George H. Rosenfield
U.S. Geological Survey
Reston, VA 22092

Analysis of Variance of Thematic Mapping Experiment Data

A weighted analysis of variance adjustment rigorously accommodates the different numbers of sample points that fall within each of the various thematic categories.

Abstract: *Thematic mapping experiments are conducted to evaluate variables such as different scales, images, algorithms, etc., that affect classification by thematic categories. The results of such experiments aid in developing specifications and techniques for producing thematic maps. A statistical technique which is useful in analyzing the results of thematic mapping experiments is analysis of variance, which operates simultaneously on measurements made from the variables under study.*

The data acquired for analyzing the accuracy of thematic mapping methods usually consist of the population of agreements and disagreements between the classifications based on field observations and the classifications interpreted from remotely sensed data, such as aerial photographs. These data are assumed to be binomially distributed. A weighted analysis of variance adjustment rigorously accommodates the different numbers of sample points that fall within each of the various thematic categories. Multiple range tests are a posteriori tests applied to population means found to be significantly different in the analysis of variance table.

As an example of the methodology, data from an experiment using three scales of land-use and land-cover mapping have been analyzed. The binomial proportions of correct interpretations have been analyzed untransformed and transformed by both the arcsine and the logit transformations. A weighted analysis of variance adjustment has been used. There is evidence of a significant difference among the three scales of mapping (1:24,000, 1:100,000, and 1:250,000) using the transformed data. Multiple range tests showed that all three scales are different for the arcsine transformed data.

A rigorous method of data analysis for thematic mapping experiments is thus tested and illustrated for the user community.

Introduction

Thematic mapping experiments are conducted to evaluate different variables that operate simultaneously and affect classification into thematic categories. Such variables might include three scales of aerial photographs; two types of images; several algorithms or equipment of digital classification; different physiographic regions; etc. Data obtained from thematic mapping experiments can be evaluated by the analysis of variance method, which has been defined by Scheffe (1959, p. 3) as a "statistical technique for analyzing measurements depending on several kinds of effects operating simultaneously, to decide which kinds of effects are important and to estimate the effects." For example, analysis of variance can simultaneously compare the classification accuracy of different categories for three scales of land-use and land-cover mapping, and can determine if a

Photogrammetric Engineering and Remote Sensing,
Vol. 47, No. 12, December 1981, pp. 1685-1692.

0099-1112/81/4712-1685$02.50/0

significant difference exists between the three scales used for that purpose.

BACKGROUND

Analysis of variance has been used very little in thematic mapping experimentation and in remote sensing studies and applications. Fitzpatrick-Lins (1978) describes an experiment in land-use and land-cover classification comparing the use and results of different scales of land-use and land-cover mapping using an application of analysis of variance. Landgrebe (1976) documented the methodology of analysis of variance in the literature on remote sensing. Bibliographic searches for other such documentation in the literature have disclosed none. Recent (as late as 1981) papers based on research in remote sensing show that analysis of variance is not being used to compare the effects of different factors under study. Review of the concepts of analysis of variance has indicated that research is required to investigate the use of this type of binomially distributed thematic mapping data in analysis of variance. It is hoped that this paper will answer certain questions regarding this type of data and place the data analysis on a sound statistical basis.

BINOMIAL DISTRIBUTION

A thematic map is usually divided into polygons, according to the categories of the theme (e.g., urban, agriculture, woodland, water, etc.) as determined from remotely sensed data or field mapping. A number of test points are selected in the polygons of each category. The interpreted category at the test points is compared with what is known from investigations on the ground. The data thus acquired results in a set of agreements and disagreements between the categories determined by ground investigation and the categories determined by interpretation or digital classification of remotely sensed data. The probability density function for the binomial distribution then depicts the probability of an exact number of successful classifications in the sample.

PURPOSE AND SCOPE

The purpose of this study was threefold: (1) to determine and document for the remote sensing researcher the minimum statistical methodology needed to perform and interpret an analysis of variance experiment; (2) to determine if there is any significant difference in the three scales of land-use and land-cover mapping, and if so, which scales are different from which others; and (3) to evaluate three recommended possible ways to preprocess the binomial proportions.

