NASA

Technical Memorandum 83871

The Influence of Autocorrelation in Signature Extraction - An Example From a Geobotanical Investigation of Cotter Basin, Mt

M.L. Labovitz, and E.J. Masuoka

DECEMBER 1981

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

THE INFLUENCE OF AUTOCORRELATION IN SIGNATURE EXTRACTION – AN EXAMPLE FROM A GEOBOTANICAL INVESTIGATION OF COTTER BASIN, MT

M. L. Labovitz and E. J. Masuoka
Code 922
Geophysics Branch
NASA/Goddard Space Flight Center
Greenbelt, Maryland 20771

ABSTRACT

The presence of positive serial correlation (autocorrelation) in remotely sensed data results in an underestimate of the variance-covariance matrix when calculated using contiguous pixels. This underestimate produces an inflation in F statistics. For a set of Thematic Mapper Simulator data (TMS), used to test the ability to discriminate a known geobotanical anomaly from its background, the inflation in F statistics related to serial correlation is between 7 and 70 times. This means that significance tests of means of the spectral bands initially appear to suggest that the anomalous site is very different in spectral reflectance and emitance from its background sites. However, this difference often disappears and is always dramatically reduced when compared to frequency distributions of test statistics produced by the comparison of simulated training sets possessing equal means, but which are composed of autocorrelated observations.

THE INFLUENCE OF AUTOCORRELATION IN SIGNATURE EXTRACTION – AN EXAMPLE FROM A GEOBOTANICAL INVESTIGATION OF COTTER BASIN, MT

INTRODUCTION

Recently researchers have demonstrated that remotely sensed data from a number of sensor systems are autocorrelated. Such conclusions have been documented for Landsat, (Coberly, 1973; Craig, 1976, 1979; Tubbs and Coberly, 1978; Tubbs, 1979; Craig and Labovitz, 1980), Thematic Mapper Simulator (TMS) (Labovitz et al., 1981) and the Viking Lander (Craig et al., 1981). Along with these findings, researchers have been studying what effects the violation of the assumption of independence has upon maximum likelihood, Bayesian classifiers and parameter estimation. This paper explores the influence of autocorrelation upon statistics in a hypothesis testing situation.

BACKGROUND AND PREVIOUS RESEARCH

Autocorrelation is a property defined for sequences, $X_1, X_2, \cdots, X_{i-1}, X_i, X_{i+1}, \cdots, Xn$ and is a measure of the randomness of the sequence. In a remote sensing context the sequence is a scan line and X_i is the digital number (DN) for the ith pixel. The autocorrelation in a scan line is measured for specific distances between pixels. For example, the autocorrelation of adjacent pixels, known as the lag 1 autocorrelation, is estimated by

$$\sum_{i=1}^{n-1} \frac{(X_i - \bar{X})(X_{i+1} - \bar{X})}{S_X^2 (n-2)}$$

where: $\bar{X}$ is the sample mean of the scan line;

S_X^2 is the sample variance of the scan line;

n is the number of pixels in the scan line.

In general, the autocorrelation can be measured for lag k (k = 1, 2, . . ., n − 1), and is estimated by

$$\sum_{i=1}^{n-k} \frac{(X_i - \bar{X})(X_{i+k} - \bar{X})}{S_X^2 (n - (k+1))}.$$

The function which maps a lag number to the degree of autocorrelation is known as the autocorrelation function (acf). The acf can be thought of as taking a sequence, sliding it incrementally along itself and calculating a correlation coefficient at every increment. The autocorrelation is, therefore, proportional to paired deviations from the mean at a specific spacing. If the acf is not zero for all lags, then there is a nonrandom pattern in the deviations; and this pattern can be expressed by a function which relates the DN for the ith pixel to DN's for the previous pixels. In this situation we speak of the system as having a memory and say that the data is autocorrelated.

Basu and Odell (1974) found that use of autocorrelated data produced biased estimates of the variance-covariance matrix. This conclusion was based upon the assumption of a equicorrelation dependence structure, i.e. the correlation between X_i and $X_j = \rho$, for all $i \neq j$. Basu et al. (1976) showed that if the simple equicorrelation had a positive coefficient, i.e. $\rho > 0$, then the bias in estimation of the variance-covariance matrix, Σ, resulted in confidence limits, about the mean, for example, which were too narrow. Intuitively, the assumption of equicorrelation is not very appealing because it implies that the correlation between pixels in a scene is the same no matter what the value of the spacing or distance between the pixels. However, the general direction of the bias appears to be correct because Cliff and Ord (1981) have described a related bias in the distribution of the t-statistics (student's t) when calculated on the estimated means from two populations, $\bar{X}_1$ and $\bar{X}_2$, when X_1 and X_2 are autocorrelated. Positive autocorrelation results in the underestimation of the standard error of the mean and consequent inflation of the t value. Gastwirth and Rubin (1971) found the presence of even slight serial correlation (autocorrelation) had a strong influence on the confidence levels of the distributions of not only the mean but also the two one-sample nonparametric tests, the sign test and the Wilcoxon test. Tubbs and Coberly (1978) and Tubbs (1979) proposed an autoregression model of order 1 AR(1) as being more realistic than equicorrelation for the dependence structure among the pixels from a Landsat scan line. The authors then went on to investigate the influence of the AR(1) structure upon estimation of frequency distribution parameters and the Bayesian and maximum likelihood classification procedures.

The authors found that the estimate of the variance-covariance matrix was no longer unbiased, that is the commonly used estimate, S, of Σ, resulted in an underestimate if calculated from adjacent pixels which were positively autocorrelated. For the Bayesian classification procedure, the authors showed that the misclassification probabilities were unaffected only when the costs of misclassifying were equal and the prior probabilities were assumed to be equal. Further, the probability of misclassification errors from a maximum likelihood classifier were appropriate only when the training set sizes were equal. Clearly, there are circumstances, including the need to obtain unbiased estimates of other parameters such as type I and type II errors which would result in experimental situations where these requirements are not met. Tubbs (1980) extended the results of Basu et al. (1976) and examined the influence of an AR(1) dependence structure (instead of equicorrelation) upon the confidence regions of the mean and variance. The author found that when the observations possessed positive serial correlation, the assumption of stochastic independence resulted in confidence intervals that are too restrictive. This meant that a $100\,(1-\alpha)$ percent confidence region calculated under the assumption of random samples was $<100\,(1-\alpha)$ percent if the data were positively serially correlated.

Craig (1976, 1979) demonstrated for Landsat data that an autoregressive moving average of order one-one, an ARIMA (1, 0, 1) (Box and Jenkins, 1970) was a more appropriate model than the AR (1) model. (We will discuss the ARIMA (1, 0, 1) model in greater detail later.) Further, Craig (1979) suggested that a sampling scheme with a minimum of 10 pixels between samples be employed in order to produce independent observations. Craig and Labovitz (1980) suggested that the ARIMA (1, 0, 1) was a general model for Landsat data and that certain physical attributes of the scene were related to the degree of autocorrelation through systematic changes in the values of the coefficients of the ARIMA (1, 0, 1).

