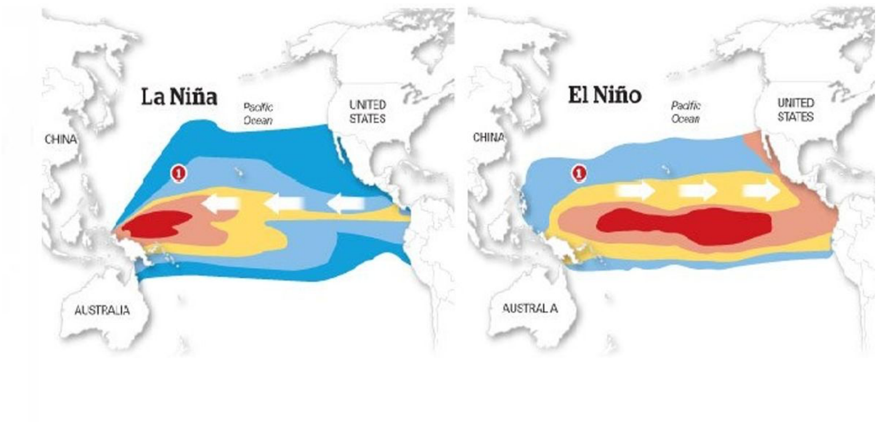
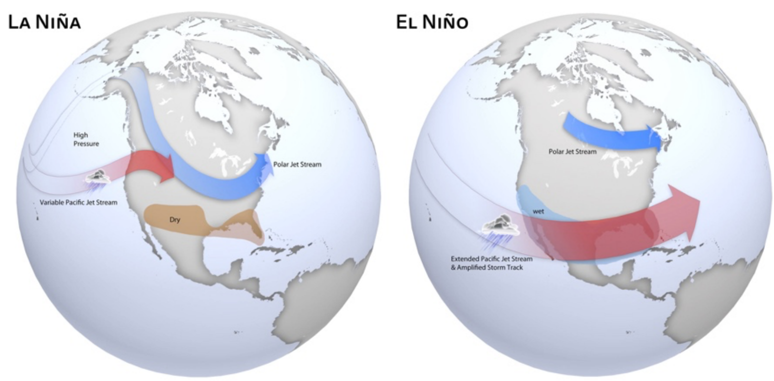
Trend Analysis
EN5423 | Spring 2024

w16_ trend_analysis_02.pdf
(Week 16)

Contents





1 Dealing with Seasonality or Multiple Site

• There are many instances where changes ***between different seasons of the year*** are a major source of variation in the response (*Y*) variable. As with other exogenous effects, seasonal variation must be compensated for or removed in order to discern the underlying trend in *Y* over time.

• If not, there may be only little power to detect trends which are truly present. A statistical test like the Mann-Kendall or OLS regression ***does not detect that there are seasonal patterns***; rather the pattern registers ***as random noise in the process***. This is because those tests are designed ***to detect monotonic or linear changes***. Because trend tests (parametric or nonparametric) are fundamentally about being able to see a trend signal ***stand out above the noise***, this seasonality will hinder our ability to truly observe the trend.

• This issue can arise in many contexts. Concentrations of a water quality constituent in surface waters typically ***show strong seasonal patterns***. Water temperatures also have a strong seasonal component and streamflow itself almost always varies across seasons. These seasonal patterns arise from seasonal variations in ***precipitation volume*** and ***air temperature***.

• Air temperature affects the precipitation type (rain versus snow) and the rate of evapotranspiration. Some of the observed seasonal variation in water quality concentration may be explained by accounting for this seasonal variation in discharge. However, seasonality often remains ***even after discharge effects have been removed*** (Hirsch and others, 1982).

• Possible additional causes of seasonal patterns include biological processes and managed activities such as ***agriculture***. For example, nutrient concentrations commonly vary with the ***seasonal application*** of fertilizers and the ***natural pattern of uptake and release by plants***.

• Other effects are the result of different sources of water contributing to the water feature (stream, lake, or pond) at different times of the year, such as snow melt or intense rainfall. The seasonal rise and fall of ground water can also influence ***water quality***.

• For example, a given discharge magnitude in one season may derive mostly from groundwater whereas the same discharge magnitude during 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.***

• Thinking in terms of trends in ***water quantity*** (streamflow or groundwater level change) it is generally the case that the ***response of the hydrologic variable*** to a given amount of precipitation can differ considerably ***across seasons***.

• This can be true because of the changes in interception by plant canopy, changes in runoff characteristics between bare fields and fields with active vegetation, and also differences in the amount of evapotranspiration across the seasons.

• In general, a rainfall-runoff relation is likely to be different at different times of the year, and ***failure to consider it will greatly diminish the power of any trend analysis***.