To illustrate the procedures, a set of data from the National Land Use and Land Cover Mapping Program of the U.S. Geological Survey was used. The set of proportions of correctly interpreted test data was subjected to three treatments of transformation as suggested by Snedecor and Cochran (1967, p. 493-495): untransformed, arcsine transformed, and logit transformed. The transformed data were tested for the three assumptions of analysis of variance: normality, homogeneity of variance, and additivity. A weighted analysis of variance was performed, and multiple range tests were applied to those computed means deemed significant.

From this research, a data analysis method and an operational computer program were developed to assist the analyst in considering the theoretical statistical concepts that are important for solving this type of problem. Copies of the program are available upon request.

SOURCE OF DATA

The data for use in this research was taken from an experiment conducted by Fitzpatrick-Lins (1978). The data are typical of those acquired in thematic mapping experiments and represent Level II land-use and land-cover mapping at three different scales (1:24,000, 1:100,000, and 1:250,000) prepared from 1:120,000-scale high-altitude aerial photographs.

Transformation and Weighting

Preprocessing of the binomial proportions p_i of each category, for entry into the analysis of variance was performed by the three different treatments as suggested by Snedecor and Cochran (1967, p. 493-495): (1) to regard the binomial proportions as already normally distributed variables (untransformed) having standard binomial variances and using weighted methods of analysis; (2) to transform the binomial proportions to equivalent angles using the arcsine $\sqrt{p}$ transformation, and treating the transformed angles as normally distributed with equivalent variances in a weighted analysis of variance; and (3) to transform the binomial proportions to its equivalent logit (Berkson, 1953) based on the logistic function with equivalent variances, also in a weighted analysis of variance.

NORMALLY DISTRIBUTED VARIABLES

Snedecor and Cochran (1967, p. 493) suggest regarding the binomial proportions p_i as normally distributed variables, where

$$p_i = (x/n)_i$$

in which x_i = the number of correctly classified test points in each category and n_i = the total number of test points in each category. The variance $\mathbf{v}(p_i)$ and weight $\mathbf{w}(p_i)$ are computed as

$$\mathbf{v}(p_i) = p_i(1 - p_i)/n_i,$$
$$\mathbf{w}(p_i) = 1/\mathbf{v}(p_i).$$

Table 1 contains the data of Fitzpatrick-Lins (1978) for the three scales of mapping: 1:24,000, 1:100,000, and 1:250,000.

ARCSINE TRANSFORMATION

Snedecor and Cochran (1967, p. 494) suggest transforming the binomial proportions p_i to equivalent angles θ_i and then treating the θ_i values as normally distributed variables. The angular transformation is performed by the expression

$$\theta_i = \text{arcsine}\ \sqrt{p_i}.$$

The angular transformation spreads out the proportions near 0 and 1 so as to increase their variance. If all of the error variance in the proportions is binomial, then the error variance in the angular scale is

$$\mathbf{v}(\theta_i) = 821/n_i.$$

Mosteller and Youtz (1961, p. 433) have tabulated transformations which had been developed by Freeman and Tukey (1950) to stabilize the variance of binomial counts for small values of n, $n \leq 50$. The Freeman-Tukey angular transformation is

$$\theta_i{}^* = 1/2 \left[\text{arcsine}\sqrt{\left(\frac{x_i}{n_i + 1}\right)} + \text{arcsine}\sqrt{\left(\frac{x_i + 1}{n_i + 1}\right)}\right]$$

When $\theta_i{}^*$ is measured in degrees, it has the asymptotic variance

$$\mathbf{v}(\theta_i{}^*) = 821/(n_i + 1/2)$$

with units of degrees squared. The asymptotic variance equation holds for a substantial range of p_i if n_i is not too small ($n_i \geq 5$). The value for the weight $\mathbf{w}(\theta_i)$ is computed as

$$\mathbf{w}(\theta_i) = 1/\mathbf{v}(\theta_i) \text{ or } 1/\mathbf{v}(\theta_i{}^*).$$

