

Protocol for Monitoring Fish Communities in Small Streams in the Heartland Inventory and Monitoring Network

SOP 9: Data Analysis, Version 2.0

Revision History Log:

Previous Version #	Revision Date	Author	Changes Made	Reason for Change	New Version #
1.0	7/6/2016	Morrison & Dodd	Description of statistical methods	Added trend analysis techniques and clarification on test assumptions for parametric tests	2.0
1.0	4/20/2017	Dodd	Added IBI for coolwater streams (Lyons et al. 1996) at EFMO	Not in version 1.0	2.0

Ecological studies based on biological, chemical, and physical data are used by resource managers to better comprehend underlying system processes and develop environmental/management policies that best serve the resource. A critical component of any long-term monitoring protocol is a consistent and systematic approach to analyzing and reporting data. This information must describe the current condition, or status, of a community, and be robust enough to detect community changes through time. This, in turn, could have substantial ecological repercussions and should be an important consideration for investigators responsible for data interpretation. Therefore, every effort should be made to collect reliable data and use statistical analyses that are straightforward and will result in confident interpretations. This SOP describes the metrics to be calculated for fish and habitat data collected from small streams, and statistical analyses for interpreting those metrics.

The data analysis process should be flexible enough to allow the use of newly developed statistical and analytical techniques and tailoring of analyses for a variety of audiences with a variety of questions about the aquatic resources at the parks. This primary audience will consist of park resource managers, superintendents, and other staff who may not have an in-depth background in statistical methods or may have limited time for evaluation of these analyses. Thus, to the extent possible, it is important that core data analysis and presentation methods are relatively straightforward to interpret, provide a standard format for evaluation of numerous variables, can be quickly updated whenever additional data become available, and work for many different types of indicators, whether univariate or multivariate. Additionally, the type and magnitude of variability or uncertainty associated with the results should be somewhat intuitive, and it may be necessary to indicate a threshold for potential management action. Primary approaches to analyzing fish data are metric estimation with use of trend analysis (parametric and non-parametric), control charts, and multivariate statistics.

Fish Metric Calculations

Biological metrics are commonly used by investigators at all levels (private, state, tribal, and federal, for example) to compare the condition of the biological community at multiple sites (Simon 1999) or examine trends over time. Barbour *et al.* (1999) define a metric as a characteristic of the biota that changes in an expected direction with increased anthropogenic disturbance. Using these characteristics (see Table 3 in SOP 4) allows investigators to determine the importance of environmental conditions, clarify which habitat factors play a large role in shaping fish communities, and identify specific sources of impairment (Karr 1981). Fish community metrics and estimated variables for analysis will be calculated at the reach scale, combining all gear types (if applicable) used and all channel type habitats (riffle, run pool) sampled in the reach. Individual fish community metrics include:

1. **Species richness** (number of species) for the entire sample reach.
2. **Simpson's Diversity Index (D)** for entire sample reach. Simpson's Diversity Index is preferable to the Shannon diversity index and will be used for data analysis because this index is independent of sample size. Simpson's Diversity Index is calculated with the following formula:

$$D = \sum_{i=1}^S ((n^2 - n)/(N^2 - N))$$

where S = number of species,

n = number of individuals of ith species, and

N = sum of n.

3. **Catch per unit effort.** Catch per unit effort (CPUE) can be calculated as either Catch per Time or Catch per Area sampled. When using electrofishing gear to collect community data, CPUE is typically calculated as Catch per Time. We will use catch per minute to obtain relative abundance. When using seining methods, CPUE is typically calculated as Catch per Area. We will use catch per square meter.
4. **Average length and weight** for each species in the reach.
5. **Percent species composition by abundance.** To calculate this metric, the individual species catch will be divided by total catch of all fish within the reach and multiplied by 100.
6. **Percent species composition by biomass.** To calculate this metric:
First, calculate individual species biomass (B_i):

$$B_i = W_i N_i$$

W_i = average weight of fish species i

N_i = number of individuals of species i

Second, calculate biomass per area for each species i (BPA_i) and total fish biomass per area (BPA_t) using the entire area of the sample reach (A):

$$BPA_i = B_i / A$$

$$BPA_t = \sum BPA_i$$

Lastly, calculate percent composition for each species (C_i)

$$C_i = BPA_i / BPA_t * 100$$

By combining multiple metrics and results for those metrics into a single index of biotic integrity (IBI), investigators can determine the overall quality of the fish community. An IBI also can be used to compare overall ecological conditions over time and between sites, providing that the selected metrics have a relation to variables responsible for impairment (Karr 1981, Barbour *et al.* 1999, Simon 1999). One of the first fish IBIs developed by Karr (1981) has been modified for use in many other states, regions, and countries (Fausch *et al.* 1984; Lyons *et al.* 1996; Hlass *et al.* 1998; Hughes and Oberdorff 1998, Simon 1999, Yoder and Smith 1999, Dauwalter *et al.* 2003, and Smogor 2005). We evaluated specific IBI's for use at parks located in regions or states that have a published IBI. To obtain an IBI, individual metrics are calculated and scored. Scores are totaled for an IBI score. The higher the IBI score the better the condition of the community.

1. For those parks located in the Ozark Highlands, the IBI by Dauwalter *et al.* (2003) will be used in streams at GWCA, PERI, and WICR. Details on calculation and scoring of metrics can be found in Dauwalter *et al.* (2003). The metrics used in this IBI are:
 - a. percent of individuals as algivorous/herbivorous, invertivorous, and piscivorous
 - b. percent of individuals with black spot or an anomaly
 - c. percent of individuals as green sunfish, bluegill, yellow bullhead, and channel catfish
 - d. percent of individuals as invertivorous
 - e. percent of individuals as top carnivores
 - f. number of darter, sculpin, and madtom species
 - g. number of lithophilic spawning species

2. For HOSP, which is located in the Ouachita Mountains region, the IBI developed by Hlass *et al.* (1998) will be used in these streams. Details on calculation and scoring of metrics can be found in Hlass *et al.* (1998). Metrics in this IBI include:
 - a. Number of native species
 - b. Number of intolerant species
 - c. Number of minnow species
 - d. Proportion of individuals as green sunfish
 - e. Proportion of individuals as top carnivores
 - f. Ratio of generalist to specialist feeders
 - g. Number of individuals in sample (catch per minute)

- h. Proportion of individuals with disease or other anomaly
3. For Dousman's Creek at EFMO, which is a cool water stream, the IBI developed by Lyons *et al.* (1996) will be used. Details on calculation and scoring of metrics can be found in Lyons *et al.* (1996). Metrics in this IBI include:
- a. Number of intolerant species
 - b. Percent of individuals that are tolerant species
 - c. Percent of individuals that are top carnivores
 - d. Percent of individuals that are stenothermal coolwater and coldwater species (native and exotic)
 - e. Percent of salmonid individuals that are Brook trout (*Salvelinus fontinalis*)
4. For the remaining parks, the IBI developed by Fausch *et al.* (1984) for the Midwest region, will be used in streams at HEHO, HOME, PIPE, and TAPR. Metrics calculated in this IBI are:
- a. Number of species
 - b. Number of darter species
 - c. Number of sunfish species
 - d. Number of sucker species
 - e. Number of intolerant species
 - f. Proportion of individuals as green sunfish
 - g. Proportion of individuals as omnivores
 - h. Proportion of individuals as insectivorous minnows
 - i. Proportion of individuals as top carnivores
 - j. Number of individuals in sample
 - k. Proportion of individuals as hybrids
 - l. Proportion of individuals with disease, tumors, fin damage, or other anomalies

Habitat and Water Quality Parameters

Physical and chemical habitat measurements will be estimated using summary statistics such as means, medians, standard errors and/or confidence intervals. In the simplest presentation of data, each parameter should be estimated in each year that data are available, and confidence intervals or standard errors calculated, where appropriate. For parameters where percentage cover class categories are used (such as substrate, embeddedness, canopy cover) the median value of the cover class will be used in calculating means and standard errors. Those parameters where presence/absence data are collected (such as fish cover and bank cover types), a percentage will be calculated for the reach.

