AgRISTARS

DC-L1-04051
JSC-17116

Domestic Crops/
Land Cover

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

March 1981

EVALUATION OF LARGE AREA CROP ESTIMATION TECHNIQUES USING LANDSAT AND GROUND-DERIVED DATA

M. L. Amis, R. K. Lennington,
M. V. Martin, W. G. McGuire, and S. S. Shen

Lockheed Engineering and Management Services Company, Inc.
1830 NASA Road 1, Houston, Texas 77058

Lyndon B. Johnson Space Center
Houston, Texas 77058

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DC-L1-04051; JSC-17116		

4. Title and Subtitle	5. Report Date
Evaluation of Large Area Crop Estimation Techniques Using Landsat and Ground-Derived Data	March 1981
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No
M. L. Amis, R. K. Lennington, M. V. Martin, W. G. McGuire, and S. S. Shen	LEMSCO-15763
	10. Work Unit No.

9. Performing Organization Name and Address	11. Contract or Grant No
Lockheed Engineering and Management Services Company, Inc. 1830 NASA Road 1 Houston, Texas 77058	NAS 9-15800
	13. Type of Report and Period Covered

12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
National Aeronautics and Space Administration Lyndon B. Johnson Space Center Houston, Texas 77058 Tech. Mon. A.L. Gronfield	Technical Report

15. Supplementary Notes

16. Abstract

This paper describes the results of the National Aeronautics and Space Administration Domestic Crops and Land Cover Classification and Clustering study on large area crop estimation using Landsat and ground truth data. One objective was to evaluate the current crop area estimation approach of the Economics and Statistics Service of the U.S. Department of Agriculture in terms of the factors that are likely to influence the bias and variance of the estimator. A second objective was to investigate alternative procedures involving replacements for the clustering algorithm, the classifier, or the regression model used in the original U.S. Department of Agriculture procedure.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Large area crop estimation, Landsat, full frame processing, regression estimator, clustering, classification, cross-validation	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	136	

*For sale by the National Technical Information Service, Springfield, Virginia 22161

EVALUATION OF LARGE AREA CROP ESTIMATION TECHNIQUES USING LANDSAT AND GROUND-DERIVED DATA

Job Order 71-352

This report describes the activities of the Domestic Crops/Land Cover project of the AgRISTARS program.

PREPARED BY

M. L. Amis, R. K. Lennington, M. V. Martin, W. G. McGuire, and S. S. Shen

APPROVED BY

T. C. Minter, Supervisor
Techniques Development Section

J. E. Wainwright, Manager
Development and Evaluation Department

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.

Under Contract NAS 9-15800

For

Earth Resources Research Division

Space and Life Sciences Directorate

NATIONAL AERONAUTICS AND SPACE ADMINSITRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

March 1981

LEMSCO-15763

PREFACE

The Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing is a 6-year program of research, development, evaluation, and application of aerospace remote sensing for agricultural resources, which began in fiscal year 1980. This program is a cooperative effort of the National Aeronautics and Space Administration, the U.S. Agency for International Development, and the U.S. Departments of Agriculture, Commerce, and the Interior.

The work which is the subject of this document was performed within the Earth Resources Research Division, Space and Life Sciences Directorate, at the Lyndon B. Johnson Space Center, National Aeronautics and Space Administration. Under Contract NAS 9-15800, personnel of Lockheed Engineering and Management Services Company, Inc., performed the tasks which contributed to the completion of this research.

CONTENTS

TABLES

1. INTRODUCTION

1.1 BACKGROUND

This report describes the results of the Domestic Crops and Land Cover classification and clustering study on area estimation. The objectives of the study are as follows:

Task 1: To understand the current crop area estimation approach of the Economics and Statistics Service (ESS) of the U.S. Department of Agriculture (USDA) in terms of the factors that are likely to influence the bias and variance of the estimators.

Task 2: To develop and evaluate alternative clustering, classification, and regression methods that could be inserted into the current ESS estimation procedure.

Task 3: To begin studies that may lead to an improved estimation procedure.

Task 1 was intended to support Task 2 by providing a working understanding of the current ESS crop estimation approach. Such understanding is needed in designing appropriate experiments for evaluating and comparing alternative components.

Consideration of these alternative methods in Task 2 was principally motivated by two factors. First, it was believed that a more theoretically based clustering algorithm would be appropriate. In particular, the CLASSY algorithm developed at the Lyndon B. Johnson Space Center (JSC) had performed well in tests and was the candidate clustering replacement. CLASSY is an adaptive maximum likelihood clustering algorithm which models the overall data distribution as a mixture of multivariate normals. In addition to its clustering properties, CLASSY can also be used to provide direct area estimates. A second factor was the belief that the Editor procedure should ideally use independent data sets for developing the regression equation and evaluating area estimates. One way to do this would be to divide the available data into training and test portions. Alternatively, this could be accomplished by generating quasi-independent segments for regression using a jackknifing technique. The

Mean Square Error (MSE) classifier lends itself well to this use, as it makes no parametric assumptions; thus, there are fewer parameters estimated with the algorithm, implying more stable parameter estimates. It was felt that, due to this robust nature, the MSE classifier would be more extendible to an independent test set.

1.2 CURRENT USDA PROCEDURE

The current USDA acreage estimation procedure comprises registration and digitization of ground truth and raw Landsat data, development of an estimator on the portions of the area of interest for which ground truth is available, and application of that estimator to the whole area of interest. This study is concerned with the procedure used in developing the estimator.

The acreage estimation procedure involves the following steps:

a. A registered raw data set for the area of interest for which ground truth is available is selected. The data could be unitemporal or multitemporal, and, usually, both are studied. The data set is separated by crop type. It is optional to remove border pixels, poorly registered fields, poorly reported fields, and pixels with extreme spectral values relative to the rest of the crop type.

b. Each ground truth crop is clustered separately, yielding a group of clusters with known cluster labels.

c. Several options are excercised and parameters specified. Among these are:

   (1) specifying the minimum and maximum number of clusters per crop type

   (2) specifying separability of clusters in spectral space

   (3) specifying percent convergence when combining clusters

   (4) specifying a priori probabilities

   (5) seeding clusters

   (6) pooling the resulting clusters

   (7) dropping clusters with small populations

   (8) not clustering crops with small populations

d. After the training set has been clustered, it is then classified; and, for each crop, a regression is performed between the ground truth and the number of pixels classified into that crop class.

e. Based on the r^2 of the regression, the percent correctly classified, and the time available, the analyst may repeat earlier steps with different parameters or options, or may drop crop types being clustered in an effort to increase the r^2 and the percent correctly classified. Thus, for each crop, a regression estimator is obtained which will predict the amount of ground truth present in the area of interest when that area is classified.

1.3 STANDARIZED PROCEDURE

For this study, it was necessary to standardize the use of the USDA crop estimation procedure so that alternative clustering and classifying components could be evaluated. The following options and parameter values were recommended through discussions with USDA analysts:

a. Specifying a priori probabilities.

b. Clustering only crop types with a minimum of 200 pure pixels available.

c. Clustering pure pixels only.

d. Removing pixels with extreme spectral values relative to others of the same ground truth crop type.

e. Specifying a minimum and a maximum number of clusters per crop type (from 1 to 15) and a minimum cluster population (from 150 to 200 pixels).

f. Specifying separability of clusters in spectral space to be in the range of 0.6 to 0.8.

g. Specifying convergence when combining clusters to be within the range of 95 to 99 percent.

With these recommendations and a desire to choose an exact procedure, the USDA Editor software was exercised on data provided by the USDA. For this data set, a priori probabilities for each crop clustered were specified as the proportion of that crop present in the training set.

Clustering was done on pure pixels only, and pixels with extreme values were omitted (clipped) from the clustering process. One analyst performed all clipping in this study for consistency. The clipping limits used for each crop type for each part of this study can be found in appendix A. To ensure a minimum of 150 to 200 pixels per cluster, the maximum number of clusters was specified as the number of pixels divided by 100, with no more than 15 clusters allowed. The minimum number of clusters was specified to be the maximum number of clusters divided by 3. A separability of 0.75 was used and was rarely reached before the clustering procedure stopped due to reaching the specified minimum number of clusters. A convergence of 95 percent provided adequate clustering without unduly increasing the computer time.

All other options in the standardized procedure were as recommended by the USDA analysts. All files generated during this study on the USDA Editor were archived on tape and will be available for at least 1 year. A list of these files is presented in appendix E. Comparison of results from this standardized procedure with the ESS Missouri analysis results are presented in section 3.

2. DATA SET

2.1 DESCRIPTION

The Landsat data used in this study included 33 segments in northwest Missouri, each having an area of approximately 1 square mile (259 hectares). These segments were all contained in strata 10 (50 to 100 percent agriculture) and had little cloud cover. Data were available for two dates: May 14 and August 3, 1979. An additional 12 segments from the August date were available but not analyzed because of heavy cloud cover. The counties represented in the Missouri data set are listed in table 2-1. The ground truth proportions in the 33 analyzed segments are given in table 2-2.

TABLE 2-1.- SEGMENTS PER COUNTY IN MISSOURI DATA SET

County	Number of segments
Daviess	5
Harrison	5
Putnam	2
Schuyler	3
Mercer	3
Caldwell	3
Gentry	4
Sullivan	4
Linn	5
Livingston	5
Grundy	3
De Kalb	3
Total	45

TABLE 2-2.- GROUND TRUTH PROPORTIONS OVER 33 SEGMENTS

Crop	Mixed			Pure*		
	Hectares	Pixels	Proportion	Hectares	Pixels	Proportion
Corn	524.7	1582	0.117	211.5	651	0.101
Winter wheat	114.0	452	.034	62.4	192	.030
Permanent pasture	1321.5	3984	.296	821.3	2527	.392
Soybeans	1073.2	3299	.245	576.9	1773	.275
Dense woodland	355.3	1136	.084	167.0	514	.080
Other hay	304.5	964	.072	192.4	592	.092
Other	681.5	2059	.153	64.2	194	.030
Total	4374.7	13476	1.000	2031.5	6443	1.000

*Pure = Poorly registered, poorly reported, and border pixels removed.

The major crops in this study are corn, soybeans, and pasture, which represent about 12, 25, and 30 percent of the crops present in each segment, respectively. Three additional crops are also studied: winter wheat (3 percent), dense woodland (8 percent) and other hay (7 percent). About 15 percent of the segments consisted of other crops, mainly wasteland. For a given crop, the minimum number of pure pixels considered for analysis was 200. The crops lumped together as "other" had well below 200, and other hay, dense woodland, and winter wheat were marginal. The best performance (as measured by percent correctly classified) in most cases was obtained for permanent pasture, in which over 2500 pure pixels were available for training.

Listed in table 2-3 are the number of pure pixels present in ground truth in each of the 33 segments used in this study, broken down by crop type. The sample mean and sample standard deviation are also listed.

The Missouri data set provided by the USDA was available at the Bolt, Beranek, Newman (BBN) remote processing center in Boston, where the USDA Editor software also resided. The same Missouri data set was placed on tape and sent to the Laboratory for Applications of Remote Sensing (LARS) at Purdue University. Software for alternate clustering and classification used in this study was located at LARS. The Missouri data sets at both BBN and LARS were identical. The following information for each pixel was provided in the data set:

a. four channel values from May 14

b. four channel values from August 3

c. ESS crop code

d. segment number

e. tract and field identification

f. Landsat row and column

g. flag indicating a border pixel

TABLE 2-3.- PURE PIXELS OF GROUND TRUTH BY SEGMENT

Segment number	Corn	Winter wheat	Permanent pasture	Soybeans	Dense woodland	Other hay
6034	5	7	45	13	0	38
6085	0	0	165	22	0	72
6015	31	19	30	81	0	8
6038	0	0	0	0	0	0
6098	0	0	310	0	0	0
6073	27	0	41	17	77	33
9046	0	0	99	0	47	36
6064	0	7	46	76	40	0
6065	29	3	28	1	0	15
6095	0	0	0	0	0	0
9061	86	0	0	183	0	0
9036	5	3	36	125	0	0
6053	0	0	40	104	4	4
6058	12	0	0	171	2	35
9057	0	0	108	17	14	45
9037	17	26	79	67	2	2
9062	8	0	251	18	30	16
9066	0	8	0	79	0	0
6045	6	0	340	13	0	0
9047	59	43	42	152	3	0
6048	53	0	7	0	62	0
9097	6	0	12	82	1	62
9096	37	0	0	26	39	28
6040	9	2	124	22	8	55
6060	0	0	0	10	3	6
6035	51	5	84	125	0	34
9016	93	0	0	85	0	0
9051	0	35	145	9	115	0
6063	22	2	31	76	16	0
6050	34	0	182	0	0	79
9052	39	19	26	137	0	0
6059	0	0	115	18	44	14
9017	22	13	141	46	7	10
Total	651	192	2527	1775	514	592
Sample mean	19.7	5.8	76.6	53.8	15.6	17.9
Sample SD*	25.2	10.8	90.1	55.8	27.3	23.4

*Standard deviation.

2.2 PHILOSOPHY OF EXPERIMENTS

The experiments were designed to understand the performance of Editor and to compare performance when alternative components are inserted. To motivate the analysis, two main variables X and Y can be considered, where X is the crop acreage derived from the classifier (or the number of classified pixels, depending upon choice of units) and Y is the corresponding ground truth acreage for the segment. In the current method of analysis, Y is regressed onto X, as if X were a fixed variable. Also, the regression is developed on the training set instead of an independent sample. Certain subtle and often overlooked features of the classified variable X are not accounted for in the current method of analysis. In particular, the values which X assumes are a function of

a. The observed spectral values and ground truth labels of the segments used to calibrate the classifier (training set). This implies that X is a random variable since the training set is picked at random.

b. The number of observations in the training set. The larger the training set, the less sensitive the classifier is to the random selection process for picking the training set.

Now let

$$X'_1, X'_2, \cdots, X'_{N_1}$$

$$X''_1, X''_2, \cdots, X''_{N_2}$$

be two sets of classifier-derived acreages over two sets of randomly selected segments. Both sets of segments are assumed to have been picked from the same population. Ideally, Y_i regressed onto X'_i, $i = 1, \cdots, N_1$, should be "about" the same as Y_i regressed onto X''_i, $i = 1, \cdots, N_2$. The following conjectures arise:

a. If X'_i is obtained by classifying the training set and X''_i is obtained by classifying an independent set, the regressions will be different for "small" training sample sizes.

b. If X'_i is obtained by classifying the training set and X''_i is obtained by classifying an independent set, then the X'_i-values will be more closely

correlated with the Y_i-values than will be the X_i''-values. This point has bearing on the relevance of the r^2 values obtained by Editor and reported by the USDA.

c. If X_i' is obtained through a jackknifing procedure and X_i'' is obtained by classifying an independent set, then the regressions of Y_i onto X_i' and Y_i onto X_i'' will be "about" the same. This conjecture is probably classifier dependent and is one of the reasons why a linear classifier was selected for study.

2.3 DESIGN OF EXPERIMENTS

In keeping with the conjectures presented in the Philosophy of Experiments section, the design of experiments was done in three levels. The first level consisted of training and testing on all 33 segments. This corresponds to the current USDA estimation procedure. In the second level, the data set was partitioned into a training set and a test set to assess the performance and the validity of the current USDA estimation procedure. Jackknifing techniques were used in the third level as a means of obtaining independent test sets which were larger than those obtainable by using a single training-and-test partitioning of the data. This experimental design was strongly influenced by the belief that the sample of segments chosen to obtain estimates is a critical part in the whole estimation process.

Many of the experiments were run in parallel as a means of comparing alternative components. That is, the standardized USDA procedure was first run on a data set, and the procedure was repeated with the only change being the use of the CLASSY clustering algorithm to generate cluster statistics which were then inserted into the Editor system. Then a corresponding analysis was performed using the MSE classifier software. One experiment was designed for the USDA Editor to specifically evaluate one particular method of estimation, namely jackknifing. A separate jackknifing experiment for the MSE classifier was designed. The experiments are described below in further detail; and, unless explicitly stated, all analysis is with multitemporal data.

2.3.1 TRAINING AND TESTING ON ALL 33 SEGMENTS

The current USDA method of training on a sample and developing the regressions on the training set was performed using all 33 segments. The following comparisons were made:

a. Comparison of unitemporal versus multitemporal - The entire estimation process was carried out for unitemporal data and for multitemporal data within the current Editor system. Summary statistics were collected. The Hotelling's T^2 test was used to determine if multitemporal data produced significantly better estimates than unitemporal. This test is described in detail in section 3.1.2.

b. Comparison of the current standarized USDA procedure versus the CLASSY clustering algorithm - The entire estimation process was performed using the standarized USDA procedure. The process was repeated but with CLASSY cluster statistics inserted into the Editor system. Summary statistics were collected. The Hotelling's T^2 test was used to determine if the use of CLASSY produced significantly better estimates on the training set than the current USDA procedure.

c. Comparison of the current standarized USDA procedure versus the MSE classifier - The entire estimation process was performed using the MSE classifier software. Summary statistics were collected. The Hotelling's T^2 test was used to determine if the use of the MSE classifier produced significantly better estimates on the training set than the current USDA procedure.

2.3.2 TRAINING ON 25 SEGMENTS AND TESTING ON 8 SEGMENTS

The data set was divided into two sets: a training set of 25 segments used to develop a classifier and a test set of 8 segments independent of the training set. The classifier developed on the training set was used to classify both the training and test sets. Regressions for the six crops of interest were developed on the training set and also on the test set. This was carried out with the standardized USDA procedure and again with CLASSY as a component of the Editor system, and finally with the MSE classifier software. Summary

statistics are presented in later sections. The following tests and comparisons were made:

a. For each of the three classification choices, an F-test was performed to determine if the regression line developed on the training set for a given crop was equal to the regression line developed on the test set. (A preliminary test for homogeneity of variance must be carried out first.) This test indicates if the regression line developed on the training set is extendible to the test set. A discussion of this test appears in section 3.2.1.

b. The current USDA clustering procedure was compared with the CLASSY clustering algorithm. The Hotelling's T^2 test was performed to determine if the use of CLASSY produced significantly better estimates on an independent set than the USDA procedure. The estimates for the independent set were obtained from the regression line, which was developed on the training set.

c. The current USDA classification procedure was compared with the MSE classification. The Hotelling's T^2 test was performed to determine if the use of the MSE classifier produced significantly better estimates on an independent set than the USDA procedure.

2.3.3 JACKKNIFING

Jackknifing techniques were used to simulate the procedure of training on a sample and developing regressions on an independent set. The following experiments were conducted.

a. Jackknifing within the Editor system

By repeating the division of the data set into training and test portions so that all segments appear exactly once in a test group, the combined test groups from all repetitions represent a quasi-independent test set. Summary statistics and regressions were obtained from this quasi-independent test set and were compared to results obtained when training and testing on all 33 segments. Details of this jackknifing are given in a later section.

b. Jackknifing with the MSE classifier

The data set was divided into sets consisting of 25 and 8 segments. The set of 25 segments was further divided into a set of 24 training segments and a set of 1 test segment. This division of the 25 segments was repeated so that each of the 25 segments appeared exactly once as a test segment. These test segments were combined to form a quasi-independent test set of 25 segments. The Hotelling's T^2 test was performed to determine if the regressions developed on the quasi-independent test set produced significantly better estimates on an independent set than did the regressions developed on the 25 segments without jackknifing. For each crop, an F-test was performed to determine if the regression line developed on the quasi-independent test set was equal to the regression line developed on an independent test set. Details of this jackknifing appear in section 5.6, Cross-Validation Procedure.

2.4 SELECTION OF TRAINING AND TEST SEGMENTS

Before training and test segments were selected, the geography, strata boundaries, and the Landsat imagery of the area covered by the 33 segments were studied. These segments, when plotted on a topographic map of Missouri, covered a rectangular area roughly 100 miles (160 kilometers,) on a side. This rectangle represented about one-eighth of the land area in Missouri. About three-quarters of the segments were 800 feet (244 meters) in elevation. The remaining segments were scattered around the perimeter at 1000 feet (305 meters).

The geography was rolling. No major urban areas were nearby. A copy of the topographic map is included in appendix B. Some cloud cover was at the edge of the scene, but very little was over the segments. A fairly uniform color distribution prevailed.

County maps with strata boundaries were provided by the USDA, with the 33 segments identified on them. All 33 segments were in strata 10. Very few were

near the strata boundaries, and the segments were evenly distributed over the counties (see table 2-1).

There were, then, two factors in the choice of training and test segments, both related to geography. It is assumed that if soil characteristics vary over the region, they will be most different between segments that are widely separated geographically. Also, segments located at the extreme edges of the region were slightly higher in elevation. Therefore, to obtain representative training segments and representative test data, each group of segments should be distributed uniformly over the geographic region covered by all 33 segments.

Eight independent test segments were chosen by laying out a uniform grid with eight boxes over the topographic map. One segment from each box was randomly selected. The remaining 25 segments constituted the training group. Listed in tables 2-4 and 2-5 are the segment number and ground truth crops present (in pixels) for the test and training groups, and the comparisons with all 33 segments.

These groups were found satisfactory, since they are both fairly representative of the total data set of 33 segments and still provide some variation.

To further validate that the 25 training segments are representative of the entire data set, the percent of the training set correctly classified when training with 25 segments and with all 33 segments are presented in table 2-6. The similarity of these results are indicative that the 25 training segments are representative.

In the Editor jackknifing experiment, it was necessary to partition the data into 11 groups of 3 segments each. This was accomplished by laying a grid with three boxes over the topographic map such that 11 segments fell into each box. Three segments were then chosen randomly, one from each box. This was repeated 10 times, obtaining 11 test groups of 3 segments each. Each corresponding training group was composed of the remaining 30 segments. Mean values for each of the 11 test and training groups can be found in appendix D.

TABLE 2-4.- PURE PIXELS OF GROUND TRUTH IN EIGHT TEST SEGMENTS

Segment no.	Corn	Winter wheat	Permanent pasture	Soybeans	Dense woodland	Other hay
6038	0	0	0	0	0	0
6048	53	0	7	0	62	0
6059	0	0	115	18	44	14
6098	0	0	310	0	0	0
9017	22	13	141	46	7	10
9037	17	26	79	67	2	2
9046	0	0	99	0	47	36
9052	39	19	26	137	0	0
Total	131	58	777	268	162	62
Mean of 8	16.4	7.2	97.1	33.5	20.3	7.8
Mean of 33	19.7	5.8	76.6	53.8	15.6	17.9
SD of 8	20.6	10.6	100.4	48.9	26.1	12.6
SD of 33	25.2	10.8	90.1	55.8	27.3	23.4

TABLE 2-5.- PURE PIXELS OF GROUND TRUTH IN 25 TRAINING SEGMENTS

Segment no.	Corn	Winter wheat	Permanent pasture	Soybeans	Dense woodland	Other hay
6015	31	19	30	81	0	8
6034	5	7	45	13	0	38
6035	51	5	84	125	0	34
6040	9	2	124	22	8	55
6045	6	0	340	13	0	0
6050	34	0	182	0	0	79
6053	0	0	40	104	4	4
6058	12	0	0	171	2	35
6060	0	0	0	10	3	6
6063	22	2	31	76	16	0
6064	0	7	46	76	40	0
6065	29	3	28	1	0	15
6073	27	0	41	17	77	33
6085	0	0	165	22	0	72
6095	0	0	0	0	0	0
9061	86	0	0	183	0	0
9036	5	3	36	125	0	0
9047	59	43	42	152	3	0
9051	0	35	145	9	115	0
9057	0	0	108	17	14	45
9016	93	0	0	85	0	0
9062	8	0	251	18	30	16
9066	0	8	0	79	0	0
9096	37	0	0	26	39	28
9097	6	0	12	82	1	62
Total	520	134	1750	1507	352	530
Mean of 25	20.8	5.4	70.0	60.3	14.1	21.2
Mean of 33	19.7	5.8	76.6	53.8	15.6	17.9
SD of 25	26.8	11.0	87.8	57.2	27.9	25.3
SD of 33	25.2	10.8	90.1	55.8	27.3	23.4

TABLE 2-6.- MULTITEMPORAL PERCENT CORRECTLY CLASSIFIED RESULTS FROM TRAINING ON 25 SEGMENTS VERSUS TRAINING ON 33 SEGMENTS

Crop	Editor		CLASSY		MSE classifier	
	Train on 25 segments	Train on 33 segments	Train on 25 segments	Train on 33 segments	Train on 25 segments	Train on 33 segments
Corn	74.24	72.57	77.12	72.31	71.36	65.61
Winter wheat	35.46	29.76	35.73	38.05	16.90	20.13
Permanent pasture	66.56	78.92	73.12	75.45	87.13	85.34
Soybeans	83.69	79.33	84.15	81.57	86.02	83.48
Dense woodland	54.55	46.65	50.19	49.74	32.91	33.98
Other hay	32.52	22.41	32.29	26.14	2.79	1.87

2.5 SAMPLE SIZE ESTIMATES

Table 2-7 gives the sample sizes needed to detect different amounts (10, 15, and 20 percent) of deviation from the average ground truth at various levels of type I error (α) and type II error (β) for each of the six crops. For each crop and for a given amount of deviation, the sample size was obtained by solving two simultaneous equations which relate the type I and type II errors to the sample size and the critical value. The equations were based on the standardized normal distribution.

TABLE 2-7.- UNIVARIATE SAMPLE SIZE ESTIMATES

Crop	Average ground truth	MSE	Δ Percent	α = 0.05			α = 0.1		
				β=0.1	β=0.2	β=0.3	β=0.1	β=0.2	β=0.3
Corn	15.909	68.165	0.1	231	167	127	177	122	88
			.15	103	74	57	79	54	40
			.20	58	42	32	45	31	22
Winter wheat	4.365	24.577	0.1	1104	797	608	846	580	421
			.15	491	355	270	376	258	187
			.20	276	200	152	212	145	106
Permanent pasture	40.047	320.887	0.1	172	124	95	132	90	66
			.15	77	55	42	59	40	29
			.20	43	31	24	33	23	17
Soybeans	32.522	128.773	0.1	105	76	58	80	55	40
			.15	47	34	26	36	25	18
			.20	27	19	15	20	14	10
Dense woodland	10.768	83.933	0.1	620	448	341	475	326	236
			.15	276	199	152	211	145	105
			.20	155	112	86	119	82	59
Other hay	9.228	92.370	0.1	929	670	511	711	488	354
			.15	413	298	228	316	217	158
			.20	233	168	128	178	122	89

Table 2-8 gives the sample sizes needed to detect different amounts (10, 15, and 20 percent) of deviation from the ground truth mean vector at type I error (α= 0.05) and various levels of type II error (β) when all six crops are considered simultaneously. For a given amount of deviation, the sample size was obtained by solving two simultaneous equations which relate the type I and type II errors to the sample size and the critical value. The equations were based on central and noncentral Hotelling's T^2 distributions.