• Techniques for dealing with seasonality fall into three major categories, as shown in table 1; one is fully nonparametric, one is a mixed procedure, and the last is fully parametric. For the upper four cells of table 1 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 but no shorter than monthly blocks. 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 four seasons, and so forth.

Table 1. General categories of options for dealing with seasonality in conducting trend tests. X is an exogenous variable such as precipitation or streamflow that may influence the variable of interest, in addition to the influence of season.

Type of trend test	Not adjusted for *X*	Adjusted for *X*
Nonparametric	Seasonal Kendall test for trend on *Y*	Seasonal Kendall trend test on residuals from loess of *Y* on *X*
Mixed	OLS regression of deseasonalized *Y* on *T*	Seasonal Kendall trend test on residuals from OLS regression of *Y* on *X*
Parametric	Multiple regression of *Y* on *T* and seasonal terms	Multiple regression of *Y* on *X*, *T*, and seasonal terms

1.1 The Seasonal Kendall

• The Seasonal Kendall test (Hirsch and others, 1982) accounts for ***seasonality by computing the Mann-Kendall test on each of m seasons separately***, and then combining the results.

• In this test, a season can be monthly, quarterly, or some other definition of time. The advantage of this approach over an ordinary Mann-Kendall approach is that in the Seasonal Kendall test ***no comparisons are made across season boundaries***.

• For example, for monthly seasons, January data are compared only with January, February only with February, and so on. The idea behind the test is that making comparisons across the seasons (say comparing the January 2012 value to the June 2008 value) is really ***not informative*** about trend and is more likely to be an expression of the differences between seasons rather than differences between years.

• The information we seek is found by only comparing data from the same season over different years. To perform the test, Kendall's S statistics, denoted S_i, are ***computed for each season*** and these are then summed to form the overall statistic S_k (eq. 1).

$$S_k = \sum_{i=1}^{m} S_i \qquad \text{Eq. (1)}$$

When the product of the 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 (eq. 2):

$$\sigma_{S_k}^2 = \sum_{i=1}^{m} \sigma_{S_i}^2 = \sum_{i=1}^{m} \frac{n_i \cdot (n_i - 1) \cdot (2n_i + 5)}{18} \qquad \text{Eq. (2)}$$

where n_i = number of years of data in season *i*. Note that the formula for the variance of S_i is exactly the same as the formula for the ordinary *S* statistic in the original Mann-Kendall test for trend.

• S_k is standardized (eq. 3) by subtracting its expectation and adding in a continuity correction and dividing by its standard deviation, σ_{S_k}, all under the assumption that ***there is no correlation among the seasons, and no ties and no cases of multiple values in a given season***.

• Also note that the value of $\sigma^2_{S_k}$ depends only on the number of seasons and the number of years of observations in each season, and not on the data values themselves.

• The resulting standardized value, Z_{S_k}, of the Seasonal Kendall test statistic S_k, is evaluated against a table of the standard normal distribution.

$$Z_{S_k} = \begin{cases} \dfrac{S_k - 1}{\sigma_{S_k}} & \text{if } S_k > 0 \\ 0 & \text{if } S_k = 0 \\ \dfrac{S_k + 1}{\sigma_{S_k}} & \text{if } S_k < 0 \end{cases} \qquad \text{Eq. (3)}$$

• The null hypothesis (no trend) is rejected at significance level α if $Z_{S_k} > Z_{crit}$, where Z_{crit} is the value of the standard normal distribution with a probability of exceedance of $\alpha/2$.

• When some of the *Y* or *T* values are tied, the formula for σ_{S_k} must be modified as discussed in chapter 8 in USGS_SMWR.*pdf*. For datasets of about 10 years or longer, the variance (σ_{S_k}) can be modified to account for serial correlation (Hirsch and Slack, 1984).

• The Seasonal Kendall Test, with and without this adjustment for serial correlation, is implemented in Pytnon libs.

• If there is variation in ***sampling frequency*** during the years of interest (meaning the sample sizes differ across the seasons), the dataset used in the trend test may need to be modified.

• If the variations 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 there happen to be no values in a particular season of a particular year, then there would be no value used for that season of that year. If, however, there is a systematic variation in the sampling frequency (for example, monthly for 7 years followed by quarterly for 5 years) then a different approach is necessary. When there is a systematic variation in sampling frequency, define the seasons on the basis of the lowest sampling frequency.

• For the 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. If the sampling frequency is reasonably consistent and there are generally two or more samples taken in each season, then the median of the available data should be used for each season of each year.

• The trend slope can be computed in a manner that is compatible with the approach used in the test. ***It is based on the Theil-Sen slope estimator, but only uses the pairwise comparisons within a given season***. The slope estimate is the median of all of these within-season slopes.

• When there is an important exogenous variable (*X*, as discussed in section previous week's *pdf*), the Seasonal Kendall test can also be used. One option is to perform the Seasonal Kendall test on residuals from a loess model of *Y* as a function of *X*.

