TECHNIQUES FOR DETECTING EFFECTS OF URBAN AND RURAL LAND-USE PRACTICES ON STREAM-WATER CHEMISTRY IN SELECTED WATERSHEDS IN TEXAS, MINNESOTA, AND ILLINOIS

By John F. Walker

U.S. GEOLOGICAL SURVEY

Open-File Report 93-130

Prepared in cooperation with the

WISCONSIN DEPARTMENT OF NATURAL RESOURCES

Madison, Wisconsin

1993

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, *Secretary*

U.S. GEOLOGICAL SURVEY
Dallas L. Peck, *Director*

For additional information write to:

District Chief
U.S. Geological Survey
6417 Normandy Lane
Madison, WI 53719

Copies of this report can be purchased from:

U.S. Geological Survey
Books and Open-File Reports
Federal Center, P.O. Box 25286
Denver, CO 80225

CONTENTS

ILLUSTRATIONS

TABLES

CONVERSION FACTORS

Multiply	***By***	***To obtain***
centimeter (cm)	0.3937	inch
square kilometer (km^2)	0.3861	square mile
cubic meter (m^3)	35.31	cubic feet
centimeter per hour (cm/hr)	0.0328	inch per hour
kilogram (kg)	2.205	pound

TECHNIQUES FOR DETECTING EFFECTS OF URBAN AND RURAL LAND-USE PRACTICES ON STREAM-WATER CHEMISTRY IN SELECTED WATERSHEDS IN TEXAS, MINNESOTA, AND ILLINOIS

By
John F. Walker

ABSTRACT

Although considerable effort has been expended during the past two decades to control nonpoint-source contamination of streams and lakes in urban and rural watersheds, little has been published on the effectiveness of various management practices at the watershed scale. This report presents a discussion of several parametric and nonparametric statistical techniques for detecting changes in water-chemistry data. The need for reducing the influence of natural variability was recognized and accomplished through the use of regression equations. Traditional analyses have focused on fixed-frequency instantaneous concentration data; this report describes the use of storm load data as an alternative.

Selected statistical techniques were applied to three urban watersheds in Texas and Minnesota and three rural watersheds in Illinois. For the urban watersheds, single- and paired-site data-collection strategies were considered. The paired-site strategy was much more effective than the single-site strategy for detecting changes. Analysis of storm load regression residuals demonstrated the potential utility of regressions for variability reduction. For the rural watersheds, none of the selected techniques were effective at identifying changes, primarily due to a small degree of management-practice implementation, potential errors introduced through the estimation of storm load, and small sample sizes. A Monte Carlo sensitivity analysis was used to determine the percent change in water chemistry that could be detected for each watershed. In most instances, the use of regressions improved the ability to detect changes.

INTRODUCTION

During the past two decades, considerable effort has been expended to control nonpoint-source contamination of lakes and streams in urban and rural watersheds of the United States. The Rural Clean Water Program (RCWP) is a Federally sponsored program developed to address agricultural nonpoint-source contamination in 20 rural watersheds across the country. The program is voluntary and provides cost-sharing and technical assistance to program participants. A similar State-sponsored program has been operating in Wisconsin since 1978.

Individual management practices, termed best-management practices (BMP's), have been evaluated at small scales for both urban and rural practices (Johnson and others, 1979; Dickey and Vanderholm, 1981; Mickelson and others, 1983; Fausey and others, 1988). However, until recently, the effectiveness of several management practices implemented at the same time in a watershed has not been documented. The National Water Quality Evaluation Project (NWQEP) was designed to evaluate the effectiveness of RCWP projects and is considering scales ranging from individual management practices to entire watershed plans (Smolen and others, 1989). The efforts undertaken by the NWQEP have been extensive and involve long-term, intensive water-chemistry sampling.

Most of the evaluation work to date has focused on instantaneous concentration data collected at fixed frequencies or on annual constituent load. Nationwide assessments of water-chemistry trends have been conducted (Smith and others, 1982; Smith and Alexander, 1983; Smith and others, 1987; and Alexander and Smith, 1988); however, the watersheds were generally quite large, and there were no conclusions drawn relating trends to management practices. Several researchers have developed guidelines for water-chemistry monitoring and have estimated the minimum change that can be detected using prescribed data-collection strategies (Spooner and others, 1987; Richards, 1989; Reckhow and Stow, 1990). The conclusion from most studies is that detection of change requires several years of intensive "before-and-after" data collection.

The purpose of this report is to describe several statistical techniques that are commonly used to detect effects of land-use practices on water chemistry and to apply these techniques to selected urban and rural watersheds. Techniques considered will include parametric and non-parametric tests of hypotheses. Statistical techniques for reducing the influence of natural variability are considered and applied to selected urban and rural watersheds. Alternative procedures for assessing the effect of land-use practices on stream-water chemistry will be compared.

TECHNIQUES FOR DETECTING EFFECTS OF LAND-USE PRACTICES ON WATER CHEMISTRY

A thorough discussion of the issues involved in selecting a technique for detecting changes in water-chemistry variables is presented by Hirsch and others (1991). The choice of a technique for detecting changes depends primarily on the type of data collected. Two basic data-collection strategies can be used to detect changes in water-chemistry data—single-site and paired-site data collection. For both types of data-collection strategies, either instantaneous concentrations or loads can be examined, depending on the objective(s) of the watershed management plan (Hirsch and others, 1991).

In this section, several statistical techniques are described and evaluated for detecting changes in water-chemistry data collected using single-site and paired-site strategies. Techniques for separating natural variability and land-use practices are considered for single-site data-collection strategies. Finally, a collection of statistical and data analysis techniques are selected for application to both urban and rural watersheds.

Statistical Techniques for Single-Site Data Collection

Single-site data-collection strategies generally are used at either a critical point in a watershed or at the confluence of a stream with another stream or lake. Two general approaches to data collection can be employed—continuous sampling at a fixed time interval or intermittent sampling during "before-and-after" periods. These two approaches dictate different statistical-analysis procedures.

For data collected at a single site, techniques are available for detecting either monotonic trends or discrete changes. The choice of technique depends, in part, on the type of trend expected but mostly on the type and frequency of data collected (Hirsch and others, 1991). Monotonic trends occur gradually through time, whereas discrete changes generally occur during a short time interval.

Two general classes of statistical approaches, parametric and nonparametric, can be applied and are selected based on assumptions regarding the underlying probability distribution of the data. Parametric techniques require a specific, assumed distribution; most parametric tests assume a Gaussian or normal distribution. Nonparametric tests, on the other hand, do not require a specific form for the underlying data distribution; these tests generally work with the ranks of the data rather than the data itself. If the form of the underlying distribution is known, parametric tests generally are more powerful than their nonparametric counterparts. If deviations from the assumed distribution exist, nonparametric tests are usually more powerful. Because water-chemistry data can contain unspecified values below detection levels and are bounded by zero, nonparametric techniques usually are more appropriate than parametric techniques. A description of some of the more effective parametric and nonparametric techniques found in the literature is given in the next section.