TABLE 2-8.- MULTIVARIATE SAMPLE SIZE ESTIMATES

Δ percent	α = 0.05			
0.1	β=0.075 46	β=0.187 36	β=0.220 34	β=0.295 30
0.15	β=0.083 24	β=0.184 20	β=0.264 18	β=0.367 16
0.20	β=0.11 16	*	*	*

*No values.

3. EVALUATION OF STANDARDIZED USDA PROCEDURE

3.1 ANALYSIS RESULTS FOR TRAINING AND TESTING ON 33 SEGMENTS

3.1.1 COMPARISON WITH USDA RESULTS

In tables 3-1, 3-2, and 3-3 are listed the omission error, the commission error, the residual MSE, and the r^2 of the regression for each crop from runs made with the standardized procedure and all 33 segments for training (columns headed JSC). Also included (when available) for comparison are figures provided by the USDA from similar runs made by USDA analysts (columns headed USDA).

The percent correctly classified is equal to one minus the omission error. From the omission and commission errors, it is apparent that the August acquisition provided better results than the May acquisition. In section 3.1.2, a multivariate statistical test was performed to determine if multitemporal data provided significantly better estimates than the August data.

3.1.2 HOTELLING'S T^2 TEST COMPARING UNITEMPORAL AND MULTITEMPORAL ESTIMATES

To compare the performance of the standardized USDA procedure using different types of data, namely unitemporal and multitemporal, a criterion to measure the performance must first be defined. The criterion adopted in this study is a vector consisting of the absolute differences between the ground truth and the regression estimate for each of the six crop types of interest. Multivariate statistical analysis techniques have been applied, because the major objective is to evaluate the performance of the procedures in classifying and estimating the crop hectarage of all six crop types simultaneously. To compare the unitemporal and multitemporal results, a test is made of the equality of the two mean vectors of the absolute differences (vectors of means of the absolute value of the differences.) If the hypothesis of equal mean vectors is rejected, the type of data yielding a smaller mean vector of absolute differences between the ground truth and the regression estimate is preferred.

TABLE 3-1.- EDITOR MULTITEMPORAL PERFORMANCE MEASURES FOR TRAINING AND TESTING ON 33 SEGMENTS

Crop	r^2		Percent correct		Omission error		Commission error		Residual MSE	
	JSC	USDA	JSC	USDA	JSC	USDA	JSC	USDA	JSC	USDA
Corn	0.80	0.82	0.73	0.67	0.27	0.33	0.37	0.24	68.2	*
Winter wheat	.38	.52	.29	.34	.71	.66	.56	.44	24.6	*
Permanent pasture	.79	.80	.79	.75	.21	.25	.46	.36	320.9	*
Soybeans	.85	.85	.79	.78	.21	.22	.33	.23	128.8	*
Dense woodland	.62	.65	.47	.47	.53	.53	.54	.44	83.9	*
Other hay	.20	.48	.22	.32	.78	.68	.60	.52	92.4	*

TABLE 3-2.- EDITOR AUGUST PERFORMANCE MEASURES FOR TRAINING AND TESTING ON 33 SEGMENTS

Crop	r^2		Percent correct		Omission error		Commission error		Residual MSE	
	JSC	USDA	JSC	USDA	JSC	USDA	JSC	USDA	JSC	USDA
Corn	0.42	0.37	0.52	0.43	0.48	0.57	0.55	0.47	197.8	*
Winter wheat	.27	.39	.34	.30	.66	.70	.68	.42	28.8	*
Permanent pasture	.74	.75	.72	.73	.27	.27	.52	.44	391.5	*
Soybeans	.75	.75	.74	.74	.26	.26	.37	.29	214.0	*
Dense woodland	.44	.44	.31	.34	.68	.66	.58	.51	125.8	*
Other hay	.03	.18	.08	.12	.92	.88	.79	.73	111.4	*

*No values.

TABLE 3-3.- EDITOR MAY PERFORMANCE MEASURES FOR TRAINING AND TESTING ON 33 SEGMENTS

Crop	r^2		Percent correct		Omission error		Commission error		Residual MSE	
	JSC	USDA	JSC	USDA	JSC	USDA	JSC	USDA	JSC	USDA
Corn	0.07	0.35	0.26	0.18	0.74	0.82	0.76	0.58	313.4	*
Winter wheat	.01	.12	.02	.09	.98	.91	.88	.66	39.0	*
Permanent pasture	.58	.74	.68	.77	.32	.23	.51	.44	648.9	*
Soybeans	.61	.63	.67	.72	.33	.28	.52	.42	326.4	*
Dense woodland	.44	.44	.33	.24	.67	.76	.65	.53	125.2	*
Other hay	.05	.20	.16	.19	.84	.81	.64	.64	109.1	*

The hypothesis is now formulated and tested as follows:

Let

$$\mu_A = \begin{bmatrix} \mu_{A1} \\ \mu_{A2} \\ \mu_{A3} \\ \mu_{A4} \\ \mu_{A5} \\ \mu_{A6} \end{bmatrix} \quad ; \quad \mu_B = \begin{bmatrix} \mu_{B1} \\ \mu_{B2} \\ \mu_{B3} \\ \mu_{B4} \\ \mu_{B5} \\ \mu_{B6} \end{bmatrix} \tag{1}$$

where

μ_{Ai} = the mean of the absolute difference between the ground truth and the regression estimate of crop i from the USDA procedure using multitemporal data,

μ_{Bi} = the mean of the absolute difference between the ground truth and the regression estimate of crop i from the USDA procedure using unitemporal data, (crop 1 is corn; crop 2 is winter wheat; crop 3 is permanent pasture; crop 4 is soybeans; crop 5 is dense woodland; crop 6 is other hay.)

It is desired to test

$$\begin{aligned} H_0&: \ \mu_A - \mu_B = 0 \\ H_1&: \ \mu_A - \mu_B \neq 0 \end{aligned} \tag{2}$$

It is assumed that a random sample of 33 segments was chosen. Classification has been performed; and ground truth, classification results, and regression estimates were obtained for each of the 33 segments. Let

$$Y_j = \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ Y_{3j} \\ Y_{4j} \\ Y_{5j} \\ Y_{6j} \end{bmatrix} \quad ; \quad \hat{Y}_{Aj} = \begin{bmatrix} \hat{Y}_{A1j} \\ \hat{Y}_{A2j} \\ \hat{Y}_{A3j} \\ \hat{Y}_{A4j} \\ \hat{Y}_{A5j} \\ \hat{Y}_{A6j} \end{bmatrix} \quad ; \quad \hat{Y}_{Bj} = \begin{bmatrix} \hat{Y}_{B1j} \\ \hat{Y}_{B2j} \\ \hat{Y}_{B3j} \\ \hat{Y}_{B4j} \\ \hat{Y}_{B5j} \\ \hat{Y}_{B6j} \end{bmatrix} \quad ; \quad j = 1,2,\cdots,33 \tag{3}$$

where

Y_{ij} = the ground truth of crop i in segment j

$\hat{Y}_{Aij}$ = the regression estimate of the ground truth of crop i in segment j from the USDA procedure using multitemporal data

$\hat{Y}_{Bij}$ = the regression estimate of the ground truth of crop i in segment j from the USDA procedure using unitemporal data, i = 1,•••,6

To test the hypothesis, the vectors formed are

$$d_j = \begin{bmatrix} |Y_{1j} - \hat{Y}_{A1j}| - |Y_{1j} - \hat{Y}_{B1j}| \\ |Y_{2j} - \hat{Y}_{A2j}| - |Y_{2j} - \hat{Y}_{B2j}| \\ |Y_{3j} - \hat{Y}_{A3j}| - |Y_{3j} - \hat{Y}_{B3j}| \\ |Y_{4j} - \hat{Y}_{A4j}| - |Y_{4j} - \hat{Y}_{B4j}| \\ |Y_{5j} - \hat{Y}_{A5j}| - |Y_{5j} - \hat{Y}_{B5j}| \\ |Y_{6j} - \hat{Y}_{A6j}| - |Y_{6j} - \hat{Y}_{B6j}| \end{bmatrix} \quad ; \quad j = 1,\cdots,33 \qquad (4)$$

The Hotelling's T^2 testing statistic is given by

$$T^2 = N\, \bar{d}' S_d^{-1}\, \bar{d} \qquad (5)$$

where

N = sample size

$\bar{d} = \frac{1}{N} \sum_{j=1}^{N} d_j$

S_d = the sample variance-covariance matrix of d_j

The computed $T^2 = 44.8324$, and $T^2_{0.05}\,(6,32) = 17.4$. Since $T^2 > T^2_{0.05}\,(6,32)$, we reject $H_0 : \mu_A - \mu_B = 0$ at the 0.05 level of significance and conclude that the mean vectors of absolute differences are not the same for multitemporal and unitemporal data. And since

$$\bar{d} = \begin{bmatrix} -5.35394 \\ -0.40394 \\ -0.81182 \\ -3.14828 \\ -1.89091 \\ -1.14818 \end{bmatrix} < \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (6)$$

it indicates that the regression estimates obtained from multitemporal data seem to be closer to the ground truth for all crop types than the regression estimates obtained from unitemporal data. It is therefore expected that multitemporal data will produce better estimation results.

3.2 EVALUATION ON AN INDEPENDENT TEST SET

3.2.1 TRAINING ON 25 SEGMENTS AND TESTING ON 8 SEGMENTS

One of the purposes of this study was to evaluate how well a classifier and the regression equations which were developed on the training set performed on an independent test set. Of the 33 segments available for analysis, 8 were chosen as a test set. The remaining 25 segments were then used in the standardized USDA procedure to train a classifier and to develop the regression equations for the six crops being studied. Performance measures of this classifier on the 25 training segments and on the 8 test segments are given in table 3-4. Also listed in this table are the r^2's (r = the correlation coefficient) and the regression MSE's for each crop in both sets of segments. This table shows that the training set had lower omission and commission errors for each crop than did the test set, with the exception of other hay. Also, the training set yielded higher r^2's than the test set, with dense woodland as the only exception. Both dense woodland and other hay are considered minor crops in this study. Finally, the overall percent correct is 57.70 for the training set as compared to 42.00 for the test set.

To determine whether the regression lines fitted to the 25 segments in the training set were appropriate for predicting ground truth in the 8 independent test segments, a two-stage F-test was performed for each crop. This test is constructed to determine if the regression line developed on the training set

TABLE 3-4.- EDITOR MULTITEMPORAL CLASSIFICATION PERFORMANCE MEASURES FOR TRAINING ON 25 SEGMENTS AND TESTING ON AN INDEPENDENT SET

Crop	25 Training segments*					8 Test segments†				
	MSE	r^2	% Correct	Omission	Commission	MSE	r^2	% Correct	Omission	Commission
Corn	28.805	0.91	74.24	25.76	31.13	202.865	0.61	54.98	45.02	42.89
Winter wheat	21.628	.50	35.46	64.54	66.05	37.275	.00	32.97	67.03	71.15
Permanent pasture	176.736	.88	66.56	33.44	44.56	1268.635	.39	51.76	48.24	47.87
Soybeans	119.426	.86	83.69	16.31	31.09	395.029	.40	71.74	28.26	63.17
Dense woodland	72.595	.66	54.55	45.45	50.12	36.662	.88	27.04	72.96	55.80
Other hay	79.541	.37	32.52	67.48	71.89	52.375	.24	39.81	60.19	88.64

*Overall % correct = 57.70

†Overall % correct = 42.00

is significantly different from the regression line developed on the test set. The structure of this test requires that the residual sum of squares for each line be pooled to form a common variance estimate. Thus, homogeneity tests for the error variances of the training and test sets must first be performed. These tests are outlined below.

Assume that the linear relationship between the ground truth and the number of pixels classified for the training set is given by

$$Y_i = a_1 + b_1 X_i + \varepsilon_{TRAIN} \tag{7}$$

and for the test set is given by

$$Y_i = a_2 + b_2 X_i + \varepsilon_{TEST} \tag{8}$$

where ε_{TRAIN} and ε_{TEST} are the random errors of the models, with variances σ^2_{TRAIN} and σ^2_{TEST}. The hypothesis for testing homogeneity of variances is stated as:

$$\begin{aligned} H_0&: \ \sigma^2_{TEST} = \sigma^2_{TRAIN} \\ H_1&: \ \sigma^2_{TEST} \neq \sigma^2_{TRAIN} \end{aligned} \tag{9}$$

The testing statistic is $F = MSE_{TEST}/MSE_{TRAIN}$, where MSE_{TEST} and MSE_{TRAIN} are the residual mean square errors obtained by separate regressions on the test set and the training set, respectively. The null hypothesis H_0 is rejected at level 0.10 if $F > F(0.95,6,23) = 2.51$ or if $F < F(0.05,6,23) = 0.260$.

If the homogeneity of variances is not rejected for a crop, then the following hypothesis is tested:

H_0: training set regression line = test set regression line
H_1: training set regression line $\neq$ test set regression line

The testing statistic is

$$F = \frac{\dfrac{SSE_{ALL} - SSE_{TRAIN} - SSE_{TEST}}{2}}{\dfrac{SSE_{TRAIN} + SSE_{TEST}}{29}} \tag{10}$$

where SSE_{TEST}, SSE_{TRAIN}, and SSE_{ALL} are the residual sums of squares obtained by separate regressions on the test set, training set, and combined test and training sets, respectively. The null hypothesis H_0 is rejected at level 0.05 if $F > F(0.95,2,29) = 3.32$. Results for these two tests for the six crops are given in table 3-5 (page 3-11). Homogeneity of variances was rejected for the major crops of corn, permanent pasture, and soybeans. Of the three remaining crops, the equality of the training set regression line and the test set regression line was rejected for the crop other hay.

In the first part of this section, F-tests were performed to determine if the regression line developed on the training set was significantly different from the regression line developed on the test set. These tests provided information as to the extendibility of the area estimation procedure which is currently being used by the USDA. In this section, an alternative method is presented with which to gain insight into the question of this extendability.

As known from regression theory, an estimator for the model variance can be obtained by summing the squared residuals from the regression and then dividing this quantity by its degrees of freedom. A similar estimator is now described. This estimator, denoted by $\hat{\sigma}^2$, is a weighted average of the 8 squared residuals obtained when the regression equation from the 25 training segments is used to predict the ground truth (Y) for the 8 test segments. The calculation is given by

$$\hat{\sigma}^2 = \frac{1}{8} \sum_{i=1}^{8} \frac{(Y_i - \hat{Y}_i)^2}{1 + \frac{1}{25} + \frac{(X_i - \bar{X})^2}{\sum_{j=1}^{25} (X_j - \bar{X})^2}} \qquad (11)$$

where

Y_i = ground truth hectarage for segment i in the test set, $i = 1,\cdots,8$

$\hat{Y}_i$ = estimated ground truth hectarage for segment i in the test set using the training regression equation as a predictor, $i = 1,\cdots,8$

X_i = number of classified pixels for segment i in the test set, $i = 1,\cdots,8$

$\bar{X}$ = the mean number of classified pixels per segment in the training set

$\sum_{j=1}^{25} (X_j - \bar{X})^2$ = the corrected sum of squares for the independent variable in the training set

And

$$E(\hat{\sigma}^2) = \frac{1}{8} \sum_{i=1}^{8} \frac{\sigma^2_{TEST} + \sigma^2_{TRAIN}\left(\frac{1}{25} + \frac{(X_i - \bar{X})^2}{\sum_{j=1}^{25} (X_j - \bar{X})^2}\right) + (EY_i - E\hat{Y}_i)^2}{1 + \frac{1}{25} + \frac{(X_i - \bar{X})^2}{\sum_{j=1}^{25} (X_j - \bar{X})^2}} \quad (12)$$

It can be seen that $E(\hat{\sigma}^2)$ depends upon the training set through its variance σ^2_{TRAIN} and $E\hat{Y}_i$. Likewise, $E(\hat{\sigma}^2)$ depends upon the test set through its variance σ^2_{TEST} and EY_i. If the training set and test set share the same regression line and $\sigma^2_{TRAIN} = \sigma^2_{TEST} = \sigma^2$ is a common variance, then $E(\hat{\sigma}^2) = \sigma^2$. No formal tests were made using $\hat{\sigma}^2$, but for a given crop if $\hat{\sigma}^2$ is considerably different from the MSE of the training set, it is an indication that $\sigma^2_{TRAIN} \neq \sigma^2_{TEST}$ and/or the training set regression line is significantly different from the test set regression line. Table 3-6 lists the MSE's of the six crops on the training set, which are unbiased estimates of the σ^2_{TRAIN}'s. Also listed is $\hat{\sigma}^2$ for each crop. For crops for which the homogeneity of variances was rejected in table 3-5, it appears that MSE's of the training set are quite different from their corresponding $\hat{\sigma}^2$'s.

TABLE 3-5.- EDITOR MULTITEMPORAL ANALYSIS: F-TESTS FOR HOMOGENEITY OF VARIANCES AND EQUALITY OF REGRESSION LINES

Crop	Computed F for homogeneity of variances	Computed F for equality of regression lines
Corn	*7.043	
Winter wheat	1.723	2.66
Permanent pasture	*7.178	
Soybeans	*3.308	
Dense woodland	.505	3.296
Other hay	.658	†3.383
Critical values	.260, 2.51	3.32

*Homogeneity of variances rejected.
†Equality of regression lines rejected.

TABLE 3-6.- EDITOR MULTITEMPORAL ANALYSIS: MEAN SQUARE ERRORS OF THE 25 TRAINING SEGMENTS AND $\hat{\sigma}^2$'s OF THE 8 TEST SEGMENTS

Crop	MSE	$\hat{\sigma}^2$
Corn	28.805	147.517
Winter wheat	21.628	43.449
Permanent pasture	176.736	1025.186
Soybeans	119.426	438.685
Dense woodland	72.595	88.132
Other hay	79.541	110.431

3.2.2 JACKKNIFING WITH THE EDITOR SYSTEM

Ideally, it is desirable to have a large sample with which to train a classifier and another large sample independent of the first with which to develop a regression line. The advantage of having a large sample is that the sampling variability is reduced as the sample size increases. When it is impossible to have a large training sample as well as a large sample with which to develop the regression line, a jackknifing procedure can be employed. The jackknifing, which is now described, simulates the method of training a classifier on a sample and then developing a regression on an independent sample.

The 33 segments were grouped into 11 sets containing 3 segments each. One set of 3 segments became the test set, while the remaining 10 sets were pooled and used to train a classifier. The test set containing three segments was then classified. This procedure was repeated 10 more times, with each set of 3 segments being the test set exactly once, and the remaining 30 segments being used to train a classifier. The 11 test sets were then combined, resulting in a sample of 33 segments, each having ground truth (Y) and a classification variable (X).

Regression equations for the 6 crops of interest were developed on this combined set of 33 segments. The regression MSE's, r^2's, and classification performance measurements are given in table 3-7 for this combined set. For comparison, the classification results obtained when all 33 segments were used to train the classifier are also given. With only one exception, the omission and commission error rates are higher in the jackknifed set than in the set where all 33 segments were used in the training. Also, the r^2's are lower in the jackknifed set. For the major crops of corn, permanent pasture, and soybeans, the decrease in r^2 is 0.15, 0.23, and 0.14, respectively. The results of this jackknifing study indicate that the r^2's reported by the USDA are overestimated.

TABLE 3-7.- EDITOR MULTITEMPORAL CLASSIFICATION PERFORMANCE MEASURES ON 33 SEGMENTS VERSUS PERFORMANCE MEASURES ON THE JACKKNIFED TEST SET

Crop	Training on all 33 segments*					Jackknifed test set†				
	MSE	$r^2=R^2$	% Correct	Omission	Com-mission	MSE	r^2	% Correct	Omission	Com-mission
Corn	68.165	0.80	72.57	27.43	36.68	83.106	0.75	67.50	32.50	37.51
Winter wheat	24.577	.38	28.76	71.24	55.78	34.538	.13	23.19	76.81	74.76
Permanent pasture	320.887	.79	78.92	21.08	45.69	680.577	.56	62.75	37.25	51.20
Soybeans	128.773	.85	79.33	20.67	32.66	243.650	.71	78.45	21.55	37.26
Dense woodland	83.933	.62	46.65	53.35	53.91	92.173	.59	48.24	51.76	59.62
Other hay	92.370	.20	22.41	77.59	60.29	113.273	.02	15.48	84.52	80.74

*Overall % correct = 57.77
†Overall % correct = 51.62

4. EVALUATION OF THE CLASSY CLUSTERING ALGORITHM

4.1 INTRODUCTION

The CLASSY clustering algorithm is an adaptive, maximum likelihood clustering procedure developed at JSC (refs. 1 to 4). The algorithm is fundamentally a density estimation algorithm which approximates the overall data distribution as a mixture of multivariate normal distributions. That is, if $\underline{x}$ is an observation vector and p is its probability density function, then

$$p(\underline{x}|m,\underline{\pi}_m) = \sum_{i=1}^{m} a_i p_i(\underline{x}|\underline{\mu}_i,\Sigma_i) \tag{13}$$

where

a_i = a priori probability of occurrence of class i

$p_i(\underline{x}|\underline{\mu}_i,\Sigma_i)$ = multivariate normal probability density function for class i with mean vector $\underline{\mu}_i$ and covariance matrix Σ_i

m = total number of classes

$\underline{\pi}_m$ = full set of parameters (i.e., $\{a_1,\cdots,a_m, \underline{\mu}_1,\cdots,\underline{\mu}_m, \Sigma_1,\cdots,\Sigma_m\}$)

Given a set of statistically independent, unlabeled sample vectors $\{\underline{x}_j\}$, the likelihood function may be formed in the following manner:

$$L(\{\underline{x}_j\}|m,\underline{\pi}_m) = \prod_{j=1}^{N} \left[\sum_{i=1}^{m} a_i p_i(\underline{x}_j|\underline{\mu}_i,\Sigma_i)\right] \tag{14}$$

where N is the total number of samples.

So far, the assumptions and equations parallel the usual maximum likelihood development. In using CLASSY, the additional assumption is that each value of the parameters m and $\underline{\pi}_m$ occurs with an a priori probability distribution $A(m,\underline{\pi}_m)$. This Bayesian formulation of the problem is taken to avoid the degenerate situation of increasing the likelihood by generating more and more clusters with smaller and smaller values of a_j. The practical limit of this process is that each class will be associated with only one data point.

In practice, the a priori probability $A(m,\underline{\pi}_m)$ has been chosen as

$$A(m,\underline{\pi}_m) = \left\{ \begin{array}{ll} \beta \prod_{i=1}^{m} c_i & \text{for } \underline{\pi}_m \in R_m \\ 0 & \text{otherwise} \end{array} \right\} \tag{15}$$

where

c_i = a constant containing normalizing factors over $\underline{\pi}_m$ space

β = overall normalization constant

R_m = finite region of $\underline{\pi}_m$ space corresponding to allowable values for the parameters

The objective of CLASSY, then, is to determine the discrete parameter m and the continuous parameter vector $\underline{\pi}_m$ so as to maximize the following function.

$$L(\{\underline{x}_j\},m,\underline{\pi}_m) = A(m,\underline{\pi}_m) \prod_{j=1}^{N} \left[\sum_{i=1}^{m} a_i p_i(\underline{x}_j | \underline{\mu}_i, \Sigma_i) \right] \tag{16}$$

The value of m and $\underline{\pi}_m$ which maximize equation (16) specify a set of distributions called clusters.

Many approaches may be taken to maximize equation (16). The approach chosen in CLASSY is to interleave the maximum likelihood iteration [designed to maximize $L(\{\underline{x}_j\},m,\underline{\pi}_m)$ with respect to the continuous parameter vector $\underline{\pi}$] with a discrete split, join, and combine process [designed to maximize $L(\{\underline{x}_j\},m,\underline{\pi}_m)$ with respect to the discrete parameter m]. Although the theoretical convergence properties of this procedure have not been examined, it is expected that, by alternating these two techniques, values of m and $\underline{\pi}$ corresponding to at least a local maximum of $L(\{\underline{x}_j\},m,\underline{\pi}_m)$ will be determined. Since the splitting and combining techniques operate around each existing cluster, and the statistics for hypotheses concerning different numbers of clusters are maintained separately, it has been observed that the final local maximum will often be global.

Necessary conditions for a maximum of $L(\{\underline{x}_j\},m,\underline{\pi}_m)$, assuming a fixed number of classes m, are well known [see Duda and Hart (ref. 5) and Wolfe (ref. 6)] and are given by the following equations:

$$p(i|\underline{x}_k,\underline{\pi}_m) = \frac{a_i p_i(\underline{x}_k|\underline{\mu}_i,\Sigma_i)}{\sum_{j=1}^{m} a_j p_j(\underline{x}_k|\underline{\mu}_j,\Sigma_j)} \tag{17}$$

$$a_i = \frac{1}{N}\sum_{k=1}^{N} p(i|\underline{x}_k,\underline{\pi}_m) \tag{18}$$

$$\underline{\mu}_i = \frac{\sum_{k=1}^{N} p(i|\underline{x}_k,\underline{\pi}_m)\underline{x}_k}{\sum_{k=1}^{N} p(i|\underline{x}_k,\underline{\pi}_m)} \tag{19}$$

$$\Sigma_i = \frac{\sum_{k=1}^{N} p(i|\underline{x}_k,\underline{\pi}_m)(\underline{x}_k - \underline{\mu}_i)(\underline{x}_k - \underline{\mu}_i)^T}{\sum_{k=1}^{N} p(i|\underline{x}_k,\underline{\pi}_m)} \tag{20}$$

where $p(i|\underline{x}_k,\underline{\pi}_m)$ is the posterior probability of class i, given the kth sample vector and the values of the parameters; and a_i, $\underline{\mu}_i$, and Σ_i, $i = 1\cdots,m$ are the elements of $\underline{\pi}_m$.

