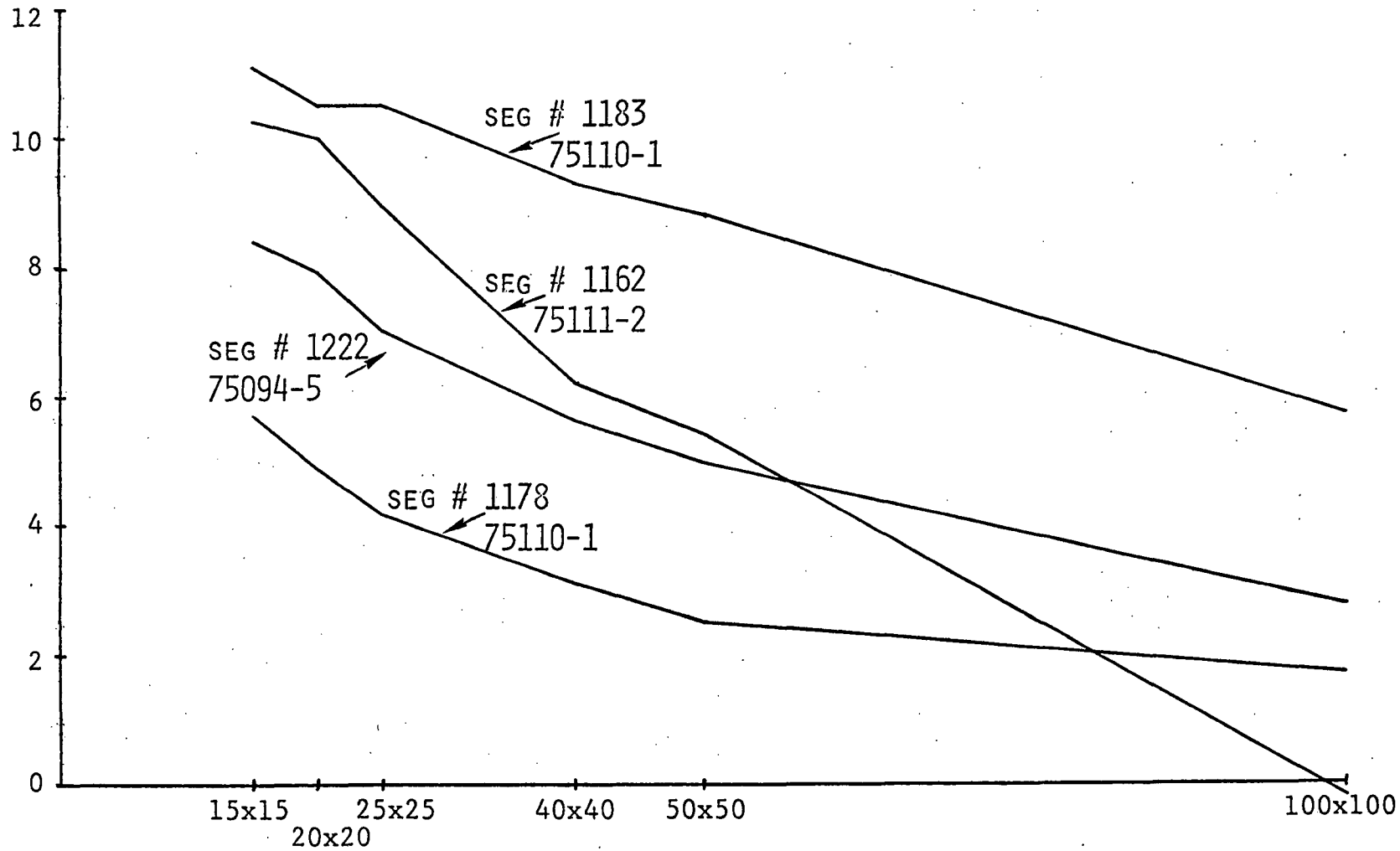
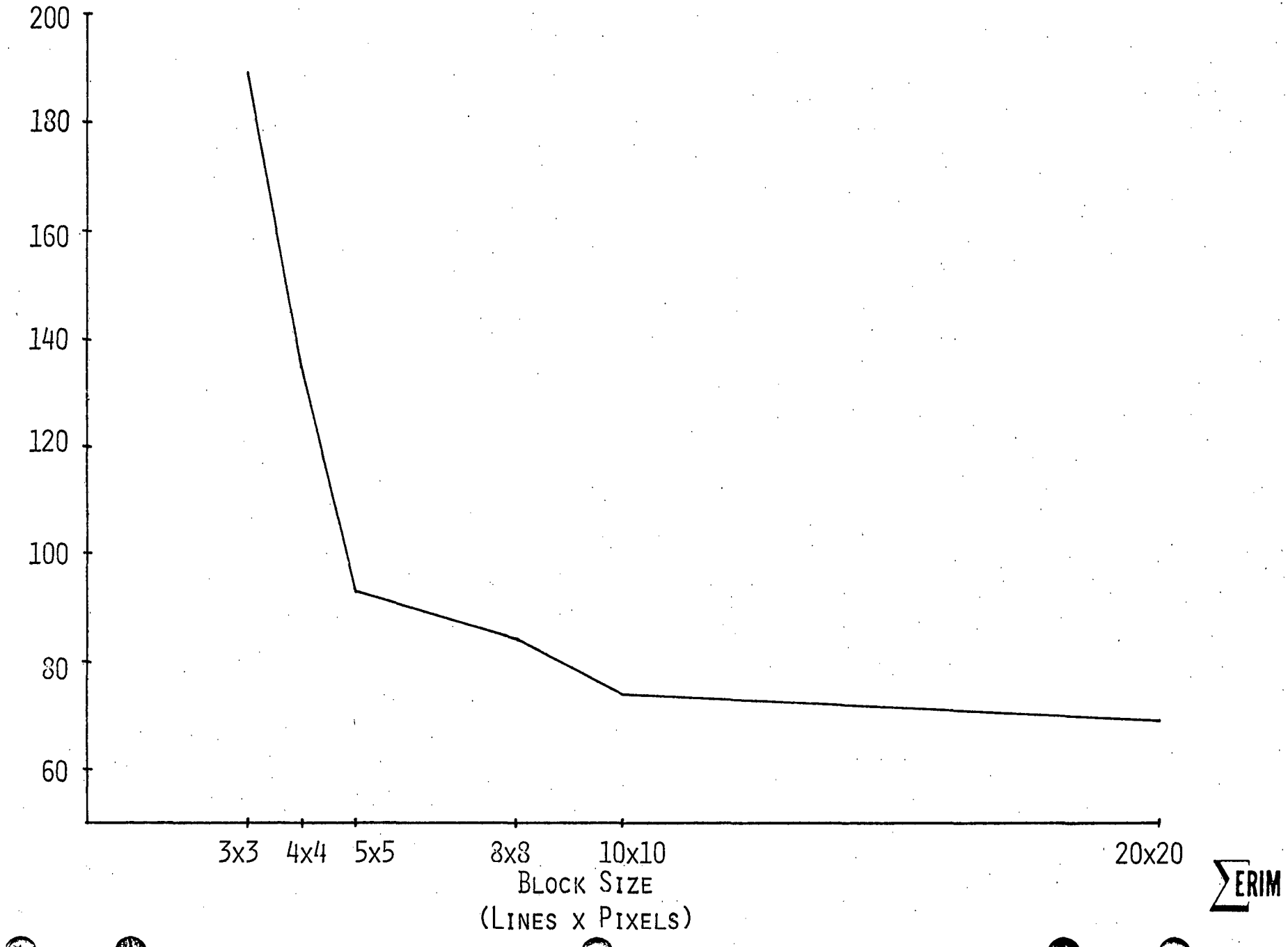


FIGURE 8B.6
COST OF SPATIALLY VARYING XSTAR HAZE CORRECTION
AS FUNCTION OF BLOCK SIZE USED




RECEIVED
DEC 1978