Lockheed Electronics Company, Inc.
A SUBSIDIARY OF LOCKHEED CORPORATION
1830 NASA Road 1, Houston, Texas 77058
Tel. 713-333-5411

JSC-14563

80-10044

NASA CR-
160365

Ref: 642-7186
Contract NAS 9-15800
Job Order 73-705-05

TECHNICAL MEMORANDUM

THE BOUNDARY PIXEL STUDY IN KANSAS AND NORTH DAKOTA

By

D. T. Register and A. L. Oña

(E80-10044) LARGE AREA CROP INVENTORY EXPERIMENT (LACIE). THE BOUNDARY PIXEL STUDY IN KANSAS AND NORTH DAKOTA (Lockheed Electronics Co.) 34 p HC A03/MF A01 CSCL 08B G3/43

N80-16394

Unclas
00044

Approved By: T. C. Minter
T. C. Minter, Supervisor
Techniques Development Section

September 1979

LEC-12826

1. Report No.
JSC-14563

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
The Boundary Pixel Study in Kansas and North Dakota

5. Report Date
September 1979

6. Performing Organization Code

7. Author(s)
D. T. Register and A. L. Oña
Lockheed Electronics Company, Inc.

8. Performing Organization Report No.
LEC-12826

9. Performing Organization Name and Address
Lockheed Electronics Company, Inc.
Systems and Services Division
1830 NASA Road 1
Houston, Texas 77058

10. Work Unit No.

11. Contract or Grant No.
NAS 9-15800

12. Sponsoring Agency Name and Address
National Aeronautics and Space Administration
Lyndon B. Johnson Space Center
Houston, Texas 77058 on
Technical Monitor: J. D. Erickson SF3

13. Type of Report and Period Covered
Technical Memorandum

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Boundary picture elements (pixels) are shown to be an important source of proportion estimation error with Land Satellite data. How boundary pixels are determined is critical to any method of handling the problem. Several methods for dealing with boundary pixels are presented along with the results from a small study of one of these methods. A general framework is outlined for viewing the boundary pixel problem, and an objective standard of comparison for evaluating the different methods is provided. This paper shows an example of how unbiased proportion estimation can be accomplished when the labeling of samples is only performed on pixels interior to fields.

17. Key Words (Suggested by Author(s))
interior pixels, boundary pixel, Procedure 1, label identification from statistical tabulation

18. Distribution Statement

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
29

22. Price*

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

JSC Form 1424 (Rev Nov 75) NASA — JSC

PREFACE

In LACIE Phase III, one of the largest sources of proportion estimation error for a class of interest in Land Satellite (Landsat) data was boundary pixels and the associated method for handling them. These are the pixels partly within the class of interest and partly not because of the coarseness of the resolution of the Landsat sensor. This memorandum describes many of the methods and approaches for dealing with the influence of boundary pixels on proportion estimation. A general framework is presented for viewing the problem, and the results of a small study are presented. The study of the effect of boundary pixels in proportion estimation is very new; almost all the methods presented are untried. They represent the most current thoughts for attacking the problem.

The principal author and project coordinator, D. T. Register, originated the statistical sampling approach that provides the general framework. A. L. Oña conducted most of the analysis of the small study in North Dakota and Kansas. Analysts B. B. Schroder, B. A. Tolbert, C. W. Haynes, and C. L. Dailey assisted in the study by examining thousands of pixels to decide whether they were boundary or not.

CONTENTS

Section | Page

1. INTRODUCTION. 1

2. IMPORTANCE OF BOUNDARY PIXELS 2

3. APPROPRIATE DEFINITIONS FOR BOUNDARY PIXELS 4

4. TECHNIQUES FOR REMOVAL OF ESTIMATION BIAS DUE TO BOUNDARY PIXELS. 6

4.1 CLUSTER-BASED TECHNIQUES 6

4.2 MAXIMUM-LIKELIHOOD-BASED TECHNIQUES. 8

4.3 STATISTICAL SAMPLING TECHNIQUE 9

4.3.1 EXAMPLE OF STRATIFIED AREAL ESTIMATION FORMULA MODIFICATION FOR BOUNDARY PIXELS 9

4.3.2 DESCRIPTION OF A PRELIMINARY STUDY OF THE TECHNIQUE. 10

4.3.3 DESCRIPTION OF THE DATA SET AND RESULTS. 12

4.3.4 PROPOSALS FOR ADDITIONAL ANALYSIS OF THE EXISTING DATA 16

4.3.5 DISCUSSION OF ASSUMPTIONS AMENABLE TO STATISTICAL TESTING. 17

4.4 MULTICATEGORY LABELING 18

5. SUMMARY 19

6. RECOMMENDATIONS 20

7. REFERENCES. 20

Appendix

DATA FROM BOUNDARY PIXEL STUDY ON KANSAS AND NORTH DAKOTA BLIND SITES A-2

TABLES

Table		Page
1	DISTRIBUTION OF BOUNDARY PIXELS AND THEIR CONTRIBUTION TO LABELING ERROR	2
2	REGRESSION COEFFICIENTS WITH THE CORRESPONDING STANDARD DEVIATIONS	16
A-1	BOUNDARY PIXEL STUDY OF KANSAS BLIND SITES	A-3
A-2	PROCEDURE 1 STUDY OF KANSAS BLIND SITES	A-4
A-3	LIST BOUNDARY PIXEL STUDY OF KANSAS BLIND SITES	A-5
A-4	BOUNDARY PIXEL STUDY OF NORTH DAKOTA BLIND SITES.	A-6
A-5	PROCEDURE 1 BOUNDARY PIXEL STUDY OF NORTH DAKOTA.	A-7
A-6	LIST BOUNDARY PIXEL STUDY OF NORTH DAKOTA	A-8

1. INTRODUCTION

A detailed examination of the segment Procedure 1 proportion estimation technique that was used in Phase III of the Large Area Crop Inventory Experiment (LACIE)(ref. 1) has shown that the largest single source for small-grain proportion error is the analyst's misidentification of picture elements (pixels) (ref. 2). Sometimes small grains were thought to be nonsmall grains (omission error), and at other times nonsmall grains were identified as small grains (commission error). The omission error rates were much higher than the commission rates; this resulted in an underestimation bias. The three major causes of analyst labeling errors were found to be (in order of importance) boundary pixels, abnormal signatures, and inadequate acquisitions. Approximately 40 percent of all mislabeled pixels were associated with boundaries. Reference 2 defined a pixel as boundary if it was spatially located so as to be only partly within a small-grain field;[1] that is, if it lay on the perimeter of a small-grain field. The intent of this memorandum is to

1. Report on the past importance of boundary pixels in Phase III and anticipate their importance for the future.
2. Discuss some possible approaches for handling boundary pixels to reduce the bias toward underestimation of the crop of interest.
3. Present the results of some experimentation into one of these approaches.
4. Make recommendations on what considerations should be used in planning a boundary pixel research program.