Using these estimated fish metrics and habitat parameters listed above, several analytical techniques may be employed to determine trends or changes in fish communities or habitat parameters and to ascertain underlying relationships between fish and environmental factors.

Basic Charts

The most basic information to provide to resource managers is a graph or chart plotting the point estimate of the respective variable over time. In general, graphs are the easiest way to communicate information to a non-technical audience. Thus, the first step in most analyses will be to plot the data and carefully examine the resulting graph. Visual examination of data is important; inferences should never be made based on test statistics alone (Anscombe 1973).

The degree of uncertainty associated with the point estimate should also be included, if available. For statistical uncertainty (i.e., random sampling error), 95% confidence intervals or standard errors should be used. For fish metrics, which are calculated at the reach scale, there is no calculation of statistical uncertainty ($N=1$). However, calculation of mean fish metric values and statistical uncertainty across sites at a park (N = number of reaches) or across years at reach (N = number of years a reach is sampled) may be warranted for specific analyses. If any other sources of error are documented and quantified (i.e., observer error, lack of precision in equipment, etc.), these could also be incorporated in such graphs as broader error bars.

If some threshold can be identified (e.g., a desired condition, a regulatory standard, etc.), this could be included and the product would be a type of control chart (see below). Any sort of relevant threshold that could be included would help to provide context to the graph.

Outliers may become obvious at this stage, depending upon how the data are plotted. Outlier tests such as Grubbs or Rosner outlier tests (Grubbs 1950, Rosner 1983) may be employed to determine whether an observation is a statistical outlier. Such outlier tests, however, determine only whether an observation is likely to have been generated from a normal distribution. Most field data from ecological studies are skewed, and do not resemble a normal distribution. Instead of using such tests, if any outliers are present, the data represented by the outliers should be verified for accuracy (see Data Management SOP 8). If no errors in data collection or entry can be found, the data points probably represent important information and should not be discarded.

Checking Assumptions and Remedial Measures

Before conducting any analysis, the assumptions of the analysis should be evaluated. In general, parametric tests will have more restrictive assumptions than non-parametric tests. To assess normality of the data, normal probability plots, Skewness and Kurtosis statistics, or the Shapiro-Wilk test can be used (Shapiro and Wilk 1965, D'Agostino *et al.* 1990, Kutner *et al.* 2005). Skewness and Kurtosis are descriptive statistics where a Skewness of 0 and a Kurtosis of 3 indicates normality of the data. The Shapiro-Wilk test provides a test statistic and *P*-value to determine normality. To test for homogeneity of variances, Bartlett's test, Levene's test, or Mauchly's test can be used (Bartlett 1937, Mauchly 1940, Levene 1960).

Graphical analysis of residuals can reveal nonlinearity of the regression function, nonconstancy of error variance, nonindependence of error terms, and autocorrelation (Kutner *et al.* 2005). A Durbin-Watson test can be employed to formally test for temporal autocorrelation (Savin and White 1977). The Durbin-Watson test provides a test statistic (D) between the values of 0 and 4. A value of or near 2 indicates no autocorrelation; a value of or near 0 indicates positive autocorrelation, whereas a value of or near 4 indicates negative autocorrelation. To test for significance of the test statistic (D), Durbin-Watson tables listing lower and upper bounds (d_L and d_U) for combinations of sample size and number of independent variables can be used. If D is less than d_L , then there is significant positive autocorrelation. If D is greater than d_U , then there is no significant autocorrelation. If D lies between d_L and d_U , the test is inconclusive. An inconclusive result does not mean there is no temporal autocorrelation, only that the test could not determine if autocorrelation exists because the sample size is too small (Kutner *et al.* 2005). The inconclusive range of the statistic grows wider as the sample size become smaller. If a test for negative autocorrelation is required, the test statistic would be $4 - D$. If $4 - D$ is less than d_L , then there is significant negative autocorrelation, etc.

Data transformations may be helpful prior to use of parametric tests that require normality and independence of the error terms (Kutner *et al.* 2005). Once the data are transformed, procedures to test for normality and autocorrelation should be re-examined. If significant autocorrelation exists, remedial measures include transformations such as the Cochrane-Orcutt procedure, the Hildreth-Lu procedure, and the First-Differences procedure (Kutner *et al.* 2005). None of these transformations is guaranteed to remove autocorrelation, however, and if significant autocorrelation is found, the best solution in most cases would be simply to apply a non-parametric test such as the modified Mann-Kendall test (see Non-Parametric Test for Trend section).

Control Charts

The construction and interpretation of control charts is covered in many texts focusing on quality control in industry (*e.g.*, Beauregard *et al.* 1992, Gyra 2001, Montgomery 2001). The application of control charts for ecological purposes is relatively straightforward. The use of control charts in environmental monitoring is discussed in texts by McBean and Rovers (1998) and Manly (2001), although not in as great detail as the texts referenced above focusing on industrial applications. Many different types of control charts could be constructed, depending upon the type of information desired. For example, control charts can be used to evaluate variables or attributes (*i.e.*, richness or number of intolerant taxa, count or frequency data) to focus on measures of central tendency or dispersion and in univariate or multivariate analyses.

Most traditional control charts assume that observations come from a normal distribution, or that data can be transformed to normality. If the observations cannot be assumed to come from a normal distribution, there are several options available beyond simple transformations of data. One option is to create subgroups of consecutive samples, and then use the subgroup averages, which will be approximately normally distributed in accordance with the central limit theorem (see Beauregard *et*

al., 1992, Montgomery 2001). It is possible to construct control charts based on other distributions (e.g., a Poisson distribution as in Atkinson *et al.* 2003) and construct analogous confidence limits as long as the distributions are known. Distribution-free confidence limits may also be calculated, although these will usually be relatively wide and less sensitive to changes (Conover 1999).

Control limits may be constructed so as to contain any desired proportion of the distribution (*i.e.*, representing $[1-\alpha]$ confidence intervals for any Type I error rate of α). In this case, choosing control limits is equivalent to specifying a critical region for testing the hypothesis that a specific observation is statistically different from the proposed centerline value. In industry, control limits are often set at a distance of 3 standard deviations on either side of the centerline (Wetherill and Brown 1991, Beauregard *et al.* 1992, Montgomery 2001). Thus, assuming a normal distribution centered at the centerline, the control limits would encompass 99.73 % of the distribution (*i.e.*, 99.7% confidence interval). To detect true changes in an ecological parameter of interest, it is crucial that the centerline value is representative of the true population parameter and the control limits account for the natural variability in the population. Therefore, caution should be taken when selecting the baseline dataset used to calculate the centerline (mean) and control limits (based on standard deviation or confidence interval). To assist in determining the appropriate baseline dataset, one approach might be to plot the sample variance for $N=3, 4, 5, 6$, etc. consecutive years of the dataset and graphically analyze for an asymptote. The location where the variance begins to asymptote would indicate the number of years where variability is no longer increasing in the dataset and thus, indicates the number of years to use as the baseline. In addition, temporal autocorrelation will affect the probabilistic interpretation of control limits based on standard deviations, and should be evaluated if such an interpretation is to be employed. Control limits could also be based on probabilistic thresholds other than confidence intervals (*e.g.*, McBean and Rovers 1998).

If control limits are based on a statistical sampling process and interpreted in relation to the Type I error rate (α), this represents a type of statistical threshold (see Figure 1), the conclusion regarding the parameter may differ from that based on a trend test, such as the Mann-Kendall test or parametric regression. This is because one is testing different hypotheses (e.g., whether a threshold is crossed versus whether a trend exists). For example, a control limit representing a Type I error rate of 0.05 could be crossed, but no significant trend found ($P > 0.05$). An advantage of control charts is that a dramatic change that results in a control limit being crossed will be noticed immediately, whereas several sampling periods may be required before a trend-line becomes significant. Control charts provide complementary information to trend tests, and both may be used simultaneously, although one must realize the potential for confusion in the reader.

