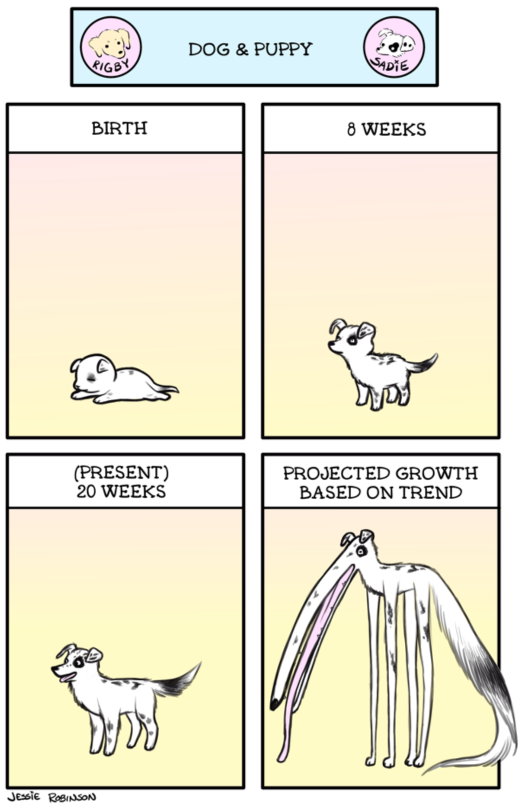
Trend Analysis

EN5423 | Spring 2024

w16_ trend_analysis_01.pdf
(Week 16)

Contents



Intro

• *Over a 20-year period, concentrations and loads of phosphorus were monitored in numerous tributaries to an estuary to determine changes in their central tendency, the confidence in these changes, and whether these changes are influenced by weather variations or regulatory actions, such as the ban on phosphorus compounds in detergents.*

• *Groundwater levels recorded over 14 years, with increased withdrawals due to new irrigation systems in the ninth year, were analyzed to determine any resultant decreases, the extent of these decreases, and the confidence in these estimates.*

• Trend analysis involves building statistical models to understand the behavior of environmental variables like discharge, solute concentrations, water levels, and temperature over time. These models include components such as regular cycles (seasonal, diurnal, tidal), patterns ***driven by exogenous variables*** (e.g., precipitation), long-term trends, and random variability. ***The focus is on determining the long-term trend component, its direction and rate of change, and estimating the uncertainty of these changes***.

• Methods for trend analysis are based on regression analysis, using ***time*** or ***other variables*** representing drivers of hydrologic change (e.g., urbanization, agricultural practices) as explanatory variables. This approach can better characterize the temporal pattern of the trend when the driver of change has a highly nonlinear time trend.

• Trend analysis considers both sudden and gradual trends, with or without accounting for time-varying natural drivers, and compares the strengths and weaknesses of various common tests.

• Spatial trends, such as variations in chloride concentration with distance from the ocean, are acknowledged but not detailed in this discussion.

1 General Structure of Trend Tests

1.1 Purpose of Trend Testing

• There are several motivations for conducting trend tests. Sometimes we anticipate a change due to known factors, such as landscape alterations or climate shifts, and we want to measure its impact on hydrologic variables.

• For instance, we might ask if the mean annual minimum daily discharge changed after removing an upstream dam or if the mean riverine load of nitrate changed after farmers began planting winter cover crops.

• In other cases, we may not have a specific hypothesis but want to determine if any significant changes have occurred due to various upstream activities. Trend analysis can guide future investigations, update engineering statistics, and categorize datasets from numerous monitoring sites by the trends observed over time.

• Trend tests help categorize sites into groups where the variable of interest has increased, decreased, or shown no clear trend. ***These groupings can be analyzed further to understand the drivers behind the trends***. The tests also estimate the ***magnitude of the trend*** and provide insights into the nature of changes, such as whether changes in water quality are more pronounced during high or low discharges. Consistent methods may be applied across multiple sites for comparability, or tailored approaches may be used for specific datasets.

• There is ***no single correct way to conduct trend tests***, and different methods may offer various advantages. When results are critical, ***using multiple methods*** can help ensure confidence. If different methods yield similar results, confidence in the findings increases. If they diverge, understanding the reasons for the differences is crucial for selecting the best approach.

• Trend analysis involves creating models that account for factors affecting the data's ***central tendency and evaluating significant changes over time***.

• Trend tests involve analyzing a series of observations over time to determine if values generally increase or decrease, and to quantify the rate of change. Results can be displayed geographically to show trends across large areas (**Figure 1**), highlighting patterns that might be missed if only statistically significant trends are shown. Trends can be analyzed using methods like the weighted regressions on time, discharge, and season (WRTDS), considering factors like distribution type, variance, and serial correlation.

• The chapter emphasizes that trend tests should provide more than just a decision to reject or not reject the null hypothesis. They should offer detailed insights, such as changes over specific periods, seasons, or discharge conditions. Emerging views suggest reporting trend analysis results with ***best estimates of trend magnitude and the probability of the trend's sign being incorrect***, helping decision-makers understand the likelihood of different trends.

• In summary, trend tests are used to ***detect and quantify changes in environmental variables over time***, considering various influencing factors and uncertainties. Different methods may be applied, and results should be presented in a way that helps inform decision-making, emphasizing practical significance and the likelihood of different trends.

