WATER RESOURCES BULLETIN
VOL. 27, NO. 3 AMERICAN WATER RESOURCES ASSOCIATION JUNE 1991

EPA/600/J-94/510

MULTIVARIATE TREND TESTING OF LAKE WATER QUALITY[1]

Jim C. Loftis, Charles H. Taylor, Avis D. Newell, and Phillip L. Chapman[2]

ABSTRACT: Multivariate methods of trend analysis offer the potential for higher power in detecting gradual water quality changes as compared to multiple applications of univariate tests. Simulation experiments were used to investigate the power advantages of multivariate methods for both linear model and Mann-Kendall based approaches. The experiments focused on quarterly observations of three water quality variables with no serial correlation and with several different intervariable correlation structures. The multivariate methods were generally more powerful than the univariate methods, offering the greatest advantage in situations where water quality variables were positively correlated with trends in opposing directions. For illustration, both the univariate and multivariate versions of the Mann-Kendall based tests were applied to case study data from several lakes in Maine and New York which have been sampled as part of EPA's long term monitoring study of acid precipitation effects.
(KEY TERMS: trend analysis; multivariate analysis; statistics; lake water quality; acid precipitation.)

INTRODUCTION

Trend analysis is a formal approach to deciding whether an apparent change in water quality is likely due to random noise or to an actual underlying change in water quality. The use of formal statistics does not generally imply that one will be able to detect trends that are not apparent from inspection of the data. However it does mean that different data analysts will be able to reach the same conclusions, given the same data and acceptance of a common set of assumptions.

Trend analysis may be defined for present purposes as a formal test of the null hypothesis that the long-term mean(s) of a given water quality variable (or vector of variables) are not changing over time. A "significant" trend is said to exist when the null hypothesis is rejected. Whether a trend is "significant," therefore, depends on its magnitude relative to the variance of observations, how long it persists, how many observations are collected, and on the significance level of the test. Unfortunately perhaps, "significance" is not defined directly in terms of the importance of a given change relative to its impact on aquatic habitat or other beneficial uses of the water resource. A change of any magnitude will be deemed statistically significant if enough samples are collected.

Trend analysis can be applied in many water quality monitoring programs where data are collected at fixed locations over a long period of time. Long-term studies of lakes (or streams) to assess the impacts of acid precipitation are a prime example. A major goal of monitoring lakes might be to determine whether apparent changes in water quality were the result of acidic deposition. However, trend analysis, as we have defined it, cannot establish cause and effect relationships and cannot, therefore,. accomplish this objective.

Trend tests can, however, serve as screening tools, indicating the likelihood of an observed change occurring as a result of natural variation when there was, in fact, no underlying change in the mean. Those changes which cause rejection of the null hypothesis would be targeted for further study from chemical, physical, and biological standpoints to establish causal relationships.

Trend analysis has become an accepted part of many monitoring programs and has been the subject of considerable research over the past dozen years. Most applications of trend analysis deal with a single water quality variable or constituent at a single location. However, attention has recently been directed to

[1]Paper No. 90060 of the *Water Resources Bulletin*. **Discussions are open until February 1, 1992.**
[2]Respectively, Professor, Dept. of Agricultural and Chemical Engineering, 100 Engineering S., Colorado State University, Ft. Collins, Colorado 80523; Senior Research Biostatistician, Merrell Dow Research Institute, 2110 E. Galbraith Rd., Cincinnati, Ohio 45215; Scientist, Mantech Environmental Technology, Inc., 200 S.W. 35th St., Corvallis, Oregon 97333; and Associate Professor, Dept. of Statistics, Colorado State University, 101 Statistics Bldg., Ft. Collins, Colorado 80523.

analysis of multivariate trends, defined here as considering more than one constituent and/or more than one location at a time.

Multivariate approaches have two advantages over the more traditional approach of performing univariate tests on several variables for the same site and period of record. If a small, yet statistically insignificant, trend is apparent in multiple variables, then, depending on the intervariable correlation structure, a multivariate approach might detect a significant trend when the variables are considered together. In addition, the multivariate test controls the overall significance level or probability of a false trend detection. Application of several univariate tests results in an inflation of the overall significance level unless the significance levels of the individual tests are adjusted, as we shall later discuss.

Recent works (Lettenmaier, 1988; Loftis *et al.*, 1991) have suggested several methods for multivariate trend analysis, comparing their performance to formulate recommendations of which is best in a given set of circumstances. Both multivariate and univariate tests were applied to selected data sets in a study of stream quality trends across the U.S. for the period 1978-1987 (Lettenmaier *et al.*, 1991). However, these papers did not examine the fundamental question of whether multivariate methods offer an important advantage over the more traditional univariate approach. We now attempt to answer this question for a particular type of monitoring program, focusing on detection of trends in lakes characterized by seasonal sampling.

We consider two classes of methods, those based on linear models or regression and those based on rank correlation or the Mann-Kendall test for trend. A simulation study is used to compare powers of trend detection for both classes of multivariate methods against the powers of corresponding univariate methods in which the nominal significance levels are adjusted using a Bonferroni inequality. For practical illustration, we then apply univariate and multivariate methods to real data, using the Mann-Kendall based tests. The case study data were collected in a study of the long-term effects of acid deposition.

The simulation study was confined to gradual, monotonic changes, simulated as linear trends occurring over 10 or 20 years. This type of change might be encountered, for example, when acid deposition rates are increased as a result of industrialization or decreased as a result of air quality control measures. We also focus on the issue of hypothesis testing and exclude the important related question of estimating the magnitude of trend.

BACKGROUND

Linear Models and Normal Theory

The concept of multivariate trend analysis is easily introduced using linear models. Let us assume that we are interested in a set of three water quality variables, say sulfate concentration, calcium concentration, and acid neutralizing capacity (ANC). We can write a set of linear equations describing seasonal variation of the variables and linear trend over time as follows:

$$y_{1,i,j} = u_1 + B_1 i + L_{1,j} + e_{1,i,j} \tag{1}$$

$$y_{2,i,j} = u_2 + B_2 i + L_{2,j} + e_{2,i,j} \tag{2}$$

$$y_{3,i,j} = u_3 + B_3 i + L_{3,j} + e_{3,i,j} \tag{3}$$

where:

$y_{k,i,j}$ = concentration of variable k in year i, season j; k=1 for sulfate; k=2 for calcium; and k=3 for ANC; i=1,2,. . . ,N (N=number of years considered); j=1, 2, 3, 4.

u_k = mean of constituent k for season 4 in absence of trend.

$L_{k,j}$ = seasonal adjustment to the mean for constituent k and season j. $L_{k,j}$ = the mean of constituent k for season j minus the mean of constituent k for season 4. $L_{k,4} = 0$ since u_4 = the mean for season 4.