---

[1]In the case of strip/fallow areas, the pixels were divided into two kinds — those whose strips were wide enough to be discernible in the Landsat imagery and those whose strips were so narrow that no stripping was discernible. If the strips were not discernible, the pixels were not considered as boundary nor in error.

2. IMPORTANCE OF BOUNDARY PIXELS

The singlemost important cause for the mislabeling of pixels in Phase III of LACIE was the boundary pixel. Table 1 presents the results of a detailed dot-by-dot examination of the labeling of type 2 dots. All the results are for Phase III of LACIE. The scope of the study included all the intensive test sites (ITSs) and a sample of blind sites drawn from each of five selected states. The results show that 66 percent more boundary pixels were labeled nonsmall grains than were labeled small grains. The percentages of error shown in table 1 do not average to the 40 percent mentioned in section 1. The percentages in the table refer to errors for which no other cause than boundary could be given. The 40 percent included all errors associated with boundaries. There were in some cases multiple causes.

TABLE 1.— DISTRIBUTION OF BOUNDARY PIXELS AND THEIR CONTRIBUTION TO LABELING ERROR

Data set	No. of segments	No. of type 2 boundary dots	No. boundary dots analysts labeled small grains	No. boundary dots analysts labeled nonsmall grains	Percentage of all pixels that are boundary	Percentage of omission errors due to boundaries	Percentage of commission errors due to boundaries
Winter ITSs	13	112	39	73	12.4	21.4	8.6
Spring ITSs	7	65	30	35	12.4	12.9	8.8
Minnesota	6	56	20	36	16.0	40.6	44.6
Montana*	10	55	28	37	6.9	21.1	29.4
North Dakota*	18	120	43	77	11.7	28.9	22.1
Oklahoma	12	80	27	53	10.3	22.1	14.3
Colorado*	6	29	11	18	7.3	25.0	100.0

*These segments contained some strip/fallow fields which were so narrow that they could not be discerned in the imagery. Pixels in these fields were not considered as boundary or in error.

The effect of boundary pixels can be expected to grow in importance in the Agricultural and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) Project as interest is directed to inventories of multiple crops in areas with smaller fields. In particular, for corn and soybeans, there will be three kinds of boundaries: corn/other, soybeans/other, and corn/soybeans, whereas in the past, only small grain/other boundaries were considered. In addition, the sizes of the fields in the corn and soybean areas are generally smaller than those for wheat in the U.S. Great Plains. The following table illustrates how the number of boundary pixels could increase as more crops are estimated in areas with smaller fields. The table lists five sites within the corn and soybean region; for each segment, the percentage of interior pixels is given.[2] For this table to demonstrate the effect of more crops, interior pixels are those which are wholly within the same category on each acquisition. Each crop was considered a separate category.

Segment	State	Percentage of pixels interior to fields
146	Kentucky	45.0
812	Missouri	46.4
824	Illinois	44.0
883	Iowa	42.1
886	Idaho	52.2

[2]These five segments were selected because of availability of data and are only for illustration. They may not be completely representative of the corn/soybean area.

3. APPROPRIATE DEFINITIONS FOR BOUNDARY PIXELS

Commonly, the term boundary pixel refers to a pixel which lies on the spatial demarcation between two categories of interest on a single Landsat acquisition. Though this definition is appropriate for visualization, it is not a working definition. A working definition must specify the method for deciding whether or not a pixel is a boundary. This point is important because the designation of pixels as boundary can and will vary considerably depending upon the method utilized. Among the possible methods which can be chosen are clustering algorithms, classification algorithms, analysts, and aircraft photographs with ground truth. Because the aircraft photographs are not routinely available for use in estimation of crop segment proportions, they will not be further considered in this memorandum as a method of determining which pixels of a scene are boundary.

There are basically two issues in regard to which pixels are boundary. The first issue can be summarized by the question: How much of a boundary does a pixel need to be before it is labeled "boundary"? The second, maybe not so obvious, issue is the following: In which Landsat acquisition is the pixel a boundary? Because of the misregistration between Landsat passes over a segment, a given pixel can migrate back and forth between and on the boundary of two fields. Thus, an adequate answer to whether or not a pixel is boundary requires specification of a particular acquisition. Sometimes, a set of acquisitions is specified. In the Label Identification from Statistical Tabulation (LIST) approach (ref. 3), a set of four acquisitions is specified. If a pixel migrates back and forth or is on the boundary of two fields as determined by an analyst in these acquisitions, then the pixel is designated a boundary. In Procedure 1, a similar approach is taken: The analyst is asked to label for type 1 dots[3] only those pixels which remain within the same field over the four (or fewer) acquisitions chosen for processing. The implication for type 1 dots is that all pixels not labeled are boundary.

---

[3] In procedure 1, the type 1 dots are first used for starting a clustering algorithm and then for labeling the resulting clusters by the nearest-neighbor rule. With type 2 dots, all are identified and used for stratified areal estimation.

Two proposals have been made for use by analysts in this determination. The first proposal is that a pixel should be called a boundary pixel if an analyst has difficulty ascertaining to which of the two contiguous categories on which the pixel lies he should assign the pixel. The second proposal is that a pixel should be called boundary if the analyst can determine that any part of the pixel is on the spatial boundary of two categories. The second proposal is more liberal than the first and will result in a larger number of boundary pixels. One must keep in mind here, however, that the importance of the number of boundary pixels is secondary to that of consistency in determination.

Several analysts were asked which of the two proposed working definitions would provide the most consistent responses for denoting pixels of a segment as boundary. The overwhelming opinion was that the liberal definition would provide more consistent responses.

The liberal definition is the one adopted for use in the study described in section 4.3.2. As for machine algorithms, such as clustering and classification, they would objectively and consistently apply the rules for determining boundaries that they are programmed to follow. Other procedures may require their own particular definitions, such as the LIST approach. In the LIST approach, whether or not a pixel is boundary is determined by an analyst's response to a given set of questions. In summary, the liberal definition appears to be an appropriate definition for an analyst to use. However, special procedures or algorithms will require their own definitions. All of these can be expected to give differing results. In addition, the designation boundary for a pixel is only applicable to a particular acquisition or set of acquisitions.

4. TECHNIQUES FOR REMOVING THE ESTIMATION BIAS DUE TO BOUNDARY PIXELS

Four basic types of techniques[4] have been suggested for handling boundary pixels: cluster-based technique, maximum-likelihood-based technique, statistical sampling technique, and multicategory labeling.

