Final Report for RD Agreement No. 97-B32
July 23, 1999

Hyperspectral Feature Analysis for Performance Forecasting

Purdue University

Larry Biehl
David Landgrebe, Ph.D.

Abstract

This report summarizes our efforts on this project for the past three years. The overall objective of this effort was to develop sound procedures for the creation of robust first and second order statistics for specified background and target classes using the HYDICE data collected for the HYMSMO program. These statistics were provided to Dr. John Kerekes of MIT/Lincoln Laboratory to support their development of the Forecasting and Analysis of Spectroradiometric System Performance (FASSP) model. Our tasks concentrated on the development, understanding and use of the second order statistics that are required for the FASSP model. We also provided support for the collection of HYDICE data over agricultural areas that were obtained to expand the range of background data available for the HYMSMO program.

The report consists of a series of five mini studies that document our efforts. The first is a study on the number of samples required to define a spectral class in terms of the first and second order statistics as a function of the classifier complexity.

The second describes the implementation of the leave-one-out covariance (looc) estimator that can be used to determine the covariance matrix for targets that have a small number of training samples available.

The third is a quantitative study of the use of the looc estimator to determine how many samples are needed for the looc estimator from a practical point of view.

The fourth summarizes the target and background statistics that have been made available to MIT/LL and the procedures that were used.

The fifth describes our support of the August 7, 1997 HYDICE flight over the agricultural areas and the ground reference information that is available. HYDICE data were also collected over the same areas on April 29, 1999. Ground information was collected for this flight also. However we did not complete processing of these data since we did not received a copy of the HYDICE data before the completion of this report.

1. Analysis of Number of Samples Needed to Define a Spectral Class in Terms of First and Second Order Statistics.

Analysis Objective

The overall objective of this effort is to support MIT/Lincoln Laboratory in their development of the Forecasting and Analysis of Spectroradiometric System Performance (FASSP) Model.

The more specific objective is to assist with analysis of and evaluate role of second order statistics in the surveillance, reconnaissance and abundance estimation of targets within a background.

Statement of the Problem

The overall task is to determine appropriate and robust method(s) to estimate the second order statistics for background and target spectral classes.

Technical Objectives

The initial objective, reported on here, is to evaluate what is currently known in the literature about the number of samples needed to define a spectral class in terms of first and second order statistics. The next objective is to generate first and second order statistics for background classes for HYDICE data sets within an accepted Terrain Categorization (TERCAT) system that will be used by the FASSP model at Lincoln Laboratory. Other objectives include investigating methods to augment the statistics for a class, such as combining samples from diverse observations (such as a.m. and p.m. observations, different altitudes, etc.) or using "non-labeled" samples [Shahshahani, 1994].

Background

The matter of the required size of a reference data set needed to properly define a data class in multispectral data is a long standing question. From a pragmatic standpoint, one would like to have a specific number conclusion at least as a rule of thumb such as 10n or 100n as reported in [Swain/Davis, 1978] and [Richards, 1993], where n is the number of channels. This can serve reasonably well in a suitably narrow and well-defined set of circumstances.

However, it is appropriate to conclude from the long standing nature of this question that there is more to the matter than may at first appear. This is because the correct answer in any instance is dependent upon a number of factors which are all interrelated. Some of these factors are (a) how information is contained in the data, (b) what type of analysis scheme is to be used, (c) what is the quality and level of detail (i. e., the number of bands and signal-to-noise ratio or bits per pixel), and (d) what classes does one wish to discriminate between and to what accuracy.

One of the early papers to show in a definitive though highly generalized way this relationship was one by [Gordon Hughes, 1968]. One of the primary results of this paper is given in Figure 1 which relates recognition accuracy and measurement complexity with the number of training samples as a parameter.

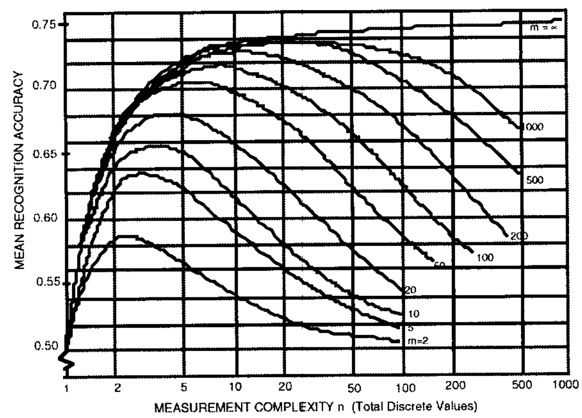


Figure 1. Mean Recognition Accuracy vs. Measurement Complexity with the number of training samples as a parameter.

The literature also indicates that there is a relationship between the number of training samples needed and the complexity of the classifier to be used. The more complex the classifier algorithm, the more samples that are needed to achieve the improved performance which more complex algorithms should provide. [Fukunaga, 1989] reports for one example circumstance that the required number of training samples is linearly related to the dimensionality for a linear classifier and to the square of the dimensionality for a quadratic classifier. [Scott, 1992] reports that the number of training samples required is exponentially related to the dimensionality for nonparametric classifiers.

Methodology/Approach

To begin to develop some quantitative information about the number of samples required to generate sounds statistics to be used in the FASSP model, a series of simulations of practical situations appropriately sampling the spectrum of relevant possibilities have been defined.

Data (pixels) for a pair of spectral classes are selected from a HYDICE data set. The data for each pair of classes is randomly assigned to two equal groups of pixels that are similar but mutually exclusive. These two groups, one for design and the other for test, then serve as the data bases to run a series of experiments in which the number of samples used for training varies and the complexity of the classifier varies. For simplicity, it is assumed that the two classes are equally probable, therefore the number of pixels in the data bases for each class is the same. A subset of the 210 channels of the HYDICE data are used. Those channels in the major water absorption regions and the very low signal to noise channels near the beginning and end of the HYDICE spectrum are not used.

Results

Data for "Grass" (13705 pixels) and "Road" (5714 pixels) background classes were selected from a portion of the Forest Radiance 1, 950824 Run05 flight line. The number of pixels assigned to the design and test data bases for each class is 2857 pixels. The channels used for this study were 6-103, 110-138, and 153-206 for a total of 181 out of the 210 channels available.

Let n be the dimensionality (the number of channels) and k be the ratio of training sample size to dimensionality. The parameter k was varied from m to 50 for dimensions of 21 and 42, m to 30 for 90 dimension and m to 15 for 181 dimensions. m represents the factor of k required for the number of samples to be equal to n+1, the minimum number of samples required to compute the covariance matrix for a class. There were not enough pixels available to run cases of k for 90 and 181 dimensions at the higher values. For each k, kn pixels from each class were randomly drawn from the design data base. All of the pixels in the test data base were used for each value of k. The design set, containing pixels drawn from the design data base, were used to estimate the sample mean vector and covariance matrix for each class. A Maximum Likelihood (ML) classifier was designed by using these estimated statistics in the quadratic ML discriminant function. Using this ML classifier to classify the design and test sets, an error rate was obtained for both sets. To obtain the expected error rate, the experiment was repeated ten times, each time randomly selecting the samples to be used from the data base and the average error rate and its standard deviation were computed. The results are shown in Figure 2. The above experiment was repeated using the Fisher Discriminant Function, a linear classifier. Those results are shown in Figure 3.

Analysis

Table 1 summarizes the values of k at which the expected error rate stabilized.

Table 1. Values of k (and number of samples) at which the expected error rate stabilized as a function of the dimension and the classifier algorithm.

Dimension	Quadratic Classifier	Linear Classifier
21	20 (420)	15 (315)
42	20 (840)	15 (630)
90	*	20 (1800)
181	*	*

*Expected error rate had not stabilized at the maximum k's possible for 90 dimensions (k=30) and 181 dimensions (k=15).

Figures 1 and 2 indicate that the expected error rate decreased as the number of dimensions increased, more so for the quadratic classifier than for the linear classifier. The results also indicate that the final expected error rates for the linear classifier are a little higher than those for the quadratic classifier as one would expect. The expected test error rate for the linear classifier was significantly less than that for the quadratic classifier for k's of 2 and lower. At the k for around 1, the test error was reduced from .2 for the quadratic classifier to around .025 for the linear classifier. Many times in 'real life' situations, the number of pixels available for training is on the order of 2 times the dimensionality or fewer. In those situations, as reported in the literature, one may do better by using a linear classifier as opposed to the quadratic classifier.