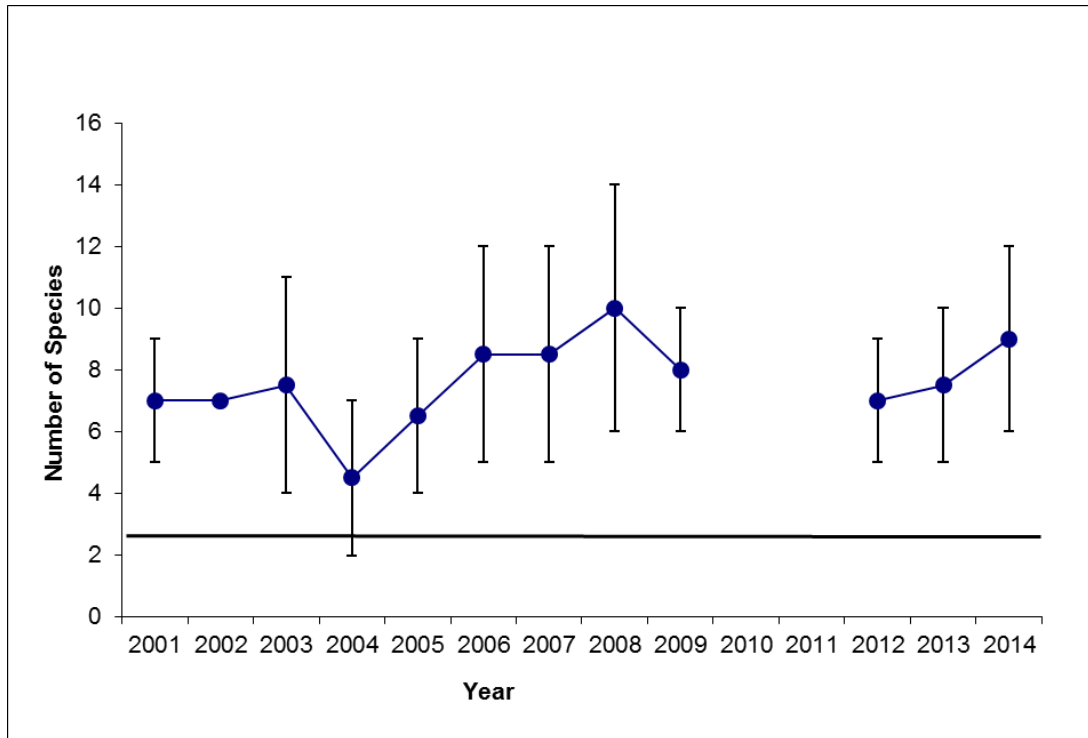


Figure 1. An example univariate control chart using a control limit corresponding to a Type I error rate (α) of 0.05. First 9 years of data were used as a baseline for computing the centerline and control limit.

It is not absolutely necessary to use values from a statistical sampling process to determine centerlines and thresholds for action. It is possible to set threshold limits based on desired management goals (e.g., diversity $\geq 75\%$), regulatory/biological thresholds (e.g., dissolved oxygen ≥ 5 mg/L), or published community integrity categories (e.g., IBI score corresponding to a “good” integrity rating; see Figure 2). It is crucial to realize that this approach has no statistical basis, and thus probabilities cannot be readily associated with the observations. This application also has a precedent in industry. Such charts, which plot observations without relevance to an underlying distribution, have been termed ‘conformance charts’. Threshold values are termed ‘action limits’ (Beauregard *et al.* 1992). If taking this approach, one should be very familiar with the system in question, and select control limits that are defensible based on the data or are based on published findings from similar systems.

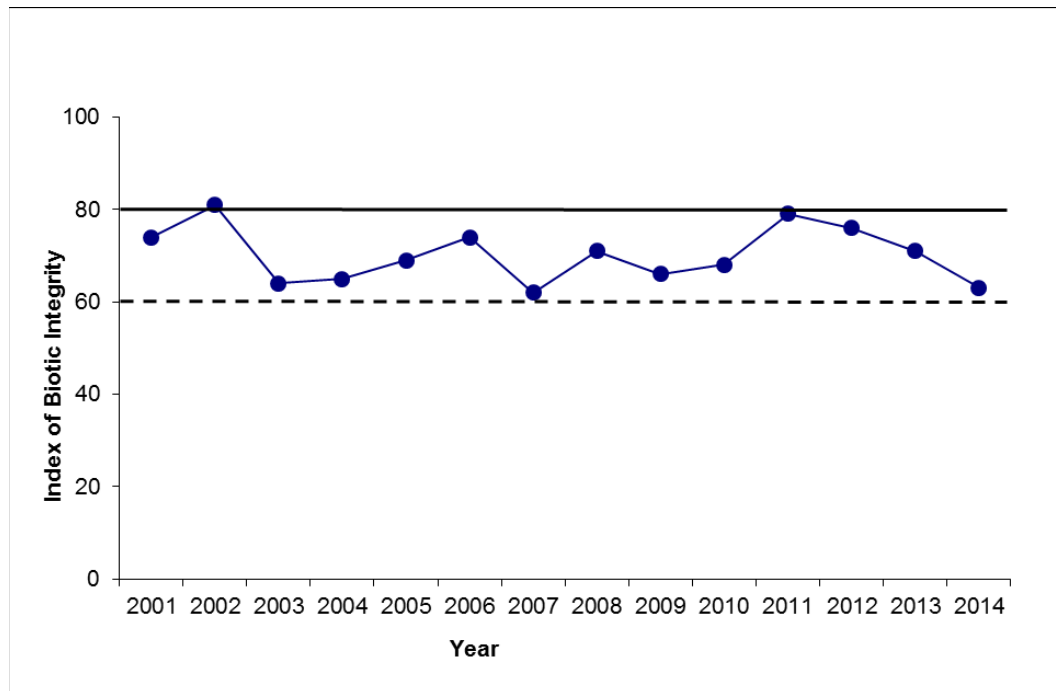


Figure 2. An example univariate control chart using rating categories as a control limit for the fish Index of Biotic Integrity. The horizontal lines represent the control limits corresponding to an integrity rating of 'good' (dashed line; IBI between 60 and 80) and 'reference/excellent' conditions (solid line; IBI ≥ 80).

Although some of the above mentioned texts discuss the use of multivariate control charts (using the Hotelling T2 statistic), this approach is only practical for a small number of variables, and assumes a multivariate normal distribution. In general, species abundances are not distributed as multivariate normal (Taylor 1961), and traditional multivariate procedures are frequently not robust to violations of this assumption (Mardia 1971, Olson 1974). Additionally, multivariate temporal autocorrelation will violate the assumption of stochasticity upon which this method is based, and a test for temporal autocorrelation can be completed using Mantel correlograms (Borcard and Legendre 2012) prior to using this method.

A type of multivariate control chart has been described for use with complex ecological communities (Anderson and Thompson 2004). A software application entitled ControlChart.exe is available for constructing these types of multivariate control charts (see Anderson and Thompson 2004). This method is essentially a distance based ordination approach that uses a bootstrapping technique to generate probabilities associated with deviations from central points. It should be noted that this software utilizes data from all years of collection to calculate percentiles by the bootstrapping method. Thus, the control limit may "shift" each year a new data point is added to the dataset.

Although control charts have potentially wide applicability, each application may be different. A generic process for control chart construction is provided below, although decisions will always have to be made and an analyst familiar with control charts should ideally be consulted.

