WATER RESOURCES BULLETIN
VOL. 24, NO. 3 AMERICAN WATER RESOURCES ASSOCIATION JUNE 1988

DATA ANALYSIS OF WATER QUALITY TIME SERIES IN LAKE ERIE[1]

Keith W. Hipel, A. Ian McLeod, and Roland R. Weiler[2]

ABSTRACT: A comprehensive data analysis study is carried out for detecting trends and other statistical characteristics in water quality time series measured in Long Point Bay, Lake Erie. In order to glean an optimal amount of useful information from the available data, the exploratory and confirmatory data anslysis stages are adhered to. To test a range of hypotheses regarding the statistical properties of the time series, a wide variety of both parametric and nonparametric techniques are employed. A particularly useful nonparametric method for discovering trends is the seasonal Mann-Kendall test.
(KEY TERMS: confirmatory data analysis; environmental impact assessment; exploratory data analysis; nonparametric tests; parametric tests; trend detection; water quality.)

INTRODUCTION

When executing a complex environmental impact assessment study, usually a wide variety of statistical tests are required in order to check a range of hypotheses regarding the statistical properties of the data. The main objective of this paper is to explain clearly how both nonparametric and parametric tests can be employed in an optimal fashion to extract, systematically, information from a set of water quality time series. In particular, the effects of industrial development at Nanticoke, Ontario, upon the nearshore Lake Erie water chemistry are examined in a comprehensive statistical study. This undertaking was carried out originally by the authors in conjunction with Acres International Limited of Niagara Falls, Ontario, for the Ministry of the Environment in the Canadian province of Ontario.

In an environmental impact assessment study, one may wish, for example, to detect and model possible trends in a set of environmental time series. Unfortunately, environmental time series often possess characteristics which do not allow the series to be easily analyzed using statistical techniques (Hirsch and Slack, 1984; Hipel and McLeod, 1989). One of the major problems with environmental time series is that there are often many missing data points among which there may be long periods of time for which no observations were taken. Hydrologic and water quality data may be nonnormally distributed and follow a distribution which is usually positively skewed. In addition, environmental data are often censored by only listing measurements below a certain level as being "less than" or measurements above a specified level as being "greater than." For instance, concentration values for metals or organic compounds which fall below the limits of detection for certain chemical tests are reported simply as less than the limits of detection. As a further major complication, one or more external interventions may significantly affect the stochastic manner in which a series behaves and thereby create a variety of trends. Because of the foregoing and other reasons, environmental data are often quite "messy."

To extract an optimal amount of information from messy environmental data, a systems design approach to data analysis can be followed. As proposed by Tukey (1977) and demonstrated by McLeod, *et al.* (1983), using water quality data from Canadian rivers, the two major steps in a statistical study consist of exploratory data analysis and confirmatory data analysis. The objective of the exploratory data analysis stage is to employ simple graphical and numerical techniques to discover important patterns and statistical characteristics such as the presence of trends. The purposes of the confirmatory data analysis stage are to confirm statistically in a rigorous fashion the presence or absence of certain properties in the data. When dealing with trends, one may also wish to characterize the significant trends according to criteria such as magnitudes, shapes, and durations. Depending upon the quantity and quality of the data being analyzed, appropriate parametric and nonparametric techniques can be employed as exploratory and confirmatory data analysis tools.

Following a discussion of data analysis in the next section, the background to the case study is described. It is then explained how appropriate statistical techniques are selected in order to detect and model trends, as well as other statistical properties, in the water quality time series measured in the Nanticoke region of Lake Erie. Some representative results are presented for demonstrating the efficacy of employing appropriate parametric and nonparametric techniques within the overall framework of the exploratory and confirmatory data analysis approach to environmental impact assessment.

[1]Paper No. 87031 of the *Water Resources Bulletin*. **Discussions are open until February 1, 1989.**
[2]Respectively, Professor, Departments of Systems Design Engineering and Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1; Associate Professor, Department of Statistical and Actuarial Sciences, The University of Western Ontario, London, Ontario, Canada N6A 5B9; and Supervisor, Hazardous Contaminants, Coordination Branch, Ontario Ministry of the Environment, Suite 100, 135 St. Clair Ave. West, Toronto, Ontario, Canada M4V 1P5.

DATA ANALYSIS

Overview

In an environmental impact assessment study, usually the investigators have a fairly clear idea, at least in a general sense, of what they want to accomplish. For example, they may wish to ascertain if increased industrialization has significantly lowered the water quality of a large lake in the industrialized region. If data are not already available, a major task would be to design a suitable data collection scheme. Assuming that observations are available for a range of water quality variables at different locations in the lake, a challenging problem is to select the most appropriate set of statistical methods than can be used in an optimal fashion for detecting and modeling trends in the data. Besides uncovering and modeling trends in water quality variables at a single site, statistical methods could be used to model the relationships among variables and trends across sites in the lake. In a large scale environmental impact assessment study, it is often necessary to use a variety of both nonparametric and parametric methods. Furthermore, in some situations it may be necessary to take into account the effects of sampling bias. This type of bias could arise, for example, by only taking measurements when specified events, such as high concentrations of certain variables or high water levels, occur. Sampling bias could also take place when samples, which are taken at the same time and location in order to gage laboratory measurement errors, are considered as being independent measurements.

As noted in the introduction, when executing a data analysis study it is recommended to carry out exploratory data analysis followed by confirmatory data analysis. Usually simple graphical methods are employed at the exploratory data analysis stage for visually detecting characteristics in the data such as trends and missing values. Some of the techniques available for use at this phase include a graph of the data against time, box-and-whisker plots (Tukey, 1977, Ch. 2), Tukey smoothing (Tukey, 1977, Ch. 7), and the autocorrelation function.

Both nonparametric and parametric tests can be utilized for hypothesis and significance testing at the confirmatory data analysis stage. Some useful parametric tests include intervention analysis (Box and Tiao, 1975; McLeod, *et al.*, 1983; Hipel and McLeod, 1989). regression analysis, and the analysis of variance. As is the case for parametric statistical techniques, a wide variety of nonparametric methods are available (see, for example, Conover, 1971; and Lehman, 1975). A particularly useful nonparametric test for detecting trends is the seasonal Mann-Kendall test (Hirsch, *et al.*, 1982; Hirsch and Slack, 1984; Van Belle and Hughes, 1984) which is defined later in this section. Because of the preponderance and proliferation of statistical methods, it is not surprising that a great number of statistical textbooks have been published and a significant number of journals print papers regarding the development and application of statistical methods. In fact, at least two major encyclopediae on statistics are now available (Kruskal and Tanur, 1978; Kotz and Johnson, 1982) and a number of informative handbooks have been written (see, for instance, Sachs, 1984).

For a parametric method, such as intervention analysis (Box and Tiao, 1975), model parameters are explicitly incorporated into the design of the mathematical model in order to capture the essential characteristics of the data being modeled. Additionally, an assumption is made regarding the distribution of the noise or innovations. When there are a few missing observations and also trends in a series caused by known interventions, a properly designed intervention model can be used to estimate the missing values and also estimate the magnitudes and shapes of the trends.