LOGIT TRANSFORMATION

Snedecor and Cochran (1967, p. 494) also suggest transforming the binomial proportion p_i to its equivalent logit Y_i where

$$Y_i = \ln(p_i/(1 - p_i)),$$

and then treating the Y_i values as normally distributed variables. The estimated variance of Y_i is approximately

$$\mathbf{v}(Y_i) = 1/(n_i p_i(1 - p_i))$$

with equivalent weight

$$\mathbf{w}(Y_i) = 1/\mathbf{v}(Y_i).$$

According to Snedecor and Cochran (1967, p. 494), if the binomial proportions p_i cover a wide range of values, then the assumption of additivity would more likely be met on the logit scale than the original p scale. In addition, the logit transformation, like the arcsine transformation, spreads out the proportions near 0 and 1, so that the scale extends from $-\infty$ to $+\infty$. On page 497 they further define logit functions for small samples (based on research by Gart and Zweifel (1967)) and state that the small sample logit is calculated as

$$Y_i = \ln((x_i + 1/2)/(n_i - x_i + 1/2))$$

with the variance

$$\mathbf{v}(Y_i) = (n_i + 1)/(x_i + 1/2)(n_i - x_i + 1/2).$$

Small samples are defined as those with values of x_i and $(n_i - x_i)$ being equal to or less than 30.

The weight value $\mathbf{w}(Y_i)$ is computed as

$$\mathbf{w}(Y_i) = 1/\mathbf{v}(Y_i).$$

TABLE 1. CATEGORIES OF LEVEL II LAND-USE AND LAND-COVER CLASSIFICATION, WITH ASSOCIATED TOTAL NUMBER OF TEST POINTS (n) IN EACH CATEGORY, AND THE NUMBER OF CORRECTLY CLASSIFIED TEST POINTS (x) IN EACH CATEGORY GIVEN FOR EACH OF THE THREE SCALES OF MAPPING

Level II Category	n[1]	1:24,000 x[1]	1:100,000 x[1]	1:250,000 x[1]
11	51	38	38	33
12	5	2	2	2
13	2	1	1	0
14	1	0	0	0
15	1	1	0	0
16	6	5	6	3
19	5	2	2	2
21	214	188	179	160
22	2	0	0	0
23	1	0	0	0
41	316	287	253	252
42	26	5	1	3
51	2	0	0	0
53	2	1	2	0
54	89	88	80	74
61	36	26	24	26
72	1	1	0	0

[1] Data for n and x from Fitzpatrick-Lins (1978).

PROCESSING OF DATA FOR ANALYSIS OF VARIANCE

Weighted analysis of variance in a two-way classification was performed on each of the three sets of data given above. These are the binomial proportions p_i treated as normally distributed variables (not transformed), and transformed to arcsines and logits for the three scales of mapping.

The weighted analysis of variance was performed using the General Linear Model (GLM) procedure of the Statistical Analysis System (SAS) computer software (Helwig, 1977, p. 33). The operational computer program has been prepared in the language of the SAS computer software to perform all the computation procedures needed for the study and explained in this paper. Copies of the computer program are available upon request.

The general method for analyzing a two-way classification is by arranging the data in the form of a row by column ($R \times C$) matrix. The columns of the matrix represent the different factors, such as each of the three scales of mapping. The rows of the matrix represent the various categories of land use and land cover. The values in the cells of the matrix are the binomial proportions either treated as normally distributed observations, as arcsine transformed observations, or as logit transformed observations. A different matrix is established for each treatment of data, and a different analysis of variance is performed.

