LARS Information Note 091968
Purdue University

LARSYSAA, A Processing System For Airborne Earth Resources Data

by D. A. Landgrebe and Staff

The system used by LARS for research in earth resources systems is embodied in a set of computer programs known as LARSYS (See Figure 1). Figure 2 shows diagramatically how a portion of this system is used by a researcher applying modern pattern recognition techniques to airborne scanner data. The purpose of this Information Note is to give examples of the types of output which can be produced for this purpose.

The chief purpose of LARSYSAH is to produce aircraft Data Storage Tapes, alphanumeric pictorial printouts and, in the future, digital display data images. LARSYSAH output is still essentially data in unreduced form. Once the Data Storage Tapes and pictorial printouts have been generated, LARSYSAA may then be used by researchers to reduce the data to useful information. As illustrated in Figure 3, LARSYSAA contains four processors, each controlled by its own supervisor. Illustrations of LARSYSAH pictorial printout output followed by examples from the LARSYSAA processors are given below. These examples were all generated using data from the University of Michigan 12-channel scanner. Correspondence between the channel numbers and spectral bands for this particular scanner are as follows:

Channel Number	Spectral Band (Microns)
1	.40- .44
2	.44- .46
3	.46- .48
4	.48- .50
5	.50- .52
6	.52- .55
7	.55- .58
8	.58- .62
9	.62- .66
10	.66- .72
11	.72- .80
12	.80-1.00

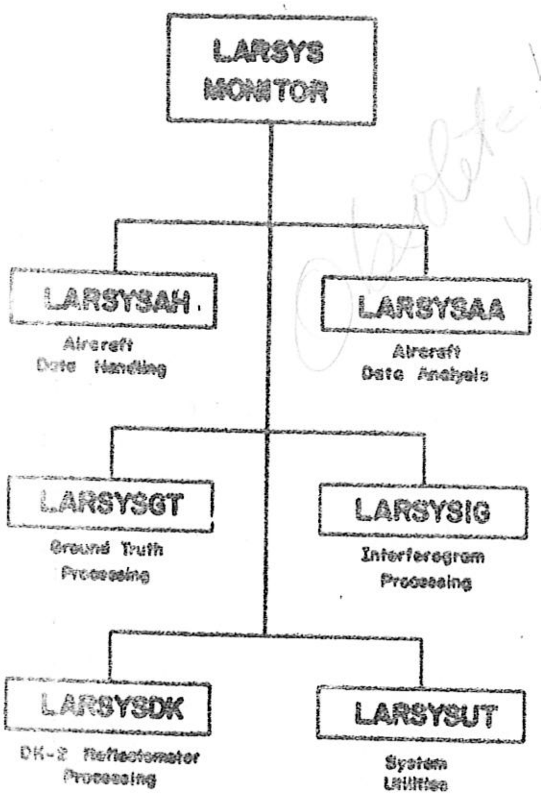


Figure 1. Diagram of LARSYS

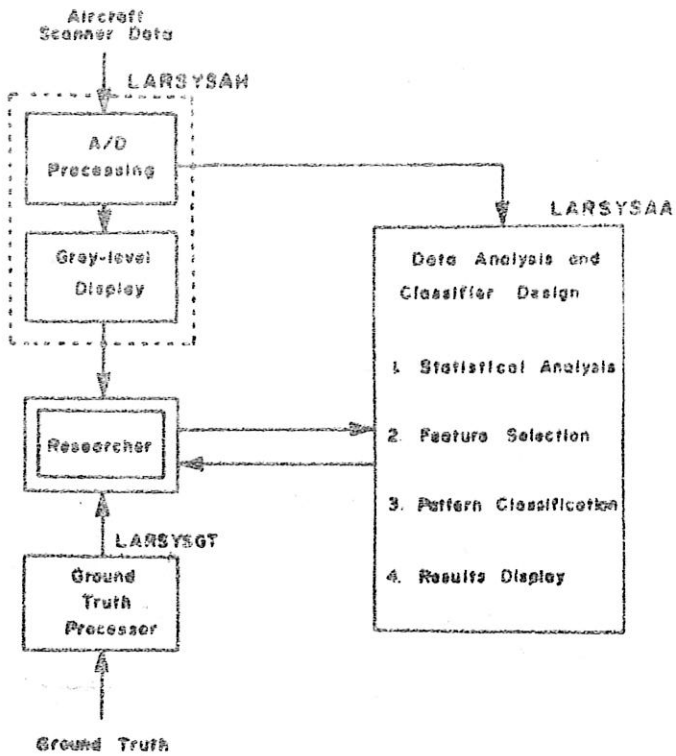


Figure 2. Data Flow in LARSYS for Aircraft Data Analysis

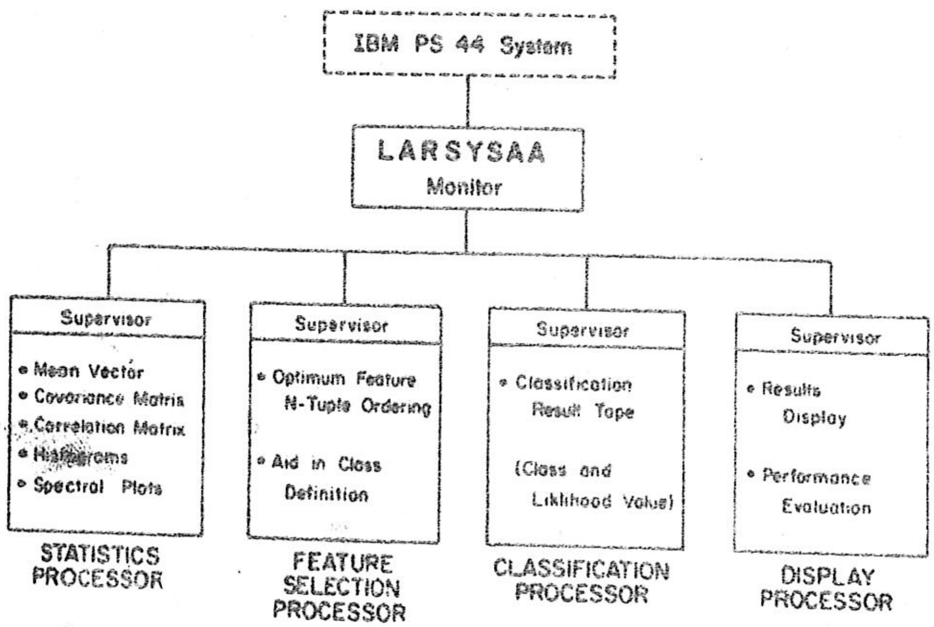


Figure 3. Orginization of LARSYSAA

Data Storage Tape

The Data Storage Tape is a digital magnetic tape produced by LARSYSAH from the aircraft analog tape. It contains the data for each resolution element stored in a packed format and has a specific address for each point in the form of a scan line number and sample number. Certain other information (run number, date, etc.) necessary for a machine controlled storage and retrieval system and data for calibration purposes derived from the aircraft analog tape is also stored in a convenient format.

Alphanumeric Pictorial Printout

Two channels of a portion of run number 26600061 are shown in pictorial printout form on the following two pages. Parameters which may be varied in this type of presentation include the spatial resolution, the radiance level-to-symbol correspondence and the symbols used.

Spatial Resolution. This particular run was digitized such that on the average there is neither underlap nor overlap of adjacent samples on the Data Storage Tape. In this case, based on the aircraft altitude and the scanner resolution, this required that every seventh scan line on the analog tape be digitized and 220 samples be made in each scan line. It is shown on the two printouts given that every other sample point of every other line on the Data Storage Tape has been printed out. Experience has shown that this choice of spatial parameters is convenient and very adequate for most purposes. Printouts of greater resolution are easily obtained, however, and an illustration of one in photographically reduced form will be given presently.

It may be noted parenthetically here that when the data is to be analyzed in image form as opposed to analyzing the spectrum of each point in quantitative form, overlap of samples is usually desirable. Underlap on the other hand may be desirable to reduce the data load when a detailed, high resolution study is not required.