Steps in constructing a univariate control chart:

1. After several years of data are available (> 5 years), plot the values of the parameter of interest (on the y-axis) against time (on the x-axis).
2. Determine a centerline value for this parameter. This centerline could represent a mean of the observations or a target value. Determining an appropriate center-line contains inherent pitfalls, and an analyst who is familiar with both control charts and the system from which the data is collected should be consulted.
3. Establish control limits. Control limits may be based on a probability distribution (i.e. confidence intervals) and thus allow one to make statistical inferences, or they may be based on target levels set by management or published findings (e.g., management goal, regulatory threshold, etc.). It is possible that only an upper control limit, or only a lower control limit, or both will be necessary, depending upon the parameter of interest and management concerns. Once again, determining appropriate control limits can be tricky, especially if statistical inferences are desired, and an analyst who is familiar with control charts and the system from which the data is collected should be consulted.
4. Continue to plot values of the parameter of interest over time as new data become available. If an observation exceeds the control limit(s), this is indicative of the potential need for management action or a more focused study.

Parametric Tests for Trend

Parametric regression may be used with monitoring data. Linear regression is commonly employed, although other regression functions should be considered. Over longer time series, it is unlikely that a simple linear function will adequately fit the data (Bart and Beyer 2012, Johnson 2012). Thus, the first step is to evaluate a graph of the data to visualize what trends may be apparent (see Basic Charts section above), and specify time periods of interest (if not the whole data series). In general, parametric regression attempts to fit a particular function to a data set (i.e., by drawing a line as close to all data points as possible). Numerous functions should be considered, including linear, exponential, polynomial, and breakpoint (for example, the “Curve Estimation” option of the “Regression” function in SPSS software allows for simultaneous comparison of a number of different types of regression functions). Decisions about model adequacy should not be made on the basis of test statistics alone, however. A large R^2 or significant F-statistic does not guarantee that the data have been fitted well (e.g., Figure 3).

The advantage of using an exponential model is that it describes change at a constant proportional rate, which provides a useful way of describing a trend. For example, if an exponential curve is fitted and the annual proportional rate of change (R) is 1.01, the variable of interest is interpreted as increasing at a rate of 1% per year. If R is 0.98, for example, the variable of interest is interpreted as decreasing at a rate of 2% per year (raise the base of the natural logarithm (e), which is equal to 2.71828, to the value of the slope parameter, b_i , to find R [i.e., $e^{b_i} = R$]). Linear and polynomial curves do not change at a constant proportional rate so describing annual change is more difficult

(Bart and Beyer 2012). Again, providing information on the magnitude of change is much more informative than merely determining ‘significance’.

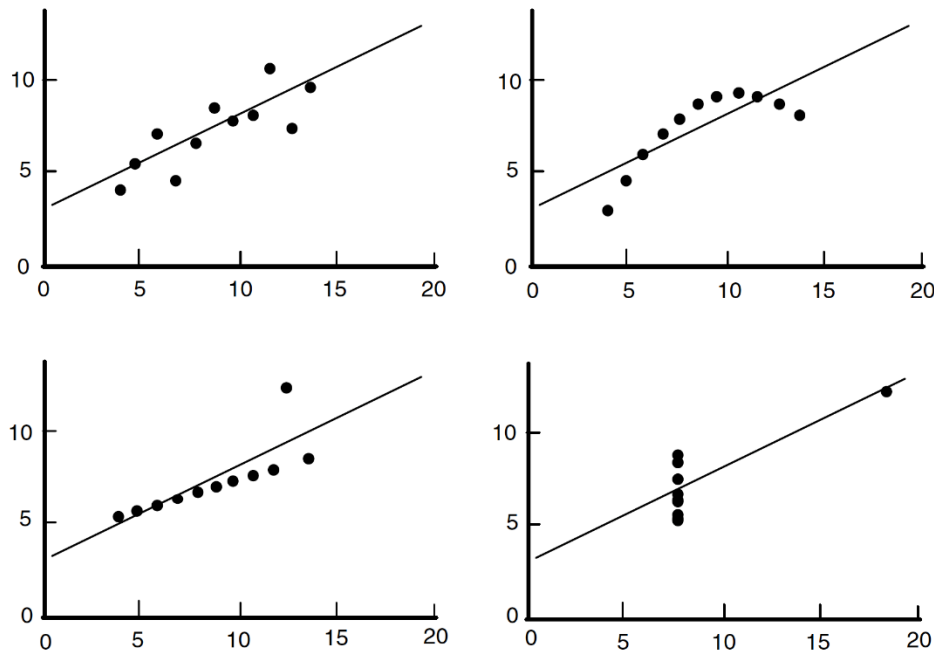


Figure 3. The data in the four graphs have exactly the same summary statistics and regression line (from Anscombe 1973).

True time series modeling (i.e., autoregressive integrated moving average [ARIMA] models) can be used for trend analysis but requires at least 50 and preferably > 100 observations (Box and Tiao 1975). Such models are complex and difficult to interpret. Thus simpler approaches are encouraged, and ultimately necessary given the available sample sizes.

Parametric regression requires that specific assumptions are met. For example, the four principal assumptions of linear regression are: (1) model form is correct (y is linearly related to x), (2) constant variance of residuals, (3) independence of errors, and (4) data are from a normal distribution. If all assumptions are not tested and met, results may be misleading (see Checking Assumptions and Remedial Measures section above). Serial correlation will violate the assumption of independent error terms. Outliers can cause serious problems because a fitted line may be pulled disproportionately toward an outlying observation. To meet the assumptions of parametric regression, data transformations are often required for ecological data. The use of parametric regression, the testing of assumptions, and potential transformations is covered in many statistics texts (e.g., Kutner *et al.* 2005) and can be implemented with many different software applications.

If a Mann-Kendall test indicates no significant trend (see description in Non-parametric Tests for Trend section), the application of parametric regression is unlikely to provide any additional useful information. If the Mann-Kendall test does indicate a significant trend, but interest is on describing

the parameters associated with the function of the relationship, then parametric regression may be warranted. Parametric regression (polynomial or breakpoint) may also be useful for a longer time series when the trend is obviously not constant over the entire data range.

Non-parametric Tests for Trends

Mann-Kendall test

A basic test for the existence of a monotonic trend in a data series is the Mann-Kendall test. It is a non-parametric analogue of parametric regression that does not require normality of the data or that the trend be linear (Mann 1945, Kendall 1975, Hirsch *et al.* 1982). If the model form specified in a regression equation were known to be correct (Y is linear with T) and the residuals were normal, then parametric regression would be optimal. If the actual situation departs, however, even to a small extent, then the Mann-Kendall test procedures will perform as well or better (Hirsch *e. al.* 1991, Helsel and Hirsch 2002).

The original Mann-Kendall test for trend is sensitive to serial correlation (see Checking Assumptions and Remedial Measures section above for serial correlation testing), however, with the probability that a significant trend will be detected being substantially higher than the stated alpha. A modified seasonal test is available (Hirsch and Slack 1984), in which the probability of detecting a significant trend (at a given alpha level) is very close to the stated alpha even when serial correlation is present, given at least 10 years of data. (Seasonal data are not necessary to employ the test.) This modified test is less powerful than the original test when data are independent, however. The modified version should be used in the presence of serial correlation.

Sen Slope (Kendall-Theil Robust Line)

This is a robust nonparametric test to determine if Y is linearly related to X. This test does not depend on normality of residuals and is not strongly affected by outliers, in contrast to parametric regression. When residuals are normally distributed, parametric regression is slightly more efficient than the Sen Slope. When residuals depart from normality (are skewed or prone to outliers), then this nonparametric test can be much more efficient than parametric regression (Helsel and Hirsch 2002). The Sen Slope, however, is affected by serial correlation, in a similar manner as parametric regression.