Craig (1981) showed that autocorrelated data, in comparison to independent observations tended to show smaller variability for short distances (contiguous pixels) and larger variation for longer distances. He also demonstrated that the inflation in the variation behaved as a function of

the coefficients in the ARIMA (1, 0, 1) model. Craig suggested three methods for compensating for autocorrelation in data – 1) filtering out the ARIMA (1, 0, 1) model; 2) resampling the data with an appropriate spacing between pixels; and 3) adjusting the classification algorithms.

Campbell (1981) used a sampling scheme to adjust his classification for the presence of autocorrelation. In this paper the author demonstrated effects of autocorrelation by comparing classification errors between classifications based on contiguous pixels and classifications based on training sets composed of randomly selected pixels. The data used was Landsat MSS data from near Richmond, VA. The results showed an almost universal improvement in classification accuracies under the random sampling alternative. This result held for data acquired during several portions of the year. Since the random sampling scheme used did not guarantee independent observations, it is possible that a further improvement might be achieved by a systematic sample with a suitable spacing.

OBJECTIVE OF STUDY

The objective of this paper is to examine what influence autocorrelation has upon the researcher's decision concerning the presence of distinguishable populations and his ability to assign explanatory power to sources of variation. When the researcher makes these decisions based upon the behavior of functions of samples (statistics) from the respective populations, the problem is recast as examination of the relationship between the degree of autocorrelation and the decision statistic.

This work arose out of ongoing research in the utility of remote sensing for geobotanical exploration. In this area of geobotany, the explorationist is trying to recognize differences in the spectrum of vegetation growing in soils containing anomalously high concentrations of stragetic metals (e.g. copper, lead, zinc, molybdenum, and chromium) as compared to vegetation growing in soils containing background concentrations of trace metals. This search is performed with the

hope that the soil is derived from a nearby ore body. If such a difference could be detected in vegetation spectra, explorationists could use this difference to "map" promising areas.

THE STUDY REGION – THE COTTER BASIN GEOBOTANICAL TEST SITE

The Cotter Basin geobotanical test site is located in Lewis and Clark County, Montana and can be found on the Stonewall Mountain Quadrangle (see Figure 1). The area is in the Northern Rocky Mountains Region and possesses considerable relief. The test site region has been studied previously by Collins et al. (1977) and Collins et al. (1980). The remotely sensed data used in their research were acquired by a prototype high resolution (500 channels from 0.400 μm to 1.0 μm) airborne spectroradiometer system (Chiu and Collins, 1978) during one overflight in 1976 and a second during 1978.

Mineralization at the Cotter Basin test site occurs in westerly striking shear zones in Precambrian sedimentary rocks (Collins et al., 1980). The mineralization is primarily sulfides of copper and lead with ancillary high concentrations of zinc, silver and gold. Soil concentrations along shear zones are as high as 10,000 parts per million (ppm) copper, 1,500 ppm lead and 400 ppm zinc (Collins et al., 1980).

Description of Data Used in the Present Study

The data used were acquired using the Landsat Thematic Mapper Simulator (TMS)-NS001-(Richard et al., 1978) mounted aboard a NASA C130 aircraft. The data came from flight line number 3, flown on August 29, 1979 commencing at 1830 GMT (12:30 p.m. MDT). The flight line was flown in a south-to-north direction at about 3.0 km (10,000 feet) above ground level. Since the instantaneous field-of-view of the NS001 is 2.5 milliradians, the pixels are approximately 7.5 m (25 feet) on a side at nadir. Table 1 gives the spectral bands for which data were collected.

Training Site Selection

For the purposes of this study two classes were defined: anomalous and background. The training site for the anomalous class was located so that it fell within the intersection of the soil and

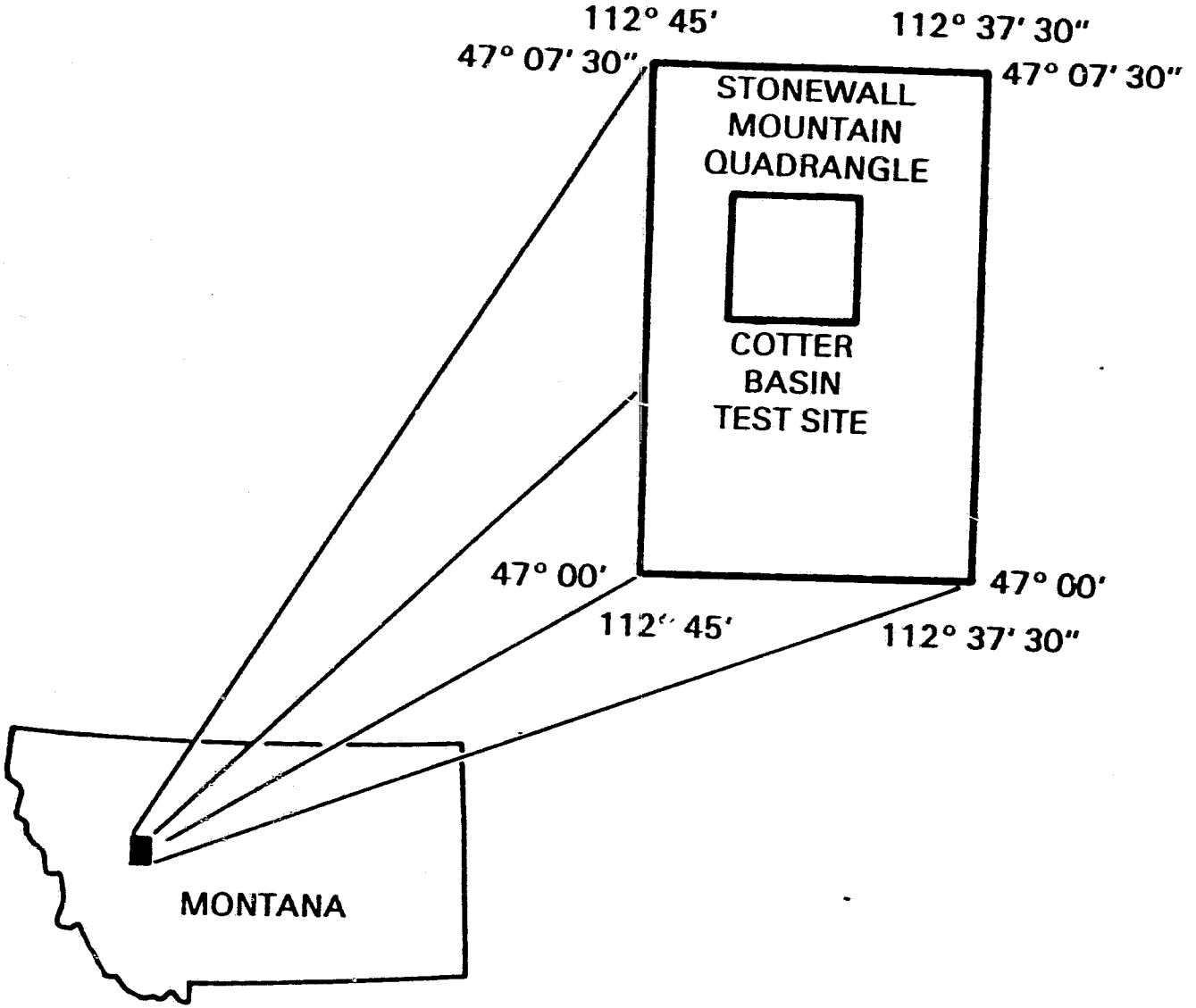


Figure 1. Location of the Cotter Basin geobotanical test site.

