N72-29421

SECTION 130

DETAILED INTERPRETATION AND ANALYSIS OF SELECTED CORN BLIGHT WATCH DATA SETS*

by

R. F. Nalepka
J. P. Morgenstern
W. L. Brown
Willow Run Laboratories
The University of Michigan
Ann Arbor, Michigan

INTRODUCTION

The Willow Run Laboratories of The University of Michigan participated in the Corn Blight Watch Experiment by collecting multispectral scanner data on a biweekly basis over thirty segments in western Indiana and by routinely processing data from fifteen of these segments. In addition to these major efforts, we also undertook a more detailed study of selected multispectral scanner data sets. A brief description of this study and the results achieved therein are the subject of this paper.

A detailed interpretation and analysis of selected corn blight data sets was undertaken in order to better define the present capabilities and limitations of agricultural remote multispectral sensing and automatic processing techniques and to establish the areas of investigation needing further attention in the development of operational survey systems. While the emphasis of this effort was directed toward the detection of various corn blight levels, problems related to the more general task of crop identification were also investigated. The goals of this effort were to: 1) investigate and define improved data preparation and processing techniques and approaches; 2) determine the relative characteristics of crop signatures and their discriminability; and 3) examine the usefulness of signatures on data sets other than those from which they were extracted.

Since the analog recognition computer (SPARC) was fully committed to the more routine aspects of processing and since the detailed interpretation and analysis required more in the way of quantitative information, our CDC 1604 digital computer was employed for this investigation.

* The work reported in this paper was supported by NASA under Contract NAS 9-9784.

Three segments in the intensive study area of Western Indiana over which multispectral scanner data were being gathered on a biweekly basis were selected as potential sites to be included in this investigation. To support this investigation, ground information, in addition to that provided by the county agents, was gathered throughout much of the scanner data collection period. This information included the location of fields, the crop planted therein, as well as the condition of the crop.

DATA PREPARATION AND PROCESSING

As already mentioned, one of the goals of this study was to investigate and define improved data preparation and processing techniques. The need for such an effort is dictated by information we have gained through our close association with multispectral scanner systems and experiences we have had in processing multispectral data. Our experiences have shown that potentially many problems may exist in each and every data set which, if not corrected, could significantly reduce the accuracy of recognition results for that data. Some of these potential problems are instrument-related while others are associated with the radiation environment and the scene being scanned. These problems include: 1) level shifts and gain changes in the recorded data resulting from instabilities in system electronics and tape speed; 2) misregistration of data between spectral bands due to unequal resolution in all the bands, the lack of optical alignment, either by design or otherwise, of the detectors in all bands, or the imperfect alignment of the tape recorder record and playback heads; 3) noisy data resulting from a combination of insufficient radiation input and lack of detector sensitivity; 4) variations in signal levels as a function of scan angle due to nonuniform angular sensitivity of the scanner, the effects of atmospheric scattering, and bidirectional reflectance effects; and 5) changes in the scene illumination level during the data collection mission.

All of the above problems could seriously affect one's ability to generate accurate classification maps and extract useful information. Without going into much detail here*, we would like to present one of our approaches to the solution of these problems. Prior to digitizing the data, problems of misregistration or skew are eliminated by aligning the data through the use of electrical delay lines. The reference signal used for alignment is one which is recorded in each band during data collection as the scan mirror views a reference source in the scanner. Level shifts are eliminated by clamping the data for each scan line in each band to the dark level signal (that signal which is generated when scanning the dark interior of the scanner housing and which produces zero radiance input to the system).

* More details will be included in the Corn Blight Watch Final Report.

Any gain changes and variations in scene illumination which would produce the same effect as changing the gain are accounted for by scaling the data in each band to the "sun sensor" signal. The sun sensor, which is scanned once for every revolution of the scan mirror, monitors the level of radiation incident upon a flat opal glass plate atop the aircraft. The problems with system noise, as well as those due to misregistration of data in the flight direction, can be significantly alleviated by taking advantage of the fact that in many cases successive scan lines overlap. Rather than not digitizing and processing all scan lines (a common approach), the lines containing largely redundant information are combined by averaging.

All of the above operations can be carried out without specific reference to the video, the data which is generated when scanning the scene. We have chosen to designate such operations as data preparation. One other operation in this category can be carried out. If the angular responsivity of the scanner is nonuniform, and this nonuniformity is known and fixed, the effect of the nonuniformity can be removed from the data. Depending, however, on what other operations are planned for eliminating angle effects, the removal of this effect may be accomplished simultaneously with the removal of the other angle effects. As mentioned earlier, these other effects are due to scattering in the atmosphere and bidirectional reflectance. We have chosen to designate as preprocessing those operations which are meant to reduce or eliminate effects on the data which originate outside the scanner.

One simple form of preprocessing which we have found to be useful assumes that the scene, over its entirety, contains an approximately equal distribution of all objects of interest at all angles. (It is felt that this assumption is valid for most areas devoted to farming.) In this approach, the average signal variation as a function of scan angle is computed for each spectral band. The average signal variation includes the nonuniform angular responsivity of the scanner if it has not already been eliminated.

The effects of scanner angular responsivity, atmospheric transmittance, and bidirectional reflectance are all multiplicative in nature. That is, the radiation incident on an object being viewed at a given angle is multiplied by its reflectance, the transmittance of the atmosphere between the object and the sensor, and the scanner responsivity at that particular angle of view to form the radiation incident on the detector. In the absence of significant path radiance (radiation scattered into the receiver by the atmosphere), which is an additive effect, much of the angular variation in the signal can be eliminated by dividing each scan line of data by the normalized averaged signal variation. Since the data about which we are concerned was gathered under relatively clear atmospheric conditions, path radiance effects would be minimal, thereby justifying this preprocessing approach.

In order to illustrate the importance of properly preparing and preprocessing the data, we now present some examples from data gathered over Segments 203 and 212. The importance of clamping the data to the dark level can be seen in Figure 1 where two histograms, one for the first minute and the other for the second and third minutes of data collection for a particular run, are presented. It is seen here that the dark level varied on the order of $\pm 3\%$ during each of the two periods examined. This variation in itself could seriously affect the ultimate discrimination capability in processing the data. Here, however, an additional shift in the mean dark level of about 5% also occurred during the run.

In Figure 2, examples are shown of the average angular signal variation for two data sets (43M 203 and 43M 212). It is clear on examining the figure that significant variations in the average angular signal occurred for both data sets with more variation present in Segment 203 data. This may be explained by examining the location of the sun during the two data collection flights. For Segment 203 data, the solar elevation was less and the solar azimuth was more easterly than that during the collection of Segment 212 data. With the sun, both lower in the sky and more nearly perpendicular to the north-to-south aircraft flight path, a larger variation of reflectance with scan angle resulted.

A more specific example of the effects of the angular signal variations is shown in Figure 3. Here, we see histograms in two spectral bands of many samples of soybeans and trees plotted as a function of their location in the scene with respect to the data collection aircraft. The effect on both soybeans and trees is a very obvious shift to higher signal value as the scan mirror rotates from east to west. An important result of this shift that is especially noticeable in spectral band 10 is the similarity of signal levels for soybeans on the east and trees on the west side of the aircraft. Obviously, this similarity would create problems in discriminating between and properly classifying both soybeans and trees independent of their location in the scene. These similarities become even more important when it is realized that of the four major object classes in this data set (corn, soybeans, pasture, and trees), soybeans exhibits the highest average signal level in all bands while trees exhibit the lowest average signal levels. Therefore, the signals for both corn and pasture fall between these extremes and one can imagine the confusion that exists between the four object classes and the effect that could be expected on the classification accuracy of the unpreprocessed data set.

The effects of preprocessing the data on the range of signal means for corn are shown in Figure 4. It is seen that a significant reduction in corn signal range was accomplished by preprocessing. Although not illustrated here, similar reductions were achieved for soybeans, pasture, and trees with the result that these crops now exhibited more unique and more easily discriminable signatures.

We have attempted, in the foregoing discussion, to illustrate the importance of careful data preparation and preprocessing in order to permit the extraction of the maximum amount of information from any multispectral scanner data set. Upon further research, the approaches and techniques described may prove not to be the best. Even so, we believe that the application of relatively simple techniques can significantly improve the ability to extract useful information from multispectral scanner data.

SIGNATURE CHARACTERISTICS AND DISCRIMINABILITY

The second goal of this study was to determine the relative characteristics of crop signatures and their discriminability. While the characteristics of signatures of the various levels of corn blight were of prime importance, the signatures of other major crops and ground covers were also important since they might affect the recognition accuracy of corn.

