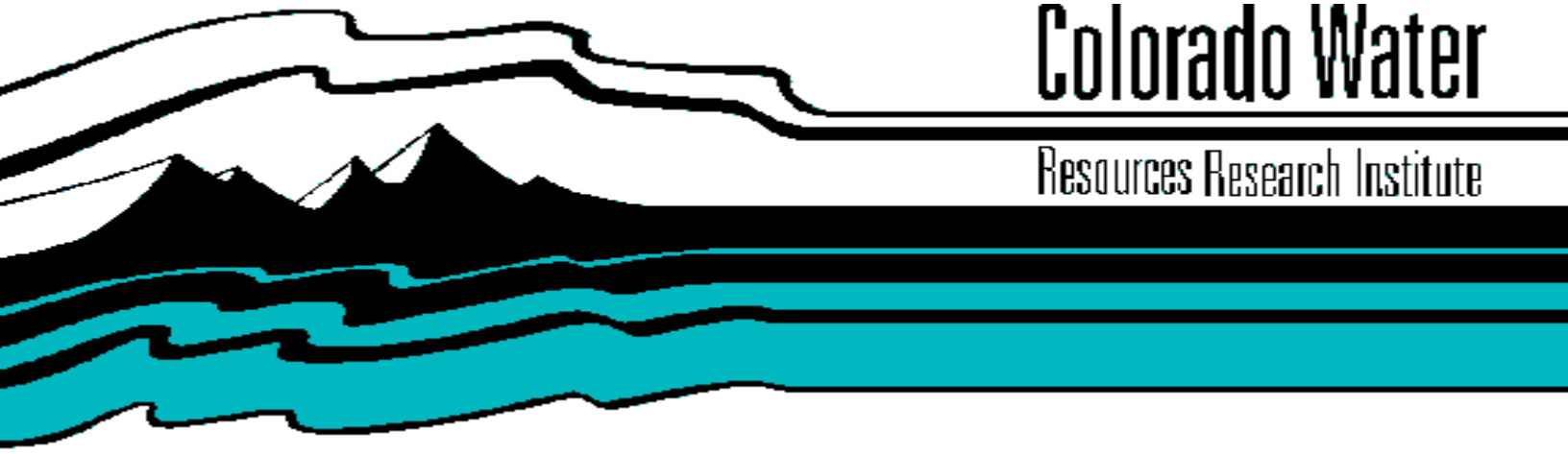
Multivariate Methods for Detecting Water Quality Trends

By Jim C. Loftis
Charles H. Taylor
Phillip L. Chapman

February 1990

Technical Report No. 56



Colorado State University

MULTIVARIATE METHODS FOR DETECTING WATER QUALITY TRENDS

by

Jim C. Loftis
Department of Agricultural and Chemical Engineering

Charles H. Taylor
and
Phillip L. Chapman
Department of Statistics

Colorado State University
Fort Collins, Colorado 80523

February 1990

Prepared for

Environmental Research Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Corvallis, Oregon 97333

COLORADO WATER RESOURCES RESEARCH INSTITUTE
Colorado State University
Fort Collins, Colorado 80523

Neil S. Grigg, Director

ABSTRACT

Several methods of testing for multivariate trend have been discussed in the statistical and water quality literature. We review both parametric and nonparametric approaches and compare their performance using synthetic data. A new method, based on a robust estimation and testing approach suggested by Sen and Puri, performed very well for serially independent observations. A modified version of the covariance inversion approach presented by Dietz and Killeen also performed well for serially independent observations. For serially correlated observations, the covariance eigenvalue method suggested by Lettenmaier was the best performer.

MULTIVARIATE METHODS FOR DETECTING WATER QUALITY TRENDS

INTRODUCTION

One of the major goals of long-term fixed station monitoring for water quality and other environmental variables is the detection of trends. Statistical methods for trend analysis of the resulting data series are well developed and fairly routinely applied (Smith and Alexander, 1987). Applications of statistical tests for trend have generally been univariate, i.e., have considered only a single water quality variable at a single monitoring location. However, routine monitoring programs typically measure several variables at several locations, and water quality processes are perhaps more logically viewed as multivariate.

Multivariate methods for trend analysis (defined for present purposes as statistical tests for trend) have been developed for a variety of applications and have recently been modified with an eye specifically toward water quality (Lettenmaier, 1988). We review the available methods, both parametric and nonparametric, and present a new approach which applies the work of Sen and Puri (1977) and combines some aspects of both parametric and nonparametric approaches. Monte Carlo simulation results, focusing on quarterly sampling, compare performance of the tests. Our discussion is limited to gradual, monotonic changes in

water quality as opposed to sudden (step) changes or changes which include both increases and decreases over the period of interest. We also concentrate on testing for significance of linear trend and touch upon the subject of estimating trend magnitudes only peripherally.

BACKGROUND

Trend detection methods for water quality time series have been extensively discussed in the literature. Montgomery and Reckhow (1984), and more recently Berryman et al. (1988) present good overviews of available techniques. Both parametric and nonparametric methods of trend analysis are important, although the latter class is certainly more widely applicable in water quality and has, therefore, received much more attention. Nevertheless, classical methods, and in particular robust extensions thereof, should not be forgotten as useful tools in water quality data analysis.

Parametric Methods

Parametric approaches are frequently based on linear models and normal theory. Transformations of the data are sometimes used to achieve resemblance of normality (McLeod et al., 1983). In both parametric and nonparametric approaches, seasonal transformations are sometimes used as well to achieve stationarity in the mean and/or variance.

Seasonal patterns in the mean may be handled more directly by the use of indicator variables in the analysis of covariance model (Neter and Wasserman, 1974). Taylor and Loftis (1989) found that analysis of covariance on the ranks of data (an approach suggested by Conover, 1980) compared favorably with other nonparametric methods of trend detection when data were lognormally distributed. Analysis of covariance on ranks, however, does not provide an estimate of the trend slope of the original data.

Multivariate extensions of analysis of covariance are relatively straightforward (Anderson, 1984). In particular, estimates of the trend slope are the same as the individual univariate estimates. Tests of significance, based on estimates of the covariance matrices, will be discussed in the methods section.

The field of robust estimation occupies a significant share of the statistical literature, and several methods appear to have promise for applications in water quality. We chose to explore an avenue suggested by Sen and Puri (1977). This method is based on the classical linear model, but uses an asymptotically distribution free aligned rank order method of estimating the parameters and testing their significance. Although Sen and Puri showed that their method was asymptotically distribution free, it remained to show that the method was robust (and powerful) with real or synthetic data records of moderate or even finite length.

Nonparametric Methods

Lettenmeier (1988) and Taylor and Loftis (1989) review the details of current nonparametric methods of trend analysis. Briefly, however, the most useful procedures are based on the Mann-Kendall test for trend (Mann, 1945 and Kendall, 1975). The popular Seasonal Kendall test (Hirsch et al., 1982) is a multivariate extension of the test in which each season is treated as a separate and independent variable.

Another multivariate version of the Mann-Kendall test, which accounts for covariance between variables, was presented by Dietz and Killeen (1981). This test was modified by Hirsch and Slack (1984) to produce a version of the Seasonal Kendall test which accounts for correlation between seasons. Lettenmaier (1988) developed a version of the test which was better suited to the analysis of multiple water quality variables and/or locations. Although the Mann-Kendall tests do not provide estimates of trend magnitude, a Seasonal Kendall slope estimator (Hirsch et al., 1982) is appropriate for application in parallel with the testing procedures.

Our study compares the last three tests above, based on the Mann-Kendall test, with two linear model based tests--classical multivariate analysis of variance (MANOVA) and the aligned rank order method of Sen and Puri. We attempt to

define situations in which tests of both groups--linear models and Mann-Kendall based tests--would be useful.

Our simulation study, like those of Hirsch and Slack (1984) and Lettenmaier (1988), includes only linear trends. While trends which are "nearly linear" are important in practice, they represent only a small subset of a broader class of general monotonic trends and of a still broader class of general changes of unspecified form. Furthermore, since a linear trend assumption should favor the linear models group, our comparison should not be extended to the more general situation. Our simulation study was also limited to "nice" data sets with no missing or censored values, complications common in real water quality records. (See Hughes and Millard, 1988, for a discussion of multiple censoring levels.)

STATISTICAL MODELS FOR ESTIMATION AND TESTING

Before discussing the actual trend estimation and testing procedures, it is useful to first discuss appropriate models for the problem. The models are then used to precisely specify the hypotheses to be tested.

Let us assume that δ water quality constituents are measured for n years at λ locations during ω seasons per year. The most general model would, therefore, have $\omega\delta\lambda$ dependent (response) variables, one for each season by constituent by location combination. To simplify this discussion we limit consideration to multiple water quality

variables at a single monitoring location. Thus there are $\omega\delta$ response variables. However, extension to the more general model is straightforward.

The dependent variables are not necessarily independently distributed. In other words, there may be correlation between constituents and between seasons, as well as cross correlations.

Our primary interest is to determine whether any of the dependent variables have a nonzero monotone trend over time. Estimates of trend magnitude are also desired.