CLASSY uses a direct functional iteration to maximize equations (19) and (20); that is, estimates for $\underline{\mu}_i$ and Σ_i are used in the right-hand side to produce improved estimates on the left-hand side. Estimates for the a priori class probabilities, a_i, are computed using an iteration scheme which has proved to converge more rapidly than the simple functional iteration using equation (18). The scheme used is described in reference 4.

The optimization of $L(\{\underline{x}_j\},m,\underline{\pi}_m)$ with respect to the discrete parameter m generates hypotheses concerning the number of clusters and the subsequent testing of these hypotheses using a likelihood ratio test. At certain points in the process of maximum likelihood iteration, it is possible to generate a hypothesis concerning the fit of a given cluster to the data; namely, either

that the data are better represented by two clusters rather than one (a split hypothesis) or that the data are better represented by combining the given cluster with another cluster (a join hypothesis). Each cluster is periodically checked throughout the program to see if either a split or a join hypothesis seems reasonable. Measures of skewness and kurtosis are compared against values expected for a single, normal distribution to see if a split hypothesis should be considered. A measure of cluster similarity is used to determine if a join hypothesis is appropriate.

Clusters may be eliminated as the result of a likelihood ratio test or if their estimated a priori class probability in the mixture falls below a set threshold. Details concerning the split, join, eliminate operations as well as the operation of the algorithm in a general may be found in references 2, 3, and 4.

4.2 DESCRIPTION OF PROCEDURE

In order to evaluate the CLASSY clustering algorithm as a replacement for the clustering algorithm currently used in the Editor system, two different experiments were performed.

In the first of these experiments, CLASSY was used to cluster the pure pixels for each of the 6 test crops in the 33 Missouri segments. Unlike the standardized USDA procedure, outlying pixels and poorly registered fields were not removed before clustering. The resultant cluster statistics for each crop were transferred to the Editor system, and all pixels in the 33-segment area were classified using Editor's maximum likelihood classifier. Regression equations relating the classified pixels to the ground truth hectarage were developed for each crop. The performance measures for classification and regression, including the percent correctly classified, the omission and commission errors, and the r^2 and MSE for regression are given in table 4-1.

In the second experiment, CLASSY was used to cluster pure pixel data for each crop contained in 25 of the available 33 Missouri segments. The remaining 8 segments were reserved as an independent test set for use in evaluating the classifier and regression equations developed using the 25 training segments.

TABLE 4-1.- CLASSY MULTITEMPORAL PERFORMANCE MEASURES FOR TRAINING AND TESTING ON 33 SEGMENTS*

Crop	MSE	r^2	% Correct	Omission	'Com-mission
Corn	23.33	0.9308	72.31	27.69	29.47
Winter wheat	22.07	.4427	38.05	61.95	58.35
Permanent pasture	239.79	.8435	75.45	24.55	45.50
Soybeans	85.95	.8877	81.57	18.43	34.01
Dense woodland	62.53	.7195	49.74	50.26	51.50
Other hay	59.45	.4845	26.14	73.86	63.05

*Overall % correct = 58.10.

The purpose of this experiment was to determine if the performance of the hectarage estimation system on an independent data set was comparable to its performance on the training data set. The percent correctly classified, the omission and commission errors of the classifier developed on the 25 training segments and those from the 8 test segments are given in table 4-2. Similarly, the MSE and r^2 for the regression equations developed separately on the 25 training segments and on the 8 test segments are also given in this table. The last column in table 4-2 is an unbiased estimate of the error variance in applying the regression equation developed on the 25 training segments to the 8 test segments. This is the same statistic described in section 3.2.1.

In section 4.3, a statistical comparison is made of the hectarage estimates obtained when using CLASSY to cluster all 33 segments with the corresponding estimates using the standardized USDA procedure. In section 4.4, a similar statistical comparison is made for the estimates obtained for the eight independent test segments. Finally, in section 4.5, a test is made to determine whether the regression line developed on the 25 training segments is statistically different from a line fitted to the 8 test segments.

4.3 COMPARISON OF CLASSY AND THE STANDARDIZED USDA PROCEDURE

To compare the performance of CLASSY and the standardized USDA procedure, the criterion defined in section 3.1.2 and the Hotelling's T^2 test on the mean vectors of absolute differences have been used.

Let

$$\mu_C = \begin{bmatrix} \mu_{C1} \\ \mu_{C2} \\ \mu_{C3} \\ \mu_{C4} \\ \mu_{C5} \\ \mu_{C6} \end{bmatrix} \qquad (21)$$

where μ_{Ci} is the mean of the absolute difference between the ground truth and the regression estimate of crop i from the CLASSY procedure.

TABLE 4-2.- CLASSY MULTITEMPORAL CLASSIFICATION PERFORMANCE MEASURES FOR TRAINING ON 25 SEGMENTS AND TESTING ON AN INDEPENDENT SET

Crop	25 Training segments*					8 Test segments†					
	MSE	R^2	% Correct	Omission	Commission	MSE	R^2	% Correct	Omission	Commission	$\hat{\sigma}^2$
Corn 70	24.06	0.9245	77.12	22.88	30.48	313.48	0.3966	55.97	44.03	48.04	227.
Winter wheat 18.24	18.22	.5808	35.73	64.27	67.34	24.48	.3432	41.76	58.24	53.09	
Permanent pasture 42	230.88	.8437	73.12	26.88	44.66	1162.13	.4449	64.20	35.80	48.84	848.
Soybeans 24	112.10	.8721	84.15	15.85	31.43	186.40	.7148	70.43	29.57	59.70	175.
Dense woodland 13	49.45	.7681	50.19	49.81	42.44	52.15	.8270	19.15	80.85	55.56	138.
Other hay 68	53.90	.5760	32.29	67.71	62.13	54.47	.2081	26.21	73.79	87.32	101.

*Overall % correct = 59.62.
†Overall % correct = 45.38.

We now test:

$$H_0: \ \mu_A - \mu_C = 0$$
$$H_1: \ \mu_A - \mu_C \neq 0 \qquad (22)$$

where μ_A is defined in section 3.1.2. The Hotelling's T^2 testing procedure for the above problem is similar to that described in section 3.1.2. The computed T^2 is 44.1959, and $T^2_{0.05}(6,32)$ is 17.4. Since $T^2 > T^2_{0.05}(6,32)$, we reject H_0: $\mu_A - \mu_C = 0$ and conclude that the mean vectors of absolute differences are not the same for the two procedures. And since

$$\bar{d} = \begin{bmatrix} 1.69696 \\ 0.36333 \\ 2.39 \\ 0.37909 \\ 0.99424 \\ 2.08121 \end{bmatrix} > \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (23)$$

indicates that the regression estimates obtained by using CLASSY seem to be closer to the ground truth than the regression estimates obtained by using the standardized USDA procedure, it is believed that the CLASSY clustering algorithm performs better than the clustering algorithm used in the standardized USDA procedure.

4.4 COMPARISON OF CLASSY AND THE STANDARDIZED USDA PROCEDURE BY WEIGHTED MEAN VECTORS

Another testing was done on the mean vectors of weighted absolute differences so that crops with larger ground truth proportions contributed more in distinguishing the difference between CLASSY and the standardized USDA procedure. The hypothesis is formulated and tested as follows:

$$H_0: \ \mu_A^* - \mu_C^* = 0$$
$$H_1: \ \mu_A^* - \mu_C^* \neq 0 \qquad (24)$$

where

$$\mu_A^* = \begin{bmatrix} \mu_{A1}^* \\ \mu_{A2}^* \\ \mu_{A3}^* \\ \mu_{A4}^* \\ \mu_{A5}^* \\ \mu_{A6}^* \end{bmatrix} \quad ; \quad \mu_C^* = \begin{bmatrix} \mu_{C1}^* \\ \mu_{C2}^* \\ \mu_{C3}^* \\ \mu_{C4}^* \\ \mu_{C5}^* \\ \mu_{C6}^* \end{bmatrix}$$

and μ_{Ai}^* = the mean of the weighted absolute difference between the ground truth and regression estimate of crop i (weighted by its ground truth pixel proportion) from the USDA procedure

μ_{Ci}^* = the mean of the weighted absolute difference between the ground truth and regression estimate of crop i (weighted by its ground truth pixel proportion) from the CLASSY procedure

The computed T^2 is 20.0823, and $T^2_{0.05}(6,32)$ is 17.4. Since $T^2 > T^2_{0.05}(6,32)$, we reject H_0: $\mu_A^* - \mu_C^* = 0$ at the 0.05 level of significance and conclude that the weighted mean vectors of absolute differences are not the same for the two procedures. And the following

$$d^* = \begin{bmatrix} 0.446021 \\ 0.026101 \\ 0.88508 \\ 0.281399 \\ 0.169621 \\ 0.149436 \end{bmatrix} > \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (25)$$

indicates again that the CLASSY clustering algorithm seems to perform better than the clustering algorithm used in the current USDA procedure.

4.5 COMPARISON OF CLASSY AND THE USDA STANDARDIZED PROCEDURE ON AN INDEPENDENT TEST SET

In this study, the 33 segments were divided into two sets. One set consisting of 25 segments is called the training set; the remaining 8 segments form the

test set. The 25 training segments were used in training the classifier and in obtaining the regression coefficients. The regression line was then applied to the eight test segments to determine how well the line predicts crop hectarage using an independent set. This study has been completed for both CLASSY and the standardized USDA procedure. To compare the performance of CLASSY and the standardized USDA procedure on an independent test set, a similar Hotelling's T^2 test as in the previous section has been done on the eight segments.

The computed T^2 is 11.035, and $T^2_{0.05}(6,7)$ is 405.92. With this sample of 8 segments, since $T^2 < T^2_{0.05}(6,7)$, there is not enough statistical evidence to reject the hypothesis that the mean vectors of absolute differences between ground truth and the regression estimate are the same for the two procedures on an independent test set. A larger independent test set would be more appropriate because the critical value $T^2_\alpha(p,N-1)$ decreases rapidly as the sample size N increases.

4.6 COMPARISON OF TRAINING AND TEST SET REGRESSION LINES

In order to determine if the regression line fitted to the 25-segment training data was appropriate for the 8 independent test segments, a two-stage F-test, as described in section 3.2, was performed. The results are presented in table 4-3. Corn and permanent pasture did not pass the homogeneity of variance test. The test for equality of regression lines indicates that the regression lines are different only for dense woodland and other hay. However, the fact that corn and permanent pasture failed the homogeneity of variance test indicates that different regression models exist for the training and the test sets for these crops.

TABLE 4-3.- CLASSY MULTITEMPORAL ANALYSIS: F-TESTS FOR HOMOGENEITY OF VARIANCES AND EQUALITY OF REGRESSION LINES

Crop	Computed F for homogeneity of variances	Computed F for equality of regression lines
Corn	*13.03	
Winter wheat	1.34	0.09
Permanent pasture	*5.03	
Soybeans	1.66	0.10
Dense woodland	1.05	†6.82
Other hay	1.01	†3.61
Critical values	.260, 2.51	3.32

*Homogeneity of variances rejected.
†Equality of regression lines rejected.

5. EVALUATION OF THE MEAN SQUARE ERROR CLASSIFIER

5.1 BACKGROUND

The MSE classifier (ref. 7) is an algorithm intended to exploit the Bayes classification rule, which assigns an observation $\underline{x} \in R^m$ to one of m populations, $\omega_1, \omega_2, \cdots, \omega_m$, in which the lowest conditional average loss is incurred in so assigning $\underline{x}$. This conditional average loss for population j is given by

$$\mu_j(\underline{x}) = \sum_{i=1}^{m} c_{ij}\, p(\omega_i|\underline{x}) \tag{26}$$

where c_{ij} is the cost incurred in assigning $\underline{x}$ to ω_j when it actually belongs in ω_i, and $p(\omega_i|\underline{x})$ is the posterior probability that $\underline{x}$ is an observation on ω_i.

If a zero cost is assumed for correct classification and equal costs of one for incorrect classification, then the optimal classification rule which minimizes total expected loss also minimizes the probability of error in classification. In this context, the cost function c_{ij} can be expressed as

$$c_{ij} = (1 - \delta_{ij}) \tag{27}$$

where

$$\delta_{ij} = 1 \text{ if } i = j$$

$$\delta_{ij} = 0 \text{ if } i \neq j$$

and

$$\mu_j(\underline{x}) = \sum_{i=1}^{m} (1 - \delta_{ij}) p(\omega_i|\underline{x})$$

$$= 1 - p(\omega_j|\underline{x})$$

Thus $\underline{x}$ is assigned to ω_i if

$$\mu_i(\underline{x}) < \mu_j(\underline{x}) \; ; \; j = 1,2,\cdots,m, \; j \neq i \tag{28}$$

or equivalently,

$$p(\omega_i|\underline{x}) > p(\omega_j|\underline{x}) \quad ; \quad j = 1,2,\cdots,m, \; j \neq i \tag{29}$$

The particular algorithm used in this experiment was developed by S. G. Thadani (ref. 7).

For a given observation $\underline{x}$, m loss functions are estimated where, for each j, $\mu_j(\underline{x})$ is approximated by

$$\hat{\mu}_j(\underline{a}_j, \underline{x}) = \underline{a}_j^T \underline{\phi}(\underline{x}) \tag{30}$$

where

$$\underline{\phi}(\underline{x}) = [\phi_1(\underline{x}), \cdots, \phi_r(\underline{x})]^T$$

is an r-dimensional vector whose components are linearly independent functions of $\underline{x}$, and

$$\underline{a}_j = (a_{j1}, \cdots, a_{jr})^T$$

is a parameter vector determined so that the following MSE is minimal:

$$E_T [\mu_j(\underline{x}) - \hat{\mu}_j(\underline{a}_j, \underline{x})]^2 \tag{31}$$

The expectation with respect to the overall mixture density function is denoted by $E_T\{\cdot\}$.

It has been shown (ref. 7) that the vector $\underline{a}_j^*$ that minimizes equation (31) is the same vector that minimizes

$$M_L^j(\underline{a}_j) = \sum_{\ell=1}^{m} p(\omega_\ell) E_\ell [\hat{\mu}_j(\underline{a}_j, \underline{x}) - c_{\ell j}]^2 \tag{32}$$

where $E_\ell\{\cdot\}$ denotes the expectation with respect to the conditional probability density function $p(\underline{x}|\omega_\ell)$; additionally, if we define

$$M_N^j(\underline{a}_j) = \frac{\epsilon |\underline{a}_j|^2}{N} + \frac{1}{N} \sum_{\ell=1}^{m} \sum_{i_\ell=1}^{N_\ell} [\hat{\mu}_j(\underline{a}_j, \underline{x}_{i_\ell}) - c_{\ell_j}]^2 \tag{33}$$

$$j = 1, \cdots, m$$

where $\underline{x}_{i_\ell}$ is the i_ℓth training sample from class ω_ℓ and $|\underline{a}_j|^2 = \underline{a}_j^T \underline{a}_j$, then if

$$E_\ell[\underline{\phi}_i(\underline{x})^T \underline{\phi}_i(\underline{x})] < \infty \quad ; \quad \begin{matrix} i = 1, \cdots, r \\ \ell = 1, \cdots, m \end{matrix} \tag{34}$$

it can be shown that

$$\lim_{N \to \infty} M_N^j(\underline{a}_j) = M_L^j(\underline{a}_j) \tag{35}$$

The approach, then, is to minimize $M_N^j(a_j)$ with respect to $\underline{a}_j$ and to use a reasonably large number of training samples. If N_i represents the number of training samples from class i and $N = N_1 + \cdots + N_m$, then

$$\underline{a}_j = \left[\sum_{k=1}^{N} \underline{\phi}(\underline{x}_k)\underline{\phi}^T(\underline{x}_k) + \varepsilon I_r\right]^{-1} \sum_{\ell=1}^{m} c_{\ell_j} \sum_{i_\ell=1}^{N_\ell} \underline{\phi}(\underline{x}_{i_\ell}) \tag{36}$$

Where the εI_r term assures that the sum is positive definite.

Since equal costs of 1 for misclassification are assumed, this minimizing vector can be expressed as

$$\underline{a}_j^* = \left[\sum_{k=1}^{N} \underline{\phi}(\underline{x}_k)\underline{\phi}^T(\underline{x}_k) + \varepsilon I_r\right]^{-1} \cdot \left[\sum_{\substack{\ell=1 \\ \ell \neq j}}^{m} \sum_{i_\ell=1}^{N_\ell} \underline{\phi}(\underline{x}_{i_\ell})\right] \tag{37}$$

5.2 DESCRIPTION OF PROCEDURE

The procedure used in testing and evaluating the MSE classifier on the USDA data set consisted of the following four tasks:

a. A determination of the most appropriate form of the classifier.

b. A simulation, using the MSE classifier, of the standardized USDA procedure in which the classifier is trained on all the data for which ground truth is available.

c. A performance evaluation of the MSE classifier on an independent test set.

d. An investigation of the efficacy of using a cross-validation procedure to produce a different regression estimator.

A detailed discussion of each of the four tasks and the results follows.

5.3 DETERMINATION OF THE CLASSIFIER

This task addressed the issues of whether pure pixels only should be used in training the classifier and whether the terms of either a linear or quadratic function should be the components of $\underline{\phi}(\underline{x})$. Separate computer runs were made in which pure pixels and then all pixels were used, respectively, for training. Additional runs were made using, first, a linear form of $\underline{\phi}(\underline{x})$ and then a quadratic form. The conclusion reached was that a classifier trained on all the pixels in the crops of interest and using a quadratic form of the vector function $\underline{\phi}(\underline{x})$ produced the best classification results.

Since no clustering is done in the algorithm, valuable information is probably lost in restricting the training set to pure pixels. In each subsequent task described, reference to training the classifier will assume the use of all pixels in the training set and the terms of a quadratic function as entries of $\underline{\phi}(\underline{x})$.

5.4 SIMULATION OF THE STANDARDIZED USDA PROCEDURE

In this task, the MSE classifier was trained on the crops of interest in the 33 segments and then used to classify all pixels in the 33 segments. For each crop, the absolute values of the residuals were compared to the corresponding USDA results using a Hotelling's T^2 test for multivariate data. This test was applied again on the same data with the exception that, for each segment, the absolute values of the residuals were weighted by the proportions of the crops of interest in that segment. A discussion of the results of both tests follows. Following this discussion is table 5-1. This table reflects the statistics collected over the 33 segments.

5.4.1 COMPARISON OF THE MEAN SQUARE ERROR CLASSIFIER AND THE STANDARDIZED USDA PROCEDURE

The Hotelling's T^2 test, as described in section 3.1.2, has been performed to compare the results from the MSE classifier with the results from the standardized USDA procedure. Stated below are the testing procedure and the inferences.

Let

$$\mu_D = \begin{bmatrix} \mu_{D1} \\ \mu_{D2} \\ \mu_{D3} \\ \mu_{D4} \\ \mu_{D5} \\ \mu_{D6} \end{bmatrix} \tag{38}$$

where μ_{Di} is the mean of the absolute difference between the ground truth and the regression estimate of crop i from the MSE classifier.

We test

$$\begin{aligned} H_0&: \ \mu_A - \mu_D = 0 \\ H_1&: \ \mu_A - \mu_D \neq 0 \end{aligned} \tag{39}$$

where μ_A is defined in section 3.1.2.

Using the ground truth and regression estimates on the 33 segments, the computed T^2 is 21.777 and $T^2_{0.05}(6,32)$ is 17.4. Since $T^2 > T^2_{0.05}(6,32)$, we reject H_0: $\mu_A - \mu_D = 0$ at the 0.05 level of significance and conclude that the mean vectors of absolute differences are not the same for the two procedures. In this case, however, we have

$$\bar{d} = \begin{bmatrix} 0.52919 \\ -0.27372 \\ -1.17319 \\ -0.34008 \\ -0.67899 \\ -4.15128 \end{bmatrix} \tag{40}$$

Thus, it cannot be concluded that one procedure is better than the other because the elements in $\bar{d}$ do not have the same sign.

5.4.2 COMPARISON OF MEAN SQUARE ERROR CLASSIFIER AND STANDARDIZED USDA PROCEDURE BY WEIGHTED MEAN VECTORS

Similar testing, as in section 4.4, was done on the mean vectors of weighted absolute differences. We test

$$H_0: \quad \mu_A^* - \mu_D^* = 0$$
$$H_1: \quad \mu_A^* - \mu_D^* \neq 0 \tag{41}$$

where μ_A^* is defined earlier and μ_D^* is defined in a similar manner as μ_A^*.

The computed T^2 is 8.2857 and $T^2_{0.05}(6,32)$ is 17.4. Since $T^2 < T^2_{0.05}(6,32)$, there is not enough statistical evidence to reject the hypothesis that the weighted mean vectors of absolute differences are the same for the two procedures.

TABLE 5-1.- MSE CLASSIFIER MULTITEMPORAL PERFORMANCE MEASURES FOR TRAINING AND TESTING ON 33 SEGMENTS*

Crop	MSE	r^2	% Correct	Omission	Commission
Corn	51.90	0.8460	65.61	0.3439	0.2412
Winter wheat	24.63	.3781	20.13	.7987	.4013
Permanent pasture	361.16	.7643	85.34	.1466	.5001
Soybeans	128.02	.8478	83.48	.1652	.3573
Dense woodland	95.15	.5733	33.98	.6602	.4765
Other hay	115.38	.0005	1.87	.9813	.4706

*Overall % correct = 57.04.

5.5 MEAN SQUARE ERROR CLASSIFIER PERFORMANCE ON AN INDEPENDENT TEST SET

The MSE classifier was trained on the same 25 segments previously referred to as the training set, and all pixels in the 33 segments were classified. The regression equation determined by the 25 points obtained in the training process was used to predict the ground truth hectares in the 8 test segments. A Hotelling's T^2 test was applied on the mean vectors of absolute differences of the ground truth and regression estimates on the 8 test segments. The test is now described.

5.5.1 COMPARISON OF THE MEAN SQUARE ERROR CLASSIFIER AND THE STANDARDIZED USDA PROCEDURE ON EIGHT TEST SEGMENTS

When the performance of the MSE classifier was compared to that of the standardized USDA procedure on an independent test set, a Hotelling's T^2 test on the mean vectors of the absolute differences was conducted on a set of eight segments. The computed T^2 was 25.1924 and $T^2_{0.05}(6,7)$ was 405.92. Since $T^2 < T^2_{0.05}(6,7)$, there is, again, not enough evidence to reject the hypothesis. A larger independent test set is needed.

5.5.2 F-TEST FOR EQUALITY OF TRAINING AND TEST REGRESSION LINES

The two-stage F-test described in section 3.2 was used to determine if the regression line fitted through the 25 points using the MSE classifier was adequate to predict the ground truth in the 8 test segments. The results are presented in table 5-2. Corn and permanent pasture failed the homogeneity of variances test. The test for equality of the regression lines was not rejected for any crop which passed the homogeneity of variances test. Following table 5-2, table 5-3 displays performance statistics compiled on the 25 training segments and on the 8 test segments.

TABLE 5-2.- MSE CLASSIFIER MULTITEMPORAL ANALYSIS: F-TESTS FOR HOMOGENEITY OF VARIANCES AND EQUALITY OF REGRESSION LINES

Crop	Computed F for homogeneity of variances	Computed F for equality of regression lines
Corn	*11.896	
Winter wheat	1.600	1.240
Permanent pasture	*4.009	
Soybeans	1.939	1.516
Dense woodland	.3979	1.724
Other hay	.6094	2.332
Critical values	.260, 2.51	3.32

*Homogeneity of variances rejected.
†Equality of regression lines rejected.

TABLE 5-3.- MSE CLASSIFIER MULTITEMPORAL PERFORMANCE MEASURES FOR TRAINING ON 25 SEGMENTS AND TESTING ON AN INDEPENDENT SET

Crop	Training on 25*					Testing on 8†					
	MSE	r^2	% Correct	Omission	Commission	MSE	r^2	% Correct	Omission	Commission	$\hat{\sigma}^2$
Corn	26.37	0.92	71.36	28.64	24.22	313.74	0.40	54.73	45.27	45.85	221.60
Winter wheat	22.09	.49	16.90	83.10	44.74	35.38	.05	32.79	67.03	36.17	34.12
Permanent pasture	316.55	.79	87.13	12.87	46.13	1269.16	.39	76.46	23.53	47.64	1028.37
Soybeans	109.97	.87	86.02	13.98	32.43	213.21	.67	74.78	25.22	59.10	209.35
Dense woodland	93.93	.56	32.91	67.09	49.55	37.38	.88	18.59	81.41	53.85	67.74
Other hay	111.99	.07	2.79	97.21	59.32	68.24	.01	2.91	97.09	70.00	59.96

*Overall % correct = 44.69.

†Overall % correct = 49.24.

5.6 CROSS-VALIDATION PROCEDURE

The current USDA procedure is to develop a regression estimator by fitting a line to the points obtained from the training data. The objective of the fourth and final task was to determine if the estimate of the ground truth hectares in the eight test segments could be improved by using a different regression line based on a procedure referred to as cross-validation. Of the 25 training segments, 1 segment is left out, and the MSE classifier is trained on the remaining 24 segments. The omitted segment is then classified as if it represented an independent test set. This process is repeated for each of the 25 segments, thus producing 25 points through which a regression line is fitted. The 8 test segments are then classified using the MSE classifier developed on all 25 segments, and the ground truth hectares for these 8 segments are predicted from the regression line.

Two tests were conducted on the results. The first was the two-stage F-test for the equality of the regression line determined in the cross-validation procedure and the regression line fitted to the eight test segments. These results are presented in table 5-4. It is noted that, in the cross-validation procedure, the hypothesis for equality of variances was rejected for woodlands in addition to the corn and pasture crops previously rejected in the noncross-validation procedure.

The final question to settle was which of the two procedures, cross-validation or noncross-validation, yielded a regression line which best predicted the ground truth for an independent set. The Hotelling's T^2 test previously discussed was used on the eight test segments. In this application,

$$H_0: \ \mu_I - \mu_{II} = 0$$
$$H_1: \ \mu_I - \mu_{II} \neq 0 \qquad (42)$$

where

μ_I = mean vector of absolute differences between the ground truth and the regression estimate on an independent set using cross-validation in obtaining the fitted line on the 25 points.

μ_{II} = mean vector of absolute differences between the ground truth and the regression estimate on an independent set without cross-validation.