Radiance Level-to-Symbol Correspondence. The analog data is to be quantized to 8-bit accuracy. Each resolution element of each

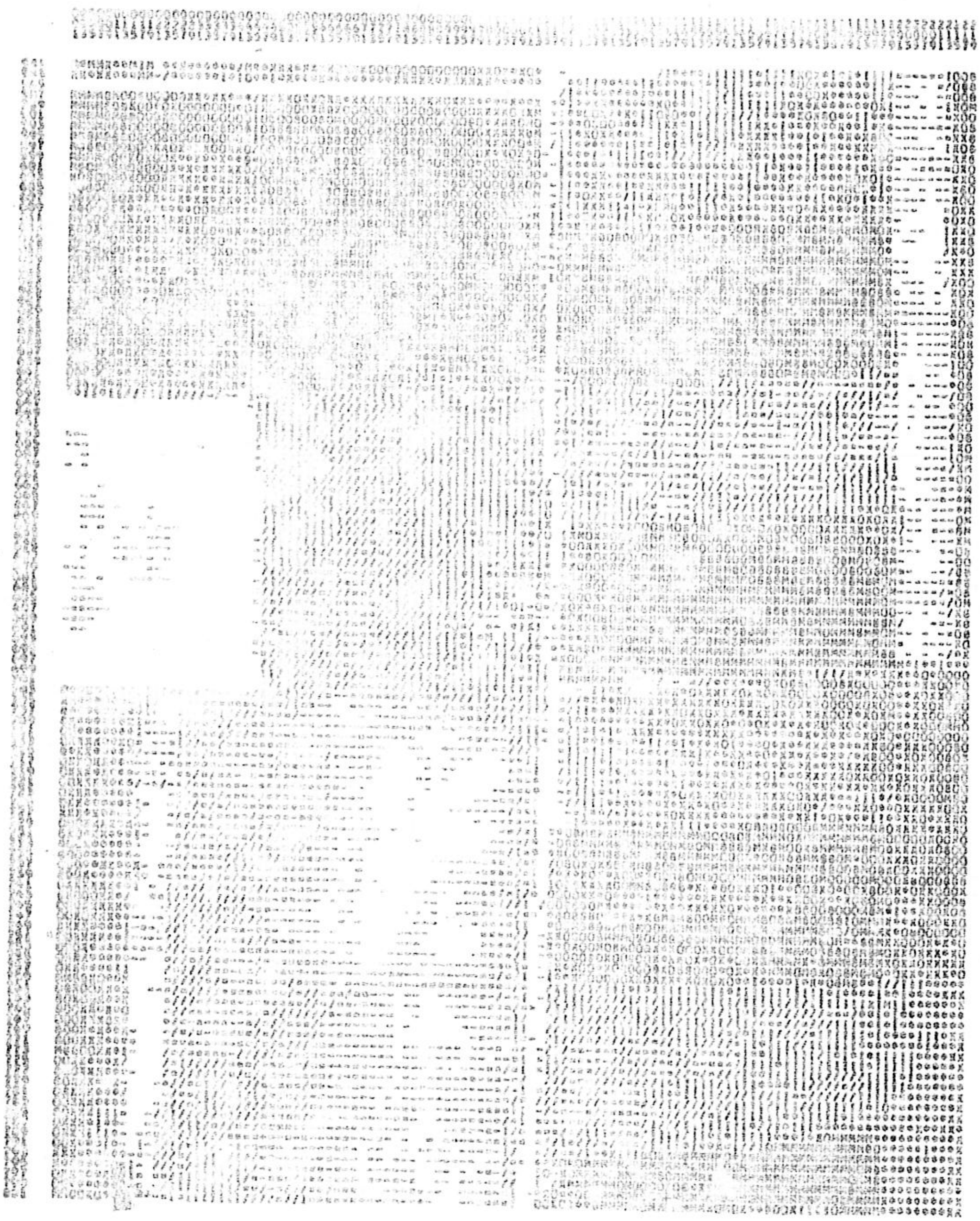
LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

RUN NUMBER ... 26600061
FLIGHT LINE .. C1
TAPE NUMBER .. 102
CHANNEL NUMBER 9

DATE .... [illegible]
TIME 1329
ALTITUDE ...... [illegible] FEET
SPECTRAL RANGE 0.62 TO 0.[illegible] MICRONS

THE CHARACTER SET USED FOR DISPLAY IS

FROM   0 TO 1[illegible] DISPLAYED AS
[illegible]
FROM 247 TO 255 DISPLAYED AS [illegible]
ILLEGAL DATA VALUE DISPLAYED AS [illegible]



Jim's printout?

spectral band will have one of 256 possible values, according to the radiance of that element in that band. Experience has shown that from 10 to 16 symbols should be used to simulate the gray scale tones in the printout. Therefore, each of the 256 levels must be assigned to one of the 16 symbols in some fashion. One speaks of assigning each symbol to its own bin, or group of gray levels.

One way to do this might be to use equal sized bins, i.e., assign levels 0 to 15 to the first symbol, 16 to 31 to the second, and so on. However, this does not usually give satisfactory results since the data is rarely if ever uniformly distributed over the entire dynamic range of radiances.

One of the unique advantages of the digital approach to image data handling is the simplicity with which the radiance level to gray scale may be varied. The photographic equivalent to this is the film density versus radiance curve and by using the above procedure, it is seen that a curve of arbitrary shape, be it linear, nonlinear or even multivalued, can be achieved.

LARSYSAH provides for three means of specifying the bin edges for each symbol: (1) use of a preassigned set, (2) assignment of an arbitrary (ad hoc) set for the current job, (3) computation of a set which will result in equal activity for each of the 16 symbols. The first of these three is the most economical insofar as both the researcher's and computer's time are concerned. Based on experience, a set of bin edges has been picked which gives satisfactory but suboptimum results for most data.

The second method may be used when the researcher knows of a good set from previous computation or if he has some more specific use of the printout in mind. The third method takes longer computationally but automatically provides the maximum contrast over the whole range of radiances. In this method the data is first histogrammed, then bin edges are set so that all bins have equal area under their portion of the histogram. It is also possible to accumulate the histogram from one area or a group of areas and use the resulting bin edges to printout another area. For the two attached printouts, the bin edges were determined from histograms of the first 950 lines of run 26600061.

The radiance level-to-symbol correspondence for a specific printout is always given in tabular form as part of the heading as shown.

Symbols Used. The set of symbols used in the accompanying printouts has been picked as a standard set based on experience and study. However, any other set may be specified as needed for specific purposes. For example, if a contour map is desired which shows areas having a certain temperature/emissivity, one could obtain it by appropriately assigned bin edges, option (2) above, and choosing blanks for all bins except the desired one.

Summary. LARSYS can produce alphanumeric pictorial printouts of variable spatial resolution using arbitrarily selected alphanumeric symbols and arbitrarily changing the range of contrasts presented.

Photographically Reduced Alphanumeric Pictorial Printouts

The illustration on the following page shows two photographically reduced pictorial printouts of run 26600061, lines 461 to 949. The area displayed includes that shown in the previous printouts. The left printout again has every other sample of every other line presented, while in the right printout every sample of every line is given. In both examples there are 10 active gray level bins (symbols). A panchromatic air photograph is given for comparison.

Output in this format is about at the crossover between viewing the data as a large number of quantitative spectra and viewing it as an image. For example, one may view the printout given as a low resolution image and also (perhaps with the aid of a hand-held magnifier) see in a photographic (rather than this mimeograph) presentation that there is a point (in Channel 9) at line number 609 and column 213 near the right edge of the wheat field which is displayed as * and therefore, from the tabulation in the heading has a relative radiance of 188-190 on a scale of 0 to 255.

Summary. Alphanumeric pictorial printouts may be photographically reduced to increase their similarity to images and make them easier to handle physically. Printouts of both greater and lesser spatial resolution than the examples given can be produced in this fashion.

Photographs of Digital Image Display Presentations

It is not possible to show an example of this type of data format since procurement of the digital display has not yet been possible. However, when it becomes available it is expected to provide images of at least studio TV quality. While the original primary purpose for the digital display was to speed the data editing function, it will in addition provide a considerable number of capabilities never before possible in man/data communication since it mates good quality image production with the flexibility provided by digital control.

Note in particular that the comments above about bin selection and gray scale control apply equally here. Additionally, images created from linear combinations of bands, reconstituted color and false color images are expected to be possible quickly.

