Forest Mapping Accuracies Are Improved Using a Supervised Nonparametric Classifier with SPOT Data

Andrew K. Skidmore and *Brian J. Turner*
Department of Forestry, Australian National University, G.P.O. Box 4, Canberra, A.C.T. 2601, Australia

ABSTRACT: A new supervised nonparametric classifier produces an image showing the empirical probability of correct classification for a pixel as well as a thematic image. This allows an analyst to visually locate those parts of the image where classification success can be improved. The algorithm was tested using SPOT XS data over a forest plantation in southeast Australia. The classifier produced thematic maps of higher accuracy than those from conventional supervised classifiers.

INTRODUCTION

TURNER ET AL. (1987) and Skidmore *et al.* (1987) outlined reasons for the apparent reluctance of foresters to embrace remotely sensed data for operational use, with the major reasons cited being imagery of poor spatial resolution, and resultant maps having poor accuracy. Though spatial resolution is improving with the new generation of satellites such as SPOT, forest cover mapping accuracies using conventional classifiers have been generally low (typically less than 80 percent mapping accuracy at 90 percent confidence levels), especially where forest types are discriminated at Anderson *et al.* (1976) level III (e.g., Strahler *et al.*, 1978; Merola *et al.*, 1983; Hame, 1984). Only a few examples of high mapping accuracies have been reported. Nelson (1981) cited accuracies of 79 to 88 percent, but only discriminated hardwood from conifer and grassland. Walsh (1980) claimed mapping accuracies of 88 percent (ranging from 85 to 95 percent per stratum). However, at each random point within a stratum, Walsh included the 25 (5 by 5) surrounding pixels into his calculation. This may have artificially increased mapping accuracy, as Kettig and Landgrebe (1976) showed that adjacent pixels are autocorrelated.

Classification strategies may be categorized into supervised or unsupervised methods. Both methods assume that the image data form separate groups in *N*-dimensional feature space (where *N*-dimensional feature space refers to the space created when *N* channels (or features) of data are each placed on orthogonal axes) and these groups can be associated with observed ground cover types. The groups of data can be described by parametric or nonparametric techniques. Parametric classification strategies assume that each group can be enclosed by a boundary, such as defined by the hyper-ellipsoid shaped decision volume of the maximum likelihood classifier (Swain and Davis, 1978; Richards, 1986). Nonparametric classifiers make no assumptions about the shape of the data distributions, except that the groups of data can be separated by some discriminant function (Nilsson, 1965) such as the linear regression functions derived from logit modeling (Strahler *et al.*, 1980).

To improve the accuracy of forest maps derived from remotely sensed data, supervised and unsupervised techniques have been combined. One methodology involves delineating training areas containing representative cover classes; the training area data are then clustered using an unsupervised strategy (Fleming, 1975; Beaubien, 1979; LaPerriere *et al.*, 1980; Thompson *et al.*, 1980; Walsh, 1980). The unsupervised cluster strategy may produce a thematic map directly, or the algorithm can generate statistics which are input to a supervised parametric classifier. Another method to improve mapping accuracies is to include data ancillary to the remotely sensed data (Strahler *et al.*, 1978; Tom and Miller, 1980; Hutchinson, 1982; Cibula and Nyquist, 1987). Hutchinson (1982) discussed the three methods of combining spectral data with other ancillary data *viz.* stratifying an image prior to classification; incorporating the ancillary data during the classification operation; and post-classification, where a classified image is modified by the ancillary data, e.g., shadow reduction. These techniques use conventional parametric and nonparametric approaches for the classification of the data, and aim to improve map accuracy through the inclusion of additional input data. The proposed nonparametric classifier does not incorporate any additional information in order to improve mapping accuracies.

Skidmore *et al.* (1988) showed that (forest plantation) cover classes could not be successfully mapped from Landsat MSS data using parallelepiped or euclidean distance classifiers. Using a general nonparametric test, they quantified the co-occurrence (or spectral overlap) of training set data distributions at each vector point in *N*-dimensional feature space for a number of forest plantation spectral classes. The main conclusion was that training area data from some plantation age classes were not co-occurring in *N*-dimensional feature space, even though the classification strategies could not separate these classes. This apparent contradiction was explained by the observation that the distributions of training area data in two-dimensional feature space (i.e., a "scatter plot") are frequently observed to not have a shape that can be approximated by a parametric classifier (e.g., a rectangular shape for a parallelepiped classifier, or an ellipsoid shape for a maximum-likelihood classifier). To improve the accuracy of the forest cover type map, a supervised nonparametric classifier based on these results has been developed.

CONSTRUCTION OF THE SUPERVISED NONPARAMETRIC CLASSIFIER

INTRODUCTION

The classifier can be generally described as follows. Training area data are collected for representative cover class areas. Each pixel for the first cover class is assigned to the cell (or vector position) in the *N*-dimensional feature space which equates with the brightness value of the pixel. The number of pixels (for the first class) that occurs in each cell is summed. Similarly, the pixels of the second cover class are summed into the cells of the *N*-dimensional feature space, but are stored as separate

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 54, No. 10, October 1988, pp. 1415–1421.