Assumptions for Analysis of Variance

Eisenhart (1947) enumerates the several assumptions underlying the analysis of variance and points out the practical importance of each. These assumptions include (1) Normality—that the errors of the observations are jointly distributed in a multivariate normal distribution; (2) Equal variances and zero correlations—that the random variables are homoscedastic (homogeneity of variance) and mutually uncorrelated; thus, they have a common variance and all covariances among them are zero; and (3) Additivity—that the true mean values of observations are simple additive functions of the row effects and the column effects. Cochran (1947) describes the adverse effects on the analysis of variance when the required assumptions are not met and gives advice on how to detect failure of the assumptions and how to avoid the more serious consequences.

NORMALITY

The KSL (Kolmogorov-Smirnov, Lilliefors) test for normality is included in the SAS computer software (Helwig, 1977, p. 101). Lilliefors (1967) has revised the standard Kolmogorov-Smirnov test for use when the mean and variance are estimated from the sample.

The KSL procedure was performed on the residuals from the adjustment for each of the three treatments of input data as normally distributed observations (ND), arcsine transformed (AT), and logit transformed (LT) for the weighted adjustment of the three scales of mapping. The results are given in Table 2.

Note that the only significant results are the departures from normality for the residuals of the arcsine and logit transformed data. All of the other results are not significant.

HOMOGENEITY OF VARIANCE

Several tests can be used to determine homogeneity of variances. Sokal and Rohlf (1969, p. 369-376) recommend Bartlett's test for more than two groups of data. However, Bartlett's test is sensitive to non-normality in the data. They also recommend the F-max test to corroborate the findings of Bartlett's test. Anderson and McLean (1974, p. 22) suggest the Burr-Foster (1972) Q-test which is not sensitive to non-normality as is the Bartlett test.

Table 2. Results of KSL Procedure Test for Normality Performed on Residuals of the Weighted Analysis of Variance for the Three Treatments of Input Data: as Normally Distributed Observations (ND), Arcsine Transformed (AT), and Logit Transformed (LT)

D-max		
n	*D-max*	*P*
ND 29	0.0608	>0.20
AT 51	0.1875*	0.01*
LT 51	0.1699*	0.01*
Skewness		
n	*g1*	*P*
ND 29	−0.548	0.206
AT 51	0.830*	0.013*
LT 51	1.073*	0.001*
Kurtosis		
n	*g2*	*P*
ND 29	0.728	0.389
AT 51	2.297*	0.000*
LT 51	2.060*	0.002*

* significant at 0.05 level
n = number of non-zero data cells
g1 = sample estimate of coefficient of skewness
g2 = sample estimate of coefficient of kurtosis
P = probability level
D-max = test criterion for normality

The test for homogeneity of variance was performed on column variances computed in the R by C matrix. The three tests for homogeneity of variance—Bartlett's test, F-max test, and the Burr-Foster Q-test—were each performed on the sets of data from each of the three treatments of the binomial proportions. For all cases, there was no reason to reject the hypothesis for homogeneity of variance at the 0.05 level.

ADDITIVITY

The Tukey (1949) one degree of freedom test for non-additivity computes the sum of squares for non-additivity with one degree of freedom. The residual sum of squares and residual degrees of freedom must be determined after the program for the analysis of variance table. The F statistic is then computed for the effects of non-additivity to residuals.

Tukey's test was applied to the same three sets of data that were used for the homogeneity of variance tests. The sum of the squares for nonadditivity (SSNA) was computed and compared with the analysis of variance table. The F value for nonadditivity was computed by the mean square for nonadditivity divided by the residual mean square. In all instances there was no reason to reject the hypothesis of additivity at the 0.05 level.

DISCUSSION OF THE ASSUMPTIONS

The analysis of variance is a robust procedure and is not highly sensitive to failures in the assumptions, especially to that of normality. According to Cochran (1947, p. 24), however, failure in the assumption of normality leads to an understatement of the probability for the tests of significance. Thus, for tests with marginal significance, too many significant results are indicated.

The tests for homogeneity of variance and for additivity indicate that both those assumptions were met by the data from all three treatments of the binomial proportions: as normally distributed observations and by the arcsine and logit transformations.