Summary. Photographs of Digital Image Display presentations can be provided as soon as the display becomes available. Color reconstitution and other types of displays should also be available shortly thereafter.

Edited Data

The chief reason for devising the pictorial printout and digital display was to enable the editing of data. That is, one must be able to extract from the Data Storage Tape the data belonging to a specific area on the ground. Refer to the previously given pictorial printouts and air photo. Note the corn field at the left edge just above the wheat field. By referring to the printout, the addresses on the Data Storage Tape of the data field (i.e., set of contiguous resolution elements rectangular in shape) in the center of this agricultural field can be determined as lines 603 to 625 and columns 13 to 33. These addresses are valid for all channels of data gathered by that (12 channel) detector set.

Once located in this fashion and extracted from the Data Storage Tape, this data can be stored in the computer for further processing, or it can be printed, punched out, or written on magnetic tape for external use by the researcher. The following page shows a printout of the above data.

Summary. LARSYS provides data edited from Data Storage Tapes in line printer, punched card, or digital magnetic tape form.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

LARSYSAA ILLUSTRATION

CORN 1

RUN NO. 26600061, FIELD 12-9

NO. OF SAMPLES- 132, FROM LINES 603 TO 625 (EVERY 2 LINES), SAMPLES 13 TO 33 (EVERY 2 SAMPLES)

EDITED FROM CHANNEL 9

LINE	SAMPLE VALUES										
	13	15	17	19	21	23	25	27	29	31	33
603	197	194	197	196	196	195	194	194	195	195	192
605	195	193	194	198	196	195	194	190	194	192	195
607	196	194	191	193	194	191	192	193	188	194	194
609	195	191	192	193	191	192	191	191	191	196	197
611	196	195	193	191	195	194	195	194	197	193	197
613	194	195	193	194	195	193	194	193	194	195	198
615	195	194	191	194	194	193	191	194	192	195	197
617	198	187	190	190	195	195	196	194	193	196	197
619	194	184	190	189	189	192	193	193	193	192	193
621	192	187	187	184	193	193	192	191	192	195	194
623	196	192	191	191	190	194	194	196	193	192	195
625	196	197	191	188	189	197	195	194	196	196	196

Histograms

One of the types of output which the LARSYSAA statistics processor can provide is the histogram of a field or a group of fields. The histograms for the data from the corn field indicated above in three arbitrarily selected bands are given on the next page. The abscissa is relative radiance (brightness), increasing to the right. (The numerical abscissa values decrease for increasing radiance because the scanner output signal is inverted.) The ordinate gives the number of resolution elements with a given relative radiance.

The page following these three histograms gives the histograms for another corn field from another flight line several miles distant. It is seen that by overlaying the first set over the second, perhaps on a light table, a quick but useful comparison between data from the two fields can be obtained. In this case they are similar in Channel 1 but different in Channels 9 and 12.

The third page of histograms shows the data from these same two fields put together as a class (i.e., a group of fields not necessarily contiguous nor even from the same flight line) and presented in histogram form. The double lobes in Channels 9 and 12 predicted from the two previous histograms are apparent, for example.

Summary. LARSYS can provide histograms of data from individual data fields and from classes (groups of data fields) in any arbitrary subsets of channels. Note that a data field is any rectangular region in the data. The whole flight line could be defined to be a data field if desired.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

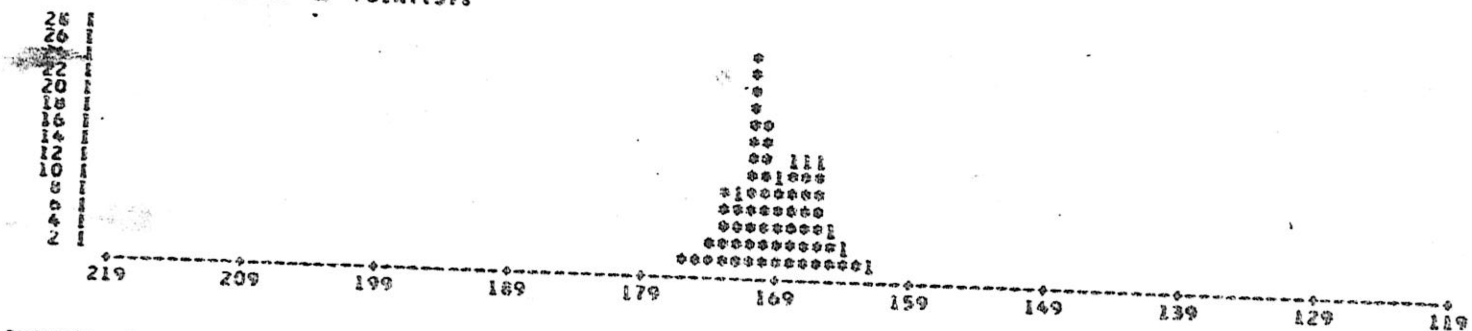
CORN 1

RUN NO. 26600061, FIELD 12-9
NO OF SAMPLES = 132, FROM LINES 603 TO 625 (EVERY 2 LINE(S)), SAMPLES 13 TO 33 (EVERY 2 SAMPLE(S))

HISTOGRAM FOR ... 12-9

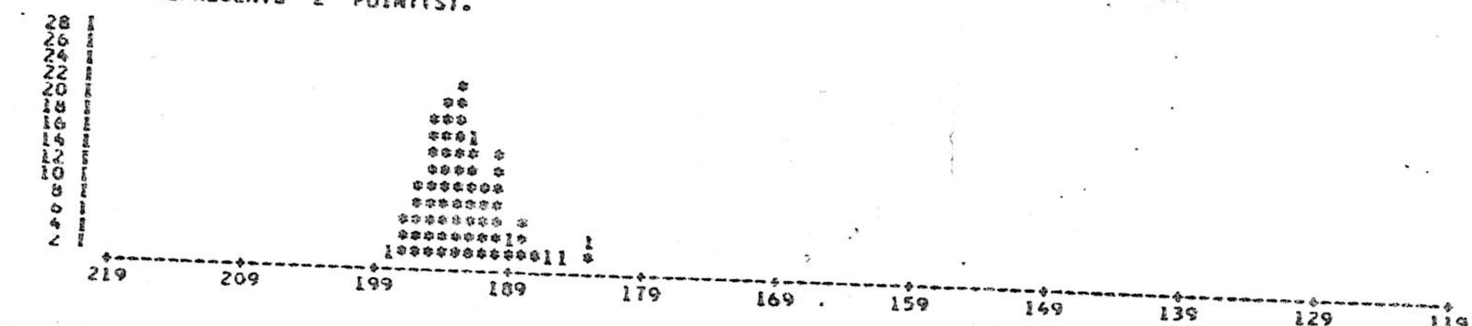
CHANNEL 1 0.40 - 0.44 MICRONS

EACH * REPRESENTS 2 POINT(S).



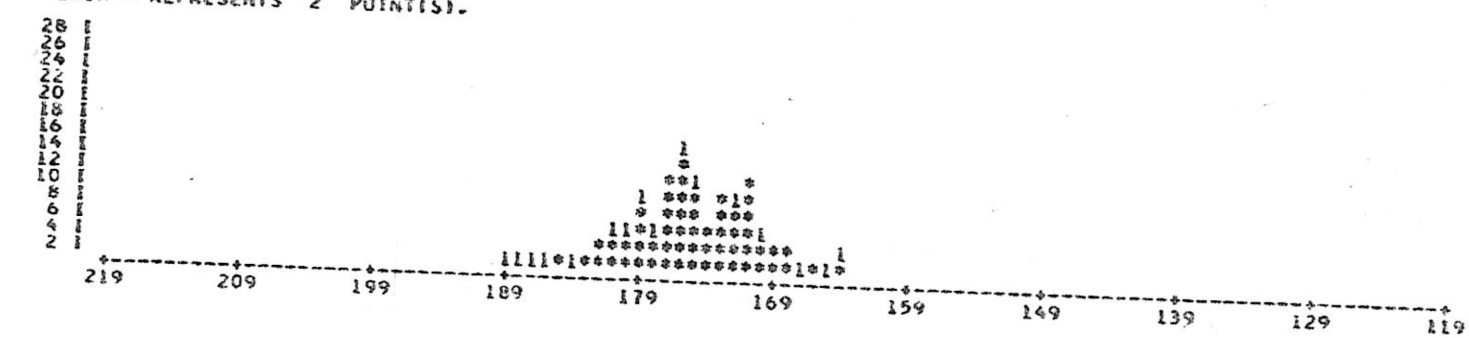
CHANNEL 9 0.62 - 0.66 MICRONS

EACH * REPRESENTS 2 POINT(S).



