NASA CR- 140200
ERIM 190100-26-T

SIGNATURE EXTENSION STUDIES
Technical Report

by

R. K. Vincent, G. S. Thomas, and R. F. Nalepka
Infrared and Optics Division

E N V I R O N M E N T A L
RESEARCH INSTITUTE
OF MICHIGAN

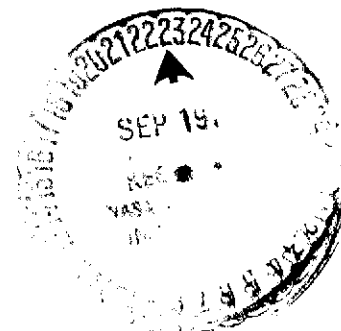FORMERLY WILLOW RUN LABORATORIES.
THE UNIVERSITY OF MICHIGAN

July 1974

prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Lyndon B. Johnson Space Center, Houston, TX 77058
Contract NAS9-9784, Task X

NOTICES

1. Report No. NASA CR-ERIM 190100-26-T	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle SIGNATURE EXTENSION STUDIES		5. Report Date
		6. Performing Organization Code
7. Author(s) R. K. Vincent, G. S. Thomas, and R. F. Nalepka, Infrared and Optics Division		8. Performing Organization Report No. 190100-26-T
9. Performing Organization Name and Address Environmental Research Institute of Michigan P. O. Box 618, Ann Arbor 48107		10. Work Unit No.
		11. Contract or Grant No. NAS 9-9784
		13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Lyndon B. Johnson Space Center Earth Observations Division Houston, TX 77058		Technical Report, 1 February 1973 through 31 January 1974
		14. Sponsoring Agency Code

15. Supplementary Notes

Dr. Andrew Potter is Technical Monitor for NASA/JSC.

16. Abstract

The importance of specific spectral regions to signature extension is explored. In the recent past, the signature extension task was focused on the development of new techniques. Tested techniques are now used to investigate this spectral aspect of the large area survey. Sets of channels were sought which, for a given technique, were the least affected by several sources of variation over four data sets and yet provided good object class separation on each individual data set.

Using sets of channels determined as part of this study, signature extension was accomplished between data sets collected over a six-day period and over a range of about 400 kilometers.

17. Key Words Remote sensing Multispectral scanning Multispectral data collection Large-area surveys Target signature analysis Signature extension	18. Distribution Statement Initial distribution is indicated at the end of this document.		
19. Security Classif. (of this report) UNCLASSIFIED	20. Security Classif. (of this page) UNCLASSIFIED	21. No. of Pages 49	22. Price

PREFACE

This report describes part of a comprehensive and continuing program of research in multispectral remote sensing of the environment from aircraft and satellites and the supporting effort of ground-based researchers in recording, coordinating, and analyzing the data gathered by these means. The basic objective of this program is to improve the utility of remote sensing as a tool for providing decision makers with timely and economical information from large geographical areas.

The feasibility of using remote sensing techniques to detect and discriminate between objects or conditions at or near the surface of the earth has been demonstrated. Applications in agriculture, urban planning, water quality control, forest management, and other areas have been developed. The thrust of this program is directed toward the development and improvement of advanced remote sensing systems and includes assisting in data collection, processing and analysis, and ground truth verification.

The research covered in this report was performed under NASA Contract NAS 9-9784. The program was directed by R. R. Legault, Director of ERIM's Infrared and Optics Division and an Institute Vice-President, and J. D. Erickson, Principal Investigator. The Institute number for this report is 190100-26-T.

The authors wish to acknowledge the administrative direction provided by Mr. R. R. Legault and Dr. Jon D. Erickson and the technical assistance given by Mr. J. P. Morgenstern and Mr. W. W. Pillars. Ms. Darlene Dickerson is thanked for her secretarial assistance.

CONTENTS

PRECEDING PAGE BLANK NOT FILMED

FIGURES

TABLES

ERIM

SIGNATURE EXTENSION STUDIES

1

INTRODUCTION AND SUMMARY

Recent nationwide shortages in resources ranging from grain to petroleum have focused attention on resource management and its importance to the nation's economy. Timely and accurate resource surveys are a prerequisite for effective resource management. Remote sensing by multispectral scanners holds the promise of providing resource survey information on an accurate and timely basis, once the techniques for making such surveys become operational. As has already been demonstrated by scores of investigations, multispectral signatures extracted from in-scene references (training sets) of various targets can be used to make automatic recognition maps of those targets for a segment of data from a given aircraft or satellite flight. However, for large-area surveys, the requirement that training sets for all targets of interest be known in each flight segment is undesirable. If the target signatures extracted from training sets of a single flight segment could be used for other flight segments separated in space and time, a great deal of time and expense could be trimmed from large-area surveys in which data are gathered by multispectral remote sensing. This procedure, called signature extension, is a highly desirable and possibly necessary prerequisite for surveying regional resources by remote sensing methods on an operational basis.

Research efforts at ERIM over the past several years have been directed toward the development and study of signature extension techniques which will bring us closer to an operational system for providing the large-area survey capability required to inventory resources. Although many of the techniques being developed are applicable to a number of disciplines, this study, because of the data sets examined, is primarily concerned with agricultural surveys. The goal of this study is and has been to develop the methods by which a small amount of ground truth, translated into target signatures, can be extended via computer processing over areas of regional size. The development of signature extension techniques is attempting to suppress four types of variations in multispectral scanner data:

(1) Scan-angle-dependent variations of received radiance within a given flight segment or satellite frame caused primarily by variations in atmospheric path radiance and atmospheric transmittance with changing optical path lengths. (These variations are much more important for aircraft data than in present-day satellite data which is collected by scanners having limited angular fields of view.)

(2) Atmospheric variations within flight segments or satellite frames caused by changes in atmospheric state. (These variations are equally important for aircraft and satellite data.)

(3) Atmospheric and solar illumination variations between flight segments and satellite frames caused by different atmospheric conditions and sun elevations at different times and places. (These variations are equally important for aircraft and satellite data.)

(4) Electronic gain changes between flight segments, possibly caused by manual gain adjustments in the aircraft system, instabilities in the data collection system, and improper calibration of the data in the aircraft and satellite systems. These variations can be larger than the naturally induced changes cited in the first three items, above.

Besides the above-listed variations, changes on the ground caused by vegetative growth, planting customs, and crop harvesting are important sources of variation between flight segments. However, at this stage of the present study, on-the-ground changes are not considered. Indeed it could be argued that such changes should not be suppressed because they may be important components of the target signature.

Previous efforts at ERIM have been directed toward developing data transformations which would suppress the four sources of "noise" enumerated above such that a representative set of field signatures from one area could be used to identify fields of the same type in distant areas (this desideratum is known as signature extension). During this report period we examined signature extension from another viewpoint by considering previously tested techniques and looking for the best set of spectral channels to use with each of those techniques. By best set we mean that subset of the original number of data channels which can be shown to be the most extendable in space and time with the highest target classification accuracies on all the data considered. This investigation into the existence of a best set of channels for large-area resource surveys was conducted using data gathered during the Corn Blight Watch Experiment of 1971. Although this study was confined primarily to agricultural targets, the implications of this research impact the design of remote sensing systems for resource management in general.

In this study it was established that not all spectral wavebands are equally extendable, that different signature extension techniques require different best sets of channels, and that the use of additional data channels can at times be detrimental to classification accuracy.

A new aircraft data-set, with which the findings of this study could be further explored, was gathered during this report period. Multispectral scanner and photographic data were collected over Lenawee County, Michigan on 4 August 1973. Ground-truth information was obtained for fields of various crops scattered throughout the county.

10

In the remainder of this report, we discuss in detail the many aspects of this study. Section 2 contains a description of the sources of data variation leading to the need for signature extension, while Section 3 provides a review of data transformations (signature extension techniques) utilized in this study. In Section 4 we describe the channel selection study —the major effort during this report period. Section 5 describes the new scanner data gathered this year, and Section 6 presents the conclusions and recommendations resulting from this study.

2

DATA VARIATION

Large-area operational surveys are going to encounter data which contains large variations. Ideally there exists a datum plane to which all data (scanner data in this case) may be reduced, and within which each signature is invariant. Signature extension methodology attempts to reduce variation within and between flight segments of data, to the point where the data seem to have been collected simultaneously.

Sources of variations in scanner signals are manifold [1]. Briefly, they may be divided into three categories: instrument sources, environmental sources, and scene-related (on-the-ground) sources of variation. Instrument-related variation is introduced both purposefully and inadvertently. In the latter case are factors such as drift and other electronic instabilities and optical-mechanical view-angle effects. In the former case, variation is induced by the recording technician who causes the scanner input signals to be within the dynamic range of the recording equipment.

Because environmental variations are dependent on the position of the sun in the sky and the state of the atmosphere, they can be large. The scene-related sources of variation can also be large, but they are not the subject of this investigation. Certainly, however, to perfect large-area operational survey techniques, scene-related variations induced by plant maturity, planting geometry, soil type, moisture condition, etc., must eventually be understood.

The basic expression

$$L_o = \rho ET + L_p$$

for the radiation being sensed in each channel by the multispectral scanner illustrates how some types of variation enter the data. The reflectance, ρ, exhibited by a target for the scanner-target-sun geometry is primarily a scene-related type of variation. E, the irradiance upon the target, is an environmental source of variation as is T, which is a measure of the atmosphere's ability to transmit radiation. The product of these factors, when added to another environmental factor, path radiance (L_p), represents the radiation observed at the scanner in each spectral band, L_o. Any of the above quantities may vary during the collection of MSS data from a given area. Consequently, even more variation is to be expected when data from different areas, collected at different times, and under different environmental conditions are to be compared.

---

1. Nalepka, R. F.,and J. P. Morgenstern, Signature Extension: An Approach to Operational Multispectral Surveys, Report No. 31650-152-T, Environmental Research Institute of Michigan, Ann Arbor, March 1973.

To correct for features in the data which are not scene-related (that is, features which overshadow and confuse the structure of the data associated with the objects of interest), a variety of methods may be used. Three methods employed in this study include the Average Signal Versus Angle (ASVA) correction, the Ratio of Adjacent Channel (RAC) transformation, and normalization. Normalization was used in conjunction with the ASVA correction but only when data from different areas were to be compared. The RAC transformation was used alone. These methods are more fully explained in Section 3, Data Transformations.

3

DATA TRANSFORMATIONS

Two signature-extension data transformations tested last year were utilized during this study. They were the Average Signal Versus Angle (ASVA correction) and the Ratio of Adjacent Channel transform. The following is a review of these signature extension techniques.

In utilizing the ASVA correction, the scene is sampled on a channel-by-channel basis to determine what angular effects are present. For a given channel an average signal is determined at each predetermined incremental angle across the scan line. This is done by calculating the average signal for each incremental scan angle over the length of the scene. At this point, the signal averages form a stepped function across the scan line in each channel. The correction function is obtained from these signal averages and better approximates a continuous one as the angle increment decreases. Final smoothing is accomplished by fitting a curve (of second order, in this case) to the corrections. The smoothed correction functions, one per data channel, may be used in one of two ways depending on what one assumes is the major source of angular variation in the scene.

Recalling the equation for the radiation being sensed by the multispectral scanner in each spectral band, $L_o = \rho E \tau + L_p$, two possible sources of major angular variation in a scene are atmospheric transmission, τ, and path radiance, L_p. The ASVA correction will provide either a multiplicative or an additive correction to account for τ or L_p, respectively. Either correction is achieved in a two-step process. In the first step the multiplicative function is made exactly 1.0 or the additive function exactly 0.0 at the reference angle. For the multiplicative correction this is accomplished by normalizing the correction function to the signal level at the reference angle. For the additive correction the data signal level at the reference angle is subtracted from the smoothed correction function. Thus no scan angle correction is applied at the reference angle in either case. The second step in achieving a scan angle correction uses the smoothed and adjusted correction function.

To reduce multiplicative variation, the data at each angle is divided by the adjusted correction function at corresponding angles. Additive effects are reduced by subtracting the adjusted correction function from the data on an angle-by-angle basis.

The ASVA method of correcting for angular variation in the scene is subject to certain constraints. The scene is assumed to have a quasi-random distribution of all object classes. This must be so if the correction is to be data-independent and truly representative of angular variation. The object classes are assumed to have similar bidirectional reflectance effects — and again this relates to the data independence of the correction. Another constraint is that this method will not correct for environmental changes that occur as the data is being collected. Any changes of this nature will be buried in the averaging process, and the resultant

14

correction function would probably not represent the angular variation existing in the scene. Time and cost are other constraints on this method since two passes through the data are required. Also, when this method is used to extend classification to other areas, additional time must be spent normalizing the data between areas. This normalization is a procedure by which the signal levels from one area are scaled to the signal levels of another area; it must be done to account for environmental differences between areas.