Conclusions

Based on these results which represent just one pair of classes which are fairly easy to separate, the k factor needed to generate sound statistics for background classes for the FASSP model should be at least 15-20 when 40 or fewer dimensions are used and more than 30 for 90 or more dimensions. A different experiment, not shown here, with grass and road classes which were not as separable as the above set indicated that error rates stabilized at k values of 20-25 for 40 or fewer dimensions.

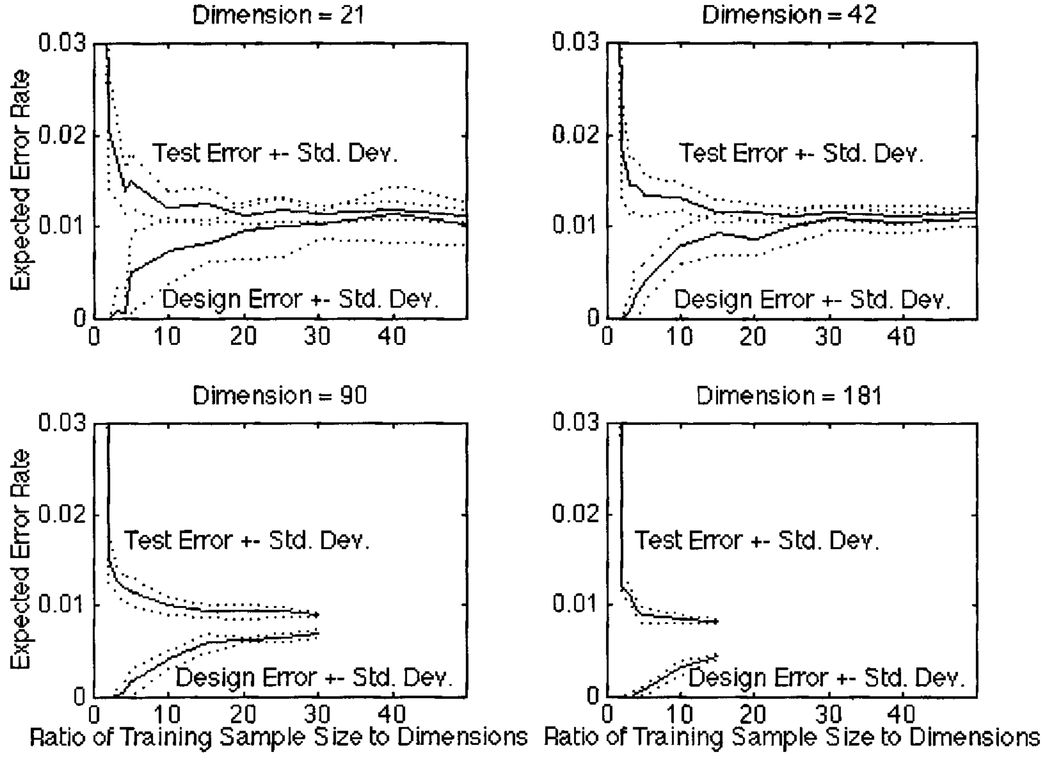


Figure 2. Error Rates vs. Training Sample Size for a Quadratic Classifier (Maximum Likelihood)

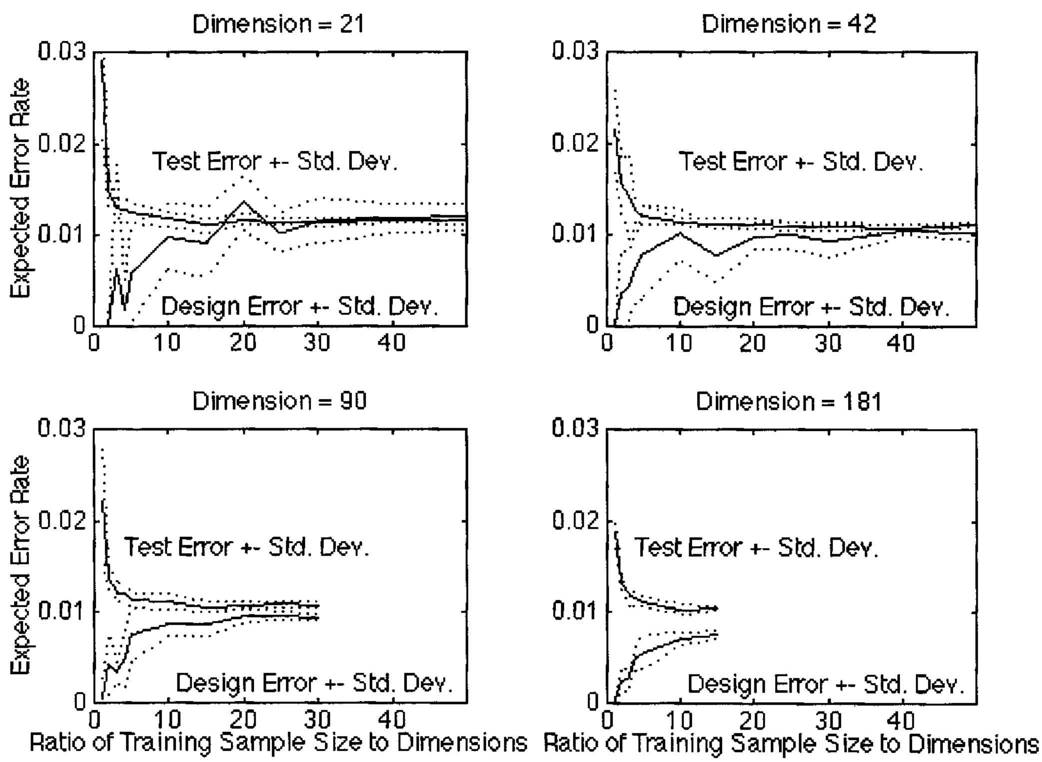


Figure 3. Error Rates vs. Training Sample Size for a Linear Classifier (Fisher Discriminant)

Recommendations

More simulation experiments need to be run for other pairs of background classes to develop a better confidence in the k factor that should be used for HYDICE data and classes. Also simulation experiments are needed which include target classes. There is very little chance of having hundreds to thousands of observations available for some of the targets which are within the background. These targets though may behave quite differently than the background classes which may reduce the need for as many samples to model their second order statistics. [Hoffbeck, 1996] summarizes and reports on some methods to estimate the covariance's of classes with limited training data that may be sufficient for some classifiers. We intend to investigate those methods as we generate the statistics for target classes for the FASSP model.

References

Fukunaga, K., Hayes, Raymond R., Effects of Sample Size in Classifier Design, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-11, No. 8, Aug-89

Hoffbeck, Joseph P. and David A. Landgrebe, "Covariance Matrix Estimation and Classification with Limited Training Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, no. 7, pp. 763-767, July 1996.

Hughes, G. F., "On the Mean Accuracy of Statistical Pattern Recognizers," *IEEE Transactions on Information Theory,* Vol. IT-14, No. 1, January 1968.

Richards, J. A. Remote Sensing Digital Image Analysis, Springer-Verlag, First Edition, 1986, Second Edition, 1993.

Scott, D. W. "Multivariate Density Estimation." New York: John Wiley & Sons, 1992.

Shahshahani, Behzad M. and David A. Landgrebe, "The Effect of Unlabeled Samples in Reducing the Small Sample Size Problem and Mitigating the Hughes Phenomenon," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, No. 5, pp. 1087-1095, September 1994.

Swain, P.H., and Davis, S. M., eds. Remote Sensing, The Quantitative Approach, McGraw-Hill, 1978.

2. Evaluate and Implement the Leave-One-Out Covariance Estimator to Determine the Covariance Matrix of Targets with Small Numbers of Training Samples.

Analysis Objective

The overall objective of this effort is to support MIT/Lincoln Laboratory in their development of the Forecasting and Analysis of Spectroradiometric System Performance (FASSP) Model.

The more specific objective is to assist with analysis of and evaluate role of second order statistics in the surveillance, reconnaissance and abundance estimation of targets within a background.