In order to lessen the number of underlying assumptions required for testing a hypothesis such as the presence of a specific kind of trend in a data set, researchers developed nonparametric tests. Because a nonparametric test is a method for testing a hypothesis whereby the test does not depend upon the form of the underlying distribution of the null hypothesis, a nonparametric test is often referred to as a distribution free or distribution independent method. Nonparametric tests were developed for use in environmental impact assessment because scientists were concerned that the statistical characteristics of messy environmental data would make it difficult to use parametric procedures.

Over the years practitioners have argued about whether nonparametric or parametric tests should be employed. An advantage of nonparametric tests is that they are distribution free and, hence, fewer assumptions have to be made about the data. On the other hand, as shown for the intervention model by McLeod, *et al.* (1983), often many difficulties with the data which appear to make the series unusable with a parametric technique, can, in fact, be overcome. Cox and Hinkley (1974, Section 6.1) describe a number of drawbacks to nonparametric tests which they say limit their practical importance. One of the limitations is that when a parametric test is appropriate, a nonparametric test cannot be as powerful as the most efficient parametric test. Additionally, the results from a nonparametric test often do not adequately describe what is happening with a data set. In order to achieve a reasonable description and understanding of the system under investigation in concise and simple terms, a parsimonious parametric model is required where each parameter describes some important aspect of the system. Keeping in mind the assets and limitations of both parametric and nonparametric tests, a pragmatic approach to data analysis may be to use whatever tests seem to be most appropriate, whether they are nonparametric or parametric tests. For example, when analyzing vast amounts of environmental data for the presence of trends, in conjunction with exploratory data analysis tools, nonparametric testing can be used to locate the data which contain trends. Subsequent to a perusal of the written historical records to find physical causes for trends in the data, intervention analysis or regression analysis can be employed for obtaining rigorous statistical statements about the types and magnitudes of the trends. Montgomery and Reckhow (1984) suggest an overall systematic procedure for determining the presence or absence of trends in environmental data. Depending upon the characteristics of the data being analyzed, they suggest various nonparametric and parametric tests which can be used. Lettenmaier (1976) compares the ability of various nonparametric and parametric tests for detecting step and linear trends. Using simulation studies, Hipel, *et al.* (1986), compare the power of the autocorrelation function at lag one and the nonparametric Mann-Kendall statistic for discovering trends.

Because nonparametric tests are usually designed to indicate the presence but not the magnitude of a given statistical characteristic, some authors consider them to be exploratory data analysis procedures. Nonetheless, nonparametric tests are designed for specifically testing certain hypotheses. Since they are utilized for hypothesis testing, within this paper nonparametric tests are categorized as confirmatory data analysis tools.

To demonstrate the usefulness of both exploratory and confirmatory data analysis, specific results are presented for box-and-whisker plots and seasonal Mann-Kendall tests, respectively, for the water quality study given later in this paper. For the convenience of the reader, these two statistical methods are now briefly defined.

Box-and-Whisker Graphs

The box-and-whisker graph is based upon what is called the five-number summary (Tukey, 1977, Ch. 2). For a given data set, the five-number summary consists of the smallest and largest values, the median and the two extreme quartiles, which are called "hinges."

To assist in characterizing extreme values, Tukey (1977) has suggested the following definitions. Let "H spread" be the difference between the two hinges, and a "step" 1.5 times the H-spread. "Inner fences" are one step outside hinges and "outer fences" are two steps outside hinges. Values between an inner fence and its neighboring outer fence are called "outside." Values beyond outer fences are referred to as "far-out." When entertaining seasonal data such as monthly or quarterly data, it is instructive to calculate a five-number summary plus outside and far-out values for each season. A convenient manner in which to display this information is to plot a "box-and-whisker" diagram for each season or month. Examples of box-and-whisker graphs are shown later in Figures 3 to 6.

Seasonal Mann-Kendall Test

Mann (1945) presented a nonparametric test for randomness against time which constitutes a particular application of Kendall's test for correlation (Kendall, 1975) commonly known as the Mann-Kendall or the Kendall t statistic. Hirsch, *et al.* (1982), defined a multivariate extension of the Mann-Kendall statistic for use with seasonal data, and, as noted by Van Belle and Hughes (1984), their test possesses some similarities to tests proposed by Jonckheere (1954) and Page (1963). Although the statistic of Hirsch, *et al.* (1982), is valid for use with data where there may be missing values and also ties, assume for the present that the time series X consists of a complete record sampled over n years where there are m seasons per year such that X is given by

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

The null hypothesis, H_0, is that for each of the m seasons the n observations are independent and identically distributed while the alternative hypothesis is there is a montonic trend. Let the matrix of ranks be denoted by

$$R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ R_{21} & R_{22} & \cdots & R_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ R_{n1} & R_{n2} & \cdots & R_{nm} \end{bmatrix}$$

where the n observations for each season or column in R are ranked among themselves. Hence, the rank of x_{jg}, which is the jth data point in the gth season, is

$$R_{jg} = [n + 1 + \sum_{i=1}^{n} sgn(x_{jg} - x_{ig})]/2 \qquad (1)$$

where

$$sgn(x) = \begin{cases} +1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

and each column of R is a permutation of $(1,2,\cdots,n)$. The Mann-Kendall test statistic for the gth season is (Hirsch, *et al.*, 1982)

$$S_g = \sum_{i=1}^{n-1} \sum_{j=k+1}^{n} sgn(x_{jg} - x_{ig}) \quad g = 1,2,\cdots,m \qquad (2)$$

The statistic S_g is asymptotically normally distributed where

$$E[S_g] = 0$$

$$\mathrm{Var}[S_g] = \sigma_g^2 = n(n-1)(2n+5)/18 \qquad (3)$$

Kendall's tau statistic can be defined for the gth season as

$$\tau_g = \frac{S_g}{\frac{1}{2}n(n-1)} = \frac{S_g}{\binom{n}{2}} \qquad (4)$$

Because τ_g is simply a multiple of S_g, the distribution of τ_g can be obtained from the distribution of S_g.

Following Hirsch, *et al.* (1982), the seasonal Mann-Kendall test statistic is

$$S' = \sum_{g=1}^{m} S_g \qquad (5)$$

which is asymptotically normally distributed where

$$E[S'] = 0$$

$$Var[S'] = \sum_{g=1}^{m} \sigma_g^2 + \sum_{\substack{g,h \\ g \neq h}} \sigma_{gh} \qquad (6)$$

Using Equation (3), $\sigma_g^2 = Var[S_g]$ can be calculated and $\sigma_{gh} = cov(S_g S_h)$. For the situation where each season is independent of each of the other seasons, the second summation in Equation (6) is zero. Based upon work by Dietz and Killeen (1981), Hirsch and Slack (1984) explain how σ_{gh} can be estimated when the seasons are not independent. Their formula for estimating σ_{gh} can be written as a function of the ranks given in Equation (1).

To handle *missing values* $sgn(x_{jg} - x_{ig})$ is defined to be zero if either x_{jg} or x_{ig} is missing. Letting n_g be the number of nonmissing observations for season g, Equation (1) is modified as

$$R_{jg} = [n_g + 1 + \sum_{i=1}^{n} sgn(x_{jg} - x_{ig})]/2 \qquad (7)$$

Consequently, the ranks of the known observations remain unchanged and each missing observation is assigned the average or midrank $(n_g + 1)/2$. As is the case when there are no missing observations, Equation (2) is used to calculate S_g and following Equation (3) the variance of S_g is determined using

$$\sigma_g^2 = n_g(n_g - 1)(2n_g + 5)/18 \qquad (8)$$

S' and its variance are determined using Equations (5) and (6), respectively. If deemed necessary, σ_{gh} can be estimated using the approach of Hirsch and Slack (1984).

