James B. Campbell
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

Spatial Correlation Effects upon Accuracy of Supervised Classification of Land Cover

The interaction between sampling procedure and natural variability within training fields can lead to noticeable errors in applications of supervised classification procedures.

Introduction

Increasing use of Landsat digital data during the past eight years has been accompanied by increasing interest in evaluation of accuracy of Landsat-based land-cover maps (e.g., Todd *et al.*, 1980)—an interest that has focused largely upon strategies for defining, then measuring, map accuracy. Less attention has been devoted to identifying and eliminating causes of error. This paper defines a specific source of error that may occur in supervised classification—error resulting from the interaction between natural variability within homogeneous land-cover parcels and the usual sampling procedures.

Abstract: *The Landsat 1 MSS imaged south central Virginia at six dates during the 1974 growing season. Data representing five rural land-cover categories (at detail corresponding roughly to Anderson's Level II) were selected in a manner that assured that the spectral values represented the same geographic areas at each date. Examination of selected parcels of forest reveal the presence of positive continuity between MSS values at adjacent pixels; the degree of continuity varies from date to date, and according to MSS bands at a given date. In at least some instances it can be domonstrated that estimates of category variances based upon values at contiguous pixels yield low values relative to those based upon random samples of the same area. These biased estimates can lead to overestimation of contrast between categories and to errors in supervised classification. Erroneously classified pixels may tend to cluster, thereby increasing the opportunity for misinterpretation of errors as genuine land-cover parcels. The amount of error varies throughout the growing season. Such error could be avoided by improved sampling strategies or by application of more sophisticated estimation procedures.*

Application of the supervised classification technique for analysis of multispectral remote sensing data requires that training areas of known identify be used to specify class signatures. Picture elements (pixels) of unknown identity are then assigned to one of several *a priori* categories. Although accurate estimation of category characteristics requires independent samples of the pixels composing each category, the usual practice of deriving estimates from groups of contiguous pixels (i.e., from the entire membership of each training field) assures that estimates will be biased if there is a positive relationship between values at neighboring pixels (Basu and Odell, 1974; Tubbs and Coberly, 1978). Because the presence of such relationships can be detected in at least some data collected by the Landsat MSS, there is reason to suspect the presence of inherent error in the usual applications of supervised classification algorithms to Landsat digital data.

Here this possibility is investigated by examination of the character of spatial variation within multitemporal Landsat MSS data of an area in central Virginia. Examination of these data indicates the presence of spatial continuity among values at neighboring MSS pixels representing oak-hickory forest, and clear changes, over time, in the allocation of variance over distance within the same study area. Trials with an alternative sampling

Photogrammetric Engineering and Remote Sensing,
Vol. 47, No. 3, March 1981, pp. 355-363.

strategy (designed to circumvent effects of positive association between adjacent values) indicate that consequent errors may attain sufficient magnitude to influence classification accuracy. There is evidence that the magnitude of these errors varies throughout the growing season, probably in response to variations in the degree of place-to-place continuity within categories.

Study Area and Data

This study examines an area located in central Virginia, about 25 miles (40 km) southwest of Richmond. This area is centered on the Sutherland, Virginia USGS 7.5-minute quadrangle, a rural area characterized by relatively small parcels of land occupied by cropland, pasture, rural residences, farm buildings, and mixed forest. Mature oak-hickory forest and oak-pine forest occur in extensive units, often mixed with small parcels of planted loblolly and shortleaf pine (Cost, 1976). One large water body and several smaller lakes and ponds are present. The area is portrayed on topographic maps at 1:24,000 (1963, revised 1974), black-and-white aerial photography at 1:23,500 (March, 1971), and high altitude color infrared photography 1:132,000 (August, 1976).

Landsat 1 MSS data of this area provide a record of seasonal land cover changes at six dates during the 1974 growing season (Table 1). Close examination of chromaticity coordinates for water pixels (Alföldi and Munday, 1977), together with frequency histograms of MSS counts, hourly weather observations (visibility and humidity), and measurements of suspended particulate matter, indicate that most of these data are free of substantive atmospheric interference. The exception are the data for June, which appear to be significantly influenced by atmospheric moisture. Despite evidence indicating degradation of the June data, they were retained for analysis without correction—a fact that should be considered in interpreting results. Because no effort was made to compare pixel values from date to date, it was not necessary to standardize the values for atmospheric effects or for differences in sun angle. Thus, MSS counts were examined in the same form as they are used for many applications of supervised classification, without efforts to correct for known or suspected radiometric errors. No pre-processing was performed on any of the data used for this study.