Table 1
TMS (NS001) Spectral Bands

Channel Number	Band Width (Micrometers)
1	0.42 – 0.52
2	0.52 – 0.60
3	0.63 – 0.69
4	0.76 – 0.90
5	1.00 – 1.30
6	1.55 – 1.75
7	2.08 – 2.35
8	10.40 – 12.50

remote sensing anomalies defined by Collins et al. (1980). The location of the anomalous training site was also constrained by the selection of two background sites such that all three sites were located on the same slope and aspect and at the same angle off-nadir with respect to the NS001 flight line. The three training sites were composed of squares of contiguous pixels, 10 pixels on a side (100 pixels per training site).

Initial Comparisons of Training Sites

Table 2 contains the sample means by training site for each of the eight bands. The results of tests of the hypotheses of mean vector equality for combinations of the anomalous and background sites is given in Table 3. We clearly reject the null hypothesis not only for comparisons of anomalous versus background sites, but also for the comparison of the two backgrounds. Thus all three sites are very significantly different from one another. Further examination of the differences in mean reflectance was accomplished by looking at univariate (by band) comparisons. Table 4 gives the results from 32 univariate tests (8 bands X 4 combinations of training sites). The F ratios formed were highly significant ($\alpha = 0.006$) for 30 of the 32 tests, indicating that the significance found in the multivariate tests was not attributable to some systematic subset of the spectral bands.

Thus, we find significant differences among all three sites for virtually all spectral bands, that is both for bands containing vegetative information as well as for bands which are insensitive to the presence of vegetation. Therefore, geobotanical sources of variation are unlikely to account for such pervasive significance. In looking for explanatory effects, we start with the testing procedure.

Table 2
Sample means for the eight TMS bands for each test site. Each mean is computed from 100 contiguous pixels.

Band	Center λ (μm)	Site		
		A	B_1	B_2
1	0.48	88.5	116.8	115.6
2	0.56	57.1	68.4	92.7
3	0.66	28.8	32.4	51.3
4	0.83	66.2	68.3	108.7
5	1.15	66.3	73.9	124.6
6	1.65	32.9	38.0	77.5
7	2.2	52.0	62.5	87.0
8	11.4	35.9	47.4	77.2

A – anomalous site
B_1 – background site 1
B_2 – background site 2

Table 3
Results from tests of null hypotheses of mean vector equality. Λ_{obs} is the observed Wilks' Lambda value, G is the related value of the F statistic, $F_{0.01}$ is the preselect critical value of $\alpha = 0.01$, and $P(F > G)$ is the probability that an F random variable with the appropriate degrees of freedom would be greater than G under the null hypotheses.

Comparison	Λ_{obs}	G	$F_{0.01}$	$P(F > G)$
A vs B_1	0.23	80.89	~ 2.51	3.03×10^{-58}
A vs B_2	0.13	165.37	~ 2.51	1.42×10^{-81}
B_1 vs B_2	0.27	65.06	~ 2.51	1.75×10^{-50}
All	0.08	92.76	~ 2.04	6.79×10^{-78}

A – anomalous site
B_1 – background site 1
B_2 – background site 2

Table 4
Results from univariate tests of mean differences in reflectance for combinations of training sites.

	Contrast[1]			
	All	A vs B_1	A vs B_2	B_1 vs B_2
Band 1	161.23 *	259.52 *	420.33 *	0.36 NS
Band 2	308.11 *	87.06 *	658.67 *	193.54 *
Band 3	278.58 *	106.62 *	338.96 *	233.39 *
Band 4	162.77 *	1.78 NS	192.55 *	187.64 *
Band 5	211.33 *	18.53 *	265.20 *	204.63 *
Band 6	265.54 *	42.88 *	312.72 *	240.00 *
Band 7	488.63 *	198.71 *	745.78 *	340.56 *
Band 8	406.02 *	297.73 *	559.91 *	275.51 *

[1]Test statistic is an F-ratio with, degrees of freedom for: two sites – 1, 198; three sites – 2, 297.

*Test significant at $\alpha = 0.05/8 = 0.006$ (Bonferroni adjustment for each contrast, (Neter and Wasserman, 1974)).

NS – Test not significant at $\alpha = 0.006$.

A – anomalous site

B_1 – background site 1

B_2 – background site 2

HYPOTHESIS TESTING PROCEDURE

The test procedure used is called multivariate analysis of variance (MANOVA) and procedes in the following manner:

1. Postulate a model

$$Y_{tj} = \mu_j + \varepsilon_{tj} \qquad (1)$$

$$(t = 1, 2, \ldots, 100;\ j = 1, 2, 3)$$

where

$\underline{Y}_{tj}$ is an 8 x 1 random vector of spectral observations of the t^{th} pixel from the j^{th} site, distributed as a multivariate normal vector deviate with mean equal to $\underline{\mu}_j$ and the covariance given in an 8 x 8 matrix Σ_j whose diagonal elements are the variances of the spectral bands and whose off diagonal elements are covariances. This is symbolized as

$$\underline{Y}_{tj} \sim MN(\mu_j, \Sigma_j); \qquad (2)$$

$\underline{\mu}_j$ is an 8 x 1 vector of means for the j^{th} site;

$\underline{\varepsilon}_{tj}$ is an 8 x 1 random vector $\sim MN(\underline{0}, \Sigma_j)$ (3)

with

$$E[\underline{\varepsilon}_{lm}\ \underline{\varepsilon}_{kn}^T] = \begin{cases} \Sigma_j, & l = k \text{ and } m = n \\ 0_{8x8}, & l \neq k \text{ or } m \neq n;\ m,n = 1, 2, 3 \end{cases}$$

This last assumption arose from the additional assumption that the error for one pixel is uncorrelated with the error for any other pixel. Thus for $l \neq k$ at one site,

$$E[\underline{\varepsilon}_{lm}\ \underline{\varepsilon}_{km}^T] = E[\underline{\varepsilon}_{lm}]\ E[\underline{\varepsilon}_{km}^T] = \underline{0} \cdot \underline{0}^T = 0_{8x8}.$$

This property is called stochastic independence and is more commonly written as

$$E[(\underline{\varepsilon}_{lm} - \underline{\mu}_{\varepsilon_m})(\underline{\varepsilon}_{km} - \underline{\mu}_{\varepsilon_m})^T] = 0_{8x8},\ l \neq k \qquad (4)$$

and (4) results when we recall $\underline{\mu}_{\varepsilon_m} = \underline{0}$. Since the $\underline{\varepsilon}_{tj}$'s are the random portion of the model, the same stochastic independence assumption applies to the $\underline{Y}_{tj}$'s. The four modeling assumptions (1) to (4) will be collectively known as Ω in further discussion.

2. Collect Data and Estimate Parameters

The observed data used has previously been described. The estimates of $\underline{\mu}_j$ and Σ_j will be denoted as $\underline{\bar{Y}}_j$ and S_j respectively and

$$\bar{Y}_j = 1/100 \sum_{t=1}^{100} Y_{tj}$$

$$S_j = 1/92 \sum_{t=1}^{100} (Y_{tj} - \bar{Y}_j)(Y_{tj} - \bar{Y}_j)^T.$$

Two other quantities are also estimated from the data. These are the within-groups (sites) sums of squares and cross products (WSSCP) matrix, W with the estimate,

$$\hat{W} = \sum_{j=1}^{3} 92\ S_j$$

and the total sums of squares and cross products (TSSCP) matrix T, estimated by

$$\hat{T} = \sum_{j=1}^{3} \sum_{t=1}^{100} (Y_{tj} - \bar{\bar{Y}})(Y_{tj} - \bar{\bar{Y}})^T$$

where $\bar{\bar{Y}} = 1/3 \sum_{j=1}^{3} \bar{Y}_j$.