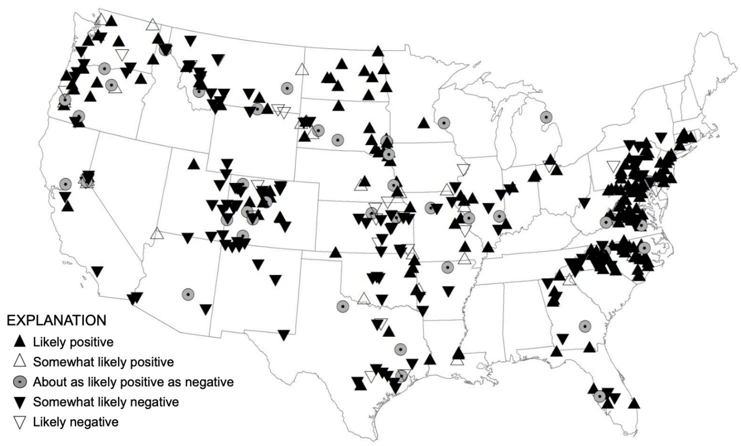


Figure 1. Map showing trend analysis results for specific conductance for the time period 1992–2002 (based on Oelsner and others, 2017). Definitions of the symbol categories are as follows. Likely positive: the likelihood that the trend is actually positive is between 0.85 and 1.0; Somewhat likely positive: the likelihood that the trend is actually positive is between 0.7 and 0.85; About as likely positive as negative: the likelihood that the trend is actually positive is between 0.3 and 0.7 and also that the likelihood that it is actually negative is between 0.3 and 0.7; Somewhat likely negative: the likelihood that the trend is actually negative is between 0.7 and 0.85; Likely negative: the likelihood that the trend is actually negative is between 0.85 and 1.0.

• In most applications of trend analysis, the null hypothesis, H_0, is that there is no trend in the central tendency of the random variable being tested. The precise mathematical definition of H_0 depends on the test that is being applied. The null hypothesis typically includes a set of assumptions related to the distribution of the data (normal versus non-normal), the type of trend (linear, monotonic, or step), and the degree of serial correlation.

• As discussed in week 02, the outcome of the test is a decision—either H_0 is rejected or not. Failing to reject H_0 does not prove there is no trend. Rather, it is a statement that given the available data and the assumptions of the particular test, ***there is not sufficient evidence to conclude that there is a trend***. The possible outcomes of a statistical test in the context of trend analysis are summarized in table 1.

Table 1. Probabilities associated with possible outcomes of a trend test.

Decision	True situation (unknown, in reality)	
	No trend	**Trend exists**
No trend	$1-\alpha$ Probability that H_0 is not rejected	(Type II error) β Probability that H_0 is not rejected
Trend	(Type I error) Significance level α Probability that H_0 is rejected	(Power) $1-\beta$ Probability that H_0 is rejected

1.2 Approaches to Trend Testing

• Five types of trend tests are presented in table 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 parametric and nonparametric***.

• ***The second factor***, shown in the columns, is whether there is some attempt to ***remove*** variation caused by other associated variables. See the headnote to table 2 for the definitions of the types of variables used here. Examples of exogenous variables might be ***precipitation*** amount when the Y variable is streamflow or water level change, or it might be river ***discharge*** when the Y variable is the concentration of some solute.

• For our purposes in this chapter, an exogenous variable is an explanatory variable that is also a random variable.

• Time may also be an explanatory variable in a trend analysis, but it is not a random variable. The reason for using exogenous variables is that they may explain a substantial part of the variance of the response variable (***for example, precipitation explains a great deal of the variance of runoff***) and by accounting for these kinds of relations the trend signal may be much more easily detected, which increases the power and accuracy of the trend analysis method.

• This ***doesn't negate the value of doing the simpler trend test without an exogenous variable***, it simply adds to our ability to discern and potentially explain the trend that is taking place. Simple trend tests (not adjusted for X) are discussed in section 2.

• Tests adjusted for X are discussed in section 3. The approaches shown in sections 2 and 3 ***assume that the data have no seasonal or other regular periodic component***. Section 4 expands on the previous sections by adjusting the methods for the presence of a ***regular periodic component***.

Table 2. Classification of five types of tests for trend.

Type of trend test	Not adjusted for the exogenous variable, X	Adjusted for the exogenous variable, X
Nonparametric	Mann-Kendall trend test of Y	Mann-Kendall trend test on residuals R from loess 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

2 Trend Tests with No Exogenous Variable

2.1 Nonparametric Mann-Kendall Test

• A simple way to construct a trend test is to determine if the central tendency of the variable of interest, Y, changes in a monotonic fashion with the time variable, T.

• Mann (1945) first suggested using the test for significance of the Kendall's τ correlation value as a test for trend, whereby the two variables being related are Y and the time variable, T. The Mann-Kendall test can be stated most generally as whether Y values tend to increase or decrease as T increases (monotonic change).

H_0: Prob $[Y_j > Y_i] = 0.5$, where time $T_j > T_i$
H_1: Prob $[Y_i > Y_i] \neq 0.5$ (two-sided test)

• In the two-sided test, we are only interested in determining, for any pair of observations in the record, if the probability that the later observation is greater than the earlier one is different from a value of 0.5.

• If the ***probability*** were greater than 0.5, that means there is a tendency for Y to increase over time, and if the ***probability*** were less than 0.5, that means there is a tendency for Y to decrease over time.

• No assumption of normality is required, but there must be no serial correlation (i.e., autocorrelation -- the values of a variable are correlated with themselves over successive time periods) in the Y values (after detrending the data) for the resulting α level of the test to be correct.

• Typically, the test is used to determine whether ***the central value or median changes over time***. As discussed in previous weeks, the results of the test do not change if the Y data are transformed by any monotonic transformation (such as logarithm or square root transformation).

• The test is performed by computing the *Kendall's S statistic* from the *Y* and *T* data pairs.

• H_0 is rejected if the value of *S* is statistically significantly different from zero. For large samples, the distribution of *S* given H_0 is approximately normal, with variance determined only by the sample size.