We now discuss two different models as possibilities for estimation and testing of a trend in the joint distribution of the dependent variables. The first and most general model has a large number of parameters and would be unwieldy for practical application to short time series. However, to obtain maximum flexibility, we used this model to generate synthetic data for evaluating the alternative trend testing procedures. The second, reduced model forms the basis of our two linear model-based procedures and has similarities to the structure assumed by the three Mann-Kendall based procedures.

Model 1:

In the first model, each water quality constituent in each season represents an individual dependent variable. Each of the $\omega\delta$ dependent variables may be correlated with every other in an arbitrary fashion. This model is,

therefore, the more relaxed in terms of assumptions on the covariance structure.

The model for observations made during year i(i=1(1)n) may be written as

$$y_i = \mu + \beta i + \varepsilon_i. \tag{1}$$

If

$j = 1,\ldots, \omega$ is the season index, and

$k = 1,\ldots, \delta$ is the constitutent index and

$$y_i = [y_{i11},\ldots,y_{ijk},\ldots,y_{i\omega\delta}]^T \tag{2}$$

is a vector of $\omega\delta$ dependent variables. Similarly,

$$\mu = [\mu_{11},\ldots,\mu_{jk},\ldots,\mu_{\omega\delta}]^T \tag{3}$$

is a vector of $\omega\delta$ intercept terms;

$$\beta = [\beta_{11},\ldots,\beta_{jk},\ldots,\beta_{\omega\delta}]^T \tag{4}$$

is a vector of $\omega\delta$ slopes (trend magnitudes) over time. The vector of $\omega\delta$ random erros,

$$\varepsilon = [\varepsilon_{i11},\ldots,\varepsilon_{ijk},\ldots,\varepsilon_{i\omega\delta}]^T, \tag{5}$$

has a zero mean and a general covariance matrix $\sum_1$, which is assumed to be constant from year to year. This model makes no assumptions about the covariance structure of the observations made on different constituents or in different seasons. (Except that $\sum_1$ must be non-negative definite.)

In scalar notation, we may represent a single observation of water quality constituent k during season j of year i as

$$y_{ijk} = \mu_{jk} + i\beta_{jk} + \varepsilon_{ijk} \qquad (6)$$

Thus each constituent, by season combination is represented by a separate straight line with a slope of β_{jk} units per year. The random error ε_{ijk} has zero mean and variance given by the corresponding diagonal element of $\sum_1$. Furthermore this error is correlated with $\varepsilon_{ij'k'}$, where $j' \neq j$ and/or $k \neq k'$. The covariance between ε_{ijk} and $\varepsilon_{ij'k'}$ is given by the appropriate off-diagonal element of $\sum_1$. Note that ε_{ijk} and $\varepsilon_{ij'k'}$ occur during the same year. However, for generating data, we also consider specific correlation between error terms in different years.

Testing for trend in this model is equivalent to testing the null hypothesis that $\beta_{jk} = 0$ for all $k = 1(1)\omega$, $j = 1(1)\delta$. Rejection of this null hypothesis is evidence that a nonzero trend exists for at least one of the constituent by season combinations.

As we mentioned earlier, this model has a large number of parameters. Suppose that two water quality constituents are measured in two seasons ($\omega = \delta = 2$). The complete model is represented by four equations of the form shown in Equation 6. There are four dependent variables, and the model parameter set consists of four intercept terms, four slopes, four variances, and $\frac{4(4-1)}{2} = 6$ intervariable

correlations for a total of 18 parameters. As the number of constituents and seasons increase, the number of parameters becomes very large very quickly. (For $\omega = \delta = 4$ there are 168 parameters, 136 of which are in $\sum_1$.)

Model 2:

The second model reduces the number of dependent variables by modeling the seasonal effect using a seasonal adjustment to the intercept term rather than separate dependent variables. Thus in this case there are only δ individual dependent variables and δ trend slopes.

Observations made during the jth season of the ith year are represented by Model 2 as

$$y_{ij} = \mu + \beta i + \Gamma_j + \epsilon_{ij} \tag{7}$$

where the indices i,j,k are as previously defined and

where $$y_{ij} = [y_{ij1},\ldots,y_{ijk},\ldots,y_{ij\delta}]^T \tag{8}$$

is a vector of δ dependent variables.

$$\mu = [\mu_1,\ldots,\mu_k,\ldots,\mu_\delta]^T \tag{9}$$

is a vector of δ intercept terms

$$\beta = [\beta,\ldots,\beta_k,\ldots,\beta_\delta]^T \tag{10}$$

is a vector of δ slopes. The vector

$$\Gamma_j = [\Gamma_{j1},\ldots,\Gamma_{jk},\ldots\Gamma_{j\delta}]^T \tag{11}$$

represent δ seasonal adjustments to the intercept vector. Since μ represents the global intercept term from averaging over seasons, only ω-1 independent seasonal adjustments are needed for each constituent. (The ω seasonal adjustments for a given constituent must sum to zero.) Also the vector,

$$\varepsilon_{ij} = [\varepsilon_{ij1},\ldots,\varepsilon_{ijk},\ldots\varepsilon_{ij\delta}]^T, \qquad (12)$$

consists of δ random error terms with zero means and covariance matrix $\sum_2$.

Model 2 assumes that the covariance matrix $\sum_2$ has a more rigid structure than that in the previous model. Specifically we assume that the covariance between ε_{ijk} and $\varepsilon_{ijk'}$ is the same for all i = 1(1)n and j = 1(1)ω (i.e., intervariable correlations are the same for all seasons and years).

Using scalar notation we can write for a single dependent variable

$$y_{ijk} = \mu_k + i\beta_k + \Gamma_{jk} + \varepsilon_{ijk} \qquad (13)$$

The slope β_k and overall intercept, μ_k, depend only on the constituent. Since the Γ_{jk} adjust the intercept terms for each season, this model in effect includes ω different parallel lines for each constituent, all having the same slope β_k. The error term ε_{ijk} is assumed to be independent of error terms corresponding to observations on the same dependent variable from different seasons and different years, but the remaining possible correlations may be nonzero.

It again follows that the variance of ε_{ijk} is the corresponding diagonal element of $\sum_2$, and the covariance between ε_{ijk} and $\varepsilon_{ijk'}$ is the corresponding off-diagonal element of $\sum_2$. Note that ε_{ijk} and $\varepsilon_{ijk'}$ correspond to measurements during the same year and season but on different constituents.

As mentioned earlier, Model 2 is applied directly in both of our linear model based tests. Of course, the Mann-Kendall based tests do not assume a linear model. However, it is appropriate to point out that our treatment of seasonal effects in the Mann-Kendall tests will more closely resemble Model 2 than Model 1. It is also appropriate to note that in the Mann-Kendall based tests, a more relaxed structure of $\sum_2$ is assumed in which covariance between error terms of different seasons in the same year is permitted.

Testing for trend in Model 2 is equivalent to testing the null hypothesis $\beta_k = 0$ for $k = 1,\ldots,\delta$. Rejection of this null hypothesis would support the alternative that a nonzero trend, averaged over all seasons, exists for at least one of the constituents, leading to a different conclusion than in Model 1 which differentiates between seasons. In applying Model 2, we effectively assume that a trend affects all seasons equally. This assumption could be relaxed by including an interaction term to adjust the slope for each season in the same manner that Γ_{jk} adjusts the intercept. For simplicity we chose not to include such a term in the present study.

To compare the complexity of models 1 and 2, recall the example with two constituents and two seasons. Model 2 is then represented by two equations of the form of Equation 3. This model has two dependent variables and 9 parameters (2 intercepts, 2 slopes, 1 seasonal adjustment for each constituent, and 3 distinct elements of $\sum_2$). One might not view this as parsimonious; still, the number of parameters does not "blow up" as rapidly with increasing ω and δ as for Model 1. For example, if $\omega = \delta = 4$, the total number of parameters is 30 (10 or which are in $\sum_2$) compared with 168 in Model 1.

THEORY

The MANOVA Procedure

Suppose each of n independent random vectors $\{y_i \mid i = 1(1)n\}$ (each of length p) is assumed to be normally distributed with a mean βx_i and a variance covariance matrix $\sum$. The v components of the design vector x_i are known. However the elements of the p x v matrix β are unknown as are the elements of the p x p matrix $\sum$. The multivariate analysis of variance (MANOVA) procedure estimates these unknowns, and tests whether certain submatrices of β are zero.

Since this method is described in detail by Chapter 8 of Anderson (1984), only a brief description of the associated testing procedure as it pertains to our specific problem is in order. To be consistent with model two,

redefine the observation vectors and the design vectors as $\{y_{ij} \mid i = 1,\ldots,n,\ j = 1,\ldots,\omega\}$ and $\{x_{ij} \mid i = 1,\ldots,n\ j = 1,\ldots,\omega,\}$ respectively, consequently, the observation vector y_{ij} corresponds to the j^{th} season of the i^{th} year. Furthermore index the δ elements (where each element corresponds to a constituent) of the vector y_{ij} as in Model 2, Equation 8. In addition partition β as $\beta = [\beta_1 \mid \beta_2]$ where β_1 and β_2 have dimensions $\delta \times \omega$ and $\delta \times 1$, respectively. Similarly, partition the design vector x_{ij} into

$$x_{ij}^T = [x_{ij1}^T \mid x_{ij2}^T] \tag{14}$$

The first element of x_{ij1} is one. However element $r = 2,\ldots,\omega$ of x_{ij1} is a one only if $j = r - 1$. Otherwise element r is zero (i.e., the r^{th} element is one only if the observation was made during season r-1. Also, if the eason is ω, then elements $2,\ldots,\omega$ of x_{ij1} are zero.) The single element of x_{ij2} is set equal to i. As a result, the k,1 element of β_1 represents an intercept term (μ_k from Equation 9) for observations of constituent variable k made during season ω. Element k,s ($s = 2(1)\omega$) represents the difference in the intercept term between that for variable k during season $j = s-1$ and that for variable k during season ω. These elements are the seasonal adjustments $\Gamma_{j,k}$ from Equation 11. Element k of β_2 represents a linear trend in variable k over years.