$e_{k,i,j}$ = error term with mean zero and unspecified distribution. Errors may be correlated between variables and over time (serial correlation).

B_k = time trend for constituent k in units per year.

The multivariate formulation considers the variables and model parameters as vectors. Thus we have **y, u, B, L,** and **e**, each a vector of dimension three, and the multivariate model is

$$\mathbf{y}_{i,j} = \mathbf{u} + \mathbf{B}_i + \mathbf{L}_j + \mathbf{e}_{i,j} \tag{4}$$

Other model forms are possible. For example, one could let each constituent by season combination be represented by a separate independent variable.

A univariate null hypothesis that any of the individual B_k equal zero can be tested using multiple linear regression and the analysis of covariance described in Taylor and Loftis (1989) (see Anderson, 1984, for a more general reference). The regression will also provide estimates of the parameters B_k, u_k, and $L_{k,j}$. Standard regression techniques are based on normality of the error terms $e_{k,i,j}$. Taylor and Loftis (1989) found that a simple extension of analysis of covariance using the ranks of data, as suggested by Conover (1980), provides a powerful nonparametric test. However, the classical parametric approach is used here.

When testing for trend in more than one constituent or location, the usual approach is to perform multiple univariate tests, each at some significance level, α, of say 5 percent. The probability of rejecting the null hypothesis when true is therefore no greater than 5 percent for any single test, but the probability of at least one rejection in K tests when all K null hypotheses are true may be much greater than α. If N=3 as in our example, the probability of at least one rejection must be no greater than $3\alpha-3\alpha^2-\alpha^3$, depending on the correlation between variables. With a = 0.050000, this probability is 0.142625.

To control the overall significance level (probability of at least one false rejection), one could apply a Bonferroni inequality and perform K univariate tests, each at a significance level of α/K (Snedecor and Cochran, 1980). When this approach is used with multiple applications of univariate analysis of covariance, we shall refer to the test as the ANOCOV/Bonferroni approach.

Alternatively, one can use a single multivariate test of the null hypothesis that all K of the trend slopes are zero, i.e.,

$$B_1, \ldots, B_k, \ldots, B_K = 0. \tag{5}$$

The multivariate test would reject the null hypothesis if any of the slopes were found to be different from zero.

A procedure for performing the multivariate test, referred to as MANOVA for multivariate analysis of variance, is described in Loftis *et al.* (1991), and more generally in Anderson (1984). Conveniently, the multivariate estimates of the parameters **B**, **u**, and **L**, are the same as the univariate estimates. A classical parametric procedure, MANOVA assumes normality of the error terms.

However, for the more general case of non-normal errors, robust procedures for estimation and testing of linear models have been developed. One of these, an aligned rank order procedure suggested by Sen and Puri (1977), was adapted to the water quality trend detection problem (Loftis *et al.*, 1991). Under the conditions studied, the Sen and Puri (SP) test was very nearly as powerful as MANOVA for normal data and as powerful as rank correlation methods (based on the Mann-Kendall test for trend) for lognormal data. The SP procedure is not pursued here as it is much more complicated and difficult to apply than either MANOVA or the Mann-Kendall based procedures described below.

Mann-Kendall Based Procedures

One of the most popular tests for trend in a single water quality variable at a single location is a rank correlation method called the seasonal Kendall (SK) test (Hirsch *et al.*, 1982). The study cited and another by Taylor and Loftis (1989), reported that the SK test compared favorably with available alternatives in terms of both power and actual significance level for normal and lognormal data with no serial dependence. A "corrected" form of the test (SKC), presented in Hirsch and Slack (1984) accounts for correlation between seasons and provides conservative significance levels in the presence of moderate serial correlation. The SKC test has low power compared to the original SK test when the data are serially independent, especially when record lengths are ten years or less.

To test for trend in more than one constituent and/or location at a time, one can use the test suggested by Dietz and Killeen (1981) from which the SKC test was developed. However, as observed by Lettenmaier (1988), this test has low power for record lengths shorter than about 20 years of monthly data. Lettenmaier referred to the Dietz and Killeen test as the covariance inversion (CI) method since computation of the test statistic involves inversion of the covariance matrix associated with the vector of seasonal Kendall test statistics. In order to improve the power performance of the test, Lettenmaier (1988) suggested a new approach, called the covariance eigenvalue (CE) method which avoids the matrix inversion step.

In a Monte Carlo study of the CE test using constant between season correlations, Lettenmaier observed significantly improved power compared to the CI test. Loftis *et al.* (1991), found similar improvement in power using first-order, autoregressive – AR(1) – serial correlation with AR parameter (lag-one correlation coefficient) values of 0.34 and smaller.

The latter authors also suggested modifications of the CI and CE tests to improve their power under serially independent observations. These modified tests, called MCI and MCE, are related to the original CI and CE tests in the same way that the SK test is related to the SKC test. More specifically the MCI, MCE, and SK tests assume that between season correlations for each constituent are zero. The CI, CE, and SKC tests account for this correlation. Correlations between constituents are still accounted for in the modified tests.

Loftis *et al.* (1991), found, not surprisingly, that the MCI and MCE tests were somewhat less powerful in detecting linear trends than methods based on linear models. However, the Mann-Kendall based procedures, MCI and MCE, are more general in that they do not assume a linear model and are simpler to compute. Of the two tests, MCI and MCE, neither seemed to have a strong advantage over the other in terms of power. Therefore, the simpler MCI test is suggested here as a useful multivariate alternative to the univariate seasonal Kendall test. A description of the test follows the paper as an appendix.

We should note here that the MCI test will produce seasonal Kendall test statistics for each of the water quality variables of concern, as well as an overall multivariate test statistic. Therefore, all of the information developed in the univariate approach is also produced in the multivariate approach. The individual SK statistics could be used, when the multivariate null hypothesis is rejected, to indicate which of the individual variables had significant trends.

CHARACTERISTICS OF LAKE QUALITY DATA

Before moving on to a Monte Carlo comparison of univariate and multivariate tests, let us first examine the characteristics of lake quality data. This examination should focus attention on the range of characteristics which would be most important to study in the simulation experiments.

The characteristics most likely to bear on the relative performance of univariate v. multivariate trend testing methods are the intervariable correlation structure and trend directions. The shape of the underlying distribution would also be important in choosing between parametric and nonparametric approaches. We considered both classes of methods in our simulation study. However, our case study applications were limited to nonparametric methods since non-normal random errors are common in water quality, and nonparametric methods are preferable for routine applications.

