Nutrient STEPS

Change-Point Analysis

N-STEPS Objectives

Provide regions, states, and tribes with support related to nutrient criteria development

■

Provide access to expert assistance with issues related to nutrient criteria development and implementation

■

Improve communication nationwide.

What Is It?

Change-point analysis is a method for identifying thresholds in relationships between two variables. More specifically, it is an analytical method that attempts to find a point along a distribution of values where the characteristics of the values before and after the point are different. In the case of a stressor-response relationship, change-point analysis can be used to identify that point along the x axis where characteristics along the y axis change – implying a shift in the average of variance or a change in slope.

Example Question: Is there a threshold in the response of the number of EPT taxa to gradients in total phosphorus?

How is it Applied to Nutrient Criteria Development?

Nutrient criteria development involves three main processes: identifying relationships between biological responses and nutrient stressors, examining these relationships, and establishing nutrient and/or biological thresholds or criteria.

Change-point analysis is a statistical method for identifying thresholds and it is essential for nutrient criteria development. More specifically, it is an analytical method that attempts to find a point along a distribution of values where the characteristics of the values before and after the point are different. In the case of algal biomass' response to total phosphorus (TP) concentrations, change-point analysis can be used to identify that TP concentration where average algal biomass shifts significantly before and after that TP concentration.

How Does It Work?

There are a few methods change-point analysis. The basic method uses deviance reduction. This approach finds the point along a distribution of points where the sum of deviances on either side of the point is lowest compared to the overall dataset deviance. The percent of error reduction associated with splitting the data is then calculated. This is an iterative process, moving along the data systematically and evaluating that point which minimizes the deviance reduction. This is essentially equivalent to a regression tree with one split and that method can be used with one predictor, limiting the process to one split in the data. However, users can select a variety of approaches to use including: non-parametric deviance reduction, least squares, cumulative summing, Bayesian estimation, etc. Bootstrap methods can be applied to estimate error around the identified threshold for non-bayesian methods.

Data Requirements

Independently collected data in the form of paired observations. These are typically continuous numeric data, although discrete numeric and categorical data can also be used. As with correlation and regression, the greater the range of environmental conditions encompassed the better.

What Should You Look For & Report?

One is looking for thresholds that result in substantial reduction in the percent error. So most techniques include some measure of error reduction and this should be reported. Some software applications use a χ^2 test to evaluate the significance of the error reduction. Confidence in the threshold should also be reported. Bayesian approaches will produce a range for the thresholds, whereas other non-parametric approaches can produce an estimate of confidence in the threshold using bootstrap re-sampling.

Pros

- Identifies clear thresholds
- Non-parametric techniques are distribution free and have limited assumptions
- Confidence intervals can be used to set criteria

Cons

- Data hungry and computationally intensive
- No ready formula is produced, so it is hard to transport the results
- Sensitive to variability in data

Alternatives

Few – thresholds analysis is not easy or well developed
Conditional probability analysis
Visual methods – loess and straight scatterplots

Citations

Qian et al. 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modelling 166:87-97