CHANNEL 12 0.80 - 1.00 MICRONS

EACH * REPRESENTS 2 POINT(S).



LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

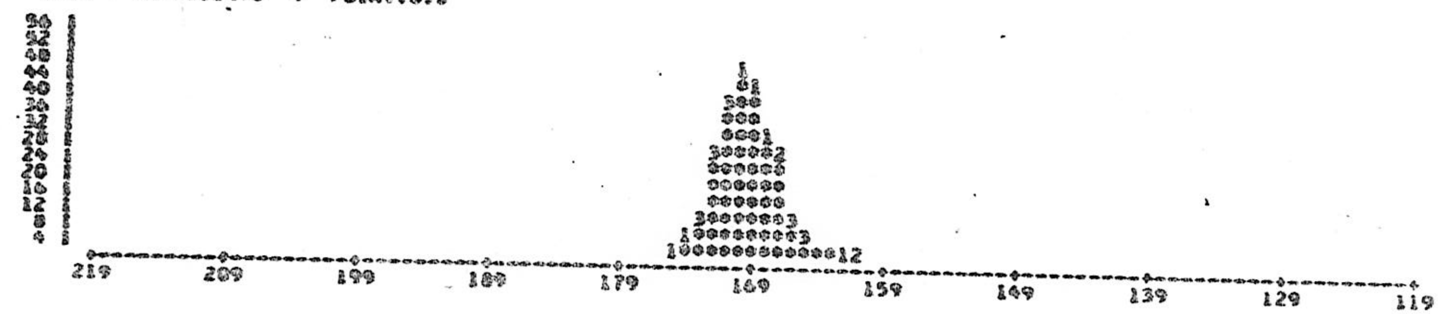
CORN 1

RUN NO. 26600001, FIELD 8-20
NO OF SAMPLES = 253, FROM LINES 669 TO 713 (EVERY 2 LINE(S)), SAMPLES 171 TO 191 (EVERY 2 SAMPLE(S))

HISTOGRAM FOR ... 8-20

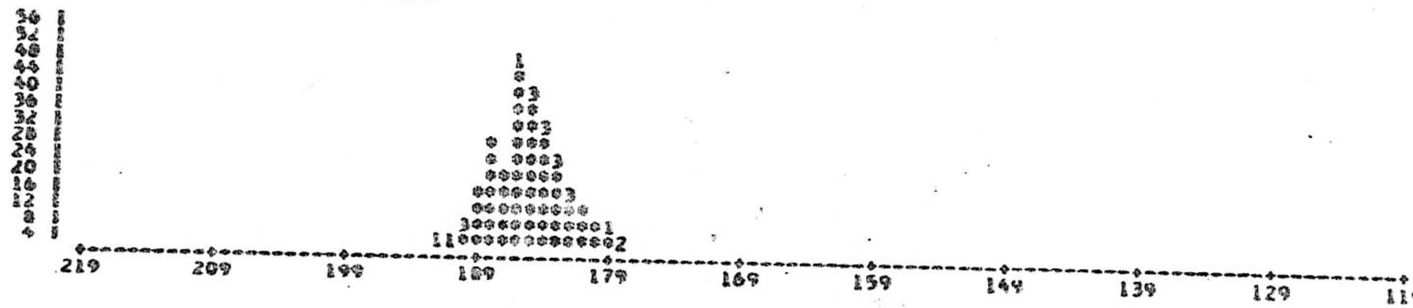
CHANNEL 1 0.40 - 0.44 MICRONS

EACH * REPRESENTS 4 POINT(S).



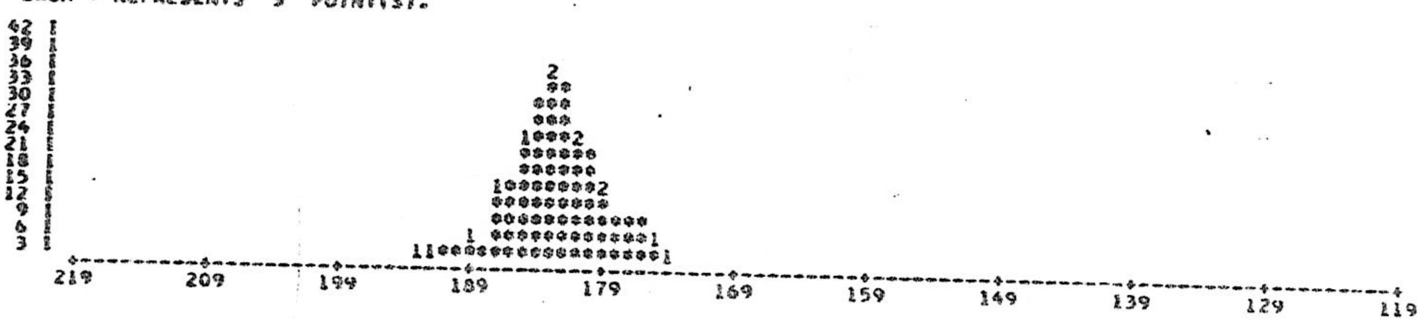
CHANNEL 9 0.62 - 0.66 MICRONS

EACH * REPRESENTS 4 POINT(S).



CHANNEL 12 0.80 - 1.00 MICRONS

EACH * REPRESENTS 3 POINT(S).



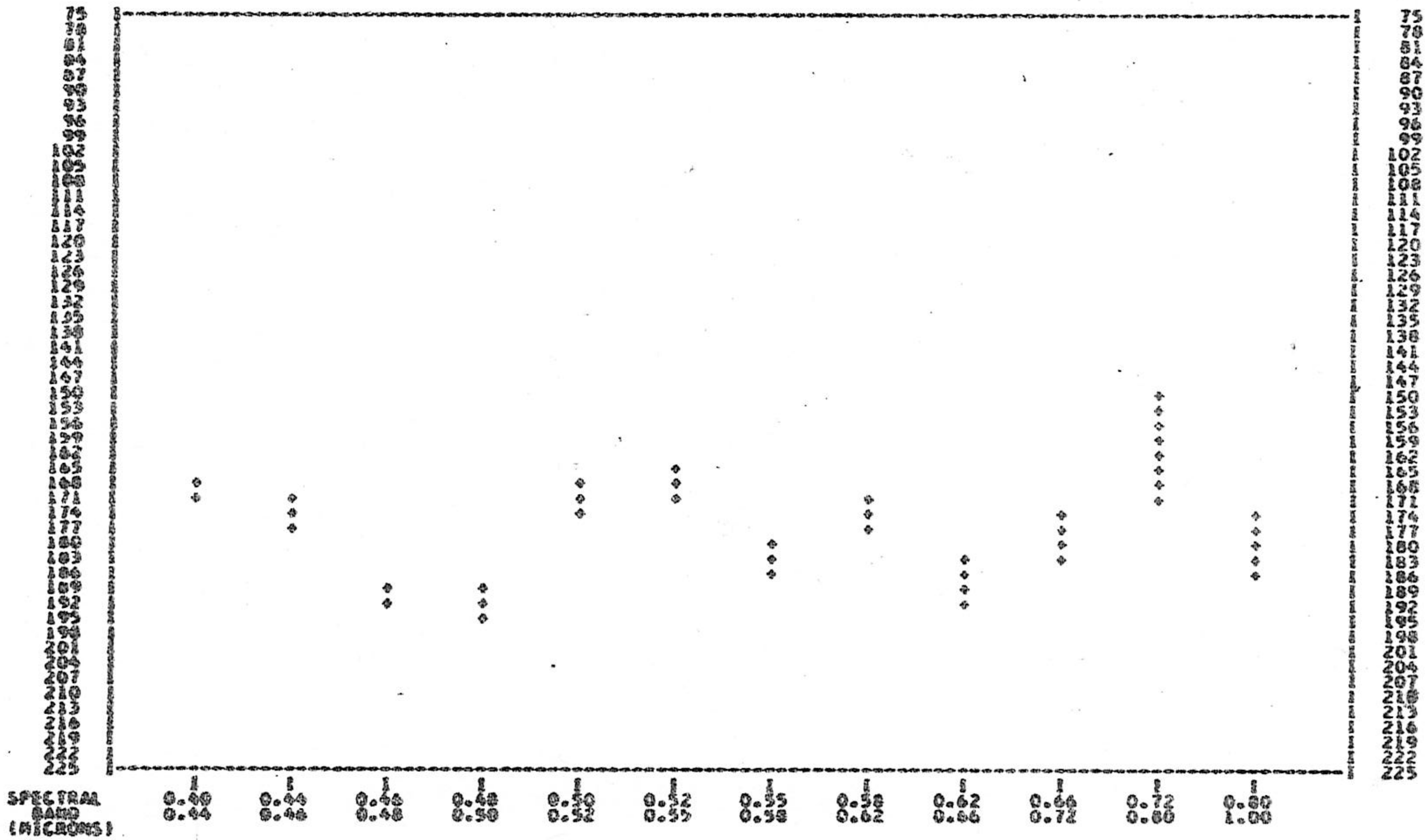
LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN I