Table 1. Landsat-1 MSS Data Used for This Study

DATE	SCENE ID
27 February 1974	1584-15152
22 April 1974	1638-15141
28 May 1974	1674-15131
15 June 1974	1692-15124
13 September 1974	1782-15092
24 November 1974	1854-15071

The April data provide the best contrast between categories, so they were examined to select training fields representing each of the four predominant land cover categories represented in the area:oak-history forest, cropland, pasture, and open water. Five separate training fields were selected for each category. Although other land-cover categories are present, they occur in parcels too small to be consistently identified on the Landsat data. Each parcel was carefully selected with the aid of topographic maps and aerial photography to form a homogeneous representation of each category and to be easily identifiable with reference to distinctive features recognizable both on the digital data and on maps and photographs. These areas, selected following the guidelines suggested by Joyce (1978), were rectangular or square in shape, and generally included at least 20 but no more than 100 pixels each (one large area included 121 pixels). Particular care was taken to avoid mixed pixels by selecting vertices well within a buffer zone several pixels wide defined to border the boundary of each category (Figure 1). Line-printer brightness maps (produced by the ORSER NMAP program) were used to identify coordinates of the vertices of each parcel; then the ORSER UMAP program was used to assess the uniformity of each area (Turner *et al.*, 1978).

After areas were located within the April data, they were again identified (in respect to size, shape, and position) by inspecting line-printer brightness maps of data for the other five dates. Field observations, and examination of maps and aerial photographs, verified that there had been no changes in land cover within these areas during the interval covered by these data. It was, therefore, possible consistently to locate the borders of

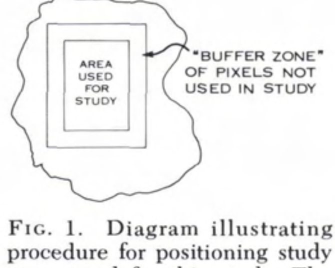

Fig. 1. Diagram illustrating procedure for positioning study areas used for this study. The shaded zone bordering outside edge of the study area represents the minimum width of a "buffer zone" of pixels intended as protection against the possibility of including border pixels in the data.

TABLE 2. DISTRIBUTION OF VARIANCE WITHIN A FOREST AREA RECORDED BY THE LANDSAT-1 MSS (BAND 5)

	Date					
	Feb	Apr	May	Jun	Sep	Nov
Mean:	16.2	24.4	17.4	23.5	12.9	12.0
Variance:	3.42	4.58	3.08	1.40	0.68	1.81
Proportion of total variance contributed by following sources of variation:						
4 by 4 cells within 8 by 8 cell:	.11	.21	.13	.06	.14	.16
2 by 2 cells within 4 by 4 cells:	.47	.38	.48	.33	.31	.34
individual observations within 2 by 2 cells:	.42	.41	.39	.61	.55	.50

Source: calculated from data described in the text using the procedure described by Mollering and Tobler (1972).

the same parcels within data for six separate dates, record coordinates for vertices, then list the MSS data for each area, each date, in all four MSS bands. Inspection of the areas in the field, on maps and aerial photographs, and in digital representation confirmed that these areas are as homogeneous, in respect to land cover and radiometric response, as one can reasonably expect parcels this size to be.

ALLOCATION OF VARIANCE OVER DISTANCE

To illustrate some characteristics of the distribution of variance over distance in digital remote sensing data, this study examines MSS values representing a relatively uniform area of oak-hickory forest at the six dates specified above. Although examination of aerial photographs of this area reveals the presence of variations in density within areas equivalent to one or two pixels, this area is typical of forest areas in this region, and the variations present are comparable to those encountered in normal applications of the supervised approach. The following analysis is based upon data from bands 5 and 6, which together form a concise example of the issues discussed here.

The MSS data are collected at regular spatial intervals; as a result, areas measuring two, four, eight, sixteen, thirty-two . . . pixels on a side form nested arrays of the kind discussed by Moellering and Tobler (1972). Because of the need to restrict this analysis to a single homogeneous category, the array examined here is limited to 64 pixel values within an array measuring eight units on a side, a size comparable to that suggested by Joyce (1978) for training areas for supervised classification. (Larger arrays might be desirable, but would incur the risk of including mixed pixels.)

