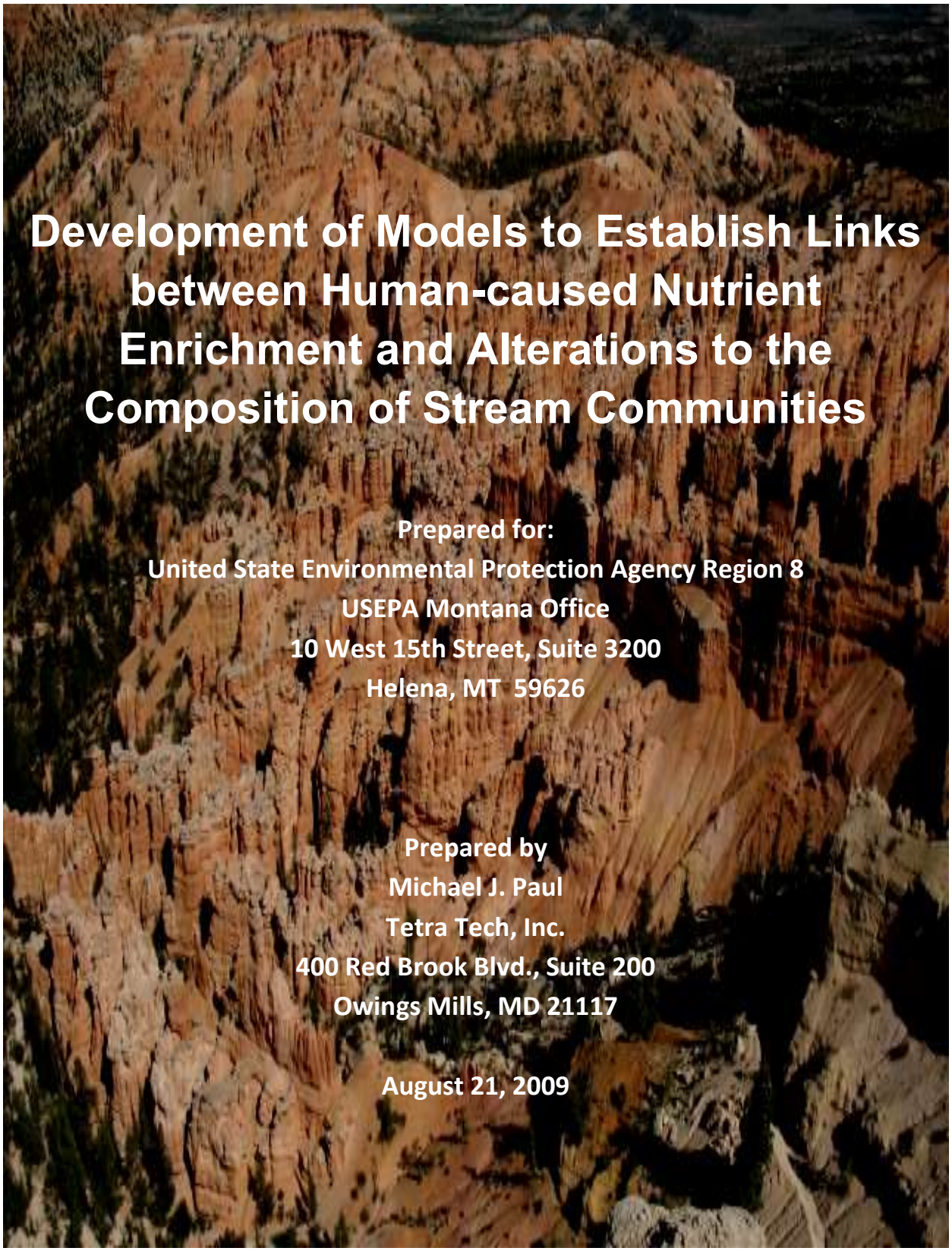


Development of Models to Establish Links between Human-caused Nutrient Enrichment and Alterations to the Composition of Stream Communities

Prepared for:
United State Environmental Protection Agency Region 8
USEPA Montana Office
10 West 15th Street, Suite 3200
Helena, MT 59626

Prepared by
Michael J. Paul
Tetra Tech, Inc.
400 Red Brook Blvd., Suite 200
Owings Mills, MD 21117

August 21, 2009

(This Page Intentionally Left Blank)

Acknowledgments

The author acknowledges the input and help of Jeff Ostermiller of UDWQ whose helping in designing and guiding the analysis was critical. Input and guidance was also provided by Tina Laidlaw of USEPA. Lastly, several colleagues at Tt including Lei Zheng, Jeff White, Erik Leppo, and Jeroen Gerritsen provided important support. Cover photo credits: M. Baker, Logan, UT

Acronyms

AWCH_AVE – Soil Water Capacity
CHL *a* – Chlorophyll *a*
CO Plateaus – Colorado Plateaus
CV - Coefficient of Variation
DNOx – Dissolved Nitrate/Nitrite
DP – Dissolved Phosphorus
Elev_WS – Watershed Elevation
E – Ephemeroptera Richness
EPT – Ephemeroptera, Plecoptera, Trichoptera
EPTPct – Percent EPT individuals
G_NU_AVE – Mean Reactive Geology
HBI - Hilsenhoff Biotic Index
HIGH_DP – High Dissolved P Class
Lat_Dec – Latitude in Decimal Degrees
LogAnnNH4 – Log_{10}Annual Ammonium
LogAnnNO3– Log_{10}Annual Total Nitrate
LogAnnDNO3 – Log_{10}Annual Dissolved Nitrate
LogAnnTP – Log_{10}Annual Total Phosphorus
LogAnnDP – Log_{10}Annual Dissolve Phosphorus
Log_Sq_km – Log_{10}Watershed Area
Long_Dec – Longitude in Decimal Degrees
LOWESS – Locally Weighted Scatterplot Smoothing
N – Nitrogen
NH4 –Ammonium
NH4TOT – Total Ammonium
NOXDISS – Dissolved Nitrate/Nitrite
NOXTOT – Total Nitrate/Nitrite
O/E – Observed to Expected taxa richness
P – Phosphorus
PCA – Principal Components Analysis
PDISS – Dissolved Phosphorus
PTOT – Total Phosphorus
PRED_ALK – Predicted Alkalinity
PRED_COND – Predicted Conductivity
PTOT – Total Phosphorus
REF – Reference Sites
RefState_JO – Reference State
RFOE – O/E score
SCR – Scraper Richness
SCRCG - Scraper-Collector Gatherer Richness
S. Rockies – Southern Rockies
Taxa – Taxa Richness
TKN – Total Kjeldahl Nitrogen
TN – Total Nitrogen
TNOx – Total Nitrate/Nitrite
TP – Total Phosphorus
UDWQ - Utah Department of Water Quality
USEPA – United States Environmental Protection Agency
UT - Utah

Table of Contents

Page

Acknowledgments iii

Acronyms iv

Table of Contents v

List of Figures vii

Background 1

Analytical Approach 2

Classification 3

Nutrient Analysis 9

Summary 24

Citations 26

Appendix 1 – Interpolated endpoints based on existing thresholds figures: annual data 29

Appendix 2 – Interpolated endpoints based on existing thresholds figures: summer data 37

Appendix 3 - Endpoints based on change-point analyses: annual data 44

Appendix 4 - Endpoints based on change-point analyses: summer data 47

List of Tables

Table 1- Coefficients of variation for all sites and just reference sites (annual and summer data) for various nutrient parameters. CVs were calculated for all data combined and then split into level III ecoregions. SA few sites were in unknown ecoregions because of missing latitude and longitude values. Means and standard deviations were calculated on log-transformed data..... 5

Table 2 – Predictors generated by Utah State used in comparing nutrient distributions and in multivariate models. 5

Table 3 – Seasonal 75th and 25th percentiles of nutrient variable medians for reference and all site populations, respectively. Also included are the medians of seasonal quartile values as annual values. All units in mg/L. 11

Table 4 – List of biological response and nutrient stressor variables that were the focus of analysis based on initial correlation screening. *Ephemeroptera richness did not show significant correlations for the summer data alone, but other response variables did. 13

Table 5 – Table of distributional statistics of biological metrics evaluated in stressor-response relationship models. Values highlighted in blue are quartiles typically used in identifying thresholds between reference and non-reference conditions in developing biological indices (see text). These values were used as thresholds for linear regression interpolation models for nutrient endpoint development. 14

Table 6 – Table of interpolated nutrient values associated with biological thresholds existing (O/E) or calculated based on reference quartiles (see text). Values are given for total nitrate (NOXTOT), dissolved nitrate (NOXDISS), and total phosphorus (PTOT). All values in mg/L. Endpoints are shown based on annual values and summer values. Abbreviations for biological responses are given in Table 4. 15

Table 7 – Change-points associated with the response of various biological responses to nutrient enrichment. All values in mg/L. 17

List of Figures

Figure 1 - Simplified diagram illustrating the causal pathway between nutrients and aquatic life use impacts. Nutrients enrich both plant/algal as well as microbial assemblages, which lead to changes in the physical/chemical habitat and food quality of streams. These effects directly impact the insect and fish assemblages. The effects of nutrients are influenced by a number of other confounding factors as well, such as light, flow, and temperature. 2

Figure 2 – Box and whisker plot of Log(Nitrate/Nitrite) annual values within different level III ecoregions. Units are mg/L..... 3

Figure 3 - Box and whisker plot of Log(Total Phosphorus) annual values within different level III ecoregions. Units are mg/L. 4

Figure 4 – Box and whisker plots of distributions of different predictor variable values between low dissolved P (N/A) and high dissolved P (Yes). This plot indicate no difference between the two nutrient regimes for latitude (Lat_Dec), Longitude (Long_Dec), soil water capacity (AWCH_AVE), and predicted alkalinity (PRED_ALK). 6

Figure 5 - Box and whisker plots of distributions of different predictor variable values between low dissolved P (N/A) and high dissolved P (Yes). This plot indicate no difference between the two nutrient regimes for mean reactive geology (G_NU_AVE), watershed areas (Log_Sq_Km), elevation (Elev_WS), and min:max flow rate ration (Hydr_Ws). 7

Figure 6 – Plot of axes 1 and 2 from a principal components analysis of predictors in Table 2. Sites are coded according to the two cluster groups generated from a cluster analysis of the same predictors. Separation along axis 1 is evident and due to differences in geology among these two classes..... 8

Figure 7 – Box and whisker plot of standardized mean annual dissolved phosphorus (logAnnDP) concentration distributions between the two cluster groups [1 and 2 (labeled 15)]. 9

Figure 8 – Box and whisker plots of seasonal total nitrate site medians for reference and all site populations. Also shown are the 75th and 25th percentiles of these populations, respectively. 10

Figure 9 - Box and whisker plots of seasonal total phosphorus site medians for reference and all site populations. Also shown are the 75th and 25th percentiles of these populations, respectively. 11

Figure 10 Plot showing two stressor-response relationships, where the response is a direct measure of designated use or can be easily linked to a designated use measure. In relationship B, a preexisting threshold of the response variable exists. For example, UDWQ O/E score thresholds (e.g., 0.74) or minimum dissolved oxygen criteria (3 to 8 mg/L depending on use class) could be the response variable and the corresponding criterion value for each used as the pre-designated threshold. The hatched line from the response threshold to the regression line is extrapolated downward to estimate the nutrient concentration associated with that threshold (b). In relationship A, no pre-existing threshold in the

response variable exists; however the response is non-linear and changes once the concentration (a) is reached. This changepoint becomes a candidate criterion. 12

Figure 11 – Simple linear regression scatterplot of log(total nitrate)(*NOXTOT) in mg/L versus taxa richness (Taxa). The solid blue lines indicate the value of total nitrate associated with the 75th percentile reference value for taxa richness (20). Also shown on the plot are the regression equation, significance, and coefficient of determination. 14

Figure 12 - Plot of HBI vs. Log_{10}(total nitrate), as well as output from a change-point analysis of HBI and Log_{10}(total nitrate). The change-point at Log_{10}(total nitrate) of -0.395 (0.403 mg/L) resulted in a 24% reduction in model error. Units in mg/L. 16

Figure 13 – Plot showing a stressor-response relationship, where the response is a direct measure of designated use or can be easily linked to a designated use measure. In this relationship A, no pre-existing threshold in the response variable exists; however the response is non-linear and the response declines consistently with any increase in stressor level observed up to point (c). 17

Figure 14 – Conceptual model of serial regressions linking nutrients and potential co-variates with macroinvertebrate response. 18

Figure 15 - Conceptual diagram of propensity score analysis. Covariates with nutrient concentration are used to generate a predicted nutrient concentration (propensity score). Sites with similar propensity scores are grouped together into bins representing similar covariate distributions. The nutrient-response relationship is then investigated per bin. 20

Figure 16 – Response of taxa richness across different propensity score bins. 21

Figure 17 – Scatterplots of total nitrate vs. taxa richness within each of the 6 propensity score bins. Nitrate concentration increased from bin 1 to 6. 22

Figure 18 - Scatterplots of total nitrate vs. EPT richness within each of the 6 propensity score bins. Nitrate concentration increased from bin 1 to 6. 23

Figure 19 - Scatterplots of total nitrate vs. EPT richness within each of the 6 propensity score bins. Nitrate concentration increased from bin 1 to 6. 24

Background

Eutrophication of the nation's surface water resources (lakes, reservoirs, rivers, streams and wetlands) due to excessive nutrients (nitrogen and phosphorus) is recognized as a significant water quality problem. The U.S. Environmental Protection Agency's (USEPA) national water quality summary reports to Congress consistently identify excessive nutrients as one of the top three leading causes of impairments of the nation's waters (along with siltation and pathogens). Natural nutrient inputs support primary producers that are essential for supporting healthy, diverse and productive ecosystems. However, excessive nutrient inputs can result in abundant growth of periphyton (diatoms and multicellular algae), macrophytes and/or phytoplankton leading to oxygen depletion, potentially toxic algal blooms, imbalances in biological community composition, an human health concerns.

As a result of the threat posed by eutrophication, USEPA published the *National Strategy for the Development of Regional Nutrient Criteria* (USEPA 1998) under section 304(a) of the Clean Water Act (66 *Federal Register* [FR] 1671). This document describes the basis and national approach for working with States in the development and adoption of nutrient criteria. Following this approach, USEPA proposed criteria for rivers, streams, lakes and reservoirs in a series of Water Quality Criteria Recommendation documents (USEPA 2000a, 2000b, 2000c, 2001a, 2001b, 2001c, 2001d, 2001e). These proposed criteria were based on nutrient ecoregions and were intended to serve as a springboard for States to develop more refined nutrient criteria, stating that "States and Tribes need to identify with greater precision the nutrient levels that protect aquatic life and recreational uses…" (USEPA 2000d, 2000e). USEPA expected States to adopt or revise the proposed nutrient ecoregion criteria into State water quality standards (WQS) by 2004 while recognizing that the time needed to develop standards and the available resources differed significantly among States. This analysis is part of Utah's effort to develop protective and scientifically defensible nutrient criteria.

It is our understanding that Utah is considering a variety of technical approaches for criteria development. This specific summary was focused on statistical analyses with existing UT ambient monitoring data to identify stressor-response relationships of utility in developing endpoints that support the criteria development effort. This analysis focuses in particular on the available aquatic life use response variables based on macroinvertebrates that are used by the state in making aquatic life use decisions. It is valuable, at this point, to summarize the conceptual model that links nutrient enrichment to effects on macroinvertebrates (Fig. 1).

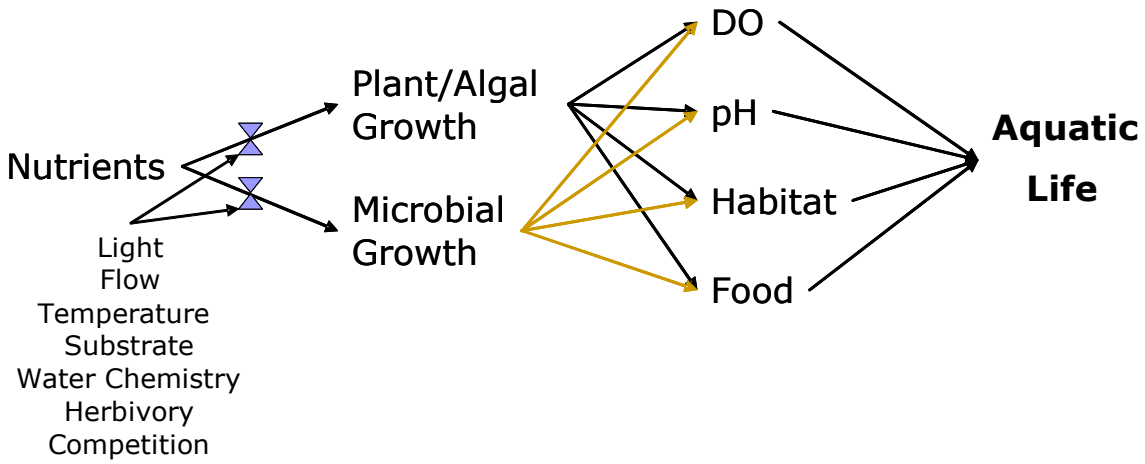


Figure 1 - Simplified diagram illustrating the causal pathway between nutrients and aquatic life use impacts. Nutrients enrich both plant/algal as well as microbial assemblages, which lead to changes in the physical/chemical habitat and food quality of streams. These effects directly impact the insect and fish assemblages. The effects of nutrients are influenced by a number of other confounding factors as well, such as light, flow, and temperature.

The experimental and other support for each of the relationships identified in this diagram can be found in freshwater texts (e.g., Allan 1995), summary papers (e.g., Carpenter et al. 1998), as well as an abundance of primary scientific literature (e.g., Elwood et al. 1981, Peterson et al. 1993, Rosemond et al. 1993, Feminella and Hawkins 1995, Hillebrand 2002, King and Richardson 2003, Slavik et al. 2004, Cross et al. 2006). There is, therefore, a well defined and scientific basis for establishing linkages between nutrients and responses along the causal pathway in Figure 1 up to and including aquatic life.

This report is focused on macroinvertebrate response variables since, to our knowledge, these represent the largest, available set of response data for use in developing stressor-response relationships linking nutrient enrichment to aquatic life use response in Utah. This is not a detailed report, rather it summarizes the general approaches, methods, and results of this preliminary analysis.

Analytical Approach

The general approach employed is consistent with EPA guidance on establishing nutrient criteria for rivers and streams (USEPA 2000d). The first step involved exploring potential classification to reduce natural variability in nutrient conditions resulting from geographic factors. After exploring classification, the second step was to explore distributional statistics of nutrients and to develop statistical stressor-response relationship models to identify candidate endpoints. We finish by reviewing analyses to strengthen support for the relationship between nutrients and aquatic life use response. Methods for each are described in order.

Classification

We worked with scientists at Utah Department of Water Quality (UDWQ) to develop a strategy for exploring potential natural classes of nutrient behavior for streams in Utah for which nutrient data were available. This process involved looking at nutrient distributions among a priori defined classifications (ecoregions) and exploring predictors of differences in nutrient concentrations within reference sites using a range of multivariate and discriminant models.

Using all the data we had in our dataset relevant to nutrients in Utah, we first examined differences in nutrient distributions by ecoregion. There was little difference among ecoregions for nitrate concentration (Fig 2) or total phosphorus concentration (Fig. 3). Non-parametric means comparison tests also indicate no significant differences among level III ecoregions.

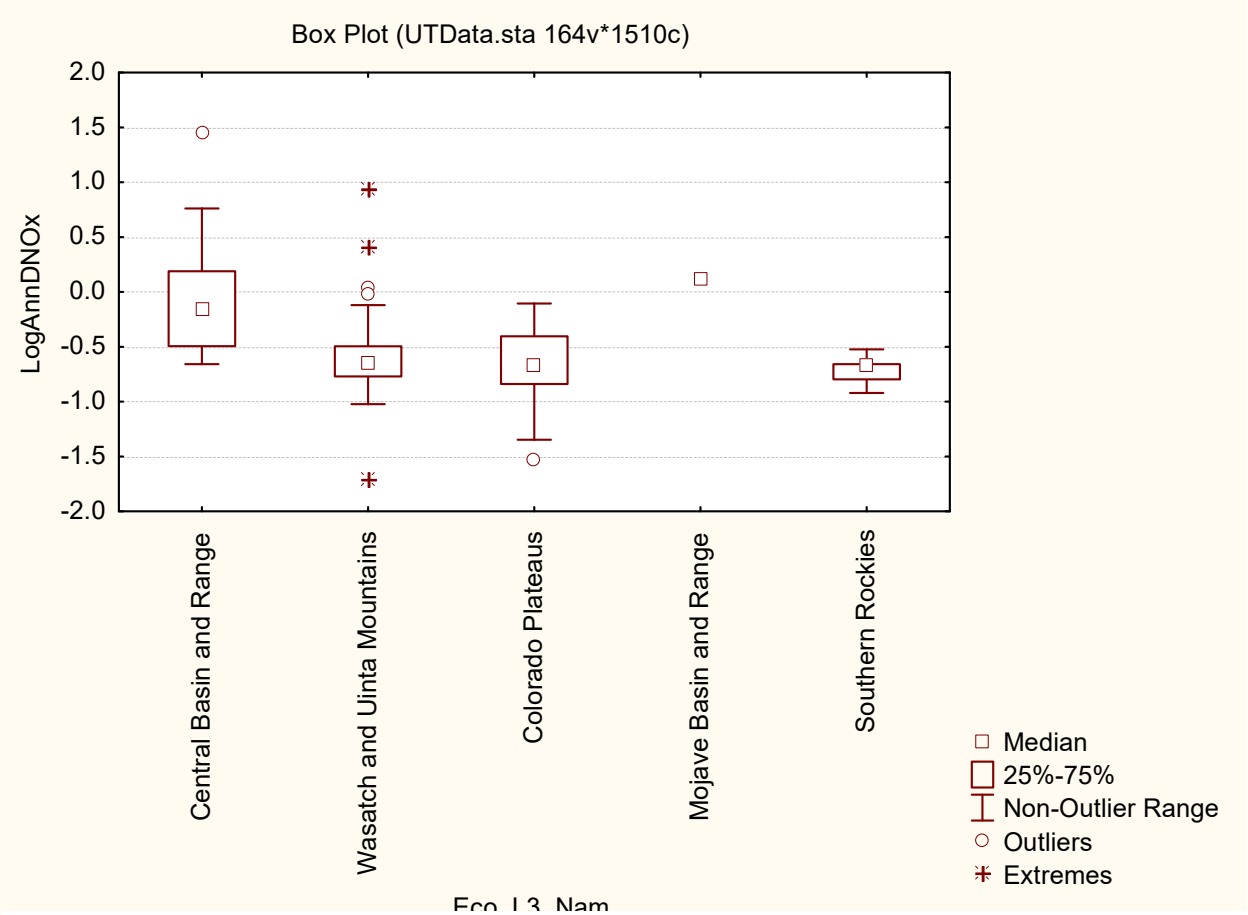


Figure 2 – Box and whisker plot of Log(Nitrate/Nitrite) annual values within different level III ecoregions. Units are mg/L.