3. Test Statistics

The SSCP associated with W is calculated about the individual test site means, while the SSCP associated with T used the grand mean over all pixels as a correction factor. Clearly if any of the μ_j are different, the deviation vectors $(Y_{tj} - \mu_j)$ of the pixels in this population will be smaller than the deviation vectors formed by the difference of the same pixels and the grand mean $\mu = \sum_{j=1}^{3} \mu_j$. This observation is the basis for the test statistic used. The determinant of a square n x n matrix A, written $|A|$, can be thought of as a measurement of the volume in n dimensional space of the hyperdimensional solid whose vertices are the columns of the matrix A. The relation between the correction factor used in the deviation vector (μ_j or μ) and the determinant function is that the volume of the matrix gets smaller if $\mu_j \neq \mu$, for at least one j. The maximum value |W| can take on then is when $\mu_j = \mu$, in this case W = T and |W| = |T|. This motivates the test statistics which is attributed to Wilks, 1932, and is known as the Wilks-Lambda or Λ, where

$$\Lambda = \frac{|W|}{|T|}.$$

Since we only have samples from the population, Λ is estimated by $\hat{\Lambda}$ with $\hat{\Lambda} = |\hat{W}| / |\hat{T}|$. Since $\hat{\Lambda}$ is a function of Y_{tj}, it is itself a statistic and under Ω and the null hypothesis, $(\Omega \cap H_o)$ a func-

tion of $\hat{\Lambda}$ is distributed approximately as an χ^2 statistic or in special cases an F statistic[1] with degrees of freedom dependent upon the number in the sample, the number of groups (sites) and the number of properties (bands). To the calculated value of the statistic $\hat{\Lambda}$, Λ_{obs} (an arbitrary but fixed real number), is applied the F statistic transformation which yields the value G. After choosing the level of type I error willing to be risked (commonly called the α level), the researcher compares G to a critical value, F_c, corresponding to the α level of the F distribution present under $\Omega \cap H_o$, the decision rule is then

$G \leq F_c$ retain null hypothesis

$G > F_c$ reject null hypothesis and conclude, at least one $\underline{\mu}_j \neq \mu$.

The hypothesis testing environment is pictured in Figure 2. The purpose of this paper is now recast as whether F_c as just constructed yields the desired decision rule when the previously stated assumption (4) under Ω is relaxed. Do we still have a Pr (type I error) = α when we allow the observation vectors to be dependent?

The presence of a dependence structure in the remote sensing data has been documented by a number of researchers whose work was summarized previously. Based on these previous papers we will assume that the TMS data follows the autoregression integrated moving average model – ARIMA (1, 0, 1) – of Box and Jenkins (1970). This model is given by

$$\dot{Y}_t = \phi_1 \dot{Y}_{t-1} - \theta_1 a_{t-1} + a_t \qquad [1]$$

where: $\dot{Y}_t = Y_t - \mu$; is the deviation of the gray scale value at pixel t from its mean;

a_t, a_{t-1} are random variables $\sim N(0, \sigma_a^2)$ and cov $(a_t, a_{t-1}) = 0$ [covariance between a_t and a_{t-1}];

ϕ_1 and θ_1 are coefficients.

Under this model we examine assumption (4) as follows:

[1] The function of $\hat{\Lambda}$ is an exact F distribution for the special cases of properties (bands) = 2,3 or groups (training sites) = 2,3 (Tatsuoka, 1970). The work herein does fall into one of these special cases.

ASSUMING 1. H_0 $\bar{\mu}_1 = \bar{\mu}_2 = \bar{\mu}_3$

2. ***OBSERVATIONS ARE STOCHASTICALLY INDEPENDENT***

3. RANDOM VARIABLES ARE JOINTLY NORMALLY DISTRIBUTED

THEN THE RANDOM VARIABLE G WHICH IS A FUNCTION OF THE WILKS-LAMBDA STATISTIC (WHICH IN TURN IS A FUNCTION OF THE RANDOM OBSERVATIONS) FOLLOWS AN APPROXIMATE F-DISTRIBUTION

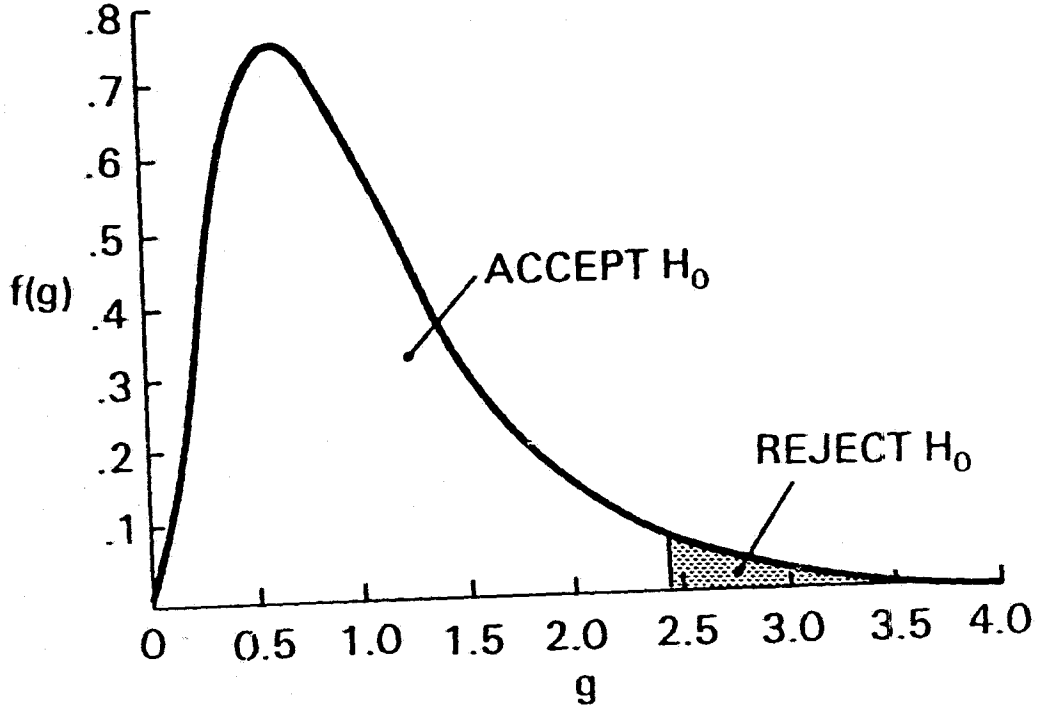


Figure 2. Setting for hypothesis testing in present study.

$$E[(Y_t - \mu)(Y_{t-1} - \mu)]$$

$= \text{cov}(\dot{Y}_t, \dot{Y}_{t-1})$ (since $\dot{Y}_t = Y_t - \mu_Y$)

$= \text{cov}(\phi_1 \dot{Y}_{t-1} - \theta_1 a_{t-1} + a_t, \dot{Y}_{t-1})$ (by [1])

$= \text{cov}(\phi_1 \dot{Y}_{t-1}, \dot{Y}_{t-1}) + \text{cov}(-\phi_1 a_{t-1}, Y_{t-1}) + \text{cov}(a_t, \dot{Y}_{t-1})$ (since cov is a linear operator)

$= \phi_1 \text{cov}(\dot{Y}_{t-1}, \dot{Y}_{t-1}) - \phi_1 \text{cov}(a_{t-1}, a_{t-1})$ (by independence of a's)

$= \phi_1 \text{var}(\dot{Y}_t) - \phi_1, \sigma_a^2$ (by stationarity)

$\neq 0$ (since $\text{var}(\dot{Y}_t) \neq 0\ \sigma_a^2$).

