LARS Technical Report 041279

Evaluation of Several Schemes for Classification of Remotely Sensed Data: Their Parameters and Performance

Donna Scholz
Nancy Fuhs
Marilyn Hixson
Tsuyoshi Akiyama

Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana 47906 USA
1979

8.0-10105

LARS Technical Report 041279

NASA CR-
160478

Evaluation of Several Schemes for Classification of Remotely Sensed Data: Their Parameters and Performance

Donna Scholz
Nancy Fuhs
Marilyn Hixson
Tsuyoshi Akiyama

(E80-10105) EVALUATION OF SEVERAL SCHEMES FOR CLASSIFICATION OF REMOTELY SENSED DATA: THEIR PARAMETERS AND PERFORMANCE (Purdue Univ.) 14 p HC A02/MF A01 CSCL 05B N80-23732 Unclas G3/43 00105

Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana 47906 USA
1979

EVALUATION OF SEVERAL SCHEMES FOR CLASSIFICATION OF REMOTELY SENSED DATA: THEIR PARAMETERS AND PERFORMANCE

Donna Scholz, Nancy Fuhs, and Marilyn Hixson

Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana

and

Tsuyoshi Akiyama[1]
Ministry of Agriculture, Forestry, and Fishery
Nishinasuno, Tochigi, Japan

ABSTRACT

Several approaches to machine analysis of remotely sensed data have been developed over the past decade, and the remote sensing data analyst is faced with selecting which analysis approach might perform best for a given problem. The overall objective of this study was to apply and evaluate several currently available classification schemes for crop identification. The approaches examined were: (1) per point Gaussian maximum likelihood classifier, (2) per point sum of normal densities classifier, (3) per point linear classifier, (4) per point Gaussian maximum likelihood decision tree classifier, and (5) texture sensitive per field Gaussian maximum likelihood classifier.

Seven agricultural data sets were used in the study and were selected to sample variability in soils, climate, and agricultural practices. Five data sets were from the U.S. Corn Belt, and two were from the U.S. Great Plains.

Test site location and classifier both had significant effects on classification accuracy of small grains; classifiers did not differ significantly in overall accuracy. The majority of the difference among classifiers was attributed to training method rather than to the classification algorithm applied. The complexity of use and computer costs for the classifiers also varied significantly.

---

[1]Dr. Akiyama was a Visiting Scientist at the Laboratory for Applications of Remote Sensing at the time this work was conducted.

I. INTRODUCTION

Over the past several years, the potential utility of remotely sensed data to survey and monitor agricultural crops and soils has been increasingly recognized. The use of a per point maximum likelihood classifier in the Corn Blight Watch Experiment during 1971 was the first attempt at large scale application of digital classification of remotely sensed multispectral data [1]. This was followed in 1973-75 by the Crop Identification Technology Assessment for Remote Sensing (CITARS) project for corn and soybeans in Indiana and Illinois using Landsat MSS data [2]. Since then extensive research has been devoted to wheat inventory with the Large Area Crop Inventory Experiment (LACIE) during 1973-78 [3]. Currently, interest has been directed toward analysis of multicrop areas, with corn and soybeans being the major crops of interest.

To support these efforts utilizing satellite remotely sensed data, several numerical analysis schemes have been developed and implemented at numerous university, business, and government facilities in the United States and abroad. Some methods developed were successful for identification of agricultural crops, while others were not as successful. The remote sensing data analyst, therefore, must determine which analysis approach or algorithm might perform best for a given problem. Numerous studies have evaluated the performance of a given classifier, but relatively few studies have objectively compared the performance of several approaches for a given problem.

The overall objective of this study was to apply several currently available classification schemes and to evaluate their performance on several agricultural data sets. The data sets were selected to include corn and soybeans; winter wheat; and spring wheat as the major crops. Classification accuracy for test fields, ease of analyst use, and computer time required were compared for the different classifiers and data sets.

II. EXPERIMENTAL APPROACH

Test sites were selected from three major data sets: CITARS data from 1973 over Illinois and Indiana [2]; LACIE data from 1976 over the U.S. Great Plains [3]; and multicrop data from 1978 [4]. An 8 x 24 kilometer (5 x 15 mile) area in Fayette County in south central Illinois was used from the CITARS data set. A 9.3 x 11.1 kilometer (5 x 6 nautical mile) area was selected in each of Foster County, North Dakota and Grant County, Kansas, from the LACIE data. Four segments, each 9.3 x 11.1 km, were selected from the multicrop data: Pottawattamie and Shelby Counties in west central Iowa, Tippecanoe County in west central Indiana, and Iroquois County in east central Illinois.

The segments sample several major crops: winter wheat in Kansas: spring wheat in North Dakota; and corn and soybeans in Indiana, Illinois and Iowa. The Corn Belt segments were located in two distinct regions to sample variability in soils, climate, and agricultural practices. Both areas are intensively cropped, with corn and soybeans being the predominant agricultural crops. Ground reference data and field maps as well as cloud-free multitemporally registered digital Landsat MSS data were available over these sites.

Four acquisition dates were selected for analysis from the most cloud-free, least noisy, and best registered acquisitions which temporally sampled the crop calendar to maximize crop development differences (Table 1). For the Corn Belt segments, an attempt was made to obtain a spring acquisition to better separate winter small grains, trees, and permanent pasture from row crops. An acquisition after corn had tasseled was included to separate corn and soybeans.

Since classification costs would be too high if all 16 bands of data were used, classifications were performed using four bands selected to maximize the average transformed divergence between pairs of classes [5]. The acquisition dates and spectral bands selected are shown in Table 2.

Five classifiers, implemented on an IBM 370/148 computer at the Laboratory for Applications of Remote Sensing (LARS), Purdue University, were selected for study:

1. CLASSIFYPOINTS is a per point Gaussian maximum likelihood classifier. It is a processor in LARSYS, a remote sensing data analysis system developed at LARS [6].

2. CLASSIFY implements a sum of normal densities maximum likelihood classification rule which first assigns each pixel into an information category and then assigns the pixel to a spectral subclass within that category. It is a processor in EODLARSYS, developed at NASA/Johnson Space Center [7].

3. MINIMUM DISTANCE is a linear classification rule which assigns each pixel to the class whose mean is closest in Euclidean distance [8]. It is a processor in LARSYS.

4. The LAYERED classifier is a multistage decision procedure [9]. It utilizes decision tree logic with an optimum subset of features at each tree node to classify each pixel using a Gaussian maximum likelihood decision rule. LAYERED is also a processor in LARSYS.

5. ECHO (Extraction and Classification of Homogeneous Objects) utilizes both spectral and local spatial information [10]. Statistical tests are used to segment the image into homogeneous regions and each region is then classified using a Gaussian maximum likelihood sample classification rule. It was also developed at LARS and is part of LARSYS.

In order to insure that differences in classification accuracies were the result of classifier differences and not training methods, the same set of training statistics was used for all classifiers. Training fields were selected to represent the classes of interest: corn, soybeans, and others in the Corn Belt segments and small grains and others in the Great Plains segments. These fields were clustered to develop means and

Table 1a. Multitemporal Data Set Composition for the Corn and Soybean Test Sites.

Corn Development Stage	Test Site				
	Fayette	Pottawattamie	Shelby	Tippecanoe	Iroquois
	Date of Landsat Acquisition				
Emergence	6/10	6/16	6/16	6/10	6/12
Pretassel	6/29, 7/17	-	-	-	-
Tasseling	8/21	7/23	7/23	7/26	8/5
Blister	-	-	8/9	-	-
Dough	-	-	-	8/21	-
Dent	-	9/6	-	-	8/31
Mature	-	9/24	9/24	9/26	9/28

Table 1b. Multitemporal Data Set Composition for the Spring and Winter Wheat Test Sites.

Wheat Development Stage	Test Sites	
	Grant	Foster
	Dates of Landsat Acquisition	
Emergence	3/13	5/26
Heading	5/15	6/30
Soft Dough	6/2	7/19
Harvest	7/8	8/24

Table 2. Spectral Bands Used in Classification.

Test Site	Landsat Acquisition Date	Spectral Bands Selected
		μm
Fayette	6/10	.6-.7
	6/29	None
	7/17	.6-.7, .8-1.1
	8/21	.6-.7
Pottawattamie	6/16	.8-1.1
	7/23	.6-.7, .8-1.1
	9/6	.7-.8
	9/24	None
Shelby	6/16	.6-.7
	7/23	.8-1.1
	8/9	.8-1.1
	9/24	.8-1.1
Tippecanoe	6/10	.6-.7, .7-.8
	7/26	.8-1.1
	8/21	.7-.8
	9/26	None
Iroquois	6/12	.7-.8
	8/15	.8-1.1
	8/31	.8-1.1
	9/28	.6-.7
Grant	3/13	.8-1.1
	5/15	.6-.7
	6/2	.6-.7
	7/8	.6-.7
Foster	5/26	.7-.8
	6/30	.7-.8
	7/19	.6-.7
	8/24	.8-1.1

covariances to define spectral subclasses for each of the classes of interest. Since CLASSIFY was designed as part of an automated analysis procedure without analyst intervention, a training method (referred to as ISOCLS) using a random selection of individual pixels to define initial cluster seeds for clustering the entire area is generally used in conjunction with that algorithm. Both training methods were used in CLASSIFY.

The Fayette County site had reference data over approximately twenty-five percent of its area, while reference data were available for the entire area for the other sites. These data were sampled to define training and test data. Half of the selected fields were used for training the classifiers, and the remaining half were set aside for testing the classification results. Training was based on 1.6% of the area in the Fayette site, and between 3.5 and 7.5% in the other sites.

III. RESULTS AND DISCUSSION

The results of the classifications (Table 3) were analyzed to assess the effects of segment and classifier on classification accuracy. Segment-to-segment variability was highly significant ($p<0.01$). Segment variability was attributed to factors other than the classifier selected, including spectral data quality and scene characteristics.

Several factors contributed to the lower classification accuracies obtained in Fayette County: (1) the quality of multitemporal registration was only marginal, (2) the acquisitions for Fayette were not as well distributed throughout the growing season as in the other counties, (3) less training data were available for the Fayette site, and (4) the training data were not as well distributed or representative.

Pottawattamie and Tippecanoe Counties had larger field sizes, accounting in part for the relatively accurate classification. Shelby County contained more confusion crops, including sorghum and spring oats, and had smaller field sizes than the other counties. Iroquois County had very few confusion crops and was almost entirely corn and soybeans. This crop distribution made it difficult to obtain training for cover types other than corn and soybeans.

There was no significant difference among classifiers in percent correct classification of corn, soybeans, or other in the five Corn Belt segments. In addition, there was no significant difference in overall accuracy using all seven segments. The sum-of-normal-densities classifier using LARSYS statistics, however, gave significantly higher small grain classification accuracy (about 2% classification improvement).

Table 4 shows the percent correctly classified averaged over all segments for the different cover types. The performance of the ECHO classifier was not as high as anticipated. This is probably due to the fact that the ECHO classifier requires the analyst to set parameters defining cell size and homogeneity factors, and the optimal settings have not been defined. Although differences were nonsignificant overall, the LARSYS training method provided a consistent improvement over the ISOCLS training method in six of the seven segments. In conclusion, given a set of training

Table 3. Comparison of Classifier Performance (Percent Correct Classification) by Test Site.

TEST SITE	CLASS	MINIMUM DISTANCE	CLASSIFY POINTS	LAYERED	ECHO	CLASSIFY Using ISOCLS Stats[1]	CLASSIFY Using LARSYS Stats[2]	TEST SITE Average
Fayette, IL								
	Corn	81.9	81.2	63.9	77.3	77.3	78.9	76.8
	Soybeans	82.0	77.0	76.8	70.7	49.7	79.0	72.5
	Other	85.5	88.6	91.3	87.8	58.8	85.6	82.9
	Overall	83.5	83.0	80.5	79.5	61.1	81.6	78.2
Pottawattamie, IA								
	Corn	98.7	97.2	95.7	98.2	93.0	98.4	96.9
	Soybeans	92.0	89.8	92.3	90.2	86.5	89.3	90.0
	Other	85.3	98.0	97.5	97.1	92.1	98.4	94.7
	Overall	94.9	94.7	94.7	95.4	90.6	95.3	94.3
Shelby, IA								
	Corn	97.1	95.1	94.5	96.1	82.8	95.9	93.6
	Soybeans	89.3	92.9	98.2	95.4	98.0	98.0	95.3
	Other	75.5	83.7	88.2	79.4	78.7	79.7	80.9
	Overall	90.0	91.7	93.3	91.5	83.9	92.1	90.4
Tippecanoe, IN								
	Corn	93.7	89.9	91.5	86.4	99.4	93.1	92.3
	Soybeans	97.6	98.2	94.9	98.0	95.1	98.4	97.0
	Other	94.3	96.7	100.0	96.7	69.9	96.7	92.4
	Overall	95.5	94.3	94.0	92.7	94.2	95.9	94.4
Iroquois, IL								
	Corn	88.1	79.5	91.0	79.3	89.9	92.8	85.1
	Soybeans	82.8	85.2	78.1	83.6	78.8	86.3	82.5
	Other	76.4	72.7	0.0	72.7	74.5	75.0	61.9
	Overall	84.9	82.1	80.5	81.2	83.6	84.2	82.8
Foster, ND								
	Small Grains	96.1	95.4	94.6	94.8	93.6	97.3	95.3
	Other	73.3	77.1	77.0	77.6	70.5	82.3	76.3
	Overall	82.7	84.7	84.3	84.8	81.3	89.3	84.5
Grant, KS								
	Small Grains	96.9	96.7	97.6	96.5	94.6	98.7	96.8
	Other	91.8	83.2	89.3	79.2	92.0	80.2	86.0
	Overall	93.1	86.5	91.4	83.5	92.6	84.8	88.6

[1] Training method generally used with CLASSIFY. Uses a random selection of individual pixels to define initial cluster seeds for clustering the entire area.

[2] Training method used with all other classifiers. Training fields were clustered to develop means and covariances to define spectral subclasses for each of the classes of interest.

Table 4. Comparison of Average Percent Correct Classification for Several Classification Approaches.

MAJOR CROPS	NO. SEGMENTS	CLASS	Classifier: MINIMUM DISTANCE	CLASSIFY POINTS	LAYERED	ECHO	CLASSIFY Using ISOCLS Stats[1]	CLASSIFY Using LARSYS Stats[2]
Corn/Soybeans	5	Corn	91.9	88.6	87.3	87.5	88.5	89.8
		Soybeans	88.7	88.6	88.1	87.6	81.6	90.2
		Other	85.4	87.9	75.4	86.7	74.8	87.1
		Overall	89.8	89.2	88.6	88.1	82.7	89.8
Small Grains	2	Small Grains	96.5	96.0	96.1	95.6	94.1	98.0
		Other	82.6	80.2	83.2	78.4	81.3	81.3
		Overall	87.9	85.6	87.8	84.2	87.0	87.0

[1] Training method generally used with CLASSIFY. Uses a random selection of individual pixels to define initial cluster seeds for clustering the entire area.

[2] Training method used with all other classifiers. Training fields were clustered to develop means and covariances to define spectral subclasses for each of the classes of interest.

statistics capable of producing high level classification results, the choice of classification algorithm for differentiation of corn and soybeans from other cover types makes relatively little difference.

Two additional comparisons of the classification schemes were considered: the ease of use of the classification method and the computer time needed for each classifier. The classification schemes varied considerably in ease of use. In increasing order of complexity the classifiers were found to be: (1) MINIMUM DISTANCE, (2) CLASSIFYPOINTS, (3) CLASSIFY, (4) ECHO, and (5) LAYERED. The MINIMUM DISTANCE and CLASSIFYPOINTS classifiers were almost identical in ease of use.

CLASSIFY was designed as part of a total analysis scheme in which participation of the analyst is minimized in the clustering and definition of training statistics with control provided by a predefined set of analysis parameters. Although the classifier itself is not extremely complex, the training procedure typically used in this scheme involves a large number of parameters about which little is known.

ECHO utilizes both temporal and spatial information. The complexity of use for ECHO arises from the necessity of setting the parameters for cell homogeneity testing and cell size. The expertise of the analyst is essential in setting the parameters with regard to data set used. The ECHO classifier is, however, one of the few available classifiers that utilize spatial as well as spectral information in the classification process.

LAYERED implements a per point Gaussian maximum likelihood decision tree logic which requires the additional step of designing the decision tree. The decision tree is designed by obtaining class means and covariance matrices for all classes and using a feature selection algorithm to determine an optimal subset of features to be used at each node of the decision tree. The time needed by the analyst to design the tree can be significant if many spectral classes and features are needed to characterize the scene of interest. Although the decision tree can become very complicated and awkward to use, this classifier is particularly well suited for use with multitemporal or multitype data sets.

Parallelling the complexity of implementation as an important variable in selecting a classification scheme is the computational cost per classification. The computer time required per square kilometer for each segment and classifier is shown in Table 5. In order of increasing cost per square kilometer for classification, not including cost for developing training statistics, were (1) MINIMUM DISTANCE (1.7 seconds), (2) ECHO (2.3 seconds), (3) LAYERED (2.3 seconds), (4) CLASSIFYPOINTS (3.8 seconds), and (5) CLASSIFY using ISOCLS statistics (11.3 seconds).

IV. CONCLUSIONS

The results of this study show little difference in the classification accuracies achieved by the five classification algorithms which were considered. However, the results for the CLASSIFY algorithm using two

Table 5. Computer CPU Time (seconds per square kilometer) Used by Each Classifier.

CLASSIFIER	TEST SITE						
	Grant	Foster	Tippecanoe	Fayette	Pottawattamie	Shelby	Iroquois
Minimum Distance	2.3	1.7	1.3	1.5	1.6	2.3	1.4
Classifypoints	6.1	3.5	2.9	3.6	2.7	3.7	3.6
Layered	3.5	2.4	1.8	3.1	1.7	1.7	2.0
ECHO	3.9	2.3	1.9	2.0	2.0	1.8	2.3
Classify(LARSYS Stats)	5.7	3.4	3.4	2.9	3.1	3.1	5.0
Classify(P-1 Stats)	10.7	12.6	12.7	8.0	12.8	8.4	14.1

different training methods did show a difference, indicating that the major variable affecting classification accuracy is not the classifier, but the training method used in generating the class statistics to be used in the classification. The most important aspect of training is that all cover types in the scene must be adequately represented by a sufficient number of samples in each spectral subclass.

The ISOCLS training algorithm was a method designed for machine automation of a large portion of the training procedure. The statistical sampling method used for selection of training data is theoretically sound, so it is possible that the lack of analyst refinement of the training statistics is seriously limiting the performance. The clusters produced by this method are of mixed cover types which may adversely affect performance.

Additional variables of interest in the study were complexity of use of the classifier and CPU cost per classification. Among the classifiers yielding similar classification accuracies, MINIMUM DISTANCE was the easiest for the analyst to use and costs the least per classification.

In conclusion, the classification performance of the five classification algorithms was found to be very similar when the same training method was utilized. The results suggest that development of representative training statistics is relatively more important for obtaining accurate classifications than selection of the classification algorithm.

V. REFERENCES

1. MacDonald, R.B., M.E. Bauer, R.D. Allen, J.W. Clifton, J.D. Erickson, and D.A. Landgrebe. 1972. Results of the 1971 Corn Blight Watch Experiment. Proc. Eighth Int. Symp. Remote Sens. Envir., Ann Arbor, Michigan, October 2-6, pp. 157-190.

2. Bizzell, R., F. Hall, A. Feiveson, M.E. Bauer, B. Davis, W. Malila and D. Rice. 1975. Results from the Crop Identification Technology Assessment for Remote Sensing (CITARS) Project. Proc. Tenth Int. Symp. Remote Sensing Envir., Ann Arbor, Michigan, October 6-10, pp. 1189-1196.

3. MacDonald, R.B., and F.G. Hall. 1978. LACIE: An Experiment in Global Crop Forecasting. Proc. Plenary Session, The LACIE Symp., Houston, Texas, October 23-26, pp. 17-48.

4. Hixson, M.M., B.J. Davis, and M.E. Bauer. 1978. Stratification and Sample Selection for Multicrop Experiments. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana. LARS Contract Report 112278.

5. Swain, Philip H., and Shirley M. Davis, ed. 1978. Remote Sensing: The Quantitative Approach. McGraw-Hill, Inc., New York.

6. Phillips, T.L., ed. 1973. LARSYS User's Manual. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.

7. Stewart, J., and P.J. Aucoin. 1978. Earth Observations Division Version of the Laboratory for Applications of Remote Sensing System (EOD-LARSYS) User Guide for the IBM 370/148 - Volume I. System Overview. NASA, Johnson Space Center, Houston, Texas JSC-13821.

8. Nilsson, N.J. 1965. Learning Machines: Foundations of Trainable Pattern-Classifying Systems. McGraw-Hill Book Co., New York.

9. Swain, P.H., C.L. Wu, D.A. Landgrebe, and H. Hauska. 1975. Layered Classification Techniques for Remote Sensing Applications. Proc. Earth Resources Survey Symp., Houston, Texas, June 9, Vol. I-b, pp. 1087-1097.

10. Kettig, R.L., and D.A. Landgrebe. 1976. Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects. IEEE Trans. Geos. Elect. 14:19-25.

VI. ACKNOLWEDGEMENTS

This work was supported in part by the NASA, Johnson Space Center, contract NAS9-15466.