The normality of the residuals is a measure of the utility of the transformations and also of how well the model fits the data. One might expect that the data which most closely meets the assumptions of the analysis of variance would be better fitted by the model. Table 2 gives the results of the normality tests on the residuals. For the three scales of mapping, the residuals meeting the normality criteria are those of the normally distributed observations from the weighted adjustment.

Thus, it might appear at first that entering the binomial proportions into the analysis of variance as normally distributed variables does not seem to cause any harm with a weighted adjustment. The weighted analysis of variance must be used if there is great disparity among the sample sizes of the various categories. However, when weights are computed for the nontransformed data, those proportions which are either zero or one each have a weight of zero. As such, these proportions have no influence on the analysis of variance except to add degrees of freedom. This amounts to throwing away the data, in a very biased manner, at both ends of the scale and leaving the middle portion to appear to be normally distributed. Furthermore, the underlying probability distribution of the observations is binomial, not normal, and a biased adjustment would result if the data at the ends of the scale were not removed by the weighting procedure.

Tests of Significance and Multiple Range Tests

Tests of significance are used to help determine if the hypothesis being tested is most probably true or false. A quantity called a test criterion is computed. If this quantity is larger than would have been expected by chance alone, then the test criterion is rejected, and the results are said to be statistically significant. For example, if the three scales of mapping are determined to be significant, then the hypothesis that they are the same is rejected.

The following types of test criteria have been proposed in the literature for use in analysis of variance for remotely sensed data.

ERROR MEAN SQUARE

Snedecor and Cochran (1967, p. 313) state that the error mean square in an analysis of variance is an unbiased estimate of the variance. Sokol and Rohlf (1969, p. 194-197) explain that the variances in analysis of variance are called mean squares because they do not estimate a population variance and that the error mean square measures the average dispersion of the items in each group around the group means but, if the groups are random samples from a homogeneous population, the error mean square should estimate the variance.

The error mean square resulting from the analysis of variance was used in the tests of significance for the means of both variables: land-use and land-cover categories and scales. For all three cases the following results are significant: the differences between the land-use and land-cover categories, and also between the three scales of mapping for each of the three treatments of binomial proportions.

ARCSINE ERROR MEAN SQUARE

Snedecor and Cochran (1967, p. 494) recommend use of the arcsine error mean square for analysis of variance with arcsine transformed data. Landgrebe (1976) applies the arcsine error mean square to the analysis of variance with infinite degrees of freedom.

Cochran (1940) discusses analysis of variance when experimental errors follow the binomial laws. He uses the chi-square test to decide if the whole of the experimental error variation is of the binomial type. He computes chi-square as the error sum of squares divided by the arcsine error mean square, with specific degrees of freedom. If the hypothesis of binomial variation is rejected, the usual analysis of variance tests should be made with the actual error mean square.

HARMONIC MEAN SQUARE ERROR

In their discussion of the method for the unweighted analysis of cell means, Snedecor and Cochran (1967, p. 475-477) treat the individual cell means as if they were all based on the same number of observations, but the error mean square used in the analysis of variance is computed from the harmonic mean of the individual cell numbers. The harmonic mean square error is not applicable in a weighted analysis of variance adjustment.

The use of the harmonic mean value with the arcsine transformed data is also a poor substitute for a weighted analysis of variance. Although this technique tries to accommodate the unequal cell frequencies in the tests of significance, the adjustment is biased, and the meaning of the results is questionable. In addition, this technique can not be used with nontransformed or logit transformed data because the variance in these cases is not constant as it is with the arcsine transformation.

MULTIPLE RANGE TESTS

Multiple range tests are described by Sokal and Rohlf (1969, p. 226-227 and p. 235-246) as a posteriori tests performed after the analysis of variance to distinguish differences between means or groups of means. They are performed only if the analysis of variance is significant. When an F-test determines that the differences between one or more of treatment means are significant, it does not specify which, if any, are not. The multiple range tests indicate which differences are significant, and which are not. The Duncan multiple range test (Steel and Torrie, 1960, p. 107-109; and Duncan, D.B., 1955) of the SAS computer software (Barr and others, 1976, p. 108) was used to perform an a posteriori comparison among the means because the analysis of variance was significant.