The nested character of these arrays permits convenient description of the allocation of variance at different scales in the arrays (that is, the total sum of squares can be partitioned into amounts assigned to each level in the hierarchy, as described by Moellering and Tobler (1972)). In the present context this procedure is of interest as a means of comparing the variation of data acquired at different dates in different spectral regions. Within the 8 by 8 cells observed at each date, there are three levels available for examination: 4 by 4 cells within the 8 by 8 cell, 2 by 2 cells within each 4 by 4 cell, and the individual observations within the 2 by 2 cells (Tables 2 and 3).

As a broad generalization, the lowest level of variation (the 2 by 2 cells) accounts for about 50 percent of the total for each date. The remaining 50 percent of the total variance is split between the two remaining scales. There is no immediately recognizable pattern to the temporal changes in the distribution of variance, although there is the possibility that the distinctive character of the June data might be the result of the atmospheric conditions mentioned previously.

These results reveal clear differences between representations of this forest area by two MSS

TABLE 3. DISTRIBUTION OF VARIANCE WITHIN A FOREST AREA RECORDED BY THE LANDSAT-1 MSS (BAND 6)

	Date					
	Feb	Apr	May	Jun	Sep	Nov
Mean:	23.2	40.1	56.3	57.1	34.9	18.0
Variance:	2.53	9.00	18.70	8.02	4.97	2.74
Proportion of total variance contributed by following sources of variation:						
4 by 4 cells within 8 by 8 cell:	.33	.12	.48	.28	.06	.09
2 by 2 cells within 4 by 4 cells:	.15	.40	.05	.63	.40	.31
individual observations within 2 by 2 cells:	.52	.48	.47	.09	.54	.60

Source: calculated from data described in the text using the procedure described by Moellering and Tobler (1972).

bands at several dates. Total variation differs from band to band and from date to date; more importantly, the allocation of variance over distance varies markedly by date and by spectral band. As a result, it seems clear that a training area of specified size has varying effectiveness in representing variation present within a category. For example, a training area measuring two units on a side applied to the band 6 data would record only a small proportion of the variation present in the May data, but would probably represent more accurately the variation present in June or September (Table 2). Note also that the effectiveness, at a given date, of such training area would vary according to the MSS band examined (Tables 2 and 3). The fact that, in practice, training fields would be much larger than the 2 by 2 cell used here as an example is no protection against this effect because it can occur at any scale. In fact we can never know beforehand if a training area of a given size will be large enough to represent accurately the variation present in a given category.

Continuity of Variation

Implicit in the preceding discussion is the concept that the set of all observations separated from each other by a specific distance x will exhibit a different degree of similarity than will the set of all observations separated from each other by a different distance $x + 1$. The degree of change in such similarity is a measure of place-to-place continuity, explicitly represented by the correlogram, which depicts changes in autocorrelation as distance ("lag") changes (Cliff and Ord, 1973). For observations spaced at uniform distances, changes in autocorrelation in relation to distance reveals changes in the mutual dependence of neighboring values at varying intervals. High positive autocorrelation values indicate that there is a close association between values of the parent distribution separated by the distance given by the product of lag value and the distance interval between observations.

Except in instances of periodic variation, natural phenomena frequently exhibit high autocorrelation values at low lags (assuming a relatively short sampling interval), declining to values at or near zero at high lags, as samples are separated by greater distances. (See Sayn-Wittgenstein (1970) for examples calculated from ground data.) If there is positive autocorrelation in the parent distribution, samples must be separated spatially if they are to provide independent information (Agterberg, 1965).

The design and operation of the MSS seem likely to produce positive autocorrelation in the digital data, in part due to the scan pattern and analog-to-digital sampling, as described by General Electric and by Slater (1979). As a result, correlograms calculated from MSS data may include instrumentally-induced positve autocorrelation in addition to whatever continuity may be naturally present in the scene. Because most supervised classification is based upon real MSS data (rather than idealized error-free data), this analysis does not include attempts to remove or adjust for such effects.