0099-1112/88/5410–1415$02.25/0

records (to the first class). This process is continued for all remaining cover classes.

Each cell in the N-dimensional feature space is now tested sequentially by the classifier. The classifier finds the class with the highest empirical probability in the cell and assigns that cover class identity to the cell. The empirical probability for the cell is calculated by dividing the number of pixels in the cell for that class by the total number of pixels tallied for all cover classes in the cell. The number of training area pixels in each land-cover class is normalized.

Any unknown pixel is classified by matching its pixel brightness value in each feature with the equivalent cell in the N-dimensional feature space, and extracting the class and empirical probability from the cell. The class and empirical probability for each pixel can then be displayed as two images. Thus, the classifier is not function based but, rather, considers each cell in the N-dimensional feature space as a separate decision rule.

Description of the Algorithm

Let $(\mathbf{X})$ describe the vector position of a pixel in N-dimensional feature space $(X_1, X_2, \ldots, X_N)$, where X_N is the brightness value or digital number (DN) in spectral band N. Training set data are generated for each i^{th} class, for $i=1,2,\ldots,j$ classes.

Let $P(i \mid \mathbf{X})$ be the probability that class i occurs at vector position $(\mathbf{X})$. Using Bayes' Theorem,

$$P(i|\mathbf{X}) = \frac{P(\mathbf{X}|i)P(i)}{P(\mathbf{X})}.$$

An estimator of $P(\mathbf{X}|i)$ is

$$P(\mathbf{X}|i) = \frac{F_i(\mathbf{X})}{F_i},$$

where $F_i(\mathbf{X})$ is the count of pixels from training sets of class i at $(\mathbf{X})$ and $F_i = \Sigma F_i(\mathbf{X})$.

Now, $P(\mathbf{X}) = \Sigma P(\mathbf{X}|j)P(j)$.

Therefore, $P(i|\mathbf{X}) = \dfrac{P(\mathbf{X}|i)\ P(i)}{\Sigma P(\mathbf{X}|j)\ (P(j)}$ (see also Geisser 1982)

$$= \frac{(F_i(\mathbf{X})/F_i)\ P(i)}{\Sigma\ (F_j(\mathbf{X})/F_j)\ P(j)}. \quad (1)$$

The total number of pixels in class i (i.e., F_i) may be normalized using the total number of pixels sampled (i.e., F_j).

$$P(i|\mathbf{X}) = \frac{(F_j/F_i)\ F_i(\mathbf{X})\ P(i)}{\Sigma\ F_j(\mathbf{X})\ P(j)}, \quad (2)$$

where $P(i)$ is the *a priori* probability for class i, in this case the relative areal extent of the classes in the image. (F_j/F_i) is the sum of all training area pixels divided by the sum of pixels in class i, that is a weighting factor to normalize training area fields of different size.

A decision rule can be generated for each vector position in feature space $(\mathbf{X})$ by allocating to the vector position $(\mathbf{X})$ the class i with the highest probability of occurrence, $P(i \mid \mathbf{X})$. That is, if $\{P(i \mid \mathbf{X}) > P(i+1 \mid \mathbf{X})\}$, then the decision rule will allocate class i and probability $P(i \mid \mathbf{X})$ to vector position $(\mathbf{X})$. In this way, a two-dimensional lookup table of vector position $(\mathbf{X})$ against probability and class number can be generated for all $(\mathbf{X})$. In the event of two (or more) classes having equal probabilities, the smallest class number is assigned to the class. Alternative strategies can include stating *a priori* which class should have preference, summing $P_i(\mathbf{X})$ in the adjacent vector positions and selecting the class with the highest sum, or randomizing the selection (i.e., flip a coin).

An unknown pixel, which is located at vector position $(\mathbf{X})$ in N-dimensional feature space, is assigned to the class and empirical probability value equated with $(\mathbf{X})$ in the lookup table.

The training sets usually contain relatively few pixels with a considerable range in brightness values and many missing values. As unfilled feature vector spaces result in unclassified image pixels, the algorithm is obviously sensitive to the degree of sparseness. To overcome the problem of sparseness, the data distributions are collapsed into a smaller brightness value range by multiplying the pixel brightness value by a factor of less than one, and truncating the result to an integer (i.e., $(\mathbf{X}') = \{\text{integer of } (\mathbf{X}.\text{f})\}$, where $0<\text{f}<1$). The result is fewer vacant feature vector spaces and more classified pixels in the image. Substituting $(\mathbf{X}')$ into Equation (2) yields $P(i \mid \mathbf{X}')$. For example, if the original data were 8 bit (0 to 255 brightness levels), than a collapsing factor of 0.5 would reduce the radiometric sensitivity to 7-bit data (0 to 127), while a collapsing factor of 0.25 would result in a 6-bit (0 to 63) data set. Obviously, if $\{\Sigma\ F_j(\mathbf{X}')\ P(j)\}$ in Equation 2 is 0, then $P(i \mid \mathbf{X}')$ will be undefined, and unknown pixels equated with the vector position $(\mathbf{X})$ will be unclassified.