Statement of the Problem

The overall task is to determine appropriate and robust method(s) to estimate the second order statistics for small, low probability target spectral classes.

Technical Objectives

The objective reported on here was to evaluate and implement an algorithm which can be used to estimate the covariance matrix of targets when the number of sample measurements for a target is less than the number of spectral channels being used. This algorithm was then used to generate the covariance matrices for targets for MIT/Lincoln Laboratory to use in their FASSP model. Another closely related objective is given an estimate of the covariance matrix of a target at one ground resolution and sun angle, what range of ground resolution and sun angles can that estimate be reliably used across.

Background

Some algorithms, such as the Gaussian Maximum Likelihood classifier, which can be used for the detection and/or identification of targets within a scene require the use of the second order statistics, e.g. the covariance matrix, to help characterize the target in addition to the mean. Also, models such as FASSP require the second order statistics of targets to predict the performance of algorithms even though the algorithm may not use the covariance matrix directly. However, many times where the number of spectral features is large and the targets are small relative to the spatial resolution of the imagery, the number of samples available to make a good estimate of the covariance matrix is small.

Ordinarily, for the Gaussian Maximum Likelihood classifier algorithm, the absolute minimum number of sample needed is p+1 samples for p-dimensional data. Practically speaking, 10-50 times p is required to provide estimates of the covariance matrix which fully utilize all of the information available from the data. (See last years report.)

Methods have been discussed in the literature which appear to be useful for designing parametric classifiers with limited training data. The methods implement a covariance estimator that performs better than the sample covariance when the training set is limited. One of these methods is to use a common covariance estimate for all classes of targets within the scene. However, the common covariance may not be optimal for all cases when there are significant differences in the covariances of the targets and the backgrounds that the targets are within. Other methods have been investigated including the leave-one-out covariance (LOOC) [Hoffbeck, 1996] and the regularized discriminant analysis (RDA) [Friedman, 1989].

The LOOC estimator employs a mixing parameter to select an appropriate mixture of the (a) sample covariance, (b) diagonal sample covariance, (c) common covariance, and (d) the diagonal common covariance. The mixture deemed appropriate is the one that achieves the best fit to the training samples in the sense that it maximizes the average likelihood of training samples that were not used in the estimate. This covariance estimator is typically non-singular when at least three samples are available regardless of the dimensions of the data, and so it can be used even when the sample covariance or common covariance estimates are singular.

RDA is a two-dimensional optimization over four-way mixtures of the sample covariance, common covariance, the identity matrix times the average diagonal element of the common covariance, and the identity matrix times the average diagonal element of the sample covariance. The two mixing parameters are called lambda and gamma, and the index that is maximized, like the LOOC, is the leave-one-out classification error. Since the index depends on the covariance estimates of the other classes, the same values of the mixing parameters are used for all classes.

Hoffbeck compared of the results of the LOOC and RDA methods using both simulated data and AVIRIS data and found the results from the tests run in the comparisons were very similar, within a percent of each other for the AVIRIS data, however, LOOC requires less computation.

Methodology/Approach

LOOC estimates were made for the targets within the Forest Radiance I scene using the pixels available in the HYDICE imagery defined by the center pixel masks that were developed for the standardized data sets. The LOOC estimates were made for three HYDICE ground resolution sizes (.75, 1.5 and 3.0 feet) and two times of day (AM and PM). Also estimates of the covariances of some of the targets were made with the LOOC technique using the spectrometer data collected during the Forest Radiance I mission. The covariance estimates of the targets for the three ground resolutions, two times of day and the spectrometer data were compared using correlation images, Bhattacharyya distance (covariance term only) and eigenvector comparisons. The estimates have been provided to John Kerekes for use in the FASSP model. The LOOC estimates of the targets should allow the capability for FASSP to use all available channels for algorithms that require the inverse of the covariance matrix since the leave-out-covariance matrix will be non-singular.

Results

Figure 1 illustrates the comparison of the correlation matrix for the cdulcans target at three ground resolutions and two times of day. Note that the correlation images have similar patterns. The major differences are in the 3.0 ground resolution data (significantly fewer samples) and between the AM and PM data. Also note the variation around the apparent reflectance mean as a function of the ground resolution. This may be due to two factors. One factor is due to the amount of target reflectance variation being averaged and the other factor is the sensor signal-to-noise ratio that is lower for lower sensor altitudes.

The next step was to compare the covariances quantitatively. The first attempt at this was to use the Bhattacharyya distance measure. It has both a mean and a covariance term; only the covariance term was used in the comparisons. The results for each target across the different altitudes and times of day indicated that the covariances were significantly different. There is concern that the Bhattacharyya measure may not be appropriate for this comparison when one has a large number of channels. Bhattacharyya distance always becomes larger as one adds more channels.

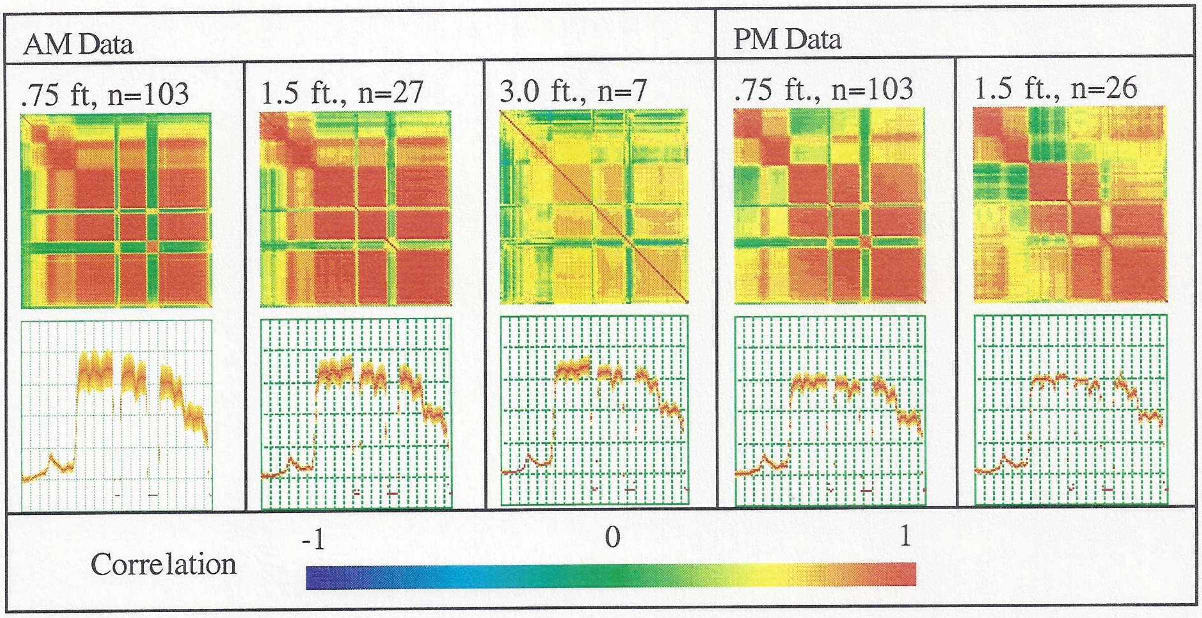


Figure 1. Illustration of correlation matrix images for cdulcans target. The y-axis scale for the apparent reflectance is –5 to 55%. The shading above and below the apparent reflectance value represents plus and minus one standard deviation. The ground resolution and number of samples for the target is given above each image.