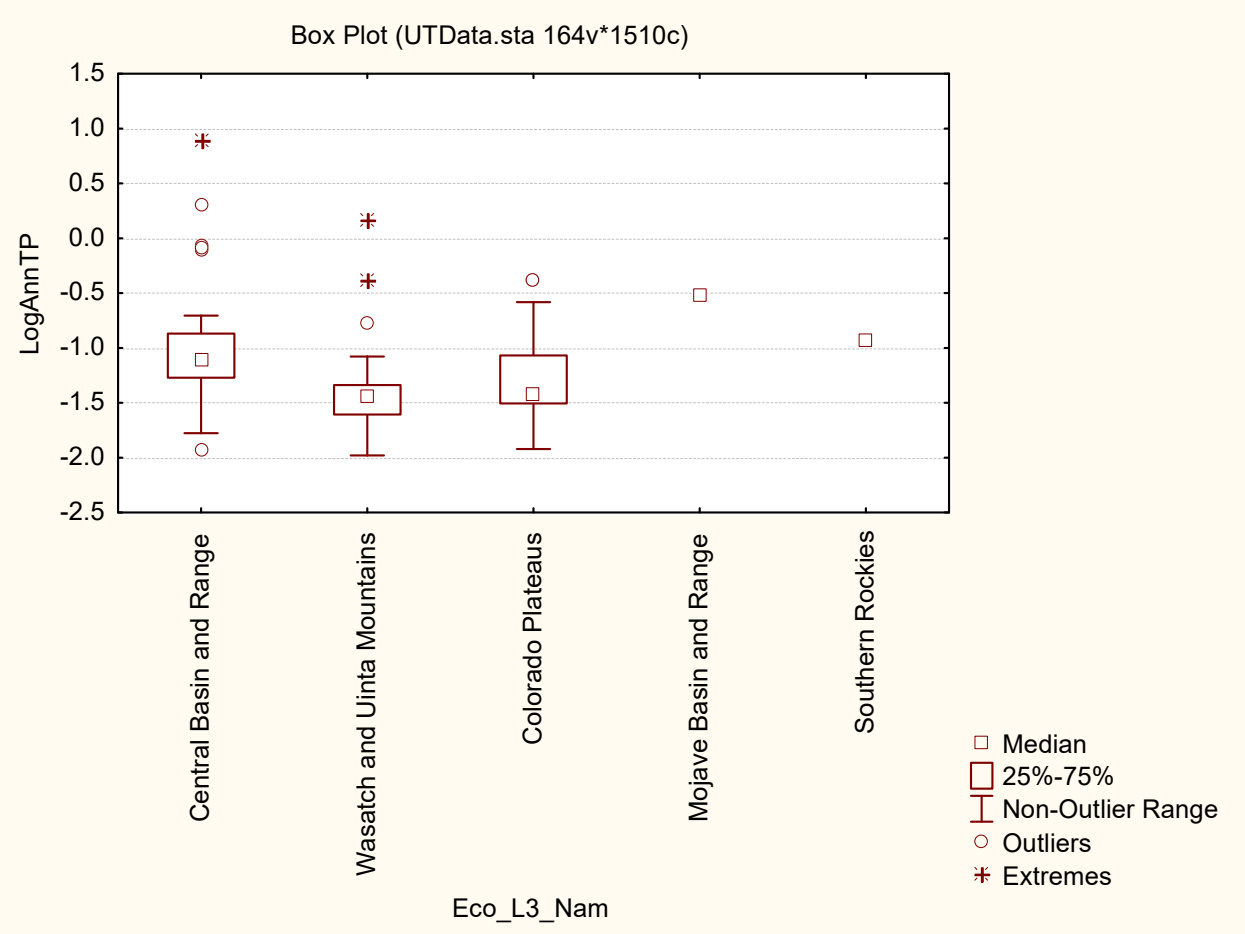


Figure 3 - Box and whisker plot of Log(Total Phosphorus) annual values within different level III ecoregions. Units are mg/L.

We also examined reductions in coefficients of variation (CV) within level III ecoregions in comparison to all regions combined to examine whether a substantial reduction in variability could be achieved by classification. There was variation in CV across ecoregions with some higher and some lower than the combined regions CV (Table 1). In general, there was little in the CV reduction information, especially among reference sites, to recommend the use of specific classes within Utah that would not result in increased variability among some resulting classes. Any improvement would be marginal, at best, based on these data.

The next set of analyses was based on standardizations of nutrient concentrations in reference sites. Annual and seasonal nutrient mean concentrations were first identified as High P or High N if they were greater than the 75th percentile of reference site nutrient distributions. We then looked at the difference in nutrient distributions between the two classes of reference streams – those that were high nutrients versus those that were typically low – based on a wide range of natural predictors (Table 2).

Table 1- Coefficients of variation for all sites and just reference sites (annual and summer data) for various nutrient parameters. CVs were calculated for all data combined and then split into level III ecoregions. SA few sites were in unknown ecoregions because of missing latitude and longitude values. Means and standard deviations were calculated on log-transformed data.

	Combined	Unknown	Central Basin	Wasatch	CO Plateaus	S. Rockies
All Sites						
LogAnnNH4	**-41.7**	**-43.1**	**-44.3**	**-29.3**	**-15.8**	**-4**
LogAnnNOx	**-98.8**	**-99.8**	**-520.9**	**-42.3**	**-44.7**	**-16.3**
LogAnnDNOx	**-109.4**	**-112.9**	**-586.7**	**-56.7**	**-47.7**	**-21.4**
LogAnnTP	**-40.5**	**-41**	**-65.2**	**-21.8**	**-32**	
LogAnnDP	**-36.5**	**-36.4**	**-64.7**	**-26.3**	**-27.5**	
Reference Sites						
LogAnnNH4	**-34.5**		**-38.5**	**-47.5**	**-13.3**	**-2.1**
LogAnnNOx	**-43.1**	**-77.8**	**-49.4**	**-20.1**	**-40.3**	
LogAnnDNOx	**-37.7**	**-6.7**	**-9.4**	**-41.5**	**-39.9**	**-20.7**
LogAnnTP	**-19.5**	**-13.6**	**-25.7**	**-10**	**-27.8**	
LogAnnDP	**-18.6**	**-9.4**		**-15.1**	**-31.7**	
Reference Sites						
LogSuNH4	**150.2**	**141.4**		**125.6**	**223.6**	
LogSuTNOx	**17.6**	**24.4**	**22.5**	**8.4**	**9.8**	
LogSuDNOx	**60.3**	**141.4**		**44.7**	**88.7**	
LogSuTP	**45.4**	**31.3**	**0**	**44**	**56.4**	
LogSuDP	**118.8**	**0**		**72.6**	**200**	

Table 2 – Predictors generated by Utah State used in comparing nutrient distributions and in multivariate models.

Physical	Temperature	Soils	Geochemistry	Hydrology
Latitude	Earliest Freeze Date	Soil water capacity	Predicted conductivity	Min:Max flow rate
Longitude	Last Freeze Date	Soil permeability	Predicted alkalinity	Mean precipitation
Elevation	Mean Air Temperature of Stream Network		Calcium content of rocks	
Watershed Area	Mean Air Temperature of Watershed		Percent carbonate geology	
Relief Ratio			Mean reactive geology	
			Percent volcanic	

geology

We examined box plots of standardized nutrient distributions between high nutrient and low nutrient reference sites for every nutrient parameter (e.g., Figs. 4 and 5). The examples shown in Figs. 4 and 5 were for dissolved phosphorus, but all the nutrient parameters showed similar results, no nutrient exhibited different distributions of potential predictors between high nutrient and low nutrient reference classes.

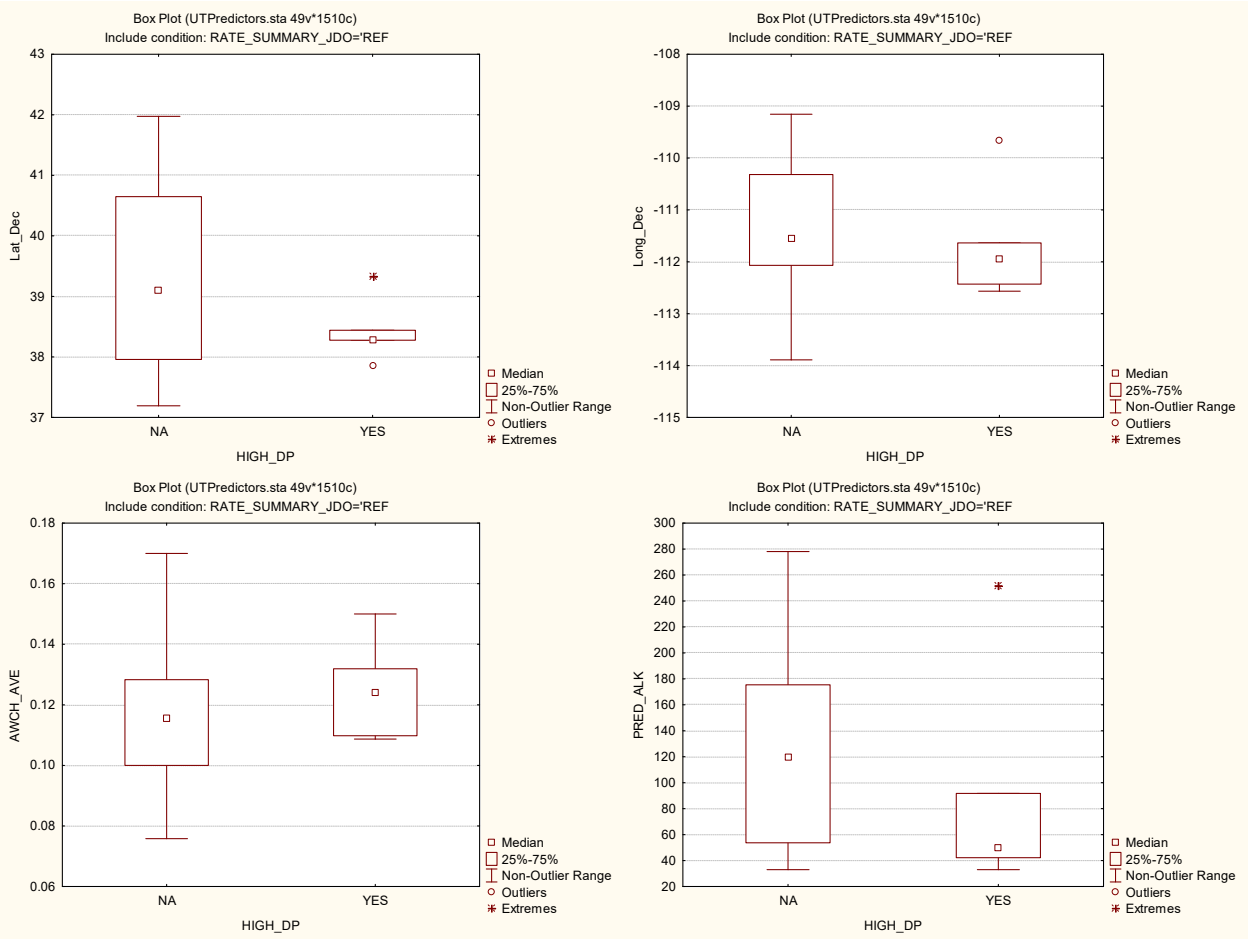


Figure 4 – Box and whisker plots of distributions of different predictor variable values between low dissolved P (N/A) and high dissolved P (Yes). This plot indicate no difference between the two nutrient regimes for latitude (Lat_Dec), Longitude (Long_Dec), soil water capacity (AWCH_AVE), and predicted alkalinity (PRED_ALK).

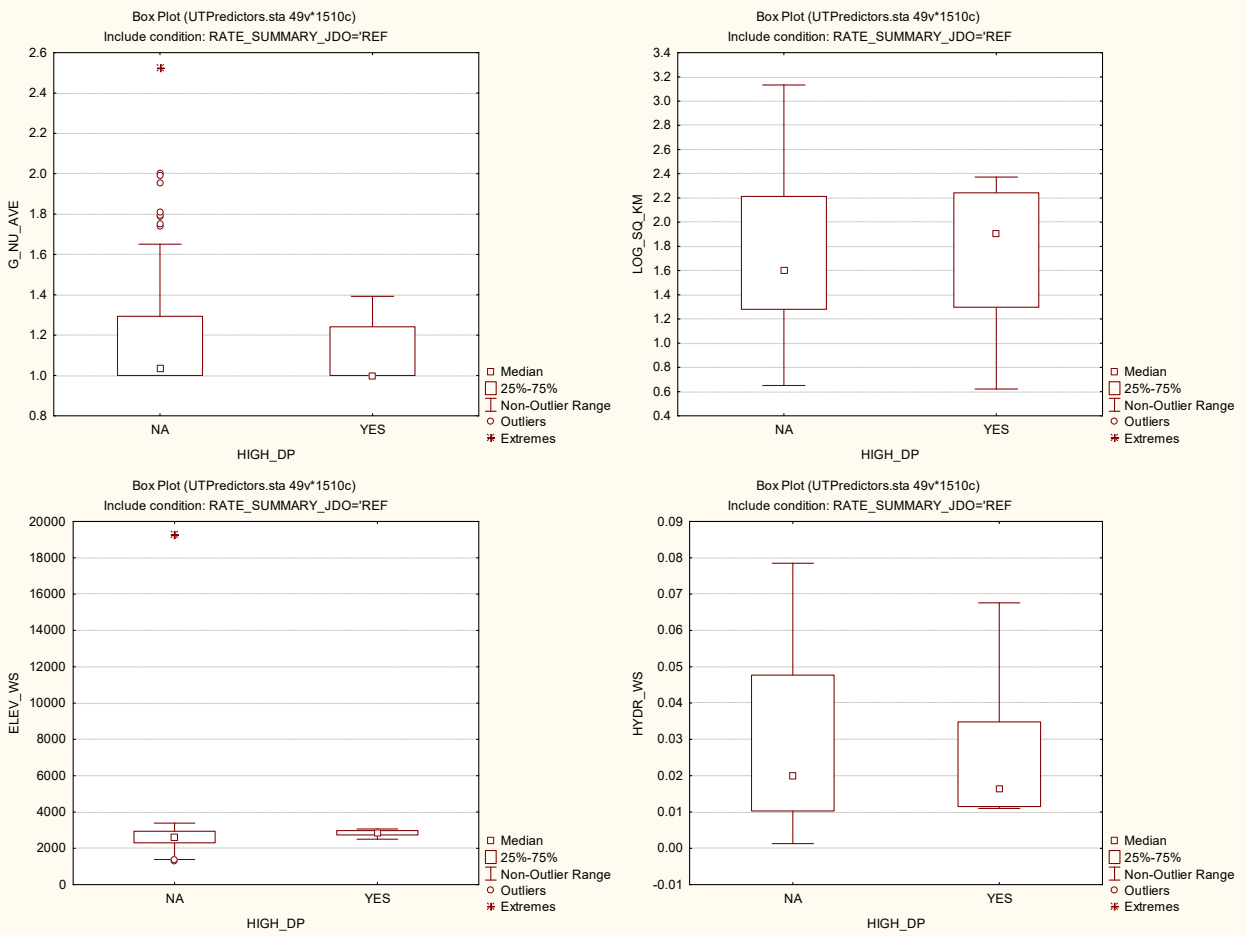


Figure 5 - Box and whisker plots of distributions of different predictor variable values between low dissolved P (N/A) and high dissolved P (Yes). This plot indicate no difference between the two nutrient regimes for mean reactive geology (G_NU_AVE), watershed areas (Log_Sq_Km), elevation (Elev_WS), and min:max flow rate ration (Hydr_Ws).

We also ran the same analysis using reference quantiles. This was performed by coding reference sites based on the quantile of reference nutrient concentrations into which they fell and attempting to identify predictors that discriminated among the 4 quantile classes. Again, there were no natural predictors that discriminated among reference nutrient classes consistently. We then ran discriminant analysis with high and low nutrient classes and reference quantiles as groups and the list of variables in Table 2 as predictors. Once again, no significant model could be developed that adequately discriminated among reference site nutrient groups.

Finally, we ran cluster analysis on the predictors in Table 2 for reference sites. We developed 2, 3, and 4 group models, but focused on the 2 group models because they showed the most promise, and this, only marginally. After defining the 2 groups with the cluster analysis, we then ran the same predictors in PCA and coded the sites by cluster group. The two clusters showed some separation with regards to the two clusters (Figure 6), and this was principally along a geologic attribute axis, reflected in predicted alkalinities and conductivities.

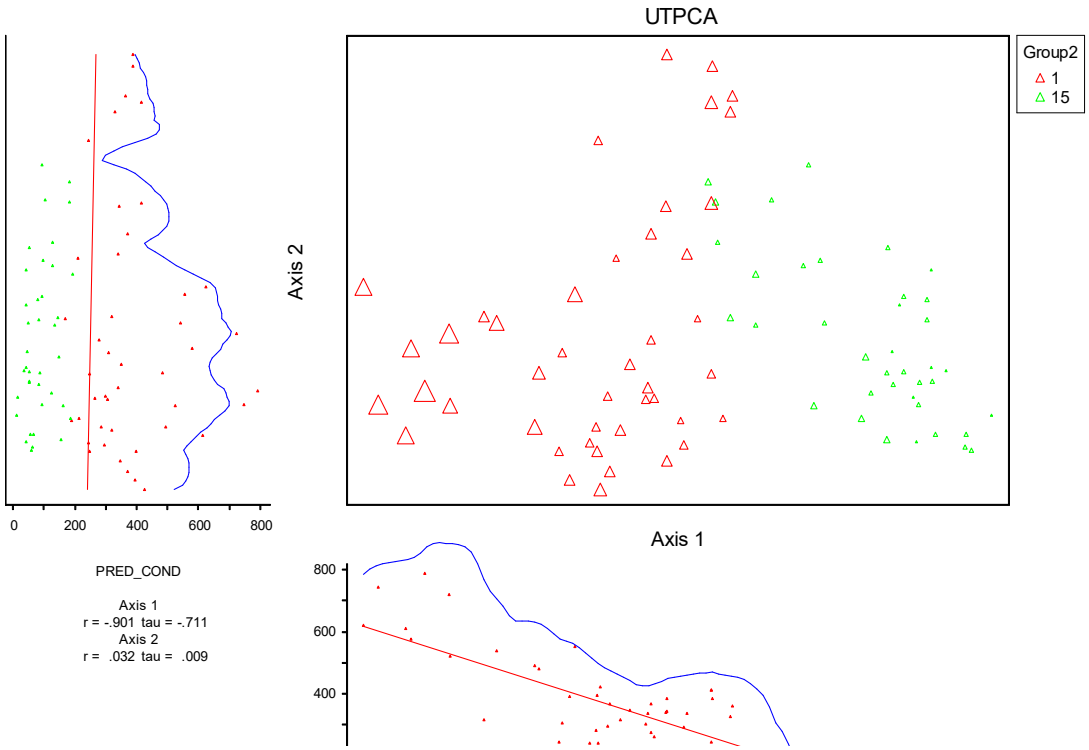


Figure 6 – Plot of axes 1 and 2 from a principal components analysis of predictors in Table 2. Sites are coded according to the two cluster groups generated from a cluster analysis of the same predictors. Separation along axis 1 is evident and due to differences in geology among these two classes.

We then examined differences in nutrient distributions within reference sites using these groupings. Only dissolved phosphorus showed any indication of separation according to the cluster groupings (Fig. 7).