• Another option is to perform the Seasonal Kendall test on residuals from an OLS regression of *Y* on *X*. The choice between the two should be based on how well the underlying assumptions of linear regression are met. If they are met reasonably well, this mixed approach of the Seasonal Kendall Test on residuals from linear regression would be an appropriate method.

1.2 Mixed Method—OLS Regression on Deseasonalized Data

• Another possible approach is to ***deseasonalize the data by subtracting seasonal medians from all data within the season, and then doing OLS regression of 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 alternatives.

• When seasonal medians are subtracted, this is equivalent to using dummy variables for $m-1$ seasons in a fully parametric regression. This approach causes the loss of $m-1$ degrees of freedom in computing the seasonal statistics, a disadvantage which can be avoided by using the fully parametric method introduced in the next section.

• However, one drawback to the fully parametric approach is that it makes ***a fairly restrictive assumption*** about the shape of the seasonal pattern (a sine wave). In some cases, hydrologic data may strongly depart from that characteristic, for example having abrupt changes between the growing season and the nongrowing season, or having distinct regular shifts from dry season to wet season. Where these more abrupt changes exist, the mixed approach may have merit.

1.3 Fully Parametric Model—Multiple Regression with Periodic Functions

• The third option for analysis of trends in seasonal data is ***to use periodic functions*** to ***describe seasonal variation***. The simplest approach, and one that is sufficient for many purposes, takes the form of equation 4.

$$Y = \beta_0 + \beta_1 \cdot \sin(2\pi T) + \beta_2 \cdot \cos(2\pi T) + \beta_3 \cdot T + \beta_4 \cdot X + \varepsilon \qquad \text{Eq. (4)}$$

where T is time in years and X is an exogenous explanatory variable such as discharge, precipitation, or level of some human activity (for example, waste discharge, basin population, or production).

• The *X* variable may be continuous or binary dummy variables as in analysis of covariance (for example, before or after the dam was removed, or before or after the treatment plant upgrade).

• 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 standard assumptions for multiple linear regression described in chapter 11 of USGS_SMWR.pdf). The residuals, ε, must be approximately normal.

• 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, *M*, at which the peak occurs. The sum of the sine and cosine terms can be re-expressed this way.

$$\beta_1 \cdot \sin(2\pi t) + \beta_2 \cdot \cos(2\pi t) = A \cdot \sin[2\pi(t + t_0)] \qquad \text{Eq. (5)}$$

Where $A = \sqrt{\beta_1^2 + \beta_2^2}$, and t_0 is the phase shift in years (the point in time when the sine wave crosses zero and has a positive slope is at time of $-t_0$ years).

• Let *M* denote the day of the year when the function reaches its maximum. It can be determined as follows

$$\text{if } \beta_1 > 0 \quad M = \frac{365.25 \cdot \left[\frac{\pi}{2} - \tan^{-1}(\beta_2/\beta_1)\right]}{2\pi} \qquad \text{Eq. (6)}$$

$$\text{if } \beta_1 < 0 \quad M = \frac{365.25 \cdot \left[-\frac{\pi}{2} - \tan^{-1}(\beta_2/\beta_1)\right]}{2\pi} + 365.25 \qquad \text{Eq. (7)}$$

There are three special cases,

$$\text{if } \beta_1 = 0 \text{ and } \beta_2 > 0 \text{ then } M = 365.25 \text{ or } 0$$
$$\text{if } \beta_1 = 0 \text{ and } \beta_2 < 0 \text{ then } M = 182.625$$
$$\text{if } \beta_1 = 0 \text{ and } \beta_2 = 0 \text{ then } M = \text{ undefined}$$

(when *M* is undefined that means there is no annual sine wave).

• Python lib, call MaxDay (from R) determines, given a pair of β_1 and β_2 values, the day of the year at which this sine wave attains its maximum value.

• Note that the same sort of trigonometric formulation can be used for other possible periodicities that might arise in water resources. The most common of these being a ***diurnal cycle*** (24 hours) and also possibly a ***weekly cycle*** (where the variable of interest is affected by human activities focused around the 7-day week).

• After including sine and cosine terms in a multiple regression to account for seasonality, the ***residuals may still show a seasonal pattern***. If this occurs, additional periodic functions with periods of half a year or some 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 in building more complex functions. For example, to use two waves, one with a period of a year and the other with a period of a half a year the following equation would be appropriate:

$$Y = \beta_0 + \beta_1 \cdot \sin(2\pi t) + \beta_2 \cdot \cos(2\pi t) + \beta_3 \cdot \sin(4\pi t) + \beta_4 \cdot \cos(4\pi t) + \varepsilon \qquad \text{Eq. (8)}$$

One way to determine how many periodic seasonal 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 legitimately settle on a model in which the t-statistics for one of the two coefficients in the pair is not significant.

• What matters is that 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.