The Sen Slope is the median annual slope of all possible pairs of values. It is expressed in the units of the variable, standardized to a year. For example, if a parameter is measured in each of 10 years, the slope is the difference between the endpoints of the fitted line divided by 10 (the number of years). Median annual slope is often compared to the median value of the variable as a measure of the magnitude of change. A statistically significant Kendall test with an annual trend of more than 1% of the median value is often considered meaningful (TimeTrends manual; <https://www.niwa.co.nz/freshwater/management-tools/water-quality-tools/analysis-of-water-quality-trends>). Providing information on the magnitude of change is much more informative than merely determining statistical significance.

The Mann-Kendall test can be conducted in conjunction with the Sen Slope; the former determines whether a significant trend exists and the latter indicates the magnitude of the trend.

Both versions of the Mann-Kendall test as well as the Sen Slope can be calculated with several software packages (e.g., TimeTrends software; see URL above for software). Unless it has been determined that no serial correlation exists, or if <10 years of data are available, the original Mann-Kendall test can be employed. Otherwise, use the modified version.

Poisson regression for count data

An option for evaluating trends with count data is the software package TRIM (Trends and Indices for Monitoring Data; Pannekoek and Van Strien 2005). This method is based on a Poisson distribution, which usually characterizes count data, and should not be used on data that are not Poisson distributed. The estimation method used in TRIM is based on a loglinear model using generalized estimating equations, which allows temporal autocorrelation to be taken into account. TRIM also predicts missing counts. Although the program is designed for use with monitoring programs consisting of multiple sites, it can be used for single sites; however, the resulting standard errors will be larger. TRIM will not assign missing values for single sites, although this can be done manually (e.g., Johnson and Soma 2012) before running the program. The program is designed for sites that are sampled annually, although it can be used for sites sampled at other regular intervals, although the interpretations would have to be adjusted accordingly. The latest version of TRIM software and manual can be downloaded at: <http://www.ebcc.info/trim.html>. The basic operation of TRIM is outlined below (applies to version 3.54).

Data preparation for TRIM

Each species must have its own data file. The data file must be an ASCII file containing one line (a record) for each combination of site and time. So, for I sites and J time points, the number of records is I x J. Each record (line) must contain at least three variables (the order is important!), separated by one or more spaces. The three variables in Table 1 reflect the minimum information necessary to run the analysis. Sites that were not sampled at particular times must be included as missing records. Other variables (covariates and weights) may also be inputted (see the TRIM manual for information on the use of these variables).

Site identifier (integer not to exceed 9 digits)	Time-point identifier (integer not to exceed 5 digits)	Count (integer in range of 0...2.109 or missing code of -1)
--	--	---

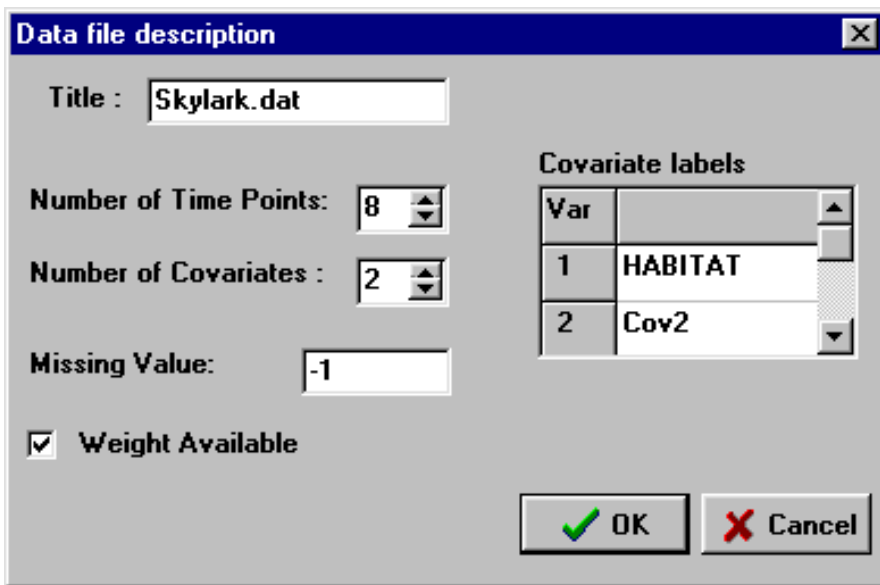
Table 1. Required format of a record within a species data file for analysis in TRIM software.

The records must be sorted by site and time so that the first J records should correspond with the J time-points for the first site and the next J records should correspond with the J time-points for another site, etc. The order of the sites is unimportant. It is only required that the records for the same site are kept together. The time-points, however, must be sorted in increasing order.

Each site has to contain at least one positive count (> 0). TRIM [v. 3.54] automatically deletes any sites with no positive counts, so this does not have to be done as part of data preparation.

Data input in TRIM

Under the “File” menu, select “New Data File”, and navigate to the appropriate data file. The “Data file description” dialog box appears (Figure 4). Input the number of time points (years). “Missing Value” should be set as “-1” if that was the code used (see above). Click “OK”. Trim produces a summary of the data (Figure 5).



The dialog box titled "Data file description" contains the following fields and controls:

- Title :** A text box containing "Skylark.dat".
- Number of Time Points:** A spin box set to "8".
- Number of Covariates :** A spin box set to "2".
- Missing Value:** A text box containing "-1".
- Weight Available:** A checked checkbox.
- Covariate labels:** A table with two rows and two columns.

Var	
1	HABITAT
2	Cov2

At the bottom right are two buttons: "OK" (with a green checkmark icon) and "Cancel" (with a red X icon).

Figure 4. Data file description dialog box for inputting data into TRIM.

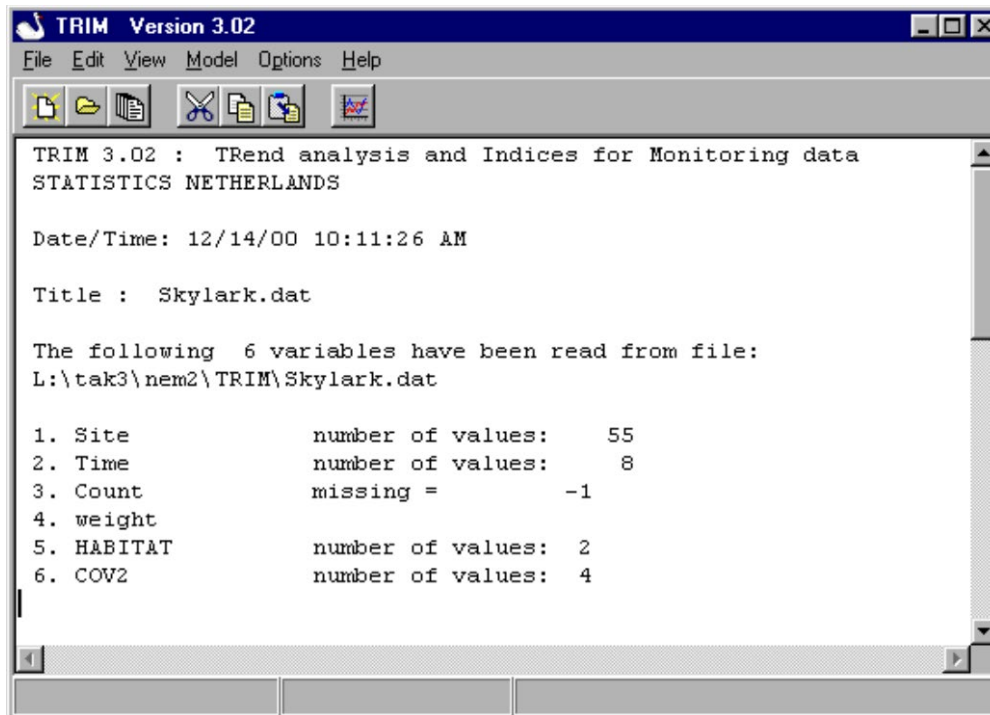


Figure 5. Main menu of TRIM with the data summary.