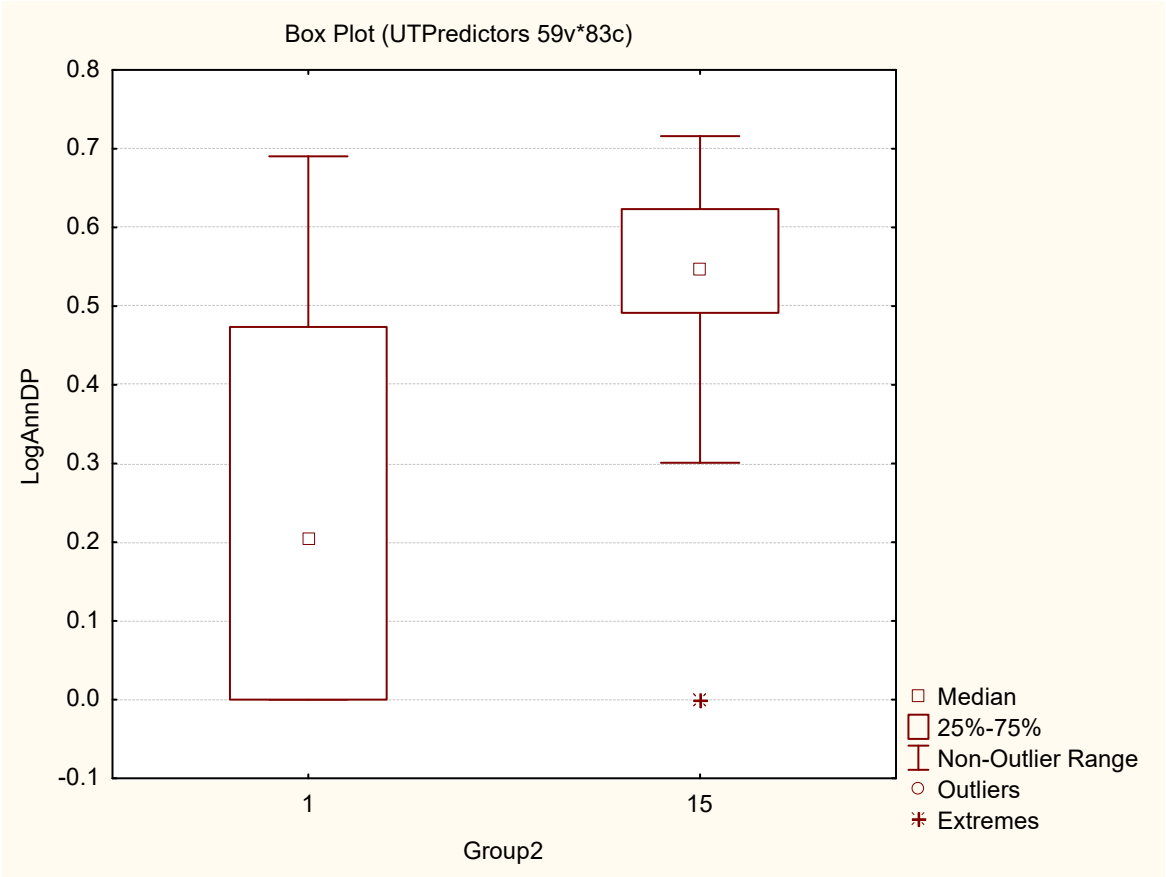


Figure 7 – Box and whisker plot of standardized mean annual dissolved phosphorus (logAnnDP) concentration distributions between the two cluster groups [1 and 2 (labeled 15)].

Given the weak relationships among natural predictors and reference site nutrient concentrations and the inability to identify a convincing basis for stream nutrient classes, we recommend against classifying streams for the basis of nutrient criteria development at this time.

Nutrient Analysis

Distribution Based Numbers

Before beginning an investigation of relationships between nutrient stressors and biological response variables, we calculated standard distributional statistics for nutrient parameters across the state. Sites were labeled according to their reference status as determined by UDWQ.

Statistics for quartiles were the focus of the analysis, after USEPA guidance (e.g.,USEPA 2000d). Seasonal quartiles were estimated for reference (75th) and all site (25th) populations, and the median of those quartiles estimated as an annual value. Graphical distributions for total nitrate and total phosphorus were developed (Figs. 8 and 9) and a tabular summary prepared (Table3).

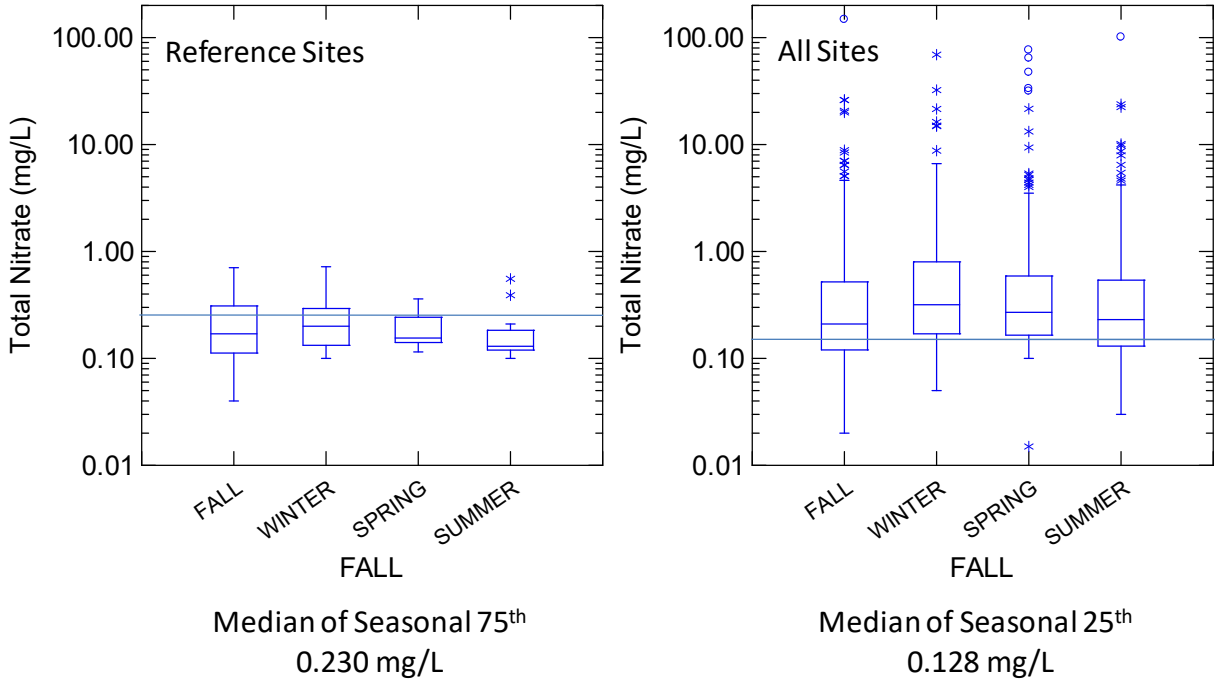


Figure 8 – Box and whisker plots of seasonal total nitrate site medians for reference and all site populations. Also shown are the 75th and 25th percentiles of these populations, respectively.

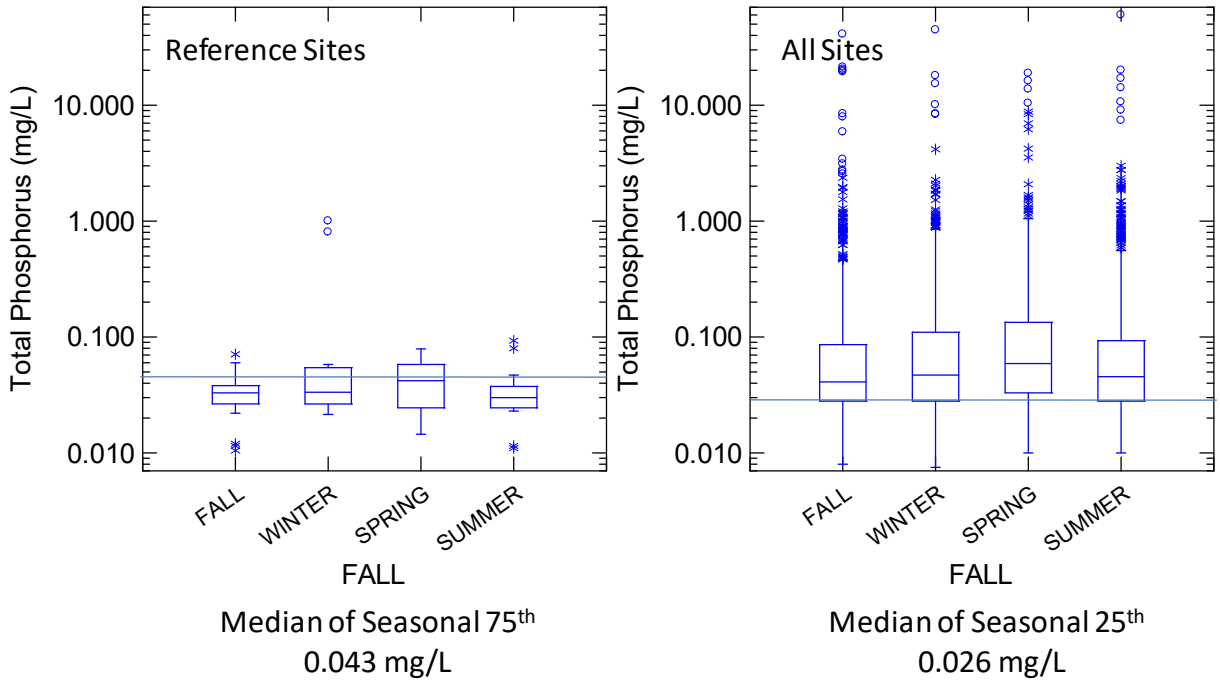


Figure 9 - Box and whisker plots of seasonal total phosphorus site medians for reference and all site populations. Also shown are the 75th and 25th percentiles of these populations, respectively.

Table 3 – Seasonal 75th and 25th percentiles of nutrient variable medians for reference and all site populations, respectively. Also included are the medians of seasonal quartile values as annual values. All units in mg/L.

	NH4	TKN	TNOx	DNOx	TP	DP
Reference Sites						
Annual	0.022		0.230	0.244	0.043	0.028
Fall 75th	0.080		0.310	0.278	0.038	0.028
Winter 75th	0.018		0.215	0.210	0.048	0.036
Spring 75th	0.000		0.245	0.598	0.058	0.029
Summer 75th	0.025		0.185	0.166	0.036	0.026
All Sites						
Annual	0.000	0.289	0.128	0.130	0.026	0.000
Fall 25th	0.000	0.195	0.100	0.130	0.024	0.000
Winter 25th	0.000	0.129	0.155	0.188	0.025	0.000
Spring 25th	0.000	0.383	0.160	0.130	0.031	0.011
Summer 25th	0.000	0.521	0.100	0.100	0.026	0.000

Annual reference site values for total nitrate and total phosphorus were 230 and 43 µg/L, respectively (seasonal 75th percentile ranges were 185-310 and 36-58 µg/L, respectively). Comparable values for the 25th percentile of all sites were 128 and 26 µg/L for total nitrate and total phosphorus, respectively (seasonal 25th percentile ranges were 100-160 and 24-31 µg/L, respectively).

Stressor-Response Analysis

Stressor-response approaches refer to a suite of analytical techniques to derive candidate endpoints by 1) identifying nutrient concentrations associated with pre-existing thresholds in response variables (interpolation) and 2) exploring and identifying thresholds in the relationships between response variables and nutrient concentrations (Figure 10).

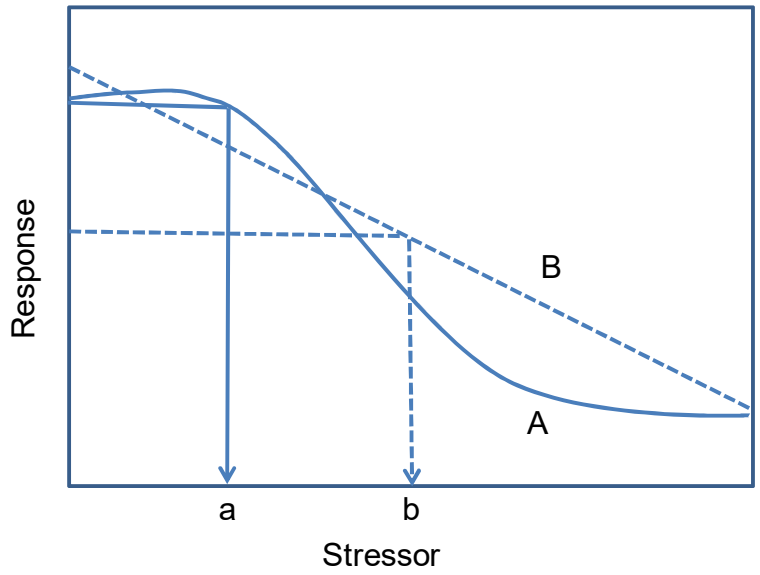


Figure 10 Plot showing two stressor-response relationships, where the response is a direct measure of designated use or can be easily linked to a designated use measure. In relationship B, a preexisting threshold of the response variable exists. For example, UDWQ O/E score thresholds (e.g., 0.74) or minimum dissolved oxygen criteria (3 to 8 mg/L depending on use class) could be the response variable and the corresponding criterion value for each used as the pre-designated threshold. The hatched line from the response threshold to the regression line is extrapolated downward to estimate the nutrient concentration associated with that threshold (b). In relationship A, no pre-existing threshold in the response variable exists; however the response is non-linear and changes once the concentration (a) is reached. This changepoint becomes a candidate criterion.

Response variables for stream nutrient endpoint development for this analysis focused on aquatic life use indicators (biological metrics and O/E scores). The value of these indicators lies in their direct linkage to aquatic life use designations. They, therefore, provide a way to connect nutrient concentrations directly to designated aquatic life use support.

Response variables were first examined with general Spearman correlation analysis and visual scatterplots with smoothing functions to identify potential relationships between biological response and nutrient variables. Locally weighted scatterplot smoothing (LOWESS) was used to fit lines representing the general relationship between the response and the nutrient stressor variables. The LOWESS technique (Cleveland 1979) models nonlinear relationships where linear methods do not perform well. LOWESS fits simple models to localized subsets of the data to construct a function that describes, essentially, the central tendency of the data.

Correlations of interest (significant correlations with highest r values) were selected and used for subsequent stressor-response analyses. Highly correlated response measures were prioritized based on using those with strongest correlation with nutrient stressors and a final list of response metrics was developed (Table 4).

Table 4 – List of biological response and nutrient stressor variables that were the focus of analysis based on initial correlation screening. *Ephemeroptera richness did not show significant correlations for the summer data alone, but other response variables did.

Biological Responses	Nutrient Stressors
Taxa Richness (Taxa)	Total Ammonium (NH4TOT)
EPT Richness (EPT)	Total Nitrate (NOXTOT)
Ephemeroptera Richness (E)*	Dissolved Nitrate (NOXDISS)
Percent EPT individuals (EPTPct)	Total Phosphorus (PTOT)
Scraper richness (SCR)	Dissolve Phosphorus (PDISS)
Scraper-Collector Gatherer Richness (SCRCG)	
Hilsenhoff Biotic Index (HBI)	
O/E Scores	

Interpolated endpoints based on existing thresholds

Once correlation analyses were complete, a set of promising stressor-response relationships were identified and the stressor-response analysis proceeded. Two specific stressor-response approaches were applied: regression analysis to identify nutrient concentration associated with existing response endpoints and change-point analysis to identify thresholds in responses to increasing nutrient concentrations not having pre-existing endpoints (Figure 10).

Linear regression is a well established technique explained in most introductory statistical texts. Linear regression models were developed for each apparently linear stressor-response relationship of interest for which identifiable or pre-existing thresholds in the response variable existed. Identifiable thresholds for macroinvertebrate metrics were generated by exploring percentile values of reference site metric distributions. A standard practice in scoring individual metrics for use in bioindicator development is to use 25th percentile metric scores of reference sites for metrics that decrease with stress (75th percentile for those that increase with stress) as thresholds between reference and non-reference sites. A similar approach is used for scoring final indices (Barbour et al. 1999). We applied these criteria to develop reference site "expectations" for individual metrics that were then used as thresholds for interpolating nutrient concentrations consistent with acceptable biological conditions (Table 5). In addition to these endpoints, O/E scores of 0.54 and 0.74 were used as response thresholds, consistent with UDWQ decision rules regarding aquatic life use evaluation based on O/E scores.

Table 5 – Table of distributional statistics of biological metrics evaluated in stressor-response relationship models. Values highlighted in blue are quartiles typically used in identifying thresholds between reference and non-reference conditions in developing biological indices (see text). These values were used as thresholds for linear regression interpolation models for nutrient endpoint development.

Variable	Descriptive Statistics (UTBioChemWatershedAnnSummCleaned.sta) Include condition: RefStat_JO="REF"						
	Valid N	Mean	Minimum	Maximum	Lower Quartile	Upper Quartile	Std.Dev.
Taxa	69	23.71739	11.00000	38.00000	20.00000	28.00000	5.75574
EPT	69	8.71739	1.00000	18.00000	6.00000	11.00000	3.74114
E	69	2.39130	0.00000	6.00000	1.00000	3.00000	1.28587
EPTPct	69	26.60851	1.07692	84.42308	12.78317	36.68262	18.85232
SCR	69	2.55797	0.00000	8.00000	1.00000	4.00000	1.60774
SCRCG	69	4.18841	0.00000	10.00000	3.00000	5.00000	1.85722
HBI	69	3.27449	1.08609	8.48214	2.11364	4.27442	1.46090

Standard regression diagnostics were used to identify and evaluate appropriate models. Linear interpolation was used to identify those nutrient concentrations associated with these specific response endpoints (e.g., biocriteria levels).

An example regression and interpolation based on a nutrient stressor-response relationship is shown in Figure 11. This same analysis was repeated for each significant regression and each nutrient-metric

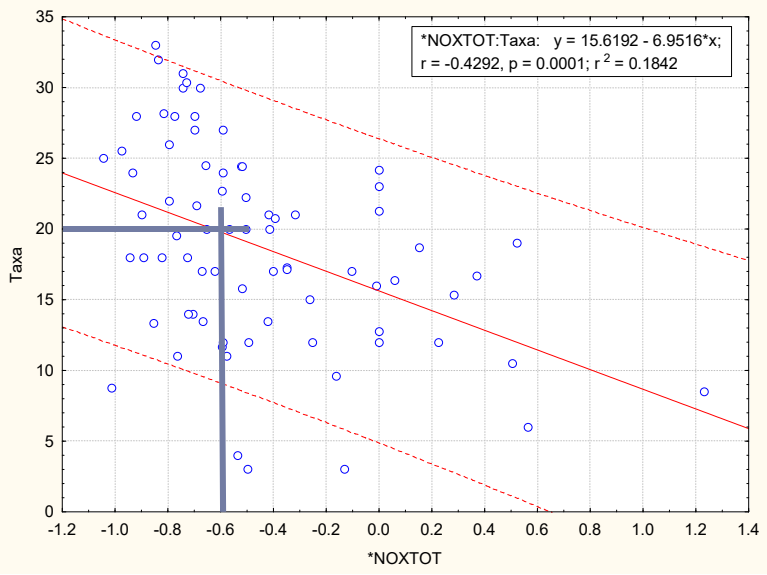


Figure 11 – Simple linear regression scatterplot of log(total nitrate)(*NOXTOT) in mg/L versus taxa richness (Taxa). The solid blue lines indicate the value of total nitrate associated with the 75th percentile reference value for taxa richness (20). Also shown on the plot are the regression equation, significance, and coefficient of determination.

response pair. The summary of interpolation values is shown in Table 6 for both annual values and summer values. The figures for each of these analyses are in Appendix 1.

Table 6 – Table of interpolated nutrient values associated with biological thresholds existing (O/E) or calculated based on reference quartiles (see text). Values are given for total nitrate (NOXTOT), dissolved nitrate (NOXDISS), and total phosphorus (PTOT). All values in mg/L. Endpoints are shown based on annual values and summer values. Abbreviations for biological responses are given in Table 4.

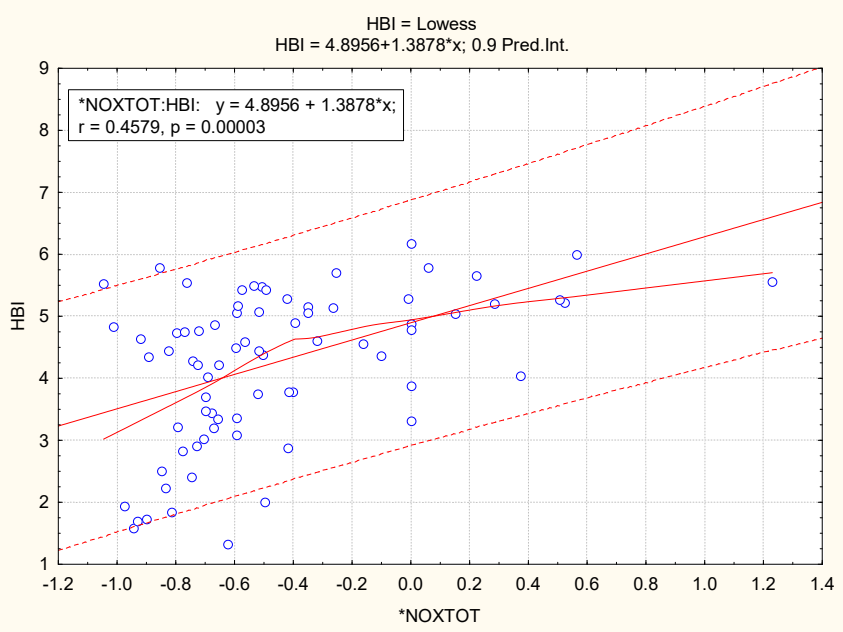
Biological Response	NOXTOT	NOXDISS	PTOT
		Annual Data	
Taxa	0.234	0.280	0.046
EPT	0.289	0.341	0.054
E	1.068	1.191	0.568
EPTPct	1.078	1.779	0.608
SCR	1.245	1.360	0.507
HBI	0.354	0.457	0.080
RFOE0.8	0.055		
RFOE0.74	0.108		
RFOE0.6	0.529		
RFOE0.54	1.040		
		Summer Data	
Taxa	0.207	0.288	
EPT	0.286	0.367	0.036
E			
EPTPct			
SCR	1.513	1.911	0.731
SCRCG	0.495	0.545	0.081
HBI	0.397	0.455	0.075
RFOE0.8	0.030	0.053	
RFOE0.74	0.070	0.124	
RFOE0.6	0.488	0.896	
RFOE0.54	1.120	2.090	

Endpoints based on change-point analysis

After exploring potential endpoints using interpolation of thresholds, we investigated whether apparent non-linear stressor-response relationships to nutrients also existed for the metric responses and conducted change-point analysis to identify thresholds in biological responses that may represent potential endpoints. For change-point analysis, nonparametric deviance reduction was used to identify thresholds in biological responses to nutrients (Qian et al. 2003, King and Richardson 2003). This technique is similar to regression tree models, which are used to generate predictive models of response variables for one or more predictors. The change-point, in our application, was the first significant split of a tree model with a single predictor variable (nutrient concentration). This first split is

that point along the stressor axis that splits the response data into the two groups that result in the greatest significant (based on a χ^2 test) reduction in deviance (Breiman et al. 1984). Output from change-point analyses included the threshold as well as the proportion of reduction in error (PRE), analogous to the regression coefficient from general linear models.