The computed T^2 is 9.9528 and $T^2_{0.05}(6,7)$ is 405.92. Since $T^2 < T^2_{0.05}(6,7)$, we conclude that there is not enough evidence to reject the hypothesis. A larger independent test set is needed.

TABLE 5-4.- MSE CLASSIFIER MULTITEMPORAL ANALYSIS: F-TESTS FOR HOMOGENEITY OF VARIANCES AND EQUALITY OF REGRESSION LINES IN CROSS-VALIDATION PROCEDURE

Crop	Computed F for homogeneity of variances	Computed F for equality of regression lines
Corn	*8.478	‡
Winter wheat	0.854	0.412
Permanent pasture	*2.980	‡
Soybeans	1.070	2.10
Dense woodland	*.255	‡
Other hay	.553	1.18
Critical values	.260, 2.51	3.32

*Homogeneity of variances rejected.
†Equality of regression lines rejected.
‡No values.

6. CALIBRATION REGRESSION APPROACH

6.1 INTRODUCTION

Statistical methods have often been illustrated with beautiful examples without adequately emphasizing the abstract ideas that underlie the methods; that is, ideas essential to correct statistical thinking. The result has been that certain problems with similar objectives appear amenable to identical statistical solutions when, in fact, intrinsic differences exist which alter considerably the details of their solutions. It is often the case that the practitioner is interested in assessing the value of some quantity which is impracticable to assess or impossible to observe directly in a given instance, the estimation being performed with the aid of a relationship between the quantity whose value is sought and another whose value can be determined directly. The curve-fitting procedure usually adopted depends on the additional assumption that the values of the independent variables are known exactly (without error) - an assumption often passed by without emphasis. This simplification of problems without explicit mention of the fact fosters misconceptions that are carried over into analysis of data, a particularly bad misconception being that the variable whose value is to be estimated automatically assumes the role of the dependent variable. The calculation and use of dosage-response curves to estimate dosage constitute an example. The dosage-response curve should be evaluated from a series of observations, with dosage as the independent variable, and the curve then used to estimate unknown dosages from observable responses.

To illustrate the aforementioned in more detail, assume that a linear relation prevails between U and V

$$\alpha_0 + \alpha_1 U + \alpha_2 V = 0 \qquad (43)$$

which may be written in the equivalent forms

$$V = \alpha + \beta U \tag{44}$$

where

$$\alpha = -\alpha_0/\alpha_2$$
$$\beta = -\alpha_1/\alpha_2$$

$$U = \gamma + \delta V \tag{45}$$

where

$$\gamma = -\alpha_0/\alpha_1$$
$$\delta = -\alpha_2/\alpha_1$$

A common impression regarding the principles of curve-fitting seems to be: If one is interested in estimating V from U, then take $\hat{V} = a + bU$ as the estimate of equation (44); if one were fitting by the method of least squares, the a and b that minimize $\Sigma(V_i - \hat{V}_i)^2$ would be found. On the other hand, if one is interested in estimating U from V, then $\hat{U} = c + dV$ is to be fitted, the values of c and d being chosen so as to make $\hat{U}$ a good fit in terms of the deviations $(U_i - \hat{U}_i)$. It does not seem to be generally realized that the fitting should be done in terms of the deviations which actually represent "error." Thus, when the research worker selects the U-values in advance, holds U to these values without error, and then observes the corresponding V-values, the errors are in the V-values. So, even if the researcher is interested in using observed values V_0 of V to estimate U, he should nevertheless fit $V = a + bU$ and then use the inverse of this relation to estimate U; i.e., $\hat{U} = (V_0 - a)/b$. Let us examine this from the viewpoint of the theory of least squares. Consider the case where the values of U are selected (or adjusted) by the research worker, and the corresponding values of V are found by observation. One can minimize $\Sigma(\hat{V}_i - V_i)^2$ and $\Sigma(\hat{U}_i - U_i)^2$, thereby obtaining the two lines, respectively:

$$V = a + bU \tag{46}$$

$$U = c + dV \tag{47}$$

$\Sigma(V_i - \hat{V}_i)^2$ minimized ANOVA I	$\Sigma(U_i - \hat{U}_i)^2$ minimized ANOVA II
Total variability of V's about their mean: $\Sigma(V_i - \bar{V})^2$	Total variability of U's about their mean: $\Sigma(U_i - \bar{U})^2$
Reduction effected by (46): $b\Sigma(U_i - \bar{U})(V_i - \bar{V})$	Reduction effected by (47): $d\Sigma(U_i - \bar{U})(V_i - \bar{V})$
Deviation about $\hat{V}$ $\Sigma(V_i - \bar{V})^2 - b\Sigma(U_i - \bar{U})(V_i - \bar{V})$	Deviation about $\hat{U}$ $\Sigma(U_i - \bar{U})^2 - d\Sigma(U_i - \bar{U})(V_i - \bar{V})$
$= \Sigma(V_i - \hat{V}_i)^2$	$= \Sigma(U_i - \hat{U}_i)^2$

The ANOVA tables are interpreted as follows. On the left, $\Sigma(V_i - \bar{V})^2$ gives a measure of the observed variability of the V-values. The second row of ANOVA I gives the portion of the observed variability of the V-values that can be attributed to the dependence of V on U, and the last row indicates the magnitude of the portion of $\Sigma(V_i - \bar{V})^2$ that must be attributed to "error" (this portion has been minimized by the fitting process). In ANOVA II, on the other hand, $\Sigma(U_i - \bar{U})^2$ represents the variability in the chosen values of U which resulted from the way in which the researcher selected them, and it should be noted that the corresponding values observed for V have in no way entered into their determination. Consequently, the apparent dependence of the U on the V, measured by the second row of ANOVA II, is a spurious dependence, and the last row of this table cannot be interpreted as being a measure of the "error" in the U-values, since it is that portion of the variability of the U-values that cannot be accounted for by the variability of the V-values. Briefly stated, when the values of U have been selected by the researcher and the corresponding V-values observed, the line obtained by minimizing $\Sigma(U_i - \hat{U}_i)^2$ is meaningless; and, accordingly, equation (46) is the only correct estimate of the postulated linear relationship between U and V. Therefore, if it is desired to reason from V to U, this must be done by means of $\hat{U} = (V_0 - a)/b$.

One example is to calibrate an instrument, say a pressure gauge. Assume that the increase in gauge marking is linearly proportional to the increase in pressure. To calibrate the gauge, one subjects it to two or more controlled

pressures (U) and notes the gauge markings (V). Using these data, the parameters are calculated, and the gauge is calibrated. The gauge is then used to determine unknown pressure (U) simply by reading the marking V_0 and obtaining $U = (V_0 - a)/b$.

The calibration problem is very general. Consider the problem of estimating the ground truth crop-type acreage of an area segment from the acreage obtained by classifying the Landsat data. To establish the relation between the two acreages, a number of segments was selected. The ground truth acreages of these segments were recorded and held as constants. The segments were then processed by a classification algorithm, and the classification acreages were obtained. In view of the fact that the ground truth acreages were controlled and the classification acreages depended upon spectral observations which can be regarded as chance occurrences and, therefore, are relatively imprecise, it seems only appropriate to consider the ground truth as the independent variable and classification acreages as the dependent variable. The ground truth acreage (U) of a new segment is then estimated from observing the classification acreage (V_0) of that segment by the equation $\hat{U} = (V_0 - a)/b$.

More detail on the application of this calibration model to the crop-type acreage estimation problem will be given in the next section.

6.2 DESCRIPTION OF NEW REGRESSION ESTIMATOR

This section provides a brief description of the regression method currently being used by the USDA to estimate the ground truth crop-type acreages. Denote the ground truth acreage by U and the acreage obtained from classifying the Landsat data by V. Using the sample of 33 segments, U is regressed onto V.

Ground truth acreages (U_i) and their corresponding classification acreages (V_i) are obtained for each of the 33 segments. The relation between U and V is then assumed to be

$$U_i = \gamma + \delta V_i + \varepsilon_i' \quad ; \quad i = 1, \cdots, 33 \tag{48}$$

where ε'_i, representing error, is distributed $N(0,\sigma'^2)$. The least squares estimators d of δ and c of γ obtained by minimizing $\Sigma(U_i - \hat{U}_i)^2$ are

$$d = \frac{\Sigma(U_i - \bar{U})(V_i - \bar{V})}{\Sigma(V_i - \bar{V})^2} \tag{49}$$

$$c = \bar{U} - d\bar{V} \tag{50}$$

The regression line is given by $U = c + dV$, and the ground truth acreage is estimated by $\hat{U} = c + dV$.

Although under this model (hereafter referred to as the current model) this estimator is unbiased with minimum variance (in the class of all unbiased linear estimators), the model does not seem to be appropriate. The reason is that, in the current model, the classification acreage (V) was considered as the fixed variable and the ground truth acreage (U) as the dependent variable; whereas, actually, the values of the ground truth acreages (U) were controlled and held as constants and only the values of the classification acreage (V) were observed and subject to error.

The calibration model is now introduced. Regressing the classification acreage (V) onto the ground truth acreage (U) gives,

$$V_i = \alpha + \beta U_i + \varepsilon_i \quad ; \quad i = 1,\cdots,n \tag{51}$$

where ε_i, representing error, is distributed as $N(0,\sigma^2)$. The least square estimators a of α and b of β obtained by minimizing $\Sigma(V_i - \hat{V}_i)^2$ are

$$b = \frac{\Sigma(U_i - \bar{U})(V_i - \bar{V})}{\Sigma(U_i - \bar{U})^2} \tag{52}$$

$$a = \bar{V} - b\bar{U} \tag{53}$$

The regression line is

$$V = a + bU \tag{54}$$

Given a classification acreage V, the ground truth acreage is estimated by

$$\hat{U}_1 = \frac{V - a}{b} \tag{55}$$

Another estimator under this model, $\hat{U}_2 = c + dV$, will also be considered where c and d are defined earlier. Note that $\hat{U}_2$ and $\hat{U}$, though having the same form c + dV, are two entirely different estimators because they are obtained under two different models. For instance, $\hat{U}_2$ is not an unbiased estimator ($\hat{U}$ is unbiased) because the classification acreages (V's) are no longer considered fixed constants under the calibration model.

We will now restrict our attention to the calibration model, which seems more appropriate than the current model in estimating the ground truth crop-type acreages. The properties of the two calibration estimators $\hat{U}_1$ and $\hat{U}_2$ are given in the next section.

6.3 THEORETICAL PROPERTIES OF THE TWO CALIBRATION ESTIMATORS

Under the calibration model

$$V_i = \alpha + \beta U_i + \varepsilon_i \quad ; \quad i = 1, \cdots, n \tag{56}$$

where ε_i is distributed as $N(0,\sigma^2)$, $\hat{U}_1$ is a maximum likelihood estimator and gives a readily interpreted analysis of variance. It may be noted here that the mean, variance, and MSE of $\hat{U}_1 = (V - a)/b$ are infinite, since there is a nonzero probability that b may be zero. The mean, variance, and MSE of $\hat{U}_2 = c + dV$ are finite for $n \geq 4$. However, it can be shown, with the help of Tchebycheff's inequality, that the probability of b lying in an interval that contains very small values, including zero, can be made very small by making $\Sigma(U_i - \bar{U})^2$ large, provided $|\sigma/\beta|$ is not large. This can be done by increasing n and choosing values of U_i that are not very close to each other. The expressions given below should be considered as corresponding to the distribution truncated for the value of b very close to zero.

Let U_0 be the quantity to be estimated and let

$$\sigma_U^2 = \frac{\sum_{i=1}^{n} (U_i - \bar{U})^2}{n - 1} \tag{57}$$

$$\theta = 1 + \frac{\sigma^2}{\beta^2 \sigma_U^2} \tag{58}$$

Then, to order $\frac{1}{n}$,

$$\text{bias } (\hat{U}_1) = \frac{\sigma^2}{(n-1)\sigma_U^2 \beta^2} (U_0 - \bar{U}) \tag{59}$$

$$\text{var } (\hat{U}_1) = \frac{\sigma^2}{\beta^2} \left[\frac{n+1}{n} + \frac{3\sigma^2}{(n-1)\sigma_U^2 \beta^2} \right] + \frac{\sigma^2}{(n-1)\sigma_U^2 \beta^2} (\bar{U} - U_0)^2 \tag{60}$$

$$\text{MSE } (\hat{U}_1) = \frac{\sigma^2}{\beta^2} \left[\frac{n+1}{n} + \frac{3\sigma^2}{(n-1)\sigma_U^2 \beta^2} \right] + \frac{\sigma^2}{(n-1)\sigma_U^2 \beta^2} (\bar{U} - U_0)^2 \tag{61}$$

$$\text{bias } (\hat{U}_2) = = \left[\frac{\sigma^2}{\beta^2 \sigma_U^2 \theta} - \frac{2\sigma^2}{n\beta^2 \sigma_U^2 \theta^3} \right] (\bar{U} - U_0) \tag{62}$$

$$\text{var } (\hat{U}_2) = \frac{\sigma^2}{\beta^2 \theta^2} \left[\frac{n+1}{n} + \frac{\sigma^2(\theta^2 - 2\theta + 6)}{(n-1)\sigma_U^2 \beta^2 \theta^2} \right] + \left[\frac{\sigma^2}{(n-1)\sigma_U^2 \beta^2 \theta^2} - \frac{2\sigma^4}{n\theta^4 \beta^4 \sigma_U^4} \right] (\bar{U} - U_0)^2 \tag{63}$$

$$\text{MSE}(\hat{U}_2) = \frac{\sigma^2}{\beta^2 \theta^2} \left[\frac{n+1}{n} + \frac{\sigma^2(\theta^2 - 2\theta + 6)}{(n-1)\sigma_U^2 \beta^2 \theta^2} \right] + \left[\frac{\sigma^2}{(n-1)\sigma_U^2 \beta^2 \theta^2} + \frac{\sigma^4}{\beta^4 \sigma_U^4 \theta^2} \left(1 - \frac{6}{n\theta^2}\right) \right] (\bar{U} - U_0)^2 \tag{64}$$

From equations (59) and (62), it is evident that both estimators are biased, but $\hat{U}_1$ is asymptotically unbiased whereas $\hat{U}_2$ is not.

$$\lim_{n\to\infty} \text{bias } (\hat{U}_1) = 0 \tag{65}$$

$$\lim_{n\to\infty} \text{bias } (\hat{U}_2) = \frac{\sigma^2}{\beta^2 \sigma_U^2 \theta} (\bar{U} - U_0) \tag{66}$$

However, both biases vanish at the point $U_0 = \bar{U}$ and may be small when U_0 lies very close to $\bar{U}$.

Berkson (ref. 8) has shown that when $|\sigma/\beta|$ is small, the asymptotic MSE of $\hat{U}_1$ is smaller than $\hat{U}_2$ except when U_0 lies very near to $\bar{U}$. Moreover, $\hat{U}_1$ is consistent whereas $\hat{U}_2$ is not. Saw (ref. 9) showed that, when U_0 lies very close to $\bar{U}$, $\hat{U}_2$ is closer than $\hat{U}_1$ to U_0; he further showed that other estimators can be obtained that may do even better than $\hat{U}_2$ in a much smaller interval. He thus found the use of $\hat{U}_2$ to be unappealing on this ground.

Applying this calibration model to the ground truth, crop-type acreage estimation problem, σ^2 and β were first estimated using the data on the 33 segments. Table 6-1 displays the two calibration estimators $\hat{U}_1$ and $\hat{U}_2$ for each of the six crops. Using the estimates of σ^2 and β and the equations given earlier, bias $(\hat{U}_1)$, bias $(\hat{U}_2)$, MSE $(\hat{U}_1)$, and MSE $(\hat{U}_2)$ were calculated for each of the six crops. These data are presented in table 6-2. It is clear that the magnitude of bias $(\hat{U}_1)$ is smaller than the magnitude of bias $(\hat{U}_2)$, and MSE $(\hat{U}_1)$ will be smaller than MSE $(\hat{U}_2)$ if U_0, the quantity we wish to estimate, is not very close to the sample mean $\bar{U}$.

TABLE 6-1.- CALIBRATION ESTIMATORS $\hat{U}_1$ AND $\hat{U}_2$

Crop	$\hat{U}_1$	$\hat{U}_2$
Soybeans	$\hat{U}_1 = \frac{V - 7.2214}{0.9548} = -7.5633 + 1.0473V$	$\hat{U}_2 = -1.4261 + 0.8870V$
Corn	$\hat{U}_1 = \frac{V - 5.8009}{0.7581} = -7.6519 + 1.3191V$	$\hat{U}_2 = -2.8899 + 1.0523V$
Permanent pasture	$\hat{U}_1 = \frac{V - 26.5319}{0.7612} = -34.8554 + 1.3137V$	$\hat{U}_2 = -19.1698 + 1.0386V$
Dense woodland	$\hat{U}_1 = \frac{V - 4.0353}{0.6769} = -5.9614 + 1.4773V$	$\hat{U}_2 = 0.3375 + 0.9211V$
Winter wheat	$\hat{U}_1 = \frac{V - 0.9228}{0.4521} = -2.0411 + 2.2119V$	$\hat{U}_2 = 1.9412 + 0.8366V$
Other hay	$\hat{U}_1 = \frac{V - 1.4935}{0.1177} = -12.6890 + 8.4962V$	$\hat{U}_2 = 4.8692 + 1.6913V$

TABLE 6-2.- BIAS AND MSE OF $\hat{U}_1$ and $\hat{U}_2$

U_0: quantity to be estimated

Crop	R^2	Bias ($\hat{U}_1$)	Bias ($\hat{U}_2$)	MSE ($\hat{U}_1$)	MSE ($\hat{U}_2$)
Soybeans	0.8469	$0.0058\ (U_0 - 32.5)$	$0.1505\ (32.5 - U_0)$	$159.3 + 0.0058\ (32.5 - U_0)^2$	$113.5 + 0.0257\ (32.5 - U_0)^2$
Corn	.7978	$.0082\ (U_0 - 15.9)$	$.1995\ (15.9 - U_0)$	$90.1 + 0.0082\ (15.9 - U_0)^2$	$56.7 + 0.0432\ (15.9 - U_0)^2$
Permanent pasture	.7906	$.0085\ (U_0 - 40.0)$	$.2067\ (40.0 - U_0)$	$428.6 + 0.0085\ (40.0 - U_0)^2$	$264.6 + 0.0462\ (40.0 - U_0)^2$
Dense woodland	.6235	$.0195\ (U_0 - 10.8)$	$.3751\ (10.8 - U_0)$	$146.6 + 0.0195\ (10.8 - U_0)^2$	$54.7 + 0.1446\ (10.8 - U_0)^2$
Winter wheat	.3782	$.0530\ (U_0 - 4.36)$	$.6240\ (4.36 - U_0)$	$77.4 + 0.0530\ (4.36 - U_0)^2$	$9.7 + 0.3933\ (4.36 - U_0)^2$
Other hay	.1991	$.1298\ (U_0 - 9.2)$	$.8041\ (9.2 - U_0)$	$659.3 + 0.1298\ (9.2 - U_0)^2$	$19.9 + 0.6500\ (9.2 - U_0)^2$

7. CONCLUSIONS AND RECOMMENDATIONS

7.1 CONCLUSIONS

With regard to the standardized USDA procedure used in this study, it seems clear that the multitemporal data produce significantly better estimates than those obtained using unitemporal data. In addition, it is clear that the current practice of evaluating the classifier and developing the regression on the same data set used to train the classifier can lead to optimistic performance estimates. With the possible exception of winter wheat and dense woodland, both of which had small populations, performance measures calculated using an independent test set and similar measures calculated using a cross-validation approach were uniformly worse than the same measures calculated on the training set. It also seems clear that the regression equation developed on the training data may not be appropriate for the test data. All crops tested, except winter wheat and dense woodland, showed differences in the regression models for lines calculated on a training set and on an independent test set.

The CLASSY clustering algorithm, when substituted for the current USDA clustering method, produced improved estimates. The estimates were significantly better than the standardized USDA procedure when testing and training were done on all 33 segments. The performance measures for the 33 segments are summarized in table 7-1.

The independent test set of eight segments was not large enough to allow the detection of any significant difference between the procedure using CLASSY and the standardized USDA procedure; however, the performance measures, as listed in table 7-2, indicate an improvement when using CLASSY clustering.

It is worthwhile to note that this improvement in performance was obtained despite the fact that CLASSY requires no decisions from an analyst concerning the number of clusters, separability thresholds, or other arbitrary parameters. In addition, CLASSY was operating with data for which the outlying observations had not been removed. Such observations were removed in the course of

TABLE 7-1.- COMPARISON OF MULTITEMPORAL PERFORMANCE MEASURES FOR TRAINING AND TESTING ON 33 SEGMENTS

Crop	Editor*					CLASSY†					MSE Classifier‡				
	MSE	r^2	% Correct	Omission	Commission	MSE	r^2	% Correct	Omission	Commission	MSE	r^2	% Correct	Omission	Commission
Corn	68.17	0.80	72.57	0.27	0.37	73.33	0.93	72.31	0.28	0.29	51.9	0.85	65.61	0.34	0.24
Winter wheat	24.58	.38	28.76	.71	.56	22.07	.44	38.05	.62	.58	24.63	.38	20.13	.80	.40
Permanent pasture	320.89	.79	78.92	.21	.46	239.79	.84	75.45	.25	.46	361.16	.76	85.34	.15	.50
Soybeans	128.77	.85	79.33	.21	.33	85.95	.89	81.57	.18	.34	128.02	.85	83.48	.17	.36
Dense woodland	83.93	.62	46.65	.53	.54	62.53	.72	49.74	.50	.52	95.15	.57	33.98	.66	.48
Other hay	92.37	.20	22.41	.78	.60	59.45	.48	26.14	.74	.63	115.38	.00	1.87	.98	.47

*Overall % correct = 57.77.

†Overall % correct = 58.10.

‡Overall % correct = 57.04.

TABLE 7-2.- COMPARISON OF MULTITEMPORAL PERFORMANCE MEASURES ON AN INDEPENDENT TEST SET

Crop	USDA Editor*					CLASSY†					MSE Classifier‡				
	$\hat{\sigma}^2$	r^2	% Correct	Omission	Commission	$\hat{\sigma}^2$	r^2	% Correct	Omission	Commission	$\hat{\sigma}^2$	r^2	% Correct	Omission	Commission
Corn	147.52	0.61	54.98	45.02	42.89	227.70	0.40	55.97	44.03	48.04	221.60	0.40	54.73	45.27	45.85
Winter wheat	43.45	.00	32.97	67.03	71.15	18.24	.34	41.76	58.24	53.09	34.12	.05	32.79	67.03	36.17
Permanent pasture	1025.19	.39	51.76	48.24	47.87	848.42	.44	64.20	35.80	48.84	1028.37	.39	76.46	23.53	47.64
Soybeans	438.69	.40	71.74	28.26	63.17	175.24	.71	70.43	29.57	59.70	209.35	.67	74.78	25.22	59.10
Dense woodland	88.13	.88	27.04	72.96	55.80	138.13	.83	19.15	80.85	55.56	67.74	.88	18.59	81.41	53.85
Other hay	110.43	.24	39.81	60.19	88.64	101.68	.21	26.21	73.79	87.32	59.96	.01	2.91	97.09	70.00

*Overall % correct = 42.00.

†Overall % correct = 45.38.

‡Overall % correct = 49.24.

executing the standardized procedure. This represents another source of subjective analyst input not needed when using CLASSY.

The MSE classifier did not produce significantly better hectarage estimates than the standardized USDA procedure when evaluated on either the training set or the independent test set. However, this classifier showed less sensitivity to the training/test degradation discussed earlier. This is evidenced by the fact that the hypothesis of equality of regression lines fitted to the training and test data sets was accepted for all crops except corn and permanent pasture, which failed the homogeneity of variances test. Also, the overall percent correct on the independent test set decreased least when using the MSE classifier. This greater extendibility might be expected due to the fewer parameters required to be estimated in using this classifier.

The calibration approach to regression points out a fundamental problem in the current regression model and suggests an alternative which has several theoretical advantages.

7.2 RECOMMENDATIONS

Several recommendations seem appropriate at the conclusion of this study. First, the use of CLASSY clustering in place of the current Editor clustering algorithm is recommended. CLASSY seems to offer a tangible improvement to the current Editor system in terms of increased performance and decreased analyst interaction.

Also, the study seems to indicate that the regression estimator may be improved. Use of a simpler classifier might make regression more extendible, and improved performance is expected using the calibration regression approach. Unfortunately, this issue is not clearly decided at this time. The recommendation is that additional research to improve regression/proportion estimation be conducted. This should include actual tests of the calibration approach as well as other alternative approaches. Such approaches include a regression model in which both ground truth and classification acreages are considered random and the use of direct proportion estimates. In an

operational setting, it is recommended that jackknifing be used to obtain more realistic performance estimates.

A final recommendation is that any future work be conducted with a larger data set, if possible. The sample size estimates reported in section 2.2 as well as our own experience in making the various tests indicate that the sample size of the Missouri data set is only marginally sufficient if testing is carried out on all the training data. The data set is not sufficiently large to achieve significant test results if it is divided into training and test portions. The sample size estimates reported in section 2.3 should serve as a guide in selecting future data sets and in designing future experiments.

8. REFERENCES

1. Lennington, R. K.; and Malek, H.: The CLASSY Clustering Algorithm - Description, Evaluation, and Comparison with the Iterative Self-Organizing Clustering System (ISOCLS). LEC-11289, March 1978.

2. Lennington, R. K.; and Rassbach, M. E.: CLASSY - An Adaptive Maximum Likelihood Clustering Algorithm. LEC-12145, May 1978, [presented at the Ninth Annual Meeting of the Classification Society (North American Branch), Clemson University (Clemson, South Carolina), May 21-23, 1978].

3. Lennington, R. K.; and Rassbach, M. E.: Mathematical Description and Program Documentation for CLASSY, An Adaptive Maximum Likelihood Clustering Method. LEC-12177 (JSC-14621), April 1979.

4. Lennington, R. K.; and Rassbach, M. E.: CLASSY - An Adaptive Maximum Likelihood Clustering Algorithm. Proceedings of Technical Sessions, Volume II, LACIE Symposium, October 1978, JSC-16015 (Houston, Texas), July 1979.

5. Duda, R. O.; and Hart, P. E.: Pattern Classification and Scene Analysis. John Wiley and Sons (New York), 1973.

6. Wolfe, J. H.: Pattern Classification by Multivariate Mixture Analysis. Multivariate Behavioral Research, vol. 5, 1970, pp. 329-350.