• There is also an estimate of the slope of such a temporal trend that is closely related to the *S* statistic and the hypothesis test. This is the *Theil-Sen estimator* (also known as the Kendall-Theil robust line).

• If *S* is positive, the Theil-Sen slope will be positive, and if *S* is negative, the Theil-Sen slope will be negative. The units of the slope estimate are the units of the *Y* variable divided by the units of time.

• For example, if the trend test is concerned with annual mean discharge, then the *Y* units might be cubic meters per second (m^3/s) and the *T* units would be year (yr), so the slope units would be $m^3/s/yr$. The example in figure 2 illustrates a dataset with a fairly strong long-term trend (the annual mean discharge of the Mississippi River at Keokuk, Iowa, 1931–2013). Let's say the selected α level for the test was 0.05.

• The two-sided *p*-value for the Mann-Kendall test is 0.000006 with $\tau = 0.34$. Thus, we can reject H_0 and conclude that there is a trend and that the trend is positive. The units of discharge are m^3/s and the units of time are years.

• The Theil-Sen robust line slope is 13.0 $m^3/s/yr$. Using the median of the discharge values and the median of the time values, the equation for the Theil-Sen robust line is

$$\hat{Q} = \beta_0 + \beta_1 \cdot T \qquad \textbf{Eq. (1)}$$

where
$\hat{Q}$ = estimate of discharge in $\frac{m^3}{s}$;
T = time in years;
$\beta_1 =$ Theil − Sen slope
$\beta_o =$ median $(Q) - \beta_1 \cdot$ median (T).

The Theil-Sen slope (β_1) is ***the median of the slopes of all pairs*** of points.
In this particular example, $\beta_o = -23{,}659$ and $\beta_1 = 13.0$. Theil-Sen line shown in figure 2 is:

$$\hat{Q} = -23{,}659 + 13.0 \cdot T$$

• One could consider doing this analysis on log-transformed discharge data. The Mann-Kendall hypothesis test for trend will have exactly the same result (*p*-value is still 0.000006) as it did without the log transformation; however, the slope will be different because of the transformation. If we use the natural logarithm for our transformation (as opposed to base 10), the slope is 0.006696 per year.

• The *p*-value will remain unchanged no matter what monotonic transformation is used (such as any one of the ladder of powers transformations). If the change is assumed to be linear with time, in this case linear in ln(Q), then that means that the trend in the original discharge units will be exponential and the ratio of the expected value from one year to the expected value for the previous year will be exp(0.006696), which is 1.006718. This result can be interpreted as an increase of 0.6718 percent from one year to the next.

• The graphical representation of the Mann-Kendall test for trend and Theil-Sen robust line on the natural log of discharge is shown in **Figure 3**. The original data are shown in **Figure 4**, but the solid line is the Theil-Sen estimate transformed back into the original units (by exponentiating the estimates from the log space model). In this case, there is relatively little difference between the two estimates (compare **Figs**. **2** and **4**), but in some cases there can be a substantial difference.

• There is ***no definitive test that will indicate which is a better estimate***; it is really a matter of judgment based on apparent curvature. In this case, either approach could be acceptable.

• One final comment on the Theil-Sen robust line: ***it should not be considered an estimate of the conditional mean of the distribution for any given year***. It is more suitable to think of it as an estimate of the conditional ***median***.

• For example, in **Figure 3**, we may say that for any given year the probability of the true value being above the Theil-Sen line is 0.5 and being below is 0.5. Because the Mann-Kendall test for trend provides the same results no matter what transformation is used on the *Y* values, it can be suitable for studies of a large number of similar datasets (for example, 50-year records of mean discharge) across a study area.

• Each dataset may suggest the need for a different ladder-of-powers transformation to make the relation ***more nearly linear***, but because the Mann-Kendall test results ***do not vary across all the possible ladder-of-powers transformations***, this step of transformations becomes unnecessary for applying the test to multiple sites. ***This is one reason that a nonparametric test such as the Mann-Kendall is well suited to a study that encompasses tests of many datasets***.

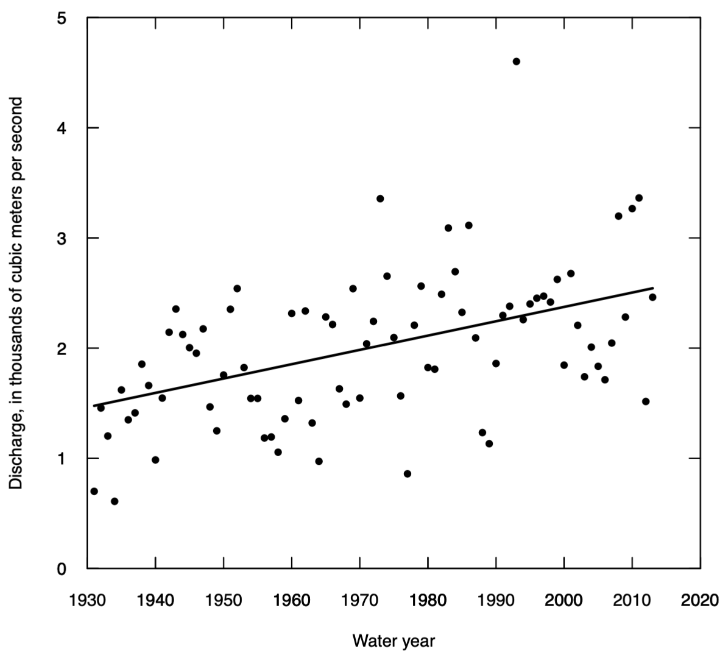


Figure 2. Plot of annual mean discharge, Mississippi River at Keokuk, Iowa, 1931–2013, shown with the Theil-Sen robust line.