One of the data sets which was studied and will be reported on here was gathered during Mission 43M over Segment 212. This data set was chosen because it included the first occurrences of high levels of corn blight and fairly complete ground information was available. More manpower was used in gathering ground information at this time since this data set was selected by the personnel of WRL and LARS as a study set to enable a check of the routine recognition processing being carried out at the two facilities.

At the time of data collection (August 17), healthy corn plants were approximately 7' - 8' tall with an estimated ground cover in most corn fields of 90 - 100%. The soybean plants were predominantly 3' - 4' tall with a ground cover of 75 - 90%. Hay fields in the segment were at various stages of maturity, with some having been cut just before the mission. Some fields surveyed early in the year as winter wheat were being used to grow hay or as pastures. Others of the fields had recently been disked. The remaining acreage within the segment consisted of farmsteads, woodlots, and pasture, with a small number of fields planted to oats.

The preparation of this data set for processing and analysis followed the approach described in the previous section. Having prepared and preprocessed the data, the next step was to extract signatures of the various crop and crop-condition classes for analysis. It was decided, for this analysis, to include only fields which appeared to be relatively uniform in the aerial photography and scanner imagery and to further limit the fields to those for which ground information was available. A large number of fields or portions of fields satisfying these criteria were selected.

The signatures from a number of the corn fields were examined using a clustering procedure (this procedure separates data points into groups or clusters and assigns data points exhibiting similar multispectral characteristics to a single cluster). If there was to be any hope of discriminating among levels of blight, the data points from fields of equal blight level should be clustered together. This, however, was not the case. Table I depicts the results of the cluster analysis with the data being assigned to six clusters. There seems to be no correlation between blight level and cluster assignment. However, a weak correlation does exist between the clusters and the location of each field in the segment (east, middle, or west) and the orientation of the rows within each field. As seen in Table I, clusters 5 and 6 include all the selected fields which were located on the East side of the segment, while clusters 3 and 4 include all those fields planted in a North-South direction. These facts seem to suggest that the angle correction applied to the data did not totally eliminate the angle effect present in scanning fields of corn and that this residual effect along with the variable effects of other stresses overshadow any detectable differences in radiation due to levels of corn blight in the range 0 - 3.

In order to check this result, additional analysis of corn blight level signatures were carried out. For this analysis, additional signatures, including some from fields of corn blight levels 4 and 5, were extracted. The mean signal values for each of the signatures (a total of 45) in each of the twelve spectral bands were plotted. Here, too, it was clear that discrimination among levels 0 - 3 would not be possible since the ranges of mean signal values overlapped almost completely in each band. This is illustrated in Figure 5 where the values in five of the spectral bands are depicted. The one somewhat promising indication offered by Figure 5 is the partial separation of blight levels 4 and 5 from levels 0 - 3 in spectral bands 8 and 9. This suggests that levels 0 - 3, 4, and 5 may be separable.

While it is necessary to have unique blight level signatures to insure the ability to discriminate among blight levels, the uniqueness among blight level signatures may not be a sufficient condition for accurate recognition if the signatures of other object classes in the scene are not also distinct. To determine whether distinct signatures existed which would permit the generation of accurate crop recognition maps, signatures were extracted from selected fields of soybeans, pasture, hay, and trees. The ranges of the mean signals in each spectral band for each of these classes was determined. This information along with that for corn is plotted in Figure 6. Upon examining this figure, it is clear that, with the exception of hay and pasture, the signal ranges for every pair of object classes exhibits no overlap in at least one spectral band. The uniqueness of the signal ranges is a good indication that a fairly accurate crop recognition map could be produced if training signatures encompassing the appropriate ranges were utilized.

A crop recognition map, a portion of which is illustrated in Figure 7, was generated. (The field identifications are shown in Figure 8). In the remainder of this section, the signatures used to train the computer and the results of the recognition process are discussed.

As previously mentioned, it was necessary that the training signatures for each object class encompass the range of variability of that class. To satisfy this requirement, a single signature was generated for each class by combining the two individual signatures for that class which fell at the upper and lower extremes of the signal range. The combined signatures, then, had a mean which was the average of the means of the extreme signatures and which included those means within ±1 standard deviation of the combined signature distribution.

Combined training signatures were calculated for corn (blight levels 0 - 3), soybeans, pasture, hay, trees and an object class designated sparse vegetation. The latter class included a wide range of sparseness from bare soil to a fairly high percentage (50%) vegetative cover. In addition to the combined signatures, the signatures from two corn fields were used to train the computer for corn blight levels 4 and 5.

It was decided that the same number of spectral bands (six) would be used in generating this recognition map as were being used in the routine processing of the Corn Blight Watch data. Six spectral bands were selected (2, 6, 8, 9, 10, and 12) which seemed to provide adequate separation between the eight object classes.

An analysis was carried out of the recognition results achieved in identifying corn fields or areas in corn fields containing specific levels of blight. The best that could be said for these results is that they were disappointing. Although most fields which were called out by the observers in the field as exhibiting high levels of blight were recognized as such, many other fields were improperly identified as containing high blight levels. The reasons for the misidentifications were many and varied, ranging from the existence of other plant deficiencies and diseases to one example of a field planted to popcorn which resembled fields containing high levels of blight. A general conclusion regarding this effort is that at this time in the growing season, enough natural variability exists in corn to prevent reliable detection of specific blight levels.

The results achieved in identifying the various object classes in the scene were very much better. The number of acres classified as belonging to each of seven categories are listed in Table II. For soybeans, the number of acres recognized was approximately 85% of the acres planted. The majority of non-soybean classifications in soybean fields were sparse vegetation. In almost all cases, these classifications were judged to be correct since they occurred in regions of the fields where the plants were stunted due to soil drainage problems. Therefore, on a point by point basis, the recognition accuracy for soybeans was actually much better than 85%. In fact, if the recognition results were interpreted on a per field basis, with all fields in which more than 50% of the points were classified as soybeans being considered 100% accurate, the accuracy of recognition exceeds 98%.

The point by point recognition accuracy for corn was significantly lower than that for soybeans. Approximately 60% of the acreage planted to corn was classified as being corn. In this case, too, the largest majority of non-corn classifications were sparse vegetation. Since the corn training signatures were extracted from uniform areas of relatively high ground cover, regions of lower cover would likely not be recognized as corn. If information concerning the likely yield of corn was required, the above approach would be reasonable. However, for a determination of the number of acres planted to corn, it is clear, at this time, that the signatures should have been extracted from areas containing a larger variety of crop conditions.

Of all the corn fields in the segment, 37 were poorly recognized. There was no apparent correlation between soil type, topography, or location of the fields in the segment but the planting density in at least 32 of these fields was lower than most of the corn fields in the segment. This is further indication that signatures should have been extracted from other areas to optimize corn recognition.

The interpretation of the corn recognition results on a per field basis improved the accuracy of a significant amount. Whereas only 60% of the corn acreage was recognized on a point by point basis, 80% was recognized on a per field basis. These results suggest that, in the future, automatic per field classification techniques should be developed.

SIGNATURE EXTENSION

One of the primary goals of the efforts being undertaken in remote sensing is the development of techniques which will enable large-area crop surveys without the need for expending a significant amount of manpower gathering ground information. If the amount of necessary ground information can be reduced, cost-effective remote crop survey systems will become a reality. To accomplish this goal, the effectiveness of spectral signatures must be extended in time and space.

The discussion which follows describes a relatively successful attempt at applying object class spectral signatures derived from one data set to another set of data gathered on a different day at a different location. In particular, data from Segment 203 of the Intensive Study Area were processed using signatures from Segment 212.

The conditions which prevailed during the two data collection flights are listed in Table III. The position of the sun in the sky, the soil types, the ground moisture, and the visibility were different for the two sets of data. Taken together, the existing conditions were different enough to produce large changes in the magnitude and spectral makeup of the scene radiance. It is changes of this sort that prevent the useful application of signatures over large areas.

The data from Segment 203 were prepared and preprocessed as described in the previous section. The preprocessing included the stabilization of the data by the sun sensor signal and the elimination of angle effects. Since both data sets included a similar distribution of objects having the same basic spectral properties, the differences in the angle-corrected signal levels were used to quantify differences in scene irradiance and atmospheric transmittance at the two locations. This information was then used to adjust the spectral signatures determined from Segment 212 data so that they could be applied to the Segment 203 data set.