The presence of serial correlation might also affect our choice of methods. Under such conditions, the choice should be based on the costs of failing to reject the null hypothesis when it is not true versus those of rejecting the null hypothesis when it is in fact true. When data are serially correlated, but there is no trend, methods which assume independence between seasons tend to reject the null hypothesis more frequently than suggested by their nominal significance levels. On the other hand, the methods which accommodate serial correlation tend to have lower power to detect real trends.

To obtain some idea of the range of between-variable correlations and serial correlation which was likely to be encountered in practice, we examined a few data records of excellent quality from Maine and New York. The data were obtained from EPA's Long Term Monitoring (LTM) project (Newell *et al.*, 1987). This EPA project has funded the continuation of water quality monitoring in various regions since 1983. Major ion chemistry, ANC, pH, conductance, DOC and Al are measured in these lakes to monitor the effects of acidic precipitation on surface water chemistry. Data from two regions, 15 lakes in the Adirondacks, and five lakes in the Tunk Mountain Watershed of Maine are considered here for estimates of between-variable correlation and serial correlation.

The Adirondack Lakes have been monitored by Charles Driscoll at the University of Syracuse since 1982 with funding from the Electrical Power Research Institute through 1985, and EPA funding since that time (Driscoll, 1991). Adirondack lakes were sampled monthly at the outlet. In order to make the Adirondack sampling schedule resemble the seasonal schedule followed in other regions, we subsampled the monthly data, selecting the observations closest to January 15, April 15, July 15, and October 15.

The Maine sites have been monitored by the U.S. Fish and Wildlife Service in cooperation with the University of Maine since 1982, with EPA funding beginning in 1983 (Kahl *et al.*, 1991). These lakes are sampled from the epilimnion in the spring, summer, and fall.

For both the Maine and New York data sets, we studied the serial and intervariable correlation structure of the seasonal series. Our objective was to roughly establish a range of correlations which would be realistic for a Monte Carlo evaluation of trend testing methods. The data were preprocessed by removing seasonal means and an estimated linear trend component in every case prior to estimation of correlations.

Serial Correlation

Lag-one autocorrelation coefficients were estimated from the case study data after preprocessing. These estimates were then compared with standard error limits to assess their significance. Of 65 series examined, each approximately six years in length, only eight has estimated values, r_1, which were outside the range of one standard error from zero. Of these eight, two were positive and six were negative. None of the estimates fell outside plus or minus two standard errors from zero.

Since the probability of r_1 falling outside one standard error is 32 percent for a serially independent series, we would not be able to infer that serial correlation is significant in the example data. Of course, with sufficiently long record lengths, we would be able to reject a null hypothesis of zero serial correlation, even if the level of correlation were very small. We should also note that we checked for serial correlation in the original monthly series from New York and found a number of significant r_1 values, typically in the range of 0.2 to 0.5.

Intervariable Correlation

We calculated intervariable correlations for several variables in the LTM data sets. The strongest relationship was observed between calcium plus magnesium and ANC, where the median correlation was 0.50. The weakest relationship was found between calcium plus magnesium and nitrate, where the median correlation was 0.17. For ANC and sulfate, the median correlation was –0.23.

Time Trends

Logic dictates that trends of interest would include simultaneous increases in some variables and decreases in others. For example, acidification of a lake would often be accompanied by increasing sulfate concentration and decreasing ANC. Examination of the case study data sets revealed that many combinations of increases and decreases are possible. We shall see several of these later in our case study applications to real data.

Consequently, it appears that we should consider as the general case unspecified trend directions for all variables. We shall formulate tests such that the null hypothesis should be rejected when any of the variables included have a trend in either direction.

MONTE CARLO STUDY

Simulation Approach, Linear Models

Our initial simulation study considered linear model based approaches only. Quarterly data for three variables at one location were generated using the procedure described in Loftis *et al.* (1990). The following five ratios of trend slopes among the three variables were considered:

1:1:1, 1:1/2:0; 1:0:0, 1:–1:0, 1:1:–1.

In the first combination, the slopes for all three variables are equal and in the same direction; in the second combination, the second variable has a slope which is one half that of the first variable; and the third variable has no trend; and so on.

Seven different intervariable correlation structures, shown in Table 1, were generated. To limit the size of the study, only normal, serially independent errors, were considered. Our two linear model based approaches, MANOVA and ANOCOV/Bonferroni, would not perform well for skewed distributions, as shown by Loftis *et al.* (1991), for the case of lognormal errors. The same study investigated the effect of serial correlation on both linear model and Mann-Kendall based tests. However, our earlier examination of the LTM data suggests that an assumption of serial independence is justified for the present study of quarterly observations of lake quality.

Each experiment involved generating 500 trivariate sequences, each 10 or 20 years in length and having a constant slope vector **B.** In each trial, the entire sequence was tested for trend under the null hypothesis of no trend in all three variables using both MANOVA and the ANOCOV/Bonferroni methods. The empirical power at a given slope is the number of rejections of the null hypothesis divided by the number of trials.

By performing the experiment over a range of slopes, indexed by the slope of the first variable, a power curve can be drawn. The empirical significance level is equal to the power when the slope vector is the zero vector. All tests were run at a nominal significance level of 5 percent.

TABLE 1. Intervariable Correlation Structures Simulated in Monte Carlo Study.

Correlation Structure	Description
1.	corr between all variables within a season = 0.2
2.	corr between all variables within a season = 0.5
3.	corr between variables 1 and 2 within a season = 0.2 corr between variables 1 and 3 within a season = –0.2 corr between variables 2 and 3 within a season = –0.2
4.	corr between variables 1 and 2 within a season = 0.0 corr between variables 1 and 3 within a season = 0.2 corr between variables 2 and 3 within a season = 0.5
5.	corr between variables 1 and 2 within a season = 0.8 corr between variables 1 and 3 within a season = –0.2 corr between variables 2 and 3 within a season = 0.2
6.	corr between variables 1 and 2 within a season = 0.8 corr between variables 1 and 3 within a season = 0.2 corr between variables 2 and 3 within a season = 0.2
7.	same as number 1 but with seasonal variance

Results of Simulations

Table 2 presents results for the ten-year simulations. A plus or minus in the table indicates that MANOVA recorded at least 25 more or fewer, respectively, rejections than the ANOCOV/Bonferroni method at the point where the curve for the more powerful test crossed 80 percent power or 400 rejections. A difference of 25 rejections corresponds to a 5 percent difference in power.

TABLE 2. Comparison of Simulated Powers for MANOVA and ANOCOV/Bonferroni Methods of Trend Detection. A plus (minus) indicates that MANOVA produced at least 25 more (fewer) rejections than the ANOCOV/Bonferroni method at 80 percent power. All tests were run on 500 quarterly sequences, each 10 years in length with normal, serially independent data. Intervariable correlation structures are shown in Table 1.