Two dimensional correlograms were calculated for the same eight by eight arrays described above (Katz and Doyle, 1963). The results are displayed as perspective block diagrams; the vertical axis represents autocorrelation values (between +1.0 and −1.0) and the two horizontal axes represent north-south and east-west directions (Figure 2). In the interests of clarity and conciseness, it is sufficient here to show only a portion of the complete autocorrelation array to represent differences between separate MSS bands and separate dates. Because the parent distributions measure only eight pixels on a side, interpretation should focus only upon the first few values; higher lags are based upon so few values that they may not accurately represent true variation of the data. A rapid decline in autocorrelation near the origin reveals an absence of place-to-place continuity (e.g., Figure 2c), whereas positive values over several lags (e.g., Figure 2b) indicate close association between values separated by relatively large distances.

The most important feature shared by these correlograms is the presence of persistent positive autocorrelation over the first few lags. (Figures 2a and 2b are representative of most of the data examined here.) This indicates the presence of place-to-place continuity over distances of several pixels, evidence that values at adjacent pixels are not independent. Interdependence of this form is probably a common, but not necessarily universal, characteristic of digital remote sensing data. (The degree of interdependence depends in part upon landscape variability and sensor resolution.) This interdependence has been exploited in the use of textural measures as classification criteria to supplement the use of spectral signatures (Jensen, 1979). Thus, in the context of textural measures, the interdependence of nearby values can be regarded as an aid to accurate classification. But in

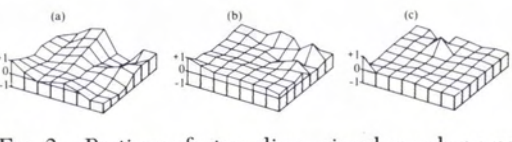

FIG. 2. Portions of a two-dimensional correlograms representing place-to-place variation of MSS values within a single forest area. The shaded portion represents negative autocorrelation; each cell represents a distance of approximately one pixel. (a) and (b) depict MSS band 6 at September and May, respectively. Compare (b) and (c) (band 4, May) to see a contrast between different bands acquired at the same data.

the usual circumstances, it may lead to rather subtle but pervasive errors in the products of supervised classification.

Effects upon Supervised Classification

The information presented above reveals the presence of interdependence between neighboring pixels—an interdependence that may vary in strength and nature through the growing season. Because most (if not all) applications of supervised classification draw upon the entire set of contiguous pixels enclosed in a training field, subsequent estimates of the means, variances, and covariances can be expected to include artifacts of these interrelationships. Depending upon specific characteristics of the landscape under consideration, errors produced by these artifacts might be trivial; or, they could be large enough to cause noticeable differences in the results of the classification process.

One way to investigate the character and magnitude of such errors is to compare values of the means, variances, covariances based upon a number of contiguous pixels (the usual procedure for finding statistics for training fields) with values based upon an equal number of pixels distributed randomly within the same training field. The randomly selected pixels should reduce the effect (if any) of positive interrelationships between adjacent pixels upon the subsequent estimates (although chance clustering of randomly selected pixels may still produce some error). Any differences in values of the means, variances, and covariances can then be ascribed to differences in the sampling procedures.

Each of the data sets discussed above (64 pixels each, represented in four bands) was sampled first be selecting a block of 16 contiguous pixels arranged in a 4 by 4 pattern, positioned randomly within the training field. Then the same 64 pixels were sampled again by randomly selecting 16 coordinate pairs to select another set of 16 values. Both sets of observations were used to derive separate estimates of means and variances in all four MSS bands, as well as covariances between bands. The results discussed below are based upon averages of three independent trials of each of these two alternative sampling strategies.

To avoid a completely subjective assessment of differences between the two procedures, the multivariate tests given by Morrison (1976) are used to compare results. The use of contiguous samples does not meet the requirements of independence and randomness required by parametric tests, so there is no basis for a rigorous interpretation of the outcome. Thus, these tests are offered here, in the absence of a more satisfactory procedure, as rather rudimentary evidence of the similarities and differences between the two sets of values.

The random sampling strategy produces values for the mean that are only slightly different from those based upon contiguous values; these differences are so small that they are unlikely to be a source of error in the classification process (Table 4). (A test for equality of vector means shows no significant difference at $\alpha = 0.05$.) However, even casual inspection of the sets of variances reveals consistent differences that seem likely to have a significant influence upon the classification process (Table 5). Variances based upon contiguous samples are consistently lower than those based upon random samples. One effect of these differences is evident in measures of statistical separability. For example, values for the normalized difference, $\bar{D}_n = (\bar{x}_1 - \bar{x}_2)/(s_1 + s_2)$, differ greatly depending upon the method used to select pixels (Table 6). The random selection process yields low values (relative to those for contiguous pixels) because differences between means, $\bar{x}_1 - \bar{x}_2$, are similar to those given by contiguous pixels, while the sums of the standard deviations, $s_1 + s_2$, are consistently higher. This effect can be seen in values of D_n for categories present in the study area (Table 6). Thus, the use of contiguous pixels to represent a land-cover category tends to underestimate variation, leading one to believe that an area is more homogeneous, and more distinct from other categories than it really is.

