

Chapter 12

Trend Analysis

Concentrations and loads of phosphorus have been observed at numerous tributaries to an important estuary over a 20-year period. Have concentrations and/or loads changed over time? Have concentrations changed when changing flow conditions are taken into account (the early years were during a very dry period), or are all changes simply due to more precipitation in the latter years? Is there an observable effect associated with a ban on phosphorus compounds in detergents which was implemented in the middle of the period of record?

Groundwater levels were recorded for many wells in a study area over 14 years. During the ninth year development of the area increased withdrawals dramatically. Is there evidence of decreasing water levels in the region's wells after versus before the increased pumpage?

Benthic invertebrate and fish population data were collected at twenty stations along one hundred miles of a major river. Do these data change in a consistent manner going downstream? What is the overall rate of change in population numbers over the one hundred miles?

Procedures for trend analysis build on those in previous chapters on regression and hypothesis testing. The explanatory variable of interest is usually time, though spatial or directional trends (such as downstream order or distance down dip) may also be investigated. Tests for trend have been of keen interest in environmental sciences over the last 10-15 years. Detection of both sudden and gradual trends over time with and without adjustment for the effects of confounding variables have been employed. In this chapter the various tests are classified, and their strengths and weaknesses compared.

12.1 General Structure of Trend Tests

12.1.1 Purpose of Trend Testing

A series of observations of a random variable (concentration, unit well yield, biologic diversity, etc.) have been collected over some period of time. We would like to determine if their values generally increase or decrease (getting "better" or "worse"). In statistical terms this is a determination of whether the probability distribution from which they arise has changed over time. We would also like to describe the amount or rate of that change, in terms of changes in some central value of the distribution such as a mean or median. Interest may be in data at one location, or all across the country. Figure 12.1 presents an example of the results of trend tests for bacteria at sites throughout the United States.

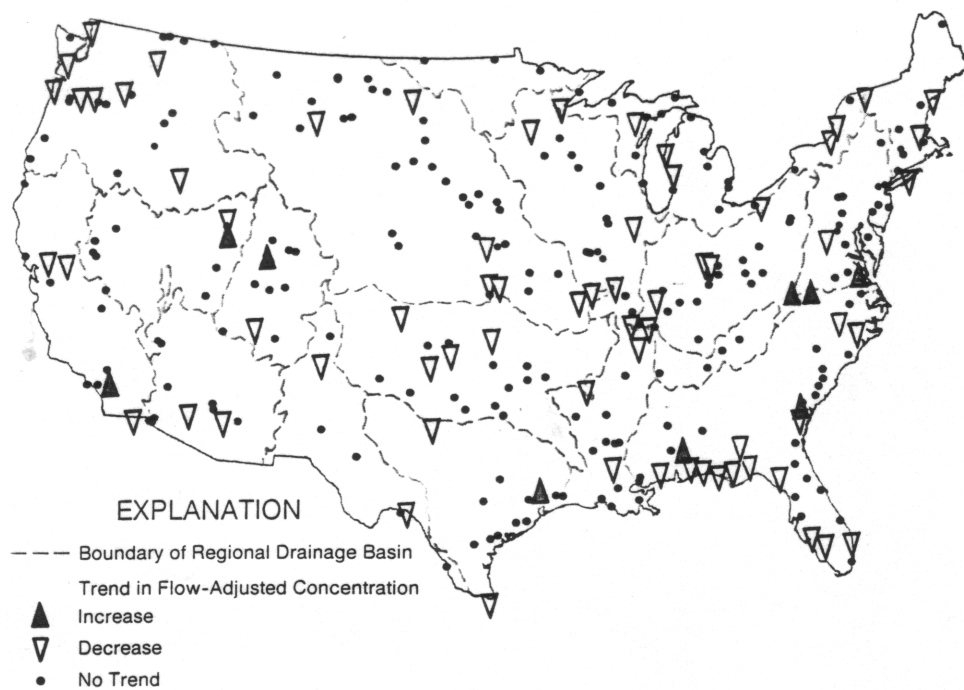


Figure 12.1 Trends in flow-adjusted concentrations of fecal streptococcus bacteria, 1974-1981 (from Smith et al., 1987).

The null hypothesis: H_0 is that there is no trend. However, any given test brings with it a precise mathematical definition of what is meant by "no trend", including a set of background assumptions usually related to type of distribution and serial correlation. The outcome of the test is a "decision" -- either H_0 is rejected or not rejected. Failing to reject H_0 does not mean that it was "proven" that there is no trend. Rather, it is a statement that the evidence available is not sufficient to conclude that there is a trend. Table 12.1 summarizes the possible outcomes of a statistical test in the context of trend analysis.

Decision	True Situation	
	No trend. H_0 true.	Trend exists. H_0 false.
Fail to reject H_0 . "No trend"	Probability = $1-\alpha$	(Type II error) β
Reject H_0 . "Trend"	(Type I error) significance level α	(Power) $1-\beta$

Table 12.1 Probabilities associated with possible outcomes of a trend test.

$$\alpha = \text{Prob}(\text{reject } H_0 | H_0 \text{ true}) \quad \text{and} \quad 1 - \beta = \text{Prob}(\text{reject } H_0 | H_0 \text{ false})$$

The power ($1-\beta$) for the test can only be evaluated if the nature of the violation of H_0 that actually exists is known. This is never known in reality (if it were we wouldn't need a test), so a test must be found which has high power for the kind of data expected to be encountered. If a test is slightly more powerful in one instance but much less powerful than its alternatives in some other reasonable cases then it should not be used. The test selected should therefore be robust -- it should have relatively high power over all situations and types of data that might reasonably be expected to occur. Some of the characteristics commonly found in water resources data, and discussed in this chapter, are:

- Distribution (normal, skewed, symmetric, heavy tailed)
- Outliers (wild values that can't be shown to be measurement error)
- Cycles (seasonal, weekly, tidal, diurnal)
- Missing values (a few isolated values or large gaps)
- Censored data (less-than values, historical floods)
- Serial Correlation

12.1.2 Approaches to Trend Testing

Five types of trend tests are presented in table 12.2. They are classified based on two factors. The first, shown in the rows of the table, is whether the test is entirely parametric, entirely nonparametric, or a mixture of procedures. The second factor (columns) is whether there is some attempt to remove variation caused by other associated variables. The table uses the following notation:

- Y = the random response variable of interest in the trend test,
- X = an exogenous variable expected to affect the value of Y,
- R = the residuals from a regression or LOWESS of Y versus X, and
- T = time (often expressed in years).

Simple trend tests (not adjusted for X) are discussed in section 12.2. Tests adjusted for X are discussed in section 12.3.

	Not Adjusted for X	Adjusted for X
Nonparametric	Mann-Kendall trend test on Y	Mann-Kendall trend test on residuals R from LOWESS of Y on X
Mixed	----	Mann-Kendall trend test on residuals R from regression of Y on X
Parametric	Regression of Y on T	Regression of Y on X and T

Table 12.2 Classification of five types of tests for trend

If the trend is spatial rather than temporal, T will be downstream order, distance downdip, etc. Examples of X and Y include the following:

- For trends in surface water quality, Y would be concentration, X would be streamflow, and R would be called the flow-adjusted concentration;
- For trends in flood flows, Y would be streamflow, X would be the precipitation amount, and R would be called the precipitation-adjusted flow (the duration of precipitation used must be appropriate to the flow variable under consideration. For example, if Y is the annual flood peak from a 10 square mile basin then X might be the 1-hour maximum rainfall, whereas if Y is the annual flood peak for a 10,000 square mile basin then X might be the 24-hour maximum rainfall).
- For trends in groundwater levels, Y would be the change in monthly water level, X the monthly precipitation, and R would be called the precipitation-adjusted change in water level.

12.2 Trend Tests With No Exogenous Variable

12.2.1 Nonparametric Mann-Kendall Test

Mann (1945) first suggested using the test for significance of Kendall's tau where the X variable is time as a test for trend. This was directly analogous to regression, where the test for significance of the correlation coefficient r is also the significance test for a simple linear regression. The Mann-Kendall test can be stated most generally as a test for whether Y values tend to increase or decrease with T (monotonic change).

$$H_0: \text{Prob } [Y_j > Y_i] = 0.5, \text{ where time } T_j > T_i.$$

$$H_1: \text{Prob } [Y_j > Y_i] \neq 0.5 \quad (2\text{-sided test}).$$

No assumption of normality is required, but there must be no serial correlation for the resulting p-values to be correct. Typically the test is used for a more specific purpose -- to determine whether the central value or median changes over time. The spread of the distribution must remain constant, **though not necessarily in the original units**. If a monotonic transformation such as the ladder of powers is all that is required to produce constant variance, the test statistic will be identical to that for the original units. For example, in figure 12.2 a lognormal Y variable is plotted versus time. The variance of data around the trend is increasing. A Mann-Kendall test on Y has a p-value **identical** to that for the data of figure 12.3 -- the logarithms of the figure 12.2 data. The logs show an increasing median with constant variance. Only the central location changes. The Mann-Kendall test possesses the useful property of other nonparametric tests in that it is invariant to (monotonic) power transformations such as those of the ladder of powers. Since only the data **or any power transformation of the data** need be distributed similarly over T except for their central location in order to use the Mann-Kendall test, it is applicable in many situations.

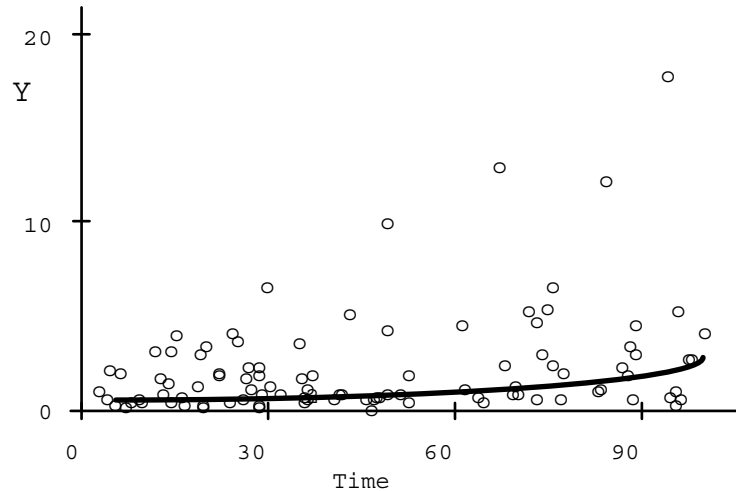


Figure 12.2 Y versus Time. Variance of Y increases over time.