Because monotonic trends occur gradually through time, statistical techniques for detecting monotonic trends work best with data sampled at a fixed time interval; in most cases, a monthly sampling schedule is suggested (Hirsch and others, 1982; Hirsch, 1988; Reckhow and Stow, 1990). One of the most powerful nonparametric tests for monotonic trend is the Kendall test (Conover, 1980; Hirsch and others, 1982), which has been modified to account for seasonality and serial correlation (Hirsch and Slack, 1984). The Spearman test (Conover, 1980), another nonparametric test for monotonic trend, has been modified to account for serial correlation (Lettenmaier, 1976) and compares favorably with the Kendall test (Berryman and others, 1988). Parametric approaches include either linear regression or analysis of variance with time as the explanatory variable (Hirsch and others, 1991).

Because discrete changes occur abruptly, statistical techniques for detecting discrete changes can be applied to data collected either continuously or sporadically. Thus, these techniques are applicable to "before-and-after" data sets where there is a gap in the data-collection activities. The nonparametric technique used in most instances is the Mann-Whitney U test (Conover, 1980; Berryman and others, 1988), which has been modified for use with data exhibiting serial correlation (Lettenmaier, 1976). The most extensively used parametric technique is the t-test (Iman and Conover, 1983).

Statistical Techniques for Paired-Site Data Collection

Paired-site strategies generally are used with "test" and "control" watersheds or in situations where there is an "inflow" and "outflow" component of the management practice. Because of problems in selecting hydrologically and climatologically similar watersheds, "test" and "control" strategies are difficult to use in rural areas. In urban areas, however, detention and retention facilities are appropriate applications for the "inflow" and "outflow" paired-site data-collection strategy.

For paired data, one compares the median (nonparametric) or mean (parametric) of the two data sets. The usual nonparametric technique is the Wilcoxon test (Conover, 1980), and the parametric equivalent is the paired t-test (Iman and Conover, 1983).

Statistical Techniques for Reducing the Influence of Natural Variability for Single-Site Data Collection

The variability of water-chemistry data often can confound statistical-analysis techniques for single-site collection strategies. For instance, differences in climatic conditions "before-and-after" BMP implementation can either mask real changes or produce false changes. Data variability can be separated into three components—measurement error, natural variability, and human-induced effects. In evaluating the effectiveness of land-use practices, one attempts to isolate the third component of variability, human-induced effects.

Several sources of measurement uncertainty exist which must be minimized to keep the effect of measurement uncertainty from dominating the overall variability in the data. The first source is sample representativeness, which depends on the sample-collection techniques. The second source is sample integrity, which depends on the sample-processing techniques. The third source is laboratory uncertainty, which depends on the analytical laboratory procedures. When load is of interest, additional uncertainty is introduced by uncertainty in estimates of discharge and in the technique used to calculate the storm load.

Climate and seasonality are the two main sources of natural data variability. Because the hydrologic cycle is driven by climatic factors (for example, precipitation, temperature, solar radiation), the spatial and temporal distribution of these factors have an effect on the variability of water-chemistry data. Additional factors, particularly in urban areas, include timing and distance from source areas. Natural seasonal variations result in part from seasonal variations in climatic factors and in part from seasonal variations in the condition of the land (frozen or unfrozen ground, condition of vegetation, and so forth). Natural variability is one of the variability components to be compensated for when evaluating water-chemistry data. Unfortunately, because land-use practices may have a seasonal component (for example, condition of soil, type of vegetation, and application of chemicals), the compensation for seasonal factors can remove a desired part of the data variability.

Hirsch and others (1982) apply statistical-analysis techniques to flow-adjusted concentrations, which are the residuals that result from a regression of constituent concentration and streamflow. A regression residual is the difference between an observed value of the dependent variable and the value estimated by the regression. On a plot of the data with the independent variable on the X-axis and the dependent variable on the Y-axis, the residual is the vertical distance between an observed point and the fitted regression line. More sophisticated regressions include seasonal and hysteresis terms in addition to the streamflow variable (Hirsch, 1988). Spooner and others (1987) apply a similar approach for decreasing external factors before determining the minimum detectable change. The regression equation "explains" that part of the water-chemistry constituent that varies with the independent variable (for example, natural variability); thus, the residuals represent the part of the water-chemistry constituent concentration caused by measurement uncertainty and other "unexplained" factors (for example, human-induced effects).

Using flow-adjusted concentrations to detect water-chemistry changes due to land-use practices in small watersheds is undesirable for two reasons. First, the land-use practice can alter discharge; in this case flow-adjusted concentrations would remove some of the desired variability. Second, for many watersheds there is considerable variability in constituent concentrations that cannot be explained adequately by discharge. Of the BMP's presently cost shared by the Wisconsin Department of Natural Resources Nonpoint Program, the following are expected to alter the quantity or temporal characteristics of discharge: contour farming, contour and field-strip cropping, field diversions, terraces, reduced-tillage systems, agricultural sedimentation basins, wetland restoration, and structural urban best-management practices. In addition, the following practices may alter the discharge characteristics, depending on site-specific conditions: grassed waterways, critical-area stabilization, grade-stabilization structures, and shoreline buffers. Several practices are not expected to affect the discharge characteristics at the basin scale; these practices include nutrient and pesticide management, shoreline and streambank stabilization, barnyard-runoff management, animal lot relocation, manure storage facilities, roofs for barnyard-runoff management and manure-storage facilities, and livestock exclusion from woodlots.

For the reasons just outlined, additional regression variables are needed to adequately reduce variability. In this light, climatic variables, particularly those related to precipitation, need to be used in the regressions. Because instantaneous variables may not be directly related to climatic variables, integration over individual storms is suggested, resulting in constituent storm loads. Thus, total constituent load is related more directly to storm precipitation characteristics, which are independent of land-use practice effects. In theory, the regressions would isolate the natural climatic variations from the measurement uncertainty and land-use practice effects. The characteristics of precipitation used for this study were total precipitation, 15- and 30-minute maximum precipitation intensities, and the Universal Soil Loss Equation (USLE) erosivity index, which is a measure of the erosive potential of a storm that combines kinetic energy and intensity into a single factor (Wischmeier and Smith, 1978):

$$EI = I_{30} \bullet \{ \Sigma P_{\Delta} \bullet [916 + 331 \bullet \log_{10} (i_{\Delta})] \}/100,$$

where EI is the erosivity index,
I_{30} is the maximum 30-minute precipitation intensity for the storm,
P_{Δ} is the precipitation amount for each 5-minute increment of the storm, and
i_{Δ} is instantaneous precipitation intensity for each 5-minute increment.

The summation is carried out over a number of small increments over the duration of the storm. Note that a maximum of 3 inches per hour is imposed on the incremental intensity (i_{Δ}). Erosivity index terms were computed for individual storms from instantaneous precipitation data. In addition to precipitation characteristics, several other variables, including seasonal terms, were considered.