SPECTRAL PLOT (MEAN PLUS AND MINUS ONE STD. DEV.) FOR TRAINING CLASS CORN I

LEGEND
+ = CLASS CORN I

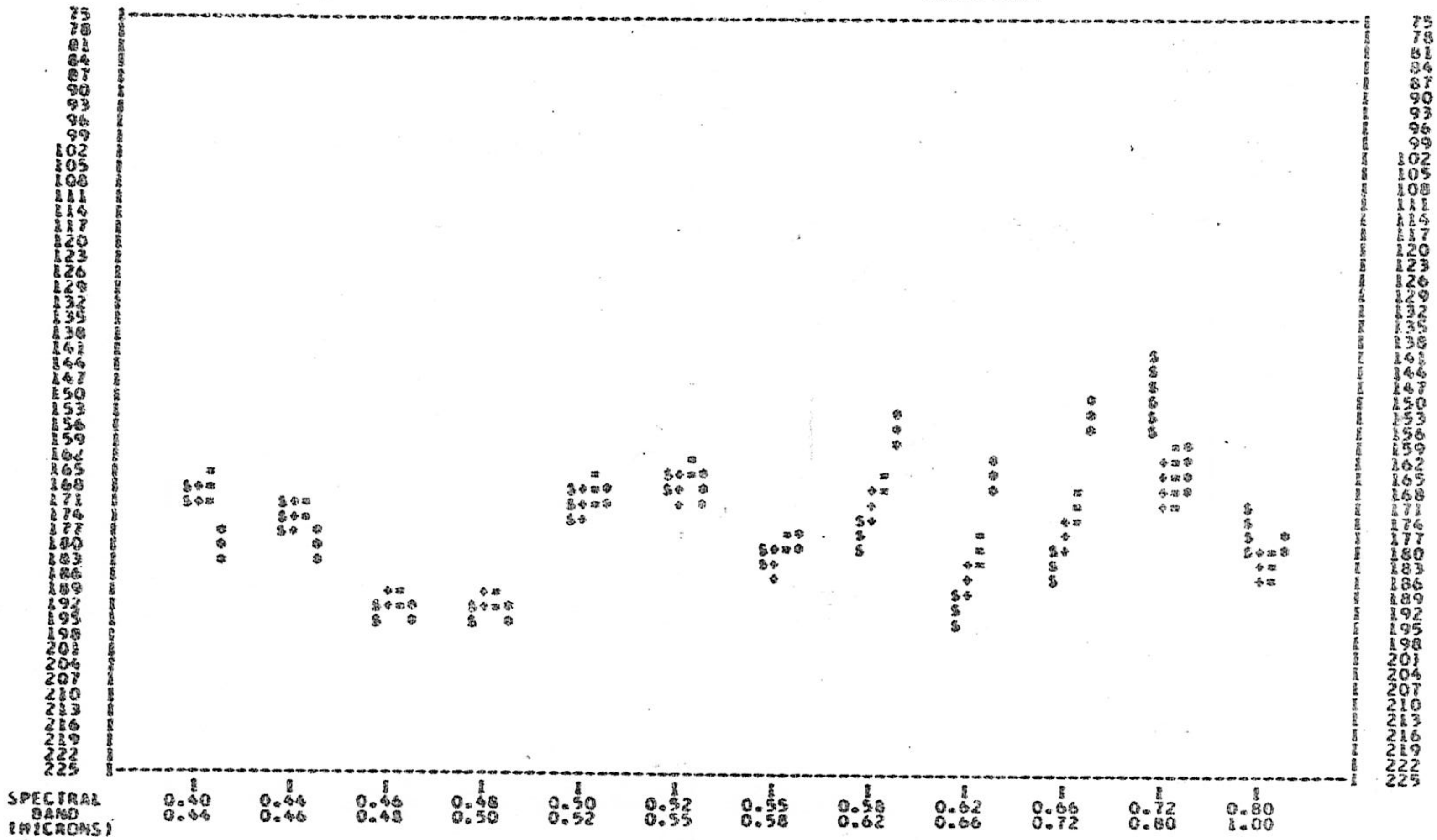


LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

SPECTRAL PLOT FOR TRAINING CLASS(ES) 1 2 3 4

LEGEND
$ = CLASS CORN A
+ = CLASS CORN B
= = CLASS SOYB
* = CLASS WHET



Statistics

The statistics processor of LARSYSAA will also provide a one-page display of statistics as illustrated on the next three pages. The first two are from the two individual corn fields used above and the third is again from these two taken together as a class.

The format is largely self-explanatory. Near the top the means and standard deviations for the data of the designated region in each spectral band are given. Below this is given the covariance matrix. The entry in the third row of the first column, which is 2.92 for the first field, for example, is the covariance between Channels 1 and 3. Since this matrix is symmetric ty definition, only that part below the major diagonal is presented.

In the lower portion of the page is presented the correlation matrix which, of course, is a normalized version of the covariance matrix. Options in the program provide for the presentation of these statistics for only a subset of these bands as desired by the researcher.

By use of another LARSYS option, the means and covariance matrix may also be obtained in punched card form. A line printer listing from a card deck obtained in this fashion is given on the page following the class statistics.

<u>Summary</u>. LARSYS can provide means, standard deviations, covariance matrices, and correlation matrices of data fields and classes in arbitrarily selected subsets of channels. LARSYS can also provide means and covariances in punched card form.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN 1

RUN NO. 26600061, FIELD 12-9
NO OF SAMPLES = 132, FROM LINES 403 TO 625 (EVERY 2 LINE(S)), SAMPLES 13 TO 33 (EVERY 2 SAMPLE(S))

THE COVARIANCE AND MEAN FOR TRAINING FIELD 12-9

	0.40-0.44	0.44-0.46	0.46-0.48	0.48-0.50	0.50-0.52	0.52-0.55	0.55-0.58	0.58-0.62	0.62-0.66	0.66-0.72	0.72-0.80	0.80-1.00
MEAN	169.39	174.33	192.80	193.45	171.08	165.83	182.29	177.59	192.54	182.98	148.94	175.16
ST DEV	2.86	3.09	1.62	1.71	2.89	2.32	2.04	2.96	2.95	3.06	8.49	4.99

COVARIANCE MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
8.18											
4.82	9.54										
2.92	2.70	2.62									
2.40	3.46	1.70	2.92								
4.77	5.67	2.99	2.85	8.37							
3.31	4.25	1.99	2.40	4.95	5.40						
2.06	3.43	1.62	2.08	3.52	3.48	4.16					
4.76	6.50	2.78	3.43	5.70	4.75	4.07	8.76				
4.81	5.90	2.67	3.60	5.26	4.12	4.09	7.27	8.69			
4.56	5.11	2.60	3.04	5.66	5.40	4.65	7.35	6.77	9.34		
-5.59	-6.29	-3.57	-5.71	-1.52	2.68	-0.43	-7.04	-10.02	-1.19	72.16	
-3.50	-3.01	-2.69	-2.91	-0.94	2.15	0.01	-3.44	-5.24	0.44	33.64	24.90