The objective is to test the null hypothesis H_0: $\beta_2 = 0$. As in the univariate regression case, this is accomplished by comparing the estimate of the covariance matrix $\sum$ under the assumption that the null hypothesis holds to the case where this assumption is relaxed. First estimate $\sum$ without the assumption of the null hypothesis as

$$\sum_{\theta} = \frac{\sum_{i,j} y_{ij} y_{ij}^T - C_\theta A_\theta^{-1} C_\theta^T}{n\omega} \tag{15}$$

where

$$C_\theta = \sum_{i,j} y_{ij} x_{ij}^T \tag{16}$$

(with dimensions $\delta \times (\omega + 1)$) and

$$A_\theta = \sum_{i,j} x_{ij} x_{ij}^T \tag{17}$$

(with dimensions $(\omega + 1) \times (\omega + 1)$). Under the null hypothesis, $\sum$ is estimated as

$$\sum_{0} = \frac{\sum_{i,j} y_{ij} y_{ij}^T - C_o A_o^{-1} C_o^T}{n\omega} \tag{18}$$

where

$$C_o = \sum_{i,j} y_{ij} x_{ij1}^T \tag{19}$$

(with dimension $\delta \times \omega$) and

where

$$A_o = \sum_{i,j} x_{ij1} x_{ij1}^T \tag{20}$$

(with dimension $\omega \times \omega$).

Define,

$$U = \frac{|\sum_\theta|}{|\sum_o|} \tag{21}$$

Then, the test statistic for testing this null hypothesis against the alternative that at least one element of β_2 is not equal to zero is

$$\pi = \frac{(1 - U)(n\omega - \omega - \delta)}{U\delta} \tag{22}$$

which has an F distribution with δ and $n\omega - \omega - \delta$ degrees of freedom. Therefore the H_0 is rejected if π is too large when compared to its critical value. We refer to the test using this statistic as the MANOVA method.

The Sen and Puri Procedure

Consider a set of w independent random vectors $\{y_i \mid i=1(1)v\}$ of length p. Each of these vectors is assumed to have an unknown cumulative distribution function

$$F_i(y_i) = P(y_i \leq y_i) = F(y_i - \alpha - \beta x_i) \tag{23}$$

where α is a vector of length p, β is a p x q matrix and x_i is a vector of length q with known elements. The elements of α and β are assumed to be unknown.

If we partion the matrix β into

$$\beta = [\beta_1 \mid \beta_2] \tag{24}$$

where β_1 and β_2 have dimensions p x q_1 and p x q_2 respectively, then Sen and Puri (1977) suggest a procedure for estimating β and testing the null hypothesis, $H_0: \beta_2 = 0$. Note that in the particular problem treated in this paper, $w = n\omega$, $p = \delta$, $q_1 = \omega-1$, and $q_2 = 1$. In addition, the observation sectors have the same form as described in the preceding section on MANOVA. Except for the removal of the first element (corresponding to the intercept term), the design vectors have the same form. In other words, $q = v-1$ where v is defined in the MANOVA section, and the present design vector consists of elements 2 through v of the former design vector.

A general discussion of this method is not available as it is for MANOVA. Therefore, in order to describe the

method in general terms, we use a single index on y_i and x_i for the remainder of this section. (Instead of indexing the time of an observation with year and season, i and j, we simply use a time i, with i = 1, n, where n = the total number of observations.

This procedure first estimates β under the null hypothesis (i.e., with $\beta_2 = 0$). The actual computations were not discussed in Sen and Puri (1977), but their estimators may be described in terms of separate univariate estimates of the p rows of β. (Note that the p^{th} row of β corresponds to the p^{th} element of y_i.) Consequently, results from other papers which discussed the univariate rank based estimators are used.

Since the estimation procedure may also be used to estimate β without the restrictions of H_0, the process will be described for a general β. Let y_{ik} be the k^{th} element of the i^{th} observation. (This could correspond to water quality constituent k at time i.) For some estimate b_k of β_k (the k^{th} row of β which corresponds to the univariate linear model

$$y_{ik} = \alpha_k + \beta_k x_i + \varepsilon_i \qquad (25)$$

where ε_i is a random error) the residual corresponding to y_{ik} is denoted as $\gamma_{ik}(b_k)$ where

$$\gamma_{ik}(b_k) = y_{ik} - b_k x_i. \qquad (26)$$

In addition define

$$R_{ik}(b_k) = \sum_{s=1}^{n} \mu(\gamma_{ik}(b_k) - \gamma_{sk}(b_k)) \qquad (27)$$

where

$$\mu(z) = \begin{cases} 1 \text{ if } z \geq 0 \\ 0 \text{ if otherwise} \end{cases} \qquad (28)$$

Thus $R_{ik}(b_k)$ is the rank of $\gamma_{ik}(b_k)$ among the $\gamma_{sk}(b_k)$. Now for each $k = 1,\ldots,p$, a set of scores, $\{a_n^k(\ell) \mid \ell = 1,\ldots,n\}$ is generated by a score function $\phi_k(\ell)$ (which is square integrable and non-decreasing on $[0,1]$) as

$$a_n^k(\ell) = \phi_k\left(\frac{\ell}{n+1}\right).$$

For the present work we chose $\phi_j(\ )$ as the identity function. Many other score functions are possible.

For a given estimate, B of β, the Sen and Puri estimation process first calculates

$$S_r(b_k) = \sum_{i=1}^{n} (x_{ir} - \bar{x}_r)\, a_n^k(R_{ik}(b_j)) \quad r = 1,\ldots,q \qquad (29)$$

where $\bar{x}_r = \frac{1}{n}\sum_{i=1}^{n} x_{ir}$.

and then

$$S^*(B) = \sum_{j=1}^{p} \sum_{r=1}^{q} |S_r(b_k)| . \qquad (30)$$

The best estimate of β is the B which minimizes $S^*(B)$. However the set $\{B \mid S^*(B) = \text{minimum}\}$ may contain an uncountable number of elements, since $S^*(B)$ is a step function. Thus the Sen and Puri estimate is actually the center of mass of this set.

Note however that minimizing $S^*(B)$ is equivalent to separately minimizing with respect to each of the p rows, $\{b_k \mid k = 1(1)p\}$, of B. But each $|S_r(b_k)|$ is a step function of q variables which makes the minimization procedure difficult. However, Jaeckel (1972) indicated during a univariate discussion, that the q x 1 vector

$$S(b_k) = [S_1(b_k).,,,,.S_q(b_k)]^T \qquad (31)$$

is essentially the gradient of

$$D(b_k) = \sum_{i=1}^{n} a_n^k(R_{ik}(b_k))\gamma_{ik}(b_k) \qquad (32)$$

Furthermore, since $D(b_k)$ is a continuous convex function of b_k whose minimum occurs in a bounded region, a point b_k which minimizes $D(b_k)$ will approximately solve $S(b_k) = 0$ which is equivalent to minimizing

$$\sum_{r=1}^{q} | S_r(b_k)| .$$

Consequently, our estimate of β is derived from the minimization of $D(b_k)$ for each k.

The actual procedure used for this minimization is suggested by Aubuchon and Hettmansperger (1984). The mechanics of this procedure are based on an asymptotic approximation of $D(b_k)$ by the quadratic function

$$D(b_k) \cdot D^*(b_j) = D(\beta_k) -$$

$$(b_k - \beta_k)^T S(\beta_k) + (2\tau)^{-1} (b_k - \beta_k)^T X^T X (b_k - \beta_k) \qquad (33)$$

(where β_k is the true value and τ is a scalar quantity) which attains a minimum at

$$b_k = \beta_k + \tau(X^TX)^{-1}S(\beta_k). \qquad (34)$$

This leads to an iterative method of estimating β_k. At the m^{th} step, the m + 1 estimate of β_k is

$$b_k^{(m+1)} = b_k^{(m)} + t^{(m)}(X^TX)^{-1}S(b_k^{(m)}) \qquad (35)$$

where $b_k^{(1)}$ is the least squares estimator of β_k (other starting estimates, such as those arrived at by LAD may also be used). But, the obvious question is the optimal value of scalar step, $t^{(m)}$.