Comparison and Selection of Techniques

Several different analyses were performed to allow a comparison of selected statistical techniques and an evaluation of the utility of using regressions to reduce the influence of natural variability for single-site data-collection strategies. Computed loads for individual storms were selected for analysis to allow the use of regressions and climatic variables to reduce the effects of natural variability. Two statistical tests—the t-test and Mann-Whitney U test—were applied to the storm load data and regression residuals. For paired-site collection strategies, two tests were used—the paired t-test and the Wilcoxon test.

APPLICATION OF TECHNIQUES TO SELECTED WATERSHEDS

The techniques selected for single-site data-collection strategies (t-test and Mann-Whitney U test) were applied to data from three urban watersheds in Texas and Minnesota and three rural watersheds in Illinois. Two techniques for paired-site collection strategies were applied to data from the urban watersheds. The watersheds were selected on the basis of the following criteria: (1) availability of data before and after BMP implementation, (2) continuity of data or computed load for individual storms, (3) reliability of continuous precipitation data or a summary of desired climatic variables for individual storms, (4) collection of suspended sediment and nutrients, and (5) watershed areas that were reasonably small (less than 100 km^2). Three urban and three rural watersheds were identified, and the data were obtained from the researchers that originated the studies. General characteristics of the watersheds are listed in table 1.

Urban Watersheds

The most common management practices in urban watersheds are retention and detention ponds. For these types of structures, the usual procedure is to monitor constituent loads at the pond inlet and outlet, which results in a paired "inflow-outflow" data set. The three urban watersheds selected all involved detention structures, and data-collection activities monitored inflow and outflow loads. To allow evaluation of a single-site data-collection strategy for urban BMP's, "before and after" data sets were simulated by splitting the data set approximately in half and using inflow load for the first part of the data to represent "before" conditions and outlet load for the second part of the data to represent "after" conditions. This was equivalent to operating a single gaging station at the outlet before and after construction of the detention structure, provided that the inlet and outlet were in close proximity to one another and that natural detention would not have existed between inlet and outlet prior to construction. For all three urban watersheds, three water-chemistry constituents were considered: suspended sediment (SS), total phosphorus (TP), and either dissolved or total lead (Pb).

Description of Watersheds

The Barton Square detention pond is one of three detention and filtering ponds that capture storm runoff from the Barton Creek Square shopping center in Austin, Texas (Welborn and Veenhuis, 1987). The pond captures runoff

Table 1. General characteristics of urban and rural watersheds selected for analysis
[km², square kilometers]

Name and location	Watershed area (km²)	Period of data collection	Sampling frequency	Number of storms	Reference
		Urban watersheds			
Barton Square detention pond (Austin, Texas)	0.32	1982-84	Event	19	Welborn and Veenhuis, 1987
Lake Ridge detention system (Woodbury, Minnesota)	1.3	1987-88	Event	15	Oberts and others, 1989
Tanner's Lake wetland system (Oakdale, Minnesota)	1.7	1988	Event	10	Oberts and others, 1989
		Rural watersheds			
Highland Silver Lake GS-1 (Madison County, Illinois)	80	1982-84	Monthly and event	58	Kelly and Davenport, 1986
Highland Silver Lake GS-2 (Madison County, Illinois)	51	1982-84	Monthly and event	34	Kelly and Davenport, 1986
Highland Silver Lake GS-3 (Madison County, Illinois)	14	1982-84	Monthly and event	36	Kelly and Davenport, 1986

from a 0.32-km² area and has a surface area of 0.008 km² and a storage capacity of 4,300 m³. The bed of the pond consists of three layers of different size material, which serves to filter the water as it infiltrates through to the outlet structure. Composite water-chemistry samples were collected and used with a continuous record of discharge to estimate individual storm loads. Continuous precipitation data were collected and used in the original study to determine total precipitation and the 15-minute maximum precipitation intensity for individual storms (Welborn and Veenhuis, 1987). Because continuous precipitation data were not obtained from the authors, the 30-minute maximum precipitation intensity and the USLE erosivity index could not be determined.

The Lake Ridge detention pond captures storm runoff from a small residential area in the Twin Cities Metropolitan Area, Minnesota (Oberts and others, 1989). The effective watershed contributing to the pond is 1.3 km²; the pond has a surface area of 0.0038 km² and a storage capacity of 8,100 m³. Composite water-chemistry samples were collected and used with a continuous record of discharge to estimate individual storm load. Continuous precipitation data were collected and used to determine total precipitation, 15- and 30-minute maximum precipitation intensities, and the USLE erosivity index for individual storms.

The Tanner's Lake wetland system captures storm runoff from another small residential area in the Twin Cities Metropolitan Area, Minnesota (Oberts and others, 1989). The system consists of a small sedimentation basin and a wetland partitioned into two parts with permeable weirs at the mid-section and downstream end of the wetland. The area of the watershed contributing to the wetland system is 1.7 km². Composite water-chemistry samples were collected and used with a continuous record of discharge to estimate individual storm load. Continuous precipitation data were collected and used to determine total precipitation, 15- and 30-minute maximum precipitation intensities, and the USLE erosivity index for individual storms.

Results for Data Collected from a Single Site

Selected statistical techniques were applied to the simulated "before-and-after" storm load data for the three urban watersheds, and the results are summarized in table 2. The Kolmogorov-Smirnov test for normality (Conover, 1980) was applied to both before and after data. Because this test failed to reject the hypothesis of normality for all cases, no transformation of the data was used for the t-test. The selected techniques were not applied to instantaneous concentration data because the storm sampling was carried out using composite samples; hence, instantaneous concentrations at a fixed interval were not available. The results of three selected techniques were comparable for most watersheds and constituents, and very few tests identified statistically significant changes for a critical level of 0.05. The exception is the Lake Ridge site, where significant changes were identified for suspended sediment and lead using the Mann-Whitney U test.

Table 2. Results of statistical tests applied to before-and-after storm load data for three urban watersheds

[SS, suspended sediment; TP, total phosphorus; Pb, dissolved lead for Barton Square, total lead otherwise]

		t-test		Mann-Whitney U test	
Constituent	Sample size	Test statistic	Significance level	Test statistic	Significance level
		Barton Square detention pond			
SS	19	-0.25	0.807	22	0.065
TP	19	.17	.865	24	.095
Pb	19	72	.483	38	.604
		Lake Ridge detention pond			
SS	15	1.1	.301	7.0	.015
TP	15	1.1	.304	16	.189
Pb	15	1.2	.261	8.5	.025
		Tanner's Lake wetland system			
SS	10	-.23	.822	10	.691
TP	10	-1.1	.319	12	1.000
Pb	10	-.38	.720	11	.841

Stepwise, multiple linear regressions were applied to the simulated "before-and-after" storm load data to investigate reducing the influence of natural variability in the data. Several independent variables were considered—total precipitation (P_{tot}); 15- and 30-minute maximum precipitation intensities (I_{15} and I_{30}); the USLE erosivity index (EI); and the trigonometric sine and cosine of time, in years (T; T=1980.5 corresponds to June 30, 1980). The sine and cosine terms represent seasonal variations. The resulting regressions are summarized in table 3; all of the regressions are significant at the 0.05 level. Note that the standard errors of the regressions are quite high, indicating a general lack of fit of the regression, and in part due to the relatively small sample sizes. While the regressions wouldn't be useful in a predictive sense, they are still useful in terms of removing some of the natural variability due to the climatic variables.