• There are also cases where the seasonal pattern may be described by a functional form that is not a simple trigonometric function. For example, see Vecchia and others (2008) for an approach used with pesticide data, which has a very specific temporal pattern throughout the year based on the timing of pesticide application.

1.4 Comparison of Methods for Dealing with Seasonality

• The Seasonal Kendall test and mixed approaches have the disadvantages of only being applicable to univariate data (either the original data or residuals from a previous analysis) and are not amenable to simultaneous analysis of multiple sources of variation. For this reason, these methods take at least two steps to compute.

• Multiple regression allows many variables to be considered easily and simultaneously by a single model. The Seasonal Kendall test has the usual advantage of nonparametrics: ***robustness against departures from normality***.

• The OLS regression of deseasonalized *Y* on *T* is perhaps ***the least robust*** because the individual seasonal datasets 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***.

• For example, ***it is unlikely that April will be very different from May***, even though the sample statistics may suggest otherwise. Multiple regression with periodic functions involves very few parameters. However, the functional form (sine and cosine terms) can become a straightjacket, constraining the seasonal pattern to a single form.

• Perhaps the annual cycle really does have abrupt breaks associated with freezing and thawing, or the growing season. Multiple regression can always use binary variables to distinguish the 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 have abrupt changes in slope, but doing this would require a good, physically based definition of the timing of the influential factors.

• 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.

1.5 Presenting Seasonal Effects

• There are many ways of characterizing ***the seasonality of a dataset*** (table 2). Any of the methods can be applied to the raw data or to residuals from a loess or OLS 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.

Table 2. Methods for characterizing seasonal patterns.

Rating	Graphical methods	Tabular methods
Best	Boxplot by season, or loess of data versus time of year	List the amplitude and peak day of cycle
Next best	None	List the 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 the seasonal means, standard deviations, or standard errors

1.6 Seasonal Differences in Trend Magnitude

• The approaches described in previous sections assume a single pattern of trend across all seasons. For example, if we are testing for trends in the mean monthly discharge, these approaches assume that the trend slope or magnitude is identical for every month.

• The reality is that trend magnitudes can ***differ*** greatly across seasons (for example, winter season discharge might have increased substantially but summer might not have changed or might even have decreased).

• ***None of the test*** statistics described above will provide any clue of these differences. This may be a gross ***over-simplification*** and can fail to reveal large differences in the trends across the different seasons.

• In fact, it is entirely possible that the Y variable exhibits a strong positive trend in the spring and summer, but a strong negative trend in the fall and winter. These changes might cancel each other out, resulting in an ***overall Seasonal-Kendall test statistic indicating little or no trend or slope for the time term in a multiple regression approach*** that suggests that there is no trend.

• This is ***not*** to suggest that single test statistics are ***never*** useful with seasonal data; many times we desire a single number to characterize what is happening at a site over the course of an entire year. Yet, when a more detailed examination of trends at an individual site is needed, it is often useful to perform and present ***the full, within-season analysis on each season***. A good approach to graphically presenting the results of such multi-season analyses is seen in **Figure 1**.

• The graph is the result of individual Mann-Kendall trend tests on the monthly discharges for the Sugar River near Brodhead, Wisconsin, for a period of 62 years. The trends in all months (expressed as $\frac{\frac{m3}{s}}{yr}$) are all positive but vary over a range from a low of 0.04 in March to a high of 0.16 in June.

• The significance level of these trends also varies greatly. Several of the months have *p*-values substantially less than 0.01, but for the month of March, it is around 0.5 (not even close to being significant).

• In the approaches using the Seasonal-Kendall test, one can also examine contrasts between the different seasonal statistics. ***Contrasting these results provides a single statistic that 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).

• The Seasonal Kendall test, which considers all 12 months, gives an overall slope of 0.093 m^3/s/yr and a *p*-value of <0.001.

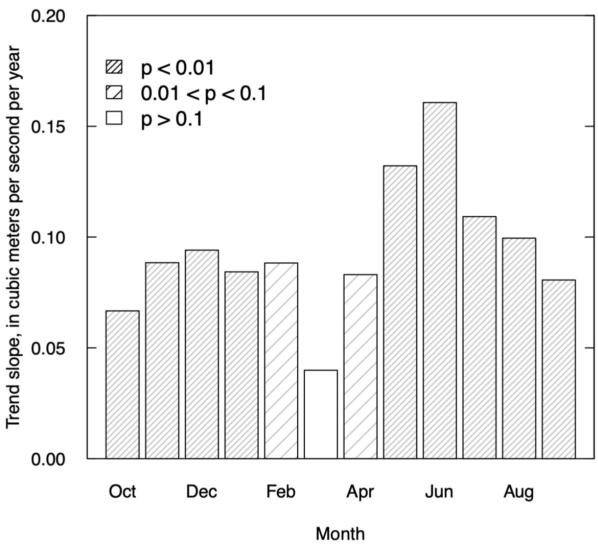


Figure1. Graph of monthly trends in discharge, Sugar River near Brodhead, Wisconsin, for water years 1952–2016. Bar heights show the Theil-Sen slope, by month, for each of the 12 months. Shading indicates the p-value for the Mann-Kendall trend test for the month.