The algorithm was written in Fortran-77 and executed on the VAX-cluster at the Australian National University. The SPIRAL geographic information system (Myers, 1986) was used for data input/output. The UNIRAS software package (European Software Contractors, 1982) was used to display the thematic images on Tektronix graphics hardware.

EVALUATION OF THE ALGORITHM

Methods

A study area of 12 by 15 km was selected over a section of the Stromlo radiata (or Monterey) pine (*Pinus radiata*) plantations near Canberra, A.C.T., Australia (see Figure 1). The plantations are even-aged stands planted between 1930 and 1986. Eight cover types were identified (*viz.* pine older than 40 years, 30 to 40 year old pine, 20 to 30 year old pine, 10 to 20 year old pine, pine less than 10 years old, urban, and water) and at least

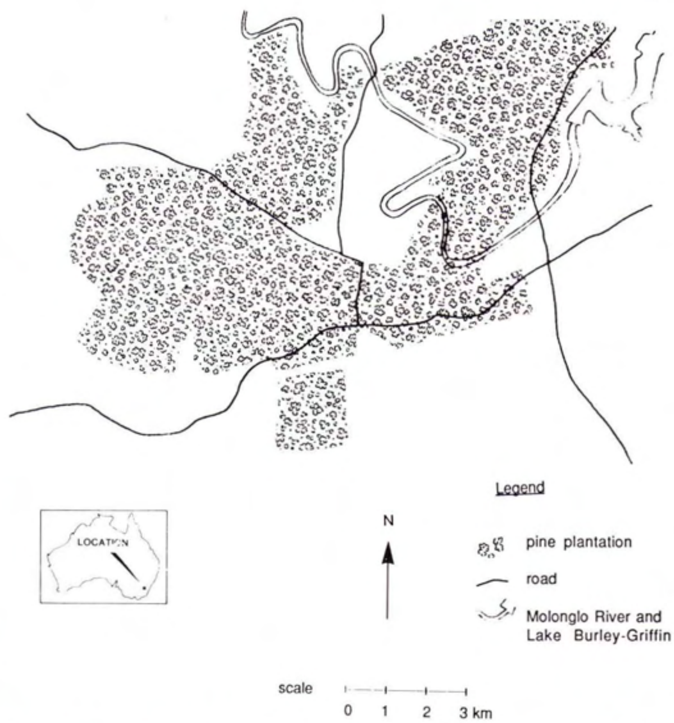


FIG. 1. Map of the study area showing the plantation boundary, urban boundary, roads, and Molonglo River.

two training areas per class were delineated. The pine cover classes had experienced different silvicultural treatments (thinning, pruning, etc.) but a history of treatments was not available, so the age cover classes could not be segmented by stand structure. The terrain varies from undulating (0 to 10 degrees) to moderately steep (approximately 25 degrees). Aerial photographs at a scale of 1:10,000 and compartment maps were available for ground truth reference.

A SPOT XS scene (K386,J421) centered over Canberra, Australia, provided the three channel remotely sensed data (Turner *et al.*, 1987). This cloud free scene was acquired on 11 September 1986, which is late winter in southeast Australia. The view angle was from the right at 4.5 degrees.

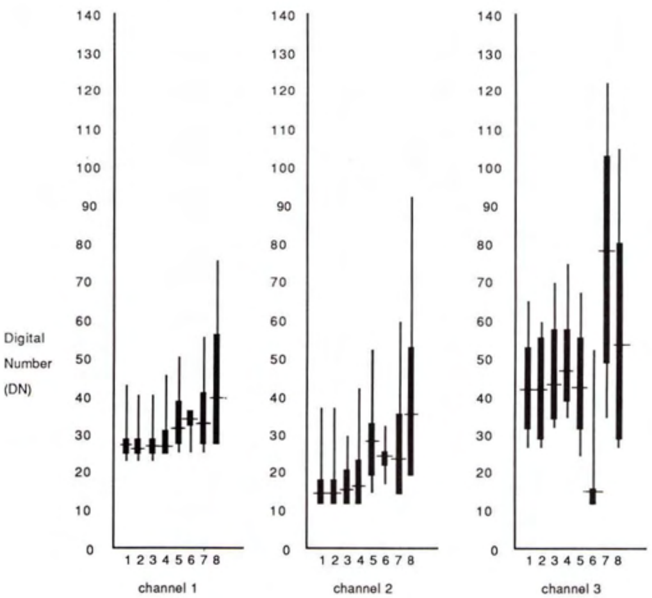
A preliminary unsupervised clustering of the area using the CLUS algorithm (Turner *et al.*, 1982) yielded the approximate areal extent of the eight cover classes. These data were used as the initial *a priori* probabilities in the classifier (i.e., '$P(i)$' in Equation 2). These probabilities were modified empirically to improve the discrimination of cover type classes in the final thematic image.