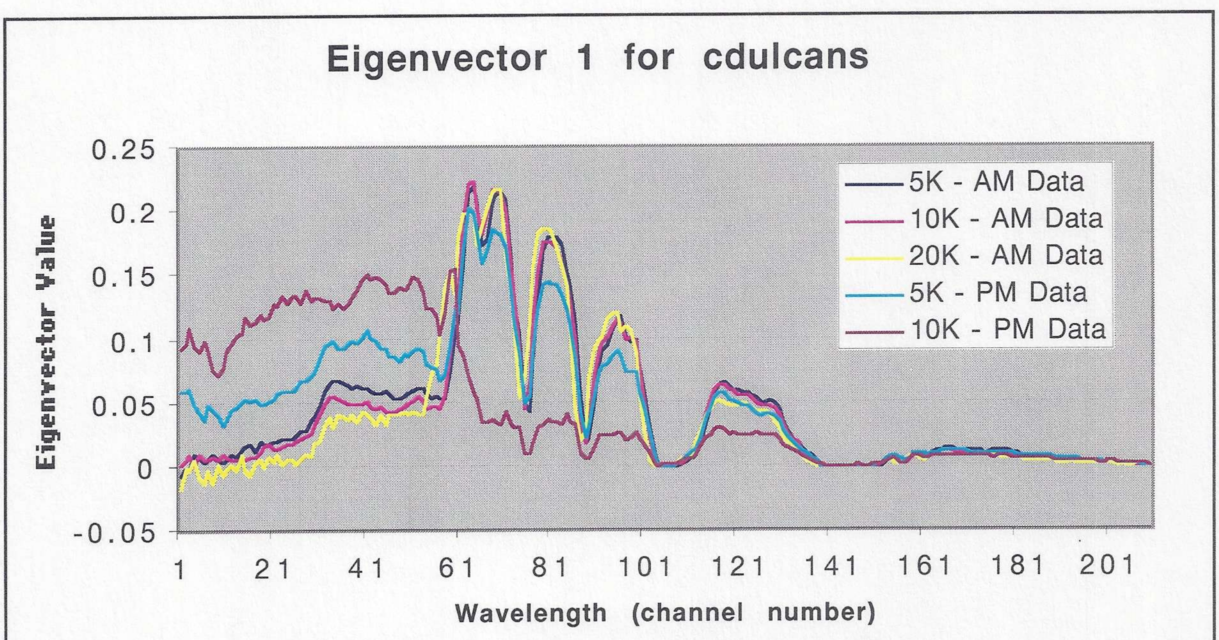


Figure 2. Comparison of the first eigenvector for the covariance matrices.

The covariance matrices were then compared using the respective eigenvectors of the covariances. The comparison for eigenvector 1 is illustrated in Figure 2. Note that the largest differences are between the AM and PM data.

Conclusions

These results indicate that even though the correlation images have similar patterns across different ground resolutions and times of day (sun angles), there may be significant difference between them, particularly the AM and PM data. This would imply that one will need to be careful how to interpret the results in modeling efforts which use the statistics for one ground resolution and time of day to another one.

Recommendations

Another step that needs to be done to compare the covariance matrix estimates is to run some classifications (or modeling trials) using the statistics generated from one ground resolution/time of day at other ground resolutions and times of day.

Also, some preliminary work has been done on significance testing relative to class statistics determination. This material would appear to have possible application to the problem of small, low probability targets. (Jeon)

References

Hoffbeck, Joseph P. and David A. Landgrebe, "Covariance Matrix Estimation and Classification with Limited Training Data," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, no. 7, pp. 763-767, July 1996. (This paper is on the web at: http://dynamo.ecn.purdue.edu/~landgreb/LOOC%20PAMI-web.pdf

J.H. Friedman, "Regularized Discriminant Analysis," Journal of the American Statistical Association, vol. 84, pp. 165-175, March 1989.

3. Quantify the Affects of Using Low Numbers of Training Samples in the Leave-One-Out Covariance Estimator.

Analysis Objective

The overall objective of this effort is to support MIT/Lincoln Laboratory in their development of the Forecasting and Analysis of Spectroradiometric System Performance (FASSP) Model.

The more specific objective is to assist with analysis of and evaluate the role of second order statistics in the surveillance, reconnaissance and abundance estimation of targets within a background.

Statement of the Problem

The overall task is to determine appropriate and robust method(s) to estimate the second order statistics for small, low probability target spectral classes.

Technical Objectives

The objective reported on here was to quantify the affects of using low numbers of training samples in the Leave-one-out covariance (LOOC) estimator for developing the statistics for target classes with a low number of available spectral samples.

Background

Some algorithms, such as the Gaussian Maximum Likelihood classifier, which can be used for the detection and/or identification of targets within a scene require the use of the second order statistics, e.g. the covariance matrix, to help characterize the target in addition to the mean. Also, models such as FASSP require the second order statistics of targets to predict the performance of algorithms even though the algorithm may not use the covariance matrix directly. However, many times where the number of spectral features is large and the targets are small relative to the spatial resolution of the imagery, the number of samples available to make a good estimate of the covariance matrix is small.

As reported on in the 1998 Program Review, the LOOC estimator can be used to estimate the covariance matrix when the number of samples available is less than the minimum required.

[Hoffbeck, 1996]. The absolute minimum number of samples needed for a sample class covariance matrix is p+1 samples for p-dimensional data. For the LOOC estimator, in theory, as few as 3 sample are all that are needed.

The LOOC estimator employs a mixing parameter to select an appropriate mixture of the (a) sample covariance, (b) diagonal sample covariance, (c) common covariance, and (d) the diagonal common covariance. The mixture deemed appropriate is the one that achieves the best fit to the training samples in the sense that it maximizes the average likelihood of training samples that were not used in the estimate. This covariance estimator is typically non-singular when at least three samples are available regardless of the dimensions of the data, and so it can be used even when the sample covariance or common covariance estimates are singular.

However, what are the affects of using this low number in practice? An experiment was conducted to measure what the affect may be.

Methodology/Approach

(1) Samples were selected for ten classes in a June 1992 hyperspectral AVIRIS scene over agricultural crops. Some of the ten classes such as "Soybeans-Drilled" and "Soybeans-North/South Rows" are very similar spectrally. Other classes such as "Wheat" are relatively easy to separate with the spectral data available. The AVIRIS scene contained 220 channels of data from .4 to 2.4 um.
(2) Random selections of the above samples were used for each of .02n, .03n, .04n, .05n, .1n, .2n, .3n, .4n, .5n, 1n, 2n, and 3n training samples, where n, the number of channels, was 220. For some classes there were only 1n training samples available.
(3) The means and LOO Covariances were computed for each of the sets of training samples selected in step 2.
(4) All samples selected in step 1 were classified using the maximum likelihood classifier and the statistics generated in step 3.
(5) Steps 2, 3 & 4 were run ten times. In other words the experiment was run ten times with a different set of training samples for each time.

Results

Figure 1 illustrates the results for four of the ten classes with .02n to 1n training samples where n is 220 channels (4 to 220 samples).