Variance/covariance matrices were calculated for the data derived from the two sampling procedures (Table 7). A test for equality of variance/covariance matrices (Morrison, 1976) indicates

Table 4. Means Produced by Two Sampling Strategies Applied to the Same Data

Sampling Strategy	MSS Band 4	5	6	7
Contiguous	24.3	17.6	55.6	32.2
Random	27.6	17.6	56.4	33.7

Source: calculated from data specified in the text. Values reported here are averages based upon three independent trials (each 16 pixels) applied to a single array of 64 pixels.

Table 5. Variances Produced by Two Sampling Strategies Applied to the Same Data

Sampling Strategy	MSS Band 4	5	6	7
Contiguous	0.40	1.01	14.00	8.92
Random	1.09	3.50	23.15	11.58

Source: calculated from data specified in the text. Values reported here are averages based upon three independent trials (each 16 pixels) applied to a single array of 64 pixels.

TABLE 6. NORMALIZED DIFFERENCES BETWEEN SELECTED LAND COVER CATEGORIES FOR SIX DATES, BASED UPON TWO SAMPLING STRATEGIES APPLIED TO THE SAME DATA

DATE & SAMPLING STRATEGY	NORMALIZED DIFFERENCES BETWEEN CATEGORIES		
	Forest-Cropland	*Forest-Pasture*	*Forest-Water*
February			
contiguous	1.14	2.10	4.20
random	1.40	2.43	4.45
April			
contiguous	2.41	3.11	5.34
random	1.84	2.34	4.32
May			
contiguous	0.21	0.34	5.29
random	0.02	0.06	4.95
June			
contiguous	1.36	0.05	4.20
random	1.22	0.30	4.81
September			
contiguous	0.99	1.39	11.54
random	1.07	1.48	13.50
November			
contiguous	2.79	0.66	7.98
random	1.91	0.50	5.07

Source: calculated from Landsat 1 MSS Data specified in the text (MSS Band 6).

that the two sampling methods yield different matrices (at $\alpha = 0.20$) for all three trials of the two sampling strategies. Because the variances and covariances, together with the vector means, form the basis in supervised classification algorithms for discrimination of categories, evidence of differences between the two matrices indicates the two sampling procedures will yield different classifications, possibly presenting the opportunity for method-produced error in the final product.

Thus, any classifier that requires information concerning class variances and covariances (e.g., maximum likelihood classifiers) will be especially susceptible to the kinds of errors discussed here. Because many of the major image-processing systems include classification algorithms of this kind (Computer Sciences Corporation, 1975; Carter *et al.*, 1977), there is ample incentive to investigate this possibility further. Here, a simple but real example illustrates the practical effects of the choice of sampling procedure. The objective is to assess the effect of sampling strategy upon the accuracy of the supervised approach applied to the discrimination of forest and pasture. Pasture is represented by 18 pixels, selected as described previously. Forest is represented first by 16 contiguous pixels, and again by 16 randomly chosen pixels, selected as described above. Then two discriminant functions were calculated—one to separate forest and pasture, with forest represented by a training field of contiguous pixels, and a second, with forest represented by training data randomly selected from the same parcel of forest.

Next, both discriminant functions were used to classify other MSS data, known to represent the same forest category at nearby sites, and similar data representing pasture. Because the identities of these pixels are known beforehand, the results form evidence of the influence of the role of sample location upon accuracy of the classification process. Results for application of the two discriminant functions to a single parcel of 121 forest pixels and 30 pasture pixels are represented in Figures 3 and 4. The discriminant function based

TABLE 7. VARIANCE/COVARIANCE MATRICES PRODUCED BY TWO SAMPLING STRATEGIES APPLIED TO THE SAME DATA

	RANDOM					CONTIGUOUS			
	(MSS BAND)					(MSS BAND)			
	4	5	6	7		4	5	6	7
4	1.09	1.21	−1.00	−0.51	4	0.40	0.21	−0.78	−0.43
5	1.21	3.50	−1.65	−1.85	5	0.21	1.01	−0.19	−1.10
6	−1.00	−1.65	23.15	12.73	6	−0.78	−0.19	14.00	9.80
7	−0.51	−1.85	12.73	11.58	7	−0.43	−1.10	9.80	8.92

Source: Calculated from Landsat 1 MSS data specified in the text.