Thus the presence of the ARIMA (1, 0, 1) implies that the observations are independent.

SETTING UP THE SIMULATION

Examination of the effect of lack of independence was accomplished through a simulation procedure. A simulation approach was adopted because the question we examined concerned the adequacy of the central F distribution when the null hypothesis was correct, but the data was dependent in the fashion described. So the significance of the value of the test statistic yield by the training set must be determined by comparison with a frequency distribution occurring under the H_o (all class mean vectors are equal), and the suitable lack of independence assumptions. No such frequency distribution is tabled. Therefore, the gist of the simulation was to use a Monte Carlo approach to generate such a frequency distribution. This was accomplished by calculating G values from randomly generated data sets whose means are identical ($\underline{\mu}_1 = \underline{\mu}_2 = \underline{\mu}_3 = \underline{0}$), validating H_o, and which possessed covariance and autocorrelation structures calculated from the Cotter Basin test sites. To provide a control, these frequency distributions were then compared with frequency distributions generated from data with the same mean and covariance structure but not possessing the ARIMA (1, 0, 1) relationship. To create this simulation a set of assumptions was needed. These assumptions dealt with the model and error structure model.

Assumptions

The simulation used the following set of assumptions:

1. The first 3 assumptions under Ω, with assumption (1) reduced to $\underline{Y}_j = \underline{\mu} + \underline{\epsilon}_j$, i.e. H_o;
2. Each band is adequately represented by an univariate ARIMA (1, 0, 1) model; a further implication of this assumption is that the sequences, formed by scan lines, are stationary in the mean and variance, i.e. possess "wide sense" stationarity;
3. The transfer function between bands has a zero value for the delay parameter. This means that any correlation across bands is in phase, or equivalently if Y_{mt} is related to $Y_{n\,t-1}$, it can be stated as the cross correlation between Y_{mt} and the ARIMA model relating Y_{nt} and $Y_{n\,t-1}$, where n and m are bands.

The Variance-Covariance Matrix of Independent Observations

The system of ARIMA equations are parameterized by the ϕ_1's, θ_1's and the variances of a_t, a NID deviate vector. Under the assumptions for the simulation we obtained an estimate of this variance-covariance matrix of the independent random vectors as follows.

Let Y_{it} and Y_{jt} be the values of the gray scale for the bands i and j at pixel t of a scan line for the same site, then

$$\begin{aligned} \dot{Y}_{it} &= \phi_{i1}\,\dot{Y}_{it-1} - \theta_{i1}\,a_{it-1} + a_{it} \\ \dot{Y}_{jt} &= \phi_{j1}\,\dot{Y}_{jt-1} - \theta_{j1}\,a_{jt-1} + a_{jt} \end{aligned} \qquad [2]$$

(By assumption of ARIMA (1, 0, 1) model).

Using Equations [2]

$$\begin{aligned} \mathrm{cov}(\dot{Y}_{it}, \dot{Y}_{jt}) &= \mathrm{cov}(\phi_{i1}\,\dot{Y}_{it-1} - \theta_{i1}\,a_{it-1} + a_{it},\ \phi_{j1}\,\dot{Y}_{jt-1} - \theta_{j1}\,a_{jt-1} + a_{jt}) \\ \text{(Expanding)} \quad &= \mathrm{cov}(\phi_{i1}\,\dot{Y}_{it-1},\ \phi_{j1}\,\dot{Y}_{jt-1}) \\ &+ \mathrm{cov}(\phi_{i1}\,\dot{Y}_{it-1}, -\theta_{j1}\,a_{jt-1}) + \mathrm{cov}(\phi_{i1}\,\dot{Y}_{t-1},\ a_{jt}) \\ &+ \mathrm{cov}(-\theta_{i1}\,a_{it-1},\ \phi_{j1}\,\dot{Y}_{jt-1}) + \mathrm{cov}(-\theta_{i1}\,a_{it-1},\ -\theta_{j1}\,a_{jt-1}) \\ &+ \mathrm{cov}(-\theta_{i1}\,a_{it-1}, a_{jt}) + \mathrm{cov}(a_{it}, \phi_{j1}\,\dot{Y}_{jt-1}) \\ &+ \mathrm{cov}(a_{it}, -\theta_{j1}\,a_{jt-1}) + \mathrm{cov}(a_{it}, a_{jt}) \end{aligned}$$

(By stationarity and coincident cross band relationships)

$$= \phi_{i1}\,\phi_{j1}\,\mathrm{cov}\,(\dot{Y}_{it}, \dot{Y}_{jt}) - \phi_{i1}\,\theta_{j1}\,\sigma_{ij} + 0 - \theta_{i1}\,\phi_{j1}\,\sigma_{ij}$$
$$+ \theta_{i1}\,\theta_{j1}\,\sigma_{ij} + 0 + 0 + 0 + \sigma_{ij}$$

where σ_{ij} is the covariance of bands i and j based on the independent random vectors. Therefore,

$$\sigma_{ij} = \frac{(1 - \phi_{i1}\,\phi_{j1})}{1 + \theta_{ij}\,\theta_{j1} - \phi_{i1}\,\theta_{j1} - \theta_{i1}\,\phi_{j1}}\ \mathrm{cov}\,(\dot{Y}_{it}, \dot{Y}_{jt}) \qquad [3]$$

and $\hat{\sigma}_{ij}$ is the estimate of σ_{ij} formed by estimating all the quantities on the right-hand side of [3].

Estimates of ϕ and θ

The coefficients of the ARIMA (1, 0, 1) model, ϕ_1 and θ_1 were estimated for each of the sites. For each test site and for the eight spectral bands, three scan lines were chosen with each scan line composed of 100 pixels incorporating the test sites and the surrounding region in the east-west direction. The ϕ_1 and θ_1 coefficients were estimated using a maximum likelihood procedure as implemented in the program by Pack et al. (1972). The value of ϕ_1 did not vary significantly over bands within one test site. However θ_1 did exhibit significant variation with spectral band over a test site and so we decided to run the simulation three times using the maximum, median and minimum value of $\hat{\phi}_1$ as these coincided with the range of values of $\hat{\theta}_1$ found at each site. The values of $\hat{\phi}_1$ and $\hat{\theta}_1$ used are given in Table 5.