The derivative of

$D(b_j^{(m)} + t(X^TX)^{-1}S(b_k^{(m)}))$ with respect to t is

$$G(t) = -a(b_k^{(m)})^T X\ (X^TX)^{-1}X^T a(b_k^{(m)}) + t(X^TX)^{-1}S(b_k^{(m)})) \quad (36)$$

where $a(b_k^{(m)})$ is a vector of length n containing the scores arising from the estimate $b_k^{(m)}$. This derivative is a non-decreasing step function. Therefore, the Illinois version of false position (Dowell and Jarratt, 1971) is used to find the optimal value of $t^{(m)}$. This two stage iteration is repeated until the percentage change in the deviance, $D(b_k)$ is arbitrarily small.

To test the null hypothesis, partition each b_k into b_{1k} and b_{2k} which have dimensions 1 x q_1 and 1 x q_2 respectively. The estimation procedure described above is carried out under the restriction that all b_{2k} must be equal to 0. Denote the best estimates of β_k under this restriction by β^*_k and calculate the following matrices:

1. M_n (of dimension p x p) where the k,k' element is defined as

$$(M_n)_{kk} = (n-1)\sum_{i=1}^{n} a_n^k\ (R_{ik}(\beta_j^*))\ a_n^k\ (R_{ik'}(\beta^*_{k'})) \quad (37)$$

$$- \bar{a}_n(k)\ \bar{a}_n(k')$$

and

2. $C = (X^TX)^{-1}$. (38)

The test statistics

$$L = \sum_{r=1}^{q} \sum_{r'=1}^{q} \sum_{k=1}^{p} \sum_{k'=1}^{p} S_r(b_k)\, S_{r'}(b_{k'})\, m^{kk'} C_{rr'}, \qquad (39)$$

(where $m^{kk'}$ is the k, k' element of M_n^{-1}) has an asympotitic chi-squared distribution with $q_2 p$ degrees of freedom under the null hypothesis. We refer to the test using this statistic as the Sen and Puri (SP) method.

Mann-Kendall Based Procedures

Before describing multivariate procedures, a brief description of the univariate Mann-Kendall test for trend is in order. Suppose that $\{y_i \mid i = 1(1)n\}$ represents a sequence of observations. Under the null hypothesis, of no trend, the Mann-Kendall procedure assumes that each of the n! possible arrangements of of these n observations in time is equally likely to occur.

Define

$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \qquad (40)$$

Under the null hypothesis

$$K = \sum_{i<j} \text{sign}\,(y_j - y_i) \qquad (41)$$

is asymptotically distributed normal with a mean of zero and a variance

$$\sigma^2 = n(n - 1)(2n + 5)/18 \qquad (42)$$

Dietz and Killeen (1981) extended these results to the multivariate case. Specifically, they considered a sequence of p - variate observations of the form $\{(y_{1i},\ldots,y_{pi}) \mid i = 1(1)n\}$. They were interested in testing the null hypothesis of all p sequences being randomly ordered against the alternative of a monotonic trend in at least one of these sequences. If K_j represents the Mann-Kendall statistic K calculated for sequence j, then let

$$\mathbf{K} = [K_1.,,,.K_p]^T \qquad (43)$$

represent the vector of such statistics. It was shown that **K** is asymptotically normally distributed with a zero mean and a variance-covariance matrix $\sum$ with elements

$$\sum_{gh} = \begin{cases} \sigma^2 & \text{if } g = h \\ (t_{gh} + r_{gh})/3 & \text{if } g \neq h \end{cases} \qquad (44)$$

where

$$t_{gh} = \sum_{i<j} \text{sign}[(X_{gj} - X_{gi})(X_{hj} - X_{hi})] \qquad (45)$$

and

$$r_{gh} = \sum_{i,j,k} \text{sign}([X_{gj} - X_{gi})(X_{hj} - X_{hk})]. \qquad (46)$$

To put this into the context of the problem considered in this paper, suppose we calculate a Mann-Kendall statistic K for each of the consituent by season combinations. These results give the asymptotic joint distribution of the $p = \delta\omega$ Mann-Kendall statistics.

Hirsch and Slack (1984) used these results to create a test for univariate trends. If only a single variable is considered at a single lake for ω seasons, then the sum of the ω Mann-Kendall statistics is normally distributed with a zero mean and a variance equal to the sum of all of the elements of the corresponding $\sum$. Since the effect of a negative trend in one season may be cancelled by a positive trend in another season, this test looks for overall trends across years.

The above results lead to three different extensions of univariate Seasonal Kendall statistics to multivariate cases. The asymptotic joint distribution of the Mann-Kendall statistics leads to an asymptotic joint distribution of the Seasonal Kendall statistics for each constituent. Specifically, arrange the $\delta\omega$ Mann-Kendall statistics of the vector K so that the statistic corresponding to j^{th} season and k^{th} constitutent is the f^{th} element of K where

$$f = j + (k-1)\omega \tag{47}$$

In a corresponding fashion arrange the elements of $\sum$. In addition define the matrix C with dimensions $\delta \times \delta\omega$.

Define the i,j element of C as

$$C_{ij} = \begin{matrix} 1 \text{ if } (i-1)\,\omega + 1 \leq j \leq i\,\omega \\ 0 \text{ if otherwise} \end{matrix} \qquad (48)$$

As a result

$$S = C\,K \qquad (49)$$

is a vector of Seasonal Kendall statistics (one for each constituent, summed over seasons). Furthermore, Γ is the variance covariance matrix of these statistics where

$$\Gamma = C\Sigma C^T \qquad (50)$$

Covariance Sum Test

The first of three Mann-Kendall based test statistics is an extension of the univariate Seasonal Kendall test. Under the null hypothesis, S is normally distributed with a zero mean and a variance covariance matrix Γ. Hence, under the null hypothesis 1^TS (1 indicates a vector with each element equal to one) is normally distributed with a zero mean and a variance of $1^T\Gamma 1$. Thus, the null hypothesis may be rejected if

$$z = 1^TS/1^T\Gamma 1 \qquad (51)$$

is large when compared to a standard normal distribution. Lettenmaier (1988) referred to this test as the "covariance sum" (CS) method and pointed out that the obvious problem

with this procedure is lack of power when the trends are of different signs.

Covariance Inversion Test

The second method is based on the work of Dietz and Killeen (1981) who suggested that

$$\psi = S^T \Gamma^{-1} S \tag{52}$$

is asymptotically distributed chi-squared with δ degrees of freedom provided Γ is of full rank. Otherwise, if the rank of Γ is $q < \delta$, then

$$\psi = S^T \Gamma^{-1} S \tag{53}$$

is asymptotically chi-squared distributed with q degrees of freedom where Γ^{-1} is a generalized inverse of Γ. Consequently, the null hypothesis may be rejected if ψ is large when compared to the appropriate chi-squared distribution. We shall follow Lettenmaier (1988) in referring to this test as the "covariance inversion" (CI) method.

Covariance Eigenvalue Test

The final method is due to Lettermaier (1988) who suggested that the step of inverting the matrix Γ as required in the CI test resulted in a test of low power. The method is derived again here for the sake of completeness. Here the distribution of

$$Q = S^T S \tag{54}$$

is considered. Johnson and Kotz (1970) show that the distribution of Q is equivalent to the distribution of

$$Q^* = \sum_{v} \tau_v Z_v^2 \tag{55}$$

where the $\{\tau_v \mid v = 1(1)\delta\}$ are the eigenvalues of Γ^* and the $\{Z_v \mid v = 1(1)\delta\}$ are standard normal random variables. Although the distribution of Q^*, is not known, the cumulants of Q^* are given in Johnson and Kotz (1970). Specifically, the s^{th} cumulant of Q^* is given by

$$\xi_s = 2^{s-1}(s - 1)!\text{trace}(\Gamma^{*s}) \tag{56}$$

(Actually, this is an asymptotic result since Γ^* is an asymptotic approximation of the variance-covariance matrix of S.)

These cumulants may be used to approximate the distribution of Q^* and therefore, Q. Johnson and Kotz (1970) suggested approximating Q^* by $\beta X^2(v,\eta)$ where $X^2(v,\eta)$ represents a non-central chi-square distribution with v degrees of freedom and a non-centrality parameter η whenever

$$2[\mu_2]^2 - \mu_1'\mu_3 \geq 0 \tag{57}$$

where μ_1' indicates the first moment and μ_2 and μ_3 indicate the second and third central moments, respectively. Note

that $\mu_1' = \xi_1$, $\mu_2 = \xi_2$ and $\mu_3 = \xi_3$, and as a result an equivalent expression of (57) is

$$2[\xi_2]^2 - \xi_1\xi_3 \geq 0 \tag{58}$$

This approximation is accomplished by equating the first three moments of Q^* with those of $\beta X^2(v,\eta)$ and solving for the values of β, v and η. This gives

$$\beta = 1/2\xi_3(2\xi_2 + \sqrt{4\xi_2 - 2\xi_1\xi_3})^{-1} \tag{59}$$

$$v = \frac{1}{2\beta}(4\xi_1 - \xi_2/\beta) \tag{60}$$

$$\eta = \frac{1}{4\beta}(\xi_2/\beta - 2\xi_1) \tag{61}$$

If (57) does not hold, then it is suggested that the distribution of Q^* should be approximated by $\alpha + \beta X^2(v)$ where $X^2(v)$ is a central chi-square distribution and the values of α, β and v are determined by equating the first three moments of Q^* with those of a random variable with a $\alpha + \beta X^2(v)$ distribution. This leads to

$$\beta = \frac{\xi_3}{4\xi_2} \tag{62}$$

$$\alpha = \xi_1 - 2\xi_2^2/\xi_3 \tag{63}$$

$$v = 8\xi_2^3/\xi_3^2 \tag{64}$$

We shall again follow Lettenmaier (1988) and refer to this method as the covariance eigenvalue (CE) test.