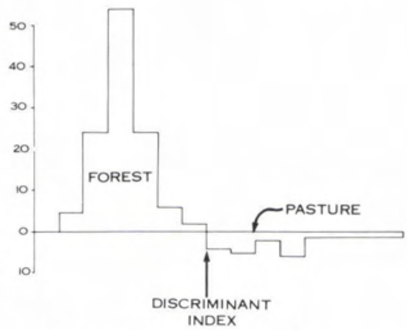

FIG. 3. Discriminant scores for 121 forest (upper side of the diagram) and 30 pasture pixels (lower side) classified using a discriminant function calculated from a training field of 16 randomly selected forest pixels and 18 pasture pixels. All pixels in both categories are correctly classified. (The abscissa represents the discriminant function line; the ordinate shows number of pixels. The discriminant index is the point on the discriminant function line that separates the two categories.)

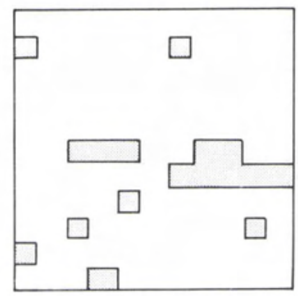

FIG. 5. Distribution of 15 misclassified pixels within an area of 121 forest pixels. Misclassified pixels are represented by the shaded areas.

upon randomly selected training data assigned all 151 pixels to their correct groups (Figure 3). The discriminant function based upon contiguous pixels misassigned 18 of the 121 pixels known *a priori* to belong to the forest category (Figure 4). These results form empirical evidence in support of the more abstract and theoretical suggestion presented above that the locations of samples used as training data may influence the effectiveness of the classification process.

Examination of the locations of these errors (Figure 5) offers additional reasons to investigate this topic further. Inspection of this pattern suggests that there is a tendency towards clustering—a conclusion that is supported by application of a test for spatial autocorrelation of dichotomous areal data (Cliff and Ord, 1973). The results indicate that we would seldom expect to encounter such a pattern in a strictly random assignment of errors to locations within this area. Although these data are insufficient basis for broad generalizations, the implication is that classification errors may cluster in space. If this suggestion is in fact true, it would be an especially insidious manifestation of error because clusters of erroneously classified pixels lend themselves to misinterpretation, whereas randomly distributed errors are more likely to be recognized as errors. For example, if one were to interpret the results represented in Figure 5 in ignorance of the true identities of the pixels, the group of seven contiguous pixels identified as pasture could reasonably be interpreted as a parcel of pasture, while isolated errors would probably be recognized as such because of their scattered locations and small sizes. This tentative suggestion is reinforced by examination of other digital classification products. For example, Mead and Meyer (1977) show digital classification maps of land cover in a northern Minnesota forest, together with manual interpretations of color infrared imagery of the same area. A qualitative, retrospective, inspection of differences between machine and manual interpretations reveals a tendency for clustering of what appear to be erroneously classified pixels. The implication is that similar errors may routinely occur in at least some supervised classification products, yet remain undetected because of their small size and scattered locations.

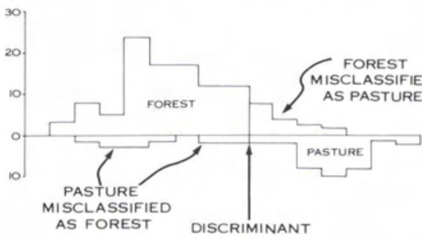

FIG. 4. Discriminant scores for 121 forest (represented on the upper side of the diagram) and 30 pasture (lower side) pixels classified using a discriminant function calculated from a training field of 16 contiguous forest pixels and 18 pasture pixels. Pixels in both categories have been misclassified. (Misclassified pixels are depicted by the dark pattern.) Compare with Figure 3.