The selected statistical techniques were applied to the residuals from the regressions presented in table 3, and the results are reported in table 4. For a substantial number of the tests, the significance levels for the regression residuals were smaller than the significance levels for the storm load data. This implies that the regressions have improved the ability of the tests to detect a discrete change. For example, the tests for the Barton Square watershed produced significant changes for suspended sediment and total phosphorus after the regressions, whereas the changes were not significant for the raw storm load data (table 2). In some cases the reverse is true; the regression residuals resulted in larger significance levels, which decreases the ability of a test to detect a significant change. For instance, the significant changes detected at the Lake Ridge site for suspended-sediment and lead storm loads are no longer significant after adjusting for natural variability.

Results for Data Collected from Paired Sites

The paired-site inflow and outflow data for the three urban watersheds were subjected to the paired t-test and Wilcoxon test to identify significant differences in the mean or median between inflow and outflow. The Kolmogorov-Smirnov test for normality (Conover, 1980) was applied to both the inflow and outflow data, and a natural logarithmic transformation was used for the t-test for cases where a significant departure from normality was detected. The two tests produced similar results, and with the exception of the t-test applied to the data for Tanner's Lake, all tests identified significant differences between the inflow and outflow (table 5).

Table 3. Results of step-wise regressions applied to before-and-after storm load data for three urban watersheds

[ln(x), natural logarithm of x; SS, suspended sediment; TP, total phosphorus; Pb, dissolved lead for Barton Square, total lead otherwise; P_{tot}, total precipitation; I_{15}, 15-minute maximum precipitation intensity; EI, erosivity index; T, date in decimal years; sin(x), trigonometric sine of x; <, less than]

Dependent variable	Independent variable(s)	Adjusted [1]R^2	Standard error [2]	Significance level
		Barton Square detention pond		
ln (SS)	I_{15} , P_{tot}	0.675	74	0.008
ln (TP)	P_{tot}	.897	29	<.001
Pb	P_{tot} , I_{15}	.660	34	.010
		Lake Ridge detention pond		
SS	EI	.961	49	<.001
TP	EI	.977	34	<.001
Pb	EI	.988	23	<.001
		Tanner's Lake wetland system		
ln (SS)	T , ln (P_{tot})	.900	35	.050
ln (TP)	T	.749	32	.037
Pb	sin (T)	.851	20	.016

[1] Fraction of variance in the dependent variable explained by the regression adjusted for degrees of freedom.

[2] Expressed as a percentage of the mean of the dependent variable.

Table 4. Results of statistical tests applied to before-and-after regression residuals for three urban watersheds

[SS, suspended sediment; TP, total phosphorus; Pb, dissolved lead for Barton Square, total lead otherwise; <, less than]

	t-test		Mann-Whitney U test	
Constituent	Test statistic	Significance level	Test statistic	Significance level
		Barton Square detention pond		
SS	4.4	<0.001	8	0.002
TP	6.0	<.001	3	<.001
Pb	.77	.451	41	.780
		Lake Ridge detention pond		
SS	1.8	.101	14	.121
TP	1.8	.101	17	.232
Pb	1.9	.095	14	.121
		Tanner's Lake wetland system		
SS	.72	.489	9	.548
TP	.76	.486	10	.690
Pb	.00	.999	10	.690

Table 5. Results of statistical tests applied to inflow-and-outflow storm load data for three urban watersheds

[SS, suspended sediment; TP, total phosphorus; Pb, dissolved lead for Barton Square, total lead otherwise; <, less than]

		t-test		Wilcoxon test	
Constituent	Sample size	Test statistic	Significance level	Test statistic	Significance level
		Barton Square detention pond			
SS	22	[1]-7.3	<0.001	-4.1	<0.001
TP	22	[1]-4.3	<.001	-2.7	.006
Pb	22	-3.5	.002	-3.3	.001
		Lake Ridge detention pond			
SS	21	[1]-8.9	<.001	-4.0	<.001
TP	21	[1]-5.3	<.001	-3.3	.001
Pb	21	[1]-4.8	<.001	-3.7	<.001
		Tanner's Lake wetland system			
SS	11	-1.4	.186	-2.9	.004
TP	11	-1.6	.135	-2.1	.033
Pb	11	-1.4	.181	-2.9	.004

[1] Test performed on logarithmic transformation of original data.

Comparison Between Results for Single-Site and Paired-Site Data

The use of regressions to compensate for natural variability for single-site data collection was predicated on the assumption that before and after data are related in a similar manner to the independent variable(s) used in the regression. Thus, the variability of the dependent variable that results from natural factors will be the same for before and after data. A separation of the before and after data around the overall regression relation will result if there is a real land-use practice effect on the data. The separation in the before and after data further emphasizes differences; hence, a statistical test applied to the residuals will have a smaller significance level than a test applied to the raw data.

Two possible explanations exist for situations where the significance level for regression residuals exceeds the significance level for the raw data. One explanation is that a significant change for the raw data results primarily from a change in a dependent variable. The regression could remove the effect of a change in the related variable; thus, the residuals may not indicate a change. A second explanation is that the independent variables affecting the natural variation are different for the before and after data.

As an example of a decrease in significance level, the simulated pre- and post-BMP data for suspended sediment at Barton Square were plotted against the two independent variables, total precipitation and 15-minute maximum precipitation intensity, in figure 1. The use of two independent variables resulted in a clear separation of the before and after data. The suspended-sediment load and the regression residuals were plotted against time in figure 2. In this case, the statistical tests were unable to identify a significant change in the storm load data (fig. 2A, table 2), whereas the use of regression substantially decreased the significance level and revealed a significant change (fig. 2B, table 4).

The suspended-sediment data for the Lake Ridge watershed illustrates the case where the independent variables are different for before and after conditions. Although the overall regression for before and after data resulted in the erosivity index (EI) as the independent variable (table 3), the before and after data do not fit a model with EI as the independent variable (fig. 3; note that the regression was calculated on untransformed data, but was plotted with a log axis for clarity). In fact, additional regression analyses revealed that the "best" regression for the before data was suspended sediment as a function of erosivity index, whereas the best regression for the after data was the log of suspended sediment as a function of EI and a seasonal term. The load and regression residuals (fig. 4) indicate that the use of a