• The Seasonal Kendall test, which considers all 12 months, gives an overall slope of 0.093 m^3/s/yr and a *p*-value of <0.001.

• The EnvStats (Millard, 2013) function kendallSeasonalTrendTest also considers the contrasts of the trends in the 12 months and does not show the trends to be heterogeneous, which means that we should not reject the null hypothesis that the trend is the same in all 12 months. The test itself will not indicate which months have different trends and which are similar, but the graphics can help sort out these differences

1.7 The Regional Kendall

• Another variation on the Mann-Kendall test for trend is the Regional Kendall test, introduced by Helsel and Frans (2006). This test is used when ***a set of Mann-Kendall trend tests are applied to data from a set of nearby monitoring locations that are expected to have correlated data***.

• It operates similarly to the Seasonal Kendall test adjusted for serial correlation. It computes the ***S statistic for each site and sums them to form an overall test statistic for the network of sites***, with the variance of this statistic modified to ***account for the cross correlations between all possible pairs of sites, using the Kendall τ correlation***.

• Suppose we are considering trends in annual mean discharges for a set of streamgages that are very close together, responding to the same precipitation events. In such a case, the cross-correlation between the sites is very high (close to 1), so the Mann-Kendall trend test *S* statistic for each site will be similar. Without adjusting for this cross-correlation, we might incorrectly conclude that the regional trend is highly significant because many sites show significance. However, the trend test results are highly dependent on each other, so the overall evidence for a trend may be no stronger than that provided by any single site.

• Examples of use include Clow (2010), Garmo and others (2014), and Archfield and others (2016). In the EnvStats package (Millard, 2013), it can be implemented using the kendallSeasonalTrendTest function, but the data must be indexed by site instead of by season.

Note: The Regional Kendall

1. Calculate the *S* Statistic for Each Site:

$$S_i = \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \text{sign}\left(x_{ik} - x_{ij}\right) \qquad \text{Eq. (9)}$$

where x_{ij} and x_{ik} are the data points for site *i*, and *sign* is the sign function.

2. Compute the Cross-Correlation Between Sites:

$$\tau_{ij} = \frac{2\left(C_{ij} - D_{ij}\right)}{n(n-1)} \qquad \text{Eq. (10)}$$

where C_{ij} is the number of concordant pairs and D_{ij} is the number of discordant pairs between sites *i* and *j*.

3. Sum the *S* Statistics to Form an Overall Test Statistic:

$$S_{\text{total}} = \sum_{i=1}^{m} S_i \qquad \text{Eq. (11)}$$

where mm is the number of sites.

4. Calculate the Variance of Each Site's *S* Statistic:

For site i, the variance $Var(S_i)$ is given by:

$$\text{Var}\ (S_i) = \frac{n_i(n_i - 1)(2n_i + 5)}{18} \quad \text{Eq. (12)}$$

5. Calculate the Covariance Between Sites:

For each pair of sites i and j, calculate the covariance $Cov(S_i, S_j)$:

$$\text{Cov}\ (S_i, S_j) = \tau_{ij}\sqrt{\text{Var}\ (S_i) \cdot \text{Var}\ (S_j)} \quad \text{Eq. (13)}$$

6. Calculate the Total Variance:

Sum the variances and covariances to get the total variance $Var(S_{total})$:

$$\text{Var}\ (S_{\text{total}}) = \sum_{i=1}^{m} \text{Var}\ (S_i) + 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \text{Cov}\ (S_i, S_j) \quad \text{Eq. (14)}$$

7. Compute the Standardized Test Statistic:

Standardize the overall test statistic using its total variance:

$$Z = \frac{S_{\text{total}}}{\sqrt{\text{Var}\ (S_{\text{total}})}} \quad \text{Eq. (15)}$$

8. Evaluate the Significance:

Compare the standardized test statistic Z to the standard normal distribution to determine the significance of the regional trend.

***Summary*:**

- ***S* Statistic**: Measures the trend at each site.
- **Kendall Tau Correlation**: Measures the correlation between sites.
- **Overall Test Statistic**: Sum of the S statistics from all sites.
- **Adjusted Variance**: Accounts for the correlation between sites, ensuring the overall test statistic is not inflated.
- **Standardized Test Statistic**: Used to assess the significance of the regional trend.

1.8 Discriminating Between Long-term Trends and Long-term Persistence

• One of the most vexing problems in the analysis of trends in streamflow, groundwater levels, or water quality is that it can be very difficult to ***distinguish between*** ***deterministic trends*** and ***long-term persistence***.