CORRELATION MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
1.00											
0.55	1.00										
0.63	0.54	1.00									
0.49	0.66	0.61	1.00								
0.58	0.64	0.64	0.58	1.00							
0.50	0.59	0.53	0.60	0.74	1.00						
0.35	0.54	0.49	0.60	0.60	0.73	1.00					
0.56	0.71	0.58	0.68	0.67	0.69	0.67	1.00				
0.57	0.65	0.56	0.71	0.62	0.60	0.68	0.83	1.00			
0.52	0.54	0.53	0.58	0.64	0.76	0.75	0.81	0.75	1.00		
-0.23	-0.24	-0.26	-0.39	-0.06	0.14	-0.02	-0.30	-0.40	-0.05	1.00	
-0.25	-0.20	-0.33	-0.34	-0.07	0.19	0.00	-0.23	-0.36	0.03	0.79	1.00

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN 1

RUN NO. 26600081, FIELD 8-20
NO OF SAMPLES = 253, FROM LINES 669 TO 713 (EVERY 2 LINE(S)), SAMPLES 171 TO 191 (EVERY 2 SAMPLE(S))

THE COVARIANCE AND MEAN FOR TRAINING FIELD 8-20

	0.40-0.44	0.44-0.46	0.46-0.48	0.48-0.50	0.50-0.52	0.52-0.55	0.55-0.58	0.58-0.62	0.62-0.66	0.66-0.72	0.72-0.80	0.80-1.00
MEAN	169.28	173.22	191.10	191.51	170.11	168.64	182.66	172.11	185.11	176.97	166.38	182.50
ST DEV	2.46	2.61	1.59	1.77	2.90	2.39	2.12	3.10	2.69	3.33	4.69	3.45

COVARIANCE MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
6.06											
2.89	6.79										
2.42	2.28	2.53									
1.56	2.81	1.42	3.12								
3.26	3.64	2.47	2.46	8.41							
2.84	4.10	2.14	2.86	4.46	5.69						
1.88	3.48	1.75	2.48	3.32	3.86	4.49					
2.85	5.18	2.77	3.60	6.27	5.88	4.92	9.60				
1.76	4.21	2.01	3.20	4.40	5.03	4.46	6.73	7.21			
1.85	4.38	2.42	3.66	5.45	5.75	5.06	8.44	7.15	11.08		
1.99	2.13	1.76	1.70	4.05	4.11	3.14	5.71	3.48	6.42	22.04	
1.93	2.44	1.69	1.35	4.27	3.79	2.34	4.81	3.05	5.23	9.10	11.91

CORRELATION MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
1.00											
0.45	1.00										
0.62	0.55	1.00									
0.36	0.61	0.51	1.00								
0.46	0.48	0.54	0.48	1.00							
0.48	0.66	0.57	0.68	0.64	1.00						
0.36	0.63	0.52	0.66	0.54	0.76	1.00					
0.37	0.64	0.56	0.66	0.70	0.80	0.75	1.00				
0.27	0.60	0.47	0.67	0.56	0.79	0.78	0.81	1.00			
0.23	0.51	0.46	0.62	0.56	0.72	0.72	0.82	0.80	1.00		
0.17	0.17	0.24	0.20	0.30	0.37	0.32	0.39	0.28	0.41	1.00	
0.23	0.27	0.31	0.22	0.43	0.46	0.32	0.45	0.33	0.45	0.56	

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN I

THE COVARIANCE AND MEAN FOR TRAINING CLASS CORN I

	0.40-0.44	0.44-0.46	0.46-0.48	0.48-0.50	0.50-0.52	0.52-0.55	0.55-0.58	0.58-0.62	0.62-0.66	0.66-0.72	0.72-0.80	0.80-1.00
MEAN	149.31	173.60	191.68	192.17	170.44	167.67	182.54	173.99	187.65	179.03	160.40	179.98
ST DEV	2.60	2.83	1.79	1.97	2.93	2.71	2.10	4.01	4.49	4.31	10.38	5.34

COVARIANCE MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
6.77											
3.57	7.99										
2.62	2.85	3.20									
1.89	3.52	2.26	3.90								
3.79	4.57	3.01	3.01	8.59							
2.92	3.43	1.01	1.44	4.00	7.36						
1.93	3.36	1.56	2.17	3.29	3.96	4.40					
3.63	6.99	4.86	5.94	7.26	2.00	4.15	16.07				
2.98	6.64	5.08	6.60	6.31	-0.01	3.69	16.09	20.17			
2.92	6.13	4.78	6.08	6.83	1.80	4.39	15.48	17.09	18.61		
-1.03	-5.13	-6.74	-8.51	-1.69	14.68	3.39	-20.44	-30.41	-19.86	107.79	
-0.11	-1.27	-2.62	-3.34	0.87	7.88	2.16	-7.10	-12.11	-6.38	46.36	28.48

CORRELATION MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
1.00											
0.49	1.00										
0.56	0.56	1.00									
0.37	0.63	0.64	1.00								
0.50	0.55	0.57	0.52	1.00							
0.41	0.45	0.21	0.27	0.50	1.00						
0.35	0.57	0.41	0.52	0.54	0.70	1.00					
0.35	0.62	0.66	0.75	0.62	0.18	0.49	1.00				
0.25	0.52	0.63	0.74	0.48	-0.00	0.39	0.89	1.00			
0.26	0.50	0.62	0.71	0.54	0.15	0.49	0.90	0.88	1.00		
-0.04	-0.17	-0.36	-0.42	-0.06	0.52	0.16	-0.49	-0.65	-0.44	1.00	
-0.01	-0.08	-0.27	-0.32	0.06	0.54	0.19	-0.33	-0.51	-0.28	0.84	1.00

```
MODULE TRAINING FIELD DECK FOR LARSYSAA
CLASS     1FIELD     2FEAT     12VECTR 1 2 3 4 5 6 7 8 910111213141516171818CAL 1
FLDINF         26600061  603   625    13    33     1  26600081   669   713   171   191      1
CLSDESC    CORN 1
FREQ    0.40  0.44   0.46   0.48   0.50   0.52   0.55   0.58   0.62   0.66   0.72   0.80
FREQ    0.44  0.46   0.48   0.50   0.52   0.55   0.58   0.62   0.66   0.72   0.80   1.00
NUPTS       385
MEAN    0.16931429E 03 0.17360258E 03 0.19168051E 03 0.19217401E 03 0.17044415E 03
MEAN    0.16767271E 03 0.16253506E 03 0.17399220E 03 0.18765454E 03 0.17902856E 03
MEAN    0.16039999E 03 0.17998181E 03
COVAR   0.67681541E 01 0.35705357E 01 0.79900970E 01 0.26241064E 01 0.28466644E 01
COVAR   0.32023535E 01 0.18930798E 01 0.35224628E 01 0.22588673E 01 0.38993235E 01
COVAR   0.37897320E 01 0.45675926E 01 0.30146637E 01 0.30136490E 01 0.85860653E 01
COVAR   0.29208326E 01 0.34321022E 01 0.10097532E 01 0.14555397E 01 0.39973011E 01
COVAR   0.73613634E 01 0.19251480E 01 0.33616266E 01 0.15568037E 01 0.21722670E 01
COVAR   0.32929783E 01 0.39594221E 01 0.44004593E 01 0.36274548E 01 0.69890823E 01
COVAR   0.48646908E 01 0.59414635E 01 0.72638865E 01 0.20026512E 01 0.41474085E 01
COVAR   0.16065033E 02 0.29786453E 01 0.66410036E 01 0.50768461E 01 0.65393366E 01
COVAR   0.63126888E 01-0.91850048E-02 0.36931343E 01 0.16091049E 02 0.20174606E 02
COVAR   0.29180803E 01 0.61285706E 01 0.47773809E 01 0.60835562E 01 0.68284225E 01
COVAR   0.18010416E 01 0.43935261E 01 0.15479389E 02 0.17088013E 02 0.18611160E 02
COVAR  -0.10348949E 01-0.51296873E 01-0.67442703E 01-0.85098953E 01-0.16911449E 01
COVAR   0.14683332E 02 0.33921871E 01-0.20436966E 02-0.30410934E 02-0.19863007E 02
COVAR   0.10778749E 03-0.11406249E 00-0.12702646E 01-0.26230106E 01-0.33405771E 01
COVAR   0.87007570E 00 0.78820543E 01 0.21607952E 01-0.70964956E 01-0.12105255E 02
COVAR  -0.63822908E 01 0.46361450E 02 0.28476227E 02
```

Feature Selection/Separability

The feature selection processor provides the capability for measuring the degree of separability of Gaussianly distributed classes and determining the optimum set of channels for doing so. This is done by calculating the statistical distance in N-dimensional space between the classes. The particular distance measure implement is known as divergence. See, for example, Marill and Green, "On the Effectiveness of Receptors in Recognition Systems," IEEE Transactions on Information Theory, January 1963. This paper also provides a relationship between numerical divergence values and the degree of separability of the classes in terms of per cent error under certain conditions. Other separability measures have been tested and could be implemented if it becomes desirable.