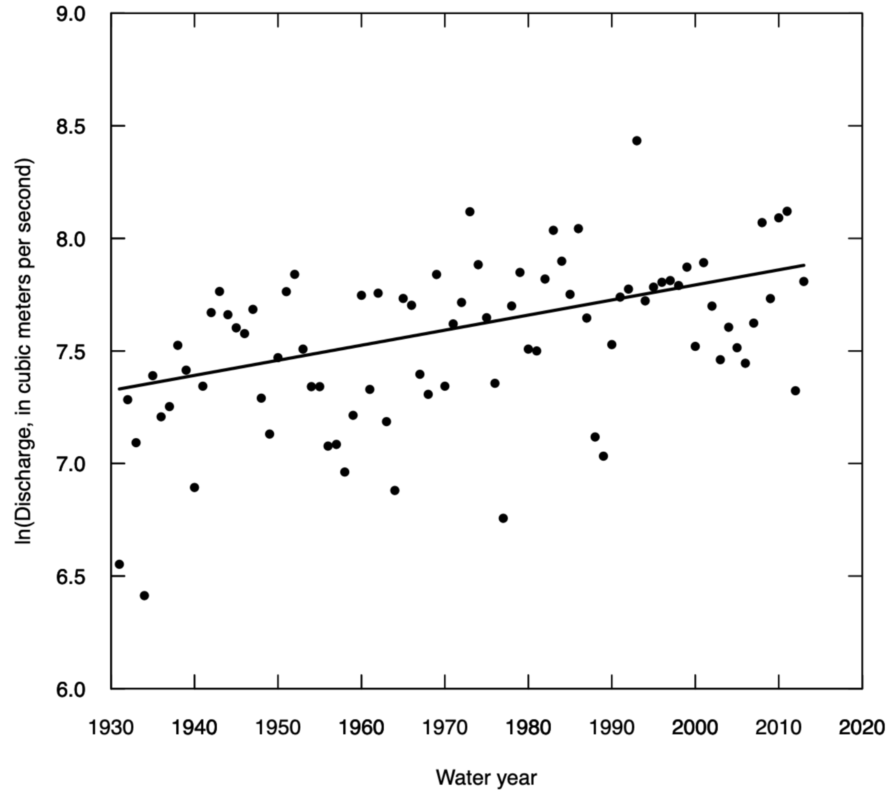


Figure 3. Plot of the natural log of annual mean discharge, Mississippi River at Keokuk, Iowa, 1931–2013, shown with the Theil-Sen robust line.

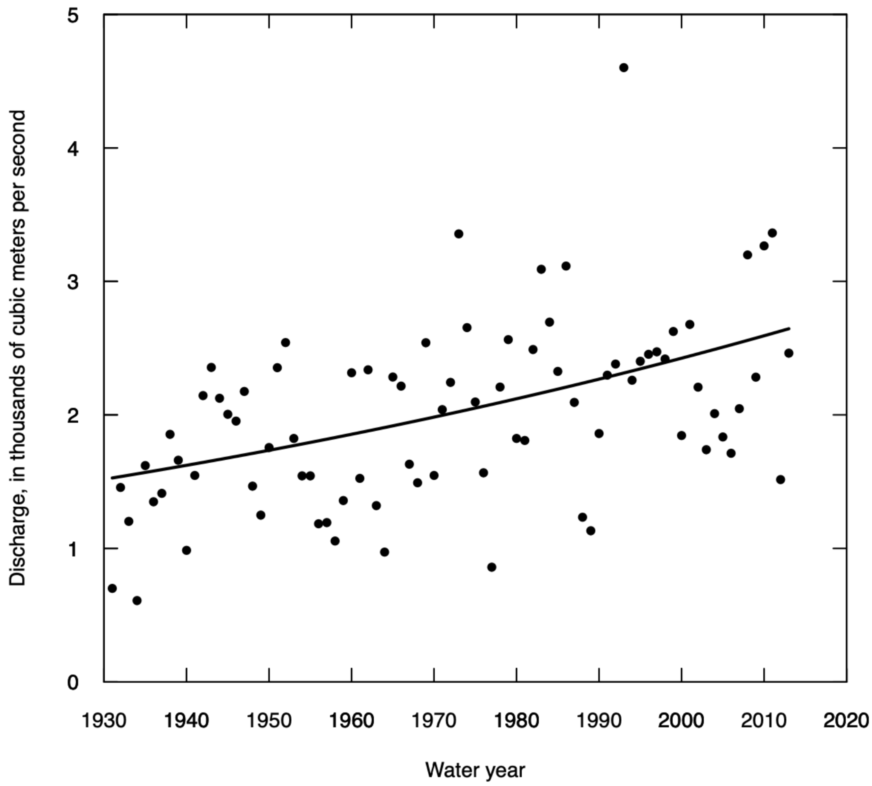


Figure 4. Plot of the annual mean discharge, Mississippi River at Keokuk, Iowa, 1931–2013, shown with the transformed Theil-Sen robust line based on slope of the natural log discharge values.

2.2 Ordinary Least Squares Regression of *Y* on Time, *T*

Ordinary Least Squares Regression of *Y* on Time, *T*

$$\hat{Q} = \beta_0 + \beta_1 \cdot T \quad \textbf{Eq. (2)}$$

• The null hypothesis is that the slope coefficient, β_1, is zero.