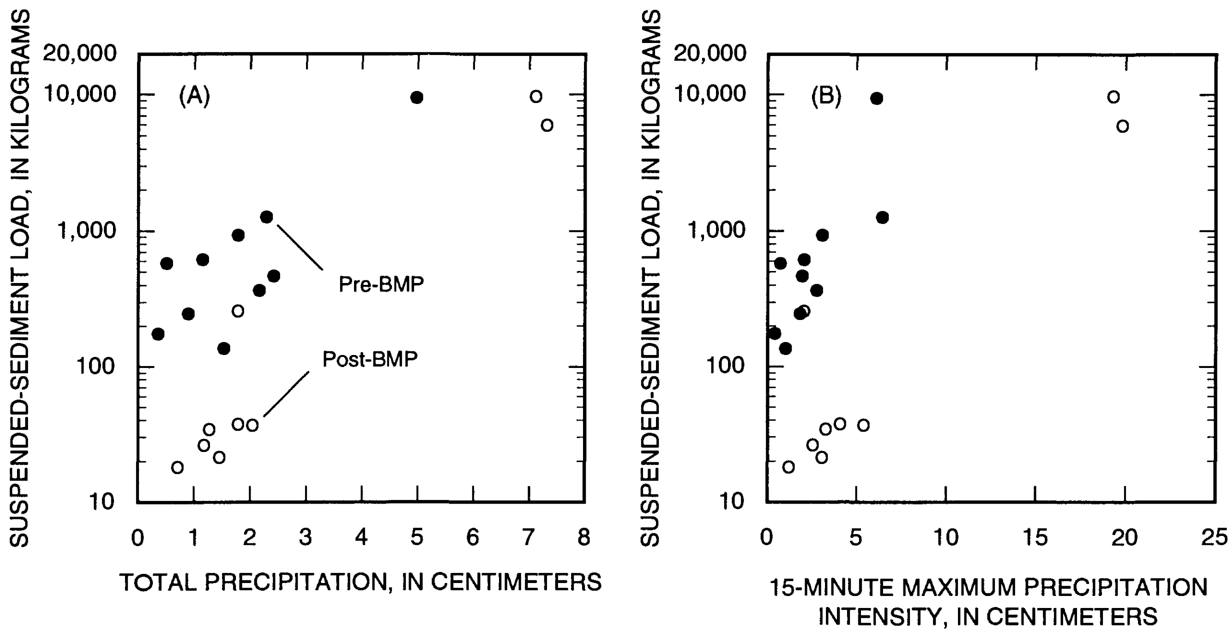

Figure 1. Relation of pre- and post-best-management-practice (BMP) suspended-sediment load and final regression variables (A) total precipitation and (B) 15-minute maximum precipitation intensity for Barton Square detention pond, Austin, Texas.

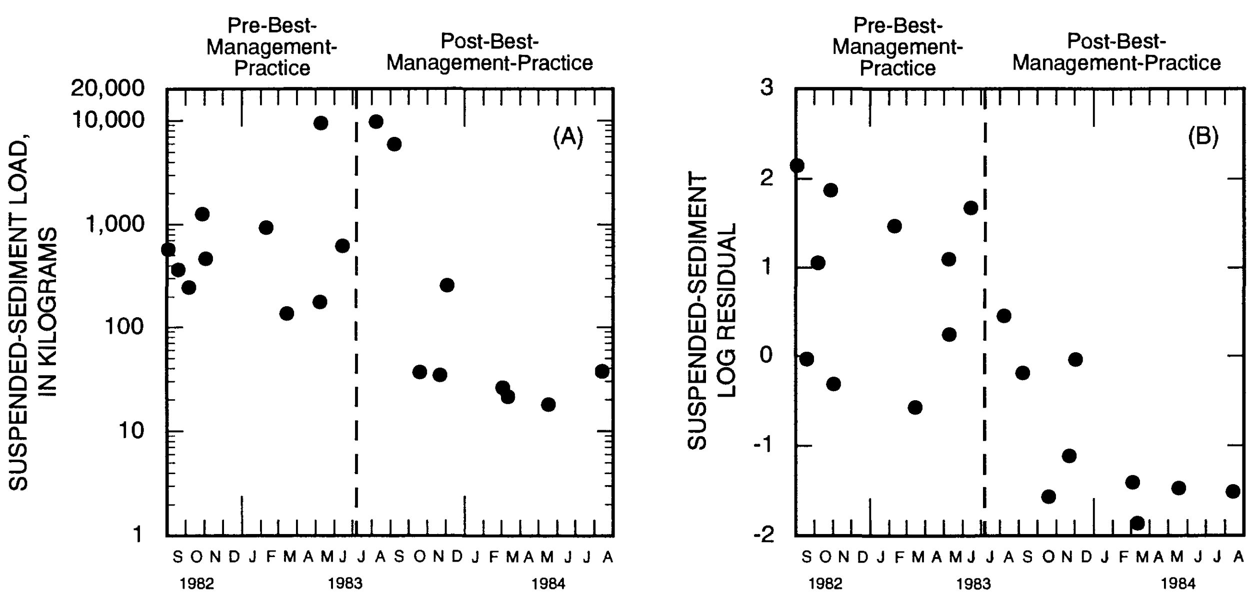

Figure 2. (A) Suspended-sediment load and (B) regression residuals for Barton Square detention pond, Austin, Texas, September 1982–August 1984.

regression in this case hindered the effectiveness of discrete change detection.

The results indicate that the paired-site data-collection strategy is more likely to identify a change than the single-site data-collection strategy (table 5). This is not surprising, because much of the variation caused by differences in storms is removed by the direct comparison of paired values. For most detention and retention facilities, it is likely that the before and after regressions are affected by different independent variables; hence, the single-site strategy would not be effective for change detection. The procedure is valid, however, for BMP's that do not drastically alter the relations between natural climatic variables and the dependent variable of interest (for example, street sweeping, catch-basin cleaning).

Rural Watersheds

Common management practices in rural agricultural watersheds include conservation tillage, contouring and strip cropping, streambank protection, and animal-waste management systems (Smolen and others, 1989). Most of the upland management systems are designed to reduce erosion from the land, which tends to decrease other water-chemistry constituents as well, particularly nutrients. Three rural watersheds, all contained within the Highland Silver Lake basin in Illinois, were selected for analysis. Because load data for individual storms was not available, suspended sediment was the only water-chemistry constituent considered.

Description of Watersheds

Highland Silver Lake is an artificial impoundment of the East Fork of Silver Creek that is used as a public-water supply for the city of Highland, Illinois and for noncontact recreation. The overall basin at the outlet of the lake is 125 km^2. Three data-collection sites (GS-1, GS-2, and GS-3) were established in three watersheds in the basin for evaluation of BMP implementation (Kelly and Davenport, 1986). The first watershed (site GS-1) is downstream of the other two watersheds, and encompasses both upstream watersheds. A variety of BMP's were implemented beginning in 1981. Because participation in the program was voluntary,