Modified Mann-Kendall Tests

All three of the Mann-Kendall based tests (CS, CI and CE) account for between-season correlation via the variance-covariance matrix Σ in Equation 44. However our formulations of the MANOVA and SP methods assume independence between seasons by using Model 2 instead of Model 1. We can, however, easily construct modified versions of the Mann-Kendall based tests which effectively assume independence between seasons and, therefore, more closely parallel the MANOVA and SP methods. We would, of course, expect these tests to have difficulty maintaining their nominal significance levels when observations are serially correlated.

The modified tests are performed simply by setting the appropriate off-diagonal elements of Σ equal to zero and proceeding as in the original test. The "appropriate" elements are those corresponding to between season covariances. The elements corresponding to between constituent covariance for the same season are left as is. We denote modified versions of the covariance eigenvalue and covariance inversion tests as MCE and MCI, respectively. We do not explore a modified version of the CS test due to its limitations as mentioned above.

MONTE CARLO STUDY

Our Monte Carlo study of the candidate trend tests simulated three water quality variables, each observed four times per year at a single location. Data were generated using Model 1 (Equation 1) discussed earlier. Trend slopes, given by the vector β, were set equal across seasons for a given water quality variable but could differ across water quality variables. Five combinations of trend slopes among the three variables were considered as follows:

$$1:1:1,\ 1:0:0,\ 1:-1:0,\ 1:1:-1,\ 1:\tfrac{1}{2}:0$$

The first combination consists of all slopes equal in magnitude and in the same direction. The second combination consists of one positive slope and two zero slopes and so on.

Random Error Generation

Normally distributed error vectors, ε, were generated for the eight correlation structures shown in Table 1. Log normal error vectors were generated using only correlation structures 1, 2 and 5.

The random error generator model had the following multivariate AR(1) form:

$$\varepsilon_i = D\ \varepsilon_{i-1} + B^T Z_i \qquad (65)$$

where ε_i is the vector of the twelve error terms for year i and Z_i is a vector of twelve independent random errors with unit variance and zero mean. The matrices D and B are both of dimension 12 x 12 and contain fixed elements which depend on the covariance and cross-covariance matrices of ε_i as follows (Salas et al., 1980).

Let

$$M_0 = E[\varepsilon_i, \varepsilon_i]^T \tag{66}$$

and

$$M_1 = E[\varepsilon_{i-1}, \varepsilon_i]^T \tag{67}$$

The matrix M_0 defines covariances between observations in the same year while M_1 defines covariances between observations one year apart. Now equations (65), (66) and (67) yield the following

$$D = M_1{}^T M_0{}^{-1} \tag{68}$$

and

$$B^T B = M_0 - M_1{}^T M_0{}^{-1} M_1. \tag{69}$$

Since M_0 is a covariance matrix it must be positive definite and since $M_0 - M_1{}^T M_0{}^{-1} M_1$ must be factored into the product of a matrix and its transpose, it must be non-negative. These conditions impose limitations on our chosen combinations of intervariable correlations for the same season and lag-one

autocorrelation. The value of $\rho = 0.34$ in Table 1 represents the upper limit for the case of intervariable correlations = 0.50.

Comparison of Empirical Powers

Empirical power curves, showing rejection of the null hypothesis *versus* trend slope, were constructed by performing each test on 500 synthetic data sequences, each 10 or 20 years in length, for each slope. For each slope combination, the plotted slope is the largest of the three. For each curve, slopes ranged from zero (providing an empirical significance level) to that which produced a power of at least 90 percent for the most powerful test. Preliminary trials were used to determine the desired range of slopes.

In nearly all cases, the power curves did not cross, i.e., the same tests had higher powers over the entire range of slopes. The power differences between tests were often quite large. However, for those cases where differences were less obvious, we needed a consistent, if hueristic, means of comparing test performance and compiling tables of results over the range of situations studied.

We concluded, somewhat subjectively, that 80 percent power is a point at which a trend test becomes useful in a practical sense and that a difference in power of 5 percent between tests could be important to users.

Consequently, we visually compared power curves at the point where the more powerful of two tests being compared crossed 80 percent power (400 rejections). We recorded a significant difference between tests only if the two curves differed by 25 or more rejections (5 percent power).

The "critical" difference of 25 rejections corresponds to two standard deviations of the difference $(p_1 - p_2)$ between binomial proportions, each estimated from 500 independent trials when the true value of p is 0.80. More rigorous tests for comparing proportions would be possible at specific slopes (Snedecor and Cochran, 1980). However, different sets of slopes were used in the various trials. Our heuristic method provides comparisons in the same power range for all cases and is adequate for identifying general patterns in relative performance of the tests.

Normal Independent Errors

Table 2 presents a comparison of four alternative tests using normal independent observations and correlation structures 1 and 2 from Table 1. The SP, MCI and MCE tests were each compared to MANOVA, the test which would logically perform best since it most closely matches the model used in the simulations. With a few exceptions, the Mann-Kendall based tests were less powerful than MANOVA. The SP test, however, was very near in power to MANOVA over the entire range of slopes. All four tests were conservative with

regard to significance level in every case. Four examples of the ten sets of power curves are shown in Figure 1.

Lognormal Independent Errors

Table 3 presents a comparison of the same four tests under lognormal independent observations. Now we use the MCI test as a standard of comparison since MANOVA no longer matches the simulation model. Under these conditions, MANOVA has low power compared to the Mann-Kendall based tests. However, the SP test performed as well or better than MCI and MCE. The MCI and MCE tests were comparable. Although the MCE test was better in the 1:1:1 cases, the simpler MCI test was slightly better overall. All four tests were again conservative with regard to significance level. Figure 2 presents four sets of power curves paralleling those in Figure 1.

Serially Correlated Errors

Table 4 compares the two Mann-Kendall tests, CI and CE, which consider between-season correlation. Six of the correlation structures shown in Table 1 were used with normally distributed errors. Intervariable correlations were 0.0, 0.2 and 0.5, and lag-one correlations were 0.0, 0.2 and 0.34. For lognormal errors only correlation structure 5 was used with intervariable correlations set at 0.5 and lag-one correlation of 0.20.

Our evaluation of the CI and CE tests was somewhat different from that of Lettenmaier (1988) in which equal between-season correlations were used rather than an AR(1) structure. As noted before the CI and CE tests do not account for correlation between observations in different years. However, Hirsch and Slack (1984) showed that in the univariate case, accounting for between-season correlation within a year could also cope with moderate levels of ARMA(1,1) serial dependence.

In all cases displayed in Table 4, the CE test was more powerful than the CI test for ten year data records, but the two tests performed about the same overall for 20-year records.

Figure 3 contains power curves for four example situations including one, 3b, where the CI test was significantly more powerful than the CE test. This situation was encountered infrequently, however. Thus our results tend to confirm Lettenmaier's (1988) assertion that the CE test represents a general improvement over CI.

The covariance sum (CS) or Seasonal Kendall test is included in the plots to reinforce the point made earlier that the test has very little power when trends of the individual variables are in opposing directions. Thus the test would be useful only in testing for homogeneous trends.

The SP test was also included in Figure 3 to illustrate that, as expected, its empirical signifiance level is much larger than the nominal level when the error vectors are

serially correlated. Both the CI and CE tests are conservative at the levels of serial correlation studied ($\rho \leq 0.34$).

Usefulness of Tests Assuming Serial Independence

Since even quarterly water quality data can be serially correlated, especially for ground water and lakes, it would be convenient to use only tests, such as CI and CE, which account for between-season correlation. However, there is a significant loss in power associated with these tests compared to these which assume independence across seasons. This is especially true for short data records. To illustrate this point, in Figure 4 we present selected comparisons of SP and MCI tests with CI and CE, all applied to serially independent data. Similar power differences are observed in most other situations.

Hirsch and Slack (1984) discuss this issue for the univariate case. They recommend that the original Seasonal Kendall test, which assumes independence across seasons, be used as a screening tool and that the corrected (for between-season correlation) test would be most useful for long records.

A similar philosophy might be appropriate in the multivarate case. Data records could initially be tested with the SP, MCI or MCE test. If the null hypothesis were rejected, then the CE test could be used. If the null hypothesis were accepted in the CE test, the "trend"

detected by the first test might be the result of serial correlation rather than a deterministic change in the mean. However, this is only one of many possible data analysis strategies, and we wish to reserve further discussion of such applications-oriented issues for the future.

Multiple Tests

When multiple constituents are observed during several seasons, the object may be to identify a trend in at least one of the constituents for at least one of the seasons. Individual hypothesis tests of trend are characterized by their size (significance level) and power for testing their particular hypothesis, and their interpretation is dependent on those characteristics. When multiple hypothesis tests are performed, the size and power of the combination of tests is usually indeterminate; therefore, conclusions formed from a combination of tests no longer carry the same interpretation as conclusions from the individual tests.