The training area data were extracted and statistics (mean, covariance matrix, range) generated. In order to visually compare the locations of the training area data in N-dimensional feature space, boxplots (Minitab, 1986) of the DN values for each cover class in the three channels were drawn (Figure 2). The boxplots visually indicate the spread of the data, the skewness of the distribution around the median, and the range within the data spread where most of the observations occur.

The proposed algorithm was tested with various collapsing factors (f) between 0.1 and 1. With a factor of $f = 0.1$, the training data sets tended to merge, lowering the empirical probabilities that pixels were correctly classified but increasing the number of classified image pixels. With no collapsing ($f = 1$), there were more unclassified image pixels but empirical probability of a correct classification was higher. Though any collapsing factor can be used, empirical testing using SPOT and Landsat MSS data indicates a value between 0.5 and 0.8 (i.e., a dynamic range decrease from 256 DN to 128 and 205 DN, respectively) ensures a reasonable compromise between reducing the number of unclassified pixels and increasing empirical probabilities. The thematic image presented here (Figure 3) was produced with a collapsing factor of 0.5. Skidmore *et al.* (1988) examined the effect of varying the collapsing factor for two classes. Spectrally similar cover classes were merged at a relatively high collapsing factor, while spectral cover classes that do not occur in close proximity in N-dimensional feature space will be merged at a lower collapsing factor. As the collapsing factor decreases towards 0, the likelihood that either of the classes will co-occur (or overlap) at a vector in N-dimensional feature space increases, until at the collapsing factor limit of 0, all classes will be merged into a single vector.

A maximum-likelihood classifier (MAXCLASS) and a Euclidean distance classifier (CLASS) contained in the ORSER package (Turner *et al.*, 1982) were executed using the same image and training areas as used in testing the proposed classifier.

A quantitative analysis of mapping accuracy for the maximum-likelihood, Euclidean distance, and supervised nonparametric classifiers was performed using the methodology proposed by Hay (1979). At least 50 pixels were randomly located within each cover class stratum on a 1:25,000-scale compartment map that was geometrically rectified with the classified images. The large number of roads within and bounding the plantation area made it easy to locate the pixels to be tested. The mapping accuracies are summarized as error matrices (Kalensky and Scherk, 1975). A fourth error matrix was generated to include those pixels sampled from the supervised nonparametric classification that had an empirical probability of more than 75 percent, to ascertain whether a relationship exists between empirical probability and overall mapping accuracy.



Key: 1= before 1949 pine; 2 = 1950-1959 pine; 3 = 1960-1969 pine; 4 = 1970-1979 pine; 5 = after 1980 pine; 6 = water; 7 = grass; 8 = urban.

Note: The median is represented by the cross-bars. The total range of the data is shown by the thin lines. The thick lines represent the range between "adjacent values" defined as: {LQ - 1.5 x S} and {UQ + 1.5 x S}, where

LQ = the lower quartile around the median

UQ = the upper quartile around the median

S = UQ - LQ. (See Minitab, 1986 for details).

FIG. 2. Boxplots of the digital number (DN) range in each channel for the training areas.

RESULTS

The boxplots of the DN values for the training areas are presented in Figure 2. The average number of pixels per training area class was approximately 900, and ranged from about 700 to 2000.

The thematic image produced by the nonparametric classifier is shown in Figure 3, and the empirical probability of correct classification is shown in Figure 4. The thematic maps output by the maximum likelihood and euclidean distance algorithms are included as Figures 5 and 6, respectively.

Error matrices for the three classification strategies, and also for pixels with a probability of more than 75 percent correct classification, are detailed in Tables 1 to 4. Note that six classes were tested, comprising the five pine age classes and a 'non-pine' class which included the grass, water, and urban classes. The overall classification mapping accuracies are summarized in Table 5.