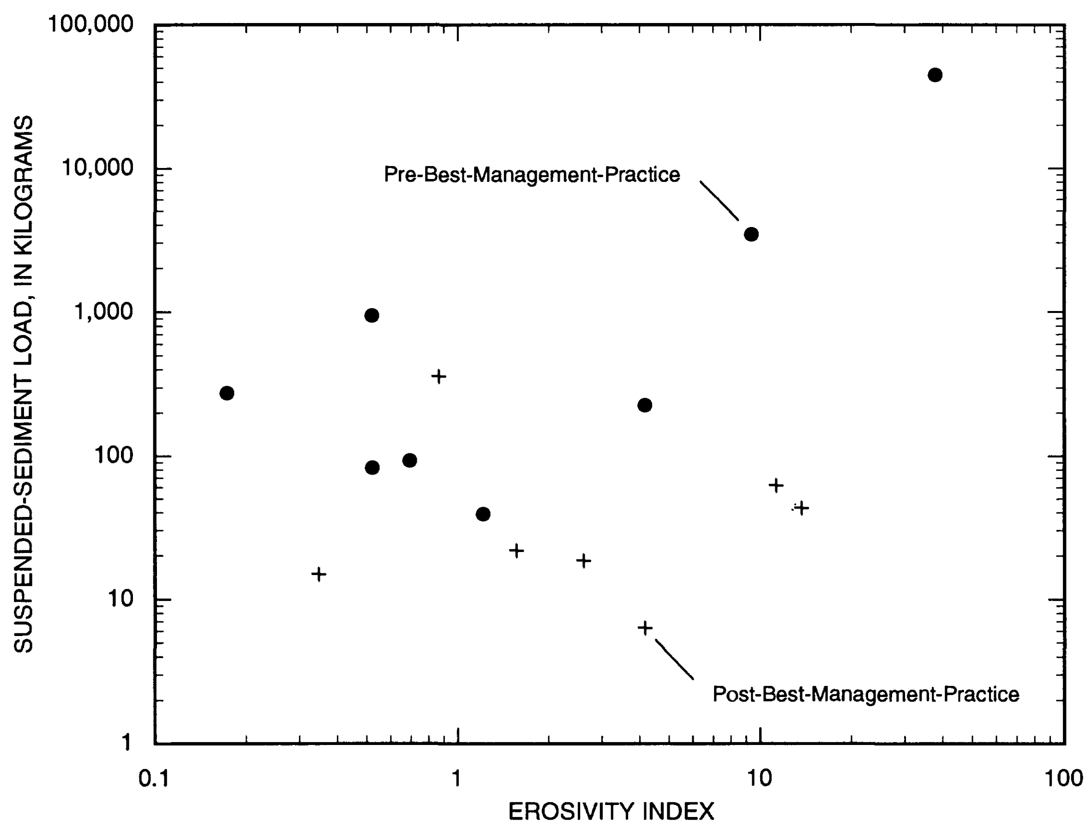

Figure 3. Relation between pre- and post-best-management-practice suspended-sediment load and Universal Soil Loss Equation Erosivity Index for Lake Ridge detention pond, Woodbury, Minnesota.

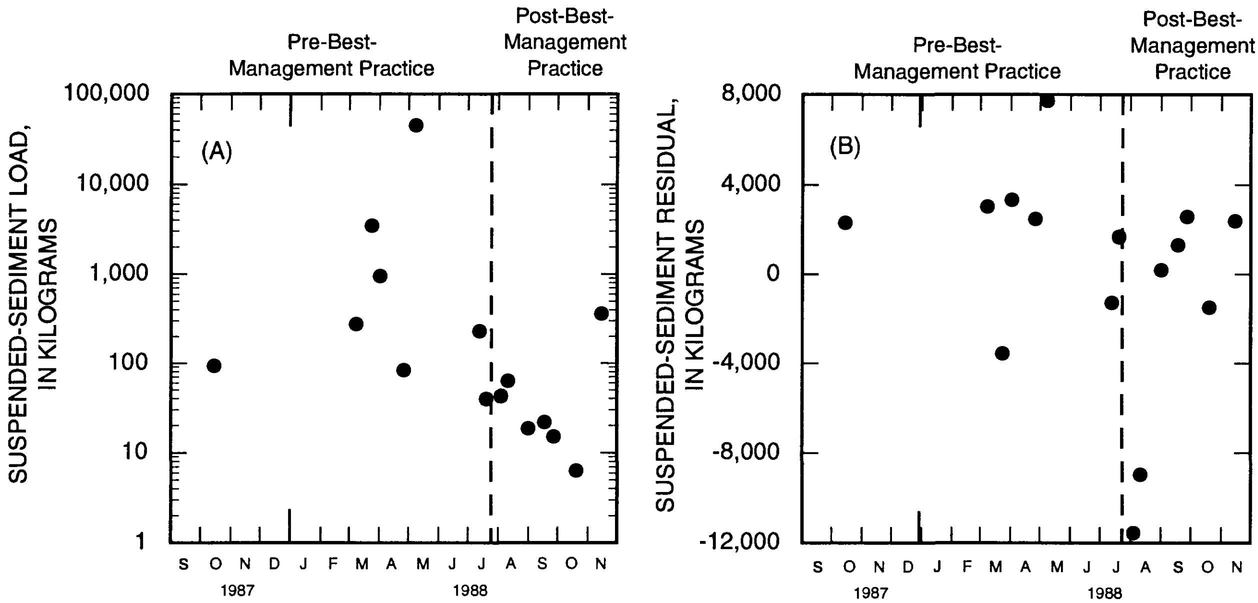

Figure 4. (A) Suspended-sediment load and (B) regression residuals for Lake Ridge detention pond, Woodbury, Minnesota, September 1987–November 1988.

implementation occurred over several years (table 6). On the basis of information presented in table 6, 1981-82 was considered as representative of "before" conditions, and 1983-84 was considered as representative of "after" conditions. Note that there is very little change in BMP implementation between the before and after periods.

Continuous streamflow was calculated from continuous measurements of stream stage and rating curves determined from periodic discharge measurements (Kelly and Davenport, 1986). Water-chemistry samples were collected bi-monthly to monthly, and more frequently during storms. Suspended-sediment concentrations were selected from the overall data set to approximate a monthly sampling strategy with occasional samples collected during storms.

Suspended-sediment loads for individual storms had to be estimated for this study. The suspended-sediment concentration data for storms were insufficient for use with the integration method (Porterfield, 1972), and adequate relations between concentration and discharge could not be developed. Therefore, suspended-sediment load for individual storms was determined by using a ratio estimator (Beale, 1962; Dolan and others, 1981) modified for variations in discharge and season. The ratio estimator was applied to three ranges of discharge and three separate periods corresponding approximately to agricultural seasons—April 1-June 30, July 1-November 30, and December 1-March 31 (Kelly and Davenport, 1986). The three discharge ranges were chosen to maximize the difference in suspended-sediment-concentration variability between the ranges. Continuous precipitation data were used to determine total precipitation, 15- and 30-minute maximum precipitation intensities, and the USLE erosivity index for individual storms.

Results for load data

Two statistical tests (t-test and Mann-Whitney U test) were applied to suspended-sediment load data for the three Highland Silver Lake watersheds; the results are reported in table 7. The results of the two tests are more or less consistent with one another; no significant trends were detected.

Step-wise, multiple linear regressions were applied to suspended-sediment load data to investigate reduction of the influence of natural variability on the data. The independent variables were identical to the ones used for the urban watersheds. The resulting regressions are summarized in table 8. All of the storm load regressions were statistically significant at the 0.05 level.