The results of Duncan's multiple range tests for the significant means from the analyses of variance are given in Table 3. Significant results were shown for the factors representing the three scales of mapping for the weighted analysis of variance. Duncan's multiple range test shows the groupings of similarity by scale to be: (1) the nontransformed data showed no differences between the scales; (2) the arcsine transformed data showed each scale to be different; and (3) the logit transformed data showed the grouping 1:24,000 and 1:100,000-scales, and 1:100,000 and 1:250,000-scales.

Significant results were also shown for the factors representing the Level II land-use and land-cover categories. In the weighted adjustment, the Duncan's multiple range test grouping was similar to the groupings by scale. The nontransformed data showed no differences between the categories, the arcsine transformed data showed no overlapping among the groupings (each grouping of categories was independent), and the logit transformed data showed intermixing among the categories.

DISCUSSION OF TESTS OF SIGNIFICANCE AND MULTIPLE RANGE TESTS

The significance of the effects among the three scales of mapping for the weighted analysis of variance indicate (1) that the accuracy of visual classification for Level II land-use and land-cover categories is different for each of the three scales, regardless of the treatment of the binomial proportions; and (2) that each Level II land-use and land-cover category cannot be visually classified equally well from each of the three scales of mapping. Duncan's multiple range tests indicate that all the categories and all the scales are grouped together for the treatment as normally distributed variables; there is no overlapping of the groups for the arcsine transformed data; and there is intermingling among the categories and among the scales for the logit transformed data. Thus, using the arcsine transformation, the three scales of

TABLE 3. RESULTS OF DUNCAN'S MULTIPLE RANGE TESTS FOR THE SIGNIFICANT MEANS FROM THE WEIGHTED ANALYSIS OF VARIANCE FOR THE THREE TREATMENTS OF THE BINOMIAL PROPORTIONS: AS NORMALLY DISTRIBUTED VARIABLES (ND), ARCSINE TRANSFORMED (AT), AND LOGIT TRANSFORMED (LT)

ND	*AT*	*LT*
categories	Level II categories	Level II categories
A – all categories	A – 15	A – 15
	B – 18, 8	A B – 11,8
	C – 6	A B C – 6
	D – 1, 16	B D C – 1, 16
	E – 14	E D C – 14
	F – 2, 7	E D F – 2, 7, 5, 17, 3
	G – 5, 17	E G F – 4, 10
	H – 3	G F – 9, 13
	I – 4, 10	G – 12
	J – 12	
	K – 9, 13	
scales	*scales*	*scales*
A – all scales	A – 1:24,000	A – 1:24,000
	B – 1:100,000	A B – 1:100,000
	C – 1:250,000	B – 1:250,000

Note: (1) The numbers shown are codes which represent the various categories of land-use and land-cover classification. (2) The letters shown represent similarities from Duncan's multiple range test. Categories or scales with the same letter designation have been found to be similar.

mapping lead to individually different accuracies in the visual interpretation of land-use and land-cover categories. But using the logit transformation, the accuracy of visual interpretation of the categories for the 1:24,000 and 1:100,000 scales are similar, and the accuracy for the 1:100,000 and 1:250,000 scales are also similar. If the desire is to be rigorous in statistical application in order to draw valid conclusions, then the data should not be treated as normally distributed variables for a weighted adjustment. A choice then must be made between accepting the results of either the arcsine or the logit transformed data. The results are valid in each case. Only the geographic analyst can make the final choice.

Analysis of Residuals

Residuals from the adjustment are analyzed to determine if the data set has met the assumption of normality, and if the model has adequately described the data. In addition, analysis of the residuals will help indicate the direction and magnitude (and possibly the source) of systematic errors which may remain in the data. The KSL procedures of the SAS computer software is used to test for normality. Searle (1971, p. 129-130) discusses examining residuals. He only gives hints to some of the available methods of analyses and references to more complete discussions. He does note two important properties of the residuals: first, that they sum to zero and, second, that their sum of squares is the error sum of squares in the analysis of variance table. Searle suggests plotting the residuals to see if they appear normally distributed. Searle also suggests plotting the residuals against the predicted values and against the independent variables. These kinds of plots might suggest that the error terms may not have constant variance or that additional terms are needed in the model. Helwig (1978, p. 63-65) indicates examples of how the SAS procedures can be used to plot predicted values and residuals against the independent variables.