DISCUSSION

The boxplots in Figure 2 show that the pine age classes are spectrally similar in all three channels, once the pine tree crowns have closed together. Pine trees greater than 10 years old had closed crowns (i.e., classes 1 to 4). The non-pine classes (urban, grass, and water) differ spectrally from each other and from the pine, particularly in channels 2 and 3. The similar pattern of the boxplots in Figure 2 indicates that channels 1 and 2 are

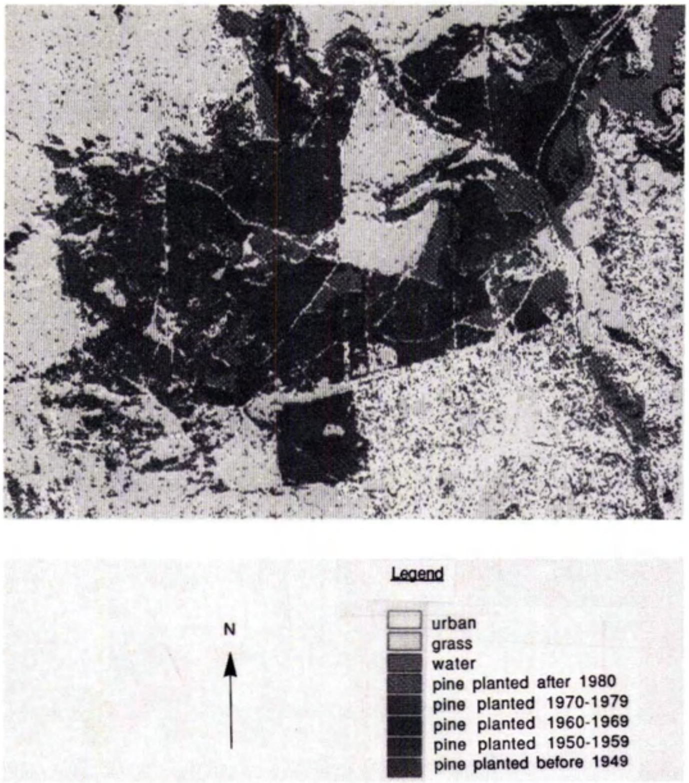


FIG. 3. Thematic map produced by the nonparametric classifier.

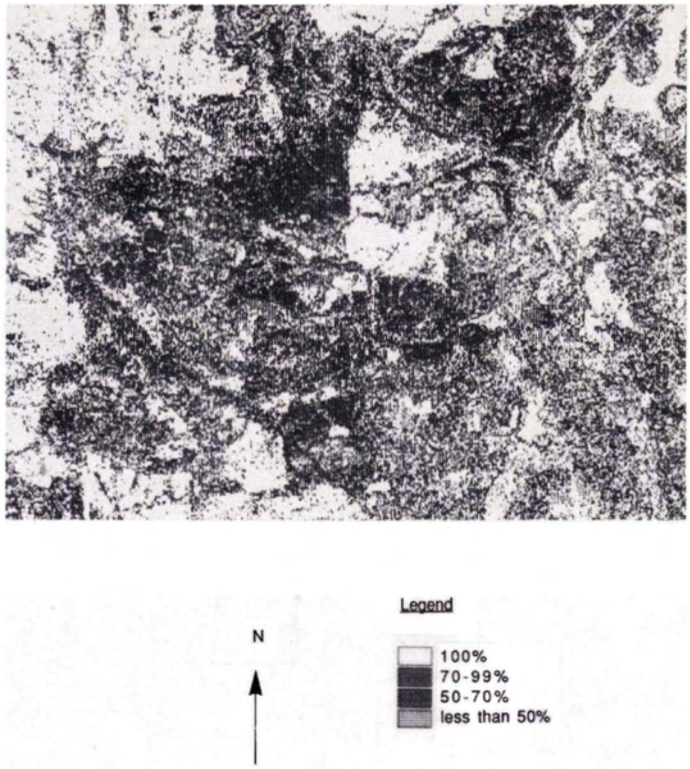


FIG. 4. The empirical probability of correct classification.

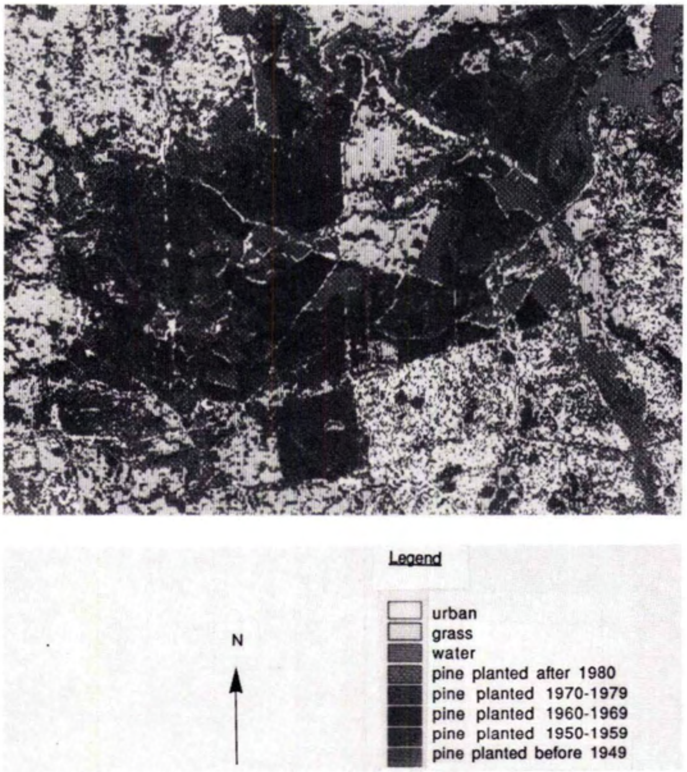


FIG. 5. The thematic image produced by the maximum likelihood classifier.