As a preliminary test of this approach, the average signal level was determined for each spectral band in the angle-corrected data for both data sets. The ratios of these signals were then computed. It was felt that these ratios could be used to make the adjustment described above. To check whether this would indeed work, signatures were extracted from a limited number of fields in Segment 203. On comparing the means of these signatures to those extracted from similar object classes in Segment 212, it was found that the ratios of the mean signatures for all objects were essentially equal to the ratios of average signal levels in each channel. Based on this limited substantiation of the approach, a recognition map was generated for Segment 203 using the adjusted signatures from Segment 212.

The spectral bands which were used in producing this map were the same as were used in processing Segment 212 (2, 6, 8, 9, 10, and 12). Figure 9 shows the ranges of the mean signals of the various ground cover types in Segment 203 for each of these bands. With the exception of band 9, the choice of spectral bands for Segment 212 appears to be also suitable for Segment 203.

A detailed analysis was made of the recognition results in all the fields in a 2-mile portion of Segment 203. The recognition map for this area is shown in Figure 10. The field identifications for this area are included in Figure 11. Rather than presenting the detailed analysis here, we will indicate the general results achieved in generating this recognition map.

The accuracy of the recognition results for Segment 203 data using signatures from Segment 212 data is somewhat less than that achieved on Segment 212 data. For corn, the detection rate was 45% while the false alarm rate was 10%. These figures for soybeans were a detection rate of 50% and a false alarm rate of 15%.

The hayfields were most poorly recognized. Only 4.6% of the area categorized as hay was correctly recognized as hay. Approximately 77% of the hay area was recognized as pasture or sparse vegetation. The recognition of trees was also not very accurate. These results were not too surprising since the tree recognition accuracy on Segment 212 was not very good. Once again, most of the false alarms were sparse vegetation.

It seems that the signature for sparse vegetation was too broad and that some modification of the signatures would have produced significant improvements in recognition accuracy for both data sets.

Considering the fact that the signals generated when viewing the two scenes were quite different, it is felt that the application of signatures from Segment 212 to recognize objects in Segment 203 was a reasonably successful effort. We believe that these results give reason for optimism regarding the feasibility of operational remote sensing crop survey systems.

TABLE I. -CLUSTER ANALYSIS RESULTS FOR 43M SEGMENT 212

1

Field	Blight Level	Location in Segment	Row Dir
EE15	2	M	E-W
EE9	2	W	E-W
EE8	3	W	E-W

2

Field	Blight Level	Location in Segment	Row Dir
RR4	0	M	E-W
RR3	1	M	E-W
UU7	1	M	E-W
UU6	1	W	E-W

3

Field	Blight Level	Location in Segment	Row Dir
NN2	0	M	E-W
UU8	1	M	E-W
ZZ7	2	W	E-W
*UU1	3	W	N-S
N1	3	M	N-S
			N-S

4

Field	Blight Level	Location in Segment	Row Dir
RR5	0	W	N-S
UU4	1	W	N-S
EE11	3	W	N-S
RR2	3	M	N-S
UU10	3	M	E-W
*UU1	3	W	N-S

5

Field	Blight Level	Location in Segment	Row Dir
SSS2	0	E	E-W
II1	1	E	E-W

6

Field	Blight Level	Location in Segment	Row Dir
SSS1	0	E	E-W

*Half of this signature clustered in Group 3 and half in Group 4.

TABLE II. -CROP CLASSIFICATION TOTALS FOR 43M SEGMENT 212

CLASS	ACRES
CORN	979
SOYBEANS	898
PASTURE	721
HAY	291
SPARSE VEGETATION	2541
TREES	534
NOT CLASSIFIED	220

TABLE III. -SUMMARY OF CONDITIONS PREVAILING IN TWO STUDY AREAS USED FOR SIGNATURE EXTENSION ANALYSIS

	Segment 203	Segment 212
Latitude	41.5°N	40.0°N
Longitude	87°W	87°W
Date of Flight	8/13/71	8/17/71
Time of Flight	1053 est	1120 est
Solar Azimuth	139°	150°
Solar Elevation	57°	60°
Drainage Conditions	Poor	Good
Days Since Last Precipitation	3	7
Amount of Last Precipitation	1.60 in.	.49 in.
Visibility	9 miles	13 miles
Cloud Cover	< 10%	< 10%
Ground Temperature	74°F	76°F
Relative Humidity	64%	48%
Dew Point	61°F	55°F
Flight Direction	N-S	N-S

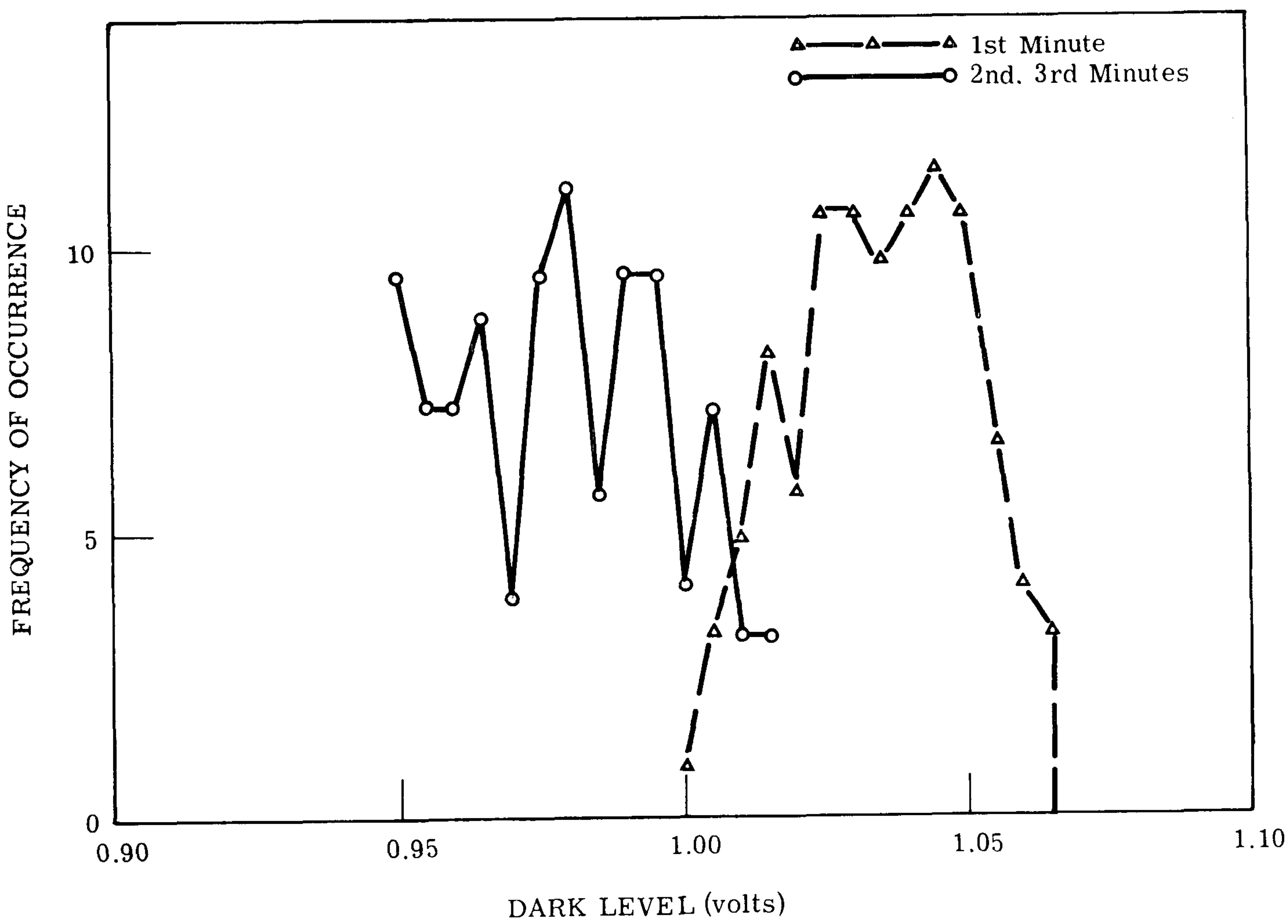
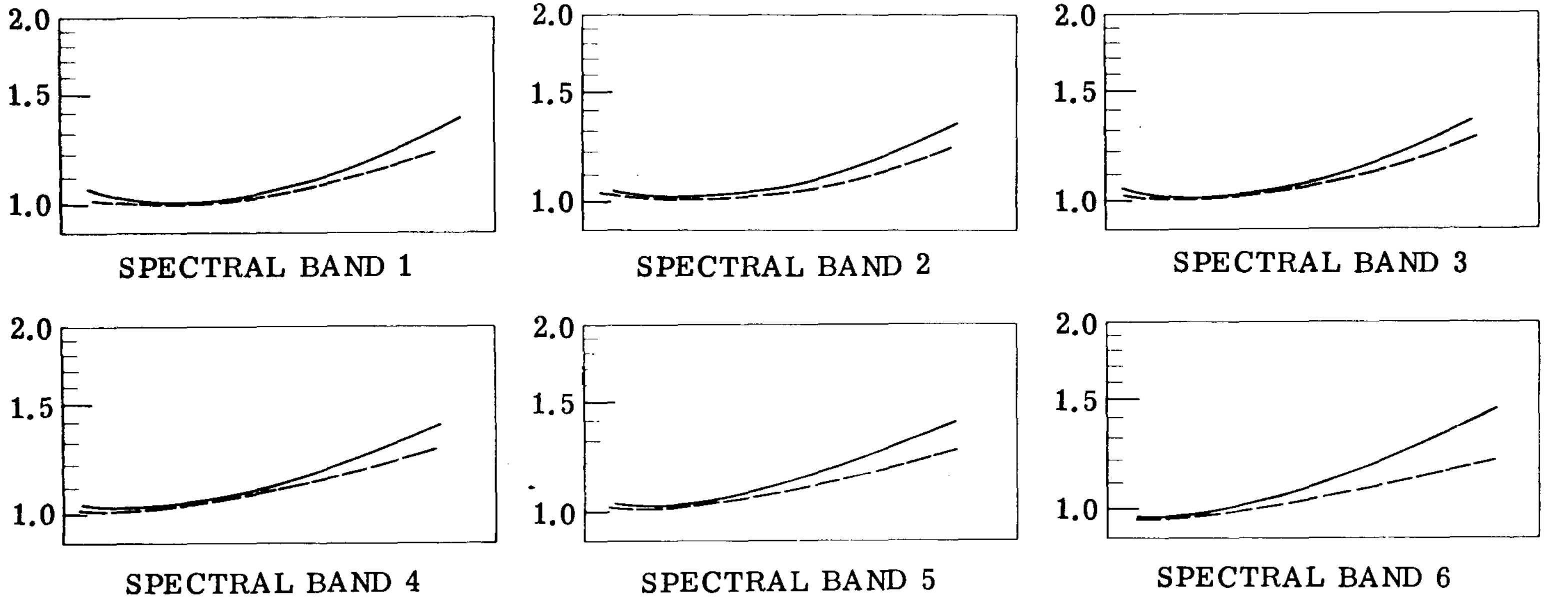