General

This report covers some of the theoretical concepts that must be considered in analyzing experimental data of this type. Regardless of what the practical results might be in any one particular application, those procedures which are based on the soundest theoretical foundation are the ones that should be followed. If the data are known to have been derived from a probability distribution that is not normal, then some type of transformation may be desirable. The question of which transformation should be used cannot be simply answered. At the least, the results of the transformation selected should be tested against the assumptions necessary for analysis of variance. If there are large discrepancies in the numbers of test points for each category, then a weighted analysis should be used. If, for some valid reason, a weighted analysis is not used and if the arcsine transformation is selected, then the arcsine error mean square using the harmonic mean value should be considered and tested in the F-test for significance. The chi-square test must be applied to the arcsine error mean square before it is selected over the error mean square from the adjustment. The theoretical advantages of the logit transformation and of the weighted adjustment should be considered and tested. Multiple range tests should be applied to separate differences between the means of factors found to be significant. In addition, the difference between the use of the arcsine or logit transformation for the binomial proportions is inconclusive as a result of only one test.

This is not the end to studies of this type. This study uses only the diagonal elements of the classification error matrices from Fitzpatrick-Lins (1978). All thematic mapping experiments performed in the past have used only the diagonal elements for analysis. The next step would be to use the entire classification error matrices in such studies. An attempt is already being made in this direction as evidenced by the presentation of Congalton (1980). His approach uses discrete multivariate analysis (Bishop *et al.*, 1975). Another approach using the entire classification error matrices would be the techniques of multivariate analysis of variance. It is hoped that some experimenters will find time for both of these approaches.

References

Anderson, V. L., and R. A. McLean, 1974. *Design of experiments:* Marcel Dekker, Incorporated, New York.

Barr, A. J., and others, 1976. *A user's guide to SAS76:* SAS Institute Incorporated, Raleigh, North Carolina.

Berkson, J., 1953. A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function: *Journal of the American Statistical Association,* v. 48, pp. 565-599.

Bishop, Y. M. M., S. E. Feinberg, and P. W. Holland, 1975. *Discrete Multivariate Analysis: Theory and Practice:* Cambridge, Massachusetts, The MIT Press.

Burr, I. W., and L. A. Foster, 1972. *A test for equality of variances:* Department of Statistics Mimeo Series No. 282, Purdue University, Lafayette, Indiana.

Cochran, W. G., 1940. The analysis of variance when experimental errors follow the Poisson or binomial laws: *Annals of Mathematical Statistics,* v. 11, pp. 335-347.

———, 1947. Some consequences when the assumptions for the analysis of variance are not satisfied: *Biometrics,* v. 3, no. 1, pp. 22-38.

Congalton, R. G., 1980. *Statistical techniques for analysis of Landsat classification accuracy:* Paper presented at meeting of the American Society of Photogrammetry, St. Louis, Missouri, March 11.

Duncan, D. B., 1955. Multiple range and multiple F tests: *Biometrics*, v. 11, pp. 1-42.

Eisenhart, C., 1947. The assumptions underlying the analysis of variance: *Biometrics*, v. 3, no. 1, pp. 1-21.

Fitzpatrick-Lins, K., 1978. Accuracy and consistency comparisons of land use and land cover maps made from high-altitude photographs and Landsat multispectral imagery: *Journal of Research*, U.S. Geological Survey, v. 6, no. 1, pp. 23-40.

Freeman, M. F., and J. W. Tukey, 1950. Transformations related to the angular and the square root: *Annals of Mathematical Statistics*, v. 21, pp. 607-611.