The program is written in a mode highly interactive with the researcher so that he may make judgments and supervise the calculations as they are proceeding. It also has many options, not all of which will be discussed here. The next page does provide a sample of the type of output provided, however. In this example problem, five classes were designated for study as follows:

Class Name	Symbol
Soybeans	S
Corn	C
Oats	O
Wheat I	W
Wheat II	M

The program was instructed to determine the 30 best sets of four channels (features) for separating, i.e., correctly classifying, these five classes. In addition, it was specified that the separability of the 4-tuple of Channels 1, 5, 10 and 12 should be displayed even if it was not in the top 30. The results are as shown. The best 4-tuple of channels is seen to be 1, 6, 10 and 12. To the right of the column labeled D(TOT) are listed the interclass divergences. For example, for Channels 1, 6, 10 and 12, that for soybeans and corn classes, S and C, is 36, while that for soybeans and oats, S and O, is 84. This means that a

*** LARSYSAA ILLUSTRATION ***

RETENTION LEVEL .. 325
MINIMUM 25
MAXIMUM 350

	FEATURES	DIJ(MIN)	D(TOT)	INTERCLASS DIVERGENCE SC (10)	SO (10)	SW (10)	SM (10)	CO (10)	CW (10)	CM (10)	OW (10)	OR (10)	WR (0)
1.	1 6 10 12	36.	1538.	36	84	194	179	120	...	347	95	133	0
2.	1 6 10 11	33.	1534.	33	81	196	184	120	...	...	90	130	0
3.	1 6 8 11												
4.	1 6 9 11	27.	1509.	27	69	201	189	83	...	...	104	136	0
5.	1 6 9 12	29.	1501.	29	73	200	175	82	...	343	111	138	0
6.	1 6 8 12	26.	1495.	26	82	170	206	93	311	...	98	159	0
7.	6 9 10 12	35.	1484.	35	66	175	153	108	...	...	116	131	0
8.	6 8 10 12	35.	1483.	35	67	160	164	109	...	...	108	140	0
9.	6 8 10 11	33.	1477.	33	60	166	183	105	...	...	104	126	0
10.	6 9 10 11	33.	1474.	33	58	179	168	105	...	...	108	123	0
11.	1 6 9 10	31.	1469.	31	54	220	192	112	...	...	71	89	0
12.	1 8 10 11	28.	1461.	28	71	166	212	95	295	...	93	151	0
13.	2 6 10 12	35.	1448.	35	76	166	154	114	...	325	95	133	0
14.	1 9 10 11												
15.	1 7 10 11	29.	1447.	29	74	174	183	101	310	...	87	139	0
16.	1 8 9 11												
17.	2 6 10 11	32.	1438.	32	69	169	158	112	...	33[illegible]	90	126	0
18.	1 6 8 10	29.	1437.	29	47	206	209	106	...	...	52	88	0
19.	1 8 10 12	30.	1436.	30	70	169	208	92	274	...	95	157	0
20.	1 7 9 11												
21.	6 8 9 11	26.	1433.	26	51	164	176	71	...	...	113	132	0
22.	1 8 9 12	26.	1432.	26	70	[illegible]	205	79	279	...	102	159	0
23.	6 8 9 12	27.	1429.	27	59	158	151	76	...	344	119	145	0
24.	1 9 10 12	27.	1428.	31	67	171	193	92	299	...	100	125	0
25.	6 10 11 12	35.	1427.	35	72	152	147	106	...	326	111	128	0
26.	2 6 9 11	26.	1426.	26	55	175	159	72	...	...	105	134	0
27.	2 6 9 12	28.	1419.	28	63	172	144	74	...	336	112	140	0
28.	1 7 10 12	32.	1415.	32	75	169	179	99	290	343	89	137	0
29.	2 6 8 11												
30.	1 10 11 12	28.	1407.	32	76	166	174	94	300	343	107	117	0
*** 34.	1 5 10 12	31.	1403.	31	72	170	181	91	279	...	89	140	0

classifier will have less difficulty separating soybeans from oats than soybeans from corn. It is inappropriate in the space available here to say much about the meaning of the size of these numbers beyond the general statement that a divergence above 400 indicates very good separability while one below about 20 would be marginal.

The entries in the column labeled D(TOT) are the sum of the interclass divergences and it is on the basis of these numbers that the feature sets are ordered. The column labeled DIJ(MIN) gives the minimum interclass divergence for each feature set for easy reference. Other options illustrated include the following. If desired, one may input a maximum divergence to be considered in ordering the feature sets. This is desirable in order to prevent one highly separable pair of classes from unduly influencing the ordering of feature sets. The maximum set in this example is 350, and the entries indicated by three dots are those exceeding this maximum.

One may also set a minimum divergence. If any interclass divergence does not exceed the minimum for a given feature set, that feature set will be deemed unacceptable and not considered further. This is indicated by the blank lines in the interclass divergence table.

Further, it is possible to apply varying weights to the interclass divergences as indicated by the numbers in parentheses under the symbols. In this example problem, even though it is necessary to define two wheat categories, it is not desired to separate them from each other. Therefore, a weight of zero has been specified for the WM interclass divergence.

Note also that the feature set specifically called for, 1, 5, 10 and 12, is displayed and marked by ***. Its rank turned out to be 34.

Summary. LARSYS assists in determining the degree of separability of classes and can provide information about the best set of features for doing so. The program does require a user with specific knowledge of divergence and this algorithm, however.

Classification and Display

The proof of separability of classes comes with a classification and this may be done using the classification and display processors of LARSYSAA. The classification processor carries out the actual classifications using a Gaussian maximum likelihood scheme and stores the results on magnetic tape. The display processor is then used to read the results from the results tape, apply various options, and display and analyze the results. Currently the display is presented only in line printer form but after procurement of the digital display, this device will be used also.

An example classification display of the area previously shown in the pictorial printouts is given on the next page. For this classification nine classes were defined, training samples from run 26600061 were selected for each class, and their statistics computed and punched using the statistics processor. Four features, Channels 1, 6, 10 and 12, were used in the classification as indicated in the heading.

After classification, the display processor was called and symbols appropriate to the classes were assigned by the user at that time. The choice of symbols is entirely arbitrary. For example, it may be desirable to display the same classification result using S for soybeans and blanks for all others in order to make more obvious the location of all soybeans.