SIMULATION PROCEDURE

Figure 3 displays the steps in the simulation procedure. The first two steps were previously described and are the estimation of descriptive statistics such as the mean vectors, variance-covariance matrix and ARIMA coefficients, some of which were inputs at later points in the simulation. The variance-covariance matrices were then adjusted for the presence of autocorrelation as previously described. This yielded a new set of variance-covariance matrices, Σ_j^*. Then using the subroutine GGNSM from the International Mathematical and Statistical Libraries (IMSL, 1979) three sets (one for each site) of 150 random vectors were generated so that the individual vectors were stochastically independent and followed a MN (0 Σ_j^*) distribution. The appropriate ARIMA

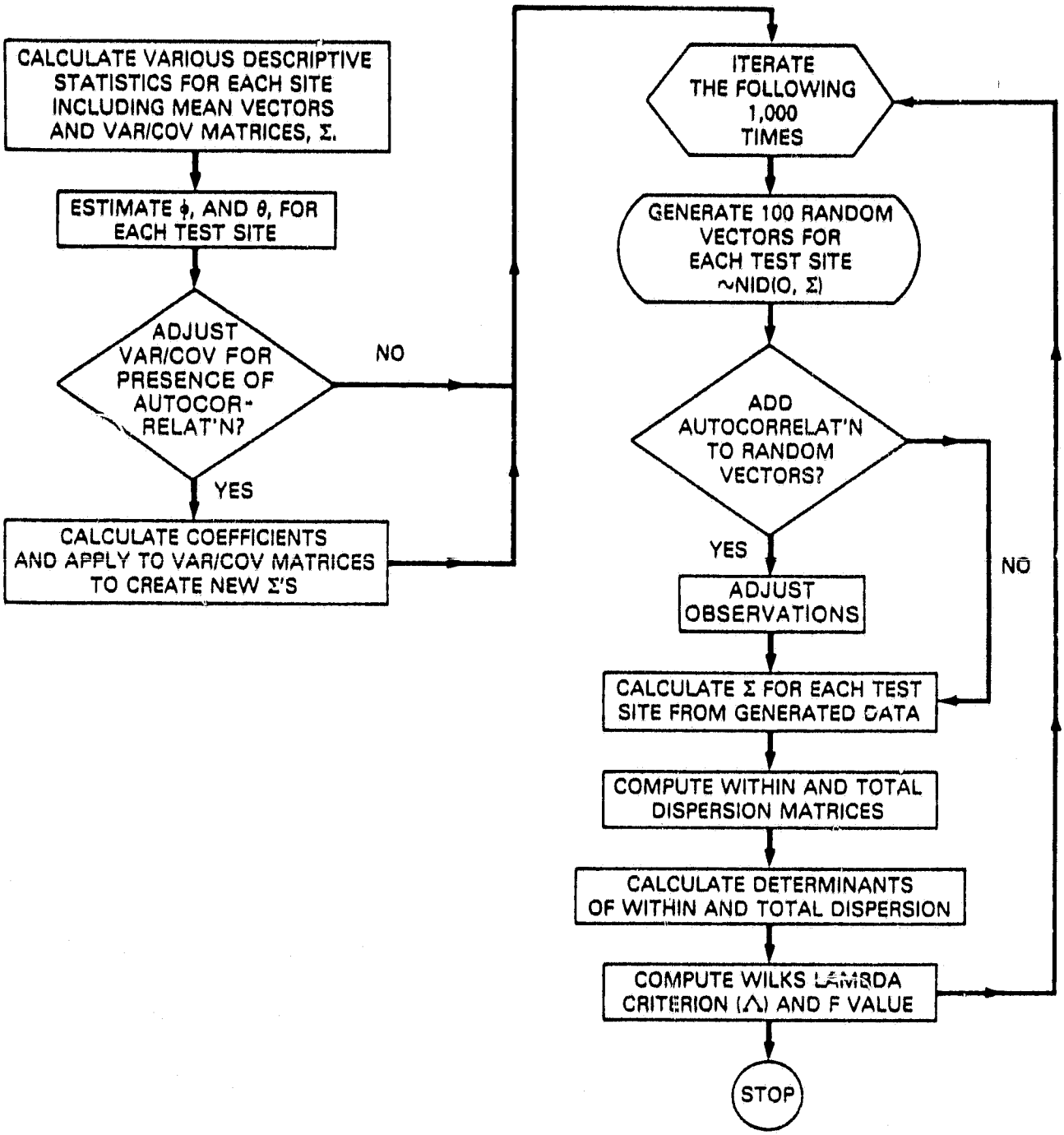


Figure 3. Flow in simulation procedure.

Table 5
Values of $\hat{\phi}_1$ and $\hat{\theta}_1$ used in simulation.

Degree of Autocorrelation		Site		
		A	B_1	B_2
Maximum	$\hat{\phi}_1$	0.948	0.927	0.926
	$\hat{\theta}_1$	0.275	-0.583	-0.112
Median	$\hat{\phi}_1$	0.907	0.895	0.901
	$\hat{\theta}_1$	0.212	0.452	-0.296
Minimum	$\hat{\phi}_1$	0.714	0.734	0.870
	$\hat{\theta}_1$	-0.259	-0.057	-0.322

A – anomalous test site
B_1 – background site 1
B_2 – background site 2

models for each site was then added to the vectors. This was accomplished by letting $\underline{a}_1 = \underline{\dot{Y}}_1$ and then using the recursive relationship formed by the ARIMA (1, 0, 1) model. The first 50 vectors were discarded to protect against transient or start up effects. From these three data sets, $\hat{W}$, $\hat{T}$ and their determinants were calculated (using the LINV2P subroutine of IMSL), followed by Λ_{obs} and G. This procedure was repeated 1000 times for each set of $\hat{\phi}_1$ and $\hat{\theta}_1$ used and yielded a frequency distribution of G. Using the same random seed string, 1000 values of G were calculated for each of the following:

1. Maximum Autocorrelation
2. Median Autocorrelation
3. Minimum Autocorrelation
4. No Autocorrelation[2]

[2]The 'no autocorrelation' iteration of the process was performed using both "adjusted" and "nonadjusted" covariance matrices. The results were identical.

RESULTS

Table 6 gives the critical or 95 and 98.75 percentile[3] values of the frequency distribution formed from the four iterations for the four comparisons (three paired site comparisons and one comparison of all three sites) plus the corresponding observed values which were calculated using the training sets selected for Cotter Basin. The same information for the comparison of all three sites is shown in Figure 4, which is the generated frequency distributions of G. The critical values for the generated data were determined from the observed cumulative frequency distribution, such that the desired proportion (either .05 or 0.0125) of the observations were greater than the critical value.

Table 6
Critical percentile values from generated frequency distributions versus G observed at Cotter Basin.

	Observed Values	Simulated Values							
		Degree of Autocorrelation							
		None		Most		Median		Least	
Comparison		α 0.05	0.0125	0.05	0.0125	0.05	0.0125	0.05	0.0125
All	92.8	2.50	3.05	216	328	102.5	140.0	28.0	35.2
A, B_1	80.9	2.00	2.80	140	196	75.2	105.6	13.5	20.4
B_1, B_2	65.0	2.00	2.68	107.5	157.5	67.2	86.4	22.0	30.5
A, B_2	165.4	2.04	2.68	122.5	162.5	60.0	79.2	25.2	33.2

A – anomalous site
B_1 – background site 1
B_2 – background site 2

The critical values from the iteration with no autocorrelation are slightly greater for all four comparisons than they would be under $\Omega \cap H_o$ (compare Tables 3 and 6). We attribute this to the unequal variance-covariance matrices from the three sites. The inflation of the critical value under the other three iterations in which autocorrelation is present is considerable. In three of the four

[3]The 98.75 percentile correspponds to an $\alpha = 0.0125$ significance level representing a Bonferroni level adjusted to yield an overall α level of 0.05 for the four comparisons within a given iteration of the simulation.