We explored the stressor-response relationships visually using LOWESS regression for the annual nutrient chemistry variables first. Of the potential relationships, only HBI and O/E showed potential threshold responses, and those only to median nitrate concentrations (e.g., Fig. 12, additional models in Appendix 3). Additionally, three models for summer nutrient variables were explored based on apparent, potential non-linear stressor-response relationships. These were O/E and total nitrate and scraper and scraper/collector-gatherer richness and total phosphorus (Appendix 4). A summary of change-point based endpoints is provided in Table



153 cases deleted due to missing data.

Split	Variable	PRE	Improvement
1	VNOXTOT	0.235	0.235

Fitting Method: Least Squares
Predicted variable: HBI
Minimum split index value: 0.050
Minimum improvement in PRE: 0.050
Maximum number of nodes allowed: 8
Minimum count allowed in each node: 5
The final tree contains 2 terminal nodes
Proportional reduction in error: 0.235

Node	from	Count	Mean	SD	Split Var	Cut Value	Fit
1	0	50	4.533	0.883	VNOXTOT	-0.395	0.235
2	1	30	4.187	0.826			
3	1	20	5.053	0.706			

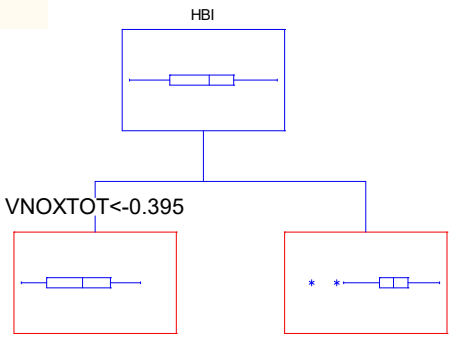


Figure 12 - Plot of HBI vs. Log_{10}(total nitrate), as well as output from a change-point analysis of HBI and Log_{10}(total nitrate). The change-point at Log_{10}(total nitrate) of -0.395 (0.403 mg/L) resulted in a 24% reduction in model error. Units in mg/L.

Table 7 – Change-points associated with the response of various biological responses to nutrient enrichment. All values in mg/L.

	Change-Points				
	Annual		Summer		
Biological Response	NOXTOT	NOXDISS	NOXTOT	NOXDISS	PTOT
Scraper Richness					0.042
Scraper Collector-Gatherer Richness					0.043
HBI	0.403	0.288			
RFOE	0.215		0.23		

An important consideration with the change-points observed in Utah that needs to inform their interpretation is the ecological phenomenon they may represent. Some change-points are identified where a lack of response to nutrients ends and nutrients begin to assert an effect on declining biological condition (e.g., Figure 10, point a). This would certainly inform the selection of an appropriately protective value. However, if there is no initial period of resistance and condition immediately responds to nutrient enrichment, then the resultant change-points may represent the point where nutrients no longer contribute to a decline in biological condition and, even worse, the point of worst possible biological condition. These change-points would be identified at the end of a response curve to a stressor (e.g., Figure 13, point c), since the statistical properties to the right of that point (flat-line) are indeed different than those to the left (declining condition).

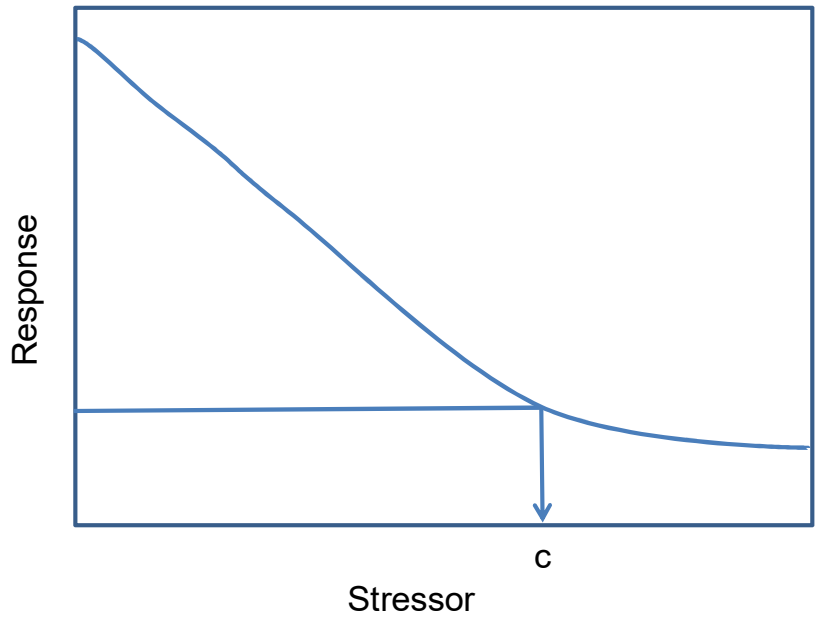


Figure 13 – Plot showing a stressor-response relationship, where the response is a direct measure of designated use or can be easily linked to a designated use measure. In this relationship A, no pre-existing threshold in the response variable exists; however the response is non-linear and the response declines consistently with any increase in stressor level observed up to point (c).

Careful consideration of change-points derived from such relationships is warranted. Such a changepoint would arguably not protect the designated use and could be considered, perhaps, an upper limit for protection. Reference expectations for the response measure might help interpret the position of such an endpoint. The changepoints described in this section all fall into this nature of change-point response. They exist at the lower end of the response curve, rather than at the upper end.

Causal Model Confirmation

We conducted two additional analyses in an attempt to better refine the causal basis for nutrient effects in UT streams. The first was a set of serial regressions and the second, an attempted propensity score analysis of existing UDWQ data.

Serial regressions

In this approach, we attempted to build a series of models (Fig. 14) along the presumptive causal pathway linking nutrients to macroinvertebrate response (Fig. 1).

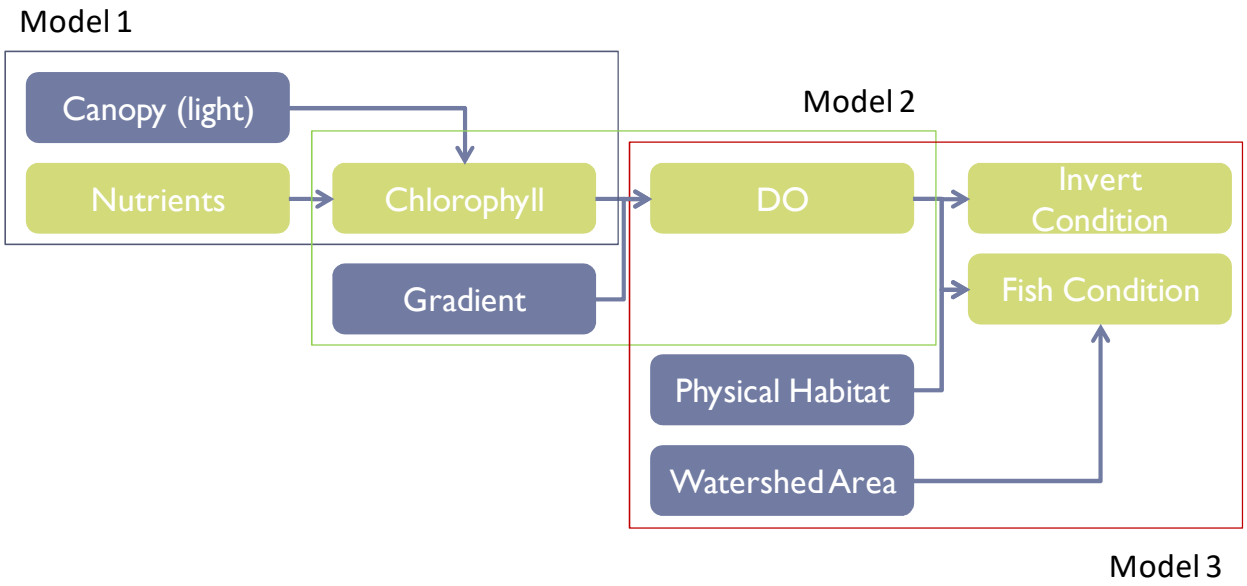


Figure 14 – Conceptual model of serial regressions linking nutrients and potential co-variates with macroinvertebrate response.

There was, unfortunately, limited chlorophyll data available. Moreover, true chlorophyll responses are difficult to characterize in streams, given the many factors that affect algal growth. Algae tend to integrate over far shorter time frames than invertebrates or fish, therefore identifying an algal condition linked to invertebrate response is difficult. We found only weak relationships for model 1 relating nutrients to chlorophyll and none of them were significant except for negative decline on chlorophyll with ammonium concentrations. We also found no relationship between chlorophyll and dissolved

oxygen grab samples (model 2). It should be noted that grab samples are not particularly accurate measures of oxygen conditions in streams. Lastly, we found a peculiar decrease in scraper richness with dissolved oxygen (model 3). Again, dissolved oxygen grab sampled were likely not an accurate measure of oxygen conditions and may even increase with algal growth if measured in the afternoon. We, therefore, could not use the serial modeling results to confirm the causal model.

Propensity score analysis

We conducted a simple propensity score analysis as an attempt to strengthen the causal basis for linking nutrients to invertebrate response (Yuan 2010). Propensity score analysis is used to investigate the strength of specific stressor-response relationships based on observational data, given the presence of covariates (Rosenbaum 2002). Nutrients co-vary with a variety of other stressors (e.g., sediment, conductivity, BOD) which make estimating the single effects of nutrients challenging, yet it is important to recognize the potential contributions of these other stressors to the observed response. The simplest way to control for effects of covariates with nutrients would be to identify sites that vary in nutrient concentration, but are similar in regards to all other stressors. In some instances this may be possible, but as the number of covariates grows, it quickly becomes difficult to do. Propensity score analysis summarizes the effects of covariates into one single variable, called the propensity score (Fig. 15). Sites with similar propensity scores are similar in regards to their covariate distribution and the effects of the factor of interest can then be explored with bins of similar propensity scores (Rosenbaum and Rubin 1983, Rosenbaum 2002, Imai and Van Dyk 2004).

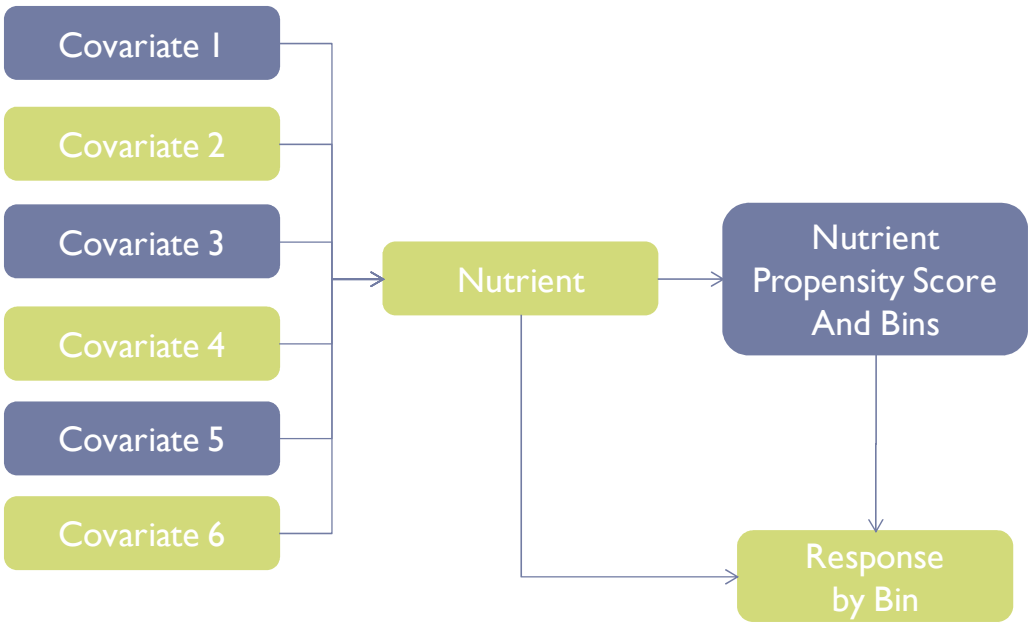


Figure 15 - Conceptual diagram of propensity score analysis. Covariates with nutrient concentration are used to generate a predicted nutrient concentration (propensity score). Sites with similar propensity scores are grouped together into bins representing similar covariate distributions. The nutrient-response relationship is then investigated per bin.

We calculated propensity scores for the Utah dataset by modeling the effects of covariates on nitrate concentration. We used the following predictors in a multiple linear regression model to predict nitrate concentration: elevation, longitude, latitude, watershed area, chloride, sulfate, bicarbonate, temperature, % agriculture, % urban, and turbidity. We split the resultant scores into 6 bins and investigated the response within each bin.

The MLR model explained 60% of the variance in nitrate concentration (Figure 16) and exhibited significant negative correlations with taxa richness, EPR richness, and positive correlations with HBI. When we explored the stressor-response relationships within bins, bin 3 exhibited the most significant decline in taxa richness and EPT richness to increasing nitrate concentration (Figs. 17 and 18). In contrast, HBI scores showed increases in almost every bin to nitrate concentration, especially bins 3 and 6, but also in 2 and 4 (Fig. 19).

Nitrate and Total Taxa Richness

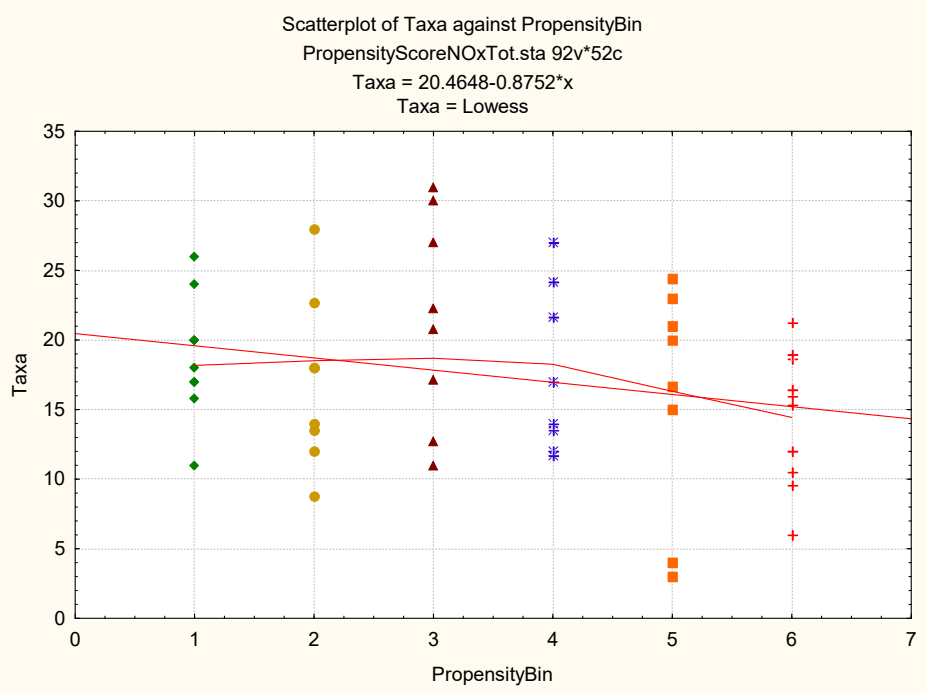


Figure 16 – Response of taxa richness across different propensity score bins.

Nitrate and Total Taxa Richness

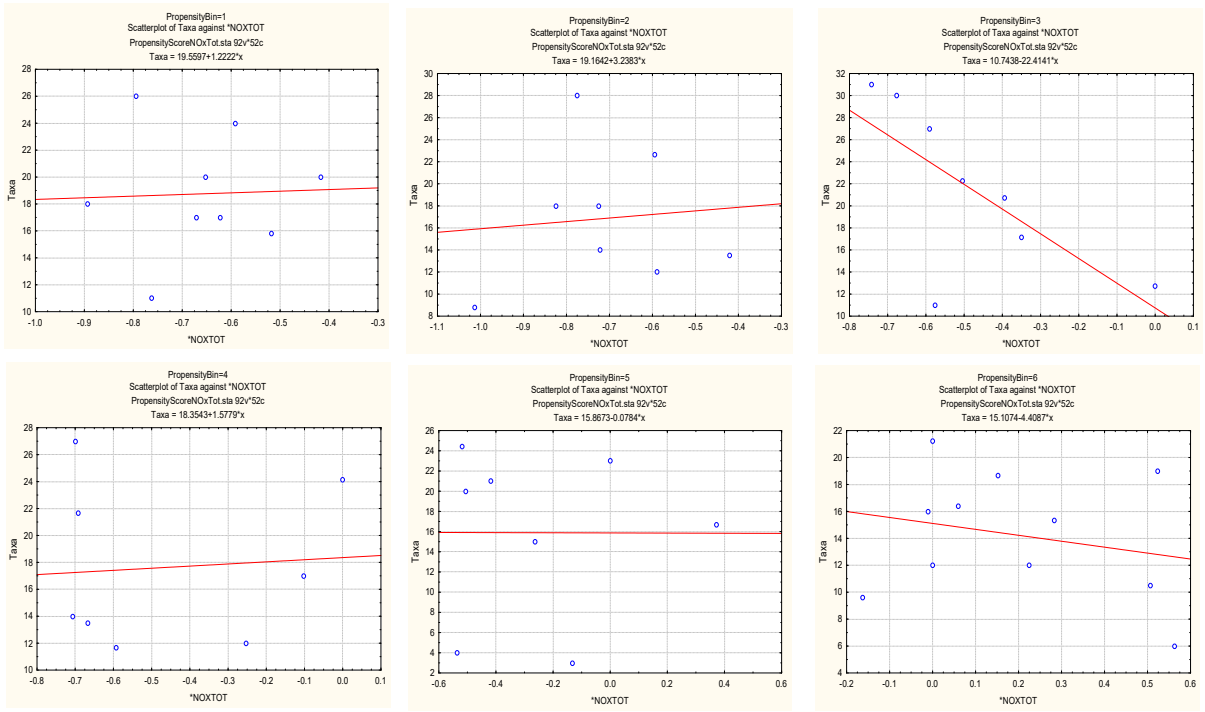


Figure 17 – Scatterplots of total nitrate vs. taxa richness within each of the 6 propensity score bins. Nitrate concentration increased from bin 1 to 6.

Nitrate and EPT Richness

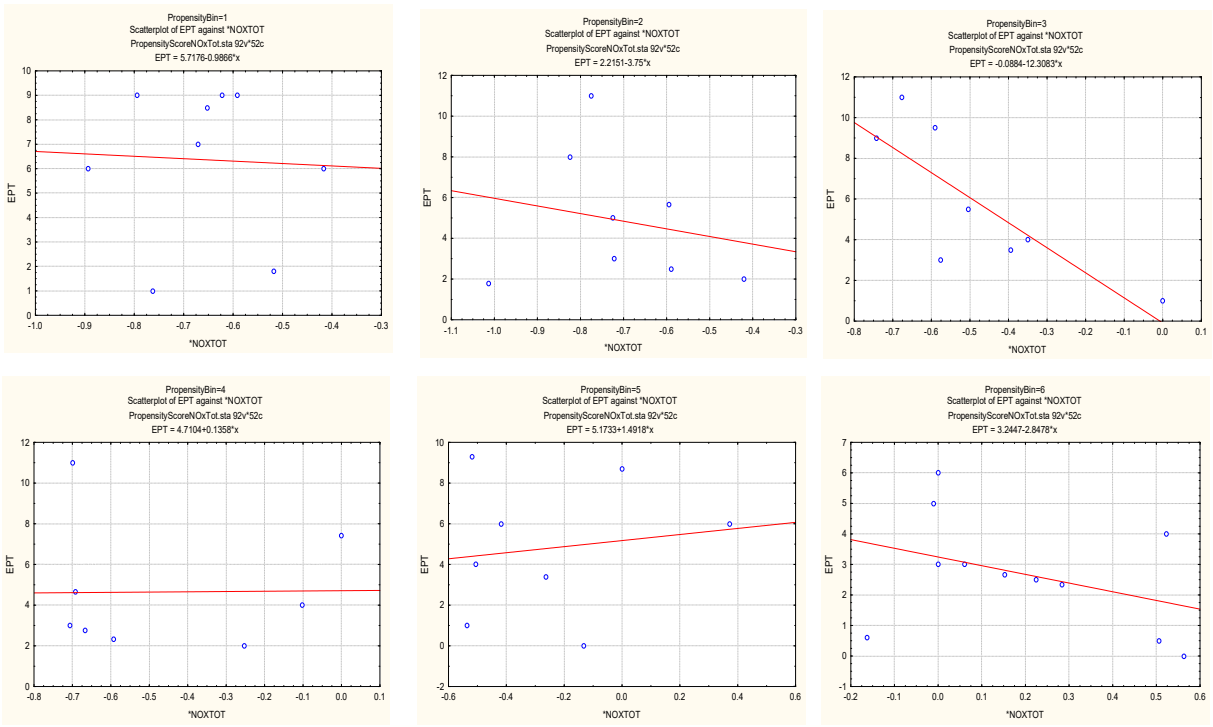


Figure 18 - Scatterplots of total nitrate vs. EPT richness within each of the 6 propensity score bins. Nitrate concentration increased from bin 1 to 6.