For censored water quality time series where some data are reported to be less than a *limit of detection,* arbitrarily fix the affected data at some constant value which is less than the limit of detection. Because nonparametric tests are based upon ranks instead of magnitudes, all censored values are interpreted as sharing the same rank which is less than the rank of all uncensored observations. Additionally, this means that handling censored data is equivalent to dealing with *ties.* Assuming, for the moment, that there are no missing values, the ranked data containing ties can be calculated using Equation (1), which automatically assigns to each of t tied values the average of the next t ranks. Following this, S_g can be determined utilizing Equation (2). The variance of S_g is

$$Var[S_g] = \sigma_g^2 = [n(n-1)(2n+5) - \sum_{j=1}^{p} t_j(t_j - 1)(2t_j + 5)]/18 \qquad (9)$$

where n is the number of years of data, p is the number of tied groups for the data x_{ig}, $i = 1,2, \cdots, n$, in season g, and t_j is the size of the jth tied group. The seasonal Mann-Kendall statistic is calculated using Equation (5) while its variance is determined by utilizing Equation (6). When there are both tied data (due to "tied" censored data and ties of actual observations) and missing values, the modifications described in this and the previous paragraph must be combined.

Before employing the seasonal Mann-Kendall statistic in Equation (5), one should examine how S_g in Equation (2) behaves for each season. Only if the same type of trend, such as an upward trend, is detected in each season, will the overall seasonal Mann-Kendall test in Equation (5) have any meaning. In other words, S' should only be calculated for a group of seasons which are expected to behave in a certain manner where hypothesis testing is done separately for this group. Besides examining the estimated S_g's across all of the seasons, a physical understanding of the problem and results from exploratory data analyses can be used to decide upon how seasons should be grouped.

Rather than using Equation (5) for combining the tests of hypotheses across seasons, another approach is to use Fisher's method (Fisher, 1970). Let the observed significance level of a one-sided test of hypothesis be denoted by SL_i. For example, because the distribution of S_g in Equation (2) is known for a given data set, one can calculate SL_g for S_g where $g = 1,2, \cdots, m$. Because of the relationship between Kendall's tau, τ_g, and S_g in Equation (4), SL_g would be the same for both τ_g and S_g in season g. When there are m independent tests, Fisher (1970, p. 99) shows that

$$-2 \sum_{i=1}^{m} \ell n SL_i \qquad (10)$$

approximately follows a χ^2 distribution with $2m$ degrees of freedom. For the situation where SL_g is the observed significance level for S_g or, equivalently, τ_g, in the gth season, the null hypothesis would be that the data for all of the seasons considered in the test come from a population where the random variables are independent and identically distributed. The alternative hypothesis is that the data across the seasons follow a monotonic trend over time. If, for example, the magnitude of the observed chi-squared variable calculated using Equation (10) is larger than the tabulated χ^2_{2m} value at a chosen significance level, one would reject the null hypothesis.

LAKE ERIE CASE STUDY

Long Point Bay is an important ecological component of Lake Erie within the Canadian province of Ontario. Its extensive areas of shoreline marsh and shallow, often weedy waters in the Inner Bay and along the north shore of Long Point, are major spawning and nursery areas of fish and continentally important staging areas for migrating water fowl. Commercial and sports fishing also possess high value on both a short-term and a long-term sustained yield basis. Therefore, deterioration of the water quality in a region having the best nearshore water quality of the three Lake Erie basins would be detrimental. A map of Long Point Bay is shown in Figure 1.

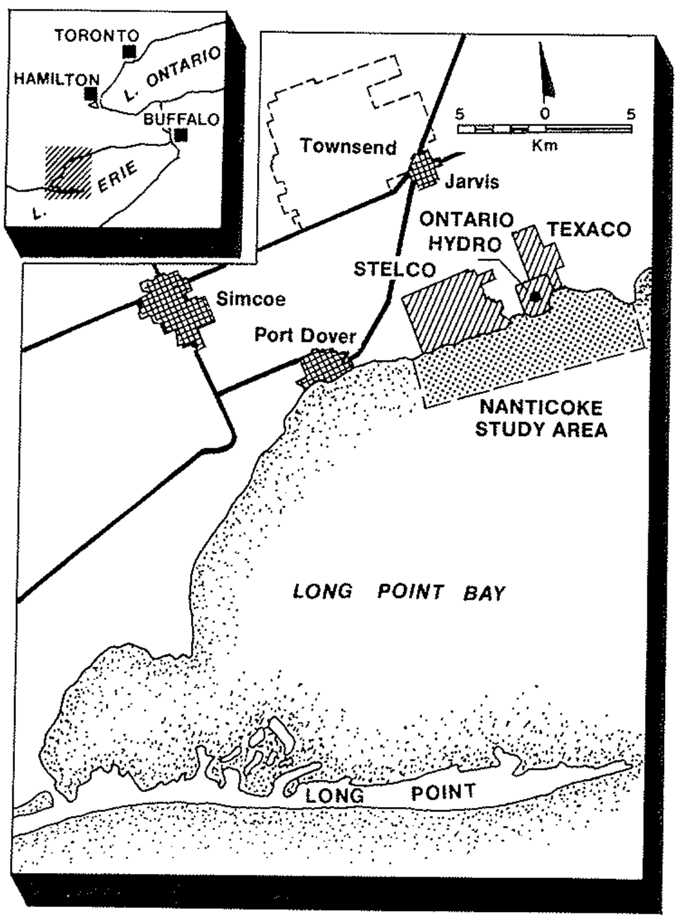

Figure 1. Location of the Lake Erie Water Quality Study.

An industrial complex which uses Long Point Bay on Lake Erie both as a water supply and a waste receiver has evolved in the area near Nanticoke since the late 1960's. In January 1972, Ontario Hydro's (the provincial company that generates almost all of the electrical power in Ontario) 4000 MW fossil-fueled thermal generating station commenced operations. Texaco Canada Inc. built a 105,000 barrel-per-day oil refinery, which began operations in November 1978. Stelco Inc. (Steel Company of Canada) chose the area for its new steel mill with an initial capacity of 1.17 million tons of steel annually. The mill operations started in May 1980. The infrastructure arising out of this development will result in an increase of population in the area which will place additional demands on the bay through an increase need for water supply and waste assimilation. The locations of the foregoing industrial activities within Long Point Bay are displayed in Figure 2.

The concern of the Ontario provincial government and the industries involved for maintenance of the water quality in the coastal zone led to the establishment of the Nanticoke Environmental Committee in 1968 to coordinate detailed investigations of the aquatic environment in the Nanticoke area. The basic objective was to establish whether the industrial complex affected the aquatic environment by monitoring any changes and to recommend, if required, appropriate mitigative action. Although the actual data gathering and analysis was carried out by the Ministry of the Environment, Ministry of Natural Resources and Ontario Hydro, both Stelco and Texaco contributed financial resources.

Selected biological, chemical, and physical measurements have been made in the region since 1968. The present program has been modified over the years and encompasses studies of currents, water temperatures, water quality, bottom fauna, phytoplankton, zooplankton, attached algae, and fish. Continuity has been achieved by using certain key sampling locations throughout the studies and maintaining the same principal study components, while making modifications to provide improvements by altering the frequency of tests and introducing new locations and study components on the basis of results obtained.

The committee has published a report (Nanticoke Environmental Committee, 1984), summarizing the study results for the period 1968-1978 and is in the process of preparing a report that covers the whole period (1968-present). These reports list the publications dealing with individual studies.

A large number of water quality variables have been measured during the study. Some of them were dropped as not being relevant to the aims of the study. The variables fall into the following groups: nutrients (phosphorus and nitrogen species), trace metals (unfiltered total), major ions (chloride, sulphate, sodium), chlorophyll, and miscellaneous (oxygen, Secchi disk, turbidity, etc.). Only the nutrients, oxygen, and Secchi disk have been measured each year. Finally, only since 1980 has the detection limit for trace metals been sufficiently low.

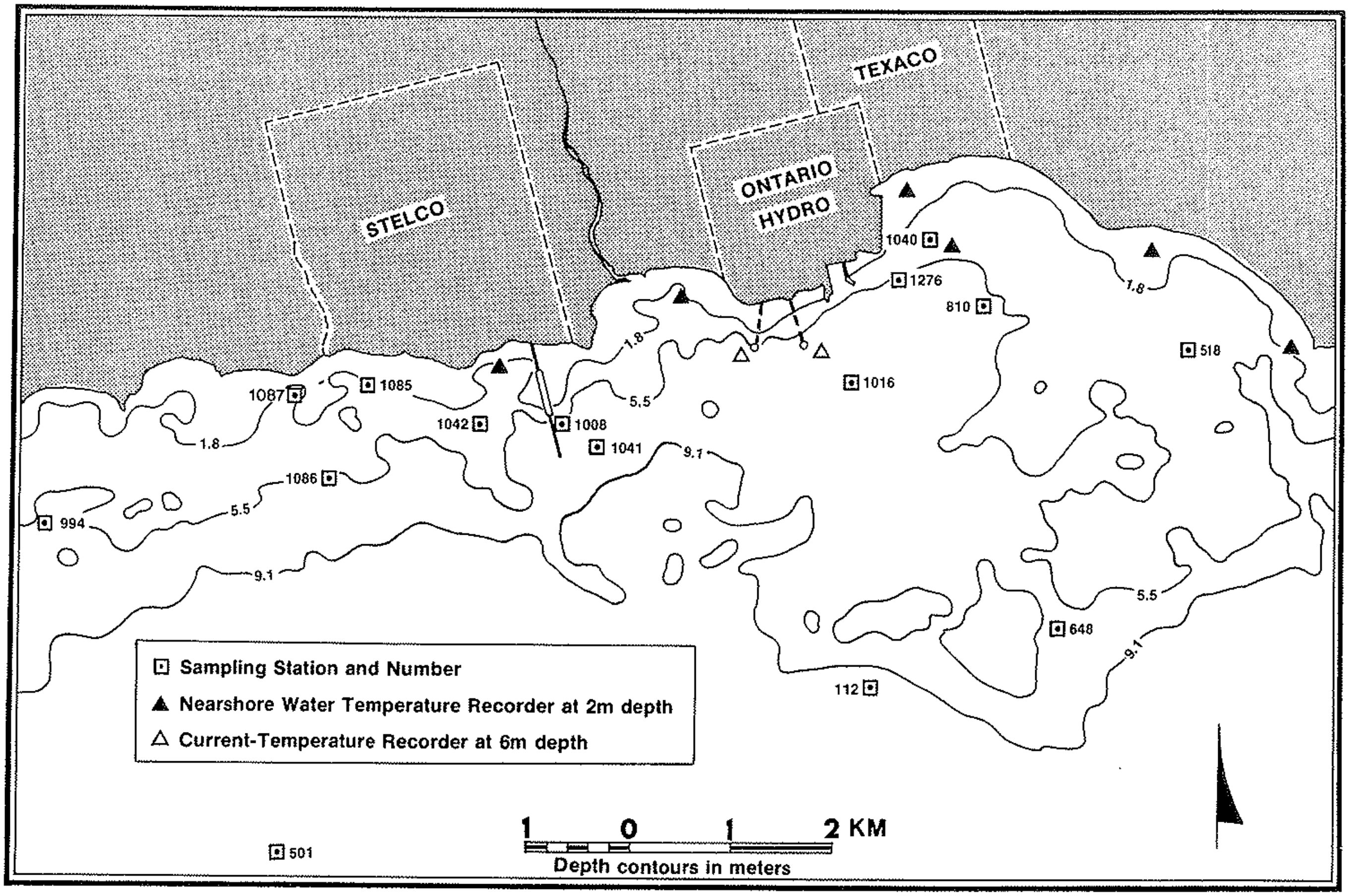

Figure 2. Sampling Stations at Long Point Bay in Lake Erie.

Water samples were taken from different depths at the stations shown in Figure 2 at intervals of two weeks to a month from 1969 to 1983. The original monitoring network consisted of eight stations. Stations 112, 501, and 648 were considered off-shore stations and each was usually sampled at two depths, near the surface and near the bottom. Stations 518, 810, 1008, and 1016 were considered near-shore stations. Stations 518 and 1008 were sampled at a depth located approximately mid-way between the surface and the bottom. Stations 810 and 1016 were sampled near the surface and near the bottom. Finally, Station 994 was considered a control station. This station was generally sampled near the surface and near the bottom.

Seven stations were added in later years of the program. In 1971, Station 1040 was established as a near shore station sampled at mid-depth, to assess the effects of the thermal plane from the generating station and of the treated effluents from the Texaso refinery. In 1975, Stations 1041 and 1042 replaced Station 1008 because the Stelco dock was built over the location of Station 1008. Stations 1085, 1086, and 1087 were added in 1978 to permit assessment of effects of Stelco treated effluent on water quality near the mouth of Centre Creek. Station 1276 was sampled in 1980 at mid-depth. Because of changes in the program only Stations 501, 994, 1085, 1086, and 1087 were sampled consistently after 1979.

The water quality samples were collected using a plastic Van Dorn bottle. The samples were transferred to specially prepared sample bottles provided by the Ministry of the Environment's Laboratory Services Branch, taken back to the laboratory and analyzed using the standard Ministry analytical techniques.

Temperature and dissolved oxygen concentrations were measured in situ with a dissolved oxygen meter. Secchi disk transparencies were measured with a black and white disk. The analytical results were entered into the Ministry's computerized database and form one of the longest unbroken time series available for water quality in Lake Erie.

DATA ANALYSIS OF THE LAKE ERIE WATER QUALITY TIME SERIES

Selection of Statistical Tests

To detect trends and uncover other statistical properties of the Lake Erie water quality time series, appropriate statistical tests must be employed. In order to have the highest probability of discovering suspected statistical characteristics which may be present in the time series, one must select the set of tests that possess the best capabilities for uncovering the specified statistical properties. To accomplish this, one must be cognizant of both the general statistical properties of the data and the main attributes of the statistical tests. For example, with respect to the characteristics of the data, one should be aware of properties such as the quantity of data, large time gaps where no measurements were taken, outliers, and data which fall below the detection limits. These properties of the data may be known in advance or else revealed using exploratory data analysis tools. From the point of view of a statistical test that can be used, one should know key facts which include the specific null and alternative hypotheses that the given statistical test is designed to check, the major distributional assumptions underlying the test, the types of samples with which the test can be used, and the kinds of measurements that can be utilized with the test. By knowing the properties of the data and the main capabilities of a wide range of both parametric and nonparametric tests, one can choose the most appropriate statistical tests for testing specified hypotheses such as the presence of trends. For the case of nonparametric tests, Conover (1971), for example, provides an informative chart for deciding upon the best nonparametric tests to utilize. Summaries and charts regarding the capabilities of parametric tests are available in many well known statistical texts. The handbook of Sachs (1984), for instance, is very helpful for locating the most appropriate parametric and nonparametric tests to employ in a given study. The tests that are eventually selected can then be used at the confirmatory data analysis stage for hypothesis testing.

The particular statistical methods used in the Lake Erie study are listed in Table 1. Notice that for each statistical method the general purpose of the technique is described and the specific reason for using it in the Lake Erie study is explained. The first four statistical methods in Table 1 constitute exploratory data analysis tools while the remaining methods are usually employed at the confirmatory data analysis stage. The nonparametric tests given in the table are marked with asterisks.

All of the statistical methods listed in Table 1 were applied to each of 14 specified water quality variables at each of the five stations, consisting of Stations 501, 810, 994, 1085, and 1086. To explain clearly how an environmental impact assessment project is carried out, some of the informative results of the Lake Erie study are now presented for the methods marked with a cross (†) in Table 1. For explaining how the techniques are used in practice, the chloride water quality (mg/ℓ) and total phosphorus (mg/ℓ) variables are used.

Representative Results

Box-and-Whisker Graphs. Figure 3 depicts box-and-whisker graphs for the total phosphorus data at Station 501 in Figure 2. In this figure the data have not been transformed using a Box-Cox transformation (Box and Cox, 1964). The upper and lower ends of a rectangle for a given month represent the two hinges and the thick line drawn horizontally within each rectangle is the value of the median. The minimum and maximum values in a particular month are the end points of the lines or "whiskers" attached to the rectangle or "box." The far-out values are indicated by a circle in Figure 3. Below each month is a number which gives the number of data points used to calculate the box-and-whisker graph above the month. The total number of observations across all the months is listed below November and December. When there are not many data points used to determine a box-and-whisker plot for a given month, any peculiarities in the plot should be cautiously considered.

For a given month in a box-and-whisker diagram, symmetric data would cause the median to lie in the middle of the rectangle and the lengths of the upper and lower whiskers would be about the same. Notice in Figure 3, for the total phosphorus data at Station 501, that the whiskers are almost entirely above the rectangles for almost all of the months and there are eight far-out values above the boxes. This lack of symmetry can at least be partially rectified by transforming the given data using natural logarithms, which is a special type of Box-Cox transformation. By comparing Figure 3 to Figure 4, where natural logarithms are taken of the total phosphorus data, the improvement in symmetry can be clearly seen. Furthermore, the natural logarithmic transformation has reduced the number of far-out entries from eight in Figure 3 to five in Figure 4.

Box-and-whisker plots can be employed as an important exploratory data analysis tool in intervention studies. If the date of the intervention is known, box-and-whisker diagrams can be constructed for each season for the data before and after the time of the intervention. These two graphs can be compared to ascertain for which seasons the intervention has caused noticeable changes. When there is sufficient data, this type of information is crucial for designing a proper intervention model to fit to the data at the confirmatory data analysis stage (McLeod, *et al.*, 1983).

For the Nanticoke data, there are two major interventions. First, Ontario Hydro built a fossil-fueled electrical generating plant which began operating in January 1972, and came into full operation by about January 1, 1976. Because not much data are available before 1972, January 1, 1976, is taken as the intervention date at which water quality measurements near the Ontario Hydro plant may be affected. Of the five stations analyzed, only Station 810 is close to the plant. For each of the water quality variables measured at Station 810, box-and-whisker plots are made for before and after the intervention date in order to qualitatively discover any possible statistical impacts of the intervention.

Second, the Steel Company of Canada (Stelco) plant came into operation about April 1, 1980. Because sites 501, 994, 1085, and 1086 are relatively close to the Stelco factory, box-and-whisker graphs are made for before and after the

TABLE 1. Statistical Methods Used in the Lake Erie Water Quality Study.

Method	General Purpose	Specific Purpose in the Study
Data listing	For each series, want to know exact values, dates of measurements, depths of measurements, and station number.	Same as under general purpose.
Graphs of the data	Visually detect main statistical characteristics of a series.	Visually detect trends and see spacing of the observations.
Tukey five number summary†	Describe how the observations in a series are distributed in each season.	Describe how measurements in a water quality series are distributed in each month.
Box-and-whisker graphs†	See a graphical display of the five number summary for each season in a series.	See a plot of the five numbers for each month in a series.
Seasonal Mann-Kendall test*†	Check for trends in a series for each season of the year.	Check for trends in a series for each month (see Equation 2).
Fisher's combination test†	Combining tests of hypotheses for one-sided tests.	Test for a trend across all the months in a series by combining the significance levels from the seasonal Mann-Kendall tests for each month into a χ^2 statistic (see Equation 10).
Wilcoxon signed rank test*	Determine whether the medians of two samples are the same.	Find out whether or not measurements taken at two different depths at exactly the same time possess the same median.
Confidence interval for the median*	Calculate a confidence interval for the difference in the medians between two samples.	Calculate 95 percent confidence interval for the difference in medians between measurements taken at two different depths at exactly the same time. If zero is not contained in the confidence interval, the two medians are significantly different from one another.
Kendall's rank correlation*	Determine whether or not two series are independent of one another (Kendall, 1975, Ch. 2).	Ascertain if measurements taken at the same time at alternate depths are correlated with one another.
Cross correlation function	Determine whether or not two series are independent of one another (Kendall, 1975, Ch. 2).	Find out if measurements taken at the same time at alternate depths are correlated with one another.
Pitman's test for equality of correlated variance	Ascertain whether or not two correlated variances are the same (Pitman, 1939).	Determine if the variances of samples taken at two different depths are the same.
One way analysis of variance	Determine if the means across k samples are significantly different from one another (Sachs, 1984, pp. 501-509). It is assumed that the k populations are normally independently distributed and have equal variances.	Find out whether or not the means among replicated samples are the same.
Kruskal-Wallis Test*	Nonparametric test to check whether or not the distributions or means across k samples are the same. The observations are assumed to be independent of one another and follow the same distribution.	Determine if the means among replicated samples are the same.
Regression analysis	Parametrically model relationships within a series and among series.	Determine the best data transformation, ascertain the components required in a regression model, estimate both the average monthly and annual values for a series, etc.

*Nonparametric test.
†Applications for this method are given in this paper.

intervention for each water quality time series at each station. Figures 5 and 6 display the box-and-whisker graphs for the natural logarithms of the total phosphorous data for before and after the Stelco intervention, respectively. The dates in brackets in the titles for Figures 5 and 6 indicate the intervals of time for which measurements were taken before and after the intervention, respectively. When these two graphs are compared, it appears that for most of the months there is a slight drop in the median level after the intervention. Using an intervention model based upon a special type of regression analysis model, a confirmatory data analysis could be executed to ascertain the magnitudes of the changes in the monthly means and if they are significant. Furthermore, one should also take into account overall changes in Lake Erie by considering measurements at locations outside of the Nanticoke region.

A third intervention in the Nanticoke region is due to the Texaco oil refinery which began production in November 1978. Because the discharge from this plant is relatively quite small, the possible effects of the Texaco intervention are not considered in this study.

Seasonal Mann-Kendall Tests. For a given water quality variable at a specified station, the seasonal Mann-Kendall test is used to detect trends in each month of the year. A detailed description of this test is given earlier using Equations (1) to (10). For the case of the chloride measurements taken at Station 501, Table 2 presents results of the seasonal Mann-Kendall and other related tests. Each entry in the table of years versus months is the median value for a given month and year. By utilizing Equation (4) for the data in a specific month across all of the years for which data are available, Kendall's tau can be determined. The observed value of S_g for each month which is calculated using Equation (2) is not displayed in the table. Because the observed τ_g value for each month is negative, this indicates that there

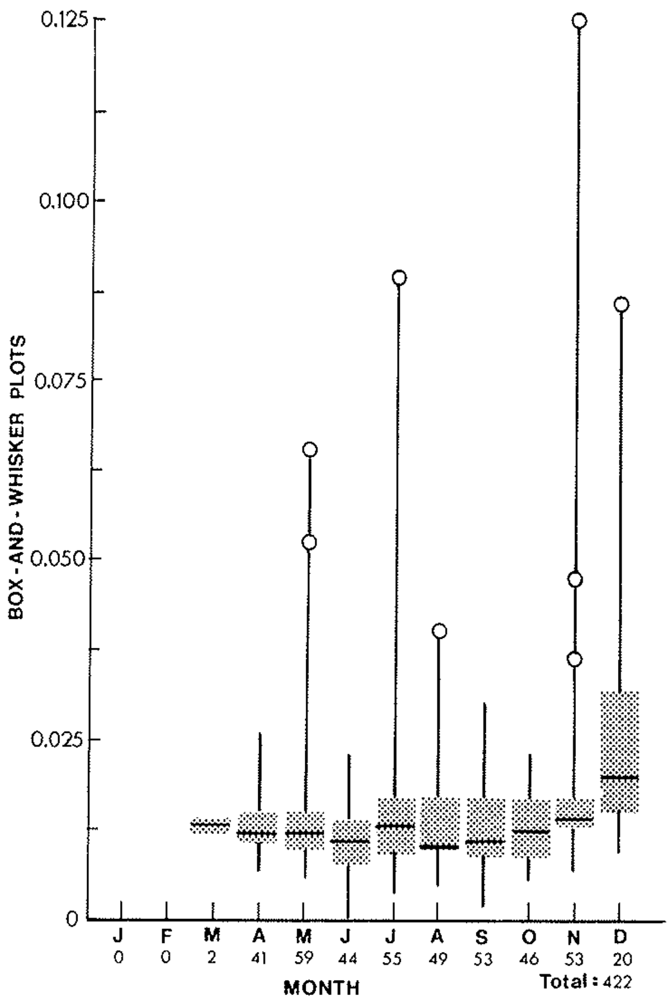

Figure 3. Box-and-Whisker Plots of the Total Phosphorous (mg/ℓ) Data at Station 501, Long Point Bay, Lake Erie, from April 22, 1969, to December 13, 1983.

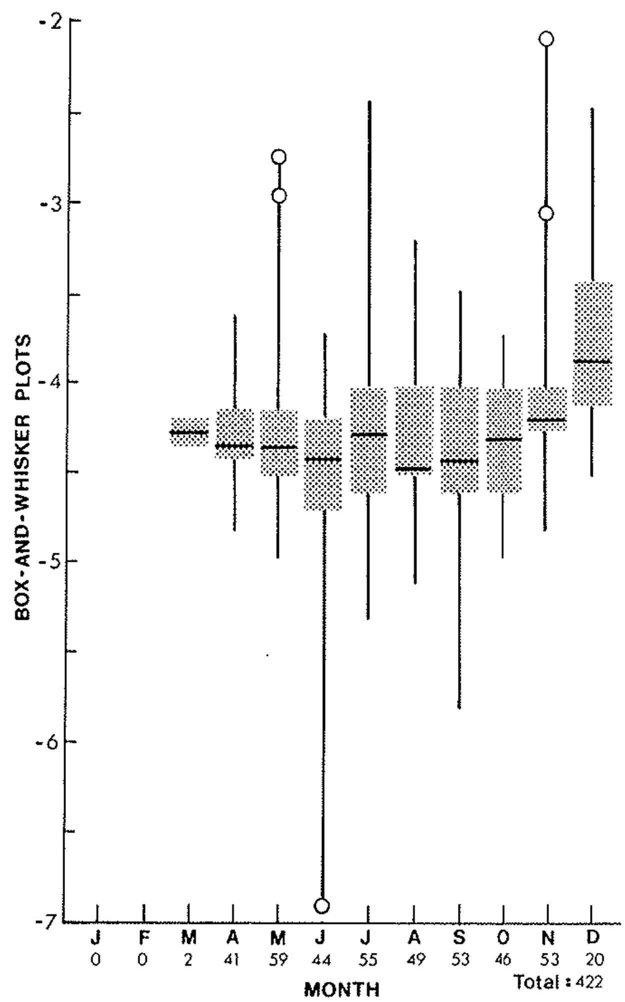

Figure 4. Box-and-Whisker Plots of the Logarithmic Total Phosphorous (mg/ℓ) Data at Station 501, Long Point Bay, Lake Erie, from April 22, 1969, to December 13, 1983.

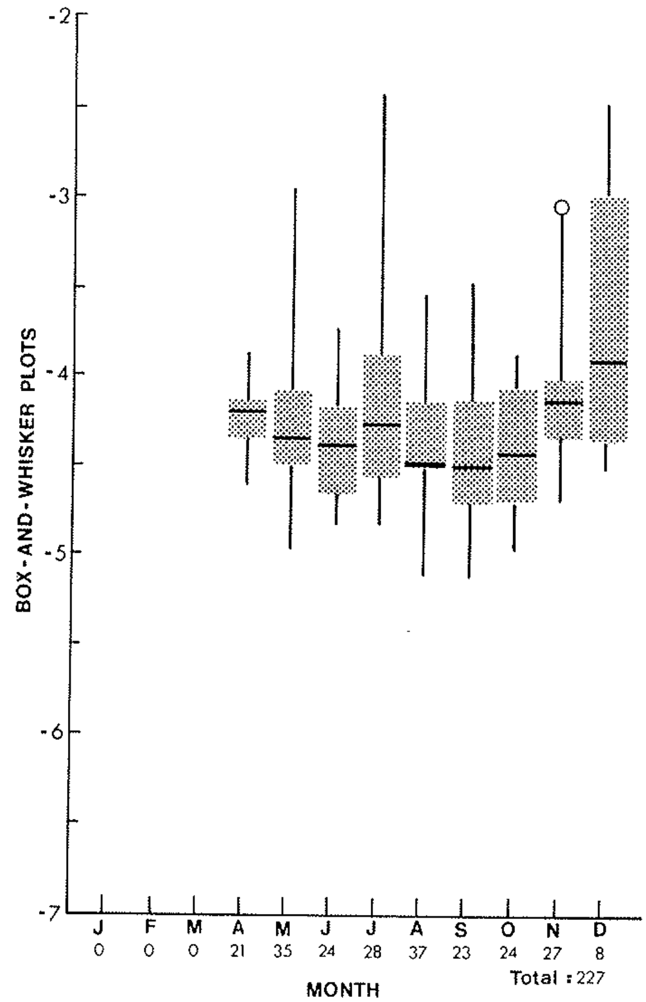

Figure 5. Box-and-Whisker Plots of the Logarithmic Total Phosphorous Data (mg/ℓ) at Station 501, Long Point Bay, Lake Erie, Before April 1, 1980 (data available from April 22, 1969, to November 19, 1979).

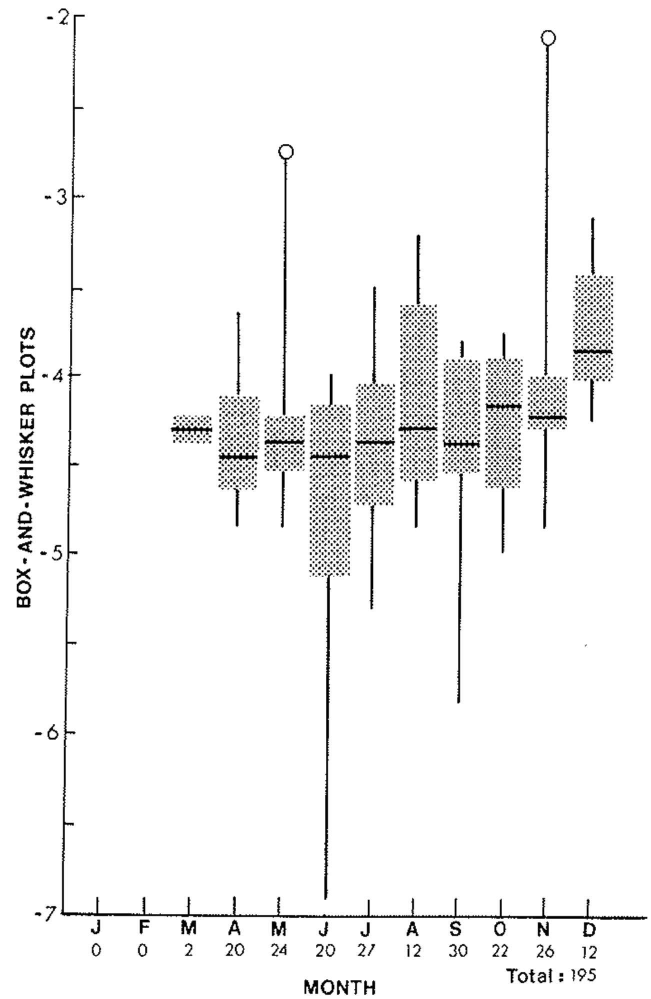

Figure 6. Box-and-Whisker Plots of the Logarithmic Total Phosphorous (mg/ℓ) Data at Station 501, Long Point Bay, Lake Erie, After April 1, 1980 (data available from April 13, 1981, to December 13, 1983).

may be a decreasing trend in each month. Consider, for instance, the month of April for which the calculated τ_g value is –0.98. Since the significance level (SL) for this month is 0.39 percent, this strongly suggests that the null hypothesis of having identically independently distributed data should be rejected in favor of accepting the alternative hypothesis of there being a monotonic decreasing trend. Notice that for each month the significance level is not greater than 5 percent and usually less than 1 percent. Consequently, one would expect that across all of the seasons, a combination test would confirm the presence of an overall decreasing trend. The seasonal Mann-Kendall test statistic in Equation (5) has a magnitude of –196.0 and a very small significance level. In addition, Table 2 shows that the significance level of Fisher's combination test in Equation (10) is also very small. Hence, both of the combination tests indicate that there is an overall trend which is decreasing due to the negative sign of S_g or τ_g in each season and also S' in Equation (5) across all of the seasons.

For each of the 14 water quality variables at each of the five stations where there are sufficient data, the seasonal Mann-Kendall test in Equations (2) or (4) is applied. Consider Table 3 which summarizes the results for chloride. Notice that at Stations 501, 810, and 994 there are obvious decreasing trends (indicated by the negative signs) for all the months for which data are available at all three sites. Except for two cases where the significance level is a = 10 percent, all of the significance levels are 5 percent or less. Consequently, these trends are significant. For a given month and station, one should certainly reject the null hypothesis that the chloride data are independently and identically distributed. Table 3 shows that sufficient data for executing

TABLE 2. Trend Analysis of Monthly Median Values Using the Seasonal Mann-Kendall Test for the Chloride Series (mg/ℓ) at Station 501, Long Point Bay, Lake Erie.

Year	Monthly Median Values Times 10											
	January	February	March	April	May	June	July	August	September	October	November	December
1970							270	260				
1971					350	250	240	235	250	250		
1972				250	250	240	240	245	240		230	
1973						235	230	235	245	240	245	
1974				230	238	200	211	225	205	210		220
1975				215	230		225	225	220	215	215	
1976				205	205	207	210	210	210	210	215	
1977				205	205	210	205	215	205		210	
1978				197	200	200	200	205	202	210	205	210
1979				195	200	205	200		195	195	190	
tau				−0.98	−0.96	−0.62	−0.93	−0.86	−0.82	−0.82	−0.88	
SL(%)	0.39			0.39	0.17	4.61	0.03	0.22	0.33	1.87	0.98	

Combination of Scores and Their Variances

Sum	Variance	SL
−196.0	5.59900×10^{2}	1.13324×10^{-16}

Fisherian Combination of the Significance Levels

CHI-SQ	DF	SL
87.01	16	0.000×10^{-1}

a seasonal Mann-Kendall test are not available for the months of January, February, and March at Stations 501, 810, and 994. Also, there are not enough observations for all of the months at Stations 1085 and 1086.

TABLE 3. Seasonal Mann-Kendall Tests for Trend in the Chloride Series (mg/ℓ) for Stations at Long Point Bay, Lake Erie.

Months	Stations				
	501	810	994	1085	1086
January					
February					
March					
April	−c	−c	−b		
May	−c	−c	−d		
June	−b	−a	−a		
July	−d	−c	−c		
August	−c	−c	−c		
September	−c	−c	−c		
October	−b	−c	−c		
November	−c	−d	−d		
December					

NOTES:
1. a, b, c, and d denote significance levels of 10, 5, 1, and 0.1 percent, respectively.
2. A blank indicates insufficient data.
3. A positive or negative value of tau is indicated by + or −.

Another water quality variable for which there are decreasing monthly trends is specific conductance for Stations 501, 810, and 994. However, as is the case for chlorophyll *a*, for most of the water quality variables across most of the months and stations, significant trends are not detected by the seasonal Mann-Kendall test.

As explained earlier, for a specified water quality variable and station one can combine the monthly Mann-Kendall test results using the combined score method in Equation (5) and the Fisherian combination in Equation (10). Table 4 summarizes these two types of combination results for the 14 water quality variables across all of the stations. Notice in Table 4 for the chloride variable that the results are highly significant for Stations 501, 810, and 994 for both combination tests (all the significance levels are d = 0.1 percent). Consequently, there are obvious trends in chloride across the months at all three sites. Due to the negative signs in Table 3, the trends are decreasing.

CONCLUSIONS

Environmental data, such as water quality time series, are often very messy. For example, water quality time series may have problems which include having missing observations, following nonnormal distributions, possessing outliers, and being short in length. Because nonparametric tests usually have less restrictive assumptions than their parametric counterparts, nonparametric tests are often ideally suited for detecting characteristics such as trends in environmental data. Furthermore, because of the increasing importance of environmental impact assessment studies in modern day society, the import of both nonparametric and parametric tests will continue to expand (Hipel and McLeod, 1989).

To demonstrate clearly the efficacy of employing nonparametric and also parametric methods in a complex environmental impact assessment study, the effects of industrial development upon water quality in Long Point Bay in Lake Erie, are systematically examined. The specific statistical methods used in the application for exploratory and confirmatory data analyses are listed in Table 1. Within the paper, the method of application and representative results are given for each of the techniques marked by a cross in Table 1. Of particular importance is the seasonal Mann-Kendall test that is used to check for the presence of trends in a range of water quality variables at different sites (see Tables 2 to 4).

TABLE 4. Combined Score Tests from Equation (5) and Fisher's Combination Results Using Equation (10) for the 14 Water Quality Variables at Long Point Bay, Lake Erie.

	Stations				
	501	810	994	1085	1086
Turbidity (FTU)	a,†	†,†	c,b	†,†	†,†
Specific Conductance (μs/cm)	d,d	d,d	d,d	†,†	†,†
Lab pH	†,†	†,†	†,†	†,d	†,†
Chloride (mg/ℓ)	d,d	d,d	d,d		
Ammonia-N (mg/ℓ)	d,a	c,b	b,†	†,†	†,†
Inorganic-N (mg/ℓ)	†,†	†,†	†,†	†,d	
Filtered Total Kjeldahl N	c,†	†,c	a,†	b,†	†,a
Kjeldahl Organic-N (mg/ℓ)	b,†	†,a	†,†	†,†	†,a
Chlorophyll a	†,†	b,†	a,†	a,†	†,d
Chlorophyll b	c,†	d,a	†,†	†,†	†,†
Phytoplankton Density					
Filtered Reactive Phosphate	d,d	b,†	d,c	†,†	†,†
Total Phosphorous (mg/ℓ)	d,a	a,b	d,c	†,†	†,†
Iron (mg/ℓ)	d,b	†,a	c,b	†,†	†,†

NOTES:

1. a, b, c, and d denote significance levels of 10, 5, 1, and 0.1 percent, respectively.
2. † denotes result not significant at the 10 percent level.
3. A blank indicates insufficient data.
4. Combined Kendall tests for trend – First Entry: Combined Score Method; Second Entry: Fisherian Combination.

ACKNOWLEDGMENTS

The authors appreciate the advice and support given by scientists such as Mr. J. W. Tanner and Dr. I. K. Hill at Acres International Limited in Niagara Falls, Ontario, as well as scientists at the Ministry of the Environment in Toronto. The statistical studies presented in this paper are part of an overall water quality study of the Nanticoke region which was carried out by Acres International Limited and funded by the Ministry of the Environment.

LITERATURE CITED

Box, G. E. P. and D. R. Cox, 1964. An Analysis of Transformations. Journal of the Royal Statistical Society, Series B, 26:211-252.

Box, G. E. P. and G. C. Tiao, 1975. Intervention Analysis with Applications to Economic and Environmental Problems. Journal of the American Statistical Association 70(349):70-79.

Conover, W. J., 1971. Practical Nonparametric Statistics. John Wiley, New York, New York.

Cox, D. R. and D. V. Hinkley, 1974. Theoretical Statistics. Chapman and Hall, London, England.

Dietz, E. J. and T. J. Killeen, 1981. A Nonparametric Multivariate Test for Monotone Trend with Pharmaceutical Applications. Journal of the American Statistical Association 76(373):169-174.

Fisher, R. A., 1970. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh, England.

Hipel, K. W. and A. I. McLeod, 1989. Time Series Modelling for Water Resources and Environmental Engineers. Elsevier, Amsterdam (in press).

Hipel, K. W., A. I. McLeod, and P. K. Fosu, 1986. Empirical Power Comparisons of Some Tests for Trend. *In:* Statistical Aspects of Water Quality Monitoring, A. H. El-Shaarawi and R. E. Kwiatkowski (Editors). Elsevier, Amsterdam, pp. 347-362.

Hirsch, R. M. and J. R. Slack, 1984. A Nonparametric Trend Test for Seasonal Data with Serial Dependence. Water Resources Research 20(6):727-732.

Hirsch, R. M., J. R. Slack, and R. A. Smith, 1982. Techniques of Trend Analysis for Monthly Water Quality Data. Water Resources Research 18(1):107-121.

Jonckheere, A. R., 1954. A Distribution-Free k-Sample Test Against Ordered Alternatives. Biometrika 41:133-145.

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

Kotz, S. and N. L. Johnson (Editors), 1982. Encyclopedia of Statistical Sciences, Volumes 1 to 9 (first volume published in 1982). Wiley, New York, New York.

Kruskal, W. H. and J. M. Tanur, 1978. International Encyclopedia of Statistics, Volumes 1 and 2. The Free Press, New York, New York.

Lehman, E. L., 1975. Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, Oakland, California.

Lettenmaier, D. P., 1976. Detection of Trends in Water Quality Data from Records with Dependent Observations. Water Resources Research 12(5):1037-1046.

Mann, H. B., 1945. Nonparametric Tests Against Trend. Econometrica 13:245-259.

McLeod, A. I., K. W. Hipel, and F. Camacho, 1983. Trend Assessment of Water Quality Time Series. Water Resources Bulletin 19(4):537-547.

Montgomery, R. H. and K. W. Reckhow, 1984. Techniques for Detecting Trends in Lake Water Quality. Water Resources Bulletin 20(1):43-52.

Nanticoke Environmental Committee, 1984. The Aquatic Environment of Long Point Bay in the Vicinity of Nanticoke on Lake Erie, 1968-1978.

Page, E. G., 1963. Ordered Hypotheses for Multiple Treatments: A Significance Test for Linear Ranks. Journal of the American Statistical Association 58:216-230.

Pitman, E. J. G., 1939. A Note on Normal Correlation. Biometrika 31:9-12.

Sachs, L., 1984. Applied Statistics, A Handbook of Techniques. Springer-Verlag, New York, New York, Second Edition.

Tukey, J., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, Massachusetts.

Van Belle, G. and J. P. Hughes, 1984. Nonparametric Tests for Trend in Water Quality. Water Resources Research 20(1):127-136.