4.1 CLUSTER-BASED TECHNIQUES

The cluster-based techniques diverge into two separate approaches. The distinctive difference between the two is the means of determining which pixels are boundary. The means may be either automatic (e.g., allowing a clustering algorithm to decide which pixels are boundary) or manual (e.g., in the case of an analyst). The first approach to be discussed is an unsupervised computer processing algorithm to detect boundary and interior pixels. The interior pixels are classified into unlabeled categories. Then, individual boundary pixels and their adjacent pixels are automatically examined to determine the categories in which they overlap. The boundary pixels are then modeled as a percentage of each adjacent category by using the mean spectral values of those categories. The analyst labels samples of the interior pixels in order to label the categories. Thus, boundary pixels would be automatically determined and allocated to the adjacent categories on the basis of the spectral data, and the analyst would be relieved of having to allocate boundaries to categories. The possible misgivings about this approach are as follows:

1. It has not been shown that an automatic boundary detection algorithm could locate boundaries with sufficient accuracy. In particular, how many actual boundaries will be overlooked, how many interior pixels will be mistakenly labeled boundary, and what detrimental effects on proportion estimation will these have?

[4]The first three techniques are suggested in the minutes of the Procedure 1 Review Conference held at Texas A&M University on July 13-14, 1978. The minutes are contained in a letter to R. P. Heydorn from T. C. Minter (dated January 18, 1979). Some of the alternatives are from a private communication with R. K. Lennington resulting from his attendance at the Conference on Maximum Likelihood Clustering held at Texas A&M University on February 8-9, 1979.

2. A basic premise of this approach is that an unsupervised clustering algorithm can partition a Landsat scene into clusters, each of high purity (same category). This is still to be demonstrated. Additionally, there is no theoretical reason to believe that omission errors would balance commission errors so that unbiased proportion estimates would result.
3. A boundary pixel may be boundary on more than two categories. Properly formulated, though, the modeling could easily be extended to three or four categories.

In the second approach, whenever a boundary pixel is encountered, the analyst would consult an unsupervised multitemporal cluster map for determining which field to assign the boundary pixel. The analyst would then label the boundary pixel according to the category of the field. This would provide an arbitrary, objective means for the allocation of boundary pixels to categories. The underlying premises are that the clusters from the chosen clustering algorithm would correspond to categories and that boundary pixels would be assigned to the most likely category by the algorithm as it assigns pixels to categories on the basis of spectral distance (some algorithms may also incorporate spatial distance).

It is envisioned that this second approach would be suitable for segment inventory procedures which use unsupervised clustering followed by either stratified areal estimation or cluster labeling with analyst-labeled samples. The procedural flow would be (1) unsupervised clustering of the image data; (2) aided by a cluster map, analyst labeling of the machine-selected pixels; and (3) another processing, such as stratified areal estimation, to provide the proportion estimates.

Possible misgivings are (1) the lack of a theoretical reason to believe that the clusters will allocate the pixels in an unbiased manner, (2) the possible dependence of the allocation on field size, (3) the proportions of each crop in the scene, (4) field shape, (5) the particular clustering algorithm used, (6) the spectral separability of the categories of interest, and (7) the extent to which boundary pixels form their own cluster. It is not clear, though, how important these considerations are. This technique has not been investigated. It could prove to be an easy, practical technique for allocating boundary pixels to categories.

4.2 MAXIMUM-LIKELIHOOD-BASED TECHNIQUES

With these techniques, an analyst would be asked to label only interior pixels. These pixels would be utilized as training samples for a maximum likelihood classification of all the pixels within the scene. Thus, the boundary pixels would have been allocated to the most likely category on the basis of probabilities computed from the labeled pixels. This technique has the same advantage as that of clustering; it offers an arbitrary, objective means for the allocation of boundary pixels to categories.

From the standpoint of an efficient use of labeled training samples, this technique would be best suited to proportion estimation of the categories of interest by the processing machine. The procedural flow would be for the analyst to (1) receive imagery and spectral aids, (2) label interior pixels sampled by some prescribed scheme, and (3) submit them for classification which would provide the estimate. Three variations on how the estimate could be produced are as follows: One variation would be to count the number of pixels classified into each category of interest. A second variation would be to compute the probability for each category of interest for each pixel and then average the probabilities over the scene. This would not provide a classification map but may provide more accurate estimation. A third variation would be the use of a boundary detection algorithm to separate the interior pixels from the boundary ones. Then, pixel counting would be used on the interior pixels and probability averaging on the boundaries. As can be easily seen, the number of possibilities here are limited only by imagination.

The possible misgivings are (1) a lack of a theoretical reason for believing that the pixels would be allocated in an unbiased manner, (2) the allocation's possible dependence on the proportions of each crop in the scene, (3) the dependence on the particular classification algorithm used, and (4) the spectral separability of the categories of interest.

4.3 STATISTICAL SAMPLING TECHNIQUE

With this technique, the analyst would no longer be required to place a single category label on a boundary pixel. Instead, he would provide two labels — one for each of the categories straddled by the boundary pixel. Each category as labeled by the analyst would then contain both interior pixels and boundary pixels. This method of labeling is best suited for use with segment inventories obtained by stratified areal estimation. The usual formulas have to be modified slightly to accommodate the boundary labels. The derivation for the two-category case follows.

4.3.1 EXAMPLE OF STRATIFIED AREAL ESTIMATION FORMULA MODIFICATION FOR BOUNDARY PIXELS

The usual formula for the proportion estimate from stratified areal estimation for two categories is the following:

$$\hat{P}_W = \frac{N_1}{\text{Base}} \frac{n_{1W}}{n_1} + \frac{N_2}{\text{Base}} \frac{n_{2W}}{n_2}$$

where

$\hat{P}_W$ = The estimated proportion for category W, one of the two categories.

N_i = The number of pixels in the segment classified into class i, where i = 1, 2.

Base = The number of pixels within the entire segment less any unidentifiable areas such as clouds.

n_i = The number of labeled samples classified into the *ith* class, where i = 1, 2.

n_{iW} = The number of samples labeled W and classified in class i, where i = 1, 2.

The required modification is based on the observation that the n_i pixels in the *ith* class actually consist of some interior pixels n_{ip} and some boundary pixels n_{iB}. Thus, $n_i = n_{iP} + n_{iB}$ for i = 1, 2. The interior pixels n_{iP} are further divided into those that are labeled W(n_{iPW}) and those labeled ϕ (the second category). These interior pixels n_{iPW} are wholly within the W category. However, the boundary pixels for the *ith* class are considered only partly within the W category; that is, a fraction β_W of the W category. Thus, in the formula, n_{iW} is replaced by $n_{iPW} + \beta_i n_{iBW}$. Thus, the modified formula becomes:

$$\hat{P}_W = \frac{N_1}{\text{Base}} \; \frac{(n_{1PW} + \beta_1 n_{1BW})}{n_1} + \frac{N_2}{\text{Base}} \; \frac{(n_{2PW} + \beta_2 n_{2BW})}{n_2}$$

In this form, the β_i is the average percentage of the boundary pixels classified into the *ith* class that should be allocated to the W category.