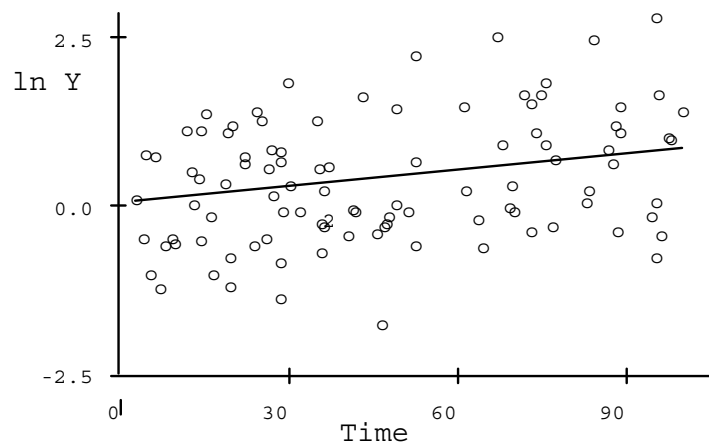


Figure 12.3 Logarithms of Y versus Time. The variance of Y is constant over time.

To perform the test, Kendall's S statistic is computed from the Y,T data pairs (see Chapter 8). The null hypothesis of no change is rejected when S (and therefore Kendall's τ of Y versus T) is significantly different from zero. We then conclude that there is a monotonic trend in Y over time.

An estimate of the rate of change in Y is also usually desired. If Y or some transformation of Y has a linear pattern versus T, the null hypothesis can be stated as a test for the slope coefficient $\beta_1 = 0$. β_1 is the rate of change in Y, or transformation of Y, over time.

12.2.2 Parametric Regression of Y on T

Simple linear regression of Y on T is a test for trend

$$Y = \beta_0 + \beta_1 \cdot T + \epsilon$$

The null hypothesis is that the slope coefficient $\beta_1 = 0$. Regression makes stronger assumptions about the distribution of Y over time than does Mann-Kendall. It must be checked for normality of residuals, constant variance and linearity of the relationship (best done with residuals plots -- see Chapter 9). If Y is not linear over time, a transformation will likely be necessary. If all is ok, the t-statistic on b_1 is tested to determine if it is significantly different from 0. If the slope is nonzero, the null hypothesis of zero slope over time is rejected, and we conclude that there is a linear trend in Y over time. Unlike Mann-Kendall, the test results for regression will not be the same before and after a transformation of Y.

12.2.3 Comparison of Simple Tests for Trend

If the model form specified in a regression equation were known to be correct (Y is linear with T) and the residuals were truly normal, then fully-parametric regression would be optimal (most powerful and lowest error variance for the slope). Of course we can never know this in any real world situation. If the actual situation departs, even to a small extent, from these assumptions then the Mann-Kendall procedures will perform either as well or better (see Chapter 10, and Hirsch et. al., 1991, p.805-806).

There are practical cases where the regression approach is preferable, particularly in the multiple regression context (see next section). A good deal of care needs to be taken to insure it is correctly applied and to demonstrate that to the audience. When one is forced, by the sheer number of analyses that must be performed (say a many-station, many-variable trend study) to work without detailed case-by-case checking of assumptions, then nonparametric procedures are ideal. They are always nearly as powerful as regression, and the failure to edit out or correctly transform a small percentage of outlying data will not have a substantial effect on the results.

Example 1

Appendix C10 lists phosphorus loads and streamflow during 1974-1985 on the Illinois River at Marseilles, IL. The Mann-Kendall and regression lines are plotted along with the data in figure 12.4. Both lines have slopes not significantly different from zero at $\alpha = 0.05$. The large load at the beginning of the record and non-normality of data around the regression line are the likely reasons the regression is considerably less significant. Improvements to the model are discussed in the next sections.

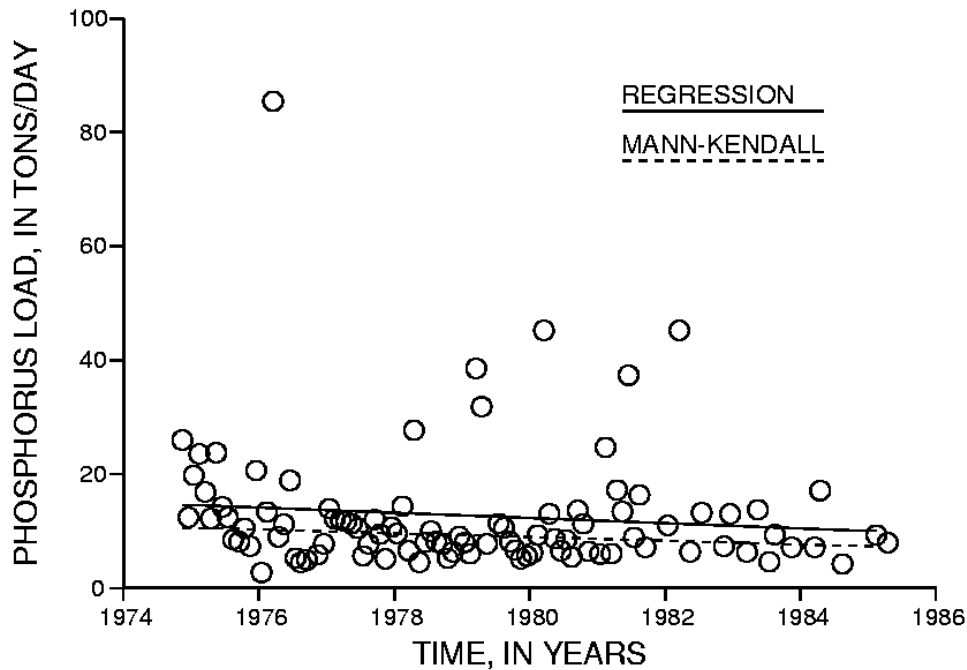


Figure 12.4 Mann-Kendall and regression trend lines (data in Appendix C10).

Regression:	Load = $16.8 - 0.46 \cdot \text{time}$	$t = -1.09$	$p = 0.28$
Mann-Kendall:	Load = $12.2 - 0.33 \cdot \text{time}$	$\text{tau} = -0.12$	$p = 0.08$.

12.3 Accounting for Exogenous Variables

Variables other than time trend often have considerable influence on the response variable Y. These "exogenous" variables are usually natural, random phenomena such as rainfall, temperature or streamflow. By removing the variation in Y caused by these variables, the background variability or "noise" is reduced so that any trend "signal" present can be seen. The ability (power) of a trend test to discern changes in Y with T is then increased. The removal process involves modelling, and thus explaining, the effect of exogenous variables with regression or LOWESS (for computation of LOWESS, see Chapter 10). This is the rationale for using the methods in the right-hand column of table 12.2.

For example, figure 12.5a presents a test for trend in dissolved solids at the James River in South Dakota. No adjustment for discharge was employed. The p-value for the test equals 0.47, so no trend is able to be seen. The Theil estimate of slope is plotted, showing the line

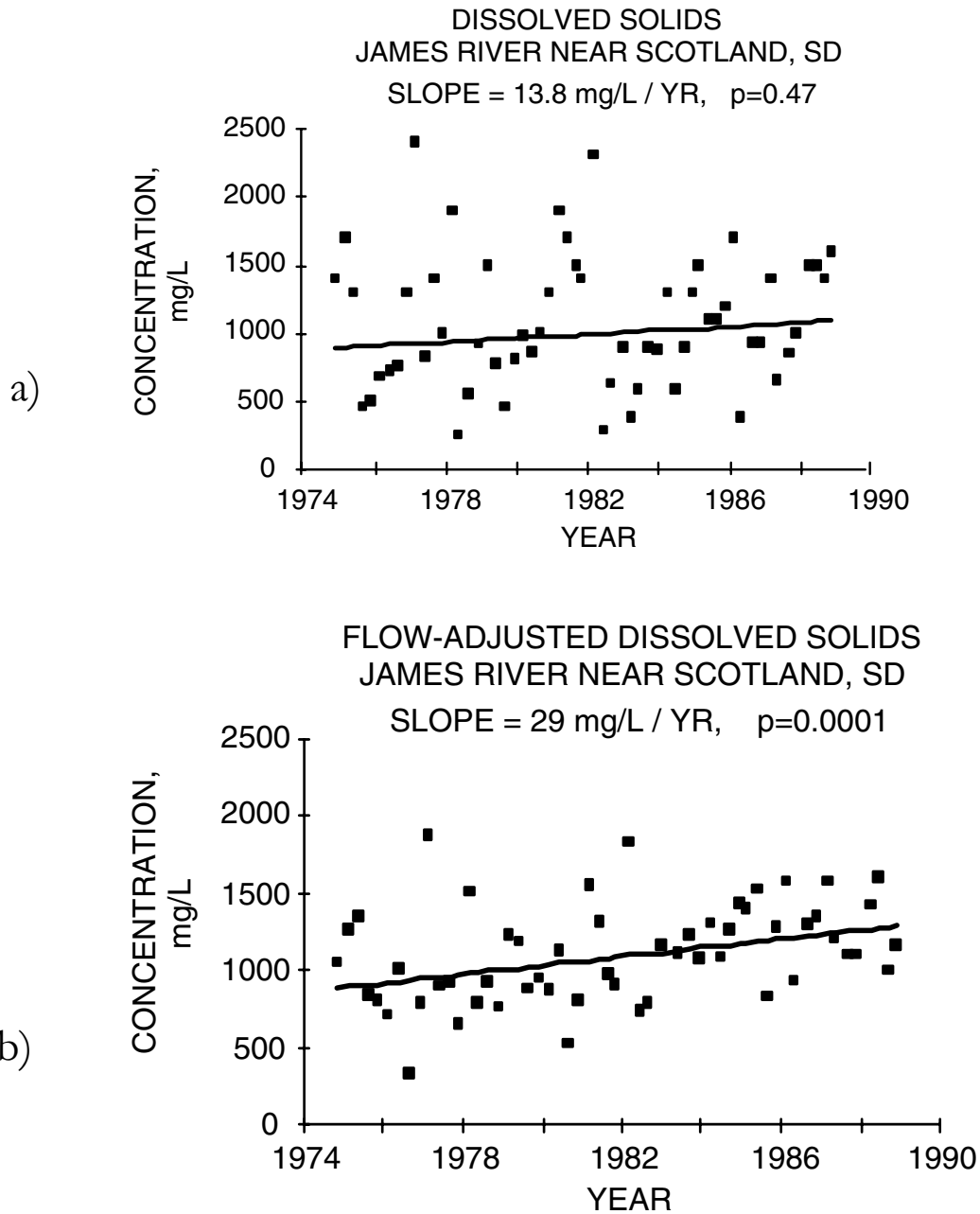
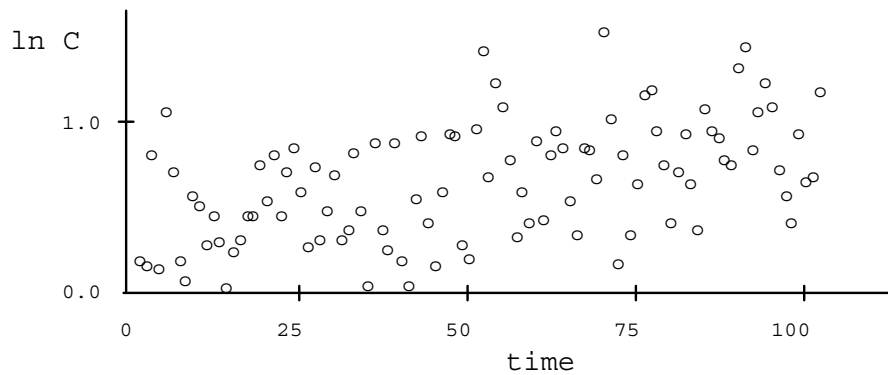