• OLS regression (eq. 2) makes stronger assumptions about the behavior of *Y* over time than does Mann-Kendall. The relation must be checked for ***normality of residuals***, ***constant variance***, and linearity of the relation (best done with residuals plots—see chap. 9). If Y is not linear with time, a transformation will likely be necessary. If all of the conditions of OLS regression are met, then the slope estimate is b1, and the t-statistic on b1 can be used to determine if the slope is significantly different from zero. Further, this test for trend has slightly more power than the nonparametric Mann-Kendall test if the conditions of OLS regression are met. But, with modest departures from normality of residuals, the Mann-Kendall test can be a good deal more powerful than regression. If the t-statistic is large (in absolute value), typically |t| >2, we can reject the null hypothesis and conclude that there is a time trend (upwards if the estimate is positive and downward if it is negative). Unlike Mann-Kendall, the test results for regression (specifically the p-value on the slope coefficient) will not be the same before and after a transformation of Y.

Note: Differences Theil-Sen estimator from OLS

The Theil-Sen estimator differs from OLS in the following ways:

1. **Robustness**: The Theil-Sen estimator is more robust to outliers than OLS. It uses the ***median*** of slopes between all pairs of points, which reduces the influence of outliers.
2. **Assumptions**: OLS assumes that the residuals (errors) are normally distributed, have constant variance, and the relationship between variables is linear. The Theil-Sen method does not make these assumptions and is nonparametric.
3. **Calculation of Slope and Intercept**: OLS calculates the slope and intercept by minimizing the sum of squared residuals. In contrast, the Theil-Sen method uses medians:
    - OLS slope (β_1) is found by minimizing the sum of squared differences between observed and predicted values.
    - Theil-Sen slope (β_1) is the median of the slopes between all pairs of data points.
    - OLS intercept (β_0) is calculated based on the mean of the data.
    - Theil-Sen intercept (β_0) is calculated using the medians of the data.
4. **Sensitivity to Data**: The OLS method is sensitive to outliers, which can significantly affect the slope and intercept. The Theil-Sen method is less sensitive to outliers due to its reliance on medians.

By using the Theil-Sen method, you obtain a robust linear fit that is less affected by anomalies in the data, providing a more reliable trend estimation in the presence of outliers or non-normal data distributions.

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 OLS regression would be optimal (most powerful and lowest error variance for the slope).

• Of course, usually we cannot know this in any actual 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 chap. 10, and Hirsch and others [1991], p. 805–806).

• There are practical cases where the ***OLS regression approach is preferable***, particularly in the multiple regression context. A good deal of care needs to be taken to ensure the regression is correctly applied and enough information is provided such that the audience is able to verify that the assumptions have been met.

• 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***. Nonparametric methods are almost always nearly as powerful as OLS regression, and failure to edit out a small percentage of bad data or correctly transform the data will not have a substantial effect on the results.

Example 1:

Milwaukee River chloride trends. Chloride concentrations sampled in the month of March in the Milwaukee River at Milwaukee, Wisconsin for the years 1978–2005 are shown in figure 12.5. Two trend tests were conducted, the Mann-Kendall test and OLS regression on time. The Theil-Sen and OLS regression lines are plotted along with the data. Using $\alpha = 0.05$, the OLS regression line is not significantly different from a slope of zero, and thus we would not reject the null hypothesis of no trend (p-value is 0.091), but the Mann-Kendall test statistic (S) is significantly different from zero (p-value is 0.017). It is interesting to note that the linear regression line is a good deal steeper than the Theil-Sen line even though it is not significant, but the Mann-Kendall test is significant. The linear regression line is heavily influenced by the high outlier value in 1999. This high value and the highly skewed distribution of the residuals are the reasons that the linear regression approach fails to provide confirmation of a trend, whereas the Theil-Sen line is unaffected by the magnitude of this high value. Later in this chapter, we will return to this dataset and consider ways that the parametric approach could be improved upon to result in a more sensitive and meaningful description of the trend, which the Mann-Kendall test strongly suggests is present.

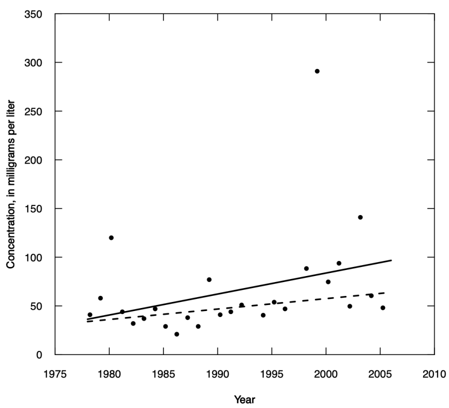


Figure 5. Graph of chloride concentration in the month of March for the Milwaukee River, at Milwaukee, Wisconsin. Solid line is the estimated trend using linear regression. The dashed line is the estimated trend using the Theil-Sen robust line. Note how the high value in 1999 has a strong influence on the linear regression but not the Theil-Sen robust line.

3 Accounting for Exogenous Variables

• Variables ***other than time*** 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 the trend signal that may exist can be seen, and the power (ability) of a trend test to discern changes in Y with T is increased***.

• Two important types of situations are relevant here. The first is the exploration of trends in some measure of water quantity. In these cases, the *Y* variable may be something like annual mean discharge, annual minimum discharge, annual maximum discharge, or some measure of change in storage in a lake or aquifer.

• ***The obvious candidate*** for an exogenous (*X*) variable would be some measure of precipitation at an appropriate spatial and temporal scale (for example, annual total precipitation averaged over several precipitation gages that span the watershed of interest).

• ***The second type of situation*** is the exploration of trends in some measure of water quality such as concentrations of solutes or sediment, a biological measure such as chlorophyll, or water temperature. In these cases, ***the obvious candidate*** for an exogenous (*X*) variable might be ***discharge*** at the time of water quality measurement or perhaps discharge during some period before the time of measurement.

• The discussion presented below applies to the latter case, but the principles discussed here could apply to either one. The process of ***removing the variation resulting from X*** involves modeling the effects of exogenous variables with OLS regression or loess (for computation of loess, see previous week's *pdf*).