4.3.2 DESCRIPTION OF A PRELIMINARY STUDY OF THE TECHNIQUE

Among the inputs needed to apply this modified formula are the classification percentages, the number of labeled test samples classified into each of the two categories, and the labels for the test samples as W, ϕ, or B. All of these are readily available after the processing of a segment. In addition, the two β's are required. The β's cannot be estimated for each segment on a real-time basis, but they can be estimated in advance using blind site data. In addition, it is possible to adjust the β's for use on a per-segment basis by making the β's a function of segment characteristics. Two considerations for adjusting the β's for each segment are to (1) select variables related to the β's and (2) quantify the variables for nonblind sites. The β's are thought to be functions of field size, shape, and separability for the category of interest. Thus, they may vary from region to region. They should, however, be fairly stable from year to year in the same area. This allows the use of the blind sites to estimate the β's for a given area. Though it was not attempted in this study, the ratio of the number of boundary pixels to the number of pixels labeled is a potential variable that is correlated well with the β's. This variable is readily available after segment dot labeling. The procedure for determining the β's for a particular area is simply to use ground-truth labels in lieu of analyst-labeled pixels and substitute the ground-truth percentage for the proportion estimate $\hat{P}_W$.

A machine processing (classification) is performed, and the results are input to the formula. This provides an equation in two unknowns β_1 and β_2 for each blind site processed. A constrained regression analysis is performed to provide unbiased estimates for β_1 and β_2. One assumption in estimating β_1 and β_2 in this manner is that they are constant for all the blind sites in the regression analysis.

A small study of this technique was conducted. Section 4.3.3 describes the data set and the results. The data are presented in the appendix. Basically, 26 blind sites in North Dakota and 19 in Kansas were used to estimate β's separately for each state. Because different definitions of boundary pixels will result in different β's, two different definitions were considered, one suitable for Procedure 1 and the other suitable for LIST.

The results indicated that the variances of the proportion estimates of small grains in the boundary pixels (β's) are unduly large. In part, this is due to the very difficult, if not impossible, task of determining the "true" proportion of small grains within a set of 209 dots. The difficulty is that even with the use of ground truth, there is and must be some subjectiveness in determining a label for some of the dots. In fact, there really is no correct label for the dots which are partially small grains and no way to measure directly the percentage of small grains in these partially small-grain pixels. Thus, the true proportion cannot be directly measured without error. Consequently, the small-grain proportion for the entire segment was used. Due to sampling variance, the small-grain ground-truth percentages for the entire segment were sometimes physically irreconcilable with the ground-truth percentages for the 209 grid dots. Specifically, either the small-grain percentage given by only the interior dots from the 209 was more than the segment proportion, or the small-grain percentage given by the number of interior dots plus all the boundary dots (treated as 100 percent small grains) in the 209 was less than the segment proportion. Further analysis to obtain more reliable estimates has not yet been undertaken.

4.3.3 DESCRIPTION OF THE DATA SET AND RESULTS

A group of analysts was asked to interpret the 209 pixel-grid intersections for 45 Phase III blind sites, 26 in Kansas and 19 in North Dakota, to identify interior and boundary pixels. The following labels were used:

Label	Definition
A	Anomalous pixel*
D	Nonagricultural area
P	Interior pixel
R	Misregistered pixel
X	Clouds and cloud shadows
1	Boundary (small grain and nonsmall grain)
2	Boundary (nonsmall grain and nonsmall grain)

*Pixel is not representative of most of the other pixels within the field; e.g., a mud puddle in a wheat field.

The following blind sites in North Dakota were used in the study:

1602 (Mountrail)	1648 (Bowman)
1604 (Renville)	1652 (Stark)
1606 (Ward)	1661 (McIntosh)
1616 (Cavalier)	1663 (Richland)
1619 (Grand Forks)	1899 (Walsh)
1622 (Ramsey)	1902 (McKenzie)
1625 (Dunn)	1903 (Mercer)
1635 (Sheridan)	1913 (Hettinger)
1637 (Stutsman)	1927 (Sargent)
1640 (Barnes)	

The following blind sites in Kansas were used:

1032 (Wichita)	1293 (Meade)
1033 (Clark)	1295 (Osborne)
1153 (Jewell)	1297 (Dickinson)
1155 (Phillips)	1885 (Rice)
1158 (Washington)	1340 (Sumner)
1166 (Lyon)	1343 (Riley)
1170 (Harper)	1346 (Geary)
1175 (Sedgwick)	1851 (Graham)
1180 (Cherokee)	1853 (Ness)
1183 (Labette)	1859 (Hamilton)
1279 (Cheyene)	1861 (Kearny)
1285 (Logan)	1864 (Stanton)
1290 (Ford)	1881 (Ellsworth)

Since individual pixels can change with multiple acquisitions from one category to another from acquisition to acquisition because of misregistration in the North Dakota blind sites, a reference date was specified; and the analysts used three selected additional acquisitions, which included ripe and harvested dates when possible, to label the 209 dots. However, in the Kansas blind sites, a reference date with the most distinct field boundaries was selected in addition to the three other acquisitions used for labeling.

The corresponding 209-dot ground-truth labels were used to identify the interior small-grain dots from the nonsmall-grain dots. The classification map from Procedure 1 was used to identify the category in which the dots were placed. The following quantities were determined: total small-grain dots classified as small grains, total small-grain dots classified as nonsmall grains, number of dots classified as small grains, number of dots classified as nonsmall grains, total small-grain/nonsmall-grain boundary dots classified as small grains, and total small-grain/nonsmall-grain boundary dots classified as nonsmall grains.

To estimate the effect of boundary pixels on the small-grain proportion (i.e., to determine the fraction of boundary pixels that should be considered small grains), the parameters β_1 and β_2 were computed from the following equation.

$$\hat{P}_{GT} = \frac{N_1}{\text{Base}}\left[\frac{n_{SG1}}{n_1} + \beta_1 \frac{n_{B1}}{n_1}\right] + \frac{N_2}{\text{Base}}\left[\frac{n_{SG2}}{n_2} + \beta_2 \frac{n_{B2}}{n_2}\right]$$

where

$\hat{P}_{GT}$ = Ground-truth small-grain proportion.

N_1 = All pixels classified as small grains.

N_2 = All pixels classified as nonsmall grains.