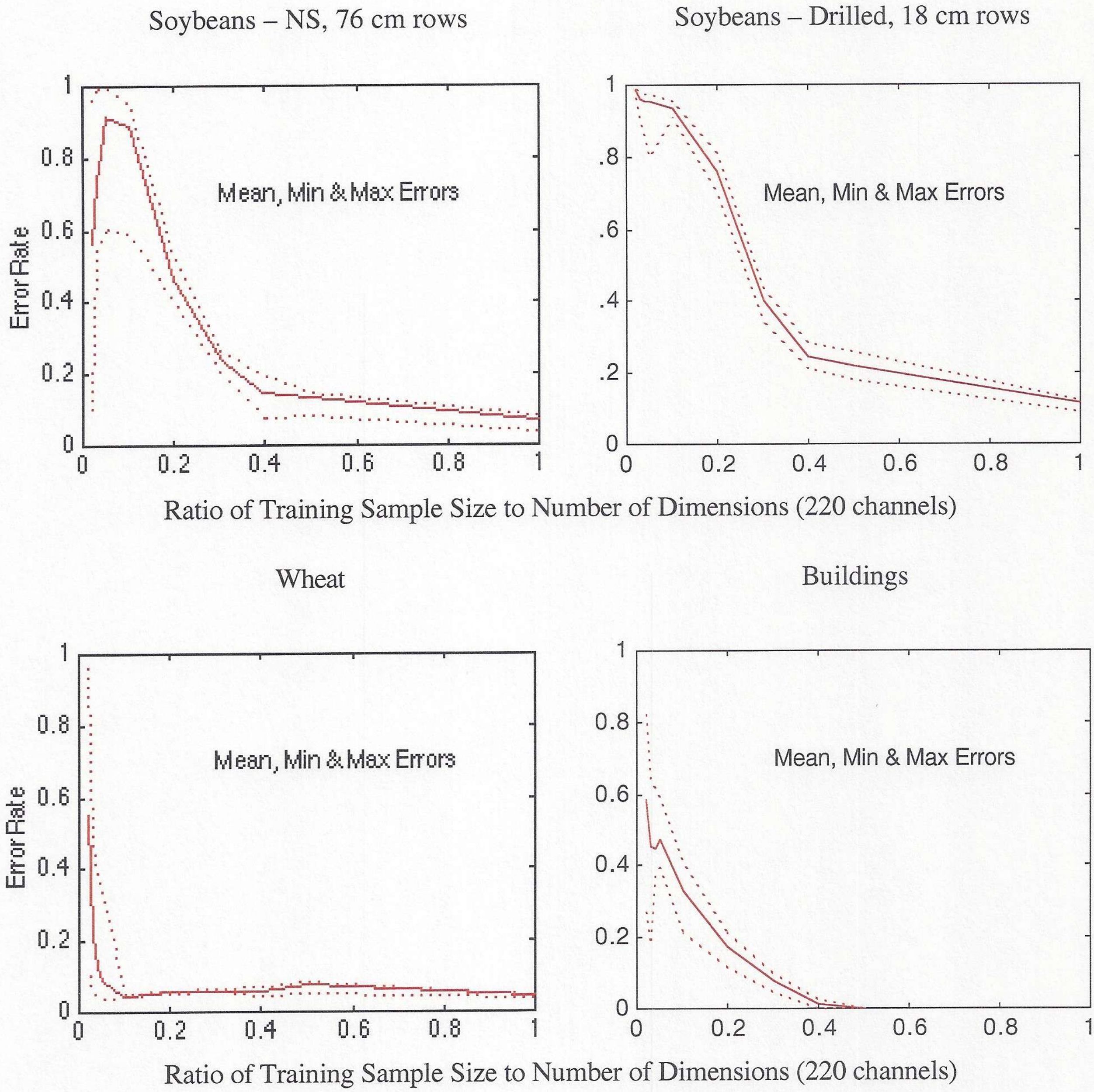


Figure 1. Illustration of the error rate found for four classes using .02n, .03n, .04n, .05n, .1n, .2n, .3n, .4n, .5n, 1n sample for training where n was 220 channels. The solid line represents the mean error rate of the ten experiments that were run. The dotted lines represent the minimum and maximum error rate found.

Note some trends that are apparent in Figure 1:

1. The error rate goes down as the number of training samples increases.
2. The variation in the error rates for a given number of training samples becomes lower as the number of training samples increase.
3. There is an inflection point in the error rate curve for the two spectrally similar soybean classes around .4n above which the error rate drops off more slowly. The inflection point for the spectrally different Wheat class is around .1n.

Conclusions

The classification of easy to separate classes (Wheat) began to breakdown, i.e. the error rate and experimental variation increased significantly, when fewer than 0.1n samples were used. The classification of difficult to separate classes (Soybeans NS and Soybeans-Drilled) began to breakdown when fewer than 0.4n samples were used. These results indicate that for a case with 220 spectral channels, one would want to have at least 20-90 samples of data to estimate the statistics for a class using the LOO Covariance estimator. Twenty samples may be sufficient if the class is spectrally different from all of the other classes in the scene. However, if there are classes that are spectrally similar then one would want to have at lease 90 samples.

Note one should not use 20 or 90 as "magic" numbers for 220 channel spectral data. These represent the results for one scene and set of experiments. However, the significant conclusion should be that 3 samples are too few for a robust estimate using the LOO Covariance estimator. Also one needs more training samples when classes within the scene are spectrally similar to each other.

Recommendations

This type of experiment should be run for other scenes and also with other classifiers.

References

Hoffbeck, Joseph P. and David A. Landgrebe, "Covariance Matrix Estimation and Classification with Limited Training Data," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, no. 7, pp. 763-767, July 1996. (This paper is on the web at: http://dynamo.ecn.purdue.edu/~landgreb/LOOC%20PAMI-web.pdf)

4. Create First and Second Order Statistics for John Kerekes of MIT/Lincoln Laboratory for Use in the FASSP Model.

Analysis Objective

The overall objective of this effort is to support MIT/Lincoln Laboratory in their development of the Forecasting and Analysis of Spectroradiometric System Performance (FASSP) Model.

Statement of the Problem

The FASSP model requires the first and second order statistics of background and target classes to predict an algorithm's performance using the HYDICE data.

Technical Objectives

The specific objective is to develop the first and second order statistics for representative background and target classes in the Forest Radiance I and Desert Radiance II HYDICE data.

Background

The statistics required for FASSP need to represent individual spectral classes that have Gaussian distributions. Therefore there may be more than one spectral class of grass, forest, etc.

Methodology/Approach

For the target statistics, the center pixel target masks developed by SITAC were used to identify the pixels to use to compute the statistics for the targets. The Leave-One-Out Covariance Estimator (described in section 2) was used since the number of pixels available for many of the targets were fewer than the number of HYDICE channels.

For the Forest Radiance I background statistics, the approach that was taken to generate the statistics was as follows:

1. Select bands outside of major water absorption regions (6-103, 110-138, 153-200); then subset as every 5th band to reduce to 35 bands. (Based on very high gain values in radiance to reflectance transform parameters and observation that the image for each channel contains at least some spatially coherent areas).

2. Select background classes. In this case six background classes were selected - soil, road, trees, grass, shaded road, shaded trees.

3. Calculate the initial mean and covariance for each class. This has been done in more than one way.

   a. For the 5,000-foot HYDICE data, selected areas within the flightline representing the above classes were used.

   b. In another experiment, the spectrometer data were used as the initial training set. In this case the diagonal of covariance matrix of spectrometer data was used as the covariance, i.e. no off-diagonal terms were used. Also the flat field correction that was given with the HYDICE to covert from radiance to apparent reflectance was used in the inverse direction to convert the spectrometer reflectance data to "apparent radiance" for use with the radiance version of the HYDICE data.

   c. In the 10,000 and 20,000 foot HYDICE data, the initial training areas were selected from unsupervised cluster maps of the areas representing the same classes as those used in the 5,000 foot data as near as possible. This was done so that the statistics found at each altitude could be compared with the theory developed by Bijan Mobasseri.

4. Assign all pixels to nearest class using a maximum likelihood classifier if the spectral threshold is less than "1/number_classes ". Otherwise treat the pixel as an unlabeled sample. This is implemented by using all pixels whose Mahalanobis distance, the exponent of the likelihood value, are less than that represented by the chi-squared statistic with P=1-1/number_classes and the degrees of freedom equal to the number of channels used in the classification. For this example, the threshold is 1/6 or .17 (P=.83) and the number of degrees of freedom is 35.

5. Go through all labeled samples and test the spatial information. If majority of pixels around belong to another class, change class ownership as long as spectral threshold is met. (If the spectral threshold is not met, assign the pixel to the unlabeled set.)

6. Scan the unlabeled samples. If surrounded spatially by labeled samples and within 3 standard deviations for each channel for the class, assign to that class. The standard deviation used is from the majority labeled pixels. The step is called checking for local homogeneity.

7. Re-compute the mean and covariance for the defined classes using the labeled samples.

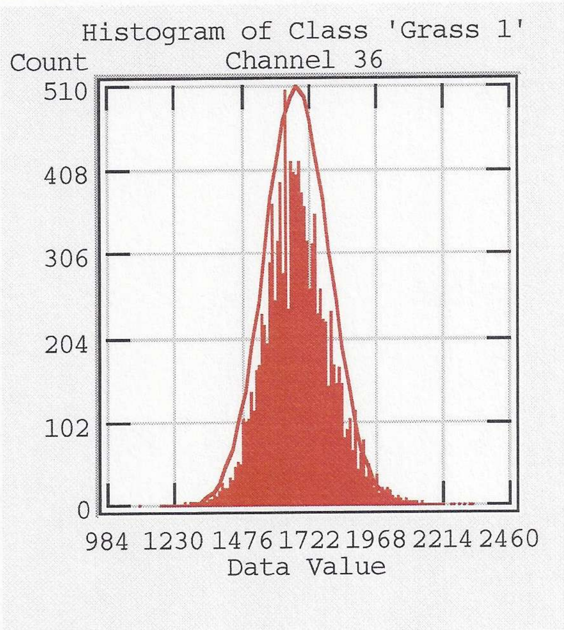
8. Loop to step five 2 or 3 times with increasing spectral thresholds of for example 0.95, 0.98 and 0.99.

9. The unlabeled samples may represent potential targets or other background classes that were not identified initially.

10. Examine the histograms of the channels for each of the resulting classes using all the pixels assigned to that class. If a histogram is definitely multimodal, then determine the data value that represents the mid-point that separates the distributions and divide the class into two sets of pixels. Use this new class, go back to step 5 and make one pass using the last spectral threshold that was used.

11. Now redo step 10 looking for other obvious multimodal distributions. Using this process in the 5,000 foot altitude data, we ended up with 4 classes of "Road", 4 classes of grass of which one we relabeled as "Bushes" and 2 classes of "Trees". We are ending up with fewer subclasses in the 10,000 and 20,000-foot data sets.

12. The final check before we release the statistics is to view the histograms of the transformed data for each spectral class to verify that the statistics unimodal. This check has indicated that the data is well represented by Gaussian distributions for each class. The density function that has been plotted on the distribution matches the actual distribution very closely for nearly all features. See Figure 1 for an example.



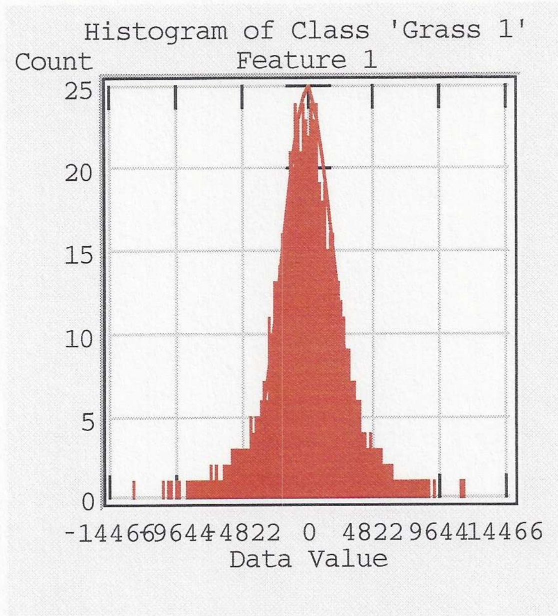


Figure 1. Example of histograms of channel 36 and 'uncorrelated' feature 1 for Grass 1 class for 5,000 foot HYDICE data.

The procedure above that we use has been coded using Matlab scripts.

For the Desert Radiance II background statistics, the approach that was taken to generate the statistics was as follows:

1. Cluster the image using an ISODATA cluster technique with 50 initial clusters and every other channel of channel set (6-101, 111-136, 154-198).

2. Evaluate the cluster map and select training samples for a Vegetation class since this class was not well represented within the cluster results.

3. Classify the image using the cluster statistics and those for the vegetation map using the maximum likelihood classifier and all of the channels in the channel set. Those cluster classes with fewer training samples than the number of channels were not used. The classification probability map was also generated.

4. Threshold all pixels in the classification map that had a probability of less than 40 percent and use this thresholded file as the training pixels for the revised spectral classes.

5. Classify the image again using the above training statistics along with a classification probability map.

6. Threshold those pixels in the classification map that have a probability of less than 4 percent and use these as the training pixels for the final classes.

7. Label the cluster classes based on the image and the information available in the pictures of the HYDICE ground collection reports.

8. This step is that same as step 12 above. The histograms of the transformed data for each spectral class were reviewed to verify that the statistics represented unimodal distributions.

Results

Statistics have been generated for the following background classes in the Desert Radiance II data set – 950626, line 06 (10,000 ft.).

Vegetation
Packed Dirt Road
Dark Gray Soil
Medium Dark Gray Soil
Gray Soil 1
Gray Soil 2
Gray Soil 3
Gray Soil 4
Gray Soil 5
Light Gray Soil
Light Brown Soil 2
Light Soil
Light Soil 2

Statistics have been generated for the following background classes in the Forest Radiance I data sets – 950824, line 05 (5,000 ft a.m.), line 07 (10,000 ft. a.m.), line 09 (20,000 ft. a.m.).

Road 1
Road 2
Road 3
Road 4
Shaded Road
Grass 1
Grass 2
Grass 3
Trees 1
Trees 2
Shaded Trees
Bushes

Statistics have been generated for the 56 target classes in the Forest Radiance I data sets – 950824, line 05 (5,000 ft a.m.), line 07 (10,000 ft. a.m.), line 09 (20,000 ft. a.m.).

The statistics were provided to MIT/LL in a predefined ASCII file format that includes the number of samples used, the wavelength description for the flight line, the log determinant that was computed with the full numeric resolution, the gain and offsets that were used to convert the data from radiance to reflectance, the means and the lower right triangular covariance matrix. The means and covariances are provided with seven decimal digits of precision.

5. Ground Reference Report for Indiana HYDICE Flight Lines HYDICE Data Collected on August 7, 1997

1. General Description of Flight Lines

Three flight lines were identified in Indiana as representative of corn/soybeans agricultural backgrounds. One, termed the Davis Farm Line, is in the east central Indiana along state road 1 north of Farmland. This flight line includes the Davis Purdue Agricultural Research Center. The second, termed the Agronomy Farm Line, is west of West Lafayette and includes the Purdue Agronomy Research Center. The third, termed the Martell Forest Line, is also west of West Lafayette and includes the Purdue Martell Forest Station. All three of the flight lines are oriented north-south. See Figures 1 - 3 for the location of the flight lines on USGS quad sheets.

The start and end coordinates (latitude and longitude) of the flight line centers are:

Davis Farm (flight line is 5.1 miles long)

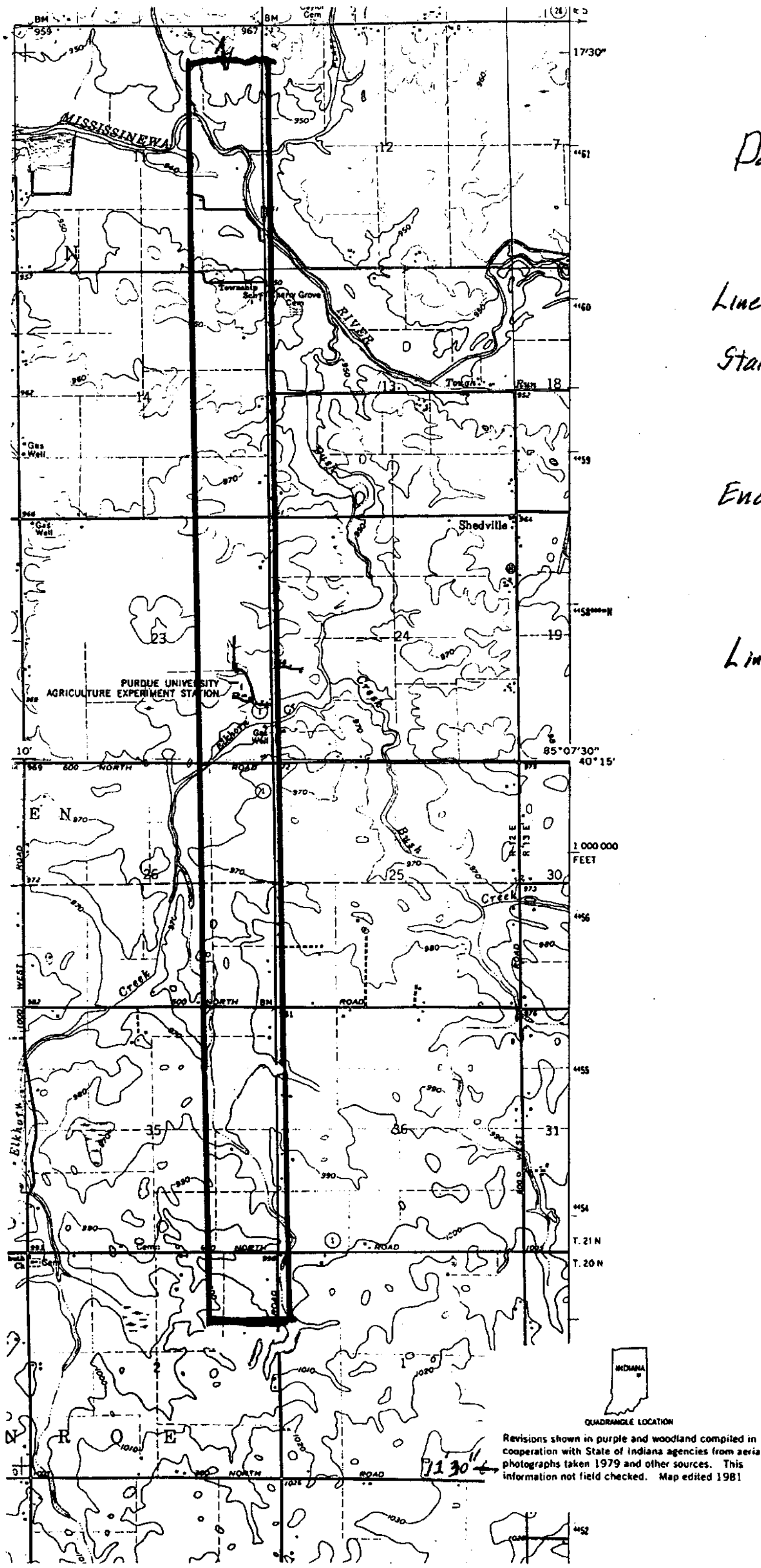
Start:	40° 13' 00" N	End:	40° 17' 30" N
	85° 09' 00" W		85° 09' 00" W