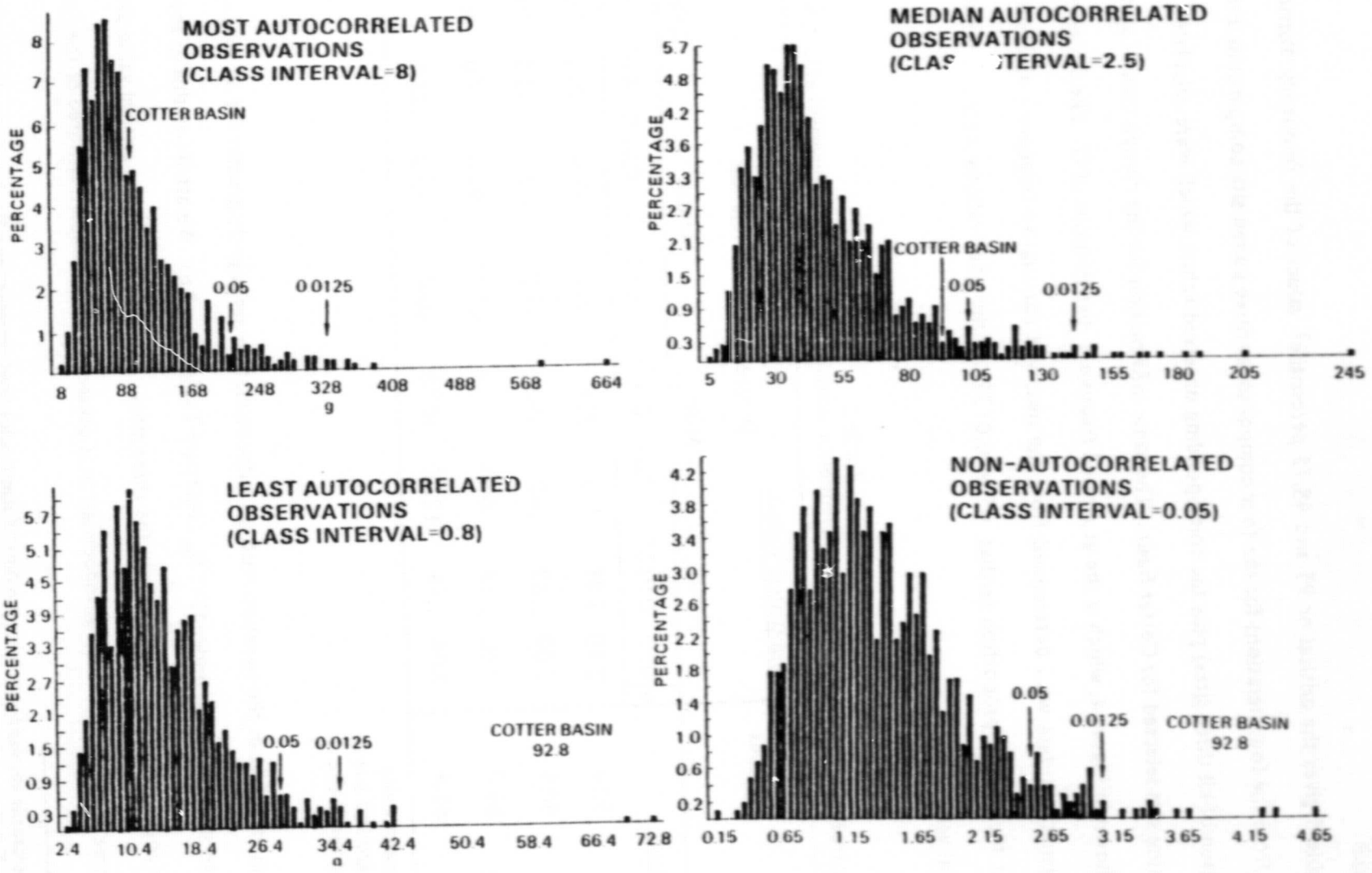
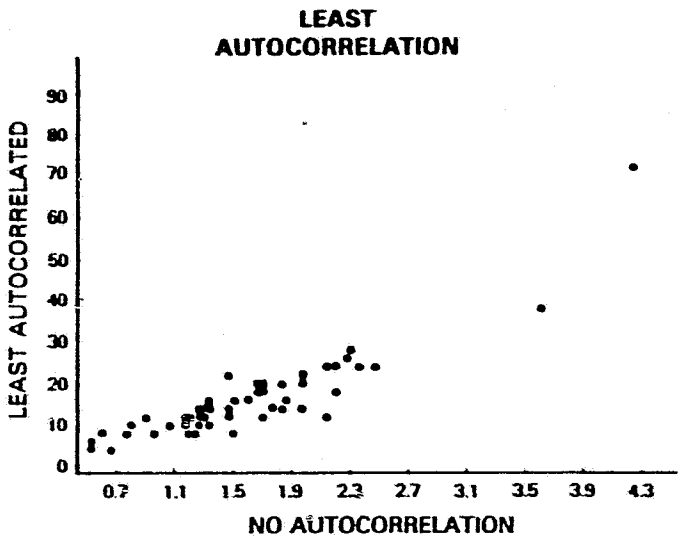


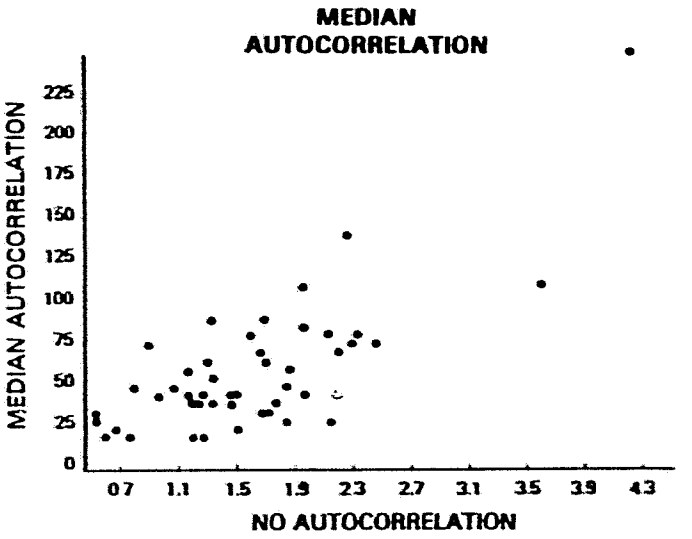
Figure 4. Frequency distribution of G values from generated data with differing degrees of autocorrelation, all three sites.

comparisons the observed value is less than the critical value of the frequency distribution generated from the "maximally" autocorrelated data. This implies that if observations were autocorrelated to the maximum degree, autocorrelation and not mean differences could account for the variation seen in the observed data (recall that the mean vectors of the generated data sets are not significantly different from 0). The fourth comparison, that between the anomalous and second background site, is just significant at the 0.0125 probability level and this result suggests a need for further examination. The important point here is that even for this comparison, the observed value of G as a test of mean differences is not, as it first appeared to be, so overwhelmingly significant when compared to the critical values derived from the simulation.