Modeling in TRIM

Choose the linear trend model to test whether a trend has happened across a number of years. Let TRIM search for substantial changepoints by using the stepwise procedure.

To do this, under the “Model” menu, select “Linear trend”. The “Linear trend” dialog box appears (Figure 6). Under “Changepoints”, click the “Select” button. The “Select Change Points” dialog box appears (Figure 7). Select all time points and move them to the “Change Points” window. Click “OK”. Back in the “Linear trend” dialog box, select the “Stepwise” option. Select the “Over dispersion” and “Serial correlation” options. “Base Time” should be set to first year of data series. Click “Run”.

A “Write Files” dialog box appears that allows for the creation of F and S files (see Trim Manual for explanation of these). These do not need to be created for basic calculation of trend. Click on the graph button to see a graph of the trend.

Note: Usually the first year of the time series is chosen as the base year. In the case of long time series, trends can be evaluated over more recent time periods; simply select the first year of interest as the base year.

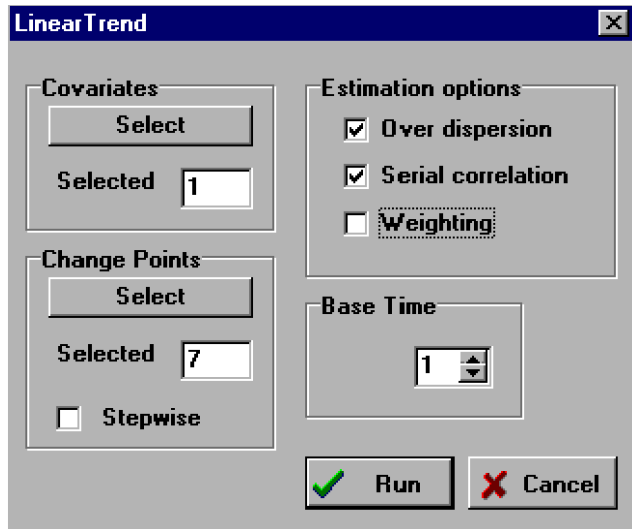


Figure 6. Linear trend dialog box in TRIM.

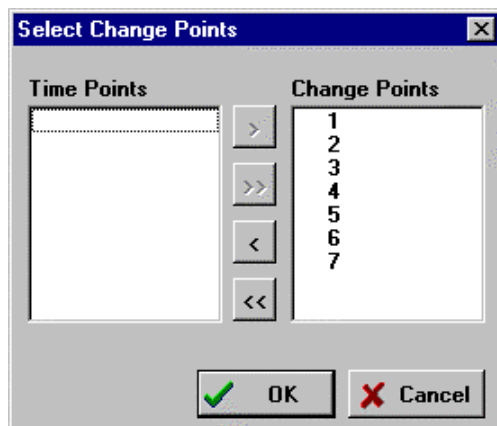


Figure 7. Selecting changepoints.

Results from TRIM

Check for “Goodness of Fit” of the model. The fit of the model is evaluated by two tests: the Chi-square test and the Likelihood Ratio test. Usually, the results of these tests are similar. If the P-value of (one of) these tests is below 0.05, the model is rejected. In case the model is rejected, one can try to find a better model by incorporating a covariate, if possible. In case of lack-of-fit, the model can still be used; TRIM converts any lack-of-fit into higher standard errors, which will make it less likely that any trend will be significant.

The overall trend is a descriptive parameter for the trend slope over the whole period studied. TRIM calculates different trends. The slope of the regression line based upon assigned indices, with intercept, should be used. This is indicated as “OVERALL SLOPE IMPUTED (recommended)” in the output. Use the multiplicative, rather than the additive slope, and associated standard error (SE). The multiplicative slope will be close to 1, the additive slope will almost always be $\ll 1$.

Interpretation of Results from TRIM

The multiplicative trend can be interpreted as the average percentage change per year. If the slope is equal to 1, there is no trend. A slope >1 indicates an increase, and a slope < 1 indicates a decline. For example, a slope of 1.08 indicates an increase of 8% per year; a slope of 0.93 indicates a decrease of 7% per year.

The TRIM manual (see Appendix C in Pannekoek and Van Strien 2005) describes how to reach an overall conclusion based on both the statistical significance and biological importance of the calculated trend. Possible qualitative interpretations include five different categories (see Table 5 and Appendices 1 and 2 in Van Strien *et al.* 2001).

TRIM (v. 3.54) reports qualitative interpretations according to statistical significance and magnitude. The multiplicative overall slope estimate in TRIM is converted into one of the following categories depending on the overall slope as well as its 95% confidence interval ($= \text{slope} \pm 1.96 * \text{the standard error of the slope}$):

Strong increase: Significant increase, more than 5% per year. Criterion: lower limit of confidence interval >1.05 .

Moderate increase: Significant increase, but not significantly more than 5% per year. Criterion: $1.00 < \text{lower limit of confidence interval} < 1.05$.

Stable: No significant increase or decline, and it is certain that trends are less than 5% per year. Criterion: confidence interval encloses 1.00 but lower limit > 0.95 and upper limit < 1.05 .

Uncertain: No significant increase or decline, but it is not certain if trends are less than 5% per year. Criterion: confidence interval encloses 1.00 but lower limit < 0.95 or upper limit > 1.05 .

Moderate decline: Significant decline, but not significantly more than 5% per year. Criterion: $0.95 < \text{upper limit of confidence interval} < 1.00$.

Steep decline: Significant decline, more than 5% per year. Criterion: upper limit of confidence interval < 0.95 .

Thus, basic reporting would include, for each species, the multiplicative trend and associated standard deviation, and the qualitative interpretation as indicated above (e.g., see Appendix 2 in Gregory *et al.* 2007 for an example)

Multiple comparisons

When conducting multiple hypothesis tests (i.e., generating multiple P values) from the same dataset, the probability of a type I error will no longer be equal to the critical value specified. The more tests that are conducted, the more likely that spurious “significant” results will be obtained. One way to correct for such multiple comparisons is the use of the sequential Bonferroni correction (Rice 1989). Other corrections are also possible. Overall, whether to correct or how to best correct for multiple comparisons is a contentious subject (e.g., Cabin and Mitchell 2000). The best reports will include

concise research hypotheses and contain as few P values as necessary, which will reduce the possibility of incorrect inferences.

Multivariate Analyses

Multivariate analysis is another frequently used analysis technique and involves methods used to explain variability in community data and to identify the environmental variables that best explain, and have an assumed responsibility for, the variability measured (Gauch 1982, Jongman *et al.* 1995, Everitt and Dunn 2001, Timm 2002). Multivariate techniques elicit a hypothesis from the biological data rather than disproving a null hypothesis. Two commonly used multivariate techniques include: ordination (such as principal components analysis, canonical correspondence analysis, and detrended correspondence analysis) and classification (such as cluster analysis). Detailed discussion of these methods can be found in several texts (Gauch 1982, Jongman *et al.* 1995, Everitt and Dunn 2001, McCune and Grace 2002, Timm 2002, Peck 2010).

Ordination

Ordination is a multivariate technique in which information from a large number of sites is represented as points in a two- or three-dimensional coordinate frame; similar sites are located near each other and dissimilar sites are farther away from each other. Ordination is useful in reducing the dimensionality of complex data sets and representing spatial or temporal relationships in a simplified visual context. Numerous ordination methods exist, although many have assumptions that may not be met with ecological data (i.e., multivariate normality), may be easily misinterpreted, or both (McCune and Grace 2002).

“Free” ordination tools are used when looking for a pattern in the response data. The two most commonly used for ecological data are Principal Components Analysis (PCA) and Nonmetric Multidimensional Scaling (NMS). PC-ORD (McCune and Grace 2002, Peck 2010) is one software option to conduct various ordination analyses.