Agronomy Farm (flight line is 4.0 miles long)

Start:	40° 27' 30" N	End:	40° 31' 00" N
	87° 00' 00" W		87° 00' 00" W

Martell Forest (flight line is 5.1 miles long)

Start:	40° 22' 30" N	End:	40° 27' 00" N
	87° 02' 09" W		87° 02' 09" W



7/28/97

Davis Research Farm

Line Center

Start: 40° 17' 30" N
85° 09' 00" W

End 40° 13' 00" N
85° 09' 00" W

Line = 5.1 miles long

Figure 1.

7/28/97

Purdue Agronomy Farm

Line Center

Start

40° 27' 30"N

87° 00' 00"W

End

40° 31' 00"N

° 00' 00"W

Line is 4.0 miles long

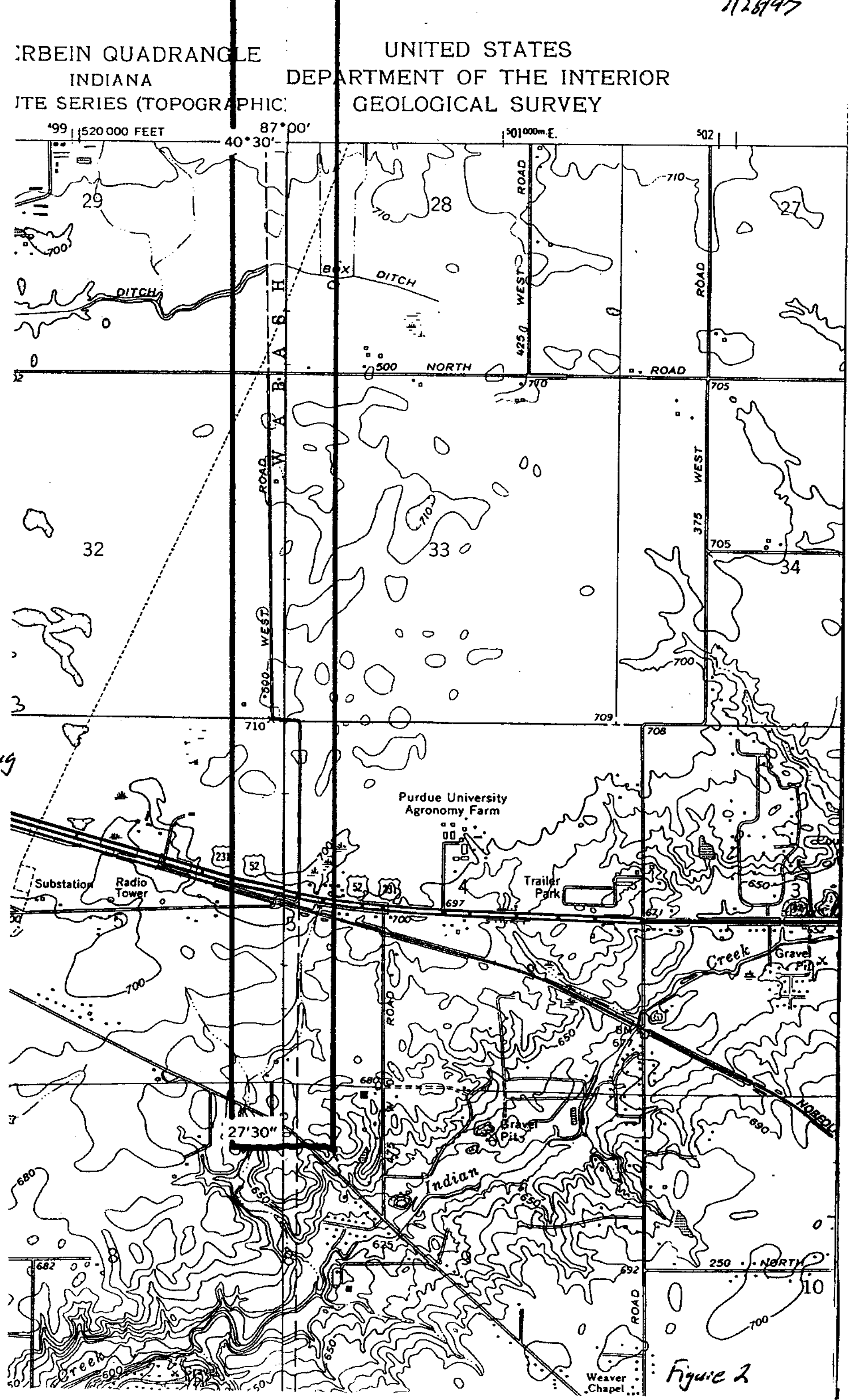
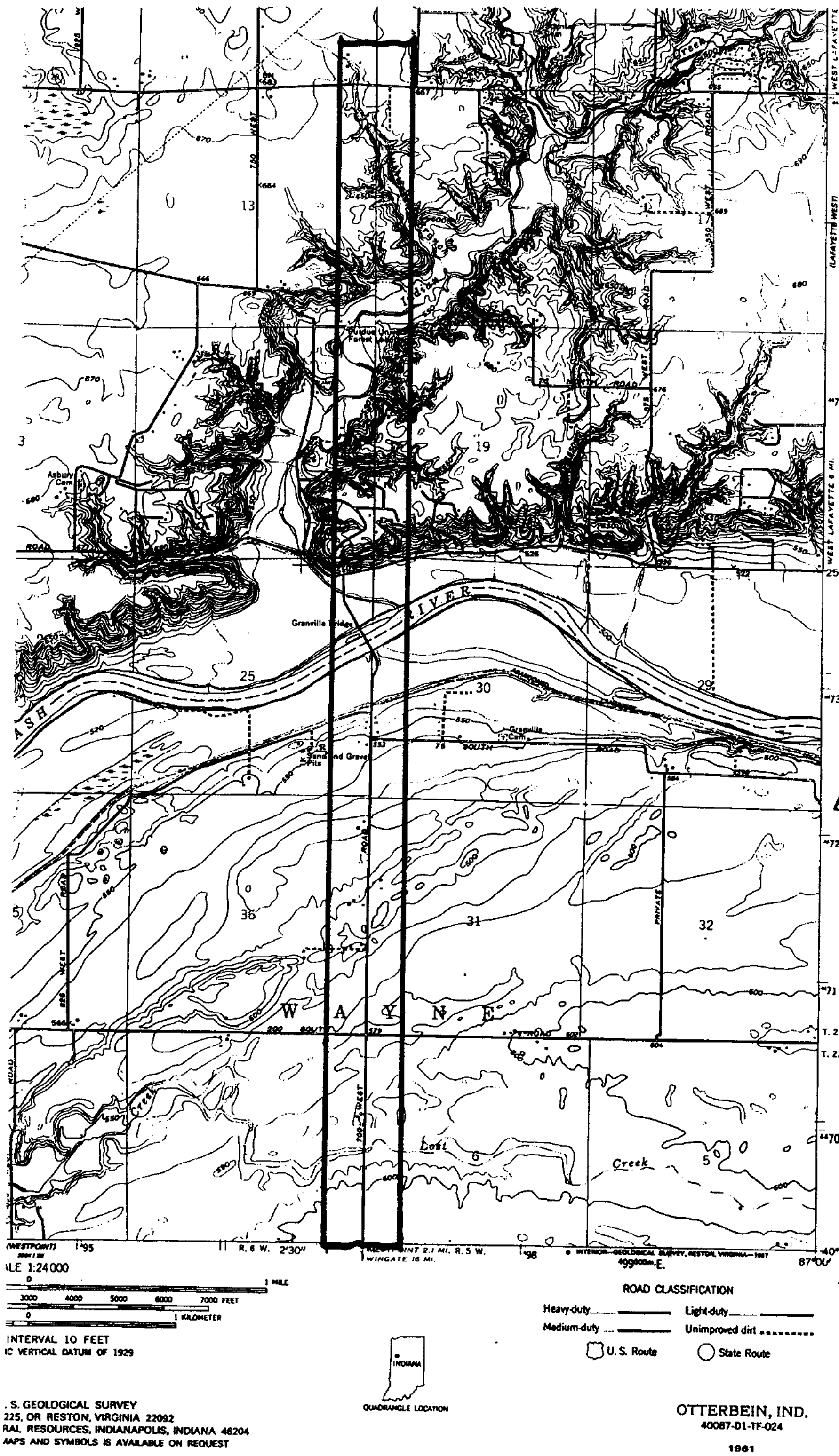