Multivariate tests provide a framework for combining information from multiple tests. The multivariate test of no trend in any variable for any season has identifiable size and power, but does not specify a particular constituent or season as having a trend. A significant result in the multivariate test may be followed by individual tests to identify trends in constituents and/or seasons. The size of the individual tests will then be controlled by the fact they are contingent on the

multivariate test. This procedure has the advantage of providing some overall control to the individual tests. However, it must be recognized that power of those individual tests is reduced, because a trend in one constituent may be obscured by its inclusion with others having no trend. Also, if a large number of constituents are considered, a large trend in one constituent may cause a significant result in the multivariate test, effectively removing error control provided by the multivariate test to the other constituents.

SUMMARY AND CONCLUSIONS

A number of methods are available for testing water quality time series for multivariate trend. We examined several of these and compared their performance under a Monte Carlo testing program, simulating linear trends. Two of the alternatives, the Sen and Puri (SP) and MANOVA tests, are based on linear models. The SP test is an aligned rank order method of estimation and testing for linear models based on Sen and Puri (1977). The test is asymptotically distribution free. We also considered three multivariate extensions of the Mann-Kendall test--the covariance sum (CS), covariance inversion (CI), and covariance eigenvalue (CE) tests discussed by Lettenmaier (1988)--along with "modified" versions of the CI and CE tests which assume independence across seasons. The CI, CE and CS tests

account for between-season correlation, providing a means to at least partially account for serial correlation.

The SP and MANOVA procedures provide estimates of trend slope directly while the other procedures do not. However, we did not explore the properties of these slope estimators.

For independent observations we found that the SP test was robust and efficient for 10 and 20 year data records. SP performed as well as MANOVA for normal data and as well as the Mann-Kendall based tests, MCI and MCE, for log normal data. Of the latter, MCI is the logical choice since it is simpler than MCE and has a slight edge in overall power performance.

For serially correlated series, the CE test appears to be the best choice all around, although its power advantage over CI disappeared for 20 year data records. The CE and CI tests were conservative under the correlation structures studied. It remains, however, to investigate the empirical significance levels of these tests under stronger levels of serial dependence than our "worst" case of AR(1) with parameter $\rho = 0.34$. The covariance sum test is, again, useful only for detecting homogeneous trends.

ACKNOWLEDGMENTS

The work described in this article was funded in part by the Colorado Experiment Station Project No. 349 and by the Environmental Protection Agency Contract No. 68-03-3439. The research has not been subjected to review by the agency, and no official endorsement should be inferred.

The authors are grateful to Dr. Dennis Lettenmaier who, through personal communication, provided much useful background information on the development of the covariance eigenvalue methods.

REFERENCES

Anderson, T. W. 1984. An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, New York.

Aubuchon, J. C. and T. P. Hettmansperger. 1984. On the the use of rank tests and estimates in the linear model. In Handbook of Statistics, Vol. 4. P.R. Krishnaiah and P.K. Sen, eds. Elsevier Science Publishers. pp. 259-274.

Berryman, D., B. Bobee, D. Cluis, and J. Haemmerli. 1988. Nonparametric tests for trend detection in water quality time series. Water Resources Bulletin. 24(3):545-556.

Conover, W. J. 1980. Practical Nonparametric Statistics. Second Edition. John Wiley and Sons, New York.

Dietz, E. J. and T. J. Killeen. 1981. A nonparametric multivariate test for monotone trend with pharmaceutical applications. J. American Statistical Assoc. 76:169-174.

Dowel, M. and P. Jarratt. 1971. A modified regula falsi method for computing the root of an equation. BIT 11, 168-174.

Hirsch, R. M. and J. R. Slack. 1984. A nonparametric test for seasonal data with serial dependence. Water Resources Research. 29(6):727-732.

Hirsch, R. M., J. R. Slack and P. A. Smith. 1982. Techniques of trend analysis for monthly water quality data. Water Resources Research. 18(1):107-121.

Hughes, J. P. and J. P. Millard. 1988. A tau-like test for trend in the presence of multiple sensoring levels. Water Resources Bulletin. 24(3):521-532.

Jaeckel, L. A. 1972. Estimating regression coefficients by minimizing the dispersion of the residuals. The Annals of Mathematical Statistics. 43(5):1449-1458.

Johnson, N. L. and S. Kotz. 1970. Continuous Univariate Distributions - 2. John Wiley and Sons, New York.

Kendall, M. G. 1975. Rank Correlation Methods. Fourth Edition. Charles Griffin, London.

Lettenmaier, D. O. 1988. Multivariate nonparametric tests for trend in water quality. Water Resources Bulletin. 24(3):505-512.

Mann, H. B. 1945. Non-parametric tests against trend. Econometrica. 13:245-259.

McLeod, A. I., K. W. Hipel and F. Comancho. 1983. Trend assessment of water quality time series. 19(4):537-547.

Montgomery, P. H. and K. H. Rockhow. 1984. Techniques for detecting trends in lake water quality. Water Resources Bulletin. 20(1):43-52.

Neter, W. E. and J. A. Wasserman. 1974. Applied Linear Statitical Models. Irwin, Inc., New York.

Salas, J. D., J. W. Delleur, V. Yevjevich, and W. L. Lane. 1980. Applied Modeling of Hydrologic Time Series. Water Resources Publications, Littleton, Colorado.

Sen, P. K. and M. L. Puri. 1977. Assymptotically distribution-free aligned rank order tests for composite hypotheses for general multivariate linear models. Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete. 39:187-195.

Smith, R. A., R. B. Alexander and M. G. Wolman. 1987. Water quality trends in the Nation's rivers. Science. 235:1607-1615.

Snedecor, G. W. and W. G. Cochran. 1980. Statistical Methods. The Iowa State University Press, Ames, Iowa.

Taylor, C. H. and J. C. Loftis. 1989. Testing for trend in lake and ground water quality time series. Water Resources Bulletin. 25(4):715-726.

TABLE 1. Alternative correlation structures for error vectors ϵ_i in Monte Carlo study of trend detection powers.

Correlation structure number	Description (all correlations are zero unless stated otherwise)
1.	corr between all variables within a season = 0.2 (lognormal and normal)
2.	corr between all variables within a season = 0.5 (lognormal and normal)
3.	corr between all variables within a season = 0.5 AR(1) lag 1 autocorrelation between seasons = 0.2
4.	corr between all variables within a season = 0.2 AR(1) lag 1 autocorrelation between seasons = 0.2
5.	corr between all variables within a season = 0.5 AR(1) lag 1 autocorrelation between seasons = 0.34
6.	corr between all variables within a season = 0.2 AR(1) lag 1 autocorrelation between seasons = 0.34

TABLE 2. Power comparison of Sen and Puri (SP), modified covariance inversion (MCI), and modified covariance eigenvalue (MCE) tests relative to MANOVA. A (+) or (-) indicates that a given test produced at least 25 more (+) or fewer (-) rejections than MANOVA at 80 percent power. A (0) indicates that the test differed from MANOVA by fewer than 25 rejections. All trials were for normal, independent data using correlation structures 1 and 2 from Table 1 and 500 replications.

Record length (yrs)	Corr. str.	Test	Slope combination 1:1:1	1:0:0	1:-1:0	1:1:-1	$1:\frac{1}{2}:0$
10	1	SP	0	0	0	0	0
		MCI	0	-	-	-	-
		MCE	0	-	-	-	-
20	1	SP	0	0	0	0	0
		MCI	0	0	-	-	-
		MCE	0	0	-	-	-
10	2	SP	0	0	0	0	0
		MCI	0	-	-	0	-
		MCE	+	-	-	-	-
20	2	SP	0	0	0	0	0
		MCI	0	-	-	0	-
		MCE	+	-	-	-	-

TABLE 3. Power comparison of Sen and Puri (SP), modified covariance eigenvalue (MCE), and MANOVA relative to modified covariance inversion (MCI). A (+) or (-) indicates that a given test produced at least 25 more (+) or fewer (-) rejections than MCI at 80 percent power. A (0) indicates that the test differed from MCI by fewer than 25 rejections. All trials were for lognormal independent observations using correlation structures 1 and 2 from Table 1 and 500 replications.

Record length (yrs)	Corr. str.	Test	Slope combinatons 1:1:1	1:0:0	1:-1:0	1:1:-1	$1:\frac{1}{2}:0$
10	1	SP	0	0	0	0	0
		MCE	+	0	0	0	0
		MANOVA	-	-	-	-	-
20	1	SP	0	0	0	0	0
		MCE	+	-	-	-	-
		MANOVA	-	-	-	-	-
10	2	SP	0	0	0	0	0
		MCE	+	-	-	-	0
		MANOVA	-	-	-	-	-
20	2	SP	0	0	0	0	0
		MCE	+	-	-	-	-
		MANOVA	-	-	-	-	-

TABLE 4. Power comparison of covariance eigenvalue (CE), test relative to covariance inversion (CI) test. A (+) or (-) indicates that the CE test produced at least 25 more (+) or fewer (-) rejections than CI of 80 percent power. A (0) indicates that CI and CE differed by fewer than 25 rejections. Both normal (N) and lognormal (LN) errors were evaluated. Correlation structures indicated correspond to those listed in Table 1.