Nitrate and HBI

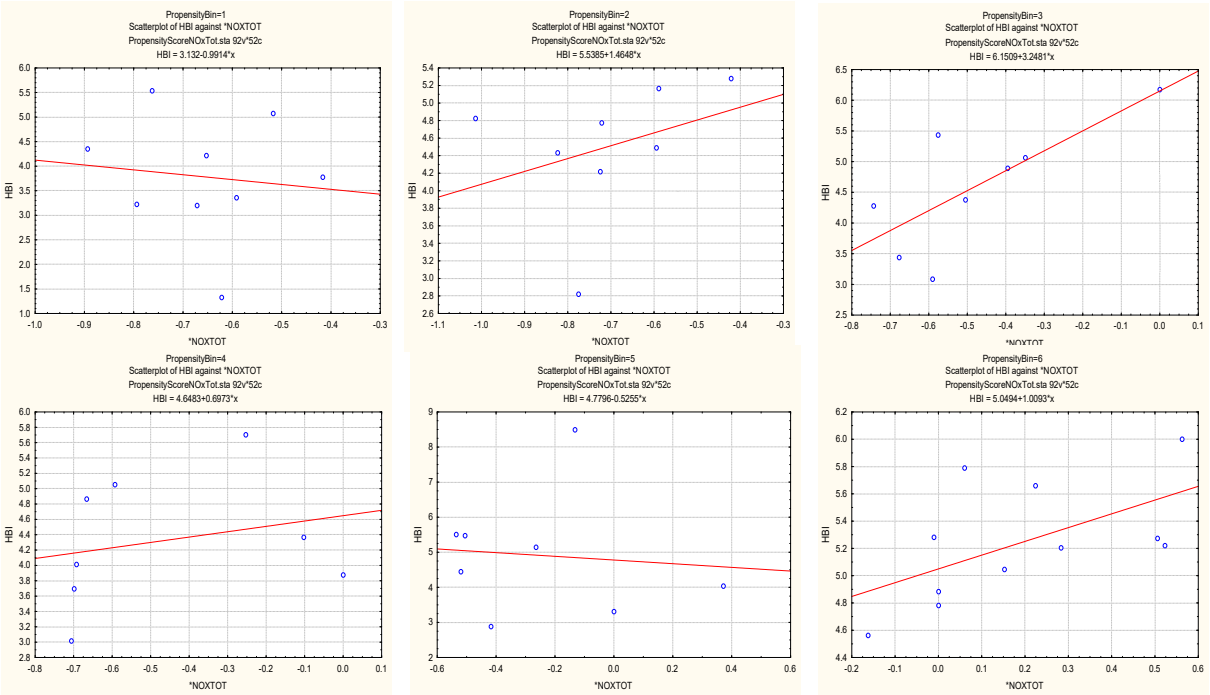


Figure 19 - Scatterplots of total nitrate vs. EPT richness within each of the 6 propensity score bins. Nitrate concentration increased from bin 1 to 6.

We interpret these results, especially the HBI results, to indicate support for an effect of nutrients on invertebrates in streams in Utah, independent of other stressors. Since those stressors are controlled for in the propensity scores and yet nitrate still exhibited a significant relationship to these response metrics, support for a deleterious effect of nutrients on invertebrate assemblages is strengthened.

Summary

- We explored classification options for streams in Utah for the purposes of potentially developing different nutrient criteria for different parts of the state. This analysis did not identify potential classes worth pursuing at this time, therefore, criteria would be developed for statewide application.
- We conducted a distribution-based analysis of nutrient concentrations based on reference site and all site populations. Quartiles for these populations were generated and reported as potential endpoints consistent with EPA guidance.
- The next analysis was developing stressor-response relationship models based on existing biological thresholds and interpolating associated nutrient values. This approach generated a

series of nutrient endpoints associated with biological conditions considered thresholds for adverse aquatic life use conditions.

- Change-point analysis was then performed for a few stressor-response relationships that exhibited potential non-linear responses. Change-points were identified for these relationships and reported.

- Lastly, two analyses were performed to strengthen the support for the presumptive effect of nutrients on aquatic life use endpoints: serial regression models and propensity score analysis. The propensity score analysis strengthened the basis for an effect of nutrients on aquatic life use responses.

- These analyses are part of a larger effort by UDWQ to develop nutrient criteria in streams. Hopefully the analyses performed here will inform and strengthen that effort, providing both scientifically defensible and protective values for the continued protection of aquatic life in Utah.

Citations

Allan, J.D. 1995. Stream Ecology: Structure and Function of Running Waters. Chapman & Hall, New York.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. 2nd ed. U.S. Environmental Protection Agency, Office of Water, Washington, DC. EPA 841-B-99-002.

Breiman, L., J.H. Friedman, R. Olshen, and C.J. Stone. 1984. Classification and Regression Trees. Wadsworth International Group, Belmont, CA.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8:559-568.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74:829-836

Cross, W.F., J.B. Wallace, A.D. Rosemond and S.L. Eggert. 2006. Whole-system nutrient enrichment increases secondary production in a detritus-based ecosystem. Ecology 87:1556-1565.

Elwood, J.W., J.D. Newbold, A.F. Trimble, and R.W. Stark. 1981. The limiting role of phosphorus in a woodland stream ecosystem: effects of P enrichment on leaf decomposition and primary producers. Ecology 62:146–158.

Feminella, J.W. and C.P. Hawkins. 1995. Interactions between stream herbivores and periphyton: A quantitative analysis of past experiments. Journal of the North American Benthological Society 14: 465–509.

Hawkins, C.P., R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000b. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecol. Applic. 10:1456-1477.

Hillebrand, H. 2002. Top-down versus bottom-up control of autotrophic biomass- A meta-analysis of experiments with periphyton. Journal of the North American Benthological Society 21: 349–369.

Imai, K. and D.A. Van Dyk. 2004. Causal inference with general treatment regimes: Generalizing the propensity score. Journal of the American Statistical Association 99: 854–866.

King, R. S., and C. J. Richardson. 2003. Integrating bioassessment and ecological risk assessment: An approach to developing numerical water-quality criteria. Environmental Management 31:795-809.

Paul, J.F. and M. E. McDonald 2005. Development of empirical, geographically specific water quality criteria: a conditional probability analysis approach. Journal of the American Water Resource Association 1211:1223.

Peterson, B.J., L. Deegan, J. Helfrich, J.E. Hobbie, M. Hullar, B. Moller, T.E. Ford, A. Hershey, A. Hiltner, G. Kipphut, M.A. Lock, D.M. Fiebig, V. McKinely, M.C. Miller, J.R. Vestal, R. Ventullo and G. Volk. 1993. Biological responses of a tundra river to fertilization. Ecology 74(3):653-672.

Qian, S.S., King, R.S., Richardson, C.J., 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modeling 166: 87–97.

Rosenbaum, P.R. 2002. Observational Studies, 2nd Edition. Springer, New York.

Rosenbaum, P.R. and D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70: 41–55.

Rosemond, A.D., P.J. Mulholland, and J.W. Elwood. 1993. Top-down and bottom-up control of stream periphyton: effects of nutrients and herbivores. Ecology 74:1264–1280.

Slavik, K., Peterson, B. J., Deegan, L. A., Bowden, W. B., Hershey, A. E., Hobbie, J. E. 2004. Long-term responses of the Kuparuk River ecosystem to phosphorus fertilization. Ecology: Vol. 85, No. 4, pp. 939–954.

USEPA. 1998. National Strategy for the Development of Regional Nutrient Criteria. EPA-822-R-98-002. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2000a. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Rivers and Streams in Nutrient Ecoregion II. EPA-822-B-00-015. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2000b. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Rivers and Streams in Nutrient Ecoregion III. EPA-822-B-00-016. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2000c. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Lakes and Reservoirs in Nutrient Ecoregion II. EPA-822-B-00-007. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2000d. Nutrient Criteria Technical Guidance Manual. Rivers and Streams. EPA-822-B-00-002. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2000e. Nutrient Criteria Technical Guidance Manual. Lakes and Reservoirs. EPA-822-B-00-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2001a. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Rivers and Streams in Nutrient Ecoregion IV. EPA-822-B-01-013. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2001b. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Rivers and Streams in Nutrient Ecoregion V. EPA-822-B-01-014. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2001c. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Lakes and Reservoirs in Nutrient Ecoregion III. EPA-822-B-01-008.U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 2001d. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Lakes and Reservoirs in Nutrient Ecoregion IV. EPA-822-B-01-009.U.S. Environmental Protection Agency, Office of Water, Washington, DC.

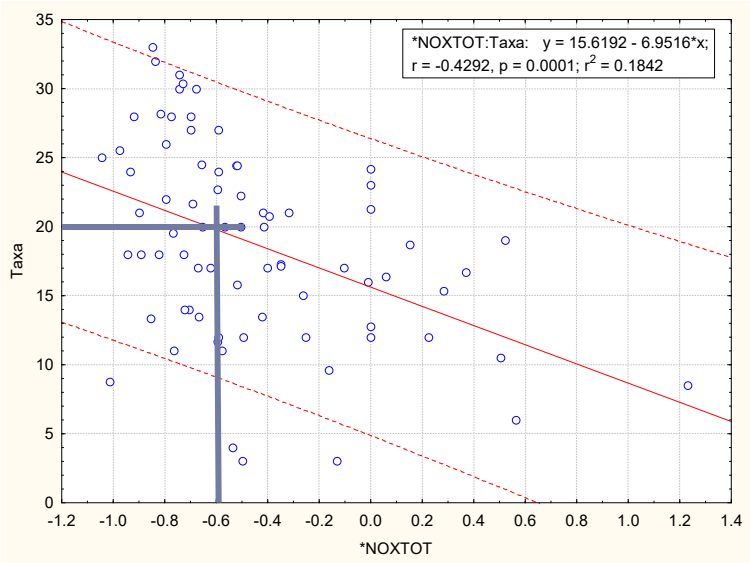
USEPA. 2001e. Ambient Water Quality Criteria Recommendations. Information Supporting the Development of State and Tribal Nutrient Criteria. Lakes and Reservoirs in Nutrient Ecoregion V. EPA-822-B-01-010. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

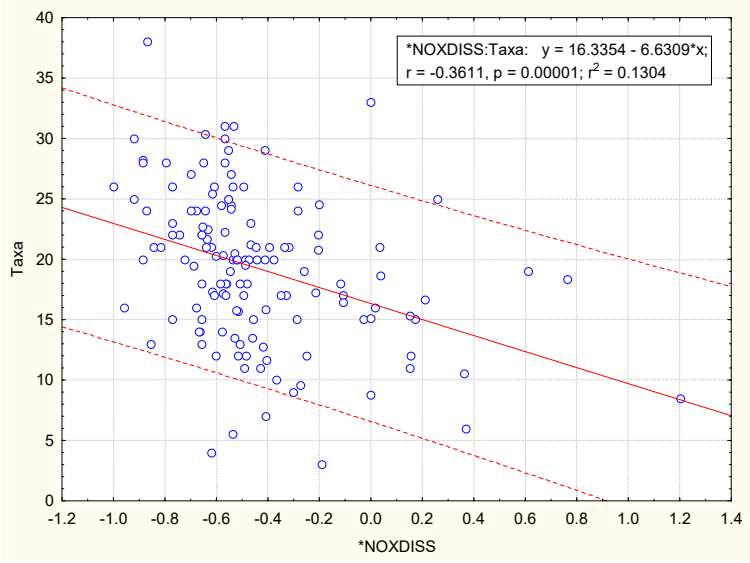
Yuan, L.L. 2010. Estimating the effects of excess nutrients on stream invertebrates from observational data. Ecological Applications 20:110–125.

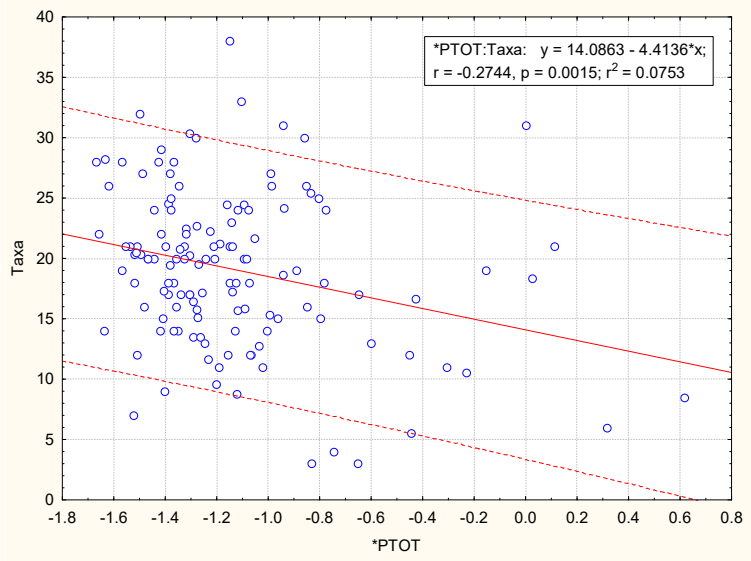
Appendix 1 – Interpolated endpoints based on existing thresholds figures: annual data

The following figures are regressions between nutrient variables [(Log_{10}total nitrate (*NOXTOT), Log_{10}dissolved nitrate (*NOXDISS), and Log_{10}total phosphorus (*PTOT)]. Response variable abbreviations are as in Table 4. Regression equations, significance, and coefficients of determination are shown. Lines are the mean regression (solid) and the 90% prediction intervals. Only significant models are shown.

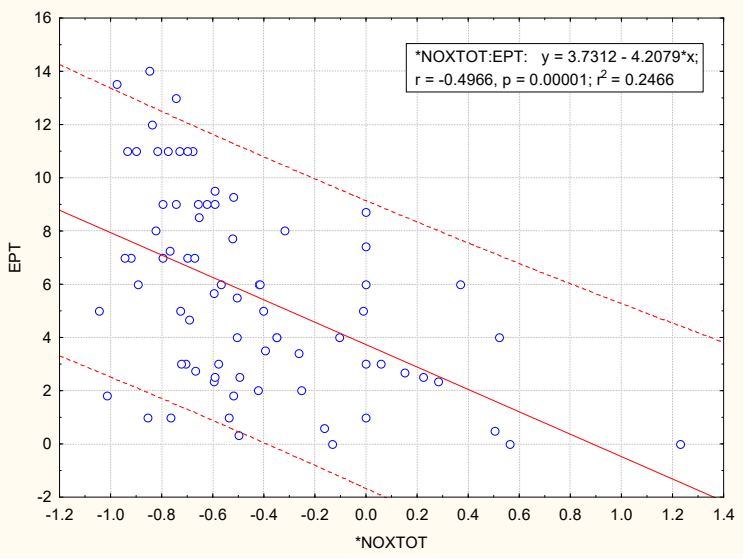
Taxa Richness

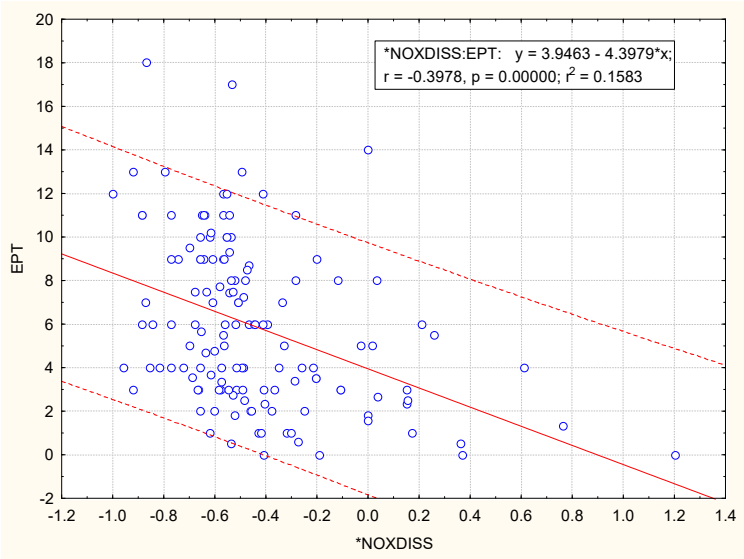


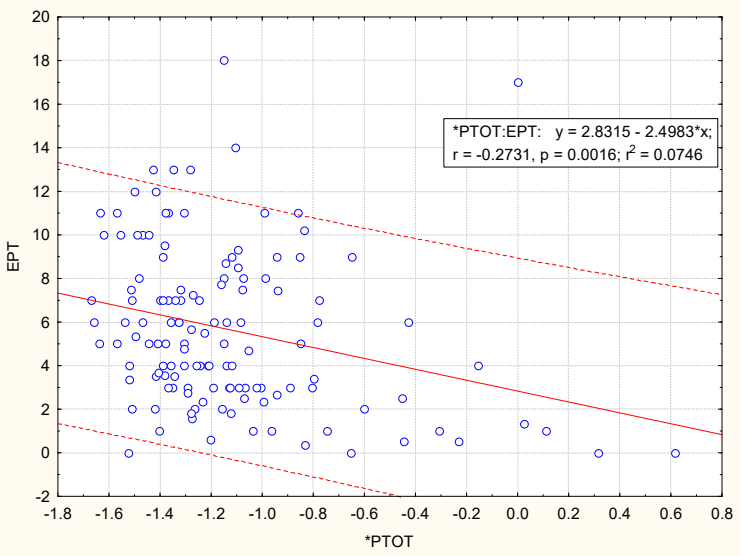




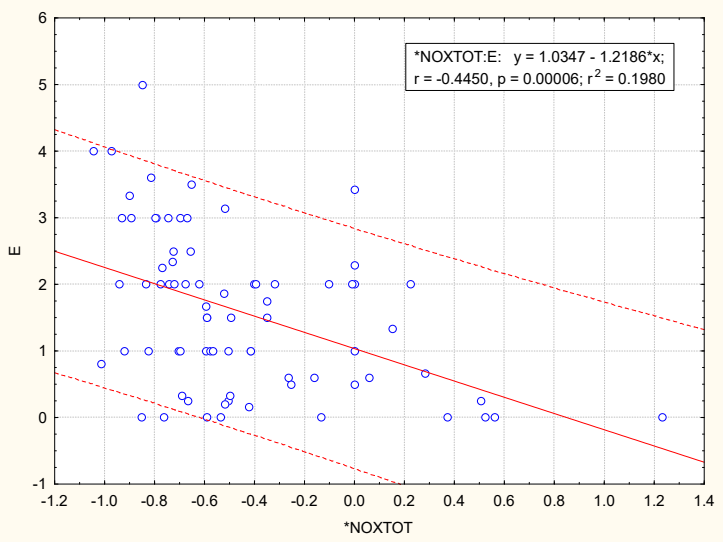
EPT Richness

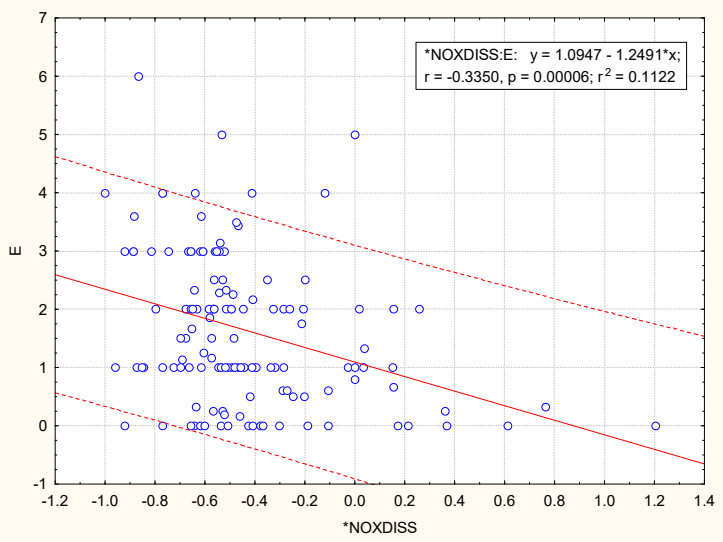


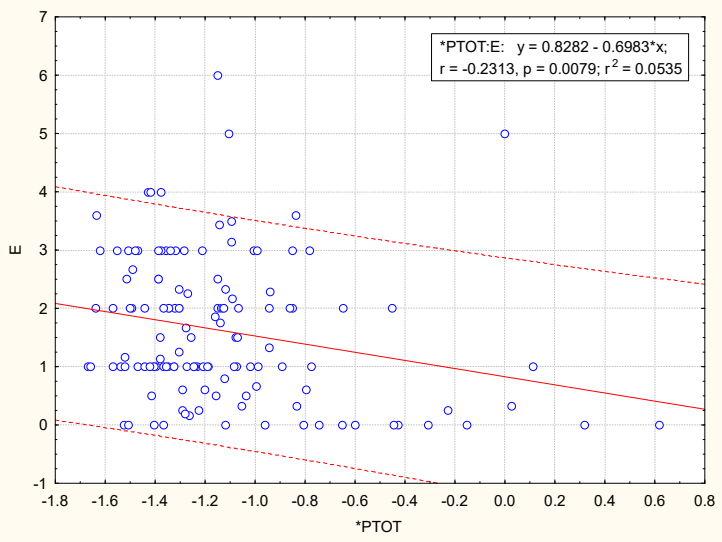




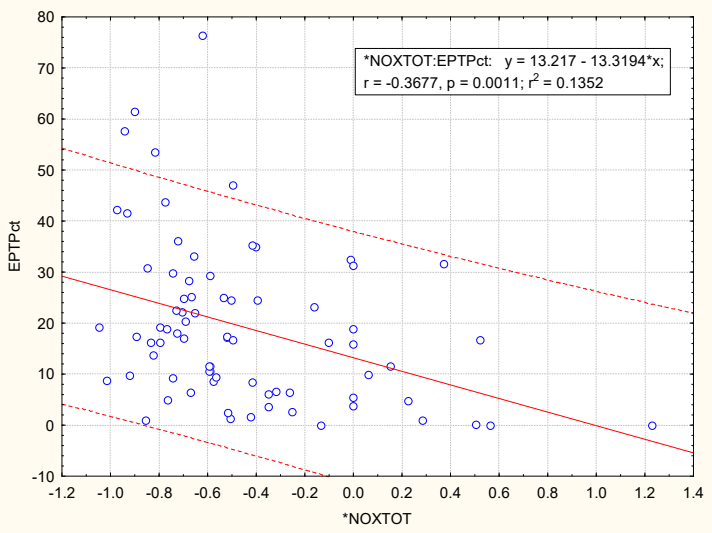
E Richness

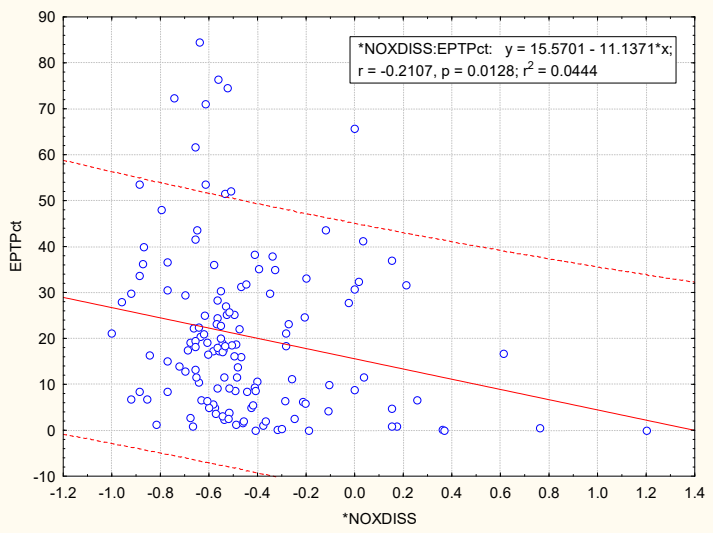

*NOXTOT:E: y = 1.0347 - 1.2186*x;
r = -0.4450, p = 0.00006; r2 = 0.1980
E
*NOXTOT