7. Thadani, S. G.: A Nonparametric Loss-Optimal Pattern Classification System. LEC-11451 (JSC-13901), April 13, 1978.

8. Berkson, J.: Estimation of a Linear Function for a Calibration Line. Technometrics, vol. 11, 1969, pp. 649-660.

9. Saw, J. C.: Letter to the Editor. Technometrics, vol. 12, 1970, p. 937.

APPENDIX A

CLIPPING LIMITS OF RADIANCE VALUES

TABLE A-1.- CORN CLIPPING LIMITS

Data type	Channels							
	2		4		6		8	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Standardized procedure								
August	12	19	42	71	*-	*-	*-	*-
May	*-	*-	*-	*-	25	45	24	60
Multitemporal	12	19	42	72	24	44	26	61
Independent test set procedure								
August	12	19	43	69	*-	*-	*-	*-
May	*-	*-	*-	*-	24	47	24	55
Multitemporal	12	19	43	68	22	47	24	57
Jackknifing training sets								
Multitemporal only								
1	12	19	43	73	21	47	24	62
2	12	19	43	70	23	45	24	61
3	12	19	43	70	23	44	24	56
4	12	19	42	70	24	44	32	61
5	12	20	41	73	24	46	24	61
6	12	21	41	73	23	45	24	61
7	12	19	41	70	23	45	24	61
8	12	19	41	71	24	42	24	61
9	12	19	41	73	24	44	24	56
10	12	19	41	73	24	44	24	61
11	12	19	41	70	24	44	24	61

*For unitemporal data, there are only four channel values.

TABLE A-2.- WINTER WHEAT CLIPPING LIMITS

Data type	Channels							
	2		4		6		8	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Standardized procedure								
August	19	33	36	65	*-	*-	*-	*-
May	*-	*-	*-	*-	15	30	52	82
Multitemporal	19	33	33	57	0	29	52	84
Independent test set procedure								
August	18	32	34	56	*-	*-	*-	*-
May	*-	*-	*-	*-	14	30	56	83
Multitemporal	19	32	34	56	14	30	59	83
Jackknifing training sets								
Multitemporal only								
1	15	36	30	58	14	32	51	84
2	18	32	34	58	14	29	52	80
3	18	33	34	57	14	29	56	79
4	18	33	35	57	14	29	55	80
5	15	33	35	58	14	35	51	84
6	16	34	34	57	14	30	59	84
7	15	33	34	58	16	31	54	78
8	17	33	34	58	14	29	55	80
9	18	33	34	58	14	29	55	80
10	18	33	34	58	14	29	52	80
11	18	32	30	58	14	29	55	83

*For unitemporal data, there are only four channel values.

TABLE A-3.- PERMANENT PASTURE CLIPPING LIMITS

Data type	Channels							
	2		4		6		8	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Standardized procedure								
August	0	26	30	73	*-	*-	*-	*-
May	*-	*-	*-	*-	16	35	45	103
Multitemporal	0	26	28	74	0	34	43	94
Independent test set procedure								
August	12	27	38	73	*-	*-	*-	*-
May	*-	*-	*-	*-	14	35	41	88
Multitemporal	12	27	38	72	12	34	41	87
Jackknifing training sets								
Multitemporal only								
1	12	26	31	75	12	39	41	94
2	12	26	31	76	13	36	43	87
3	12	24	33	72	9	35	18	98
4	12	26	31	76	9	37	18	104
5	12	28	31	77	9	42	18	103
6	12	28	33	76	9	37	18	104
7	12	26	33	76	9	38	18	104
8	12	26	32	76	9	35	18	103
9	12	26	33	76	9	34	18	102
10	12	26	30	76	9	35	18	104
11	12	27	37	78	9	38	18	104

*For unitemporal data, there are only four channel values.

TABLE A-4.- SOYBEANS CLIPPING LIMITS

Data type	Channels							
	2		4		6		8	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Standardized procedure								
August	0	22	47	99	*-	*-	*-	*-
May	*-	*-	*-	*-	23	48	31	65
Multitemporal	0	22	50	98	25	48	28	62
Independent test set procedure								
August	12	21	48	105	*-	*-	*-	*-
May	*-	*-	*-	*-	22	47	28	65
Multitemporal	12	24	44	105	21	47	28	63
Jackknifing training sets								
Multitemporal only								
1	11	23	47	98	21	48	28	63
2	11	23	46	101	22	44	28	64
3	11	23	46	98	23	47	28	64
4	11	23	47	105	22	45	28	65
5	11	24	47	98	18	48	28	67
6	11	23	47	105	15	47	28	72
7	11	22	39	105	22	46	29	64
8	11	22	49	98	22	45	28	61
9	11	22	46	98	23	45	28	65
10	11	22	47	98	24	46	30	63
11	12	23	47	98	24	47	28	65

*For unitemporal data, there are only four channel values.

TABLE A-5.- DENSE WOODLAND CLIPPING LIMITS

Data type	Channels							
	2		4		6		8	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Standardized procedure								
August	0	23	43	69	*-	*-	*-	*-
May	*-	*-	*-	*-	14	28	54	72
Multitemporal	10	23	44	68	0	30	51	73
Independent test set procedure								
August	10	19	48	67	*-	*-	*-	*-
May	*-	*-	*-	*-	16	27	53	74
Multitemporal	10	19	48	70	15	28	54	74
Jackknifing training sets								
Multitemporal only								
1	10	24	42	70	14	28	51	74
2	10	22	42	70	14	29	51	72
3	10	22	42	67	14	30	52	74
4	10	22	42	70	14	28	50	74
5	10	24	43	68	14	29	52	74
6	10	24	42	70	14	29	52	74
7	10	20	47	70	14	28	55	74
8	10	22	45	68	14	29	54	74
9	10	24	42	70	14	29	52	74
10	11	24	43	70	14	29	52	72
11	10	24	44	70	14	30	54	74

*For unitemporal data, there are only four channel values.

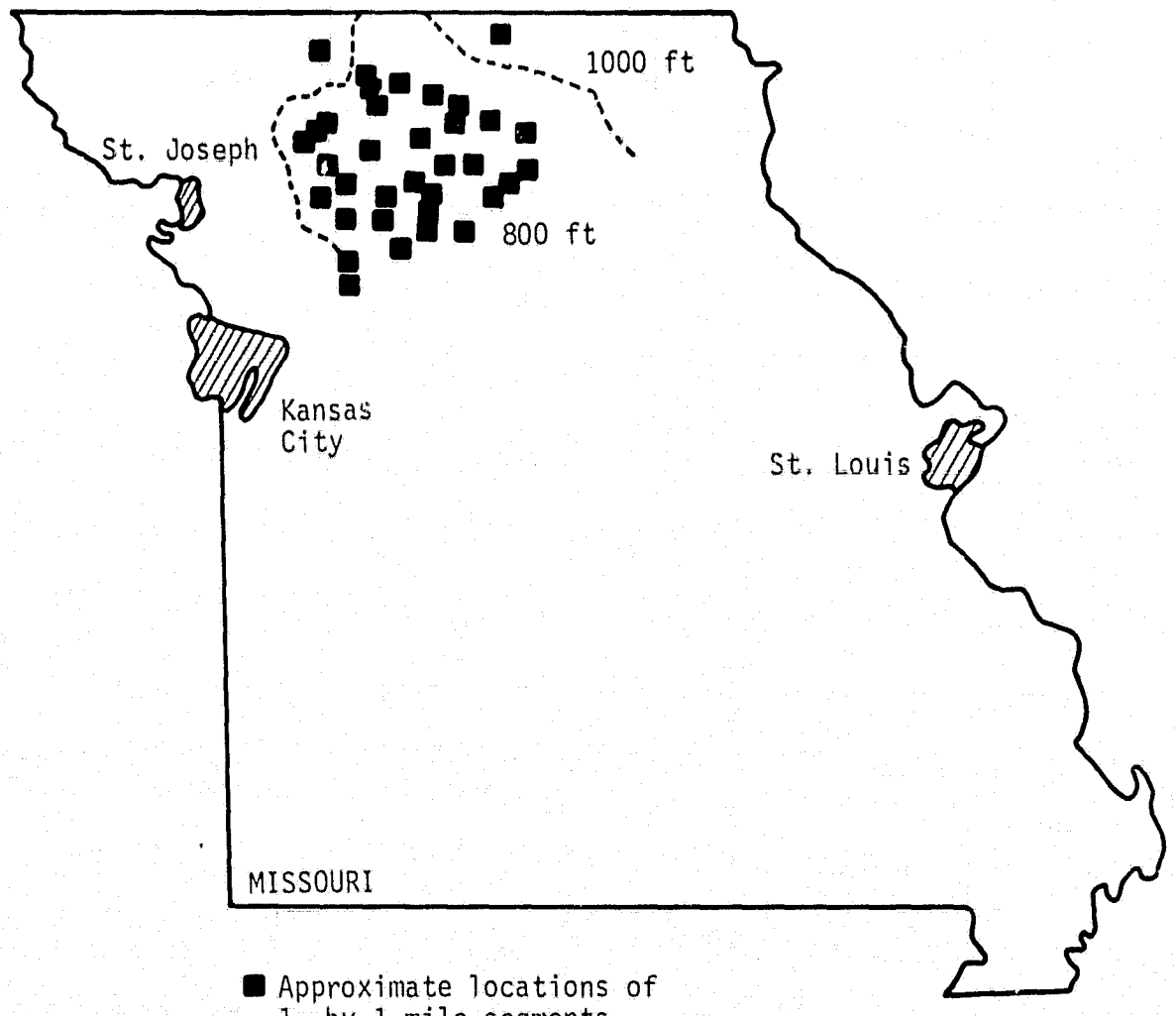
TABLE A-6.- OTHER HAY CLIPPING LIMITS

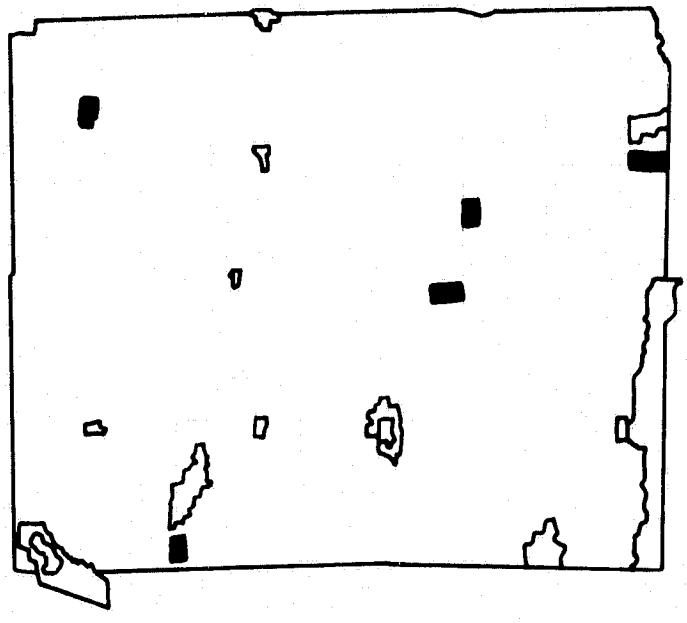
Data type	Channels							
	2		4		6		8	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Standardized procedure								
August	0	25	38	68	*-	*-	*-	*-
May	*-	*-	*-	*-	15	33	56	91
Multitemporal	13	27	38	68	0	32	56	91
Independent test set procedure								
August	13	28	37	68	*-	*-	*-	*-
May	*-	*-	*-	*-	13	32	54	99
Multitemporal	13	26	38	69	12	32	49	100
Jackknifing training sets								
Multitemporal only								
1	11	28	37	68	12	32	52	92
2	11	27	37	68	12	32	50	100
3	13	27	33	75	12	32	50	90
4	12	28	38	68	12	32	50	100
5	11	28	37	69	14	32	50	93
6	11	29	37	75	12	32	50	100
7	11	28	33	68	12	33	49	100
8	14	28	37	67	13	29	56	91
9	13	28	33	68	12	32	50	98
10	11	24	39	68	12	31	50	93
11	11	28	39	68	12	32	50	90

*For unitemporal data, there are only four channel values.

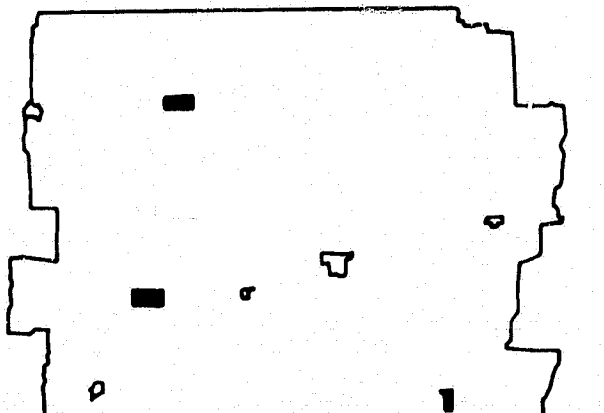
APPENDIX B

MAP OF MISSOURI WITH SEGMENT LOCATIONS

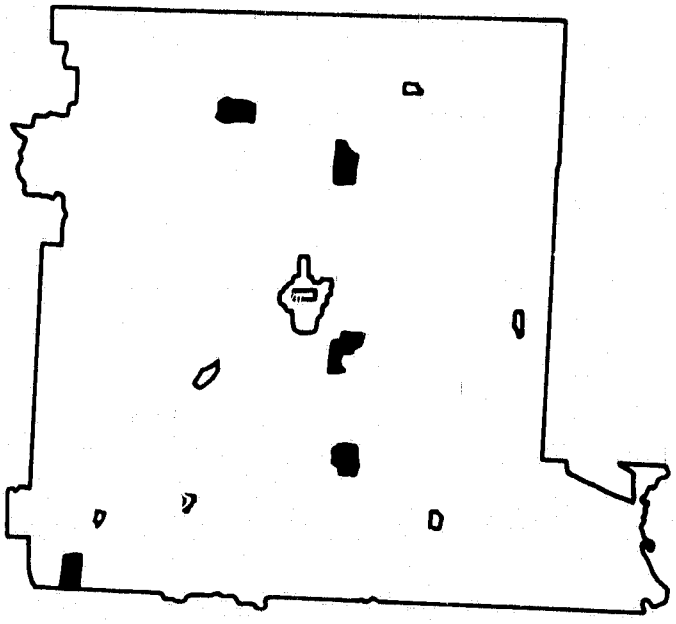

1000 ft
St. Joseph
800 ft
Kansas City
St. Louis
MISSOURI
Approximate locations of 1- by 1-mile segments



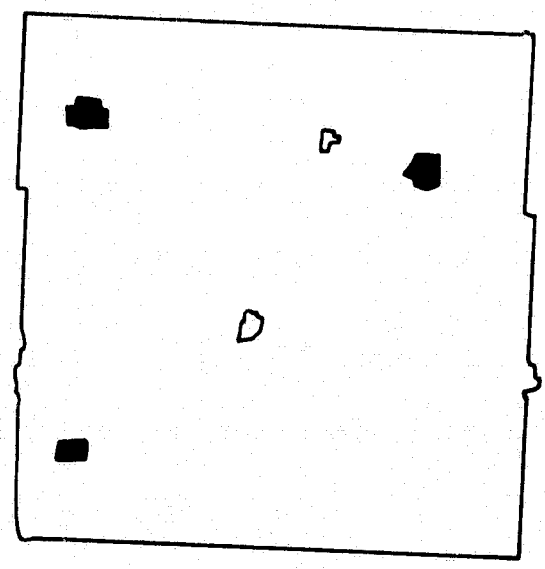
Linn County



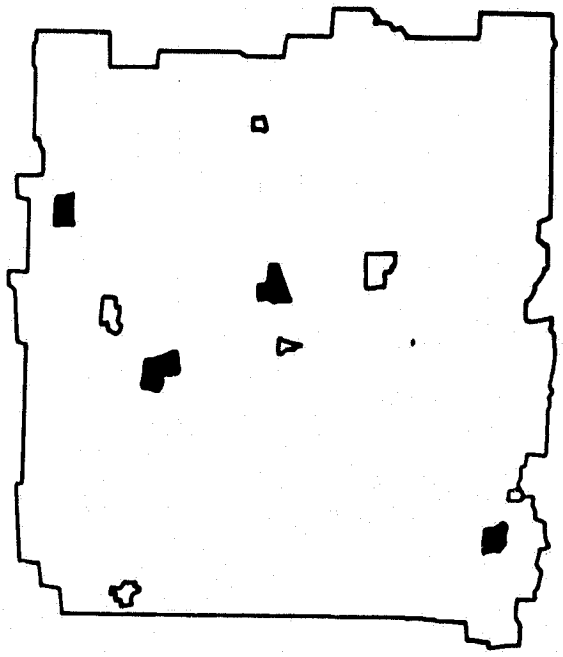
De Kalb County



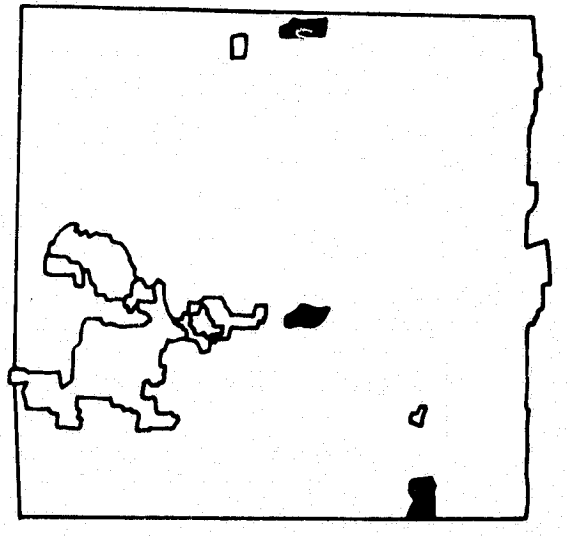
Livingston County



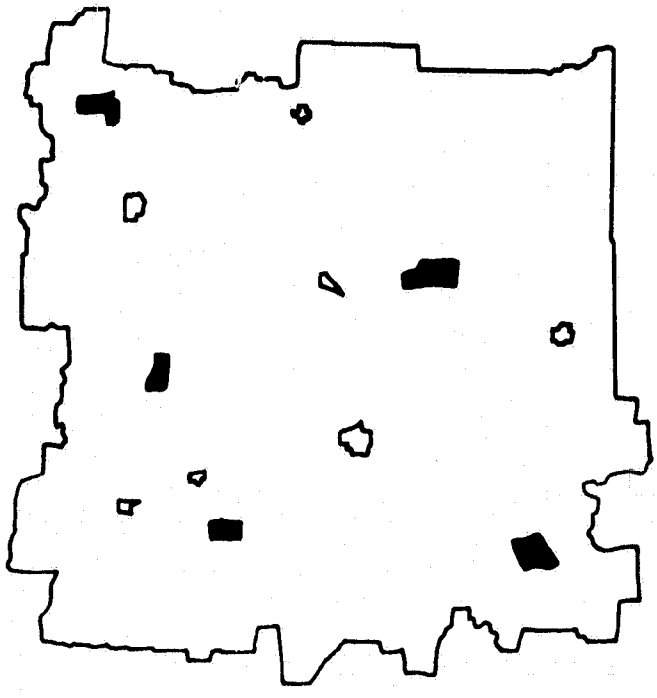
Mercer County



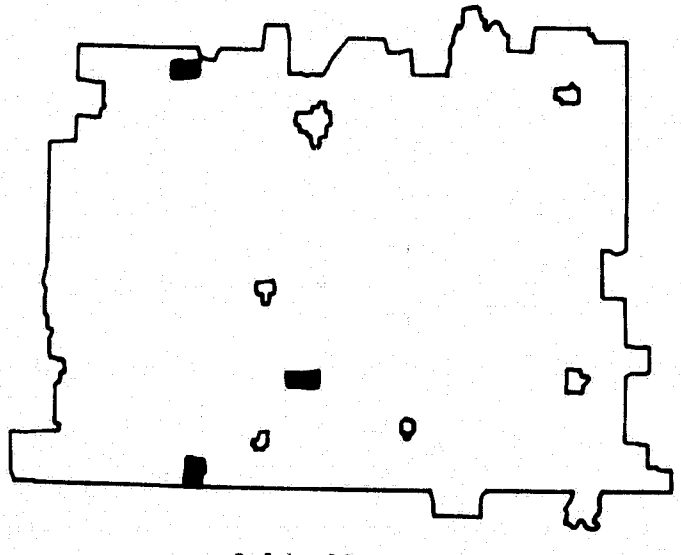
Gentry County



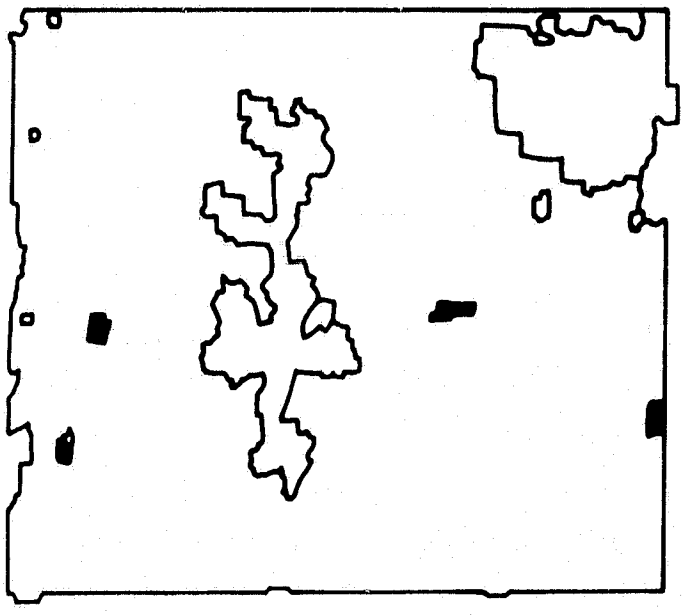
Grundy County



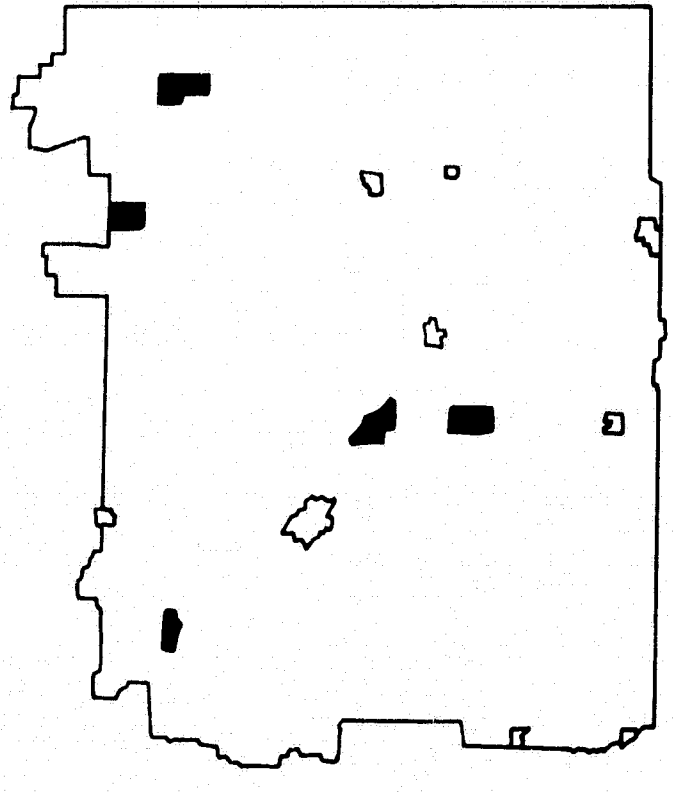
Daviess County



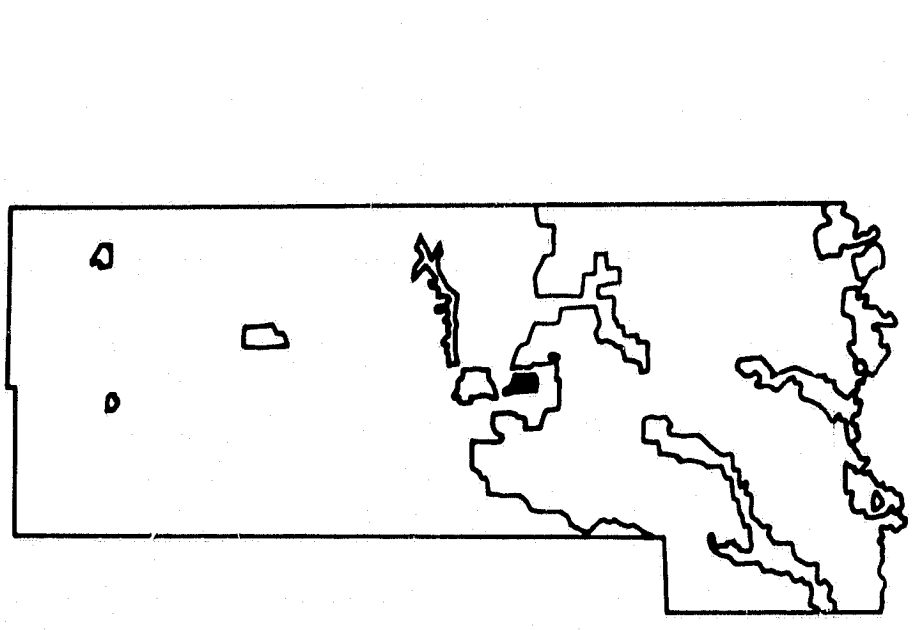
Caldwell County



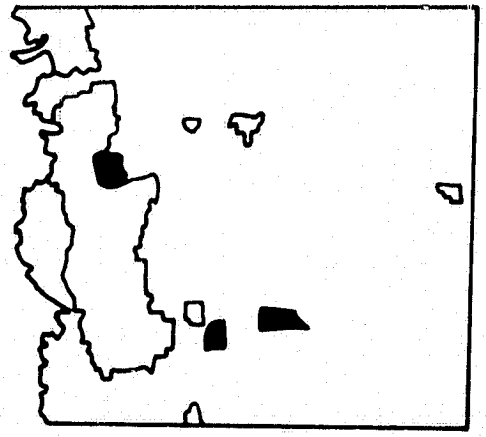
Sullivan County



Harrison County



Putnam County



Schuyler County

APPENDIX C

NUMBER OF CLUSTERS GENERATED

TABLE C-1.- NUMBER OF CLUSTERS GENERATED - TRAINING AND TESTING ON ALL 33 SEGMENTS

Crop	JSC*	USDA†	CLASSY (pure)	CLASSY (mixed)
Corn	2	7	6	7
Winter wheat	1	5	2	1
Permanent pasture	5	15	7	6
Soybeans	5	10	5	8
Dense woodland	1	7	3	4
Other hay	1	7	4	5

*At BBN.

