

COMPUTER CLASSIFICATION OF REMOTELY SENSED MULTISPECTRAL IMAGE DATA BY EXTRACTION AND CLASSIFICATION OF HOMOGENEOUS OBJECTS

R. L. Kettig

D. A. Landgrebe



**School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907**

TR-EE 75-24

May 1975

The work reported in this report was conducted under sponsorship of
NASA Grant NGL-15-005-112 and NASA Contract NAS9-14016.

COMPUTER CLASSIFICATION OF REMOTELY
SENSED MULTISPECTRAL IMAGE DATA BY
EXTRACTION AND CLASSIFICATION OF
HOMOGENEOUS OBJECTS

R. L. Kettig

D. A. Landgrebe

TR-EE 75-24

May 1975

School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907

The work in this report was conducted under sponsorship of NASA Grant
NGL-15-005-112 and NASA Contract NAS9-14016.

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
LIST OF FIGURES	vi
ABSTRACT	ix
CHAPTER 1 INTRODUCTION	1
1.1 Pattern Recognition Systems	1
1.2 Remote Sensing of Earth's Resources	5
1.3 Related Work	9
CHAPTER 2 CLASSIFICATION	14
2.1 Statistical Model of Multi-Spectral Scanner Data	14
2.2 No-Memory Classification	18
2.2.1 Maximum A-Posteriori Probability (MAP) Strategy	18
2.2.2 Maximum Likelihood (ML) Strategy	19
2.2.3 Generalized Maximum Likelihood (GML) Strategy	20
2.2.4 Probability of Error for the GML Strategy	21
2.3 Sample Classification	23
2.3.1 Minimum Distance (MD) Strategy - A Structured Approach to Classification	23
2.3.2 MAP and ML Sample Classification	24
2.3.3 GML Sample Classification	27
2.3.4 Maximum Likelihood vs. Minimum Distance .	28
CHAPTER 3 IMAGE PARTITIONING	35
3.1 Partitioning Logic	35
3.2 Unsupervised Mode	39
3.2.1 Unsupervised Annexation	39
3.2.2 Unsupervised Cell Selection	47
3.3 Supervised Mode	48
3.3.1 Supervised Annexation	48
3.3.2 Supervised Cell Selection	52
3.3.3 Alternative Partitioning Logic	53

CHAPTER 4	CLASSIFICATION RESULTS	55
4.1	Analysis Schemes	56
4.2	Run 71052800 - Crop Identification	61
4.3	Run 72064412 - Classification of Satellite Data	86
4.4	Run 71052501 - Forest Cover Mapping	98
4.5	Run 72032803 - Classification of Satellite Data	111
4.6	General Observations	124
CHAPTER 5	CONCLUSION	129
5.1	Summary	129
5.2	Recommendations for Further Work	131
LIST OF REFERENCES	133
APPENDICES		
Appendix A:	Spatial Correlation	136
Appendix B:	The Compound Decision Approach	152
Appendix C:	Composite Classes	155
Appendix D:	Field Listings	162

LIST OF TABLES

Table	Page
4.2.1 Relative Influence of Each Class on Overall Performance - Run 71052800	68
4.2.2 Matrix of Average Error Rates (%) for Eight Combinations of Significance Levels - Run 71052800	84
4.2.3 Matrix of Overall Error Rates (%) for Eight Combinations of Significance Levels - Run 71052800	84
4.3.1 Relative Influence of Each Class on Overall Performance - Run 72064412	90
4.3.2 Matrix of Average Error Rates (%) for Seven Combinations of Significance Levels - Run 72064412	97
4.3.3 Matrix of Overall Error Rates (%) for Seven Combinations of Significance Levels - Run 72064412	97
4.4.1 Relative Influence of Each Class on Overall Performance - Run 71052501	103
4.4.2 Classes Ordered By Generalized Variance.	103
4.4.3 Matrix of Average Error Rates (%) for Four Combinations of Significance Levels (MV mode) - Run 71052501	109
4.4.4 Matrix of Overall Error Rates (%) for Four Combinations of Significance Levels (MV mode) - Run 71052501	109
4.4.5 Matrix of Average Error Rates (%) for Six Combinations of Significance Levels (MUV mode) - Run 71052501	110

4.4.6 Matrix of Overall Error Rates (%) for Six
Combinations of Significance Levels
- Run 71052501 110

4.5.1 Relative Influence of Each Class on Overall
Performance - Run 72032803. 115

4.5.2 Classes Ordered According to Generalized
Variance - Run 72032803. 115

4.5.3 Effect of Parameters on Performance
(unsupervised mode). 121

Appendix
Table

D1 Test Areas for Run 71052800 163

D2 Test Areas for Run 72064412 168

D3 Reference Areas for Run 71052501. 171

D4 Test Areas for Run 72032803 175

LIST OF FIGURES

Figure		Page
2.3.3.1	Two Normal Densities With Low Separability . . .	29
2.3.3.2	A Bound on Conditional Probability of Error for the Maximum Likelihood Strategy as a Function of the Sample Size.	30
3.1.1	Basic Flow Chart for a Two-Level, Conjunctive Partitioning Algorithm	38
4.2.1	Correspondence Between Channel Number and Spectral Band for Aircraft Scanner	62
4.2.2	Classification Performance vs. Processing Scheme - Run 71052800.	67
4.2.3	Performance By Class (Run 71052800).	70
4.2.4	Gray-Scale-Coded Classification Map - Produced by CLASSIFYPOINTS	76
4.2.5	Gray-Scale-Coded Classification Map - Produced by ECHO	76
4.2.6	Logogrammatic Classification Maps.	77
4.2.7	Class Improvement vs. Log-Generalized Variance	79
4.2.8	Effect of Annexation Threshold (t) on Average Performance - Run 71052800	81
4.2.9	Effect of Annexation Threshold (t) on Overall Performance - Run 71052800	83
4.2.10	Error Rate and CPU Time for Four Classification Schemes - Run 71052800.	85
4.3.1	Correspondence Between Channel Number and Spectral Band for LANDSAT-1 Scanner.	87

4.3.2	Classification Performance vs. Processing Scheme - Run 72064412.	89
4.3.3	Performance By Class (Run 72064412)	91
4.3.4	Effect of Annexation Threshold (t) on Average Performance - Run 72064412	95
4.3.5	Effect of Annexation Threshold (t) on Overall Performance - Run 72064412	96
4.3.6	Error Rate and CPU Time for Four Classification Schemes - Run 72064412.	99
4.4.1	Classification Performance vs. Processing Scheme - Run 71052501.	101
4.4.2	Performance By Class (Run 71052501)	104
4.4.3	Effect of Annexation Threshold (t) on Average Performance - Run 71052501	107
4.4.4	Effect of Annexation Threshold (t) on Overall Performance - Run 71052501	108
4.4.5	Error Rate and CPU Times for Five Classification Schemes - Run 71052501.	112
4.5.1	Classification Performance vs. Processing Scheme - Run 72032803.	114
4.5.2	Performance By Class (Run 72032803)	116
4.5.3	Effect of Annexation Threshold (t) on Average Performance - Run 72032803	119
4.5.4	Effect of Annexation Threshold (t) on Overall Performance - Run 72032803	120
4.5.5	Error Rate and CPU Time for Four Classification Schemes - Run 72032803.	123
Appendix		
Figure		
A1	Spatial Correlation Coefficients - Run 71052501 Region: lines 501-800, columns 1-100.	142
A2	Spatial Correlation Coefficients - Class FORAGE - Run 71052501	143

A3	Spatial Correlation Coefficients - Class DECIDUOUS - Run 71052501	144
A4	Spatial Interchannel Correlation Coefficients - Class FORAGE - Run 71052501	145
A5	Spatial Interchannel Correlation Coefficients - Class DECIDUOUS - Run 71052501.	146
A6	Spatial Correlation Coefficients - Run 72032803 Region: lines 260-559, columns 998-1097.	147
A7	Spatial Correlation Coefficients - Class CORN - Run 72032803	148
A8	Spatial Correlation Coefficients - Class TOWN - Run 72032803	149
A9	Spatial Interchannel Correlation Coefficients - Class CORN - Run 72032803	150
A10	Spatial Interchannel Correlation Coefficients - Class TOWN - Run 72032803	151
C1	A Case of Spatial Texture on the Order of the Pixel Size	157
C2	Texture in One Spatial Dimension Only.	158
C3	Two Densities and Their Compound Density for Spatial Texture on the Order of a Pixel.	161

ABSTRACT

A method of classification of digitized multispectral images is developed and experimentally evaluated on actual earth resources data collected by aircraft and satellite. The method is designed to exploit the characteristic dependence between adjacent states of nature that is neglected by the more conventional simple-symmetric decision rule. Thus contextual information is incorporated into the classification scheme. The principle reason for doing this is to improve the accuracy of the classification. For general types of dependence this would generally require more computation per resolution element than the simple-symmetric classifier. But when the dependence occurs in the form of "redundance", the elements can be classified collectively, in groups, thereby reducing the number of classifications required. Thus a potential exists for increased, rather than decreased, efficiency.

Basically, the method can be thought of as an image partitioning transformation that delineates (extracts) the statistically homogeneous groups (samples) of elements and a

sample classifier that classifies them. Various possibilities are considered for both operations.

The main result is that a combination of the two is found which consistently provided the lowest error rates, rivaling those obtained when ground observational data was used to delineate the samples manually. The relative efficiency of this method depends largely on the complexity of the classification task. For relatively complex classification, the time saved by sample classification more than compensates for the extra time required for partitioning. But for relatively simple classification the simple-symmetric classifier is faster. Of course in the latter case, efficiency is not as great a consideration since the total CPU time involved is much less than in the former case.

CHAPTER 1

INTRODUCTION

The general objective of this thesis is to advance the state of the art of pattern recognition as it is applied in remote sensing technology. This chapter opens with a discussion of pattern recognition and remote sensing systems that leads up to the specific problem under investigation. In the process much of the prevalent terminology is introduced. Other work that is related to this problem is discussed in Section 1.3.

1.1 Pattern Recognition Systems

Man's most abundant source of information about a scene is the radiant electromagnetic energy which emanates from it. The information is embodied in the spatial, spectral, and temporal variations (patterns) of the radiance. The general process of extracting information from patterns (radiance or otherwise) is known as pattern recognition. The most common form of pattern recognition is "classification", the assignment of an observed pattern to one of several prespecified categories (classes). This requires a certain degree of experience; i.e. the recognition system must know the possible classes and have

some sort of unique characterization for each one. Typically this experience is "learned" from representative "training" patterns (or sets of patterns) that are supplied as references for each class. In the simplest case, each set of patterns is a complete characterization of the class it represents. Then classification is a straightforward matter of comparison. More generally, a statistical characterization might be the only adequate approach, and the training patterns might be used to estimate statistical quantities. Classification then becomes a problem in statistical decision theory.

Of course it is not always possible to prespecify the categories that a pattern might belong to. This is often true in scene analysis, where the number of possibilities can be enormous. Then pattern recognition can take the form of "description". In general, pattern recognition can involve both classification and description. A complex scene composed of relatively simple objects is often described by classifying the objects and recording their relative positions and orientations in the scene. This description might be considered the final result, or it might in turn be used to classify the scene itself.

All systems that extract information from a scene consist of a data collection system and a data processor. The purpose of the data collection system is to reduce the scene to a manageable number of measurements (features) without losing the desired information. Further reduction

(feature selection) is often possible in the processor. The choice of features obviously depends upon the information that is desired, and conversely the information that can be extracted depends upon the choice of features. Most collection systems are similar in many ways to the human eye, which forms features by "sampling" the spatial, spectral, and temporal dimensions, thereby converting a scene into series of electrical pulses. Spatial sampling can be accomplished by forming an image of the scene on an array of detectors (electrical or chemical) or by scanning the image with an electrical detector. The resolution element of such a system is the projection of the detector back through the optical system onto the scene. It is commonly called a "pixel", short for picture element. The overall system resolution depends on both the pixel size and the interval between samples, which are normally about equal.

Spectral sampling is accomplished by measuring the radiance of each resolution element with detectors (channels) that are sensitive to different spectral bands. A prism, grating, or interference filter is often used to separate the radiant energy spectrally before detection. Temporal sampling is accomplished merely by taking spatial and spectral samples at discrete times.

Depending on the type of information that is desired, one can emphasize or de-emphasize a particular dimension by sampling it relatively many or relatively few times. A

single black-and-white photograph, for example, emphasizes spatial information since it is created by sampling only once spectrally and once temporally. A color photograph contains three spectral samples and thus emphasizes both spatial and spectral information. The extent to which a pattern is sampled falls under the category of "measurement complexity". Under-sampling results in loss of information, but over-sampling results in an excess of data to process. Technically, the data dimensionally increases faster than its intrinsic dimensionality.

"Data dimensionality" refers to the dimension of the measurement or observation space, in which a sampled pattern can be considered an observation of a multi-dimensional random variable. The probability density of this random variable is a function of N variables (dimensions), where N is the number of measurements. The "intrinsic" dimensionality of a random variable (X) is the minimum dimension that another random variable (Y) can have if X is uniquely related to Y . Thus it is the minimum number of measurements that could be used to convey the same information as X if the relationship were known. Over-sampling increases the data dimensionality, but the individual measurements tend to be more highly correlated causing the information conveyed per measurement to decrease.

The information that can be extracted from an image is also limited by the sophistication of the processor which

must handle the data. Just as the necessary measurement complexity depends on the information being sought, so does the method of processing. The human mind is often an extremely good processor, particularly when the information is of primarily a spatial nature. For this purpose the data is presented in visual image form, which is known as an "image-oriented" processing system. By contrast, in a "numerically-oriented" system the decision-making element is a computer, and the visual image plays little or no part. Advantages of the computerized approach are its high load (volume) capacity, comparatively low cost under high load, and capacity to handle high measurement complexity.

1.2 Remote Sensing of Earth's Resources

An important subject before the engineering and scientific community at the present time is the processing of scenes which represent tracts of the earth's surface as viewed from above. A typical scene consists primarily of regular and/or irregular regions arranged in a patchwork manner and each containing one class of surface cover type. These homogeneous regions are the "objects" in the scene. A basic processing goal is to locate and classify the objects and produce a description of the scene in terms of tabulated results and/or a "type-map". As in other image processing applications, the locations and spatial features (e.g. size, shape, orientation) of objects are revealed by changes in average spectral properties that occur at boundaries. But

unlike most other applications, the spatial features of an object often have only a weak relationship to its class. Research has shown, however, that many classes can be distinguished reasonably well on the basis of their spectral features, using statistical pattern classification techniques. Current research is directed toward use of temporal features as well, but not in this investigation.

Our interest is in the numerically-oriented system approach to processing these scenes. The input to the system is in the form of digitized multi-spectral scanner (MSS) data stored on magnetic tape. A typical multi-spectral scanner samples the spectral dimension and one spatial dimension. The second spatial dimension is provided by the motion of the platform which carries the scanner over the region of interest, generating a raster-type scan. The temporal dimension is provided by rescanning the region at different times.

Computer classification of MSS data is typically done by applying a "simple symmetric" decision rule to each pixel. This means that each pixel is classified individually on the basis of its spectral measurements alone. A basic premise of this technique is that the objects of interest are large compared to the size of a pixel. Otherwise a large proportion of pixels would be composites of two or more classes, making statistical pattern classification unreliable; i.e. the prespecified categories would be inadequate to describe the actual states

of nature. (For later reference we shall call this "Premise A".) Since the sampling interval is usually comparable to the pixel size (to preserve system resolution), it follows that each object is represented by an array of pixels. This suggests a statistical dependence between consecutive states of nature, which the simple symmetric classifier fails to exploit. To reflect this property, we shall refer to simple symmetric classification as "no-memory" classification.

One method for dealing with dependent states is to apply the principles of compound decision theory or sequential compound decision theory. Abend [1] points out that a sequential procedure can be implemented relatively efficiently when the states form a low-order Markov chain. However the prospect is considerably less attractive when they form a Markov mesh, which is a more suitable model for two-dimensional scenes. Furthermore, estimation of the state transition probabilities could be another significant obstacle to implementation of such a procedure. A short appendix on the compound decision approach is included in this thesis.

The compound decision formulation is a powerful approach for handling very general types of dependence. This suggests that perhaps by tailoring an approach more directly to the problem at hand, one can obtain similar results with considerable simplification. A distinctive characteristic of the spatial dependence in MSS data is redundance; i.e. the probability of transition

from state i to state j is much greater if $j=i$ than if $j \neq i$, because the sampling interval is small compared to the size of an object. This suggests the use of an "image partitioning" transformation to delineate the arrays of statistically similar pixels before classifying them. Since each homogeneous array represents a statistical "sample" (a set of observations from a common population), a "sample classifier" could then be used to classify the objects. In this way, the classification of each pixel in the sample is a result of the spectral properties of its neighbors as well as its own. Thus its "context" in the scene is used to provide better classification. The acronym ECHO (extraction and classification of homogeneous objects) designates this general approach.

A characteristic of both no-memory and compound decision techniques is that the number of classifications which must be performed is much larger than the actual number of objects in the scene. When each classification requires a large amount of computation, even the no-memory classifier can be relatively slow. An ECHO technique would substantially reduce the number of classifications, resulting in a potential increase in speed (decrease in cost). Whether or not this potential is realized depends on the efficiency of the partitioning operation.

The goal of the current investigation is to further the development of the ECHO concept. In particular, various processing options are devised, implemented, and tested on a

wide variety of data sets. Input parameters are varied to determine their effect, and performance comparisons are made using no-memory classification as a norm.

1.3 Related Work

The recent literature contains numerous references to image partitioning algorithms. Robertson [2] divides them into two main categories. "Boundary seeking" algorithms characteristically attempt to exploit object contrast. These techniques include local gradient [3,4], template matching [5], two-dimensional function fitting [6], clustering [4,7], and gradients estimated from variable-sized neighborhoods [8]. Two of these have been implemented with digitized multispectral imagery.

Anuta [4], investigated a multivariate extension of a two-dimensional gradient operator. The gradient operator of a unispectral image maps each pixel into a number which reflects the average positive difference between that pixel and its neighbors. The multivariate operator sums these numbers over all spectral features for each pixel. Since the differences are generally larger for boundary pixels than for non-boundary pixels, thresholding this sum (for each pixel) at the "proper" level provides a boundary enhanced version of the original image. This technique is relatively fast, but it has several serious problems. First, it is inherently noisy, which is typical of differentiation techniques. It is also very sensitive to

the threshold level used. Furthermore the boundaries derived by this technique often fail to close upon themselves. For example, a boundary line may become discontinuous or fade out completely, leaving the objects ambiguously defined. In special cases where the object shape is restricted [3,9], the true boundaries can sometimes be deduced, but in general they cannot. This may not be a serious drawback for applications such as image registration, but closed boundaries are necessary for sample classification. This particular problem is common to all the boundary seeking algorithms mentioned above.

Wacker [7] developed an algorithm for MSS data which performs a cluster analysis (unsupervised classification) of a small region of the image and then scans the result for the presence of a boundary. The estimated boundary structure for the entire image is obtained simply by taking the union of the boundaries found in all such regions. This is a much more time-consuming process, but it is less noisy and less sensitive to input parameters. Of course it suffers from the same open boundary problem as the other boundary seeking algorithms.

The other category of image partitioning algorithms can be called "object seeking" algorithms, which characteristically exploit the internal regularity (homogeneity) of the objects. As the name implies, an object seeking algorithm always produces well-defined samples (and thus closed boundaries as well). There are two

opposite approaches to object seeking, which we shall call conjunctive and disjunctive. A conjunctive algorithm begins with a very fine partition and simplifies it by progressively merging adjacent elements together that are found to be similar according to certain statistical criteria [10,11]. A disjunctive algorithm begins with a very simple partition and subdivides it until each element satisfies a criterion of homogeneity. For example, Robertson's algorithm [2,12] is based on the premise that if a region contains a boundary, splitting the region arbitrarily will usually produce two subregions with significantly different statistical characteristics.

Early work in the application of sample classification to MSS data was reported by Huang [13]. His method of "polling" requires classification of the individual pixels in the sample and is thus relatively inefficient. Wacker and Landgrebe [14] investigated the "minimum distance approach" using parametric and non-parametric methods. Both studies relied on manual definition of the object boundaries, based on actual surface (ground) observations, to locate the samples that were classified.

We combined Rodd's conjunctive partitioning algorithm with a minimum distance sample classifier and observed an improvement in classification accuracy over conventional no-memory classification, but processing time was increased [15]. Gupta and Wintz [16] added a test of second order statistics to Rodd's first order test, but obtained

essentially the same results as the first order test at greater cost in processing time. Robertson [2,12] implemented a disjunctive partitioning algorithm with the same minimum distance classifier. He obtained about the same classification accuracy as conventional no-memory classification with an order of magnitude increase in processing time. This points to one essential difference between the disjunctive and conjunctive approaches. With a disjunctive approach, every time a region is divided new sample statistics must be calculated from raw data. With a conjunctive approach, every time two regions are merged the statistics for the resultant region can be obtained merely by "pooling" the statistics of the original two subregions. This results in a significant computational advantage for the conjunctive approach.

The current investigation is devoted to further development of the conjunctive approach. A much faster sample classifier is proposed and tested. This problem is discussed in Chapter 2. New statistical criteria are proposed as well as new object seeking logic in Chapter 3. Extensive test results appear in Chapter 4, comparing different algorithms against each other and against conventional no-memory classification. The main result is that the stability, classification accuracy, and speed of the ECHO technique have been greatly improved. Compared to the no-memory classifier, consistently lower error rates are observed using an ECHO approach, and for a reasonably

complex classification its efficiency exceeds that of the conventional method.

CHAPTER 2

CLASSIFICATION

The motivation for object extraction is to enable faster and more accurate classification of the pixels within the object. In Section 2.3 we discuss the classification algorithms that accomplish this. They are based on a certain model of the objects to be classified, which is described next.

2.1 Statistical Model of Multi-Spectral Scanner Data

As we have indicated, a typical scene consists primarily of objects whose boundaries form a partition of the scene. The partition is generally unknown at the outset, but we can at least assume that it is relatively coarse compared to the size of a pixel. Each object in the scene belongs to some class. For representation purposes, each class is divided into one or more "subclasses". They are also called "spectral classes" (as opposed to "informational classes") to indicate that they can be distinguished spectrally although it may not be useful to do so. Let W_{ij} denote the j th subclass of the i th class. Let F denote an object (represented by an array of pixels), and let \underline{x} denote a pixel in some object. (The underbar is used

to indicate a q -dimensional variable ($\underline{X} \in R^q$), where q henceforth denotes the number of spectral channels.) Then $F \in W_{ij}$ denotes the event that F belongs to the subclass W_{ij} . The a-priori probability of this event is denoted by $P(F \in W_{ij})$. In accordance with Section 1.2, we ignore any statistical dependence of this event on the spatial features of F . If there were a strong, known dependence then it could be used to help classify F , but that is not our intention. A consequence of this assumption is that $P(\underline{X} \in W_{ij}) = P(F \in W_{ij})$, and we denote both quantities simply by $P(W_{ij})$.

The pixels within a given object of a given spectral class are completely characterized by their class-conditional, joint, probability distribution function. For no-memory classification, such a complete model is unnecessary; only the marginal distribution of each pixel is required. Furthermore, the pixels within a single object are usually assumed to have a common (i.e. stationary) marginal distribution, which is due to the homogeneity of the types of objects typically encountered in remote sensing applications. Although the data is digitized, it is convenient to represent this q -variate distribution by a continuous-parameter probability density function (pdf) which, for subclass W_{ij} , is denoted by $p(\underline{X}=\underline{x} | \underline{X} \in W_{ij})$ or simply by $p(\underline{x} | W_{ij})$. (The vertical bar indicates conditional probability).

Two pixels in spatial proximity to one-another are unconditionally correlated, with the degree of correlation decreasing as the distance between them increases. Much of this correlation is attributable to the effect of dependent states, discussed in Section 1.2, which is the effect we wish to exploit. For simplicity we shall ignore other sources of correlation. Thus we assume that pixels within the same object are class-conditionally independent; i.e. each object is a "simple" sample from one of the spectral class populations. Then the joint pdf of the pixels can be expressed as just the product of their marginal pdf's. This approximation leads to fast, effective (though suboptimal) processing algorithms, but theoretical predictions based on this simplified model should be interpreted cautiously. This aspect of modeling is discussed at greater depth in Appendix A.

It is possible to express other statistical characteristics in terms of the ones above. If W_i denotes the i th class, then

$$P(\underline{X} \in W_i) = P(\bigcup_j \underline{X} \in W_{ij}) = \sum_j P(W_{ij}) \quad 2.1.1$$

where U denotes the union of events. The pdf of \underline{X} , conditional on this event, is given by

$$p(\underline{x}|W_i) = \frac{1}{P(W_i)} \sum_j p(\underline{x}|W_{ij})P(W_{ij}) \quad 2.1.2$$

This equation defines the representation of a class in terms of its subclasses. The unconditional pdf can be written in

two ways:

$$p(\underline{x}) = \sum_i \sum_j p(\underline{x}|W_{ij})P(W_{ij}) = \sum_i p(\underline{x}|W_i)P(W_i) \quad 2.1.3$$

Within this framework, all that is required to complete the statistical model for a given scene (or class of scenes) is to specify the spectral classes that are present and assign an a-priori probability and conditional pdf to each. Of course the true distributions are assigned by nature, and the accuracy of the model depends on how well we can estimate them. Fortunately we are usually able to obtain estimates of the class-conditional pdf's based on training samples taken directly from the data set. For this we usually rely on actual surface (ground) observations or manual photo-interpretation to locate areas representing each class of cover-type. For the purpose of classifier design, we assume that the size of each sample is sufficiently large that the error in the corresponding distribution estimate is negligible. The subject of training is discussed further in Chapter 4.

The distribution estimates can be parametric or non-parametric in general. It has been found that the multi-variate normal (MVN) distribution is a reasonable model for MSS data [17]; i.e. $p(\underline{x}|W_{ij}) = N(\underline{M}_{ij}, \underline{C}_{ij}; \underline{x})$, where

$$N(\underline{M}, \underline{C}; \underline{x}) = (|2\pi\underline{C}| \exp((\underline{x}-\underline{M})' \underline{C}^{-1} (\underline{x}-\underline{M})))^{-\frac{1}{2}} \quad 2.1.4$$

(Note that $(\underline{x}-\underline{M})'$ denotes the transpose of vector $(\underline{x}-\underline{M})$.)

It follows that if $\underline{x} \in W_{ij}$,

$$E(\underline{x}) = \underline{M}_{ij}$$

$$E((\underline{X}-\underline{\mu}_{ij})(\underline{X}-\underline{\mu}_{ij})') = \underline{C}_{ij}$$

where $E(\cdot)$ denotes statistical expectation. Thus $\underline{\mu}_{ij}$ and \underline{C}_{ij} are the mean vector and covariance matrix of the subclass distribution. Note that in order to obtain a parametric estimate of a MVN distribution, it is only necessary to estimate its first and second order moments. This is the approach that we will use.

2.2 No-Memory Classification

In order to introduce certain concepts that will be useful later, we now review some common techniques of no-memory classification including (in one case) a discussion of a bound on the probability of error.

2.2.1 Maximum A Posteriori Probability (MAP) Strategy

Let \underline{x} be a pixel, as before. Under the hypothesis that $\underline{x} \in W_i$, the pdf of \underline{x} is $p(\underline{x}=\underline{x}|\underline{x} \in W_i)$, which is given by equation 2.1.2. Assuming that this function is accurately known, the hypothesis is "simple". The goal of classification is to devise a strategy for choosing one of the possible classes (hypotheses) based on \underline{x} , the observed value of \underline{x} ; i.e. we must specify a function, $W(\underline{x})$, which maps \underline{x} into the set of possible classes. We can maximize the probability of a correct decision by always choosing the class, W_i , which has the maximum a posteriori probability, $P(\underline{x} \in W_i | \underline{x}=\underline{x})$. To show this we merely write the probability of a correct decision in the following form:

$$P(\underline{X} \in W(\underline{X})) = \int_{\underline{X} \in R^q} P(\underline{X} \in W(\underline{x}) | \underline{X}=\underline{x}) p(\underline{X}=\underline{x}) d\underline{x} \quad 2.2.1.1$$

It is apparent that this quantity is maximized with respect to the decision function by adopting the MAP decision rule. To implement this strategy we use the mixed form of Bayes rule to write

$$P(\underline{X} \in W_i | \underline{X}=\underline{x}) = \frac{p(\underline{X}=\underline{x} | \underline{X} \in W_i) P(W_i)}{p(\underline{X}=\underline{x})} \quad 2.2.1.2$$

The denominator is independent of i , so we need only to seek the i which maximizes the numerator. In other words, for a given observation, \underline{x} , $W(\underline{x})$ is chosen such that

$$p(\underline{X}=\underline{x} | \underline{X} \in W(\underline{x})) P(W(\underline{x})) = \max_i p(\underline{X}=\underline{x} | \underline{X} \in W_i) P(W_i) \quad 2.2.1.3$$

This result can also be obtained as a special case of Bayes decision rule for minimum risk when a "zero-one" loss function is assumed (i.e. when the risk equals the probability of error). Thus it is often referred to as "Bayes classifier".

2.2.2 Maximum Likelihood (ML) Strategy

When all the classes are equiprobable, the MAP decision rule reduces to

$$p(\underline{X}=\underline{x} | \underline{X} \in W(\underline{x})) = \max_i p(\underline{X}=\underline{x} | \underline{X} \in W_i) \quad 2.2.2.1$$

As a function of i , the statistic $p(\underline{X}=\underline{x} | \underline{X} \in W_i)$ is called the likelihood function, so this decision rule is called the maximum likelihood strategy.

The ML strategy is usually a reasonable approach even when the classes are not equiprobable. In particular, the MAP strategy tends to discriminate against classes whose a-priori probability is low; i.e. it encourages a relatively large conditional probability of error when a "rare" class occurs in order to minimize the overall error probability. Thus when one is interested in classifying the less abundant classes (as well as the more abundant classes) with reasonable accuracy, the MAP strategy may not be as desirable as one which makes more errors but distributes them more equitably among the classes. With the ML strategy, the conditional probability of error when the i th class occurs depends only on the degree of statistical "separability" (or "distance") between class i and the other classes. It is independent of the a-priori probability of class i .

2.2.3 Generalized Maximum Likelihood (GML) Strategy

Often the a-priori subclass probabilities are unknown. Then the hypothesis that $\underline{x} \in W_i$ is a composite hypothesis; i.e. $p(\underline{x}|W_i) = \sum_j A_{ij} p(\underline{x}|W_{ij})$ where the coefficients are unknown. Of course we know that $A_{ij} \geq 0$ and $\sum_j A_{ij} = 1$. A procedure that has been found to be useful in this situation is to form maximum likelihood estimates of the unknown parameters under each hypothesis. Then the unknowns are replaced by their estimated values, and a hypothesis is selected by the ML strategy. We will refer to this

procedure as the generalized maximum likelihood strategy. The resultant decision rule can be simply expressed in the following form:

$$p(\underline{x}|V(\underline{x})) = \max_i \max_j p(\underline{x}|W_{ij}) \quad 2.2.3.1$$

where $V(\underline{x})$ maps \underline{x} into the set of spectral classes. Then $W(\underline{x})$ is simply defined to be the informational class containing $V(\underline{x})$.

We note that the GML strategy is equivalent to a ML strategy over the set of spectral classes. Thus when all spectral classes are equiprobable, it maximizes the probability of classifying the observation into the correct one.

2.2.4 Probability of Error For The GML Strategy

Let V_i denote the i th spectral class, and let E be the event that \underline{x} is classified into the wrong spectral class.

Then

$$P(E) = \sum_j P(E|\underline{X} \in V_j)P(V_j) \quad 2.2.4.1$$

If E_{ij} is the event that V_i produces a larger likelihood statistic than V_j , then

$$P(E|V_j) = P\left(\bigcup_{i \neq j} E_{ij} | V_j\right) \leq \sum_{i \neq j} P(E_{ij} | V_j) \quad 2.2.4.2$$

Thus it is of some interest to investigate the pairwise error probabilities.

Let $F_{ij}(T) = P(R_{ij}(\underline{X}) > T | \underline{X} \in V_j) = P(L_{ij}(\underline{X}) > \ln(T) | \underline{X} \in V_j)$, where $R_{ij}(\underline{X})$ and $L_{ij}(\underline{X})$ are the random variables:

$$R_{ij}(\underline{X}) = \frac{p(\underline{X}|V_i)}{p(\underline{X}|V_j)} = \text{likelihood ratio} \quad 2.2.4.3$$

$$L_{ij}(\underline{X}) = \ln R_{ij}(\underline{X})$$

Then $P(E_{ij}|V_j) = F_{ij}(1)$. Unfortunately the conditional distribution functions of $R_{ij}(\underline{X})$ and $L_{ij}(\underline{X})$ are not usually explicitly available. But, if we can find the moment generating function, $\phi_{ij}(u)$, corresponding to the conditional distribution of $L_{ij}(\underline{X})$ given $\underline{X} \in V_j$, then we can bound $F_{ij}(T)$ as follows:

$$F_{ij}(T) \leq T^{-u} \phi_{ij}(u), \quad 0 \leq u \quad 2.2.4.4$$

Furthermore

$$F_{ij}(T) \leq T^{u-1} \phi_{ij}(u), \quad u \leq 1$$

This is known as the Chernoff bound [18].

By definition:

$$\phi_{ij}(u) = E(\exp(uL_{ij}(\underline{X})) | \underline{X} \in V_j) \quad 2.2.4.5$$

When the subclasses are MVN, the expectation can be explicitly evaluated [18]. The result is:

$$\phi_{ij}(u) = \sqrt{\frac{|C_j|^u |C_i|^{1-u}}{|uC_j + (1-u)C_i|} \exp(-u(1-u)(\underline{M}_i - \underline{M}_j)'(uC_j + (1-u)C_i)^{-1}(\underline{M}_i - \underline{M}_j))} \quad 2.2.4.6$$

Substituting into 2.2.4.4 provides the desired bound. In particular, for $u = .5$ we have:

$$F_{ij}(T) \leq \phi_{ij}(.5) / \sqrt{T} \quad 2.2.4.7$$

We note in passing that $-\ln \phi_{ij}(.5)$ is simply the Bhattacharyya "distance" between subclasses V_i and V_j .

Combining equations 2.2.4.1, 2.2.4.2, and 2.2.4.7 gives an expression for a bound on $P(E)$:

$$P(E) \leq \sum_j P(V_j) \sum_{\substack{i \\ i \neq j}} \phi_{ij} \quad (2.2.4.8)$$

By dropping the terms for which V_i and V_j are in the same class, this becomes a bound on the total probability of error.

2.3 Sample Classification

For the purposes of this section we can assume that the partition of the scene is known and we simply want to classify the objects. (In Chapter 3 we discuss conjunctive partitioning algorithms for actually estimating the partition.) We shall treat each object separately, thus ignoring any contextual information resulting from spatial relationships of objects. So we observe a set (sample) of q -dimensional random variables, $X = (X_1, \dots, X_n)$, from a common population, and our goal is to classify them.

2.3.1 Minimum Distance (MD) Strategy

- A Structured Approach to Classification

A structured approach is one in which the basic form of the processor is simply assumed, perhaps leaving certain parameters or options to the discretion of the user. A reasonable procedure is to choose some characteristic that differs from class to class, measure it for the sample to be classified, and select the class whose characteristic most closely matches this observation. Under our assumption of

2.3.1

simple samples, each class is completely characterized by a known q -dimensional pdf. Therefore, in MD classification, the n data vectors are used to estimate the pdf of the population, and the class is selected whose pdf is closest to this estimate as measured by some appropriately defined "distance measure" on the set of density functions. Ideally one would like to choose the density estimator and distance measure in some optimum manner, but in practice the best guidelines are provided by experimental investigations [14]. Note that a possible drawback of the MD strategy is that the sample size (n) must not be too small to obtain meaningful density estimates.

When spatial correlation is introduced into the model (Appendix A), each class is only partially characterized by a simple q -dimensional pdf. Although perhaps not as effective as a higher dimensional pdf would be, it is still a reasonable and valid characteristic for distinguishing between classes. In fact if the spatial correlation is class-invariant (such as that induced by the scanner), the q -dimensional pdf might be just as effective as the higher dimensional one.

2.3.2 M.A.P. and M.L. Sample Classification

In contrast to the MD strategy, the MAP strategy is a completely non-structured approach. The decision rule is determined solely by the criterion of minimum error rate with no a-priori restrictions. Of course a greater degree

of statistical information is also required (the a-priori class probabilities). We can obtain the MAP decision rule by direct extension of 2.2.1.3 if we consider X as a qn -dimensional random variable to be classified. Let x be the set of variates (x_1, \dots, x_n) and the event $X=x$ be defined as the joint event $X_i=x_i, i=1, \dots, n$. Then, under the hypothesis $X \in W_i$, the pdf of X is

$$\begin{aligned} p(X=x|W_i) &= \frac{1}{P(W_i)} \sum_j P(W_{ij}) p(X=x|W_{ij}) & 2.3.2.1 \\ &= \frac{1}{P(W_i)} \sum_j P(W_{ij}) \prod_{m=1}^n p(X_m=x_m|W_{ij}) \end{aligned}$$

The MAP decision rule can be stated as follows:

$$p(X=x|W(x))P(W(x)) = \max_i p(X=x|W_i)P(W_i) \quad 2.3.2.2$$

There is no minimum sample size required to implement this strategy. For $n=1$ it simply reduces to MAP no-memory classification (2.2.1.3).

Note that we have represented the joint pdf of a sample in terms of the marginal pdf of one pixel. When spatial correlation is present, this is no longer a fully adequate representation. But as in the case of MD classification, it still provides a useful statistic for distinguishing classes while avoiding the complexities of more rigorous representations.

So far we have tacitly assumed that the decision rule must assign the same class to all the pixels in the sample. With this type of strategy, either all the pixels are classified correctly or all are misclassified. Thus the MAP

decision rule maximizes the average number of times that all the pixels in X are classified correctly. But performance is generally measured by just the average number of pixels in X that are classified correctly each time. We can show that the MAP decision rule maximizes this criterion also. Any decision rule that we adopt must assign a class to \underline{X}_i for any event $X=x$. We denote this mapping by $W_i(x)$. Let $Z(\cdot)$ be an indicator function, i.e. a zero-one random variable which assumes the value 1 if and only if the event specified in the argument actually occurs. The number of elements correctly classified in the sample is given by the random variable

$$N = \sum_{i=1}^n Z(\underline{X}_i \in W_i(X)) \quad 2.3.2.3$$

$$\begin{aligned} E(N) &= \sum_{i=1}^n E(Z(\underline{X}_i \in W_i(X))) = \sum_{i=1}^n P(\underline{X}_i \in W_i(X)) \\ &= \sum_{i=1}^n \int_{x \in R^{qn}} P(\underline{X}_i \in W_i(x) | X=x) p(X=x) dx \end{aligned}$$

The integration implied here is a qn -dimensional one. Note that the event $\underline{X}_i \in W_i(x)$ is equivalent to $X \in W_i(x)$, so all terms of this summation are identical, with the possible exception of the decision function. Thus the decision function which maximizes one term also maximizes the others. This confirms that the optimum decision rule assigns the same class to all the elements of the field. Denoting this decision function by $W(x)$, we have

$$E(N) = n \int P(X \in W(x) | X=x) p(X=x) dx \quad 2.3.2.4$$

This, of course, is maximized by the MAP strategy (2.3.2.2).

The ML strategy follows directly from the MAP strategy by dropping the a-priori probabilities. The result is

$$p(X=x|W(x)) = \max_i p(X=x|W_i) \quad 2.3.2.5$$

2.3.3 G.M.L. Sample Classification

We can obtain the GML decision rule by direct extension of 2.2.3.1. The result is

$$p(X=x|V(x)) = \max_i \max_j p(X=x|W_{ij}) = \max_i p(X=x|V_i) \quad 2.3.3.1$$

We can also bound the probability of error for classifying simple samples. The analysis of Section 2.2.4 carries over directly when \underline{X} is replaced by X and the moment generating function is recomputed as follows:

$$R_{ij}(X) = \prod_{m=1}^n R_{ij}(X_m) \quad 2.3.3.2$$

$$L_{ij}(X) = \sum_{m=1}^n L_{ij}(X_m)$$

This is a sum of independent, identically distributed random variables. Thus

$$\begin{aligned} E(\exp(uL_{ij}(X)) | X \in V_j) &= \prod_{m=1}^n E(\exp(uL_{ij}(X_m)) | X_m \in V_j) \quad 2.3.3.3 \\ &= (E(\exp(uL_{ij}(X)) | X \in V_j))^n \\ &= (\phi_{ij}(u))^n \end{aligned}$$

Equation 2.2.4.7 becomes

$$F_{ij}(T) = P(R_{ij}(X) > T | X \in V_j) \ll (\phi_{ij}(.5))^n / \sqrt{T} \quad 2.3.3.4$$

It is a property of moment generating functions that $\phi_{ij}(u) \leq \phi_{ij}(0) = 1$, so this bound is an exponentially decreasing function of n when $\phi_{ij}(.5) \neq 1$, or equivalently when the Bhattacharyya distance is non-zero. Thus the probability of error for the GML sample classification strategy is bounded by a sum of exponentially decreasing functions of the sample size.

To illustrate how powerful this bound can be we now consider a simple example. Suppose that the i th and j th spectral class densities are as depicted in Fig. 2.3.3.1. The mean vectors are equal, which results in a high degree of "overlap". Therefore the Bhattacharyya distance is only 0.11, and $\phi_{ij}(.5) = \sqrt{0.8} = 0.8944$. The actual conditional error rate, $F_{ij}(1)$, for no-memory classification ($n=1$), is 50%, which represents very poor performance. This implies that a "polling" classifier also has a 50% error rate regardless of the sample size. But Figure 2.3.3.2 shows how the GML performance improves as the sample size increases. For a sample of just 40 observations the error rate is practically insignificant. Although we probability cannot expect such dramatic performance in practice (due to the idealizations of our model), this still provides a strong motivation for our effort to apply sample classification to MSS data.

2.3.4 Maximum Likelihood vs. Minimum Distance

Let $X = (X_1, \dots, X_n)$ be a simple sample from a MVN

SEPARABILITY MEASURES

Divergence = 1.125

Transformed Divergence = 262

Bhattacharyya Distance = 0.11



Figure 2.3.3.1 Two Normal Densities With Low Separability

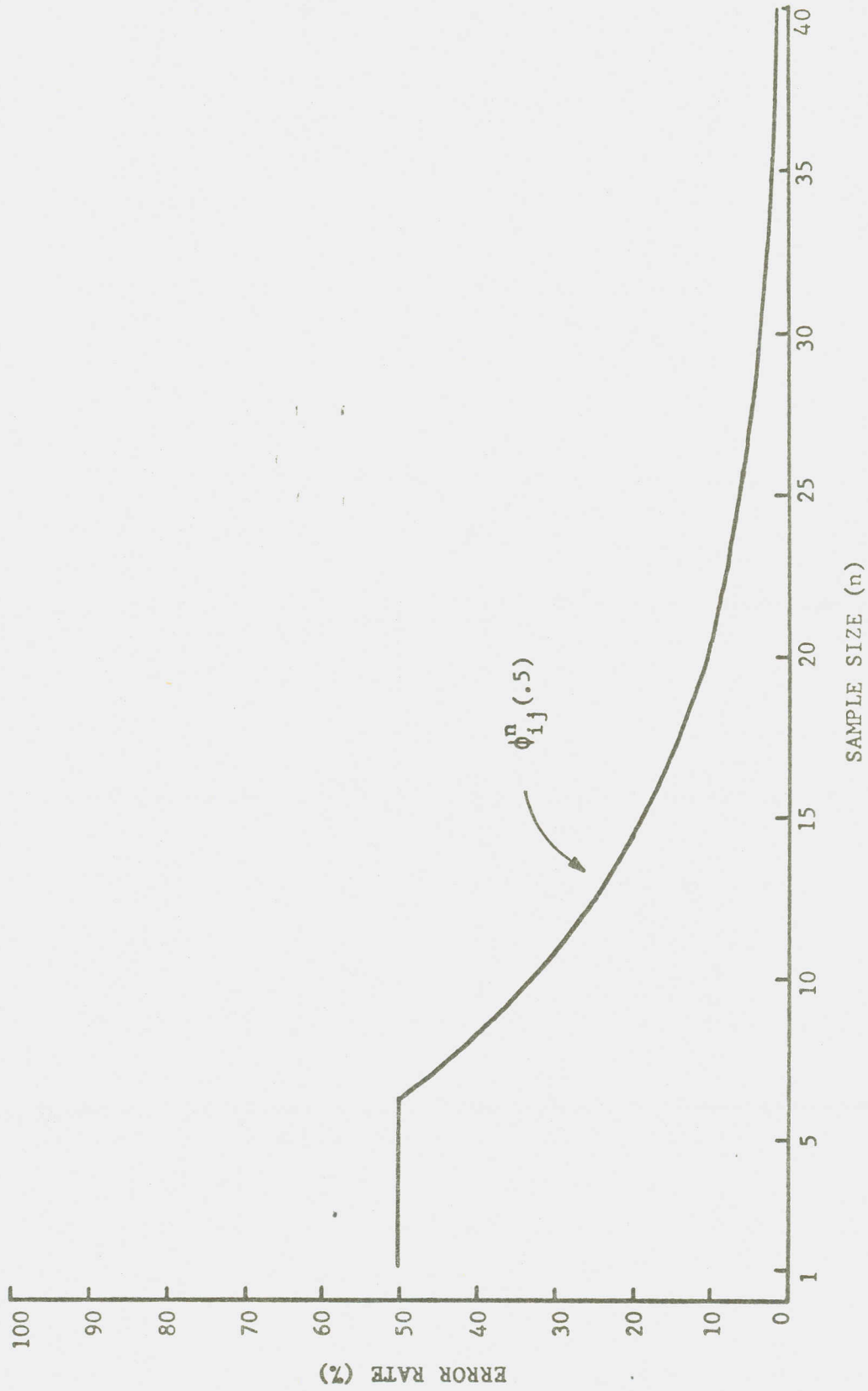


Figure 2.3.3.2 A Bound on Conditional Probability of Error for the Maximum Likelihood Strategy as a Function of the Sample Size

population, and define the statistics

$$\underline{S}_1 = \sum_{i=1}^n \underline{X}_i \quad 2.3.4.1$$

$$\underline{S}_2 = \sum_{i=1}^n \underline{X}_i \underline{X}_i'$$

The maximum likelihood estimates of the mean vector and covariance matrix are:

$$\underline{M} = \underline{S}_1/n \quad 2.3.4.2$$

$$\underline{C} = \frac{1}{n} \sum_{i=1}^n (\underline{X}_i - \underline{M})(\underline{X}_i - \underline{M})' = \underline{S}_2/n - \underline{M} \underline{M}'$$

The corresponding density estimate is given by equation 2.1.4.

Two popular distance measures are the Bhattacharyya distance and the divergence. If W_i is a class with density $N(\underline{M}_i, \underline{C}_i; \underline{x})$ then the Bhattacharyya distance between this and $N(\underline{M}, \underline{C}; \underline{x})$ is given by

$$B = .25 \left(\ln \frac{1}{|\underline{C}| |\underline{C}_i|} \frac{(|\underline{C} + \underline{C}_i|/2)^2}{|\underline{C}| |\underline{C}_i|} + (\underline{M} - \underline{M}_i)' (\underline{C} + \underline{C}_i)^{-1} (\underline{M} - \underline{M}_i) \right) \quad 2.3.4.3$$

and the divergence can be efficiently calculated from

$$D = .5 \operatorname{tr}((\underline{C}^{-1} + \underline{C}_i^{-1})(\underline{C} + \underline{C}_i + (\underline{M} - \underline{M}_i)(\underline{M} - \underline{M}_i)')) - 2q \quad 2.3.4.4$$

Computationally, D is faster than B, requiring about $2q(q+2)$ multiplications plus 1 matrix inversion per class for each $(\underline{S}_1, \underline{S}_2)$ pair classified. In addition to this, B requires a determinant and a logarithm. (This does not include quantities such as $|\underline{C}_i|$ which can be computed once and saved.) However B appears to provide an advantage in terms of classification accuracy, based on experimental evidence

2.3.4

[14]. And its direct relationship to the Chernoff bound gives B some intuitive appeal as well.

In order for the ML strategy to be computationally competitive with D and B, the likelihood function must be expressed in terms of \underline{S}_1 and \underline{S}_2 as follows:

$$p(X|W_i) = \prod_{j=1}^n N(\underline{M}_i, \underline{C}_i; \underline{X}_j) \quad 2.3.4.5$$

$$= (|2\pi\underline{C}_i|)^n \exp\left(-\frac{1}{2} \sum_{j=1}^n (\underline{X}_j - \underline{M}_i)' \underline{C}_i^{-1} (\underline{X}_j - \underline{M}_i)\right)$$

$$\ln p(X|W_i) = -0.5(n \ln|2\pi\underline{C}_i| + \sum_{j=1}^n (\underline{X}_j' \underline{C}_i^{-1} \underline{X}_j - 2\underline{M}_i' \underline{C}_i^{-1} \underline{X}_j + \underline{M}_i' \underline{C}_i^{-1} \underline{M}_i))$$

The quadratic term yields

$$\sum_{j=1}^n \underline{X}_j' \underline{C}_i^{-1} \underline{X}_j = \sum_{j=1}^n \text{tr}(\underline{C}_i^{-1} \underline{X}_j \underline{X}_j') = \text{tr}(\underline{C}_i^{-1} \sum_{j=1}^n \underline{X}_j \underline{X}_j')$$

so

$$\ln p(X|W_i) = -0.5 \text{tr}(\underline{C}_i^{-1} \underline{S}_2) + \underline{M}_i' \underline{C}_i^{-1} \underline{S}_1 - 0.5n(\underline{M}_i' \underline{C}_i^{-1} \underline{M}_i + \ln|2\pi\underline{C}_i|) \quad 2.3.4.6$$

which can be computed with just $.5q(q+5)$ multiplications, once the non-data-dependent quantities have been initialized. Thus the ML strategy requires only 25%-50% as many multiplications as D and no matrix inversions or determinants.