The other data transform we employed in this study was the Ratio of Adjacent Channels. The basis for this transform is the high correlation between undesirable multiplicative effects in adjacent spectral regions. If additive effects can be ignored, the ratio of two adjacent channels closely approximates the ratio of target reflectances in those spectral channels. Also, the transformed data are not scan-angle-dependent. For this transformation, discrimination between object classes depends upon relative spectral characteristics between adjacent bands. If the relationships between adjacent bands change, object classes will become confused, i.e., separability of object classes will be lost.

Use of this transform is advantageous in that no a priori knowledge of the data is required; the targets may be distributed in any manner. Scan angle effects of the multiplicative variety are cancelled in the ratioing process. Ratio signatures from one area should be extendable in a straightforward fashion to other areas.

4
CHANNEL SELECTION STUDY

This section describes the data used in this project as well as the preparations and modifications performed upon it. Also described is the search — performed via computer and processing techniques utilizing the above data — for a set of spectral wavebands which would provide not only good discrimination between object classes from one area but also good discrimination between those same object classes at other locations dissimilar in many respects to the first.

4.1 DESCRIPTION OF THE DATA AND DATA PREPARATION

Most of the data employed for this investigation were also the subject of last year's investigation. Another segment was added, however, which allowed the maximum spatial range of the survey to be extended to nearly three times what it had been. Still in use from the previous year were the Corn Blight Watch Experiment intensive study segments 204, 203 and 212. These segments represented spatial and temporal separations of four days and about 128 km.

A total of three data sets were generated for each original segment of data. The first set consisted of data which had been calibrated to radiance (see below) but not modified by any so-called signature extension technique. This data was designated "Untransformed." The second set was similar to the first but besides being calibrated it was scan-angle-corrected by the ASVA method. The third set consisted of calibrated data as modified by the Ratio of Adjacent Channel transform.

All data considered in this study were calibrated to a reference lamp contained within the scanner. This attempt was suboptimal because accurate calibration figures for the three-element near infrared detector were not known and because the reference lamp itself had not been calibrated for some time prior to these flights. This calibration eliminated variations in signal levels between flights due to gain changes (both intentional and unintentional) in the scanner and its peripheral equipment. Any variations between flight segments still remaining in the signal levels after calibration should thus be caused by either the environment or the scene. Calibration of the data was critical to two of the data types considered in this study: the no-transformation (Untransformed) case and the Ratio of Adjacent Channel transformation case. For these two cases, any uncorrected gain change between flights in a channel used for classification could result in gross recognition inaccuracies. Calibrating the data to a reference lamp was the simplest and most all-inclusive means of eliminating the effect of known, and perhaps unknown, gain changes between flights. In the third case considered in this study, the ASVA-correction case, calibration data were utilized, as with the other two cases.

In addition, normalization was employed with the ASVA correction between flight segments. The signal level in each channel averaged over the whole flight segment length, for a given reference angle (near the center of the scan line),was computed for each flight segment. To obtain a normalization factor for a given data channel of segment 203, the averaged signal level at the reference angle for segment 203 (the reference flight segment) was divided by the average signal level at the reference angle for segment 204. Normalization factors were determined for all twelve data channels. Each channel in segment 203 was then scaled by its appropriate normalization factor. This procedure, from determination of normalization factors through scaling of data channels by those factors, was repeated for segments 212 and 227.

Models for eliminating environmental and scene-related sources of variation in scanner data were not utilized in this study. Nor was normalization applied to data in the Untransformed or the Ratio of Adjacent Channel transform cases. These remain as avenues for further study.

Signatures had previously been extracted for seven major object classes (corn, soy, pasture, cut hay, trees, sparse vegetation, and growing hay) for segment 204, and for three of these seven (viz., corn, soy and trees) for segments 203 and 212. These data were modified for use in this investigation in several ways.

The first modification was the incorporation of new scanner calibration data not available when segments 204, 203 and 212 were digitized. The new calibration values covered the seven-month period (June through December 1971) which included the four days over which all the data used in this report was collected. Calibration allowed the conversion of arbitrary voltages from the data tapes to absolute radiance. The data from the three segments were adjusted by appropriate scaling factors on a channel-by-channel basis to effect proper calibration. In this attempt to reduce the data to an absolute reference plane, certain pitfalls had to be acknowledged. In the first place, the error attached to calibration algorithms and methodology could be large. This was found as part of a study [2] performed this year on another task which examined the error factor in the calibration process. Among other things, this study found that some ERIM multispectral scanner channels were more susceptible to error than others. We also realized that the time period between updates of scanner calibration was too lengthy. It was thought that more frequent calibration of the scanner (perhaps monthly rather than semi-annually) would give more meaning to the scanner calibration values used to calibrate data collected between scanner calibrations.

2. Stewart, S., D. Christenson, and L. Larsen, Systematic Monitoring of M-7 Scanner Performance and Data Quality, Report No. 190100-23-T, Environmental Research Institute of Michigan, Ann Arbor (in press).

17

A second modification was the addition of another Corn Blight Watch Experiment intensive study segment (227) to the body of data considered in this study. Segment 227 was added to allow the testing of signature extension techniques over a wider variety of environmental and scene-related conditions than had been considered in the past. Its latitude was nearly 2 degrees farther south than the closest of the other three segments. This represented a distance of about 400 km. Table 1 compares some statistics for the four segments at the time of data collection. As may be seen, a variety of conditions are represented in the data. The range in solar zenith angle between the four segments was expected to provide significant differences in the amount of radiation reaching the ground and, hence, being sensed by the scanner from the individual segments. The multispectral scanner signals from Segment 227 were scaled to radiance using the same radiance calibration values as were used in updating the calibration of Segments 204, 203 and 212.

A third modification was the generation of signatures for the seven object classes for all four segments. Fields were selected for signature extraction (and technique testing) throughout each segment. An attempt was made to encompass a representative variety of each object class within each segment. For example, when considering, among all possible soybean fields, which ones to select for signature extraction, an attempt was made to choose an assortment of soybean fields displaying a range of tones and textures on false-color IR imagery of the segment under consideration.

TABLE 1. COMPARISON OF DATA COLLECTION CONDITIONS
FOR FOUR SEGMENTS

	203	204	212	227
Date	8/13/71	8/13/71	8/17/71	8/12/71
Time	10:53 EST	10:25 EST	11:20 EST	12:00 EST
Latitude	41°36'	41°12'	40°03'	38°04'
Longitude	86°32'	87°28'	86°49'	87°01'
Solar Azimuth	127°	118°	137°	147°
Solar Zenith	53°	48°	57°	66°
Haze	Moderate	Moderate	Moderate	Light

Signatures generated in this manner were used for within-segment classification in testing ASVA-corrected, Ratio-of-Adjacent-Channel, and Untransformed data. An additional set of signatures was extracted from segment 204. This second set of signatures, extracted from a one-mile-square area centered on nadir near the beginning of segment 204, was used to classify object classes in segments 203, 212 and 227. In the case of actual large-area surveys, such extension techniques — requiring only a limited amount of ground truth — would be preferred from time and cost standpoints.

18

4.2 APPROACH

This section presents a summary of the work done this year in selecting a "best set" of channels for a given signature extension data transformation method. A "best set" of channels is that group of spectral bands necessary to produce the best classification results over the entire data set under consideration. Results were measured in terms of classification accuracy—i.e., the number of pixels within test fields correctly classified expressed as a percentage of the total number of pixels within the fields. The test fields were of an agricultural nature; hence, it should be kept in mind that the sets of channels discussed here and later are best from an agricultural viewpoint. It is probable that for signature extension over agricultural areas at other times of the year (that is, at other stages of the plant growth cycle), or, for signature extension over non-agricultural areas of interest, quite different "best sets" of channels may be required.

The concept of a best set of channels for large-area agricultural surveys has, until now, been largely unexplored. Since two methods already existed which had demonstrated capabilities in signature extension—namely the ASVA correction and Ratio of Adjacent Channel data transform—these methods were utilized to explore another facet of signature extension: Which of the spectral channels chosen on the basis of spectral contrasts between targets are most extendable in space and time?

Our investigation into best sets of channels was conducted in two phases. The first phase was designed to select sets of channels which, on the basis of spectral contrasts, were determined by a computer program to be the best for discriminating among the seven targets being considered. Sets of channels were chosen for individual Corn Blight Watch intensive study segments as well as combinations of segments. The second phase entailed testing the sets of channels chosen in the first phase. Testing methodology was the same as that used in the past to test signature extension techniques. Spectral signatures were extracted from training fields and utilized to classify the test fields.

4.2.1 PHASE ONE

Phase One involved an analysis of signatures extracted for seven object classes. The signature of any one object class is composed of the mean radiances of the scanner data training sets of that class as detected in each of twelve data channels (see Table 2) and a covariance matrix defining the radiance correlations between channels. The signatures of corn, soybeans, trees, pasture, cut hay, growing hay and bare soil were used as input to a computer program called STEPLIN. Program STEPLIN analyzed the relative separability of the seven target signatures and, on the basis of relative spectral contrasts, ranked the data channels with regard to their discriminatory ability. The first channel chosen was the one that generated the minimum average pairwise probability of misclassification (APPM). The second channel

FORMERLY WILLOW RUN LABORATORIES, THE UNIVERSITY OF MICHIGAN

TABLE 2. SCANNER CHANNEL WAVEBANDS

Multiplexer Channel	Multispectral Scanner Waveband[*] (in μm)
1	0.46-0.49
2	0.48-0.51
3	0.50-0.54
4	0.52-0.57
5	0.54-0.60
6	0.58-0.65
7	0.61-0.70
8	0.72-0.92
9	1.00-1.40
10	1.50-1.80
11	2.00-2.60
12	9.30-11.70

[*]These wavebands span the 50% points of the response function for their particular detector.

chosen was that one which, along with the first-chosen channel, provided the largest reduction in the APPM. The remainder were selected in similar fashion until all of the twelve available channels had been "prioritized." This was done for each of the four segments for each of three data forms: Untransformed, ASVA-corrected, and Ratio-of-Adjacent-Channel transformed data.

A further step taken in Phase One had important consequences. The signatures for each target were combined from various pairs of flight segments. Segment 204 had been utilized as the reference segment throughout this signature extension task. That is, the signatures from Segment 204 were extended to other segments. This tradition was maintained as signatures from Segment 204 were combined with signatures from the other segments. An example may serve to illustrate this: the corn signature from Segment 204 was combined with the corn signature from Segment 203, the soybean signature from 204 with the soybean signature from 203, and so on for the other targets. The result was a new signature (a new set of means and a new covariance matrix) for each of the seven target types which represented the mean signal levels and the total variance of both signatures. Since the input signatures were representative of the variation of a target type within an entire segment, it was thought that the combination of the signatures would be representative of the variation of a target type within two segments. A final step was the combination of signatures from all four segments into one four-segment signature containing the variation present in the seven target types for all four segments. The combined signatures served as another input for STEPLIN.

The first six channels as ranked by STEPLIN have typically been used for classification purposes. Past experience, as well as present STEPLIN results, have shown that there is little benefit to using more than six. Plots of APPM versus number of channels (see Figures 1 through 8, for example) typically level off at about six channels, meaning that after more than about six channels, there is no appreciable decrease in the average pairwise probability of misclassification. Because of the variability, visible in the plots, as to exactly when the curves level off or arrive at some predetermined minimum, and because of the fact that STEPLIN was working with signatures that were averages based upon some sample of the data, a criterion was established to allow the comparison of data sets, of data types, and data channels: this criterion is the pairwise probability of misclassification as generated by STEPLIN. The channels as ranked by STEPLIN were divided into two groups: those which provided at least a one percent decrease in average pairwise probability of misclassification, and those which did not.

Thus the program STEPLIN was used to investigate two aspects of channel selection: (1) which channels are best for a particular application (data type and segment), and (2) how many (or few) produce at least a one percent decrease in the APPM for a particular application.