• ***A deterministic trend*** in hydrology is a change in the central tendency of a process that comes about because of human activity in the watershed (for example, increasing water consumption, increasing urbanization, decreasing point source loadings owing to improved waste treatment).

• ***Long-term persistence (LTP)*** is a change in the central tendency of a hydrologic variable that is a result of the chaotic behavior of the land-atmosphere-ocean system, that often manifests itself as quasi-periodic oscillation (an oscillation that may have a characteristic length but does not have a regular and predictable periodicity) that exist across a wide range of time scales from years to centuries or millennia.

• LTP has been recognized for many decades (see Hurst, 1951; Mandelbrot and Wallis, 1968; Klemeš, 1974; Koutsoyiannis, 2002); the literature shows that hydrologic time series exhibit persistence at all time scales and that this persistence goes well beyond what can be modeled by simple auto-regressive moving-average (ARMA) processes (Box and others, 2015).

• The annual mean discharge data for the Big Sioux River shown in **Figure 2** provides an excellent example of a highly persistent hydrologic time series. ***There is no obvious cause for this nearly century-long increase*** in streamflow that can be related to water or land management actions.

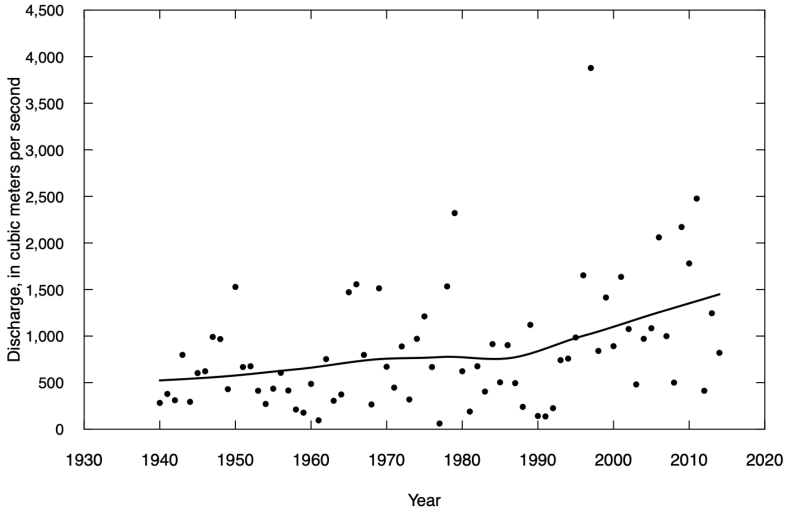


Figure 2. Graph of annual peak discharge, Red River of the North, at Grand Forks, North Dakota, 1940–2014, and a loess smooth of the data (using the default span value of 0.75).

• Thus, we cannot consider it to be a deterministic trend. Rather, it appears to be driven by a particular combination of climate variations at large temporal and spatial scale, along with the properties of the regional groundwater system.

• Long-term persistence (***as well as short term serial correlation***) ***will inflate the variance of a trend-test statistic***. This means that for a stationary, but highly persistent process, the variance of the trend-test statistic will be larger than what it would be under the standard assumptions that ***the random variable is independent and identically distributed***.

• Thus, type I errors have a higher probability than the nominal type I error rate for the trend test being used. Although this text does not address these issues in any detail, it is nonetheless an important concept when thinking about the statistics of water resources and has been the subject of extensive discussion in the literature.

• Cohn and Lins (2005) provide an excellent overview of the difficulties encountered in attempting to distinguish between ***long-term persistence and trend***. They reference many important papers that deal with these issues and offer some potential solutions in terms of hypothesis testing.

• Being able to distinguish between deterministic trend and long-term persistence is important because it provides guidance about how to understand what the future will be like.

• If an apparent trend is largely a ***deterministic*** one, then if we can successfully build a simulation model of how the drivers of that deterministic trend influence the variable of interest, then ***we can base our forecast*** of the future on forecasts of that driving variable.

• But if the apparent trend is simply a manifestation of long-term persistence then ***we have no real ability to forecast*** the variable of interest because the drivers of the phenomena could just as easily continue at their current levels or reverse themselves tomorrow and take the system in a new direction.

• Most hydrologic variables are responding to a mix of these two types of drivers, and assessing that mixture is critical to projecting future conditions. One way that long term persistence affects trend analysis is in the selection of the time period over which a trend is assessed.

• The following example data illustrates the problem. Annual peak discharges of the Red River of the North at Grand Forks, North Dakota from 1940–2014, and a loess smooth of the dataset are shown in **Figure 2**.

• It is obvious from visual inspection, and confirmed by trend test results, that the null hypothesis of no trend is easily rejected. However, if we look further back in the historical record to the first recorded observation in 1882, we come away with a very different impression (**Figure. 3**).