Record length (yrs)	Corr. str.	Test	Slope combinatons 1:1:1	1:0:0	1:-1:0	1:1:-1	$1:\frac{1}{2}:0$
10	1	N	+	+	+	+	+
20	1	N	+	0	0	0	0
10	2	N	+	+	+	+	+
20	2	N	+	-	-	-	0
10	5	N	+	+	+	+	+
20	5	N	+	0	-	0	0
10	6	N	+	+	+	+	+
20	6	N	+	0	0	0	+
10	7	N	+	+	+	+	+
20	7	N	-	0	-	0	0
10	8	N	+	+	+	+	+
20	8	N	+	0	0	0	+
10	5	LN	+	+	+	+	+
20	5	LN	+	0	0	0	0

LIST OF FIGURES

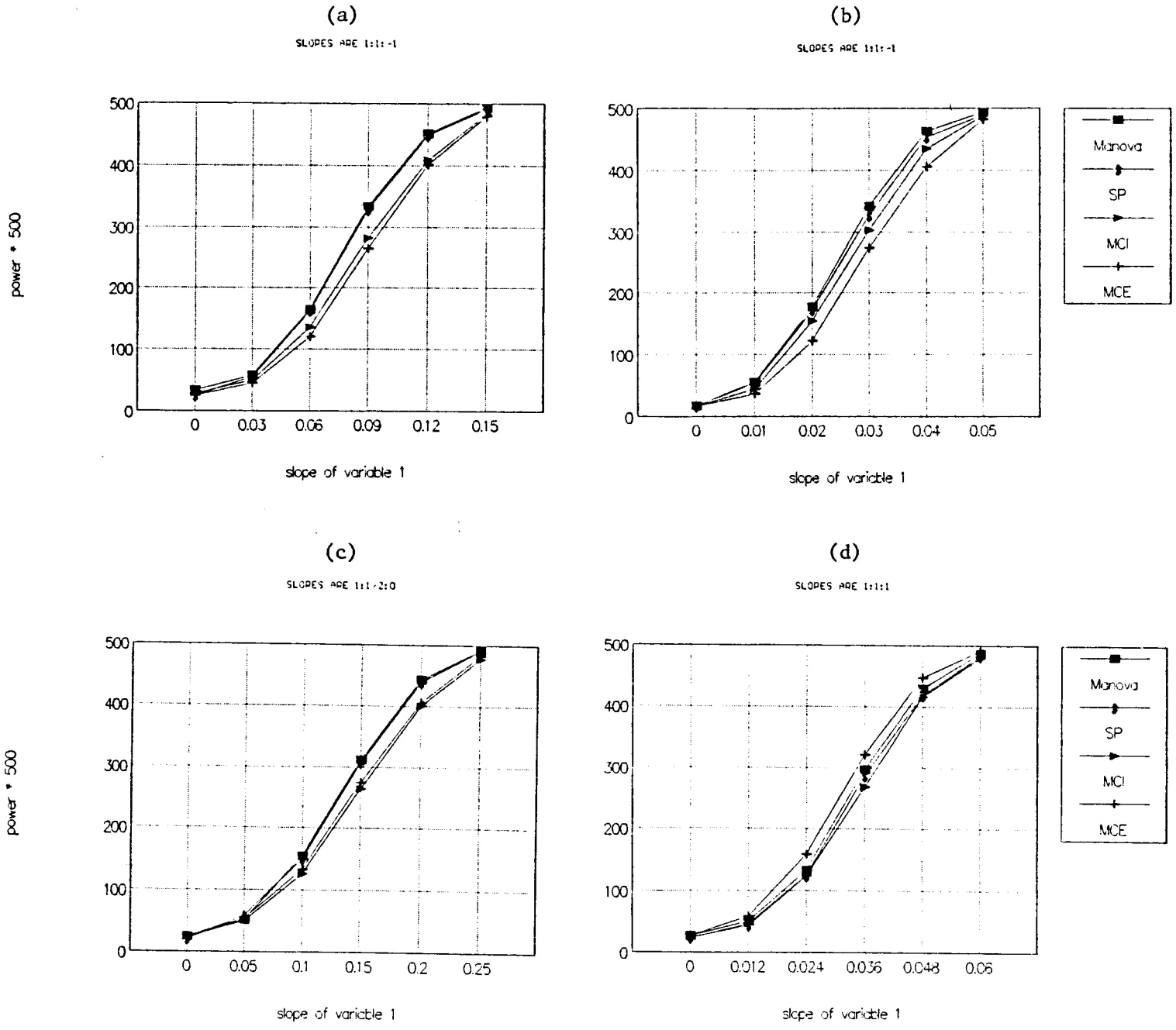

(a)
SLOPES ARE 1:1:-1
power * 500
500
400
300
200
100
0
0
0.03
0.06
0.09
0.12
0.15
slope of variable 1
(b)
SLOPES ARE 1:1:-1
500
400
300
200
100
0
0
0.01
0.02
0.03
0.04
0.05
slope of variable 1
Manova
SP
MCI
MCE
(c)
SLOPES ARE 1:1:2:0
power * 500
500
400
300
200
100
0
0
0.05
0.1
0.15
0.2
0.25
slope of variable 1
(d)
SLOPES ARE 1:1:1
500
400
300
200
100
0
0
0.012
0.024
0.036
0.048
0.06
slope of variable 1
Manova
SP
MCI
MCE

Figure 1.

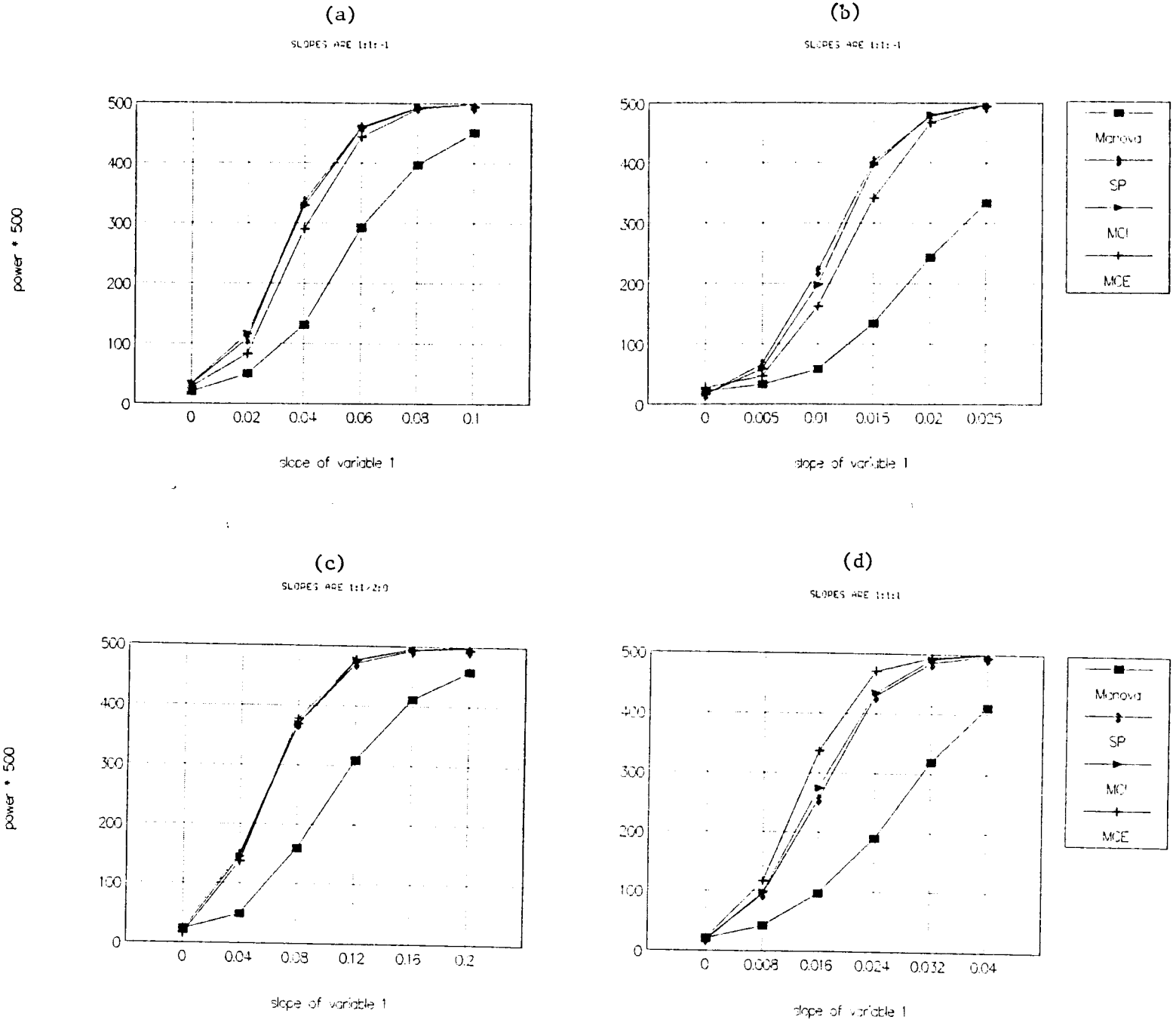

(a)
SLOPES ARE 1:1:-1
(b)
SLOPES ARE 1:1:-1
(c)
SLOPES ARE 1:1/2:0
(d)
SLOPES ARE 1:1:1
power * 500
slope of variable 1
Manova
SP
MCI
MCE
0
100
200
300
400
500
0.02
0.04
0.06
0.08
0.1
0.005
0.01
0.015
0.02
0.025
0.12
0.16
0.2
0.008
0.016
0.024
0.032
0.04

Figure 2.