It is interesting to express equation 2.3.4.6 in terms of \underline{M} and \underline{C} . Substituting for \underline{S}_1 and \underline{S}_2 from equations 2.3.4.1 and simplifying, provides:

$$\frac{1}{n} \ln p(X|W_i) = -.5(\ln|2\pi\tilde{C}_i| + \text{tr}(\tilde{C}_i^{-1}(\tilde{C} + (\underline{M} - \underline{M}_i)(\underline{M} - \underline{M}_i)')))) \quad 2.3.4.7$$

which we shall denote by $L_i(\underline{M}, \tilde{C})$. By adopting the ML strategy, one is essentially using this quantity as a measure of the "similarity" of sample X to class W_i , just as B and D are used to measure their "dissimilarity". Therefore, $-L_i(\underline{M}, \tilde{C})$ can be interpreted as a measure of dissimilarity between the distributions $N(\underline{M}, \tilde{C}; \underline{x})$ and $N(\underline{M}_i, \tilde{C}_i; \underline{x})$. However it is not a distance measure in the sense of Wacker and Landgrebe [14], because it satisfies none of the three basic properties of distance measures; i.e. if $f(i, j)$ is a "distance" between distributions $N(\underline{M}_i, \tilde{C}_i; \underline{x})$ and $N(\underline{M}_j, \tilde{C}_j; \underline{x})$ then

1. $f(i, j) \geq 0$ 2.3.4.8
2. $f(i, i) = f(j, j) = 0$
3. $f(i, j) = f(j, i)$

One can force compliance with properties 1 and 2 by adding a bias term as follows:

$$\begin{aligned} d(i, j) &= -L_i(\underline{M}_j, \tilde{C}_j) + L_j(\underline{M}_j, \tilde{C}_j) \quad 2.3.4.9 \\ &= .5(\ln \frac{|\tilde{C}_i|}{|\tilde{C}_j|} + \text{tr}(\tilde{C}_i^{-1}(\tilde{C}_j + (\underline{M}_j - \underline{M}_i)(\underline{M}_j - \underline{M}_i)')) - q) \end{aligned}$$

which can be recognized as a form of one of the Kullback-Leibler numbers [19]. Since the bias term is independent of i (the class number), use of this criterion is still equivalent to the ML strategy, as long as $|\tilde{C}_j| > 0$. Also, the quantity $d(i, j) + d(j, i)$ is equivalent to the divergence, which satisfies all three distance measure properties.

The ML strategy has other compelling properties besides computational efficiency. On theoretical grounds, for the idealized conditions we have stated, it is the optimum strategy (for minimum error rate) when the a-priori class probabilities are equal. Also, the Chernoff bound for ML no-memory classification can be extended to provide an error bound for ML sample classification. Experimentally, under non-idealized conditions, the ML strategy does appear to be slightly better than MD (using B) on the whole, although it is not consistently better. The experimental results appear in Chapter 4. Wacker's experimental results for Kullback-Leibler numbers [14] also lend some support to this observation.

Another important property is that for small sample sizes the ML strategy does not break down as do the MD strategies. For a sample size of 1, it merely reduces to no-memory classification. Finally, the summation in equation 2.3.4.5 is distributed as chi-squared with nq degrees of freedom when W_i is the correct hypothesis for sample X . Therefore it can be used to construct a significance test of this hypothesis. This is useful for detecting samples that belong to none of the specified classes.

CHAPTER 3

IMAGE PARTITIONING

Once the partition is known, powerful techniques are available for classifying the individual objects. Thus, when the partition is unknown, an image partitioning algorithm offers an attractive alternative to no-memory classification. For reasons discussed in Section 1.3, the algorithms considered here are the type we refer to as conjunctive object-seeking. We have previously described this approach as a progressive merging of adjacent elements which are found to be similar according to some statistical criterion. Thus an algorithm consists of statistical tests applied in some logical sequence. The "logical sequence" is the subject of Section 3.1, and the rest of the chapter surveys some possible test criteria.

3.1 Partitioning Logic

In general it is not possible to design an error-free partitioning algorithm. First of all, there is a certain amount of ambiguity in defining the "true" partition due to real effects such as pixels that overlap physical boundaries or ambiguity in the physical boundaries themselves. Secondly, two main types of decision errors can occur,

leading to: (1) false boundaries, and (2) missed boundaries. Also the combined effect of these two errors can produce "approximate" boundaries, which is not actually a well-defined category due to the ambiguity of the true partition. Since object size and shape are not used as classification features, Type-1 errors are generally much less likely to lead to misclassifications than are Type-2 errors. This philosophy accounts for certain simplifications in the partitioning logic.

The basic approach that we have adopted (due to Rodd [11]) consists of two "levels" of tests. Initially the pixels are divided, by a (hypothetical) grid, into small groups of four (for example). At the first level of testing, each group becomes a unit called a "cell", provided that it satisfies a relatively mild criterion of homogeneity. Those groups that are rejected are assumed to overlap a boundary and their individual pixels are usually classified by the no-memory method. These groups are referred to as "singular" cells. At this level it is usually desirable to maintain a fairly low rejection rate to reflect the relatively high a-priori probability of a group being homogeneous. The goal at this level is essentially the same as the goal of the boundary seeking techniques discussed in Section 1.3, i.e. to detect as many pixels as possible that lie along boundaries without requiring that the ones detected form closed contours or even be connected.

SEPARABILITY MEASURES

Divergence = 1.125

Transformed Divergence = 262

Bhattacharyya Distance = 0.11

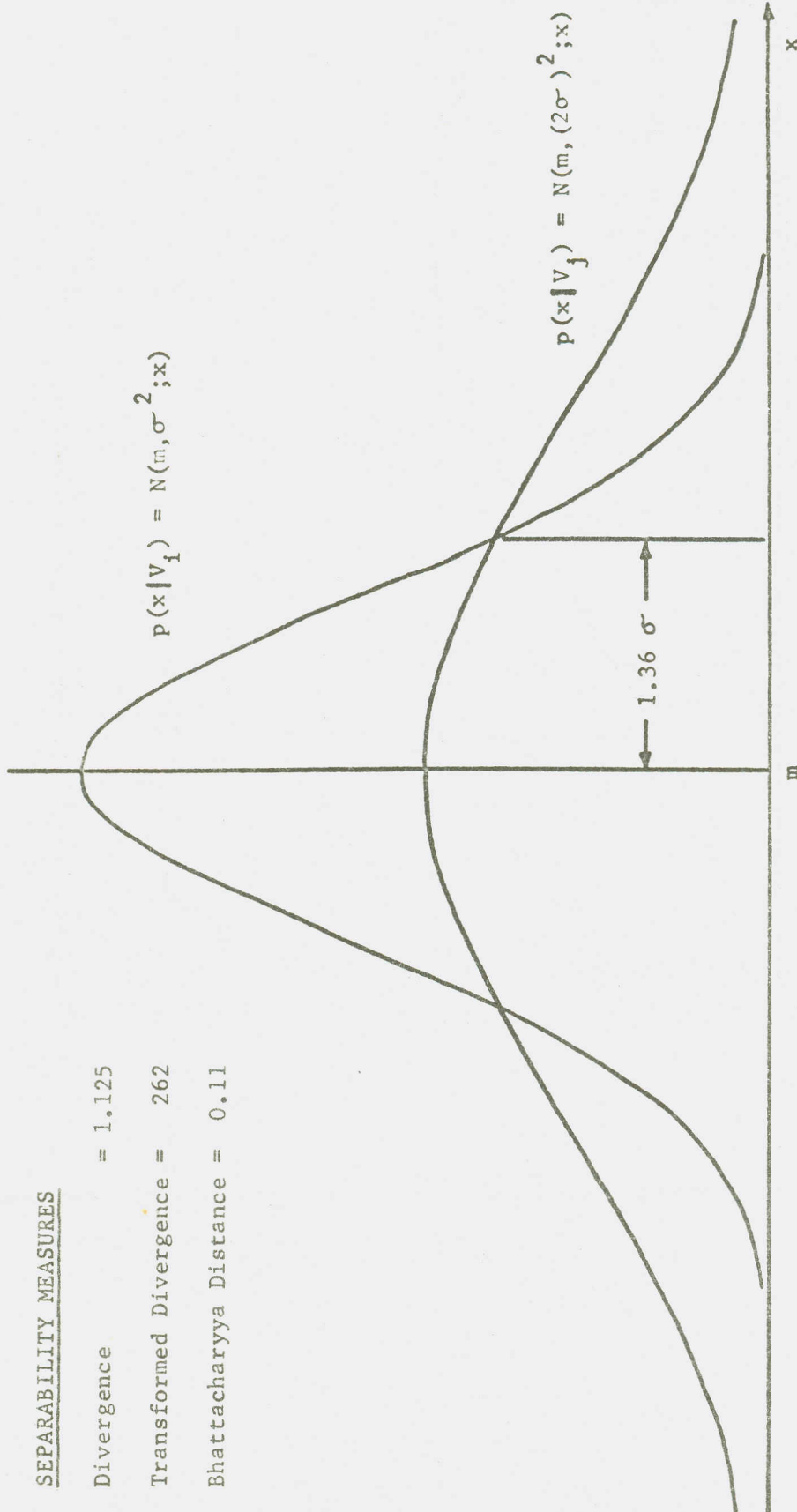


Figure 2.3.3.1 Two Normal Densities With Low Separability

It is a property of moment generating functions that $\phi_{ij}(u) \leq \phi_{ij}(0) = 1$, so this bound is an exponentially decreasing function of n when $\phi_{ij}(.5) \neq 1$, or equivalently when the Bhattacharyya distance is non-zero. Thus the probability of error for the GML sample classification strategy is bounded by a sum of exponentially decreasing functions of the sample size.

To illustrate how powerful this bound can be we now consider a simple example. Suppose that the i th and j th spectral class densities are as depicted in Fig. 2.3.3.1. The mean vectors are equal, which results in a high degree of "overlap". Therefore the Bhattacharyya distance is only 0.11, and $\phi_{ij}(.5) = \sqrt{0.8} = 0.8944$. The actual conditional error rate, $F_{ij}(1)$, for no-memory classification ($n=1$), is 50%, which represents very poor performance. This implies that a "polling" classifier also has a 50% error rate regardless of the sample size. But Figure 2.3.3.2 shows how the GML performance improves as the sample size increases. For a sample of just 40 observations the error rate is practically insignificant. Although we probability cannot expect such dramatic performance in practice (due to the idealizations of our model), this still provides a strong motivation for our effort to apply sample classification to MSS data.

2.3.4 Maximum Likelihood vs. Minimum Distance

Let $X = (\underline{X}_1, \dots, \underline{X}_n)$ be a simple sample from a MVN

$$E(N) = n \int P(X \in W(x) | X=x) p(X=x) dx \quad 2.3.2.4$$

This, of course, is maximized by the MAP strategy (2.3.2.2).

The ML strategy follows directly from the MAP strategy by dropping the a-priori probabilities. The result is

$$p(X=x|W(x)) = \max_i p(X=x|W_i) \quad 2.3.2.5$$

2.3.3 G.M.L. Sample Classification

We can obtain the GML decision rule by direct extension of 2.2.3.1. The result is

$$p(X=x|V(x)) = \max_i \max_j p(X=x|W_{ij}) = \max_i p(X=x|V_i) \quad 2.3.3.1$$

We can also bound the probability of error for classifying simple samples. The analysis of Section 2.2.4 carries over directly when \underline{X} is replaced by X and the moment generating function is recomputed as follows:

$$R_{ij}(X) = \prod_{m=1}^n R_{ij}(X_m) \quad 2.3.3.2$$

$$L_{ij}(X) = \sum_{m=1}^n L_{ij}(X_m)$$

This is a sum of independent, identically distributed random variables. Thus

$$\begin{aligned} E(\exp(uL_{ij}(X)) | X \in V_j) &= \prod_{m=1}^n E(\exp(uL_{ij}(X_m)) | X_m \in V_j) \quad 2.3.3.3 \\ &= (E(\exp(uL_{ij}(X)) | X \in V_j))^n \\ &= (\phi_{ij}(u))^n \end{aligned}$$

Equation 2.2.4.7 becomes

$$F_{ij}(T) = P(R_{ij}(X) > T | X \in V_j) \ll (\phi_{ij}(.5))^n / \sqrt{T} \quad 2.3.3.4$$

decision rule maximizes the average number of times that all the pixels in X are classified correctly. But performance is generally measured by just the average number of pixels in X that are classified correctly each time. We can show that the MAP decision rule maximizes this criterion also. Any decision rule that we adopt must assign a class to \underline{X}_i for any event $X=x$. We denote this mapping by $W_i(x)$. Let $Z(\cdot)$ be an indicator function, i.e. a zero-one random variable which assumes the value 1 if and only if the event specified in the argument actually occurs. The number of elements correctly classified in the sample is given by the random variable

$$N = \sum_{i=1}^n Z(\underline{X}_i \in W_i(X)) \quad 2.3.2.3$$

$$\begin{aligned} E(N) &= \sum_{i=1}^n E(Z(\underline{X}_i \in W_i(X))) = \sum_{i=1}^n P(\underline{X}_i \in W_i(X)) \\ &= \sum_{i=1}^n \int_{x \in R^{qn}} P(\underline{X}_i \in W_i(x) | X=x) p(X=x) dx \end{aligned}$$

The integration implied here is a qn -dimensional one. Note that the event $\underline{X}_i \in W_i(x)$ is equivalent to $X \in W_i(x)$, so all terms of this summation are identical, with the possible exception of the decision function. Thus the decision function which maximizes one term also maximizes the others. This confirms that the optimum decision rule assigns the same class to all the elements of the field. Denoting this decision function by $W(x)$, we have

2.3.2

of statistical information is also required (the a-priori class probabilities). We can obtain the MAP decision rule by direct extension of 2.2.1.3 if we consider X as a qn -dimensional random variable to be classified. Let x be the set of variates (x_1, \dots, x_n) and the event $X=x$ be defined as the joint event $X_i=x_i, i=1, \dots, n$. Then, under the hypothesis $X \in W_i$, the pdf of X is

$$\begin{aligned} p(X=x|W_i) &= \frac{1}{P(W_i)} \sum_j P(W_{ij}) p(X=x|W_{ij}) & 2.3.2.1 \\ &= \frac{1}{P(W_i)} \sum_j P(W_{ij}) \prod_{m=1}^n p(X_m=x_m|W_{ij}) \end{aligned}$$

The MAP decision rule can be stated as follows:

$$p(X=x|W(x))P(W(x)) = \max_i p(X=x|W_i)P(W_i) \quad 2.3.2.2$$

There is no minimum sample size required to implement this strategy. For $n=1$ it simply reduces to MAP no-memory classification (2.2.1.3).

Note that we have represented the joint pdf of a sample in terms of the marginal pdf of one pixel. When spatial correlation is present, this is no longer a fully adequate representation. But as in the case of MD classification, it still provides a useful statistic for distinguishing classes while avoiding the complexities of more rigorous representations.

So far we have tacitly assumed that the decision rule must assign the same class to all the pixels in the sample. With this type of strategy, either all the pixels are classified correctly or all are misclassified. Thus the MAP

2.3.1

simple samples, each class is completely characterized by a known q -dimensional pdf. Therefore, in MD classification, the n data vectors are used to estimate the pdf of the population, and the class is selected whose pdf is closest to this estimate as measured by some appropriately defined "distance measure" on the set of density functions. Ideally one would like to choose the density estimator and distance measure in some optimum manner, but in practice the best guidelines are provided by experimental investigations [14]. Note that a possible drawback of the MD strategy is that the sample size (n) must not be too small to obtain meaningful density estimates.

When spatial correlation is introduced into the model (Appendix A), each class is only partially characterized by a simple q -dimensional pdf. Although perhaps not as effective as a higher dimensional pdf would be, it is still a reasonable and valid characteristic for distinguishing between classes. In fact if the spatial correlation is class-invariant (such as that induced by the scanner), the q -dimensional pdf might be just as effective as the higher dimensional one.

2.3.2 M.A.P. and M.L. Sample Classification

In contrast to the MD strategy, the MAP strategy is a completely non-structured approach. The decision rule is determined solely by the criterion of minimum error rate with no a-priori restrictions. Of course a greater degree

Combining equations 2.2.4.1, 2.2.4.2, and 2.2.4.7 gives an expression for a bound on $P(E)$:

$$P(E) \leq \sum_j P(V_j) \sum_{\substack{i \\ i \neq j}} \phi_{ij} \quad (2.2.4.8)$$

By dropping the terms for which V_i and V_j are in the same class, this becomes a bound on the total probability of error.

2.3 Sample Classification

For the purposes of this section we can assume that the partition of the scene is known and we simply want to classify the objects. (In Chapter 3 we discuss conjunctive partitioning algorithms for actually estimating the partition.) We shall treat each object separately, thus ignoring any contextual information resulting from spatial relationships of objects. So we observe a set (sample) of q -dimensional random variables, $X = (X_1, \dots, X_n)$, from a common population, and our goal is to classify them.

2.3.1 Minimum Distance (MD) Strategy

- A Structured Approach to Classification

A structured approach is one in which the basic form of the processor is simply assumed, perhaps leaving certain parameters or options to the discretion of the user. A reasonable procedure is to choose some characteristic that differs from class to class, measure it for the sample to be classified, and select the class whose characteristic most closely matches this observation. Under our assumption of

2.2.4

$$R_{ij}(\underline{X}) = \frac{p(\underline{X}|V_i)}{p(\underline{X}|V_j)} = \text{likelihood ratio} \quad 2.2.4.3$$

$$L_{ij}(\underline{X}) = \ln R_{ij}(\underline{X})$$

Then $P(E_{ij}|V_j) = F_{ij}(1)$. Unfortunately the conditional distribution functions of $R_{ij}(\underline{X})$ and $L_{ij}(\underline{X})$ are not usually explicitly available. But, if we can find the moment generating function, $\phi_{ij}(u)$, corresponding to the conditional distribution of $L_{ij}(\underline{X})$ given $\underline{X} \in V_j$, then we can bound $F_{ij}(T)$ as follows:

$$F_{ij}(T) \leq T^{-u} \phi_{ij}(u), \quad 0 \leq u \quad 2.2.4.4$$

Furthermore

$$F_{ij}(T) \leq T^{u-1} \phi_{ij}(u), \quad u \leq 1$$

This is known as the Chernoff bound [18].

By definition:

$$\phi_{ij}(u) = E(\exp(uL_{ij}(\underline{X})) | \underline{X} \in V_j) \quad 2.2.4.5$$

When the subclasses are MVN, the expectation can be explicitly evaluated [18]. The result is:

$$\phi_{ij}(u) =$$

$$\sqrt{\frac{|C_j|^u |C_i|^{1-u}}{|uC_j + (1-u)C_i|} \exp(-u(1-u)(\underline{M}_i - \underline{M}_j)'(uC_j + (1-u)C_i)^{-1}(\underline{M}_i - \underline{M}_j))}$$

2.2.4.6

Substituting into 2.2.4.4 provides the desired bound. In particular, for $u = .5$ we have:

$$F_{ij}(T) \leq \phi_{ij}(.5) / \sqrt{T} \quad 2.2.4.7$$

We note in passing that $-\ln \phi_{ij}(.5)$ is simply the Bhattacharyya "distance" between subclasses V_i and V_j .

At the second level, an individual cell is compared to an adjacent "field", which is simply a group of one or more connected cells that have previously been merged. If the two samples appear statistically similar by some appropriate criterion, then they too are merged. Otherwise the cell is compared to another adjacent field or becomes a new field itself. By successively "annexing" adjacent cells, each field expands until it reaches its natural boundaries, where the rejection rate abruptly increases, thereby halting further expansion. The field is then classified by a sample classifier, and the classification is assigned to all its pixels.

This approach has the important advantage that it can be implemented "sequentially"; i.e. raw data need be accessed only once and in the same order that it is stored on tape. This is important for practical, rather than theoretical, considerations. The flow chart in Figure 3.1.1 indicates how it can be done. In this chart, the top of the scene is referred to as north, and the general processing sequence is from north to south.

A possible drawback of the approach described above is that in certain hypothetical situations, Type-1 errors are a certainty. For example, a U-shaped object would develop as two separate fields which expand southward and eventually meet at the base of the U. But since no provision is made for merging such fields, a false boundary between them will

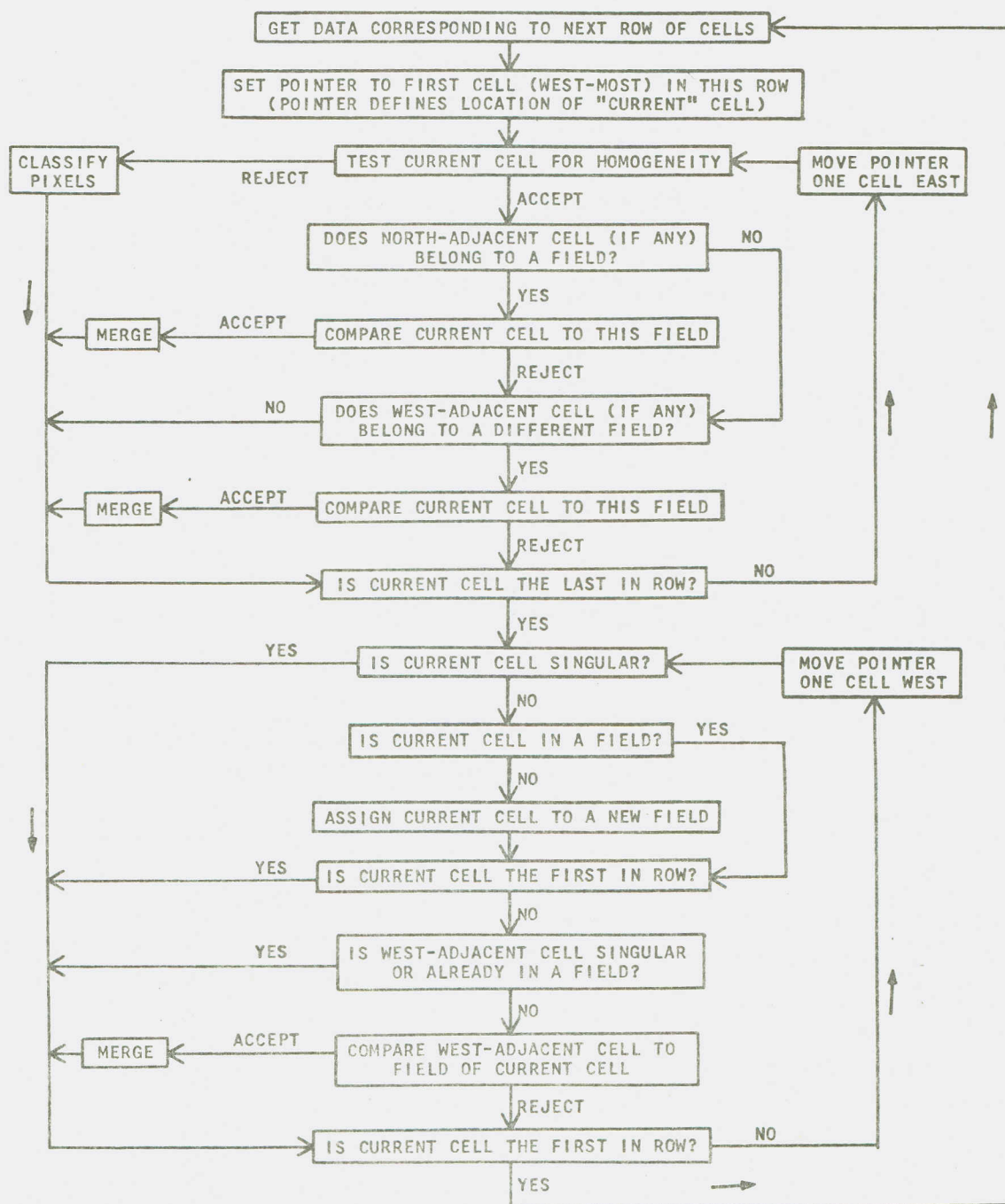


Figure 3.1.1 Basic Flow Chart for a Two-Level, Conjunctive Partitioning Algorithm

3.1

result. Such a provision can of course be made, but only at the cost of additional complexity. This does not appear to be warranted by the relatively harmless nature of an isolated Type-1 error. Thus the false boundary actually results from a design simplification rather than a true decision error.

Many modifications (both large and small) to the basic flow chart are, of course, possible. For example, the Level-1 test can be removed from the loop if performed in advance and intermediate results saved on tape. (This is particularly useful in a research environment.) Another modification is described in Section 3.3.3. It involves comparing a cell to as many as three different fields at once, instead of one-at-a-time.

3.2 Unsupervised Mode

In order to implement the sequential approach we must specify two test criteria corresponding to the two levels. In this section we consider ways to do this "unsupervised"; i.e. the test criteria are independent of specific knowledge of the spectral class distributions. Note that our usage of this term is analogous, but not identical, to the conventional usage.

3.2.1 Unsupervised Annexation

Let $X = (X_1, \dots, X_n)$ represent the pixels in a group of one or more cells which have been merged by successive annexations. Let $Y = (Y_1, \dots, Y_m)$ represent the pixels in an

adjacent, non-singular cell. Since both X and Y will have satisfied certain criteria of homogeneity, we assume that each is a sample from a MVN population. Let f and g represent the corresponding density functions. It is desired to test the (null) hypothesis that $f = g$. This is a composite hypothesis, since it does not specify f and g . The "likelihood ratio procedure" [20] provides an effective statistic for testing this hypothesis. Van Trees [21] refers to it as the "generalized likelihood ratio". Let

$$H_0(x,y) = \{p(x,y|f,g): g=f, f \in \Omega\}$$

$$H_1(x,y) = \{p(x,y|f,g): f \in \Omega, g \in \Omega, g \neq f\}$$

where $p(x,y|f,g)$ is the conditional joint density of X and Y evaluated at $x \in \mathbb{R}^{nq}$ and $y \in \mathbb{R}^{mq}$, and Ω is a set of MVN density functions. The assumption of class-conditional independence enables us to express the joint density of pixels as the product of their marginal densities. Thus:

$$p(x,y|f,g) = p(x|f) p(y|g) = \left(\prod_{i=1}^n f(x_i) \right) \left(\prod_{i=1}^m g(y_i) \right)$$

The generalized likelihood ratio is defined by:

$$\Lambda = \frac{\sup H_0(X,Y)}{\sup H_1(X,Y)} = \frac{\max_{f \in \Omega} p(X|f) p(Y|f)}{\max_{\substack{f \in \Omega \\ g \in \Omega \\ g \neq f}} p(X|f) p(Y|g)} \quad 3.2.1.1$$

For an "unsupervised" approach to partitioning we take Ω to be the following set of functions of $\underline{x} \in \mathbb{R}^q$:

$$\Omega = \left\{ N(\underline{M}, \underline{C}; \underline{x}): \underline{M} \in \mathbb{R}^q, \underline{C} = \text{symmetric, positive-definite} \right\}$$

Since for any $f \in \Omega$ there exists a $g \in \Omega$ that is arbitrarily

3.2.1

close to f , the condition " $g \neq f$ " can be dropped from the denominator of 3.2.1.1. Therefore:

$$\begin{aligned} \Lambda &= \frac{\max N(\underline{M}, \underline{C}; X) N(\underline{M}, \underline{C}; Y)}{\max N(\underline{M}_X, \underline{C}_X; X) N(\underline{M}_Y, \underline{C}_Y; Y)} \leq 1 \\ &= \frac{\max N(\underline{M}_X, \underline{C}; X) N(\underline{M}_Y, \underline{C}; Y)}{\max N(\underline{M}_X, \underline{C}_X; X) N(\underline{M}_Y, \underline{C}_Y; Y)} \cdot \frac{\max N(\underline{M}, \underline{C}; X) N(\underline{M}, \underline{C}; Y)}{\max N(\underline{M}_X, \underline{C}; X) N(\underline{M}_Y, \underline{C}; Y)} \\ &= \Lambda_2 \cdot \Lambda_1 \end{aligned}$$

where

$$N(\underline{M}, \underline{C}; X) = \prod_{i=1}^n N(\underline{M}, \underline{C}; X_i)$$

$$N(\underline{M}, \underline{C}; Y) = \prod_{i=1}^m N(\underline{M}, \underline{C}; Y_i)$$

and in each case the maximization is with respect to the mean vectors and covariance matrices.

Anderson [22] shows that:

$$\Lambda_1 = (|A|/|B|)^{N/2} \quad 3.2.1.2$$

$$\Lambda_2 = (|A_X/n|^n |A_Y/m|^m / |A/N|^N)^{.5} \quad 3.2.1.3$$

where

$$N = n + m$$

$$\bar{X} = \sum_{i=1}^n X_i/n \quad \bar{Y} = \sum_{i=1}^m Y_i/m$$

$$A_X = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})' \quad A_Y = \sum_{i=1}^m (Y_i - \bar{Y})(Y_i - \bar{Y})'$$

(In order to assure non-singular matrices with probability 1, we need $n > q < m$.) [22]

$$\Lambda = A_X + A_Y$$

$$\underline{M} = (n\bar{X} + m\bar{Y})/N$$

$$\underline{B}_x = \sum_{i=1}^n (X_i - \underline{M})(X_i - \underline{M})' = A_x + n(\bar{X} - \underline{M})(\bar{X} - \underline{M})'$$

$$\underline{B}_y = \sum_{i=1}^m (Y_i - \underline{M})(Y_i - \underline{M})' = A_y + m(\bar{Y} - \underline{M})(\bar{Y} - \underline{M})'$$

$$\underline{B} = \underline{B}_x + \underline{B}_y = \underline{A} + \frac{mn}{N}(\bar{X} - \bar{Y})(\bar{X} - \bar{Y})'$$

Anderson also suggests the following modification:

$$\lambda = \lambda_1 \cdot \lambda_2$$

where λ_1 and λ_2 are obtained from Λ_1 and Λ_2 by replacing the number of pixels in each sample by the number of degrees of freedom; i.e. replace n by $n-1$, m by $m-1$, and N by $N-2$ in formulas 3.2.1.2 and 3.2.1.3. In either case, the statistics are invariant with respect to a linear transformation on the data vectors. It follows that their distributions under the null hypothesis are independent of the actual MVN population from which the samples are drawn.

The test procedure is to compare λ with some decision threshold $T < 1$, which depends in general upon n and m . The hypothesis is accepted if $\lambda > T$ and rejected if $\lambda < T$. In the unsupervised mode, T is determined by specifying the desired "size" (significance level) of the test, because the power of the test is indeterminate. In order to do this, however, the distribution of λ must be tabulated. Under the null hypothesis λ_1 and λ_2 are independently distributed [22], so we can simplify the distribution theory and accomplish the same objective by the following procedure. Pick significance levels s_1 and s_2 such that

3.2.1

$s = 1 - (1 - s_1)(1 - s_2)$, where s is the desired significance level. Test λ_i using a threshold T_i such that $P(\lambda_i < T_i | H) = s_i$, $i=1,2$, where H denotes the event that the null hypothesis is true. The null hypothesis is rejected if either test produces a rejection. Thus the effective size of the test is given by

$$\begin{aligned} 1 - P(\lambda_1 \geq T_1, \lambda_2 \geq T_2 | H) &= 1 - P(\lambda_1 \geq T_1 | H) P(\lambda_2 \geq T_2 | H) \\ &= 1 - (1 - s_1)(1 - s_2) = s \end{aligned}$$

as desired. This procedure gives us complete freedom to pick the ratio $s_2/s_1 \geq 0$. As this ratio increases, the power of the test against the alternative " $\underline{\mu}_x = \underline{\mu}_y, \underline{C}_x \neq \underline{C}_y$ " increases, and the power against the alternative " $\underline{\mu}_x \neq \underline{\mu}_y, \underline{C}_x = \underline{C}_y$ " decreases.

We now review the distribution theory that is needed to implement these tests. There is a transformation of λ_1 which, given H , has an F-distribution with q and $(N-q-1)$ degrees of freedom [23]. It is given by

$$F_1 = \frac{(|B| - 1)(N - q - 1)}{|A|} \quad 3.2.1.4$$

Thus the test for a significant difference between the mean vectors can be implemented by computing F_1 and comparing it to a threshold t_1 determined by the relation $P(F_1 > t_1 | H) = s_1$. Alternatively, it can be shown [18] that

$$\frac{(|B| - 1)(N - 2)}{|A|} = (N - 2) \frac{nm}{N} (\bar{X} - \bar{Y})' \tilde{A}^{-1} (\bar{X} - \bar{Y}) = T^2 \quad 3.2.1.5$$

which (given H) has a T^2 distribution with $N-2$ degrees of freedom [22]. T^2 is Hotelling's generalization of the Student-t statistic, which is commonly used to test the

hypothesis that the means of two univariate normal distributions are equal given that the variances are equal.

The following transformation of λ_2 has the F-distribution (given H) [23]:

$$G = -2 \ln \lambda_2 \quad 3.2.1.6$$

$$g = \frac{2a^2 + 3a - 1}{6(a+1)} \left\{ \frac{1}{n-1} + \frac{1}{m-1} - \frac{1}{N-2} \right\}$$

$$u = \frac{(a-1)(a+2)}{6} \left\{ \frac{1}{(n-1)^2} + \frac{1}{(m-1)^2} - \frac{1}{(N-2)^2} \right\} - g^2$$

$$v = \frac{a(a+1)}{2}$$

$$w = \frac{v+2}{|u|}$$

$$Z = \begin{cases} \frac{(1-g-v/w)G}{v}, & u > 0 \\ \frac{(1-g+2/w)G}{w}, & u < 0 \end{cases}$$

$$F_2 = \begin{cases} Z, & u > 0 \\ \frac{wZ}{v(1-Z)}, & u < 0 \end{cases}$$

F_2 has an F-distribution with v and w degrees of freedom. Thus the test is implemented by computing F_2 and comparing it to a threshold t_2 determined by the relation

$$P(F_2 > t_2 | H) = s_2.$$

Due to the complexity of the test for different covariance matrices, it may be desirable to rely only on the difference in mean vectors (assuming that a difference exists). A common approach is to simply assume that all covariance matrices are equal, thereby eliminating the need to test the statistic F_2 ; i.e. let $s_2=0$. The test of F_1 is

probably fairly robust under departures from this assumption anyway [23]. An alternative approach, which does not require this assumption, is to use the Behrens-Fischer statistic defined as follows [22]:

$$Z_i = Y_i - X_i(m/n)^{1/2}, \quad i=1,2,\dots,m \leq n. \quad 3.2.1.7$$

$$\bar{Z} = \sum_{i=1}^m Z_i / m$$

$$D = \sum_{i=1}^m (Z_i - \bar{Z})(Z_i - \bar{Z})' / (m-1)$$

$$T^2 = m(\bar{X} - \bar{Y})' D^{-1} (\bar{X} - \bar{Y})$$

This has a T^2 -distribution with $m-1$ degrees of freedom, or equivalently the statistic

$$F_3 = \frac{T^2(m-q)}{(m-1)q} \quad 3.2.1.8$$

has an F-distribution with q and $m-q$ degrees of freedom under the hypothesis that $\underline{M}_x = \underline{M}_y$. Thus the test is implemented by comparing F_3 to a threshold t_3 which satisfies the relation $P(F_3 > t_3 | \underline{M}_x = \underline{M}_y) = s$.

These multivariate tests all have the same weakness as MD classification, namely the problem of estimating a MVN density from a relatively small sample (sometimes known as the "dimensionality" problem). This led to the constraint $n > q$, a condition which is often not met. Even when the condition is met, poor estimates can result, leading to decision errors. One approach to this problem is to reduce q by deleting features. It is well-known, for example, that a subset of features used to train a classifier from small

3.2.1

training samples can sometimes produce better classification results than the full set [24]. With this approach, however, one is faced with the problem of choosing the subset.

Another approach is to base the decision on the q , univariate, marginal distributions; i.e. simply consider the data in one spectral channel at a time. This has been termed a "multiple univariate" (MUV) approach. In each channel we test the univariate hypothesis that the means and variances of the two samples are equal. Since the boundaries may be strong in some spectral channels and weak in others, we accept the null hypothesis only if the univariate hypothesis is accepted in all q channels. Besides avoiding the dimensionality problem, the MUV procedure requires less computation and simpler distribution theory. However, it must be pointed out that in situations where class separability is primarily a multivariate effect, the MV procedure may be more advantageous.

In order to obtain the univariate tests we can follow the same development that led to the multivariate tests except that $q=1$ and A_x , A_y , A , and B are just one-dimensional matrices (scalars). Thus equations 3.2.1.4 and 3.2.1.5 simplify as follows:

$$F_1 = T^2 = \frac{(N-2)nm}{N} \frac{(\bar{X}-\bar{Y})^2}{A} \quad 3.2.1.9$$

This has an F-distribution with 1 and $(N-2)$ degrees of freedom, under the null hypothesis. Equivalently we can say

3.2.1

that the statistic

$$\sqrt{\frac{(N-2)nm}{NA}} (\bar{X}-\bar{Y}) \quad 3.2.1.10$$

has a Student-t distribution with N-2 degrees of freedom.

The statistic λ_2 simplifies to

$$\lambda_2 = \sqrt{(1+K)^{N-2} f(r)} \quad 3.2.1.11$$

where

$$K = (m-1)/(n-1)$$

$$r = K A_x/A_y$$

$$f(r) = \frac{r^{n-1}}{(r+K)^{N-2}}$$

The statistic r has an F-distribution with $n-1$ and $m-1$ degrees of freedom, and it is independent of F_1 under the null hypothesis [22]. But since $f(\cdot)$ is not monotonic, two thresholds must be determined in order to implement a test on this statistic. For a significance level s_2 , the thresholds T' and T'' must satisfy

$$P(r < T' \text{ or } r > T'' | H) = s_2 \quad 3.2.1.12$$

$$f(T') = f(T'')$$

Alternatively one could resort to the transformation in 3.2.1.6.

3.2.2 Unsupervised Cell Selection

"Cell selection" refers to the Level-1 test, which is used to detect cells that apparently overlap boundaries. Such cells frequently exhibit abnormally large sample variances. Thus a possible criterion for a cell is to

3.2.2

require that the sample variance in each spectral channel fall below some reasonable threshold. A similar approach is to form the ratio of the square root of the sample variance to the sample mean and compare it to a threshold (which we shall call "c"). This criterion has the advantage of being independent of the scale of the data.

A possible multivariate approach is to place an upper limit on the sample generalized variance, $|A_y/m|$, that any cell (Y) can have. This is equivalent to placing a lower limit on the value of the statistic $\max_{M,C} N(\underline{M}, \underline{C}; Y)$. But again we mention that the dimensionality problem seriously weakens the MV approach. It can cause very poor estimation of the generalized variance and increase the chance of a decision error.

3.3 Supervised Mode

In this section we develop a way to "supervise" the sequential partitioning process, using the known spectral class distributions. Our approach is based on the same composite hypothesis testing procedure as the unsupervised approach. The effect of the spectral class distributions is to greatly simplify each hypothesis, but paradoxically the resultant test criterion is much more complicated. Fortunately, much of the computation can be done "sequentially", i.e. relying on previous saved results.

3.3.1 Supervised Annexation

Let X and Y be samples from a field and an adjacent

3.3.1

cell as in Section 3.2.1. We follow the same development as in that section, except that for a supervised approach to partitioning we take Ω as:

$$\Omega = \{p(\underline{x}|V_i): i=1,2,\dots,k\}$$

where k is the number of spectral classes. Note that this is a considerably more restrictive condition than before. The corresponding generalized likelihood ratio statistic is:

$$\Lambda = \frac{\max_i (p(X|V_i) p(Y|V_i))}{\max_{\substack{i,j \\ j \neq i}} (p(X|V_i) p(Y|V_j))} \quad 3.3.1.1$$

Note that this is a multivariate statistic without the constraint $m > q$ that was necessary in the unsupervised mode. However the maxima in formula 3.3.1.1 cannot be expressed in a simple analytic form as in 3.2.1.1. They can only be obtained by exhaustive search. Furthermore, the distribution of 3.3.1.1 is unknown under either hypothesis, because it depends on the true classes of X and Y . But in return we gain a statistic which should be more "sensitive" to the presence or absence of a boundary. This should produce better performance and make the specification of a decision threshold less critical. In fact, the experimental results in Chapter 4 indicate that the threshold need not be a function of n , the current size of sample X , in order to obtain good results. Furthermore, the results tend to be fairly stable over several orders of magnitude of threshold variation. Thus we will find it convenient to represent the decision threshold as

3.3.1

$$T = 10^{-t}, \quad t \geq 0$$

3.3.1.2

Unlike the unsupervised approach, a constant decision threshold of $T=1$ does not imply that the null hypothesis is always rejected. But it does lead to the same final result when the GML strategy is used to classify the objects. This is because Λ can exceed 1 only if X and Y would be classified the same by the GML rule anyway. Consequently, the only practical values of T are those between 0 and 1. Lacking any distribution theory to provide guidance in choosing a suitable threshold in this range, we shall rely instead on an empirical approach.

Calculation of the generalized likelihood ratio criterion can be greatly simplified by the following measures:

1. Change the denominator of Λ to $\max_{i,j} (p(X|V_i)p(Y|V_j))$. The only effect of this change is to cause the value of Λ to saturate at an upper limit of 1. It does not affect the value of Λ when $\Lambda < 1$. Since T is always less than 1, the change cannot affect any decisions. The simplification that it affords is that Λ can now be written as follows:

$$\Lambda = \frac{\max_i (p(X|V_i)p(Y|V_i))}{(\max_i p(X|V_i))(\max_j p(Y|V_j))}$$

which is simpler to compute.

2. Compare $\ln(\Lambda)$ to $\ln(T)$ instead of Λ to T .

3.3.1

$$\ln \Lambda = \max_i (\ln p(X|V_i) + \ln p(Y|V_i)) - \max_i (\ln p(X|V_i)) - \max_i (\ln p(Y|V_i)) \quad 3.3.1.3$$

Besides converting multiplications into (faster) additions, the quantity $\ln p(Y|V_i)$ can be efficiently computed by formula 2.3.4.6. The quantity $\ln p(X|V_i)$ can be obtained by other means, as we shall see.

3. Assume for the moment that the log-likelihood function of X is already available in a storage array G (say); i.e.

$$G(i) = \ln p(X|V_i), \quad i=1,2,\dots,k$$

and J is an integer such that $G(J) = \max_i G(i)$. Then compute the first term of $\ln(\Lambda)$ using an intermediate storage array g (say), as follows:

$$g(i) = G(i) + \ln p(Y|V_i), \quad i=1,2,\dots,k.$$

and j is an integer such that $g(j) = \max_i g(i)$. So:

$$\ln(\Lambda) = g(j) - G(J) - \max_i \ln p(Y|V_i)$$

If the indication is to merge Y and X , then the new log-likelihood function of the field is just the sum of the log-likelihood functions of X and Y . Therefore we can simply update G and J as follows:

$$G(i) = g(i), \quad i=1,2,\dots,k$$

$$J = j$$

and the preliminary assumption is justified. Thus we avoid using formula 2.3.4.6 to compute $\ln p(X|V_i)$.

When the point is reached that the field X stops expanding, it must be classified. This would normally require the sample mean vector and autocorrelation (or

covariance) matrix, which would have to be continually updated as cells are added to the field. Recall however that the GML strategy is

$$p(X|V(X)) = \max_i p(X|V_i) = \max_i \exp(G(i)) = \exp(G(J))$$

Therefore $V(X) = V_J$, so no additional updating or computation is required to classify the field if the GML strategy is used.

3.3.2 Supervised Cell Selection

A useful statistic for cell selection is

$$Q_j(Y) = \text{tr}(\underline{C}_j^{-1} \sum_{i=1}^m Y_i Y_i') - 2 \underline{M}_j' \underline{C}_j^{-1} \sum_{i=1}^m Y_i + m \underline{M}_j' \underline{C}_j^{-1} \underline{M}_j$$

where j is such that

$$\ln p(Y|V_j) = \max_i \ln p(Y|V_i) = \max_i \{-.5(m \cdot \ln |2\pi \underline{C}_i| + Q_i(Y))\}$$

The decision rule is to accept the hypothesis that Y is homogeneous if $Q_j(Y) < c$, where c is a prespecified threshold. Otherwise the hypothesis is rejected. This criterion has the particular advantage that it tends to reject not only inhomogeneous cells, but "unrecognizable" cells as well. (Unrecognizable cells are those which represent spectral classes that the classifier has not been trained to recognize.) Another advantage of this criterion is that its use of the log-likelihood function makes it especially compatible with the supervised annexation criterion and the GML sample classifier.

As a final note, the distribution function $P(Q_j(Y) > c | Y \in V_j)$ is chi-squared with mq degrees of freedom.

3.3.2

This can be used to provide initial guidance in choosing c .

3.3.3 Alternative Partitioning Logic

The logic of Figure 3.1.1.1 compares a cell to the north, west, and east-adjacent fields (if necessary) seeking a "match". If a match is found, the merge takes place immediately without regard to whether it is the "best" match or not. Another approach that is used is to compare the cell to all three fields at once (if that many distinct adjacent fields exist) and attempt to determine the best match. In the supervised mode a match is determined by comparing the likelihood ratio to a fixed threshold, so a reasonable definition for the best one is the field for which this ratio is largest. Normally the east-adjacent field would not exist at the time the other two comparisons are made, so its likelihood ratio is supplied by "looking ahead"; i.e. the east-adjacent cell is compared to its north-adjacent field and if they match, the current cell is compared to their union to obtain the likelihood ratio.

This approach has not been used in the unsupervised mode, mainly because of the difficulty of determining the best match. A logical approach would be to choose the field for which the null hypothesis is "least rejectable"; e.g. choose the field which maximizes the minimum significance level at which the null hypothesis would not be rejected. In other words, if λ_{ij} represents the observed value of λ_i for the j th field, then the field is chosen for which

3.3.3

$\min_{i=1,2} P(\lambda_i < \lambda_{ij} | H)$ is maximum. The difficulty in actually

doing this is that the complete distribution function of λ_i would be required. Generally it is available for only a few isolated significance levels.

CHAPTER 4

CLASSIFICATION RESULTS

Experimental results obtained in the investigation of multispectral image partitioning techniques are presented in this chapter. Several different data sets of markedly different characteristics are classified by these techniques. In many respects they represent a cross-section of MSS data. Both low altitude aircraft data and 930 km high LANDSAT-1 data are included. The ground resolution varies from 4.6m to 80m, and the size of physical objects varies from just a few pixels to thousands of pixels. Data representation is 8 bits for aircraft data and 6 bits for LANDSAT. Spectral resolution varies from 0.02 to 2.40 micrometers, while the number of spectral channels available varies from 12 to 4. The actual number of channels used for analysis varies between 3 and 6. The number of spectral classes representing ground cover types varies from 5 to 17, and the number of informational classes varies from 5 to 11.

The results are grouped by data set rather than by analysis technique to facilitate the comparison of different analyses of a given data set. In order to provide a quantitative measure of comparison, only data sets are used for

which a substantial number of "test areas" are available. By comparing the results of a given analysis on a point-by-point basis with the desired result in each test area, one can obtain an estimate of the accuracy (or inaccuracy) of the analysis. The larger and more numerous these test areas, the better this estimate will be. Thus one analysis technique is regarded as being better than another if it tends to achieve fewer misclassifications in the test areas.

Relative error rate is an important measure of a classification scheme, but it is not the only consideration. Obviously speed is a desirable attribute. Although CPU times are compared in this chapter, it is important to remember that efficient coding has a lot to do with speed, and no claim is made that the research programs used here are optimized. A less tangible consideration is the amount of effort and experience required to use a particular analysis scheme. The schemes considered in this investigation were designed with simplicity in mind, requiring a minimum of user input. The results in this chapter will help to assess the degree of experience needed to provide this input, and they provide a data base of experience from which to draw.

4.1 Analysis Schemes

Within the framework of Chapter 3, an analysis scheme is specified by choosing:

1. A Level-1 option and associated parameters (threshold and cell size)

2. A Level-2 option and associated parameter(s) (significance levels or threshold)
3. A sample classifier option.

It would be a hopeless, and probably pointless, task to try to investigate all the possible combinations of these three options. Instead the less logical combinations were arbitrarily eliminated in order to concentrate more effort on evaluation of the remaining ones. Consequently only wholly unsupervised and wholly supervised methods are used for the partitioning phase of processing. (No "hybrid" combinations of Level-1 and Level-2 options are considered.) Thus partitioning is done in either the unsupervised or supervised "mode". Furthermore, in the unsupervised mode, no hybrids of MV and MUV tests are used to test first and second order statistics at Level-2. At Level-1, only the MUV ratio test (described in Section 3.2.2) is used in the unsupervised mode. Although the Behrens-Fischer test requires only one significance level to be specified by the user, it tests only first order statistics, provides unattractively few degrees of freedom, and requires a substantial amount of computation. Consequently it was eliminated as an option. Due to the advantages (enumerated in Chapter 2) of maximum likelihood sample classification over minimum distance classifiers, it was the logical choice for the classifier option. The results of this chapter are based on the ML strategy. When subclasses are necessary the generalized ML strategy is actually used, although it is

referred to simply as a ML classifier.

After all the above simplifications, we basically are left with four schemes to evaluate and compare:

1. unsupervised MUV partitioning and ML sample classification
2. unsupervised MV partitioning and ML sample classification
3. supervised partitioning and ML sample classification
4. conventional ML no-memory classification.

Furthermore the cell size for the first three schemes was eliminated as a variable by fixing it at a constant 2x2 pixels, which is the minimum size that can be used in the unsupervised mode of partitioning. This choice appears to provide a reasonable compromise between speed and resolution for MSS data.

A common element of all four schemes is the process of "training". This is the process by which each main class is modeled statistically with the aid of data vectors (patterns) known to belong to that class. If the training data for a class exhibits a multimodal structure, then it is usually divided into two or more subclasses, each corresponding to a mode. This serves two purposes. (1) It enables each subclass to be modeled approximately by a MVN distribution which is completely characterized by a mean vector and covariance matrix. These can be estimated easily from the data vectors assigned to that subclass. (2) Data from a single physical object is usually reasonably unimodal and symmetrical in distribution. Often, those objects which

4.1

are multimodal can be divided into a few smaller objects which are unimodal. Although multimodal training data may be representative of a particular main class as a whole, it is not representative of the individual objects which compose that class. Since it is the individual objects that must be dealt with, the definition of unimodal subclasses is a logical step to take. In other words, $p(x|W_i)$ (eqn. 2.3.2.1) cannot be expressed in terms of $p(\underline{x}|W_i)$ (eqn. 2.1.2). Each component (mode) of $p(\underline{x}|W_i)$ must be known.

The training and test data for a given scene compose a set of labeled observations which we shall refer to as "reference data". There are many possible methods of using a finite amount of reference data to train a classifier and estimate its error rate. Theoretically the best training (i.e. the lowest error rate) is obtained by using all the available reference data for training. If the same data is used for testing, this is called the "C-method". Theoretical and experimental results indicate that, for the Bayes classifier at least, the C-method produces an optimistic (negatively biased) estimate of the error rate; but the bias and variance of the estimate decrease roughly as the reciprocal of the number of observations used [18]. In contrast, the "U-method" requires test data to be independent of training data. The most common procedure (called "sample partitioning") is to use a relatively small proportion (p) of the reference data for training and the remainder for testing. In this case the error estimate is

unbiased and its variance decreases as the reciprocal of the number of test data; but the actual error rate tends to be larger than with the C-method, and its variance is p^{-2} times larger than the error variance by the C-method.

The interpretation of results is usually somewhat easier for the C-method, because the question of whether or not the training is "representative" of the test data does not arise. For comparative purposes our interest is in relative (rather than absolute) performance, so the bias induced by the C-method tends to cancel out. There is no reason to believe that the bias would be significantly greater for one scheme than for the others.

On the other hand the U-method is routinely used in conventional analysis work where absolute performance is emphasized. Effective representation is obtained in most cases by using a fairly large training data set that consists of observations drawn from the same general regions as the test data. For some of the data sets used in our investigation, reasonably good training statistics are available from previous conventional analyses. By using available training we obtain results with minimum effort, and the results relate directly to those obtained by conventional methods.