Figure 2

7/28/97



Purdue
Martell
Forest
Start:
40° 27' 00"N
87° 02' 09"W

End
40° 22' 30"N
87° 02' 09"W

Line 3 5.1 miles
long

Figure 3

2. Ground Data Collected

A. Spectral Data

Terry Hemmer of SITAC came to Purdue on August 4th with an ASD spectrometer system and two ground reflectance reference panels to use to support the HYDICE flights. A site was prepared for the reference panels and a boom was made to be used on the top of a van to mount the spectrometer to allow efficient spectrometer data collection over the panels and 'short' scenes such as grass, soybeans, soil, etc. Arrangements were made for a boom truck to be available during the morning of August 7 to collect spectrometer data over 'tall' scenes such as corn canopies. Several hundred observations were collected over the following canopies/targets during August 7 and 8:

August 7
- Corn (14:35 - 17:15 GMT)
- HYDICE black reflectance reference panel (21:15 - 21:25 GMT)
- HYDICE white reflectance reference panel (21:15 - 21:25 GMT)
- Soybeans (21:25 - 21:30 GMT)

August 8 (14:23 - 16:30 GMT)
- Soybeans
- Alfalfa
- Wheat stubble
- Foxtail
- Mowed grass,
- Soil
- Gravel road

All of the ASD spectral measurements were made in the Agronomy Farm flight line.

A 30 cm. square Spectralon reflectance standard and a 60 cm square painted Barium Sulfate reflectance standard were used to calibrate the measurements to reflectance. The Barium Sulfate reflectance standard was calibrated against the Spectralon standard. The Barium Sulfate reflectance standard was much easier to use with the boom truck data collection operation. Measurements were made of the reflectance standards every 10-12 minutes. The reflectance measurements can be linearly interpolated between these time to have good estimates from the reflectance standard at the time that the canopy/target measurements were collected.

Boom Truck Data Collection. The ASD spectrometer system was around 9 meters above the ground when making measurements over the corn. The corn canopy was 1.8-2.1 meters tall. The ASD system was 1 meter above the reflectance standards when making measurements of them. A platform with 2 lab jacks was used to level the reflectance standards. The bottom of the bucket of the boom was covered with a black cloth to reduce any reflection interactions between the bucket and the reflectance standards. The sky was clear during the data collection over the corn until 16:45 GMT. Cumulus continued to build until data collection was stopped by 17:30 because of the varying illumination conditions. The size of the field of view (5 degrees) of the ASD system was .8 meters over the corn canopy and 9 cm. over the reflectance standards. Measurements were made at several locations within each corn plot to account for the variations within the canopy and the small size of the field of view.

Van Data Collection. A wooden boom was mounted on top of a full size van to position the ASD system 2 meters above the ground and 1.9 meters away from the side of the van. This system was used to make the reflectance measurements over the HYDICE reflectance reference panels and all non-corn canopies/targets. A portable platform with two lab jacks was used to level and position the reflectance standards approximately 1 meter below the ASD system when making reflectance standard measurements. The size of the field of view (5 degrees) of the ASD system was 18 cm over the canopies/targets and 9 cm. over the reflectance standards. Measurements were made at 5 to 10 locations over each canopy/target.

B. Meteorological Measurements

A chart of the total incidence for the day was recorded in the north end of the Martell Forest flight line. This chart is given in Figure 4.

The plan was to obtain measurements to compute the optical depth during the time of the aircraft over flight. These measurements were not collected because they required steady illumination conditions before and after the flight that we did not have. There were many clouds passing over and near the sun near the time of the HYDICE over flight.

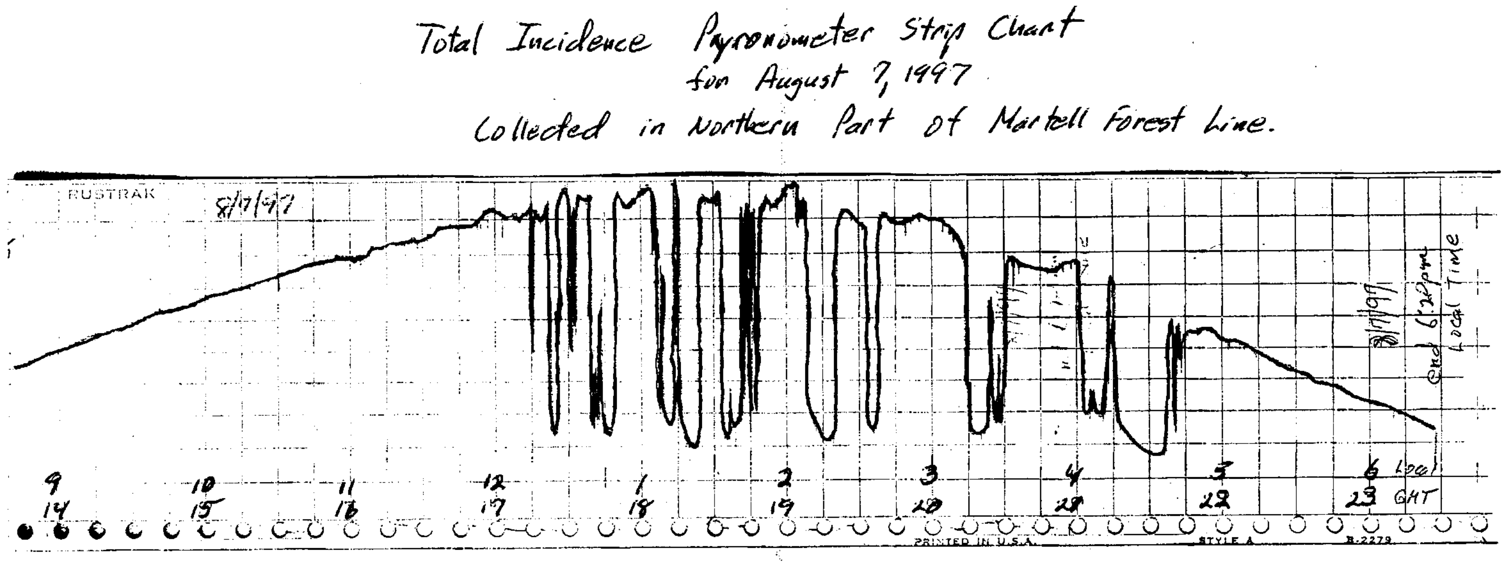


Figure 4. Total Incidence Pyronometer data.

3. Ground Reference Maps

Larry Biehl, Paul Carter and Patrick Willis have collected information about the ground cover for the three flight lines between July 31 and August 18.

The full report contains maps, pictures and annotations of the cover types for many fields within the HYDICE flightline. The entire report, 42 pages, can be obtained from Larry Biehl.