Base = The number of pixels within the entire segment less any unidentifiable areas such as clouds.

n_{SG1} = Total small-grain dots classified as small grains.

n_{SG2} = Total small-grain dots classified as nonsmall grains.

n_1 = Number of dots classified as small grains.

n_2 = Number of dots classified as nonsmall grains.

n_{B1} = Total small-grain/nonsmall-grain dots classified as small grains.

n_{B2} = Total small-grain/nonsmall-grain dots classified as nonsmall grains.

β_1 = Fraction of the boundary pixels which are small grains, given that the boundary pixels are classified as small grains.

β_2 = Fraction of the boundary pixels which are small grains, given that the boundary pixels are classified as nonsmall grains.

To solve for the parameters β_1 and β_2, a linear regression on X_1 and X_2 on Y was fitted through the origin with the following model:

$$Y = \beta_1 X_1 + \beta_2 X_2$$

where

$$Y = P_{GT} - \left(\frac{N_1}{\text{Base}}\right)\left(\frac{n_{SG1}}{n_1}\right) - \left(\frac{N_2}{\text{Base}}\right)\left(\frac{n_{SG2}}{n_2}\right)$$

$$X_1 = \left(\frac{N_1}{\text{Base}}\right)\left(\frac{n_{B1}}{n_1}\right)$$

and

$$X_2 = \left(\frac{N_2}{\text{Base}}\right)\left(\frac{n_{B2}}{n_2}\right)$$

Two sets of data were analyzed: Procedure 1 data in which A, D, P, and R labels were considered interior pixels and the 1 and 2 labels were considered boundary pixels; and LIST data, in which D, P, and 2 labels were considered interior pixels and the A, R, and 1 labels were considered boundary pixels. Pixels labeled X were excluded from the analyses.

Table 2 presents the regression coefficients along with the corresponding standard deviations separately for each labeling method and state. Results indicated that the fraction of small grains in the boundary pixels that are classified as small grains is 72 percent for Procedure 1 and 49 percent for LIST. For Procedure 1 in North Dakota, 72 percent of the boundary pixels classified as small grains was small grains, and 19 percent of the boundary pixels classified as small grains was small grains. However, in Kansas, 72 percent of the boundary pixels classified as small grains was small grains, and 3 percent of the boundary pixels classified as nonsmall grains was small grains. Extremely large standard deviations of the regression coefficients reflect the variability in the 209-dot sampling and the variability due to the number of segments within the state. For example, segment 1663 in North Dakota (Procedure 1 data) had a ground-truth wheat percentage of 51.84. There were 97 interior small-grain dots and five boundaries. The random estimate for wheat from the interior pixels was (97/209) × 100 = 46.41 percent, and the random estimate for wheat from interior and boundary pixels was (102/209) × 100 = 48.80 percent. The ground-truth percentage (51.84 percent) is greater than the percentage of wheat from the random estimate (48.80 percent) with the interior and the boundary pixels, indicating that the boundary pixels would need to be considered as greater than 100 percent small grains. This variability is due to random sampling.

TABLE 2.— REGRESSION COEFFICIENTS WITH THE CORRESPONDING STANDARD DEVIATIONS

State	β_1	Standard deviation of β_1	β_2	Standard deviation of β_2	R^2 coefficient of determination
Procedure 1					
North Dakota	0.72656	0.54149	0.19814	0.34859	0.44986
Kansas	0.72404	0.99781	0.03050	0.30608	0.09065
LIST					
North Dakota	0.49496	0.19963	0.43250	0.16310	0.78673
Kansas	0.58680	0.49548	0.54625	0.47139	0.35887

4.3.4 PROPOSALS FOR ADDITIONAL ANALYSIS OF THE EXISTING DATA

One possible way of reducing the variance of the estimated β's would be to incorporate prior knowledge into the estimation process. In this case, when the ground-truth proportion P_{GT} is lower than the random sample estimate for pure small grains $\hat{P}_L$, then $\hat{P}_L$ could be used in place of P_{GT} for estimation of the β's. Likewise, when P_{GT} is greater than the random sample estimate for pure small grains plus all boundary small-grain $\hat{P}_U$, the $\hat{P}_U$ should be used in place of P_{GT} in the regression. The resulting regression on the truncated proportions will perhaps bias the estimates of the β's. However, they will have smaller variances and will dampen the effect of sampling anomalies.

Another interesting analysis that could be performed is to use the Procedure 1 type 2 dots to obtain a simple random sample proportion for small grains with the analyst labels, which include designations of which pixels are boundary to perform the regression estimation for the β's. These β's would estimate the analyst's opinion of the amount of small grains in the boundary pixels for each category of machine classification. Small-grain proportion estimates could then be calculated using the estimated β's and would perhaps demonstrate a smaller variance on the proportion estimate than the Procedure 1 estimate although there should be no improvement in the bias.

Both of these would provide proportion estimates with equivalent bias and perhaps lower variances than the Procedure 1 estimates. Superior approaches to estimation of the β's exist, but they entail the collection of new data from additional follow-on studies. Recommendations are presented in section 5.

4.3.5 DISCUSSION OF ASSUMPTIONS AMENABLE TO STATISTICAL TESTING

The statistical sampling technique for handling boundary pixels represents an entire class of techniques corresponding to different assumptions about β_1 and β_2 for two categories. The following table lists five assumptions on the β's and explains how they correspond to methods sometimes proposed for handling boundaries.

Assumption	Correspondence
$\beta_1 = \beta_2 = 1/2$	Each boundary pixel is considered as containing 50 percent of the category of interest regardless of its classification.
$\beta_1 = \beta_2 = \beta$, where β is a constant	Each boundary pixel is considered as containing a fraction β of the category of interest regardless of its classification.
$\beta_1 \neq \beta_2$, where β_1 and β_2 are constants	The boundary pixels classified into the *ith* class are considered as containing a fraction β_i for the category of interest. The fraction will be different depending on its classification but will otherwise remain constant.

Assumption	Correspondence
$\beta_i = \dfrac{n_{iPW}}{n_{iPW} + n_{iP\phi}}$	The fraction β_i of the category of interest in the boundary pixels which are classified into the *ith* category is considered to be the same as the ratio of interior pixels for the category of interest to all interior pixels classified into the *ith* class.
β_i is subjectively determined and varies from segment to segment and from analysis to analysis	This is the present method; by this method, an analyst is forced to make a determination (label) on each boundary pixel.

One of the biggest advantages of the statistical sampling technique is that assumptions like these concerning the β's can be tested.

As was discussed earlier, the definition of boundary pixels is critical to any method that attempts to handle them. Thus, assumptions which may be suitable for some definitions may not be suitable for others. A main point to be made, though, is that the statistical sampling method offers a standard against which to measure all other methods for handling boundary pixels. No matter what definition is chosen for boundary pixels by a candidate method, the statistical sampling method can use that definition and provide a minimum variance unbiased estimate of the average fraction of a boundary pixel that contains the category of interest (even conditioned on the classification, if one is given). Hence, a best estimate for β_i is available and can be used for hypothesis testing to facilitate the comparison of methods for handling boundary pixels.