For those data sets where previous training is unavailable or inadequate, neither the C-method nor the U-method is used. Instead, the available test areas are sampled at an interval sufficient to provide a reasonably

4.1

large training set. Like the C-method this method is simple, it produces representative training, and it eliminates human bias in selecting the training set. It also induces considerably less bias into the error estimate than the C-method does. Of course, once the training set is obtained, feature selection and subclass definition may have to be done before training is complete. An example of this process is described in the next section.

4.2 Run 71052800 - Crop Identification

MSS data collected over a particular region at a particular time and stored on digital magnetic tape is catalogued by "run" number. Run 71052800 is a set of 12 channel data collected over flightline 221 in Indiana on August 12, 1971 during the 1971 Corn Blight Watch Experiment [25]. The correspondence between channel numbers and spectral bands is indicated in Figure 4.2.1. Channels 1-7 cover the visible portion of the spectrum, 8-11 lie in the reflective IR portion, and 12 is a thermal IR sensor. It is evident from the figure that there is a considerable amount of redundancy in the coverage of the visible spectrum. In other words, the data in channels 1-7 will tend to be strongly correlated, causing the information in these channels to be rather redundant, at least more so than in the IR channels.

The area covered by this run is a rectangular strip of agricultural land about 1.6 km wide and 13.8 km long. It is

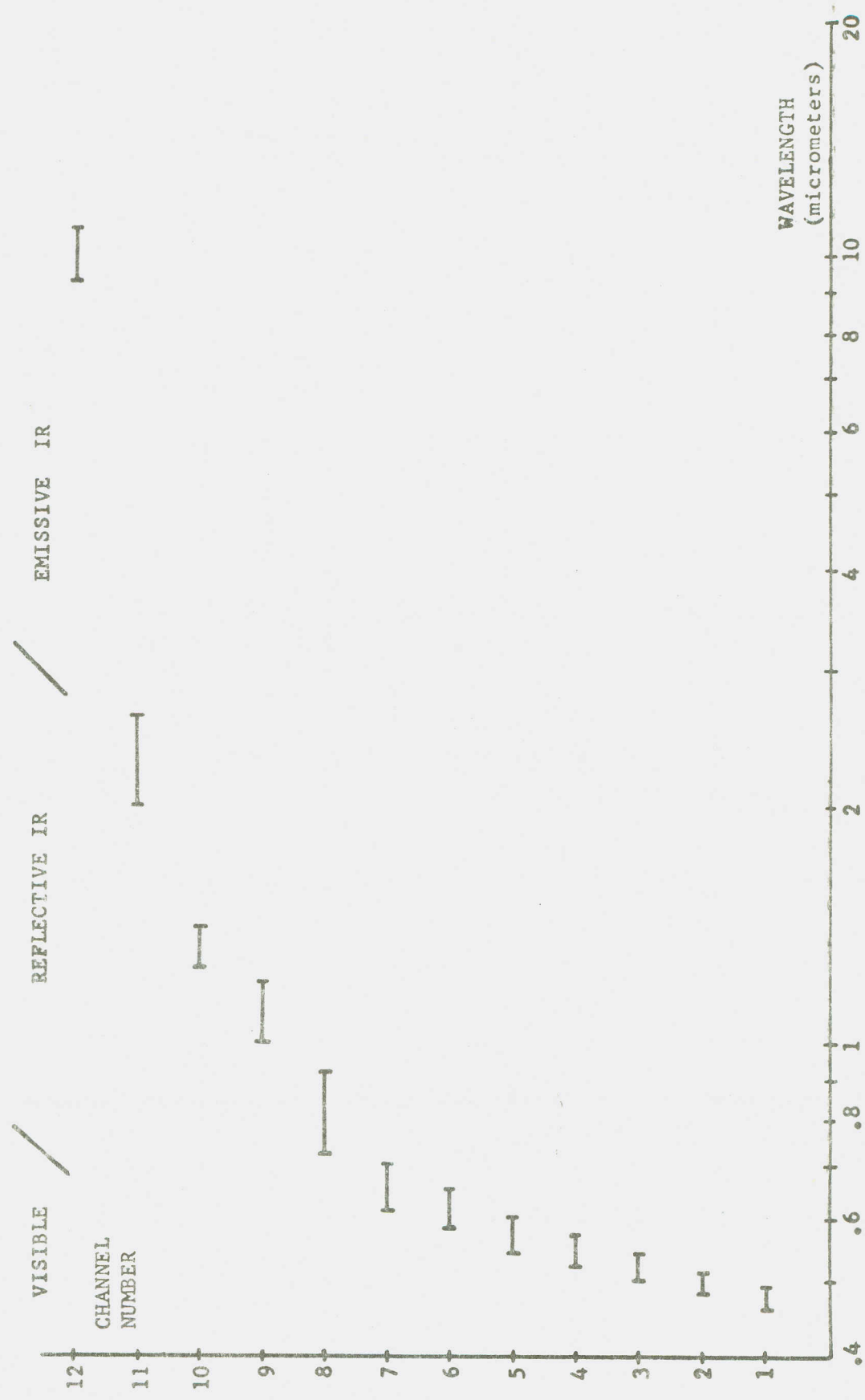


Figure 4.2.1 Correspondence Between Channel Number and Spectral Band for Aircraft Scanner

sampled 222 times along its width and 1374 times along its length. The scanner was carried by aircraft at an altitude of 1524m with an instantaneous field-of-view of three milliradians.

This data set was chosen for analysis for several reasons. (1) A large number of test areas, containing 84,855 pixels, were available from a previous crop identification study [25]. (2) The complexity of the classification is high, providing an opportunity to see how well the new techniques perform in such a situation. (3) The data set contains a combination of some very challenging classes and intermediate classes, as well as some easy classes to identify. The 11 main classes are: corn, soy, wheat (mostly harvested), rye, hay, lespedeza (a grass), pasture, wooded pasture, forest, idle fields, and non-farm. The latter two categories tend to be "catch-all" types which are characteristically difficult to identify by conventional methods. The reason for this will be discussed later.

No previous training statistics were available for this data set, so training data was obtained by sampling the test areas as per Section 4.1. The resultant ratio of the amount of test data to the amount of training data is approximately 5:1. The LARS system (LARSYS) STATISTICS processor [26] was used to compute the statistics of each main class, and SEPARABILITY [26] was used to compute transformed divergence values for every pairwise combination of main classes and every combination of 6 channels. The purpose of this is

"feature selection"; i.e. due to high correlation between channels, it is usually possible to find a subset of the available channels that discriminates between the main classes almost as well as the complete set. Typically only 3 or 4 of the 12 available channels are used to analyze aircraft scanner data, resulting in a large time savings. For the present study, on the basis of the transformed divergence results, the best set of 6 channels is (2,4,10,11,12) plus either channel 8 or channel 9. (Recalling the discussion of Figure 4.2.1, this result is not surprising.) Channel 8 maximizes the average transformed divergence (averaged over all class pairs), and channel 9 maximizes the minimum transformed divergence for any pair of classes. The difference between the two is slight, so channel 8 was selected arbitrarily. Based on histograms of the training data it was decided to subdivide some of the training classes. The LARSYS CLUSTER processor [26] was used to cluster these classes into 2 or 3 modes and SEPARABILITY was used to determine the divergence between modes. Histograms, cluster quotients, and divergence values were examined to determine if the modes of each class were distinct, and if two modes were not distinct, they were recombined (pooled) into a single mode. The final result is 2 subclasses each for corn, soybeans, lespedeza, and idle, 3 subclasses of pasture, and only 1 "subclass" for each of the remaining classes, a total of 17 spectral classes. The LARSYS MERGESTATISTICS processor was used to merge them into

4.2

a single LARSYS statistics deck. Expecting to further reduce the number of channels needed, SEPARABILITY was again applied. It was found that although most class pairs can be distinguished on the basis of some set of 4 or fewer channels, there is no one set of 4 channels which can adequately do this for all class pairs. Thus it was decided to do 6-channel classification. This ended the training phase of the analysis.

Next the data set was classified by a number of different schemes. The LARSYS CLASSIFYPOINTS processor [26] was used to perform ML no-memory classification. The LARSYS SAMPLECLASSIFY processor [26] was used to perform minimum distance (Bhattacharyya) sample classification of the test areas. In the latter case the processor is essentially given a-priori knowledge of the boundaries. Maximum likelihood sample classification of the test areas was accomplished by modifying the SAMPLECLASSIFY software. To avoid confusion this processor will be referred to as "SAMPLECLASSIFY (ML)", and the minimum distance version will be referred to as "SAMPLECLASSIFY (MD)".

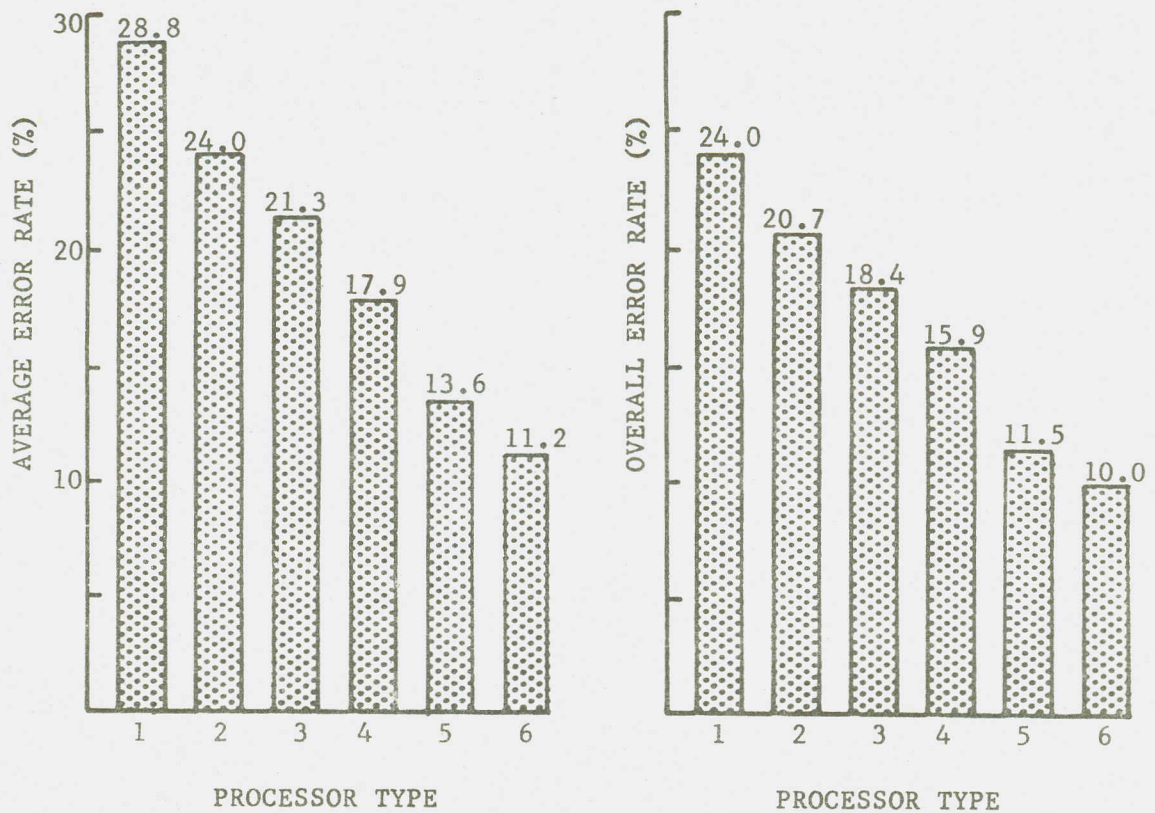
Unsupervised, MUV partitioning and supervised partitioning schemes were implemented using LARSYS-compatible processors that were developed specifically for this investigation. The unsupervised, MV version cannot be used for this data set, because it requires that the number of channels be less than the cell size, which is not the case. The results of all these

analyses are presented in the remainder of this section.

Figure 4.2.2 shows classification performances achieved by the various processors indicated above. Note that processor #2 is equivalent to cell selection without annexation. Thus comparing the results of #2 with the results of #1 (CLASSIFYPOINTS) gives a good indication of the effectiveness of Level-1 alone. And comparing the results of #2 with the results of #3 and #4 indicates the effectiveness of just the annexation (Level-2) phase of processing.

Also note that processor #5 should give about the same results as if the entire partitioning phase were done flawlessly. Thus one can think of the results of #5 as a performance "goal". This goal, however, is not a strict bound (more on this later).

Both "average" and "overall" error rates are shown in Figure 4.2.2. The former is just a straight average (over all classes) of the observed class-conditional error rates. The latter is a weighted average, where the error rate of each class is weighted by the proportion of test data in that class. These proportions are given in Table 4.2.1. Assuming that these proportions coincide roughly with the actual proportions of the classes in the data set, then the overall error rate can be taken as an estimate of the unconditional probability of error.



Processor Key:

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, $t=0$
- #3 Optimum Unsupervised Partitioning
- #4 Optimum Supervised Partitioning
- #5 SAMPLECLASSIFY (ML)
- #6 SAMPLECLASSIFY (MD)

Figure 4.2.2 Classification Performance vs. Processing Scheme
- Run 71052800

Table 4.2.1 Relative Influence of Each Class on Overall Performance - Run 71052800

<u>Class</u>	<u>Percentage of Total Test Pixels</u>
Corn	43.1
Soy	22.4
Wheat	17.6
Idle	4.9
Non-Farm	4.3
Lespedeza	3.0
Pasture	1.6
Hay	1.5
Wooded Pasture	0.8
Rye	0.5
Forest	0.4

4.2

The class-conditional error rates are given in Figure 4.2.3. The results are grouped by class for easy comparison. Several observations are worthy of mention at this point.

Observation 1

Both the unsupervised and supervised modes are effective at reducing the error rate. As expected, the supervised mode has a fairly consistent advantage. It performed better for 8 of the 11 classes, and its average and overall error rates are lower. The actual reduction in error rate due to the supervised mode is 10.9% (average) and 8.1% (overall).

Observation 2

As one would expect, the relative effectiveness of the ECHO approach is highly class dependent. The effect varies from slight degradation for some classes to vast improvement for others.

Observation 3

The classes where the greatest gains are made are wheat, wooded pasture, idle, and non-farm. It has already been observed that the latter two are "catch-all" categories which are typically difficult to identify using CLASSIFYPOINTS. The reason for this is that such classes tend to have relatively broad probability density functions which overlap with those of other classes but at a lower likelihood level. Recalling the case of Figure 2.3.3.1, the

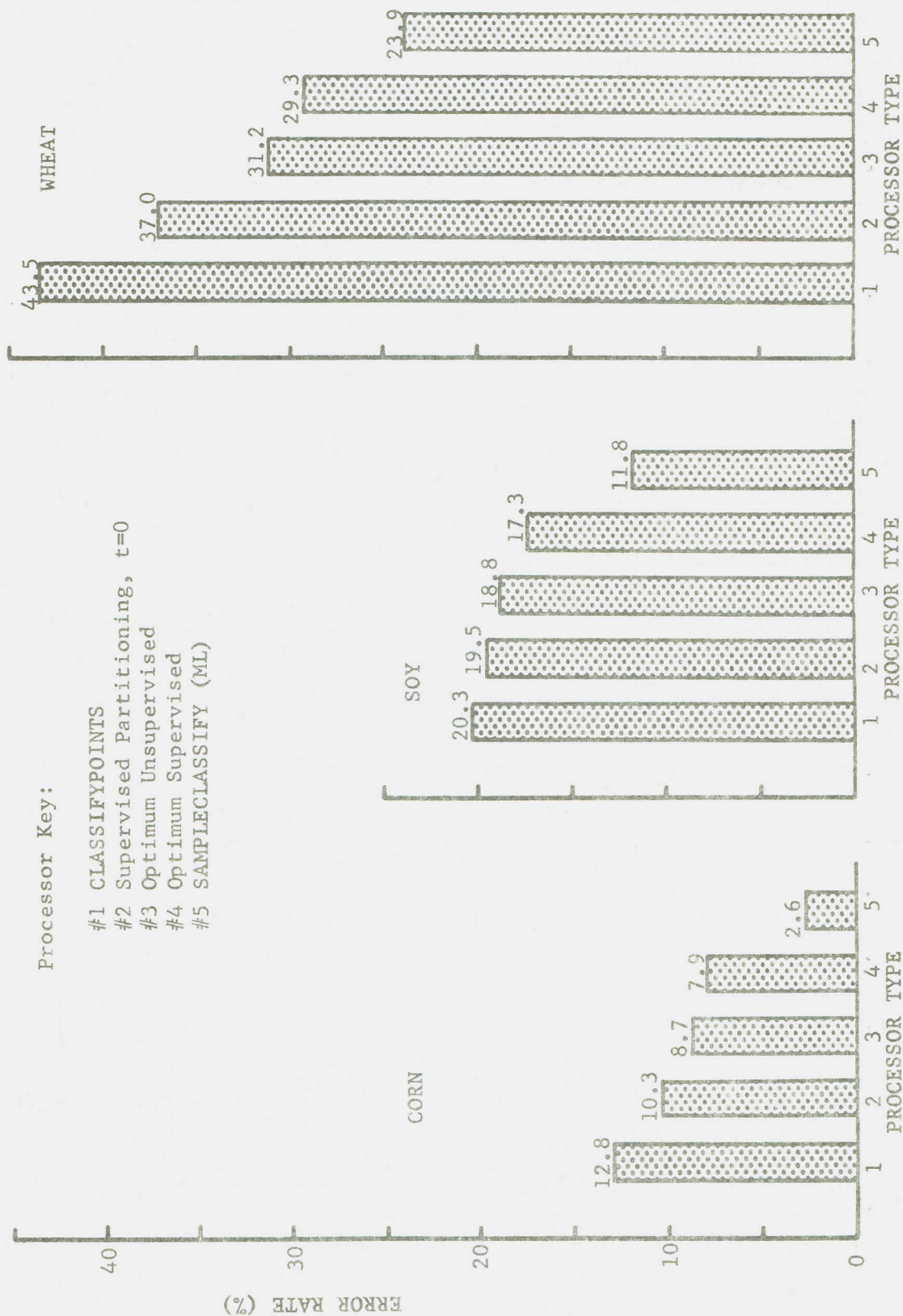


Figure 4.2.3 Performance By Class (Run 71052800)

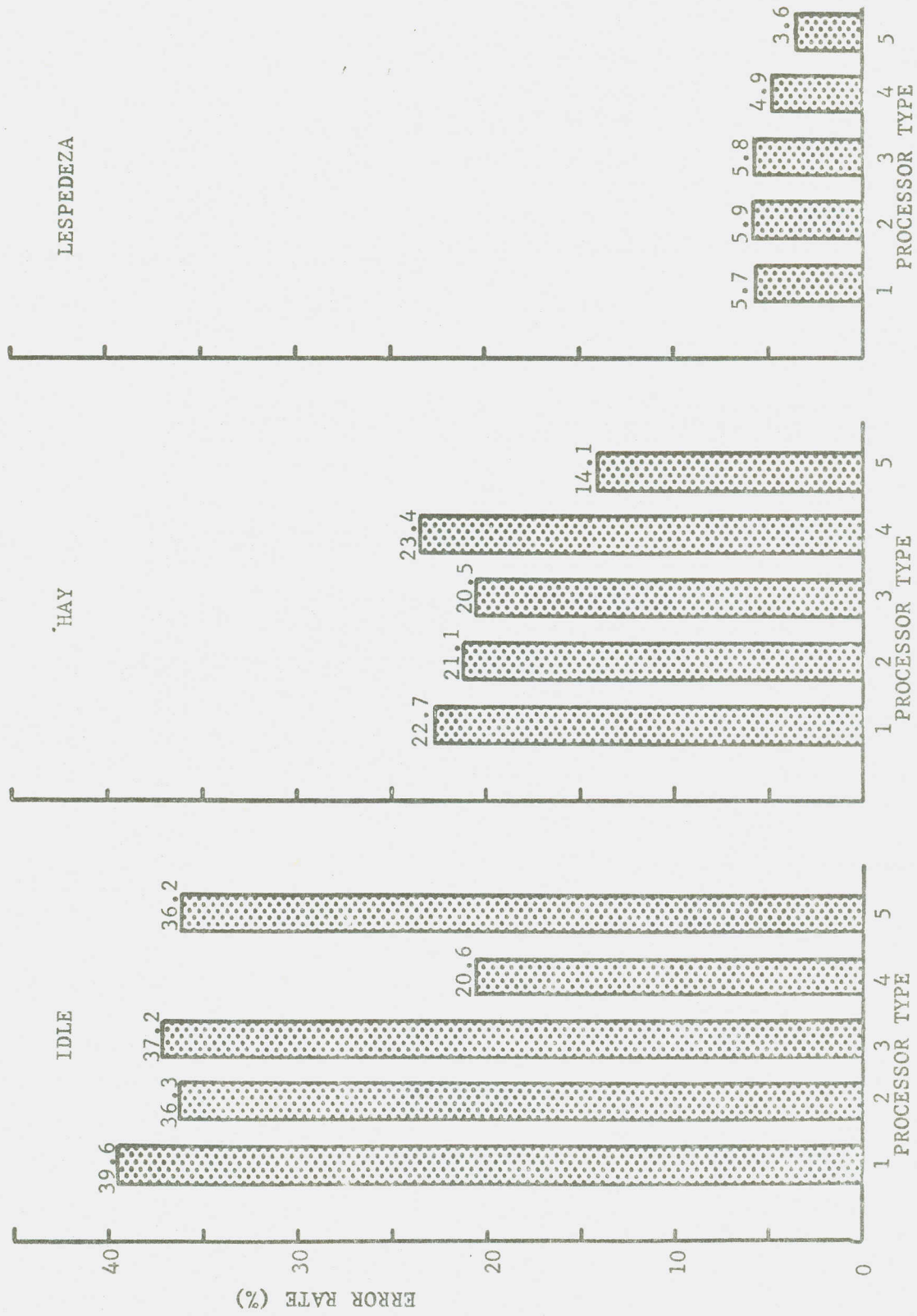


Figure 4.2.3, continued

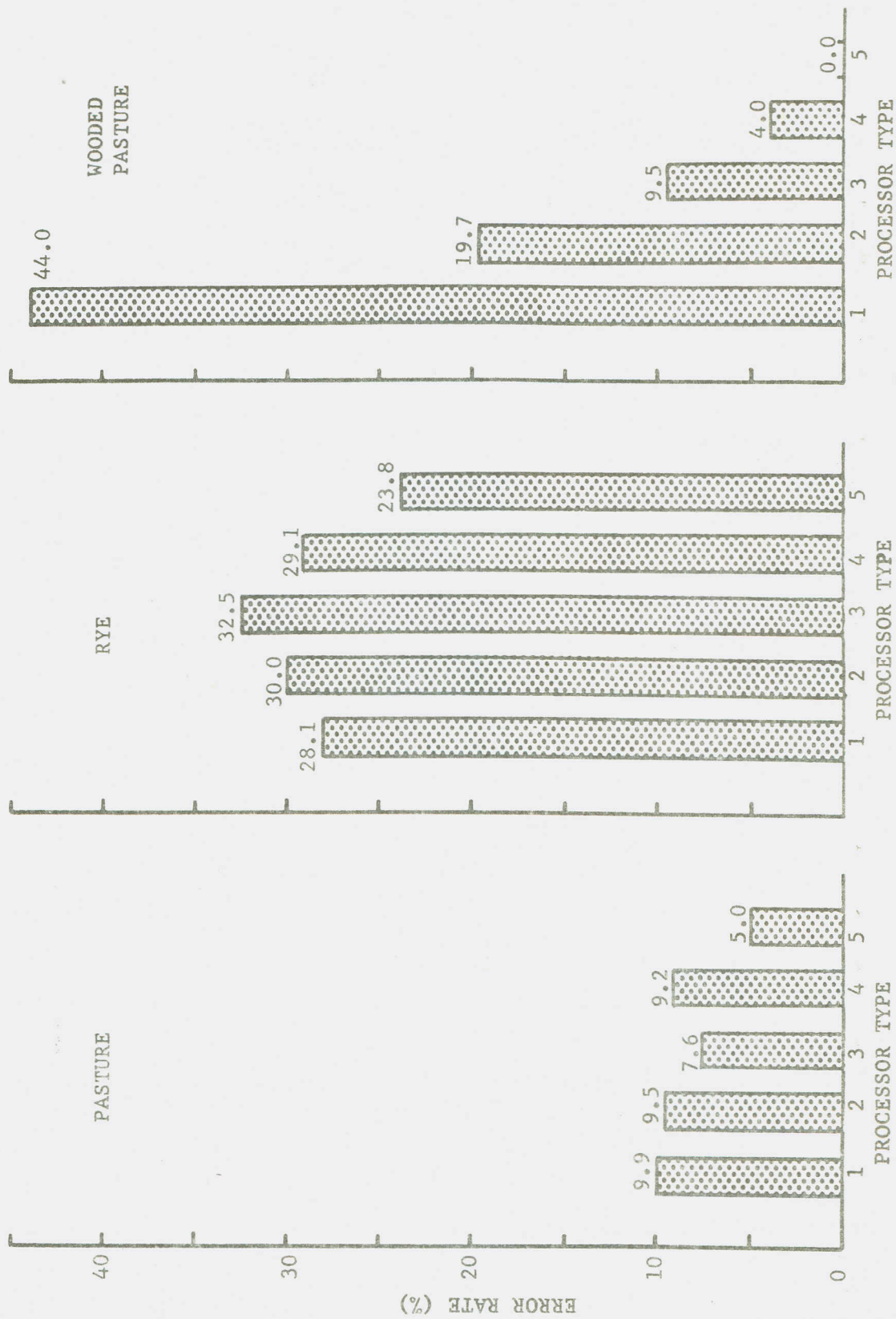


Figure 4.2.3, continued

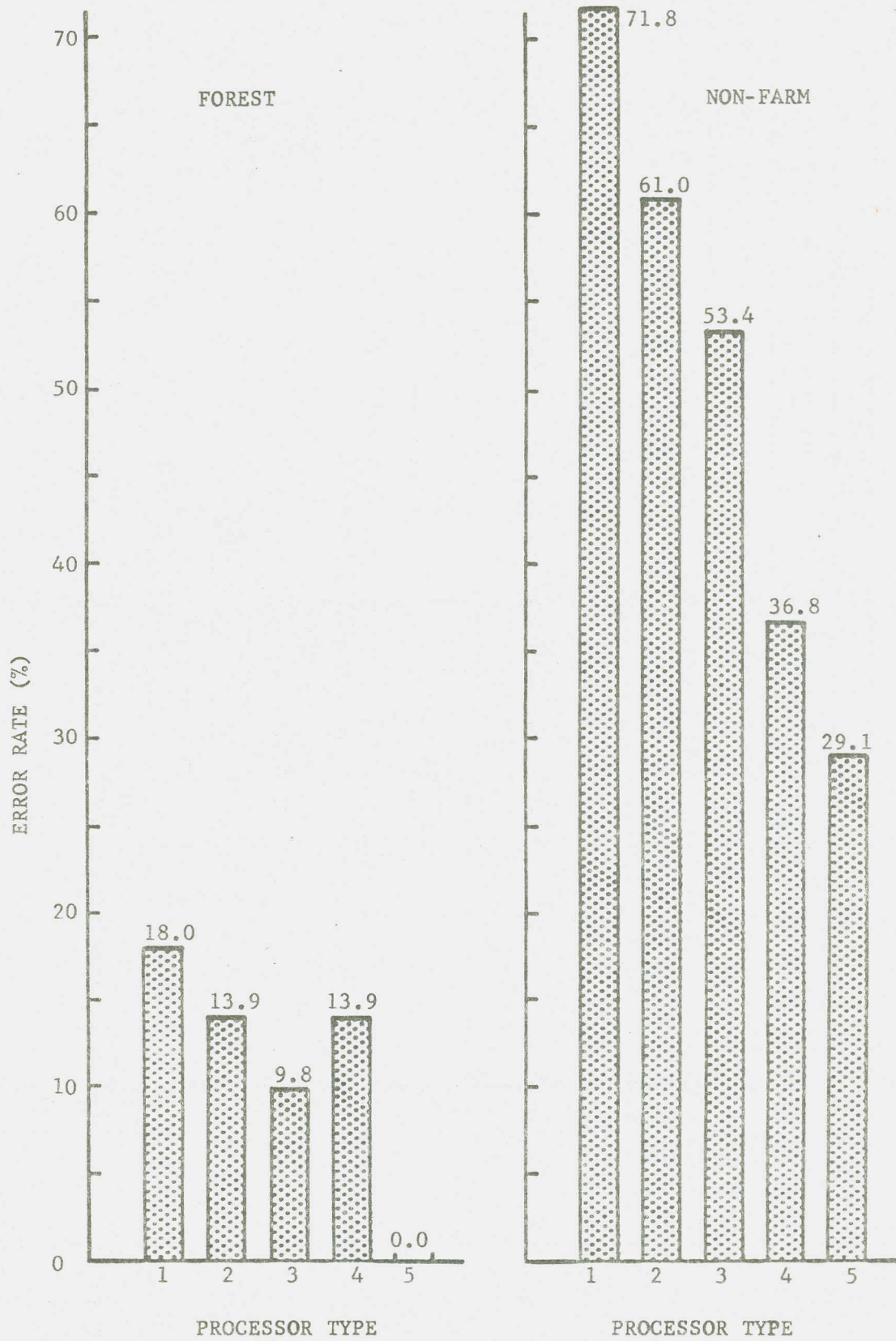


Figure 4.2.3, continued

4.2

conditional error rate for the "broad" distribution is 50%, while the conditional error rate for the "narrow" distribution is just 17.4%. If other classes overlap the "tails" of the broad distribution, then this discrepancy becomes even greater. But when a sample of data from the broad distribution is made available for classification, it usually consists of a mixture of values both near and far from the mean. This makes it possible for the classifier to determine the correct classification of the sample.

The wooded pasture class also has a relatively broad distribution due to its composition and spatial texture. Note that this does not necessarily imply that wooded pasture is statistically inhomogeneous or bimodal. We refer to it as a "compound" class. (See Appendix C.) In this particular case it might at first appear that the method of using test areas to evaluate performance is biased in favor of sample classifiers. In other words, the large error rate observed for CLASSIFYPOINTS may be due to the assignment of many test points to the classes forest and pasture, which may actually be accurate labels for those particular pixels. This argument is invalid on several counts. First of all, the number of test points classified as forest or pasture accounts for only 12 of the 44 percent error rate. The reduction in error rate brought about by supervised partitioning is 40 percent, or 3.3 times as great as the maximum possible error attributable to this cause.

4.2

Secondly, whenever CLASSIFYPOINTS classifies an area in a "salt and pepper" manner, the information is highly unreliable. If the area actually were that way, Premise A (Section 1.2) would be violated. Thirdly, even if valid point-by-point classification were possible, most analysts are not interested in the actual classification of each individual pixel. Instead their goal is to produce a "type map" which consists of a partition of the region with a general label assigned to each element of the partition. An element containing a mixture of trees and pasture for example would be labeled "wooded pasture".

These points are illustrated in Figures 4.2.4, 4.2.5, and 4.2.6. Figure 4.2.4 shows a section of Run 71052800 (lines 101-300) that has been classified by CLASSIFYPOINTS. Each class has been assigned a gray level and displayed electronically to form the image. The "classification noise" is readily apparent. In contrast to this, Figure 4.2.5 shows the same section as classified by ECHO (supervised). The random errors have, for the most part, been eliminated. This map is much closer to the desired result than is the CLASSIFYPOINTS output. Figure 4.2.6 shows the centers of these two maps in greater detail. Each class is represented by an assigned symbol (or blank), and each symbol represents one pixel. The four rectangular areas are test areas designated as wooded pasture. This class is displayed as a blank space to emphasize the



Figure 4.2.4 Gray-Scale-Coded Classification Map
- Produced by CLASSIFYPOINTS

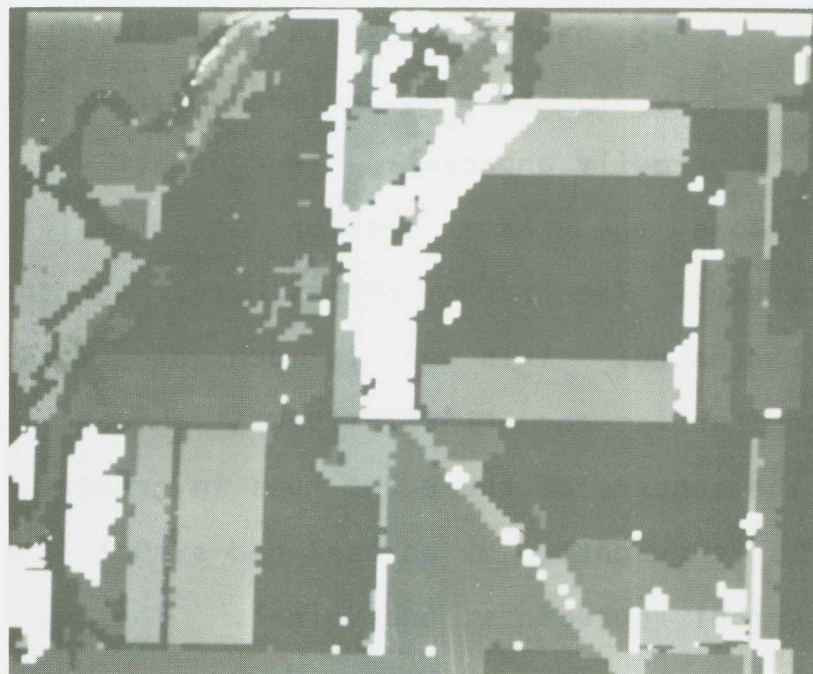


Figure 4.2.5 Gray-Scale-Coded Classification Map
- Produced by ECHO



CLASSIFYPOINTS



ECHO

<u>SYMBOL</u>	<u>CLASS</u>	<u>SYMBOL</u>	<u>CLASS</u>
W	Wheat	I	Idle
=	Lespedeza	*	Forest
-	Pasture	.	Corn, Soy, Rye, Hay
blank	Wooded Pasture	/	Non-Farm

Figure 4.2.6 Logogrammatic Classification Maps

4.2

contrast between it and the others. The diversity of symbols in the test areas testifies to the inadequacy of CLASSIFYPOINTS for classifying such textured regions. Most of this confusion is avoided by the ECHO technique.

The wheat class too has a broad distribution, probably due to the fact that the wheat is mostly harvested. Whatever the cause, it adds further support to the argument that classes with broad distributions tend to benefit the most by sample classification. To clarify this point further, the classification improvement is plotted in Figure 4.2.7 vs. the common logarithm of the generalized variance. In the case of a class with subclasses the average generalized variance is used. For this data, the correlation between these quantities is 0.81.

Of course a broad distribution does not necessarily imply that partitioning and sample classification will produce dramatic improvements over CLASSIFYPOINTS. For example, another class may have about the same distribution, in which case no classification scheme can reliably distinguish between them. Or the class may be so unlike any other class that CLASSIFYPOINTS leaves no room for improvement. Also the broad distribution may be caused by inadequate training (i.e. not representative), in which case accurate classification may be impossible until the training is corrected. Obviously the mechanisms which affect classification performance in a multidimensional,



Figure 4.2.7 Class Improvement vs. Log-Generalized Variance

multiclass, multi-subclass situation are very complex. Observation 3 provides only limited insight into the overall process.

Observation 4

The supervised ECHO results for class idle actually surpass the performance "goal" set by processor #5. A conceivable explanation for this is that idle test areas may actually consist of several physical objects containing different subclasses of idle. Since the ECHO processor can classify such objects separately, it can actually provide an advantage over SAMPLECLASSIFY (ML) which must classify each test area as a whole.

Observation 5

As expected, processor #2 (Level-1 partitioning) can provide a fairly significant degree of improvement on its own. Again the effect is strongly class dependent. The effect would probably be much greater if not for the correlation that exists between adjacent pixels.

The main parameter that is required for the supervised mode is the annexation threshold, t . Figure 4.2.8 shows how the average error varies for seven values of t . Of these, the optimum value is $t=5$, although all values tried gave significantly better performance than CLASSIFYPOINTS. The Level-1 threshold, being of much lesser importance, was not varied in this study. It was previously established at $c=90$ by processing a small subregion of the data set several

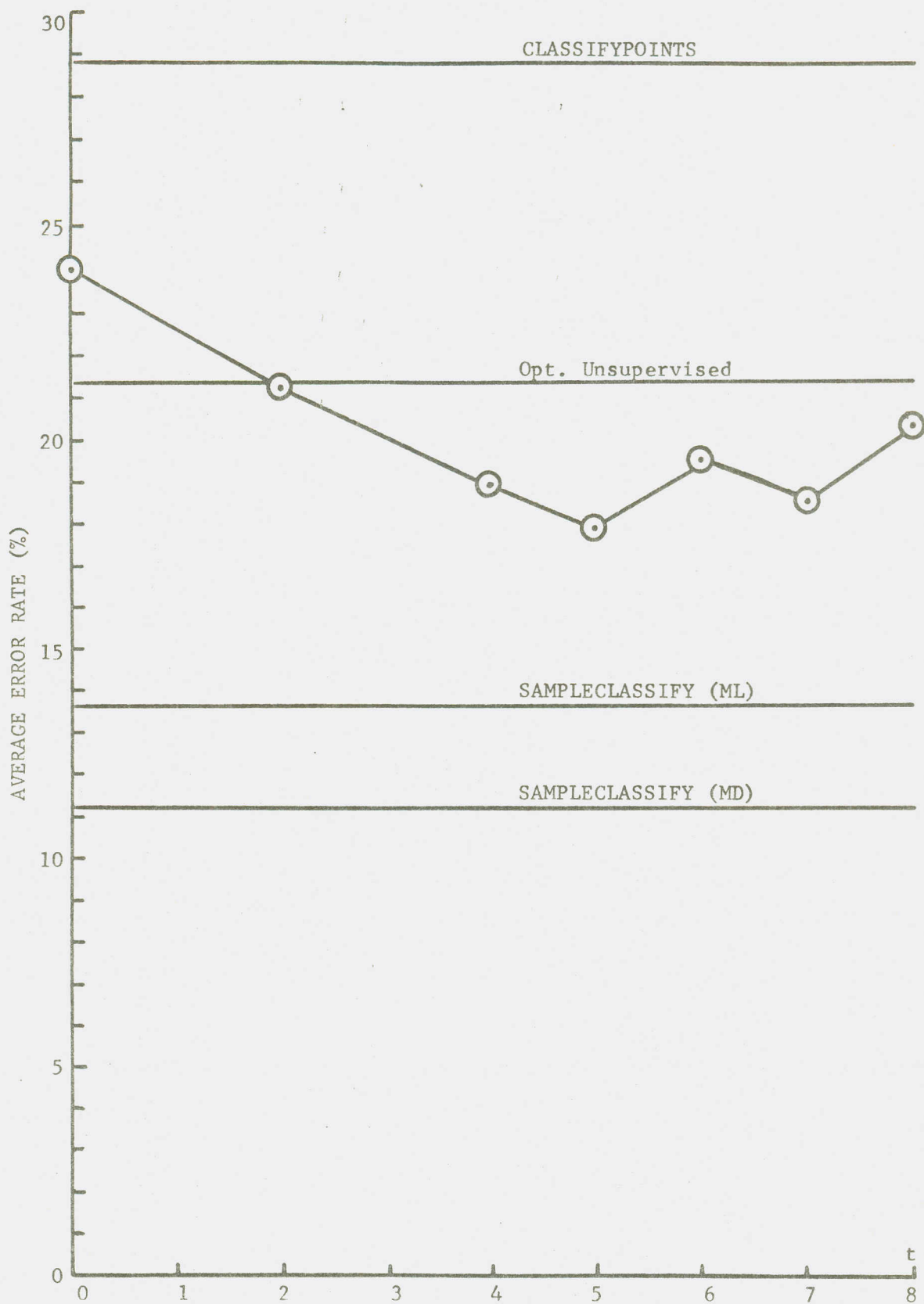


Figure 4.2.8 Effect of Annexation Threshold (t) on Average Performance
- Run 71052800

times while varying c . This value provides a sufficiently low rejection rate that the occurrence of singular cells is limited mainly to patterns resembling boundary lines (as desired). Thus the classification results are not necessarily optimized with respect to c , but they are believed to be near that optimum.

Figure 4.2.9 shows the behavior of the overall error vs. t . It is very similar to the average error except that the results are shifted downward due to the heavy influence of the corn class. Its minimum also occurs at $t=5$.

The analogous results for the unsupervised mode cannot be presented as easily because the performance is a function of two variables, the significance levels. For the same reason, the optimum performance cannot be determined as easily. Tables 4.2.2 and 4.2.3 give average and overall error rates for eight different combinations of significance levels. The Level-1 threshold was maintained at a constant $c = 0.25$. Cells found to be singular at this level were classified as small samples rather than as individual pixels; i.e. "cell-splitting" was not in effect. The best results occur at about $s_2 = .005$ or $.001$ and $s_1 = .001$. Possibly a lower value of s_1 would produce better results, but this is beyond the capability of the current processor.

Figure 4.2.10 shows how the processors compare with regard to both error rate and CPU time. In terms of time, the unsupervised mode is the fastest by far because it performs

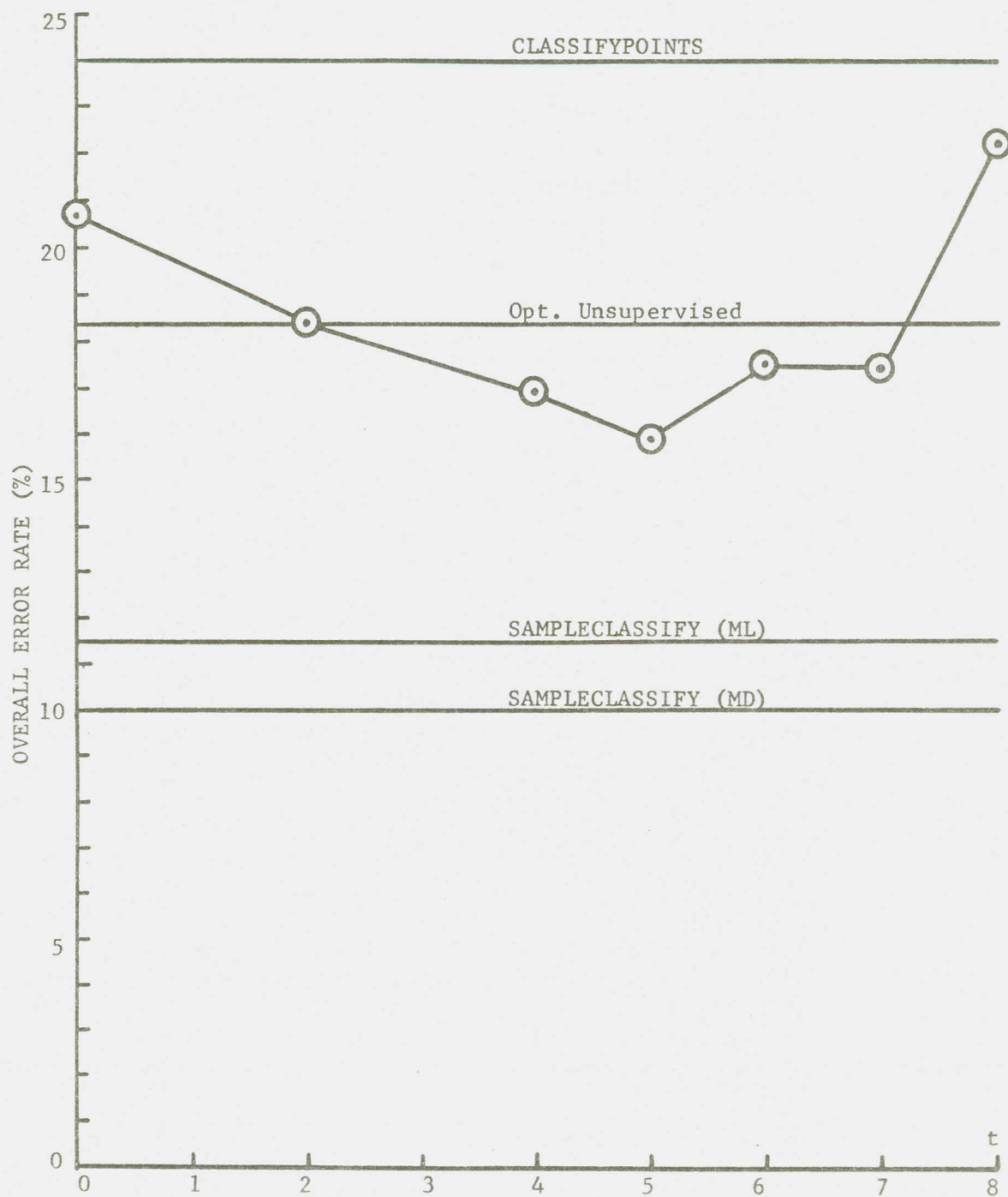


Figure 4.2.9 Effect of Annexation Threshold (t) on Overall Performance
- Run 71052800

Table 4.2.2 Matrix of Average Error Rates (%) for Eight Combinations of Significance Levels
- Run 71052800

$s_2 \backslash s_1$.001	.005	.025	.100
.000		24.6	24.3	
.001	21.0	23.4		
.005	21.3	21.9		23.9
.025		21.8		

Table 4.2.3 Matrix of Overall Error Rates (%) for Eight Combinations of Significance Levels
- Run 71052800

$s_2 \backslash s_1$.001	.005	.025	.100
.000		20.1	20.1	
.001	18.7	19.7		
.005	18.4	19.3		20.4
.025		19.5		

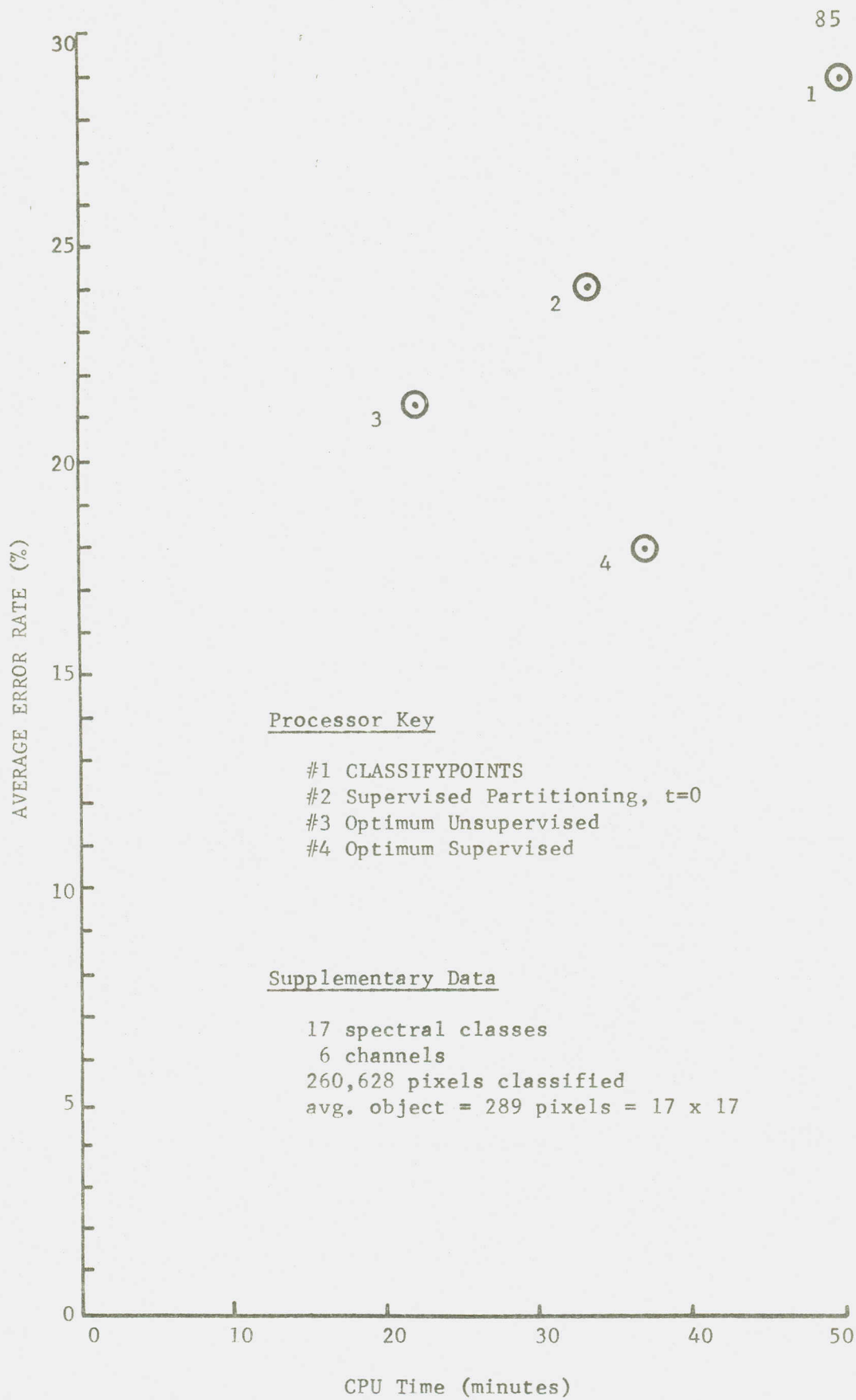


Figure 4.2.10 Error Rate and CPU Time for Four Classification Schemes
 - Run 71052800

4.2

the fewest classifications. CLASSIFYPOINTS is the slowest because it performs the most. The supervised mode is in between and provides the lowest error rate.

One of the factors influencing CPU time is the size (in pixels) of an average object, since the larger the objects the greater the number of pixels that can be classified at one time. A rough indication of this factor is obtained by dividing the number of test areas into the total number of test pixels. As indicated on Figure 4.2.10, for this data set an average test area (object) is equivalent to a square, 17 pixels wide.

4.3 Run 72064412 - Classification Of Satellite Data

Three LANDSAT passes over a region in Indiana on different dates were combined to produce this data set. Only data from the first date, August 25, 1972, is used for analysis in this study. Four spectral channels are available on LANDSAT-1. The spectral bands are indicated in Figure 4.3.1. The instantaneous field of view for the three visible band channels is 86 microradians. The region covered by the data set is a rectangle 45.1 km wide and 53.1 km long. It is sampled 804 times along its width and 673 times lengthwise, for a total of 541,092 pixels. A 21.4 km by 43.5 km subregion (containing 210,100 pixels) was analyzed. This region was previously the subject of a study of strip mine activity (unpublished). The analyst provided both training statistics and test areas.* Briefly, the

* Courtesy of John Berkebile, LARS.

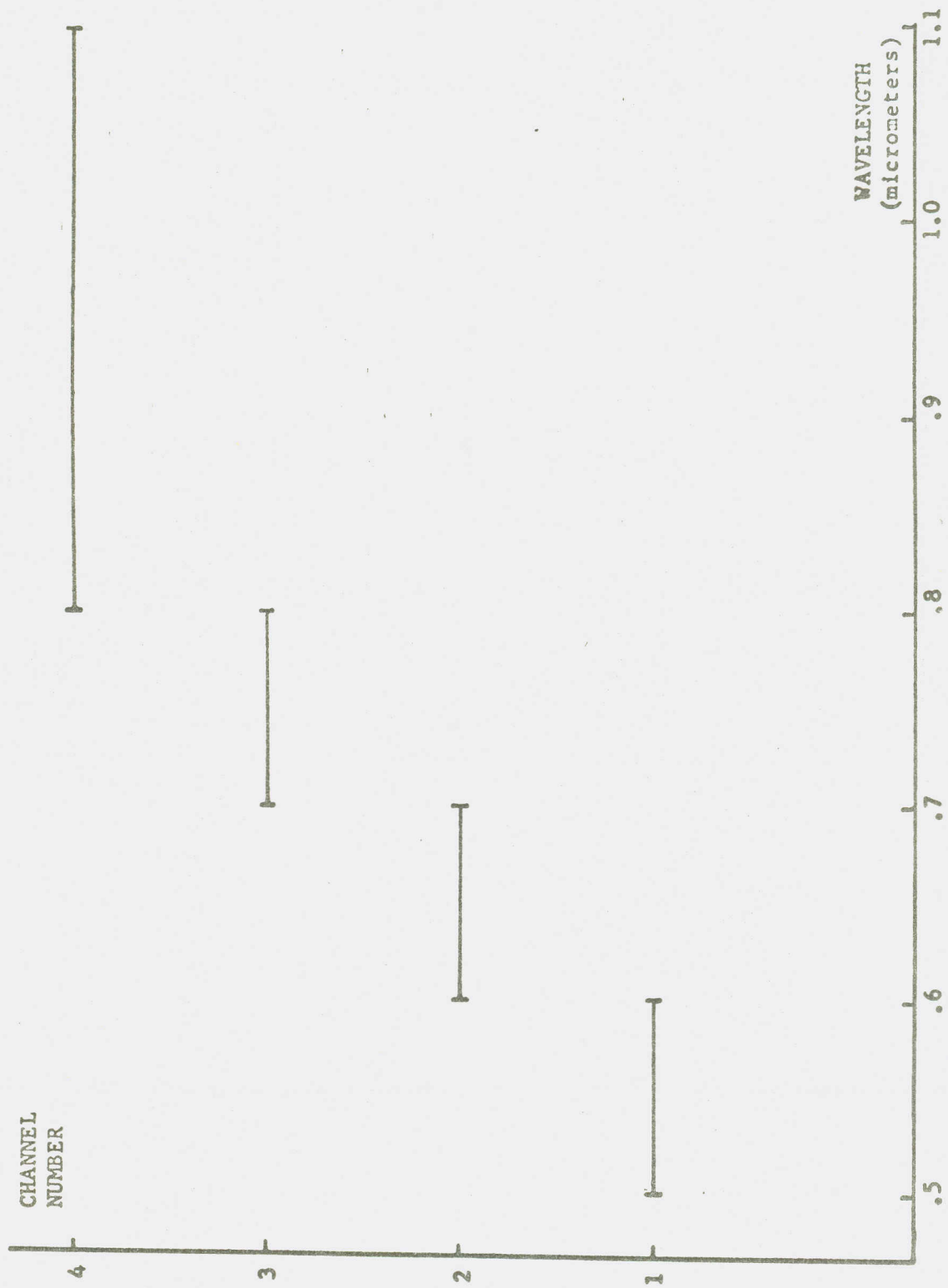
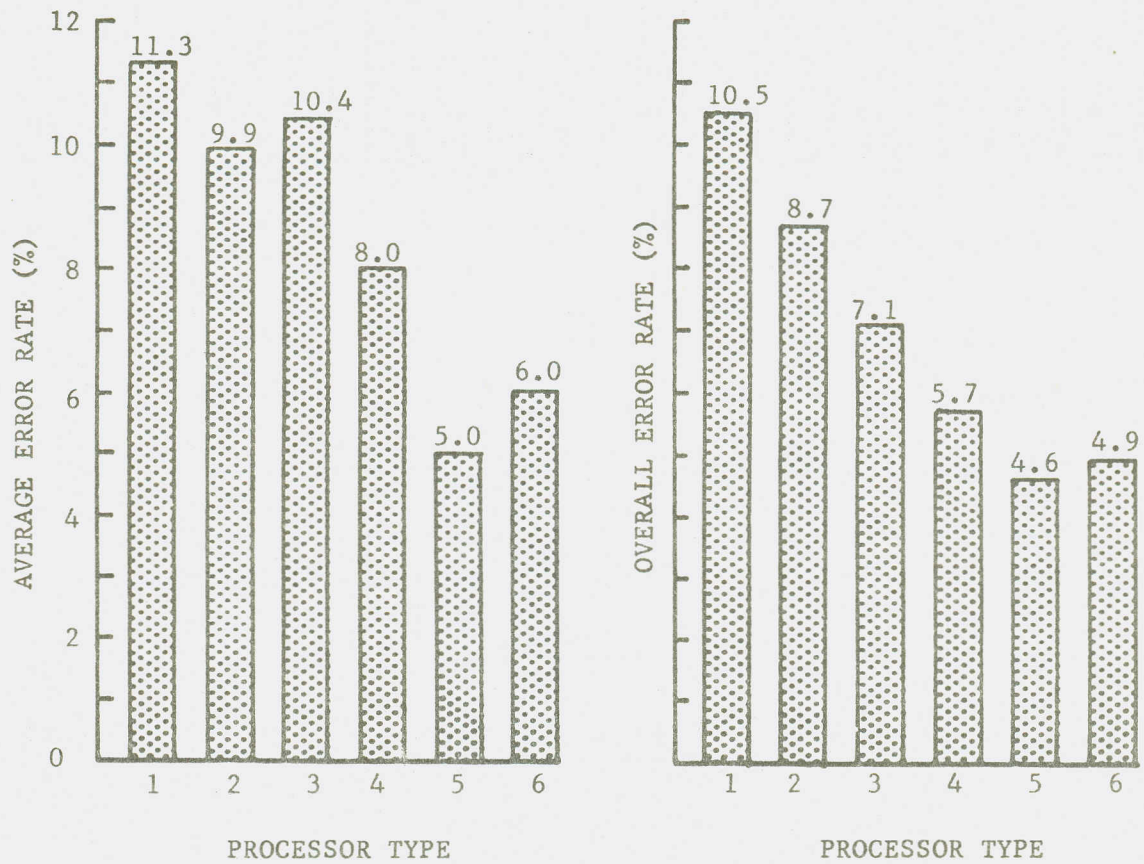


Figure 4.3.1 Correspondence Between Channel Number and Spectral Band for LANDSAT-1 Scanner

4.3

method of training was to perform both a manual analysis of designated training areas, using maps and aerial photography, and an unsupervised clustering analysis of the MSS data corresponding to each training area. The manual analysis was used to associate each cluster class with an informational class, and the statistics of the cluster classes became the training statistics. Training and test areas have no pixels in common. The number of test pixels is 9512, a fairly large number for a LANDSAT analysis. The main classes are agriculture, forest, recent-mine, pit (containing various grades of water), revegetated mine, residential, and clouds. The numbers of subclasses are 3, 2, 4, 2, 1, 1, and 1 respectively, a total of 14. Test areas were supplied only for the first 5 of 7 main classes. All 4 channels are used in the analysis.

Figure 4.3.2 shows the average and overall error rates of the various processors. The weights for overall error are listed in Table 4.3.1. The class-conditional error rates are given in Figure 4.3.3. As for Run 71052800, the number of channels used is too large to permit unsupervised, MV processing. On the whole, the results appear quite similar to those of Run 71052800, in spite of the considerable differences between the two runs. Both the average and overall error rates are significantly reduced by the ECHO techniques, with the supervised mode providing a consistent advantage over the unsupervised mode.



Processor Key:

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, $t=0$
- #3 Optimum Unsupervised Partitioning
- #4 Optimum Supervised Partitioning
- #5 SAMPLECLASSIFY (ML)
- #6 SAMPLECLASSIFY (MD)

Figure 4.3.2 Classification Performance vs. Processing Scheme
- Run 72064412

Table 4.3.1 Relative Influence of Each Class on Overall Performance - Run 72064412

<u>Class</u>	<u>Percentage of Total Test Pixels</u>
Forest	42.1
Agriculture	24.6
Recent Mine	19.1
Revegetated Mine	11.0
Pit	3.2

PROCESSOR KEY

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, t=0
- #3 Optimum Unsupervised
- #4 Optimum Supervised
- #5 SAMPLECLASSIFY (ML)

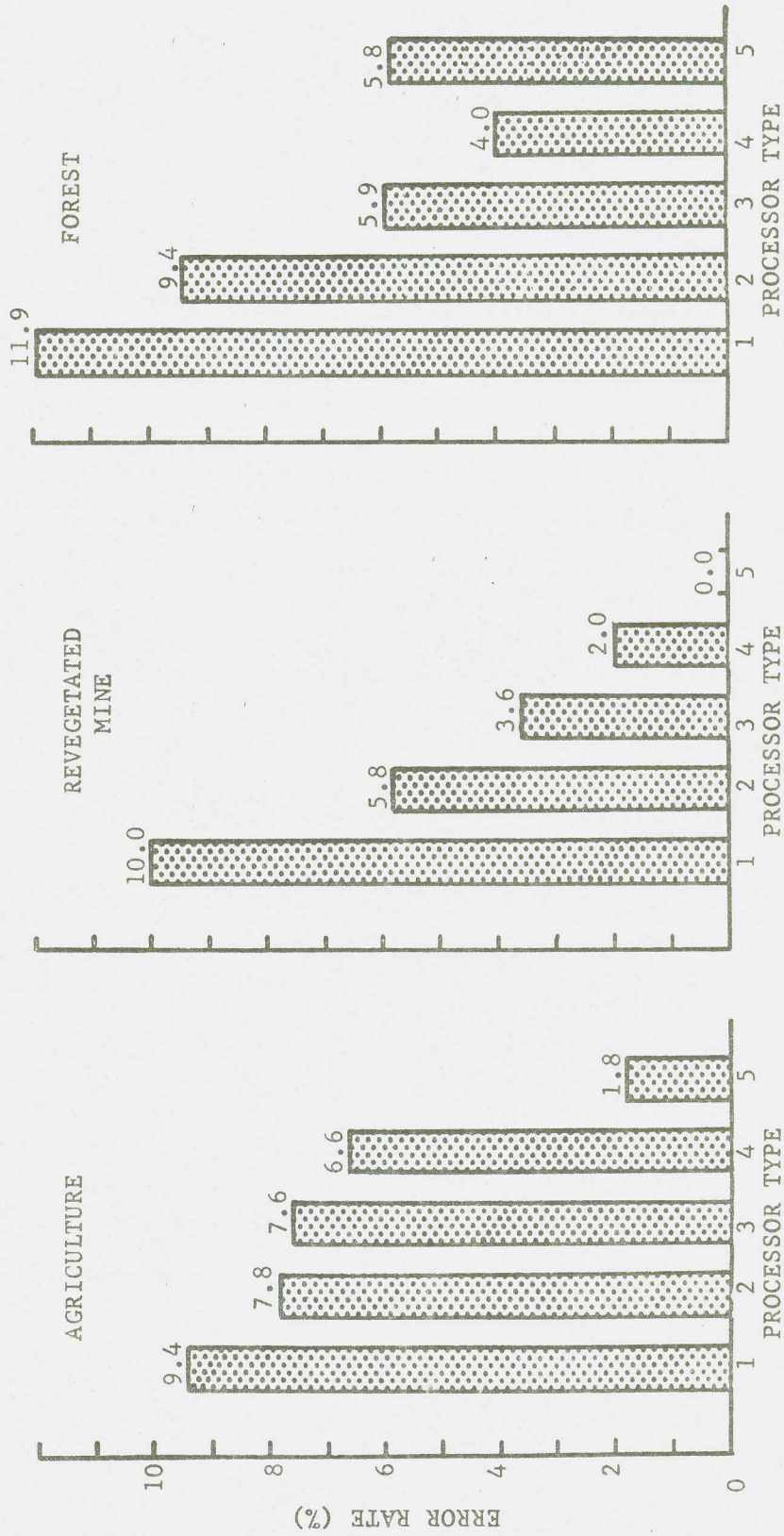


Figure 4.3.3 Performance By Class (Run 72064412)

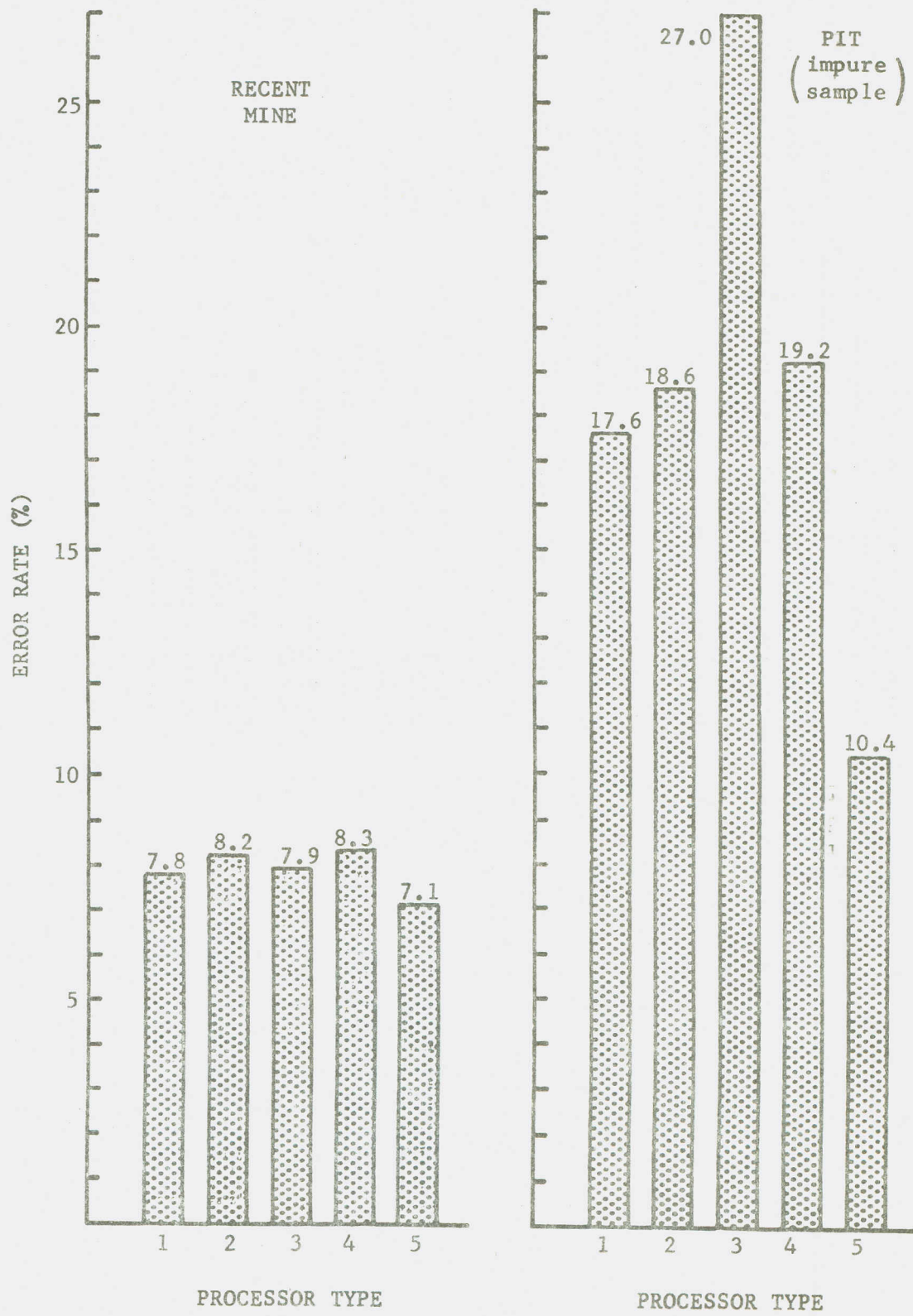


Figure 4.3.3, continued

4.3

The behavior of the pit class is misleading and requires further discussion. Normally water is one of the easiest classes to identify using CLASSIFYPOINTS, yet the 17.6% error rate is by far the highest of any class. Therefore it is apparent that something is wrong with either the original training statistics or test areas. As it turns out, Premise "A" (Section 1.2) has been violated, and this has caused the pit test areas to contain many pixels that overlap both pit and mine or revegetated mine classes. Consequently the pit class results are not truly indicative of performance, but they are included here for the sake of completeness. This accounts for the relatively high average error rate of the unsupervised mode, which is still lower than the average CLASSIFYPOINTS error rate. With the exception of this class, it can again be said that no class is significantly degraded, while some are greatly improved.

The recent-mine class is another that bears comment. Notice that processor #5 performs only slightly better than CLASSIFYPOINTS. Consequently one cannot expect partitioning to improve the accuracy, since the classifier seems unable to use the information effectively. The root of the problem is that about half of the test areas in this class are actually labeled "partially revegetated mine", and some of these appear as revegetated mine to the sample classifier. Apparently these categories are spectrally too similar to distinguish reliably even on a sample basis. Confusion of

this sort is also the source of almost all the recent-mine misclassifications that CLASSIFYPOINTS makes, so there is very little other type of error for the sample classifier to correct.

The performance of the supervised mode is again plotted for seven values of the annexation threshold, t (Figures 4.3.4 and 4.3.5). Again the optimum value is $t=5$. Notice that the overall error approaches both SAMPLECLASSIFY results quite closely. For this study the Level-1 threshold was held at a constant $c=55$. This choice reflects the lower number of channels and higher incidence of singular cells as compared to the aircraft data.

The effect of c was briefly investigated when for $t=4$, the values $c=55$ and $c=80$ were compared. The higher value causes slight improvements in the agriculture and forest classes and slight declines in the others. The overall error rate is unchanged however. The effects are not large enough to be of any serious concern here.

The performance of the unsupervised mode is given in Tables 4.3.2 and 4.3.3 for seven different combinations of significance levels. Of these, the best combination is $s_2=.001$ and $s_1=.005$. The Level-1 threshold was maintained between 0.20 and 0.25, and cell-splitting was not in effect. For comparison, the data was also processed with cell-splitting in effect using the "optimum" s_1 and s_2 above. The pit class performance improved to about the same level as the

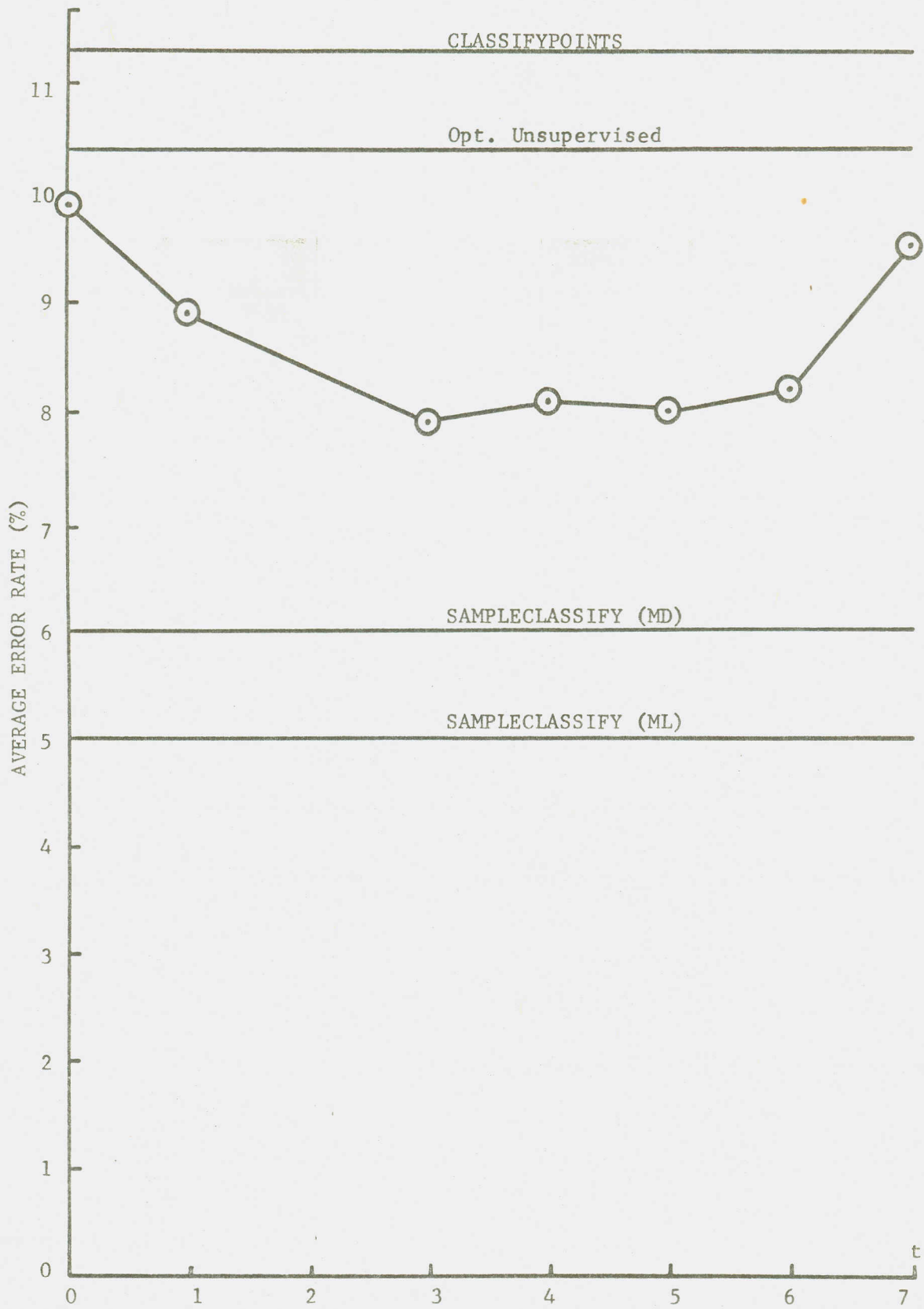


Figure 4.3.4 Effect of Annexation Threshold (t) on Average Performance
- Run 72064412

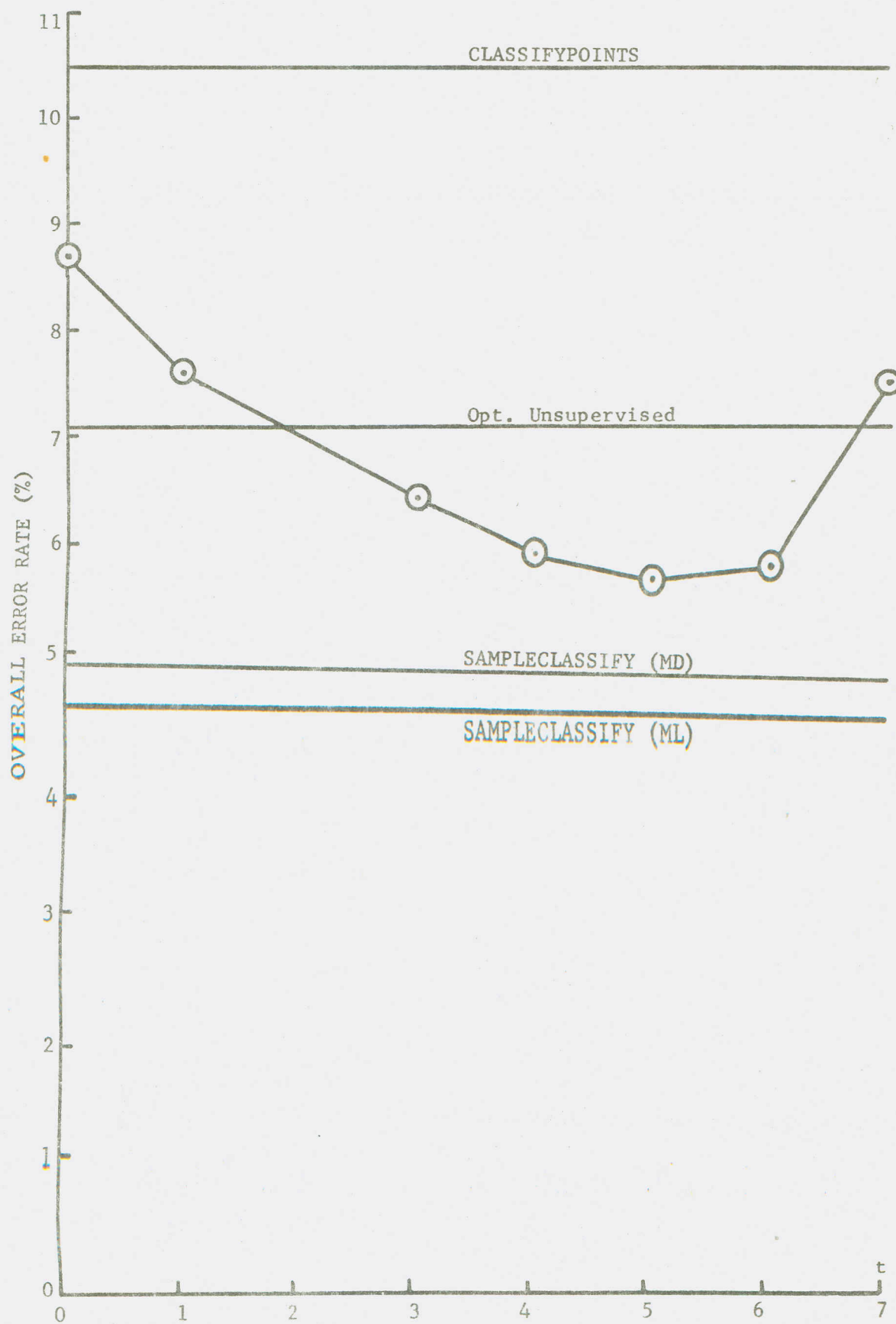


Figure 4.3.5 Effect of Annexation Threshold (t) on Overall Performance
- Run 72064412

Table 4.3.2 Matrix of Average Error Rates (%) for Seven Combinations of Significance Levels
- Run 72064412

$s_2 \backslash s_1$.001	.005	.025	.100
.000	13.1	11.8		
.001	12.0	10.4	10.9	11.8
.005		11.1		

Table 4.3.3 Matrix of Overall Error Rates (%) for Seven Combinations of Significance Levels
- Run 72064412

$s_2 \backslash s_1$.001	.005	.025	.100
.000	12.3	9.5		
.001	9.5	7.1	7.9	8.9
.005		8.0		

other processors, but this is overshadowed by degraded performance in the other classes. In contrast to this, a non cell-splitting version of the supervised mode was tried with $t=4$ and $c=55$, and the class-by-class performance was uniformly worse than with the cell-splitting version. This might imply that the supervised Level-1 test is more effective than the unsupervised Level-1 test or simply that a better Level-1 threshold exists (for the unsupervised test) than the one used here. The evidence is inconclusive on this minor point.

Figure 4.3.6 shows how the processors compare with regard to both overall error rate and CPU time. As before, the unsupervised mode is fastest, and the supervised mode is most accurate. Due to the greatly reduced number of pixels per physical object compared to aircraft data, the difference in speed for the three processors is much less significant.

4.4 Run 71052501 - Forest Cover Mapping

Corn Blight Watch Experiment flightline 218 is a 1.6 x 16.1 km strip of land in southwestern Indiana. In contrast to the relatively flat flightline 221, it is on a "maturely dissected, westward sloping plateau characterized by abundant stream valleys and a well-integrated drainage system." "Most of the land area is in slope, with flat, narrow ridge tops and steep valley walls." [27] Consequently row crops (corn and soybeans) are in the

4.4

minority compared to forage crops. About 60% of the land is covered by hardwood forest with a few, small stands of white pine. MSS data was collected over this region on the same date and by the same means as Run 71052800 and labeled 71052501. The region was sampled 1605 times lengthwise and 222 times across its width. In contrast to 71052800, a sun angle correction transformation was applied to the data. This data set was previously the subject of a forest cover mapping study [27], and the analyst's training statistics and test areas are used in the present investigation. Six main classes are considered: deciduous forest, coniferous forest, water, forage, corn, and soy. A composite class (forest and forage) was deleted by the analyst, who recognized the inability of CLASSIFYPOINTS to handle such data adequately. Our previous results on wooded pasture indicate that this would have been unnecessary if ECHO techniques had been available to him. In contrast to the previously described analyses, the training is as simple as possible, involving no clustering or subclasses. The available reference data (53,516 pixels) was simply divided into non-overlapping test and training areas at a ratio of about 13:1. Based on the transformed divergences, the best set of 3 channels is (6,10,12), which is used in the present investigation.

Figure 4.4.1 shows the average and overall performances of the various processors. The weights for overall

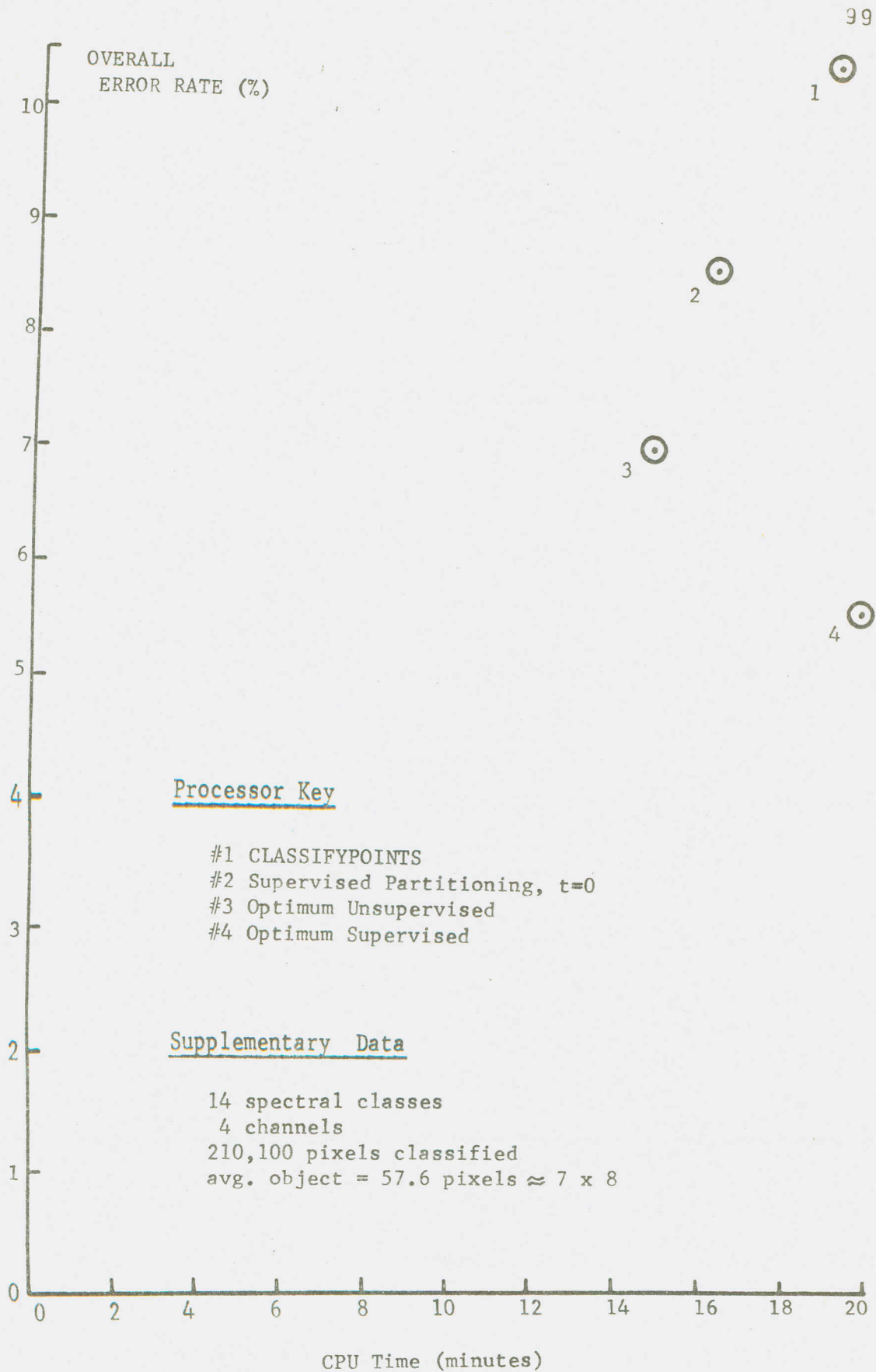


Figure 4.3.6 Error Rate and CPU Time for Four Classification Schemes
- Run 72064412

PROCESSOR KEY

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, t=0
- #3 Optimum MV Unsupervised
- #4 Optimum MUV Unsupervised

- #5 Optimum Supervised Partitioning
- #6 SAMPLECLASSIFY (ML)
- #7 SAMPLECLASSIFY (MD)

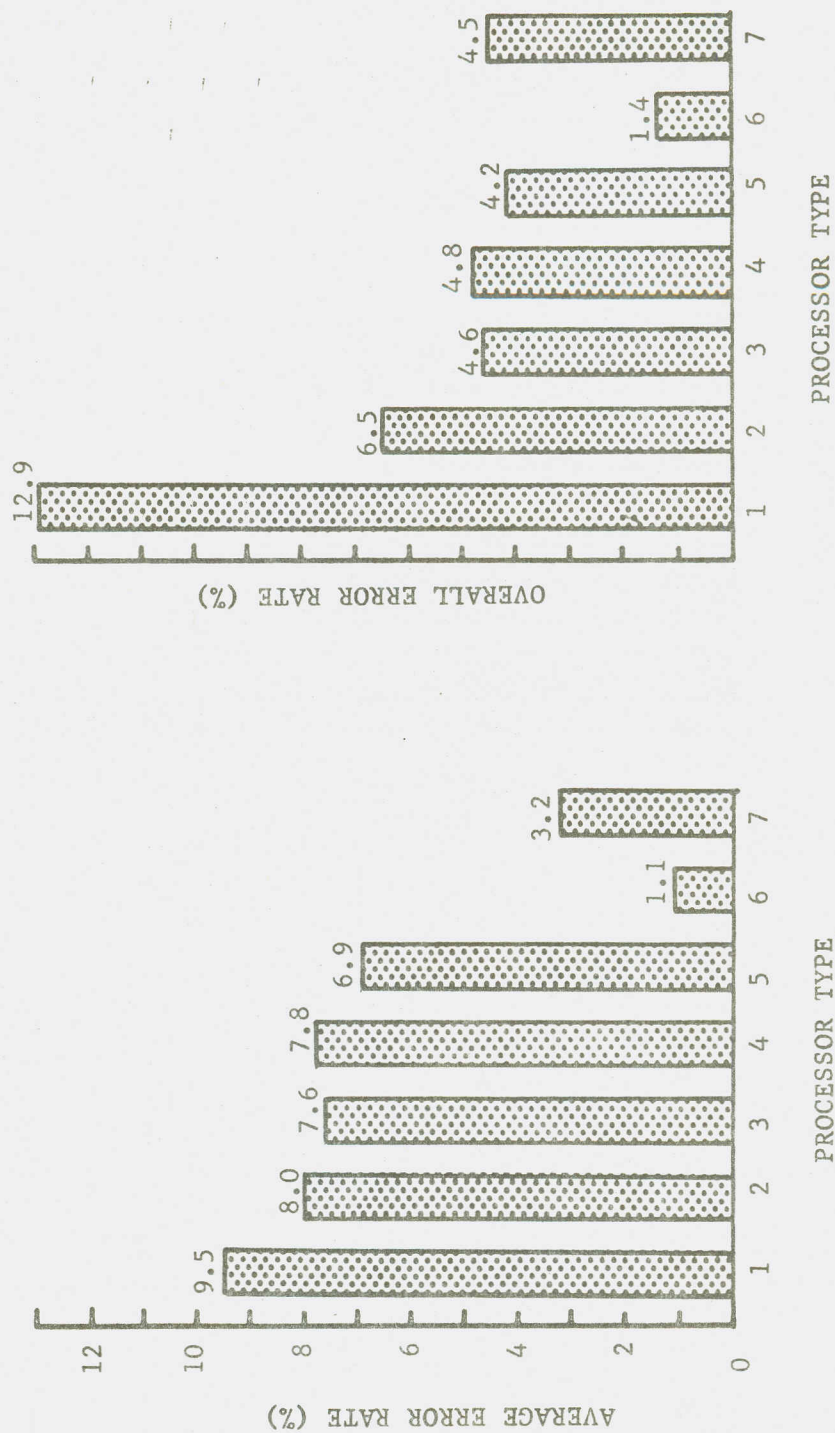


Figure 4.4.1 Classification Performance vs. Processing Scheme - Run 71052501

4.4

performance are given in Table 4.4.1. The original analyst has indicated that these are in roughly the same proportion as the actual occurrence of the classes in the data set. Thus the overall error rates are estimates of the actual probabilities of error. The class-conditional error rates are shown in Figure 4.4.2. Several observations can be made:

Observation 1

Again the average and overall error rates are significantly reduced by the ECHO techniques, with the supervised mode providing the greatest advantage.

Observation 2

This is the only data set analyzed to which the MV version of the unsupervised mode (processor #3) is applicable for the number of channels deemed necessary. Its performance can be described as erratic. It performed much better than CLASSIFYPOINTS for the deciduous class and much better than all processors (except SAMPLECLASSIFY) for the forage class. However it had an especially difficult time distinguishing between soybeans and corn.

Observation 3

With the exception of the coniferous class, the supervised mode (processor #5) again equaled or greatly improved upon the performance of CLASSIFYPOINTS. And the unsupervised, MUV mode did almost as well. In regard to the "apparently" poor coniferous performance, there are

Table 4.4.1 Relative Influence of Each Class on Overall Performance - Run 71052501

<u>Class</u>	<u>Percentage of Total Test Points</u>
Deciduous Forest	64.8
Forage	23.6
Corn	5.4
Soy	5.4
Water	0.7
Coniferous Forest	0.2

Table 4.4.2 Classes Ordered By Generalized Variance

<u>Class</u>	<u>Common Logarithm of Generalized Variance</u>	<u>Improvement Over CLASSIFYPOINTS</u>
Forage	8.02	+9.7 %
Deciduous Forest	6.06	+9.9 %
Water	6.01	0 %
Soy	5.78	+0.5 %
Corn	5.01	+0.3 %
Coniferous Forest	4.33	-4.6 %

PROCESSOR KEY

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, t=0
- #3 Optimum MV Unsupervised
- #4 Optimum MUV Unsupervised
- #5 Optimum Supervised
- #6 SAMPLECLASSIFY (ML)

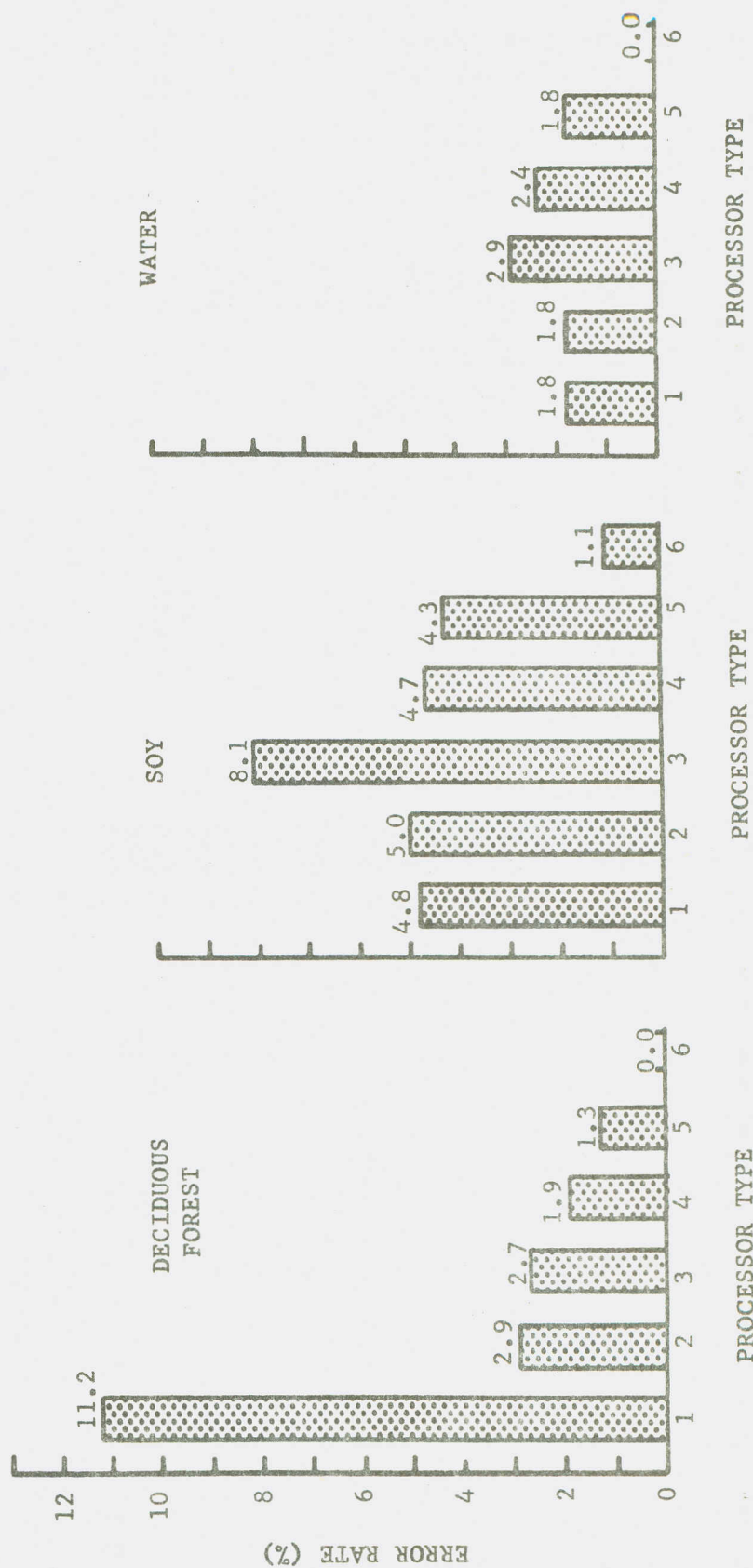


Figure 4.4.2 Performance By Class (Run 71052501)

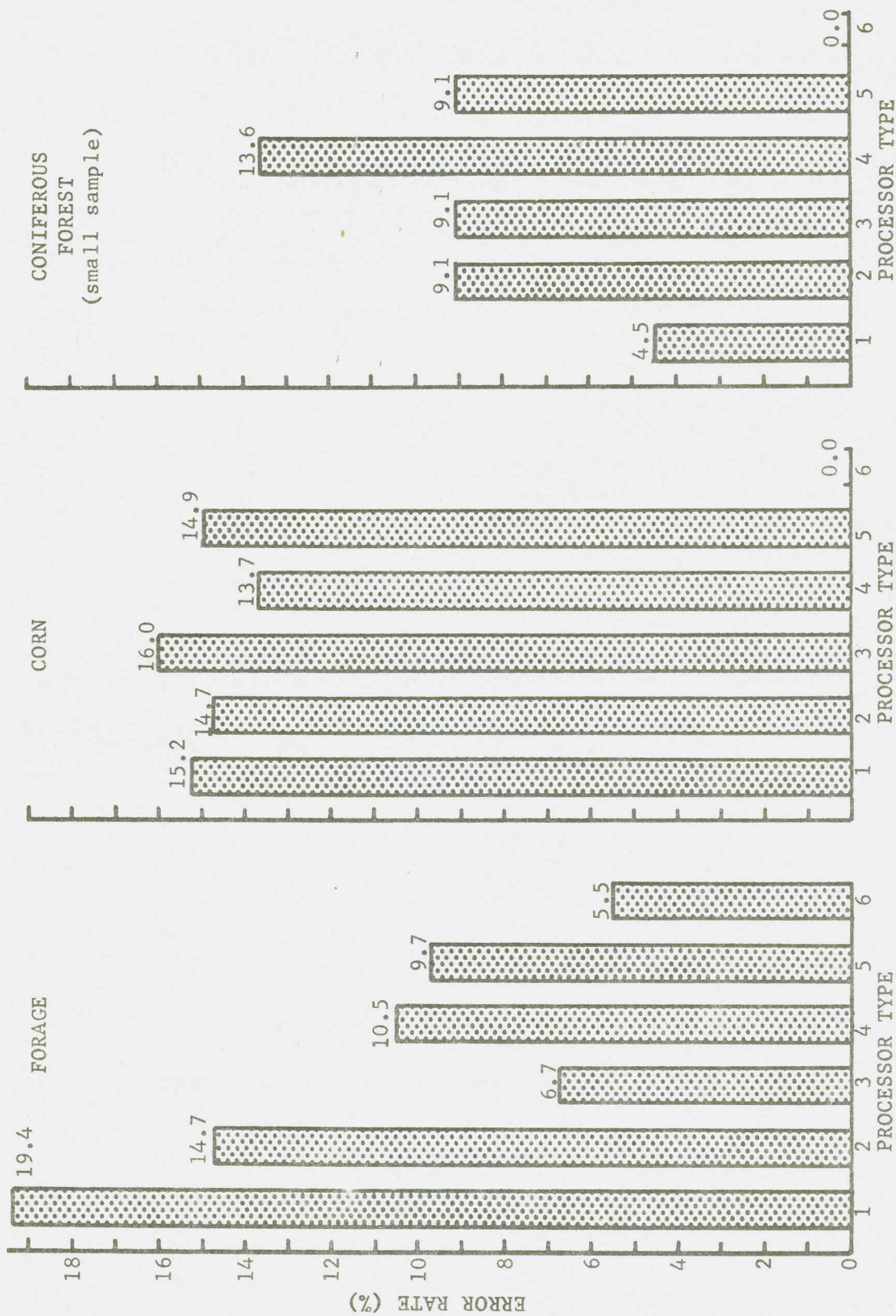


Figure 4.4.2, continued

4.4

extenuating circumstances to consider. It has already been mentioned that only a few small stands of white pine are known of in the entire data set. The original analyst used most of these pixels to ensure good training statistics, leaving only four areas, averaging only 22 pixels (or 5 1/2 cells) each, for testing purposes. With such an extremely small sample, a difference of 4.6% is entirely insignificant. It represents the misclassification of only 1 cell. Thus the coniferous results are included here only for completeness. No conclusion can be drawn from them.

Observation 4

The classes which benefit the most by the new techniques are deciduous forest and forage. Once again these are the classes with the largest generalized variances, as can be seen from Table 4.4.2. The correlation coefficient for this data is .81, which is coincidentally the same as for Figure 4.2.7.

The performance of the supervised mode is plotted for five values of the annexation threshold, t (Figures 4.4.3 and 4.4.4). The optimum performance occurs at the value $t=4$. At this value, the overall error rate is reduced from 12.9 percent to just 4.2 percent. For this study the Level-1 threshold was held at $c=60$, so again the results are not necessarily optimized with respect to c .

The performances of the unsupervised modes are given in Tables 4.4.3 through 4.4.6 for various combinations of

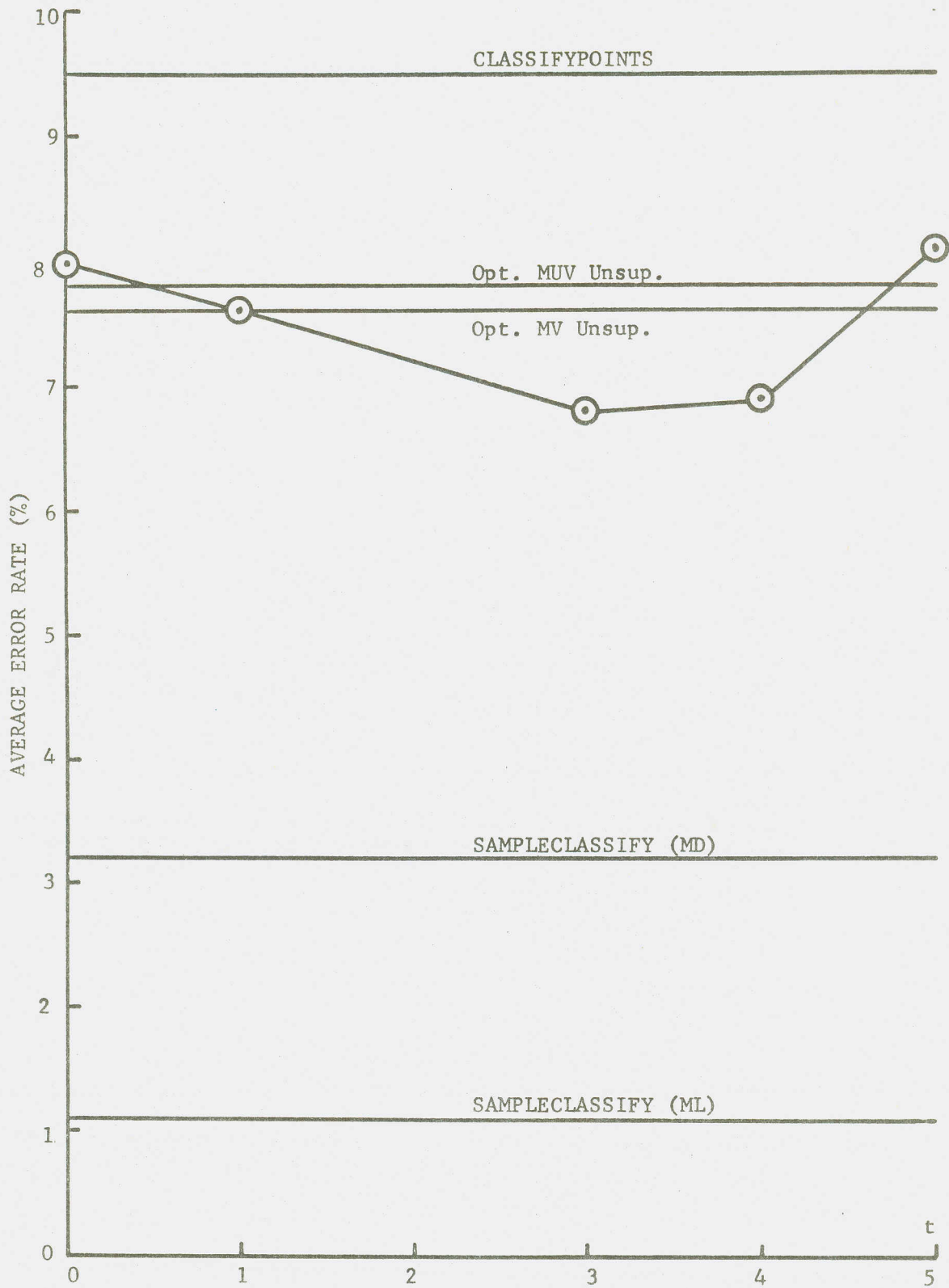


Figure 4.4.3 Effect of Annexation Threshold (t) on Average Performance
- Run 71052501

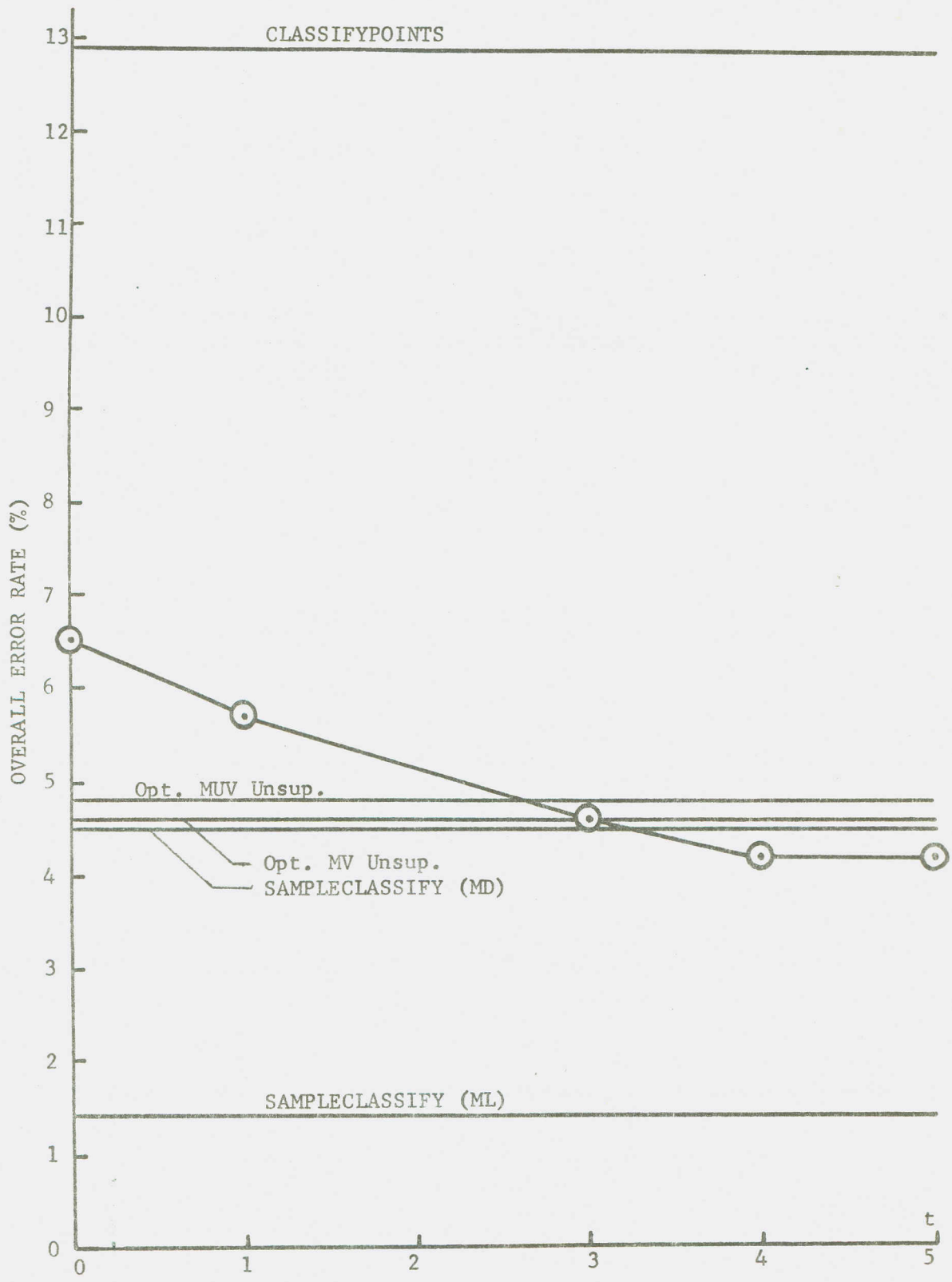


Figure 4.4.4 Effect of Annexation Threshold (t) on Overall Performance
 - Run 71052501

Table 4.4.3 Matrix of Average Error Rates (%) for Four Combinations of Significance Levels (MV mode)
- Run 71052501

$s_2 \backslash s_1$.025	.050	.100
.000	11.3	9.2 7.6*	11.6
.001	14.1		

Table 4.4.4 Matrix of Overall Error Rates (%) for Four Combinations of Significance Levels (MV mode)
- Run 71052501

$s_2 \backslash s_1$.025	.050	.100
.000	5.6	4.7 4.6*	5.9
.001	6.0		

* result of "cell-splitting"

Table 4.4.5 Matrix of Average Error Rates (%) for Six Combinations of Significance Levels (MUV mode)
- Run 71052501

$s_2 \backslash s_1$.005	.025	.100
.000	24.8	9.6	8.4 7.8*
.001	11.4	9.1	
.005		9.3	

Table 4.4.6 Matrix of Overall Error Rates (%) for Six Combinations of Significance Levels
- Run 71052501

$s_2 \backslash s_1$.005	.025	.100
.000	7.3	4.5	4.8 4.8*
.001	6.5	4.9	
.005		5.4	

* result of "cell-splitting"

4.4

significance levels. The study was done using a constant Level-1 threshold of 0.45. This unusually large value is a reflection of the narrow autocorrelation functions of the forest class (Appendix A). Cell splitting was found to reduce the average error rate some, but it has little or no effect on the overall error rate. The best results for the MUV version occur at about $s_2 = 0$ and $s_1 = .100$, while the MV version is best at $s_2 = 0$ and $s_1 = .050$.

Figure 4.4.5 shows how the processors compare with regard to both overall error rate and CPU time. The small number of classes and channels has reduced the time required to do a single classification to such a low point that a processor's speed depends on considerations other than the number of classifications that it must perform. Thus the speed advantage of the unsupervised modes has disappeared.

4.5 RUN 72032803 - Classification of Satellite Data

This LANDSAT-1 data set covers a 66.3 km long by 111 km wide region in Illinois on August 9, 1972. The region was sampled 837 times lengthwise and 1920 times along its width. The subregion analyzed is 40.3 km by 46.3 km and contains 406,400 pixels. This area was chosen because it contains a large set of reference data (15,067 pixels), corresponding to actual surface observations [28]. The 5 main classes are corn, soy, other-agriculture (including alfalfa, grass, oats, pasture, and wheat), woods, and town. Training data was obtained by sampling the test areas as indicated in

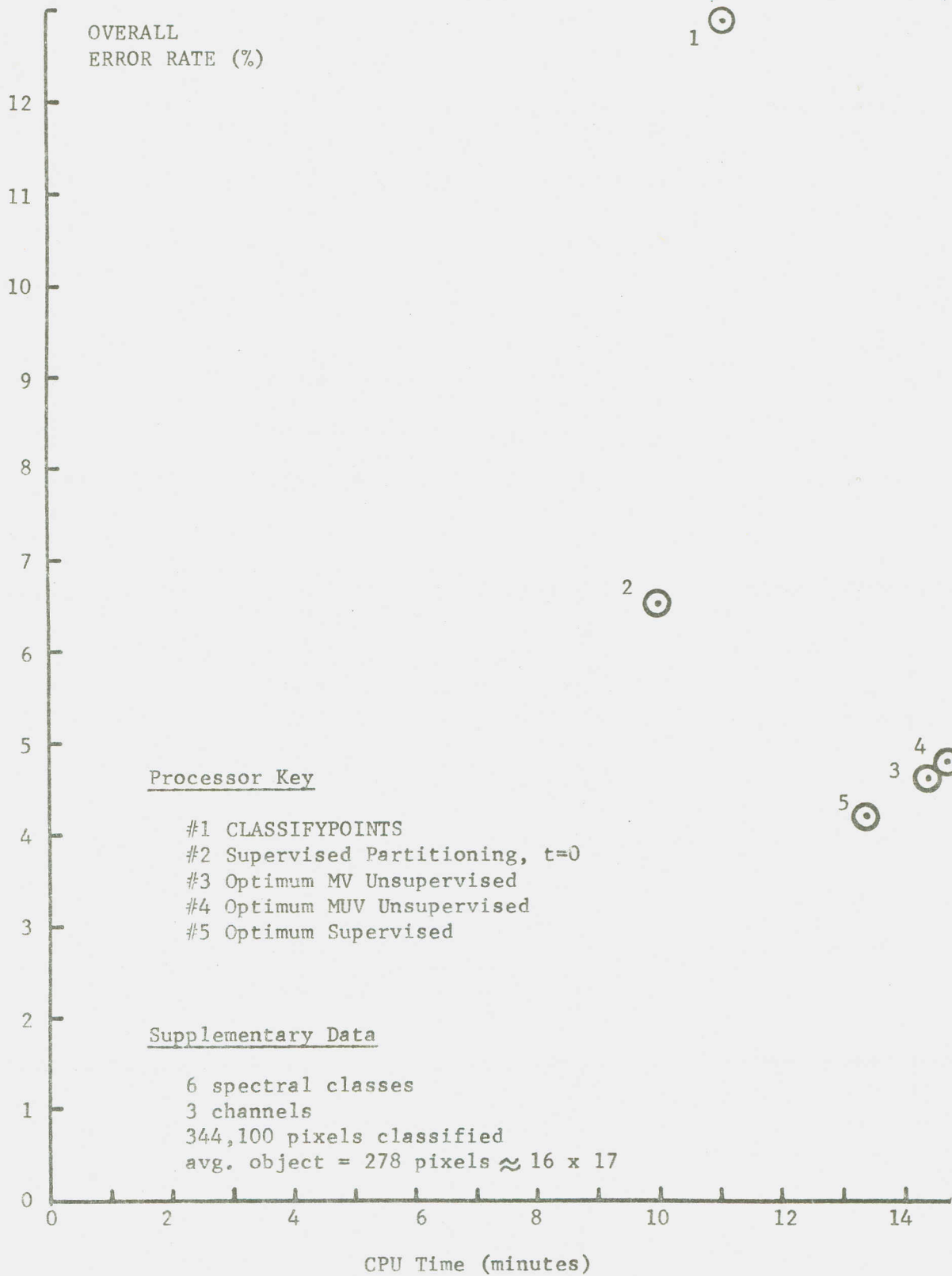


Figure 4.4.5 Error Rate and CPU Times for Five Classification Schemes
- Run 71052501

Section 4.1. The resultant ratio of test data to training data is about 5:1. Examination of the training data indicated that no subclasses were needed. All four channels were used in the analysis, because the separability of the classes is fairly low.

Figure 4.5.1 shows the average and overall performances achieved by the various classification schemes. The weights for overall error are given in Table 4.5.1. The class-conditional error rates appear in Figure 4.5.2. Several observations are worthy of mention:

Observation 1

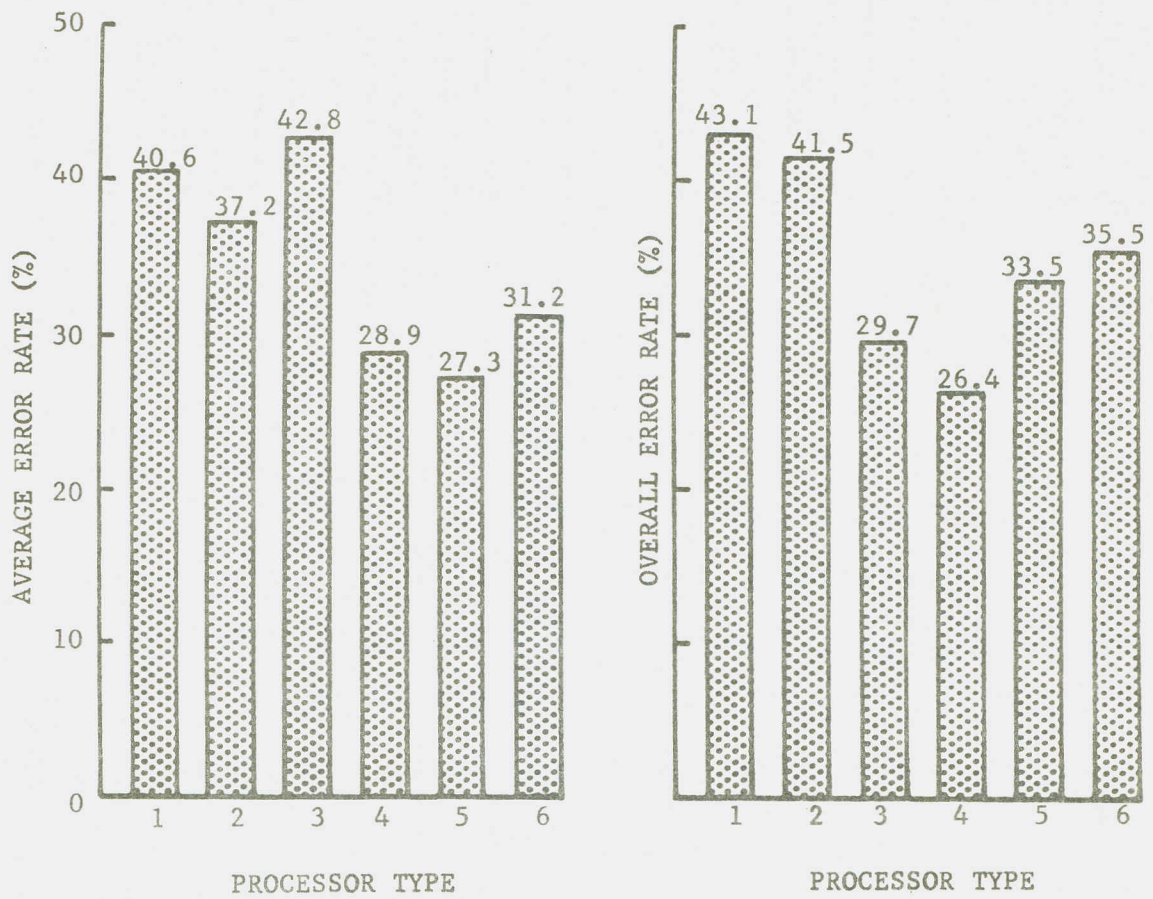
Once again supervised ECHO greatly reduced the average and overall error rates and performed as well or much better than CLASSIFYPOINTS on every class. On three of the five classes, both the supervised and unsupervised modes out-performed even SAMPLECLASSIFY (ML).

Observation 2

The unsupervised mode did very well in terms of overall error rate, but its average error rate is slightly greater than that of CLASSIFYPOINTS due to unusually poor performance in the "woods" class. The reason for this behavior is unknown, but it suggests that the unsupervised mode tends to be less stable than the supervised mode.

Observation 3

This is the only data set in which a "town" class has been included. We would anticipate a fairly broad



Processor Key

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, $t=0$
- #3 Optimum Unsupervised Partitioning
- #4 Optimum Supervised Partitioning
- #5 SAMPLECLASSIFY (ML)
- #6 SAMPLECLASSIFY (MD)

Figure 4.5.1 Classification Performance vs. Processing Scheme
- Run 72032803

Table 4.5.1 Relative Influence of Each Class on Overall Performance - Run 72032803

<u>Class</u>	<u>Percentage of Total Test Pixels</u>
Corn	67.1
Soy	18.8
Other-Ag.	10.2
Town	2.9
Woods	0.9

Table 4.5.2 Classes Ordered According to Generalized Variance - Run 72032803

<u>Class</u>	<u>Common Logarithm of Generalized Variance</u>	<u>Improvement (sup. mode) Over CLASSIFYPOINTS</u>
Other-Ag.	4.38	+ 14.0
Town	3.63	+ 21.7
Soy	3.50	+ 1.6
Corn	2.61	+ 21.3
Woods	2.27	0.0

PROCESSOR KEY

- #1 CLASSIFYPOINTS
- #2 Supervised Partitioning, t=0
- #3 Optimum Unsupervised
- #4 Optimum Supervised
- #5 SAMPLECLASSIFY (ML)

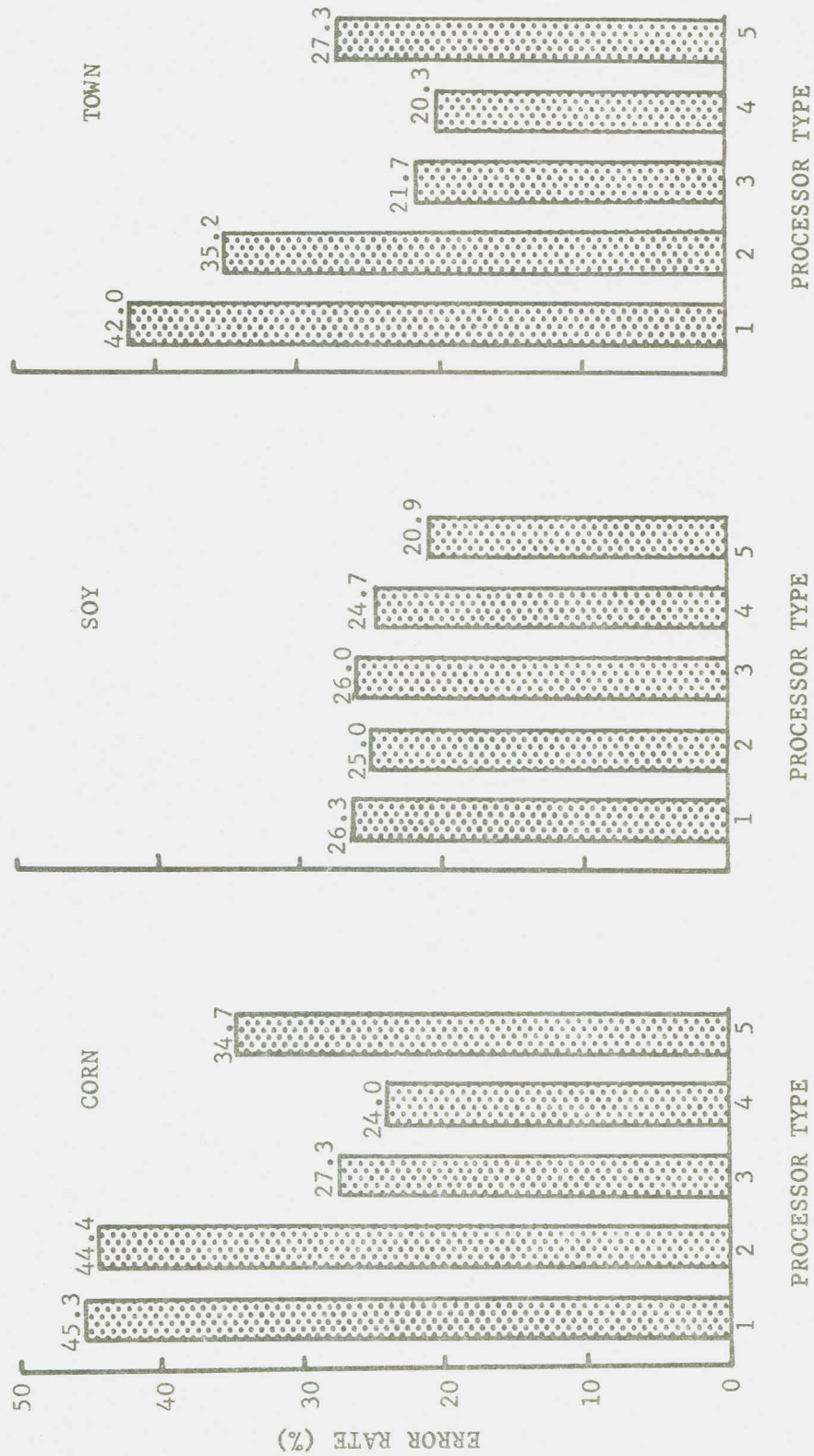


Figure 4.5.2 Performance By Class (Run 72032803)

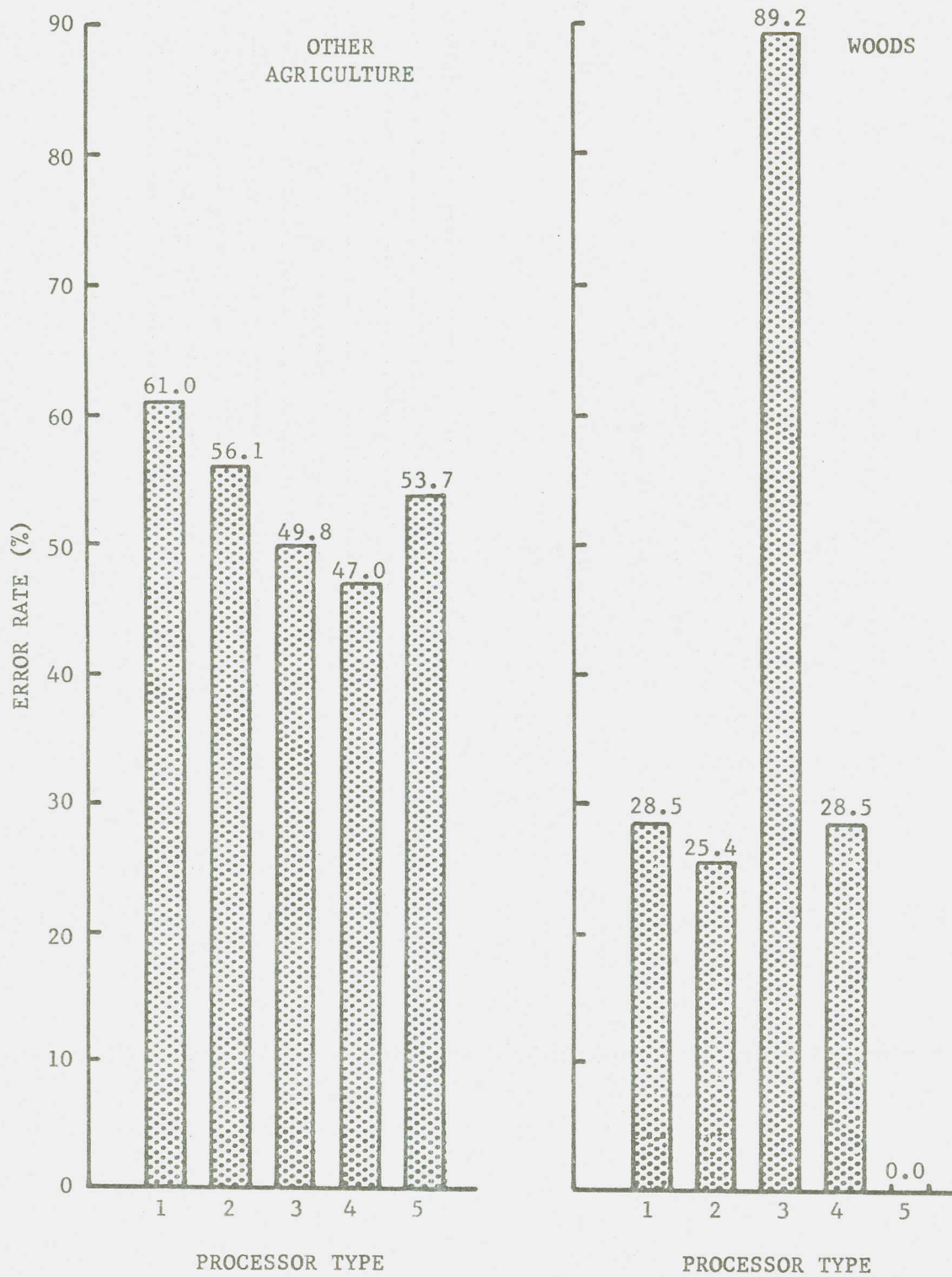


Figure 4.5.2, continued

distribution for this class and therefore a potentially large improvement in classification accuracy by the ECHO technique. The results bear this out. Table 4.5.2 shows that the town class ranks second in terms of generalized variance and first in terms of classification improvement. But unlike the aircraft data sets, the overall correlation between these quantities is fairly low.

The performance of the supervised mode is plotted vs. t in Figures 4.5.3 and 4.5.4. The best results occur at $t=3$ and $t=4$. The Level-1 threshold was held constant at $c=50$ for this study, and cell-splitting was not used. The effect of c was briefly investigated for $t=3$ by trying the values $c=30$ and $c=80$. Only class "other" was significantly affected. For $c=30$ its error rate increased 5.4% and for $c=80$ it increased just 0.3%. Thus performance again appears to be quite stable with respect to c .

Table 4.5.3 shows the average and overall error rates of the unsupervised mode for seven combinations of parameter values. Cell splitting was not used. The best performance occurs for $s_2=0$, $s_1=.005$ and $c=.25$ or $.20$. $c=.25$ produces the lowest overall error, and $c=.20$ produces the lowest average error. The value $.25$ was finally judged as "best" for this data for two reasons. (1) The value $.20$ produces an excessive number of singular cells. The value $.25$ produces a pattern of singular cells more closely resembling boundaries, as desired. (2) The high average error rate for

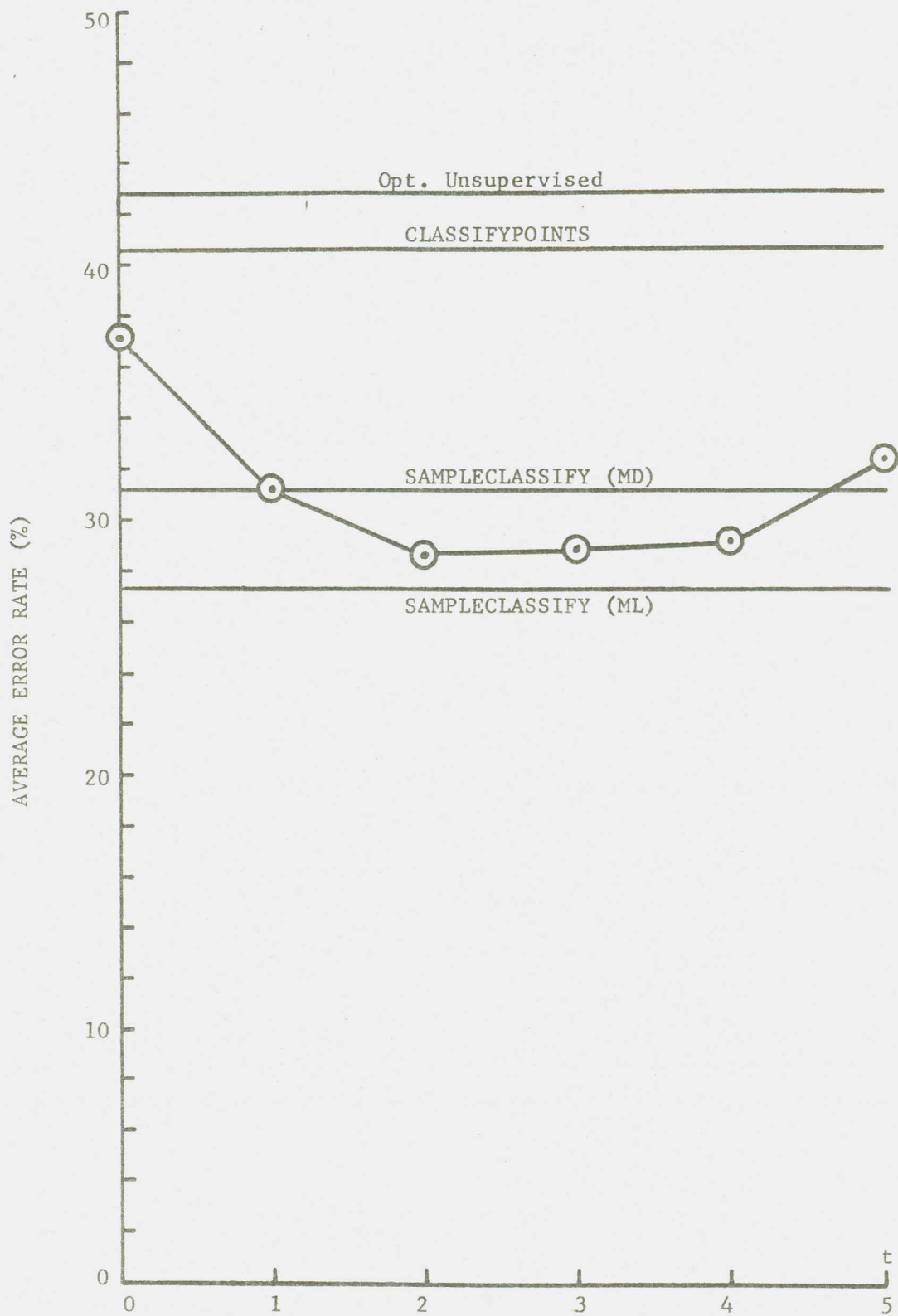


Figure 4.5.3 Effect of Annexation Threshold (t) on Average Performance
- Run 72032803

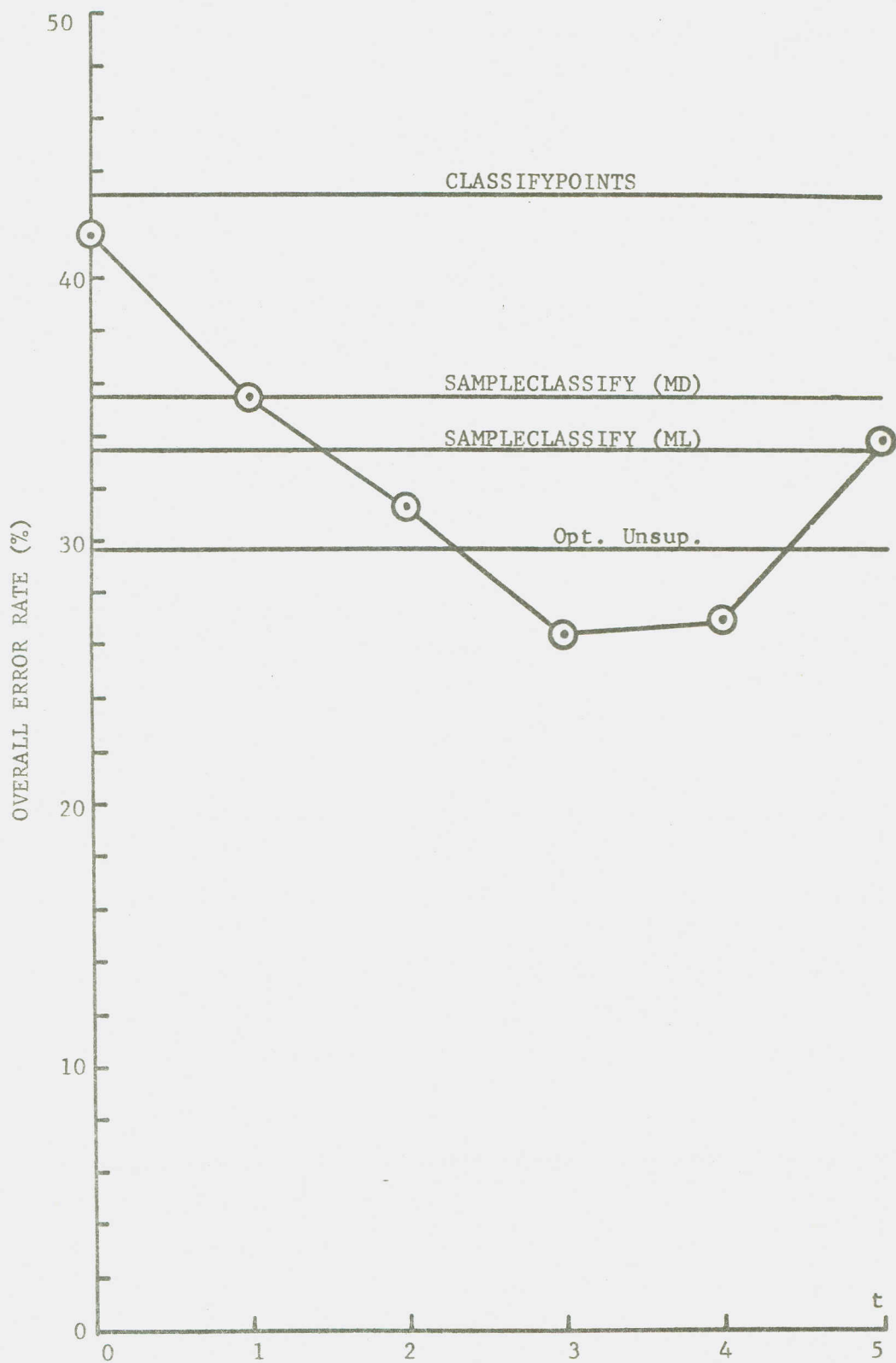


Figure 4.5.4 Effect of Annexation Threshold (t) on Overall Performance
- Run 72032803

Table 4.5.3 Effect of Parameters on Performance
(unsupervised mode)

<u>s₂</u>	<u>c</u>	<u>s₁</u>	<u>Avg. Error</u>	<u>Overall Error</u>
.001	.25	.005	40.3%	38.8%
0	.25	.001	36.2	35.5
0	.25	.005	42.8	29.7
0	.25	.025	41.7	36.1
0	.20	.001	42.5	32.3
0	.20	.005	35.2	32.1
0	.15	.005	38.9	35.2

$c=.25$ is due solely to the behavior of the minority class "woods". (All the other classes are at their optimum performances for this value.) This behavior is believed to be an anomalous effect attributable mainly to the small amount of test data available for woods (.9%). This belief is supported by the unusual observation that for $s_1=.025$, $.005$, and $.001$, the error rates for this class are 56.9%, 89.2%, and 48.5% respectively. A similar effect also occurred in the supervised mode. For $t=2, 3, 4$, and 5 , the error rates are 14.6%, 28.5%, 31.5%, and 9.2% respectively.

Figure 4.5.5 compares the overall error and CPU times of the various processors. In this case the partitioning schemes require more CPU time than the non-partitioning schemes. This can be attributed to the low number of spectral classes combined with a small number of pixels per physical object.

We also note that processor #2 performed only one-fourth as many actual classifications as #1 (since the cell size was 4), and yet they required the same CPU time. The same effect occurred on Run 71052501. This indicates a significant margin for improvement of the efficiency (speed) of the supervised mode. The cause of the inefficiency is that the ECHO processors are coded in FORTRAN, whereas the classification subroutine in CLASSIFYPOINTS is coded in assembler language for optimum efficiency.

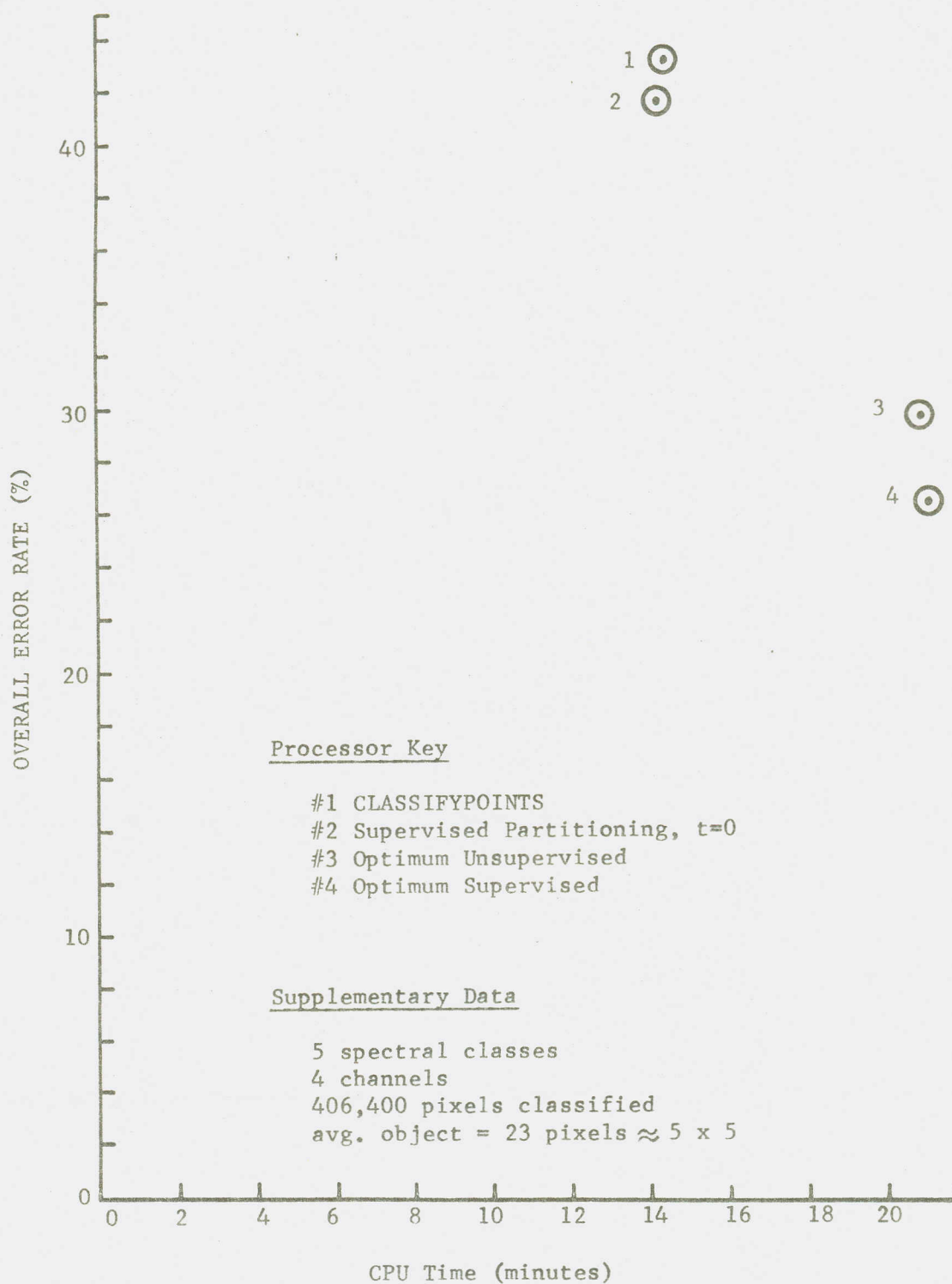


Figure 4.5.5 Error Rate and CPU Time for Four Classification Schemes
- Run 72032803

4.6 General Observations

Having studied four significantly different data sets, it is now possible to make some general observations. Strictly speaking, our observations (and therefore our conclusions) apply only to these data sets. But due to the consistency of the results, it is reasonable to expect other data sets to exhibit similar behavior as long as they are reasonably similar to these. This assumes, of course, that representative training data is available and that training is done in accordance with the assumptions of the model.

Observation 1

First of all it is apparent that the new techniques are effective for a wide range of variables including classification parameters (number of classes, subclasses, channels, etc.) and data parameters (spatial, spectral, and measurement resolution and spatial correlation).

Observation 2

Equally important is the stability of the performance with respect to the processor's input parameters. For the supervised mode the performance is very stable. The main input parameter is the annexation threshold, t . On all four data sets all values tried gave better performance than CLASSIFYPOINTS. Furthermore the value $t=4$ gave optimal, or near optimal, performance in every case. There is some evidence to suggest that as the number of channels increases, so does the optimum value of t .

For the unsupervised MUV mode the main input parameters are the significance levels s_1 and s_2 , in that order. The results seem to be fairly stable with respect to them, but their optimum values tend to be elusive. For the four data sets the optimum value of s_1 varies among .001, .005, and .100. s_2 is usually effective at around .001 or 0. Presently there is no way to predict the optimum combination, which means that in order to use this mode one will generally have to settle for suboptimal results. In most cases however, the results are still somewhat better than CLASSIFYPOINTS.

A secondary parameter is the cell selection threshold c . A suitable value can generally be obtained by processing a small subregion of the data set with different thresholds until one is found which produces a pattern of singular cells that resembles object boundaries (as opposed to random noise). In the unsupervised mode $c=.25$ most frequently produced this effect, but when spatial correlation was low, a larger value was needed. In the supervised mode, c depends on the number of channels, q . As a rule of thumb, $c=15q$ appears to be a reasonable empirical guideline (at least for $3 \leq q \leq 6$). But a larger value ($c=20q$) was required when the spatial correlation was low.

Observation 3

Another type of stability is the sensitivity of the processor to the particular characteristics of a class. For

example, CLASSIFYPOINTS is sensitive to the variance of the class pdf. It tends to become "unstable" when a class with a relatively broad distribution is encountered. Since a single pixel contains no information about the variance of its class, CLASSIFYPOINTS is better suited for distinguishing classes that differ in mean rather than in variance.

The partitioning schemes, on the other hand, use both the mean and variance of the data in an object in order to classify it. Consequently they tend to be much more stable in this respect. But in return they should be somewhat sensitive to the typical size of the objects of a particular class. One would expect them to become unstable as the size of the objects approaches the cell size. Some evidence of this was detected (e.g. pit class), but the problem was avoided for the most part by choosing a relatively small cell width (2 pixels). If the object size approaches this value, then Premise A is violated and CLASSIFYPOINTS also becomes unstable, as observed for the pit class.

Observation 4

The main advantage of the unsupervised mode appears to be speed when classification time, as opposed to partitioning time, is the limiting factor. In other words, when the classification is relatively complex, involving many spectral classes and many channels, one can save a significant amount of time by extracting and classifying

4.6

objects rather than classifying on a point-by-point basis. This is particularly true when the number of classes is large, because the time required for unsupervised partitioning is independent of the number of classes. The largest number of spectral classes considered in this investigation is 17, for which the unsupervised mode required less than half the time that CLASSIFYPOINTS required. It is not uncommon to see analysts using many more than this, a trend attributable to the increasing size of areas classified (encompassing more classes) and to the increasing use of unsupervised clustering techniques for subclass definition.

It seems advisable at this point to inject a word of caution against the use of clustering beyond that which is necessary to achieve reasonably unimodal training classes. "Creating" more subclasses reduces the number of pixels available per subclass for training. Thus the chance that each subclass is representative of an actual spectral class in the data set is reduced. Also the narrow distributions that result cause the classifier performance to be more sensitive to non-stationary class statistics over a large area. Finally, as the number of subclasses increases, the effective size of the objects in the scene decreases, thus reducing the potential advantage of a sample classifier.

Observation 5

The MV version of the unsupervised mode is not generally

applicable to all situations because of the relation between cell size and number of channels. In one case in which it was applicable it performed erratically, but its overall performance was as good as the MUV version. Although cases probably exist for which the MV version provides a significant advantage over the MUV version, none have yet been encountered.

Observation 6

The supervised mode consistently provided the lowest average and overall error rates (excluding SAMPLECLASSIFY), and for complex classifications it is more efficient than CLASSIFYPOINTS. (This is in spite of the fact that it is programmed considerably less efficiently than CLASSIFYPOINTS.) For the four data sets studied, the average reduction in class-average and overall error rates from the rates provided by CLASSIFYPOINTS is 7.1% and 9.6% respectively. For comparison, the corresponding values for SAMPLECLASSIFY (ML) are 10.8% and 9.9%. For SAMPLECLASSIFY (MD) they are 9.7% and 8.9%. Thus we conclude that the ML strategy is an effective means of sample classification for MSS data, and that supervised partitioning is an effective means of applying it when the true partition is unknown.

CHAPTER 5

CONCLUSION

5.1 Summary

The general problem that we have investigated is the application of sample classification techniques to MSS data. The purpose of this is to incorporate "memory" into the classification process, thereby improving performance and reducing the number of items that must be individually classified. To begin, we modeled the objects in the scene as simple samples from multivariate normal populations. Then we motivated the investigation by showing that the classification scheme which achieves the minimum error rate when object boundaries are known, is a particular type of sample classification (MAP). Furthermore, the closely related strategy (GML) that we use in practice has an error rate that is upper bounded by a sum of exponentially decreasing functions of the sample size.

When object boundaries are unknown, the technique that is commonly used is to classify each resolution element independently of the others (no-memory classification). This of course is suboptimal, because spatially adjacent states of nature are usually strongly dependent. However an

optimal strategy (Appendix B) would require a complete statistical description of this dependence as well as a large number of classifications and a large amount of computation per classification. As an intermediate approach we chose to exploit the particular nature of the dependence by applying an image partitioning transformation to the data set prior to actual classification. In particular we focused our attention on what we call the conjunctive object-seeking approach. The basic algorithm that we implemented requires two "levels" of statistical tests that are applied in a logical sequence in order to "merge" adjacent elements of the scene that are spectrally similar. The likelihood ratio procedure led us to multivariate tests of first and second order statistics using criteria whose distribution functions are known under the null hypothesis. This enables us to relate the size of the test to the decision threshold. The power of the test depends on the alternative hypothesis. Due to the "dimensionality problem" the "multiple-univariate" versions of these tests are actually more useful in practice.

The likelihood ratio procedure can also be adapted to provide a supervised mode of operation. The test statistic is more complex, but it is multivariate and yet free of the dimensionality problem. Experimental results indicate that it adds a significant measure of stability to the processor's performance. Consequently it consistently

5.1

provided the lowest class-average and overall error rates, approaching those attained by direct sample classification of the test areas. Also, visual inspection of the results indicated that the classification map produced by this method is much closer to the "type-map" form that is usually desired than is the map produced by no-memory classification.

In terms of efficiency, the partitioning schemes varied from better to worse than the no-memory classifier to which they were compared, depending on the complexity of the classification. But given comparable programming efficiency, it appears that this balance would be shifted significantly in favor of the partitioning approach.

5.2 Recommendations for Further Work

Parameter Selection

One would expect that the optimum values of input parameters such as s_1 , s_2 , and t are statistically dependent in some complex way on factors such as class separability, spatial correlation, number of channels, average object size, and cell size. Therefore it could be beneficial to investigate the possible use of such information to predict the optimum input parameters. As we have noted, this appears to be needed more in the unsupervised mode than in the supervised mode.

Use of Texture

In this investigation we have restricted ourselves to

5.2

the assumption that pixels are class-conditionally uncorrelated. However in Appendix A we have shown not only that they are correlated but that the correlation function is class-dependent. We have also indicated how a sample classifier could be designed to efficiently exploit this dependence for improved discrimination between classes. This effect could be investigated through direct sample classification of test areas, similar to the investigation done by Wacker and Landgrebe. If the degree of improvement proves to be significant, then it is likely that the performance of the partitioning schemes can also be significantly improved by redesigning them to exploit spatial correlation.

LIST OF REFERENCES

1. Abend, K., "Compound Decision Procedures for Pattern Recognition", Proc. NEC, 22, pp. 777-780, 1966.
2. Robertson, T.V., K.S. Fu, P.H. Swain, "Multispectral Image Partitioning", Ph.D. Thesis #25970, School of Electrical Engineering, Purdue University, West Lafayette, In., August, 1973.
3. Pingle, K.K., "Visual Perception By A Computer", Automatic Interpretation and Classification of Images, A. Grasselli (ed.), Academic Press, N.Y., pp. 277-284, 1969.
4. Anuta, P.E., "Spatial Registration of Multispectral and Multitemporal Imagery Using Fast Fourier Transform Techniques", IEEE Trans. Geoscience Electronics, Vol. GE-8, No. 4, pp. 353-368, Oct. 1970.
5. Prewitt, J.M.S., "Object Enhancement and Extraction", Picture Processing and Psychopictorics, B.S. Lipkin and A. Rosenfeld (eds.), Academic Press, N.Y., pp. 113-114, 1970.
6. Hueckel, M.H., "An Operator Which Locates Edges in Digital Pictures", Journal ACM, Vol. 18, No. 1, pp. 113-125, Jan. 1971.
7. Wacker, A.G. and D.A. Landgrebe, "Boundaries in Multispectral Imagery By Clustering", IEEE Symposium on Adaptive Processes, Decision, and Control (9th), Univ. of Texas (Austin), Dec. 1970.
8. Rosenfeld, A., M. Thurston, Y.H. Lee, "Edge and Curve Detection: Further Experiments", IEEE Trans. Computers, Vol. C-21, No. 7, pp. 677-715, July 1972.
9. Griffith, A.K., "Edge Detection in Simple Scenes Using A-Priori Information", IEEE Trans. Computers, Vol. C-22, No. 4, April 1973.

10. Muerle, J.L. and D.C. Allen, "Experimental Evaluation of Techniques for Automatic Segmentation of Objects in a Complex Scene", Pictorial Pattern Recognition, G.C. Cheng, R.S. Ledley, D.K. Pollock, A. Rosenfeld (eds.), Thompson Book Co., Washington, D.C., pp. 3-13, 1968.
11. Rodd, E.M., "Closed Boundary Field Selection in Multi-spectral Digital Images", IBM Publication No. 320.2420, Jan. 1972.
12. Robertson, T.V., "Extraction and Classification of Objects in Multispectral Images", Proceedings of the Conference on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, In., Section 3B, pp. 27-34, Oct. 1973.
13. Huang, T., "Per Field Classifier for Agricultural Applications", LARS Information Note 060569, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., June 1969.
14. Wacker, A.G. and D.A. Landgrebe, "The Minimum Distance Approach to Classification", Ph.D. Thesis #24733, School of Electrical Engineering, Purdue University, West Lafayette, In., Jan. 1972.
15. Kettig, R.L. and D.A. Landgrebe, "Automatic Boundary Finding and Sample Classification of Remotely Sensed Multispectral Data", LARS Information Note 041773, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., April 1973.
16. Gupta, J.N. and P.A. Wintz, "Closed Boundary Finding, Feature Selection, and Classification Approach to Multi-Image Modeling", LARS Information Note 062773, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., June 1973.
17. Crane, R.B., W.A. Malila, W. Richardson, "Suitability of the Normal Density Assumption for Processing Multispectral Scanner Data", IEEE Trans. Geoscience Electronics, Vol. GE-10, No. 4, pp. 158-165, Oct. 1972.
18. Fukunaga, K., Introduction to Statistical Pattern Recognition, Academic Press, N.Y., 1972.
19. Kailath, T., "The Divergence and Bhattacharyya Distance Measures in Signal Selection", IEEE Trans. Communication Technology, Vol. COM-15, No. 1, pp. 52-60, Feb. 1967.
20. Lehmann, E.L., Testing Statistical Hypotheses, Wiley & Sons Inc., N.Y., 1959.

21. Van Trees, H.L., Detection, Estimation, and Modulation Theory, Part 1, Wiley & Sons Inc., N.Y., 1968.
22. Anderson, T.W., An Introduction to Multivariate Statistical Analysis, Wiley & Sons Inc., N.Y., 1958.
23. Cooley, W.W. and P.R. Lohnes, Multivariate Data Analysis, Wiley & Sons Inc., N.Y., 1971.
24. Duda, R.O. and P.E. Hart, Pattern Classification and Scene Analysis, Wiley & Sons Inc., N.Y., 1973.
25. "1971 Corn Blight Watch Experiment, Final Report", National Aeronautics and Space Administration, Lyndon B. Johnson Space Center, Houston, Texas, June 1973.
26. Phillips, T.L. (ed.), "LARSYS Version 3 User's Manual", Vol. 2, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., June 1973.
27. Coggeshall, M.E. and R.M. Hoffer, "Basic Forest Cover Mapping Using Digitized Remote Sensor Data and Automatic Data Processing Techniques", LARS Information Note 030573, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., March 1973.
28. Bauer, M.E. and J.E. Cipra, "Identification of Agricultural Crops By Computer Processing of ERTS [LANDSAT] MSS Data", LARS Information Note 030173, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., March 1973.

General References

29. Abend, K., T.J. Harley, L.N. Kanal, "Classification of Binary Random Patterns", IEEE Trans. Information Theory, IT-11, No. 4, pp. 538-544, Oct. 1965.
30. Fisz, M., Probability Theory and Mathematical Statistics, Wiley & Sons Inc., N.Y., 1963.
31. Rulon, P.J., and W.D. Brooks, "On Statistical Tests of Group Differences", Handbook of Measurement and Assessment in Behavioral Sciences, Ch. 2, D.K. Whitla (ed.), Reading, Mass., Addison-Wesley, 1968.
32. Swain, P.H., "Pattern Recognition: A Basis for Remote Sensing Data Analysis", LARS Information Note 111572, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, In., Nov. 1972.

APPENDIX A
SPATIAL CORRELATION

In Chapter 2 we assumed, for simplicity, that pixels within the same object are statistically independent observations from some subclass population. Then the joint pdf of the pixels can be expressed in terms of the marginal pdf of a single pixel. A more general approach is to allow for some correlation to exist between each pixel (\underline{x} , say) and the other pixels in the object that lie within some "neighborhood" of \underline{x} . Spatial correlation can be inherent in the object, and it can also be induced by the scanner. For example, a line scanner induces a certain amount of correlation along each scan line because its bandwidth is constrained to reduce detector noise. Also, one commonly finds that adjacent pixels actually overlap. In LANDSAT-1 data, for example, the overlap is 29%. Due to the position invariance of these effects and to the homogeneity of the objects, it is reasonable to assume that the class-conditional, joint pdf of the pixels in the neighborhood of \underline{x} is invariant with respect to the position of \underline{x} within the object (as we previously did for the marginal pdf of \underline{x} alone). This of course neglects any

non-stationary effects of the scanner such as banding, variable sun-angle, non-uniform sampling, and d.c. drift, much of which can, in principle, be corrected for in the data. Another reasonable assumption for most types of objects is "transpose symmetry"; i.e. the joint pdf of a spatial array of pixels is invariant with respect to the row or column transpose of the array. A possible consequence of this is discussed later.

Assuming (as before) that the pixels in an object are jointly MVN, then all that is required to specify their pdf is a mean vector and covariance function matrix (i.e. interchannel covariance matrix as a function of displacement). If $\underline{X}(m,n)$ represents the data vector for the scene element whose spatial coordinates are m and n (line and column numbers), then we can estimate this matrix for a given object by computing the following average:

$$\begin{aligned} \underline{C}(m,n) &= \frac{1}{N} \sum_{k,l} (\underline{X}(m+k,n+l) - \underline{M}(m,n)) (\underline{X}(k,l) - \underline{M}(0,0))' & A1 \\ &= \frac{1}{N} \sum_{k,l} \underline{X}(m+k,n+l) \underline{X}'(k,l) - \underline{M}(m,n) \underline{M}'(0,0) \end{aligned}$$

where

$$\underline{M}(m,n) = \frac{1}{N} \sum_{k,l} \underline{X}(m+k,n+l)$$

and N is the number of terms in the summation. To measure "local" characteristics (e.g. for a particular spectral class), the summations include only pixels from a single object. To measure the characteristics of a larger region, the summations extend over that region. These measurements

have been made for several data sets, and some typical results are presented in Figures A1-A5 (aircraft data) and Figures A6-A10 (satellite data). The quantity $R_{ij}(m,n)$ which appears in these figures is the correlation coefficient, which is related to the covariance by

$$R_{ij}(m,n) = C_{ij}(m,n) / \sqrt{C_{ii}(0,0) C_{jj}(0,0)} \quad A2$$

where $C_{ij}(m,n)$ is the i,j th element of matrix $C(m,n)$. This normalization provides easier comparison of the various functions. Local correlation was measured within the test areas to ensure that only pixels from the same object were used. The results were averaged over all the test areas for a given class to obtain the final estimate for that class. Regional correlation is generally greater than local correlation for a given displacement, because it includes the effect of dependent states as well as the class-conditional (local) correlation effects.

The measurements indicate that intraclass spatial correlation is a significant effect. Naturally the strongest correlation occurs between adjacent pixels, and the effect diminishes rapidly to a fairly low level. The "knee" in the curve generally occurs at a displacement of about 2 pixels. An important point is illustrated by Figures A2 and A3, which compare the classes "deciduous forest" and "forage". We observe a definite class dependency for the spatial correlation functions. The relatively narrow correlation functions for forest indicate

a broader spatial frequency spectrum, thus a faster rate of change, than the forage class. The implication is that this characteristic can be measured and used by a sample classifier to distinguish these two classes. Thus the "texture" of a sample contains potentially useful information about its identity. This comes as no surprise, but it is not always obvious how to exploit such information in a numerically-oriented pattern recognition system. When spectral information is also available and both are observed in a multivariate measurement space, even the human system may be unable to use all the information contained in the sample. We shall now briefly discuss how the classifiers described in Chapter 2 can be generalized to accomplish such classification.

To design a true MAP or ML classifier for a spatially correlated sample, $X = (X_1, \dots, X_n)$, is conceptually a straightforward matter, given the mean vector and covariance function matrix of each class. These can be used to construct the $n \times n$ -dimensional matrix:

$$Q_i = E(Q_i(X) | X \in W_i) \quad A3$$

where

$$Q_i(X) = ((X_1 - M_i)', \dots, (X_n - M_i)')' ((X_1 - M_i)', \dots, (X_n - M_i)')$$

bearing in mind the spatial arrangement of the pixels. Then for the hypothesis $X \in W_i$, the log-likelihood function evaluated at X is given by

$$p(X | X \in W_i) = -.5(\ln |2\pi Q_i| + \text{tr}(Q_i^{-1} Q_i(X))) \quad A4$$

For large samples, this rapidly becomes too cumbersome to be practical. Also, a covariance matrix would have to be constructed for each class for each unique spatial arrangement of pixels to be classified.

For large samples, with relatively small correlation distances (small neighborhoods), a more practical approach would be to implement a minimum distance decision rule based on just the marginal joint pdf of a neighborhood, which contains all the information that characterizes a class population. In fact, assuming transpose symmetry holds, this information is contained in the marginal, joint pdf of just one quadrant of the neighborhood. To estimate this pdf from a sample, we estimate the mean vector and covariance function matrix according to formula A1. But we construct only the covariance matrix of an array of pixels corresponding to one quadrant of a neighborhood, thereby avoiding both problems associated with the maximum likelihood approach.

We note that when assuming transpose symmetry, the covariance matrix estimate can be improved for any given displacements, (m,n) , by averaging $\hat{C}(m,n)$ and $\hat{C}(-m,n)$ to obtain the final estimate.

Also, the above strategy can be easily modified (for simplicity) by truncating the tail of the intraclass correlation function as desired, thereby reducing the dimension of the required covariance matrix. Since the knee

in the correlation function generally occurs at about 2 pixels, this might be a reasonable correlation distance to use.

Finally, the choice of distance measure is arbitrary. However, an interesting possibility is the use of $-L_i(\underline{M}, \underline{C})$ (from Section 2.3.4), appropriately modified for the higher dimensional space. We have seen that it is essentially the same as using a Kullback-Leibler number, but computationally it is much more efficient. Due to its relationship to the likelihood function and its lack of distance measure properties, a strategy that uses this criterion can perhaps best be described as a modified maximum likelihood strategy (modified to avoid an nq -dimensional matrix).

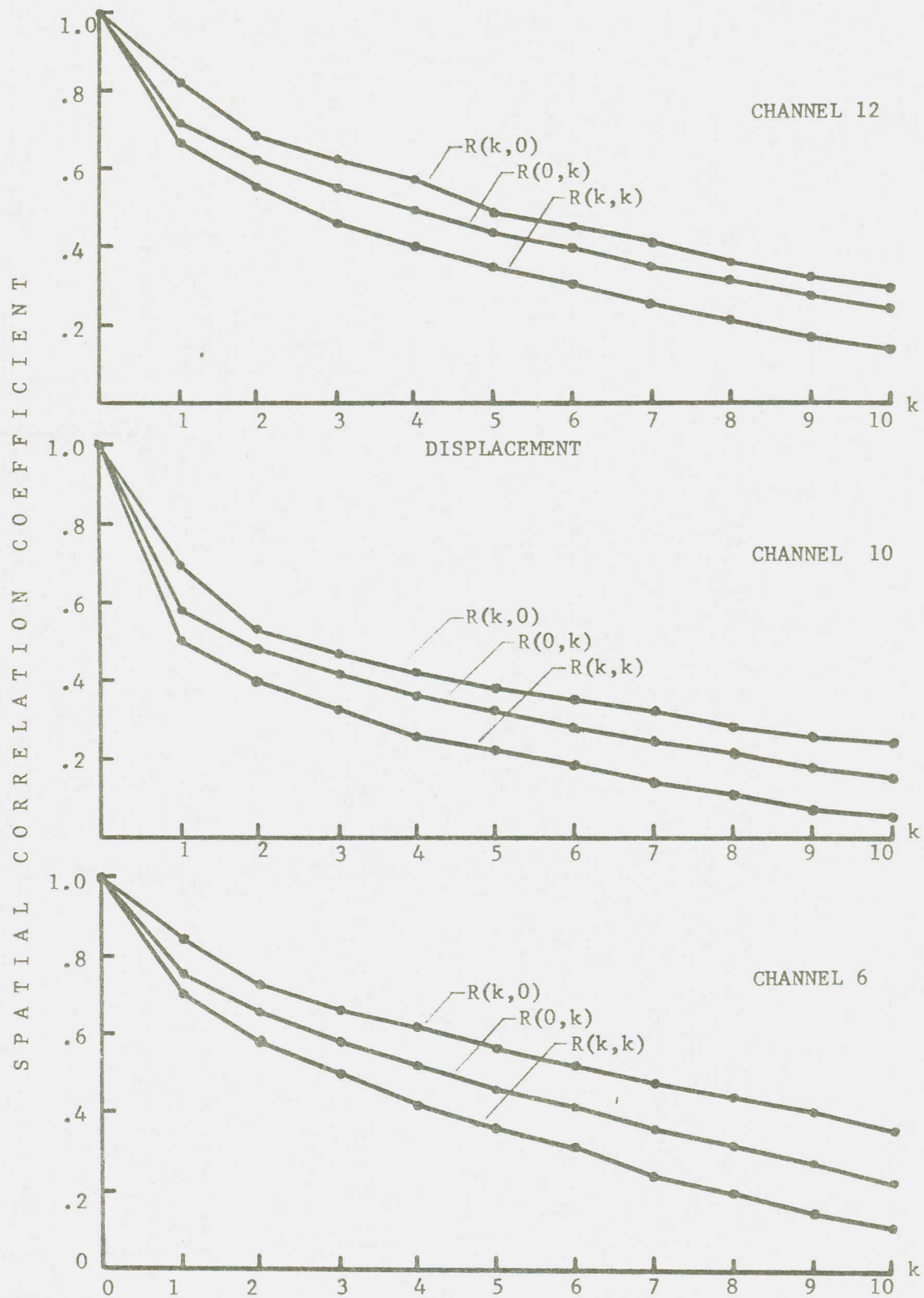


Figure A1 Spatial Correlation Coefficients - Run 71052501
Region: lines 501-800, columns 1-100

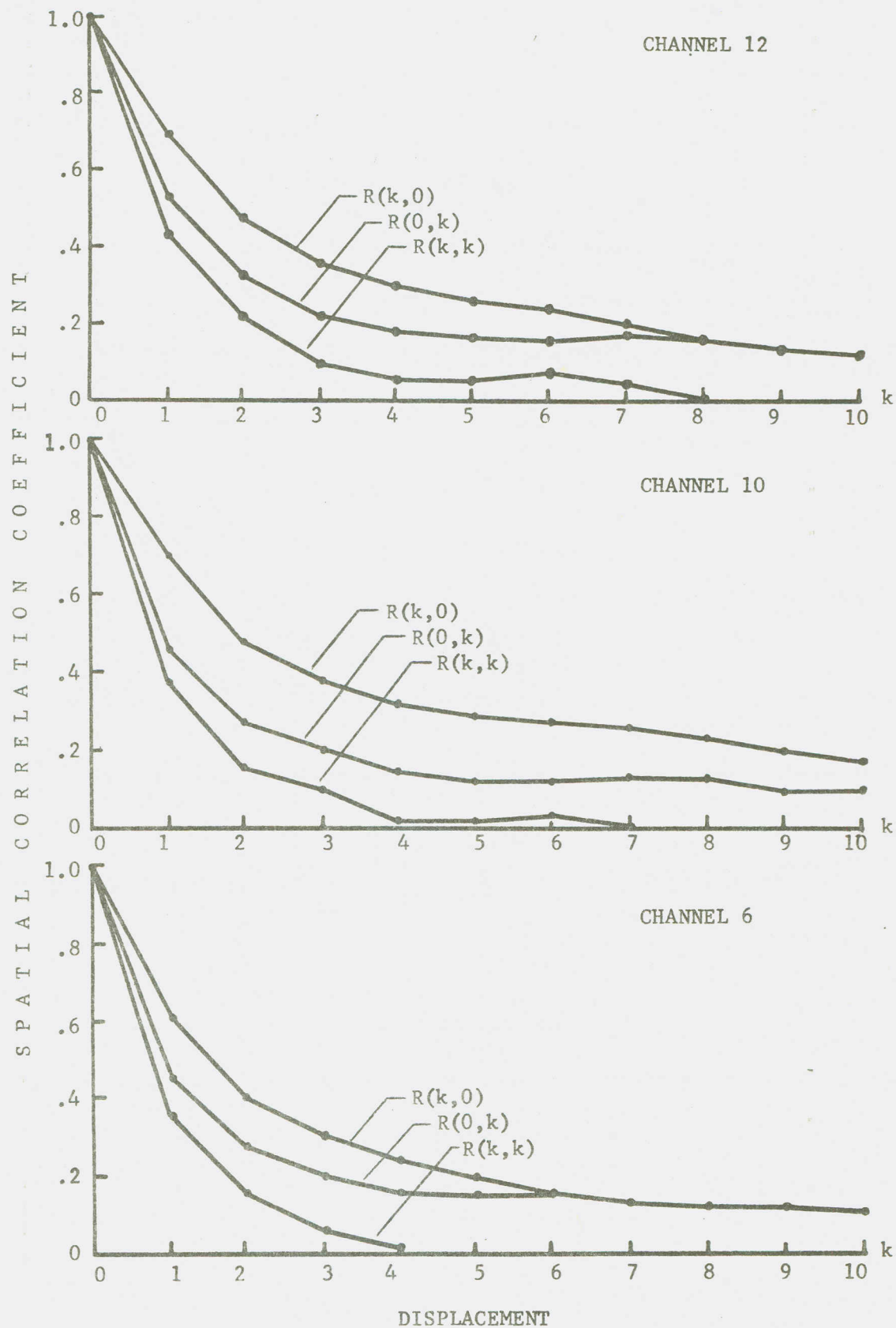


Figure A2 Spatial Correlation Coefficients - Class FORAGE
- Run 71052501

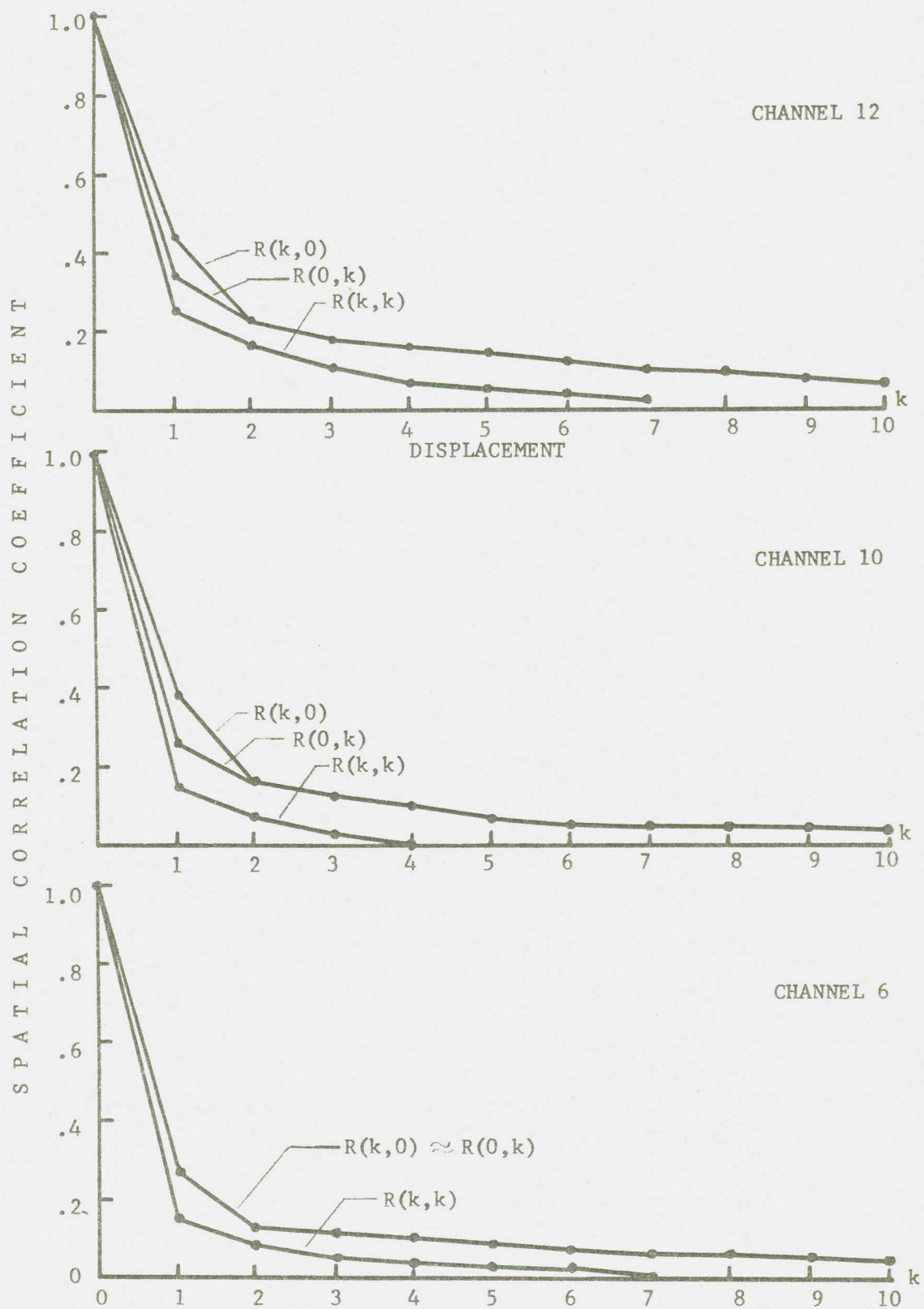


Figure A3 Spatial Correlation Coefficients - Class DECIDUOUS
- Run 71052501

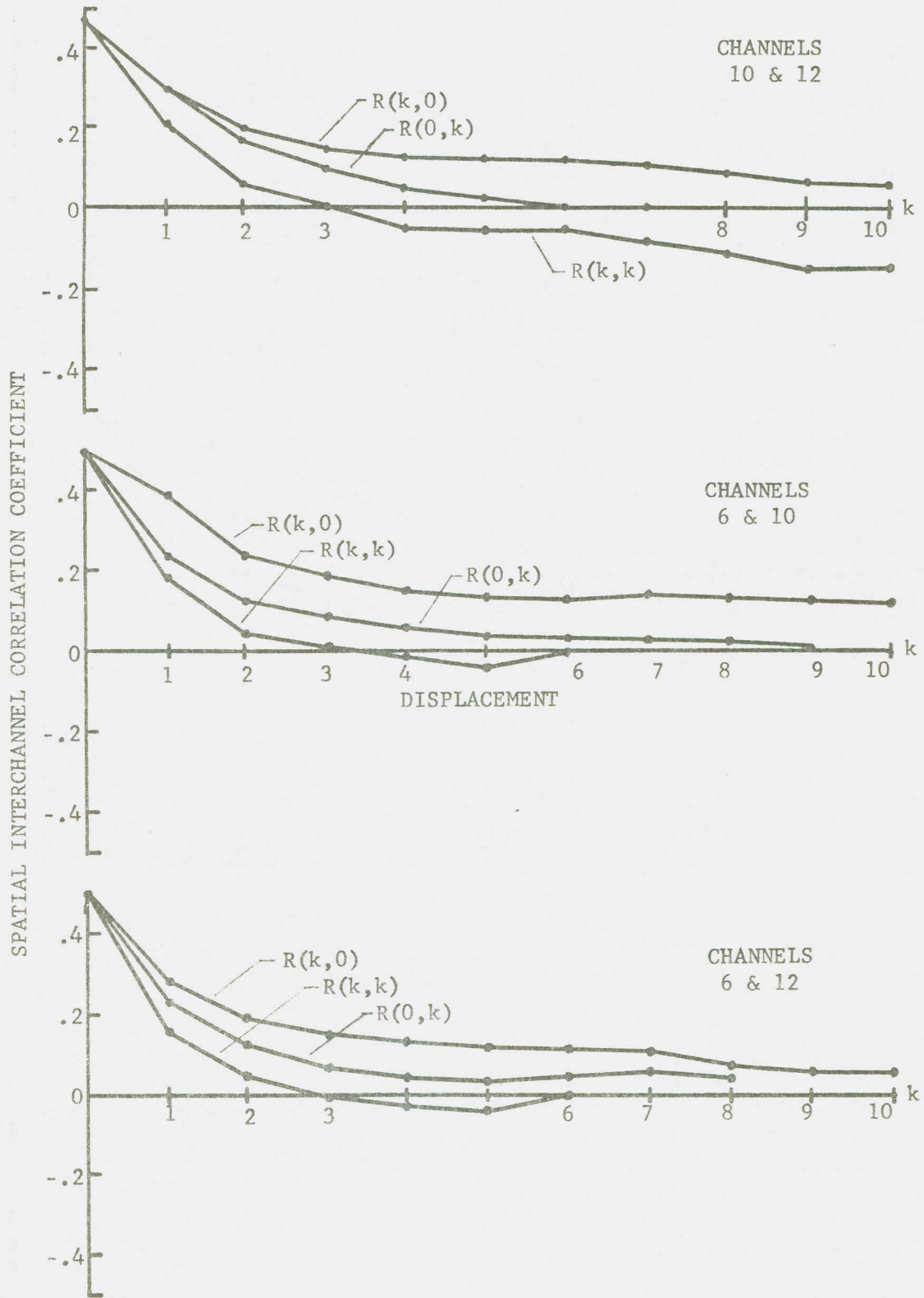


Figure A4 Spatial Interchannel Correlation Coefficients
 - Class FORAGE - Run 71052501

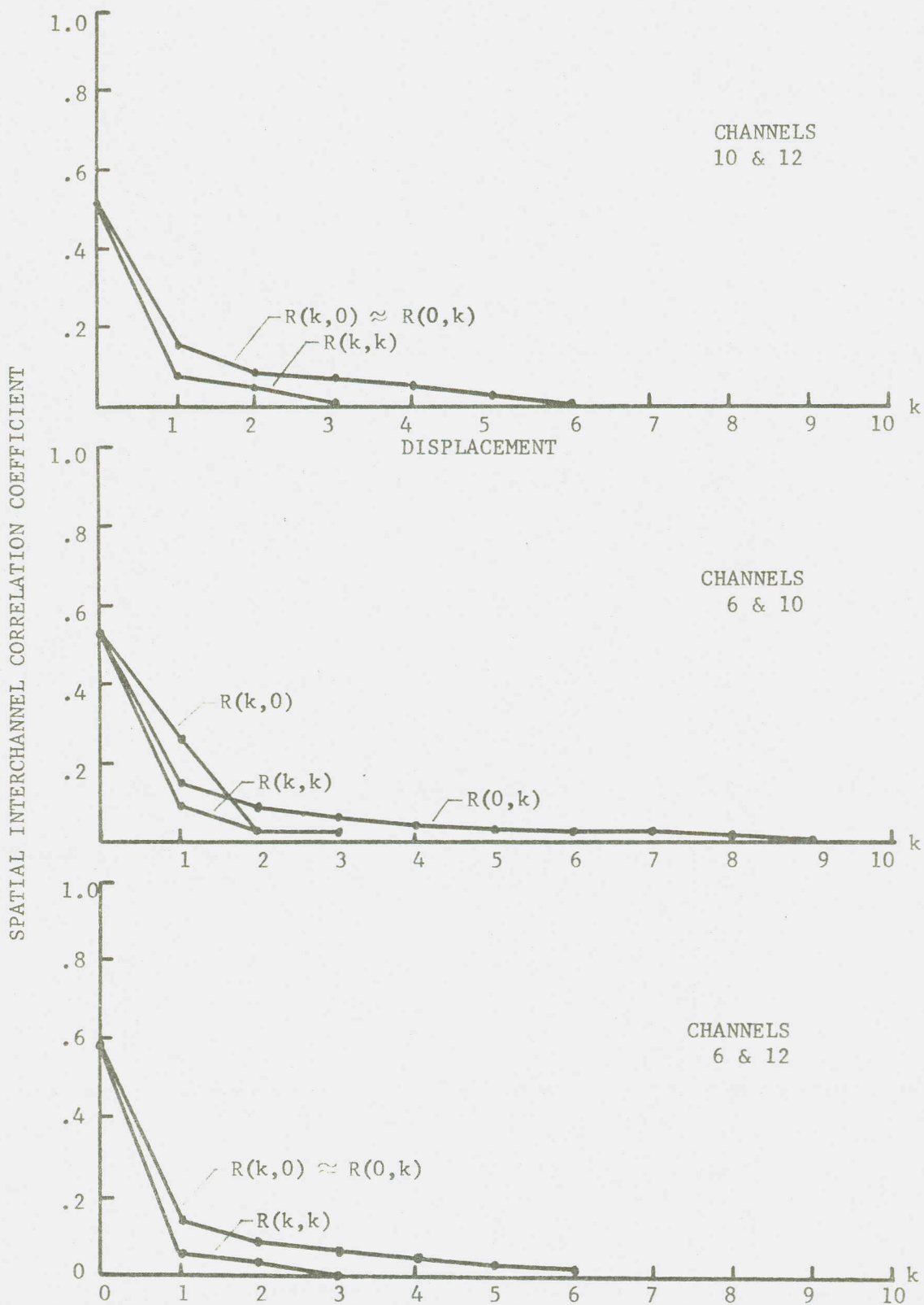


Figure A5 Spatial Interchannel Correlation Coefficients
 - Class DECIDUOUS - Run 71052501

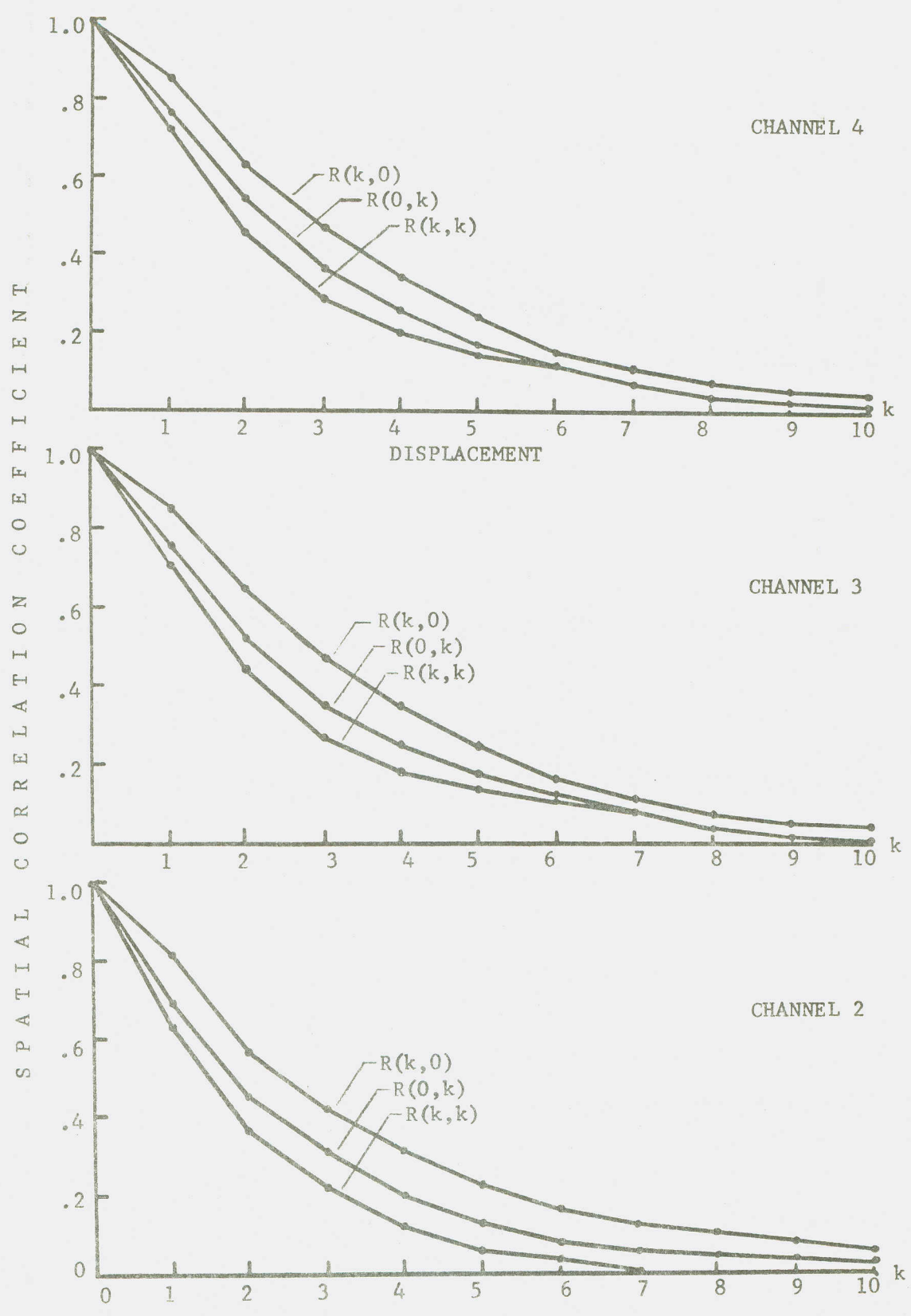


Figure A6 Spatial Correlation Coefficients - Run 72032803
Region: lines 260-559, columns 998-1097

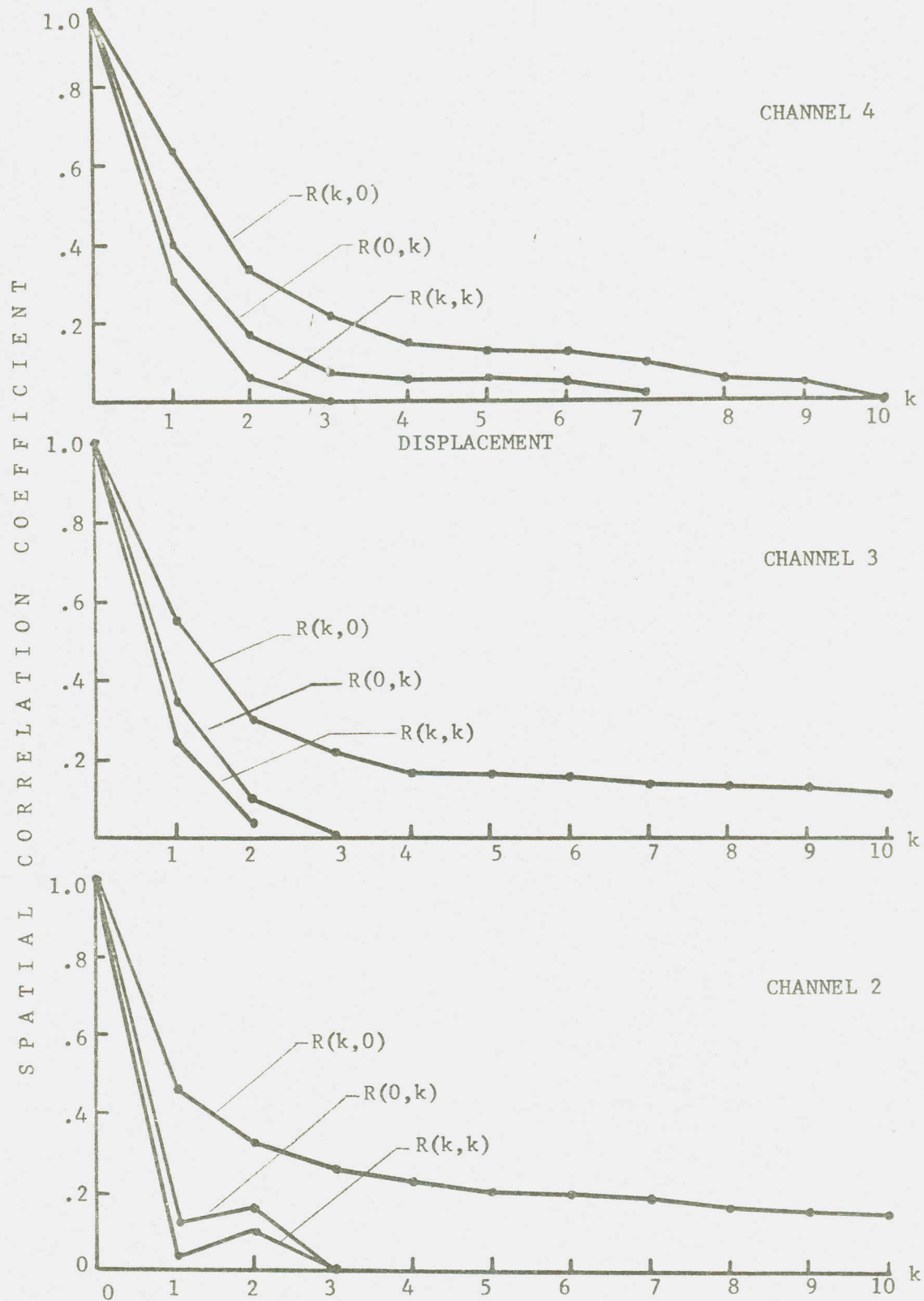


Figure A7 Spatial Correlation Coefficients - Class CORN
 - Run 72032803

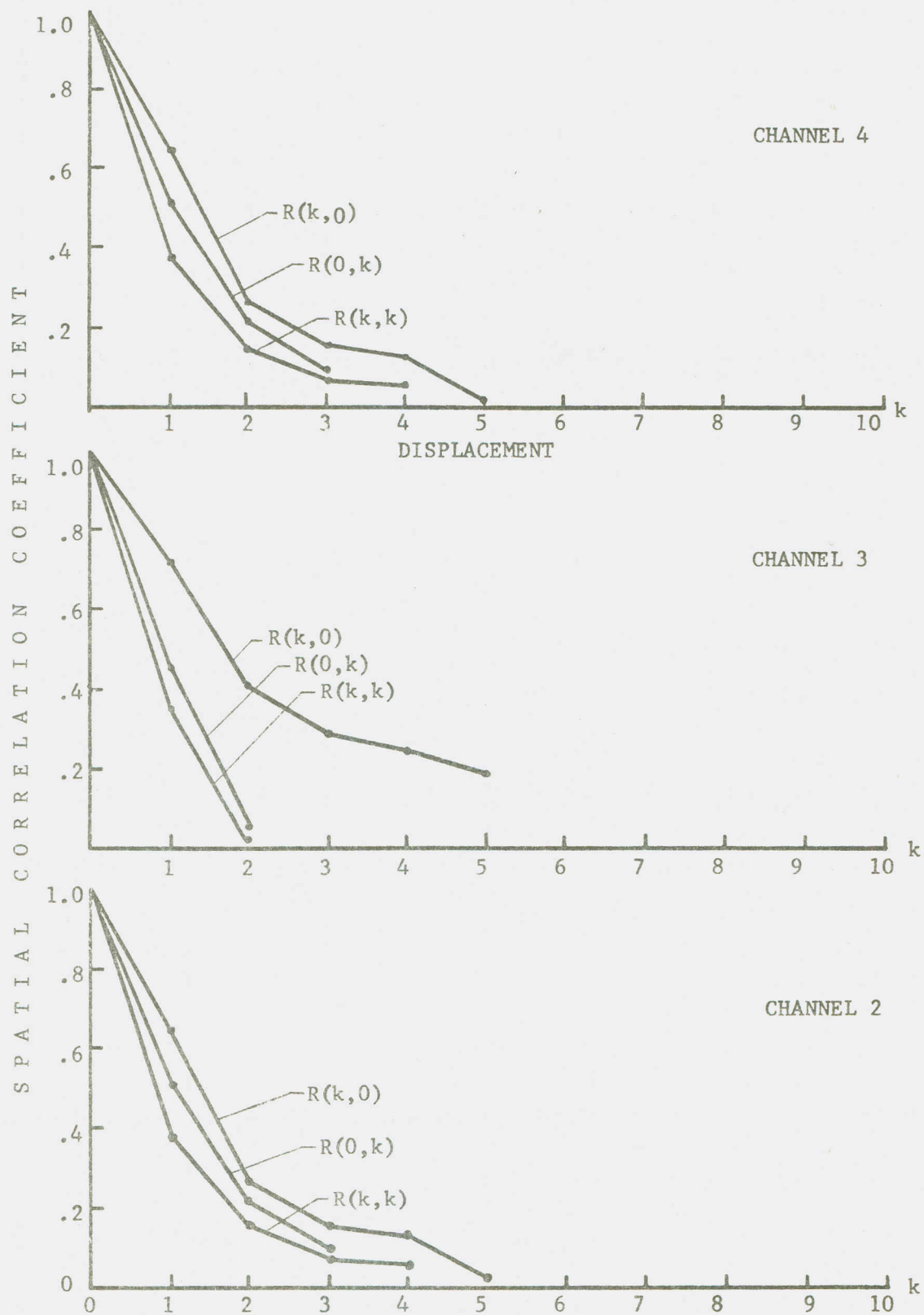


Figure A8 Spatial Correlation Coefficients - Class TOWN
- Run 72032803

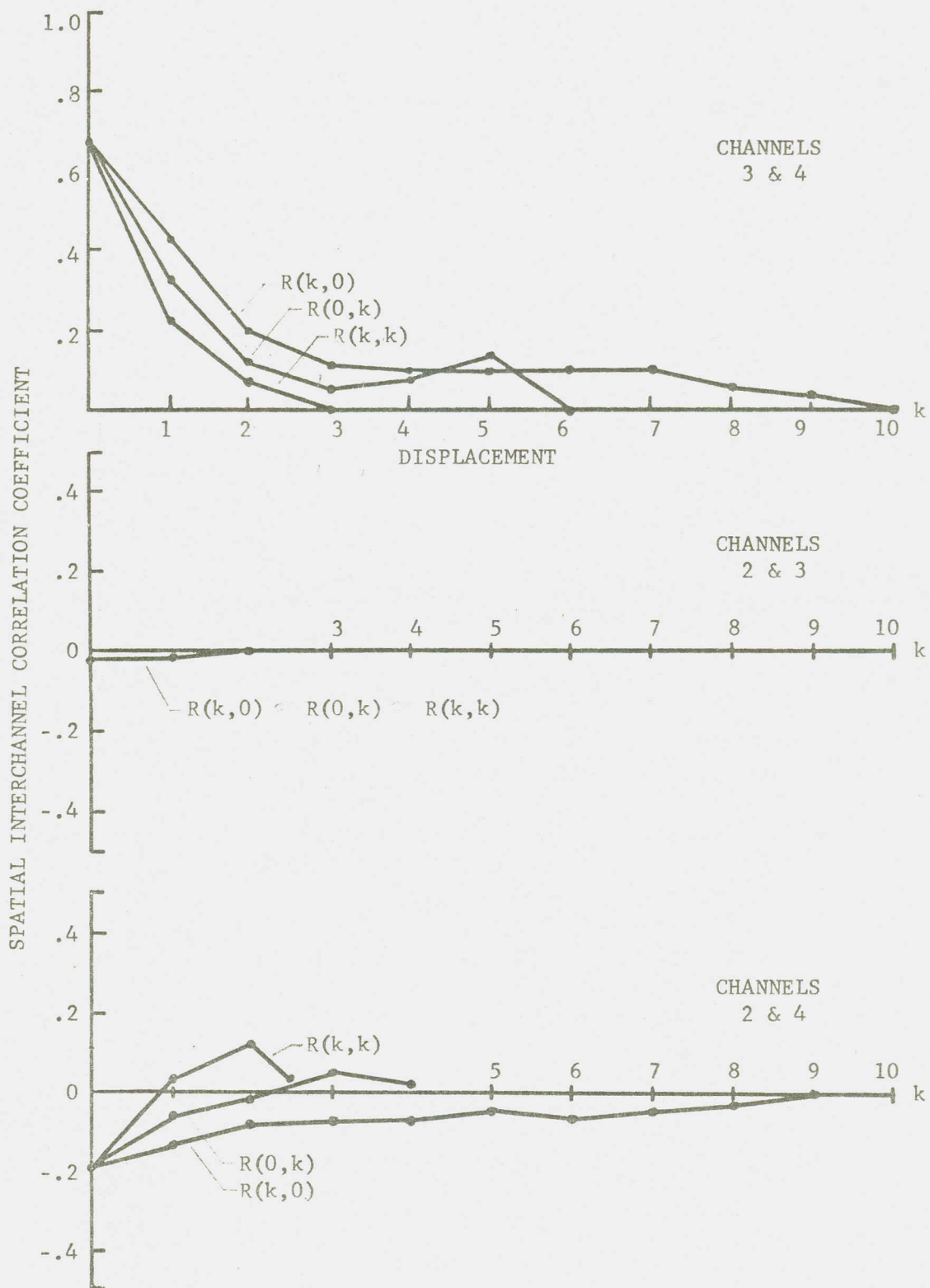


Figure A9 Spatial Interchannel Correlation Coefficients - Class CORN
 - Run 72032803

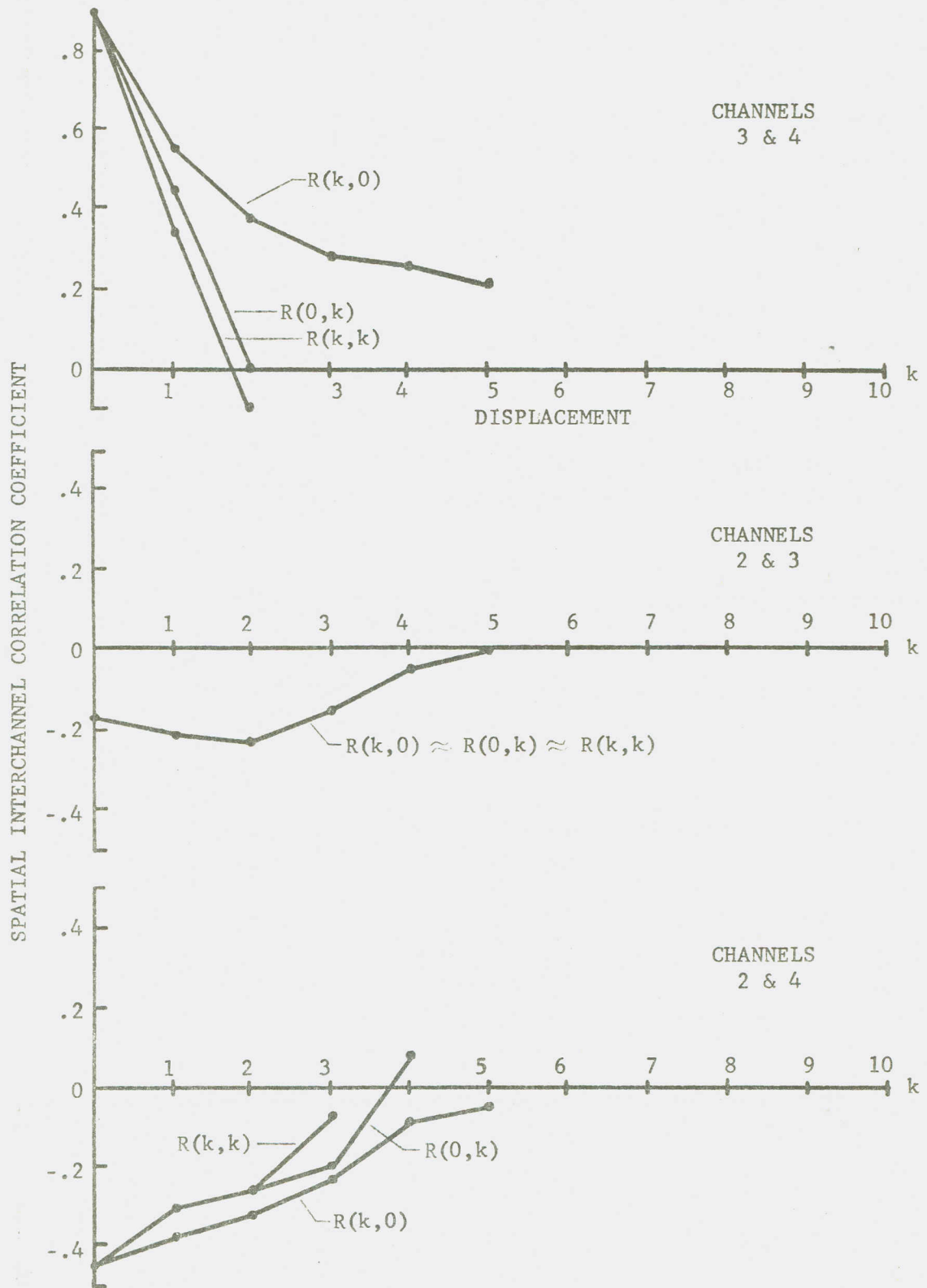


Figure A10 Spatial Interchannel Correlation Coefficients - Class TOWN
 - Run 72032803

APPENDIX B
THE COMPOUND DECISION APPROACH

Let $X^n = (\underline{X}_1, \dots, \underline{X}_n)$ denote a set of pixels (q-dimensional random variables) with dependent states (classes). As in Section 2.1 we assume class-conditional independence. The goal is to classify each pixel into the set of classes $W = (W_1, \dots, W_m)$ (i.e. classify X^n into W^n) such that the expected number of misclassified pixels is minimized. As shown in Section 2.3.2, this is accomplished by decision functions $W_i(x^n)$ such that

$$P(\underline{X}_i \in W_i(x^n) | X^n = x^n) = \max_j P(\underline{X}_i \in W_j | X^n = x^n), \quad i=1, \dots, n \quad B1$$

A simpler, though suboptimal procedure is to classify the pixels "sequentially" without "looking ahead"; i.e. let the ith decision function be based only on the observations $x^i = (\underline{X}_1, \dots, \underline{X}_i)$. This is equivalent to setting $n=i$ in equation B1. Making this change and applying Bayes rule provides:

$$p(X^i = x^i | \underline{X}_i \in W_i(x^i)) P(W_i(x^i)) = \max_j p(X^i = x^i | \underline{X}_i \in W_j) P(W_j) \quad B2$$

Before proceeding we shall shorten the notation by expressing the right hand side of B2 as:

$$\max_{A_j \in W} p(x^i | A_j) P(A_j) \quad B3$$

where A_j denotes both a class and the event that \underline{X}_i is a

random variable from that class. Similarly A^i will denote a vector of i classes and the event that these are the true classes of the set X^i . Defining $j=i-1$ and applying the law of total probability to B3 provides:

$$\begin{aligned} \max_{A_i} \sum_{A^j \in W^j} p(x^i | A_i, A^j) P(A_i | A^j) P(A^j) & \quad B4 \\ & = \max_{A_i} \sum_{A^j} p(x_i | A_i) P(A_i | A^j) p(x^j | A^j) P(A^j) \end{aligned}$$

Defining the quantity $Q_i(A_1, \dots, A_i) = p(x^i | A^i) P(A^i)$, it is apparent that B4 can be computed recursively as:

$$\max_{A_i} \sum_{A^j} Q_i(A_1, \dots, A_i) \quad B5$$

where

$$Q_i(A_1, \dots, A_i) = p(x_i | A_i) P(A_i | A^j) Q_j(A_1, \dots, A_j) \quad B6$$

but the number of terms in the summation grows as m^j .

If the states form a first order Markov chain, B4 reduces to simply:

$$\max_{A_i} R_i(A_i) \quad B7$$

where

$$R_i(A_i) \triangleq p(x_i | A_i) \sum_{A^j} p(x^j | A^j) P(A^j)$$

but it can be computed recursively as:

$$R_i(A_i) = p(x_i | A_i) \sum_{A_j} P(A_i | A_j) R_j(A_j) \quad B8$$

For a k th order Markov chain ($k > 1$) we define $j=i-k$, and B4 reduces to:

$$\max_{A_i} p(x_i | A_i) \sum_{A_{i-1}} p(x_{i-1} | A_{i-1}) \sum_{A_{i-2}} \dots \sum_{A_{j+1}} R_i(A_i, \dots, A_{j+1}) \quad B9$$

where

$$\begin{aligned}
 R_i(A_i, \dots, A_{j+1}) &\stackrel{\Delta}{=} p(x_{j+1} | A_{j+1}) \sum_{A_j} p(x^j | A_j) P(A_i) & \text{B10} \\
 &= p(x_{j+1} | A_{j+1}) \sum_{A_j} P(A_i | A_{i-1}, \dots, A_j) R_{i-1}(A_{i-1}, \dots, A_j)
 \end{aligned}$$

The number of summation terms in B9 and B10 is only m^k .

To obtain this relatively efficient classifier we have assumed a sequential approach, Markov chain dependence between states, and class-conditionally independent data vectors. In spite of these simplifications it is clear that the computational and memory requirements of the compound decision approach are considerably greater than the no-memory approach.

APPENDIX C
COMPOSITE CLASSES

Let $p(x|W_1)$ and $p(x|W_2)$ be two univariate normal class densities. Let W_3 be a third class that is a composite (spatially) of the two ground cover-types represented by these densities. The problem is to determine the density of this class. We shall assume that the conditional density of a pixel containing 100*a% of class 1 and 100(1-a)% of class 2 is also normal with mean $M(a) = aM_1 + (1-a)M_2$ and variance $V(a) = aV_1 + (1-a)V_2$, where M_i and V_i represent the mean and variance of the ith class. Thus

$$p(x|W_3) = \int_0^1 N(M(a), V(a); x) p(a) da \quad C1$$

for some distribution $p(a)$ which depends on the overall proportion of each class and the "texture" of the composite.

Assume for example that the overall proportion of each class is 0.5, and consider three different cases of texture:

Case 1 - Maximum Variance

If the texture of the composite is very coarse compared to the size of a pixel, then a pixel will usually contain either one class or the other, not both. $p(a)$ is approximated by $.5(\delta(a) + \delta(a-1))$, where $\delta(\cdot)$ represents the

Dirac delta function. Thus $p(x|W_3)$ is approximately $.5(p(x|W_1)+p(x|W_2))$, which of course is bimodal in general. We refer to this as a "mixture" class, because the constituents retain their individual characteristics. The normal way to handle such a class is to cluster the data and treat each mode as a subclass. Thus a mixture of W_1 and W_2 cannot be treated as a distinct third class by a Gaussian classifier. (Of course, post-processing could be applied to the classifier output to search for such a mixture, if desired.)

Case 2 - Minimum Variance

If the texture of the composite is very fine, then $p(a)$ is approximated by $\delta(a-.5)$, and $p(x|W_3)$ is approximately normal with mean $(M_1+M_2)/2$ and variance $(V_1+V_2)/2$. This can (and should) be treated as a distinct third class. We refer to it as a "compound" class, because the constituents lose their individuality.

Case 3 - Intermediate Variance

When the texture of the composite is on the same order as the pixel size (e.g. Figures C1 and C2), a randomly selected pixel can contain any proportions of the two classes. As a first-order approximation to $p(a)$ we consider the strictly periodic pattern in Figure C1 and let the coordinates (X,Y) of the pixel be uniformly distributed random variables over the area covered by the pattern. Assume for simplicity that the orientation of the pixel

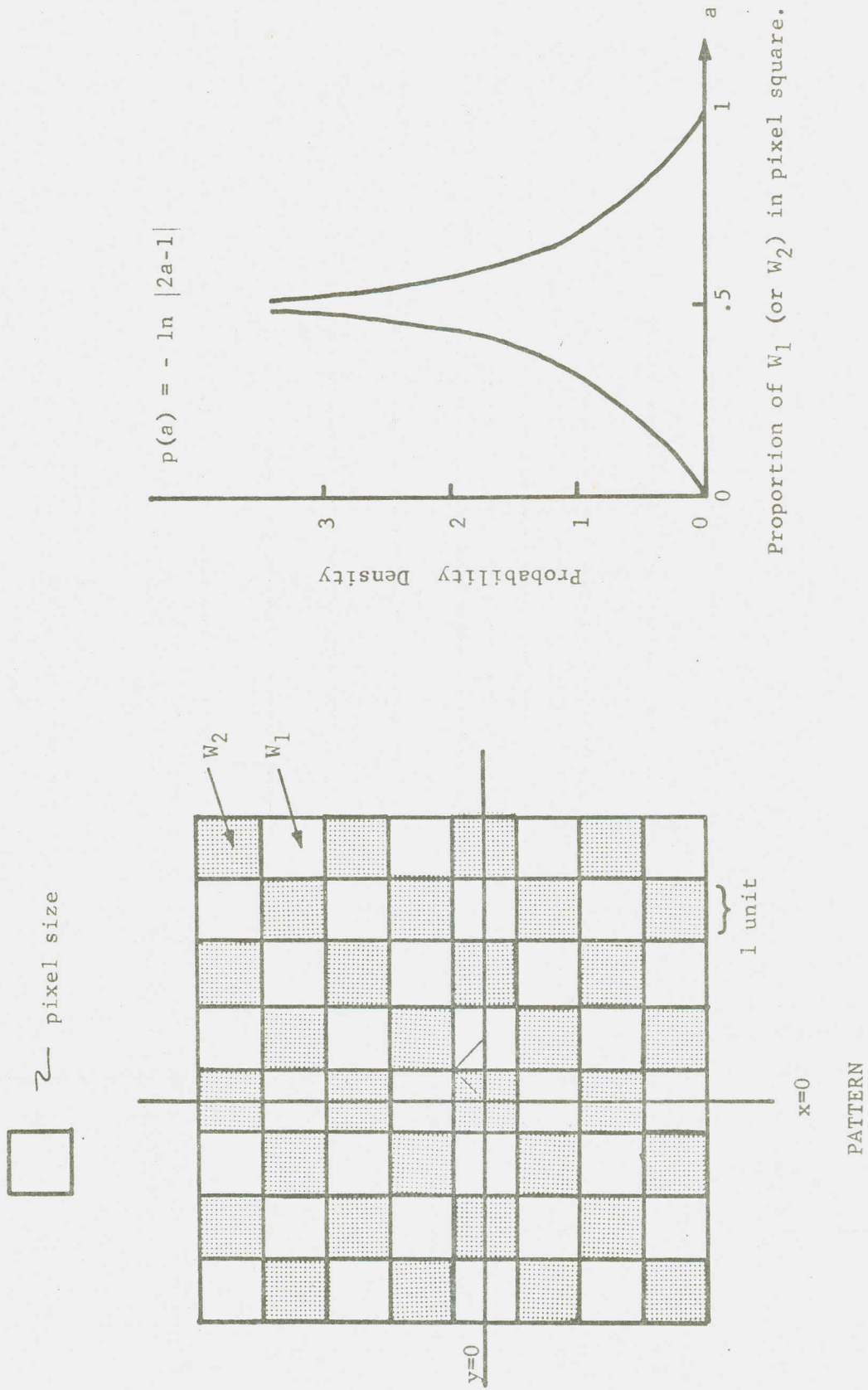
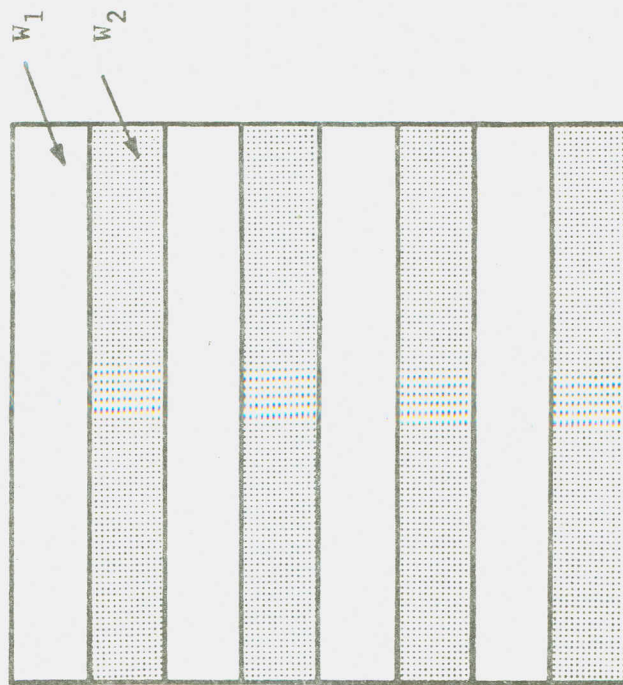
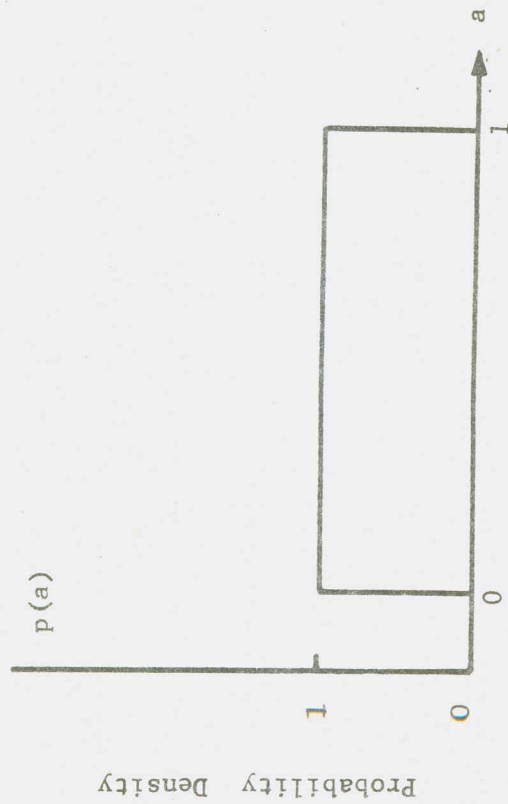


Figure C1 A case of spatial texture on the order of the pixel size.



PATTERN



Proportion of W_1 (or W_2) in pixel square.

Figure C2 Texture in One Spatial Dimension Only

square is the same as the square elements in the pattern. Symmetry considerations require that $p(a)$ be unchanged if the point (X,Y) is uniformly distributed over the region $0 \leq x \leq 1, 0 \leq y \leq \min(x,1-x)$. In this region $a(x,y) = x-2xy+y$.

By transformation of random variables it follows that:

$$p(a) = -\ln |2a-1|, \quad 0 \leq a \leq 1 \quad C2$$

which is shown in Figure C1. The variance of this distribution is intermediate between Cases 1 and 2.

Figure C3 shows two hypothetical class densities, $p(x|W_1)$ and $p(x|W_2)$, and the density of their composite, $p(x|W_3)$ (obtained by numerical evaluation of formula C1 using formula C2). As in Case 2, $p(x|W_3)$ forms a distinct, unimodal class, but its variance is larger due to the spread of $p(a)$.

Another interesting case is the pattern shown in Figure C2. Here the texture is primarily one-dimensional. (This tendency can be observed for instance in a few of the objects in Figure 4.2.4.) When the detector size matches the line width, $p(a)$ is uniform on the interval $0 \leq a \leq 1$. Its variance is greater than that of formula C2, which results in the composite class $p(x|W_4)$ shown in Figure C3. Important points to notice are that it is unimodal, relatively broad, and it forms a distinct compound class. As M_2-M_1 increases relative to the variance of W_1 and W_2 , $p(x|W_4)$ tends toward uniform on the interval (M_1, M_2) . Thus, if other classes lie in this interval, observations from the

compound class will tend to be misclassified at a high rate by a no-memory classifier. Of course, much less confusion should result if classification is done on a sample basis.

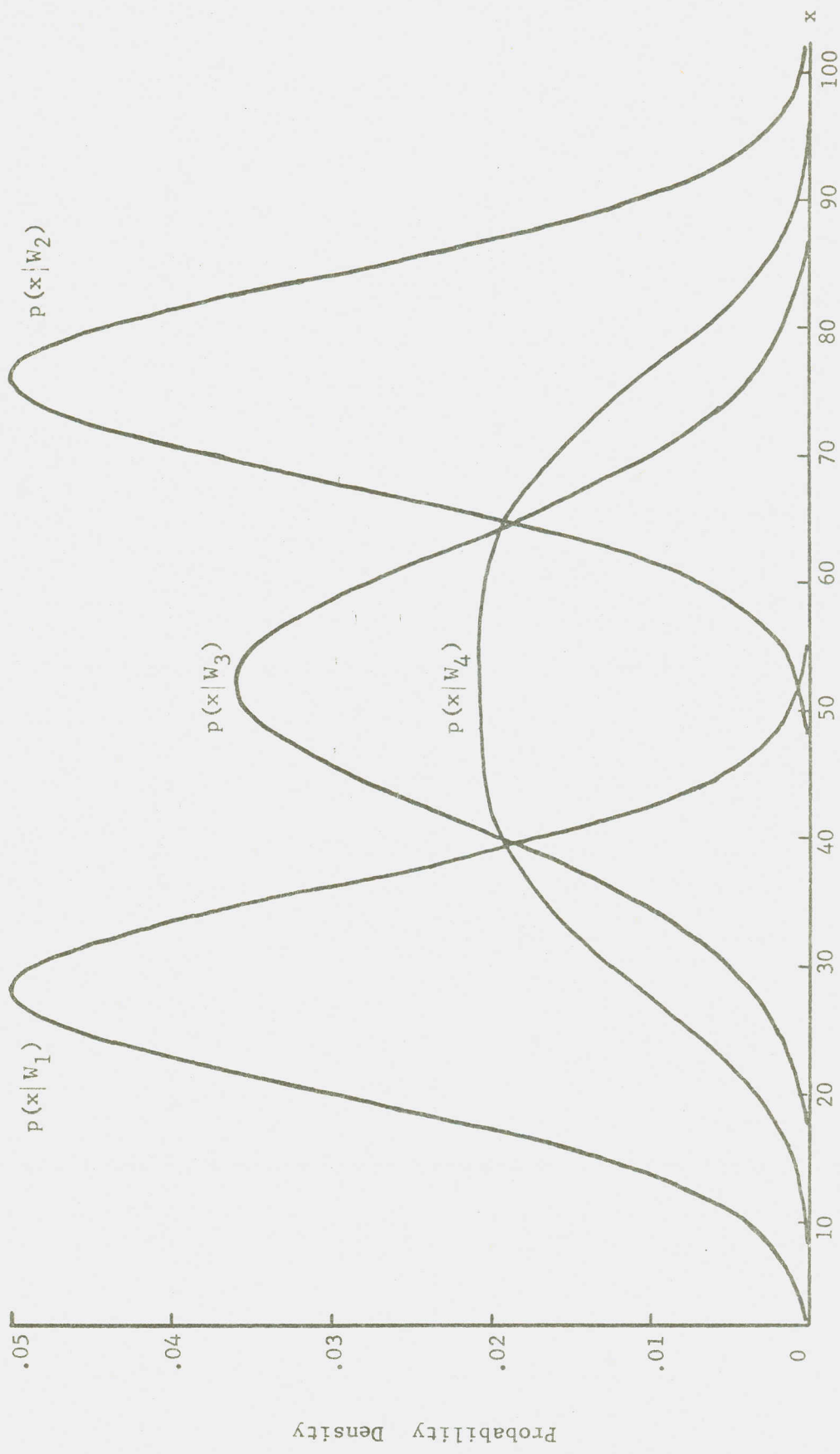


Figure C3 Two Densities and Their Compound Density for Spatial Texture on the Order of a Pixel.

APPENDIX D

FIELD LISTINGS

TABLE D1 TEST AREAS FOR RUN 71052800

RUN NUMBER	ARFA DESIGNATION	FIRST LINE	LAST LINE	LINE INT	FIRST COL	LAST COL	COL INT	COVER TYPE
71052800	H-15	113	124	1	20	45	1	CORN
71052800	H-15	127	131	1	34	43	1	CORN
71052800	H-14	136	139	1	22	29	1	CORN
71052800	H-14.0	142	152	1	21	43	1	CORN
71052800	G-10	227	239	1	130	140	1	CORN
71052800	G-11	233	239	1	111	117	1	CORN
71052800	G-11	242	262	1	111	124	1	CORN
71052800	G-10	242	248	1	139	145	1	CORN
71052800	G-11	254	262	1	127	136	1	CORN
71052800	G-11	265	288	1	111	145	1	CORN
71052800	Y-1	291	317	1	196	207	1	CORN
71052800	H-1	292	319	1	23	93	1	CORN
71052800	Y-1	292	298	1	181	193	1	CORN
71052800	Y-1	301	308	1	190	193	1	CORN
71052800	Y-3	311	319	1	181	184	1	CORN
71052800	H-1	322	352	1	23	54	1	CORN
71052800	G-15	324	352	1	83	85	1	CORN
71052800	G-3	364	383	1	99	116	1	CORN
71052800	T-4	365	384	1	157	160	1	CORN
71052800	G-3	386	408	1	104	136	1	CORN
71052800	D-8	388	417	1	47	94	1	CORN
71052800	D-8	389	396	1	40	44	1	CORN
71052800	F-2	425	429	1	190	199	1	CORN
71052800	F-2	432	439	1	180	204	1	CORN
71052800	F-2	442	446	1	166	210	1	CORN
71052800	P-7	452	463	1	157	178	1	CORN
71052800	P-3	453	475	1	30	33	1	CORN
71052800	P-6	453	455	1	139	142	1	CORN
71052800	P-7	458	463	1	148	154	1	CORN
71052800	P-3	466	482	1	38	42	1	CORN
71052800	P-7	466	480	1	140	179	1	CORN
71052800	E-1	486	492	1	140	145	1	CORN
71052800	N-10	486	503	1	120	134	1	CORN
71052800	E-1	495	502	1	145	149	1	CORN
71052800	E-1	505	510	1	151	153	1	CORN
71052800	D-1	515	520	1	193	196	1	CORN
71052800	D-8	517	541	1	154	178	1	CORN
71052800	D-1	519	524	1	190	192	1	CORN
71052800	D-2	522	526	1	196	198	1	CORN
71052800	D-1	524	532	1	183	186	1	CORN
71052800	D-2	529	532	1	190	193	1	CORN
71052800	D-2	536	539	1	184	186	1	CORN
71052800	N-11	538	574	1	145	146	1	CORN
71052800	CC-3	539	603	1	66	69	1	CORN
71052800	N-2	579	601	1	20	22	1	CORN
71052800	CC-5	609	628	1	68	79	1	CORN
71052800	W-3	613	659	1	147	170	1	CORN
71052800	X-4	634	739	1	60	97	1	CORN
71052800	Q-23	662	729	1	149	162	1	CORN
71052800	Q-24	663	729	1	165	180	1	CORN
71052800	Q-2	663	697	1	183	192	1	CORN
71052800	X-4	707	739	1	100	114	1	CORN
71052800	Q-9	732	740	1	149	172	1	CORN
71052800	X-10	745	801	1	45	61	1	CORN
71052800	X-9	745	761	1	68	81	1	CORN
71052800	F-12	746	800	1	29	40	1	CORN
71052800	X-9	753	761	1	84	92	1	CORN
71052800	Q-20	755	765	1	96	115	1	CORN
71052800	Q-6	755	807	1	121	126	1	CORN
71052800	Q-3	756	792	1	141	145	1	CORN
71052800	Q-20	764	784	1	86	92	1	CORN
71052800	Q-20	768	772	1	95	104	1	CORN
71052800	Q-6	768	772	1	112	118	1	CORN

TABLE D1, CONTINUED

71052800	Q-6	775	784	1	97	118	1	CORN
71052800	Q-21	786	801	1	79	83	1	CORN
71052800	Q-6	787	800	1	88	118	1	CORN
71052800	Q-6	789	799	1	129	133	1	CORN
71052800	Q-10	803	846	1	27	37	1	CORN
71052800	Q-21	803	823	1	68	98	1	CORN
71052800	Q-21	804	811	1	52	65	1	CORN
71052800	Q-19	804	817	1	142	147	1	CORN
71052800	Q-10	817	846	1	40	54	1	CORN
71052800	Q-19	819	823	1	133	140	1	CORN
71052800	Q-14	839	842	1	142	147	1	CORN
71052800	Q-10	840	846	1	74	85	1	CORN
71052800	Q-14	848	850	1	150	154	1	CORN
71052800	Q-10	849	859	1	28	94	1	CORN
71052800	Q-14	854	858	1	159	162	1	CORN
71052800	Q-14	861	865	1	165	168	1	CORN
71052800	Q-18	881	891	1	104	118	1	CORN
71052800	Q-18	894	944	1	88	129	1	CORN
71052800	Q-17	903	913	1	136	148	1	CORN
71052800	Q-22	905	943	1	75	83	1	CORN
71052800	Q-22	916	943	1	70	72	1	CORN
71052800	Q-17	925	943	1	134	172	1	CORN
71052800	Q-17	937	941	1	173	182	1	CORN
71052800	R-4	941	944	1	28	58	1	CORN
71052800	R-8	949	961	1	44	54	1	CORN
71052800	R-3	950	1005	1	31	37	1	CORN
71052800	R-2	950	1005	1	40	41	1	CORN
71052800	R-8	964	973	1	44	49	1	CORN
71052800	AA-2	965	969	1	156	164	1	CORN
71052800	AA-1	976	985	1	178	187	1	CORN
71052800	V-6	989	997	1	140	144	1	CORN
71052800	V-1	1008	1023	1	173	184	1	CORN
71052800	Z-1	1014	1037	1	121	134	1	CORN
71052800	S-15	1041	1049	1	37	55	1	CORN
71052800	S-13	1052	1065	1	37	57	1	CORN
71052800	S-13	1053	1066	1	63	94	1	CORN
71052800	S-6	1074	1101	1	130	143	1	CORN
71052800	S-2	1083	1098	1	155	184	1	CORN
71052800	S-6	1105	1121	1	139	142	1	CORN
71052800	S-6	1105	1160	1	145	146	1	CORN
71052800	S-4	1120	1134	1	166	185	1	CORN
71052800	S-6	1126	1165	1	152	157	1	CORN
71052800	S-10	1156	1198	1	106	127	1	CORN
71052800	BB-1	1204	1217	1	95	178	1	CORN
71052800	BB-4	1219	1231	1	184	187	1	CORN
71052800	BB-1	1220	1230	1	163	172	1	CORN
71052800	BB-1	1226	1240	1	73	88	1	CORN
71052800	I-2	112	128	1	153	181	1	SOY
71052800	J-1	113	127	1	96	130	1	SOY
71052800	H-8	207	224	1	80	89	1	SOY
71052800	H-2	229	288	1	19	34	1	SOY
71052800	G-9	262	274	1	160	175	1	SOY
71052800	D-6	263	275	1	181	208	1	SOY
71052800	G-9	277	284	1	171	176	1	SOY
71052800	G-8	278	288	1	149	157	1	SOY
71052800	K-1	323	355	1	141	163	1	SOY
71052800	E-6	482	510	1	199	209	1	SOY
71052800	E-3	482	488	1	172	178	1	SOY
71052800	N-9	511	514	1	102	109	1	SOY
71052800	CC-4	532	602	1	73	76	1	SOY
71052800	N-4	581	631	1	12	16	1	SOY
71052800	X-2	633	714	1	117	142	1	SOY
71052800	X-5	635	708	1	28	55	1	SOY
71052800	X-7	712	740	1	13	55	1	SOY
71052800	X-2	717	722	1	136	140	1	SOY
71052800	X-2	717	739	1	118	126	1	SOY

TABLE D1, CONTINUED

71052800	Q-25	755	784	1	130	132	1	SOY
71052800	M-4	787	795	1	171	181	1	SOY
71052800	M-4	787	802	1	150	168	1	SOY
71052800	M-4	805	807	1	153	163	1	SOY
71052800	Q-12	864	874	1	133	145	1	SOY
71052800	R-5	867	884	1	81	91	1	SOY
71052800	R-5	867	896	1	67	78	1	SOY
71052800	R-5	868	934	1	28	64	1	SOY
71052800	AA-7	950	1004	1	75	112	1	SOY
71052800	AA-4	950	977	1	129	133	1	SOY
71052800	Z-4	1012	1033	1	70	96	1	SOY
71052800	S-11	1041	1056	1	127	137	1	SOY
71052800	S-11	1061	1067	1	122	127	1	SOY
71052800	S-9	1073	1080	1	49	58	1	SOY
71052800	S-9	1073	1091	1	61	79	1	SOY
71052800	S-7	1077	1084	1	110	115	1	SOY
71052800	S-7	1093	1104	1	118	122	1	SOY
71052800	S-9	1136	1153	1	99	103	1	SOY
71052800	S-7	1138	1157	1	137	139	1	SOY
71052800	S-5	1139	1196	1	175	187	1	SOY
71052800	BB-3	1237	1245	1	127	140	1	SOY
71052800	BB-3	1251	1263	1	112	145	1	SOY
71052800	I-1	112	128	1	139	148	1	WHEAT
71052800	H-19	113	172	1	75	89	1	WHEAT
71052800	H-19	114	117	1	69	73	1	WHEAT
71052800	H-19	120	130	1	64	73	1	WHEAT
71052800	H-19	133	172	1	57	72	1	WHEAT
71052800	H-19	151	172	1	50	54	1	WHEAT
71052800	H-25	153	203	1	132	185	1	WHEAT
71052800	H-25	163	171	1	125	129	1	WHEAT
71052800	H-25	174	203	1	116	129	1	WHEAT
71052800	H-10	190	203	1	34	41	1	WHEAT
71052800	D-5	228	246	1	181	208	1	WHEAT
71052800	G-7	228	247	1	149	176	1	WHEAT
71052800	H-28	231	288	1	78	85	1	WHEAT
71052800	H-5	245	288	1	90	105	1	WHEAT
71052800	D-5	249	259	1	180	208	1	WHEAT
71052800	G-7	250	258	1	155	176	1	WHEAT
71052800	G-5	292	319	1	140	175	1	WHEAT
71052800	K-2	323	353	1	168	192	1	WHEAT
71052800	T-2	358	363	1	150	161	1	WHEAT
71052800	T-6	359	371	1	181	198	1	WHEAT
71052800	U-1	373	383	1	201	205	1	WHEAT
71052800	D-8	421	429	1	50	55	1	WHEAT
71052800	F-3	437	447	1	139	152	1	WHEAT
71052800	F-3	437	439	1	155	160	1	WHEAT
71052800	E-5	483	510	1	191	197	1	WHEAT
71052800	C-2	488	519	1	47	52	1	WHEAT
71052800	A-1	489	504	1	15	29	1	WHEAT
71052800	A-1	490	504	1	8	12	1	WHEAT
71052800	O-6	547	566	1	159	173	1	WHEAT
71052800	CC-1	558	630	1	37	44	1	WHEAT
71052800	Q-3	589	608	1	167	181	1	WHEAT
71052800	N-3	610	631	1	20	22	1	WHEAT
71052800	W-2	613	628	1	174	181	1	WHEAT
71052800	W-2	639	659	1	174	181	1	WHEAT
71052800	Q-8	745	751	1	96	124	1	WHEAT
71052800	M-1	745	784	1	168	171	1	WHEAT
71052800	Q-9	765	783	1	78	82	1	WHEAT
71052800	M-1	768	784	1	174	180	1	WHEAT
71052800	M-6	822	846	1	151	161	1	WHEAT
71052800	M-6	827	835	1	139	148	1	WHEAT
71052800	M-6	829	843	1	164	169	1	WHEAT
71052800	Q-13	846	853	1	138	146	1	WHEAT
71052800	Q-13	855	862	1	147	154	1	WHEAT
71052800	Q-13	864	869	1	157	161	1	WHEAT

TABLE D1, CONTINUED

71052800	AA-3	948	961	1	146	178	1	WHFAT
71052800	AA-3	949	960	1	181	186	1	WHEAT
71052800	V-2	1018	1030	1	160	166	1	WHFAT
71052800	S-3	1103	1115	1	154	184	1	WHEAT
71052800	D-7	279	285	1	181	207	1	IDLE
71052800	K-5	322	352	1	202	208	1	IDLE
71052800	D-3	514	520	1	182	185	1	IDLE
71052800	CC-2	551	629	1	49	62	1	IDLE
71052800	X-11	765	801	1	65	74	1	IDLE
71052800	Q-5	816	823	1	102	121	1	IDLE
71052800	Q-5	826	836	1	87	118	1	IDLE
71052800	Q-11	853	870	1	115	126	1	IDLE
71052800	Q-15	878	881	1	127	132	1	IDLE
71052800	Q-15	884	888	1	135	140	1	IDLE
71052800	Q-15	898	900	1	151	154	1	IDLE
71052800	AA-6	950	1003	1	117	125	1	IDLE
71052800	AA-8	962	1006	1	66	69	1	IDLE
71052800	AA-8	972	979	1	60	63	1	IDLE
71052800	R-7	983	1006	1	56	63	1	IDLE
71052800	S-8	1073	1088	1	87	90	1	IDLE
71052800	S-8	1073	1105	1	93	101	1	IDLE
71052800	S-8	1106	1107	1	97	100	1	IDLE
71052800	S-8	1124	1147	1	114	120	1	IDLE
71052800	S-8	1126	1138	1	108	111	1	IDLE
71052800	G-16	324	352	1	77	80	1	HAY
71052800	G-18	340	343	1	57	73	1	HAY
71052800	F-5	406	411	1	192	199	1	HAY
71052800	N-6	497	501	1	78	85	1	HAY
71052800	N-8	565	603	1	81	82	1	HAY
71052800	X-3	634	703	1	101	112	1	HAY
71052800	X-3	726	740	1	142	145	1	HAY
71052800	H-20	147	152	1	94	107	1	LESPEDFZA
71052800	H-20	155	159	1	94	100	1	LESPEDFZA
71052800	H-26	207	223	1	118	186	1	LESPEDFZA
71052800	H-3	229	288	1	38	46	1	LESPEDFZA
71052800	M-2	746	784	1	150	164	1	LESPEDFZA
71052800	M-8	822	827	1	173	180	1	LESPEDFZA
71052800	M-8	830	832	1	179	184	1	LESPEDFZA
71052800	M-7	852	855	1	166	171	1	LESPEDFZA
71052800	H-22	183	202	1	93	96	1	PASTURE
71052800	H-12	188	206	1	5	10	1	PASTURE
71052800	H-22	205	223	1	93	100	1	PASTURE
71052800	H-6	228	240	1	97	106	1	PASTURE
71052800	H-27	228	288	1	67	73	1	PASTURE
71052800	D-9	368	384	1	23	26	1	PASTURE
71052800	C-3	488	507	1	56	73	1	PASTURE
71052800	C-3	510	512	1	56	64	1	PASTURE
71052800	E-2	484	491	1	149	169	1	RYE
71052800	E-2	504	511	1	157	169	1	RYE
71052800	D-4	518	526	1	206	210	1	RYE
71052800	V-5	985	995	1	148	156	1	RYE
71052800	H-23	133	150	1	123	127	1	WOODPAST
71052800	H-23	155	162	1	112	118	1	WOODPAST
71052800	H-23	174	202	1	101	111	1	WOODPAST
71052800	H-23	205	222	1	104	113	1	WOODPAST
71052800	X-1	719	740	1	130	133	1	FOREST
71052800	R-6	862	866	1	99	107	1	FOREST
71052800	V-7	975	977	1	148	155	1	FOREST
71052800	V-7	978	980	1	139	145	1	FOREST
71052800	V-4	1018	1022	1	151	154	1	FOREST

TABLE D1, CONTINUED

71052800	S-14	1021	1032	1	58	61	1	FOREST
71052800	Z-2	1030	1039	1	104	115	1	FOREST
71052800	H-16	156	158	1	21	37	1	NON-FARM
71052800	H-17	161	172	1	43	46	1	NON-FARM
71052800	H-21	166	173	1	93	95	1	NON-FARM
71052800	H-11	176	181	1	31	35	1	NON-FARM
71052800	H-11	184	189	1	28	29	1	NON-FARM
71052800	H-11	192	201	1	18	22	1	NON-FARM
71052800	H-7	227	235	1	89	91	1	NON-FARM
71052800	G-12	252	258	1	146	151	1	NON-FARM
71052800	G-13	347	353	1	71	73	1	NON-FARM
71052800	G-1	349	354	1	97	100	1	NON-FARM
71052800	G-1	355	359	1	100	115	1	NON-FARM
71052800	P-2	422	428	1	22	44	1	NON-FARM
71052800	P-2	431	449	1	37	46	1	NON-FARM
71052800	P-2	454	473	1	48	56	1	NON-FARM
71052800	P-4	478	483	1	31	34	1	NON-FARM
71052800	B-1	508	561	1	9	29	1	NON-FARM
71052800	B-1	516	525	1	67	73	1	NON-FARM
71052800	B-1	523	535	1	55	64	1	NON-FARM
71052800	B-1	532	539	1	45	52	1	NON-FARM
71052800		540	541	1	47	52	1	NON-FARM
71052800	B-1	540	545	1	33	46	1	NON-FARM
71052800	O-7	546	566	1	178	180	1	NON-FARM
71052800	O-5	548	574	1	151	153	1	NON-FARM
71052800	B-1	563	631	1	27	32	1	NON-FARM
71052800	B-1	564	568	1	9	22	1	NON-FARM
71052800	O-1	569	576	1	178	182	1	NON-FARM
71052800	O-2	579	586	1	178	182	1	NON-FARM
71052800	W-1	631	636	1	179	181	1	NON-FARM
71052800	Q-1	702	709	1	184	190	1	NON-FARM
71052800	O-10	733	739	1	176	183	1	NON-FARM
71052800	Q-7	744	751	1	128	133	1	NON-FARM
71052800	X-8	744	749	1	85	92	1	NON-FARM
71052800	Q-4	837	845	1	121	126	1	NON-FARM
71052800	DD-1	990	995	1	193	194	1	NON-FARM
71052800	DD-1	998	1004	1	194	202	1	NON-FARM
71052800	S-1	1072	1077	1	151	166	1	NON-FARM

TABLE D2 TEST AREAS FOR RUN 72064412

RUN NUMBER	ARFA DESIGNATION	FIRST LINE	LAST LINE	LINE INT	FIRST COL	LAST COL	COL INT	COVER TYPE
72064412	AG	59	62	1	307	317	1	AGRICULTURE
72064412	AG	68	77	1	426	432	1	AGRICULTURE
72064412	AG	74	78	1	369	375	1	AGRICULTURE
72064412	AG	79	87	1	262	265	1	AGRICULTURE
72064412	AG	84	90	1	273	285	1	AGRICULTURE
72064412	AG	100	110	1	274	283	1	AGRICULTURE
72064412	AG	106	124	1	234	244	1	AGRICULTURE
72064412	AG	111	114	1	245	257	1	AGRICULTURE
72064412	AG	113	119	1	276	288	1	AGRICULTURE
72064412	AG	126	130	1	252	267	1	AGRICULTURE
72064412	AG	164	174	1	586	591	1	AGRICULTURE
72064412	AG	180	187	1	369	373	1	AGRICULTURE
72064412	AG	194	199	1	264	267	1	AGRICULTURE
72064412	AG	197	205	1	233	238	1	AGRICULTURE
72064412	AG	202	209	1	247	253	1	AGRICULTURE
72064412	AG	212	217	1	296	302	1	AGRICULTURE
72064412	AG	217	221	1	275	280	1	AGRICULTURE
72064412	AG	219	222	1	443	454	1	AGRICULTURE
72064412	AG	229	232	1	395	403	1	AGRICULTURE
72064412	AG	247	251	1	268	274	1	AGRICULTURE
72064412	AG	250	253	1	488	497	1	AGRICULTURE
72064412	AG	257	264	1	236	240	1	AGRICULTURE
72064412	AG	257	263	1	291	298	1	AGRICULTURE
72064412	AG	283	289	1	501	506	1	AGRICULTURE
72064412	AG	300	302	1	514	521	1	AGRICULTURE
72064412	AG	308	315	1	314	319	1	AGRICULTURE
72064412	AG	319	326	1	302	306	1	AGRICULTURE
72064412	AG	320	324	1	366	369	1	AGRICULTURE
72064412	AG	350	354	1	233	238	1	AGRICULTURE
72064412	AG	376	380	1	400	404	1	AGRICULTURE
72064412	AG	384	388	1	524	530	1	AGRICULTURE
72064412	AG	400	404	1	538	542	1	AGRICULTURE
72064412	AG	406	411	1	353	356	1	AGRICULTURE
72064412	AG	416	418	1	372	375	1	AGRICULTURE
72064412	AG	422	424	1	524	536	1	AGRICULTURE
72064412	AG	430	433	1	385	394	1	AGRICULTURE
72064412	AG	460	464	1	265	274	1	AGRICULTURE
72064412	AG	460	467	1	495	500	1	AGRICULTURE
72064412	AG	461	464	1	293	298	1	AGRICULTURE
72064412	AG	464	470	1	320	327	1	AGRICULTURE
72064412	AG	479	482	1	452	455	1	AGRICULTURE
72064412	AG	488	491	1	397	402	1	AGRICULTURE
72064412	AG	499	503	1	487	492	1	AGRICULTURE
72064412	AG	508	510	1	328	333	1	AGRICULTURE
72064412	AG	541	545	1	483	490	1	AGRICULTURE
72064412	AG	563	567	1	431	440	1	AGRICULTURE
72064412	AG	573	577	1	352	358	1	AGRICULTURE
72064412	AG	578	585	1	326	331	1	AGRICULTURE
72064412	AG	580	586	1	348	352	1	AGRICULTURE
72064412	AG	582	588	1	438	443	1	AGRICULTURE
72064412	AG	586	591	1	510	515	1	AGRICULTURE
72064412	F-S	66	68	1	236	247	1	FOREST
72064412	F-T	85	96	1	367	403	1	FOREST
72064412	F-T	86	91	1	322	326	1	FOREST
72064412	F-T	89	98	1	313	321	1	FOREST
72064412	F-T	89	101	1	424	434	1	FOREST
72064412	F-T	90	103	1	330	357	1	FOREST
72064412	F-T	92	101	1	266	270	1	FOREST
72064412	F-T	94	104	1	232	236	1	FOREST
72064412	F-T	94	98	1	275	283	1	FOREST
72064412	F-T	97	107	1	361	375	1	FOREST
72064412	F-T	106	112	1	489	499	1	FOREST

TABLE D2, CONTINUED

72064412	F-S	132	136	1	340	360	1	FOREST
72064412	F-T	132	156	1	557	577	1	FOREST
72064412	F-S	139	147	1	486	493	1	FOREST
72064412	F-T	157	164	1	557	570	1	FOREST
72064412	F-T	170	181	1	542	553	1	FOREST
72064412	F	202	205	1	379	386	1	FOREST
72064412	F	213	221	1	380	385	1	FOREST
72064412	F-T	254	257	1	270	275	1	FOREST
72064412	F	321	335	1	579	598	1	FOREST
72064412	F-T	339	348	1	320	325	1	FOREST
72064412	F	379	388	1	572	578	1	FOREST
72064412	F	412	429	1	572	591	1	FOREST
72064412	F-S	436	443	1	305	310	1	FOREST
72064412	F-T	456	469	1	236	240	1	FOREST
72064412	F-T	456	469	1	584	590	1	FOREST
72064412	F-T	460	467	1	509	514	1	FOREST
72064412	F-S	514	522	1	384	391	1	FOREST
72064412	F-T	547	551	1	433	447	1	FOREST
72064412	F-S	556	565	1	268	285	1	FOREST
72064412	F-T	581	588	1	477	480	1	FOREST
72064412	RM	57	72	1	484	499	1	RECENT-MINE
72064412	RM	68	72	1	436	439	1	RECENT-MINE
72064412	PRVGM	81	87	1	496	501	1	RECENT-MINE
72064412	RM	153	157	1	354	359	1	RECENT-MINE
72064412	RM	166	174	1	382	391	1	RECENT-MINE
72064412	PRVGM	170	176	1	518	525	1	RECENT-MINE
72064412	PRVGM	172	174	1	317	324	1	RECENT-MINE
72064412	RM	175	179	1	388	394	1	RECENT-MINE
72064412	PRVGM	181	189	1	498	505	1	RECENT-MINE
72064412	PRVGM	185	188	1	224	229	1	RECENT-MINE
72064412	RM	189	191	1	327	334	1	RECENT-MINE
72064412	PRVGM	192	194	1	230	236	1	RECENT-MINE
72064412	RM	236	242	1	325	329	1	RECENT-MINE
72064412	PRVGM	331	335	1	453	459	1	RECENT-MINE
72064412	PRVGM	339	346	1	460	469	1	RECENT-MINE
72064412	PRVGM	362	369	1	302	309	1	RECENT-MINE
72064412	PRVGM	366	373	1	315	321	1	RECENT-MINE
72064412	PRVGM	378	383	1	343	349	1	RECENT-MINE
72064412	PRVGM	385	389	1	333	337	1	RECENT-MINE
72064412	PRVGM	404	413	1	267	270	1	RECENT-MINE
72064412	PRVGM	404	408	1	302	315	1	RECENT-MINE
72064412	PRVGM	407	415	1	254	259	1	RECENT-MINE
72064412	RM	412	415	1	289	293	1	RECENT-MINE
72064412	RM	414	416	1	279	285	1	RECENT-MINE
72064412	RM	423	428	1	316	320	1	RECENT-MINE
72064412	RM	440	445	1	323	330	1	RECENT-MINE
72064412	RM	445	465	1	388	393	1	RECENT-MINE
72064412	RM	445	447	1	515	521	1	RECENT-MINE
72064412	RM	466	469	1	385	392	1	RECENT-MINE
72064412	RM	480	485	1	294	313	1	RECENT-MINE
72064412	PRVGM	492	494	1	297	315	1	RECENT-MINE
72064412	PRVGM	497	500	1	299	310	1	RECENT-MINE
72064412	PRVGM	506	521	1	306	311	1	RECENT-MINE
72064412	TIP	58	62	1	465	467	1	PIT
72064412	W-S	62	63	1	519	520	1	PIT
72064412	W-S	64	65	1	534	535	1	PIT
72064412	TIP	76	79	1	329	332	1	PIT
72064412	W-S	80	80	1	407	408	1	PIT
72064412	W-S	80	80	1	534	536	1	PIT
72064412	W-S	81	81	1	409	410	1	PIT
72064412	W-C	151	154	1	298	299	1	PIT
72064412	W-C	153	156	1	516	518	1	PIT
72064412	W-C	186	190	1	315	320	1	PIT
72064412	TIP	236	239	1	416	418	1	PIT
72064412	TIP	240	243	1	410	413	1	PIT

TABLE D2, CONTINUED

72064412	TIP	245	249	1	416	417	1	PIT
72064412	TIP	248	249	1	315	320	1	PIT
72064412	TIP	251	252	1	324	326	1	PIT
72064412	W-C	251	256	1	447	448	1	PIT
72064412	TIP	390	396	1	288	291	1	PIT
72064412	W-C	391	393	1	320	322	1	PIT
72064412	TIP	392	397	1	292	302	1	PIT
72064412	TIP	394	396	1	303	308	1	PIT
72064412	W-C	469	471	1	419	420	1	PIT
72064412	W-C	565	568	1	408	411	1	PIT
72064412	RVGM	74	78	1	391	398	1	REVEG. MINE
72064412	RVGM	76	79	1	383	388	1	REVEG. MINE
72064412	RVGM	115	127	1	354	355	1	REVEG. MINE
72064412	RVGM	149	152	1	488	496	1	REVEG. MINE
72064412	RVGM	150	159	1	467	470	1	REVEG. MINE
72064412	RVGM	161	166	1	521	526	1	REVEG. MINE
72064412	RVGM	164	170	1	471	477	1	REVEG. MINE
72064412	RVGM	176	181	1	444	453	1	REVEG. MINE
72064412	RVGM	192	196	1	438	445	1	REVEG. MINE
72064412	RVGM	209	215	1	528	531	1	REVEG. MINE
72064412	RVGM	217	220	1	480	487	1	REVEG. MINE
72064412	RVGM	221	225	1	342	349	1	REVEG. MINE
72064412	RVGM	224	228	1	323	333	1	REVEG. MINE
72064412	RVGM	240	244	1	349	352	1	REVEG. MINE
72064412	RVGM	263	266	1	405	417	1	REVEG. MINE
72064412	RVGM	453	455	1	465	470	1	REVEG. MINE
72064412	RVGM	459	461	1	338	343	1	REVEG. MINE
72064412	RVGM	464	467	1	345	348	1	REVEG. MINE
72064412	RVGM	464	467	1	458	462	1	REVEG. MINE
72064412	RVGM	467	471	1	464	467	1	REVEG. MINE
72064412	RVGM	532	538	1	345	351	1	REVEG. MINE
72064412	RVGM	536	548	1	357	362	1	REVEG. MINE
72064412	RVGM	543	550	1	302	303	1	REVEG. MINE
72064412	RVGM	549	553	1	331	336	1	REVEG. MINE
72064412	RVGM	556	559	1	321	328	1	REVEG. MINE
72064412	RVGM	565	567	1	312	318	1	REVEG. MINE
72064412	RVGM	570	573	1	264	282	1	REVEG. MINE
72064412	RVGM	575	579	1	308	321	1	REVEG. MINE

TABLE D3 REFERENCE AREAS FOR RUN 71052501

RUN NUMBER	ARFA DESIGNATION	FIRST LINE	LAST LINE	LINE INT	FIRST COL	LAST COL	COL INT	AREA TYPE
71052501	TRAINING	173	188	1	186	192	1	DECIDUOUS
71052501	TRAINING	519	525	1	026	032	1	DECIDUOUS
71052501	TRAINING	046	049	1	067	070	1	CONIFEROUS
71052501	TRAINING	048	048	1	076	078	1	CONIFEROUS
71052501	TRAINING	052	053	1	066	067	1	CONIFEROUS
71052501	TRAINING	055	055	1	073	077	1	CONIFEROUS
71052501	TRAINING	068	072	1	052	052	1	CONIFEROUS
71052501	TRAINING	333	339	1	083	088	1	CONIFEROUS
71052501	TRAINING	396	399	1	109	126	1	CONIFEROUS
71052501	TRAINING	146	148	1	108	112	1	WATER,POND
71052501	TRAINING	151	155	1	113	119	1	WATER,POND
71052501	TRAINING	155	159	1	120	129	1	WATER,POND
71052501	TRAINING	155	156	1	023	024	1	WATER,POND
71052501	TRAINING	227	228	1	052	053	1	WATER,POND
71052501	TRAINING	337	338	1	108	111	1	WATER,POND
71052501	TRAINING	831	834	1	011	014	1	WATER,POND
71052501	TRAINING	843	846	1	004	008	1	WATER,POND
71052501	TRAINING	847	849	1	002	005	1	WATER,POND
71052501	TRAINING	877	879	1	041	045	1	WATER,POND
71052501	TRAINING	1174	1174	1	197	202	1	WATER,RIVER
71052501	TRAINING	1175	1175	1	011	020	1	WATER,RIVER
71052501	TRAINING	1184	1185	1	030	032	1	WATER,RIVER
71052501	TRAINING	1216	1217	1	068	073	1	WATER,RIVER
71052501	TRAINING	1217	1219	1	075	080	1	WATER,RIVER
71052501	TRAINING	1239	1245	1	077	080	1	WATER,RIVER
71052501	TRAINING	1273	1282	1	074	076	1	WATER,RIVER
71052501	TRAINING	1290	1294	1	082	085	1	WATER,RIVER
71052501	TRAINING	1291	1293	1	140	143	1	WATER,RIVER
71052501	TRAINING	32	35	1	50	57	1	FORAGE,PAST
71052501	TRAINING	56	67	1	188	195	1	FORAGE,PAST
71052501	TRAINING	105	111	1	189	197	1	FORAGE,STUB
71052501	TRAINING	115	127	1	187	193	1	FORAGE,STUB
71052501	TRAINING	196	204	1	180	192	1	FORAGE,PAST
71052501	TRAINING	335	339	1	197	204	1	FORAGE,PAST
71052501	TRAINING	405	419	1	211	215	1	FORAGE,PAST
71052501	TRAINING	506	510	1	073	086	1	FORAGE,PAST
71052501	TRAINING	564	574	1	205	217	1	FORAGE,PAST
71052501	TRAINING	650	654	1	048	056	1	FORAGE,PAST
71052501	TRAINING	678	681	1	173	185	1	FORAGE,STUB
71052501	TRAINING	817	825	1	100	118	1	FORAGE,HAY
71052501	TRAINING	853	859	1	144	171	1	FORAGE,HAY
71052501	TRAINING	1143	1152	1	047	058	1	FORAGE,HAY
71052501	TRAINING	1189	1193	1	184	208	1	FORAGE,STUB
71052501	TRAINING	1202	1209	1	184	208	1	FORAGE,STUB
71052501	TRAINING	1275	1284	1	109	125	1	FORAGE,STUB
71052501	TRAINING	1449	1452	1	127	142	1	FORAGE,HAY
71052501	TRAINING	661	670	1	174	181	1	CORN
71052501	TRAINING	681	685	1	137	143	1	CORN
71052501	TRAINING	683	685	1	153	167	1	CORN
71052501	TRAINING	698	703	1	141	143	1	CORN
71052501	TRAINING	1250	1257	1	93	100	1	CORN
71052501	TRAINING	1275	1282	1	93	100	1	CORN
71052501	TRAINING	1403	1412	1	194	213	1	CORN
71052501	TRAINING	1481	1485	1	033	042	1	CORN
71052501	TRAINING	1487	1512	1	049	051	1	CORN
71052501	TRAINING	1492	1506	1	033	042	1	CORN
71052501	TRAINING	498	500	1	099	117	1	SOY
71052501	TRAINING	1027	1031	1	005	025	1	SOY

TABLE D3, CONTINUED

71052501	TRAINING	1045	1048	1	124	128	1	SOY
71052501	TRAINING	1191	1193	1	043	046	1	SOY
71052501	TRAINING	1253	1274	1	129	135	1	SOY
71052501	TRAINING	1473	1479	1	078	090	1	SOY
71052501	TFST	1	025	1	061	157	1	DECIDUOUS
71052501	TFST	43	051	1	122	156	1	DECIDUOUS
71052501	TFST	48	088	1	161	179	1	DECIDUOUS
71052501	TFST	72	83	1	113	137	1	DECIDUOUS
71052501	TFST	77	107	1	001	022	1	DECIDUOUS
71052501	TFST	88	110	1	107	138	1	DECIDUOUS
71052501	TFST	91	122	1	160	178	1	DECIDUOUS
71052501	TFST	98	134	1	073	093	1	DECIDUOUS
71052501	TFST	116	161	1	006	019	1	DECIDUOUS
71052501	TFST	123	144	1	115	128	1	DECIDUOUS
71052501	TFST	149	158	1	059	103	1	DECIDUOUS
71052501	TFST	150	190	1	212	222	1	DECIDUOUS
71052501	TFST	191	209	1	123	160	1	DECIDUOUS
71052501	TFST	220	261	1	200	222	1	DECIDUOUS
71052501	TFST	245	286	1	140	155	1	DECIDUOUS
71052501	TFST	246	262	1	177	199	1	DECIDUOUS
71052501	TFST	256	295	1	043	054	1	DECIDUOUS
71052501	TFST	297	316	1	025	064	1	DECIDUOUS
71052501	TFST	312	325	1	120	184	1	DECIDUOUS
71052501	TFST	317	361	1	019	040	1	DECIDUOUS
71052501	TFST	339	352	1	166	192	1	DECIDUOUS
71052501	TFST	342	366	1	093	133	1	DECIDUOUS
71052501	TFST	371	390	1	115	188	1	DECIDUOUS
71052501	TFST	386	403	1	42	58	1	DECIDUOUS
71052501	TFST	386	404	1	060	068	1	DECIDUOUS
71052501	TFST	391	402	1	145	172	1	DECIDUOUS
71052501	TFST	408	429	1	160	183	1	DECIDUOUS
71052501	TFST	431	455	1	141	148	1	DECIDUOUS
71052501	TFST	437	446	1	149	153	1	DECIDUOUS
71052501	TFST	464	491	1	199	210	1	DECIDUOUS
71052501	TFST	466	475	1	001	027	1	DECIDUOUS
71052501	TFST	508	526	1	099	126	1	DECIDUOUS
71052501	TFST	519	540	1	033	040	1	DECIDUOUS
71052501	TFST	575	578	1	177	181	1	DECIDUOUS
71052501	TFST	575	602	1	161	176	1	DECIDUOUS
71052501	TFST	631	647	1	190	222	1	DECIDUOUS
71052501	TFST	635	649	1	058	073	1	DECIDUOUS
71052501	TFST	657	677	1	107	126	1	DECIDUOUS
71052501	TFST	785	826	1	136	203	1	DECIDUOUS
71052501	TFST	814	854	1	031	051	1	DECIDUOUS
71052501	TFST	858	870	1	180	222	1	DECIDUOUS
71052501	TFST	935	974	1	046	111	1	DECIDUOUS
71052501	TFST	1007	1066	1	183	211	1	DECIDUOUS
71052501	TFST	1079	1098	1	099	130	1	DECIDUOUS
71052501	TFST	1197	1218	1	094	115	1	DECIDUOUS
71052501	TFST	1354	1369	1	189	222	1	DECIDUOUS
71052501	TFST	344	348	1	074	077	1	CONIFEROUS
71052501	TFST	409	411	1	145	150	1	CONIFEROUS
71052501	TFST	413	420	1	141	144	1	CONIFEROUS
71052501	TFST	512	514	1	144	149	1	CONIFEROUS
71052501	TFST	294	295	1	158	159	1	WATER, POND
71052501	TFST	834	837	1	004	010	1	WATER, POND
71052501	TFST	1191	1192	1	164	167	1	WATER, RIVER
71052501	TFST	1195	1200	1	161	162	1	WATER, RIVER
71052501	TFST	1206	1215	1	163	166	1	WATER, RIVER
71052501	TFST	1213	1214	1	058	063	1	WATER, RIVER
71052501	TFST	1214	1227	1	007	007	1	WATER, RIVER
71052501	TFST	1216	1223	1	162	165	1	WATER, RIVER
71052501	TFST	1219	1230	1	083	086	1	WATER, RIVER
71052501	TFST	1224	1238	1	161	164	1	WATER, RIVER

TABLE D3, CONTINUED

71052501	TFST	1252	1259	1	080	082	1	WATER, RIVER
71052501	TFST	1261	1267	1	157	159	1	WATER, RIVER
71052501	TFST	1279	1282	1	153	155	1	WATER, RIVER
71052501	TFST	1285	1287	1	149	150	1	WATER, RIVER
71052501	TFST	1297	1299	1	121	126	1	WATER, RIVER
71052501	TFST	10	15	1	023	030	1	FORAGE, PAST
71052501	TFST	94	99	1	185	192	1	FORAGE, STUB
71052501	TFST	121	142	1	215	222	1	FORAGE, PAST
71052501	TFST	135	140	1	163	174	1	FORAGE, PAST
71052501	TFST	141	144	1	153	164	1	FORAGE, PAST
71052501	TFST	164	169	1	179	184	1	FORAGE, PAST
71052501	TFST	164	173	1	170	176	1	FORAGE, STUB
71052501	TFST	182	187	1	115	135	1	FORAGE, PAST
71052501	TFST	205	211	1	002	007	1	FORAGE, PAST
71052501	TFST	216	219	1	128	134	1	FORAGE, PAST
71052501	TFST	230	237	1	122	133	1	FORAGE, PAST
71052501	TFST	235	241	1	157	164	1	FORAGE, PAST
71052501	TFST	272	280	1	211	222	1	FORAGE, PAST
71052501	TFST	288	297	1	124	132	1	FORAGE, PAST
71052501	TFST	291	296	1	140	149	1	FORAGE, PAST
71052501	TFST	293	302	1	213	215	1	FORAGE, STUB
71052501	TFST	300	308	1	153	157	1	FORAGE, PAST
71052501	TFST	330	337	1	047	051	1	FORAGE, PAST
71052501	TFST	350	357	1	145	160	1	FORAGE, PAST
71052501	TFST	382	399	1	197	208	1	FORAGE, PAST
71052501	TFST	408	417	1	051	058	1	FORAGE, PAST
71052501	TFST	420	431	1	208	215	1	FORAGE, PAST
71052501	TFST	426	429	1	027	035	1	FORAGE, PAST
71052501	TFST	440	451	1	122	131	1	FORAGE, STUB
71052501	TFST	447	462	1	167	172	1	FORAGE, STUB
71052501	TFST	458	467	1	113	119	1	FORAGE, PAST
71052501	TFST	463	469	1	175	182	1	FORAGE, PAST
71052501	TFST	471	483	1	173	184	1	FORAGE, PAST
71052501	TFST	475	495	1	056	067	1	FORAGE, PAST
71052501	TFST	486	496	1	151	159	1	FORAGE, PAST
71052501	TFST	491	501	1	76	94	1	FORAGE, PAST
71052501	TFST	499	508	1	023	028	1	FORAGE, PAST
71052501	TFST	499	509	1	042	052	1	FORAGE, PAST
71052501	TFST	529	532	1	108	120	1	FORAGE, STUB
71052501	TFST	532	550	1	088	102	1	FORAGE, PAST
71052501	TFST	543	551	1	127	140	1	FORAGE, PAST
71052501	TFST	559	565	1	24	35	1	FORAGE, PAST
71052501	TFST	564	570	1	127	165	1	FORAGE, PAST
71052501	TFST	572	580	1	048	055	1	FORAGE, PAST
71052501	TFST	582	591	1	206	222	1	FORAGE, PAST
71052501	TFST	619	628	1	084	099	1	FORAGE, PAST
71052501	TFST	634	647	1	137	147	1	FORAGE, PAST
71052501	TFST	649	663	1	133	169	1	FORAGE, PAST
71052501	TFST	649	655	1	195	210	1	FORAGE, PAST
71052501	TFST	689	704	1	110	113	1	FORAGE, STUB
71052501	TFST	693	704	1	115	125	1	FORAGE, PAST
71052501	TFST	696	703	1	088	093	1	FORAGE, STUB
71052501	TFST	700	703	1	178	203	1	FORAGE, PAST
71052501	TFST	725	733	1	022	032	1	FORAGE, STUB
71052501	TFST	881	888	1	106	132	1	FORAGE, PAST
71052501	TFST	908	916	1	029	039	1	FORAGE, PAST
71052501	TFST	954	968	1	169	181	1	FORAGE, PAST
71052501	TFST	969	976	1	153	172	1	FORAGE, PAST
71052501	TFST	985	990	1	106	130	1	FORAGE, PAST
71052501	TFST	1016	1021	1	060	079	1	FORAGE, STUB
71052501	TFST	1021	1039	1	032	041	1	FORAGE, PAST
71052501	TFST	1062	1068	1	156	172	1	FORAGE, STUB
71052501	TFST	1075	1084	1	146	163	1	FORAGE, PAST
71052501	TFST	1089	1098	1	059	083	1	FORAGE, PAST
71052501	TFST	1123	1129	1	117	133	1	FORAGE, PAST
71052501	TFST	1131	1140	1	123	147	1	FORAGE, PAST

TABLE D3, CONTINUED

71052501	TFST	1131	1142	1	047	058	1	FORAGE,HAY
71052501	TFST	1162	1175	1	166	181	1	FORAGE,PAST
71052501	TFST	1179	1186	1	085	102	1	FORAGE,PAST
71052501	TFST	1196	1206	1	134	146	1	FORAGE,PAST
71052501	TFST	1199	1209	1	173	183	1	FORAGE,STUB
71052501	TFST	1221	1252	1	061	069	1	FORAGE,STUB
71052501	TFST	1252	1274	1	109	125	1	FORAGE,STUB
71052501	TFST	1313	1317	1	164	180	1	FORAGE,PAST
71052501	TFST	1343	1356	1	050	059	1	FORAGE,PAST
71052501	TFST	1360	1384	1	166	180	1	FORAGE,PAST
71052501	TFST	1390	1400	1	184	222	1	FORAGE,PAST
71052501	TFST	1402	1415	1	176	189	1	FORAGE,STUB
71052501	TFST	1403	1416	1	130	150	1	FORAGE,HAY
71052501	TFST	1455	1464	1	118	131	1	FORAGE,STUB
71052501	TFST	1474	1483	1	001	010	1	FORAGE,HAY
71052501	TFST	1482	1489	1	140	210	1	FORAGE,PAST
71052501	TFST	1496	1501	1	107	126	1	FORAGE,HAY
71052501	TFST	1499	1517	1	161	174	1	FORAGE,HAY
71052501	TFST	1510	1516	1	133	148	1	FORAGE,HAY
71052501	TFST	650	660	1	174	181	1	CORN
71052501	TFST	698	700	1	154	169	1	CORN
71052501	TFST	707	714	1	134	146	1	CORN
71052501	TFST	924	934	1	159	169	1	CORN
71052501	TFST	1212	1248	1	172	193	1	CORN
71052501	TFST	1417	1438	1	157	167	1	CORN
71052501	TFST	1482	1512	1	052	082	1	CORN
71052501	TFST	1482	1492	1	085	103	1	CORN
71052501	TFST	1503	1512	1	152	157	1	CORN
71052501	TFST	1510	1513	1	196	203	1	CORN
71052501	TFST	294	318	1	218	222	1	SOY
71052501	TFST	706	711	1	190	194	1	SOY
71052501	TFST	921	926	1	109	124	1	SOY
71052501	TFST	1057	1059	1	139	157	1	SOY
71052501	TFST	1101	1112	1	158	171	1	SOY
71052501	TFST	1131	1136	1	032	043	1	SOY
71052501	TFST	1137	1155	1	017	042	1	SOY
71052501	TFST	1140	1165	1	010	012	1	SOY
71052501	TFST	1144	1161	1	001	003	1	SOY
71052501	TFST	1156	1165	1	017	023	1	SOY
71052501	TFST	1175	1179	1	030	033	1	SOY
71052501	TFST	1186	1189	1	039	042	1	SOY
71052501	TFST	1192	1196	1	013	026	1	SOY
71052501	TFST	1198	1208	1	15	28	1	SOY
71052501	TFST	1215	1218	1	046	051	1	SOY
71052501	TFST	1220	1222	1	011	030	1	SOY
71052501	TFST	1258	1264	1	017	029	1	SOY
71052501	TFST	1420	1442	1	129	140	1	SOY
71052501	TFST	1420	1429	1	143	152	1	SOY
71052501	TFST	1431	1443	1	143	150	1	SOY
71052501	TFST	1466	1478	1	033	045	1	SOY
71052501	TFST	1477	1479	1	049	076	1	SOY
71052501	TFST	1508	1514	1	107	122	1	SOY
71052501	TFST	1537	1544	1	162	180	1	SOY

TABLE D4 TEST AREAS FOR RUN 72032803

RUN NUMBER	AREA DESIGNATION	FIRST LINE	LAST LINE	LINE INT	FIRST COL	LAST COL	COL INT	COVER TYPE
72032803	135	268	271	1	1434	1439	1	CORN
72032803	138	268	273	1	1470	1486	1	CORN
72032803	132	272	279	1	1381	1389	1	CORN
72032803	134	275	279	1	1427	1433	1	CORN
72032803	133	282	284	1	1383	1391	1	CORN
72032803	139	287	292	1	1460	1465	1	CORN
72032803	141	297	302	1	1480	1487	1	CORN
72032803	143	316	319	1	1452	1456	1	CORN
72032803	144	320	323	1	1453	1457	1	CORN
72032803	148	322	325	1	1483	1489	1	CORN
72032803	156	329	336	1	1381	1387	1	CORN
72032803	146	333	338	1	1460	1484	1	CORN
72032803	145	334	336	1	1452	1455	1	CORN
72032803	154	338	344	1	1431	1432	1	CORN
72032803	152	339	343	1	1440	1444	1	CORN
72032803	151	340	344	1	1454	1458	1	CORN
72032803	160	350	352	1	1418	1427	1	CORN
72032803	162A	355	359	1	1443	1447	1	CORN
72032803	0-154	356	359	1	1248	1256	1	CORN
72032803	0-157	358	360	1	1265	1267	1	CORN
72032803	159A	359	364	1	1402	1407	1	CORN
72032803	158A	360	364	1	1392	1398	1	CORN
72032803	162B	360	362	1	1444	1448	1	CORN
72032803	0-155	360	363	1	1249	1258	1	CORN
72032803	0-153	361	367	1	1229	1235	1	CORN
72032803	0-151	363	369	1	1209	1216	1	CORN
72032803	159B	365	369	1	1404	1411	1	CORN
72032803	158B	365	369	1	1394	1399	1	CORN
72032803	0-160	365	368	1	1259	1263	1	CORN
72032803	166A	366	375	1	1447	1453	1	CORN
72032803	0-158	368	371	1	1246	1251	1	CORN
72032803	0-119A	369	372	1	1170	1173	1	CORN
72032803	0-170B	369	371	1	1276	1278	1	CORN
72032803	167	371	376	1	1415	1431	1	CORN
72032803	0-163	371	373	1	1261	1264	1	CORN
72032803	0-119B	373	377	1	1171	1174	1	CORN
72032803	172	374	381	1	1394	1402	1	CORN
72032803	164	375	378	1	1474	1481	1	CORN
72032803	0-166	375	377	1	1263	1266	1	CORN
72032803	166B	376	381	1	1450	1458	1	CORN
72032803	0-118A	376	379	1	1156	1161	1	CORN
72032803	0-118B	376	378	1	1162	1169	1	CORN
72032803	168	377	386	1	1422	1433	1	CORN
72032803	0-140	377	379	1	1190	1191	1	CORN
72032803	0-123	379	381	1	1179	1180	1	CORN
72032803	0-167	379	382	1	1264	1274	1	CORN
72032803	0-76	380	383	1	1135	1137	1	CORN
72032803	0-122A	381	384	1	1164	1177	1	CORN
72032803	0-171	382	389	1	1281	1288	1	CORN
72032803	165	383	387	1	1482	1488	1	CORN
72032803	0-170A	383	388	1	1275	1276	1	CORN
72032803	0-80 A	384	387	1	1147	1148	1	CORN
72032803	0-22	385	390	1	1091	1094	1	CORN
72032803	0-122B	385	386	1	1173	1183	1	CORN
72032803	0-28	385	386	1	1116	1119	1	CORN
72032803	0-79	386	392	1	1133	1141	1	CORN
72032803	0-169	386	387	1	1253	1262	1	CORN
72032803	0-25	387	388	1	1102	1105	1	CORN
72032803	0-145	387	389	1	1205	1213	1	CORN
72032803	170	388	389	1	1411	1419	1	CORN
72032803	0-80 B	388	393	1	1149	1150	1	CORN
72032803	0-168	388	392	1	1246	1249	1	CORN
72032803	0-29	390	393	1	1119	1122	1	CORN

TABLE D4, CONTINUED

72032803	0-11 A	391	392	1	1049	1050	1	CORN
72032803	0-13 A	391	392	1	1057	1057	1	CORN
72032803	0-83 A	391	394	1	1154	1157	1	CORN
72032803	0-147	391	392	1	1206	1211	1	CORN
72032803	0-173	391	394	1	1259	1259	1	CORN
72032803	169	392	398	1	1415	1423	1	CORN
72032803	0-26	392	401	1	1106	1114	1	CORN
72032803	0-124	392	395	1	1179	1185	1	CORN
72032803	0-11 B	393	394	1	1050	1050	1	CORN
72032803	0-13 B	393	394	1	1058	1058	1	CORN
72032803	0-23	393	396	1	1095	1096	1	CORN
72032803	0-150	393	394	1	1219	1222	1	CORN
72032803	0-7 A	395	396	1	1023	1024	1	CORN
72032803	0-11 C	395	396	1	1051	1051	1	CORN
72032803	0-13 C	395	396	1	1059	1059	1	CORN
72032803	0-83 R	395	401	1	1157	1158	1	CORN
72032803	0-148	395	396	1	1207	1211	1	CORN
72032803	0-7 B	397	398	1	1024	1025	1	CORN
72032803	0-149A	397	403	1	1193	1201	1	CORN
72032803	0-174	397	398	1	1257	1267	1	CORN
72032803	0-4	398	401	1	1014	1020	1	CORN
72032803	0-176	398	401	1	1270	1278	1	CORN
72032803	0-7 C	399	400	1	1026	1026	1	CORN
72032803	0-125	399	401	1	1184	1189	1	CORN
72032803	0-16 A	400	403	1	1055	1061	1	CORN
72032803	0-178	401	405	1	1253	1254	1	CORN
72032803	0-215	401	405	1	1287	1298	1	CORN
72032803	0-126	402	406	1	1165	1167	1	CORN
72032803	0-15	403	407	1	1041	1044	1	CORN
72032803	0-9 C	404	405	1	1034	1037	1	CORN
72032803	0-85	404	407	1	1157	1162	1	CORN
72032803	0-149B	404	407	1	1195	1200	1	CORN
72032803	0-9 B	405	408	1	1028	1031	1	CORN
72032803	0-16 B	405	406	1	1057	1066	1	CORN
72032803	0-9 A	406	409	1	1020	1026	1	CORN
72032803	0-5	407	412	1	1015	1017	1	CORN
72032803	0-3	409	411	1	1000	1001	1	CORN
72032803	0-180	409	412	1	1247	1252	1	CORN
72032803	0-9 D	410	412	1	1028	1031	1	CORN
72032803	0-17 A	410	412	1	1051	1061	1	CORN
72032803	0-34	410	411	1	1118	1121	1	CORN
72032803	0-210	412	414	1	1291	1295	1	CORN
72032803	0-17 B	413	417	1	1052	1055	1	CORN
72032803	0-32	413	415	1	1104	1107	1	CORN
72032803	0-33	413	418	1	1112	1115	1	CORN
72032803	0-181	413	414	1	1226	1234	1	CORN
72032803	0-211A	413	417	1	1283	1288	1	CORN
72032803	0-9 F	414	416	1	1030	1033	1	CORN
72032803	0-20	414	416	1	1066	1069	1	CORN
72032803	0-209	415	419	1	1292	1301	1	CORN
72032803	0-18	417	419	1	1046	1049	1	CORN
72032803	0-21	418	421	1	1076	1078	1	CORN
72032803	0-211B	418	420	1	1285	1290	1	CORN
72032803	0-208A	421	422	1	1300	1305	1	CORN
72032803	0-184	422	425	1	1230	1232	1	CORN
72032803	0-208B	422	423	1	1293	1298	1	CORN
72032803	0-87	424	432	1	1160	1163	1	CORN
72032803	0-202A	425	432	1	1283	1292	1	CORN
72032803	0-207	425	429	1	1296	1308	1	CORN
72032803	0-186	427	432	1	1210	1215	1	CORN
72032803	0-189	427	430	1	1233	1237	1	CORN
72032803	0-138A	429	432	1	1175	1179	1	CORN
72032803	0-188	431	432	1	1221	1227	1	CORN
72032803	0-206	432	435	1	1298	1300	1	CORN
72032803	0-66	433	438	1	1106	1115	1	CORN
72032803	0-93	433	439	1	1171	1173	1	CORN
72032803	0-138B	433	439	1	1177	1180	1	CORN

TABLE D4, CONTINUED

72032803	O-202B	433	436	1	1291	1295	1	CORN
72032803	O-86	434	435	1	1133	1134	1	CORN
72032803	O-92 B	434	437	1	1162	1167	1	CORN
72032803	O-204B	434	437	1	1288	1289	1	CORN
72032803	O-193	435	436	1	1226	1230	1	CORN
72032803	O-92 A	436	438	1	1153	1161	1	CORN
72032803	O-127	436	438	1	1185	1189	1	CORN
72032803	O-72 B	437	437	1	1128	1130	1	CORN
72032803	O-192	437	438	1	1220	1223	1	CORN
72032803	O-202C	437	441	1	1293	1296	1	CORN
72032803	O-205	437	440	1	1300	1301	1	CORN
72032803	O-72 A	438	438	1	1123	1125	1	CORN
72032803	O-132	438	439	1	1201	1203	1	CORN
72032803	O-204A	438	441	1	1289	1290	1	CORN
72032803	O-89	439	441	1	1136	1143	1	CORN
72032803	O-99	439	443	1	1164	1169	1	CORN
72032803	O-131A	439	440	1	1195	1199	1	CORN
72032803	O-199	439	443	1	1265	1272	1	CORN
72032803	O-98	440	447	1	1157	1161	1	CORN
72032803	O-198	440	447	1	1258	1261	1	CORN
72032803	O-71	441	442	1	1127	1132	1	CORN
72032803	O-131B	441	442	1	1195	1203	1	CORN
72032803	O-69	442	444	1	1119	1120	1	CORN
72032803	O-70	442	443	1	1123	1124	1	CORN
72032803	O-130	442	448	1	1181	1191	1	CORN
72032803	O-48	443	445	1	1073	1080	1	CORN
72032803	O-94	443	446	1	1138	1145	1	CORN
72032803	O-67 B	444	446	1	1109	1111	1	CORN
72032803	O-101	444	445	1	1172	1176	1	CORN
72032803	O-100	445	446	1	1166	1170	1	CORN
72032803	O-67 A	447	450	1	1111	1112	1	CORN
72032803	O-106A	448	450	1	1176	1178	1	CORN
72032803	O-195	448	449	1	1245	1249	1	CORN
72032803	O-37	450	450	1	1043	1054	1	CORN
72032803	O-105	450	459	1	1165	1174	1	CORN
72032803	O-74	451	458	1	1062	1070	1	CORN
72032803	O-106B	451	455	1	1178	1179	1	CORN
72032803	O-194B	452	455	1	1238	1242	1	CORN
72032803	O-194A	453	457	1	1233	1237	1	CORN
72032803	O-38	454	455	1	1046	1049	1	CORN
72032803	O-39	455	460	1	1054	1058	1	CORN
72032803	O-104A	455	462	1	1138	1143	1	CORN
72032803	O-36	456	457	1	1034	1041	1	CORN
72032803	O-59	458	460	1	1103	1109	1	CORN
72032803	O-134	460	461	1	1186	1190	1	CORN
72032803	O-41	461	465	1	1050	1053	1	CORN
72032803	O-104B	463	469	1	1139	1147	1	CORN
72032803	O-135	463	468	1	1188	1191	1	CORN
72032803	O-45	473	479	1	1057	1072	1	CORN
72032803	O-51	474	476	1	1084	1090	1	CORN
72032803	O-137	475	478	1	1192	1200	1	CORN
72032803	O-44	476	480	1	1048	1051	1	CORN
72032803	O-109	478	486	1	1144	1154	1	CORN
72032803	O-54 A	479	481	1	1091	1093	1	CORN
72032803	O-108	481	487	1	1130	1139	1	CORN
72032803	O-54 B	482	485	1	1093	1094	1	CORN
72032803	O-117	484	490	1	1182	1191	1	CORN
72032803	O-114	487	489	1	1160	1170	1	CORN
72032803	O-113B	488	491	1	1155	1158	1	CORN
72032803	O-113A	489	491	1	1150	1154	1	CORN
72032803	O-112	490	491	1	1141	1148	1	CORN
72032803	O-111B	492	492	1	1132	1137	1	CORN
72032803	O-111A	493	494	1	1127	1133	1	CORN
72032803	O-58	495	497	1	1119	1124	1	CORN
72032803	4A	524	526	1	1450	1453	1	CORN
72032803	4B	527	529	1	1451	1454	1	CORN
72032803	2A	529	530	1	1434	1437	1	CORN

TABLE D4, CONTINUED

72032803	2B	531	534	1	1436	1439	1	CORN
72032803	3A	531	533	1	1429	1433	1	CORN
72032803	3B	534	536	1	1430	1434	1	CORN
72032803	75A	534	536	1	1716	1717	1	CORN
72032803	2C	535	538	1	1437	1441	1	CORN
72032803	3C	537	539	1	1431	1435	1	CORN
72032803	75B	537	539	1	1717	1718	1	CORN
72032803	115	537	539	1	1744	1751	1	CORN
72032803	72	539	543	1	1691	1698	1	CORN
72032803	3D	540	543	1	1432	1436	1	CORN
72032803	66	542	550	1	1648	1657	1	CORN
72032803	77A	542	546	1	1736	1741	1	CORN
72032803	6	544	545	1	1441	1445	1	CORN
72032803	5	545	546	1	1433	1438	1	CORN
72032803	70C	546	550	1	1690	1694	1	CORN
72032803	61	546	552	1	1610	1619	1	CORN
72032803	65	546	552	1	1635	1643	1	CORN
72032803	77B	547	551	1	1738	1743	1	CORN
72032803	7B	548	550	1	1434	1438	1	CORN
72032803	8A	548	550	1	1442	1452	1	CORN
72032803	70B	548	555	1	1678	1689	1	CORN
72032803	7A	549	551	1	1429	1433	1	CORN
72032803	70A	550	555	1	1668	1677	1	CORN
72032803	8B	551	553	1	1443	1448	1	CORN
72032803	9B	551	554	1	1438	1440	1	CORN
72032803		552	552	1	1527	1528	1	CORN
72032803	47	552	557	1	1524	1526	1	CORN
72032803	117	552	555	1	1749	1751	1	CORN
72032803	9A	553	555	1	1431	1437	1	CORN
72032803	48	553	559	1	1527	1530	1	CORN
72032803	62	554	556	1	1613	1623	1	CORN
72032803	74A	555	557	1	1711	1717	1	CORN
72032803	11	558	561	1	1446	1457	1	CORN
72032803	74B	558	560	1	1712	1719	1	CORN
72032803	63A	559	561	1	1614	1617	1	CORN
72032803	63B	559	560	1	1618	1622	1	CORN
72032803	10	560	564	1	1435	1443	1	CORN
72032803	13A	567	572	1	1436	1446	1	CORN
72032803	120	567	569	1	1755	1763	1	CORN
72032803	13	573	577	1	1438	1448	1	CORN
72032803	44	573	577	1	1489	1494	1	CORN
72032803	121	575	576	1	1750	1753	1	CORN
72032803	79	577	580	1	1725	1731	1	CORN
72032803	13	578	581	1	1440	1450	1	CORN
72032803	17B	581	582	1	1461	1466	1	CORN
72032803	43B	581	584	1	1489	1495	1	CORN
72032803	114A	581	583	1	1742	1746	1	CORN
72032803	13	582	585	1	1441	1451	1	CORN
72032803	17A	582	583	1	1455	1460	1	CORN
72032803	114B	582	585	1	1736	1741	1	CORN
72032803	49B	583	586	1	1556	1560	1	CORN
72032803	21	584	589	1	1465	1468	1	CORN
72032803	49A	584	587	1	1551	1555	1	CORN
72032803	43A	585	590	1	1487	1497	1	CORN
72032803	18B	587	587	1	1446	1453	1	CORN
72032803	124	587	594	1	1763	1770	1	CORN
72032803	18A	588	588	1	1442	1448	1	CORN
72032803	82	590	594	1	1685	1695	1	CORN
72032803	83	592	599	1	1670	1675	1	CORN
72032803	22B	595	597	1	1451	1456	1	CORN
72032803	22A	596	598	1	1446	1450	1	CORN
72032803	125	599	602	1	1758	1762	1	CORN
72032803	24A	601	605	1	1456	1458	1	CORN
72032803	24B	602	606	1	1452	1455	1	CORN
72032803	24C	603	607	1	1449	1451	1	CORN
72032803	26	607	614	1	1453	1461	1	CORN
72032803	127	608	611	1	1769	1772	1	CORN

TABLE D4, CONTINUED

72032803	42	616	622	1	1478	1481	1	CORN
72032803	101A	616	618	1	1717	1720	1	CORN
72032803	100	618	625	1	1705	1712	1	CORN
72032803	128	618	619	1	1773	1777	1	CORN
72032803	101B	620	622	1	1718	1721	1	CORN
72032803	28	627	628	1	1471	1476	1	CORN
72032803	27	628	630	1	1458	1468	1	CORN
72032803	88	631	634	1	1655	1658	1	CORN
72032803	99A	631	632	1	1694	1696	1	CORN
72032803	99R	633	634	1	1695	1696	1	CORN
72032803	103A	633	634	1	1722	1723	1	CORN
72032803	93A	635	636	1	1670	1675	1	CORN
72032803	103B	635	636	1	1723	1724	1	CORN
72032803	92	636	640	1	1665	1666	1	CORN
72032803	93R	637	640	1	1671	1673	1	CORN
72032803	103C	637	638	1	1724	1725	1	CORN
72032803	89	639	643	1	1645	1648	1	CORN
72032803	60	641	642	1	1631	1636	1	CORN
72032803	32	645	650	1	1479	1482	1	CORN
72032803	58	645	648	1	1610	1612	1	CORN
72032803	31	648	661	1	1470	1475	1	CORN
72032803	97	649	651	1	1676	1680	1	CORN
72032803	53	650	656	1	1581	1586	1	CORN
72032803	52A	652	654	1	1570	1574	1	CORN
72032803	33	652	661	1	1483	1492	1	CORN
72032803	107	653	655	1	1714	1719	1	CORN
72032803	52R	655	657	1	1571	1575	1	CORN
72032803	112A	656	660	1	1732	1738	1	CORN
72032803	51	657	662	1	1544	1548	1	CORN
72032803	110B	659	660	1	1715	1717	1	CORN
72032803	110A	660	661	1	1711	1714	1	CORN
72032803	109A	661	662	1	1703	1708	1	CORN
72032803	112B	661	664	1	1734	1740	1	CORN
72032803	109B	662	663	1	1698	1702	1	CORN
72032803	34	665	670	1	1487	1498	1	CORN
72032803	L-86	684	686	1	1340	1342	1	CORN
72032803	L-82	686	689	1	1320	1323	1	CORN
72032803	L-81	687	690	1	1313	1317	1	CORN
72032803	L-119	688	691	1	1342	1347	1	CORN
72032803	L-80	688	690	1	1308	1310	1	CORN
72032803	L-77	688	689	1	1285	1293	1	CORN
72032803	L-79	690	692	1	1300	1305	1	CORN
72032803	L-118	691	693	1	1350	1353	1	CORN
72032803	L-122	693	698	1	1321	1326	1	CORN
72032803	L-123	695	699	1	1317	1319	1	CORN
72032803	L-91	696	700	1	1301	1306	1	CORN
72032803	L-74	696	698	1	1272	1273	1	CORN
72032803	L-90	697	698	1	1293	1298	1	CORN
72032803	L-112	698	700	1	1349	1356	1	CORN
72032803	L-87	700	701	1	1273	1276	1	CORN
72032803	L-92	701	704	1	1296	1299	1	CORN
72032803	L-111C	702	705	1	1352	1355	1	CORN
72032803	L-111R	703	705	1	1347	1350	1	CORN
72032803	L-68	703	704	1	1257	1262	1	CORN
72032803	L-129	705	707	1	1329	1330	1	CORN
72032803	L-111A	705	707	1	1342	1344	1	CORN
72032803	L-67 B	705	706	1	1252	1254	1	CORN
72032803	L-51	705	707	1	1201	1215	1	CORN
72032803	L-128	706	709	1	1319	1323	1	CORN
72032803	L-67 A	706	707	1	1247	1250	1	CORN
72032803	L-127	707	711	1	1312	1315	1	CORN
72032803	L-114	707	708	1	1365	1369	1	CORN
72032803	L-60	707	709	1	1231	1234	1	CORN
72032803	L-32	707	712	1	1184	1193	1	CORN
72032803	L-93	707	709	1	1296	1301	1	CORN
72032803	L-108C	708	711	1	1352	1355	1	CORN
72032803	L-94	708	710	1	1291	1294	1	CORN

TABLE D4, CONTINUED

72032803	L-108B	709	712	1	1345	1349	1	CORN
72032803	L-54	709	710	1	1200	1220	1	CORN
72032803	L-108A	711	713	1	1335	1347	1	CORN
72032803	L-52	711	713	1	1207	1217	1	CORN
72032803	L-109C	712	715	1	1352	1355	1	CORN
72032803	L-49	712	714	1	1200	1204	1	CORN
72032803	L-109B	713	717	1	1347	1349	1	CORN
72032803	L-5	713	718	1	1102	1110	1	CORN
72032803	L-96 A	714	716	1	1285	1289	1	CORN
72032803	L-22	714	716	1	1158	1165	1	CORN
72032803	L-96 B	715	717	1	1281	1283	1	CORN
72032803	L-109A	715	717	1	1342	1346	1	CORN
72032803	L-73	716	717	1	1259	1273	1	CORN
72032803	L-33	717	718	1	1174	1182	1	CORN
72032803	L-107	718	721	1	1330	1334	1	CORN
72032803	L-103A	718	720	1	1292	1293	1	CORN
72032803	L-37	719	720	1	1187	1199	1	CORN
72032803	L-2	719	721	1	1091	1097	1	CORN
72032803	L-100	720	723	1	1283	1285	1	CORN
72032803	L-103B	721	722	1	1294	1295	1	CORN
72032803	L-102	722	725	1	1288	1293	1	CORN
72032803	L-1	722	728	1	1078	1086	1	CORN
72032803	L-103C	723	724	1	1296	1297	1	CORN
72032803	L-19	723	726	1	1134	1143	1	CORN
72032803	L-38	724	728	1	1183	1186	1	CORN
72032803	L-26	724	728	1	1162	1172	1	CORN
72032803	L-40	725	727	1	1190	1192	1	CORN
72032803	L-17	725	728	1	1120	1129	1	CORN
72032803	L-41 B	726	728	1	1199	1208	1	CORN
72032803	L-41 A	727	729	1	1195	1198	1	CORN
72032803	L-7	730	743	1	1103	1106	1	CORN
72032803	L-44	732	741	1	1195	1198	1	CORN
72032803	L-4	732	734	1	1078	1083	1	CORN
72032803	L-29	736	738	1	1166	1173	1	CORN
72032803	L-25	740	751	1	1144	1162	1	CORN
72032803	L-30	741	744	1	1169	1178	1	CORN
72032803	L-46	743	746	1	1187	1193	1	CORN
72032803	L-13 C	744	744	1	1133	1135	1	CORN
72032803	L-13 A	745	745	1	1126	1128	1	CORN
72032803	L-13 B	745	745	1	1130	1131	1	CORN
72032803	L-48	748	750	1	1199	1211	1	CORN
72032803	L-8	751	757	1	1094	1104	1	CORN
72032803	L-11	751	761	1	1131	1138	1	CORN
72032803	L-31	752	757	1	1153	1166	1	CORN
72032803	L-10	755	764	1	1117	1124	1	CORN
72032803	136	262	266	1	1453	1463	1	SOY
72032803	137A	269	272	1	1462	1466	1	SOY
72032803	137B	273	276	1	1463	1468	1	SOY
72032803	142	307	313	1	1450	1457	1	SOY
72032803	150	332	335	1	1487	1490	1	SOY
72032803	147	338	341	1	1489	1491	1	SOY
72032803	157	346	351	1	1383	1392	1	SOY
72032803	161	347	352	1	1437	1446	1	SOY
72032803	O-156	361	366	1	1243	1246	1	SOY
72032803	O-152	362	368	1	1221	1225	1	SOY
72032803	163	363	368	1	1462	1467	1	SOY
72032803	O-159	366	368	1	1254	1256	1	SOY
72032803	173	372	376	1	1407	1411	1	SOY
72032803	O-165	373	375	1	1270	1272	1	SOY
72032803	O-139	374	376	1	1184	1188	1	SOY
72032803	O-141	377	381	1	1194	1196	1	SOY
72032803	O-142	381	383	1	1188	1193	1	SOY
72032803	O-75	381	385	1	1122	1124	1	SOY
72032803	O-143	382	383	1	1185	1186	1	SOY
72032803	O-27	386	388	1	1108	1111	1	SOY
72032803	O-81 A	387	390	1	1144	1145	1	SOY

TABLE D4, CONTINUED

72032803	0-24	388	389	1	1098	1098	1	SOY
72032803	0-172B	389	391	1	1266	1268	1	SOY
72032803	0-172A	389	392	1	1269	1272	1	SOY
72032803	0-172C	390	393	1	1262	1265	1	SOY
72032803	0-81 B	391	396	1	1146	1147	1	SOY
72032803	0-146	391	393	1	1201	1202	1	SOY
72032803	0-175	392	394	1	1254	1256	1	SOY
72032803	0-12 A	392	393	1	1054	1055	1	SOY
72032803	0-8	394	401	1	1030	1034	1	SOY
72032803	0-12 B	394	395	1	1055	1056	1	SOY
72032803	0-31	407	409	1	1110	1114	1	SOY
72032803	0-182	413	414	1	1222	1223	1	SOY
72032803	0-19	415	417	1	1058	1062	1	SOY
72032803	0-183	417	419	1	1222	1224	1	SOY
72032803	0-187	428	432	1	1204	1206	1	SOY
72032803	0-203A	435	439	1	1284	1285	1	SOY
72032803	0-190	436	436	1	1212	1216	1	SOY
72032803	0-90	437	440	1	1146	1148	1	SOY
72032803	0-203B	439	442	1	1286	1287	1	SOY
72032803	0-68	443	444	1	1115	1116	1	SOY
72032803	0-95	448	450	1	1139	1142	1	SOY
72032803	0-35	450	452	1	1034	1040	1	SOY
72032803	0-197	450	452	1	1261	1265	1	SOY
72032803	0-102A	453	457	1	1152	1154	1	SOY
72032803	0-62	458	464	1	1122	1125	1	SOY
72032803	0-102B	458	460	1	1150	1155	1	SOY
72032803	0-60 B	461	461	1	1110	1112	1	SOY
72032803	0-40	461	467	1	1042	1045	1	SOY
72032803	0-60 A	462	462	1	1106	1108	1	SOY
72032803	0-115	463	470	1	1181	1184	1	SOY
72032803	0-43 B	467	468	1	1060	1065	1	SOY
72032803	0-43 A	468	470	1	1055	1059	1	SOY
72032803	0-136	470	472	1	1190	1195	1	SOY
72032803	0-50	476	478	1	1078	1082	1	SOY
72032803	0-53	480	483	1	1084	1089	1	SOY
72032803	0-46	481	482	1	1057	1057	1	SOY
72032803	0-110B	488	488	1	1137	1141	1	SOY
72032803	0-110A	489	489	1	1132	1136	1	SOY
72032803	0-56	491	492	1	1110	1114	1	SOY
72032803	0-57	496	498	1	1113	1116	1	SOY
72032803	76A	540	542	1	1729	1733	1	SOY
72032803	67R	542	547	1	1665	1671	1	SOY
72032803	116A	542	545	1	1746	1748	1	SOY
72032803	67A	543	548	1	1660	1664	1	SOY
72032803	76B	543	548	1	1731	1733	1	SOY
72032803	73A	546	548	1	1697	1699	1	SOY
72032803	116B	546	549	1	1747	1749	1	SOY
72032803	73B	549	551	1	1699	1701	1	SOY
72032803	73C	552	554	1	1701	1703	1	SOY
72032803	12A	554	554	1	1450	1454	1	SOY
72032803	73D	555	557	1	1703	1704	1	SOY
72032803	12B	555	555	1	1446	1448	1	SOY
72032803	68	555	561	1	1645	1652	1	SOY
72032803	118	558	561	1	1751	1754	1	SOY
72032803	14A	565	567	1	1448	1453	1	SOY
72032803	14B	568	569	1	1449	1453	1	SOY
72032803	122A	578	581	1	1758	1761	1	SOY
72032803	80	579	581	1	1715	1720	1	SOY
72032803	122B	582	585	1	1760	1762	1	SOY
72032803	20A	585	590	1	1455	1459	1	SOY
72032803	131	585	586	1	1753	1754	1	SOY
72032803	20B	588	590	1	1456	1460	1	SOY
72032803	19B	589	593	1	1451	1454	1	SOY
72032803	123A	589	590	1	1755	1758	1	SOY
72032803	19A	590	594	1	1444	1450	1	SOY
72032803	123B	590	591	1	1752	1754	1	SOY
72032803	25A	592	594	1	1458	1465	1	SOY

TABLE D4, CONTINUED

72032803	25B	595	598	1	1459	1466	1	SOY
72032803	126A	598	601	1	1766	1769	1	SOY
72032803	25C	599	603	1	1461	1467	1	SOY
72032803	84	602	609	1	1669	1671	1	SOY
72032803	126B	602	605	1	1767	1769	1	SOY
72032803	85	611	612	1	1661	1667	1	SOY
72032803	41	619	621	1	1486	1495	1	SOY
72032803	40	622	629	1	1493	1497	1	SOY
72032803	29	625	629	1	1479	1483	1	SOY
72032803	38	625	633	1	1503	1507	1	SOY
72032803	102	628	629	1	1720	1723	1	SOY
72032803	30	631	637	1	1474	1485	1	SOY
72032803	39	633	637	1	1513	1516	1	SOY
72032803	91	637	639	1	1657	1661	1	SOY
72032803	104	638	641	1	1719	1720	1	SOY
72032803	54	640	643	1	1585	1587	1	SOY
72032803	96	645	647	1	1674	1678	1	SOY
72032803	95C	648	650	1	1668	1670	1	SOY
72032803	106	648	650	1	1720	1723	1	SOY
72032803	95D	651	652	1	1669	1671	1	SOY
72032803	108	653	654	1	1723	1725	1	SOY
72032803	35	655	657	1	1496	1497	1	SOY
72032803	50	658	663	1	1530	1533	1	SOY
72032803	111A	658	659	1	1724	1726	1	SOY
72032803	36	659	665	1	1520	1524	1	SOY
72032803	111B	660	661	1	1725	1727	1	SOY
72032803	111C	662	663	1	1726	1727	1	SOY
72032803	113	670	673	1	1729	1733	1	SOY
72032803	L-85	684	687	1	1333	1337	1	SOY
72032803	L-84	685	687	1	1329	1331	1	SOY
72032803	L-120	690	692	1	1335	1339	1	SOY
72032803	L-115	692	695	1	1358	1361	1	SOY
72032803	L-75	695	696	1	1276	1277	1	SOY
72032803	L-124	697	699	1	1313	1314	1	SOY
72032803	L-89 B	697	700	1	1286	1290	1	SOY
72032803	L-89 A	698	700	1	1280	1285	1	SOY
72032803	L-65	698	700	1	1250	1251	1	SOY
72032803	L-113	699	702	1	1361	1364	1	SOY
72032803	L-69	699	699	1	1264	1268	1	SOY
72032803	L-63	699	702	1	1242	1244	1	SOY
72032803	L-59	701	703	1	1231	1240	1	SOY
72032803	L-53	703	705	1	1218	1221	1	SOY
72032803	L-57	703	706	1	1225	1225	1	SOY
72032803	L-50	704	708	1	1197	1197	1	SOY
72032803	L-71	706	707	1	1259	1269	1	SOY
72032803	L-61	706	708	1	1237	1238	1	SOY
72032803	L-126	708	712	1	1305	1308	1	SOY
72032803	L-106B	712	717	1	1325	1328	1	SOY
72032803	L-106C	712	716	1	1330	1333	1	SOY
72032803	L-95 B	712	713	1	1299	1302	1	SOY
72032803	L-97	713	714	1	1277	1279	1	SOY
72032803	L-95 A	713	715	1	1292	1298	1	SOY
72032803	L-36	714	716	1	1191	1197	1	SOY
72032803	L-106A	714	719	1	1316	1322	1	SOY
72032803	L-105	715	721	1	1308	1312	1	SOY
72032803	L-104A	717	719	1	1303	1304	1	SOY
72032803	L-101	719	720	1	1287	1291	1	SOY
72032803	L-104B	720	722	1	1304	1306	1	SOY
72032803	L-99 A	721	723	1	1280	1281	1	SOY
72032803	L-99 B	724	725	1	1281	1282	1	SOY
72032803	L-3	724	726	1	1092	1102	1	SOY
72032803	L-42	730	730	1	1184	1188	1	SOY
72032803	L-27 B	730	730	1	1171	1175	1	SOY
72032803	L-27 A	731	731	1	1165	1170	1	SOY
72032803	L-16	731	732	1	1121	1130	1	SOY
72032803	L-6	731	740	1	1092	1097	1	SOY
72032803	L-43	733	739	1	1186	1189	1	SOY

TABLE D4, CONTINUED

72032803	L-28	733	733	1	1165	1172	1	SOY
72032803	L-14	738	742	1	1126	1133	1	SOY
72032803	L-47	743	745	1	1197	1206	1	SOY
72032803	L-12	747	748	1	1127	1135	1	SOY
72032803	L-9	748	752	1	1109	1121	1	SOY
72032803	0-164	374	376	1	1258	1259	1	OTHER,ALFAL
72032803	0-120	376	376	1	1178	1180	1	OTHER,ALFAL
72032803	0-78	388	390	1	1126	1130	1	OTHER,ALFAL
72032803	0-30	390	390	1	1116	1117	1	OTHER,ALFAL
72032803	0-10 A	391	395	1	1037	1046	1	OTHER,ALFAL
72032803	0-14	396	396	1	1062	1062	1	OTHER,ALFAL
72032803	0-6 A	397	398	1	1022	1022	1	OTHER,ALFAL
72032803	0-10 B	398	400	1	1039	1048	1	OTHER,ALFAL
72032803	0-6 R	399	400	1	1023	1023	1	OTHER,ALFAL
72032803	0-214	404	405	1	1279	1284	1	OTHER,ALFAL
72032803	0-1	406	406	1	1001	1002	1	OTHER,ALFAL
72032803	0-213	407	408	1	1275	1285	1	OTHER,ALFAL
72032803	0-212	410	412	1	1278	1286	1	OTHER,ALFAL
72032803	0-185	423	426	1	1224	1226	1	OTHER,ALFAL
72032803	0-201B	445	448	1	1282	1286	1	OTHER,ALFAL
72032803	0-201A	446	449	1	1278	1281	1	OTHER,ALFAL
72032803	0-55	467	468	1	1096	1102	1	OTHER,ALFAL
72032803	71A	541	543	1	1676	1680	1	OTHER,ALFAL
72032803	71B	544	546	1	1677	1681	1	OTHER,ALFAL
72032803	81A	584	585	1	1716	1720	1	OTHER,ALFAL
72032803	81B	586	587	1	1717	1721	1	OTHER,ALFAL
72032803	90A	638	639	1	1651	1654	1	OTHER,ALFAL
72032803	90B	640	641	1	1652	1655	1	OTHER,ALFAL
72032803	59B	642	643	1	1623	1627	1	OTHER,ALFAL
72032803	59A	643	644	1	1619	1622	1	OTHER,ALFAL
72032803	105	644	644	1	1719	1722	1	OTHER,ALFAL
72032803	98	660	662	1	1674	1674	1	OTHER,ALFAL
72032803	64A	546	549	1	1628	1630	1	OTHER,GRASS
72032803	64B	550	554	1	1630	1631	1	OTHER,GRASS
72032803	119	563	566	1	1753	1753	1	OTHER,GRASS
72032803	15A	571	572	1	1450	1451	1	OTHER,GRASS
72032803	15B	573	574	1	1451	1452	1	OTHER,GRASS
72032803	16A	576	578	1	1452	1453	1	OTHER,GRASS
72032803	16B	579	580	1	1453	1454	1	OTHER,GRASS
72032803	55B	642	644	1	1594	1597	1	OTHER,GRASS
72032803	55A	643	645	1	1590	1593	1	OTHER,GRASS
72032803	56B	647	649	1	1594	1599	1	OTHER,GRASS
72032803	56A	648	651	1	1590	1593	1	OTHER,GRASS
72032803	57	650	654	1	1604	1608	1	OTHER,GRASS
72032803	0-161	364	366	1	1267	1269	1	OTHER,AGRIC
72032803	0-162	369	372	1	1267	1270	1	OTHER,AGRIC
72032803	0-77	381	381	1	1140	1144	1	OTHER,AGRIC
72032803	0-121	382	385	1	1158	1162	1	OTHER,AGRIC
72032803	0-144	388	393	1	1193	1197	1	OTHER,AGRIC
72032803	0-177	395	397	1	1246	1250	1	OTHER,AGRIC
72032803	0-2	401	405	1	1004	1007	1	OTHER,AGRIC
72032803	0-84 B	402	406	1	1127	1131	1	OTHER,AGRIC
72032803	0-84 A	403	407	1	1123	1126	1	OTHER,AGRIC
72032803	0-179A	408	410	1	1254	1256	1	OTHER,AGRIC
72032803	0-179B	411	412	1	1255	1257	1	OTHER,AGRIC
72032803	0-128	427	433	1	1191	1199	1	OTHER,AGRIC
72032803	0-88 B	428	430	1	1169	1171	1	OTHER,AGRIC
72032803	0-88 A	429	431	1	1166	1168	1	OTHER,AGRIC
72032803	0-73	433	435	1	1122	1128	1	OTHER,AGRIC
72032803	0-129	435	437	1	1191	1193	1	OTHER,AGRIC
72032803	0-63	444	444	1	1087	1089	1	OTHER,AGRIC
72032803	0-133R	453	456	1	1183	1184	1	OTHER,AGRIC
72032803	0-61	459	460	1	1114	1118	1	OTHER,AGRIC
72032803	0-116	473	478	1	1185	1187	1	OTHER,AGRIC
72032803	0-52	482	484	1	1072	1074	1	OTHER,AGRIC
72032803	0-47	485	487	1	1053	1055	1	OTHER,AGRIC

TABLE D4, CONTINUED

72032803	46	534	536	1	1517	1522	1	OTHER, AGRIC
72032803	78R	552	554	1	1744	1745	1	OTHER, AGRIC
72032803	78A	553	555	1	1740	1743	1	OTHER, AGRIC
72032803	23R	599	600	1	1453	1457	1	OTHER, AGRIC
72032803	23A	600	601	1	1446	1452	1	OTHER, AGRIC
72032803	129A	620	621	1	1768	1769	1	OTHER, AGRIC
72032803	129B	621	621	1	1765	1767	1	OTHER, AGRIC
72032803	130	632	634	1	1778	1781	1	OTHER, AGRIC
72032803	L-116	686	689	1	1356	1359	1	OTHER, AGRIC
72032803	L-83	686	688	1	1326	1326	1	OTHER, AGRIC
72032803	L-117	689	690	1	1350	1352	1	OTHER, AGRIC
72032803	L-78	692	694	1	1289	1294	1	OTHER, AGRIC
72032803	L-76	693	695	1	1280	1284	1	OTHER, AGRIC
72032803	L-125	697	699	1	1310	1310	1	OTHER, AGRIC
72032803	L-88	699	701	1	1278	1278	1	OTHER, AGRIC
72032803	L-64	699	701	1	1246	1248	1	OTHER, AGRIC
72032803	L-70	701	703	1	1265	1269	1	OTHER, AGRIC
72032803	L-58	702	704	1	1228	1228	1	OTHER, AGRIC
72032803	L-66 B	703	703	1	1253	1255	1	OTHER, AGRIC
72032803	L-66 A	704	704	1	1247	1251	1	OTHER, AGRIC
72032803	L-62	705	707	1	1242	1242	1	OTHER, AGRIC
72032803	L-20	706	717	1	1143	1150	1	OTHER, AGRIC
72032803	L-56	707	709	1	1226	1226	1	OTHER, AGRIC
72032803	L-110	707	708	1	1334	1339	1	OTHER, AGRIC
72032803	L-55	708	710	1	1223	1223	1	OTHER, AGRIC
72032803	L-98	708	711	1	1276	1278	1	OTHER, AGRIC
72032803	L-72	709	711	1	1271	1273	1	OTHER, AGRIC
72032803	L-35	715	717	1	1185	1188	1	OTHER, AGRIC
72032803	L-23	719	721	1	1160	1169	1	OTHER, AGRIC
72032803	L-18	719	719	1	1138	1141	1	OTHER, AGRIC
72032803	L-21	721	724	1	1148	1155	1	OTHER, AGRIC
72032803	L-34	721	721	1	1176	1180	1	OTHER, AGRIC
72032803	L-39	722	723	1	1189	1200	1	OTHER, AGRIC
72032803	L-24	726	729	1	1154	1158	1	OTHER, AGRIC
72032803	L-45	741	741	1	1188	1191	1	OTHER, AGRIC
72032803	O-82 A	383	385	1	1150	1154	1	OTHER, PAST
72032803	O-82 B	386	388	1	1151	1155	1	OTHER, PAST
72032803	O-82 C	389	389	1	1152	1156	1	OTHER, PAST
72032803	O-196	452	455	1	1245	1251	1	OTHER, PAST
72032803	69A	551	554	1	1663	1665	1	OTHER, PAST
72032803	69R	555	557	1	1664	1666	1	OTHER, PAST
72032803	69C	558	560	1	1666	1668	1	OTHER, PAST
72032803	86A	611	613	1	1677	1681	1	OTHER, PAST
72032803	140	295	297	1	1462	1468	1	OTHER, WHEAT
72032803	155	319	325	1	1387	1393	1	OTHER, WHEAT
72032803	149	327	330	1	1485	1489	1	OTHER, WHEAT
72032803	171	386	389	1	1398	1406	1	OTHER, WHEAT
72032803	O-200	448	451	1	1267	1273	1	OTHER, WHEAT
72032803	87A	629	630	1	1662	1665	1	OTHER, WHEAT
72032803	87B	631	633	1	1663	1665	1	OTHER, WHEAT
72032803	95A	642	644	1	1666	1668	1	OTHER, WHEAT
72032803	95B	645	647	1	1667	1669	1	OTHER, WHEAT
72032803	OGLF-WTR	338	344	1	1105	1109	1	WOODS
72032803	OGLF-WTR	343	348	1	1121	1130	1	WOODS
72032803	OGLF-WTR	352	358	1	1131	1135	1	WOODS
72032803	SYCAMORE	351	361	1	1595	1605	1	TOWN
72032803	DEKALB	445	455	1	1545	1555	1	TOWN
72032803	DEKALB-M	455	465	1	1525	1535	1	TOWN
72032803	HINCKLEY	643	650	1	1786	1795	1	TOWN