Figure 12.5 Trend tests a) before adjustment for flow. b) after adjustment for flow.
(from Hirsch et al., 1991)

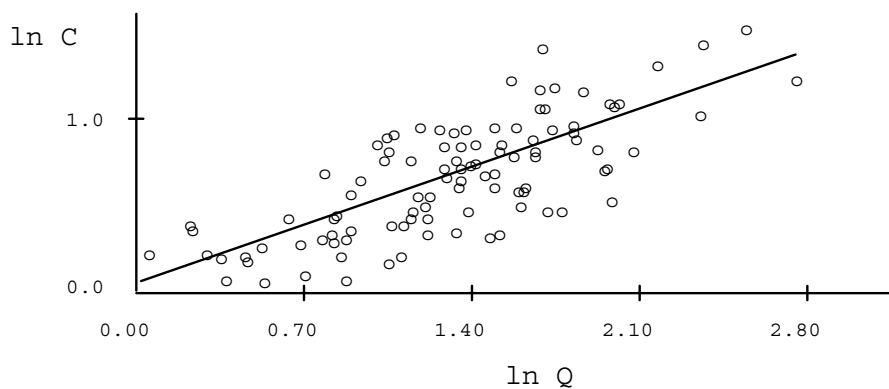
to be surrounded by a lot of data scatter. In figure 12.5b, the same data are plotted after using regression to remove the variation due to discharge. Note that the amount of scatter has

decreased. The same test for trend now has a p-value of 0.0001; for a given magnitude of flow, dissolved solids are increasing over time.

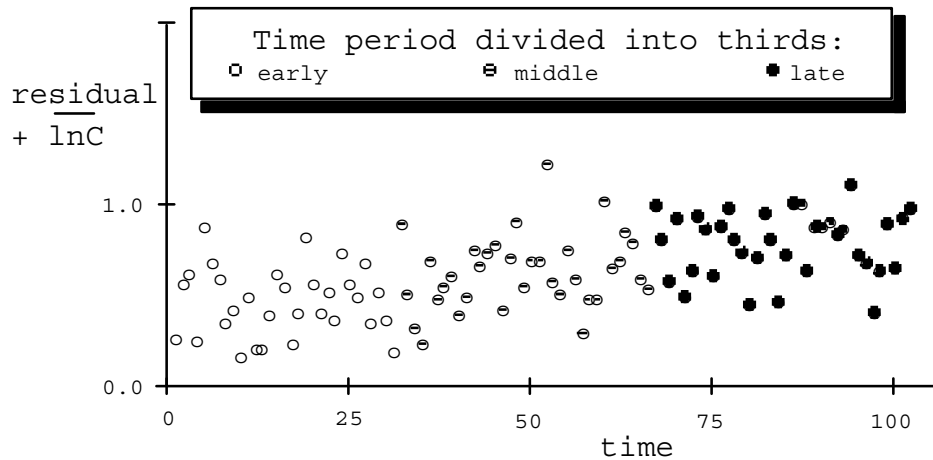
When removing the effect of one or more exogenous variables X , the probability distribution of the X s is assumed to be unchanged over the period of record. Consider a regression of Y versus X (figures 12.6a and 6b). The residuals R from the regression describe the values for the Y variable "adjusted for" exogenous variables (figure 12.6c). In other words, the effect of other variables is removed by using residuals -- residuals express the variation in Y over and above that due to the variation caused by changes in the exogenous variables. A trend in R implies a trend in the relationship between X and Y (figure 12.6d). This in turn implies a trend in the distribution of Y over time while accounting for X . However, if the probability distribution of the X s has changed over the period of record, a trend in the residuals may not necessarily be due to a trend in Y .



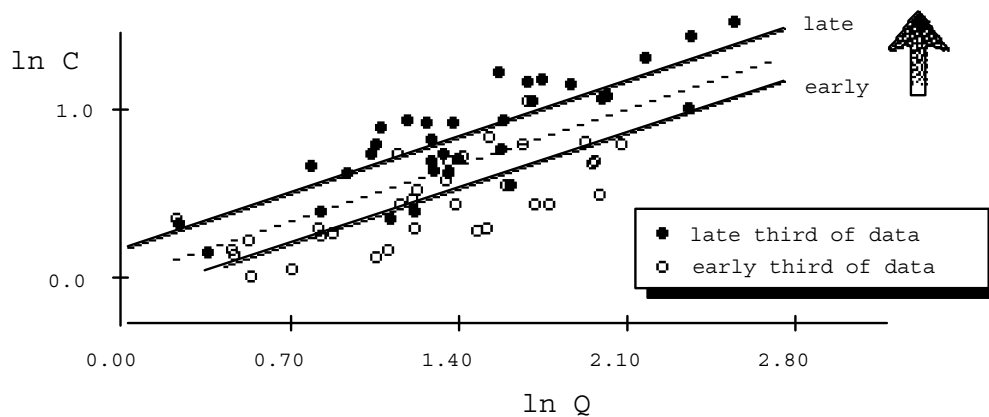
12.6a. Log of concentration vs. time. Trend is somewhat difficult to see.



12.6b. Ln of concentration vs. exogenous variable: ln of streamflow (Q).
Strong linear relation shown by regression line.
Expect higher concentrations at higher flows.



12.6c. Residuals from 12.6b regression over time. Trend much easier to detect than in 12.6a, as effect of Q has been removed by using residuals.



12.6d. Trend in fig. 6c can also be seen as an increase in the $\ln C$ vs $\ln Q$ relationship over time. For a given value of Q , the value for C increases over time.

What kind of variable is appropriate to select as an exogenous variable? It should be a measure of the driving force behind the process of interest, but must be free of changes in human manipulation. Thus a streamflow record that spans a major reservoir project, new diversion, or new operating policy on an existing system would be unacceptable, due to human alteration of the probability distribution of X during the period of interest. A streamflow record which reflects some human influence is acceptable, provided that the human influence is consistent over the period of record. Where human influence on streamflow records makes them unacceptable as X variables, two major alternatives exist. The first is to use flow at a nearby unaffected station which could be expected to be correlated with natural flow at the site of interest. The other alternative is to use weather-related data: rainfall over some antecedent period, or model-generated streamflows resulting from a deterministic watershed model that is driven by historical weather data.

Where Y is streamflow concentration, a great deal of the variance in Y is usually a function of river discharge. This comes about as a result of two different kinds of physical phenomena. One is dilution: a solute may be delivered to the stream at a reasonably constant rate (due to a point source or ground-water discharge to the stream) as discharge changes over time. The result of this situation is a decrease in concentration with increasing flow (see figure 12.7). This is typically seen in most of the major dissolved constituents (the major ions). The other process is wash-off: a solute, sediment, or a constituent attached to sediment can be delivered to the stream primarily from overland flow from paved areas or cultivated fields, or from streambank erosion. In these cases, concentrations as well as fluxes tend to rise with increasing discharge (see fig. 12.8). Some constituents can exhibit combinations of both of these kinds of behavior. One example is total phosphorus. A portion of the phosphorus may come from point sources such as sewage treatment plants (dilution effect), but another portion may be derived from surface wash-off and be attached to sediment particles (see fig. 12.9). The resulting pattern is an initial dilution, followed by a stronger increase with flow due to washoff.

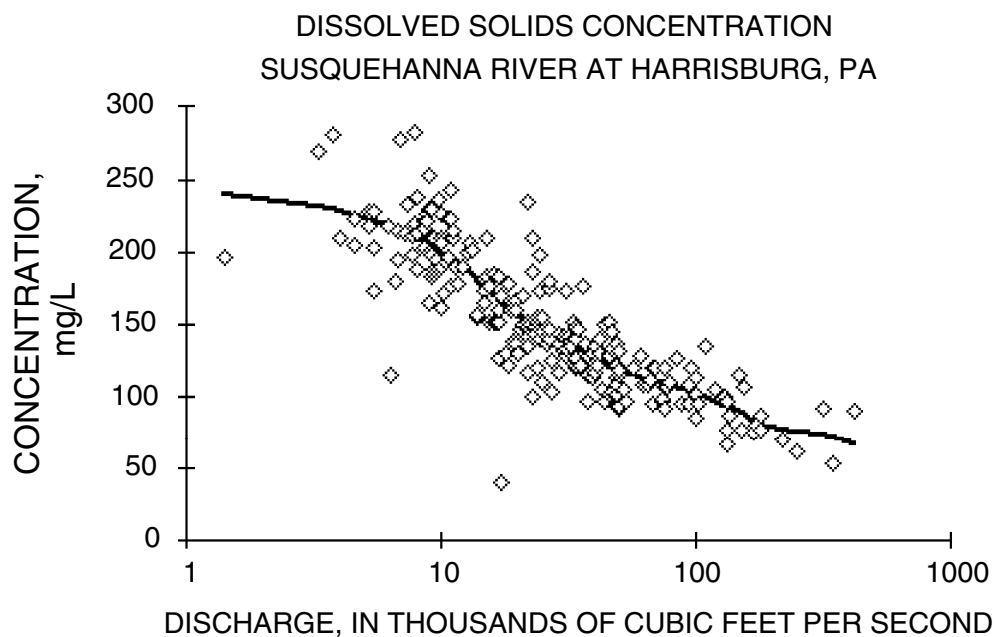


Figure 12.7 Dilution of dissolved solids with discharge (from Hirsch et al., 1991).