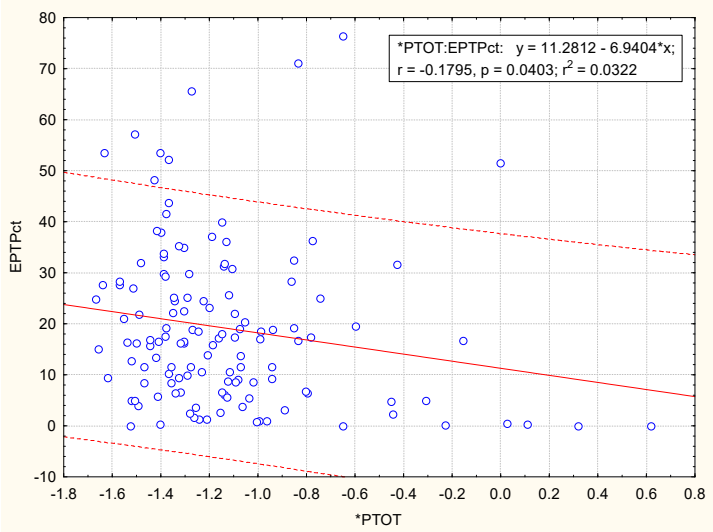

*NOXDISS:E: y = 1.0947 - 1.2491*x;
r = -0.3350, p = 0.00006; r2 = 0.1122
E
*NOXDISS


*PTOT:E: y = 0.8282 - 0.6983*x;
r = -0.2313, p = 0.0079; r2 = 0.0535
E
*PTOT

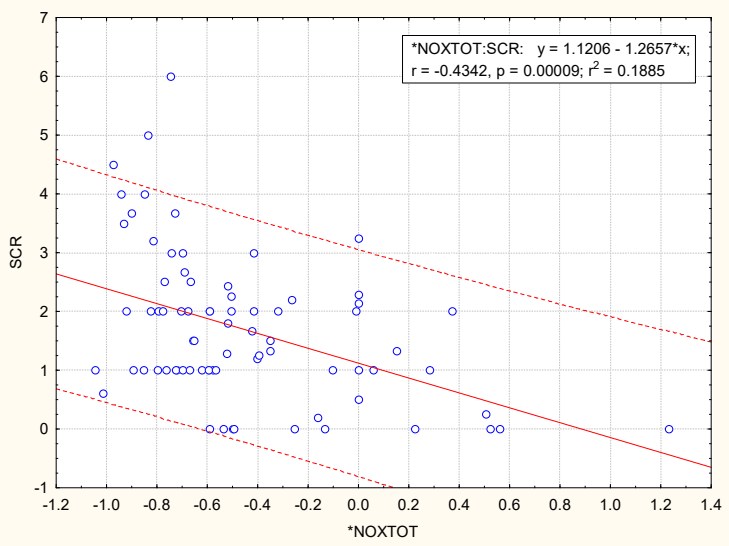
Percent EPT Individuals

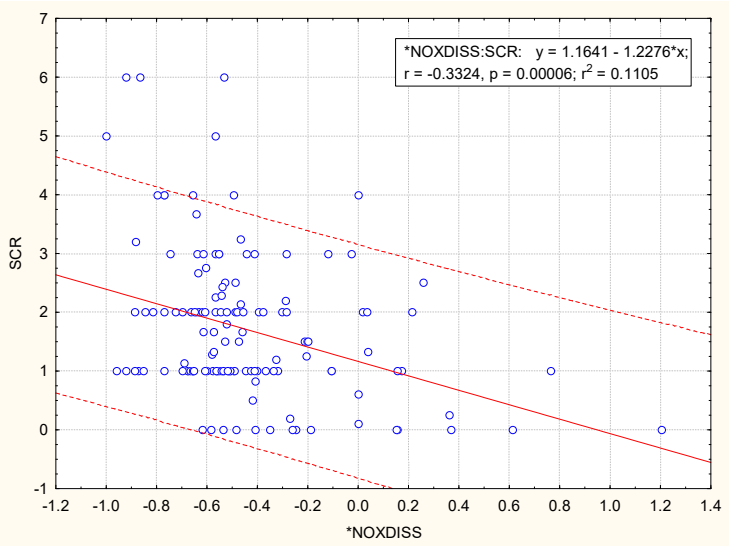


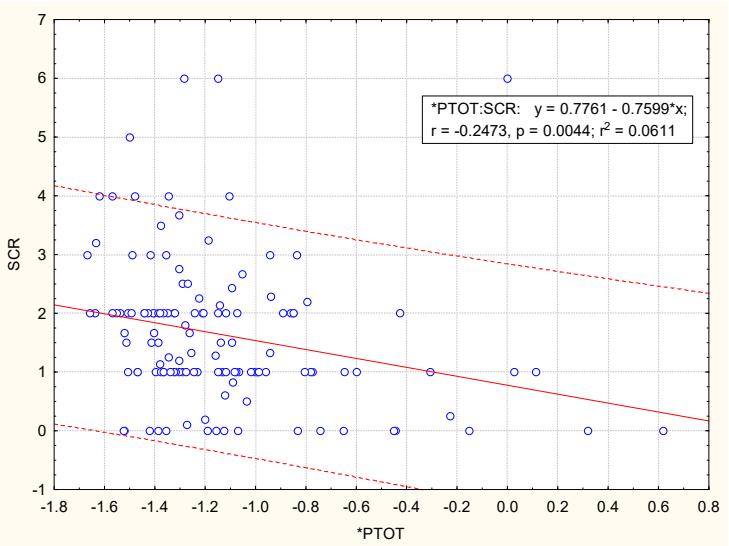




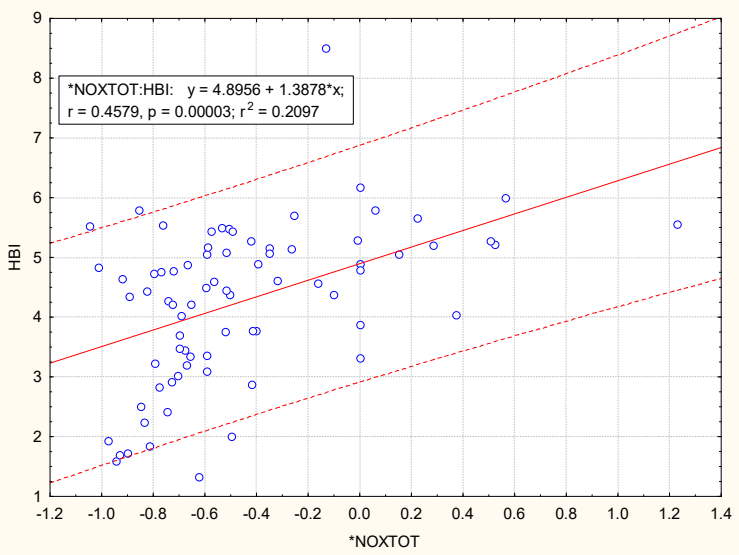
Scraper Richness

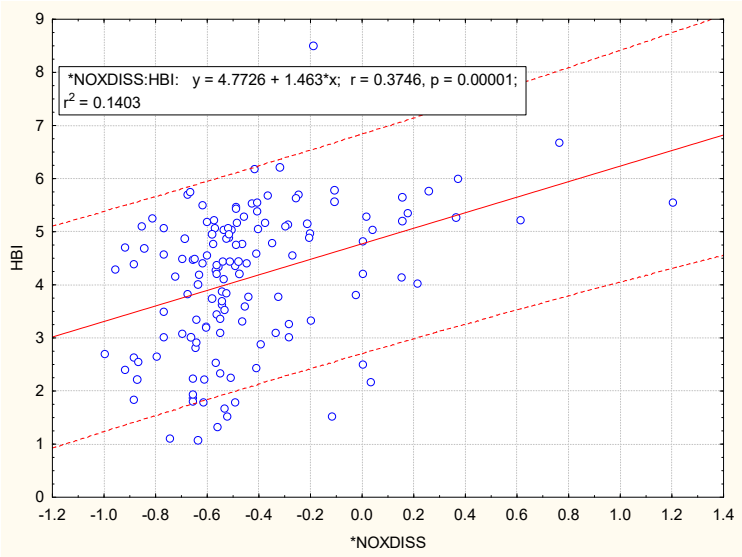


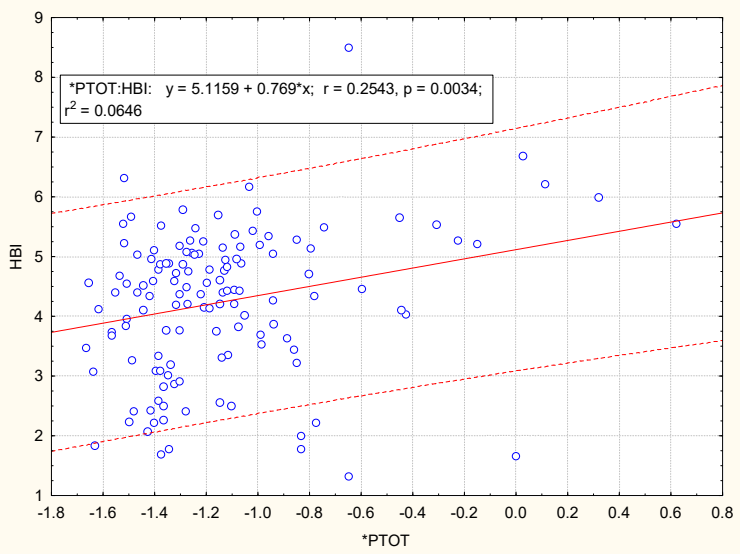




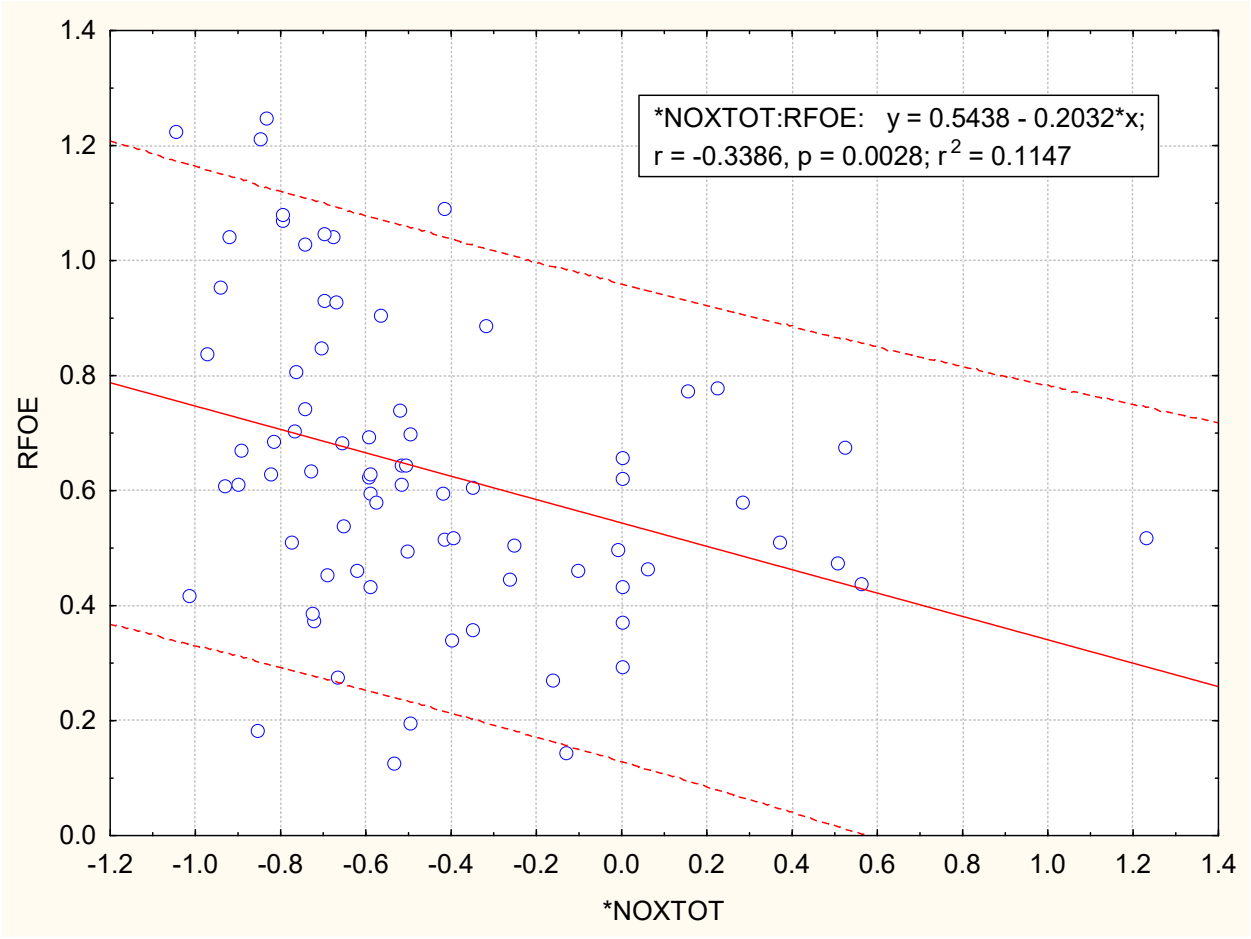
HBI


*NOXTOT:HBI: y = 4.8956 + 1.3878*x;
r = 0.4579, p = 0.00003; r2 = 0.2097
HBI
*NOXTOT


*NOXDISS:HBI: y = 4.7726 + 1.463*x; r = 0.3746, p = 0.00001;
r2 = 0.1403
HBI
*NOXDISS


*PTOT:HBI: y = 5.1159 + 0.769*x; r = 0.2543, p = 0.0034;
r2 = 0.0646
HBI
*PTOT

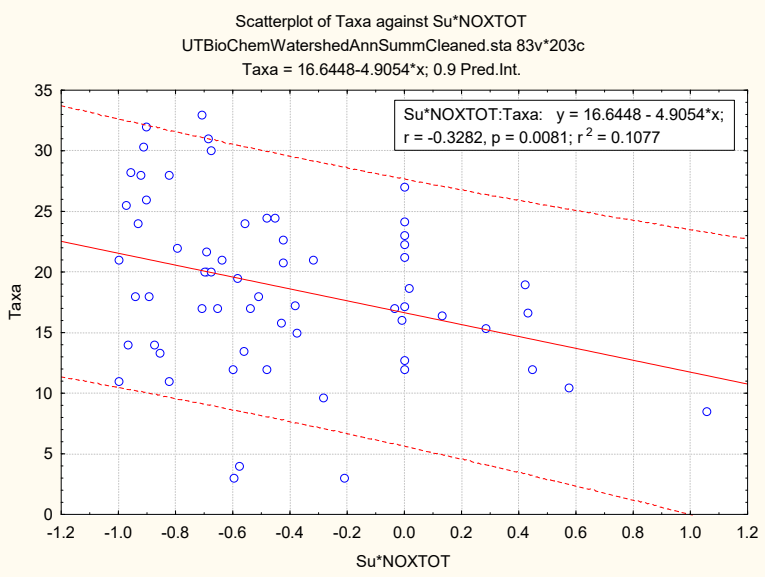
OE Model

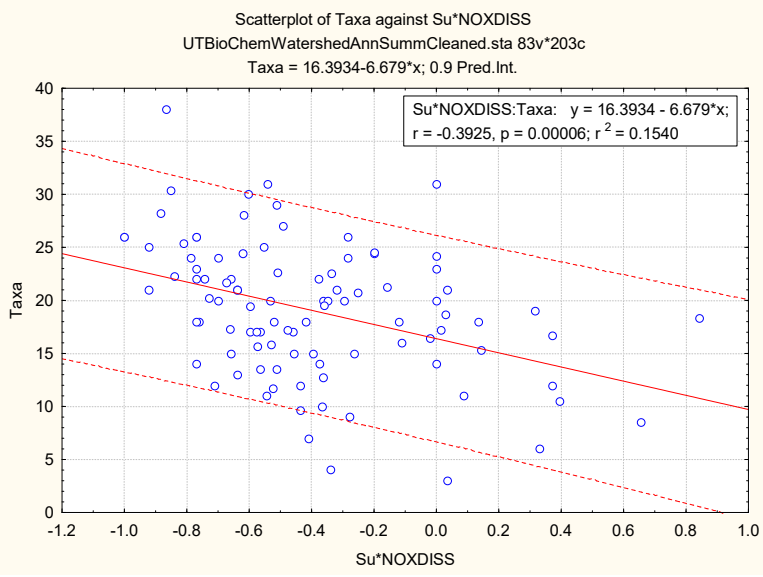

*NOXTOT:RFOE: y = 0.5438 - 0.2032*x;
r = -0.3386, p = 0.0028; r 2 = 0.1147
RFOE
*NOXTOT
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4

Appendix 2 – Interpolated endpoints based on existing thresholds figures: summer data

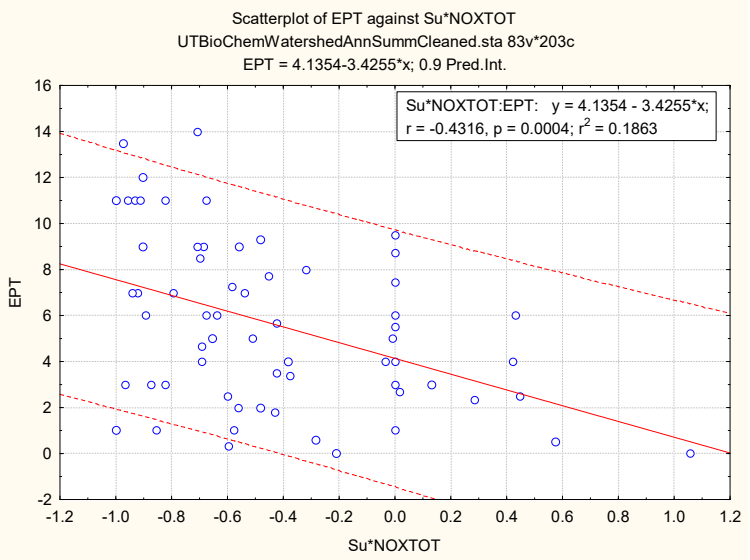
The following figures are regressions between summer nutrient variables [(Log_{10}total nitrate (Su*NOXTOT), Log_{10}dissolved nitrate (Su*NOXDISS), and Log_{10}total phosphorus (Su*PTOT)]. Response variable abbreviations are as in Table 4. Regression equations, significance, and coefficients of determination are shown. Lines are the mean regression (solid) and the 90% prediction intervals. Only significant models are shown.