FIGURE 1. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING SIGNATURES FROM
SEGMENT 204

FIGURE 2. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING SIGNATURES FROM
SEGMENT 203

23
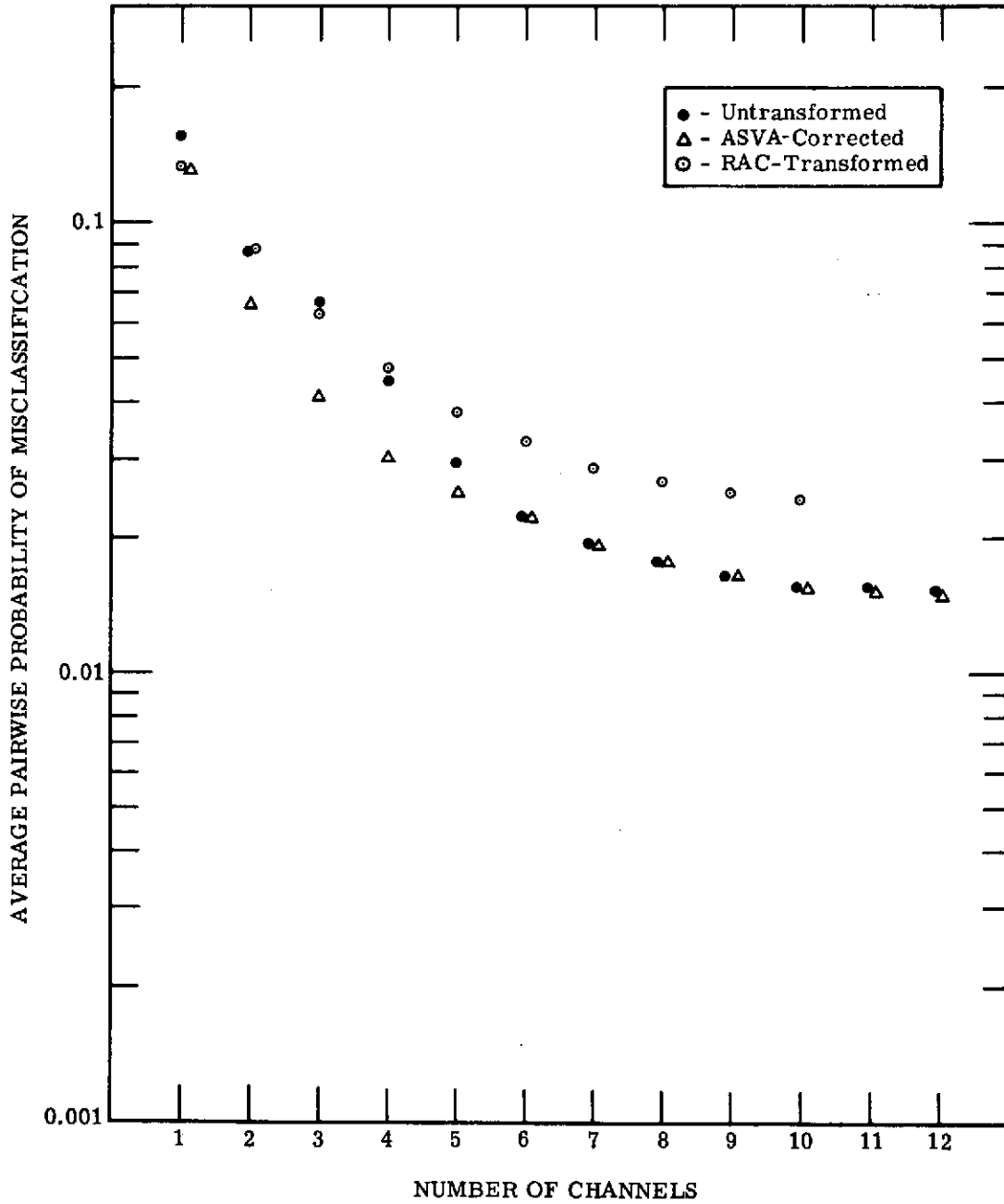
FIGURE 3. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING SIGNATURES FROM
SEGMENT 212
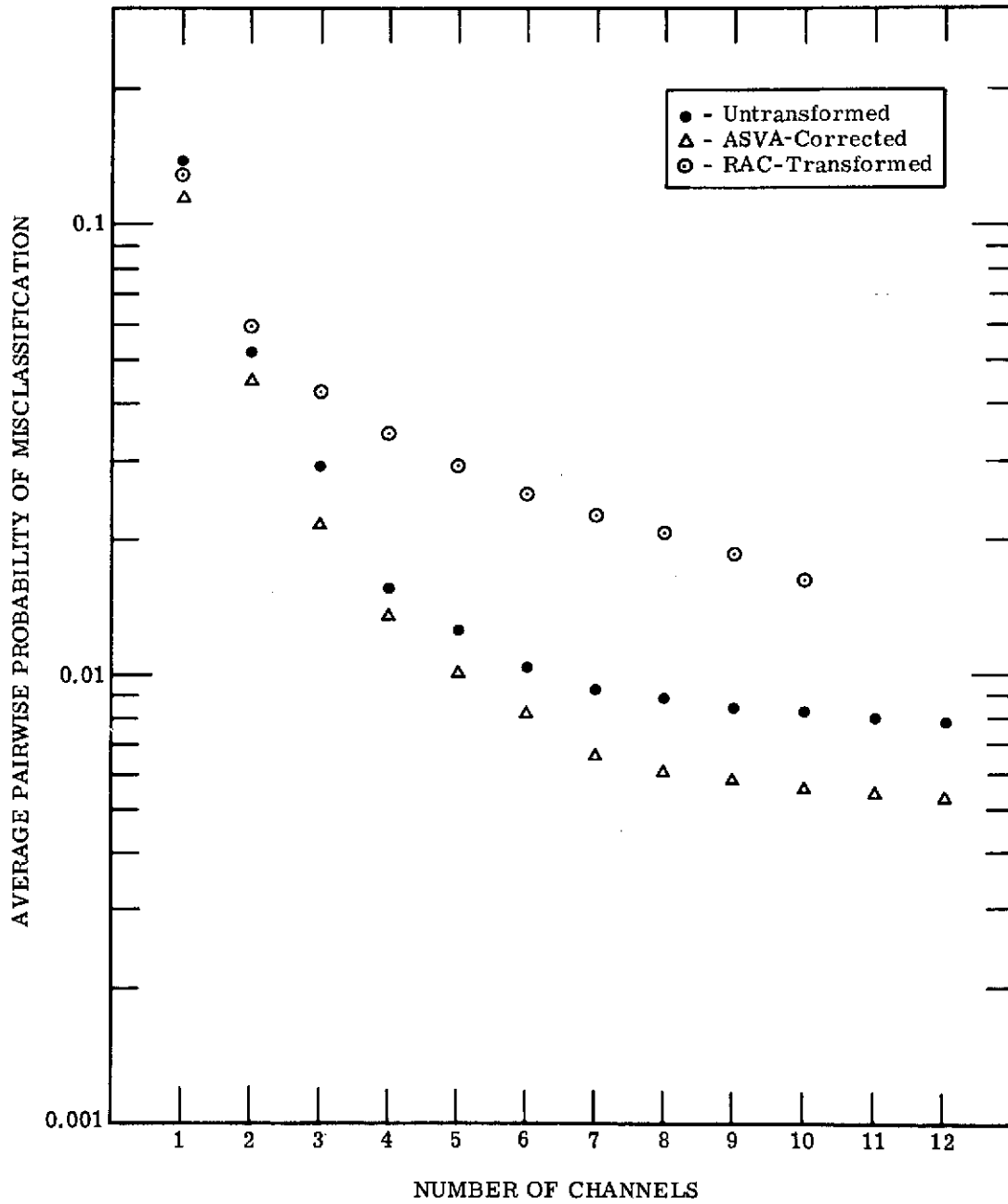
FIGURE 4. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING SIGNATURES FROM
SEGMENT 227

25
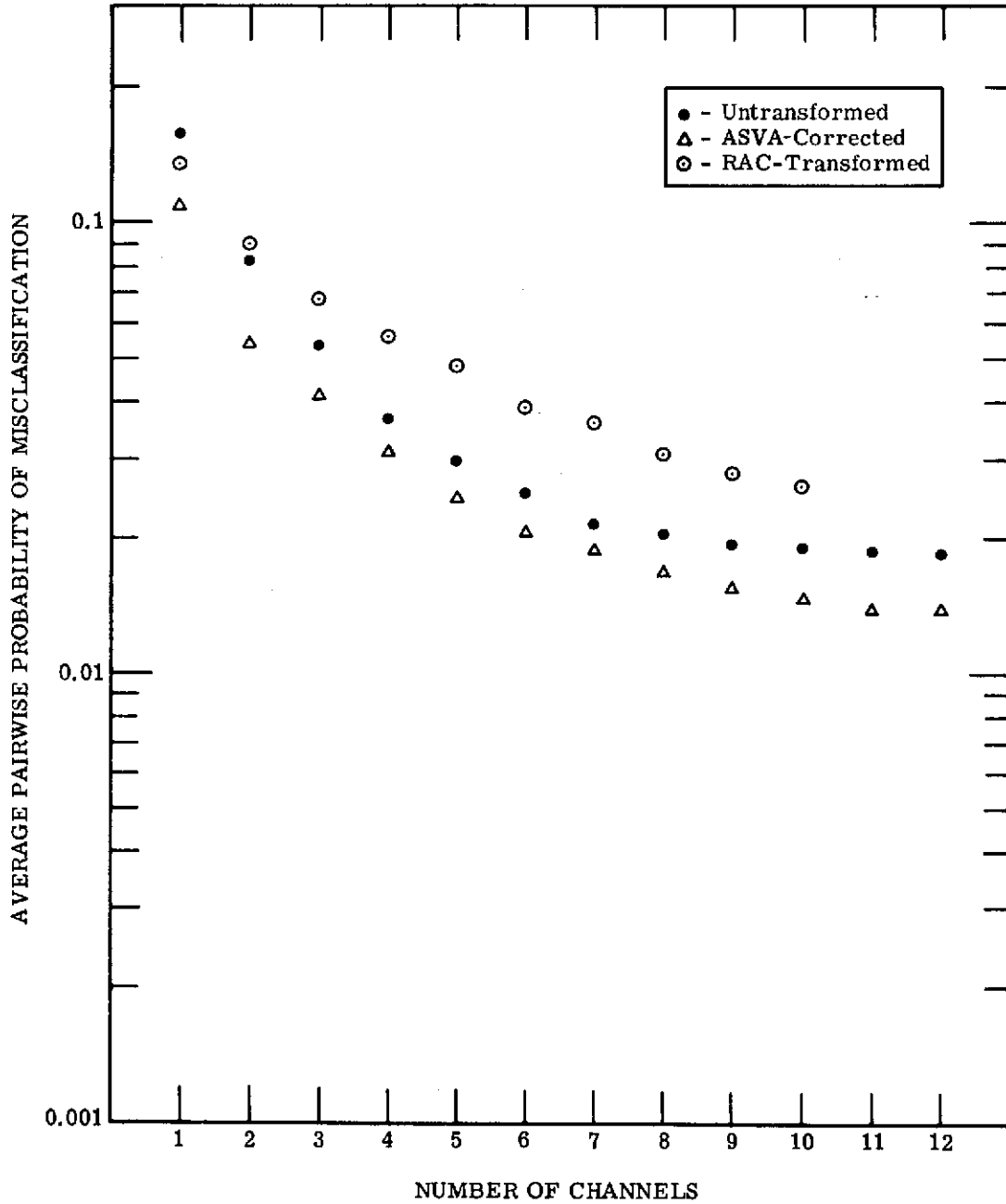
FIGURE 5. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING COMBINED SIGNATURES
FROM SEGMENTS 204 AND 203

26

FIGURE 6. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING COMBINED SIGNATURES
FROM SEGMENTS 204 AND 212

27

FIGURE 7. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR THREE DATA TYPES USING COMBINED SIGNATURES
FROM SEGMENTS 204 AND 227