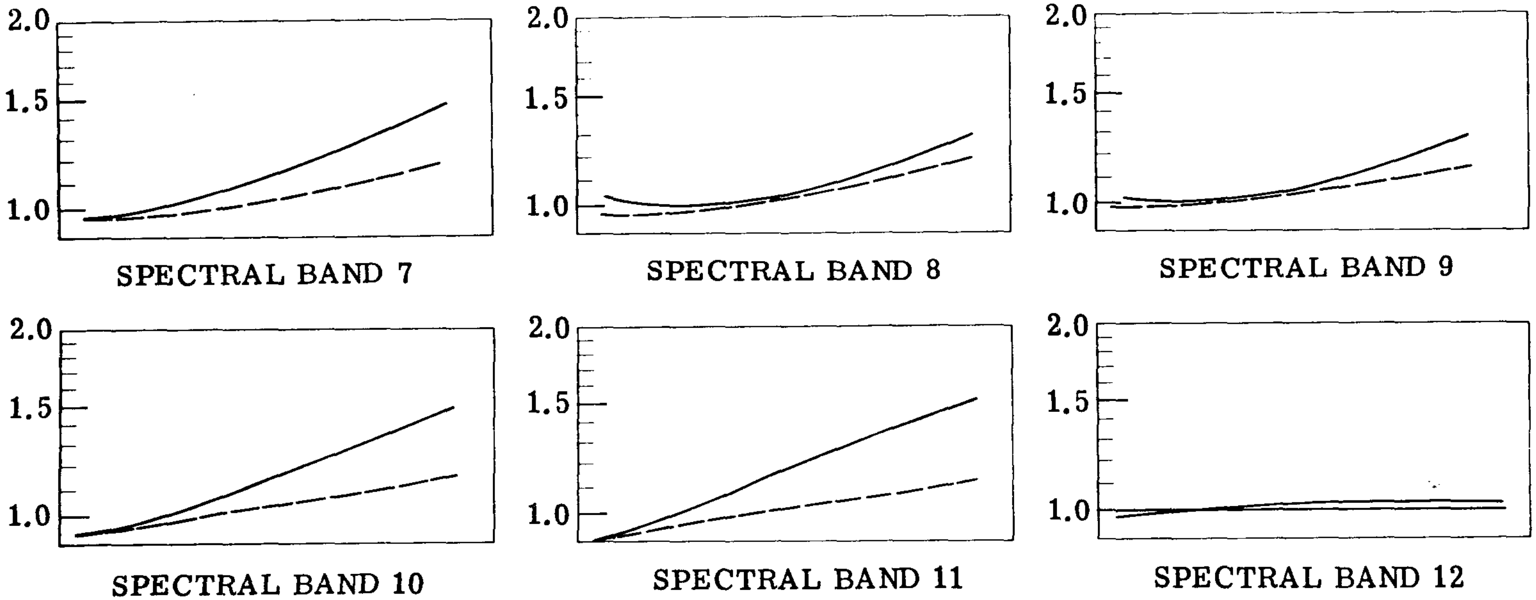
FIGURE 1. DARK LEVEL VARIATIONS IN UNCLAMPED DATA (ANALOG) FOR SEGMENT 203, MISSION 43



Key:

Solid Line: Segment 203
Broken Line: Segment 212

FIGURE 2(a). AVERAGE ANGULAR VARIATION OF SIGNAL FOR MISSION 43



Key:

Solid Line: Segment 203
Broken Line: Segment 212

FIGURE 2(b). AVERAGE ANGULAR VARIATION OF SIGNAL FOR MISSION 43

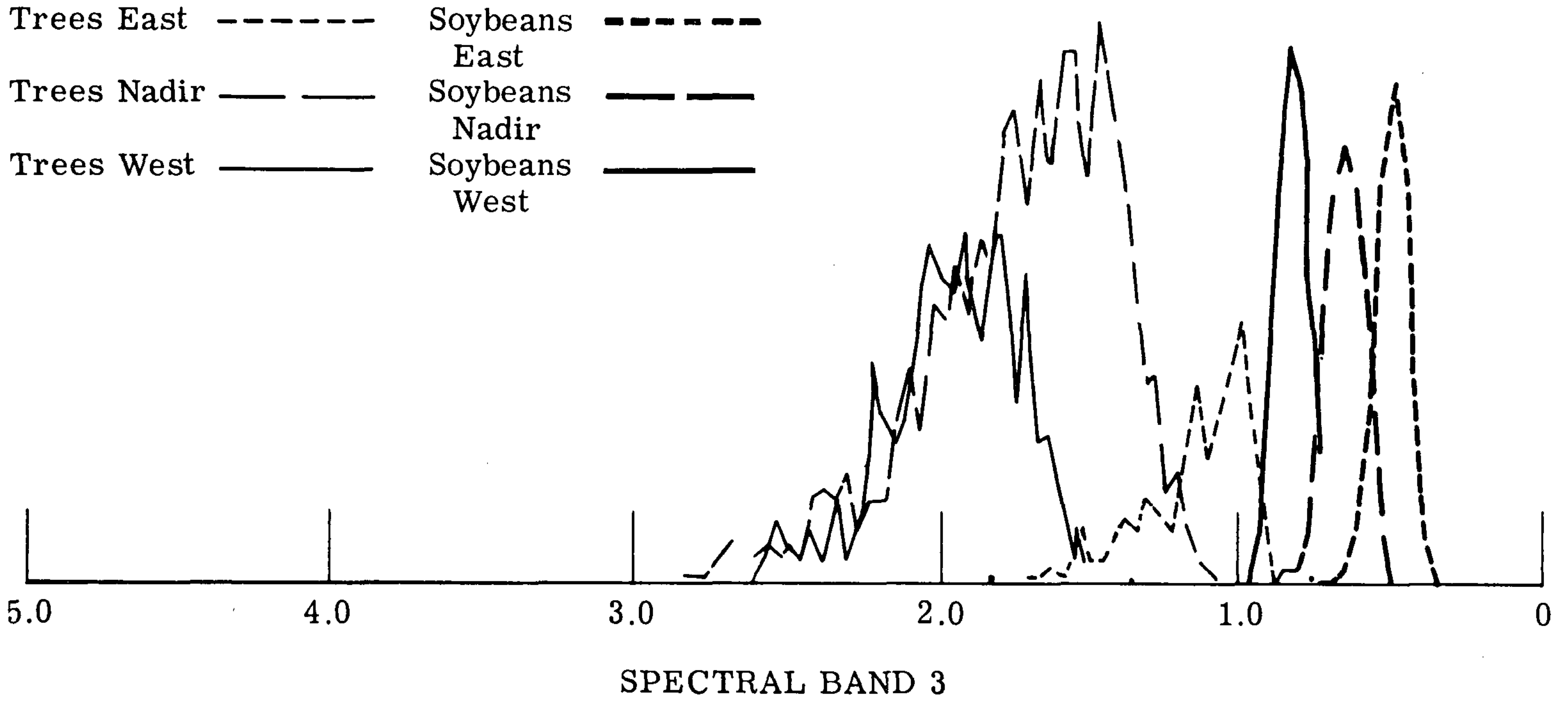


FIGURE 3a. HISTOGRAMS OF SIGNALS FROM SOYBEANS AND TREES AS A FUNCTION OF LOCATION IN SCENE

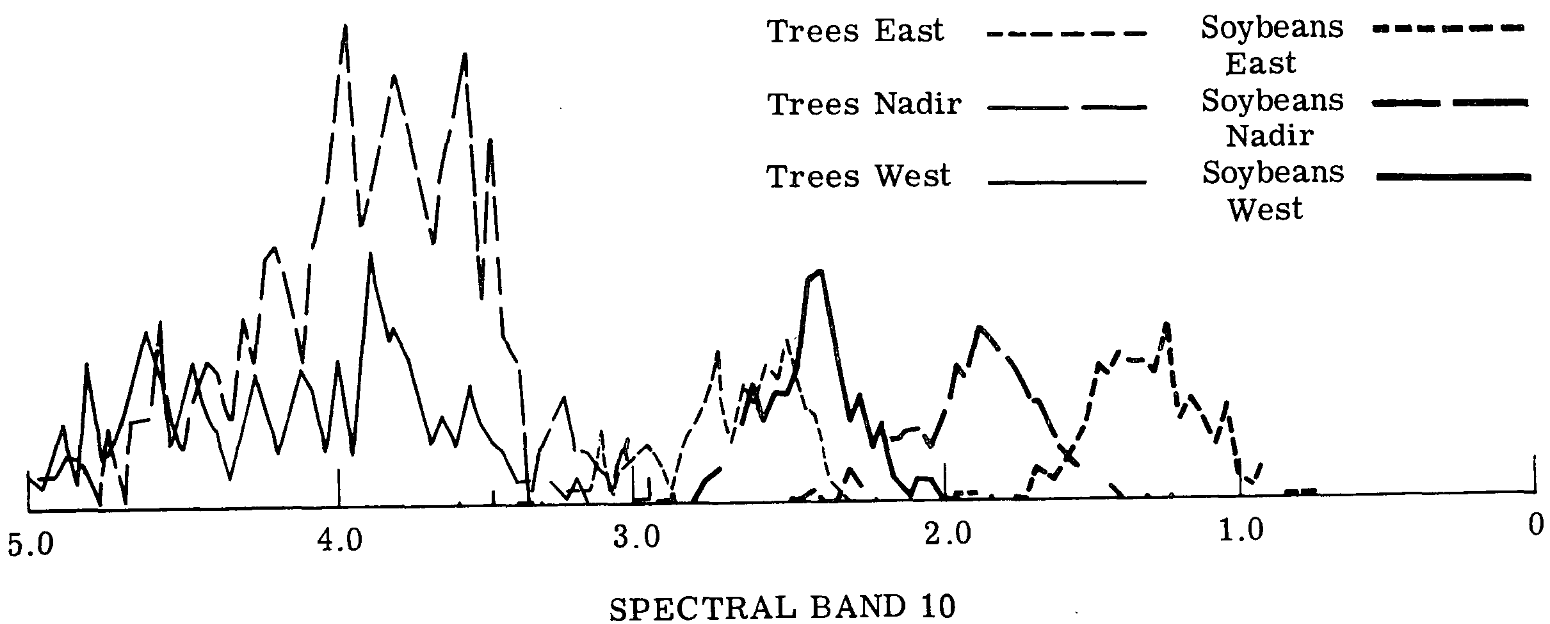


FIGURE 3(b). HISTOGRAMS OF SIGNALS FROM SOYBEANS AND TREES AS A FUNCTION OF LOCATION IN SCENE