PCA is a linear modeling technique and is most commonly used for reducing the number of habitat variables (see Sergeant *et al.* 2016) to a fewer, less redundant dataset that can then be used in other ordination procedures. For PCA, the response data set must be multivariate normally distributed and should have less than 20% of the dataset as zero values (see McCune and Grace 2002 and Peck 2010 for test assumptions). PCA results in ordination scores that help determine differences in sample units (usually sample sites or locations) based on your response variables (usually habitat parameters) and gives the percentage of variance explained by each axis in the PCA ordination.

Nonmetric multidimensional scaling (NMS; Kruskal 1964, Mather 1976) is a robust, non-parametric technique that does not have the restrictive assumptions of other ordination methods. It doesn't require the data be multivariate normal or have a unimodal form and can be employed with datasets that contain a large number of zero values. McCune and Grace (2002) describe it as “the most generally effective ordination method for ecological community data” and state that it “should be the method of choice” among ordination techniques.

NMS orders sample units in ordination space so that their interpoint distances reflect the redundant pattern of covariation in the original data (Peck 2010). NMS analyses (and ordinations in general) are considered to be primarily exploratory and descriptive, rather than confirmatory (i.e., usually used to create hypotheses prior to more in-depth data collection), and can be useful to graphically display spatial or temporal patterns in two or three dimensions.

When conducting a NMS analyses in PC-ORD (see McCune and Grace 2002 and Peck 2010) or any other software, the following general guidelines are recommended. Prior to analyses, transformations or relativizations of data may be useful in increasing the resolution of any existing patterns. For example, transformations may be helpful when there exists a high degree of variation in abundance among species. Relativizations may be useful if habitat variables are measured in different units and on scales of different magnitudes. A general relativization by columns is often useful in such a case, with the value of p chosen for the relativization set as 1, which is appropriate for the Sørensen distance measure (McCune and Grace 2002). One note of caution in using NMS analyses is that runs from different starting configurations may produce different results, resulting in unstable solutions. Thus, multiple preliminary runs should be conducted from different starting configurations (random number seeds) to ensure that they converge on a stable solution. Preliminary solutions should be obtained using autopilot mode on the “slow and thorough” setting and run at least three times with different random number seeds, to determine the appropriate dimensionality and evaluate the consistency of results (Peck 2010). The Sørensen similarity is the recommended distance measure (McCune and Grace 2002). Conduct the final NMS analysis using a large number of runs (≥ 250) with real data, at the dimensionality determined by the preliminary solutions. Select the option to automatically rotate the axes to orthogonal principal axes; this eigenanalysis procedure is a rigid rotation designed to make the axes independent of each other (Peck 2010).

The stress score for a NMS solution should always be reported, which is an indicator of how misleading the solution could be. Lower stress scores are desired. Most ecological data sets have stress between 10 and 20, and values below 15 are usually considered satisfactory (McCune and Grace 2002). The dimensionality and instability of the solution, as well as coefficients of determination (r^2), should also be reported. Because NMS is not based on partitioning of variance, the coefficients of determination do not represent variance explained by the ordination, but rather a comparison of how well the distances between points in ordination space represent distances in the original n -dimensional space, and the relative contribution of each axis (McCune and Grace 2002). The trajectory of each site may be plotted using the successional vector overlay option in PC-ORD. PerMANOVA, a distance-based nonparametric multivariate analysis of variance (Anderson 2001), may be used to determine significant differences among sites and years.

NMS can be used to calculate correlation coefficients of habitat variables with ordination axes, and overlay habitat vectors on the ordination graphs, producing joint plots. The squared correlation coefficients (i.e., coefficients of determination) express the amount of variation in position along an axis that is explained by a given habitat variable. Such correlations should primarily be used for descriptive purposes, and the use of P -values is discouraged in such cases because of lack of independence of ordination scores, and the effect of sample size on P -values (McCune and Grace

2002). Nevertheless, correlation coefficients can be interpreted in terms of *P*-values, although these must be calculated outside of PC-ORD.

NMS is a powerful and complex analysis technique, with many options. The type and number of transformations or relativizations made will affect the results, as will the selection of distance measure and dimensionality. There are no absolutely right or wrong solutions, just more or less useful ones. McCune and Grace (2002) and Peck (2010) should be studied carefully before embarking on any NMS analysis.

Guided ordination tools are used when the objective is to use additional data to guide or summarize the patterns in the response data. Weighted Average ordination (WA) uses a weighting factor for each species and multiplies it by the species abundances. Examples of possible weighting factors are species tolerance values, monetary values (for recreationally important species), or even ordination scores from other ordination techniques. The results of the WA procedure is a set of ordination scores for each sample unit that help identify which sample unit has the highest abundance of the most important species (e.g., most intolerant species, species with highest economic value).

Canonical Correspondence Analysis (CCA) is another guided ordination tool and allows for simultaneous ordering of the sample units and species by using the linear relationship between the co-occurrence of the species data and a set of explanatory variables. This procedure is a combination of reciprocal averaging and multiple linear regression, is designed for species data that does not contain more than 20% zeros, and requires a unimodal distribution of the data (normality is not required). CCA produces two sets of ordination scores (one for sample units and one for species) that indicate how species composition changes in relation to the explanatory variables and produces canonical coefficients that indicate how well the explanatory variables explain the variation in the response data. CCA may be appropriate when one is interested only in community structure that is related to measured habitat variables, and when species responses to habitat are unimodal (hump-shaped). Otherwise, NMS is preferred and will be less likely to produce misleading results (see McCune and Grace 2002).

Redundancy Analysis (RA) is similar to CCA, but uses PCA ordination rather than reciprocal averaging and requires that all response variables be normally distributed. This guided ordination tool is used mainly for habitat datasets when needing to reduce the number of environmental variables prior to use of other ordination techniques. This method should be used with caution because the underlying linear model for RA is usually not appropriate for community analyses (McCune and Grace 2002). See McCune and Grace (2002) and Peck (2010) for assumptions and details on performing guided ordination (WA, CCA, RA) analyses.

Classification

Classification analyses are used when looking for groups within the response dataset, typically to evaluate spatial or species patterns in the dataset. Ordination arranges sample units along a continuum, while classification appoints sample units (can be sample sites or species) to specific groups. Cluster Analysis (CA) and two-way Cluster Analysis are types of classification tools. CA produces a dendrogram that shows the grouping of sample units, when they were grouped, and at what level of similarity (i.e. at what level of information remains) they were combined into

groupings. In order to identify biologically significant/important groups in the data, it is necessary to apply a decision rule based on the objectives of the analysis (is it better to have few vs. many groups), determination of naturally distinct groups that make biological sense, and consideration of the level of information remaining in the heterogeneity of the dataset. Two-way CA allows for assignment of both sample units and species into groups simultaneously and is practical for use with smaller datasets. McCune and Grace (2002) and Peck (2010) should be consulted for assumptions and details on performing classification analyses.

References Cited

- Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral ecology* 26: 32-46.
- Anderson, M. J. and A. A. Thompson. 2004. Multivariate control charts for ecological and environmental monitoring. *Ecological Applications*. 14:1921-1935.
- Anscombe, F. A., 1973, *Graphs in Statistical Analysis: American Statistician*. 27: 17-21.
- Atkinson, A. J., R. N. Fisher, C. J. Rochester, and C. W. Brown. 2003. Sampling design optimization and establishment of baselines for herpetofauna arrays at the Point Loma Ecological Reserve. U.S. Geological Survey. 39 p.
- Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. EPA 841-B-99-002. Office of Water. U.S. Environmental Protection Agency. Washington, D.C. [variously paginated].
- Bart, J. and H. L. Beyer. 2012. Analysis options for estimating status and trends in long-term monitoring. *Design and Analysis of Long-term Ecological Monitoring Studies*. R. A. Gitzen, J. L. Millsaugh, A. B. Cooper and D. S. Licht. Cambridge, UK, Cambridge University Press, pp. 253-278.
- Bartlett, M. S. 1937. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London Series A* 160:268-282.
- Beauregard, M. R., R. J. Mikulak, and B. A. Olson. 1992. A practical guide to statistical quality improvement: Opening up the statistical toolbox. New York. Van Nostrand Reinhold. 469 p.
- Borcard, D. and P. Legendre. 2012. Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. *Ecology*. 93(6): 1473-1481.
- Box, G. E. P., and G. C. Tiao. 1975. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association* 70: 70-79.
- Cabin, R.J. and Mitchell, R. J. 2000. To Bonferroni or not to Bonferroni: When and how are the questions. *Bulletin of the Ecological Society of America*, 81:246-248.