4.4 MULTICATEGORY LABELING

As was mentioned earlier in section 2, the analyst tends to be conservative in labeling a category of interest, especially for the boundary pixels. In the past, analysts were instructed to identify small-grain pixels and label any other category as nonsmall grains. The intent of this approach was to label the small grains accurately. This produced high accuracies for those pixels identified as small grains. Whenever a pixel was labeled small grains in Phase III, the probabilities of being correct were 91.1 percent (ref. 1). The analyst did not label all the small grains as small grains. In fact, in Phase III, the analyst was able to identify only 78.6 percent of the small grains correctly. One of the major sources of this small-grain omission error was the preponderance of boundary pixels, which are of course partly small grains being identified as nonsmall grains.

Thus, if the reason for conservative labeling is that the analyst is only concerned with one category of interest, then the obvious response is to have the analyst identify all categories within the segment. This approach was adopted for North Dakota in August of the transition year and later extended to all the spring-grain states. Current plans are to use this philosophy of identifying all categories in the corn and soybean experiments. No results are as yet available on the effectiveness of this change to multi-labeling. The cost has increased from the standpoint of interpretation time and additional materials required for the analysis. At this time, the approach appears to have a significant potential value in reducing the proportion estimation bias due to boundary pixels and labeling interior pixels. Though not related to boundary pixels, an added feature is the study of other labeling error sources. Whenever an omission error occurs, it will be known as to what category the analyst was confusing with the category of interest. This will greatly aid analyst training and feedback.

5. SUMMARY

Boundary pixels have been shown to be highly important as sources of proportion estimation error with Landsat data. The particular definition chosen for deciding which pixels are boundary has been found to be critically important to any method of handling the problem. Some of the many methods for handling boundary pixels are described in this paper. Two characteristics are common to all the methods. First, there must be a procedure for deciding which pixels to call boundary; and second, there must be an objective rule for processing them. This processing may be counting, performing a regression analysis on their tabulations, or assigning them to classes based on spectral measurements. One of the proposed methods, the statistical sampling approach to handling boundary pixels, was shown to be a standard by which any of the other methods could be compared objectively and provided a general framework for viewing the problem. This method also provided a way of producing unbiased proportion estimates from labeling techniques that can be applied only to interior pixels.

6. RECOMMENDATIONS

The following considerations are recommended for use in planning a boundary pixel research program:

a. Conduct an additional analysis of the existing data.

b. Conduct a study to determine the actual stability of the β's of the statistical sampling method and establish how they might vary with such factors as the fraction of boundary pixels.

c. Test and evaluate several machine algorithms on how well they can handle boundary pixels using the statistical sampling method as a standard.

d. Investigate over larger geographical areas and with a greater number of segments the feasibility of having analysts label only interior pixels and simply denote boundary pixels as "boundary."

e. Evaluate the dependence of the β's on field size and shape.

7. REFERENCES

1. Detailed Analysis Procedures for LACIE Phase III. LACIE-00720 (JSC-11693), Aug. 1977.

2. Clinton, N. James: LACIE Phase III Labeling Error Characterization: Final Report. LEC-13012 (JSC-14745), Mar. 1979.

3. Pore, M.D.; and Abotteen, R. A.: A Programmed Labeling Approach to Image Interpretation. Presented at LACIE Symposium (Houston, Texas), Oct. 1978. [To be published in the Proceedings of the LACIE Symposium.]

APPENDIX

DATA FROM BOUNDARY PIXEL STUDY ON KANSAS AND NORTH DAKOTA BLIND SITES

APPENDIX

DATA FROM BOUNDARY PIXEL STUDY ON KANSAS AND NORTH DAKOTA BLIND SITES

The tabulation of results for the small boundary pixel study in Kansas and North Dakota are presented in tables A-1 through A-6. Tables A-2 and A-3 contain the raw tabulations of Procedure 1 and LIST for Kansas and A-5 and A-6 for North Dakota. The transformed values amenable for modeling using the statistical sampling approach are in table A-1 for Kansas and A-4 for North Dakota.

TABLE A-1.– BOUNDARY PIXEL STUDY OF KANSAS BLIND SITES

Segment no.	Procedure 1			LIST		
	X_1	X_2	Y	X_1	X_2	Y
1032	1.7308	1.0318	4.2006	4.3270	4.1272	7.2295
1033	0.0000	4.8520	1.6170	0.4902	1.9408	-1.7794
1153	0.4124	21.5556	7.6945	0.2062	5.8788	1.8157
1155	0.0000	6.1815	2.9983	0.9037	4.7550	1.9527
1158	0.0000	0.0000	0.9189	0.6484	0.9598	1.7358
1166	0.4861	1.9120	13.9335	0.4861	1.9120	16.3316
1170	3.5329	10.6872	-0.9777	9.0846	2.2265	4.7522
1175	1.1024	7.3872	6.0505	2.7560	12.4659	11.3977
1180	1.6268	7.3236	-0.7518	3.2536	6.8928	9.8254
1183	1.1606	1.8664	-0.9446	6.3833	3.2662	6.3719
1279	0.6506	1.7136	-3.5894	5.8554	4.2840	3.9796
1285	0.9338	10.5952	0.5601	1.4007	1.4448	3.3909
1290	0.0000	1.0380	-7.0656	9.0080	0.5190	3.8812
1293	0.6973	0.9126	-2.4968	0.6973	1.8252	-2.0405
1295	4.8650	9.1466	5.6433	8.2705	7.2210	10.5389
1297	0.7197	4.1931	-26.1042	0.7197	3.2613	-22.5891
1340	2.2328	6.8775	-1.5304	5.5820	7.7945	13.6204
1343	3.6030	2.3785	-0.8845	4.2035	1.9028	-2.5377
1346	0.0000	1.4223	8.0594	0.6231	0.4741	7.2602
1851	2.0884	3.5992	4.1042	1.0442	1.3497	6.4981
1853	8.7244	31.1483	3.2513	8.2112	4.6490	5.9622
1859	0.9526	2.8926	0.2775	2.3815	1.9284	-2.1330
1861	0.4919	4.2471	0.1832	1.9676	2.3595	4.5303
1864	2.2076	4.1778	-0.6148	4.4152	7.8914	0.6644
1881	4.5990	8.0495	6.5900	2.5550	5.6820	0.2845
1885	5.5627	6.4120	14.4641	8.0912	6.8700	13.6435
$r_{X_1X_2}$ = 0.60633[a]				$r_{X_1X_2}$ = 0.22873[b]		

[a]The correlation coefficient is significantly different from zero at the 1-percent level.

[b]The correlation coefficient is not significantly different from zero.