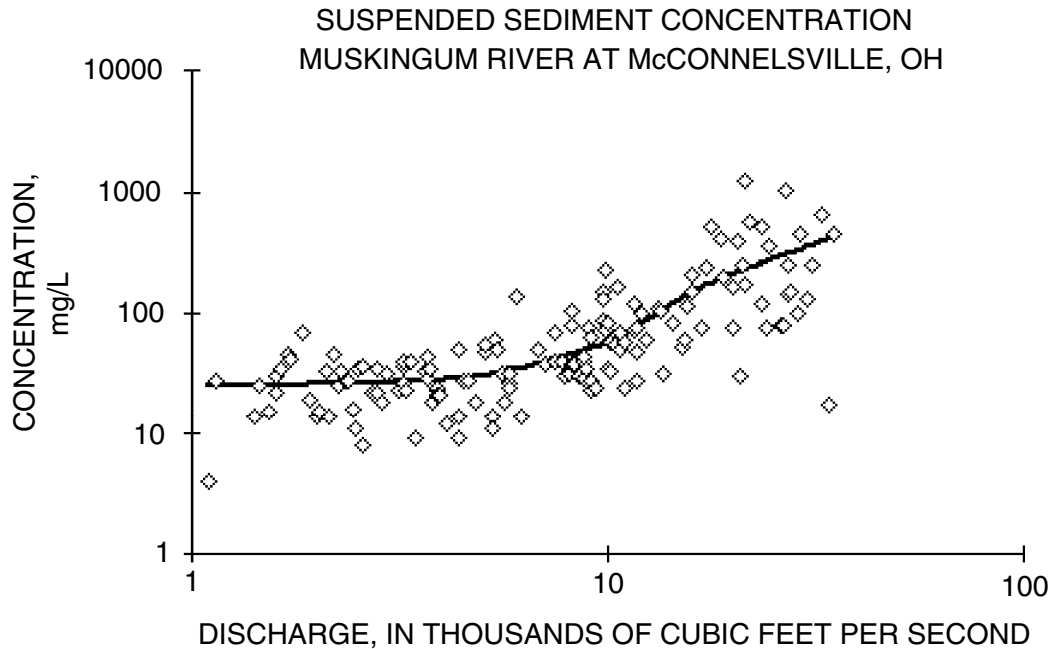


Figure 12.8 Washoff of suspended sediment with discharge (from Hirsch et al., 1991).

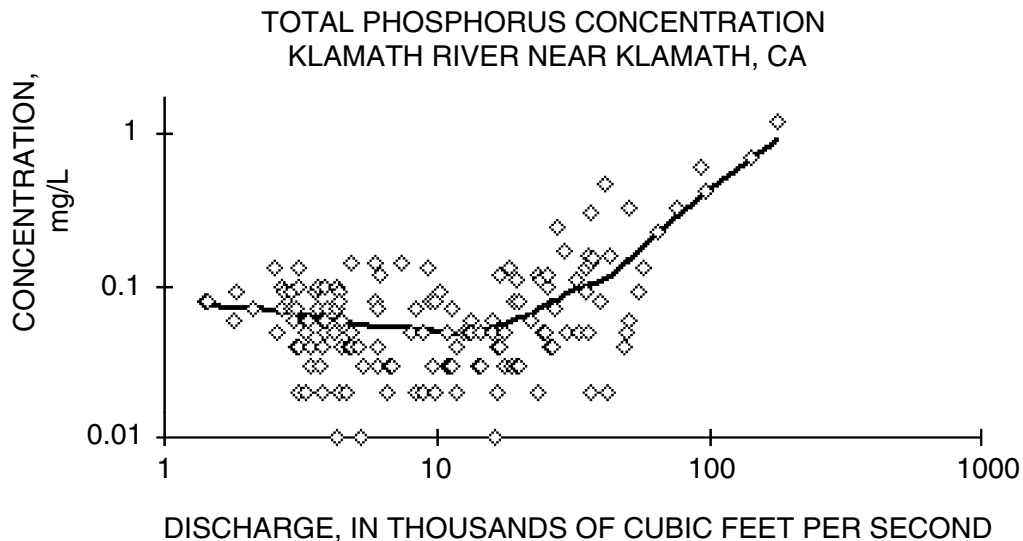


Figure 12.9 Dilution and subsequent washoff of total phosphorus as discharge increases (from Hirsch et al., 1991).

12.3.1 Nonparametric Approach

The smoothing technique LOWESS (LOcally WEighted Scatterplot Smooth) describes the relationship between Y and X without assuming linearity or normality of residuals. It is a robust description of the data pattern. Numerous smooth patterns result whose form changes as the

smoothing coefficient is altered. The LOWESS pattern chosen should be smooth enough that it doesn't have several local minima and maxima, but not so smooth as to eliminate true changes in slope. Given the LOWESS fitted value \hat{Y} the residuals R are computed as

$$R = Y - \hat{Y} .$$

Then the Kendall S statistic is computed from the R,T data pairs, and tested to see if it is significantly different from zero. The test for S is the test for trend.

12.3.2 Mixed approach: Mann-Kendall on Regression Residuals

To remove the effect of X on Y prior to performing the Mann-Kendall test, a linear regression could be computed between Y and one or more Xs:

$$Y = \beta_0 + \beta_1 \cdot X + \varepsilon$$

Unlike LOWESS, the adequacy of the regression model (is β_1 significant, should X be transformed due to lack of linearity or constant variance?) must be checked. When all is OK, the residuals R are computed as observed minus predicted values:

$$R = Y - b_0 - b_1 \cdot X .$$

Then the Kendall S statistic is computed from the R,T data pairs, and tested to see if it is significantly different from zero.

The Mann-Kendall test on residuals is a hybrid procedure -- parametric removal of effects of the exogenous variables, followed by a nonparametric test for trend. Care must be taken to insure that the model of Y versus X is reasonable (residuals should have no extreme outliers, Y is linear with X, etc.). The fully nonparametric alternative using LOWESS given in 12.3.1 avoids the necessity for close checking of assumptions.

Alley (1988) showed that this two-stage procedure resulted in lower power than an alternative which is analogous to the partial plots of Chapter 9. His "adjusted variable Kendall test" performs the second stage as a Mann-Kendall test of R versus T^* rather than R versus T, where T^* are the residuals from a regression of T versus X:

$$T = b_0 + b_1 \cdot X + T^*$$

In this way the effect of a drift in X over time is removed, so that the R versus T^* relationship is totally free of the influence of X. This test is a Mann-Kendall test on the partial plot of Y versus T, having removed the effect of all other exogenous variable(s) X from both Y and T by regression.

12.3.3 Parametric Approach

Consider the multiple regression of Y versus time T and one or more Xs:

$$Y = \beta_0 + \beta_1 \cdot T + \beta_2 \cdot X + \varepsilon .$$

The null hypothesis for the trend test is that $\beta_1 = 0$. Therefore the t-statistic for β_1 tests for trend. This test simultaneously compensates for the effects of exogenous variables by including them in the model. No two-stage process is necessary. The model must be checked for

adequacy – for the correct form of relationship (linear in the Xs and T), normality of residuals, and that b_2 is significantly different from zero. If b_1 is significantly different from zero (based on the t-statistic) then the null hypothesis of no trend is rejected, and we conclude that there is a linear trend in Y over T.

12.3.4 Comparison of Approaches

In general, the power and efficiency of any procedure for detecting and estimating the magnitude of trends will be aided if the variance of the data can be decreased (figure 12.5). This can be done by removing discharge effects either simultaneously or in stages. Simultaneous modelling of trend and discharge has a small but distinct advantage over the equivalent stagewise method (Alley, 1988). Thus parametric multiple regression has more power than a stagewise regression. The adjusted Kendall test has a similar advantage over the Mann-Kendall test on residuals R versus unadjusted T. We presume that a Mann-Kendall test of R on T* where both are computed using LOWESS (Y on X and T on X) would have similar advantages over the unadjusted method in section 12.3.1, though no data exists on this to date.

More important is whether the adjustment process should be conducted using a parametric or nonparametric method. The choice between regression and LOWESS should be based on the quality of the regression fit. LOWESS and linear regression fits of phosphorus concentration and stream discharge are compared for Klamath River in Figure 12.10. LOWESS would be a sensible alternative here due to the nonlinearity of the relationship. In studies where many data sets are being analyzed, and individualized checking of multiple models is impractical, LOWESS is the method of choice. It is also valuable when transformation of Y to achieve normality is not desirable. Where detailed model checking is practical and where high-quality parametric models can be constructed, multiple regression provides a one-step process with maximum efficiency. It and the adjusted Kendall method should be used over stagewise procedures.

All methods incorporating exogenous X variables discussed thus far assume that the change in the X,Y relationship over time is a parallel shift -- a change in intercept, no change in slope (see figure 12.6d). Changes in both (a rotation) are certainly possible. However it will not be possible to classify all such changes as uptrends or downtrends. For example, if the X,Y relationship pivots counterclockwise over time, then for high X there is an uptrend in Y and for low X there is a downtrend in Y. There is no simple way to generalize the Mann-Kendall test on residuals to identify such situations.

However, regression could be used as follows:

$$Y = \beta_0 + \beta_1 \cdot X + \beta_2 \cdot T + \beta_3 \cdot X \cdot T + \epsilon$$

The X•T term is an interaction term describing the rotation. One could compare this model to the "no trend" model (one with no T or X•T terms) using an F test. It can also be compared to

the simple trend model (one with an X and a T term but no X•T term) using a partial F test. The result will be selection of one of three outcomes: no trend, trend in the intercept, or trend in slope and intercept (rotation).

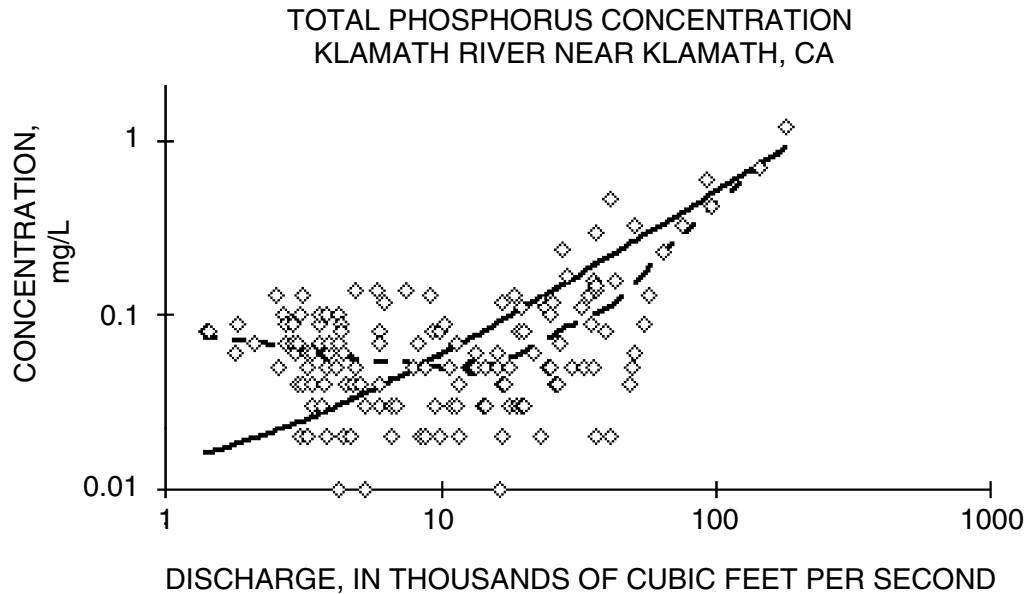


Figure 12.10 Comparison of LOWESS (dashed line) and linear regression (solid line) fits of concentration to stream discharge. From Hirsch, et al. (1991).