- Conover, W. J. 1999. Practical Nonparametric Statistics. John Wiley & Sons, Inc., New York, NY.
- D'Angostino, R. B., A. Belanger, R. B. D'Angostino, Jr. 1990. A suggestion for using powerful and informative tests of normality. *American Statistician* 44:316-321.
- Dauwalter, D. C., E. J. Pert, and W. E. Keith. 2003. An index of biotic integrity for fish assemblages in Ozark Highland Streams of Arkansas. *Southeastern Naturalist*. 2: 447-468.
- Everitt, B. S. and G. Dunn. 2001. Applied Multivariate Data Analysis, 2nd edition. 342 p.
- Fausch, K.D., J.R. Karr, and P.R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. *Transactions of the American Fisheries Society* 113:39-55.
- Gauch Jr., H. G. 1982. Multivariate analysis in community ecology. Cambridge, U.K. Cambridge University Press. 298 p.
- Gregory R. D., Reif J., Fornasari L., Burfield I. J., Chylarecki P., Van Strien A., Vorisek P., Jiguet F. D. R., Gmelig Meyling A. W. 2007. Population trends of widespread woodland birds in Europe. *Ibis* 149: 78-97.
- Grubbs, F.E. 1950. Sample criteria for testing outlying observations. *Annals of Mathematical Statistics*. 21: 27-58.
- Gyrna, F. M. 2001. Quality planning and analysis: From product development through use. New York, McGraw-Hill Irwin. 752 p.
- Helsel, D. R. and R. M. Hirsch, 2002, Statistical Methods in Water Resources, *USGS Techniques of Water Resources Investigations, Book 4, Chapter A3*, 510 p.
- Hirsch, R. M., J. R. Slack, and R. A. Smith. 1982. Techniques for trend assessment for monthly water quality data. *Water Resources Research* 18:107-121.
- Hirsch, R. M., and J. R. Slack. 1984. A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research* 20:727-732.
- Hirsch, R. M., R. B. Alexander, and R. A. Smith, 1991, Selection of methods for the detection and estimation of trends in water quality, *Water Resources Research* 27: 803-813.
- Hlasek L. J., W. L. Fisher, and D. J. Turton. 1998. Use of the index of biotic integrity to assess water quality in forested streams of the Ouachita Mountains Ecoregion, Arkansas. *Journal of Freshwater Biology*. 13(2): 181-192.
- Hughes, R. M. and T. Oberdorff. 1998. Applications of IBI concepts and metrics to waters outside the United States and Canada, in Simon, T.P., ed., *Assessment and approaches for estimating biological integrity using fish assemblages*. Boca Raton, Florida. Lewis Press. p. 79-93.

- Johnson, D. H. 2012. Monitoring that matters. Design and Analysis of Long-term Ecological Monitoring Studies. R. A. Gitzen, J. L. Millspaugh, A. B. Cooper and D. S. Licht. Cambridge, UK, Cambridge University Press, pp. 54-72.
- Johnson D. H. and M.B. Soma. 2012. The treatment of missing data in long-term monitoring programs. Pages 298-312 in Gitzen RA, Millspaugh J. J., Cooper A. B., and Licht D. S., eds. *Design and Analysis of Long-term Ecological Monitoring Studies*. Cambridge University Press.
- Jongman, R. H. G., C. J. F. ter Braak, and O. F. R. van Tongeren. 1995. Data analysis in community and landscape ecology. Cambridge, Cambridge University Press. 299 p.
- Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries. 6: 21–27.
- Kendall, M. G. 1975. Rank correlation methods, 4th edition. Charles Griffin, London.
- Kruskal, J. B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29: 1-27.
- Kutner, M. H, C. J. Nachtsheim, J. Neter, and W. Li. 2005. Applied linear statistical models, 5th edition. McGraw-Hill/Irwin, New York, NY.
- Levene, H. 1960. Robust tests for equality of variances. Pages 278-292 in I. Olkin, editor. Contributions to probability and statistics. Stanford University Press, Palo Alto, California.
- Lyons, J., L. Wang, and T.D. Simonson. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. North American Journal of Fisheries Management 16:241-256.
- Manly, B. F. J. 2001. Statistics for environmental science and management. Boca Raton, Florida. Chapman & Hall/CRC. 326 p.
- Mann, H. B. 1945. Non-parametric tests against trend. Econometrica 13:163-171.
- Mardia, K. V. 1971. The effect of nonnormality on some multivariate tests and robustness to nonnormality in the linear model. Biometrika 58: 105-121.
- Mather, P. M. 1976. Computational methods of multivariate analysis in physical geography. London, J. Wiley and Sons.
- Mauchly, J. W. 1940. Significance test for sphericity of a normal n-variate distribution. The Annals of Mathematical Statistics 11: 204-209.
- McBean, E. A. and F. A. Rovers. 1998. Statistical procedures for analysis of environmental monitoring data and risk assessment. Upper Saddle River, New Jersey. Prentice Hall PTR. 320 p.

- McCune, B. and J. B. Grace. 2002. Analysis of ecological communities. MjM Software Design, Oregon.
- Montgomery, D. C. 2001. Introduction to statistical quality control. New York, John Wiley & Sons. Inc. 776 p.
- Olson, C. L. 1974. Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association. 69: 894-908.
- Pannekoek J, Van Strien AJ. 2005. TRIM 3 Manual (TRENDS and Indices for Monitoring data). Statistics Netherlands. Technical report.
- Peck, J. E. 2010. Multivariate analysis for community ecologists: step-by-step using PC-ORD. MjM Software Design, Gleneden Beach, Oregon.
- Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution, 43: 223-225.
- Rosner, B. 1983. Percentage points for a generalized ESD many-outlier procedure. Technometrics. 25(2): 165-172.
- Savin, N. E. and K. J. White. 1977. The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors. Econometrica 45(8):1989-1996.
- Sergeant, C. J., E. N. Starkey, K. K. Bartz, M. H. Wilson, and F. J. Mueter. 2016. A practitioner's guide for exploring water quality patterns using principal components analysis and Procrustes. Environmental Monitoring and Assessment. 188: 249, 15p.
- Shapiro, S. S. and M. B. Wilk. 1965. An analysis of variance test for normality. Biometrika 52: 591-599.
- Simon T.P. 1999. Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities. CRC Press, Inc., Boca Raton, FL.
- Smogor, R. 2005. Draft manual for interpreting Illinois fish IBI scores. Illinois Environmental Protection Agency, Bureau of Water, Surface Water Section.
- Taylor, L. R. 1961. Aggregation, variance, and the mean. Nature. 189: 732-735.
- Timm, N. H. 2002. Applied multivariate analysis. University of Pittsburgh Press. 693 p.
- Van Strien AJ, Pannekoek J, Gibbons DW. 2001. Indexing European bird population trends using results of national monitoring schemes: a trial of a new method. Bird Study 48: 200-213.
- Wetherill, G. B., D. W. Brown. 1991. Statistical Process Control - Theory and Practice, Chapman & Hall, London.
- Yoder, C.O., and M.A. Smith. 1999. Using fish assemblages in a state biological assessment and criteria program: essential concepts and considerations. Pages 17-56 in T.P. Simon, editor.

Assessing the sustainability and biological integrity of water resources using fish communities.
CRC Press, Inc., Boca Raton, FL.