TABLE A-2.— PROCEDURE 1 STUDY OF KANSAS BLIND SITES[a]

Segment no.	$\hat{p}_{BC}$	N_1	N_2	Base	n_1	n_2	n_{SG1}	n_{SG2}	n_{B1}	n_{B2}
1032	38.7	8635	14038	22673	88	120(1)	69	9(1)	4	2
1033	8.9	779	21923	22702	7	199(3)	1	14	0	10
1153	23.5	235	22558	22793	5	202	3	31(2)	2	44
1155	12.7	1226	18153	19379	7	197(3)	6	9	0	13
1158	20.1	3090	19603	22693	21	180(2)	17	17	0	0
1166	22.1	1103	21585	22688	10	199	5	12	1	4
1170	63.0	14208	8492	22700	124	84(1)	110	19	7	24
1175	43.9	4872	17788	22660	39	170	36	39	2	16
1180	26.0	4608	17924	22732	26	183	17	30	2	17
1183	15.1	4459	18137	22596	34	172(3)	18	12	2	4
1279	30.4	7245	15479	22724	49	159(1)	45	11	1	4
1285	18.1	4782	17978	22760	45	164	19	18	2	22
1290	41.5	9752	11681	21433	101	105(3)	94	12	0	2
1293	12.5	1945	10182	12127	23	184(2)	13	13(2)	1	2
1295	42.5	7527	15225	22752	68	139(2)	53	23(2)	10	19
1297	2.1	2285	20392	22677	14	193(1)	12	42(1)	1	9
1340	56.8	6459	16229	22688	51	156(2)	47	70	4	15
1343	8.7	4448	16712	21160	35	166(8)	12	5	6	5
1346	11.5	849	21858	22707	6	203(1)	4	2	0	3
1851	22.5	4622	18076	22698	39	177(2)	18	20	4	8
1853	28.7	7924	14780	22704	68	140(1)	36	15	17	67
1859	29.5	5523	17209	22732	51	157(1)	32	29	2	6
1861	35.3	10274	12427	22701	92	116(1)	57	15	1	9
1864	35.8	6350	12520	18860	61	143(5)	50	19	4	9
1881	23.8	4629	18015	22644	40	168(1)	17	18	9	17
1885	54.3	11367	11334	22701	99	109(1)	67	13	11	14

[a]Figures in parentheses represent the number of thresholded pixels.

$\beta_1 = 0.72404$ $\qquad s_{\beta_1} = 0.99781$

$\beta_2 = 0.03050$ $\qquad s_{\beta_2} = 0.30608$

TABLE A-3.– LIST BOUNDARY PIXEL STUDY OF KANSAS BLIND SITES[a]

Segment no.	$\hat{P}_{BC}$	N_1	N_2	Base	n_1	n_2	n_{SG1}	n_{SG2}	n_{B1}	n_{B2}
1032	38.7	8635	14038	22673	88	120(1)	62	9	10	8
1033	8.9	779	21923	22702	7	199(3)	1	21	1	4
1153	23.5	235	22558	22793	5	202	3	43	1	12
1155	12.7	1226	18153	19379	7	197(5)	4	15	1	10
1158	20.1	3090	19603	22693	21	180(2)	15	18	1	2
1166	22.1	1103	21585	22688	10	199	4	8	1	4
1170	63.0	14208	8492	22700	124	84(1)	96	22	18	5
1175	43.9	4872	17788	22660	39	170	33	31	5	27
1180	26.0	4808	17924	22732	26	183	13	13	4	16
1183	15.1	4459	18137	22596	34	172(3)	7	10	11	7
1279	30.4	7245	15479	22724	49	159(1)	36	7	9	10
1285	18.1	4782	17978	22760	45	164	15	16	3	3
1290	41.5	9752	11681	21433	101	105(3)	72	10	20	1
1293	12.5	1945	10182	12127	23	184(2)	13	12	1	4
1295	42.5	7527	15225	22752	68	139(2)	37	29	17	15
1297	2.1	2285	20392	22677	14	193(1)	11	36	1	7
1340	56.8	6459	16229	22688	51	156(2)	33	54	10	17
1343	8.7	4448	16712	21160	35	166(8)	10	11	7	4
1346	11.5	849	21858	22707	6	203(1)	3	5	1	1
1851	22.5	4622	18076	22698	39	177(2)	16	17	2	3
1853	28.7	7924	14780	22704	68	140(1)	28	18	16	10
1859	29.5	5523	17209	22732	51	157(1)	32	34	5	4
1861	35.3	10274	12427	22701	92	116(1)	52	11	4	5
1864	35.8	6350	12520	18860	61	143(5)	46	21	8	17
1881	23.8	4629	18015	22644	40	168(1)	21	27	5	12
1885	54.3	11367	11334	22701	99	109(1)	65	17	16	15

[a]Figures in parentheses represent the number of thresholded pixels.

$\beta_1 = 0.58680$ $\quad$ $s_{\beta_1} = 0.49548$

$\beta_2 = 0.54625$ $\quad$ $s_{\beta_2} = 0.47139$

TABLE A-3.— LIST BOUNDARY PIXEL STUDY OF KANSAS BLIND SITES[a]

Segment no.	$\hat{P}_{BC}$	N_1	N_2	Base	n_1	n_2	n_{SG1}	n_{SG2}	n_{B1}	n_{B2}
1032	38.7	8635	14038	22673	88	120(1)	62	9	10	8
1033	8.9	779	21923	22702	7	199(3)	1	21	1	4
1153	23.5	235	22558	22793	5	202	3	43	1	12
1155	12.7	1226	18153	19379	7	197(5)	4	15	1	10
1158	20.1	3090	19603	22693	21	180(2)	15	18	1	2
1166	22.1	1103	21585	22688	10	199	4	8	1	4
1170	63.0	14208	8492	22700	124	84(1)	96	22	18	5
1175	43.9	4872	17788	22660	39	170	33	31	5	27
1180	26.0	4808	17924	22732	26	183	13	13	4	16
1183	15.1	4459	18137	22596	34	172(3)	7	10	11	7
1279	30.4	7245	15479	22724	49	159(1)	36	7	9	10
1285	18.1	4782	17978	22760	45	164	15	16	3	3
1290	41.5	9752	11681	21433	101	105(3)	72	10	20	1
1293	12.5	1945	10182	12127	23	184(2)	13	12	1	4
1295	42.5	7527	15225	22752	68	139(2)	37	29	17	15
1297	2.1	2285	20392	22677	14	193(1)	11	36	1	7
1340	56.8	6459	16229	22688	51	156(2)	33	54	10	17
1343	8.7	4448	16712	21160	35	166(8)	10	11	7	4
1346	11.5	849	21858	22707	6	203(1)	3	5	1	1
1851	22.5	4622	18076	22698	39	177(2)	16	17	2	3
1853	28.7	7924	14780	22704	68	140(1)	28	18	16	10
1859	29.5	5523	17209	22732	51	157(1)	32	34	5	4
1861	35.3	10274	12427	22701	92	116(1)	52	11	4	5
1864	35.8	6350	12520	18860	61	143(5)	46	21	8	17
1881	23.8	4629	18015	22644	40	168(1)	21	27	5	12
1885	54.3	11367	11334	22701	99	109(1)	65	17	16	15

[a]Figures in parentheses represent the number of thresholded pixels.