• Using the same dataset used in the example above, we will consider all three of the options mentioned in the "Adjusted for the exogenous variable, X" column in **Table 2**.

• Consider a regression of *Y* versus *X*. The residuals (*R*) from this regression express the variation in ***Y* not explained by *X***. A trend in *R* implies a trend in the relation between *X* and *Y*.

• This in turn implies a trend in the distribution of *Y*, ***but this conclusion may not hold if there is a trend in X***. What kind of variable is appropriate to select as an exogenous (*X*) variable?

• Here, we use the term exogenous to indicate that it is a particular kind of explanatory variable that is ***itself a random variable which is externally driven*** rather than being driven by some human activity that may also be driving variations in *Y*, the variable for which we are doing the trend test.

• This exogenous variable should be a measure of a ***driving force behind the process of interest***, but it must be relatively ***free of changes owing to human manipulation***.

• For a water-quality trend study, the streamflow record at (or near) the site where the water quality data were collected is an obvious choice for an exogenous (*X*) variable.

• However, if the streamflow record ***being used*** includes a time span that covers a period both ***prior to and after*** major upstream water management changes, then the streamflow record would be unacceptable as a ***choice of an exogenous random variable*** because the probability distribution of X has likely changed substantially during the period of interest.

• Examples of such changes include the completion of a major dam, removal of a major dam, initiation of a major new diversion in or out of the watershed, or a major change in operating policy of a water resource system.

• A streamflow record which reflects some subtle human influence is acceptable, provided that the effect 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 streamgage 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.

• Of course, as landscape manipulations (such as artificial drainage), regional groundwater depletion, or global greenhouse gas concentrations increase over time, it becomes impossible to say that any hydrologic or climatic variable is free of human manipulation. ***Decisions to use climate records or streamflow records as exogenous variables in a trend analysis involve a trade-off***.

• The use of exogenous variables is very helpful in reducing the unexplained variance in the variable of interest (Y) and thereby increasing our ability to discern and describe the trend. ***However, as time periods get longer and climate, landscape, or groundwater storage changes get stronger, the use of such exogenous variables becomes problematic***.

• To do a trend study of a random variable, Y, we need to be confident that the observed trend in Y is a function of a shift in the X-Y relation and not simply a function of a trend in X.

• Resolution of this trade-off will be a **challenge** to water-related trend studies for the foreseeable future. Where Y is a concentration (of a solute or particulate matter), 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 processes. **One process** is dilution: a solute may be delivered to the stream at a reasonably constant rate (for example, effluents from a point source or groundwater discharge to the stream) as discharge changes over time. The result of this situation is a decrease in concentration with increasing flow; 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***. 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.

• The resulting pattern is an initial dilution at the low end of the discharge range, followed by an increase with discharge at higher values of discharge.

• The Milwaukee River chloride record exhibits this kind of nonmonotonic behavior in the *X*-*Y* relation, as illustrated in **Figure 6**. Subsections 3.1., 3.2., and 3.3. consider three types of approaches to trend testing using an exogenous variable. All three are applied to the dataset presented in the previous example (Milwaukee River chloride data). We will consider a dataset that consists of a single chloride concentration value for the month of March from each year, and the associated daily discharge value on that sampling date for the Milwaukee River for the 26 years from 1978 through 2005.

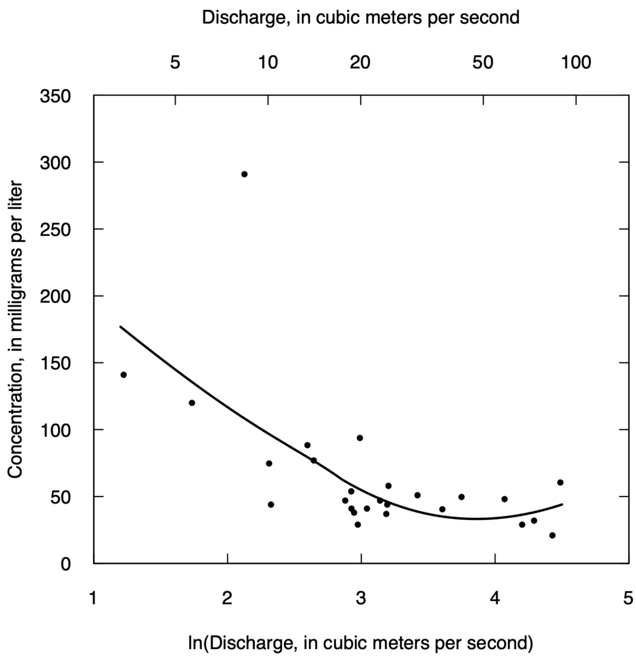


Figure 6. Graph of the relation between chloride concentration and the natural log of discharge, Milwaukee River at Milwaukee, Wisconsin, for samples collected in March 1978–2005. Solid line is the loess smooth. Residuals are the vertical differences between the data points and the line.

3.1 Mann-Kendall Trend Test on Residuals, *R*, from Loess of *Y* on *X*

• The first of the three approaches to trend testing is ***to remove the influence of discharge*** on the chloride data and then do a Mann-Kendall test for trend in the residuals.

• The relation between chloride concentration and discharge is clearly evident in **Figure 6**. In this case, the exogenous variable is the natural log of discharge, the smoothing method is the loess function, and the span value is 1.0 (the default value for span is 0.75). See previous week's *pdf* for a discussion of loess and setting the span.

• The resulting scatterplot with the superimposed loess curve indicates that these are reasonable choices (the curve follows the bulk of the data and doesn't have any jagged oscillations).

• The decision to transform the explanatory variable and the choice of the span value for the loess are judgement calls ***on the part of the analyst***.

