National Nonpoint Source Monitoring Program

Tech NOTES 6

November 2011

Donald W. Meals, Jean Spooner, Steven A. Dressing, and Jon B. Harcum.
2011. Statistical analysis for monotonic trends, Tech Notes 6, November 2011. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA, 23 p. Available online at *https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoring-technical-notes*.

Through the National Nonpoint Source Monitoring Program **(NNPSMP)**, states monitor and evaluate a subset of watershed projects funded by the Clean Water Act Section 319 Nonpoint Source Control Program.

The program has two major objectives:

1. To scientifically evaluate the effectiveness of watershed technologies designed to control nonpoint source pollution
2. To improve our understanding of nonpoint source pollution

NNPSMP Tech Notes is a series of publications that shares this unique research and monitoring effort. It offers guidance on data collection, implementation of pollution control technologies, and monitoring design, as well as case studies that illustrate principles in action.

Statistical Analysis for Monotonic Trends

Introduction

The purpose of this technical note is to present and demonstrate the basic analysis of long-term water quality data for trends. This publication is targeted toward persons involved in watershed nonpoint source monitoring and evaluation projects such as those in the National Nonpoint Source Monitoring Program (NNPSMP) and the Mississippi River Basin Initiative, where documentation of water quality response to the implementation of management measures is the objective. The relatively simple trend analysis techniques discussed below are applicable to water quality monitoring data collected at fixed stations over time. Data collected from multiple monitoring stations in programs intentionally designed to document response to treatment (e.g., paired-watershed studies or above/below-before/after with control) or using probabilistic monitoring designs may need to apply other techniques not covered in this technical note.

Trend analysis can answer questions like: "Are streamflows increasing as urbanization increases?" or "Have nutrient loads decreased since the TMDL was implemented?"

Trend Analysis

For a series of observations over time—mean annual temperature, or weekly phosphorus concentrations in a river—it is natural to ask whether the values are going up, down, or staying the same. Trend analysis can be applied to all the water quality variables and all sampling locations in a project, not just the watershed outlet or the receiving water.

Broadly speaking, trends occur in two ways: a gradual change over time that is consistent in direction (monotonic[1]) or an abrupt shift at a specific point in time (step trend). In watershed monitoring, the questions might be "Are streamflows increasing as urbanization increases?" [a monotonic trend] or "Did nonpoint source nutrient loads decrease after the TMDL was implemented in 2002?" [a step trend]. When a monitoring project involves widespread implementation of best management practices (BMPs), it is usually desirable

[1] Linear trends are a subset of monotonic trends.

to know if water quality is improving: “Have suspended sediment concentrations gone down as conservation tillage adoption has gradually increased?” [a monotonic trend] or “Has the stream macroinvertebrate community improved after cows were excluded from the stream with fencing in 2005?” [a step trend]. If water quality is improving, it is also important to be able to state the degree of improvement.

Trend analysis has advantages and disadvantages for the evaluation of nonpoint source projects, depending on the specific situation (Table 1). Simple trend analysis may be the best—or only—approach to documenting response to treatment in situations where treatment was widespread, gradual, and inadequately documented, or where water quality data are collected only at a single watershed outlet station. For data from a short-term (e.g., 3 years) monitoring project operated according to a paired-watershed design (Clausen and Spooner 1993), analysis of covariance (ANCOVA) using data from the control watershed may be more appropriate than trend analysis to evaluate response to treatment because it directly accounts for the influences of climate and hydrology in a short-term data set. In contrast, for a long data record from a single watershed outlet station, trend analysis may be the best approach to evaluate gradual change resulting from widespread BMP implementation in the watershed in the absence of data from a control site.

Table 1. Advantages and disadvantages of simple trend analysis as the principal approach for evaluation of nonpoint source monitoring projects.

Advantages	Disadvantages
Can be done on data from a single monitoring station	Usually requires long, continuous data record
Does not require calibration period	Difficult to account for variability in water quality data solely related to changes in land treatment or land management
Applicable to large receiving waterbodies that may be subject to many influences	Not as powerful as other watershed monitoring designs that have baseline (or pre-BMP data) with controls (e.g., control watershed or upstream data), especially with small sample sizes
Useful for BMPs that develop slowly or situations with long lag times	Provides no insight into cause(s) of trend

The application of trend analysis to evaluate the effects of a water quality project depends on the monitoring design. Data from a watershed project that uses an upstream/downstream or before/after study design where intensive land treatment occurs over a short period generating an abrupt or step change may be evaluated for a step trend using a variety of parametric and nonparametric tests including the two sample t-test, paired t-test, sign test, analysis of (co)variance, or Kruskal-Wallis test. In general, these tests are most applicable when the data can be divided into logical groups.

On the other hand, data from long-term, fixed-station monitoring programs where gradual responses such as those due to incremental BMP implementation or continual urbanization are of most interest, are more aptly evaluated with monotonic trend analyses that correlate the response variable (i.e., pollutant concentration or load) with time or other independent variables. These types of analyses are useful in situations where vegetative BMPs like the riparian buffers implemented in the Stroud Preserve NNPSMP project (Newbold et al. 2008) must mature, resulting in gradual effects expressed over time. Trend analysis of data collected in a large receiving waterbody such as a lake or estuary may be the principal way of evaluating large, complex watershed programs. The Chesapeake Bay Program has conducted trend analyses since the early 1990s to detect and quantify water quality responses in the Bay to nutrient reduction actions to measure progress toward Bay restoration goals (CBP 2007). While the examples in this technical note focus on detection of changes in concentration of individual pollutants with respect to time, these tools can also be used when evaluating the relationship between variables such as chlorophyll and nutrients.

Trend analysis needs to account for the variability in water quality data that can be due to many factors, including:

- Seasonal cycles;
- Diurnal cycles;
- Variations in hydrology and weather;
- Human activities and management;
- Measurement error;
- Natural variability; and
- Actual trends

> Statistical trend analysis can help to identify trends and estimate the rate of change, but will not provide much insight in attributing a trend to a particular cause. Interpreting the cause of a trend requires knowledge of the watershed, and a specific study design.

The task of trend analysis is to characterize and account for other sources of variation and to identify and quantify the actual trend in a statistically rigorous way.

It is important to recognize some other limitations of trend analysis. Trend analysis is more effective with longer periods of record. Short monitoring periods and small sample sizes make documentation of trends more difficult. Most importantly, the statistical methods discussed below can help identify trends and estimate the rate of change, but will not provide much insight in attributing a trend to a particular cause. Interpreting the cause of a trend requires knowledge of hydrologic processes, land use, and human activities in the watershed. Establishing causality requires a different study design.

Finally, in looking for trends in water quality, it is important to recognize that some increasing or decreasing patterns in water quality especially over short time periods are **not** trends. Many water quality variables exhibit seasonality as a result of temperature, precipitation, and flow. A snapshot of water quality data from a few months may suggest an increasing trend, while examination of an entire year shows this "trend" to be part of a regular cycle associated with temperature, precipitation, or cultural practices. Autocorrelation—the tendency for the value of an observation to be similar to the observation immediately before it—may also be mistaken for a trend over the short term. Changes in sampling schedules, field methods, personnel, or laboratory practices may also cause shifts in data that could be erroneously interpreted as trends. Characterization of project data through exploratory data analysis (see ***Tech Note #1***) will help recognize and account for such features in a dataset.

All increasing or decreasing patterns in water quality are not trends. Characterize your data to avoid misinterpreting seasonal cycles, autocorrelation, or changes in monitoring methods as significant trends.

General Considerations

Is a simple trend analysis appropriate?

The first step in trend analysis is to decide if it is an appropriate tool for answering the questions you have about project data. Effective trend analysis requires a fairly long sequence of data collected at a fixed location, collected by consistent methods, with few long gaps. It has been suggested that five years of monthly data are the minimum for monotonic trend (continuous rate of change, increasing or decreasing) analysis; for a step trend (abrupt shift up or down), at least two years of monthly data before and after treatment are required (Hirsch 1988). These time frames are only guidelines; longer periods of record and/or more intensive sampling frequency would generally provide a greater sensitivity to detect smaller changes. Trend analysis is best suited for a situation where the land treatment program has been successful in implementing BMPs over an extensive portion of the critical area, implementation occurs over several years, and water quality change is expected to be gradual.

The water resource type, project design, type of land treatment, and implementation schedule largely determine the type of trend to be expected. Most of the trend analysis techniques discussed in this publication apply to the evaluation of a monotonic trend, the kind of change that might be expected in response to gradual, widespread implementation of BMPs. Step trends may occur in response to an abrupt change in the watershed, such as the completion of a detention pond or a ban on winter manure application. To properly evaluate a step trend, it is critical to have a solid *a priori* hypothesis concerning when the step change took place; examination of the data themselves to search for the best place to locate a shift is inappropriate. Although techniques exist for testing for step trends, in many cases a two-sample test (e.g., t-test of before vs. after) may be a better choice when an abrupt change at a specific point in time is expected.

Explore the data first

Before beginning trend analysis, define the question that needs to be answered and then conduct exploratory data analysis (EDA) on the data set (see ***Tech Note #1***). EDA will often give preliminary indications of trends and set the stage for further trend analysis. Use EDA to evaluate how well the data satisfy assumptions of parametric statistical analysis (normal distribution, constant variance, and independence), evaluate the effectiveness of transformations, and characterize relationships between variables. EDA can reveal important explanatory variables (covariates) like flow or precipitation that drive dependent variables at this point. Some trend analysis techniques can account for covariates.

Evaluate the data set for significant missing observations, such as a year-long interruption in the middle of a 7-year program. Some techniques are sensitive to gaps in data collection. If a long gap exists in the data, step trend procedures (e.g., assessing the difference in sample means between the two periods using a two-sample t-test) may be more appropriate than the monotonic trend analysis techniques discussed below. Although there is no specific decision rule, Helsel and Hirsch (1992) advise using step trend rather than monotonic trend analysis if a data gap is greater than one-third of the total record.

Select variables

Trend analysis can be applied to all the water quality variables and all sampling locations in the project. In large projects tracking many variables at many stations, this can be a daunting task. If full analysis is not feasible, there are several options. First, a subset of monitored variables can be selected, focusing on those expected to be most responsive to land treatment or those that directly relate to water quality impairment. Alternatively, it may be possible to use an index that combines information from a number of variables, such as the Index of Biotic Integrity (IBI) for stream fish communities (Karr 1981), or the Oregon Water Quality Index (OWQI) that integrates measurements of temperature, dissolved oxygen, BOD, pH, ammonia+nitrate nitrogen, total phosphate, total solids, and fecal coliform (Cude 2005). Third, overall water quality trends have been efficiently assessed and presented by conducting trend analysis on principal components as surrogate variables for individual water quality constituents (Ye and Zou 1993).

Data reduction and flow adjustment

Before proceeding to trend analysis tests, it may be necessary or beneficial to perform some preliminary data reduction. Transformations may be necessary to satisfy assumptions for parametric analysis. If sampling has been collected regularly at very frequent intervals, the data can be aggregated to standard periods (e.g., from daily observations to monthly means or medians). Adjusting data because of changing sampling frequency (i.e., weekly in years 1–5, monthly in years 6–10) requires subsampling from the higher frequency data to create data of the same frequency as the lower frequency to preserve constant variance. For example, do not compute monthly averages from weekly data in the early part of the record to combine with monthly data collected in the more recent part of the record. Rather, randomly choose one sample per month from the weekly data to construct a consistent data record of monthly samples. On the other hand, aggregating data, by computing monthly means or medians from weekly data throughout the period of record will reduce autocorrelation.

The flow-weighted or time-weighted mean concentrations are common methods to aggregate data collected with high frequency (Richards and Baker 1993). Flow-weighted mean concentration (FWMC) can be defined as:

$$\mathbf{FWMC} = \frac{\sum \mathbf{c_i q_i t_i}}{\sum \mathbf{q_i t_i}}$$

where c_i is the concentration of the i^{th} sample, q_i is the instantaneous flow associated with the i^{th} sample, and t_i is the time associated with the i^{th} sample. In other words, the FWMC is calculated by dividing the total pollutant load by the total flow volume over a given time period. The FWMC can be thought of as pollutant load normalized for flow or a flow-proportional concentration.

The time-weighted mean concentration (TWMC) can be defined as:

$$\mathbf{TWMC} = \frac{\sum \mathbf{c_i t_i}}{\sum \mathbf{t_i}}$$

In a fixed-frequency sampling program, the TWMC would be identical to the **arithmetic mean** of the observed concentrations.

Because much of the variance in nonpoint source pollutant concentrations may result from variation in streamflow, flow adjustment is a common technique to prepare for trend analysis. Removing this source of variance from the data makes subsequent trend tests more powerful and prevents the identification of a trend in concentration when it is the result of correlation with flow. When flow effects are removed from a record of concentrations, the test performed becomes a test for a time trend in the flow-adjusted concentrations versus time.

Flow adjustment is a common technique to prepare for trend analysis. Removing this source of variance from the data makes subsequent trend tests more powerful and prevents the identification of a trend in concentration when it is the result of correlation with flow.

A regression of concentration against some function of discharge is computed and the residuals (the differences between observed concentrations and concentrations predicted from the regression, i.e., flow-adjusted concentrations) are then tested for trend. Examples of this analysis are found in Hirsch et al. (1991) and Helsel and Hirsch (1992). This technique requires that a relationship exists between concentration and discharge. For this procedure to be valid, the streamflow distribution must be stationary, i.e., be itself free of trend. If the distribution of streamflow has changed over the period of record (e.g., because of diversions, detention ponds, or stormwater BMPs), then residuals analysis or any other flow-adjustment technique should not be used. Presence or absence of trend in flow can be verified through knowledge of changes in watershed hydrology or by independent analysis of trends in the streamflow record itself. Where streamflows are not stationary, it may be possible to remove the effects of varying hydrologic conditions on the concentration variable by using some appropriate measure of basin precipitation as a covariate or account for hydrologic changes by other trend analysis techniques.

Alternatively, because land treatment effects are generally expected to change the relationship between concentrations and flow, an analysis of covariance will usually be appropriate.

Graphing

Before proceeding to intensive numerical analysis, it is useful to re-examine the time series plots developed earlier in the process of exploratory data analysis. Visual inspection of a time series plot is the easiest way to look for a trend, but data variability may obscure a trend. Visualization of trends in noisy data can be clarified by various data smoothing techniques. Plotting moving averages or medians, for example, instead of raw data points, reduces apparent variation and may reveal general tendencies. Spreadsheets like Excel can display a moving-average trend line in time-series scatterplots with adjustable averaging periods. A more complex smoothing algorithm, such as *LOWESS* (***LO***cally ***WE***ighted ***S***catterplot ***S***moothing), can reveal patterns in very large datasets that would be difficult to resolve by eye. LOWESS is computationally intensive (see Helsel and Hirsch 1992), but computer programs exist that make the procedure relatively easy to accomplish.

Note, however, that visualization has limitations because people tend to focus on outliers, strong seasonal variation can mask trends in a variable of interest, and gradual trends are difficult to detect by eye alone. Additionally, simple visualization cannot adequately quantify the magnitude of a trend. Visualization is not a substitute for the hypothesis testing discussed below.

Monotonic Trend Analysis

A number of statistical tests are available to identify and quantify monotonic trends in a way that is defensible and repeatable. Statistical trend analysis is a hypothesis-testing process. The null hypothesis (H_O) is that there is no trend; each test has its own parameters for accepting or rejecting H_O. Failure to reject H_O does not prove that there is not a trend, but indicates that the evidence is not sufficient to conclude with a specified level of confidence that a trend exists.

Table 2 lists some trend tests available for different circumstances, including adjustments for a covariate and the presence of seasonality. The tests are further divided into parametric, nonparametric, and mixed types. Parametric tests are considered more powerful and/or sensitive to detect significant trends than are nonparametric tests, especially with a small sample number. However, unless the assumption of normal distribution for parametric statistics is met, it is generally advisable to use a nonparametric test (Lettenmaier 1976, Hirsch et al. 1991, Thas et al. 1998). Both parametric and nonparametric tests require constant variance and independence. Methods for testing assumptions of distribution, constant variance, and independence required for parametric linear regressions are discussed in detail in USEPA (1997a). Nonparametric tests provide higher statistical power in case of nonnormality and are robust against outliers and large data gaps.

Table 2. Classification of tests for trend (adapted from Helsel and Hirsch 1992)

	Type of test	Not Adjusted for covariate (X)	Adjusted for covariate (X)
No seasonality	Parametric	Linear regression of Y on t	Multiple linear regression of Y on X and t
	Mixed	–	Mann-Kendall on residuals from regression of Y on X
	Nonparametric	Mann-Kendall	Mann-Kendall on residuals from LOWESS of Y on X
Seasonality	Parametric	Linear regression of Y on t and periodic functions	Multiple linear regression of Y on X, t, and periodic functions
	Mixed	Regression of deseasonalized Y on t	Seasonal Kendall on residuals from regression of Y on X
	Nonparametric	Seasonal Kendall on Y	Seasonal Kendall on residuals from LOWESS of Y on X

Y = dependent variable of interest; X = covariate, t = time

These tests will be discussed below, with emphasis on linear regression, Mann-Kendall, and seasonal Kendall procedures. For more detailed information, consult the references listed at the end of this technical note.

Tests without covariates (Y versus time)

Parametric test: Linear regression of Y on t (Example 1, p. 18).

If project data satisfy all the assumptions necessary for linear regression (Y is linearly related to t, residuals are normally distributed, residuals are independent, and variance of residuals is constant), a simple linear regression of Y on time is a test for linear trend:

$$Y = \beta_0 + \beta_1 t + \varepsilon$$

The null hypothesis is that the slope coefficient $\beta_1 = 0$. The t-statistic on β_1 is tested to determine if it is significantly different from zero. If the slope is nonzero, the null hypothesis is rejected and it can be concluded that there is a linear trend in Y over time, with rate equal to β_1. Missing values are allowed. In some cases, it might have been necessary to log transform the data to satisfy the above regression assumptions. In this case, the trend slope will be expressed in log units. A linear trend in log units is an exponential trend in original units. This can be expressed in percent per year to make the trend easier to interpret. If β_1 is the estimated slope of the linear trend in $\log_{10}$ units, then the percentage change over any given year is $(10^{\beta 1} - 1)*100$. When there is no trend, the slope is zero and the equation results in zero percent change (i.e., $\beta_1 = 0$).

Nonparametric test: Mann-Kendall (Example 2, p. 19)

If the data do not conform to a normal distribution, the Mann-Kendall test can be applied. This test evaluates whether y values tend to increase or decrease over time through what is essentially a nonparametric form of monontonic trend regression analysis. The Mann-Kendall test analyzes the sign of the difference between later-measured data and earlier-measured data. Each later-measured value is compared to all values measured earlier, resulting in a total of n(n-1)/2 possible pairs of data, where n is the total number of observations. Missing values are allowed and the data do not need to conform to any particular distribution. The Mann-Kendall test assumes that a value can always be declared less than, greater than, or equal to another value; that data are independent; and that the distribution of data remain constant in either the original units or transformed units (Helsel and Hirsch 1992). Because the Mann-Kendall test statistics are invariant to transformations such as logs (i.e., the test statistics will be the same value for both raw and log-transformed data), the Mann-Kendall test is applicable in many situations.

To perform a Mann-Kendall test, compute the difference between the later-measured value and all earlier-measured values, $(y_j - y_i)$, where j>i, and assign the integer value of 1, 0, or –1 to positive differences, no differences, and negative differences, respectively. The test statistic, S, is then computed as the sum of the integers:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sign}(y_j - y_i)$$

Where $\operatorname{sign}(y_j - y_i)$, is equal to +1, 0, or –1 as indicated above.

When S is a large positive number, later-measured values tend to be larger than earlier values and an upward trend is indicated. When S is a large negative number, later values tend to be smaller than earlier values and a downward trend is indicated. When the absolute value of S is small, no trend is indicated. The test statistic τ can be computed as:

$$\tau = \frac{S}{n(n-1)/2}$$

which has a range of −1 to +1 and is analogous to the correlation coefficient in regression analysis. The null hypothesis of no trend is rejected when S and τ are significantly different from zero. If a significant trend is found, the rate of change can be calculated using the Sen slope estimator (Helsel and Hirsch 1992):

$$\beta_1 = \text{median}\left(\frac{y_j - y_i}{x_j - x_i}\right)$$

for all $i < j$ and $i = 1, 2, \ldots, n-1$ and $j = 2, 3, \ldots, n$; in other words, computing the slope for all pairs of data that were used to compute S. The median of those slopes is the Sen slope estimator.

Tests accounting for covariates

Variables other than time usually influence the behavior of water quality variables. These covariates are usually natural phenomena such as precipitation, temperature, or streamflow. By removing the variation caused by these explanatory variables, the noise may be reduced and a trend revealed. Correction for hydrologic and meteorologic variability is essential in both parametric and nonparametric trend tests to determine if the statistically significant trends are due to processes and transport changes such as land use changes, or to artifacts of system variability.

Selection of an appropriate covariate is critical. It should be a measure of the driving force behind the behavior of the variable of interest, but must not be subject to human manipulation during the course of the project, i.e., must not be changed by BMPs or the land treatment program. In nonpoint source monitoring, much of the variance in concentration data is usually a function of runoff and streamflow; thus, natural streamflow is a commonly used covariate in trend analysis. However, streamflow should not be used as a covariate if the land treatment program itself affects streamflow, such as with urban stormwater infiltration practices or conservation tillage. In such cases, precipitation may be a good choice for a covariate.

In deciding whether or not to remove the variation caused by flow from a data set, consider project objectives and the nature of the land treatment program. If a land treatment program has caused a measurable change in the watershed flow regime, such a change may in fact be a desired outcome and the resulting trend in both flow and pollutant concentration may be important to detect and quantify. Removing variation caused by flow may risk reducing the magnitude of any trend in concentration alone below detection level, considering other noise in the system. On the other hand, failure to account for a trend in flow that is not associated with the land treatment program may result in showing a trend in concentration where none exists. It is generally advisable to test the covariate data set independently for trend before proceeding.

Parametric: Multiple linear regression of Y on X and t

Multiple linear regression can be used to account for the effects of other variables such as flow, land management, or other water quality characteristics on a response variable. Multiple regression includes covariates in trend analysis in a single step. Appropriate covariates are those that are correlated with the water quality variable Y and adjust for changes in climate to better isolate trends due to BMPs. Consider multiple regression of concentration (Y) versus time (t) and flow (Q):

$$Y = \beta_0 + \beta_1 t + \beta_2 Q + \varepsilon$$

The test accounts for the effects of the covariate by including them in the regression model. The null hypothesis for the trend test is $\beta_1 = 0$; the t-statistic for β_1 tests for trend. If the coefficient β_2 for the covariate is not significantly different from zero, the effect of the covariate is not significant and a simple regression model of Y on t should be used. An exception to this would be the case where flow is increasing over time and the effects of increasing flow are already accounted for in the time component; in such a case, flow might still be logically included in the regression model even if β_2 is not different from zero. It should be emphasized that as for simple linear regression, the assumptions that Y is linearly related to t and Q, that residuals are normally distributed and independent, and that variance of residuals is constant must be satisfied to use this test properly.

Mixed: Mann-Kendall on residuals from regression of Y on X

This is a hybrid test that includes removal of covariate effects by a parametric procedure, followed by a nonparametric test for trend. If a reasonable linear regression can be obtained (i.e., residuals have no extreme outliers, Y is approximately linear with X), the regression between Y and one or more Xs (i.e., $Y = \beta_0 + \beta_1 X + \varepsilon$) can remove the effect of X prior to performing the Mann-Kendall test for trend.

The residuals (R) from the regression model are computed as observed minus predicted values:

$$R = Y - \beta_0 + \beta_1 X$$

Then the Kendall S statistic is computed on the R-time data pairs and tested to see if it differs significantly from zero. If assumptions for parametric statistics are seriously violated, a fully nonparametric alternative (e.g., using LOWESS) should be selected to estimate the relationship between Y and X as described in the next section.

Nonparametric: Mann-Kendall on residuals from LOWESS of Y on X

The LOWESS smoothing technique describes the relationship between Y and a covariate X without assuming linearity or normality of residuals. Applying LOWESS smoothing to a scatterplot of X and Y is roughly analogous to regression, without forcing a straight line. Given the LOWESS fitted value Y', the residuals (R) are computed as:

$$R = Y - Y'$$

Then, the Kendall S statistic is computed on the R–t data pairs and tested to see if it differs significantly from zero.

If the distribution of the data is unknown or known to violate parametric assumptions, this procedure should be used instead of the parametric or mixed tests.

Seasonality

Water quality data often show seasonal patterns that require trend analysis techniques that go beyond simply controlling for the effects of flow.

Frequently, changes between seasons are a major source of variation in water quality data because land management and use change with the seasons. Most concentrations in surface waters show strong seasonal patterns. Seasonal variation in streamflow is an important component of this seasonality, but biological processes (e.g., enhanced survival of fecal microorganisms in colder water temperatures, release of nitrogen through decomposition) and management activities (e.g., fertilizer applications, tillage) often contribute to seasonal variation. Thus, some techniques beyond controlling for the effects of a flow covariate are often necessary for water quality trend analysis.

Some trend analysis techniques require you to define a "season" in advance. Examination of box plots of data by season or other graphical displays may help identify reasonable divisions. In general, seasons should be just long enough so that some data are available for most of the seasons in most years of monitoring. If data are collected or aggregated on a monthly frequency, for example, seasons should be defined representing each of the 12 months. If data are considered in quarterly blocks, there should be four seasons. In agricultural settings, it may make sense to consider either two or four "seasons": cropping and non-cropping, or non-cropping, seed preparation, cropping, and harvest.

Parametric: Linear regression of Y on X, t, and periodic functions

Periodic functions like sine and cosine can be used to describe cyclic seasonal variations in a multiple regression model, with or without covariates. For an annual cycle:

$$Y_\tau = \beta_0 + \beta_1 \sin(2\pi t/n) + \beta_2 \cos(2\pi t/n) + \beta_3 t + \text{other terms} + \varepsilon_\tau$$

Where: t=1,2,3...N (N=total number of samples)
n = number of samples per year (e.g., 12 for monthly data, 52 for weekly data)
note: a "DATE" variable can be used instead of 't' with n=365.25 because 'DATE' is a daily value.

Where "other terms" are covariates such as flow, precipitation, or other influences. The trend test is conducted by determining if the slope coefficient on t (β_3) differs significantly from zero. This test assumes that the sine and cosine terms realistically simulate annual seasonal cycles. Of course, the usual assumptions of parametric regression must be met. If variability introduced by strong seasonality (e.g., extremely dry or wet season) is high enough to cause violation of parametric assumptions, it may become necessary to break out data by season before conducting trend analysis.

Mixed: Seasonal Kendall on residuals from regression of Y on X *and* Regression of deseasonalized Y on t

Two hybrid procedures may be used to account for seasonality. First, the seasonal Kendall test can be applied to residuals from a simple linear regression of Y versus X. This approach should only be used when the relationship of Y and X complies with the appropriate assumptions for parametric statistics.

Second, the data can be "deseasonalized" by subtracting seasonal medians or some other measure of seasonal effect from all the data within the season. The deseasonalized data is then regressed against time (Montgomery and Reckhow 1984). Although this technique has the advantage of producing a description of seasonality (seasonal medians), it has generally low statistical power.

Nonparametric: Seasonal Kendall on Y (Example 3, p. 21)

The seasonal Kendall test statistic is computed by performing a Mann-Kendall calculation for each season, then combining the results for each season. For monthly seasons, January observations are compared only to other January observations, etc. No comparisons are made across seasonal boundaries. The Seasonal Kendall test is highly robust and relatively powerful, and is often the recommended method for most water quality trend monitoring.

The S_k statistic is computed as the sum of the S from each season:

$$S_k = \sum_{i=1}^{m} S_i$$

where S_i is the S from the i^{th} season and m is the number of seasons.

The seasonal statistics are summed and a Z statistic is computed; consult other sources for the method of calculating Z_{Sk} (e.g., Helsel and Hirsch 1992, USEPA 1997b). If the number of seasons and years are sufficiently large (seasons * years $\geq$ 25), the Z value may be compared to standard normal tables to test for a statistically significant trend. For fewer seasons/years, the applicability of standard normal tables has not been evaluated. An estimate of the trend slope for Y over time can be computed as the median of all slopes between data pairs within the same season using a generalized version of the Sen slope estimator described above. Consult other sources for the method of calculation (e.g., Helsel and Hirsch 1992, USEPA 1997b).

Emerging trend analysis techniques

A recent paper by Hirsch et al. (2010) called for a "next generation" of trend analysis techniques in response to the observations that new and longer monitoring data sets exist, new questions about the effectiveness of control efforts, and the availability of new statistical tools. The authors identified seven critical attributes for the next generation of trend analysis:

- Focuses on revealing the nature and magnitude of change, rather than strict hypothesis testing;
- Does not assume that the flow-concentration relationship is constant over time;
- Makes no assumptions that seasonal patterns repeat exactly over the period of record, but allow the shape of seasonality to evolve over time;
- Allows the shape of an estimated trend to be driven by the data and not constrained to follow a specific form such as linear or quadratic; trend patterns should be allowed to differ for different seasons or flow conditions;
- Provides consistent results describing trends in both concentration and load;
- Provides not only estimates of trends in concentration and flux but also trend estimates where the variation in water quality due to variation in streamflow has been statistically removed; and
- Includes diagnostic tools to assist in understanding the nature of the changes that have taken place over time, e.g., to identify particular times of year or hydrologic conditions during which water quality changes are most pronounced.

The authors propose and demonstrate an experimental trend analysis technique called *Weighted Regressions on Time, Discharge, and Season* (WRTDS) that addresses these critical attributes. While a presentation of this approach is beyond the scope of this Tech Note, the reader is referred to the original paper for additional information.

Step Trends

Monotonic trend analysis may not be appropriate for all situations. Other statistical tests for discrete changes (step trends) should be applied where a known discrete event (like BMP implementation over a short period) has occurred. Testing for differences between the "before" and "after" conditions is done using two-sample procedures such as t-tests and analysis of covariance (parametric techniques) and nonparametric alternatives such as the rank-sum test, Mann-Whitney test, and the Hodges-Lehmann estimator of step-trend magnitude (Helsel and Hirsch 1992, Walker 1994).

Monitoring Program Design and Trend Analysis

Trend analysis is effective with data sampled continuously at fixed-time intervals. If you are presently designing your watershed monitoring program, here are key points to consider if you plan to use trend analysis to evaluate your project:

- Use consistent sampling locations throughout the monitoring period;
- Operate the monitoring program continuously, starting before implementation and continuing after implementation;

- Use consistent field and laboratory procedures;
- Collect data on important covariates to help explain variations in water quality; and
- Monitor land treatment, land use, and other nonpoint source-related activities in your watershed to provide information to help you interpret observed trends.

Statistical tools for trend analysis

Trend tests, especially nonparametric tests like the Mann-Kendall and seasonal Kendall are computationally intensive and are impractical to apply manually in most cases. Unfortunately, statistical software packages that calculate Mann-Kendall and other nonparametric analyses are less common than those that perform parametric tests. The table below lists some examples of software that will run some or all of the nonparametric tests discussed in this publication and web sites to visit for more information. *Practical Stats* (see Further Reading and Resources, below) provides a useful review of the capabilities of low-cost statistical software at: ***http://www.practicalstats.com/aes/aes/DownloadsAES_files/Evaluation2.pdf***

Package Name	Web Site URL
Add-ins for MS Excel:	
XLStat	***http://www.xlstat.com/en/home/***
Analyse-it	***http://www.analyse-it.com/***
WinStat	***http://www.winstat.com/***
StatistiXL	***http://www.statistixl.com/default.aspx***
Fast Statistics	***http://www.fatesoft.com/excel/***
Stand-alone statistical software:	
ChemStat	***http://www.pointstar.com/ChemStat/chemstat.asp***
JMP	***http://www.jmp.com/***
MINITAB	***http://www.minitab.com/products/minitab/default.aspx***
R	***http://www.r-project.org/***
S-Plus	***http://www.msmiami.com/directory.cfm?CategoryID=42*** [add-in programs for nonparametric trend analysis from USGS software library: ***http://water.usgs.gov/software/S-PLUS/***]
SAS	***http://www.sas.com/technologies/analytics/statistics/index.html***
SPSS	***http://www.spss.com/spss/***
SYSTAT	***http://www.systat.com/products/Systat/***

Stand-alone Windows programs for running Mann-Kendall and Seasonal Kendall tests have been published by USGS and are available for free download at ***http://pubs.usgs.gov/sir/2005/5275/***. An example of a custom-made spreadsheet calculator for running Mann-Kendall tests on quarterly data can be found at ***http://www.in.gov/idem/4213.htm*** (Indiana Department of Environmental Management 2011).

References

Chesapeake Bay Program. 2007. Chesapeake Bay 2007 health & restoration assessment. ***http://www.chesapeakebay.net/content/publications/cbp_26038.pdf*** [Accessed 9-15-2011]

Clausen, J.C. and J. Spooner. 1993. Paired watershed study design. 841-F-93-009. U.S. Environmental Protection Agency, Office of Water, Washington, DC

Cude, C. 2005. Specific examples of trend analysis using the Oregon water quality index. Oregon Department of Environmental Quality, Laboratory Division, ***http://www.deq.state.or.us/lab/wqm/trendex.htm*** [Accessed 9-15-2011]

Helsel, D.R. and R.M. Hirsch. 1992. *Statistical methods in water resources*. Studies in Environmental Science 49. New York: Elsevier. (available on-line as a pdf file at: ***http://water.usgs.gov/pubs/twri/twri4a3/***) [Accessed 9-15-2011].

Helsel, D.R., D.K. Mueller, and J.R. Slack. 2005. Computer program for the Kendall family of trend tests. USGS Scientific Investigations Report 2005-5275, U.S. Geological Survey, Reston, VA. ***http://pubs.usgs.gov/sir/2005/5275/*** [Accessed 9-15-2011]

Hirsch, R.M. 1988. Statistical methods and sampling design for estimating step trends in surface water quality. *Water Resour. Res*. 24:493-503.

Hirsch, R.M., R.B. Alexander, and R.A. Smith.1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resour. Res*. 27:803-813.

Hirsch, Robert M., Douglas L. Moyer, and Stacey A. Archfield, 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. *J. Am. Water Resour. Assoc.* 46(5):857-880.

Indiana Department of Environmental Management. 2011. Mann-Kendall spreadsheet. ***http://www.in.gov/idem/4213.htm*** [Accessed 9-15-2011].

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6(6):28-30.

Lettenmaier. D.P. 1976. Detection of trends in water quality data from records with dependent observations. *Water Resour. Res*. 12:1037-1046.

Meals, D.W. 2001. *Lake Champlain Basin agricultural watersheds section 319 national monitoring program project, final project report: May, 1994–September, 2000*. Vermont Dept. of Environmental Conservation, Waterbury, VT, 227 p.

Montgomery, R.H. and K.H. Reckhow. 1984. Techniques for detecting trends in lake water quality. *Water Resour. Bull*. 20(1):43-52.

Newbold, J.D., S. Herbert, B.W. Sweeney and P. Kiry. 2008. Water quality functions of a 15-year-old riparian forest buffer system. *NWQEP Notes* No. 130. NCSU Water Quality Group, NC State University, Raleigh, NC.

Richards, R.P. and D.B. Baker. 1993. Trends in nutrient and suspended sediment concentrations in Lake Erie tributaries, 1975-1990. *J. Great Lakes Res*. 19(2):200-211.

Thas O., L. Van Vooren, and J.P. Ottoy. 1998. Nonparametric test performance for trends in water quality with sampling design applications. *J. American Water Resour. Assoc*. 34(2):347-357.

US EPA. 1997a. Linear regression for nonpoint source pollution analysis. EPA-841-B-97-007. Office of Water, Washington, DC.

US EPA. 1997b. Monitoring guidance for determining the effectiveness of nonpoint source control projects. EPA 841-B-96-004. Office of Water, Washington, DC.

Walker, J.F. 1994. Statistical techniques for assessing water quality effects of BMPs. *J. Irr. Drainage Eng. ASCE*. 120(2):334-347.

Ye, Y-S and S. Zou. 1993. Relating trends of principal components to trends of water-quality constituents. *Water Resour. Bull*. 29(5):797-806.

Further Reading and Resources

There is far more to trend analysis than is covered in this technical note. For more details on the techniques and calculations discussed, examples, and information on other approaches, consult these additional sources:

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

Gilbert, R.O. 1987. *Statistical methods for environmental pollution monitoring*. New York: Van Nostrand Reinhold.

Grabow, G.L., L.A. Lombardo, D.E. Line, and J. Spooner. 1999. Detecting water quality improvement as BMP effectiveness changes over time: use of SAS for trend analysis. *NWQEP Notes* No. 95. NCSU Water Quality Group, NC State University, Raleigh, NC.

Hirsch, R.M, J.R. Slack, and R.A. Smith. 1982. Techniques of trend analysis for monthly water quality data. *Water Resour. Res*. 18:107-121.

Practical Stats, offering statistical training for environmental sciences and natural resources, with emphasis on applied environmental statistics, nondetects data analysis, and multivariate analysis. ***http://www.practicalstats.com/index.html*** [Accessed 9-15-2011].

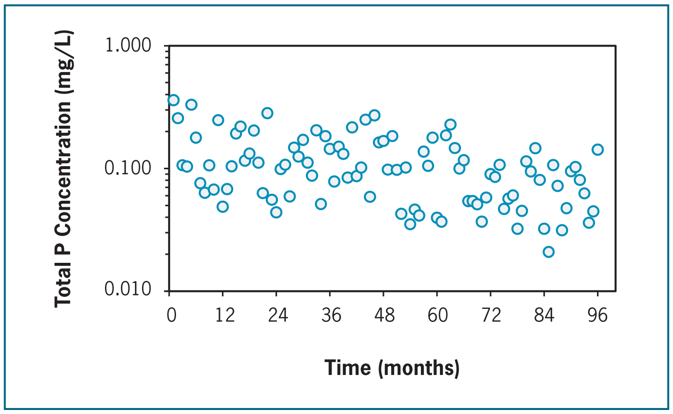
Rheem, S. and G.I. Holtzman. 1990. A SAS program for seasonal Kendall trend analysis of monthly water quality data. Proceedings, Sixteenth Annual SAS Users Group International (SUGI) Conference, February 17-20, 1991, New Orleans.

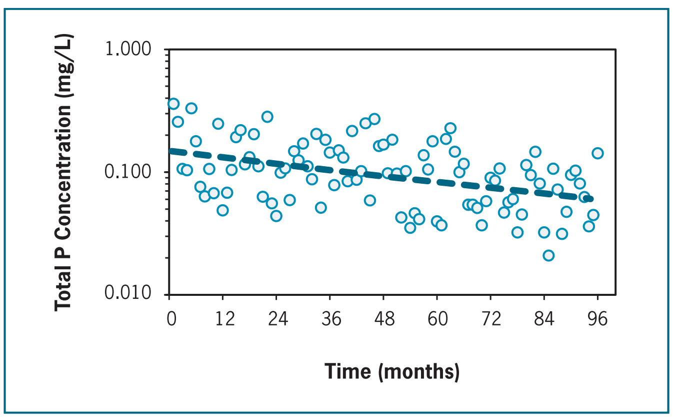
Richards, R.P. and D.B. Baker. 2002. Trends in water quality in LEASEQ rivers and streams (northwestern Ohio), 1975-1995. *J. Environ. Qual*. 31(1): 90 - 96.

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

Trend analysis example 1: Simple linear regression

- Eight years of monthly total phosphorus concentration data from Samsonville Brook, a stream draining a Vermont agricultural watershed
- Data satisfy assumptions for regression after log transformation
  - Normal distribution
  - Constant variance
  - Independence (low autocorrelation)





Simple linear regression
(using Excel or any basic statistical package)

Log [TP] = −0.8285 – 0.00414(Time)

$r^2 = 0.18$ $F = 21.268$ $P \leq 0.001$

Rate of change:
Slope of log-transformed data = −0.00414

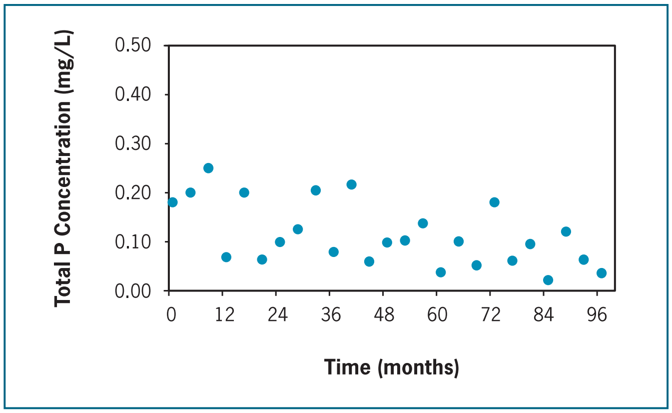
$(10^{-0.00414} - 1)*100 = -0.95\%$/month or ~11%/yr

This result suggests that total P concentrations have decreased significantly over the period at a rate of approximately 11% a year.

Note: data used in this example are taken from the Vermont NMP Project, *Lake Champlain Basin agricultural watersheds section 319 national monitoring program project*, 1993 – 2001 (Meals 2001).

Trend analysis example 2: Mann-Kendall

Month (n=25)	[TP] (mg/L)
1	0.180
5	0.200
9	0.250
13	0.068
17	0.201
21	0.063
25	0.099
29	0.125
33	0.205
37	0.078
41	0.216
45	0.059
49	0.098
53	0.102
57	0.137
61	0.037
65	0.100
69	0.051
73	0.180
77	0.060
81	0.095
85	0.021
89	0.120
93	0.063
97	0.035



- Eight years of quarterly mean total phosphorus concentration data from Samsonville Brook, a stream draining a Vermont agricultural watershed
- Data satisfy assumptions for constant variance and independence, but are not normally-distributed without transformation

The Mann-Kendall trend test for this example may be evaluated in two ways. First, in a manual calculation, use the formulas below. The value of S (sum of the signs of differences between all combinations of observations) can be determined either manually or by using a spreadsheet to compare combinations, create dummy variables (–1, 0, and +1), and sum for S.

$$\textbf{Mann – Kendall} \qquad S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}\,(y_j - y_i) = -106$$

$$\tau = \frac{S}{n(n-1)/2} = \frac{-106}{300} = -0.353 \longrightarrow \textbf{decreasing trend}$$

Calculating Z_S as $(S \pm 1)/\sigma_s$ where $\sigma_s = \sqrt{(n/18) \times (n-1) \times (2n+5)} = 42.817$

$$Z = \frac{-105}{42.817} = -2.454 \qquad \text{(USEPA 1997b)}$$

This Z statistic is significant at P = 0.014, indicating a significant trend, i.e., we are 98.6% confident there is a decreasing trend in TP. See USEPA (1997b) for the calculation of σ_S when there are ties among the data.

To estimate the rate of change, use the Sen slope estimator

$$\beta_1 = \textbf{median}\left(\frac{y_j - y_i}{x_j - x_i}\right)$$

211 individual slopes −0.00945 to +0.00766
median slope = −0.0011 mg/L/month = −0.013 mg/L/yr

This result suggests that total P concentration decreased significantly over the period at a rate of about 0.013 mg/L/yr.

Alternatively, use a statistics computer program to run the Mann-Kendall procedure. For example, using the USGS program for the Kendall family of tests (Helsel et al. 2005), set up a text data input file specifying the Mann-Kendall test (test #4) without flow adjustment ("0") or seasons (blanks) and name the data input file ("MKexample2.txt") as:

```
4 0  MKexample 2
1    0.180
5    0.200
9    0.250
.
.
97   0.035
```

The output from the program gives the same results as shown above, including the estimated slope of the trend (-0.0011) computed by the Sen slope estimator above:

```
          Kendall's tau Correlation Test
            US Geological Survey, 2005

Data set:        MK Example 2

The tau correlation coefficient is -0.353
     S =  -106.
     z =  -2.454
     p =  0.0141

The relation may be described by the equation:
     Y = 0.15412  +  -0.1125E-02 * X
```

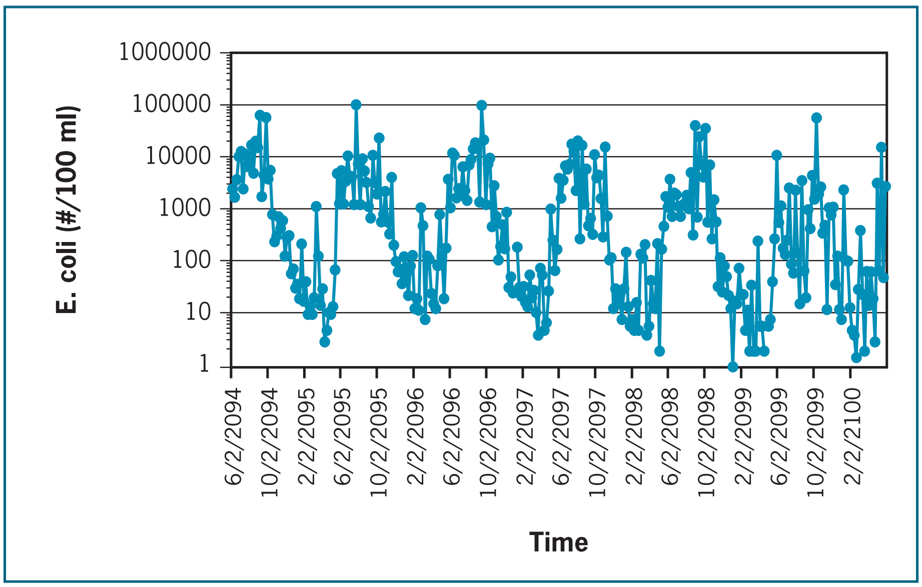
Note: data used in this example are taken from the Vermont NMP Project, *Lake Champlain Basin agricultural watersheds section 319 national monitoring program project*, 1993 – 2001 (Meals 2001).

Helsel, D.R., D.K. Mueller, and J.R. Slack. 2005. Computer program for the Kendall family of trend tests. USGS Scientific Investigations Report 2005-5275, U.S. Geological Survey, Reston, VA. *http://pubs.usgs.gov/sir/2005/5275/*

Trend analysis example 3: Seasonal Kendall

- Six years of weekly *E. coli* data from a stream draining Godin Brook, a Vermont agricultural watershed
- Data satisfy assumptions for constant variance and independence, but are not normally-distributed without transformation
- Data display high degree of seasonality to the eye (low *E. coli* counts in winter, high counts in summer) due to influence of water temperature on bacteria survival and to grazing season

Raw data plotted:

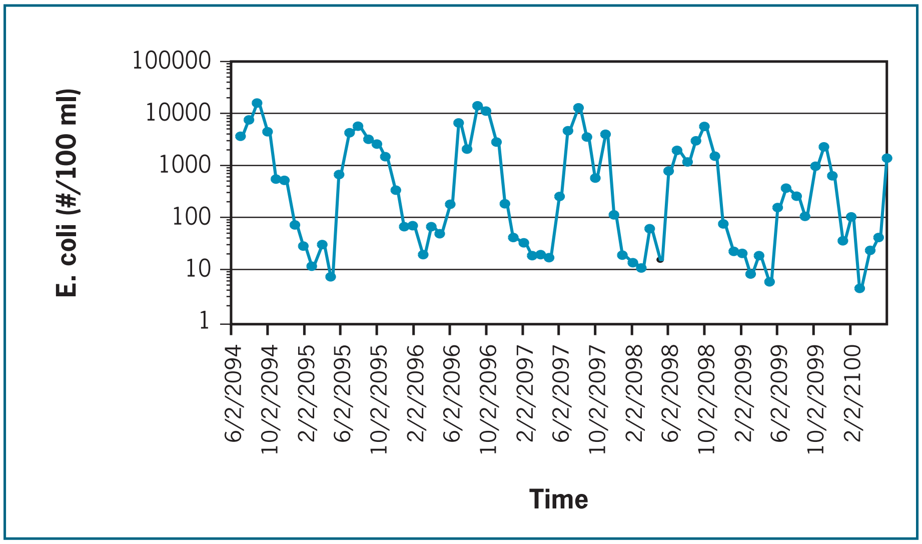


Data aggregated to monthly median values:

	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
1994	3750	7725	16350	4600	565	535	74					
1995	4400	5900	3300	2663	1530	345	69	29	12	31	8	688
1996	6788	2125	14500	11450	2900	190	43	72	20	69	50	185
1997	4825	13250	3635	592	4100	116	20	33.5	19	20	18	262
1998	2025	1200	3083	5825	1563	78	23	14	11	63	16	807.5
1999	378	265	109	1000	2360	653	37	21	8.5	19	6	161
2000								106	4.5	24	42	1432

(values represent *E. coli*/100 ml)

Monthly median data plotted still show a strong seasonal cycle:



As in Example 2, the Seasonal Kendall trend test may be computed manually, using the formulas below, either by hand or using a spreadsheet:

The Mann-Kendall statistic (S_i) is calculated for each month; the seasonal Kendall statistic S_k calculated as sum of monthly S_i:

$$S_k = \sum_{i=1}^{m} S_i = -48 \qquad \text{suggesting a downward trend}$$

Z_{Sk} is estimated as $Z_{Sk}=(S+1)/\sigma_{Sk}$ where $\sigma_{Sk} = \sqrt{\sum_{i=1}^{m} (n_i/18) \times (n_i - 1) \times (2n_i + 5)}$

$$= 18.439$$

$$Z = \frac{-47}{18.439} = -2.549 \qquad \text{(USEPA 1997a)}$$

This Z statistic is significant at P = 0.011, indicating a significant decreasing trend.

To use the Sen slope estimator, calculate slopes between all possible pairs within each season, rank all slope estimates, and find the median:

$$\beta_1 = \text{median} \left(\frac{y_j - y_i}{x_j - x_i} \right)$$

180 individual slopes –13,050 to +11,200
median slope = –5.8 *E. coli*/100 ml/yr

This result suggests that *E. coli* counts have decreased significantly over the period at an approximate rate of 6 *E. coli*/100 ml/yr.

Alternatively, use a statistics computer program to run the Seasonal Kendall procedure. For example, using the USGS program for the Kendall family of tests (Helsel et al. 2005), set up a text data input file specifying the Seasonal Kendall test with data as year, season, and value (test #2) without flow adjustment ("0"), seasons (ignored for this type of input data) and name the data input file (SKexample3.txt) as:

```
2  0    SK Example 3
1994    6    3750
1994    7    7725
1994    8    16350
.
.
.
2000    3    24
2000    4    42
2000    5    1432
```

The output from the program gives the same results as shown above, including the estimated slope of the trend (–5.75) computed by the Sen slope estimator above:

```
                 Seasonal Kendall Test for Trend
                   US Geological Survey, 2005

Data set:        SK Example 3

The record is 7 complete water years with 12 seasons per
     year beginning in water year 1994.

The tau correlation coefficient is -0.267
     S =  -48.
     z =  -2.549
     p =  0.0108
     p =  0.2003 adjusted for correlation among seasons
      (such as serial dependence)

The adjusted p-value should be used only for data with more
     than 10 annual values per season.

The relation may be described by the equation:
     Y = 246.1  +  -5.750 * Time
     where Time = Year (as a decimal) - 1993.75 (beginning
     of first water year)
```

Note: data used in this example are taken from the Vermont NMP Project, *Lake Champlain Basin agricultural watersheds section 319 national monitoring program project*, 1993 – 2001 (Meals 2001).

Helsel, D.R., D.K. Mueller, and J.R. Slack. 2005. Computer program for the Kendall family of trend tests. USGS Scientific Investigations Report 2005-5275, U.S. Geological Survey, Reston, VA. ***http://pubs.usgs.gov/sir/2005/5275/***