12.4 Dealing With Seasonality

There are many instances where changes between different seasons of the year are a major source of variation in the Y variable. As with other exogenous effects, seasonal variation must be compensated for or "removed" in order to better discern the trend in Y over time. If not, little power may be available to detect trends which are truly present. We may also be interested in modeling the seasonality to allow different predictions of Y for differing seasons.

Most concentrations in surface waters show strong seasonal patterns. Streamflow itself almost always varies greatly between seasons. This arises from seasonal variations in precipitation volume, and in temperature which in turn affects precipitation type (rain versus snow) and the rate of evapotranspiration. Some of the observed seasonal variation in water quality may be explained by accounting for this seasonal variation in discharge. However, seasonality often remains even after discharge effects have been removed (Hirsch et al. 1982). Possible additional causes of seasonal patterns include biological activity, both natural and managed activities such as agriculture. For example, nutrient concentrations vary with seasonal application of fertilizers and the natural pattern of uptake and release by plants. Other effects are due to different

sources of water dominant at different times of the year, such as snow melt versus intense rainfall. Seasonal rise and fall of ground water can also influence water quality. A given discharge in one season may derive mostly from ground water while the same discharge during the another season may result from surface runoff or quick flow through shallow soil horizons. The chemistry and sediment content of these sources may be quite different.

Techniques for dealing with seasonality fall into three major categories (table 12.3). One is fully nonparametric, one is a mixed procedure, and the last is fully parametric. In the first two procedures it is necessary to define a "season". In general, seasons should be just long enough so that there is some data available for most of the seasons in most of the years of record. For example, if the data are primarily collected at a monthly frequency, the seasons should be defined to be the 12 months. If the data are collected quarterly then there should be 4 seasons, etc. Tests for trend listed in table 12.2 have analogs which deal with seasonality. These are presented in table 12.3.

	Not Adjusted for X	Adjusted for X
Nonparametric	Seasonal Kendall test for trend on Y (Method I)	Seasonal Kendall trend test on residuals from LOWESS of Y on X (Method I)
Mixed	Regression of deseasonalized Y on T (Method II)	Seasonal Kendall trend test on residuals from regression of Y on X (Method I)
Parametric	Regression of Y on T and seasonal terms (Method III)	Regression of Y on X, T, and seasonal terms (Method III)

Table 12.3 Methods for dealing with seasonal patterns in trend testing

12.4.1 The Seasonal Kendall Test

The seasonal Kendall test (Hirsch et al., 1982) accounts for seasonality by computing the Mann-Kendall test on each of m seasons separately, and then combining the results. So for monthly "seasons", January data are compared only with January, February only with February, etc. No comparisons are made across season boundaries. Kendall's S_i statistic for each season are summed to form the overall statistic S_k .

$$S_k = \sum_{i=1}^m S_i \quad [12.1]$$

When the product of number of seasons and number of years is more than about 25, the distribution of S_k can be approximated quite well by a normal distribution with expectation equal to the sum of the expectations (zero) of the individual S_i under the null hypothesis, and variance equal to the sum of their variances. S_k is standardized (eq. 12.2) by subtracting its expectation $\mu_k = 0$ and dividing by its standard deviation σ_{S_k} . The result is evaluated against a table of the standard normal distribution.

$$Z_{S_k} = \begin{cases} \frac{S_k - 1}{\sigma_{S_k}} & \text{if } S_k > 0 \\ 0 & \text{if } S_k = 0 \\ \frac{S_k + 1}{\sigma_{S_k}} & \text{if } S_k < 0 \end{cases} \quad [12.2]$$

$$\text{where } \mu_{S_k} = 0, \\ \sigma_{S_k} = \sqrt{\sum_{i=1}^m (n_i/18) \cdot (n_i - 1) \cdot (2n_i + 5)}, \text{ and}$$

n_i = number of data in the i th season.

The null hypothesis is rejected at significance level α if $|Z_{S_k}| > Z_{\text{crit}}$ where Z_{crit} is the value of the standard normal distribution with a probability of exceedance of $\alpha/2$. When some of the Y and/or T values are tied the formula for σ_{S_k} must be modified, as discussed in Chapter 8. The significance test must also be modified for serial correlation between the seasonal test statistics (see Hirsch and Slack, 1984).

If there is variation in sampling frequency during the years of interest, the data set used in the trend test may need to be modified. If variations in sampling frequency are random (for example if there are a few instances where no value exists for some season of some year, and a few instances when two or three samples are available for some season of some year) then the data can be collapsed to a single value for each season of each year by taking the median of the available data in that season of that year. If, however, there is a systematic trend in sampling frequency (monthly for 7 years followed by quarterly for 5 years) then the following type of approach is necessary. Define the seasons on the basis of the lowest sampling frequency. For that part of the record with a higher frequency define the value for the season as the observation taken closest to the midpoint of the season. The reason for not using the median value in this case is that it will induce a trend in variance, which will invalidate the null distribution of the test statistic.

An estimate of the trend slope for Y over time T can be computed as the median of all slopes between data pairs within the same season (figure 12.11). Therefore no cross-season slopes contribute to the overall estimate of the Seasonal Kendall trend slope.

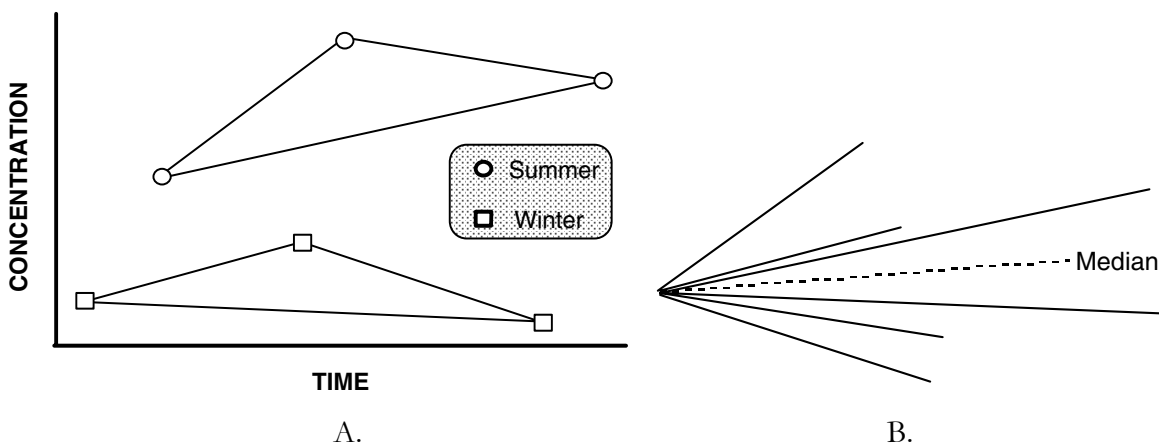


Figure 12.11 A. All pairwise slopes used to estimate the Seasonal Kendall trend slope (two seasons -- compare with figure 10.1).
B. Slopes rearranged to meet at a common origin

To accommodate and model the effects of exogenous variables, directly follow the methods of section 12.3 until the final step. Then apply the Seasonal Kendall rather than Mann-Kendall test on residuals from a LOWESS of Y versus X and T versus X (R versus T^*).

12.4.2 Mixture Methods

The seasonal Kendall test can be applied to residuals from a regression of Y versus X , rather than LOWESS. Keep in mind the discussion in the previous section of using adjusted variables T^* rather than T . Regression would be used only when the relationships exhibit adherence to the appropriate assumptions.

A second type of mixed procedure involves deseasonalizing the data by subtracting seasonal medians from all data within the season, and then regressing these deseasonalized data against time. One advantage of this procedure is that it produces a description of the pattern of the seasonality (in the form of the set of seasonal medians). However, this method has generally lower power to detect trend than other methods, and is not preferred over the other alternatives. Subtracting seasonal means would be equivalent to using dummy variables for $m-1$ seasons in a fully parametric regression. Either use up $m-1$ degrees of freedom in computing the seasonal statistics, a disadvantage which can be avoided by using the methods of the next section.

12.4.3 Multiple Regression With Periodic Functions

The third option is to use periodic functions to describe seasonal variation. The simplest case, one that is sufficient for most purposes, is:

$$Y = \beta_0 + \beta_1 \cdot \sin(2\pi T) + \beta_2 \cdot \cos(2\pi T) + \beta_3 \cdot T + \text{other terms} + \epsilon \quad [12.3]$$

where "other terms" are exogenous explanatory variables such as flow, rainfall, or level of some human activity (e.g. waste discharge, basin population, production). They may be continuous, or binary "dummy" variables as in analysis of covariance. The trend test is conducted by determining if the slope coefficient on T (β_3) is significantly different from zero. Other terms in the equation should be significant and appropriately modeled. The residuals ϵ must be approximately normal.

Time is commonly but not always expressed in units of years. Table 12.4 lists values for $2\pi T$ for three common time units: years, months and day of the year.

The expression	$2\pi T$	= $6.2832 \cdot t$	when t is expressed in years.
		= $0.5236 \cdot m$	when m is expressed in months.
		= $0.0172 \cdot d$	when d is expressed in day of year.

Table 12.4 Three values for $2\pi T$ useful in regression tests for trend

To more meaningfully interpret the sine and cosine terms, they can be re-expressed as the amplitude A of the cycle (half the distance from peak to trough) and the day of the year D_p at which the peak occurs:

$$\beta_1 \cdot \sin(2\pi t) + \beta_2 \cdot \cos(2\pi t) = A \sin[2\pi(t + t_0)] \quad [12.4]$$

where $A = \sqrt{\beta_1^2 + \beta_2^2}$ [12.5]

The phase shift $t_0 = \tan^{-1}(\beta_2 / \beta_1)$,

$t_0' = t_0 \pm 2\pi$ if necessary to get t_0 within the interval $0 < t_0 < 2\pi = 6.2832$

and $D_p = 58.019 \cdot (1.5708 - t_0')$ [12.6]

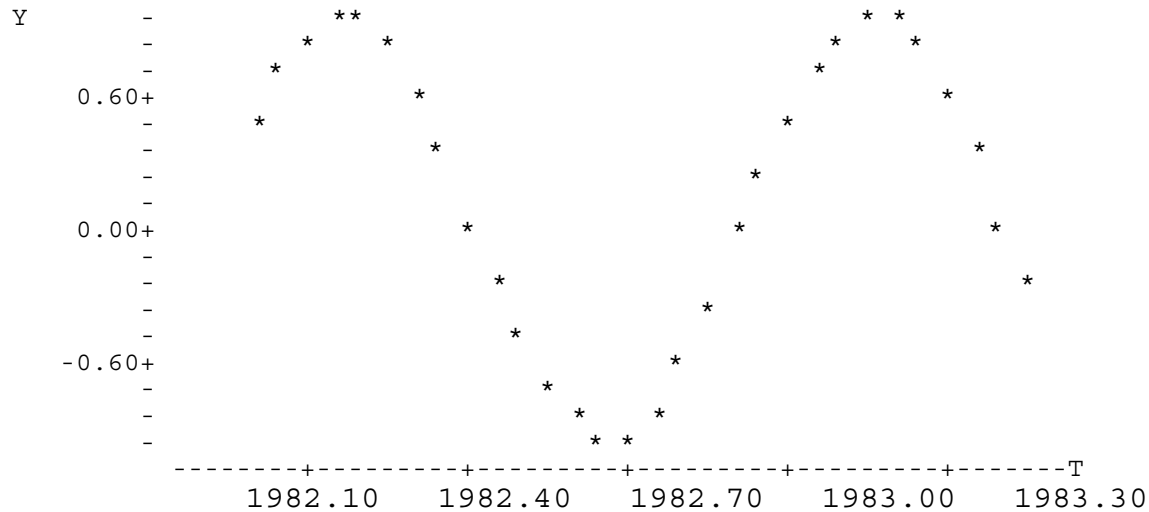
Example 2: Determining peak day and amplitude


Figure 12.12 Data showing seasonal (sine and cosine) pattern.