TEMPORAL VARIATIONS IN ACCURACY

Given the seasonal variations in the amount and allocation of variation documented in Tables 2 and 3, it seems likely that errors attributable to the use of contiguous samples could vary in magnitude

TABLE 8. ACCURACIES OF TWO SAMPLING PROCEDURES IN CLASSIFYING 250 FOREST PIXELS AT SIX DATES

DATE	RANDOM SAMPLES (% correct)	CONTIGUOUS SAMPLES (% correct)	ACCURACY DIFFERENCE ASCRIBED TO SAMPLING PROCEDURE*
FEB	100	93	−7
APR	100	100	0
MAY	100	85	−15
JUN	62	64	+2
SEP	94	89	−5
NOV	99	99	0

* Negative values signify the decrease in accuracy (relative to randomly located pixels) attributable to the use of blocks of contiguous pixels for training fields; the positive value indicates the single instance in which the contiguous pixels yielded a more accurate classification than did the random pixels.

Source: calculated following the procedure described in the text, based upon 250 Landsat MSS pixels.

during the growing season. To study this possibility, the procedure described above was used to classify MSS data collected at all six dates considered for this study. For each date, 250 forest pixels were classified using both contiguous and random training data; a tabulation of differences in accuracies indicates the varying effect of the choice of sampling strategy (Table 8). (As before, test data consisted of pixels of known identity *not* used to derive discriminant functions.)

The results follow the pattern observed for the May data (Table 8); erroneously classified pixels tend to cluster (not illustrated). Other results were not observed in the May example. The randomly selected training data produced some error in June, September, and November. These errors exhibit the same tendancy for clustering noted previously. The June data produced a much lower degree of accuracy than did any other date, and it is the only date for which random samples were less accurate than contiguous samples. These distinctive features of the June data could possibly be related to the previously mentioned atmospheric degradation of the June data, but there is no specific evidence to support this hypothesis. In general, there does not appear to be any clear seasonal pattern to the errors attributable to sampling strategy.

SUMMARY

Even in this very simple example it seems impossible to identify the sources of classification error and to assign responsibility in a given situation. It is clear that classification errors under the circumstances described here are related to the interaction between numerous scene-related elements that operate simultaneously and with varying influence throughout the growing season. Contrast between categories (Table 6), internal variability within categories (Tables 2 and 3), and seasonal variations in discriminating power of individual MSS bands can be identified as important scene-related variables contributing to variations in accuracy. In addition, the examples presented here demonstrate that the interaction between sampling procedure and natural variability within training fields can lead to noticeable errors in applications of supervised classification procedures. The evidence discussed here is far too limited to form the basis for firm conclusions, yet it does seem sufficient to warrant further examination.

To form a framework for continued research, a set of related hypotheses are offered:

- Homogeneous land-cover categories recorded by multi-spectral remote sensing data (specifically, those gathered by the Landsat MSS) exhibit positive spatial autocorrelation. Positive continuity is probably caused by some combination of (a) inherent continuity of the spatial variation of the landscape at the resolution of the sensor, (b) instrument design and operation, and (c) data processing algorithms.
- Because these data are often sampled using blocks of contiguous pixels to derive estimates of category means, variances, and covariances, the usual training fields may inaccurately or inefficiently represent characteristics of land-cover categories.
- Under these circumstances, use of the entire membership of each training field may lead to errors in applications of supervised classification.
- There is reason to suspect that such errors may tend to cluster spatially. Clustered errors may be especially potent in degrading the usefulness of a classification product because spatially isolated errors can sometimes be recognized as errors, but spatially aggregated errors might be interpreted as genuine land-cover parcels.
- Seasonal variations in the amount and allocation of variance within training fields may lead to seasonal variations in the accuracy of resulting classifications. Thus, the problems listed above may vary in significance throughout the growing season.

Confirmation, modification, or rejection of these possibilities would require analysis of many more data from differing environments observed at several dates. Examination of such data eventually may develop statistical evidence that could form the basis for guidelines specifying the sizes, numbers, and locations of training fields, and the most efficient procedures for selecting numbers and placement of pixels within training fields. Basu and Odell (1974) and Tubbs and Coberly (1978) suggest modification of classification algorithms (based upon assumed characteristics of the autocorrelation structure) to adjust for non-independence of training samples. Because it seems likely that autocorrelation structures vary by land-cover category and by season, it may be more practical to revise sampling procedures than to attempt to devise generally applicable classification algorithms.