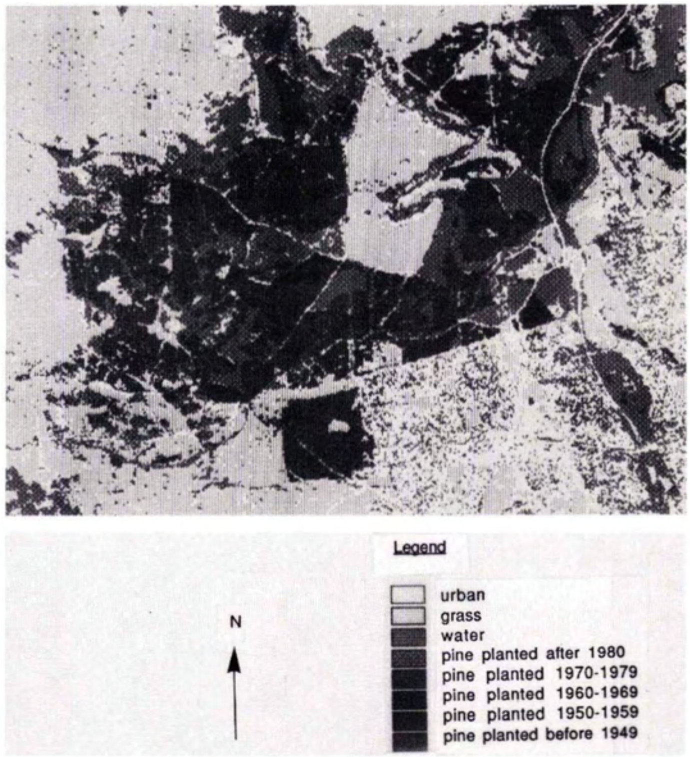


FIG. 6. The thematic image produced by the Euclidean distance classifier.

correlated, with channel 2 having lower DN values than channel 1. This was confirmed by inspecting the correlation matrices for the six classes, which showed a correlation coefficient between channels 1 and 2 of greater than 0.90 for every class except for the pine class less than 10 years old. The boxplots also show that the classes are skewed towards the lower values in channels

TABLE 1. ERROR MATRIX FOR THE SUPERVISED NONPARAMETRIC CLASSIFIER

	Number of pixels							Omissions		
Class	I	II	III	IV	V	VI	VII	Total	No.	%
I	29	6	14	6				55	26	47
II	5	22	14	8			1	50	28	56
III	1		41	12				54	13	24
IV			4	43		3		50	7	14
V	2		1	6	39	2		50	11	22
VI				1	9	46	1	57	11	19
Total no. of pixels	37	28	74	76	48	51	2	316	96	
Overall classification accuracy*	70%									

Table Legend: I = greater than 40 year old pine
II = 30 to 40 year old pine
III = 20 to 30 year old pine
IV = 10 to 20 year old pine
V = pine younger than 10 years old
VI = non-pine (including water, grass and urban classes)
VII = unclassified

*Ratio of the sum of correctly classified pixels in all classes to the sum of the total number of pixels tested.

TABLE 2. ERROR MATRIX FOR THE MAXIMUM-LIKELIHOOD CLASSIFIER

	Number of pixels							Omissions		
Class	I	II	III	IV	V	VI	VII	Total	No.	%
I	30	9	11	2	1	2		55	25	45
II	9	9	16	10	6			50	41	82
III	19	10	11	11	2	1		54	43	79
IV	1		3	39		7		50	11	22
V	2			3	45			50	5	10
VI				1	14	42		57	15	26
Total no. of pixels	61	28	41	66	68	52		316	142	
Overall classification accuracy	56%									

TABLE 3. ERROR MATRIX FOR THE SUPERVISED EUCLIDEAN DISTANCE CLASSIFIER

	Number of pixels							Omissions		
Class	I	II	III	IV	V	VI	VII	Total	No.	%
I	26	6	16	5	1		1	55	29	53
II	5	8	15	16	5		1	50	42	84
III	23		19	11		1		54	35	64
IV			11	36		3		50	14	28
V	4			3	42		1	50	8	16
VI	1				14	27	15	57	30	53
Total no. of pixels	59	14	61	71	62	31	18	316	160	
Overall classification accuracy	50%									

1 and 2. The high spectral variance is also obvious, especially for channel 3.