The data in figure 12.12 were generated using coefficients $b_1 = 0.8$ and $b_2 = 0.5$. From equations 12.4 through 12.6, the amplitude $A = 0.94$, $t_0 = 0.559$, $t_0' = 0.559$, and so the peak day = 59 (February 28).

After including sine and cosine terms in a multiple regression to account for seasonality, the residuals may still show a seasonal pattern in boxplots by season, or in a Kruskal-Wallis test by season. If this occurs, additional periodic functions with periods of 1/2 or 1/3 or other fractions of a year (multiple cycles per year) may be used to remove additional seasonality. Credible explanations for why such cycles might occur are always helpful. For example, the following equation may be appropriate:

$$Y = \beta_0 + \beta_1 \cdot \sin(2\pi t) + \beta_2 \cdot \cos(2\pi t) + \beta_3 \cdot \sin(4\pi t) + \beta_4 \cdot \cos(4\pi t) + \text{other terms} + \varepsilon$$

One way to determine how many terms to use is to add them, two at a time, to the regression and at each step do an F test for the significance of the new pair of terms. As a result one may, very legitimately, settle on a model in which the t-statistics for one of a pair of coefficients is not significant, but as a pair they are significant. Leaving out just the sine or just the cosine is not a sensible thing to do, because it forces the periodic term to have a completely arbitrary phase shift, rather than one determined by the data.

12.4.4 Comparison of Methods

The Mann-Kendall and mixed approaches have the disadvantages of only being applicable to univariate data (either original units or residuals from a previous analysis) and are not amenable to simultaneous analysis of multiple sources of variation. They take at least two steps to

compute. Multiple regression allows many variables to be considered easily and simultaneously by a single model.

Mann-Kendall has the usual advantage of nonparametrics: robustness against departures from normality. The mixed method is perhaps the least robust because the individual seasonal data sets can be quite small and the estimated seasonal medians can follow an irregular pattern. In general this method has far more parameters than either of the other two methods and fails to take advantage of the idea that geophysical processes have some degree of smoothness in the annual cycle. That is: it is unlikely that April will be very different from May, even though the sample statistics may suggest that this is so.

Regression with periodic functions takes advantage of this notion of smoothness and thereby involves very few parameters. However, the functional form (sine and cosine terms) can become a "straight jacket". Perhaps the annual cycles really do have abrupt breaks associated with freezing and thawing, or the growing season. Regression can always use binary variables as abrupt definitions of season ($G=1$ for "growing season", $G=0$ otherwise). Observations can be assigned to a season based on conditions which may vary in date from year to year, and not just based on the date itself. Regression could also be modified to accept other periodic functions, perhaps ones that are much more squared off. To do this demands a good physically-based definition of the timing of the influential factors, however.

All three methods provide a description of the seasonal pattern. Regression and mixed methods automatically produce seasonal summary statistics. However, there is no difficulty in providing a measure of seasonality consistent with Mann-Kendall by computing seasonal medians of the data after trend effects have been removed.

12.4.5 Presenting Seasonal Effects

There are many ways of characterizing the seasonality of a data set (table 12.5). Any of them can be applied to the raw data or to residuals from a LOWESS or regression that removes the effects of some exogenous variable. In general, graphical techniques will be more interpretable than tabular, although the detail of tables may sometimes be needed.

	GRAPHICAL METHODS	TABULAR METHODS
BEST	Boxplots by season, or LOWESS of data versus time of year	List the amplitude and peak day of the cycle
NEXT BEST		List of seasonal medians and seasonal interquartile ranges, or list of distribution percentage points by season
WORST	Plot of seasonal means with standard deviation or standard error bars around them	List of seasonal means, standard deviations, or standard errors

Table 12.5 Rating of methods for dealing with seasonality

12.4.6 Differences Between Seasonal Patterns

The approaches described above all assume a single pattern of trend across all seasons. This may be a gross over-simplification and can fail to reveal large differences in behavior between different seasons. It is entirely possible that the Y variable exhibits a strong trend in its summer values and no trend in the other seasons. Even worse, it could be that spring and summer have strong up-trends and fall and winter have strong down-trends, cancelling each other out and resulting in an overall seasonal Kendall test statistic stating no trend. Another situation might arise where the X-Y relationship (e.g. rainfall-runoff, flow-concentration) has a substantially different slope and intercept for different seasons.

No overall test statistic will provide any clue of these differences. This is not to suggest they are not useful. Many times we desire a single number to characterize what is happening in a data set. Particularly when dealing with several data sets (multiple stations and/or multiple variables), breaking the problem down into 4 seasons or 12 months simply swamps us with more results than can be absorbed. Also, if the various seasons do show a consistent pattern of behavior there is great strength in looking at them in one analysis. For example in a seasonal Kendall analysis each month viewed by itself might show a positive S value, none of which is significant, but the overall Seasonal Kendall test could be highly significant. Yet in detailed examinations of individual stations it is often useful to perform and present the full, within-season analysis on each season. Figure 12.13 is a good approach to graphically presenting the results of such multi-season analyses.

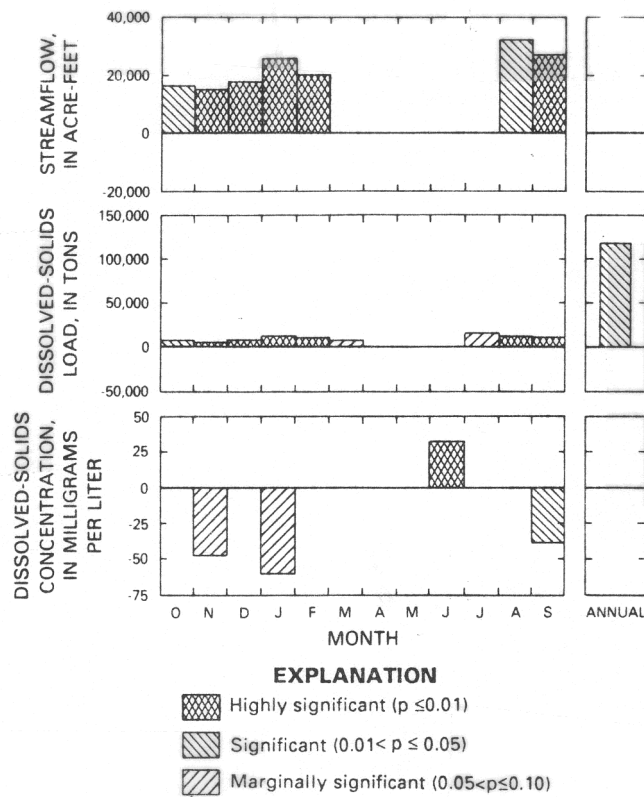


Figure 12.13 Illustration of seasonal and annual step trends on the Green River (from Liebermann and others, 1989)

In the approaches using Method I above, one can also examine "contrasts" between the different seasonal statistics. This provides a single statistic which indicates whether the seasons are behaving in a similar fashion (homogeneous) or behaving differently from each other (heterogeneous). The test for homogeneity is described by van Belle and Hughes (1984).

For each season i ($i=1,2,\dots,m$) compute $Z_i = S_i / \sqrt{\text{Var}(S_i)}$. Sum these to compute the "total" chi-square statistic, then compute "trend" and "homogeneous" chi-squares:

$$\chi^2_{(\text{total})} = \sum_{i=1}^m Z_i^2 \tag{12.7}$$

$$\chi^2_{(\text{trend})} = m \cdot \bar{Z}^2 \quad \text{where } \bar{Z} = \frac{\sum_{i=1}^m Z_i}{m} \tag{12.8}$$

$$\chi^2_{(\text{homogeneous})} = \chi^2_{(\text{total})} - \chi^2_{(\text{trend})} \tag{12.9}$$

The null hypothesis that the seasons are homogeneous with respect to trend ($\tau_1 = \tau_2 = \dots = \tau_m$) is tested by comparing $\chi^2_{(\text{homogeneous})}$ to tables of the chi-square distribution with $m-1$ degrees of freedom. If it exceeds the critical value for the pre-selected α , reject the null hypothesis and conclude that different seasons exhibit different trends.

12.5 Use of Transformations in Trend Studies

Water resources data commonly exhibit substantial departures from a normal distribution. Surface-water concentration, load, and flow data are often positively skewed, with many observations lying close to a lower bound of zero and a few observations one or more orders of magnitude above the lower values. If only a test for trend is of interest, then the decision to make some monotonic transformation of the data (to render them more nearly normal) is of no consequence provided that a nonparametric test is used. Nonparametric trend tests are invariant to monotonic power transformations (such as the logarithm or square root). In terms of significance levels the test results will be identical whether the test was applied to the raw data or the transformed data.

The decision to transform data is, however, highly important in terms of any of the procedures for removing the effects of exogenous variables (X), for computing significance levels of a parametric test (figure 12.14), and for computing and expressing slope estimates. Trends which are nonlinear (say exponential or quadratic) will be poorly described by a linear slope coefficient, whether from regression or a nonparametric method. It is quite possible that negative predictions may result for some values of time or X. By transforming the data so that the trend is linear, a Mann-Kendall or regression slope can later be re-expressed back into original units. The resulting nonlinear trend will better fit the data than the linear expression, even though their nonparametric significance tests are identical. Thus, it may be appropriate to run analyses on transformed Y values, even if the analysis is a nonparametric one.