Slope	Intervariable Correlation Structure Number						
Combination	1	2	3	4	5	6	7
1:1:1	0	–	+	0	+	0	0
1:0:0	0	+	+	0	+	+	0
1:–1:0	+	+	+	+	+	+	+
1:1:–1	+	+	–	+	+	+	+
1:1/2:0	0	+	+	+	+	0	0

This method of comparison is consistent with an interpretation of the empirical powers as estimates of binomial proportions, p. The 5 percent difference represents two standard errors of the difference between two binomial proportions, each estimated from 500 independent trials when the true value of the binomial parameter, p, is 0.80. We suggest that this method of comparison is objective and consistent, although heuristic.

MANOVA was found to have significantly greater power than the ANOCOV/Bonferroni method in 24 of the 35 cases studied and significantly lower power in two cases. In all cases, the empirical significance levels of both methods were less than or equal to nominal levels. Interestingly, the slope combination which showed the smallest advantage for MANOVA was 1:1:1. The slope combination which showed the most consistent advantage for MANOVA was 1:–1:0. The intervariable correlation structure which showed the smallest advantage for MANOVA was #1, while correlation structure #5 showed the greatest advantage.

In 20-year simulations (not shown), MANOVA had significantly higher power for the 1:1:1 slope combination with correlation structure #7. Otherwise, the results are the same as for ten years.

In Figures 1, 2, and 3, we compare empirical power curves for MANOVA and the ANOCOV/Bonferroni method, using selected slope combinations for each intervariable correlation structure. From the figures, interesting results emerge.

For correlation structure #1, the two power curves are very close together for slope combinations 1:1:1 and 1:0:0 (Figure 1a, b). However, for correlation structure #2, the ANOCOV/Bonferroni method has higher power than MANOVA for 1:1:1 slopes and lower power than MANOVA for 1:0:0 slopes (Figure 2a, b). For correlation structure #4 (Figure 3a, b), MANOVA is more powerful in both slope combinations.

In Figure 3b the Bonferroni method shows much lower power than MANOVA. This situation occurred in several other cases, but the reverse, MANOVA having much lower power, was never observed.

Overall, MANOVA has a definite advantage over the Bonferroni method. Although there are a very few cases (notably those of homogeneous trend) where the Bonferroni method has slightly higher power, these represent the exception rather than the general rule. Since it would be prudent to assume non-homogeneous trends for real situations, the multivariate approach would be the method of choice.

Mann-Kendall Based Tests

We later performed a similar simulation study of the MCI test in comparison to application of multiple seasonal Kendall (SK) tests, the latter having significance levels adjusted via the Bonferroni inequality.

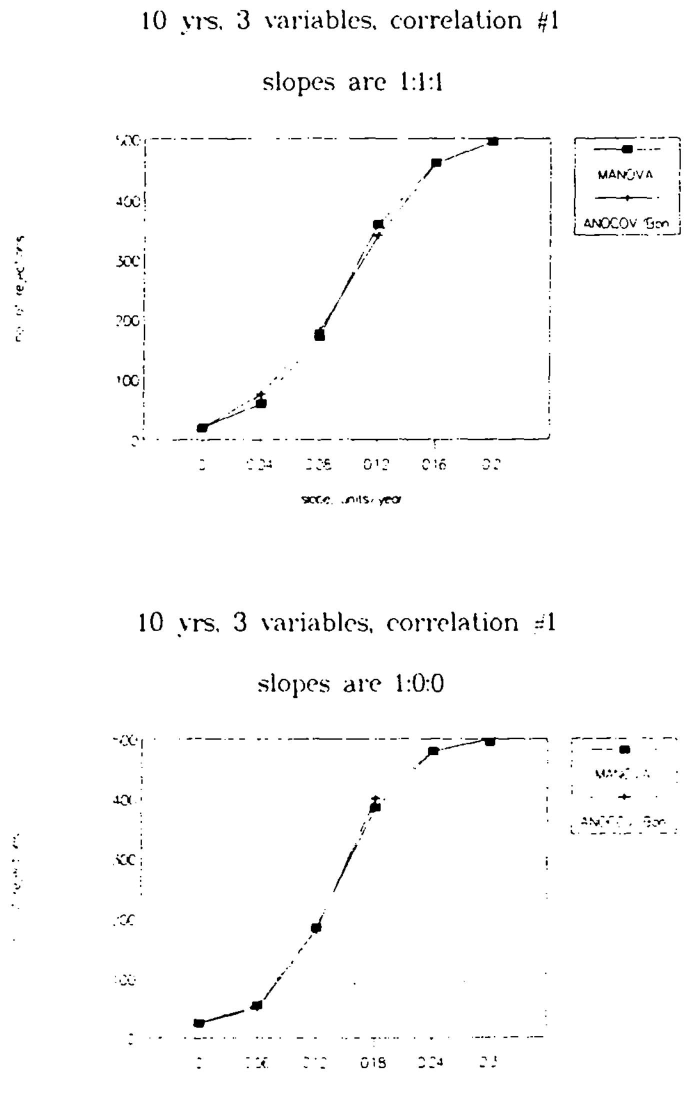

Figure 1a, b. Power Curves from Monte Carlo Comparison of MANOVA and ANOCOV/Bonferroni Methods. Each point represents 500 simulations of three variables over 10 years (quarterly) with normal data. Correlation structure #1 is described in Table 1. The slope plotted on the horizontal axis is the largest of the three and is given in units per year.