The two statistical techniques were applied to the residuals from the regressions presented in table 8, and the

Table 6. Best-management-practice implementation history for Highland Silver Lake GS-1, GS-2, and GS-3 watersheds, Madison County, Illinois
[Data from Kelly and Davenport, 1986]

Best management practice	Percentage of total watershed in indicated practice			
	1981	1982	1983	1984
	GS-1 watershed			
Chisel plow, spring, 20-percent residue	47	40	30	42
Chisel plow, fall, 20-percent residue	28	28	32	21
Conservation tillage, 20-percent residue	15	18	21	18
Conservation tillage, 30-percent residue	3.1	6.2	7.9	9.4
No till, 80-percent residue	.2	.9	1.5	3.0
Contouring	.4	.7	1.1	1.1
Terraces	.7	.8	.9	1.0
Pasture management	.6	.7	.8	.8
Livestock exclusion	.8	1.0	1.0	1.0
Animal-waste system	0	0	.1	.1
	GS-2 watershed			
Chisel plow, spring, 20-percent residue	47	40	30	42
Chisel plow, fall, 20-percent residue	28	28	32	21
Conservation tillage, 20-percent residue	15	18	21	18
Conservation tillage, 30-percent residue	2.8	5.9	7.6	9.1
No till, 80-percent residue	.3	1.4	1.8	3.5
Contouring	.4	.8	1.2	1.2
Terraces	.9	.9	.9	1.1
Pasture management	.7	.8	1.0	1.0
Livestock exclusion	.8	1.0	1.0	1.0
Animal-waste system	0	0	0	.1
	GS-3 watershed			
Chisel plow, spring, 20-percent residue	50	44	31	43
Chisel plow, fall, 20-percent residue	27	30	35	23
Conservation tillage, 20-percent residue	14	16	20	16
Conservation tillage, 30-percent residue	2	2.2	5	7.6
No till, 80-percent residue	0	0	1.0	2.6
Contouring	.5	.5	.5	.5
Terraces	.5	.5	.5	.5
Pasture management	.6	.6	.6	.6
Livestock exclusion	.8	.8	.8	.8
Animal-waste system	0	0	0	0

Table 7. Results of statistical tests applied to before-and-after suspended-sediment storm load data for Highland Silver Lake GS-1, GS-2, and GS-3 watersheds, Madison County, Illinois

	Test statistic			Probability level		
Statistical test	GS-1	GS-2	GS-3	GS-1	GS-2	GS-3
T-test	[1]-0.17	[1]0.23	[1]-0.46	0.869	0.817	0.645
Mann-Whitney U test	400	130	150	.821	.687	.617

[1] Test performed on logarithmic transformation of original data.

Table 8. Results of step-wise regressions applied to suspended-sediment storm load data for Highland Silver Lake GS-1, GS-2, and GS-3 watersheds, Madison County, Illinois
[SS, suspended sediment; sin(x), trigonometric sine of x; cos(x), trigonometric cosine of x; P_{tot}, total precipitation; I_{30}, 30-minute maximum precipitation intensity; T, date in decimal years; <, less than]

Watershed	Sample size	Dependent variable	Independent variable(s)	Adjusted [1]R^2	Standard error[2]	Significance level
GS-1	58	SS	P_{tot}, I_{30}	0.690	72	<0.001
GS-2	34	SS	P_{tot}, sin(T), cos (T)	.671	78	<.001
GS-3	36	SS	P_{tot}, I_{30}	.593	97	<.001

[1] Fraction of variance in the dependent variable explained by the regression and adjusted for degrees of freedom.

[2] Expressed as a percentage of the mean of the dependent variable.

Table 9. Results of statistical tests applied to before-and-after suspended-sediment storm load regression residuals for Highland Silver Lake GS-1, GS-2, and GS-3 watersheds, Madison County, Illinois

	Test statistic			Probability level		
Statistical test	GS-1	GS-2	GS-3	GS-1	GS-2	GS-3
t-test	0.49	0.52	0.85	0.625	0.609	0.403
Mann-Whitney U test	380	100	120	.559	.184	.203

results are given in table 9. For the suspended-sediment storm load data, the use of regressions improved the significance levels of every test.

The failure of the statistical tests to detect significant changes was probably the result of little BMP implementation in the three watersheds. By use of the expressions for minimum detectable change developed by Spooner and others (1987) for flow-adjusted data, the change in instantaneous concentration for each of the three rural watersheds would have to be nearly 60 percent to be detectable. The degree of actual BMP implementation (table 6) would not be sufficient to result in detectable changes.

A limited Monte Carlo sensitivity analysis was used to determine the level of discrete changes in storm load that could be detected. For a given watershed, N sample values (storm load and associated independent variables) were drawn at random (with replacement) from the pre-BMP data and used to represent pre-BMP conditions. Additional

N sample values were drawn at random (with replacement) from the pre-BMP data, and the dependent variable was decreased by a fixed percentage to represent a discrete change for post-BMP conditions. The appropriate regressions were calculated, and statistical tests applied to the constructed data set. This sequence of steps was repeated 100 times for each combination of sample size and percentage decrease for each watershed. The statistical tests were applied at a significance level of 0.05. If at least 95 of the 100 replications resulted in a significant change, that particular change level (percentage decrease in the dependent variable) was considered significant.

The results of the sensitivity analysis are given in table 10. The minimum level of change resulting in at least 95-percent change detection at the 0.05 significance level is presented. For both tests examined, the results indicate that the regressions were extremely useful in aiding detection of changes in storm load. Note that the minimum detectable change decreased with sample size and reached a minimum value approximately equal to the standard error of the pre-BMP regression. This indicates "background noise" in the system and the absolute minimum level of change that can be detected.

Table 10. Minimum detectable change determined from results of Monte Carlo sensitivity analysis for suspended-sediment storm load data

[Values in percent; >, greater than]

Sample size	T-test		Mann-Whitney U test	
	Raw data	Residuals	Raw data	Residuals
	Highland Silver Lake GS-1 watershed [1]			
10	> 90	> 90	> 90	> 90
20	> 90	80	90	90
30	90	60	90	60
40	80	50	80	50
50	70	50	70	50
	Highland Silver Lake GS-2 watershed [2]			
10	> 90	> 90	> 90	> 90
20	> 90	70	90	70
30	80	50	80	50
40	80	40	70	40
50	70	40	70	40
	Highland Silver Lake GS-3 watershed [3]			
10	> 90	> 90	> 90	>90
20	> 90	80	> 90	90
30	> 90	70	90	70
40	90	50	90	50
50	80	50	80	50

[1] Regression standard error for pre-best-management-practice conditions = 52 percent.
[2] Regression standard error for pre-best-management-practice conditions = 51 percent.
[3] Regression standard error for pre-best-management-practice conditions = 48 percent.