ACKNOWLEDGMENTS

Data for this study were acquired at NASA's Goddard Space Flight Center, Greenbelt, Maryland under the sponsorship of a fellowship awarded in 1979 by NASA and the American Association for Engineering Education. The author would like to thank Ms. Elizabeth Middleton and other staff of the Eastern Regional Remote Sensing Applications Center for their advice and assistance. In addition, three anonymous reviewers provided constructive criticism. The author, of course, assumes all responsibility for errors, and emphasizes that the findings are his own and not necessarily those of sponsoring organizations.

REFERENCES

Agterberg, F. P., 1965. The Technique of Serial Correlation Applied to Continuous Series of Element Concentration Values in Homogeneous Rocks, *Journal of Geology,* Vol. 73, pp. 142-154.

Alföldi, Thomas T., and John C. Munday, 1977. Progress Toward a Landsat Water Quality Monitering System, *4th Canadian Symposium on Remote Sensing,* Canadian Aeronautics and Space Institute, Ottawa, pp. 325-340.

Basu, J. P., and P. L. Odell, 1974. Effects of Intraclass Correlation Among Training Samples on the Misclassification Probabilities of Bayes' Procedure, *Pattern Recognition,* Vol. 6, pp. 13-16.

Carter, Virginia, Frederic Billingsley, and Jeannine Lamar, 1977. *Summary Tables for Selected Digital Image Processing Systems.* USGS Open File Report 77-414 (NTIS 277801), 45 pp.

Cliff, A. D., and J. K. Ord, 1973. *Spatial Autocorrelation,* Pion, London, 178 pp.

Computer Sciences Corporation, 1975. *Image Processing Applications Software.* (3000-38700-01TM), Goddard Space Flight Center, Greenbelt, MD, 98 pp.

Cost, Noel D., 1976. *Forest Statistics for the Coastal Plain of Virginia.* USDA Forest Service Bulletin SE-34, Southeast Forest Experiment Station, Asheville, NC, 33 pp.

General Electric Space Division (no date). *Landsat 3 Reference Manual.* Valley Forge Space Center, Philadelphia, PA, 189 pp.

Jensen, John R., 1979. Spectral and Textural Features to Classify Elusive Land Cover at the Urban Fringe, *The Professional Geographer,* Vol. 31, pp. 400-409.

Joyce, Armond T., 1978. *Procedures for Gathering Ground Truth Information for a Supervised Approach to a Computer-Implemented Land Cover Classification of Landsat-Acquired Multispectral Scanner Data.* NASA Reference Publication 1015, Lyndon B. Johnson Space Center, Houston, TX, 43 pp.

Katz, Y. H., and W. L. Doyle, 1964. *Automatic Pattern Recognition of Meteorological Satellite Cloud Photography,* Memorandum RM-3412-NASA, The Rand Corporation, Santa Monica, CA 89 pp.

Mead, Roy A., and Merle P. Meyer, 1977. Landsat Digital Data Application to Forest Vegetation and Land Use Classification in Minnesota, *Machine Processing of Remotely Sensed Data Symposium,* IEEE, Piscataway, NJ, pp. 270-280.

Moellering, H., and W. Tobler, 1972. Geographical Variances, *Geographical Analysis,* Vol. 4, pp. 34-50.

Morrison, Donald F., 1976. *Multivariate Statistical Methods,* 2nd. ed., McGraw-Hill, NY, 415 pp.

Sayn-Wittgenstein, L., 1970. Patterns of Spatial Variation in Forests and Other Natural Populations, *Pattern Recognition,* Vol. 2, pp. 245-253.

Slater, Philip N., 1979. A Re-examination of the Landsat MSS, *Photogrammetric Engineering and Remote Sensing,* Vol. 45, pp. 1479-1485.

Todd, W. J., D. G. Gehring, and J. F. Haman, 1980. Landsat Wildland Mapping Accuracy, *Photogrammetric Engineering and Remote Sensing,* Vol. 46, pp. 509-520.

Tubbs, J. D., and W. A. Coberly, 1978. Spacial Correlation and its Effect Upon Classification Results in Landsat. *Proceedings of the Twelfth International Symposium on Remote Sensing of Environment,* Environmental Research Institute of Michigan Ann Arbor, pp. 775-781.

Turner, B. J., D. N. Applegate, and B. F. Merembeck, 1978. *Satellite and Aircraft Multispectral Scanner Digital Data User Manual.* Office for Remote Sensing of Earth Resources, University Park, PA, 405 pp.

(Received 29 May 1980; revised and accepted 20 September 1980)