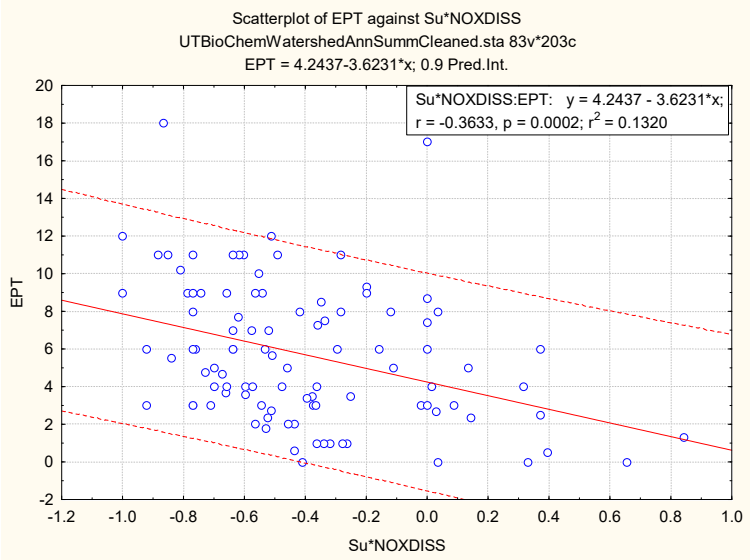
Taxa Richness

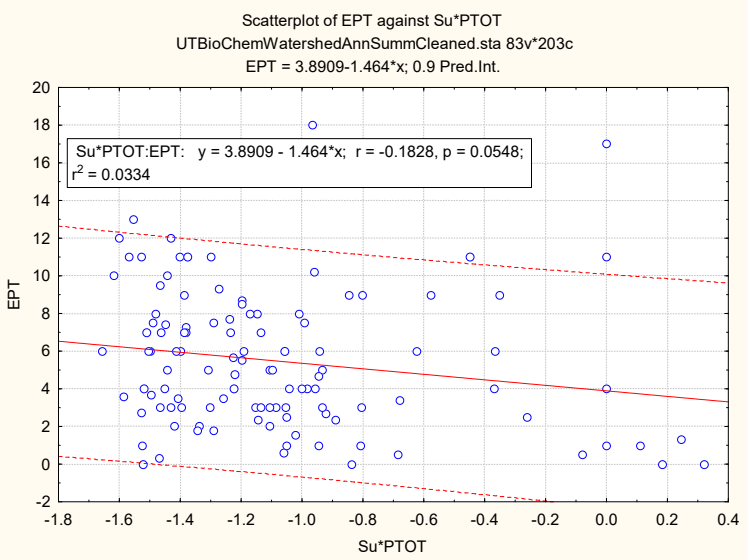




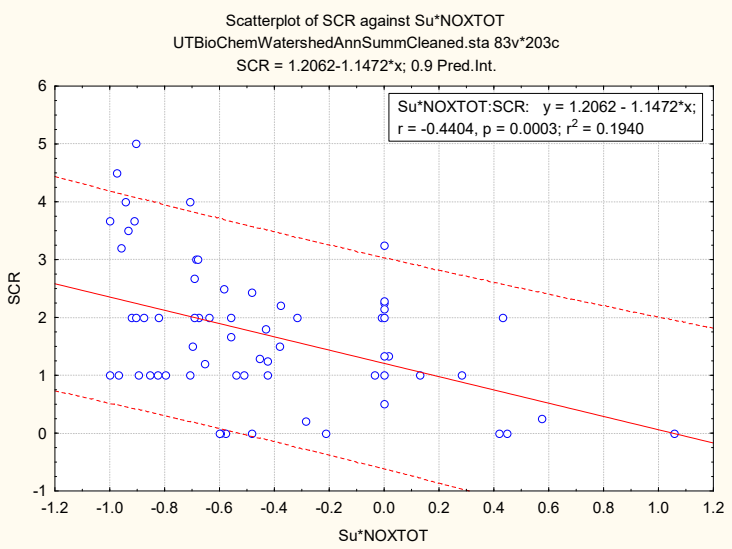
EPT Richness

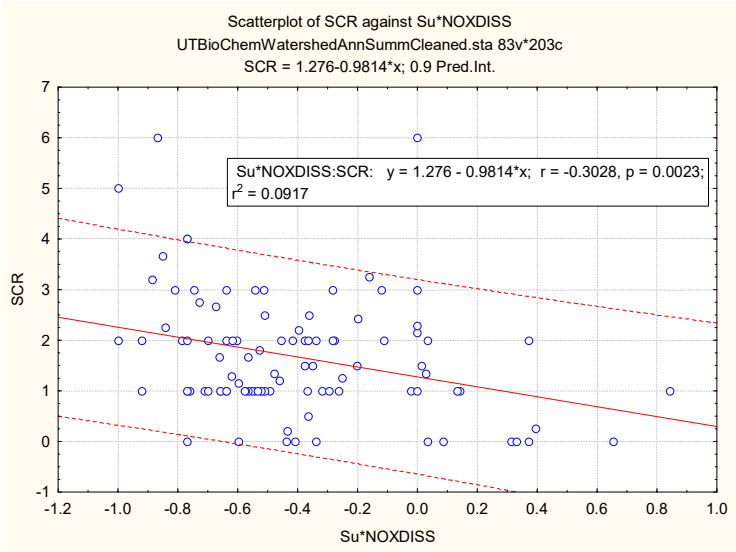


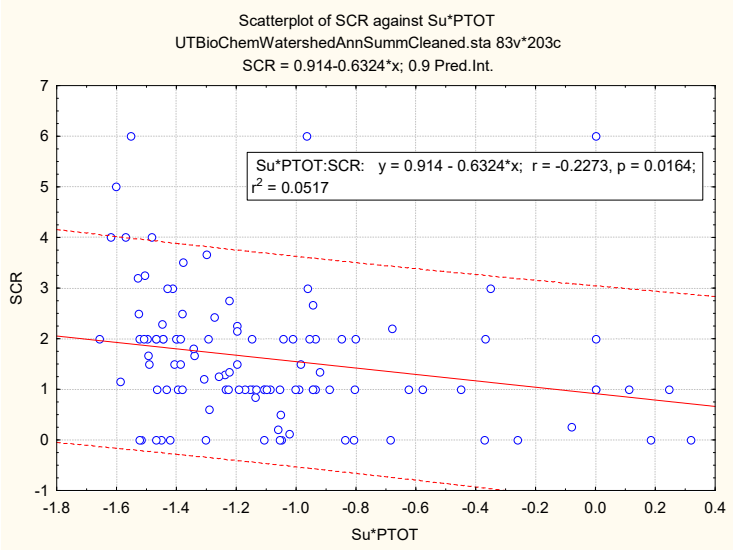




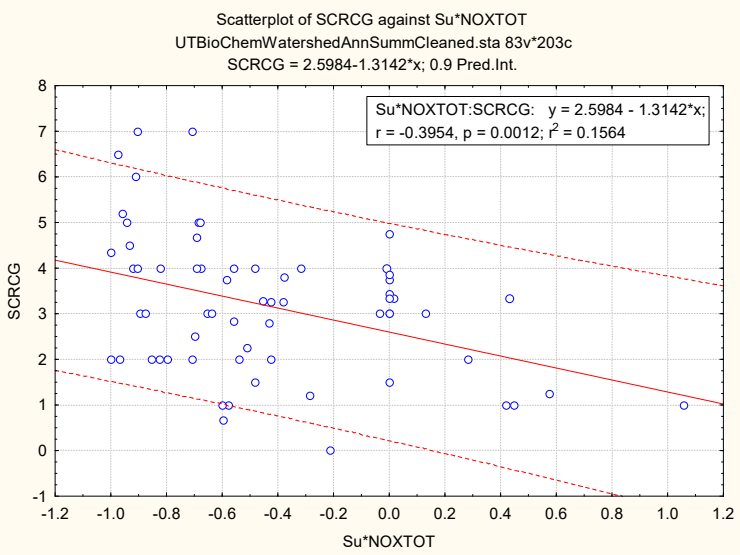
Scraper Richness

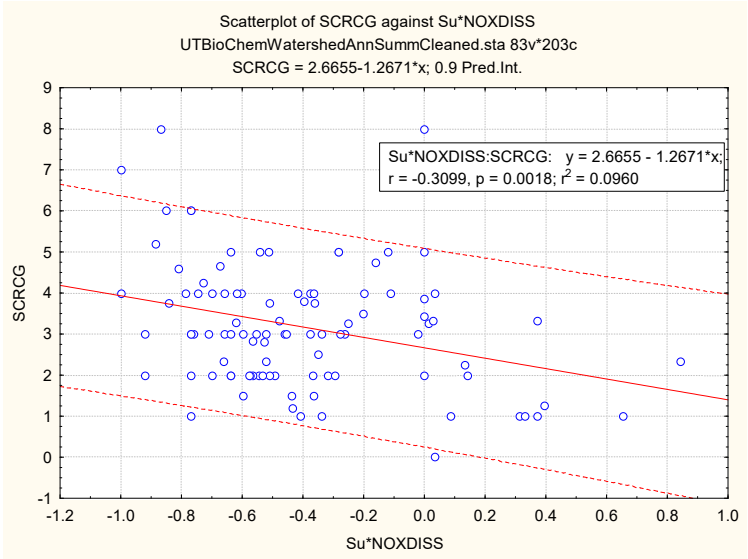


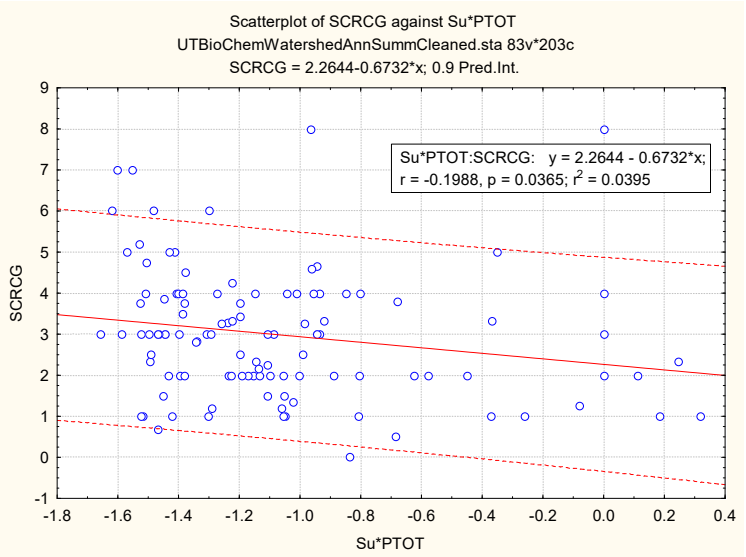




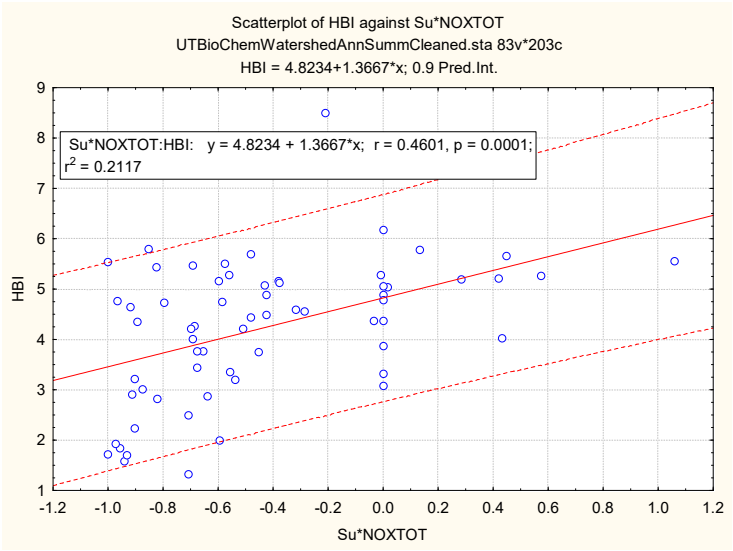
Scraper+Collector/Gatherer Richness

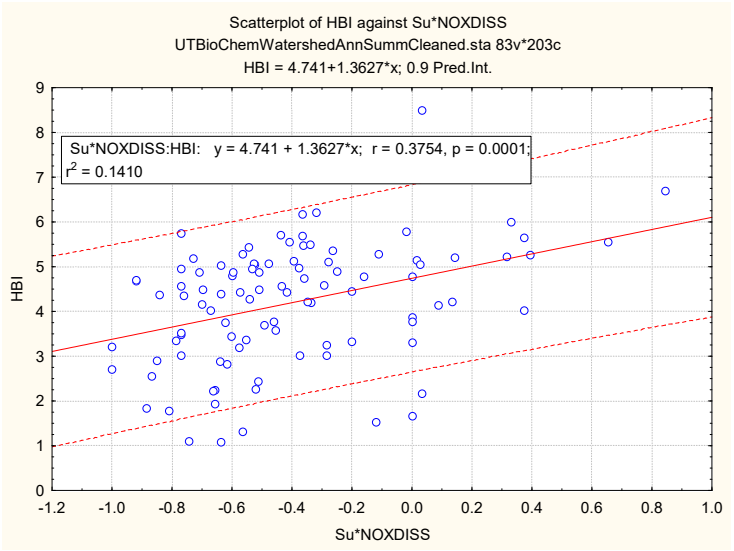


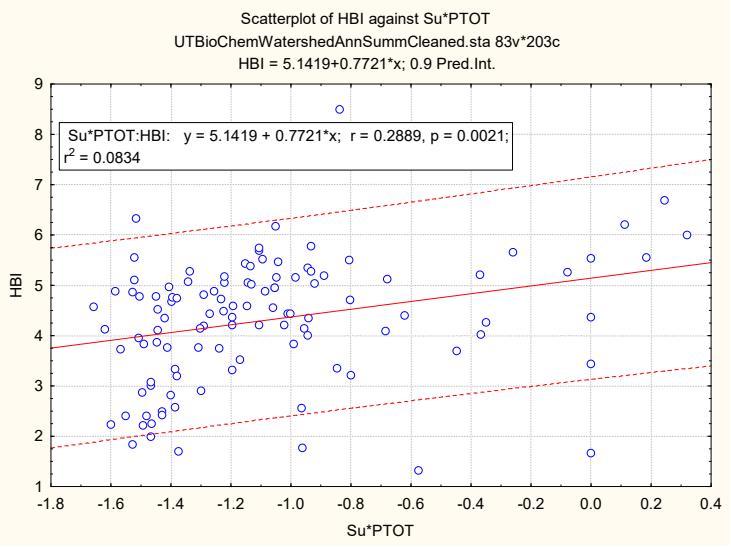




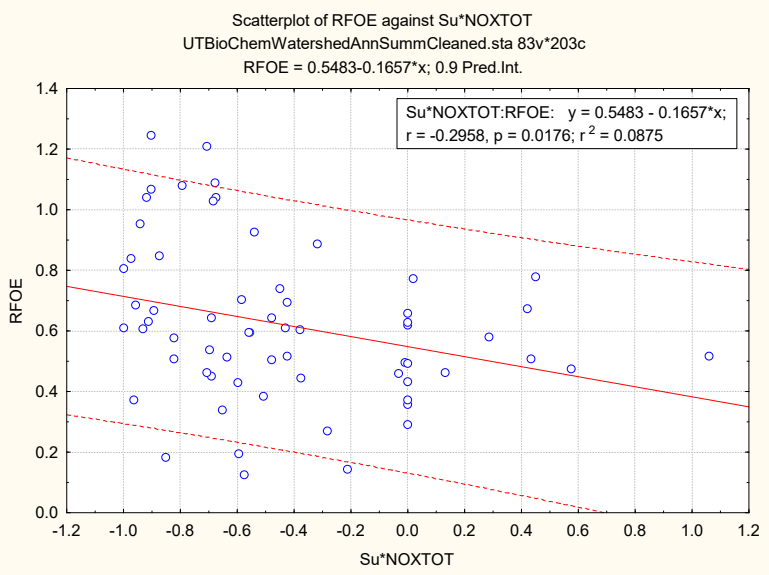
HBI

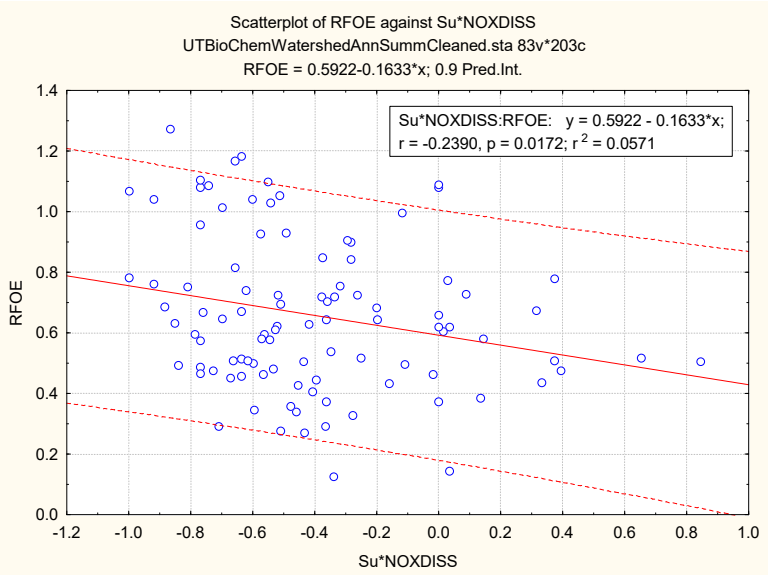

Scatterplot of HBI against Su*NOXTOT
UTBioChemWatershedAnnSummCleaned.sta 83v*203c
HBI = 4.8234+1.3667*x; 0.9 Pred.Int.
Su*NOXTOT:HBI: y = 4.8234 + 1.3667*x; r = 0.4601, p = 0.0001; r2 = 0.2117
HBI
Su*NOXTOT


Scatterplot of HBI against Su*NOXDISS
UTBioChemWatershedAnnSummCleaned.sta 83v*203c
HBI = 4.741+1.3627*x; 0.9 Pred.Int.
Su*NOXDISS:HBI: y = 4.741 + 1.3627*x; r = 0.3754, p = 0.0001; r2 = 0.1410
HBI
Su*NOXDISS


Scatterplot of HBI against Su*PTOT
UTBioChemWatershedAnnSummCleaned.sta 83v*203c
HBI = 5.1419+0.7721*x; 0.9 Pred.Int.
Su*PTOT:HBI: y = 5.1419 + 0.7721*x; r = 0.2889, p = 0.0021; r2 = 0.0834
HBI
Su*PTOT

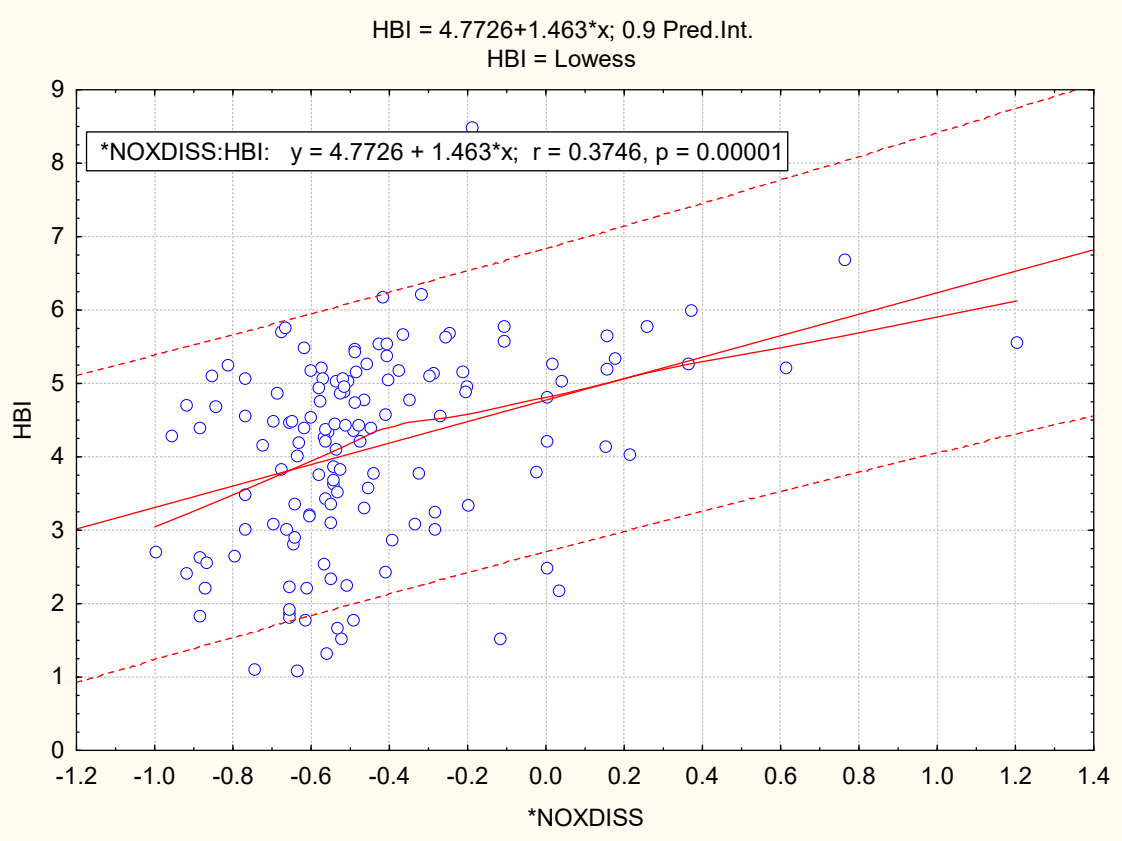
O/E Model


Scatterplot of RFOE against Su*NOXTOT
UTBioChemWatershedAnnSummCleaned.sta 83v*203c
RFOE = 0.5483-0.1657*x; 0.9 Pred.Int.
Su*NOXTOT:RFOE: y = 0.5483 - 0.1657*x;
r = -0.2958, p = 0.0176; r 2 = 0.0875
RFOE
Su*NOXTOT
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2


Scatterplot of RFOE against Su*NOXDISS
UTBioChemWatershedAnnSummCleaned.sta 83v*203c
RFOE = 0.5922-0.1633*x; 0.9 Pred.Int.
Su*NOXDISS:RFOE: y = 0.5922 - 0.1633*x;
r = -0.2390, p = 0.0172; r 2 = 0.0571
RFOE
Su*NOXDISS
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-1.2
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0

Appendix 3 - Endpoints based on change-point analyses: annual data

The following figures are scatterplots indicating non-linear relationships between annual nutrient variables [(Log_{10}total nitrate (Su*NOXTOT), Log_{10}dissolved nitrate (Su*NOXDISS)] and biological responses. Response variable abbreviations are as in Table 4 and regression plots are as in Appendix 1. Non-linear LOWESS fits were added as curvilinear solid lines to visualize non-linearities (tension = 0.8). Also shown are results of change-point analysis indicating percent reduction in model error and a visual tree plot with change-point in nutrient concentration indicated in log units (back-transformed units).