SUMMARY AND CONCLUSIONS

Several common statistical techniques for detecting changes in water-chemistry data were selected from the literature, including parametric and nonparametric tests of hypotheses for discrete changes. Regression analyses were used to decrease the influence of natural variability. Storm loads were introduced as an alternative to periodic instantaneous concentrations. The techniques were applied to three urban and three rural watersheds; for the urban watersheds, single- and paired-site data-collection strategies were used.

The paired-site collection strategy was found to be more efficient than single-site techniques for detecting changes in the three urban watersheds. For single-site strategies, the use of regressions revealed changes that were not detected using the raw data. Differences in significant regression variables for "before-and-after" data hindered the ability of the regressions to properly reduce data variability. For urban management practices with distinct inflow and outflow components (for example, detention and retention facilities), the paired-site techniques were more effective than the single-site techniques. For management practices that do not drastically alter the relations between water chemistry and climatological variables (for example, street sweeping, catch-basin cleaning), single-site techniques applied to storm load regression residuals showed some promise and warrant further investigation.

None of the statistical techniques applied to the three rural watersheds were successful in detecting changes. This may be because of the relatively small degree of BMP implementation in the watersheds, because of the potential increased measurement error introduced through uncertain estimates of storm load, and limited sample sizes. The use of regressions with storm load data improved the significance levels of the statistical tests somewhat, helped decrease data variability, and helped to isolate measurement error and anthropogenic effects.

A Monte Carlo sensitivity analysis was used to examine the minimum change needed to detect significant discrete changes for the rural watershed data. In most instances, the minimum detectable change was smaller when regressions were used, which demonstrates the utility of decreasing the influence of natural variability.

The results presented in this report reveal the need for further research in two main areas. The application of nonparametric tests to regression residuals for storm load data needs to be explored further, particularly to develop techniques for estimating the minimum detectable change for known or estimated "before" conditions. Second, all of the techniques discussed need to be further assessed through application to additional data sets, particularly ones with extensive BMP implementation where there is a greater likelihood of significant changes.

REFERENCES CITED

Alexander, R.B., and Smith, R.A., 1988, Trends in lead concentration in major U.S. rivers and their relation to historical changes in gasoline-lead consumption: Water Resources Bulletin, v. 24, no. 3, p. 557-569.

Beale, E.M.L., 1962, Some uses of computers in operational research: Industrielle Organisation, v. 31, p. 27-28.

Berryman, David, Bobee, Bernard, Cluis, Daniel, and Haemmerli, John, 1988, Nonparametric tests for trend detection in water quality time series: Water Resources Bulletin, v. 24, no. 3, p. 545-556.

Conover, W.J., 1980, Practical nonparametric statistics (2d ed.): New York, John Wiley & Sons, 493 p.

Dickey, E.C., and Vanderholm, D.H., 1981, Vegetative filter treatment of livestock feedlot runoff: Journal of Environmental Quality, v. 10, no. 3, p. 279-284.

Dolan, D.M., Yui, A.K., and Geist, R.D., 1981, Evaluation of river load estimation methods for total phosphorus: Journal of Great Lakes Research, v. 7, no. 3, p. 207-214.

Fausey, N.R., Edwards, W.M., Owens, L.B., and White, R.K., 1988, Subsurface drained infiltration beds for feedlot runoff management: Transactions of the American Society of Agricultural Engineers, v. 31, no. 1, p. 98-101.

Hirsch, R.M., 1988, Statistical methods and sampling design for estimating step trends in surface-water quality: Water Resources Bulletin, v. 24, no. 3, p. 493-503.

Hirsch, R.M., Alexander, R.B., and Smith, R.A., 1991, Selection of methods for the detection and estimation of trends in water quality: Water Resources Research, v. 27, no. 5, p. 803-813.

Hirsch, R.M., and Slack, J.R., 1984, A nonparametric trend test for seasonal data with serial dependence: Water Resources Research, v. 20, no. 6, p. 727-732.

Hirsch, R.M., Slack, J.R., and Smith, R.A., 1982, Techniques of trend analysis for monthly water quality data: Water Resources Research, v. 18, no. 1, p. 107-121.

Iman, R.L., and Conover, W.J., 1983, A modern approach to statistics: New York, John Wiley and Sons, 497 p.

Johnson, H.P., Baker, J.L., Shrader, W.D., and Laflen, J.M., 1979, Tillage system effects on sediment and nutrients in runoff from small watersheds: Transactions of the American Society of Agricultural Engineers, v. 22, p. 1110-1114.

Kelly, M.H., and Davenport, T.E., 1986, Water resource data and trend analysis for the Highland Silver Lake comprehensive monitoring and evaluation project, Madison County, Illinois, Phase 4: U.S. Environmental Protection Agency Report IEPA/WPC/86-001, 198 p.

Lettenmaier, D.P., 1976, Detection of trends in water quality data from records with dependent observations: Water Resources Research, v. 12, no. 5, p. 1037-1046.

Mickelson, S.K., Baker, J.L., and Laflen, J.M., 1983, Managing corn residue to control soil and nutrient losses: American Society of Agricultural Engineers Paper No. 83-2161, 17 p.

Oberts, G.L., Wotzka, P.J., and Hartsoe, J.A., 1989, The water quality performance of selected urban runoff treatment systems: Metropolitan Council Report 590-89-0629, 80 p.

Porterfield, George, 1972, Computation of fluvial-sediment discharge: U.S. Geological Survey Techniques of Water-Resources Investigations book 3, chap. C2, 66 p.

Reckhow, Kenneth, and Stow, Craig, 1990, Monitoring design and data analysis for trend detection: Lake and Reservoir Management, v. 6, no. 1, p. 49-60.

Richards, R.P., 1989, Evaluation of some approaches for estimating non-point pollution loads for unmonitored areas: Water Resources Bulletin, v. 25, no. 4, p. 891-904.

Smith, R.A., and Alexander, R.B., 1983, A statistical summary of data from the U.S. Geological Survey's national water quality networks: U.S. Geological Survey Open-File Report 83-533, 30 p.

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

Smith, R.A., Hirsch, R.M., and Slack, J.R., 1982, A study of trends in total phosphorus measurements at NASQAN stations: U.S. Geological Survey Water-Supply Paper 2190, 34 p.

Smolen, M.D., Humenik, F.J., Brichford, S.L., Spooner, Jean, Lanier, Alicia, Bennett, T.B., Coffey, S.W., and Adler, K.J., 1989, NWQEP 1988 Annual Report: Status of agricultural nonpoint source projects: U.S. Environmental Protection Agency Report 506/9-89/002, 169 p.

Spooner, Jean, Jamieson, C.J., Maas, R.P., and Smolen, M.D., 1987, Determining statistically significant changes in water pollutant concentrations: Lake and Reservoir Management, v. III, p. 195-201.

Welborn, C.T., and Veenhuis, J.E., 1987, Effects of runoff controls on the quantity and quality of urban runoff at two locations in Austin, Texas: U.S. Geological Survey Water-Resources Investigations Report 87-4004, 101 p.

Wischmeier, W.H., and Smith, D.D., 1978, Predicting rainfall erosion losses - a guide to conservation planning: U.S. Department of Agriculture Handbook No. 537, 58 p.