One way to ensure that the fitted trend line will not predict negative values is to take a log transformation of the data prior to trend analysis. The trend slope will then be expressed in log units. A linear trend in log units translates to an exponential trend in original units, which can then be re-expressed in percent per year to make the trend easier to interpret. If b_1 is the estimated slope of a linear trend in natural log units then the percentage change from the beginning of any year to the end of that year will be $(e^{b_1} - 1) \cdot 100$. If slopes in original units are preferred, then instead of multiplying by 100, multiply by some measure of central tendency in the data (mean or median) to express the slope or step-trend in original units.

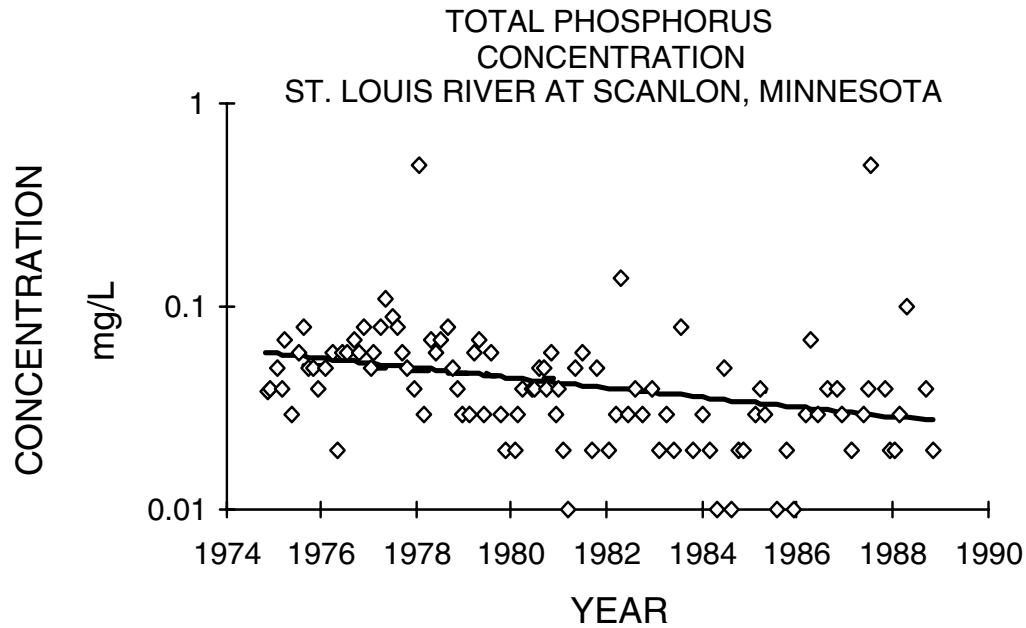


Figure 12.14 No trend evident in concentration ($p=0.432$ for regression slope). After log transformation, there is a statistically significant decline ($p=0.001$). Regression on logs shown as solid line. (from Hirsch et al., 1991).

In general, more resistant and robust results can be obtained if log transformations are used for variables that typically have ranges of more than an order of magnitude. With variations this large, transformations should be used in conjunction with both parametric and nonparametric tests. However, in multiple record analyses the decision to transform should be made on the basis of the characteristics of the class of variables being studied, not on a case-by-case basis. Variables on which log transforms are typically helpful include: flood flows, low flows, monthly or annual flows in small river basins, concentrations of sediment, total concentration (suspended plus dissolved) for a constituent when the suspended fraction is substantial (for example phosphorus and some metals), concentrations or counts of organisms, concentrations of substances that arise from biological processes (such as chlorophyll), and downstream load for virtually any constituent.

Some argue that data should always be transformed to normality, and parametric procedures computed on the transformed data. Transformations to normality are not always possible, as some data are non-normal due not to skewness but to heavy tails of the distribution (see Schertz and Hirsch, 1985). The strongest argument for transformations are that regression methods allows simultaneous consideration of the effects of multiple exogenous variables along with temporal trend. Such simultaneous tests are more difficult with nonparametric techniques. Multivariate smoothing methods are available (Cleveland and Devlin, 1988) which at least allow removal of multiple exogenous effects in one step, but they are not implemented yet in any commercial software.

The weakest situations for parametric techniques on transformed data are for analyses of multiple data sets. The transformation appropriate to one data set may not be appropriate to another. If different transformations are used on different data sets then comparisons among results is difficult, if not impossible. Also, there is an element of subjectivity in the choice of transformation. The argument of the skeptic that: "You can always reach the conclusion you want if you manipulate the data enough" is not without merit. The credibility of results is enhanced if a single statistical method is used for all data sets in a study, and this is next to impossible with the several judgements of model adequacy required for parametric methods. Nonparametric procedures are therefore well suited to multi-record trend analysis studies. In analyses of individual records, use of transformations with parametric methods can be very appropriate.

12.6 Monotonic Trend versus Two Sample (Step) Trend

Study of long term changes in hydrologic variables can be carried out in either of two modes. Up to this point "monotonic trends" were discussed, gradual and continuing changes over time. The Mann-Kendall test and regression are the two basic tools used in this case. The other mode compares two non-overlapping sets of data, an "early" and "late" period of record. Changes between the periods are called "step trends", as values of Y step up or down from one time period to the next. Testing for differences between these two groups involves procedures similar or identical to those described in other chapters, including the rank-sum test, two-sample t -tests, and analysis of covariance. Each of them also can be modified to account for seasonality.

The basic parametric test for step trends is the two-sample t -test. See Chapter 5 for its computation. The magnitude of change is measured by the difference in sample means between the two periods. Helsel and Hirsch (1988) discuss the disadvantages of using a t -test for step trends on data which are non-normal -- loss of power, inability to incorporate data below the detection limit, and an inappropriate measure of the step trend size. The primary nonparametric alternative is the rank-sum test and associated Hodges-Lehmann (H-L) estimator of step-trend magnitude (see Chapter 5, and Hirsch, 1988). The H-L estimator is the median of all possible differences between data in the "before" and "after" periods. Table 12.6 summarizes the step-trend approaches not considering seasonality and 12.7 summarizes those which consider seasonality. The rank-sum test can be implemented in a seasonal manner just like the Mann-Kendall test, called the seasonal rank-sum test. It computes the rank-sum statistic separately for each season, sums the test statistics, their expectations and variances, and then evaluates the overall summed test statistic. The H-L estimator can be similarly modified by considering only data pairs within a given season.

	Not Adjusted for X	Adjusted for X
Nonparametric	Rank-sum test on Y	Rank-sum test on residuals from LOWESS of Y on X
Mixed	---	Rank-sum test on residuals from regression of Y on X
Parametric	Two sample t-test	Analysis of covariance of Y on X and group

Table 12.6 Step-trend tests (two-sample) which do not consider seasonality (note "group" refers to a dummy variable 0 for "before" and 1 for "after")

	Not Adjusted for X	Adjusted for X
Nonparametric	Seasonal rank-sum test on Y	Seasonal rank-sum test on residuals from LOWESS of Y on X
Mixed	Two-sample t test on deseasonalized Y	Seasonal rank-sum test on residuals from regression of Y on X
Parametric	Analysis of covariance of Y on seasonal terms and group	Analysis of covariance of Y on X, seasonal terms, and group

Table 12.7 Step-trend tests (two-sample) which do consider seasonality (note "group" refers to a dummy variable 0 for "before" and 1 for "after")

Step trend procedures should be used in two situations. The first is when the record (or records) being analyzed are naturally broken into two distinct time periods with a relatively long gap between them. There is no specific rule to determine how long the gap should be to make this the preferred procedure. If the length of the gap is more than about one-third the entire period of data collection, then the step trend procedure is probably best (see figure 12.15). In general, if the within-period trends are small in comparison to the between-period differences, then step-trend procedures should be used.

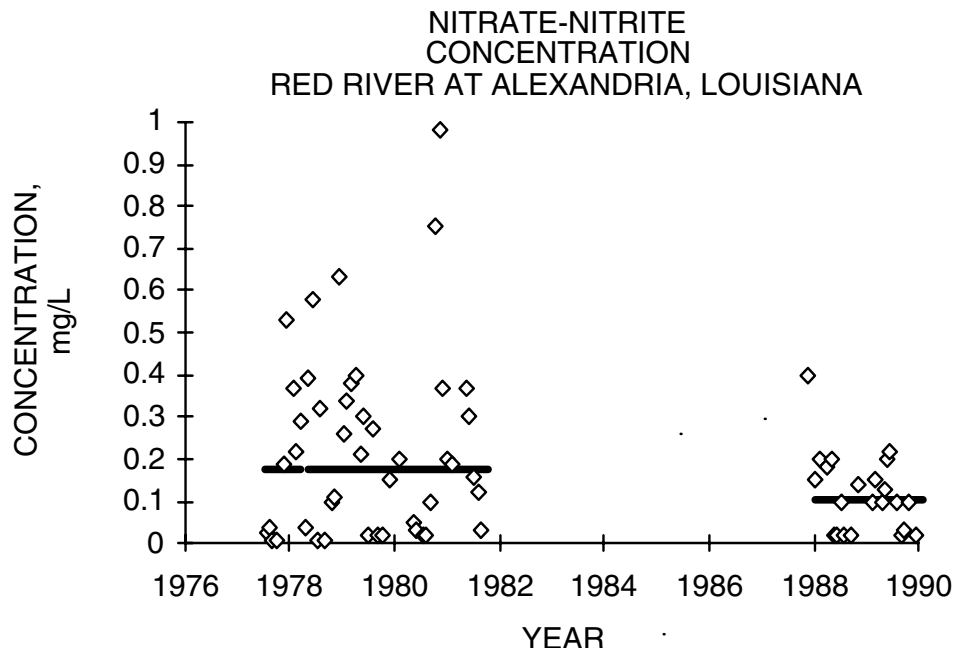


Figure 12.15 Significant ($p=0.085$) step trend as measured by rank-sum test. Solid lines are group medians. Monotonic trend test is not significant ($p=0.167$). Modified from Hirsch et al., 1991.

The second situation to test for step-trend is when a known event has occurred at a specific time during the record which is likely to have changed water quality. The record is first divided into "before" and "after" periods at the time of this known event. Example events are the completion of a dam or diversion, the introduction of a new source of contaminants, reduction in some contaminant due to completion of treatment plant improvements, or the closing of some facility (figure 12.16). It is imperative that the decision to use step-trend procedures not be based on examination of the data (i.e. the analyst notices an apparent step but had no prior hypothesis that it should have occurred), or on a computation of the time which maximizes the difference between periods. Such a prior investigation biases the significance level of the test, finding changes which are not really there. Step-trend procedures require a highly specific situation, and the decision to use them should be made prior to any examination of the data.