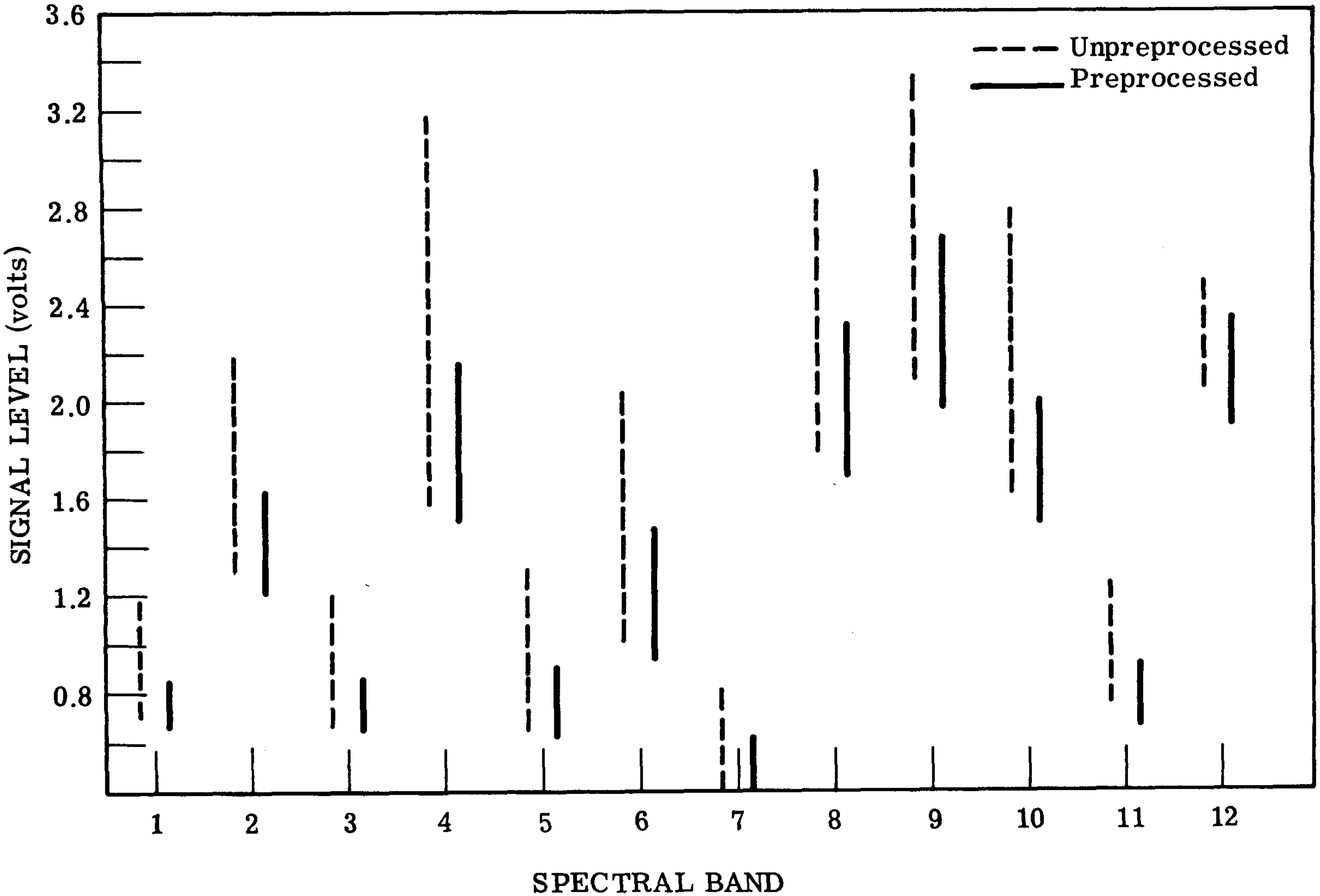


FIGURE 4. RANGE OF SIGNAL MEANS FOR CORN

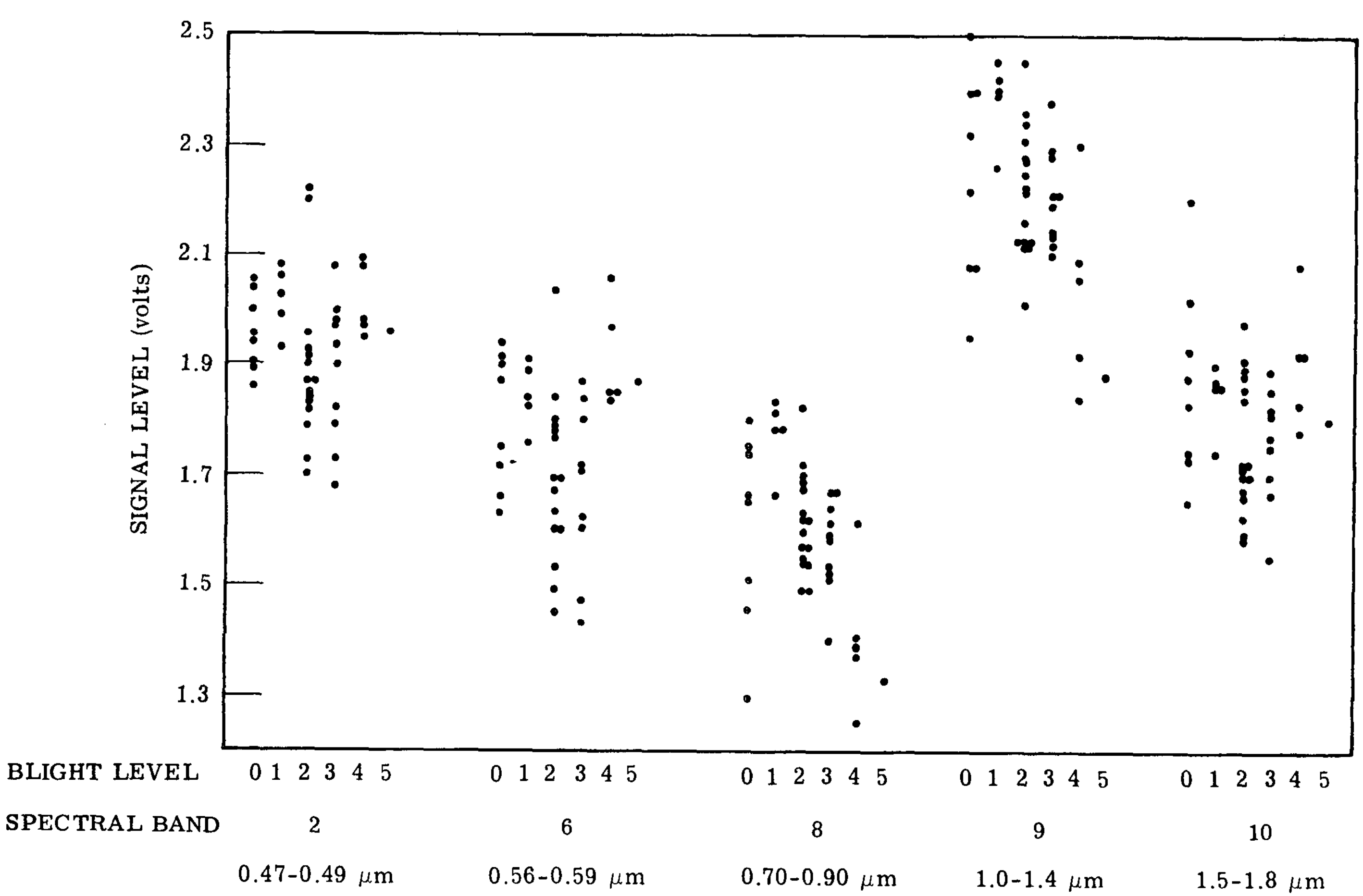


FIGURE 5. VARIATION IN MEAN SIGNALS OF CORN FIELDS FOR SELECTED BANDS BY BLIGHT LEVEL

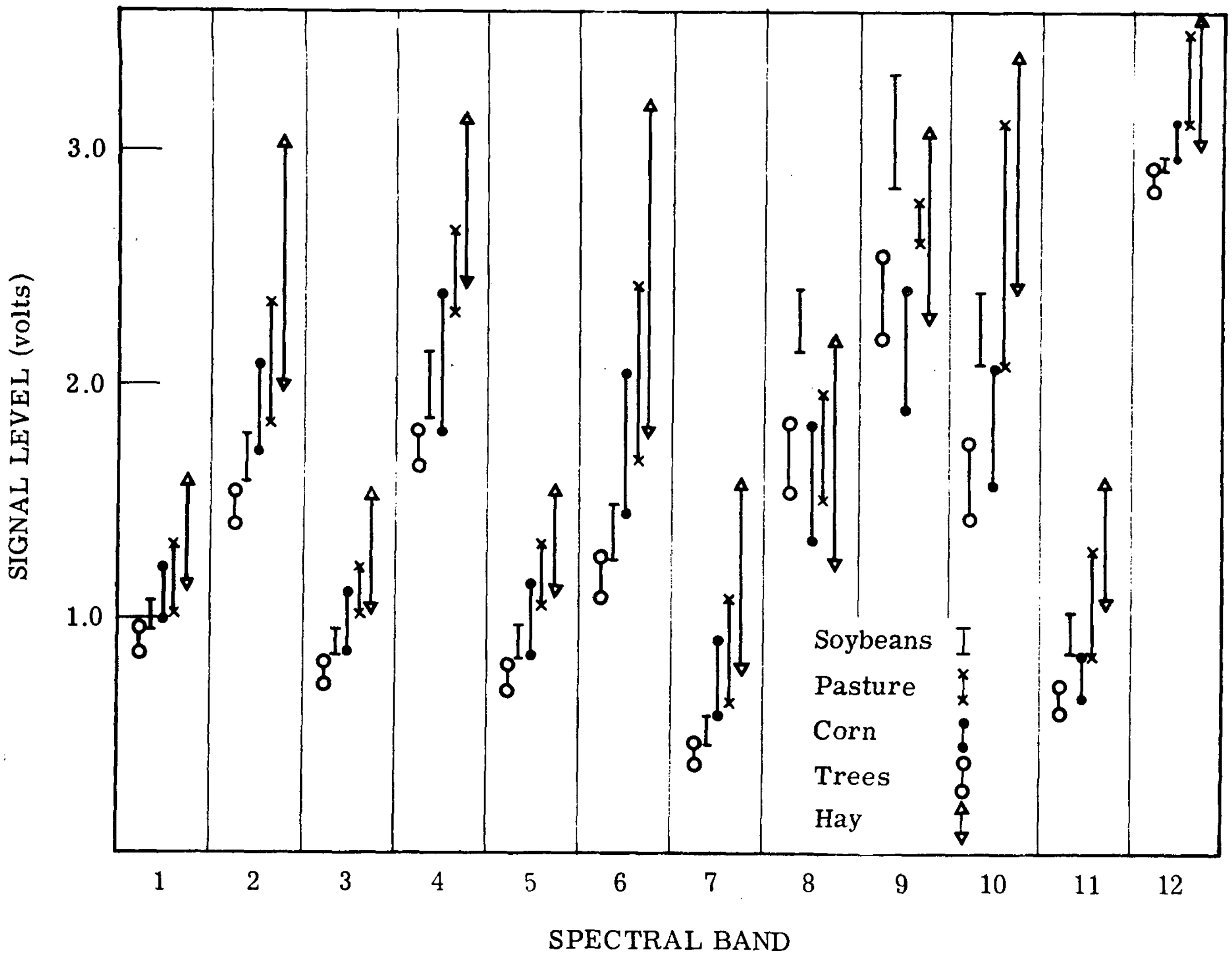


FIGURE 6. SIGNAL RANGE PER CHANNEL FOR 5 OBJECT CLASSES FOR SEGMENT 212, MISSION 43

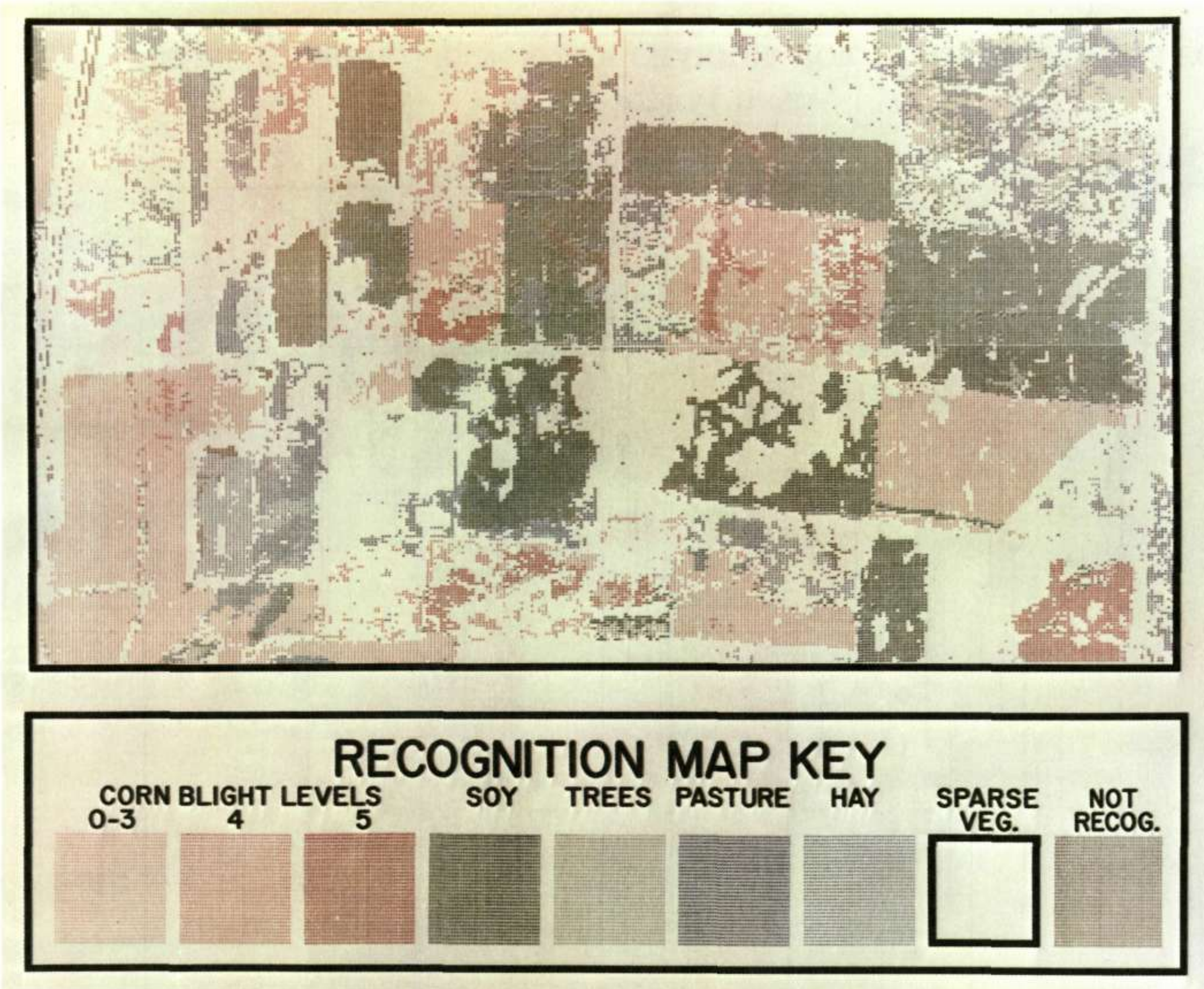