†On Illiac.

APPENDIX D

MEAN PURE GROUND TRUTH PIXELS PER SEGMENT FOR JACKKNIFED TRAINING AND TEST SETS

TABLE D-1.- JACKKNIFING PROCEDURE - MEAN PURE GROUND TRUTH PIXELS PER SEGMENT IN THE 11 TRAINING AND TEST SETS

Training/ test sets	Corn		Winter wheat		Permanent pasture		Soybeans		Dense woodland		Other hay	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
1	21	12	6	9	76	80	55	39	17	0	16	39
2	21	9	6	0	73	117	59	6	15	26	19	11
3	21	9	6	3	78	58	57	26	14	29	18	17
4	19	30	6	1	83	12	49	103	17	0	20	0
5	21	4	6	0	79	49	49	97	16	7	17	28
6	21	8	5	11	73	110	54	55	16	11	19	6
7	18	39	5	14	71	130	54	55	15	22	20	0
8	20	17	6	0	80	45	55	43	16	16	19	6
9	17	48	6	1	81	28	52	73	16	13	18	13
10	20	19	5	12	72	119	56	28	13	44	17	26
11	20	20	5	11	75	94	52	67	15	17	19	8

APPENDIX E

ARCHIVED FILES

APPENDIX E

ARCHIVED FILES

Listed below are the file naming conventions used in naming files created by Editor.

FRAME.NAMES	list of Landsat scenes used
TASK2	33 segments split into two groups, 25 for training and 8 for testing
TASK3	jackknifing procedure using Editor
No task number	training and testing done on all 33 segments
CLASSY	CLASSY clustering algorithm used instead of Editor clustering
TRAIN	training set file
TEST	test set file
MAY	file using May acquisition only
AUG	file using August acquisition only
MTEMP or BILL	file using both acquisitions
SEGTOT or SGT	segment total file
EST or ESP or ESTPAR	estimated parameter file
STAT	statistics file
ISTAT	inverted statistics file
PACK	automatically packed file
CAT	categorized file
SCAT	list of files for scattergramming
TBL or TABLE	table file
CLIPPED	clipped file
APRIOR or PUR	prior probabilities specified

N-GRPS	N clusters
SEGS	list of segment numbers
NB	all pixels available
-NB	border pixels removed
CORN	corn pixels only
PERMANENTPASTURE or PASTURE	permanent pasture pixels only
DENSEWOODLAND or WOODS	dense woodland pixels only
OTHERHAY or HAY	other hay pixels only
WINTERWHEAT or WHEAT	winter wheat pixels only
SOYBEANS	soybean pixels only

Examples:

TASK3/TRAIN/5/PACK.-CORN/CLIPPED is a packed file of clipped corn pixels used as training data in the fifth of 11 jackknifing runs. (There are 30 segments of corn pixels in this file.)

BILL/WOODS.SGT is the segment totals file for crop dense woodland using multitemporal data when testing and training on all 33 segments.

TASK2/TEST/CLASSY/CAT. is the categorized file resulting from classification of the 8 test segments after CLASSY was used to cluster the 25 training segments.

<JSC>FRAME.NAMES;2

<JSC>TASK2/TRAIN.SEGS;2

<JSC>TASK2/TEST.SEGS;2

```
<JSC>MAY/WHEAT.ESTPAR;1
<JSC>MAY/WHEAT.SEGTOT;1
<JSC>MAY/TABLE.;4
<JSC>MAY/STAT.-WINTERWHEAT/CLIP;1
<JSC>MAY/STAT.15-GRPS/CLIP/EP;1
<JSC>MAY/STAT.-OTHERHAY/CLIP;1
<JSC>MAY/STAT.-DENSEWOODLAND/CLIP;1
<JSC>MAY/STAT.-SOYBEANS/CLIP;1
<JSC>MAY/STAT.-PERMANENTPASTURE/CLIP;1
<JSC>MAY/STAT.-CORN/CLIP;1
<JSC>MAY/SOYBEANS.ESTPAR;1
<JSC>MAY/SOYBEANS.SEGTOT;1
<JSC>MAY/PERMPAST.ESTPAR;1
<JSC>MAY/PERMPAST.SEGTOT;1
<JSC>MAY/PACK.-PERMANENTPASTURE/CLIP;1
<JSC>MAY/PACK.-OTHERHAY/CLIP;1
<JSC>MAY/PACK.-DENSEWOODLAND/CLIP;1
<JSC>MAY/PACK.-SOYBEANS/CLIP;1
<JSC>MAY/PACK.-WINTERWHEAT/CLIP;1
<JSC>MAY/PACK.-CORN/CLIP;1
<JSC>MAY/PACK.-OTHERHAY/M079;1
<JSC>MAY/PACK.-DENSEWOODLAND/M079;1
<JSC>MAY/PACK.-SOYBEANS/M079;1
<JSC>MAY/PACK.-PERMANENTPASTURE/M079;1
<JSC>MAY/PACK.-WINTERWHEAT/M079;1
<JSC>MAY/PACK.-CORN/M079;1
<JSC>MAY/PACK.-NB/M079;1
<JSC>MAY/PACK.NB/M079;1
<JSC>MAY/ISTAT.15-GRPS/PUR;1
<JSC>MAY/ISTAT.15-GRPS/EP;1
<JSC>MAY/HAY.ESTPAR;1
<JSC>MAY/HAY.SEGTOT;1
<JSC>MAY/DENSEWOODS.ESTPAR;1
<JSC>MAY/DENSEWOODS.SEGTOT;1
<JSC>MAY/CORN.ESTPAR;1
<JSC>MAY/CORN.SEGTOT;1
<JSC>MAY/CAT.;1
<JSC>MAY/TABLE.;3
```

```
<JSC>AUG/SOYBEANS.-ESTPARAMETER;2
<JSC>AUG/SOYBEANS.-SEGTOTAL;2
<JSC>AUG/SAVTABLE.;2
<JSC>AUG/CAT.NB-22GPS/PUE;1
<JSC>AUG/CAL.PRIOR;1
<JSC>AUG/WINTERWHEAT.-ESTPARAMETER;1
<JSC>AUG/WINTERWHEAT.-SEGTOTAL;1
<JSC>AUG/STAT.-22GPS/EP;1
<JSC>AUG/STAT.-OTHERHAY/CLIP;1
<JSC>AUG/STAT.-DENSEWOODLAND/CLIP;1
<JSC>AUG/STAT.-SOYBEANS/CLIP;1
<JSC>AUG/STAT.-PERMANENTPASTURE/CLIP;1
<JSC>AUG/STAT.-WINTERWHEAT/CLIP;1
<JSC>AUG/STAT.-CORN/CLIP;1
<JSC>AUG/SOYBEANS.-ESTPARAMETER;1
<JSC>AUG/SOYBEANS.-SEGTOTAL;1
<JSC>AUG/SAVTABLE.;1
<JSC>AUG/PERMPAST.-SEGTOTAL;1
<JSC>AUG/PACK.-SOYBEANS/CLIP;1
<JSC>AUG/PACK.-OTHERHAY/CLIP;1
<JSC>AUG/PACK.-DENSEWOODLAND/CLIP;1
<JSC>AUG/PACK.-PERMANENTPASTURE/CLIP;1
<JSC>AUG/PACK.-WINTERWHEAT/CLIP;1
<JSC>AUG/PACK.-CORN/CLIP;1
<JSC>AUG/PACK.-NB/M079;1
<JSC>AUG/PACK.NB/M079;1
<JSC>AUG/OTHERHAY.-ESTPARAMETER;1
<JSC>AUG/OTHERHAT.-SEGTOTAL;1
<JSC>AUG/ISTAT.-22GPS/PUR;1
<JSC>AUG/ISTAT.-22GPS/EP;1
<JSC>AUG/DENSEWD.-ESTPARAMETER;1
<JSC>AUG/DENSEWD.-SEGTOTAL;1
<JSC>AUG/CORN.-ESTPARAMETER;1
<JSC>AUG/CORN.-SEGTOTAL;1
```

```
 <JSC>BILL/WOODS.ESP;1
 <JSC>BILL/WOODS.SGT;1
 <JSC>BILL/WHEAT.ESP;1
 <JSC>BILL/WHEAT.SGT;1
 <JSC>BILL/TBL.;2
 <JSC>BILL/SOYBEANS.ESP;1
 <JSC>BILL/SOYBEANS.SGT;1
 <JSC>BILL/PASTURE.ESP;1
 <JSC>BILL/PASTURE.SGT;1
 <JSC>BILL/PACK.-OTHERHAY/CLIPPED;1
 <JSC>BILL/PACK.-DENSEWOODLAND/CLIPPED;1
 <JSC>BILL/PACK.-SOYBEANS/CLIPPED;1
 <JSC>BILL/PACK.-PERMANENTPASTURE/CLIPPED;1
*B10
 <JSC>BILL/PACK.-WINTERWHEAT/CLIPPED;1
 <JSC>BILL/PACK.-CORN/CLIPPED;1
 <JSC>BILL/PACK.-NB/M079;1
 <JSC>BILL/PACK.NB/M079;1
 <JSC>BILL/ISTAT.1GGRPS/APRIOR;2
 <JSC>BILL/HAY.ESP;1
 <JSC>BILL/HAY.SGT;1
 <JSC>BILL/CORN.ESP;1
 <JSC>BILL/CORN.SGT;1
 <JSC>BILL/CAT.;3
```

```
 <JSC.CLASSY/WOODS.EST/PUREPIX;1
 <JSC>CLASSY/WHEAT.EST/PUREPIX;1
 <JSC>CLASSY/TABLE.PUREPIX;1
 <JSC>CLASSY/STAT.-WINTERWHEAT/PUREPIX;1
 <JSC>CLASSY/STAT.-OTHERHAT/PUREPIX;1
 <JSC>CLASSY/STAT.-DENSEWOODLAND/PUREPIX;1
↑33
 <JSC>CLASSY/STAT.-SOYBEANS/PUREPIX;1
 <JSC>CLASSY/STAT.-CORN/PUREPIX;1
 <JSC>CLASSY/STAT.-PASTURE/PUREPIX;1
<JSC>CLASSY/SOYBEANS.EST/PUREPIX;1
<JSC>CLASSY/SOYBEANS.SGT/PUREPIX;1
<JSC>CLASSY/SGT.-WHEAT/PUREPIX;1
<JSC>CLASSY/SGT.PUREPIX;1
<JSC>CLASSY/PASTURE.EST/PUREPIX;1
<JSC>CLASSY/PASTURE.SGT/PUREPIX;1
<JSC>CLASSY/ISTAT.-27GRPS/PUREPIX/PRIORS;1
↑033
<JSC>CLASSY/ISTAT.-27-GRPS/PUREPIX;1
<JSC>CLASSY/HAY.EST/PUREPUIX;1
<JSC>CLASSY/HAY.SGT/PUREPIX;1
<JSC>CLASSY/CORN.EST/PUREPIX;1
<JSC>CLASSY/CAT.PUREPIX/PRIORS;1
<JSC>CLASSY/WOODS.EST;1
<JSC>CLASSY/WOODS.SGT;1
<JSC>CLASSY/WHEAT.EST;1
<JSC>CLASSY/WHEAT.SGT;1
<JSC>CLASSY/TABLE.;1
<JSC>CLASSY/STAT.FIXED;1
<JSC>CLASSY/SOYBEANS.EST;1
<JSC>CLASSY/SOYBEANS.SGT;1
<JSC>CLASSY/PASTURE.EST;1
<JSC>CLASSY/PASTURE.SGT;1
<JSC>CLASSY/PACK.NB/INCL/BORDERPIXELS;1
<JSC>CLASSY/ISTAT.PUR;1
<JSC>CLASSY/HAY.EST;1
<JSC>CLASSY/HAY.SGT;1
<JSC>CLASSY/CORN.EST;1
<JSC>CLASSY/CORN.SGT;1
<JSC>CLASSY/CAT.NB/PUR;1
```

```
 <JSC>TASK2/TRAIN/MAY/WINTERWHEAT.ESP;1
 <JSC>TASK2/TRAIN/MAY/WINTERWHEAT.SGT;1
 <JSC>TASK2/TRAIN/MAY/SOYBEANS.ESP;1
 <JSC>TASK2/TRAIN/MAY/SOYBEANS.SGT;1
 <JSC>TASK2/TRAIN/MAY/PERMANENTPASTURE.ESP;1
*0821
 <JSC>TASK2/TRAIN/MAY/PERMANENTPASTURE.SGT;1
*0821
 <JSC>TASK2/TRAIN/MAY/PACK.-OTHERHAY/CLIP;1
821
 <JSC>TASK2/TRAIN/MAY/PACK.-DENSEWOODLAND/CLIP;1
*d  20821
 <JSC>TASK2/TRAIN/MAY/PACK.-SOYBEANS/CLIP;1
*821
 <JSC>TASK2/TRAIN/MAY/PACK.-PERMANENTPASTURE/CLIP;1
* and  20821
 <JSC>TASK2/TRAIN/MAY/PACK.-WINTERWHEAT/CLIP;1
* 20821
 <JSC>TASK2/TRAIN/MAY/PACK.-CORN/CLIP;1
 <JSC>TASK2/TRAIN/MAY/PACK.-WINTERWHEAT/M079;1
* 20821
 <JSC>TASK2/TRAIN/MAY/PACK.-NB/M079;1
 <JSC>TASK2/TRAIN/MAY/PACK.NB/M079;1
 <JSC>TASK2/TRAIN/MAY/PACK.-OTHERHAY/M079;2
*821
 <JSC>TASK2/TRAIN/MAY/PACK.-DENSEWOODLAND/M079;2
*d  20821
 <JSC>TASK2/TRAIN/MAY/PACK.-SOYBEANS/M079;2
*821
 <JSC>TASK2/TRAIN/MAY/PACK.-PERMANENTPASTURE/M079;2
* and  20821
 <JSC>TASK2/TRAIN/MAY/PACK.-CORN/M079;2
 <JSC>TASK2/TRAIN/MAY/OTHERHAY.ESP;1
 <JSC>TASK2/TRAIN/MAY/OTHERHAY.SGT;1
 <JSC>TASK2/TRAIN/MAY/DENSEWOODLAND.ESP;1
*1
 <JSC>TASK2/TRAIN/MAY/DENSEWOODLAND.SGT;1
*1
 <JSC>TASK2/TRAIN/MAY/CORN.ESP;1
 <JSC>TASK2/TRAIN/MAY/CORN.SGT;1
 <JSC>TASK2/TRAIN/MAY/CAT.;1
 <JSC>TASK2/TRAIN/MAY/TBL.;1
 <JSC>TASK2/TRAIN/MAY/STAT.16-GRPS;1
 <JSC>TASK2/TRAIN/MAY/ISTAT.16-GRPS/CLIP;1
```

```
<JSC>TASK2/TEST/MAY/TBL.;1
<JSC>TASK2/TEST/MAY/CAT.;1
<JSC>TASK2/TEST/MAY/WINTERWHEAT.ESP;1
<JSC>TASK2/TEST/MAY/WINTERWHEAT.SGT;1
<JSC>TASK2/TEST/MAY/SOYBEANS.ESP;1
<JSC>TASK2/TEST/MAY/SOYBEANS.SGT;1
<JSC>TASK2/TEST/MAY/PERMANENTPASTURE.ESP;1
%021
<JSC>TASK2/TEST/MAY/PERMANENTPASTURE.SGT;1
%021
<JSC>TASK2/TEST/MAY/PACK.NB/MQ79;1
<JSC>TASK2/TEST/MAY/OTHERHAY.ESP;1
<JSC>TASK2/TEST/MAY/OTHERHAY.SGT;1
<JSC>TASK2/TEST/MAY/DENSEWOODLAND.ESP;1
<JSC>TASK2/TEST/MAY/DENSEWOODLAND.SGT;1
<JSC>TASK2/TEST/MAY/CORN.ESP;1
<JSC>TASK2/TEST/MAY/CORN.SGT;1
```

```
 <JSC>TASK2/TRAIN/AUG/WINTERWHEAT.ESP;1
 <JSC>TASK2/TRAIN/AUG/WINTERWHEAT.SGT;1
 <JSC>TASK2/TRAIN/AUG/SOYBEANS.ESP;1
 <JSC>TASK2/TRAIN/AUG/SOYBEANS.SGT;1
 <JSC>TASK2/TRAIN/AUG/PERMANENTPASTURE.ESP;1
*0821
 <JSC>TASK2/TRAIN/AUG/PERMANENTPASTURE.SGT;1
*0821
 <JSC>TASK2/TRAIN/AUG/PACK.-OTHERHAY/CLIP;1
*821
 <JSC>TASK2/TRAIN/AUG/PACK.-CORN/CLIP;1
 <JSC>TASK2/TRAIN/AUG/PACK.-DENSEWOODLAND/CLIP;1
*d  20821
 <JSC>TASK2/TRAIN/AUG/PACK.-WINTERWHEAT/CLIP;1
* 20821
 <JSC>TASK2/TRAIN/AUG/PACK.-PERMANENTPASTURE/CLIP;1
* and  20821
 <JSC>TASK2/TRAIN/AUG/PACK.-SOYBEANS/CLIP;1
*821
 <JSC>TASK2/TRAIN/AUG/PACK.-OTHERHAY/M079;1
*821
 <JSC>TASK2/TRAIN/AUG/PACK.-DENSEWOODLAND/M079;1
*d  20821
 <JSC>TASK2/TRAIN/AUG/PACK.-SOYBEANS/M079;1
*821
 <JSC>TASK2/TRAIN/AUG/PACK.-PERMANENTPASTURE/M079;1
* and  20821
 <JSC>TASK2/TRAIN/AUG/PACK.-WINTERWHEAT/M079;1
* 20821
 <JSC>TASK2/TRAIN/AUG/PACK.-CORN/M079;1
 <JSC>TASK2/TRAIN/AUG/PACK.-NB/M079;1
 <JSC>TASK2/TRAIN/AUG/PACK.NB/M079;1
 <JSC>TASK2/TRAIN/AUG/OTHERHAY.ESP;1
 <JSC>TASK2/TRAIN/AUG/OTHERHAY.SGT;1
 <JSC>TASK2/TRAIN/AUG/DENSEWOODLAND.ESP;1
*1
 <JSC>TASK2/TRAIN/AUG/DENSEWOODLAND.SGT;1
*1
 <JSC>TASK2/TRAIN/AUG/CORN.ESP;1
 <JSC>TASK2/TRAIN/AUG/CORN.SGT;1
 <JSC>TASK2/TRAIN/AUG/TBL.;1
 <JSC>TASK2/TRAIN/AUG/STAT.1G-GRPS;1
 <JSC>TASK2/TRAIN/AUG/ISTAT.1G-GRPS/CLIP;1
 <JSC>TASK2/TRAIN/AUG/CAT.;1
```

```
<JSC>TASK2/TEST/AUG/WINTERWHEAT.SGT;1
<JSC>TASK2/TEST/AUG/WINTERWHEAT.ESP;2
<JSC>TASK2/TEST/AUG/SOYBEANS.ESP;1
<JSC>TASK2/TEST/AUG/SOYBEANS.SGT;1
<JSC>TASK2/TEST/AUG/PERMANENTPASTURE.ESP;1
-021
<JSC>TASK2/TEST/AUG/PERMANENTPASTURE.SGT;1
-021
<JSC>TASK2/TEST/AUG/PACK.NB/M070;1
<JSC>TASK2/TEST/AUG/OTHERHAY.ESP;1
<JSC>TASK2/TEST/AUG/OTHERHAY.SGT;1
<JSC>TASK2/TEST/AUG/DENSEWOODLAND.ESP;1
<JSC>TASK2/TEST/AUG/DENSEWOODLAND.SGT;1
<JSC>TASK2/TEST/AUG/CORN.ESP;1
<JSC>TASK2/TEST/AUG/CORN.SGT;1
<JSC>TASK2/TEST/AUG/CAT.;1
<JSC>TASK2/TEST/AUG/TBL.;1
```

```
 <JSC>TASK2/TRAIN/MTEMP/TBL1.;1
 <JSC>TASK2/TRAIN/MTEMP/STAT.16-GRPS;1
 <JSC>TASK2/TRAIN/MTEMP/ISTAT.16-GRPS/CLIP;1
*0027
 <JSC>TASK2/TRAIN/MTEMP/CAT.;1
 <JSC>TASK2/TRAIN/MTEMP/WINTERWHEAT.ESP;1
*1
 <JSC>TASK2/TRAIN/MTEMP/WINTERWHEAT.SGT;1
*1
 <JSC>TASK2/TRAIN/MTEMP/SOYBEANS.ESP;1
 <JSC>TASK2/TRAIN/MTEMP/SOYBEANS.SGT;1
 <JSC>TASK2/TRAIN/MTEMP/PERMANENTPASTURE.ESP;1
* 20821
 <JSC>TASK2/TRAIN/MTEMP/PERMANENTPASTURE.SGT;1
* 20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-OTHERHAY/CLIP;1
*20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-DENSEWOODLAND/CLIP;1
*and  20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-SOYBEANS/CLIP;1
*20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-PERMANENTPASTURE/CLIP;1
*9  and  20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-WINTERWHEAT/CLIP;1
*d  20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-CORN/CLIP;1
*1
 <JSC>TASK2/TRAIN/MTEMP/PACK.-OTHERHAY/M079;1
*20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-DENSEWOODLAND/M079;1
*and  20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-SOYBEANS/M079;1
*20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-PERMANENTPASTURE/M079;1
*9  and  20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-WINTERWHEAT/M079;1
*d  20821
 <JSC>TASK2/TRAIN/MTEMP/PACK.-CORN/M079;1
*1
 <JSC>TASK2/TRAIN/MTEMP/PACK.-NB/M079;1
 <JSC>TASK2/TRAIN/MTEMP/PACK.NB/M079;1
 <JSC>TASK2/TRAIN/MTEMP/OTHERHAY.ESP;1
 <JSC>TASK2/TRAIN/MTEMP/OTHERHAY.SGT;1
 <JSC>TASK2/TRAIN/MTEMP/DENSEWOODLAND.ESP;1
*821
 <JSC>TASK2/TRAIN/MTEMP/DENSEWOODLAND.SGT;1
*821
 <JSC>TASK2/TRAIN/MTEMP/CORN.ESP;1
 <JSC>TASK2/TRAIN/MTEMP/CORN.SGT;1
```

```
 <JSC>TASK2/TEST/MTEMP/TBL.;2
 <JSC>TASK2/TEST/MTEMP/CAT.;1
 <JSC>TASK2/TEST/MTEMP/WINTERWHEAT.ESP;1
 <JSC>TASK2/TEST/MTEMP/WINTERWHEAT.SGT;1
 <JSC>TASK2/TEST/MTEMP/SOYBEANS.ESP;1
 <JSC>TASK2/TEST/MTEMP/SOYBEANS.SGT;1
 <JSC>TASK2/TEST/MTEMP/PERMANENTPASTURE.ESP;1
*20821
 <JSC>TASK2/TEST/MTEMP/PERMANENTPASTURE.SGT;1
*20821
 <JSC>TASK2/TEST/MTEMP/PACK.NB/M079;1
 <JSC>TASK2/TEST/MTEMP/OTHERHAY.ESP;1
 <JSC>TASK2/TEST/MTEMP/OTHERHAY.SGT;1
 <JSC>TASK2/TEST/MTEMP/DENSEWOODLAND.ESP;1
*21
 <JSC>TASK2/TEST/MTEMP/DENSEWOODLAND.SGT;1
*21
 <JSC>TASK2/TEST/MTEMP/CORN.SGT;1
 <JSC>TASK2/TEST/MTEMP/CORN.ESP;1
```

```
 <JSC>TASK2TEST/CLASSY/SOYBEANS.EST;1
 <JSC>TASK2/TRAIN/CLASSY/WOODS.EST;1
 <JSC>TASK2/TRAIN/CLASSY/WOODS.SGT;1
 <JSC>TASK2/TRAIN/CLASSY/WHEAT.SGT;1
 <JSC>TASK2/TRAIN/CLASSY/WHEAT.EST;1
 <JSC>TASK2/TRAIN/CLASSY/SOYBEANS.EST;1
 <JSC>TASK2/TRAIN/CLASSY/SOYBEANS.SGT;1
 <JSC>TASK2/TRAIN/CLASSY/PASTURE.EST;1
 <JSC>TASK2/TRAIN/CLASSY/PASTURE.SGT;1
 <JSC>TASK2/TRAIN/CLASSY/HAY.EST;1
 <JSC>TASK2/TRAIN/CLASSY/HAY.SGT;1
 <JSC>TASK2/TRAIN/CLASSY/CORN.EST;1
 <JSC>TASK2/TRAIN/CLASSY/CORN.SGT;1
 <JSC>TASK2/TRAIN/CLASSY/CAT.;1
 <JSC>TASK2/TEST/CLASSY/WOODS.EST;1
 <JSC>TASK2/TEST/CLASSY/WOODS.SGT;1
 <JSC>TASK2/TEST/CLASSY/WHEAT.EST;1
 <JSC>TASK2/TEST/CLASSY/WHEAT.SGT;1
 <JSC>TASK2/TEST/CLASSY/SOYBEANS.SGT;1
 <JSC>TASK2/TEST/CLASSY/PASTURE.EST;1
 <JSC>TASK2/TEST/CLASSY/PASTURE.SGT;1
 <JSC>TASK2/TEST/CLASSY/HAY.EST;1
 <JSC>TASK2/TEST/CLASSY/HAY.SGT;1
 <JSC>TASK2/TEST/CLASSY/CORN.EST;1
 <JSC>TASK2/TEST/CLASSY/CORN.SGT;1
 <JSC>TASK2/TEST/CLASSY/CAT.;1
 <JSC>CLASSY/TRAIN/TABLE.;1
 <JSC>CLASSY/TEST/TABLE.;1
 <JSC>TASK2/TRAIN/CLASSY/ISTAT.-24GRPS/PUR;1
#0853
 <JSC>TASK2/TRAIN/CLASSY/ISTAT.-24GRPS/EP;1
#853
 <JSC>TASK2/TRAIN/CLASSY.BOUT;1
 <JSC>TASK2/TEST/CLASSY.BOUT;1
 <JSC>TASK2/CLASSY/STAT.-OTHER;1
 <JSC>TASK2/CLASSY/STAT.-PERMANENT;1
 <JSC>TASK2/CLASSY/STAT.-SOYBEANS;1
 <JSC>TASK2/CLASSY/STAT.-SOYBEANS;2
 <JSC>TASK2/CLASSY/STAT.-CORN;1
 <JSC>TASK2/CLASSY/STAT.-WINTER;1
 <JSC>TASK2/CLASSY/STAT.-DENSE;1
```

```
<JSC>TASK3/TRAIN/1.SEQS;4


 <JSC>TASK3/TRAIN/1/PACK.-OTHERHAY/CLIP;1
*7
 <JSC>TASK3/TRAIN/1/PACK.-DENSEWOODLAND/CLIP;1
* 20827
 <JSC>TASK3/TRAIN/1/PACK.-PERMANENTPASTURE/CLIP;1
*nd  20827
 <JSC>TASK3/TRAIN/1/PACK.-SOYBEANS/CLIP;1
*7
 <JSC>TASK3/TRAIN/1/PACK.-WINTERWHEAT/CLIP;1
*0827
 <JSC>TASK3/TRAIN/1/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/1/PACK.-OTHERHAY/M079;1
*7
 <JSC>TASK3/TRAIN/1/PACK.-DENSEWOODLAND/M079;1
* 20827
 <JSC>TASK3/TRAIN/1/PACK.-SOYBEANS/M079;1
*7
 <JSC>TASK3/TRAIN/1/PACK.-PERMANENTPASTURE/M079;1
*nd  20827
 <JSC>TASK3/TRAIN/1/PACK.-WINTERWHEAT/M079;1
*0827
 <JSC>TASK3/TRAIN/1/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/1/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/1/PACK.NB/M079;2


 <JSC>TASK3/TRAIN/1/WINTERWHEAT.ESP;1
 <JSC>TASK3/TRAIN/1/WINTERWHEAT.SGT;1
 <JSC>TASK3/TRAIN/1/TBL.;1
 <JSC>TASK3/TRAIN/1/STAT.16-GRPS;1
 <JSC>TASK3/TRAIN/1/STAT.-OTHERHAY/CLIP;1
*7
 <JSC>TASK3/TRAIN/1/STAT.-DENSEWOODLAND/CLIP;1
* 20827
 <JSC>TASK3/TRAIN/1/STAT.-SOYBEANS/CLIP;1
*7
 <JSC>TASK3/TRAIN/1/STAT.-PERMANENTPASTURE/CLIP;1
*nd  20827
 <JSC>TASK3/TRAIN/1/STAT.-WINTERWHEAT/CLIP;1
*0827
 <JSC>TASK3/TRAIN/1/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/1/SOYBEANS.ESP;1
 <JSC>TASK3/TRAIN/1/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/1/PERMANENTPASTURE.ESP;1
*27
 <JSC>TASK3/TRAIN/1/PERMANENTPASTURE.SGT;1
*27
 <JSC>TASK3/TRAIN/1/OTHERHAY.ESP;1
 <JSC>TASK3/TRAIN/1/OTHERHAY.SGT;1
 <JSC>TASK3/TRAIN/1/ISTAT.16-GRPS/CLIP;1
 <JSC>TASK3/TRAIN/1/ISTAT.16-GRPS;1
 <JSC>TASK3/TRAIN/1/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TRAIN/1/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TRAIN/1/CORN.ESP;1
 <JSC>TASK3/TRAIN/1/CORN.SGT;1
 <JSC>TASK3/TRAIN/1/CAT.;1
```

```
 <JSC>TASK3/TEST/1/WINTERWHEAT.SGT;1
 <JSC>TASK3/TEST/1/WINTERWHEAT.ESP;1
 <JSC>TASK3/TEST/1/TBL.;1
 <JSC>TASK3/TEST/1/SOYBESNS.ESP;1
 <JSC>TASK3/TEST/1/SOYBEANS.SGT;1
 <JSC>TASK3/TEST/1/PERMANENTPASTURE.ESP;1
*7
 <JSC>TASK3/TEST/1/PERMANENTPASTURE.SGT;1
*7
 <JSC>TASK3/TEST/1/PACK.NB/M079;1
 <JSC>TASK3/TEST/1/OTHERHAY.ESP;1
 <JSC>TASK3/TEST/1/OTHERHAY.SGT;1
 <JSC>TASK3/TEST/1/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TEST/1/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TEST/1/CORN.ESP;1
 <JSC>TASK3/TEST/1/CORN.SGT;1
 <JSC>TASK3/TEST/1/CAT.;1
 <JSC>TASK3/TEST/1.SEGS;2
```

```
 <JSC>TASK3/TRAIN/2.SEGS;2
 <JSC>TASK3/TRAIN/2/WINTERWHEAT.ESP;1
 <JSC>TASK3/TRAIN/2/WINTERWHEAT.SGT;1
 <JSC>TASK3/TRAIN/2/TBL.;1
 <JSC>TASK3/TRAIN/2/STAT.16-GRPS;1
 <JSC>TASK3/TRAIN/2/STAT.-OTHERHAY/CLIP;1
*3
 <JSC>TASK3/TRAIN/2/STAT.-DENSEWOODLAND/CLIP;1
* 20833
 <JSC>TASK3/TRAIN/2/STAT.-SOYBEANS/CLIP;1
*3
 <JSC>TASK3/TRAIN/2/STAT.-PERMANENTPASTURE/CLIP;1
*nd  20833
 <JSC>TASK3/TRAIN/2/STAT.-WINTERWHEAT/CLIP;1
*0833
 <JSC>TASK3/TRAIN/2/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/2/SOYBEANS.ESP;1
 <JSC>TASK3/TRAIN/2/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/2/PERMANENTPASTURE.ESP;1
*33
 <JSC>TASK3/TRAIN/2/PERMANENTPASTURE.SGT;1
*33
 <JSC>TASK3/TRAIN/2/PACK.-SOYBEANS/CLIP;1
*3
 <JSC>TASK3/TRAIN/2/PACK.-OTHERHAY/CLIP;1
*3
 <JSC>TASK3/TRAIN/2/PACK.-DENSEWOODLAND/CLIP;1
* 20833
 <JSC>TASK3/TRAIN/2/PACK.-PERMANENTPASTURE/CLIP;1
*nd  20833
 <JSC>TASK3/TRAIN/2/PACK.-WINTERWHEAT/CLIP;1
*0833
 <JSC>TASK3/TRAIN/2/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/2/PACK.-OTHERHAY/M079;1
*3
 <JSC>TASK3/TRAIN/2/PACK.-DENSEWOODLAND/M079;1
* 20833
 <JSC>TASK3/TRAIN/2/PACK.-SOYBEANS/M079;1
*3
 <JSC>TASK3/TRAIN/2/PACK.-PERMANENTPASTURE/M079;1
*nd  20833
 <JSC>TASK3/TRAIN/2/PACK.-WINTERWHEAT/M079;1
*0833
 <JSC>TASK3/TRAIN/2/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/2/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/2/PACK.NB/M079;1
 <JSC>TASK3/TRAIN/2/OTHERHAY.ESP;1
 <JSC>TASK3/TRAIN/2/OTHERHAY.SGT;1
 <JSC>TASK3/TRAIN/2/ISTAT.16-GRPS/CLIP;1
 <JSC>TASK3/TRAIN/2/ISSTAT.16-GRPS;1
 <JSC>TASK3/TRAIN/2/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TRAIN/2/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TRAIN/2/CORN.ESP;1
 <JSC>TASK3/TRAIN/2/CORN.SGT;1
 <JSC>TASK3/TRAIN/2/CAT.;1
 <JSC>TASK3/TRAIN/2.SCAT-FILES;1
 <JSC>TASK3/TRAIN/2.SEGS;1
```

```
  <JSC>TASK3/TEST/2/WINTERWHEAT.ESP;1
  <JSC>TASK3/TEST/2/WINTERWHEAT.SGT;1
  <JSC>TASK3/TEST/2/TBL.;1
  <JSC>TASK3/TEST/2/SOYBEANS.ESP;1
  <JSC>TASK3/TEST/2/SOYBEANS.SGT;1
  <JSC>TASK3/TEST/2/PERMANENTPASTURE.ESP;1
**3
  <JSC>TASK3/TEST/2/PERMANENTPASTURE.SGT;1
**3
  <JSC>TASK3/TEST/2/PACK.NB/M079;1
  <JSC>TASK3/TEST/2/OTHERHAY.ESP;1
  <JSC>TASK3/TEST/2/OTHERHAY.ESP;2
  <JSC>TASK3/TEST/2/OTHERHAY.SGT;1
  <JSC>TASK3/TEST/2/OTHERHAY.SGT;2
  <JSC>TASK3/TEST/2/DENSEWOODLAND.ESP;1
  <JSC>TASK3/TEST/2/DENSEWOODLAND.SGT;1
  <JSC>TASK3/TEST/2/CORN.ESP;1
  <JSC>TASK3/TEST/2/CORN.SGT;1
  <JSC>TASK3/TEST/2/CAT.;1
  <JSC>TASK3/TEST/2.SCAT-FILES;1
  <JSC>TASK3/TEST/2.SEGS;2
```

```
 <JSC>TASK3/TRAIN/3/WINTERWHEAT.ESP;1
 <JSC>TASK3/TRAIN/3/WINTERWHEAT.SGT;1
 <JSC>TASK3/TRAIN/3/TBL.;1
 <JSC>TASK3/TRAIN/3/STAT.IG-GRPS;1
 <JSC>TASK3/TRAIN/3/STAT.-OTHERHAY/CLIP;1
*3
 <JSC>TASK3/TRAIN/3/STAT.-DENSEWOODLAND/CLIP;1
* 20833
 <JSC>TASK3/TRAIN/3/STAT.-SOYBEANS/CLIP;1
*3
 <JSC>TASK3/TRAIN/3/STAT.-PERMANENTPASTURE/CLIP;1
*nd  20833
 <JSC>TASK3/TRAIN/3/STAT.-WINTERWHEAT/CLIP;1
*0833
 <JSC>TASK3/TRAIN/3/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/3/SOYBEANS.ESP;1
 <JSC>TASK3/TRAIN/3/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/3/PERMANENTPASTURE.ESP;1
*33
 <JSC>TASK3/TRAIN/3/PERMANENTPASTURE.SGT;1
*33
 <JSC>TASK3/TRAIN/3/PACK.-OTHERHAY/CLIPO;1
*33
 <JSC>TASK3/TRAIN/3/PACK.-DENSEWOODLAND/CLIP;1
* 20833
 <JSC>TASK3/TRAIN/3/PACK.-SOYBEANS/CLIP;1
*3
 <JSC>TASK3/TRAIN/3/PACK.-PERMANENTPASTURE/CLIPO;1
*and  20833
 <JSC>TASK3/TRAIN/3/PACK.-WINTERWHEAT/CLIP;1
*0833
 <JSC>TASK3/TRAIN/3/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/3/PACK.-OTHERHAY/MO79;1
*3
 <JSC>TASK3/TRAIN/3/PACK.-DENSEWOODLAND/MO79;1
* 20833
 <JSC>TASK3/TRAIN/3/PACK.-SOYBEANS/MO79;1
*3
 <JSC>TASK3/TRAIN/3/PACK.-PERMANENTPASTURE/MO79;1
*nd  20833
 <JSC>TASK3/TRAIN/3/PACK.-WINTERWHEAT/MO79;1
*0833
 <JSC>TASK3/TRAIN/3/PACK.-CORN/MO79;1
 <JSC>TASK3/TRAIN/3/PACK.-NB/MO79;1
 <JSC>TASK3/TRAIN/3/PACK.NB/MO79;1
 <JSC>TASK3/TRAIN/3/OTHERHAY.ESP;1
 <JSC>TASK3/TRAIN/3/OTHERHAY.SGT;1
 <JSC>TASK3/TRAIN/3/ISTAT.IG-GRPS/CLIP;1
 <JSC>TASK3/TRAIN/3/ISTAT.IG-GRPS;1
 <JSC>TASK3/TRAIN/3/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TRAIN/3/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TRAIN/3/CORN.ESP;1
 <JSC>TASK3/TRAIN/3/CORN.SGT;1
 <JSC>TASK3/TRAIN/3/CAT.;1
 <JSC>TASK3/TRAIN/3.SCAT-FILES;1
 <JSC>TASK3/TRAIN/3.SEQS;2
```

```
 <JSC>TASK3/TEST/3/WINTERWHEAT.ESP;1
 <JSC>TASK3/TEST/3/WINTERWHEAT.SGT;1
 <JSC>TASK3/TEST/3/TBL.;1
 <JSC>TASK3/TEST/3/SOYBEANS.ESP;1
 <JSC>TASK3/TEST/3/SOYBEANS.SGT;1
 <JSC>TASK3/TEST/3/PERMANENTPASTURE.ESP;1
*3
 <JSC>TASK3/TEST/3/PERMANENTPASTURE.SGT;1
*3
 <JSC>TASK3/TEST/3/PACK.NB/M079;1
 <JSC>TASK3/TEST/3/OTHERHAY.ESP;1
 <JSC>TASK3/TEST/3/OTHERHAY.SGT;1
 <JSC>TASK3/TEST/3/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TEST/3/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TEST/3/CORN.ESP;1
 <JSC>TASK3/TEST/3/[illegible]N.SGT;1
 <JSC>TASK3/TEST/3/CAT.;1
 <JSC>TASK3/TEST/3.SCAT-FILES;1
 <JSC>TASK3/TEST/3.SEGS;2  .
```

```
 <JSC>TASK3/TRAIN/4/WINTERWHEAT.ESP;1
 <JSC>TASK3/TRAIN/4/WINTERWHEAT.SGT;1
 <JSC>TASK3/TRAIN/4/TBL.;1
 <JSC>TASK3/TRAIN/4/STAT.16-GRPS;1
 <JSC>TASK3/TRAIN/4/STAT.-OTHERHAY/CLIP;1
*2
 <JSC>TASK3/TRAIN/4/STAT.-DENSEWOODLAND/CLIP;1
* 20842
 <JSC>TASK3/TRAIN/4/STAT.-SOYBEANS/CLIP;1
*2
 <JSC>TASK3/TRAIN/4/STAT.-PERMANENTPASTURE/CLIP;1
*nd  20842
 <JSC>TASK3/TRAIN/4/STAT.-WINTERWHEAT/CLIP;1
*0842
 <JSC>TASK3/TRAIN/4/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/4/SOYBEANS.ESP;1
 <JSC>TASK3/TRAIN/4/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/4/PERMANENTPASTURE.ESP;1