A similar result held true for the first three comparisons under the median iteration, that is none of the comparisons were significant at an $\alpha = 0.0125$ level. Only under the minimum iteration did all four comparisons appear significant at the $\alpha = 0.0125$ level. Even for this iteration the critical values were inflated by at least time times.

The relationship between the simulated data possessing varying degrees of autocorrelation and the uncorrelated data is summarized in Table 7. For each comparison, 50 sets of values of G were randomly selected. Each set contained four values of G, one from each iteration; all values of G in a set were based on the same random number seed. Figure 5 uses the data generated from the comparison of all three sites to display the values of G from the three autocorrelated iterations plotted individually against their corresponding non-autocorrelated values. The inflation in the value of G, as was indicated earlier, increases with the degree of autocorrelation. The slope (Table 7), as the inflation multiplier, varied between 7 and 70. Clearly r^2 must increase as the degree of autocorrelation is decreased since G based on autocorrelation data approaches the value of G based on non-autocorrelated data. This is an outcome of the common generation procedure.





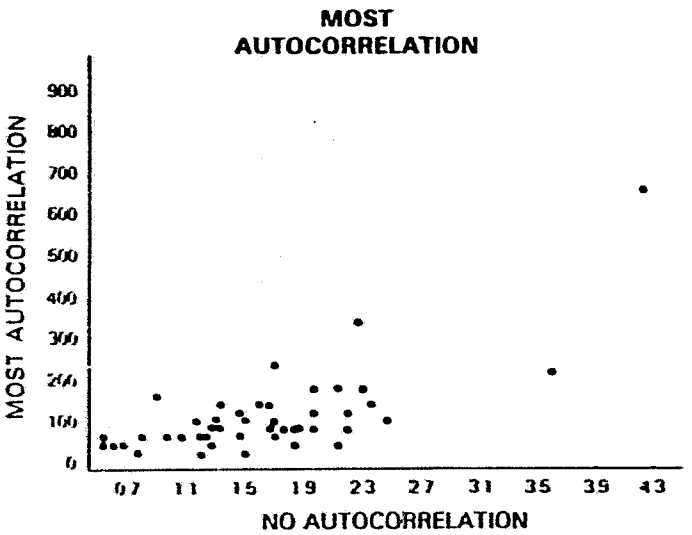


Figure 5. Plots of G values from autocorrelated observations versus G values from non-autocorrelated observations from all sites.

Table 7
Relationship between F statistics calculated from observations possessing differing degree of autocorrelation. Abscissa F values are calculated from corresponding non-autocorrelated data sets.

Comparison	Y	Slope	Intercept	R^2
All	Most	67.85	2.46	0.401
	Median	30.24	6.39	0.418
	Least	11.00	-0.81	0.783
A, B_1	Most	47.18	12.94	0.397
	Median	28.15	4.10	0.534
	Least	6.42	0.22	0.820
B_1, B_2	Most	36.23	13.88	0.374
	Median	20.72	9.12	0.360
	Least	9.80	0.72	0.757
A, B_2	Most	39.76	15.29	0.391
	Median	18.64	8.82	0.374
	Least	10.31	1.78	0.692

A – anomalous site
B_1 – background site 1
B_2 – background site 2

SUMMARY AND CONCLUSIONS

For a geobotanical test site at Cotter Basin, Montana, we have examined the MANOVA hypothesis testing situation for differences in the mean vectors, replacing the assumption of independence among the pixels (random vectors) in the training set by the assumption of an ARIMA (1, 0, 1) model for sequence of pixels. The examination was accomplished by a simulation which results in frequency distributions of variables which should be distributed as F under the null hypothesis that the mean vectors are equal.

Using TMS data acquired over Cotter Basin and previous research in this area, we selected test sites from a geobotanical anomaly and two suitably constrained background area. From initial comparisons of the mean vectors, we concluded that the test sites were very significantly different from one another based on reflectance and emittance in the eight TMS bands. Using estimates of

the covariances and autocorrelation developed from the test sites, data sets with mean vectors equal to zero were generated. The generated data was then manipulated to form the test statistics. Comparison of the frequency distributions of these test statistics with the initial result led us to conclude that the variation in mean reflectance from the test sites could be explained by the presence of autocorrelation in the data and not differences related to the presence of a geobotanical anomaly. From this result and those of other researchers, we believe it is a grevious, indeed fatal error, to use contiguous pixels as the training or testing set with most conventional statistical procedures; and conclusions drawn from analyses which do not observe this constraint should be treated circumspectly. This claim includes both parametric and nonparametric procedures. If procedures requiring indpendent observations are needed then a grid sample [often with a spacing $\geq$ 10 pixels (Craig, 1978; Labovitz and Masuoka, 1981)] or a suitable filtering algorithm should be used to remove the autocorrelation.

ACKNOWLEDGMENTS

We would like to acknowledge earlier research on this site and data set performed by our colleagues Charles Schnetzler and Alicia Siegrist. We would also like to thank Dr. Schnetzler for directing the location of training sites. R. Bell, J. C. Griffiths, J. R. Irons and C. C. Schnetzler are thanked for their thoughtful review of this manuscript. Lynne Hooker is acknowledged for assistance in preparation of the manuscript.

REFERENCES

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

Basu, R., Basu, J. P. and Lewis, T. O., 1976, "Effect of Interclass Correlation on Confidence Sets Based on Chi-Squared Statistics," IEEE Trans. Systems Man and Cybernetics, Vol. SMC-6, no. 4, pp. 445-448.

Box, G. E. P. and Jenkins, G. M., 1970, "Time Series Analysis: Forecasting and Control, 2nd Edition," Holden-Day, NY, 553 p.

Campbell, J. B., 1981, "Spatial Correlation Effects Upon Accuracy of Supervised Classification of Land Cover," Photogrammetric Engineering and Remote Sensing, V. 47, no. 3, pp. 553-563.

Chieu, H. Y., and Collins, W. E., 1978, "A Spectroradiometer for Airborne Remote Sensing," Photogrammetric Engineering and Remote Sensing, V. 44, no. 4, pp. 507-517.

Cliff, A. D. and Ord, J. L., 1981, "The Effects of Spatial Autocorrelation of Geographical Modeling," in Future Trends in Geomathetmatics, R. G. Craig and M. L. Labovitz, eds., Pion Ltd., London, pp. 108-137.

Coberly, W. A., 1973, Personel Communications, Memorandum from Mathematical Physics Branch of Mission Planning and Analysis Division, NASA/JSC, Houston, TX, (as reported in Basu et al., 1976)

Collins, W. E., Chang, S., and Kuo, J. T., 1980, "Detection of Hidden Mineral Deposits by Airborne Spectral Analysis of Forest Canopies," Final Report to NASA, Contract NSG-5222, 61 p.

Collins, W. E., Raines, G. L., and Canney, F. C., 1977, "Airborne Spectroradiometer Discrimination of Vegetation Anomalies Over Sulfide Mineralization – A Remote Sensing Technique," Geological Society of America Abstract with Programs, V. 9, no. 7, pp. 932-933.