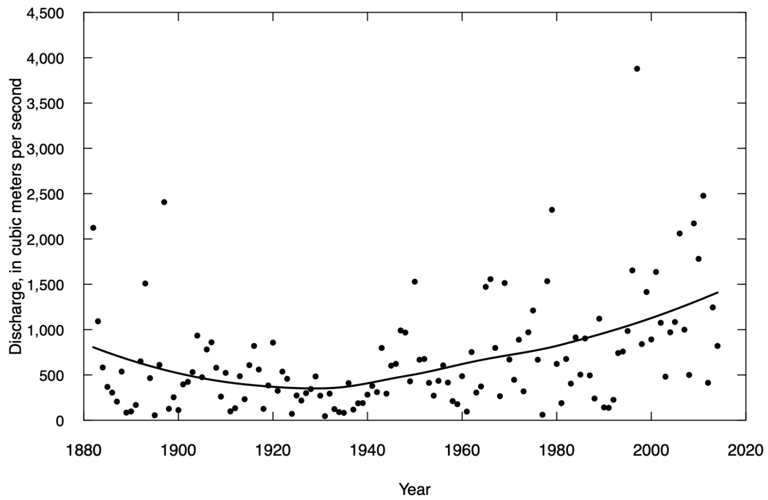


Figure 3. Graph of annual peak discharge, Red River of the North, at Grand Forks, North Dakota, 1882–2014, and a loess smooth of the data (using the default span value of 0.75).

• From the 133-year time scale perspective we may conclude that we have some type of quasi-periodic oscillation and our observations constitute ***something less than one full cycle*** of this oscillation.

• This conclusion should inform us that very large swings in flood magnitude over multidecadal periods are something we should expect in this watershed. Results of statistical tests (on either the shorter or longer version of this dataset) that include adjustment for serial correlation, including possible LTP, will indicate the presence of a statistically significant trend, but with an attained significance level (p-value) that is much larger than would arise from a standard method that assumed independence (such as Mann-Kendall or OLS regression approaches).

• Regardless of the specific p-value reported the records shown here tell us that over the next few decades flood magnitudes are unlikely to be similar to those observed in the mid-20th century.

• As a consequence, management actions should ***consider the possibility of the recent trend continuing for some years into the future***, while recognizing ***that a reversal of this trend could certainly happen at any time***.

• Admittedly this case is a rather extreme example, however, quasi-periodic oscillations are well known to exist in climate and hydrology time series (typically associated with phenomena such as **El Niño** or the Atlantic Multidecadal Oscillation).

• The analyst should always consider the possibility that what is seen as a ***deterministic trend is merely one limb of such an oscillation***.

• Where strong persistence is evident in the data, planners must be prepared for a future that is quite different from the recent past. ***In the limit, using a null hypothesis that includes oscillatory or long-term persistent processes***, it becomes impossible to ever reject the hypothesis that the process is highly persistent but stationary.

• Vogel and others (2013) stated the problem succinctly, "Our ability to distinguish stochastic persistence from deterministic trends is in its infancy (Cohn and Lins, 2005)… Earth systems evolve over space and time, thus new theory and practical algorithms are needed to address long term social and physical drivers and feed-backs. New exploratory and statistical tools are needed to sharpen our insights into the emergent properties of such systems, and to guide modeling and prediction."

• Loftis and others (1991) also offer a useful perspective on this problem. They note that "… a process which is stationary over long times … may contain short-term runs which would be important from a management standpoint."

• There are methods of adjusting parametric and nonparametric trend tests to partly account for the effects of ***serial correlation***. These methods are designed to achieve the nominal type I error rate (for example $\alpha = 0.05$) when the data arise from a stationary process that follows a specified temporally dependent process.

• These adjustments are helpful, but in practice the process parameters are generally estimated from the data, and the presence of a deterministic trend in the data will influence these adjustments.

• For example, in the case of the Mann-Kendall test, an adjustment for autocorrelation was developed by Yue and others (2002) and has been implemented in the zyp package in R (Bronaugh and Werner, 2013). The adjustment is referred to as "prewhitened nonlinear trend analysis" and variations and power analysis of the method have been discussed and debated in the literature.

• Other recent literature relevant to this problem include Zhang and others (2000), Matalas and Sankarasubramanian (2003), Yue and others (2003), Bayazit and Önöz (2007), Hamed (2008), Önöz and Bayazit (2012), and Wang and others (2015).

• Fortunately, as was demonstrated by Cohn and Lins (2005), even if we are dealing with a process that had no serial correlation or LTP there is only a very small loss of power associated with using a trend test designed for a serially correlated or persistent process versus using a test designed for white noise.

• The WRTDS method, mentioned above, uses a block bootstrap method for computing statistical significance of trends (Hirsch and others, 2015). This method accounts for the influence of serial correlation at time scales of ***about 200 days or less***, but does not account for longer term persistence such as decadal or multidecadal persistence or quasi-periodic oscillations.