If there is no prior hypothesis of a time of change or if records from a variety of stations are being analyzed in a single study, monotonic trend procedures are most appropriate. In multiple record studies, even when some of the records have extensive but not identical gaps, the monotonic trend procedures are generally best because comparable periods of time are more easily examined among all the records. In fact, the frequent problem of multiple starting dates, ending dates, and gaps in a group of records presents a significant practical problem in trend analysis studies. In order to correctly interpret the data, records examined in a multiple station study must be concurrent. For example it is pointless to compare a 1975-1985 trend at one

station to a 1960-1980 trend at another. The difficulty arises in selecting a period which is long enough to be meaningful but does not exclude too many shorter records.

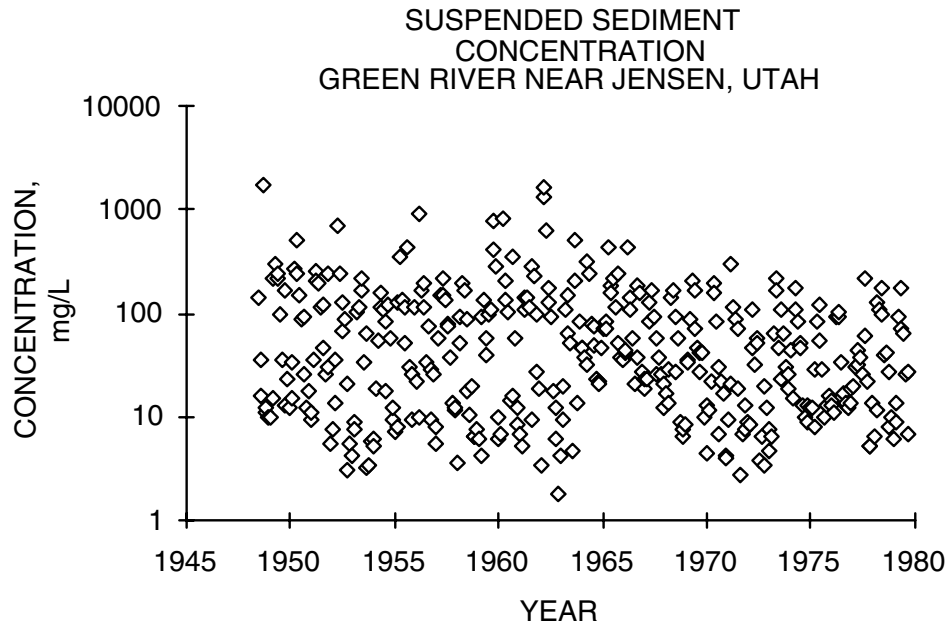


Figure 12.16 A weakly significant ($p=0.105$) reduction in suspended sediment after completion of the Flaming Gorge reservoir (located 93 miles upstream of the station) in late 1962 as measured by a rank-sum test. From Hirsch et al., 1991.

A further difficulty involves deciding just how complete a record must be to be included in the analysis. For example, if the study is for 1970-1985 and there is a record that runs from 1972 through 1985 it is probably prudent to include it in the study. A one- or two-year gap in the middle of the record should not disqualify a station from the analysis. More difficult are questions such as inclusion of a 1976-1984 record, or inclusion of a record that covers 1970-1975 and 1982-1985. One reasonably objective rule for deciding whether to include a record is:

- 1) divide the study period into thirds (three periods of equal length),
- 2) determine the coverage in each period (e.g. if the record is generally monthly, count the months for which there are data),
- 3) if any of the thirds has less than 20 percent of the total coverage then the record should not be included in the analysis. See Schertz (1990) for an application of these kinds of approaches.

12.7 Applicability of Trend Tests With Censored Data

Censored samples are records in which some of the data are known only to be "less than" or "greater than" some threshold (see Chapter 13). The two most common examples in hydrology are constituent concentrations less than the detection limit and floods which are known to be less than some threshold of perception (e.g. the annual flood of 1887 was not sufficiently large that local record keepers bothered to record the maximum stage). The existence of censored values complicates the use of the previously discussed parametric procedures and all of the procedures involving removal of the effect of an exogenous variable. Any arbitrary choice of a value to represent the censored values (e.g., zero or the reporting limit) can give inaccurate results for hypothesis tests and biased estimates of trend slopes (Helsel, 1990).

A parametric approach to the detection of trends in censored data is the estimation of the parameters of a linear regression model relating Y to T , or Y to T and X , through the method of maximum likelihood estimation (MLE), also referred to as Tobit estimation (Hald, 1949; Cohen, 1950). These effects can be modeled simultaneously in this approach as can be done in a conventional multiple regression. Because the MLE method assumes a linear model with normally distributed errors, transformations (such as logarithms) of Y and X are frequently required to make the data more nearly normal and improve the fit of the MLE regression. Failure of the data to conform to these assumptions will tend to lower the statistical power of the test, and give unreliable estimates of the model parameters. The Type I error of the test is, however, relatively insensitive to violations of the normality assumption.

An extension of the MLE method was developed by Cohen (1976) to provide estimates of regression model parameters for data records with multiple censoring levels. An adjusted MLE method for multiply-censored data that is less biased in certain applications than the MLE method of Cohen (1976) was also recently developed by Cohn (1988). The availability of multiply-censored MLE methods is noteworthy for the analysis of lengthy water-quality records with censored values since these records frequently have multiple reporting limits that reflect improvements in the accuracy of analytical methods (and reductions in reporting limits) with time. Similarly, the multiply-censored case can arise in flood studies in that some very old portion of a flood record may contain estimates of only the very largest floods while a more recent part of the record (when flood plain development was more intense and record keeping more complete) may contain estimates of floods exceeding a more moderate threshold.

The Mann-Kendall test can be used without any difficulty when only one censoring threshold exists. Comparisons between all pairs of observations are possible. All the "less thans" are less than the other values and are considered to be tied with each other. Thus the S statistic and τ are easily computed using the tie correction for the standard deviation (see Chapter 8) in the large-sample approximation. Equation 8.4 for the corrected standard deviation is repeated here.

$$\sigma_S = \sqrt{\frac{[n(n-1)(2n+5) - \sum_{i=1}^n t_i(i-1)(2i+5)]}{18}} \quad [12.10]$$

When more than one detection limit exists, the Mann-Kendall test can not be performed without further censoring the data. Consider the data set: <1, <1, 3, <5, 7. How can a <1 and <5 be compared? A 3 and a <5? These ambiguities make the test impossible to compute. The only way to perform a Mann-Kendall test is to censor and recode the data at the highest detection limit. Thus, these data would become: <5, <5, <5, <5, 7. There is certainly a loss of information in making this change, and a loss of power to detect any trends which may exist. Other nonparametric tests incorporating multiple detection limits are briefly discussed in chapter 13. None can be performed when the change in detection limits is a function of the process being tested for, such as time. When this occurs, the change in censoring limit over time will induce the test to find a change in Y over time more often than it actually occurs (α will actually be larger than that which was stated for the test).

Although the sign of the estimated Theil trend slope is accurate for highly censored data records, the magnitude of the slope estimate is likely to be in error. Substitution of an arbitrarily chosen value between zero and the reporting limit for all censored values will likely give biased estimates of the trend slope. While the amount of bias cannot be stated precisely, the presence of only a few nondetected values in a record (less than about five percent) is not likely to affect the accuracy of the trend slope magnitude significantly.

Table 12.8 classifies monotonic trend tests for data sets with censored data.

	Not Adjusted for X	Adjusted for X
Nonparametric	Mann-Kendall test for trend on Y	No test yet available
Fully Parametric	Tobit regression of Y on T	Tobit regression of Y on X and T

Table 12.8 Classification of monotonic trend tests for censored data

Exercises

- 12.1 During the period 1962-1969 the Green River Dam was constructed about 35 miles upstream of a gaging station on the Green River at Munfordville, Kentucky. It is a flood control dam, thus it regulates flow (changes the flow duration curve) but has little or no effect on total annual runoff from this 1660 square-mile basin.

The question is this--over the period of record 1952-1972 (which includes pre-dam, transition, and regulated periods) has there been a monotonic trend in sediment transport? The data available in Appendix C17 are the year, suspended sediment load in thousands of tons per year, and the annual discharge in cfs-days.

Using each of the four trend analysis approaches described in the chapter, what would you conclude about suspended sediment trends?

- 12.2: Seasonal Kendall Test with censored data

The following data are dissolved lead concentrations (in mg/L) for the Potomac River at Chain Bridge at Washington, D.C.

	<u>Winter</u>	<u>Spring</u>	<u>Summer</u>	<u>Fall</u>
1973	–	4	3	–
1974	–	3	2	6
1975	4	5	5	11
1976	2	3	6	4
1977	16	2	18	17
1978	26	7	4	–
1979	5	3	9	6
1980	<2	<2	<2	<2
1981	5	<2	<2	<2
1982	<2	2	–	2
1983	2	3	<2	5
1984	3	5	<2	<2
1985	<2	3	–	<2

Compute Kendall's S and Var [S] for each season and perform a Mann-Kendall test for each season, reporting p-values. Then compute the Seasonal Kendall statistic and its variance and report the p-value for the Seasonal Kendall test.

- 12.3 The Maumee River, Ohio is a major tributary to Lake Erie. The control of phosphorus inputs to the lake has been a major concern since the early 1970's. In Appendix C18 are 133 observations of instantaneous load and streamflow for the Maumee River from the 1972 through 1986. The variables in the data set are TIME (decimal time in years), MONTH (month of the year), Q (discharge in thousands of cfs), and LOAD (total phosphorus load in tons per day).

Using the various approaches listed in this chapter, describe the trend in the data set. Try each of the approaches listed in table 12.2. For the nonparametric approaches compute the seasonal-Kendall test.

Then using the model you select as most appropriate, estimate loads for the following four cases:

$$(Q, \text{TIME}) = (11, 1972.5), (1, 1972.5), (11, 1986.5), (1, 1986.5).$$

- 12.4 Major ion chemistry of groundwater may be altered as it remains in contact with aquifer materials. One common alteration is ion exchange of calcium with sodium. Carbon-14 can be used as a measure of the age of groundwater, with C-14 (modern carbon) decreasing with increasing age. Compute the p-value and determine whether there is a trend to lower calcium concentrations with increasing age for the following data:

	<u>%Carbon-14</u>	<u>Calcium, in % meq</u>		<u>%Carbon-14</u>	<u>Calcium, in % meq</u>
(old)	12.4	20.16		48.9	16.56
	19.5	0.46		52.0	22.88
	23.9	4.18		55.6	12.64
	25.3	28.95		57.7	14.26
	28.2	20.00		58.1	13.37
	31.5	0.57		66.2	35.22
	33.1	1.84		67.1	24.43
	33.4	13.99	(young)	71.8	60.24
	38.6	18.02			

- 12.5 A well screened in the middle aquifer in New Jersey is investigated to determine if groundwater levels have declined over time as a result of pumpage. The data are given in Appendix C19 (A. Pucci, written communication, U.S. Geological Survey, Trenton, NJ). Conduct a trend analysis to determine whether a decline has occurred. Estimate the rate of decline over the period of record.