Given the spectral similarity of the cover classes to be discriminated, the utility of the proposed nonparametric classifier in improving mapping accuracies becomes apparent (Table 5). The nonparametric classifier yielded a higher mapping accuracy (70 percent) than the maximum likelihood (56 percent) or the Euclidean distance (50 percent) classifiers using identical training area data. Tables 1 to 3 show that the classifiers had the lowest mapping accuracies when discriminating between the older pine age classes (i.e., greater than 20 years old). Figure 2 shows that these classes were spectrally the most similar. Nevertheless, the supervised nonparametric classifier achieved adequate class mapping accuracies for the 20 to 30 year old pine, and had a comparatively better class mapping accuracy for the 30 to 40 year old pine. The class mapping accuracies for the more than 40 year old pine was similar for all three classifiers.

As part of the mapping accuracy check, an error matrix was formed using only those pixels with an empirical probability greater than 75 percent. This threshold was chosen to provide

TABLE 4. ERROR MATRIX FOR PIXELS GENERATED BY THE SUPERVISED NONPARAMETRIC CLASSIFIER WITH AN EMPIRICAL PROBABILITY OF GREATER THAN 75 PERCENT

	Number of pixels							Omissions		
Class	I	II	III	IV	V	VI	VII	Total	No.	%
I	27		4					31	4	13
II		13	3					16	3	19
III			12	1		1		14	2	14
IV	1			13		1		15	2	13
V				2	30			32	2	6
VI					6	44	1	51	7	13
Total no. of pixels	28	13	19	16	36	46	1	159	20	
Overall classification accuracy	87%									

TABLE 5. SUMMARY OF MAPPING ACCURACY RESULTS

Classifier	Overall mapping accuracy (%)
supervised nonparametric	70
supervised nonparametric (>75% empirical probability)	87
maximum likelihood	56
supervised Euclidean distance	50

a subsample of pixels with a high empirical probability. The resulting mapping accuracy was 87 percent, which was better than any other classifier including the nonparametric classifier error matrix where the empirical probability ranged from 0 to 100 percent (Table 1). However, the smaller sample size (less than 50 pixel samples per cover type) meant that the precision of the accuracy figures was lower. This is consistent with preliminary results presented by Skidmore (1987), where high mapping accuracies (90 percent) were obtained from Landsat MSS data while discriminating between six forest types in Pennsylvania, using pixels that had a 100 percent empirical probability of correct classification.

Those areas having a low empirical probability of correct classification can be visually isolated (Figure 4). For example, in the present study, pine older than 40 years and the 30 to 40 year old pine classes tended to have an empirical probability of correct classification of less than 50 percent. To improve class mapping accuracy, new training areas may be selected, and old training areas checked for consistency in delineating the class. Then, if the empirical probability of correct classification stays low, it may be concluded that classes are spectrally similar and could perhaps be merged into one class.

The superior mapping accuracy result obtained with the proposed classifier is due to the lack of assumptions concerning the shape or distribution of decision volumes which are implicit in other classification strategies. Each vector in N-dimensional intensity space is treated as a separate decision rule by the proposed classifier. The supervised nonparametric classifier requires *a priori* probabilities as an integral part of the classification, and these prior probabilities can be modified easily to improve the thematic map. Although *a priori* probabilities could be included in maximum likelihood and Euclidean distance classifiers (Duda and Hart, 1973), it is not commonly done, so in this study the *a priori* probabilities were assumed to be constant.

Strahler *et al.* (1980) stated that an advantage with their proposed nonparametric logit classification model is that it can accept continuous data (i.e., Landsat data) or categorical data (i.e., a soil map) as independent variables to model (by regression) the probability that a pixel is a member of a given class. The decision rule described here for the nonparametric classifier is based upon ordered statistics, and so any data types may be included as features in the classification without violating statistical assumptions.

The algorithm required approximately four times the CPU time compared with the maximum-likelihood classifier, for 200,000 pixels. The CPU time requirements appeared to increase proportionally to $n.\log(n)$, where n is the number of pixels. As additional channels are added, the number of vector spaces (or bins) would increase. For example, using TM data, 256^7 (7×10^{16}) bins would be required, many of which would be empty. In this case some feature reduction technique, such as principal component analysis, could be used. The improved mapping accuracies obtained with the proposed classifier have to be offset against the higher computational expense.

ACKNOWLEDGMENTS

The assistance of Dr. W. Myers of The Pennsylvania State University in suggesting some computational enhancements to the program algorithm, and Dr. D. Jupp of the Commonwealth Scientific and Industrial Research Organisation, Canberra, Australia, in clarifying the mathematical proof of the algorithm, is appreciated. The Forestry Commission of N.S.W., Australia, provided study leave funding for the senior author while this work was undertaken.

REFERENCES

Anderson, J. R., E. E. Hardy, J. T. Roach, and R. E. Witmer, 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*: USGS Professional Paper 964, 28 p.

Beaubien, J., 1979. Forest type mapping from Landsat digital data. *Photogrammetric Engineering and Remote Sensing* 45(8):1135–1144.