$\hat{\beta}_1 = 0.58680$ $\qquad s_{\beta_1} = 0.49548$

$\hat{\beta}_2 = 0.54625$ $\qquad s_{\beta_2} = 0.47139$

TABLE A-4.— BOUNDARY PIXEL STUDY OF NORTH DAKOTA BLIND SITES

Segment no.	Procedure 1			LIST		
	X_1	X_2	Y	X_1	X_2	Y
1602	11.6679	11.7234	13.2698	19.7847	9.9198	10.9025
1604	3.7890	12.0624	3.2374	11.3670	27.1404	18.2598
1606	5.0643	7.6320	-0.4161	5.0643	8.1090	0.1466
1616	3.1044	5.3364	3.0688	8.7958	11.1175	10.3936
1619	1.8248	1.0210	0.7811	3.1934	3.0630	3.6269
1622	1.8040	1.0308	2.6032	8.5690	4.1232	7.4998
1625	0.0000	2.8566	-1.7298	0.0000	4.7610	-1.5924
1635	0.0000	0.5266	-3.4571	0.0000	0.5266	-4.6316
1637	0.0000	0.4727	-3.5637	0.0000	0.4727	-2.6183
1640	1.1018	1.3014	-2.3832	4.9581	6.9408	4.9781
1648	2.5242	5.0670	5.6236	3.7863	8.1072	6.6370
1652	3.2697	6.3817	-0.0360	7.4736	15.2179	7.1847
1661	0.9490	4.4838	2.0305	0.4745	0.0000	-2.4059
1663	0.9918	1.4196	4.5095	0.4959	1.4196	3.5858
1889	1.3968	0.0000	0.7601	12.1056	3.0402	8.3741
1902	0.0000	3.8112	2.9232	0.0000	0.4764	1.0176
1903	1.6761	5.6280	5.9761	8.9392	8.9110	12.5011
1913	0.5604	2.7060	11.8000	2.8020	5.9532	14.2834
1927	0.0000	0.9390	-3.4258	0.5146	2.8170	-2.9112
	$r_{X_1X_2} = 0.77586^a$			$r_{X_1X_2} = 0.58203^a$		

[a]The correlation coefficient is significantly different from zero at the 1-percent level.

TABLE A-5.– PROCEDURE 1 BOUNDARY PIXEL STUDY OF NORTH DAKOTA[a]

Segment no.	$\hat{P}_{GT}$	N_1	N_2	Base	n_1	n_2	n_{SG1}	n_{SG2}	n_{B1}	n_{B2}
1602	38.24	6534	6092	12626	102	107	43	7	23	26
1604	52.42	5930	16784	22714	62	147	44	61	9	24
1606	32.94	3937	10637	14574	48	153(8)	33	31	9	16
1616	66.76	11408	11320	22728	97	112	87	42	6	12
1619	52.69	11848	10931	22779	114	94(1)	97	15	4	2
1622	50.28	12271	10400	22671	120	89	92	12	4	2
1625	21.54	1725	20890	22615	14	194(1)	13	34	0	6
1635	16.03	2697	6808	9505	70	136(3)	13	27	0	1
1637	35.77	3992	14552	18544	36	166(7)	31	44	0	1
1640	52.10	10872	11810	22682	87	120(2)	80	24	2	3
1648	20.29	5453	17282	22735	57	150(2)	18	14	6	10
1652	30.62	5090	17609	22699	48	158(3)	32	32	7	13
1661	40.82	6489	13330	19819	69	135(5)	45	35	2	9
1663	51.84	7759	14915	22674	69	139(1)	63	34	2	3
1899	59.33	15205	7469	22674	144	65	116	9	3	0
1902	8.64	315	22445	22760	1	207(1)	0	12	0	8
1903	17.35	2792	19923	22715	22	187	17	4	3	12
1913	29.89	2675	20051	22726	42	163(4)	24	21	2	5
1927	31.36	7207	15378	22585	62	154(2)	53	16	0	2

[a]Figures in parentheses represent the number of thresholded pixels.

$\beta_1 = 0.72656$ $\quad$ $s_{\beta_1} = 0.54149$

$\beta_2 = 0.19814$ $\quad$ $s_{\beta_2} = 0.34859$

TABLE A-6.– LIST BOUNDARY PIXEL STUDY OF NORTH DAKOTA[a]

Segment no.	$\hat{P}_{GT}$	N_1	N_2	Base	n_1	n_2	n_{SG1}	n_{SG2}	n_{B1}	n_{B2}
1602	38.24	6534	6092	12626	102	107	45	10	39	22
1604	52.42	5930	16784	22714	62	147	31	42	27	54
1606	32.94	3937	10637	14574	48	153(8)	32	31	9	17(1)
1616	66.76	11408	11320	22728	97	112	78	36	17	25
1619	52.69	11848	10931	22779	114	94(1)	93	13(1)	7	6
1622	50.28	12271	10400	22671	120	89	80	13	19	8(1)
1625	21.54	1725	20890	22615	14	194(1)	11	36	0	10(1)
1635	16.03	2697	6808	9505	70	136(3)	12	30	0	1
1637	35.77	3992	14552	18544	36	166(7)	31	42	0	1
1640	52 10	10872	11810	22682	87	120(2)	69	21	9	16
1648	20.29	5453	17282	22735	57	150(2)	18	12	9	16
1652	30.62	5090	17609	22699	48	158(3)	26	23	16	31(2)
1661	40.82	6489	13330	19819	69	135(5)	47	42(3)	1	0
1663	51.84	7759	14915	22674	69	139(1)	62	37(1)	1	3
1899	59.33	15205	7469	22674	144	65	104	5	26	6
1902	8.64	315	22445	22760	1	207(1)	0	16	0	1
1903	17.35	2792	19923	22715	22	187	7	2	16	19
1913	29.89	2675	20051	22726	42	163(4)	19	19	10	11(1)
1927	31.36	7207	15378	22585	62	145(2)	52	16	1	6

[a]Figures in parentheses represent the number of thresholded pixels.

$\beta_1 = 0.49496$ $\qquad s_{\beta_1} = 0.19963$

$\beta_2 = 0.43250$ $\qquad s_{\beta_2} = 0.16310$