28
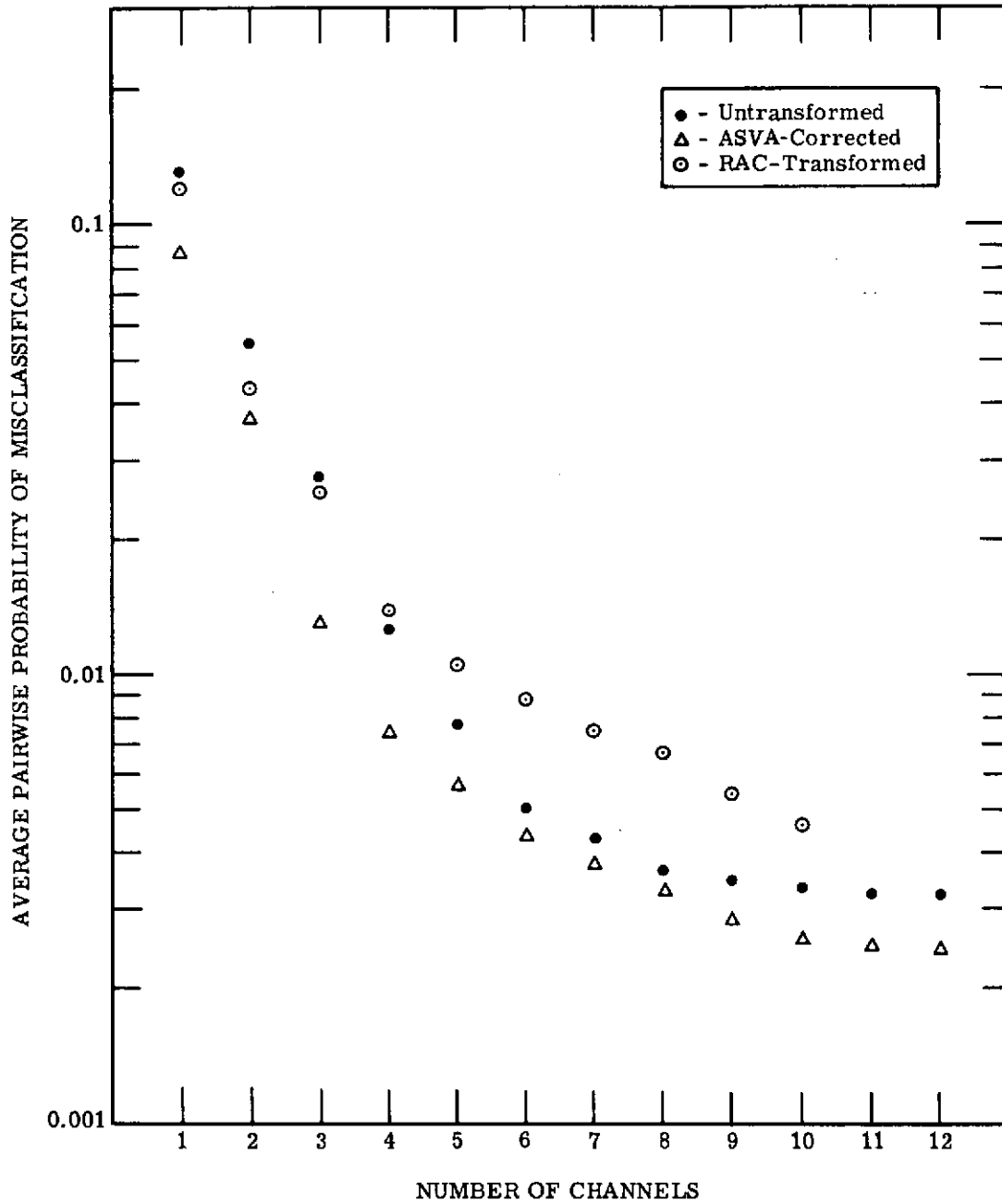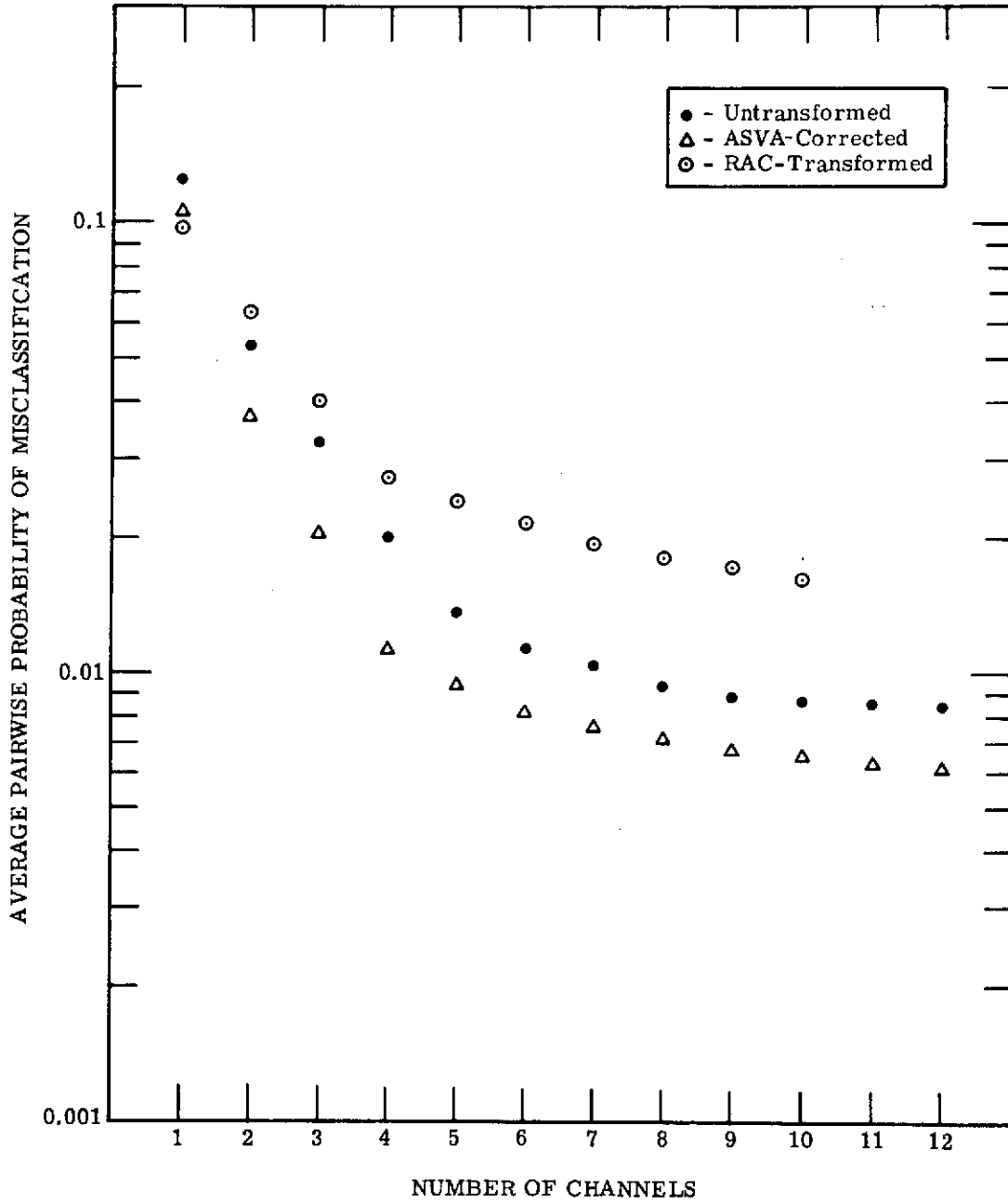
FIGURE 8. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS NUMBER OF CHANNELS FOR THREE DATA TYPES USING COMBINED SIGNATURES FROM SEGMENTS 203, 204, 212 AND 227
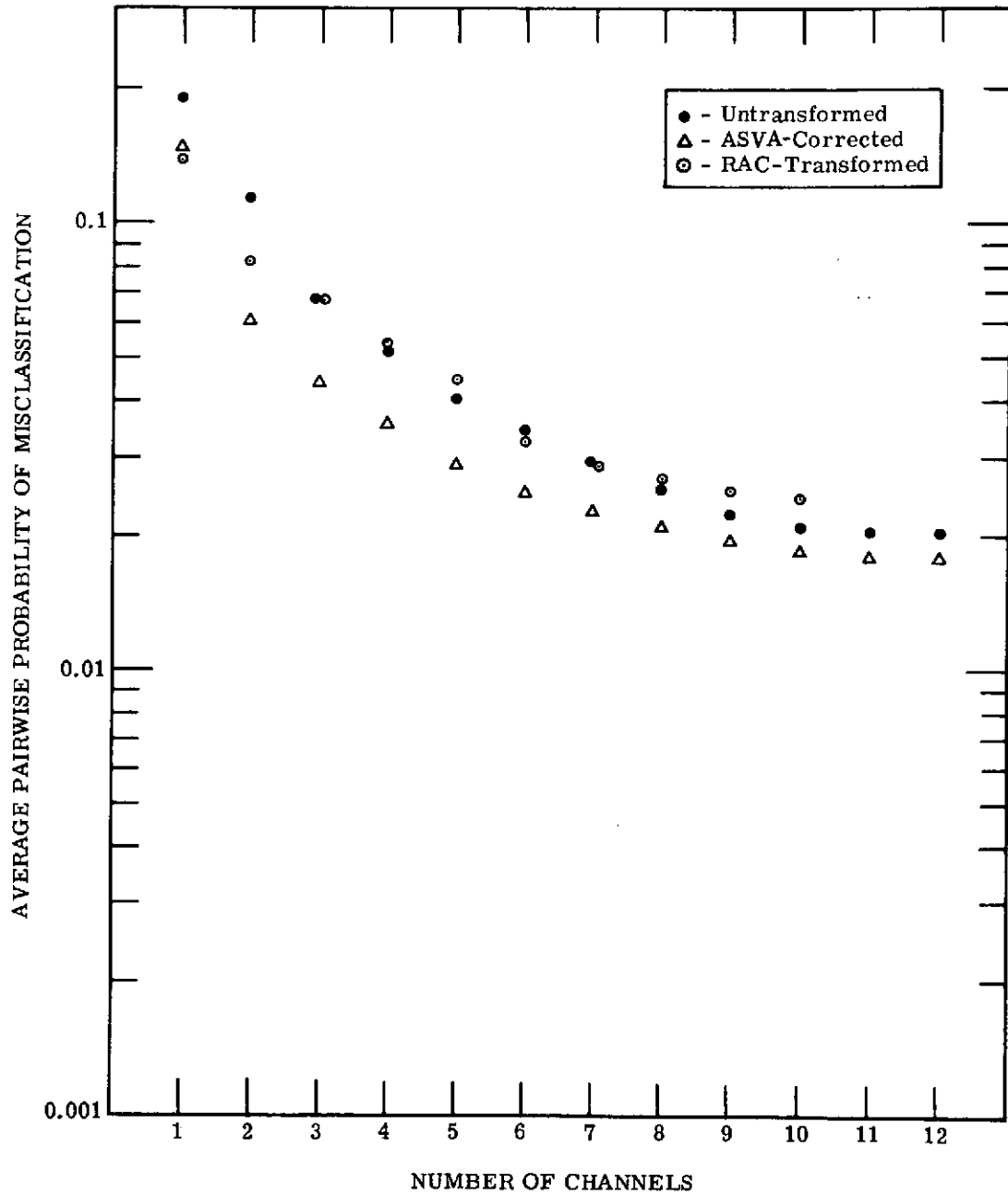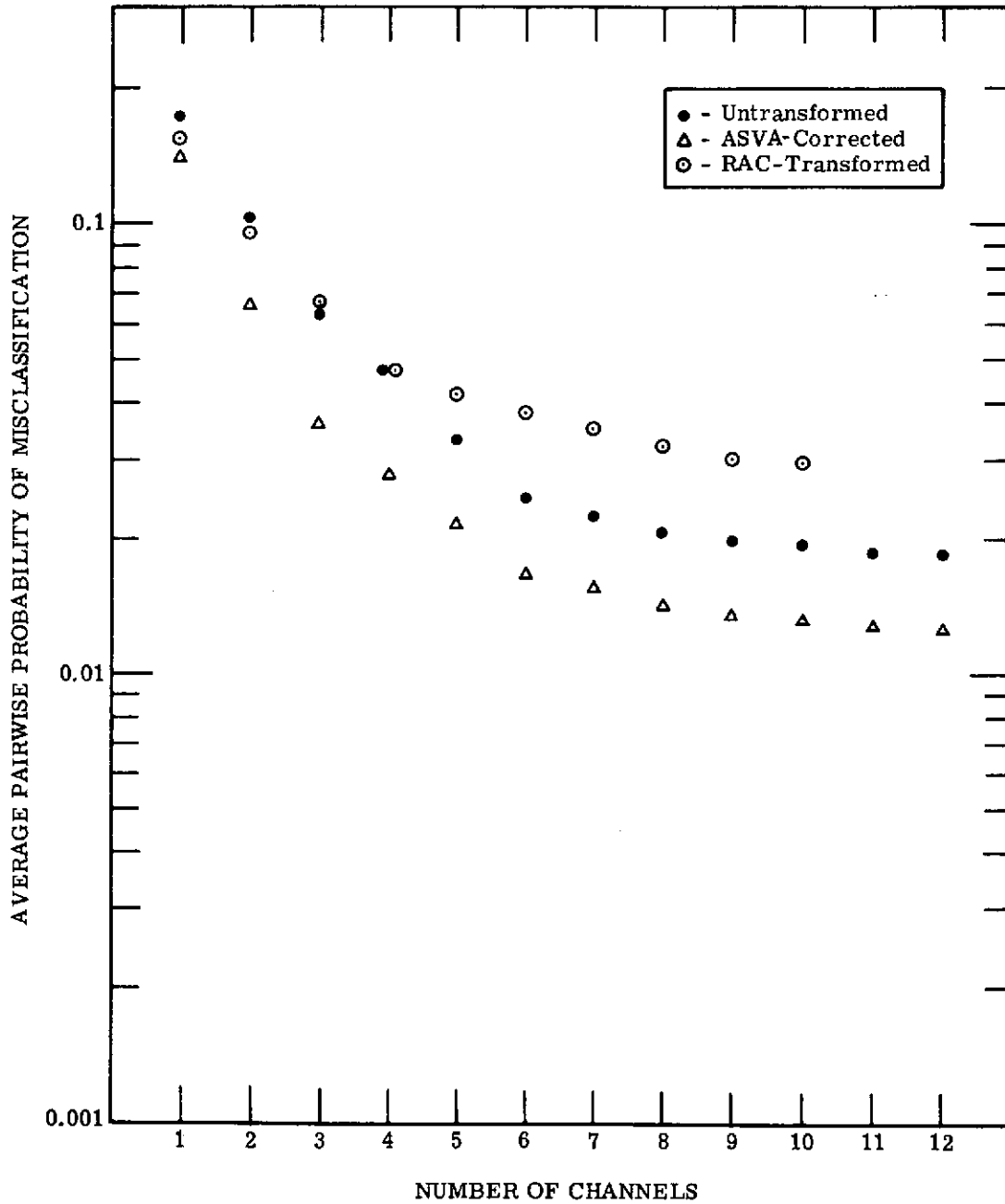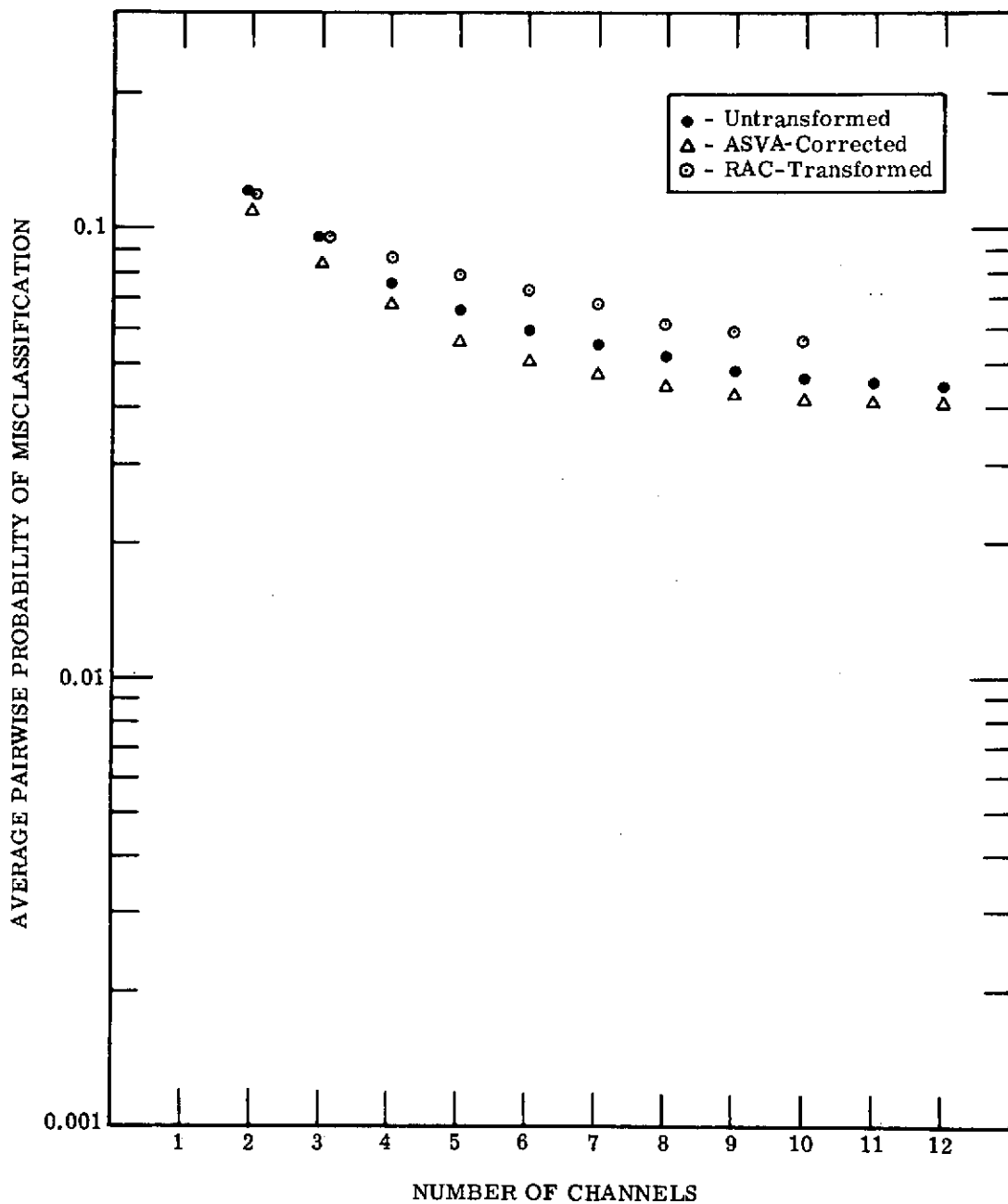
4.2.2 PHASE TWO