Cibula, W. G., and M. O. Nyquist, 1987. Use of topographic and climatological models in a geographic data base to improve Landsat MSS classification for Olympic National Park: *Photogrammetric Engineering and Remote Sensing* 53(1):67–75.

Duda, R. O., and P. E. Hart, 1973. *Pattern Classification and Scene Analysis*. Wiley and Sons, New York.

European Software Contractors, 1982. *Uniras Users Manuals for GEOPAK, RASPAK and GIMAGE*. Gentofte, Denmark.

Fleming, M. D., 1975. *Computer aided analysis of Landsat MSS data: a comparison of three approaches including a modified clustering approach*. Purdue Univ., LARS Information Note 072475, 9p.

Geisser, S., 1982. *Bayesian discrimination. Handbook of Statistics*, (P. R. Krishnaiah and L. N. Karnal, editors) Volume 2, Chapter 4. North-Holland, Amsterdam.

Hame, T., 1984. Landsat-aided forest site type mapping. *Photogrammetric Engineering and Remote Sensing* 50(8):1175–1183.

Hay, A. M., 1979, Sampling design to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing* 45(4):529–533.

Hutchinson, C. F., 1982. Techniques for combining Landsat and auxillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing* 48(1):123–130.

Kalensky, Z., and L. R. Sherk, 1975. Accuracy of forest mapping from Landsat CCT's. *Proc. 10th Int. Symp. on Remote Sensing of the Environment* Vol II:1159–1163.

Kettig, R. L., and D. A. Landgrebe, 1976. Classification of multispectral image data by extraction and classification of homogeneous objects. *I.E.E.E. Tran. Geo. Elect.* GE–14:19–46.

LaPerriere, A. J., P. C. Lent, W. C. Gassaway, and F. A. Nodler, 1980. Use of Landsat data for moose habitat analysis in Alaska. *J. Wildl. Management* 44(4):881–887.

Merola, J. A., R. A. Jaynes, R. O. Harniss, 1983. Determination of aspen/conifer forest mixes from multitemporal Landsat digital data. *Proc. 17th Int. Symp. on Remote Sensing of Environment*, Vol II:883–893.

Minitab, 1986. *Minitab reference manual - release 5.* Minitab Inc., State College, Pennsylvania.

Myers, W. L., 1986. *SPIRAL Steps and System Structure:* Office of Remote Sensing of Earth Resources, Pennsylvania State University, University Park, Pa. 16802, U.S.A., Publ. no LW8607.

Nelson, R. F., 1981, A comparison of two methods for comparing forest land. *Int. J. Remote Sensing* 2(1):49–60.

Nilsson, N. J., 1965. *Learning Machines.* McGraw-Hill, New York.

Richards, J. A., 1986. *Remote Sensing - Digital Analysis.* Springer-Verlag, New York.

Skidmore, A. K., 1987, A supervised nonparametric classifier to improve forest mapping accuracy. *Fourth Australasian Remote Sensing Conference,* Adelaide, Australia. Vol. II:573:583.

Skidmore, A. K., G. B. Wood, and K. R. Shepherd, 1987. Remotely sensed digital data in forestry: a review. *Aust. For.* 50(1):40–53.

Skidmore, A. K., G. W. Forbes, and D. Carpenter, 1988. A test of overlap in multispectral digital data. *Int. J. Remote Sensing* 9(4):777–785.

Strahler A. H., T. L. Logan, and N. A. Bryant, 1978, Improving forest cover classification accuracy from Landsat by incorporating topographical information. *Proc. 12th Int. Symp. on Remote Sensing of Environment* Vol II:927–942.

Strahler, A. H., J. E. Estes, P. F. Maynard, F. C. Mertz, and D. A. Stow, 1980. Incorporating collateral data in Landsat classification and modelling procedures. *Proc. 14th Int. Symp. on Remote Sensing of Environment* Vol II:1009–1026.

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

Thompson, D. G., G. H. Kassen, and J. Cihlar, 1980. Caribou habitat mapping in the southern district of Keewatin, N. W. T.: An application of digital Landsat data. *J. Appl. Ecol.* 17:125–138.

Tom, C. H., and L. D. Miller, 1980. Forest site index mapping and modeling. *Photogrammetric Engineering and Remote Sensing* 46(2):1385–1596.

Turner, B. J., W. L. Myers, and G. B. Baumer, 1982. *The ORSER Remote Sensing Analysis System: A User's Manual.* ORSER, Pennsylvania State University, University Park, Pennsylvania. Res. Publ. 109/OR.

Turner, B. J., D. M. Moore, and A. K. Skidmore, 1988. Forest management application of SPOT data in Australia. *Proc., SPOT-1:Image Utilization, Assessment, Results International Conference*, Paris, France. (In press).

Walsh, S. J., 1980. Coniferous tree species mapping using Landsat data. *Remote Sensing of Env.* 9:11–26.

(Received 9 March 1988; revised and accepted 1 July 1988)