Gart, J. J. and J. R. Zweifel, 1967. On the bias of various estimators of the logit and its variance with application to quantal bioassay: *Biometrics*, v. 54, nos. 1 and 2, pp. 181-187.

Helwig, J. T. [ed.], 1977. *SAS supplemental library user's guide:* SAS Institute Incorporated, Raleigh, North Carolina.

———, 1978. *SAS introductory guide:* SAS Institute, Incorporated, Raleigh, North Carolina.

Landgrebe, D. A., 1976. *Final report*, NASA contract NAS9-14016: Laboratory for applications of remote sensing, Purdue University, West Lafayette, Indiana 47096 (NAS9-14016, T-1039, MA-129TA).

Lilliefors, H. W., 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown: *Journal of the American Statistical Association*, v. 62, no. 318, pp. 399-402.

Mosteller, F., and C. Youtz, 1961. Tables of the Freeman-Tukey transformation for the binomial and Poisson distributions: *Biometrika*, v. 48, nos. 3 and 4, pp. 433-440.

Scheffe, H., 1959. *The analysis of variance:* John Wiley and Sons, Incorporated, New York.

Searle, S. R., 1971. *Linear models:* John Wiley and Sons, New York.

Snedecor, G. W., and W. G. Cochran, 1967. *Statistical methods:* Sixth Edition, The Iowa State University Press, Ames, Iowa.

Sokal, R. R., and J. Rohlf, 1969. *Biometry:* W. H. Freeman and Company, San Francisco, California.

Steel, R. G. D., and J. H. Torrie, 1960. *Principles and procedures of statistics:* McGraw-Hill Book Company, Incorporated, New York.

(Received 20 September 1980; revised and accepted 13 May 1981)

Forthcoming Articles

Youssef I. Abdel-Aziz, Nonphotogrammetric Plotters for Photogrammetric Applications.

Youssef I. Abdel-Aziz, Accuracy of the Normal Case of Close-Range Photogrammetry.

M. E. O. Ali and *A. J. Brandenberger,* Analytical Triangulation of Space Photography.

Herman Arp, Jean Claude Griesbach, and *Joseph P. Burns,* Mapping in Tropical Forests: A New Approach Using the Laser APR.

B. Brisco and *R. Protz,* Manual and Automatic Crop Identification with Airborne Radar Imagery.

Paul J. Curran, Multispectral Photographic Remote Sensing of Green Vegetation Biomass and Productivity.

W. Frobin and *E. Hierholzer,* Calibration and Model Reconstruction in Analytical Close-Range Stereophotogrammetry. Part I: Mathematical Fundamentals.

W. Frobin and *E. Hierholzer,* Calibration and Model Reconstruction in Analytical Close-Range Stereophotogrammetry. Part II: Special Evaluation Procedures for Rasterstereography and Moire Topography.

A. Grün, The Accuracy Potential of the Modern Bundle Block Adjustment in Aerial Photogrammetry.

Charles F. Hutchinson, Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement.

S. Kondo, R. Ohba, and *K. Murata,* The Location of Three-Dimensional Linear Objects by Using Multiple Projections.

Thomas M. Lillesand, Douglas E. Meisner, Anne LaMois Downs, and *Richard L. Deuell,* Use of GOES and TIROS/NOAA Satellite Data for Snow-Cover Mapping.

B. K. Quirk and *F. L. Scarpace,* A Comparison between Aerial Photography and Landsat for Computer Land-Cover Mapping.

G. H. Rosenfield, K. Fitzpatrick-Lins, and *H. S. Ling,* Sampling for Thematic Map Accuracy Testing.

Horst H. Schöler, Production of Photogrammetric Stereopairs by Beam Division in Single-Lens Cameras.

H. Schreier, L. C. Goodfellow, and *L. M. Lavkulich,* The Use of Digital Multi-Date Landsat Imagery in Terrain Classification.

Kenneth Watson, Susanne Hummer-Miller, and *Don L. Sawatzky,* Registration of Heat Capacity Mapping Mission Day and Night Images.