*42
 <JSC>TASK3/TRAIN/4/PERMANENTPASTURE.SGT;1
*42
 <JSC>TASK3/TRAIN/4/PACK.-OTHERHAY/CLIP;1
*2
 <JSC>TASK3/TRAIN/4/PACK.-DENSEWOODLAND/CLIP;1
* 20842
 <JSC>TASK3/TRAIN/4/PACK.-PERMANENTPASTURE/CLIP;1
*nd  20842
 <JSC>TASK3/TRAIN/4/PACK.-WINTERWHEAT/CLIP;1
*0842
 <JSC>TASK3/TRAIN/4/PACK.-SOYBEANS/CLIP;1
*2
 <JSC>TASK3/TRAIN/4/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/4/PACK.-OTHERHAY/M079;1
*2
 <JSC>TASK3/TRAIN/4/PACK.-DENSEWOODLAND/M079;1
* 20842
 <JSC>TASK3/TRAIN/4/PACK.-SOYBEANS/M079;1
*2
 <JSC>TASK3/TRAIN/4/PACK.-PERMANENTPASTURE/M079;1
*nd  20842
 <JSC>TASK3/TRAIN/4/PACK.-WINTERWHEAT/M079;1
*0842
 <JSC>TASK3/TRAIN/4/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/4/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/4/PACK.NB/M079;1
 <JSC>TASK3/TRAIN/4/OTHERHAY.ESP;1
 <JSC>TASK3/TRAIN/4/OTHERHAY.SGT;1
 <JSC>TASK3/TRAIN/4/ISTAT.16-GRPS/CLIP;1
 <JSC>TASK3/TRAIN/4/ISTAT.16-GRPS;1
 <JSC>TASK3/TRAIN/4/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TRAIN/4/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TRAIN/4/CORN.ESP;1
 <JSC>TASK3/TRAIN/4/CORN.SGT;1
 <JSC>TASK3/TRAIN/4/CAT.;1
 <JSC>TASK3/TRAIN/4.SCAT-FILES;1
 <JSC>TASK3/TRAIN/4.SEGS;2
```

```
<JSC>TASK3/TEST/4/WINTERWHEAT.ESP;1
<JSC>TASK3/TEST/4/WINTERWHEAT.SGT;1
<JSC>TASK3/TEST/4/TBL.;1
<JSC>TASK3/TEST/4/SOYBEANS.ESP;1
<JSC>TASK3/TEST/4/SOYBEANS.SGT;1
<JSC>TASK3/TEST/4/PERMANENTPASTURE.ESP;1
*2
<JSC>TASK3/TEST/4/PERMANENTPASTURE.SGT;1
*2
<JSC>TASK3/TEST/4/PACK.NB/M079;1
<JSC>TASK3/TEST/4/OTHERHAY.ESP;1
<JSC>TASK3/TEST/4/OTHERHAY.SGT;1
<JSC>TASK3/TEST/4/DENSEWOODLAND.SGT;1
<JSC>TASK3/TEST/4/CORN.ESP;1
<JSC>TASK3/TEST/4/CORN.SGT;1
<JSC>TASK3/TEST/4/CAT.;1
<JSC>TASK3/TEST/4.SCAT-FILES;1
<JSC>TASK3/TEST/4.SEGS;2
<JSC>TASK3/TEST/4/DENSEWOODLAND.ESP;1
```

```
<JSC>TASK3/TRAIN/5.SCGS;2
<JSC>TASK3/TRAIN/5/WOODS.BOUT;1
<JSC>TASK3/TRAIN/5/WHEAT.BOUT;1
<JSC>TASK3/TRAIN/5/STAT.16-GRPS;1
<JSC>TASK3/TRAIN/5/STAT.-DENSEWOODLAND/CLIP;1
  20844
<JSC>TASK3/TRAIN/5/STAT.-WHEAT/CLIP;1
<JSC>TASK3/TRAIN/5/STAT.-SOYBEANS/CLIP;1
4
<JSC>TASK3/TRAIN/5/STAT.-PASTURE/CLIP;1
<JSC>TASK3/TRAIN/5/STAT.-HAY/CLIP;1
<JSC>TASK3/TRAIN/5/STAT.-CORN/CLIP;2
<JSC>TASK3/TRAIN/5/SOYBEANS.BOUT;1
<JSC>TASK3/TRAIN/5/PASTURE.BOUT;1
<JSC>TASK3/TRAIN/5/PACK.-NB/M079;1
<JSC>TASK3/TRAIN/5/PACK.-DENSEWOODS/CLIP;1
844
<JSC>TASK3/TRAIN/5/PACK.-OTHERHAY/CLIP;1
4
<JSC>TASK3/TRAIN/5/PACK.-SOYBEANS/CLIP;1
4
<JSC>TASK3/TRAIN/5/PACK.-PASTURE/CLIP;1
<JSC>TASK3/TRAIN/5/PACK.-WHEAT/CLIP;1
<JSC>TASK3/TRAIN/5/PACK.-CORN/CLIP;1
<JSC>TASK3/TRAIN/5/ISTAT.16-GRPS;1
<JSC>TASK3/TRAIN/5/HAY.BOUT;1
<JSC>TASK3/TRAIN/5/CORN.BOUT;2
<JSC>TASK3/TRAIN/5/PACK.-OTHERHAY/M079;1
*4
<JSC>TASK3/TRAIN/5/PACK.-DENSEWOODLAND/M0379;1
*  20844
<JSC>TASK3/TRAIN/5/PACK.-SOYBEANS/M079'1
*4
<JSC>TASK3/TRAIN/5/PACK.-PASTURE/M079;1
<JSC>TASK3/TRAIN/5/PACK.-WINTERWHEAT/M079;1
*0844
<JSC>TASK3/TRAIN/5/PACK.-CORN/M079;1
<JSC>TASK3/TRAIN/5/PACK.NB/M079;1
<JSC>TASK3/TRAIN/5/COMMAND.BOUT;1
<JSC>TASK3/TRAIN/5/WINTERWHEAT.ESP;1
<JSC>TASK3/TRAIN/5/WINTERWHEAT.SGT;1
<JSC>TASK3/TRAIN/5/TBL.;1
<JSC>TASK3/TRAIN/5/SOYBEANS.ESP;1
<JSC>TASK3/TRAIN/5/SOYBEANS.SGT;1
<JSC>TASK3/TRAIN/5/SENSEWOODLAND.ESP;1
<JSC>TASK3/TRAIN/5/PERMANENTPASTURE.ESP;1
*53
<JSC>TASK3/TRAIN/5/PERMANENTPASTURE.SGT;1
*53
<JSC>TASK3/TRAIN/5/OTHERHAY.ESP;1
<JSC>TASK3/TRAIN/5/OTHERHAY.SGT;1
<JSC>TASK3/TRAIN/5/ISTAT.16-GRPS/CLIP;1
<JSC>TASK3/TRAIN/5/DENSEWOODLAND.SGT;1
<JSC>TASK3/TRAIN/5/CORN.ESP;1
<JSC>TASK3/TRAIN/5/CORN.SGT;1
<JSC>TASK3/TRAIN/5/CAT.;1
```

```
 <JSC>TASK3/TEST/5/WINTERWHEAT.SGT;1
 <JSC>TASK3/TEST/5/TBL.;1
 <JSC>TASK3/TEST/5/SOYBEANS.ESP;1
 <JSC>TASK3/TEST/5/SOYBEANS.SGT;1
 <JSC>TASK3/TEST/5/PERMANENTPASTURE.ESP;1
#3
 <JSC>TASK3/TEST/5/PERMANENTPASTURE.SGT;1
#3
 <JSC>TASK3/TEST/5/OTHERHAY.ESP;1
 <JSC>TASK3/TEST/5/OTHERHAY.SGT;1
 <JSC>TASK3/TEST/5/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TEST/5/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TEST/5/CORN.ESP;1
 <JSC>TASK3/TEST/5/CORN.SGT;1
 <JSC>TASK3/TEST/5/CAT.;1
 <JSC>TASK3/TEST/5.SEGS;2
 <JSC>TASK3/TEST/5/PACK.NB/M079;1
 <JSC>TASK3/TEST/5/COMMAND.BOUT;1
```

```
 <JSC>TASK3/TRAIN/G/WOODS.BOUT;1
 <JSC>TASK3/TRAIN/G/WHEAT.BOUT;1
 <JSC>TASK3/TRAIN/G/STAT.17-GRPS;1
 <JSC>TASK3/TRAIN/G/STAT.-WOODS/CLIP;1
 <JSC>TASK3/TRAIN/G/STAT.-WHEAT/CLIP;1
 <JSC>TASK3/TRAIN/G/STAT.-SOYBEANS/CLIP;1
*4
 <JSC>TASK3/TRAIN/G/STAT.-PASTURE/CLIP;1
 <JSC>TASK3/TRAIN/G/STAT.-OTHERHAY/CLIP;1
*4
 <JSC>TASK3/TRAIN/G/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/G/SOYBEANS.BOUT;1
 <JSC>TASK3/TRAIN/G/PASTURE.BOUT;1
 <JSC>TASK3/TRAIN/G/PACK.-OTHERHAY/CLIP;1
*4
 <JSC>TASK3/TRAIN/G/PACK.-WOODS/CLIP;1
 <JSC>TASK3/TRAIN/G/PACK.-SOYBEANS/CLIP;1

 <JSC>TASK3/TRAIN/G/PACK.-PERMANENTPASTURE/CLIP;1
ind  20844
 <JSC>TASK3/TRAIN/G/PACK.-WINTERWHEAT/CLIP;1
*0844
 <JSC>TASK3/TRAIN/G/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/G/PACK.-NB/M079;2
 <JSC>TASK3/TRAIN/G/HAY.BOUT;1
 <JSC>TASK3/TRAIN/G/CORN.BOUT;1
 <JSC>TASK3/TRAIN/G/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TRAIN/G/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TRAIN/G.SEGS;2
 <JSC>TASK3/TRAIN/G/PACK.NB/M079;2
 <JSC>TASK3/TRAIN/G/ISTAT.17-GRPS;2
 <JSC>TASK3/TRAIN/G/COMMAND.BOUT;1
 <JSC>TASK3/TRAIN/G/PACK.-OTHERHAY/M079;1
*4
 <JSC>TASK3/TRAIN/G/PACK.-DENSEWOODLAND/M079;1
* 20844
 <JSC>TASK3/TRAIN/G/PACK.-SOYBEANS/M079;1
**4
 <JSC>TASK3/TRAIN/G/PACK.-PERMANENTPASTURE/M079;1
**nd  20844
 <JSC>TASK3/TRAIN/G/PACK.-WINTERWHEAT/M079;1
**0844
 <JSC>TASK3/TRAIN/G/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/G/WINTERWHEAT.SGT;1
 <JSC>TASK3/TRAIN/G/TBL.;1
 <JSC>TASK3/TRAIN/G/TBL.;2
 <JSC>TASK3/TRAIN/G/SOYBEANS.ESP;1
 <JSC>TASK3/TRAIN/G/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/G/PERMANENTPASTURE.ESP;1
*53
 <JSC>TASK3/TRAIN/G/PERMANENTPASTURE.SGT;1
*53
 <JSC>TASK3/TRAIN/G/OTHERHAY.ESP;1
 <JSC>TASK3/TRAIN/G/OTHERHAY.SGT;1
 <JSC>TASK3/TRAIN/G/ISTAT.17-GRPS/CLIP;1
 <JSC>TASK3/TRAIN/G/CORN.ESP;1
 <JSC>TASK3/TRAIN/G/CORN.SGT;1
 <JSC>TASK3/TRAIN/G/CAT.;1
```

```
 <JSC>TASK3/TEST/6/PACK.-NB/M079;1
 <JSC>TASK3/TEST/6/PACK.NB/M079;2
 <JSC>TASK3/TEST/6/COMMAND.BOUT;1
 <JSC>TASK3/TEST/6/WINTERWHEAT.ESP;1
 <JSC>TASK3/TEST/6/WINTERWHEAT.SGT;1
 <JSC>TASK3/TEST/6/TBL.;1
 <JSC>TASK3/TEST/6/SOYBEANS.ESP;1
 <JSC>TASK3/TEST/6/SOYBEANS.SGT;1
 <JSC>TASK3/TEST/6/PERMANENTPASTURE.ESP;1
*3
 <JSC>TASK3/TEST/6/PERMANENTPASTURE.SGT;1
*3
 <JSC>TASK3/TEST/6/OTHERHAY.SGT;1
 <JSC>TASK3/TEST/6/OTHERHAY.ESP;1
 <JSC>TASK3/TEST/6/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TEST/6/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TEST/6/CORN.ESP;2
 <JSC>TASK3/TEST/6/CORN.SGT;2
 <JSC>TASK3/TEST/6/CAT.;1
 <JSC>TASK3/TEST/6.SEGS;2
```

```
 <JSC>TASK3/TRAIN/7/PACK.NB/M079;1
 <JSC>TASK3/TRAIN/7/WOODS.BOUT;1
 <JSC>TASK3/TRAIN/7/WHEAT.BOUT;1
 <JSC>TASK3/TRAIN/7/STAT.-1GGRPS/CLIP;1
 <JSC>TASK3/TRAIN/7/STAT.-WOODS/CLIP;1
 <JSC>TASK3/TRAIN/7/STAT.-WINTERWHEAT/CLIP;1
*0844
 <JSC>TASK3/TRAIN/7/STAT.-SOYBEANS/CLIP;1
*4
 <JSC>TASK3/TRAIN/7/STAT.-PASTURE/CLIP;1
 <JSC>TASK3/TRAIN/7/STAT.-HAY/CLIP;1
 <JSC>TASK3/TRAIN/7/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/7/SOYBEANS.BOUT;1
 <JSC>TASK3/TRAIN/7/PASTURE.BOUT;1
 <JSC>TASK3/TRAIN/7/PACK.-OTHERHAY/M079;1
*4
 <JSC>TASK3/TRAIN/7/PACK.-DENSEWOODLAND/M079;1
* 20844
 <JSC>TASK3/TRAIN/7/PACK.-SOYBEANS/M079;1
*4
 <JSC>TASK3/TRAIN/7/PACK.-PERMANENTPASTURE/M079;1
*nd  20844
 <JSC>TASK3/TRAIN/7/PACK.-WINTERWHEAT/M079;1
*0844
 <JSC>TASK3/TRAIN/7/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/7/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/7/PACK.-OTHERHAY/CLIP;1
*4
 <JSC>TASK3/TRAIN/7/PACK.-WOODS/CLIP;1
 <JSC>TASK3/TRAIN/7/PACK.-SOYBEANS/CLIP;1
*4
 <JSC>TASK3/TRAIN/7/PACK.-PASTURE/CLIP;1
 <JSC>TASK3/TRAIN/7/PACK.-WHEAT/CLIP;1
 <JSC>TASK3/TRAIN/7/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/7/ISTAT.-1GGRPS/CLIP;1
 <JSC>TASK3/TRAIN/7/HAY.BOUT;1
 <JSC>TASK3/TRAIN/7/CORN.BOUT;1
 <JSC>TASK3/TRAIN/7/COMMAND.BOUT;1
 <JSC>TASK3/TRAIN/7/WINTERWHEAT.ESP;1
 <JSC>TASK3/TRAIN/7/WINTERWHEAT.SGT;1
 <JSC>TASK3/TRAIN/7/TBL.;1
 <JSC>TASK3/TRAIN/7/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/7/SOYBEANS.ESP;1
 <JSC>TASK3/TRAIN/7/PERMANENTPASTURE.ESP;1
*53
 <JSC>TASK3/TRAIN/7/PERMANENTPASTURE.SGT;1
*53
 <JSC>TASK3/TRAIN/7/OTHERHAY.ESP;1
 <JSC>TASK3/TRAIN/7/OTHERHAY.SGT;1
 <JSC>TASK3/TRAIN/7/ISTAT.-1GGRPS/WITHNAMES;1
*20853
 <JSC>TASK3/TRAIN/7/ISTAT.-1GGRPS/PUR;1
 <JSC>TASK3/TRAIN/7/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TRAIN/7/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TRAIN/7/CORN.ESP;1
 <JSC>TASK3/TRAIN/7/CORN.SGT;1
 <JSC>TASK3/TRAIN/7/CAT.;1
 <JSC>TASK3/TRAIN/7.SEGS;2
```

```
 <JSC>TASK3/TEST/7/WINTERWHEAT.ESP;1
 <JSC>TASK3/TEST/7/WINTERWHEAT.SGT;1
 <JSC>TASK3/TEST/7/TBL.;3
 <JSC>TASK3/TEST/7/SOYBEANS.ESP;1
 <JSC>TASK3/TEST/7/SOYBEANS.SGT;1
 <JSC>TASK3/TEST/7/PERMANENTPASTURE.SGT;1
*3
 <JSC>TASK3/TEST/7/PERMANENTPASTURE.ESP;1
*3
 <JSC>TASK3/TEST/7/OTHERHAY.ESP;1
 <JSC>TASK3/TEST/7/OTHERHAY.SGT;1
 <JSC>TASK3/TEST/7/DENSEWOODLAND.ESP;1
 <JSC>TASK3/TEST/7/DENSEWOODLAND.SGT;1
 <JSC>TASK3/TEST/7/CORN.ESP;1
 <JSC>TASK3/TEST/7/CORN.SGT;1
 <JSC>TASK3/TEST/7/CAT.;1
 <JSC>TASK3/TEST/7.SEGS;2
<JSC>TASK3/TEST/7/PACK.NB/M079;1
<JSC>TASK3/TEST/7/COMMAND.BOUT;1
<JSC>TASK3/TEST/7/PACK.-NB/M079;1
```

```
 <JSC>TASK3/TRAIN/8.SEGS;2
  <JSC>TASK3/TRAIN/8/WOODS.BOUT;1
  <JSC>TASK3/TRAIN/8/WHEAT.BOUT;1
  <JSC>TASK3/TRAIN/8/STAT.-15GRPS/CLIP;1
  <JSC>TASK3/TRAIN/8/STAT.-OTHERHAY/CLIP;1
 *4
  <JSC>TASK3/TRAIN/8/STAT.-DENSEWOODLAND/CLIP;1
 * 20844
  <JSC>TASK3/TRAIN/8/STAT.-SOYBEANS/CLIP;1
 *4
  <JSC>TASK3/TRAIN/8/STAT.-PERMANENTPASTURE/CLIP;1
 *nd  20844
  <JSC>TASK3/TRAIN/8/STAT.-WINTERWHEAT/CLIP;1
 *0844
  <JSC>TASK3/TRAIN/8/STAT.-CORN/CLIP;1
  <JSC>TASK3/TRAIN/8/SOYBEANS.BOUT;1
  <JSC>TASK3/TRAIN/8/PASTURE.BOUT;1
  <JSC>TASK3/TRAIN/8/PACK.-OTHERHAY/M079;1
 *4
  <JSC>TASK3/TRAIN/8/PACK.-DENSEWOODLAND/M079;1
 * 20844
  <JSC>TASK3/TRAIN/8/PACK.-SOYBEANS/M079;1
 *4
  <JSC>TASK3/TRAIN/8/PACK.-PERMANENTPASTURE/M079;1
 *nd  20844
  <JSC>TASK3/TRAIN/8/PACK.-WINTERWHEAT/M079;1
 *0844
  <JSC>TASK3/TRAIN/8/PACK.-CORN/M079;1
  <JSC>TASK3/TRAIN/8/PACK.-OTHERHAY/CLIP;1
 *4
  <JSC>TASK3/TRAIN/8/PACK.-DENSEWOODLAND/CLIP;1
 * 20844
  <JSC>TASK3/TRAIN/8/PACK.-SOYBEANS/CLIP;1
 *4
  <JSC>TASK3/TRAIN/8/PACK.-PERMANENTPASTURE/CLIP;1
 *nd  20844
  <JSC>TASK3/TRAIN/8/PACK.-CORN/CLIP;1
  <JSC>TASK3/TRAIN/8/PACK.-WINTERWHEAT/CLIP;1
 *0844
  <JSC>TASK3/TRAIN/8/PACK.-NB/M079;1
  <JSC>TASK3/TRAIN/8/ISTAT.-15GRPS/CLIP;1
  <JSC>TASK3/TRAIN/8/HAY.BOUT;1
  <JSC>TASK3/TRAIN/8/CORN.BOUT;1
  <JSC>TASK3/TRAIN/8/COMMAND.BOUT;1
  <JSC>TASK3/TRAIN/8/WHEAT.BOUT;1
  <JSC>TASK3/TRAIN/8/STAT.-WINTERWHEAT/CLIP;1
 *0844
  <JSC>TASK3/TRAIN/8/PACK.NB/M079;1
  <JSC>TASK3/TRAIN/8/WINTERWHEAT.ESP;1
  <JSC>TASK3/TRAIN/8/WINTERWHEAT.SGT;1
  <JSC>TASK3/TRAIN/8/TBL.;1
  <JSC>TASK3/TRAIN/8/SOYBEANS.ESP;1
  <JSC>TASK3/TRAIN/8/SOYBEANS.SGT;1
  <JSC>TASK3/TRAIN/8/PERMANENTPASTURE.ESP;1
 *53
  <JSC>TASK3/TRAIN/8/PERMANENTPASTURE.SGT;1
 *53
```

```
<JSC>TASK3/TRAIN/8/OTHERHAY.SGT;1
<JSC>TASK3/TRAIN/8/OTHERHAY.ESP;1
<JSC>TASK3/TRAIN/8/ISTAT.-15GRPS/WITHNAMES;1
*20853
<JSC>TASK3/TRAIN/8/ISTAT.-15GRPS/PUR;1
<JSC>TASK3/TRAIN/8/DENSEWOODLAND.ESP;1
<JSC>TASK3/TRAIN/8/DENSEWOODLAND.SGT;1
<JSC>TASK3/TRAIN/8/CORN.ESP;1
<JSC>TASK3/TRAIN/8/CORN.SGT;1
<JSC>TASK3/TRAIN/8/CAT.;1
```

```
<JSC>TASK3/TEST/8/PACK.-NB/M079;1
<JSC>TASK3/TEST/8/WOODS.EST;1
<JSC>TASK3/TEST/8/WOODS.SGT;1
<JSC>TASK3/TEST/8/WHEAT.EST;1
<JSC>TASK3/TEST/8/WHEAT.SGT;1
<JSC>TASK3/TEST/8/TABLE.;1
<JSC>TASK3/TEST/8/PACK.NB/M079;1
<JSC>TASK3/TEST/8/COMMAND.BOUT;1
<JSC>TASK3/TEST/8/SOYBEANS.EST;1
<JSC>TASK3/TEST/8/SOYBEANS.SGT;1
<JSC>TASK3/TEST/8/PASTURE.EST;1
<JSC>TASK3/TEST/8/PASTURE.SGT;1
<JSC>TASK3/TEST/8/HAY.EST;1
<JSC>TASK3/TEST/8/HAY.SGT;1
<JSC>TASK3/TEST/8/CORN.EST;1
<JSC>TASK3/TEST/8/CORN.SGT;1
<JSC>TASK3/TEST/8/CAT.;1
<JSC>TASK3/TEST/8.SEGS;2
```

```
 <JSC>TASK3/TRAIN/9.CEGS;2
 <JSC>TASK3/TRAIN/9/STAT.16-GRPS;1
 <JSC>TASK3/TRAIN/9/STAT.-OTHERHAY/CLIP;1
*4
 <JSC>TASK3/TRAIN/9/STAT.-DENSEWOODLAND/CLIP;1
* 20844
 <JSC>TASK3/TRAIN/9/STAT.VSOYBEANS/CLIP;1
*4
 <JSC>TASK3/TRAIN/9/STAT.-PERMANENTPASTURE/CLIP;1
*nd  20844
 <JSC>TASK3/TRAIN/9/STAT.-WINTERWHEAT/CLIP;1
*0844
 <JSC>TASK3/TRAIN/9/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/9/PACK.NB/M079;1
 <JSC>TASK3/TRAIN/9/ISTAT.16-GRPS;1
 <JSC>TASK3/TRAIN/9/COMMAND.BOUT;1
 <JSC>TASK3/TRAIN/9/WOODS.EST;1
 <JSC>TASK3/TRAIN/9/WOODS.SGT;1
 <JSC>TASK3/TRAIN/9/WHEAT.EST;1
 <JSC>TASK3/TRAIN/9/WHEAT.SGT;1
 <JSC>TASK3/TRAIN/9/TABLE.;1
 <JSC>TASK3/TRAIN/9/SOYBEANS.EST;1
 <JSC>TASK3/TRAIN/9/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/9/PASTURE.EST;1
 <JSC>TASK3/TRAIN/9/PASTURE.SGT;1
 <JSC>TASK3/TRAIN/9/ISTAT.16-GRPS/CLIP;1
 <JSC>TASK3/TRAIN/9/HAY.EST;1
 <JSC>TASK3/TRAIN/9/HAY.SGT;1
 <JSC>TASK3/TRAIN/9/CORN.EST;1
 <JSC>TASK3/TRAIN/9/CORN.SGT;1
 <JSC>TASK3/TRAIN/9/CAT.;1
 <JSC>TASK3/TRAIN/9/WOODS.BOUT;1
 <JSC>TASK3/TRAIN/9/WHEAT.BOUT;1
 <JSC>TASK3/TRAIN/9/SOYBEANS.BOUT;1
 <JSC>TASK3/TRAIN/9/PASTURE.BOUT;1
 <JSC>TASK3/TRAIN/9/PACK.-OTHERHAY/M079;1
*4
 <JSC>TASK3/TRAIN/9/PACK.-DENSEWOODLAND/M079;1
* 20844
 <JSC>TASK3/TRAIN/9/PACK.-SOYBEANS/M079;1
*4
 <JSC>TASK3/TRAIN/9/PACK.-PERMANENTPASTURE/M079;1
*nd  20844
 <JSC>TASK3/TRAIN/9/PACK.-WINTERWHEAT/M079;1
*0844
 <JSC>TASK3/TRAIN/9/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/9/PACK.-OTHERHAY/CLIP;1
*4
 <JSC>TASK3/TRAIN/9/PACK.-DENSEWOODLAND/CLIP;1
* 20844
 <JSC>TASK3/TRAIN/9/PACK.-SOYBEANS/CLIP;1
*4
 <JSC>TASK3/TRAIN/9/PACK.-PERMANENTPASTURE/CLIP;1
*nd  20844
 <JSC>TASK3/TRAIN/9/PACK.-WINTERWHEAT/CLIP;1
*0844
 <JSC>TASK3/TRAIN/9/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/9/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/9/HAY.BOUT;1
 <JSC>TASK3/TRAIN/9/CORN.BOUT;1
```

```
<JSC>TASK3/TEST/9.SEGS;2
<JSC>TASK3/TEST/9/PACK.-NB/MD79;1
<JSC>TASK3/TEST/9/PACK.NB/MD79;1
<JSC>TASK3/TEST/9/COMMAND.BOUT;1
<JSC>TASK3/TEST/9/WOODS.EST;1
<JSC>TASK3/TEST/9/WOODS.SGT;1
<JSC>TASK3/TEST/9/WHEAT.EST;1
<JSC>TASK3/TEST/9/WHEAT.SGT;1
<JSC>TASK3/TEST/9/TABLE.;1
<JSC>TASK3/TEST/9/SOYBEANS.EST;1
<JSC>TASK3/TEST/9/SOYBEANS.SGT;1
<JSC>TASK3/TEST/9/PASTURE.EST;1
<JSC>TASK3/TEST/9/PASTURE.SGT;1
<JSC>TASK3/TEST/9/HAY.SGT;1
<JSC>TASK3/TEST/9/HAY.EST;1
<JSC>TASK3/TEST/9/CORN.EST;1
<JSC>TASK3/TEST/9/CORN.SGT;1
<JSC>TASK3/TEST/9/CAT.;1
```

```
 <JSC>TASK3/TRAIN/10/WOODS.BOUT;1
 <JSC>TASK3/TRAIN/10/WHEAT.BOUT;1
 <JSC>TASK3/TRAIN/10/SOYBEANS.BOUT;1
 <JSC>TASK3/TRAIN/10/PASTURE.BOUT;1
 <JSC>TASK3/TRAIN/10/PACK.-OTHERHAY/CLIP;1
!44
 <JSC>TASK3/TRAIN/10/PACK.-DENSEWOODLAND/CLIP;1
!  20844
 <JSC>TASK3/TRAIN/10/PACK.-SOYBEANS/CLIP;1
!44
 <JSC>TASK3/TRAIN/10/PACK.-PERMABNENTPASTURE/CLIP;1
! and  20844
 <JSC>TASK3/TRAIN/10/PACK.-WINTERWHEAT/CLIP;1
*20844
 <JSC>TASK3/TRAIN/10/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/10/HAY.BOUT;1
 <JSC>TASK3/TRAIN/10/CORN.BOUT;1
 <JSC>TASK3/TRAIN/10/PACK.-OTHERHAY/M079;1
*44
 <JSC>TASK3/TRAIN/10/PACK.-DENSEWOODLAND/M079;1
*  20844
 <JSC>TASK3/TRAIN/10/PACK.-SOYBEANS/M079;1
*44
 <JSC>TASK3/TRAIN/10/PACK.-PERMANENTPASTURE/M079;1
*and  20844
 <JSC>TASK3/TRAIN/10/PACK.-WINTERWHEAT/M079;1
*20844
 <JSC>TASK3/TRAIN/10/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/10/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/10/WOODS.EST;1
 <JSC>TASK3/TRAIN/10/WOODS.SGT;1
 <JSC>TASK3/TRAIN/10/WHEAT.EST;1
 <JSC>TASK3/TRAIN/10/WHEAT.SGT;1
 <JSC>TASK3/TRAIN/10/TABLE.;1
 <JSC>TASK3/TRAIN/10/STAT.-15GRPS/EP;1
 <JSC>TASK3/TRAIN/10/STAT.-OTHERHAY/CLIP;1
*44
 <JSC>TASK3/TRAIN/10/STAT.-DENSEWOODLAND/CLIP;1
*  20844
 <JSC>TASK3/TRAIN/10/STAT.-SOYBEANS/CLIP;1
*44
 <JSC>TASK3/TRAIN/10/STAT.-PERMANENTPASTURE/CLIP;1
*and  20844
 <JSC>TASK3/TRAIN/10/STAT.-WINTERWHEAT/CLIP;1
!*20844
 <JSC>TASK3/TRAIN/10/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/10/SOYBEANS.EST;1
 <JSC>TASK3/TRAIN/10/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/10/PASTURE.EST;1
 <JSC>TASK3/TRAIN/10/PASTURE.SGT;1
 <JSC>TASK3/TRAIN/10/PASTURE.BOUT;1
 <JSC>TASK3/TRAIN/10/PACK.NB/M079;1
 <JSC>TASK3/TRAIN/10/ISTAT.-15GRPS/EP;1
 <JSC>TASK3/TRAIN/10/HAY.EST;1
 <JSC>TASK3/TRAIN/10/HAY.SGT;1
 <JSC>TASK3/TRAIN/10/CORN.EST;1
 <JSC>TASK3/TRAIN/10/COMMAND.BOUT;1
 <JSC>TASK3/TRAIN/10/CAT.;1
 <JSC>TASK3/TRAIN/10.SEGS;2
 <JSC>TASK3/TRAIN/10/ISTAT.-15GRPS/PUR;1
 <JSC>TASKK3/TRAIN/10/CORN.SGT;1
```

```
<JSC>TASK3/TEST/10/WOODS.EST;1
<JSC>TASK3/TEST/10/WOODS.SGT;1
<JSC>TASK3/TEST/10/WHEAT.EST;1
<JSC>TASK3/TEST/10/WHEAT.SGT;1
<JSC>TASK3/TEST/10/TABLE.;1
<JSC>TASK3/TEST/10/SOYBEANS.EST;1
<JSC>TASK3/TEST/10/SOYBEANS.SGT;1
<JSC>TASK3/TEST/10/PASTURE.EST;1
<JSC>TASK3/TEST/10/PASTURE.SGT;1
<JSC>TASK3/TEST/10/PACK.NB/M079;1
<JSC>TASK3/TEST/10/HAY.SGT;1
<JSC>TASK3/TEST/10/HAY.SGT;2
<JSC>TASK3/TEST/10/CORN.EST;1
<JSC>TASK3/TEST/10/CORN.SGT;1
<JSC>TASK3/TEST/10/COMMAND.BOUT;1
<JSC>TASK3/TEST/10/CAT.;1
<JSC>TASK3/TEST/10.SEGS;2
<JSC>TASK3/TEST/10/HAY.EST;1
<JSC>TASK3/TEST/10/PACK.-NB/M079;1
```

```
 <JSC>TASK3/TRAIN/11/TABLE.;2
 <JSC>TASK3/TRAIN/11/STAT.-1GGRPS/CLIP;1
 <JSC>TASK3/TRAIN/11/STAT.-DENSEWWODLAN/CLIP;1
* 20844
 <JSC>TASK3/TRAIN/11/STAT.-WINTERWHEAT/CLIP;1
*20844
 <JSC>TASK3/TRAIN/11/STAT.-SOYBEANS/CLIP;1
*44
 <JSC>TASK3/TRAIN/11/STAT.-PERMANENTPASTURE/CLIP;1
*and  20844
 <JSC>TASK3/TRAIN/11/STAT.-OTHERHAY/CLIP;1
*44
 <JSC>TASK3/TRAIN/11/STAT.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/11/SOYBEANS.BOUT;1
 <JSC>TASK3/TRAIN/11/PASTURE.BOUT;1
 <JSC>TASK3/TRAIN/11/PACK.-OTHERHAY/CLIP;1
*44
 <JSC>TASK3/TRAIN/11/PACK.-DENSEWOODS/CLIP;1
*0844
 <JSC>TASK3/TRAIN/11/PACK.-SOYBEANS/CLIP;1
*44
 <JSC>TASK3/TRAIN/11/PACK.-PERMANENTPASTURE/CLIP;1
*and  20844
 <JSC>TASK3/TRAIN/11/PACK.-WINTERWHEAT/CLIP;1
·*20844
 <JSC>TASK3/TRAIN/11/PACK.-CORN/CLIP;1
 <JSC>TASK3/TRAIN/11/ISTAT.-1GGRPS/CLIP;1
**4
 <JSC>TASK3/TRAIN/11/HAY.BOUT;1
<JSC>TASK3/TRAIN/11/CORN.BOUT;1
 <JSC>TASK3/TRAIN/11/WOODS.BOUT;1
 <JSC>TASK3/TRAIN/11/WHEAT.BOUT;1
 <JSC>TASK3/TRAIN/11/PACK.-OTHERHAY/M079;1
*44
 <JSC>TASK3/TRAIN/11/PACK.-SOYBEANS/M079;1
*44
 <JSC>TASK3/TRAIN/11/PACK.-DENSEWOODLAND/M079;1
*  20844
 <JSC>TASK3/TRAIN/11/PACK.-PERMANENTPASTURE/M079;1
*and  20844
 <JSC>TASK3/TRAIN/11/PACK.-WINTERWHEAT/M079;1
*20844
 <JSC>TASK3/TRAIN/11/PACK.-CORN/M079;1
 <JSC>TASK3/TRAIN/11/PACK.-NB/M079;1
 <JSC>TASK3/TRAIN/11/WOODS.EST;1
 <JSC>TASK3/TRAIN/11/WOODS.SGT;1
 <JSC>TASK3/TRAIN/11/WHEAT.EST;1
 <JSC>TASK3/TRAIN/11/WHEAT.SGT;1
 <JSC>TASK3/TRAIN/11/TABLE.;1
 <JSC>TASK3/TRAIN/11/SOYBEANS.EST;1
 <JSC>TASK3/TRAIN/11/SOYBEANS.SGT;1
 <JSC>TASK3/TRAIN/11/PASTURE.EST;1
 <JSC>TASK3/TRAIN/11/PASTURE.SGT;1
 <JSC>TASK3/TRAIN/11/PACK.NB/M079;1
 <JSC>TASK3/TRAIN/11/ISTAT.-1GGRPS/PUR;1
 <JSC>TASK3/TRAIN/11/HAY.EST;1
 <JSC>TASK3/TRAIN/11/HAY.SGT;1
 <JSC>TASK3/TRAIN/11/CORN.EST;1
 <JSC>TASK3/TRAIN/11/CORN.SGT;1
 <JSC>TASK3/TRAIN/11/COMMAND.BOUT;1
 <JSC>TASK3/TRAIN/11/CAT.;1
 <JSC>TASK3/TRAIN/11.SEGS;2
```

```
<JSC>TASK3/TEST/11/WOODS.EST;1
<JSC>TASK3/TEST/11/WOODS.SGT;1
<JSC>TASK3/TEST/11/WHEAT.EST;1
<JSC>TASK3/TEST/11/WHEAT.SGT;1
<JSC>TASK3/TEST/11/TABLE.;1
<JSC>TASK3/TEST/11/SOYBEANS.EST;1
<JSC>TASK3/TEST/11/SOYBEANS.SGT;1
<JSC>TASK3/TEST/11/PASTURE.EST;1
<JSC>TASK3/TEST/11/PASTURE.SGT;1
<JSC>TASK3/TEST/11/PACK.NB/MO79;1
<JSC>TASK3/TEST/11/HAY.EST;1
<JSC>TASK3/TEST/11/HAY.SGT;1
<JSC>TASK3/TEST/11/CORN.EST;1
<JSC>TASK3/TEST/11/CORN.SGT;1
<JSC>TASK3/TEST/11/COMMAND.BOUT;1
<JSC>TASK3/TEST/11/CAT.;1
<JSC>TASK3/TEST/11.SEGS;2
```


☆ U.S. GOVERNMENT PRINTING OFFICE: 1981— 771-014/1068