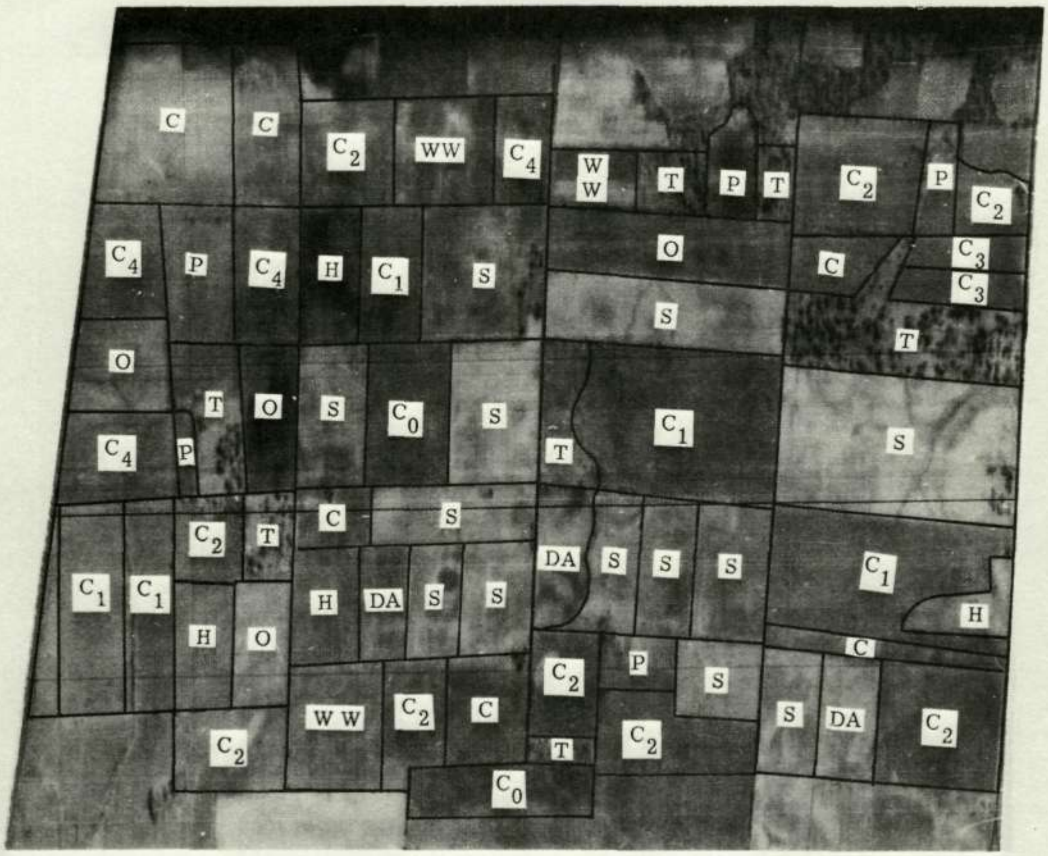


FIGURE 7. DIGITAL RECOGNITION MAP OF A PORTION OF 1971 CORN BLIGHT WATCH SEGMENT 212 (MISSION 43)



Key: C_n = Corn with n the Designated Blight Level
S = Soybeans
WW = Winter Wheat
DA = Diverted Acres
P = Pasture
T = Trees
H = Hay
O = Oats

FIGURE 8. SCANNER IMAGERY OF SPECTRAL BAND 1.0-1.4 μm FOR SEGMENT 212, MISSION 43. FIELD BOUNDARIES AND COVERS ARE INDICATED.