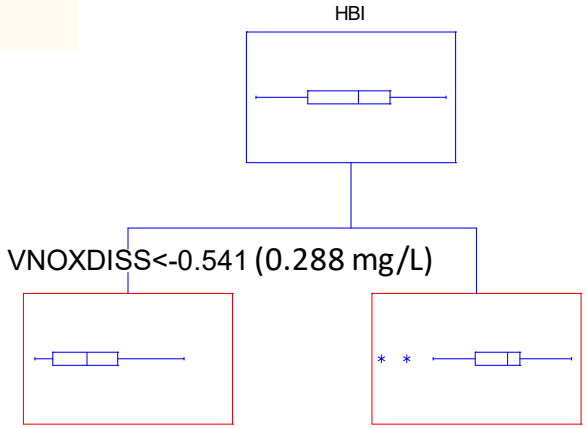


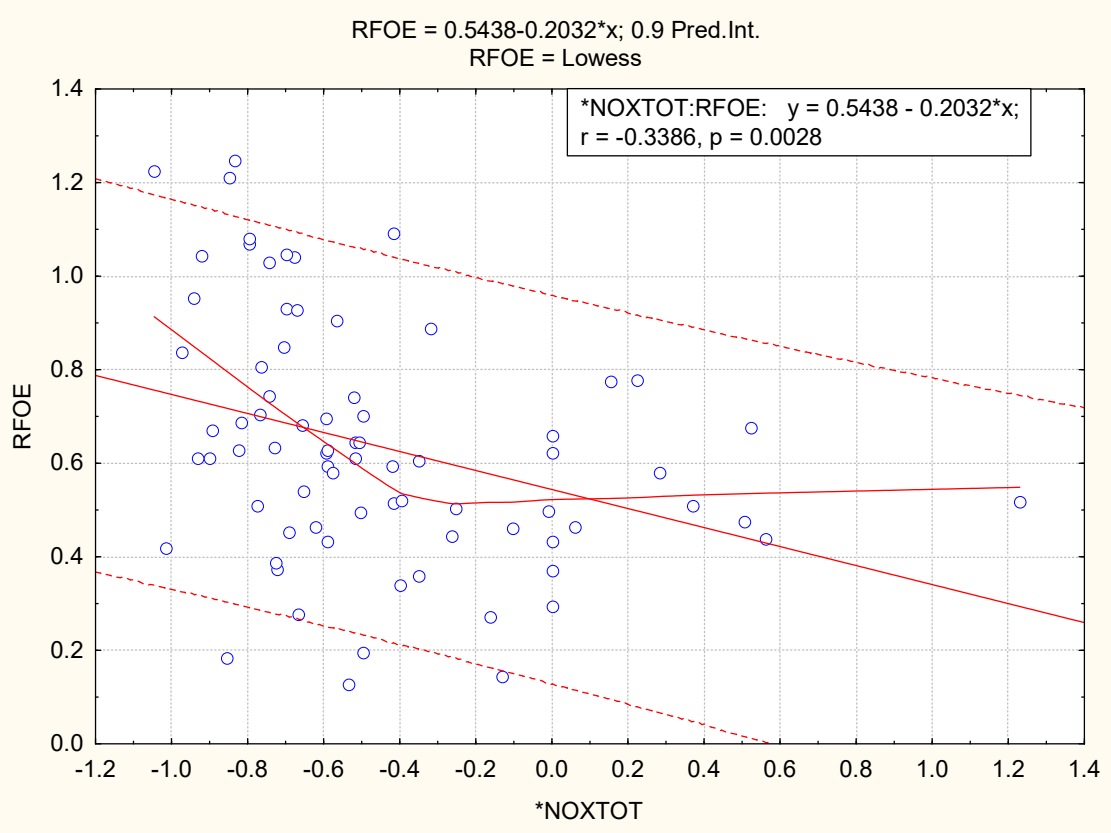
153 cases deleted due to missing data.

Split	Variable	PRE	Improvement
1	VNOXDISS	0.314	0.314

Fitting Method: Least Squares
Predicted variable: HBI
Minimum split index value: 0.050
Minimum improvement in PRE: 0.050
Maximum number of nodes allowed: 4
Minimum count allowed in each node: 5
The final tree contains 2 terminal nodes
Proportional reduction in error: 0.314

Node	from	Count	Mean	SD	Split Var	Cut Value	Fit
1	0	50	4.533	0.883	VNOXDISS	-0.541	0.314
2	1	18	3.880	0.722			
3	1	32	4.901	0.748			



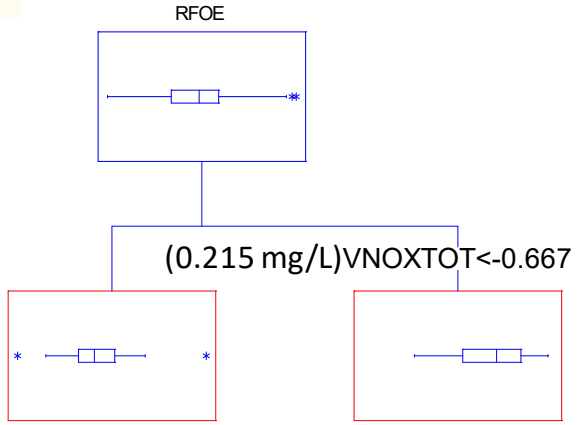


153 cases deleted due to missing data.

Split	Variable	PRE	Improvement
1	VNOXTOT	0.254	0.254

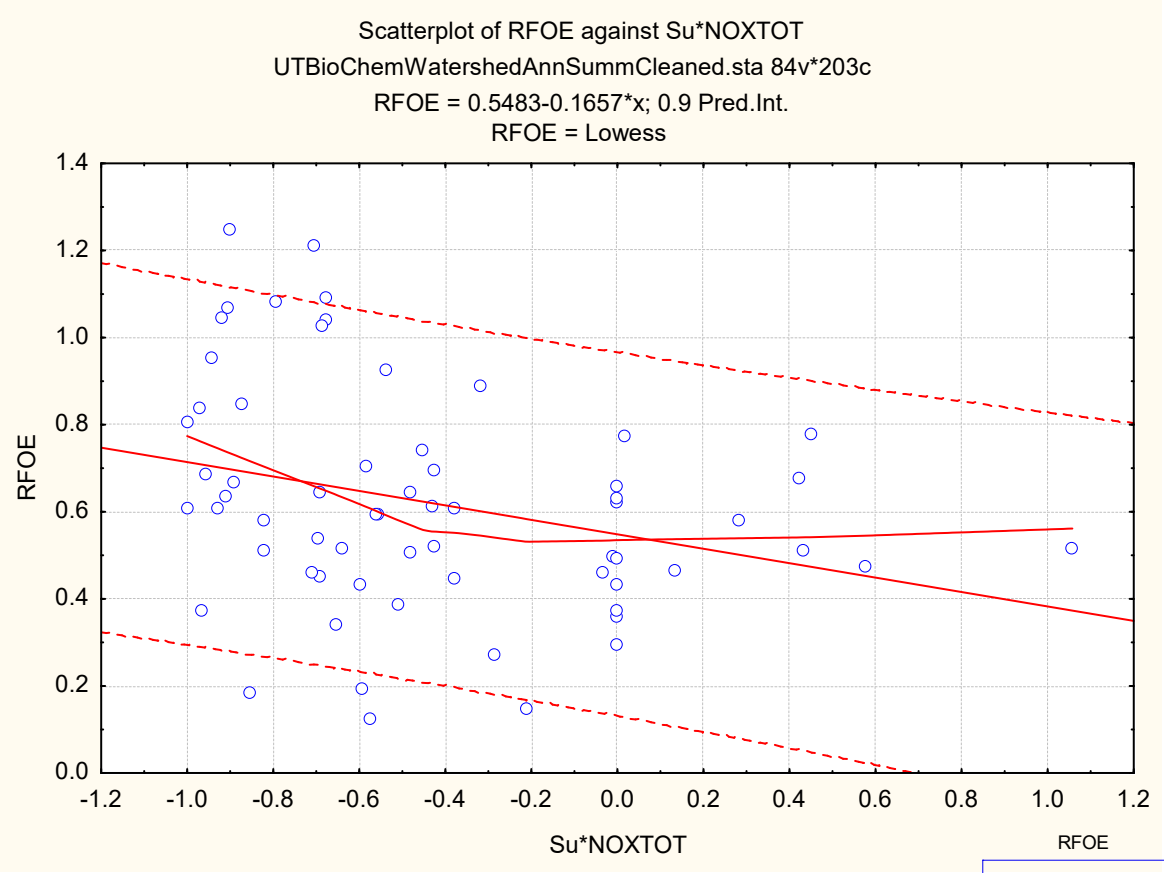
Fitting Method: Least Squares
Predicted variable: RFOE
Minimum split index value: 0.050
Minimum improvement in PRE: 0.050
Maximum number of nodes allowed: 4
Minimum count allowed in each node: 5
The final tree contains 2 terminal nodes
Proportional reduction in error: 0.254

Node	from	Count	Mean	SD	Split Var	Cut Value	Fit
1	0	50	0.599	0.215	VNOXTOT	-0.667	0.254
2	1	13	0.779	0.221			
3	1	37	0.535	0.175			



Appendix 4 - Endpoints based on change-point analyses: summer data

The following figures are scatterplots indicating non-linear relationships between summer nutrient variables [(Log_{10}total nitrate (Su*NOXTOT) and Log_{10}total phosphorus (Su*PTOT)] and biological responses. Response variable abbreviations are as in Table 4 and regression plots are as in Appendix 1. Non-linear LOWESS fits were added as curvilinear solid lines to visualize non-linearities (tension = 0.8). Also shown are results of change-point analysis indicating percent reduction in model error and a visual tree plot with change-point in nutrient concentration indicated in log units (back-transformed units).

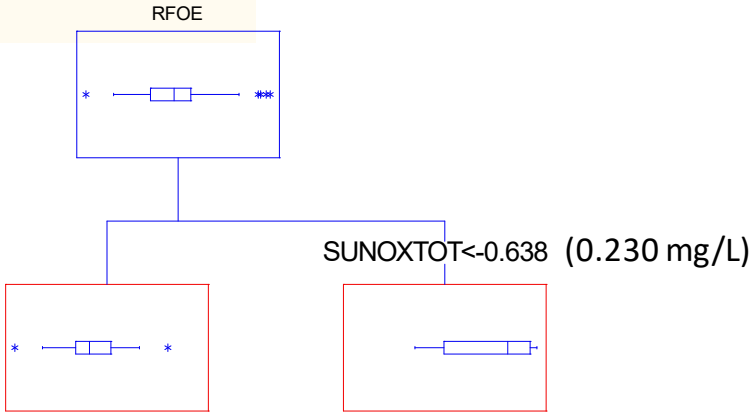


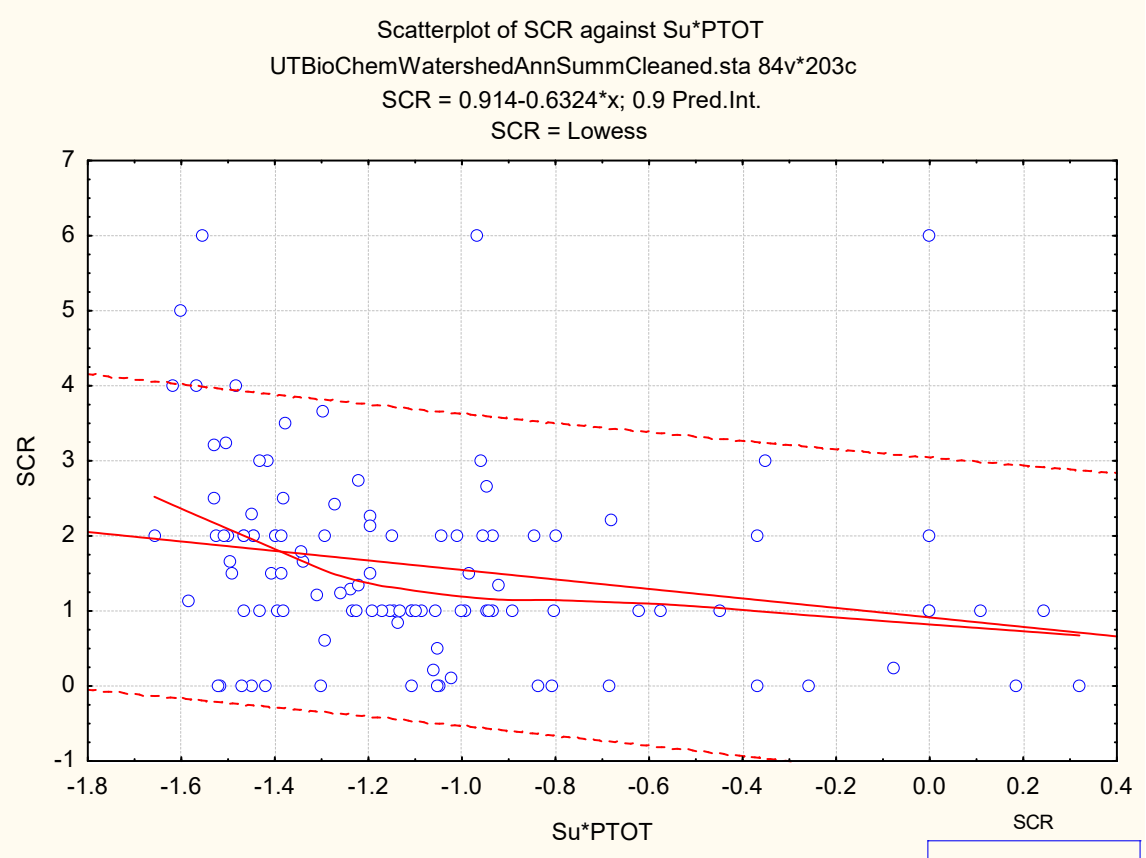
165 cases deleted due to missing data.

Split	Variable	PRE	Improvement
1	SUNOXTOT	0.323	0.323

Fitting Method: Least Squares
Predicted variable: RFOE
Minimum split index value: 0.050
Minimum improvement in PRE: 0.050
Maximum number of nodes allowed: 22
Minimum count allowed in each node: 5
The final tree contains 2 terminal nodes
Proportional reduction in error: 0.323

Node	from	Count	Mean	SD	Split Var	Cut Value	Fit
1	0	38	0.601	0.222	SUNOXTOT	-0.638	0.323
2	1	8	0.843	0.257			
3	1	30	0.537	0.164			



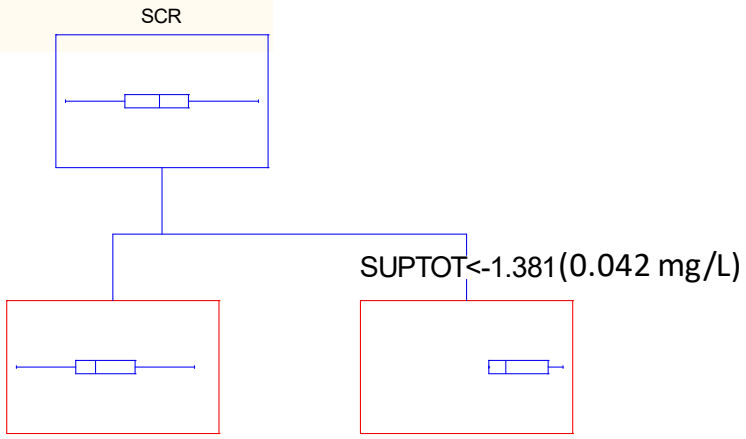


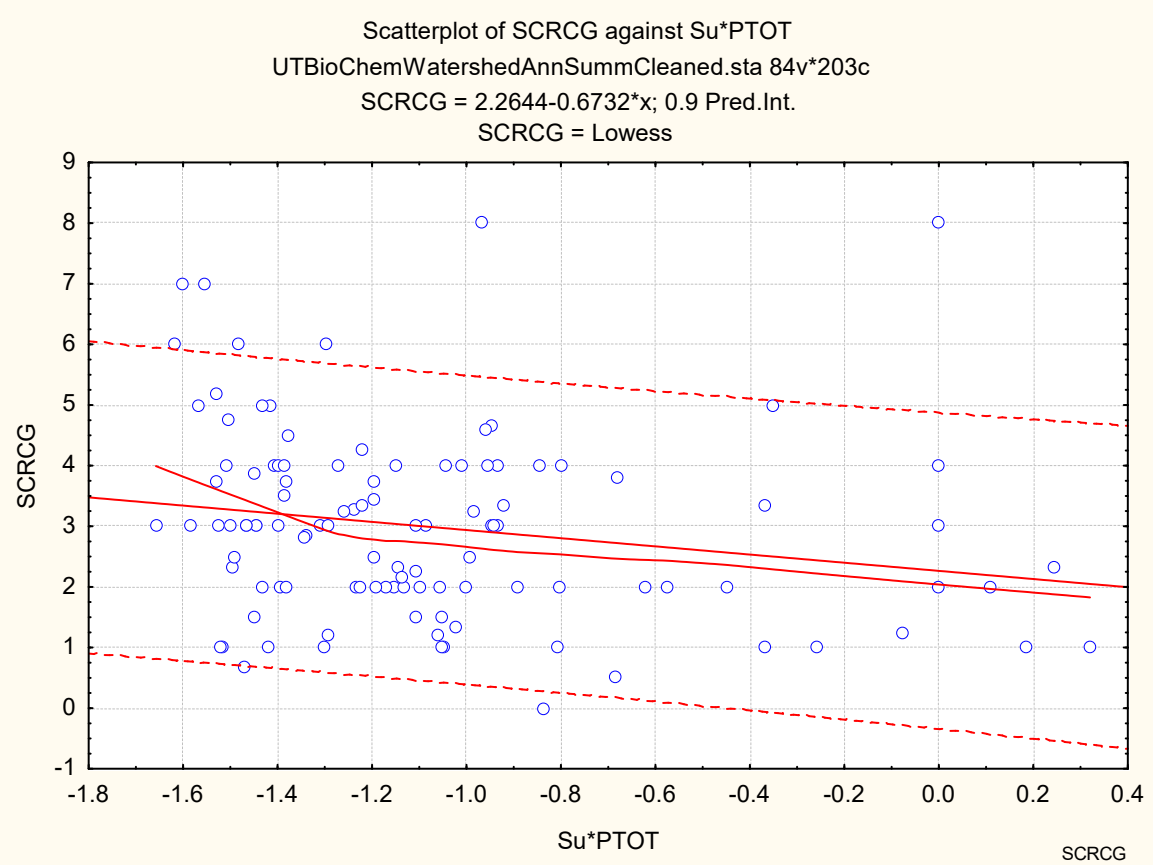
167 cases deleted due to missing data.

Split	Variable	PRE	Improvement
1	SUPTOT	0.201	0.201

Fitting Method: Least Squares
Predicted variable: SCR
Minimum split index value: 0.050
Minimum improvement in PRE: 0.050
Maximum number of nodes allowed: 4
Minimum count allowed in each node: 5
The final tree contains 2 terminal nodes
Proportional reduction in error: 0.201

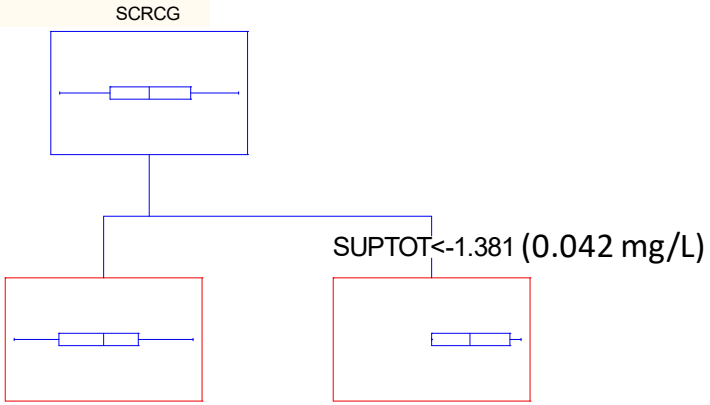
Node	from	Count	Mean	SD	Split Var	Cut Value	Fit
1	0	36	1.526	0.893	SUPTOT	-1.381	0.201
2	1	5	2.507	0.583			
3	1	31	1.367	0.836			





167 cases deleted due to missing data.

Split	Variable	PRE	Improvement
1	SUPTOT	0.103	0.103

Fitting Method: Least Squares
Predicted variable: SCRCG
Minimum split index value: 0.050
Minimum improvement in PRE: 0.050
Maximum number of nodes allowed: 4
Minimum count allowed in each node: 5
The final tree contains 2 terminal nodes
Proportional reduction in error: 0.103

Node	from	Count	Mean	SD	Split Var	Cut Value	Fit
1	0	36	3.021	1.143	SUPTOT	-1.381	0.103
2	1	5	3.921	0.942			
3	1	31	2.876	1.118