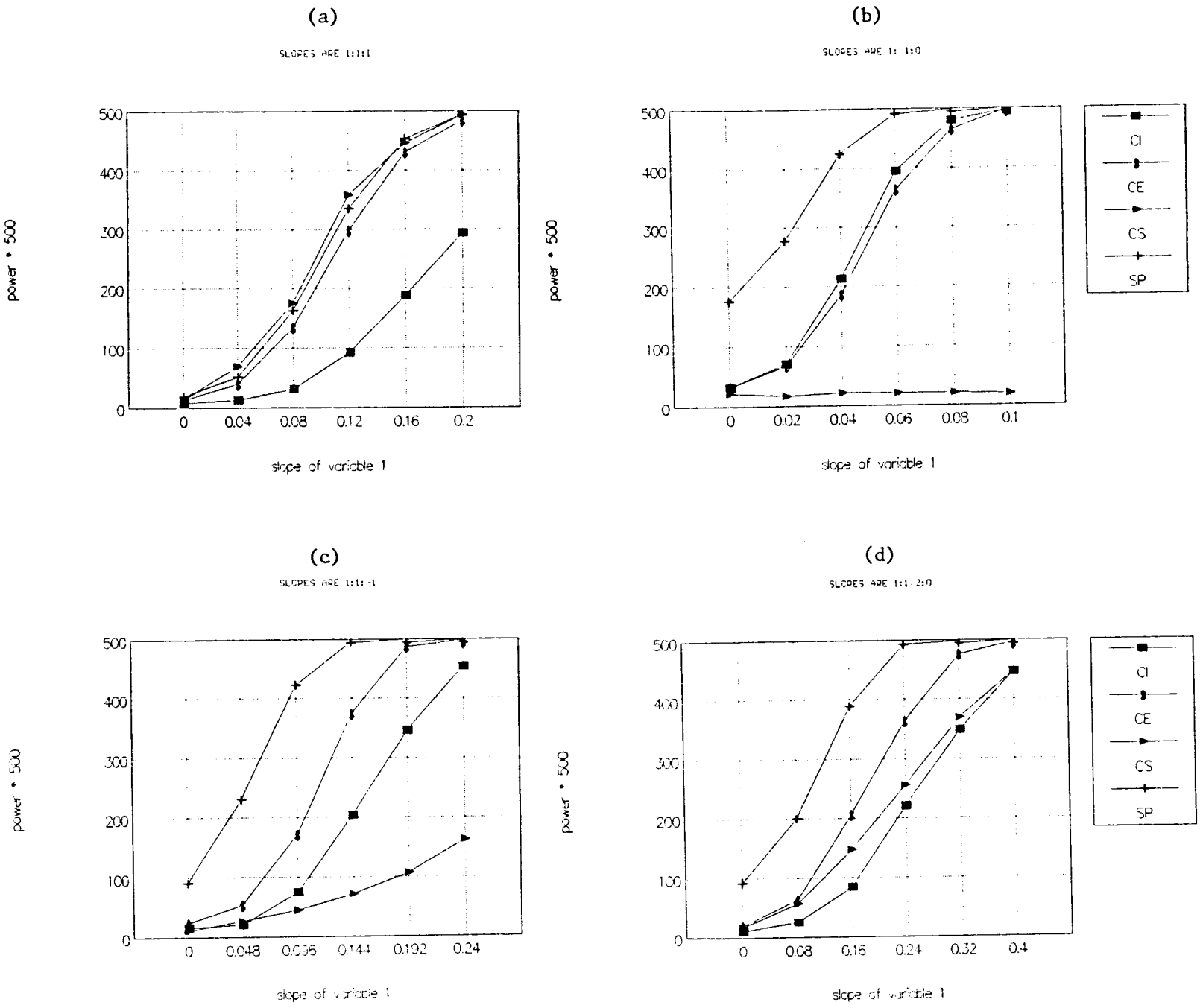

(a)
SLOPES ARE 1:1:1
power * 500
slope of variable 1
(b)
SLOPES ARE 1: 1:0
power * 500
slope of variable 1
CI
CE
CS
SP
(c)
SLOPES ARE 1:1: -1
power * 500
slope of variable 1
(d)
SLOPES ARE 1:1: 2:0
power * 500
slope of variable 1
CI
CE
CS
SP

Figure 3.

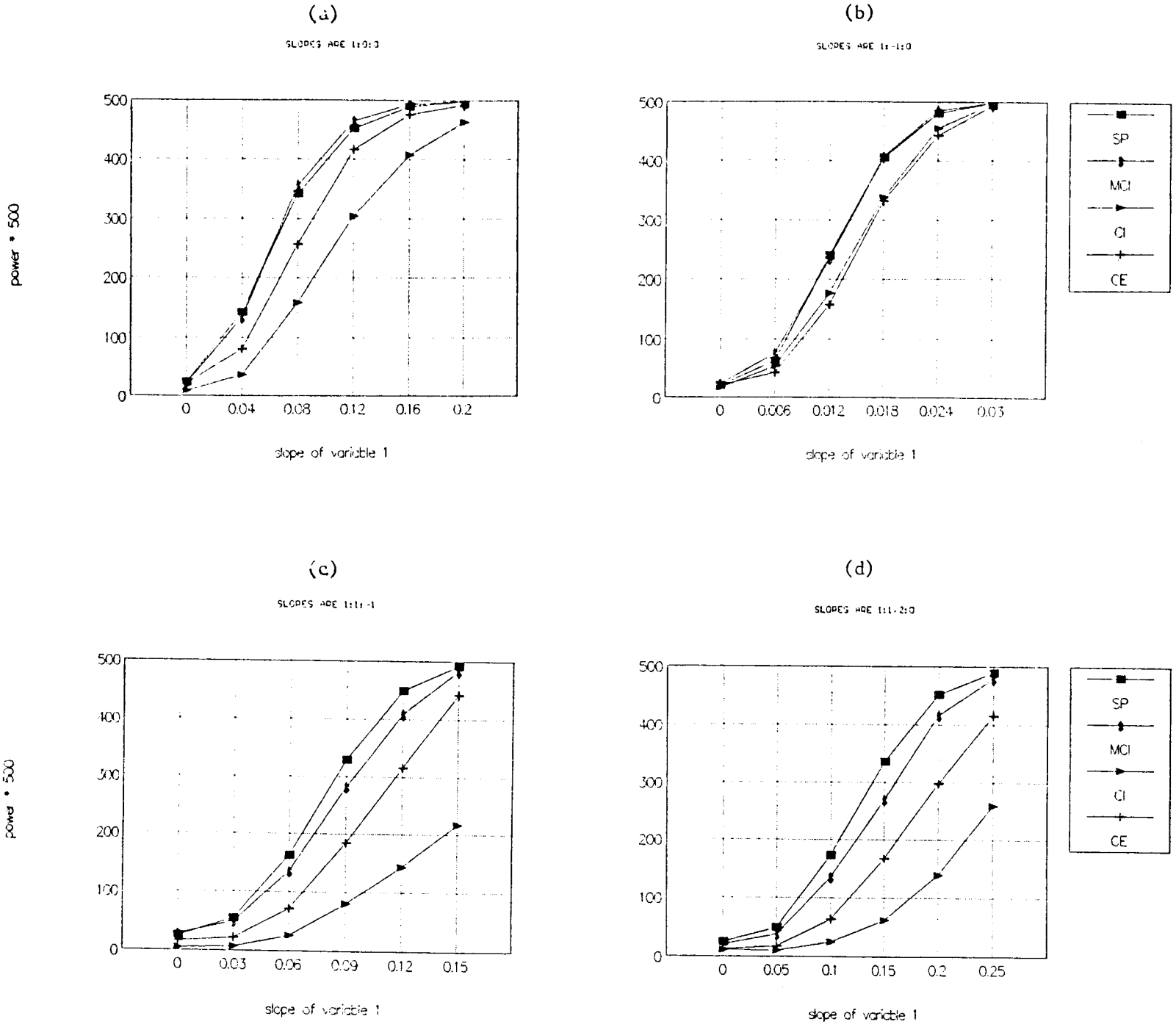

(a)
SLOPES ARE 1:0:0
(b)
SLOPES ARE 1:-1:0
(c)
SLOPES ARE 1:1:-1
(d)
SLOPES ARE 1:1.2:0
power * 500
slope of variable 1
0
0.04
0.08
0.12
0.16
0.2
0.006
0.012
0.018
0.024
0.03
0.09
0.15
0.05
0.1
0.25
100
200
300
400
500
SP
MCI
CI
CE

Figure 4.