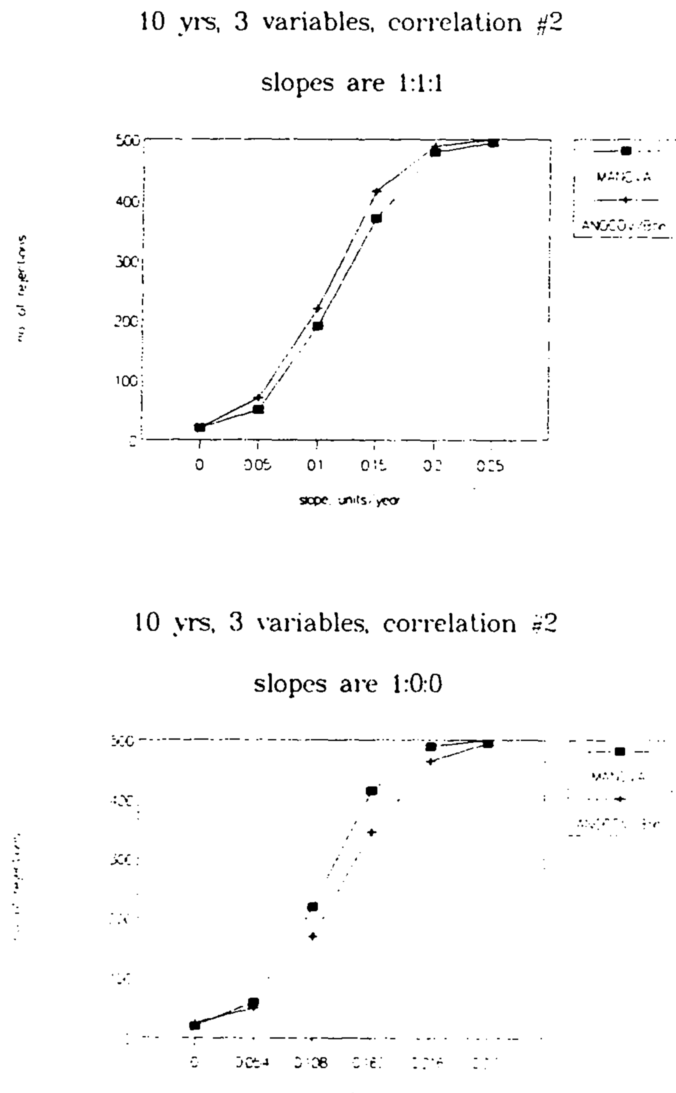

Figure 2a, b. Power Curves from Monte Carlo Comparison of MANOVA and ANOCOV/Bonferroni Methods. Each point represents 500 simulations of three variables over 10 years (quarterly) with normal data. Correlation structure #2 is described in Table 1. The slope plotted on the horizontal axis is the largest of the three and is given in units per year.

The data generation procedure was the same as that of the linear models study, again using normal errors. However, since the Mann-Kendall procedures use only the ranks or relative magnitudes of the data, exactly the same results would have been obtained using log-normal errors or any other distribution.

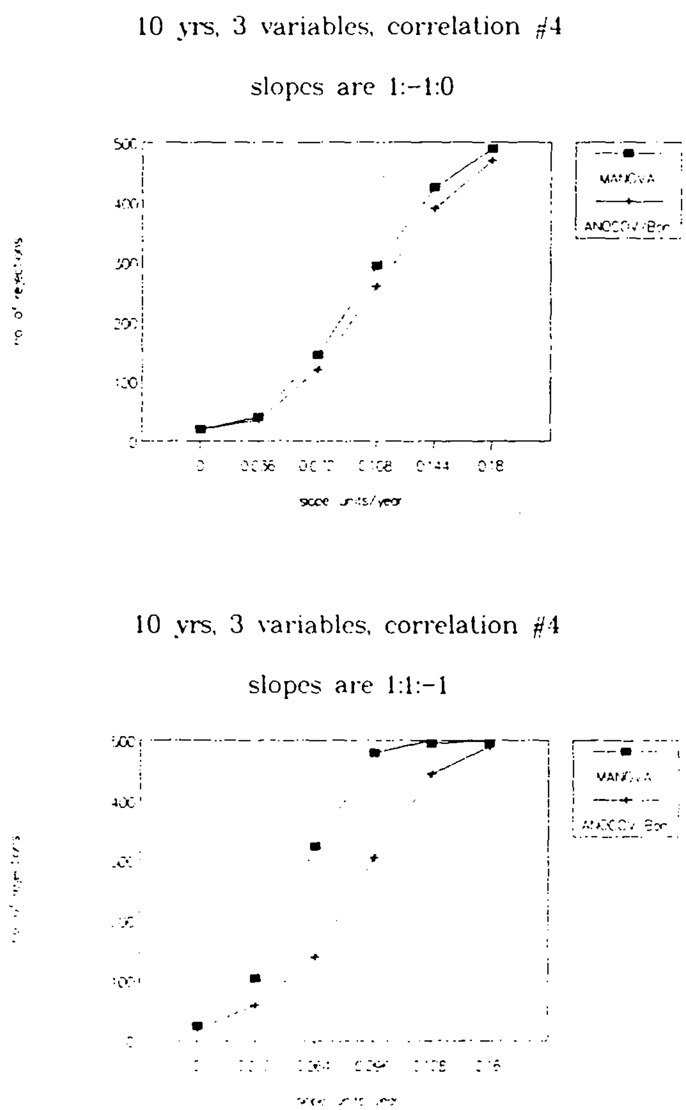

Figure 3a, b. Power Curves from Monte Carlo Comparison of MANOVA and ANOCOV/Bonferroni Methods. Each point represents 500 simulations of three variables over 10 years (quarterly) with normal data. Correlation structure #4 is described in Table 1. The slope plotted on the horizontal axis is the largest of the three and is given in units per year.

For three water quality variables at a single location, we evaluated six different slope patterns for intervariable correlation structures #1 and #2 (from Table 1) and eight slope patterns for correlation structure #6. We then extended the study to five water quality variables, considering only three slope patterns and correlation structures #1 and #2.

Results of the study are summarized in Table 3 using the same format as Table 2 for the linear model results. The results are quite similar for both the linear model and Mann-Kendall based approaches. Here again, the multivariate method has a definite power advantage overall. The MCI test had higher power by our previous criterion in 15 of 26 cases studied, while the SK/Bonferroni method had higher power in only one case.

TABLE 3. Comparison of Simulated Powers for the Modified Covariance Inversion (MCI) and SK/Bonferroni Methods of Trend Detection Using Three and Five Variables. A plus (minus) indicates that the MCI test produced at least 25 more (fewer) rejections than the SK/Bonferroni method at 80 percent power. All tests were run on 500 quarterly sequences, each 10 years in length with normal, serially independent data. Intervariable correlation structures are shown in Table 1.

Slope Combination	Intervariable Correlation Structure Number		
	1	2	3
1:1:1	0	0	0
1:0:0	0	+	+
1:0:1	0	+	+
1:-1:0	+	+	+
1:1:-1	+	+	+
1:1:0			0
1:-1:1			+
1:0:-1			+
1:1:1:1:1	0	–	
1:1:-1:-1:0	+	+	
1:-1:0:0:0	+	+	

When all of the trends are in the same direction (1:1:1 slope pattern), the two approaches give similar performance, and the Bonferroni method sometimes has a slight advantage. An example is presented in Figure 4a for five variables with correlation structure #1 and homogeneous trend. However, when some of the trends are in opposing directions, the multivariate method generally has a large power advantage. Figure 4b, for the 1:1:-1:-1:0 slope combination, is an example.