• One could also try an OLS regression fit for the purpose of computing residuals, but an examination of **Figure 6** provides ample evidence that ***OLS would be problematic because of the curvature in the relation at higher discharges***. With OLS, the residuals would not be independent of the explanatory variable.

• Based on this loess fit, residuals are computed and they can be plotted against time (**Figure. 7**). One thing that is clear at a glance is that the distribution of the residuals contains one very extreme outlier. Based on this fact, the use of a nonparametric approach to trend testing is probably a good approach. For the null hypothesis that the residuals are trend-free, we apply the Mann-Kendall test using a two-sided $\alpha = 0.1$ and the result is that we should reject the null hypothesis (the attained p-value is 0.052).

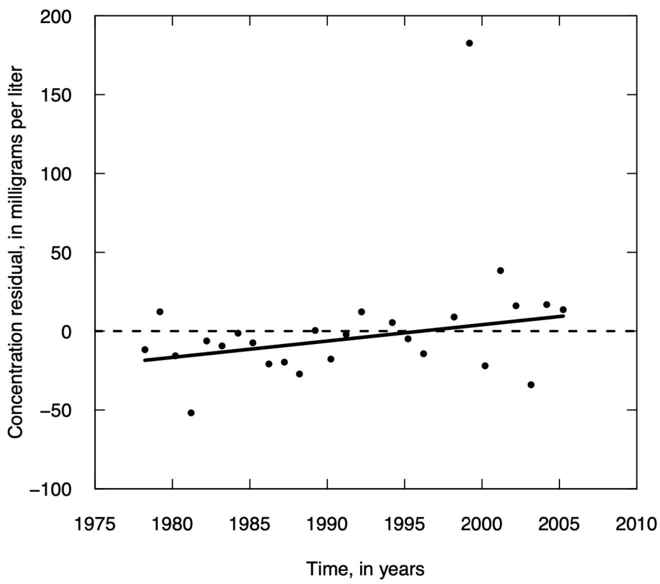


Figure 7. Graph of concentration residuals (from a loess fit of concentration as a function of the log of discharge) versus time for chloride concentrations in the Milwaukee River at Milwaukee, Wisconsin, for samples collected in March from 1978 through 2005. Solid line shows the Theil-Sen line, with a slope of 1.03 milligrams per liter per year. Dashed line is residual = 0 for all years.

3.2 Mann-Kendall Trend Test on Residuals, *R*, from Loess of *Y* on *X*

• Consider a mixed parametric and nonparametric approach to test for trends. In this approach, we use OLS regression of Y on X to obtain the residuals, which we then test for trend using a Mann-Kendall trend test.

• In this particular case, exploration of various forms of the regression model of concentration as a function of discharge suggests that an appropriate model would take the form $ln(C) = \beta 0 + \beta 1 \cdot ln(Q) + R$, where C is concentration, Q is discharge, and R is the residual.

Note: OLS Regression

1) The exact equation to calculate the residuals for Ordinary Least Squares (OLS) regression is:

$R_i = Y_i - \hat{Y}_i$

2) The OLS regression equation is:

$\hat{Y}_i = \beta_0 + \beta_1 X_i$

3) the residual for the ii-th observation is:

$R_i = Y_i - (\beta_0 + \beta_1 X_i)$

4) Calculate the means of *X* and *Y*:

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$

5) Calculate the slope β_1:

$\beta_1 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$

6) Calculate the intercept β_0:

$\beta_0 = \bar{Y} - \beta_1 \bar{X}$

• Using this type of model, with the dependent variable being a log transformation of the variable of interest, means that the residuals are no longer in the original units (mg/L) but are now residuals expressed in log space.

• We can then proceed to do a test for trend on these residuals using the Mann-Kendall test. The results are shown in **Figure 8**; the trend slope is ***0.0232/yr***, and the *p*-value for the test is 0.0082, which is a stronger indicator of trend than was attained from the previous test where the trend was computed on loess residuals.

• With this particular dataset, the use of linear regression to ***compute the residuals is problematic*** because the small size of the dataset makes it difficult to verify the soundness of the approach. ***In this case, the loess approach to computing the residuals should probably be given more credence***. The fact that in this case the trend is expressed in log units will be discussed later in next week's *pdf* file. Alley (1988) showed that this type of two-stage procedure resulted in lower power than an alternative, which is analogous to the partial plots of previous week's lectur. His "adjusted variable Kendall test" performs the second stage as a Mann-Kendall test of *R* versus e^* rather than *R* versus *T*, where e^* are the residuals from an OLS regression of *T* versus *X*.

$$T = b_0 + b_1 \cdot X + e^* \quad \textbf{Eq. (3)}$$

• ***In this way the effect of a drift in X over time is removed, so that the R versus e* relation is totally free of the influence of X***. This test is a Mann-Kendall test on the partial residuals of *Y* versus *T*, having removed the effect of all other exogenous variable(s) *X* from both *Y* and *T* by regression. For more discussion of the partial Mann-Kendall test see Libiseller and Grimvall (2002).