Note: Quasi-periodic oscillations (QPOs)
Quasi-periodic oscillations (QPOs) refer to fluctuations in a system that exhibit periodic behavior but ***without a fixed, regular interval***. In other words, these oscillations have a characteristic timescale but the ***period between peaks or troughs is not constant***, and the exact length of the cycle can vary. This variability means that while QPOs have a general tendency to repeat over certain intervals, ***the timing is not strictly predictable***. In the context of hydrology and climate systems, QPOs are often associated with large-scale climate phenomena that affect weather and water patterns over extended periods. Examples of such phenomena include: 1. **El Niño-Southern Oscillation (ENSO):** o Period: Approximately 2 to 7 years. o ENSO is a climate pattern that involves fluctuating ocean temperatures in the central and eastern equatorial Pacific, affecting global weather and climate. 2. **Atlantic Multidecadal Oscillation (AMO):** o Period: Approximately 60 to 80 years. o AMO is a climate cycle affecting the sea surface temperature of the North Atlantic Ocean, with impacts on weather patterns in North America, Europe, and other regions. 3. **Pacific Decadal Oscillation (PDO):** o Period: Approximately 20 to 30 years. o PDO is a long-term ocean fluctuation of the Pacific Ocean, influencing the climate in the Pacific Basin and North America. The exact period of these oscillations can vary due to the complex interactions within the Earth's climate system. Therefore, while researchers can identify characteristic timescales for these phenomena, the precise timing and duration of each cycle can differ.

• The problem of distinguishing deterministic trends from persistence is a serious issue in the analysis of water quality, where records are generally much shorter and have much lower sampling frequencies than those for discharge.

• An added difficulty regarding water quality trends is that there are so few datasets of long duration (say greater than 30 years) that also have a high sampling frequency (at least a few dozen samples per year).

• Without such long-term high-frequency datasets it will continue to be difficult to sort out natural long-term variations, shorter-term serial correlation, and trends owing to changes in human activity. The advent of multiyear records from monitors that measure water quality at time steps such as 15 minutes or 1 hour could be very helpful in sorting out some of these issues (see for example, Godsey and others [2010]).

• These issues of distinguishing deterministic trends from persistent or quasi-periodic oscillations of hydrologic processes remains a challenge for the water resources community. We know that trends are ubiquitous in hydrologic time series and it is difficult to sort out the relative roles of human activities on the landscape, human driven changes in the global atmosphere, and chaotic behavior of the land-atmosphere-ocean system that ***would exist in the absence of human activity***.

• Awareness of all of these drivers of change cause us to question the long-standing statistical foundations of hydrology (in other words, time series that are independent and identically distributed). Milly and others (2008) stated the problem this way: "…we assert that stationarity is dead and should no longer serve as a central default assumption in water-resource risk assessment and planning. Finding a suitable successor [to stationarity] is crucial…"

• Describing trends is an important scientific goal in support of hazard mitigation, water resources planning, and evaluation of water quality improvement strategies. Statistical tools will need to continue to evolve and improve (see Salas and others, 2018). Such tools need to be cognizant of the atmospheric and watershed processes that drive the changes that would exist even in the absence of human interventions.

• There is no simple solution to this problem of distinguishing persistence from deterministic trends. The best advice to the analyst is to be cognizant of these issues in designing their analysis and explaining the meaning of the results.

1.9 Final Thoughts About Trend Assessments

• Conveying trend results is most effective through graphics, focusing on capturing major patterns of change. Analysts aim to maximize the signal-to-noise ratio ***by removing variations from exogenous factors like seasonality, enhancing the clarity of the signal***. Thoughtful trend analysis protocols are crucial for identifying unexpected behaviors and verifying the accuracy of deterministic models. Surprises in trend analyses can improve understanding and predictive capabilities, essential for effective management.

• As time series data lengthens, it is vital to recognize nonmonotonic trends and identify the timing and magnitude of reversals. The goal is to align observed trends with natural or human-driven forces such as regulations, land use changes, and atmospheric conditions. Analyzing multiple datasets is key to achieving this, but practical problems arise when dealing with multiple starting and ending dates and gaps in records. Records must be concurrent for accurate interpretation.

• Trend results should not be filtered by significance levels alone, as trends that are ***not statistically significant can still be valuable***. Analysts should quantify the confidence in trend direction and magnitude rather than exclude results based on significance.

• Decision-makers need to understand the likelihood of making wrong decisions based on trends. Analysts should also characterize the nature of the trend, its rate of change, and any changes in slope or direction. This helps verify predictive models, evaluate water management strategies, and provide early warnings of emerging issues. Hydrologic systems are noisy, and the analyst's role is to make trend signals clear and understandable for the public and decision-makers, ultimately improving models and tools for future water resource planning and management.