Analysis of the simulation results for correlation structure #6 indicates a relationship between power advantage and the strength of (positive) intervariable correlation. In this structure, the first and second variables have noise terms with a correlation coefficient of 0.8 while the other intervariable correlations are 0.2. Comparing the results for the 1:0:-1 and 1:-1:0 patterns in Figure 5a, b, we find that the MCI test shows a larger power advantage, compared to the

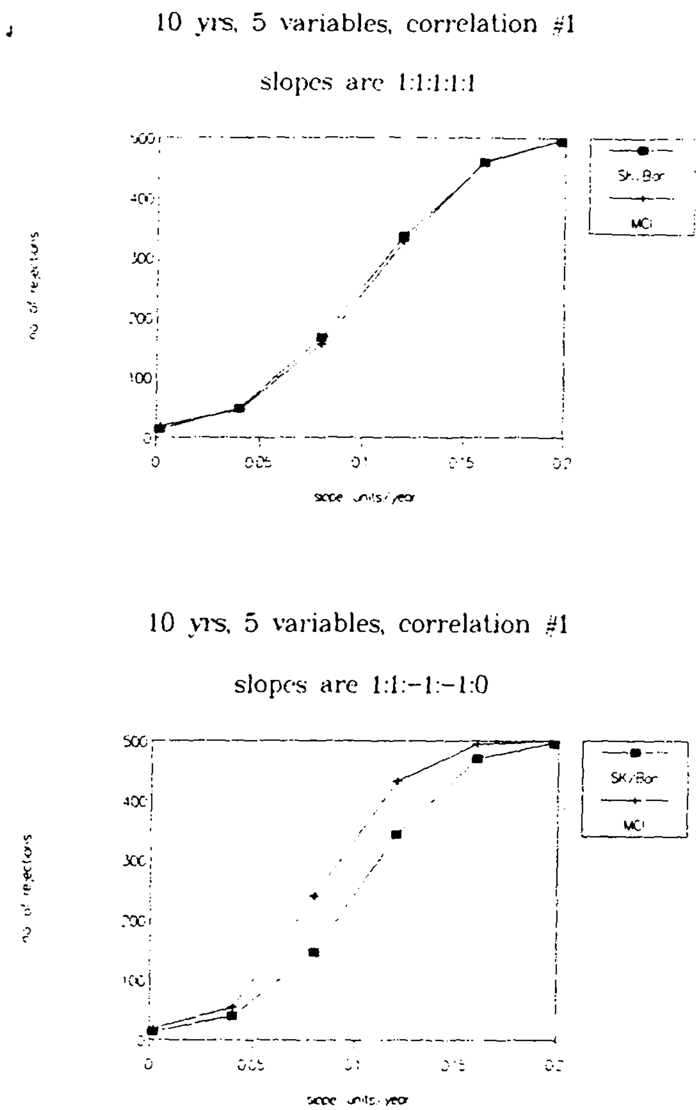

Figure 4a, b. Power Curves from Monte Carlo Comparison of MCI Test and SK/Bonferroni Test. Each point represents 500 simulations of five variables over 10 years (quarterly) with normal data. Correlation structure #1 is described in Table 1. The slope plotted on the horizontal axis is the largest of the three and is given in units per year.

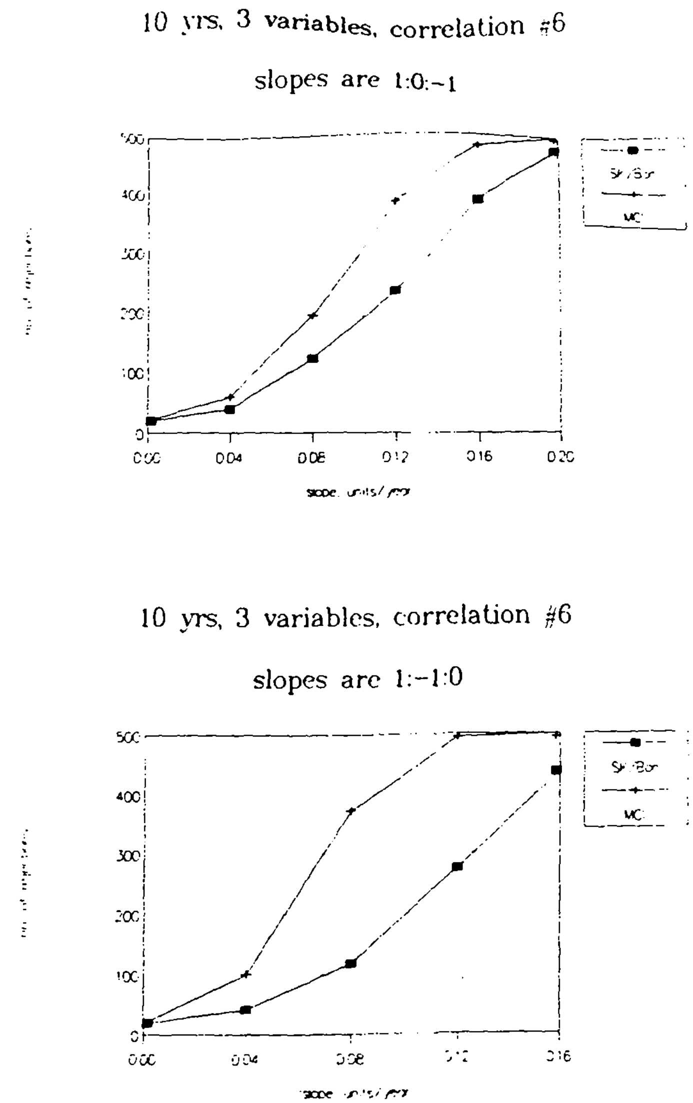

Figure 5a, b. Power Curves from Monte Carlo Comparison of MCI Test and SK/Bonferroni Test. Each point represents 500 simulations of three variables over 10 years (quarterly) with normal data. Correlation structure #6 is described in Table 1. The slope plotted on the horizontal axis is the largest of the three and is given in units per year.

SK/Bonferroni method, when opposing trends occur coincident with larger positive correlation.

APPLICATIONS TO LTM DATA

As illustrative case studies, we applied both the multivariate MCI test and univariate SK/Bonferroni tests to the LTM data from Maine and New York. These applications were performed for the sole purpose of comparing the performance of the two tests on real data. Detailed trend results and interpretation of the Adirondack and Maine data are presented in Driscoll (1991) and Kahl *et al.* (1991), respectively. The MCI test is a logical choice over MANOVA for routine applications since it is less restrictive in its assumptions regarding the distribution of the noise terms and of the functional form of the trend. In keeping with an overall monitoring objective of evaluating the effects of acid precipitation on lake water quality, we considered a set of three variables: ANC, sulfate, and calcium.

Each test considered only one location (lake) at a time and was performed over the entire length of record. The Maine data consisted of three observations per year (April, July, and October) while the New York data were a quarterly subsample (January, April, July, and October) of the original monthly series.