In Phase Two we tested the sets of best channels generated by Phase One. For determining classification ability in this investigation, we employed the same test fields, or a similar assortment, that we used last year (only segment 227 was not then in use). Using a variety of classification runs, we first classified each segment according to its own signatures and best set of six channels. Second, each segment other than 204 was classified with signatures and best six channels from segment 204.

4.3 RESULTS

In the following subsections we discuss the results of Phases One and Two. Phase One revealed information about the character of the four segments as well as providing a means for evaluating various signature extension methods. The particular spectral bands necessary to produce at least a certain level of inter-object class separation are also discussed.

Phase Two was intended to verify the results of Phase One. It was realized early in this phase that there were differences between the data sets not revealed in the preceding phase. After a clarification of these differences, various subsets of the best channels as determined in Phase One were used to extend the signatures extracted from one segment to the other three segments.

4.3.1 PHASE ONE

The results of Phase One were twofold. First, sets of channels were determined which, for a particular data type and flight segment, provided more separability of object classes than other channels also considered. Second, the number of channels that met the one-percent criterion was established.

Table 3 presents the results of the STEPLIN channel-selection investigation. The channels are ordered as STEPLIN ranked them with the first or "best" channel leftmost. Within individual segments and considering only those channels satisfying the one-percent criterion, the near-infrared channels predominated over any particular spectral band (or color) in maximizing inter-object class separation. The only exception was segment 203 when the ASVA correction was used; but even here the near-infrared region was at least as well represented as any other spectral region. For the relationship between RAC channels and multispectral scanner channels, see Table 4. Pairwise combinations of segment 204 with other segments and the four-segment combination yielded results similar to the individual-segment case. This dependence on the infrared may be explained in part by the reduced impact of atmospheric scattering as a source of variation in this region. The infrared channels were also less susceptible to noise because they spanned relatively broad regions of the spectrum in comparison to the visible bands.

TABLE 3. STEPLIN CHANNEL-SELECTION BASED ON SEVEN OBJECT CLASSES

Segments	Untransformed	ASVA-Corrected	RAC-Transformed
S-204	11 6 4 8 ∧ 1 9	10 4 8 ∧ 1 9 6	7 3 9 8 ∧ 2 10
S-203	7 8 11 12 ∧ 9 4	6 11 12 ∧ 8 3 9	7 8 9 10 ∧ 6 3
S-212	7 11 9 4 8 ∧ 1	5 9 8 1 ∧ 11 2	7 8 3 9 10 ∧ 2
S-227	10 9 7 4 ∧ 11 12	10 9 2 ∧ 8 11 4	7 3 9 ∧ 8 10 4
S-204 +S-203	6 11 8 4 ∧ 1 12	10 5 8 11 ∧ 7 9	7 9 2 3 ∧ 8 10
S-204 +S-212	7 8 4 10 12 ∧ 1	5 8 10 7 9 ∧ 11	7 3 9 10 ∧ 2 8
S-204 +S-227	7 4 12 8 10 ∧ 1	10 9 4 ∧ 1 8 12	7 9 2 8 ∧ 10 1
S-204 +S-203 +S-212 +S-227	6 8 4 ∧ 10 11 1	3 8 10 4 12 ∧ 6	7 3 9 ∧ 8 4 10

The caret (∧) indicates channels between which the decrease in average pairwise probability of misclassification was less than 1% (the reduction in APPM increases leftwards from the caret).

TABLE 4. RELATIONSHIP BETWEEN RAC CHANNELS
AND MULTISPECTRAL SCANNER CHANNELS

RAC Channel	MSS Channels	MSS Wavebands (μm)
1	1/2	0.46-0.49/0.48-0.51
2	2/3	0.48-0.51/0.50-0.54
3	3/4	0.50-0.54/0.52-0.57
4	4/5	0.52-0.57/0.54-0.60
5	5/6	0.54-0.60/0.58-0.65
6	6/7	0.58-0.65/0.61-0.70
7	8/7	0.72-0.92/0.61-0.70
8	8/9	0.72-0.92/1.00-1.40
9	9/10	1.00-1.40/1.50-1.80
10	10/11	1.50-1.80/2.00-2.60

Non-Adjacent Ratio Channel		
11	9/3	1.00-1.40/0.50-0.54
12	11/1	2.00-2.60/0.46-0.49

Although the channels selected by STEPLIN to maximize separation between object classes varied from segment to segment for a given data type, the most consistent "best" channel selection was observed for the RAC transform. This meant that a "best set" of channels selected from one segment on the basis of the RAC transform would probably work well in classifying the other segments. Applied to the four segments, the RAC transform utilized the smallest subset (seven) of the twelve data channels based on the one-percent improvement criterion.

Two ratios were common to all four segments. One was the ratio of a near infrared band (0.72-0.92 μm) to a red band (0.61-0.70 μm). This ratio is quite sensitive to the amount of ground covered by vegetation, because it enhances the contrasts between vegetative and non-vegetative targets. The other common ratio involved two near infrared bands (1.0-1.4 μm/ 1.5-1.8 μm). These channels were independent enough that the ratio contained useful information, but we are not certain what physical parameter(s) affect this ratio the most. The amount of water in plant leaves is one possibility.

Two additional ratios were common to three out of four segments (not the same three, however). One was the ratio of two green bands (0.50-0.54 μm/0.52-0.57 μm). The choice of this ratio indicated that the quality of the green chlorophyll reflectance region was different among the various object classes. Whether this difference resulted from such factors as plant leaf structure or geometry, or possibly planting density, is not known for certain. The other ratio involved two near-infrared bands (0.72-0.92 μm/1.00-1.40 μm) as did one of the ratios common to all four segments. Again, little was known about what parameter(s) this ratio measured. The Untransformed case showed some uniformity. Channels 4, 7, 8 and 11 occurred more often than any others above the one percent criterion for the individual segments. None of these channels was selected consistently for all segments, however. The top x* channels chosen for use with the ASVA correction showed a lack of uniformity.

With regard to the number of channels required to meet the one-percent criterion, between three and four channels were sufficient for most segments and transforms. Segment 212 was the only one to require more than four channels and it did so for the Untransformed and RAC transform cases only. Segment 212 required four channels after the ASVA correction, whereas the other three segments required only three channels each. This indicated that the object classes for segment 212 were more difficult to distinguish. Examination of the signatures for the seven object classes showed that for the Untransformed segment 212, there was no single channel in which all seven classes were well separated. Every channel had a pair of classes with nearly identical means (and standard deviations averaged about 10 percent of the mean for all channels and object classes). The ratio signatures for segment 212 presented a different case, with much more grouping around certain means in individual ratio channels.

---

*See Table 5.

Some groups of two, three, and four object classes had nearly identical means and there were other groups with slightly different means. Even though the standard deviations for the Ratio of Adjacent Channel case were very small (averaging 3 to 4% of the mean), it was obvious that most of the ratio channels were not separating the various vegetation classes. This is understandable since most vegetation classes exhibit similar spectral shapes in much of the spectrum.