In addition to assigning each resolution element to one of the nine categories, a threshold capability is also provided in the display processor. That is, the classification of each element is compared to a user-assigned threshold and if the element classification does not exceed the threshold; that is, if the element doesn't look sufficient like a member of the class to which it has been tentatively assigned even though this is the most likely class, then final classification of the element is declined and that element is assigned to a null category and displayed as a blank. Different thresholds may be assigned to each of the classes individually if desired.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705007300
CLASSIFICATION DATE .. JULY 5, 1968

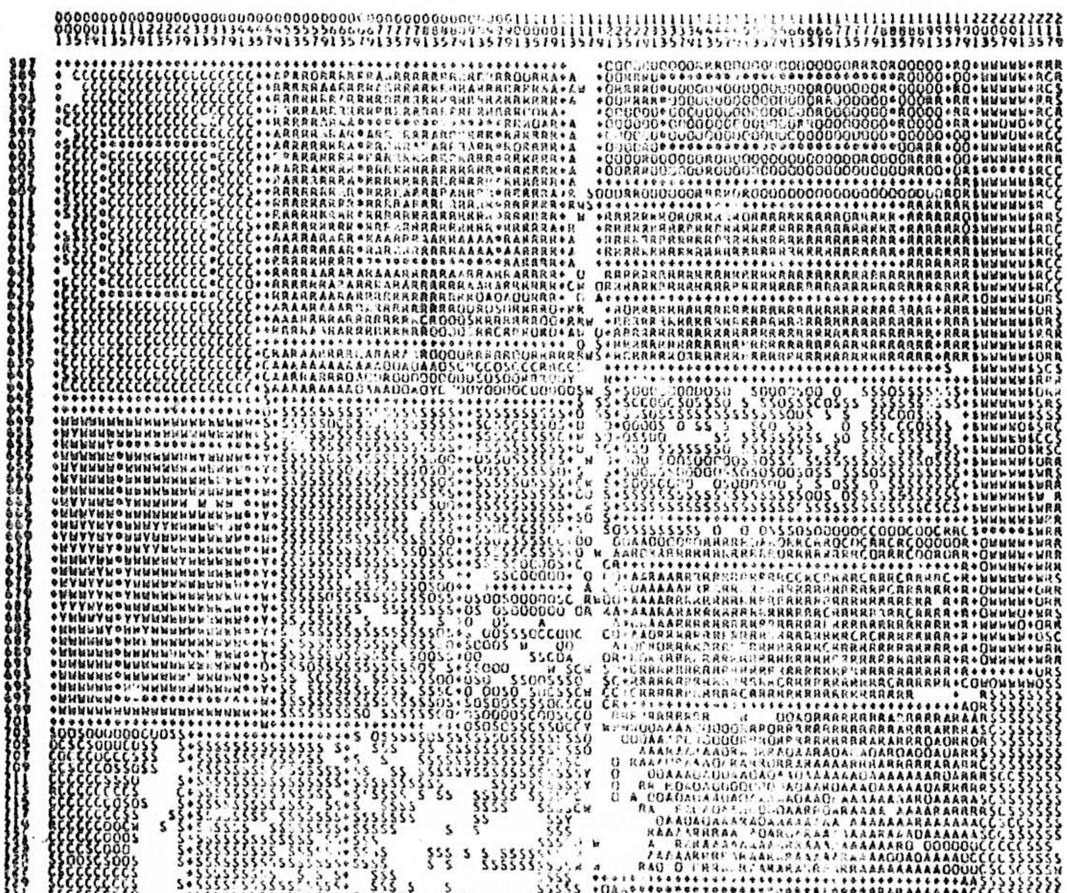
RUN NUMBER-----26600061	DATE ----- 5/28/66
FLIGHT LINE---- C1	TIME------1229
TAPE NUMBER---- 102	ALTITUDE-- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	14.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
H	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00



TOTAL NUMBER OF SAMPLED POINTS = 11600

While by studying such a display, one can obtain a satisfactory qualitative determination of the classifier performance, a more quantitative analysis of results is usually desirable. To facilitate this analysis, the display processor is able to receive the addresses of additional data fields, called test fields, together with a designation of the class to which each belongs and tabulate the performance in these fields. The areas so designated in this classification are outlined with + on the display and the tabulation is given on the page following the display. The same type of tabulation except giving the performance on a per class basis is given on the next page. In addition to the test data fields, the program can produce the same type of outlining and tabulation for the training samples.

<u>Summary</u>. LARSYS produces Gaussian maximum likelihood classifications stored on tape for repeated reference using an arbitrary number of classes and an arbitrary number of features. It also displays the results with a number of options and provides a quantitative tabulation of the results. Again these two processors require an especially knowledgeable user.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705807300
CLASSIFICATION DATE .. JULY 5, 1968

RUN NUMBER------26600061 DATE------ 6/28/66
FLIGHT LINE---- C1 TIME------1229
TAPE NUMBER---- 102 ALTITUDE--- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	14.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
W	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST FIELDS

	CLASS	NO OF SAMPS	PCT. CORCT	SOYB	CORN	OATS	WHEA	RED	ALFA	RYE	SOIL	THRS
				NO OF SAMPLES CLASSIFIED INTO								
7-27	SOYB	407	63.4	258	12	84	0	0	0	0	0	53
12-7	SOYB	513	88.9	456	4	31	0	0	0	0	0	22
12-2	SOYB	150	79.3	119	11	18	0	0	0	0	0	2
12-3	SOYB	752	89.2	671	8	0	0	0	0	0	0	73
1-23	SOYB	546	97.3	531	4	0	1	0	0	0	10	0
12-9	CORN	588	94.0	25	553	1	0	1	0	0	0	8
7-1	OATS	370	84.9	0	0	314	0	56	0	0	0	0
7-2	WHEA	260	93.5	0	0	17	243	0	0	0	0	0
12-10	WHEA	546	90.1	0	0	0	492	0	0	48	0	6
12-8	RED	713	80.2	2	3	27	0	572	109	0	0	0
1-29	RED	128	96.9	0	0	4	0	124	0	0	0	0
7-28	RED	175	98.9	0	0	2	0	173	0	0	0	0
	RED	385	86.8	0	17	5	0	334	24	0	0	5
7-24	ALFA	190	93.2	0	0	2	0	11	177	0	0	0
7-24	ALFA	266	83.8	0	4	19	0	18	223	0	0	2
	TOTAL	5984		2062	616	524	736	1289	533	48	10	171

OVERALL PERFORMANCE = 87.5

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705607300
CLASSIFICATION DATE .. JULY 5, 1968

RUN NUMBER------26600061   DATE------ 6/28/66
FLIGHT LINE---- C1   TIME------1229
TAPE NUMBER---- 102   ALTITUDE--- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	16.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
W	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST FIELDS

	CLASS	NO OF SAMPS	PCT. CORCT	SOYB	CORN	OATS	WHEA	RED	ALFA	RYE	SOIL	THRS
				NO OF SAMPLES CLASSIFIED INTO								
7-27	SOYB	407	63.4	258	12	84	0	0	0	0	0	53
12-7	SOYB	513	88.9	456	4	31	0	0	0	0	0	22
12-2	SOYB	150	79.3	119	11	18	0	0	0	0	0	2
12-3	SOYB	752	89.2	671	8	0	0	0	0	0	0	73
1-23	SOYB	546	97.3	531	4	0	1	0	0	0	10	0
12-9	CORN	588	94.0	25	553	1	0	1	0	0	0	8
7-1	OATS	370	84.9	0	0	314	0	56	0	0	0	0
7-2	WHEA	260	93.5	0	0	17	243	0	0	0	0	0
12-10	WHEA	546	90.1	0	0	0	492	0	0	48	0	6
12-8	RED	713	80.2	2	3	27	0	572	109	0	0	0
1-29	RED	128	96.9	0	0	4	0	124	0	0	0	0
7-28	RED	175	98.9	0	0	2	0	173	0	0	0	0
	RED	385	86.8	0	17	5	0	334	24	0	0	5
7-24	ALFA	190	93.2	0	0	2	0	11	177	0	0	0
7-24	ALFA	266	83.8	0	4	19	0	18	223	0	0	2
	TOTAL	5989		2062	616	524	736	1289	533	48	10	171

OVERALL PERFORMANCE = 87.5

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705807300
CLASSIFICATION DATE .. JULY 5, 1968

RUN NUMBER------26600061 DATE------ 6/28/66
FLIGHT LINE---- C1 TIME------1229
TAPE NUMBER---- 102 ALTITUDE-- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	14.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
M	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOYB	CORN	OATS	WHEA	RED	ALFA	RYE	SOIL	THRS
1	SOYB	2368	85.9	2035	39	133	1	0	0	0	10	150
2	CORN	586	94.0	25	553	1	0	1	0	0	0	8
3	OATS	370	84.9	0	0	314	0	56	0	0	0	0
4	WHEA	806	91.2	0	0	17	735	0	0	48	0	6
5	RED	1401	85.9	2	20	38	0	1203	133	0	0	5
6	ALFA	456	87.7	0	4	21	0	29	400	0	0	2
	TOTAL	5989		2062	616	524	736	1289	533	48	10	171

OVERALL PERFORMANCE = 87.5

AVERAGE PERFORMANCE BY CLASS = 88.3