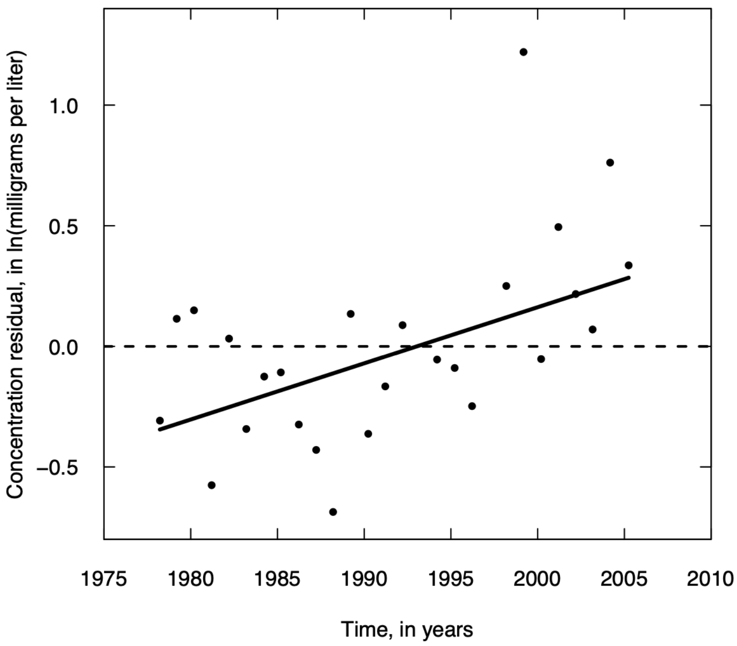


Figure 8. Graph of log concentration residuals versus time for chloride concentrations in the Milwaukee River at Milwaukee, Wisconsin, for samples collected in March, from 1978 through 2005. Solid line shows the Theil-Sen line, with a slope of 0.023 log units/year. Dashed line is residual = 0 for all years.

3.3 Regression of *Y* on *X* and *T*

• The regression of *Y* on *T* is an entirely parametric approach to evaluating trends in *Y* adjusted for some exogenous random variable, *X*.

• This approach uses multiple linear regression to do in a single step what the preceding two approaches did sequentially. If we knew that the dataset had the right sort of characteristics for multiple regression, ***then this would be the most powerful and least biased approach to the problem***. The characteristics that would be most important are (1) that the relations be linear, and (2) that the errors be normally distributed and homoscedastic.

• Of course, we ***can't determine these things for certain***, but effective use of model building and checking can help to identify if the data at least approximate these characteristics. The building of a multiple regression model for this data needs to be done with care considering the many issues discussed in previous chapters on regression, particularly those related to issues of transformation of explanatory variables and dependent variables and concerns about the behavior of the residuals. The appropriate regression model for this dataset is

$$\ln(C) = \beta_o + \beta_1 \cdot T + \beta_2 \cdot \ln(Q) + \varepsilon \quad \textbf{Eq. (4)}$$

where C is concentration, T is time in years, and Q is discharge in m³/s.

• The R^2 for the fitted model is 0.63, and the overall F-test for the model shows it to be highly significant ($p < 0.0001$). We can summarize the fitted coefficients, their t-statistics, and their p-values in Table 3.

Table 3. Model coefficients and their t-statistics and p-values.

Regression results	b_0	b_1	b_2
Coefficient value	−48.27	0.0270	−0.465
t-statistic	−2.90	3.23	−5.31
p-value	0.008	0.003	<0.001

• Viewing this as a trend test, we can focus directly on the β_1 coefficient. We can see that the coefficient value is very nearly the same as the one we found in the previous test (***Mann-Kendall on residuals from the ln(C) versus ln(Q) linear regression***).

• We also see that it is statistically significant ($p = 0.003$). This is exactly what we would hope to be the case, ***that this simultaneous approach gives similar results to those that come from the sequential parametric approach***.

• Our conclusion for this example then is, for any particular value of discharge, the expected value of the natural log of concentration is increasing over time and is doing so at a rate of 0.027 per year.

• This is equivalent to a 2.74 percent per year increase (exp(0.027) = 1.0274). The statement is accurate but not particularly easy to picture or understand. However, Figure 9 provides a simple way to illustrate the fitted model.

• The data are categorized into three groups (low, medium, and high discharges) using different symbols for each group. Then the fitted regression is evaluated at three example discharges (15, 30, and 60 m³/s). The trends are linear in the natural log of concentration, but the graph is presented in concentration units so these linear trends in the logs become exponential trends in concentration.

• ***The curves represent a median estimate of concentration for each of the three example discharges*** because the model is built on the assumption that the errors around the regression line are normal with zero mean and constant variance. When the regression lines are transformed to concentration units, they still represent the conditional median of concentration for the given discharge and year, but they do not represent the conditional mean (see the discussion in previous week's *pdf* regarding retransformation of regression results).

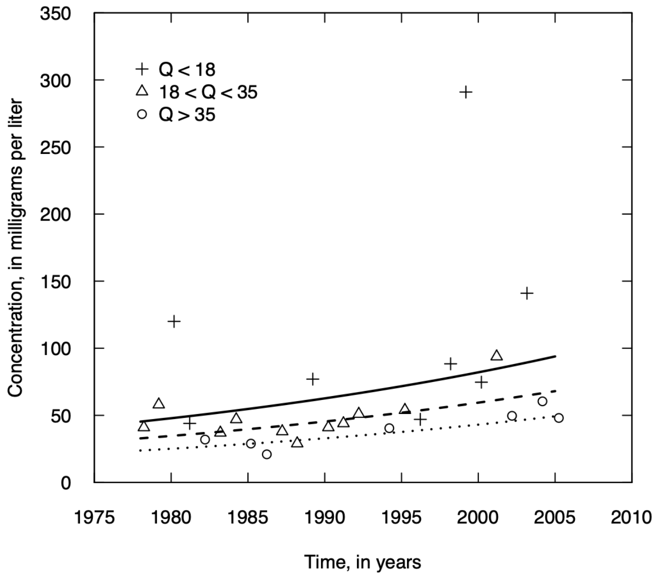


Figure 9. Graph of curves that represent median estimates of chloride concentration as a function of time, from the Milwaukee River at Milwaukee, Wisconsin. Solid line is at discharge (Q) = 15 cubic meters per second (m^3/s), dashed line is at Q = 30 m^3/s, and dotted line is at Q = 60 m^3/s. Symbols represent data points based on discharge in m^3/s.