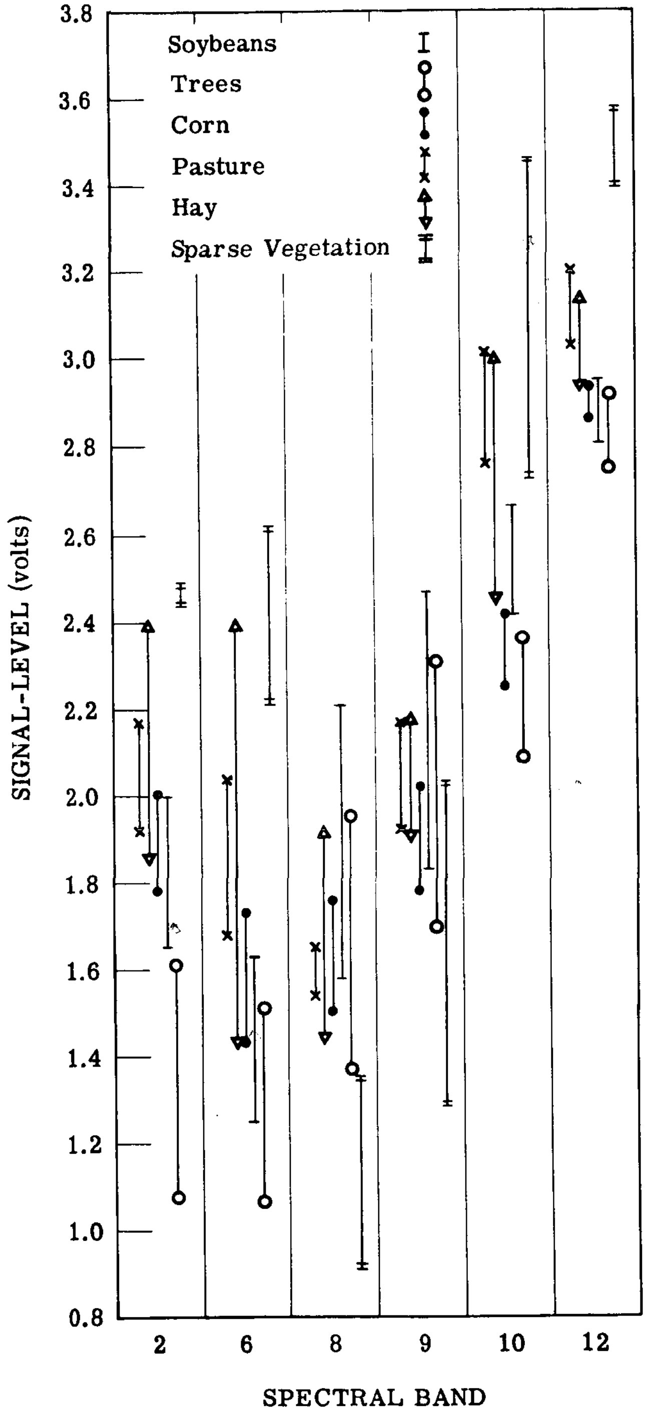


FIGURE 9. RANGE OF MEAN SIGNALS OF GROUND COVER TYPES FOR SELECTED CHANNELS FOR SEGMENT 203, MISSION 43

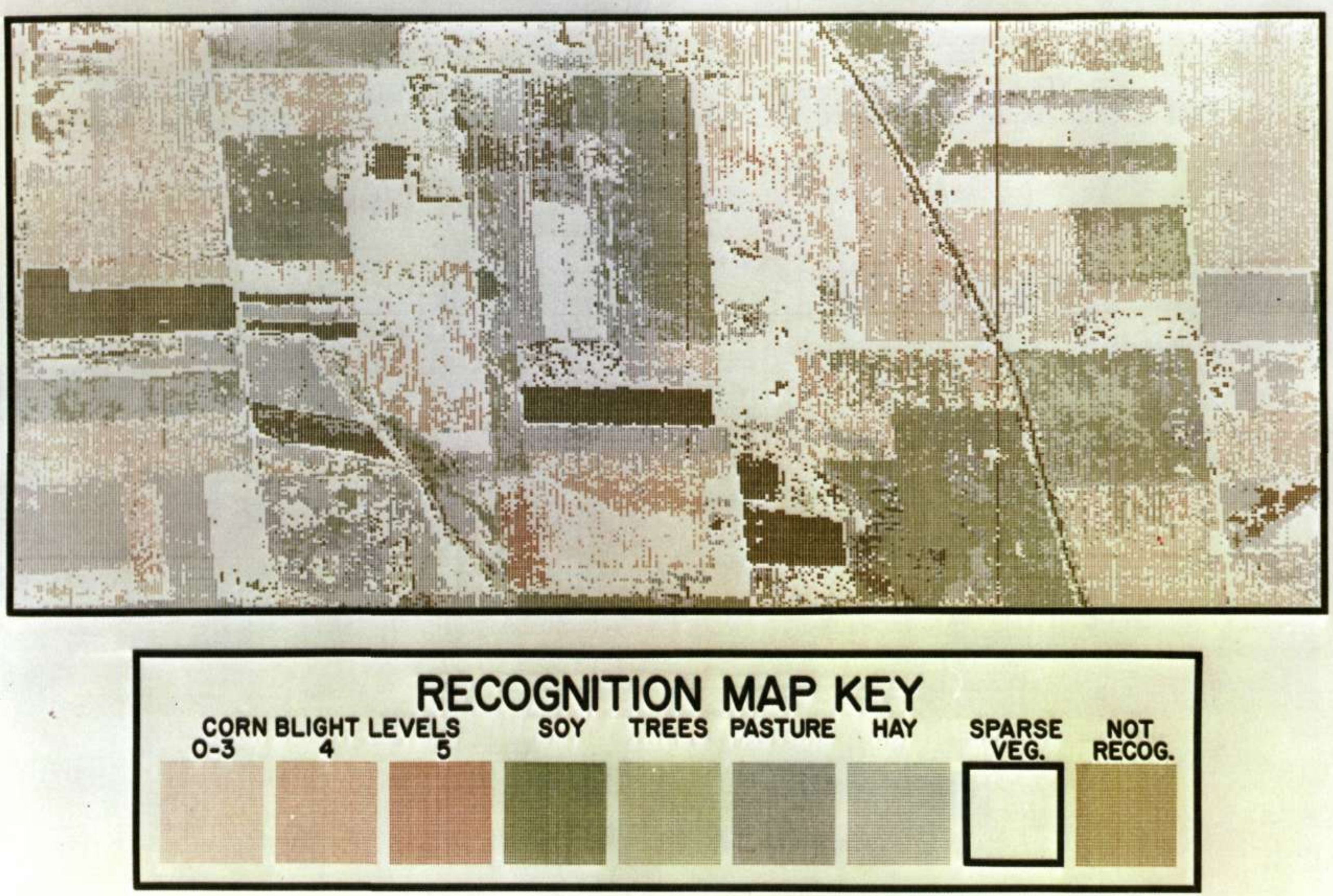
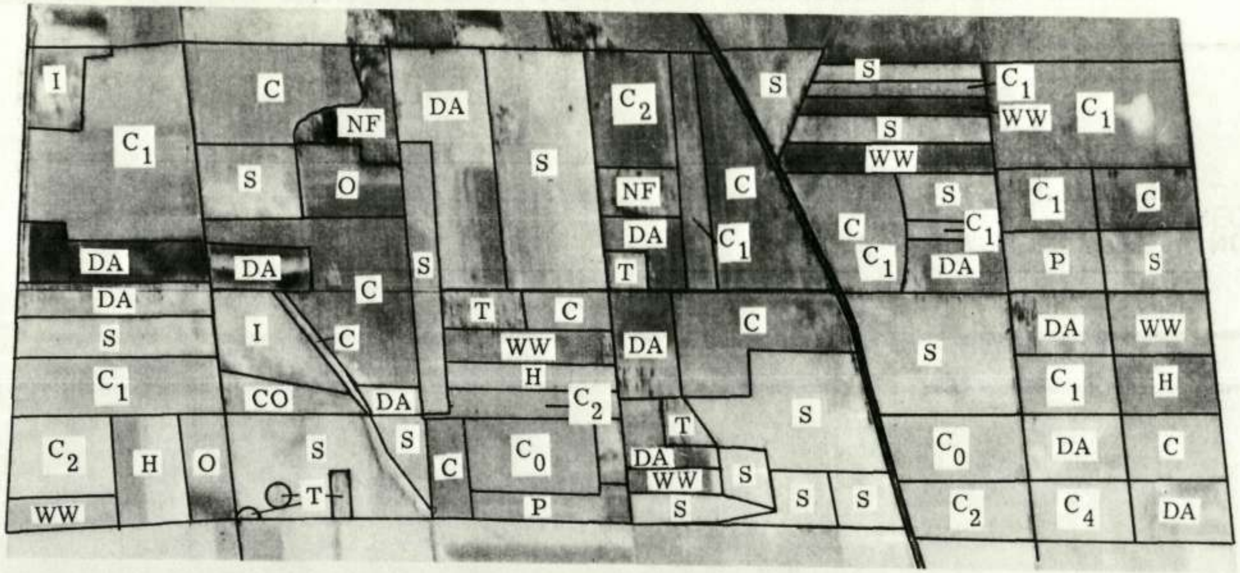


FIGURE 10. DIGITAL RECOGNITION MAP OF A PORTION OF 1971 CORN BLIGHT WATCH SEGMENT 203 (MISSION 43)



Key: C_n = Corn with n the Designated Blight Level
S = Soybeans
WW = Winter Wheat
DA = Diverted Acres
P = Pasture
T = Trees
H = Hay
I = Idle
O = Oats
NF = Non Farm

FIGURE 11. SCANNER IMAGERY OF SPECTRAL BAND 1.0-1.4 μm FOR SEGMENT 203, MISSION 43. FIELD BOUNDARIES AND COVERS ARE INDICATED.