Unlike the simulation study, this application to real data is uncontrolled in the sense that we do not know the true values of population parameters such as intervariable correlations, serial correlation structure, or underlying distribution. Furthermore, since we do not know the true trend magnitudes, we cannot say whether our tests obtain the "correct" result in a given case. However, based on the earlier simulations, we would expect the MCI test to detect trends more often than the SK/Bonferroni test.

Table 4 presents the trend testing results. The directions of the trends in the individual variables are indicated in the three columns following the lake identifier. These were determined by seasonal Kendall tests using a nominal significance level of 0.20 to obtain high power. The last two columns give the rejection significance levels (sometimes called the "p" values) for both the MCI and SK/Bonferroni tests.

The rejection significance level for each test should be interpreted in the following way. If the test were performed at a nominal (*a priori*) significance level greater than the given rejection level, then the test would have resulted in rejection of the null hypothesis of no trend, and vice versa.

In 16 of 20 cases, the MCI test produced a rejection significance level equal to or smaller than that of the SK/Bonferroni test, suggesting that the MCI test would tend to reject the null hypothesis more often. If we choose a specific significance level, say 0.10, we see that the two tests would have given similar results. There are three cases in which the MCI test would have rejected the null hypothesis when the Bonferroni approach did not: lakes NY4, NY5, and ME1. There is one case where the reverse is true, lake NY8.

There was considerable variety in the patterns of slope directions, supporting the general assumption of non-homogeneous trend. In some cases, opposing trends were noted in ANC and sulfate concentration, and, in others, the two variables moved in the same direction over time. Figure 6a, b presents examples of the two situations. For the Maine Lake of Figure 6a, ANC and sulfate have opposing trends. In the New York Lake data shown in Figure 6b, ANC and sulfate have trends in the same direction.

TABLE 4. Results of Applying Trend Tests to Case Study LTM Data. For individual variables "+" indicates an increasing trend, "–" indicates a decreasing trend, and "0" indicates no significant trend at 0.20 significance level using a univariate SK test. Rejection significance levels are given for the SK test adjusted via the Bonferroni inequality and for the multivariate MCI (modified covariance inversion) test. All Maine lake observations are from the epilimnion. The New York data sets consisted of six years of quarterly data, and the Maine sets consisted of seven years of observations three times per year.

Lake ID Number	Variables: ANC	Variables: Sulfate	Variables: Calcium	Rejection Significance Levels: SK/Bonferroni	Rejection Significance Levels: MCI
			NEW YORK LAKES		
NY1	–	–	–	0.019	0.021
NY2	0	–	+	0.418	0.116
NY3	0	–	+	0.034	0.026
NY4	–	–	+	0.146	0.067
NY5	–	–	0	0.146	0.014
NY6	0	0	0	1.000	0.647
NY7	–	–	0	0.223	0.114
NY8	–	–	0	0.092	0.113
NY9	–	0	0	0.223	0.164
NY10	0	–	+	0.011	0.009
NY11	–	–	+	0.011	0.013
NY12	0	–	+	0.223	0.110
NY13	–	–	+	0.011	0.010
NY14	–	–	–	0.334	0.217
NY15	–	–	0	0.019	0.012
			MAINE LAKES		
ME1	+	0	0	0.136	0.071
ME2	+	–	+	0.005	0.005
ME3	+	0	0	0.009	0.005
ME4	+	0	0	0.026	0.056
ME5	+	–	0	0.005	0.002

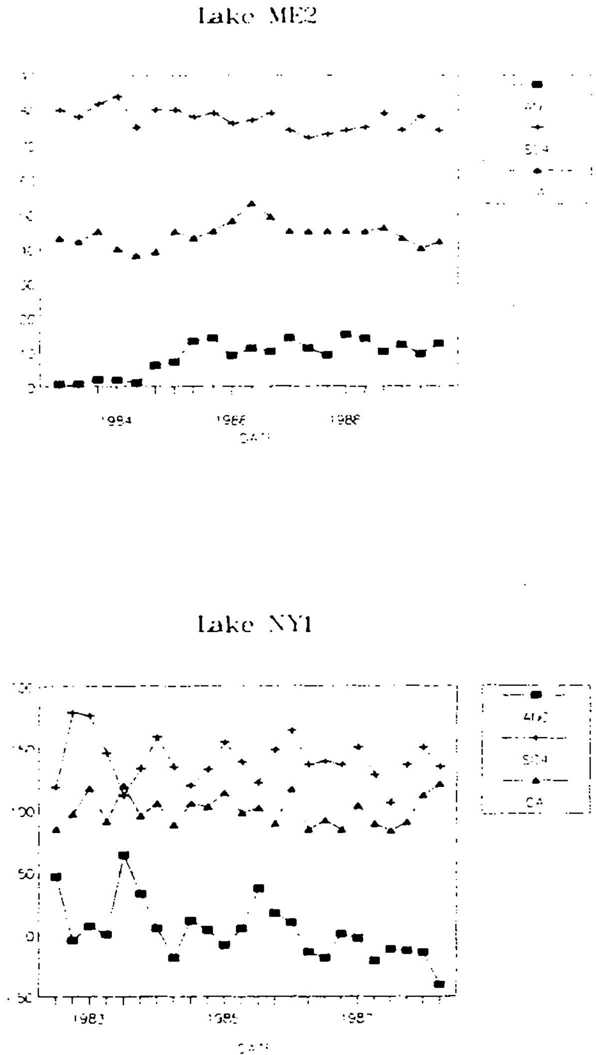

Figure 6a, b. Time Series Plots for ANC, Sulfate Concentration, and Calcium Concentration for the LTM Lakes ME2 and NY1.

SUMMARY AND CONCLUSIONS

Simulation and case study applications were used to compared the performance of univariate and multivariate methods of testing for trend in seasonal water quality data. Monte Carlo studies suggested that multivariate analysis of variance (MANOVA) and the modified covariance inversion (MCI) tests were generally more powerful than their univariate counterparts applied using the Bonferroni inequality.

This conclusion applied over the studied range of between-variable correlations, both positive and negative. Multivariate methods had the greatest power advantage in the case of positive correlations between variables with opposing slope directions. The Bonferroni methods performed better in only a few cases where trends were homogeneous across variables. However, non-homogeneous trends should be assumed for most real applications.

Application of the MCI and seasonal Kendall (SK) tests to lake quality data from EPA's Long Term Monitoring program in Maine and New York yielded comparable results. The three variables selected to evaluate acid deposition impacts exhibited a variety of patterns in trend directions.