Plots of average pairwise probability of misclassification versus numbers of channels (see Figures 1 through 8) were examined on the basis of the one-percent criterion. A general trend emerged as to which data form does better or worse than the others. In all cases the Ratio of Adjacent Channels had a higher average pairwise probability of misclassification compared with the other two data forms. The ASVA-correction transform produced slightly lower APPM than did the Untransformed case. Overall, all three data forms produced results within a 3% APPM range of one another. Figures 1 through 4 also illustrate the relative ease of object class separability among the four segments. Segment 204 achieved a lower APPM than any of the other segments for a given number of channels. The rate of improvement of the APPM with each additional channel is greater for this segment than any other. These characteristics indicated that the seven object classes, as represented by corresponding signatures containing much of the variety of those classes, were most separable in segment 204. Segment 212 exhibited characteristics (Figure 3) which reinforced the observation made above that the object classes were more difficult to distinguish for this segment. Segments 203 and 227 lay between the other two on the separability continuum.

Values of APPM achieved for best sets of six and best sets of x channels (see Table 5) for any segment or data form revealed that the best x performed nearly as well as the best six in separating the seven object classes, even though x < 6. As indicated in the table, both sets realized APPMs on the order of 0.03. This low value for probability of misclassification indicates that the object classes signatures are separable to a high degree, if the signatures input to STEPLIN are truly representative of the variations of their respective object classes.

When best sets of six and best sets of x channels determined from segment 204 signatures were used with signatures from each other segment as input to STEPLIN, values of APPM were generated which were very similar to those achieved using each segment's own best sets of channels. This indicated that the channels selected from segment 204 as the best six or the best x did very well in separating the object class signatures of the other segments. This observation was reflected also in the similarities between the best channels of the four segments for a given data type as discussed earlier.

Two non-adjacent ratios were also considered in the Phase One effort. One was the ratio of a near-infrared band to a green band (1.0 to 1.4 μm/0.50 to 0.54 μm), designated as ratio 11, and the other was the ratio of a near-infrared band to a blue band (2.0 to 2.6 μm/0.46 to 0.49 μm),

TABLE 5. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION
WITH 6 OR x* CHANNELS

| | | | Probabilities Based on Signatures from: | | | |
			S-203	S-212	S-227	S-204
Untransformed						
	Segment z**:	Best 6 channels	0.011	0.022	0.010	--
	Segment z**:	Best x channels	0.020	0.030	0.016	--
	Segment 204:	Best 6 channels	0.013	0.022	0.012	0.005
	Segment 204:	Best x channels	0.024	0.045	0.020	0.013
ASVA-Corrected						
	Segment z:	Best 6 channels	0.008	0.022	0.008	--
	Segment z:	Best x channels	0.020	0.030	0.021	--
	Segment 204:	Best 6 channels	0.015	0.023	0.010	0.004
	Segment 204:	Best x channels	0.025	0.050	0.031	0.013
Ratio of Adjacent Channels						
	Segment z:	Best 6 channels	0.022	0.033	0.025	--
	Segment z:	Best x channels	0.027	0.037	0.042	--
	Segment 204:	Best 6 channels	0.022	0.033	0.027	0.009
	Segment 204:	Best x channels	0.035	0.047	0.034	0.014

*x - refers to the number of channels after which a 1% decrease in the pairwise probability of misclassification was not attained.

**z - refers to the segment which supplied the signatures for this table.

designated as ratio 12. Ratio 11 was ranked first by STEPLIN for two of the segments (204 and 212) and was not in the top six ratio channels for the other two segments. Ratio 12 was ranked third for segment 203 and second for the other three segments. Plots of APPM versus number of channels for segments 204 and 212 (see Figs. 9 and 10) were made for the 10-ratio and 12-ratio channel cases. Examination shows that the non-adjacent ratios improved the object class separability for these two segments. Environmental variation, however, will hamper the straightforward application of non-adjacent ratios over large areas. Use of this type of ratio will probably require periodic normalization of the non-adjacent ratio signal levels using some well distributed in-scene reference.

The results of Phase One of the investigation can be summarized as follows:

(1) If signatures are representative of object classes within individual flight segments, then four scanner channels should perform well in the classification process for untransformed and ASVA-transformed data, and six scanner channels (two being redundant in the four ratios considered) should perform well for the Ratio of Adjacent Channel transform.

(2) In general, the ASVA correction achieved the lowest APPM and the RAC transform the highest for the three data types. Considering either 'best' six or 'best x' channels, all three types generated values which were acceptably low ($< 5\%$). The relative values of APPM reached by the data types can be compared by some criterion to evaluate the relative effectiveness of the different data types.

(3) The wavebands selected to provide maximum separability of certain object classes for one area can provide good separation of the same object classes from a different area.

(4) Near-infrared data channels predominated in providing inter-object class separability for all three data types.

4.3.2 PHASE TWO RESULTS

Phase Two was originally intended to be simply a check on Phase One, but an analysis of the results revealed an important lesson: more channels can yield worse classification results if the signatures are not truly representative of all the variation within an object class. The "more channels mean higher classification accuracies" principle, which is a fundamental precept in the classification of small areas, cannot be generalized in the consideration of large-area surveys.

In Phase Two, we performed classifications for seven object classes, using sets of best six channels for individual segments as selected by STEPLIN, i.e., each segment was classified with its own signatures and best set of six channels for each of the three data transformation cases. The results are given in the first four columns of Table 6. The ASVA-corrected data

36

FIGURE 9. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR RATIOS OF ADJACENT CHANNELS AND RATIOS OF
ADJACENT AND NON-ADJACENT CHANNELS FOR SIGNATURES FROM SEGMENT
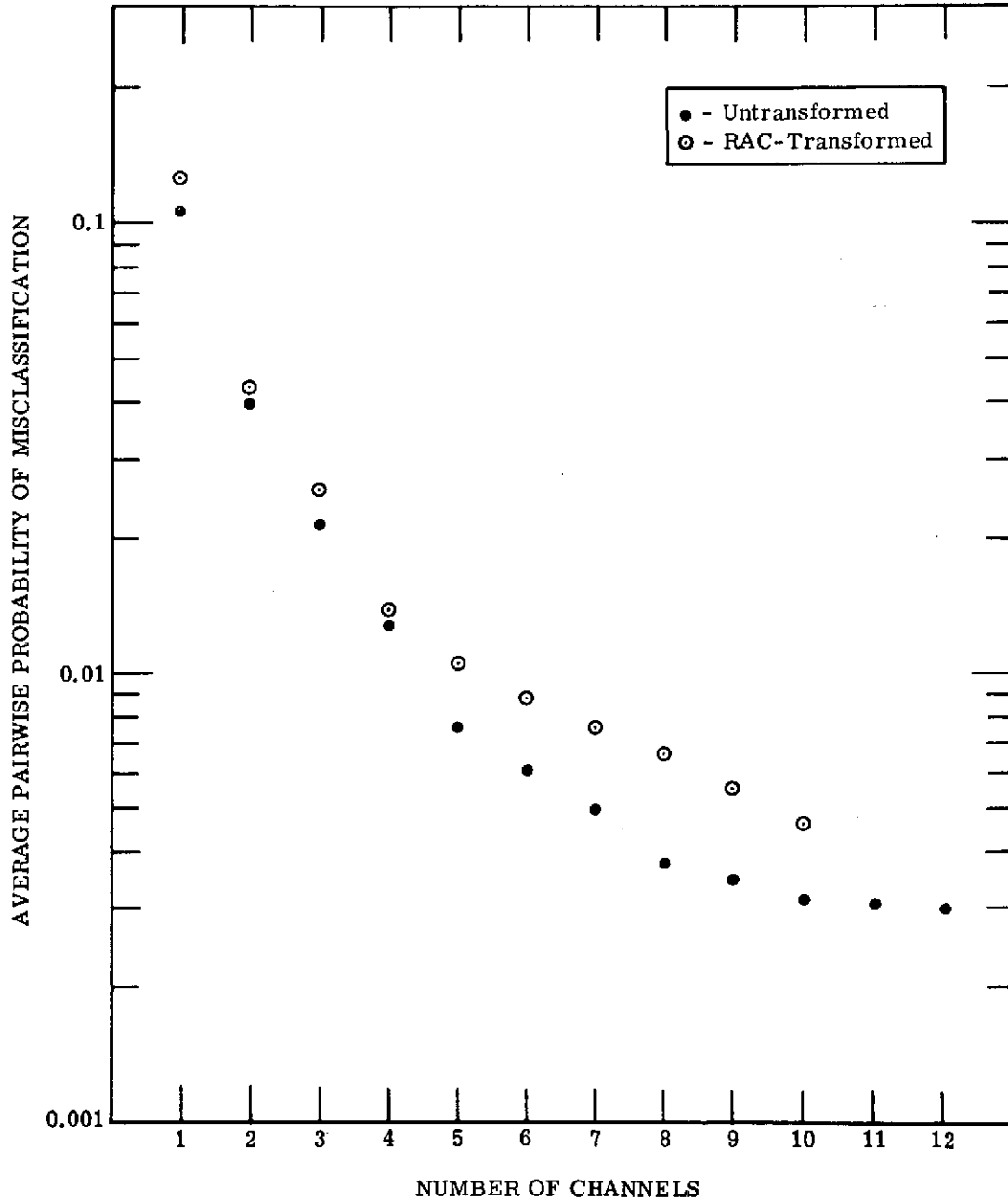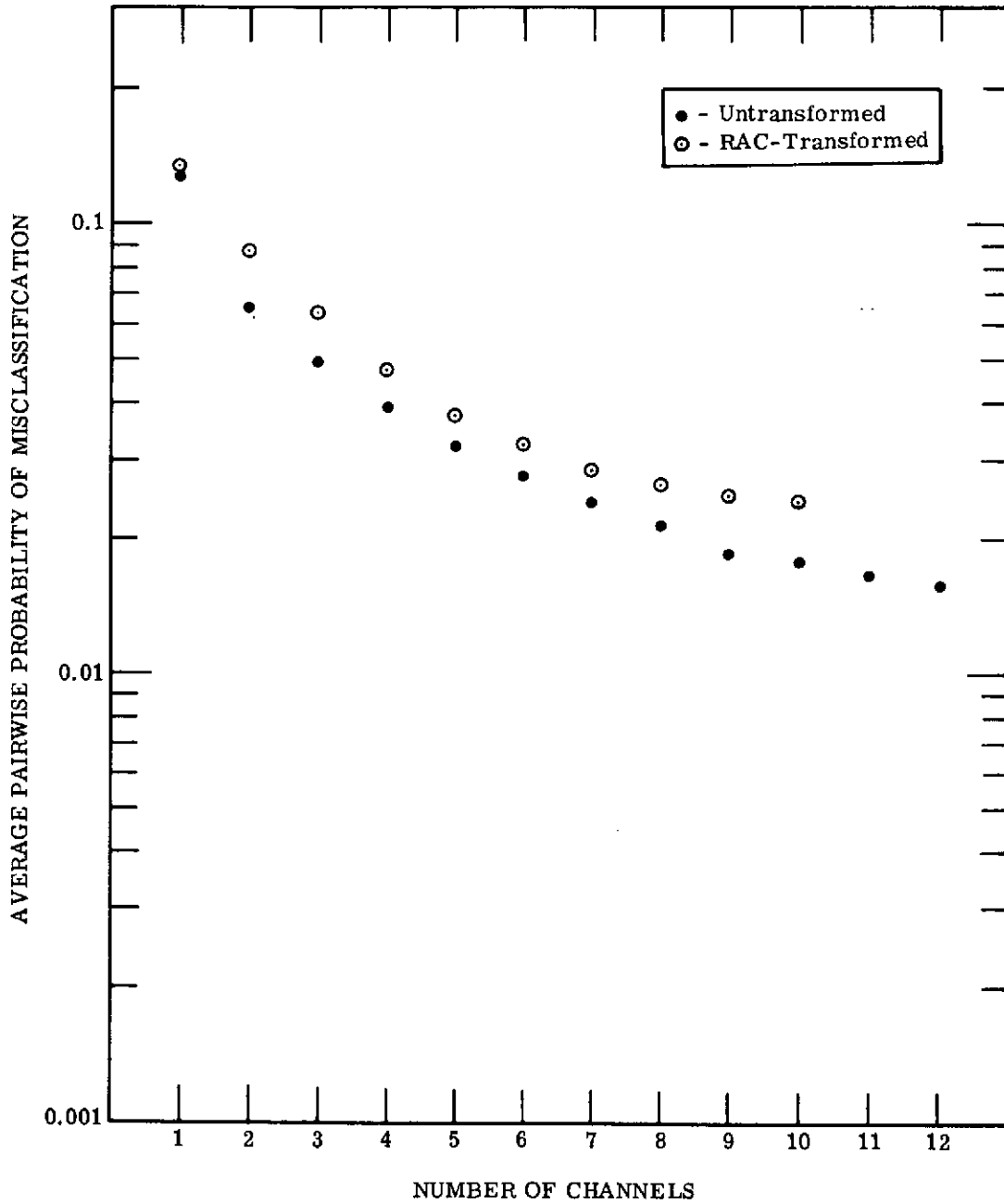204

FIGURE 10. AVERAGE PAIRWISE PROBABILITY OF MISCLASSIFICATION VERSUS
NUMBER OF CHANNELS FOR RATIOS OF ADJACENT CHANNELS AND RATIOS OF
ADJACENT AND NON-ADJACENT CHANNELS FOR SIGNATURES FROM SEGMENT
212

TABLE 6. RECOGNITION ACCURACIES FOR 3 OF 7 OBJECT CLASSES RESULTING FROM 6-CHANNEL
CLASSIFICATION OF UNTRANSFORMED, ASVA-CORRECTED, AND RAC-TRANSFORMED DATA

UNTRANSFORMED

Object Class	Segment 203/203 Middle	Edge	204/204 Middle	Edge	212/212 Middle	Edge	227/227 Middle	Edge	203/204 Middle	Edge	212/204 Middle	Edge	227/204 Middle	Edge
Corn	99.6	94.8	96.6	97.7	96.9	98.0	93.4	--	1.8	0.7	27.8	32.6	0.0	--
Soybeans	89.9	88.9	80.4	82.7	83.0	92.5	92.9	--	0.8	0.6	19.0	28.2	0.0	--
Trees	93.7	86.4	96.6	4.1	91.1	88.7	94.9	--	62.5	9.3	84.1	56.2	0.3	--

ASVA-CORRECTED

Object Class	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge
Corn	95.8	92.5	97.7	91.9	94.0	97.2	93.2	--	94.4	92.0	98.0	99.0	88.8	--
Soybeans	93.3	94.6	88.3	98.9	85.1	96.6	91.5	--	65.4	43.9	62.1	65.4	8.7	--
Trees	96.3	97.2	98.1	84.2	96.2	97.0	98.4	--	95.3	96.8	95.1	87.6	94.0	--

RAC-TRANSFORMED

Object Class	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge	Middle	Edge
Corn	95.2	92.7	95.5	97.6	92.8	97.3	89.6	--	0.5	0.1	0.8	2.7	0.0	--
Soybeans	84.2	81.4	85.7	75.0	80.3	85.2	85.6	--	0.1	0.0	0.9	2.8	0.0	--
Trees	90.4	84.5	96.1	65.3	85.0	82.5	91.5	--	89.9	95.2	11.2	31.2	4.3	--

† - The numerator supplied the test fields; the denominator supplied the object class signatures and set of channels.

were the most accurately classified, followed in order by the Untransformed case and the Ratio of Adjacent Channel case, in relative agreement with Phase One results. The average accuracy on any segment, taken over three object classes of major interest (corn, soybeans and trees), was in excess of 86% for the Ratio transform, greater than 90% for untransformed data, and greater than 91% for ASVA-corrected data when only fields in the middle of the scanned scene (near nadir) were considered. [The relatively high accuracies for side fields of the three object classes for all segments (with in-segment signatures) for untransformed data is to be expected, since the procedure of selecting training fields from a wide range of scan angles was followed.] Trees on the sides of segment 204 were anomalously low in classification accuracy (4.1%), primarily because of an inadvertent exclusion of trees from the side of segment 204 in the training sets which generated the signatures used to classify segment 204. This points out the high susceptibility of the untransformed data to scan angle effects, especially when the object classes are not sampled from throughout the scene. This is not to say, however, that sampling methodology alone will compensate for scan angle effects. Indeed, combining training sets which are in regions strongly influenced by scan angle effects with training sets from regions not so affected could yield a resultant signature which would not be representative of either training region. The ASVA correction and the RAC transformation partially correct for this effect. The fair results from the Ratio of Adjacent Channel transformation and the good results using the ASVA correction for trees on the sides of segment 204 again show the improvement in data quality made by these transforms.

The sets of signatures which had been generated from training sites within a one-mile-square area centered on the middle of the scan line of segment 204 were used to classify the other three segments. The sets of best six channels selected by STEPLIN for use on the three data forms of segment 204 were used with corresponding data forms of segments 203, 212 and 227. The classification results, shown in the last three columns of Table 6, are surprising in several respects.

The ASVA correction, as expected, performed well, extending segment 204 signatures with equal facility to each of the three segments. Corn and trees were quite accurately classified, though soybeans were not classified as well as anticipated. The results from the Untransformed and RAC transform cases, however, were unexpectedly low. Environmental and instrumental variations as well as difficulties with calibration between flight segments evidently caused the poor results for these two data types. These between-segment variations were apparently compensated for, however, by the inter-segment normalization of the ASVA correction. The RAC transform case was especially surprising, since the STEPLIN results of Phase One for that transformation had revealed the most uniform set of best channels over all segments. Closer examination of the signal levels in the six best ratio channels for segment 204 revealed that ratio 8 (0.72-0.92 μm/1.0-1.4 μm) was very different for segment 204

than for the other three segments. The mean signal levels in ratio 8 for segments 203, 212 and 227 were not within three standard deviations of the mean in ratio 8 for segment 204 (in most cases they exceed four standard deviations). Signatures for several object classes for the Untransformed case from each segment were examined in order to re-evaluate environmental differences between the segments. Relative to segment 204, the other three segments showed significant differences in the apparent amount of radiation reaching the scanner for a given object class. The largest differences occurred in the infrared bands (channels 8 through 11), with channel 9 (1.0-1.4 μm) predominating. The Ratio of Adjacent Channels transform possibly ran afoul of wavelength-dependent changes in illumination. However, ratio channel 8 also bridged the eight-band spectrometer and the three-element near-infrared detector (using one band from each); hence, it could have succumbed to inaccuracies in the calibration of the two sensor packages. Other ratio channels showed differences between segments, but none approached the magnitude of Ratio channel 8. These results were not predictable on the basis of the STEPLIN output.

At this point it was thought desirable to do more than simply accept the poor results and relevant explanations for Untransformed and RAC-transformed data. Therefore, classification was performed again, much the same as before, but using the best three channels as chosen by STEPLIN for each data type from the four-segment combined signature case. These channels were selected for several reasons. Utilizing three channels avoided the use of those channels which were subject to large inter-area variations in the cases of Untransformed and RAC-transformed data (inter-segment normalization had greatly reduced this problem for the ASVA-corrected data, of course). For all data types except the ASVA correction, the channels used were also the set of channels which satisfied the one-percent criterion based on APPM and therefore allowed an evaluation of the criterion as a means for determining the number of channels to be used for classification. The combined four-segment signatures were used as a basis for channel selection because within these signatures was contained some measure of the inter-segment object class variability.

A further step taken was to consider only four object classes (corn, soybeans, trees, and bare soil) during classification. This was done to reduce possible interference between ambiguous object classes such as pasture and hay and the four object classes of interest. The ambiguity in an object class such as pasture or hay stems from the fact that they tend to be catchall categories for a variety of types of vegetation and hence tend to have spectral signatures with relatively large variances. Such signatures can overlap others and thereby cause confusion during classification. The increase in classification accuracies due to these two changes was significant (see Table 7). Gains in classification accuracies on the order of 70% to 90% were common when the Untransformed and RAC-transformed signatures from segment 204

41

TABLE 7. THREE-CHANNEL RESULTS IN PERCENT CORRECTLY CLASSIFIED FOR 3 OF 4 OBJECT CLASSES USING SEGMENT 204 SIGNATURES ON THE OTHER THREE SEGMENTS

UNTRANSFORMED (Channels 4, 6, 8)

	Segment 203		Segment 212		Segment 227	
	Middle	Side	Middle	Side	Middle	Side
Corn	96.9	87.7	97.8	95.2	73.1	--
Soybeans	60.3	57.0	74.5	95.0	94.1	--
Trees	87.4	60.5	78.1	71.8	78.1	--

ASVA-CORRECTED (Channels 3, 8, 10)

Corn	98.5	92.2	98.5	99.3	89.2	--
Soybeans	79.9	59.0	81.8	70.7	94.6	--
Trees	94.3	96.9	93.2	97.0	94.1	--

RATIO OF ADJACENT CHANNELS (Channel 3, 7, 9 ratios:
3/4, 8/7, and 9/10, respectively)

Corn	95.1	98.4	89.0	97.6	70.1	--
Soybeans	61.2	77.3	93.0	99.6	93.8	--
Trees	60.0	61.7	46.8	25.3	17.7	--

were extended to the other segments using only three channels. Some improvement also oc-curred with three-channel classification in one object class (soybeans) of the ASVA-corrected data. However, the ASVA-corrected data tended to produce similar results with three- and six-channel classifications. The impact of reducing the number of object classes from seven to four was evaluated for all three data types.

Signatures from segment 204 for the four object classes were extended to segment 203 using sets of best six channels, as shown in Table 8. The classification accuracies shown in this table should be compared with the fifth column of Table 6. The Untransformed and Ratio of Adjacent Channel classification accuracies were improved by less than 5% (absolute) by reducing the number of object classes from seven to four. Corn and trees showed no improve-ment with the ASVA correction, but soybeans improved 15%-20%. Next, four-object-class signatures from segment 204 were extended to segments 212 and 227 using the ASVA correction, with the results shown in Table 9. These results were similar to the segment 203 ASVA re-sults shown in Table 5. In summary, the six-channel, four-object-class results for segment 203 (with 204 signatures) differed from the three-channel, four-object-class results for the same segment by 20% to 90% in the Untransformed case (three channels doing better than six), 0% to 4% in the ASVA-correction case (6 channels doing better than 3), and -36% to +96% in the Ratio of Adjacent Channels case. The negative percentage reflects the fact that the six-channel results for trees on segment 203 using the RAC transform were better than the three-channel results, whereas the classification accuracies for corn and soybeans were much higher when three rather than six channels were used. In all but one instance (trees in segment 203 with the Ratio of Adjacent Channel transform), a reduction in the number of channels used improved results in the Untransformed and Ratio of Adjacent Channel data cases. There was not much difference between six- and three-channel classification results for segments 203 and 212 using the ASVA correction, although there was a noteworthy improvement in soybeans in segment 227 with three-channel classification. Classification accuracy was not expected to improve very much using three channels versus six with the ASVA correction. This was because inter-segment normalization had been utilized with the ASVA correction. The fact that there was only about 4% difference in classification accuracy between the six-channel, four-object-class results and the three-channel, four-object-class results verified these ex-pectations.

The results of Phase Two of the investigation can be summarized as follows:

(1) Ability to classify object classes does not necessarily improve with an increase in the number of data channels used in the classification process.

(2) Signature extension was achieved over a spatial span on the order of 400 km and under a variety of atmospheric and solar conditions.

(3) The Average Signal versus Scan Angle correction, in general, produced results of greater accuracy than those achieved using the Untransformed or Ratio of Adjacent Channel transformed data.

(4) For Untransformed or Ratio of Adjacent Channel transform data to be more successful in extending signatures with a minimum of prior knowledge from one area to another, some way must be found to eliminate non-scene-related, inter-area variations. Good results with the Average Signal versus Scan Angle correction showed the beneficial effects of normalization (which is one way of eliminating inter-area variation).

(5) It is advisable to look for best sets of channels on the basis of both inter-object class separation (spectral contrast) and extendability in data sets where all non-scene-related variation cannot be corrected for by preprocessing.

TABLE 8. SIX-CHANNEL CLASSIFICATION RESULTS USING SEGMENT
204 SIGNATURES ON SEGMENT 203 FOR 3 OF 4 OBJECT CLASSES

		Untransformed		ASVA-Corrected		Ratio of Adjacent Channels Case	
		Middle	Side	Middle	Side	Middle	Side
S-203	Corn	6.5	1.0	96.9	91.8	3.6	1.7
	Soybeans	0.5	0.7	81.7	62.7	0.1	0.0
	Trees	66.2	11.8	95.5	97.0	93.9	98.3

TABLE 9. SIX-CHANNEL CLASSIFICATION RESULTS USING
SEGMENT 204 SIGNATURES ON SEGMENTS 212 AND 227
(ASVA-CORRECTED) FOR 3 OF 4 OBJECT CLASSES

		ASVA-Corrected Data	
		Middle	Side
S-212	Corn	98.2	99.5
	Soybeans	75.0	82.8
	Trees	95.6	96.4
S-227	Corn	88.9	no
	Soybeans	25.6	side fields
	Trees	94.8	

44

5

NEW DATA

On 4 August 73, a data collection flight was made over Lenawee County, Michigan. Adrian, the county seat, is centrally located in the county and lies at about 42°N latitude and 84°W longitude. Agriculture is the county's primary land use. Corn and soybeans were the two major crops in 1973. Topographically the county varies from east to west. The eastern portion tends to be flatter and without the lakes that are present in the west. Multispectral scanner and photographic coverage was obtained using an ERIM aircraft and remote sensing equipment. The mission over the county was flown in eleven passes at a 3.81-km altitude. Ten passes, each 38.4 km long, were made in alternate North and South flight directions covering the county from West to East. An additional pass was flown East to West. The data were gathered between the hours of 9 and 12 EDT (8 to 11 EST) which nearly spanned the morning clear sky window. The clear sky window is bounded on the one side by early morning ground fogs and low light levels and on the other side by late morning cloud buildup.

Ground truth information was gathered by several ERIM personnel and consisted of photographs and Agricultural Stabilization and Conservation Service (ASCS) field reports. Kodak Ektachrome slide film was used to take roadside photographs of about 150 fields. This photography was accomplished over a three-day period and followed the data collection flight by about 10 days. The ASCS records did not cover the entire county; neither was there a complete record of the crops within any one section. ASCS field information was obtained for sections scattered throughout three flight lines, one on the West side, one in the middle and one on the East side of the county. (See Table 10 for numbers and kinds of fields identified in each of the three flight lines.)

Scanner data were inspected by means of an oscilloscope for noise and other gross anomalies, and A/D parameters were established. All 12 MSS data channels were digitized. Smoothing (data averaging) was accomplished during the digitization to reduce data redundancy from oversampling and to reduce the effect of noise in the data.

Photographic imagery gathered from the aircraft was of three varieties: black and white panchromatic, color, and false-color infrared. This photographic record is essential for the strategy of developing more comprehensive large-area survey techniques.

Purposes behind gathering these data were manifold. In the main, this data collection flight was designed to be operational in nature. It followed the pattern of a routine data-gathering mission. As a result the data contain the kinds of variability that can be expected in operational large-area surveys. The data were collected over a two and one-half hour period beginning relatively early in the morning (8:26 EST) and covering hours when change in solar elevation

was near maximum. This should result in changes in illumination (from the first flight pass to the last) greater than those encountered in the Corn Blight data, which were gathered over a two-hour period centered on about 11:30 EST. Data gathered from a single county should tend to show less variation in the effect of latitude on plant maturity than was possible in the Corn Blight data. Since the data were gathered in one morning, soil and plant moisture conditions were expected to be more constant. Also, the multispectral scanner was calibrated two days after the data were collected. Thus all flight passes will utilize the same timely calibration data.

Some additional benefit from this data-set might result from the fact that it represents complete coverage of a regional-size political unit and that it was gathered over a short time frame in the manner of a large-area survey.

TABLE 10. NUMBERS OF FIELDS IN THREE FLIGHT LINES FOR
SEVERAL OBJECT CLASSES IDENTIFIED FROM ASCS RECORDS

Object Class	Number of Fields		
	Flight Line No. 2 (West)	Flight Line No. 6	Flight Line No. 9 (East)
Corn	61	93	120
Soybeans	45	43	98
Wheat	17	37	55
Hay	14	19	7
Oats	5	14	7
Pasture	3	6	1

6

CONCLUSIONS AND RECOMMENDATIONS

It has been shown that in large-area surveys, periodic checking and correcting may be a necessity when processing the data in order to compensate for the many different types of variation which can be expected to influence the character of the data. In this report we have described several means by which this may be accomplished. Of importance are normalization, scan angle correcting, data transforming,and proper selection of wavebands to use in the classification process. In fact, waveband selection was the primary subject of this study.

We learned that a subset of MSS wavebands selected to provide maximum inter-object class separation for one area can provide good separation of the same object-classes in a different area. Although the translation of object-class separability into classification accuracies was greatly hampered by inter-area data variability, we found that high classification accuracies could be maintained by two means: (1) through normalization, such as is employed when performing signature extension on data corrected by the ASVA method, and (2) by eliminating from the classification process those data channels which show the most variability with regard to inter-area differences. Thus it was established that a best set of channels must not only provide satisfactory inter-object class separation but must, in addition, be extendable (maintain suitable signal levels) from one area to another.

The extendability of spectral wavebands over time and space has not been fully explored. This aspect of signature extension, as well as others, should be investigated via data collection methods that approximate to a high degree those used to routinely gather large-area survey data. One data-set gathered utilizing such methodology already exists at ERIM; it is for a regional-size political unit — Lenawee County.

In the future, signature extension must concern itself more deeply with such parameters as waveband suitability and number. It is also desirable to explore new and beneficial techniques for data correction. The availability of data gathered under operational conditions will be extremely helpful in conducting such research.

DISTRIBUTION LIST

NASA Johnson Space Center
Earth Observations Division
Houston, Texas 77058

ATTN: Dr. A Potter/TF3	(4)
ATTN: Mr. Robert MacDonald	(1)
ATTN: Mr. B. Baker/TF3	(8)
ATTN: Mr. R. Dean Bretton/TF53	
Chief Earth Resources Research Data Facility	(8)
ATTN: Mr. A. H. Watkins/HA	(1)

NASA Johnson Space Center
Facility & Laboratory Support Branch
Houston, Texas 77058

ATTN: Mr. D. Riley BB631/B4 (1)

NASA/Johnson Space Center
Computation & Flight Support
Houston, Texas 77058

ATTN: Mr. Eugene Davis/FA (1)

NASA Headquarters
Washington, D.C. 20546

ATTN: Mr. C. W. Mathews (1)

U.S. Department of Agriculture
Agricultural Research Service
Washington, D.C. 20242

ATTN: Dr. Robert Miller (1)

U.S. Department of Agriculture
Soil & Water Conservation Research Division
P.O. Box 267
Weslaco, Texas 78596

ATTN: Dr. Craig Wiegand (1)

U.S. Department of Interior
Geological Survey
Washington, D.C. 20244

ATTN: Dr. James R. Anderson (1)

Director, Remote Sensing Institute
South Dakota State University
Agriculture Engineering Building
Brookings, South Dakota 57006

ATTN: Mr. Victor I. Myers (1)

U.S. Department of Interior
Fish & Wildlife Service
Bureau of Sport Fisheries & Wildlife
Northern Prairie Wildlife Research Center
Jamestown, North Dakota 58401

ATTN: Mr. Harvey K. Nelson (1)

U.S. Department of Agriculture
Forest Service
240 W. Prospect Street
Fort Collins, Colorado 80521

ATTN: Dr. Richard Driscoll (1)

U.S. Department of Interior
Geological Survey
Water Resources Division
500 Zack Street
Tampa, Florida 33602

ATTN: Mr. A. E. Coker (1)

U.S. Department of Interior
Director, EROS Program
Washington, D.C. 20244

ATTN: Mr. J. M. Denoyer (1)

Earth Resources Laboratory, GS
Mississippi Test Facility
Bay St. Louis, Mississippi 39520

ATTN: Mr. R. O. Piland, Director (1)
ATTN: Mr. Sid Whitley (1)

U.S. Department of Interior
Geological Survey
GSA Building, Room 5213
Washington, D.C. 20242

ATTN: Mr. W. A. Fischer (1)

NASA Wallops
Wallops Station, Virginia 23337

ATTN: Mr. James Bettle (1)

Purdue University
Purdue Industrial Research Park
1200 Potter
West Lafayette, Indiana 47906

ATTN: Dr. David Landgrebe (1)
ATTN: Dr. Philip Swain (1)
ATTN: Mr. Terry Phillips (1)

U.S. Department of Interior
EROS Office
Washington, D.C. 20242

ATTN: Dr. Raymond W. Fary (1)

U.S. Department of Interior
Geological Survey
801 19th Street, N.W.
Washington, D.C. 20242

ATTN: Mr. Charles Withington (1)

U.S. Department of Interior
Geological Survey
801 19th Street, N.W.
Washington, D.C. 20242

ATTN: Mr. M. Deutsch (1)

U.S. Geological Survey
801 19th Street, N.W., Room 1030
Washington, D.C. 20242

ATTN: Dr. Jules D. Friedman (1)

U.S. Department of Interior
Geological Survey
Federal Center
Denver, Colorado 80225

ATTN: Dr. Harry W. Smedes (1)

U.S. Department of Interior
Geological Survey
Water Resources Division
901 S. Miami Ave.
Miami, Florida 33130

ATTN: Mr. Aaron L. Higer (1)

University of California
School of Forestry
Berkeley, California 94720

ATTN: Dr. Robert Colwell (1)

School of Agriculture
Range Management
Oregon State University
Corvallis, Oregon 97331

ATTN: Dr. Charles E. Poulton (1)

U.S. Department of Interior
EROS Office
Washington, D.C. 20242

ATTN: Mr. William Hemphill (1)

Chief of Technical Support
Western Environmental Research Laboratories
Environmental Protection Agency
P.O. Box 15027
Las Vegas, Nevada 89114

ATTN: Mr. Leslie Dunn (1)

ERIM

NASA, Langley Research
Mail Stop 470
Hampton, Virginia 23365

ATTN: Mr. William Howle (1)

U.S. Geological Survey
Branch of Regional Geophysics
Denver Federal Center, Building 25
Denver, Colorado 80225

ATTN: Mr. Kenneth Watson (1)

NAVOCEANO, Code 7001
Naval Research Laboratory
Washington, D.C. 20390

ATTN: Mr. J. W. Sherman, III (1)

U.S. Department of Agriculture
Administrator
Agricultural Stabilization and Conservation Service
Washington, D.C.

ATTN: Mr. Kenneth Frick (1)

Pacific Southwest Forest & Range Experiment Station
U.S. Forest Service
P.O. Box 245
Berkeley, California 94701

ATTN: Mr. R. C. Heller (1)

Pacific Southwest Forest & Range Experiment Station
U.S. Forest Service
P.O. Box 245
Berkeley, California 94701

ATTN: Dr. P. Weber (1)

NASA/Johnson Space Center
Mission Planning & Analysis Division
Houston, Texas 77058

ATTN: Mr. H. G. De Vezin/FM8 (1)

University of Texas at Dallas
Box 688
Richardson, Texas 75080

ATTN: Dr. Patrick L. Odell (1)

Department of Mathematics
University of Houston
Houston, Texas 77004

ATTN: Dr. Henry Decell (1)

Institute for Computer Services and Applications
Rice University
Houston, Texas 77001

ATTN: Dr. M. Stuart Lynn (1)

U.S. National Park Service
Western Regional Office
450 Golden Gate Avenue
San Francisco, California 94102

ATTN: Mr. M. Kolipinski (1)

U.S. Department of Agriculture
Statistical Reporting Service
Washington, D.C. 20250

ATTN: D. H. VonSteen/R. Allen (2)

U.S. Department of Agriculture
Statistical Reporting Service
Washington, D.C. 20250

ATTN: Mr. H. L. Trelogan, Administrator (1)

Department of Watershed Sciences
Colorado State University
Fort Collins, Colorado 80521

ATTN: Dr. James A. Smith (1)

Lockheed Electronics Co.
16811 El Camino Real
Houston, Texas 77058

ATTN: Mr. R. Tokerud (1)

TRW System Group
Space Park Drive
Houston, Texas 77058

ATTN: Dr. David Detchmendy (1)

IBM Corporation
1322 Space Park Drive
Houston, Texas 77058

ATTN: Dr. D. Ingram (1)

S&D—DIR
Marshall Space Flight Center
Huntsville, Alabama 35812

ATTN: Mr. Cecil Messer (1)

Code 168-427
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, California 91103

ATTN: Mr. Fred Billingsley (1)

NASA/Johnson Space Center
Technical Library Branch
Houston, Texas 77058

ATTN: Ms. Retha Shirkey/JM6 (4)

NASA Headquarters
Washington, D.C. 20546

ATTN: Mr. W. Stoney/ER (1)
ATTN: Mr. Leonard Jaffe/ER (1)
ATTN: Mr. M. Molloy/ERR (1)
ATTN: Mr. G. Thorley/ERR (1)

Ames Research Center
National Aeronautics and Space Administration
Moffett Field, California 94035

ATTN: Dr. I. Poppoff (1)

Goddard Space Flight Center
National Aeronautics and Space Administration
Greenbelt, Maryland 20771

ATTN: Mr. W. Nordberg, 620 (1)
ATTN: Mr. W. Alford, 563 (1)

Lewis Research Center
National Aeronautics and Space Administration
21000 Brookpark Road
Cleveland, Ohio 44135

ATTN: Dr. Herman Mark (1)

John F. Kennedy Space Center
National Aeronautics and Space Administration
Kennedy Space Center, Florida 32899

ATTN: Mr. S. Claybourne/FP (1)

NASA/Langley
Mail Stop 214
Hampton, Virginia 23665

ATTN: Mr. James L. Raper (1)