Based on these results we can make a very positive recommendation for general application of multivariate approaches. There are very few cases where the univariate methods perform better, and, in those cases, the power advantage is small. Therefore, it should not be necessary to examine a data set for between-variable correlation structure or trend directions before choosing the multivariate approach. Of the two multivariate methods studied, we recommend the MCI test for routine applications because of its robust performance (as demonstrated elsewhere for rank-based methods in general) and its simplicity.

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 subject to review by the agency, and on official endorsement should be inferred.

LITERATURE CITED

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

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

Driscoll, C. T., 1991. Processes Regulating Seasonal Variations in the Water Chemistry of Adirondack Lakes. Water, Air and Soil Pollution (in preparation).

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.

Kahl, J. S., T. A. Haines, S. A. Norton, and R. B. Davis, 1991. Recent Trends in the Acid-Base Status of Surface Waters in Maine, U.S.A. Water, Air, and Soil Pollution (in press).

Lettenmaier, D. P., 1988. Multivariate Nonparametric Tests for Trend in Water Quality. Water Resources Bulletin 24(3):505-512.

Lettenmaier, D. P., E. R. Hooper, C. Wagoner, and K. B. Faris, 1991. Trends in Stream Quality in the Continental United States, 1978-1987. Water Resources Research 27(3):327-339.

Loftis, J. C., C. H. Taylor, and P. L. Chapman, 1991. Multivariate Tests for Trend in Water Quality. Water Resources Research 27(7):1419-1429.

Newell, A. D., C. F. Powers, and S. J. Christie, 1987. Analysis of Data from Long Term Monitoring of Lakes. EPA/600/4-87/014, U.S. Environmental Protection Agency, Washington, D.C.

Sen, P. K. and M. L. Puri, 1977. Asymptotically Distribution-Free Aligned Rank Order Tests for Composite Hypotheses for General Multivariate Linear Models. Zeischrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete 39:187-195.

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 Times Series. Water Resources Bulletin 25(4):715-726.

APPENDIX

THE MODIFIED COVARIANCE INVERSION (MCI) TEST FOR TREND

Before describing the multivariate procedure, a brief description of the univariate Mann-Kendall test for trend is in order.

$$\text{Define} \quad \mathrm{sign}(x) = \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ if } x = 0 \\ -1 \text{ if } x < 0 \end{cases} \tag{A1}$$

Under the null hypothesis of no trend

$$K = \sum_{i<j} sign(y_j - y_i) \tag{A2}$$

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

$$\sigma^2 = n(n-1)(2n+5)/18 \tag{A3}$$

Dietz and Killeen (1981) extended these results to the multivariate case. Specifically, they considered a sequence of p- variate observations of the form $\{(y_{1i}, \ldots, y_{pi}): i=1, \ldots, n\}$. 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

$$K = [K_1, \ldots, K_p]^T \tag{A4}$$

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

$$\Sigma_{gh} = \begin{cases} \sigma^2 & if\ g=h \\ (t_{gh}+r_{gh})/3 & if\ g \neq h \end{cases} \tag{A5}$$

where:

$$t_{gh} = \sum_{i<j} sign\left[\left(x_{gj} - x_{gi}\right)\left(x_{hj} - x_{hi}\right)\right] \tag{A6}$$

and

$$r_{gh} = \sum_{i,j,k} sign\left[\left(x_{gj} - x_{gi}\right)\left(x_{hj} - x_{hk}\right)\right] \tag{A7}$$

Hirsch and Slack (1984) used these results to generate a test for univariate trends. If only a single variable is considered at a single lake for omega 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 variance-covariance matrix Σ.

The above results lead to extensions of univariate seasonal Kendall statistics to multivariate cases. We shall consider δ water quality variables for ω seasons at a single location. 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 δω Mann-Kendall statistics of the vector K so that the statistic corresponding to the season j and constituent k is element f of K where

$$f = j + (k-1)\,\omega. \tag{A8}$$

In a corresponding fashion, arrange the elements of Σ. In addition, define the matrix **C** with dimension δ by δω.

Define the i,j element of C as

$$C_{ij} = \begin{cases} 1\ if\ (i-1)\,\omega + 1 \le j \le i\omega \\ 0\ if\ otherwise \end{cases} \tag{A9}$$

As a result

$$\mathbf{S} = \mathbf{CK} \tag{A10}$$

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

$$\Gamma = \mathbf{C\Sigma C^T} \tag{A11}$$

The "covariance inversion" test (a term used by Lettenmaier, 1988) is based on the work of Dietz and Killeen (1981) who suggested that

$$\psi = S^T \Gamma^{-1} S \qquad \text{(A12)}$$

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

$$\psi = S^T \Gamma^{-1} S \qquad \text{(A13)}$$

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 statistic.

The covariance inversion test accounts for between-season correlation via the variance-covariance matrix Σ in Equation (A5). However our formulations of the MANOVA and seasonal Kendall methods assume independence between seasons. We can easily construct a modified version of the covariance inversion test which effectively assumes independence between seasons and, therefore, more closely parallels the MANOVA and seasonal Kendall methods.

The modified test is 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.

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/J-94/510	2.	3. RECIPIENT'S ACCESSION NO. PB95-148722
4. TITLE AND SUBTITLE Multivariate trend testing of lake water quality		5. REPORT DATE
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) [1]J.C.Loftis, [2]C.H.Taylor, [3]A.D.Newell [1]P.L.Chapman		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS [1]CSU, FT.Collins, CO, [2]Merrel Dow Rsch Inst.,Cincinnati, OH, [3]ManTech, ERL-Corvallis,OR		10. PROGRAM ELEMENT NO.
		11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS US EPA ENVIRONMENTAL RESEARCH LABORATORY 200 SW 35th Street Corvallis, OR 97333		13. TYPE OF REPORT AND PERIOD COVERED Journal Article
		14. SPONSORING AGENCY CODE EPA/600/02

15. SUPPLEMENTARY NOTES

1991 Water Resources Bulletin, 27(3) 461-473

16. ABSTRACT

Multivariate methods of trend analysis offer the potential for higher power in detecting gradual water quality changes as compared to multiple applications of univariate tests. Simulation experiments were used to investigate the power advantages of multivariate methods for both linear model and Mann-Kendall based approaches. The experiments focused on quarterly observations of three water quality variables with no serial correlation and with several different intervariable correlation structures. The multivariate methods were generally more powerful than the univariate methods, offering the greatest advantage in situations where water quality variables were positively correlated with trends in opposing directions. For illustration, both the univariate and multivariate versions of the Mann-Kenadall based tests were applied to case study data from several lakes in Maine and New York which had been sampled as part of EPA's long term monitoring study of acid precipitation effects.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
trend analysis, multivariate analysis, statistics, lake water quality, acid precipitation, probability, population space		
18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES 13
	20. SECURITY CLASS *(This page)*	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE