Refinement of Stream Nutrient Impairment Thresholds in New Mexico

New Mexico Environment Department
Surface Water Quality Bureau

JUNE 1, 2016

Overview

This document provides a summary of the process the New Mexico Environment Department (NMED) Surface Water Quality Bureau (SWQB) initiated and helped to complete in order to refine thresholds for plant nutrients in perennial, wadeable streams in New Mexico. This effort was necessary to apply the *State of New Mexico Standards for Interstate and Intrastate Surface Waters* narrative standard for plant nutrients found at 20.6.4.13 NMAC:

> ***E. Plant Nutrients:*** *Plant nutrients from other than natural causes shall not be present in concentrations that will produce undesirable aquatic life or result in a dominance of nuisance species in surface waters of the state.*

Narrative criteria must be translated to numeric thresholds for consistent impairment, NPDES permit limit, and TMDL budget determinations. This project follows United States Environmental Protection Agency (USEPA) nutrient criteria guidance (USEPA 2010) and Empirical Approaches for Nutrient Criteria Derivation (USEPA 2009). The goal at this time is to define thresholds for application of New Mexico's narrative nutrient water quality standard rather than numeric water quality criteria. Analysis was conducted to refine nutrient thresholds using regional data, defined reference conditions, relationships between cause and response variables, and a verified classification system.

Staff from Tetra Tech, Inc.; USEPA Region 6; USEPA's Office of Water, Office of Science and Technology; and SWQB worked as a team to complete this project. With input and directions from the workgroup Ben Jessup of Tetra Tech conducted most of the analysis and drafted the final report (Jessup et al. 2015). The work was supported by USEPA's Office of Water, Office of Science and Technology, through the Nutrient Scientific Exchange and Partnership System (N-STEPS) administered by USEPA's National Nutrient Criteria Program.

The results of this project will be used to revise the current perennial, wadeable stream nutrient assessment protocol (NMED/SWQB 2015) for development of subsequent Integrated Lists. Revision of the assessment protocol and associated thresholds was needed to better define nutrients from "other than natural causes" and to associate nutrient concentrations with the impairment of designated uses and to identify thresholds that filter out impaired systems.

The analysis consisted of two major approaches: reference conditions and stressor-response relationships. The reference condition approach derived candidate thresholds from distributions of nutrient concentrations from least disturbed sites which are the best estimate of "natural" conditions. Stressor-response analyses derived candidate thresholds by defining the relationships between total nitrogen (TN) and total phosphorus (TP) concentrations (i.e., causal variables) and response variables. Diatom and benthic macroinvertebrate metrics, and dissolved oxygen (DO) and chlorophyll a (chl-a) concentrations and metrics were the response variables used in this anaysis. Response variables represent the relative integrity of the aquatic community and indicate when designated aquatic life uses are being protected and not producing "undesirable aquatic life" or "dominance of nuisance species."

Data were collected within New Mexico and in the surrounding areas through NMED and national monitoring programs, including the National Rivers and Streams Assessment (NRSA), the Wadeable Streams Assessment (WSA), and Environmental Monitoring and Assessment Program (EMAP). A GIS analysis of sites and their catchments was conducted to characterize environmental conditions for use in disturbance gradient designations and site classification. All data were compiled in a relational database. Screening sites for data integrity and completeness resulted in 663 sites with nutrient data for one or more samples collected between 1990 and 2012. Other types of data (diel DO, chl-a, macroinvertebrates, and diatoms) were available for subsets of those sites. All data were screened for outlier values and nutrient values were standardized to common detection limits.

The reference site and disturbance gradient analysis of 542 sites resulted in 31% of sites being identified as least disturbed reference sites. Analysis of reference sites was used to determine site classes based on nutrient conditions. For nitrogen, concentrations were associated with average watershed land slope, and three nutrient classes were identified as TN Flat, TN Moderate, and TN Steep sites. For phosphorus, soil TP and volcanic geology were important in addition to land slope, resulting in three different nutrient classes identified as TP High-Volcanic, TP Flat-Moderate, and TP Steep. Frequency distributions of nutrient conditions in reference sites were used to derive TN, TP, and Delta DO candidate thresholds for each site class.

Correlation and other multivariate techniques supported the major linkages between nutrients, chl-a, DO, diatoms, and macroinvertebrates. Chl-a relationships supported some causal linkages between nutrients and DO but were too weak and variable to support its use as indicator of nutrient impairment. Regression interpolations and change-point analysis for macroinvertebrate, diatom, and DO metrics in response to nutrient concentration resulted in multiple candidate TN and TP thresholds in each site class. For each nutrient and site class, candidate thresholds from all analyses were evaluated and the selected values are shown in the table below:

Table 1. Proposed TN, TP, and DO thresholds

Nutrient Site Class	TN (mg/L)	TP (mg/L)	Delta DO (mg/L)
TN Flat	0.65	-	-
TN Moderate	0.37	-	-
TN Steep	0.30	-	-
TP High-volcanic	-	0.084	5.02
TP Flat-Moderate	-	0.061	4.08
TP Steep	-	0.030	1.79

This document provides a summary of the nutrient threshold development process and is excerpted from the final report Prepared by Tetra Tech, Inc. in cooperation with NMED, USEPA, and the N-STEPS Program (Jessup et al. 2015). The entire 100+ page final report details the results of each analysis, and is available on the SWQB web site at: https://www.env.nm.gov/swqb/Nutrients/.

Table of Contents

1. Selecting and Evaluating Data

Water quality data were compiled from multiple sources: NMED monitoring programs, the National Rivers and Streams Assessment (NRSA) and the Wadeable Streams Assessment program (WSA) (Table 2). Most monitoring took place in the critical low flow index period from August to November. Samples used in this project were collected between 1990 and 2013. The study area included the state boundaries of New Mexico as well as regions in adjacent states in a level 3 ecoregion that also exist within New Mexico. For most ecoregions, sites within 50 miles of New Mexico were included. To the north and east of New Mexico, sites from further away (up to 150 miles) were included.

NMED provided four types of data for this analysis: nutrients and other chemical parameters, periphyton, macroinvertebrates, and diel dissolved oxygen. Periphyton data consisted of both biomass as benthic chlorophyll a (benthic chl-a) and diatom community composition (diatoms). NMED monitored four of USEPA's recommended primary water quality variables in New Mexico streams (TN, TP, benthic chl-a, and turbidity) plus a number of secondary variables including DO concentration, DO percent saturation, specific conductance and pH (NMED/SWQB 2008). NMED collected and processed samples in accordance with methods documented in an USEPA approved Quality Assurance Project Plan (QAPP) https://www.env.nm.gov/swqb/QAPP/index.html and associated Standard Operating Procedures (SOP) https://www.env.nm.gov/swqb/SOP/index.html . For the NRSA projects, benthic macroinvertebrates, periphyton, and both benthic and sestonic chl-a data were available (USEPA 2004). For the WSA data, benthic macroinvertebrates were the only biological data available. All data were compiled in a single relational database (Microsoft Access), though data from each source were maintained in separate database tables (Tetra Tech 2014). Screening sites for data integrity and completeness resulted in 663 sites with nutrient data for one or more samples collected between 1990 and 2012. Other types of data (diel DO, chl-a, macroinvertebrates, and diatoms) were available for subsets of those sites.

1.1 Water Quality Data

The NMED nutrient records were most numerous and complete, while other types of data were relatively sparse. The NMED nutrient database included more than 7,000 records of nutrient concentrations from 883 stream sites. Four primary water quality variables were monitored in New Mexico streams (TN, TP, benthic chl-a, and turbidity) plus a number of secondary variables including DO concentration, DO percent saturation, and pH. TN was calculated by adding total Kjeldahl nitrogen (TKN), and nitrate plus nitrite (NO3+NO2). All nutrient data were screened for outlier values then standardized to common detection limits (0.03 mg/L for TP and 0.1 for both NO3+NO2 and TKN). Based on review of scientific literature, limited data analysis, and review of analysis of similar data sets, half detection limit substitutions were used for all analyses (Antweiler and Taylor 2008, Helsel 2010). Approximately 150 NMED nutrient records were identified as outliers and removed from analytical data sets. Many outliers were associated with storm flows and fire runoff.

Table 2. Data summary by source.

NMED:	*883 valid sites in NM with water chemistry (targeted sampling design)* *Multiple samples per site (approximately 7352 samples)* *Years 1990 - 2012* *Chemistry, site & habitat characteristics (partial data depending on site and visit)* *Benthic macroinvertebrate samples in 202 sites (440 samples)* *Periphyton (diatoms) in 212 sites* *Benthic chl-a in 146 sites* *Diel dissolved oxygen data in 175 sites*
NRSA:	*88 sites, each with a single visit (probabilistic sampling design)* *Years 2008 - 2009* *44 sites in NM, others within 50-150 miles of NM* *Chemistry, benthic & sestonic chl-a, periphyton (diatoms), site & habitat characteristics at each site*
WSA:	*56 sites, each with a single visit (probabilistic sampling design)* *Years 2000 - 2004* *10 sites in NM, others within 50-150 miles of NM* *Chemistry, benthic macroinvertebrates, site & habitat characteristics at each site*

1.2 Chlorophyll a

Of the NMED wadeable stream sites with nutrient data, 174 also had benthic chl-a data (including 35 with benthic macroinvertebrate data as well). These samples were collected between 2004 and 2011 in the months of August to November. Chl-a data were also collected for 50 NRSA sites, including both benthic and water column measures.

1.3 Periphyton Data

Periphyton data in and around New Mexico were collected by NMED and the NRSA. NMED collected roughly 212 diatom samples from 2002 to 2008 mostly in the fall sampling season (August - November). Samples were collected using a targeted richest habitat sampling method (NMED 2014). Periphyton data from 69 NRSA sites in and around New Mexico were added to a single periphyton database. Approximately 68 diatom metrics were calculated including metrics and taxa attributes described by Porter et al. (2008), Stevenson et al. (2008), and periphyton indices developed by Potapova and Charles (2007). Potential bias that might be introduced by different sampling protocols was investigated by comparing metric distributions.

1.4 Benthic Macroinvertebrates

NMED macroinvertebrate samples were collected using primarily four different methods, including reachwide, EMAP targeted riffle, kicknet from riffles, and Hess from riffles. Biomonitoring samples were collected in accordance with the USEPA Rapid Bioassessment Protocol (RBP) (Barbour et al. 1999), the NMED Standard Operating Procedures (SOP) (NMED/SWQB 2015) and/or modified USEPA EMAP macroinvertebrate sampling method (Peck et al. 2006). Opportunities to aggregate samples collected by different methods were explored and samples from multiple methods were pooled when the results of each method overlapped in stressor-response bi-plots. Samples methods with non-overlapping data points in the bi-plots were not used. NRSA and WSA benthic data were collected with consistent reachwide or targeted riffle methods (Peck et al. 2006) and were summarized as metrics in spreadsheet format. In the WSA and NRSA datasets, 56 and 40 benthic samples (respectively) matched chemistry samples.

1.5 Diel Dissolved Oxygen

Diel dissolved oxygen data were collected by NMED in stream sites throughout New Mexico between June and October (mostly August and September) from 2001 to 2012. These data were collected along with pH, specific conductance, temperature, and turbidity using multi-parameter, continuous recording sondes with recording periods of at least 48 hours and recording intervals ranging from 15 – 60 minutes. The data from approximately 200 spreadsheets were combined into a single data set so that metrics could be calculated efficiently. Data were checked for errors and data points or whole records were revised or eliminated if they were perceived to be inconsistent. There were four record sets with minimum DO greater than 10 mg/L that were removed as outliers.

After QC, statistics on 175 diel DO records were calculated. Diel DO statistics were related to nutrients in 133 sites. Diel DO and chl-a measurements coincided in 64 sites. Numerous metrics were calculated for each DO record including overall minimum DO, maximum daily fluctuation (Delta DO), gross primary production (GPP), ecosystem respiration (ER), and standard distribution statistics. NMED provided metrics on the maximum productivity (Pmax) and respiration (Rmax) in each data set based on 2, 3, and 4 hour intervals. The 4-hour interval was used for this analysis. Distribution metrics were also calculated for DO percent saturation data. In addition, system metabolism was calculated as GPP and ER, which accounted for temperature, elevation, and estimates or derivations of barometric pressure, nighttime regression, and light exposure. The calculations were carried out in R software using code provided by Dr. Robert Hall (Department of Zoology and Physiology, University of Wyoming, Laramie, WY).

2. Defining Human Disturbance Gradient

Site characteristics were observed or measured in the field, or derived from GIS analysis. The observed or measured data were recorded during site visits and included physical habitat assessments, channel dimensions, slope, canopy cover, riparian vegetation, riparian integrity, substrate characterizations, flow, and more. Each data source (NMED, NRSA, and WSA) included somewhat different variables for the observed and measured site characteristics. For NMED, habitat and flow variables were not collected at each site, but were often associated with benthic macroinvertebrate samples.

GIS analysis was conducted on a subset of 660 sites prioritized based on data completeness and potential reference site status. The information summarized from the GIS analysis includes land use, human activities, and environmental characteristics that are appropriate for reference site designation and site classification. These data include information on the setting of the sampling site and surrounding areas, such as ecoregion, average land slope, land use types and intensity, roads and road crossings, population density, watershed area, and more (Table 3).

Table 3. Variables used in GIS analysis.

Variable	Description
Point Values	
Stream Slope	NHD Plus join with flowline attributes table
Stream Order	NHD Plus join with NHDFlowlineVAA table
Elevation (cm)	NHD Plus DEM files
Designated Use	RAD 305b Assessed Segments joined with ATTAINS data
Precipitation	PRISM
Temperature	PRISM
Level 3 and 4 Ecoregions	USEPA Ecoregions
Geology	USGS Integrated Geological Map
Watershed Values	
Road density	Attila tool and TIGER 2000 files
Number of road/stream crossings	ARCGIS tools
Land Slope	ARCGIS Spatial Analysis Slope tool
Land Use and Cover	Attila tool and NLCD 2006 data
Canopy Density	Attila tool and NLCD 2001 Canopy data

Reference sites were needed for characterizing the nutrient conditions in the absence of substantial disturbance. This allowed exploration of natural variation in nutrient concentrations across the study area and for derivation of potential nutrient thresholds from distributions of nutrient values in the least disturbed sites. Stream classification and reference site designations hinged upon each other to characterize nutrient conditions relative to both natural and disturbance gradients.

Land use coverage and human activity in the catchments were examined for appropriate thresholds to indicate different levels of disturbance. Development and agricultural land uses indicate catchment scale intensity of disturbance. Both pasture and crops were considered agricultural uses. Forest, water, and wetland are usually undisputable natural land covers. However, the "natural-ness" of scrub/shrub, grassland, and barren coverages are uncertain because they could be due to human activities or natural environmental factors, especially in more arid areas. Therefore, the known disturbances were emphasized. Road density and the density of road-stream crossings were used as a surrogate for intensity of human activity in the watershed. Known human activities in the catchment (dams, NPDES permits, Superfund sites, and mines) were used to qualify reference sites. Information on these activities were available as counts in the catchment, densities (counts/catchment area), and distance to the sampling site.

For each reference criteria variable, thresholds were established for five disturbance categories from reference to extremely stressed sites (Table 4). The thresholds were derived from distribution statistics for each criterion in all sites. The percentiles were used as guidelines for establishing thresholds, but subjective adjustments were made to arrive at feasible values and adequate numbers of sites in each disturbance category. Five disturbance categories were defined from best to worst conditions: Reference, Near-Reference, Other, Stressed, and Extremely Stressed. To receive reference status, a site must not fail any of the Stressed criteria and must pass at least 7 of the 8 Reference criteria. Near-Reference sites did not fail any of the

Stressed criteria and passed at least 7 of the 8 Near-Reference criteria. Stressed and Extremely Stressed sites failed at least 2 of the Stressed or Extremely Stressed criteria, respectively. Sites that did not fall into any of these categories were classified as Other.

Table 4. Reference and stressed site criteria, based on distributions of values over all 660 sites.

Variable	Reference	Near Reference	Stressed	Extreme Stress
Urban Index (% cover)	0.01	0.02	1	2
Agricultural index (% cover)	0.1	0.5	4	5
Road Density (mi/mi^2)	1	1.4	3	5
Road Crossing Density (#/mi^2)	1	1.25	2	5
Dam Density (#/mi^2)	0	0.005	0.03	0.05
NPDES Density (#/mi^2)	0	0.01	0.1	0.2
Superfund Density (#/mi^2)	0	na	0.01	na
Mine Density (#/mi^2)	0.05	0.1	5	10
Dam Distance (mi)	na	na	1	0.5
NPDES Distance (mi)	na	na	1	0.5
Superfund Distance (mi)	na	na	2	1
Mine Distance (mi)	na	na	0.5	0.25

Because the NMED staff are familiar with site conditions that may not be reflected in the GIS data, they reviewed the reference designations indicated through empirical analyses and made changes to designations based on knowledge of the sites. For example GIS coverages do not reflect the intensity of grazing. Most of the designations (83%) assigned by numeric site criteria based on GIS analysis of land use coverage and human activity in the catchments were confirmed during the NMED review.

The reference site analysis and disturbance gradient designations resulted in 20% of sites identified as least disturbed reference sites. Another 11% were designated as near-reference. This is a reasonable proportion of reference sites because sites with potential for least-disturbed reference status were targeted when selecting sites for sampling and analysis. Smaller percentages of sites were designated as stressed (7%) or extremely stressed (6%). The remaining sites were designated as "other", having moderate levels of disturbance. The reference and near-reference sites were combined and used as the lease disturbed sites for further analysis after confirming that the nutrient distributions were similar.

3. Forming Site Classes

Natural gradients that affect potential nutrient and biological response indicators were examined. Appropriate statistical methods (e.g., principal components analysis, correlation analysis, and examination of bi-plots and distributions, etc.) were used on minimal disturbed sites to develop a stream classification scheme that captures environmental variability. Aggregate ecoregions used in the EMAP-West study (Stoddard et al. 2005)—Mountains, Plains, and Xeric—were considered as a starting point for stream classification, and considered along with other categorical and continuous variables. The classification scheme developed for sediment assessments – Mountains, Foothills, and Xeric areas (Jessup et al. 2014) was also tested. Additional classification categories and variables were examined, including Level III and IV ecoregions (Griffith 2006), geology, latitude, longitude, stream order, elevation, drainage area, average land slope in the catchment, average annual precipitation, average annual temperature, width/depth ratio, entrenchment ratio, sinuosity, channel substrate, and stream slope.

3.1 Principal components analysis

Principal components analysis (PCA) was used as a tool for selecting site classification variables (Table 5). The PCA was run in two configurations: reference and near-reference sites, and all sites. Nutrient-related axes that were correlated with biotic variables were examined to gain insight into potential scaling or classification variables that would minimize biological variability and thus focus biological responses on disturbances. Variables were transformed as needed to approximate normal distributions using logarithmic and Arcsine-Square Root transformations. Ecoregion designations and other classification variables were entered as binary code (true or false).

Table 5. Classification variables.

Code	Description	Type
Latitude	Latitude	Continuous
Longitude	Longitude	Continuous
Elev_m	Elevation (m)	Continuous
DrAreaMi2	Drainage area (square miles) (log transformed)	Continuous
LndSlpAvgpct	Average land slope (%)	Continuous
PrecipAvg30	30 year average precipitation (mm)	Continuous
TempAvg30	30 year average air temperature (C)	Continuous
NMEDnutClass	NMED existing nutrient classes	Categorical
MFX	Mountain, Foothill, and Xeric classes	Categorical
Ecoreg3	Level 3 ecoregion	Categorical
GeolRockType1	Geologic rock type	Categorical

3.2 Correlation analysis

Correlation analysis was used to describe single factor relationships between nutrients and environmental variables in reference sites. In contrast to the PCA, the correlation analysis was always limited to reference sites to emphasize the effects of natural site conditions instead of disturbance levels. The non-parametric Spearman rank order correlation coefficient was used because it is less sensitive to skewed distributions. The relationships suggested by PCA and correlations were examined in box plots and bi-plots. Bi-plots were used to show patterns of relationships between variables and to highlight tertiary attributes of the relationships such as reference status, ecoregion, or other covariates.

Preliminary analysis indicated nutrient data could be pooled across data sources (NMED, NRSA, and WSA) and that the reference and near-reference sites should be combined for the remaining classification exercises. Existing classification schemes based on level 2 ecoregions (Stoddard et al. 2005) or sediment regions (Jessup et al. 2014) showed insufficient separation of nutrient distributions and a determination of a new classification scheme specific to nutrient condition was warranted.

3.3 Classification and Regression Tree

Classification and Regression Tree (CART) models (also called recursive partitioning) account for variation in a dependent variable by progressively splitting samples into two bins that best partition the total variation among samples. This process forms a prediction tree based on a series of binary splits in the data. The first split occurs at the value of the predictor variable that most efficiently (as measured by the mean within-group standard deviation [SD]) partitions overall variation of the dependent variable into two groups. CART then partitions each of these two groups, if justified, into two smaller groups or nodes in the same manner, although the partitioning variable may differ. CART models were built with the R routine, 'rpart' (version 3.0.2; R Development Core Team, http://www.r-project.org/), for both TP and TN. The splits can inform site classification – giving variables and thresholds that partition the data by nutrient levels. At the end of each branch of the tree, TN or TP values are predicted as the average for that group.

A random forest routine (R: random Forest) was conducted to find the most important variables in 500 runs of CART using random subsets of the data for each run. Importance can be used to confirm the selection of variables in the predictive CART model. At first, only continuous variables (Table 5) were used to predict splits relative to site average TP and TN (log transformed) in reference and near-reference sites. Categorical variables were added to the models to determine whether existing classifications were as strong as the quantitative variables.

3.3.1 Phosphorus

Phosphorus values were first partitioned by longitude in the CART models both with and without categorical variables. The split was defined at longitude -108.13 which is the approximate watershed boundary between the Rio Grande and Gila River basins. Additional splits were based on average land slope, latitude, and precipitation. The random forest analysis suggested that the most important classifying variables were longitude (importance measure =

3.04), land slope (2.73), latitude (2.32), and precipitation (2.09). Land slope was the first split of CART analyses conducted separately for sites in western or eastern classes. Steeper sites have lower TP, in general (Figure 1). A CART analysis forcing land slope as the classification variable in all sites resulted in a split threshold of 29%.

NMED reviewed the initial classification analysis resulting in classes defined by longitude and land slope. The longitude split appeared to be driven by the large number of reference sites in the higher background TP watersheds in SW New Mexico. Higher background TP was suspected of being related to volcanic geology, but the geologic designations alone could not explain differences in reference TP. While all of the high TP reference sites were in volcanic regions, other volcanic formations did not have high background TP. Instead, specific basins (8-digit HUCs) were identified with high TP in reference sites. These watersheds correspond to those shown to have high soil TP (Woodruff et al. 2015).

Creation of the TP High-Volcanic site class resulted in 3 site classes for TP (Table 6). The TP High-Volcanic site class had more homogenous reference TP values than the class based on longitude and average land slope. The TP High-Volcanic site class includes the following basins: the Upper Gila, the Upper Gila-Mangas, the San Francisco, the Mimbres, and the San Antonio/Conejos. The San Antonio/Conejos is the only basin that is not in southwest NM. It is a volcanic region along the central section of the northern border of New Mexico. The following smaller basins (12-digit HUCs) were excluded though they are in the Upper Gila basin: Diamond, Taylor and Beaver Creeks (HUCs 150400010404, 150400010406, 150400010402, 150400010403, 150400010305, and 150400010302). The Jemez basin was suspected of being part of this class, but was not because background TP levels were not as high as in other TP High-Volcanic sites.

Sites not in the TP High-Volcanic class were separated into two classes based on 29% average catchment land slope. The TP Steep class has sites with slopes greater than 29% and background TP concentrations that were the lowest of the three classes. The TP Flat-Moderate class has flatter landscapes, though three basins with marginally flat sites (<31.8% land slope) were included because background TP concentrations were higher than typical TP Steep sites. These exceptions included drainages in the Vallecitos, Pajarito and Sulphur/Redondo basins (HUCs 130202020204, 130202010204, and 130202020202).

Table 6. Site classes for TP.

TP High-Volcanic –The class includes all sites in the San Antonio and Conejos, the Upper Gila, Upper Gila-Mangas, San Francisco, and Mimbres basins. In the Upper Gila basin, it excludes sites in the Diamond, Taylor and Beaver Creek sub-basins (HUCs 150400010404, 150400010406, 150400010402, 150400010403, 150400010305, and 150400010302).
TP Flat-Moderate - This class includes all sites less than or equal to 29% average land slope and not in the TP High-Volcanic site class. It also includes sites in three drainages of the Jemez basin, the Vallecitos, Pajarito, and Sulphur/Redondo sub-basins (HUCS 130202020204, 130202010204, and 130202020202).

TP Steep - The Steep class includes all sites with average land slopes greater than 29% and not in the TP High-Volcanic site class.

These three classes had significantly different TP values (Figure 1) based on the non-parametric Kruskal-Wallis test ($p<0.01$ for all comparisons).

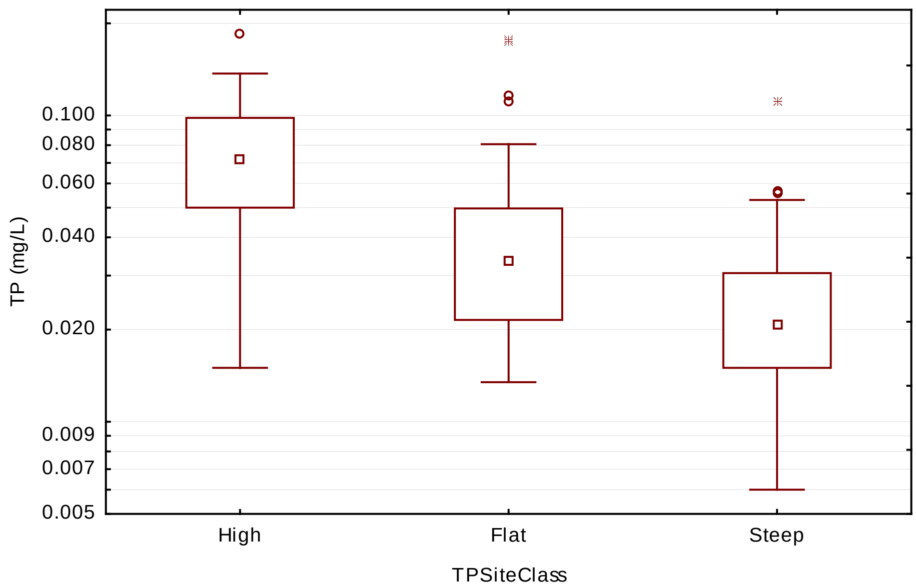


Figure 1. Total Phosphorus (TP) concentrations in reference or near-reference sites by potential site classes for TP. Sample sizes are 55, 76, and 48, in the order displayed.

3.3.2 Nitrogen

Similar to phosphorus, total nitrogen values were first partitioned by longitude in the CART models both with and without categorical variables. Importance coefficients in the random forest analysis were as follows: longitude (1.92), land slope (1.76), precipitation (1.57), latitude (1.44), and temperature (1.28). The split for longitude at -105.2, divided both ecoregions and watersheds. Land slope was explored as a classification variable because it was important in the random forest analysis. A CART analysis with all sites forcing only land slope in the model resulted in 2 splits at 15% and 32%. The western streams with the flattest landscapes were represented by only 12 sites. Because land slope appears to partition TN values as well as or better than longitude *and* land slope, classes were based on land slope alone (Table 7). This classification scheme resulted in distinct TN values within the classes (Figure 2).

Table 7. Site classes for TN.

TN Flat - TN Flat sites have average catchment land slopes less than 15%
TN Moderate - TN Moderate sites have average catchment land slopes from 15% to 32%

TN Steep - TN Steep sites have average catchment land slopes greater than 32%

The TN values were significantly different based on the non-parametric Kruskal-Wallis test. The differences in relation to the Flat class (p<0.001) were greater than the difference between the Moderate and Steep groups (p=0.03).

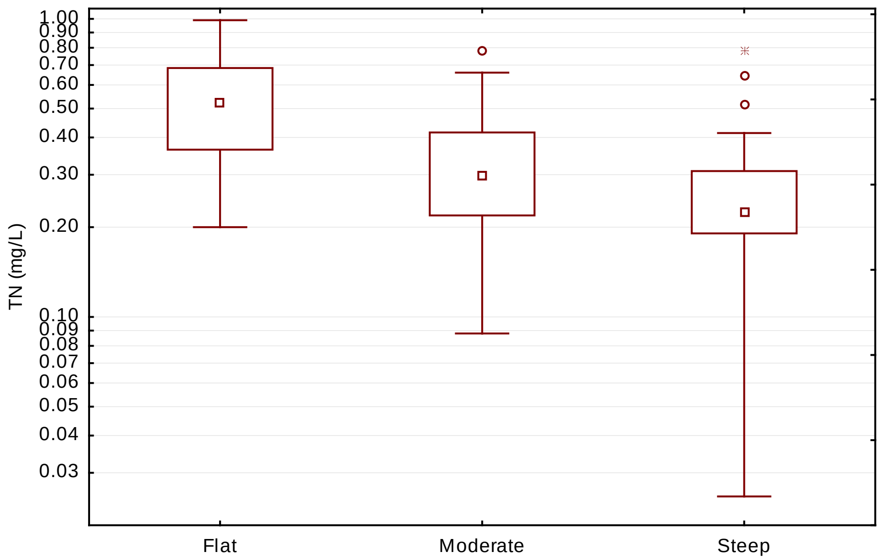


Figure 2. Total Nitrogen (TN) concentrations in reference and near-reference sites by potential site classes for TN. Sample sizes are 31, 95, and 51, in the order displayed.

Different site classes for TN and TP were not anticipated when classifying sites to partition nutrient variability. In the independent analyses for each nutrient, similar classification variables, longitude and land slope, were identified. Though identical site classes for TN and TP were attempted, separate classes for TN and TP were more precise and are appropriate for application of numeric nutrient thresholds.

4. Frequency Distributions

Once sites were divided into the different classes, the frequency distributions of these classes were analyzed. Non-parametric quantiles were calculated from frequency distributions of nutrient concentrations of reference and near reference sites divided into site classes. The distributions were based on median TN and TP concentrations at each site. NMED preferred to use the median values as the best representation of site conditions because the data were log normally distributed and the median was a better estimate of the central tendency of the data. Also, mean values can be biased by a few extreme values. The frequency statistics included data from all data sources for nutrients.

Frequency distributions of reference sites resulted in quantiles that were considered as candidate thresholds for TN and TP in data subsets by site class. Within site classes, the median, 75th, 80th, 85th, and 90th quantiles of the concentrations were determined to characterize the combined reference and near-reference sites (Table 8). To illustrate the validity of using reference quantiles to derive thresholds for the New Mexico data sets, distributions of TP, TN and benthic chl-a were plotted by site class and reference status. Nutrient concentrations generally increased with increasing disturbance. Confidence intervals (90%) were calculated for each quantile using 1001 bootstrap iterations. Analysis was conducted using R software.

NMED chose the 90th quantile to represent a starting point for candidate thresholds. Quantile selection for is dependent upon the data, and the certainty one has that they accurately reflect reference conditions. For this analysis, there was a high level of confidence in reference site selection and the 75th quantile did not seem to include naturally enriched systems. In most cases, the 90th quantile was more closely aligned with the benthic macroinvertebrate and diatom change point analyses, and is hence assumed protective of the applicable designated aquatic life use(s). However, lower quantiles were selected when the 90th quantile did not align with stressor response thresholds (highlighted in Table 8).

Benthic chl-a concentrations were evaluated in the TP classes and had only one observation per site. For benthic chl-a, median values in Stressed and Highly Stressed categories were consistently higher than medians in Reference and Near-Reference categories, though the stressed categories were represented by fewer than five samples in all but the TP Flat-Moderate class. The uneven distribution of benthic chl-a in the TN site classes did not allow for this type of analysis.

Table 8. Frequency distribution statistics for median TP and TN reference sites. The recommended candidate thresholds are highlighted.

	TP (mg/L)			TN (mg/L)		
Quantile	Lower 90% CI	Value	Upper 90% CI	Lower 90% CI	Value	Upper 90% CI
	TP High-Volcanic (N=55)			TN Flat (N=30)		
50th	0.049	0.058	0.071	0.38	0.47	0.56
75th	0.072	**0.084**	0.09	0.55	0.61	0.67
80th	0.08	0.088	0.104	0.56	0.62	0.7
85th	0.084	0.092	0.106	0.59	**0.65**	0.84
90th	0.089	0.105	0.114	0.62	0.69	0.85
	TP Flat-Moderate (N=76)			TN Moderate (N=96)		
50th	0.016	0.025	0.033	0.23	0.25	0.28
75th	0.034	0.041	0.05	0.33	0.35	0.37
80th	0.036	0.048	0.058	0.35	**0.37**	0.41
85th	0.043	0.054	0.061	0.36	0.40	0.45
90th	0.051	**0.061**	0.069	0.38	0.42	0.51
	TP Steep (N=48)			TN Steep (N=53)		
50th	0.015	0.015	0.015	0.18	0.20	0.21
75th	0.015	0.015	0.018	0.21	0.23	0.27
80th	0.015	0.016	0.023	0.22	0.25	0.3

85th	0.015	0.018	0.035	0.24	0.28	0.33
90th	0.016	**0.030**	0.053	0.26	**0.30**	0.34

Quantiles of Delta DO and Pmax4hr diel DO measures in reference and near-reference sites were similar in the High-Volcanic and TP Flat-Moderate site classes and were relatively lower in the TP Steep site class (Table 9). Delta DO and Pmax4hr increased with increasing stress in the streams, especially in the TP Flat-Moderate site class.

Table 9. Frequency distribution statistics for diel DO statistics in valid reference and near reference sites. The recommended candidate threshold (90th quantile) is shown in bold-type.

Quantile	Delta DO			Pmax4hr		
	Lower 90% CI	Value	Upper 90% CI	Lower 90% CI	Value	Upper 90% CI
	TP High-Volcanic			TP High-Volcanic		
50th	1.93	2.17	3.03	0.176	0.304	0.439
75th	2.17	3.27	4.29	0.331	0.501	0.648
90th	3.13	**5.02**	7.24	0.460	**0.635**	0.720
	TP Flat-Moderate			TP Flat-Moderate		
50th	1.22	2.28	3.52	0.148	0.208	0.393
75th	2.28	3.06	3.98	0.296	0.501	0.678
90th	3.52	**4.08**	7.26	0.493	**0.682**	1.200
	TP Steep			TP Steep		
50th	1.10	1.13	1.57	0.095	0.105	0.196
75th	1.10	1.57	2.37	0.105	0.186	0.490
90th	1.40	**1.79**	2.37	0.126	**0.284**	0.490

5. Evaluating Estimated Stressor-Response Relationships

Step 4 of Empirical Approaches for Nutrient Criteria Derivation (USEPA 2009) is Evaluating Estimated Stressor-Response relationships. A conceptual model is helpful in defining stressor-response relationships. Conceptual models were developed to represent known relationships between changes in TN and TP concentrations, biological effects, and attainment of designated uses. Conceptual nutrient models are well established and were not reconstructed for this study. Instead, the conceptual model published by USEPA (2010) was used as a standard that is applicable in New Mexico streams (Figure 3). The conceptual model shows intricate pathways of effects. It illustrates interactions that might be effective though our analytical data set is insufficient to account for them. Analyses that compared indirect elements in the conceptual model (e.g., relating nutrient concentrations to macroinvertebrate responses) relied on the validity of the intermediate linkages. The conceptual model provides a means of communicating the current state of knowledge regarding the effects of TN and TP in aquatic systems and is an important tool for guiding causal analyses.

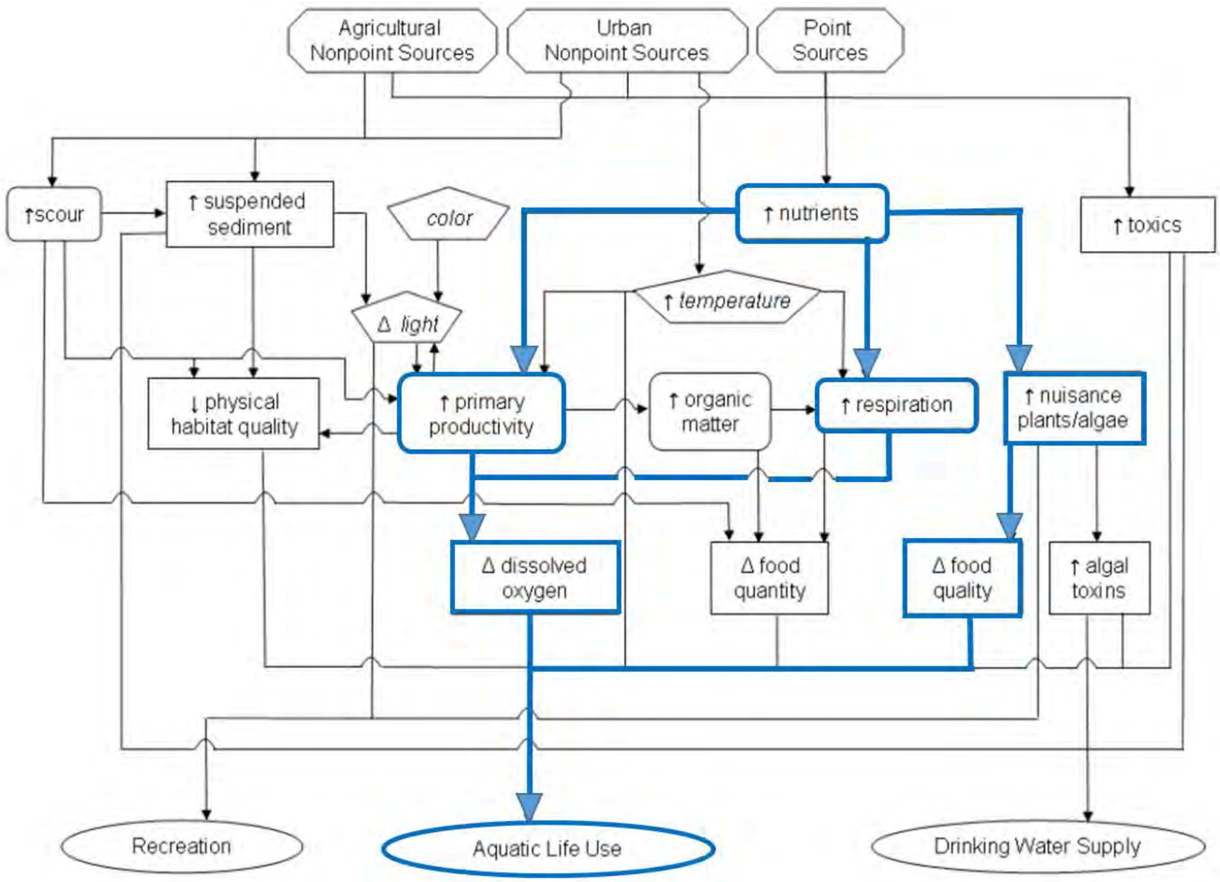


Figure 3. Conceptual diagram linking sources of human disturbance with designated uses through pathways that include nutrients (from USEPA 2010).

The linkages and relationships explored through stressor-response analysis are listed below. These relationships were analyzed as one-to-one stressor-response relationships, and as indirect relationships (e.g., nutrient-macroinvertebrate responses without intermediate links). The indirect relationships between nutrients – DO and nutrients – macroinvertebrate metrics were explored in addition to the direct nutrient – chl-a – DO – macroinvertebrate relationships.

1. ↑ Nutrients = ↑ Chlorophyll a
2. ↑ Chlorophyll a = Δ DO dynamics
3. ↑ Nutrients = Δ Diatom Metrics
4. Δ DO dynamics = Δ Macroinvertebrate Metrics
5. ↑ Nutrients = Δ DO dynamics

5.1 Correlations and Interactions

To strengthen understanding of the primary linkages in the conceptual model (nutrients - chl-a, chl-a - DO dynamics, nutrients - diatom metrics, and DO - macroinvertebrate metrics), each relationship was explored in further detail. Relationships were also examined to determine whether consistent modifiers could be factored out to refine the general relationship. The explorations included correlation analysis, multiple regression, CART, random forest, and graphic displays. Nutrient concentrations collected within 30 days of the response variable were averaged. Censored data were assigned a value of half the detection limit and retained for these analyses. In addition, Spearman rank correlation analysis for TN and TP and water quality variables (NO_3+NO_2, TKN, pH, conductivity, turbidity, temperature, and DO) showed positive relationships between the nutrients and between nutrients and other water quality measures; with pH and conductivity being more strongly correlated to TN, and turbidity and temperature being more strongly correlated with TP.

5.1.1 Chlorophyll a

A Spearman rank-order correlation analysis was conducted with TN and TP against benthic chl-a concentrations. For benthic chl-a, a total of 192 valid samples from NMED and NRSA sites were included in the analysis. On average, NRSA benthic chl-a concentration was less than NMED chl-a as NMED used a targeted richest habitat sampling, so the data were analyzed separately by source. Of all the Spearman correlations, only TP and chl-a in NMED TP High-Volcanic and all sites were significantly correlated (Table 11). The lack of significant correlations does not support the use of benthic chl-a as a reliable indicator of nutrient enrichment.

Table 11. Sample sizes (N) and Spearman rank correlation coefficients (rho) for benthic chl-a by nutrient and site class. Significant correlations ($p<0.05$) are marked with an asterisk (*).

	NMED		NRSA	
TP (Site Classes)	N	rho	N	rho
All sites	140	0.17*	50	-0.16
TP High-Volcanic	23	0.48*	13	-0.29
TP Flat-Moderate	90	0.16	28	-0.29
TP Steep	27	0.05	7	0.94
TN (Site Classes)				
All sites	142	0.02	50	0.09
TN Flat	26	0.04	24	0.28
TN Moderate	94	-0.04	21	0.13
TN Steep	22	-0.04	5	N/A

Positive correlations were expected between nutrients and benthic chl-a. However, a number of biotic and abiotic variables influence algal biomass accrual (Marks et al 2000). For example, hydrologic disturbances have a strong influence on algal biomass (Biggs 2000; Peterson et al. 1994). Other modifying factors (e.g., canopy cover, stream flow, drainage area, temperature, turbidity, sampling variability, etc.) may be confounding the assumed relationship between nutrients and benthic chl-a. Canopy cover, stream flow, elevation, land use types, drainage area, latitude, longitude, conductivity, temperature, pH, and turbidity were investigated. However the dataset did not have sufficient information to account for canopy cover and stream flow influences or other confounding factors.

5.1.2 Diel Dissolved Oxygen

A Spearman rank correlation analysis was conducted with TN, TP, and benthic chl-a against diel DO statistics. Production (Pmax4hr) and respiration (Rmax4hr) were negatively correlated to each other (Spearman rho = -0.92, p<0.05) as were gross primary production (GPP), and ecosystem respiration (ER) (Spearman rho = -0.55, p<0.05). TP was positively correlated with productivity measures and maximum daily change in DO (DeltaDO) (Table 12). Both TP and TN were negatively correlated with minimum DO (rho = -0.18, p<0.05). TN was also negatively correlated with Rmax4hr and positively correlated with Pmax4hr and DeltaDO (p=0.06). Benthic chl-a was positively correlated to DeltaDO, Pmax4hr, and ER. The correlation with ER was expected to be negative, as it was with Rmax4hr. The bi-plots show weak positive relationships between the nutrients and DeltaDO and Pmax4hr and weak negative relationships between nutrients and minimum DO. Minimum DO and chl-a are not strongly related. However, the relationships of chl-a with Pmax4hr and DeltaDO are somewhat stronger, supporting the causal linkage between chl-a and DO.

Table 12. Spearman correlation coefficients for TN, TP, and benthic chl-a versus diel DO statistics; minimum DO (DO_min), maximum daily DO change (DeltaDO), 4 hour maximum production (Pmax4hr), 4 hour maximum respiration (Rmax4hr), gross primary production (GPP), and ecosystem respiration (ER). Asterisk (*) denotes significant correlations (p<0.05).

	DO_min	DeltaDO	Pmax4hr	Rmax4hr	GPP	ER
TN	-0.18*	0.17	0.17	-0.19*	0.05	0.11
TP	-0.18*	0.30*	0.29*	-0.31*	0.19*	0.01
Benthic chl-a	-0.11	0.28*	0.38*	-0.25*	0.09	0.38*
DO_min		-0.58*	-0.50*	0.45*	-0.31*	0.04
DeltaDO			0.91*	-0.90*	0.53*	0.12
Pmax4hr				-0.92*	0.62*	0.08
Rmax4hr					-0.62*	-0.06
GPP						-0.55*

5.1.3 Diatoms

A Spearman rank-order correlation analysis of nutrient values associated with diatom metrics was conducted. Sixty-eight (68) diatom metrics were correlated with TN, TP, and potential modifying factors in 151 NMED sites and 49 NRSA sites. The analysis started with NMED sites only and then addressed NRSA sites and combined data sets. The data in the analysis were limited to one sample per site when nutrient and diatom samples were collected within 30 days of each other. Diatoms appear to be more sensitive to TP than to TN, based on the number of significant correlations. The fewest significant relationships were between the metrics and chl-a. Based on significant correlations and metric types, eight responsive metrics were selected for continued analysis in stressor-response analyses (Table 13).

Table 13. Diatom metrics showing responsiveness in correlation analysis and used in stressor-response analysis.

Metric code	Metric description	Metric type
wa_OptCat_DisTotMMI	Multi-metric index of disturbance	Weighted average, general disturbance
wa_OptCat_L1DisTot	Sum of disturbances	Weighted average, general disturbance
wa_OptCat_L1Ptl	Western EMAP TP score	Weighted average, TP
wa_OptCat_LNtl	Western EMAP TN score	Weighted average, TN
wa_OptCat_NutMMI	Western EMAP multi-metric index	Weighted average, nutrients
pi_NAWQA_TN_1	% TN tolerant diatoms	Percent Individuals, TN
pi_Ptpv_TP_all_Hi	% high TP diatoms, all regions	Percent Individuals, TP
x_Shan_e	Shannon-Wiener Diversity Index	Taxa diversity

5.1.4 Benthic Macroinvertebrates

A Spearman rank-order correlation analysis of nutrients, diel DO, and benthic chl-a associated with benthic macroinvertebrate metrics was conducted. The samples were limited to one per site when nutrient or chl-a and macroinvertebrate samples were collected within 30 days of each other. For diel DO statistics, the analysis was limited to DO and macroinvertebrate samples collected within the same season (within 80 days). Distributions of metric values collected with different sampling methods were overlapping in stressor-response biplots. Therefore, all wadeable stream samples were pooled, including multiple methods and data from NMED, WSA, and NRSA. Only early kicknet samples from NMED and low gradient samples from NRSA were eliminated. The data screening resulted in 438-440 samples for TP and TN, respectively, from 313 sites. For diel DO and benthic chl-a samples, there were 76 and 193 samples, respectively. Dissolved oxygen grab samples were collected along with macroinvertebrate samples with greater frequency than diel DO samples. However, because the variability inherent to DO over time, the DO grab data were not used.

5.1.4.1 Benthic Macroinvertebrate DO Correlations

The minimum DO (DO_min) and maximum 4 hour productivity (Pmax4hr) had the highest numbers of significant correlations in all sites (15-16 of 63 metrics, each). Other DO statistics had fewer significant correlations (GPP: 5, ER: 4, and Rmax4hr: 10). The strongest correlations for DO_min were positive with shredder taxa, shredder percent, and Plecoptera percent (Spearman rho = 0.39, 0.34, and 0.31, respectively). Other metrics with high positive correlations (rho = 0.30) included Ephemeroptera taxa, Beck's index, and intolerant percent. Gastropod percent and the HBI were negatively correlated (rho = 0.30). With productivity

(Pmax4hr), the strongest metric correlations were negative (rho = -0.37 - -0.42) with Plecoptera taxa, Plecoptera percent, Beck's index, and intolerant taxa.

5.1.4.2 Benthic Macroinvertebrate Chlorophyll Correlations

Benthic chl-a was correlated to 14 macroinvertebrate metrics in all sites. The strongest correlations were with the Shannon-Wiener diversity index and intolerant taxa (Spearman rho = 0.37 and -0.34, respectively). A positive correlation with Shannon-Wiener diversity and a positive correlation with Trichoptera taxa and percent Trichoptera indicates that higher benthic chl-a increases some aspects of the macroinvertebrate assemblage diversity.

5.1.4.3 Benthic Macroinvertebrate Nutrient Correlations

For TN, 42 of the 63 metrics were significantly ($p<0.05$) correlated with nutrient concentration in all sites. Fewer metrics were significantly correlated in each site class, with 28, 25, and 6 in the TN Flat, TN Moderate, and TN Steep classes, respectively. Only predator taxa and % predators were correlated in all site classes. The strongest correlations (all negative) were in the TN Flat and TN Moderate classes for the total taxa metric, EPT taxa metric, Beck's Biotic Index (weighted richness of sensitive taxa), and the clinger taxa metric. The relatively unresponsive metrics in the steep site class might be due to lower nutrient concentrations in general, with fewer values >0.5mg/L TN than in the TN Flat and TN Moderate site classes.

For TP, fewer correlations were significant and the strength was weaker, compared to the TN correlations. There were 21 significant ($p<0.05$) correlations when including all site classes. Correlations varied among site classes. In the TP High-Volcanic, TP Flat-Moderate, and TP Steep classes, 17, 12, and 6 metrics were significantly correlated, respectively. Macroinvertebrates appear to be more responsive to TN than to TP. The positive correlation between TN and TP was not very strong in this data set (Spearman rho = 0.28, $p < 0.05$).

Several benthic macroinvertebrate metrics were related to nutrients, benthic chl-a, and DO, but only ten were selected for ongoing analyses to simplify interpretation of the stressor-response relationships. Responses for 19 candidate metrics were qualified as strongly positive, positive, negative, or strongly negative in relation to multiple measures of nutrients, benthic chl-a, and DO. The ratings were based on correlation coefficients in all sites and in the individual site classes or methods. Ten benthic macroinvertebrate metrics with consistent and strong correlations were identified (Table 14, bold font).

Table 14. Qualitative response trends for macroinvertebrate metrics to nutrients, benthic chl-a, and DO. The trends of responses were negative (Neg) or positive (Pos). Stronger relationships (more significant correlations in site classes) are shown in bold type.

		TN	TP	Chl-a	DO[a]	Overall
Richness	Total Taxa	Neg	Neg	Pos	Mix	Mix
	1 EPT Taxa	**Neg**	**Neg**	Mix	**Neg**	Neg
	2 Ephemeroptera Taxa	**Neg**	**Neg**	Neg	**Neg**	**Neg**

		TN	TP	Chl-a	DO[a]	Overall
	3 Plecoptera Taxa	**Neg**	**Neg**	**Neg**	**Neg**	**Neg**
	Trichoptera Taxa	**Neg**	**Neg**	Pos	Mix	Mix
	Shannon-Wiener Index	**Neg**	**Neg**	**Pos**	Mix	Mix
Composition	**4 EPT percent**	Neg	Mix	Mix	Mix	Neg/Mix
	Ephem percent	Mix	Mix	Mix	Mix	Mix
	5 Pleco percent	**Neg**	**Neg**	**Neg**	**Neg**	**Neg**
	Trich percent	**Neg**	**Neg**	**Pos**	Mix	Mix
	6 NonIn percent	**Pos**	Pos	Pos	Pos	**Pos**
Tolerance	**7 Intolerant Taxa**	**Neg**	**Neg**	**Neg**	**Neg**	**Neg**
	8 Toler percent	**Pos**	Pos	Pos	**Pos**	**Pos**
Feeding Group	Cllct percent	**Pos**	**Pos**	**Neg**	Mix	Mix
	Scrap percent	**Neg**	**Neg**	**Pos**	Mix	Mix
	Shred percent	**Neg**	Neg	**Neg**	**Neg**	**Neg**
	9 Shredder Taxa	**Neg**	**Neg**	**Neg**	**Neg**	**Neg**
Habit	Brrwr percent	**Pos**	**Pos**	**Neg**	Pos	Mix
	10 Clngr percent	**Neg**	**Neg**	Mix	Mix	Neg

a: The DO measures characterized in the qualitative correlations were Pmax4hr and GPP, which gave opposite responses compared to Rmax4hr and minimum DO.

As with diatom metrics, the most responsive macroinvertebrate metrics were selected for continuing analysis of stressor-response effects. These included 10 metrics that are commonly used in bioassessments, had consistent and strong correlations, and represent different attributes of the community. In the conceptual model, macroinvertebrates respond directly to minimum DO conditions, which are related to chl-a. Other measures of DO (Pmax4hr and DeltaDO) were also related to macroinvertebrate metrics, though these were not specified in the conceptual model. However, in these analyses, the strongest relationship was directly between nutrients and macroinvertebrate metrics (bypassing DO or chl-a). This might be due to a larger data set for nutrients relative to datasets for DO or chl-a. The intermediate stressors (chl-a and DO) showed trends that support the causal model. Different macroinvertebrate sampling methods were indistinct in biplots of stressors and metrics. Therefore, data from multiple sampling methods were pooled in stressor-response analyses.

5.2 Regression Interpolation

When a clear linear relationship is evident between a nutrient concentration and a response variable with an existing threshold, then a nutrient concentration can be associated with the response threshold through intersection with the linear regression. Since no response thresholds were established, the 25th or 75th quartile of the response metrics in reference sites was used to represent a protective threshold. The high-small multi-metric macroinvertebrate condition index (Jacobi et.al. 2006) was not used, though it had an associated threshold of impact, because it could only be applied in a limited number of sites.

Ten macroinvertebrate and eight diatom metrics were regressed against TN and TP. The regressions included all nutrient and biological samples that were taken from the same site within a 30 day window. Data from all site classes were used to derive the regression equations because this assured a complete nutrient gradient on the x-axis and showed more significant relationships. This resulted in a mean regression slope that is not as steep as the effective slope observed as a regression of the upper quantiles of the data. Shallow slopes of the regression equations result in large changes in interpolated nutrient values for each incremental change of the reference metric value.

The reference 25th (or 75th) quantile of metric values in each nutrient site class were interpolated to a nutrient value on the x-axis. These quartiles of reference observations were selected as the critical values to represent reference expectations. The nutrient values associated with reference metric quartiles were interpolated by substituting the critical metric value as y in the equation and solving for x. The results were only considered as candidate nutrient thresholds if the regression equation was significant ($p<0.05$) and the interpolated nutrient value was within the range of observed values. Any values extrapolated beyond the observed range in each site class were disregarded. The resulting valid candidate thresholds ranged from 0.13 to 3.26 mg/L for TN and from 0.003 to 1.74 mg/L for TP (Table 15). Median valid candidate threshold values were calculated. Regression interpolation of TN and TP from a critical, minimum DO of 5 or 6 mg/L did not yield valid results.

Table 15. Candidate thresholds derived from regression interpolations on selected macroinvertebrate and diatom metrics. Values in gray font were not valid because they did not have significant regression equations or were outside of the observed range of values in the site classes.

	TN (mg/L)			TP (mg/L)		
	TN Flat	TN Moderate	TN Steep	TP High-Volcanic	TP Flat-Moderate	TP Steep
EPTTax	3.70	0.18	0.43	0.11	3.39	0.11
EphemTax	2.27	0.62	0.62	1.00	211	1.00
PlecoTax	3.26	3.26	3.26	1.61	1.61	0.15
IntolTax	3.48	0.53	0.85	0.88	6.22	0.33
Toler percent	501	0.29	0.33	0.22	3.11	0.017
EPT percent	398	0.13	21.47	59102	31.91	491758
Pleco percent	0.49	0.49	0.49	1.74	1.74	0.80
NonIn percent	28.33	0.21	0.37	1.46	0.281	0.003
ShredTax	2.28	2.28	0.64	56.47	56.47	0.60
Clngr percent	108	0.88	2.44	13.22	5.01	0.50
BMI Medians	2.28	0.49	0.46	0.11	1.61	0.11
wa_OptCat_DisTotMMI	10.35	0.36	0.19	0.042	0.168	0.028
wa_OptCat_L1DisTot	18.26	0.30	0.19	0.024	0.358	0.027
wa_OptCat_L1Ptl	7.45	0.43	0.29	0.068	0.145	0.029
wa_OptCat_LNtl	10.49	0.33	0.18	0.057	0.311	0.054
wa_OptCat_NutMMI	9.26	0.32	0.23	0.047	0.193	0.025

	TN (mg/L)			TP (mg/L)		
	TN Flat	TN Moderate	TN Steep	TP High-Volcanic	TP Flat-Moderate	TP Steep
pi_NAWQA_TN_1	1.28	4.36	5.32	0.457	0.129	0.010
pi_Ptpv_TP_all_Hi	7.98	0.25	0.69	0.083	0.152	0.011
x_Shan_e	2.26	16.63	161700	9.272	7.272	0.012
Diatom Medians	1.28	0.33	0.21	0.052	0.168	0.026
Median of all valid interpolated values	2.27	0.33	0.35	0.063	0.237	0.025
Reference 90th quantile	0.69	0.42	0.30	0.105	0.071	0.054
Maximum in site class	3.44	2.63	0.75	0.22	1.82	0.12

For the regression interpolation of DO statistics on nutrient concentrations, both nutrient concentrations and DO stats were log transformed. The regression equations were calculated with all sites classes combined. Regression equations in the three individual TP site classes resulted in non-significant regressions in the TP High-Volcanic and TP Steep site classes. However, in the TP Steep class, the relationships between TP and both Delta DO and Pmax4hr were negative and significant ($p<0.05$). The negative relationships were only seen in the TP Steep site class. Equations for the TP Flat-Moderate class were similar to those in all sites, so we emphasized results from equations for all site classes combined (Table 16). Regression interpolation in the TP High-Volcanic and TP Steep site classes were at the extreme high and low (respectively) ends of the range of observed values.

Table 16. Candidate thresholds for DO statistics derived from regression interpolations on reference 90th quantile nutrient concentrations. Values in gray font were not valid because they were at the extremes of the range of values.

	Delta DO			Pmax4hr		
	TP High-Volcanic	TP Flat-Moderate	TP Steep	TP High-Volcanic	TP Flat-Moderate	TP Steep
TN	16.39	3.34	1.13	3.23	0.47	0.13
TP	12.63	4.06	0.92	2.36	0.56	0.08

5.3 Change-point Analysis

The change-point is the point along an environmental gradient (nutrient concentrations) at which there is a high degree of change in the response variable (macroinvertebrate, diatom, or DO metrics). The nonparametric deviance reduction method for identifying change-points (Qian et al. 2003, King and Richardson 2003) works well when the response is stepped, or drastically changing at a recognizable point along nutrient concentration gradient. With this method, the data are divided into two groups, above and below a potential nutrient threshold, where each group is internally similar and the difference among groups is high. One caveat of the change-point analysis is that a change-point may be identified, and even determined to be statistically significant, when the change-point value is actually only an artifact of the analysis and not an indication of a change in system properties (Qian and Cuffney 2012, Daily et al. 2012). The methods can find change-points, even in datasets with nearly straight line relationships between X and Y. It has been well established that nutrient concentrations limit algal growth as well as species composition. Therefore, it is reasonable to believe an ecological threshold does exist between certain periphyton metrics and nutrient concentrations. The changepoint results were qualified using three assessment measures: valuation of the 95th quantile regression line, the relative size of the confidence interval around the change-point, and coincidence of an appropriate slope in the LOWESS regression line at the change-point. Confidence intervals were calculated for each change-point to illustrate the possible ranges of change-points.

Change-points were identified for both TN and TP from 10 macroinvertebrate metrics, 8 diatom metrics, and 2 DO measures (Table 17). Change-points for chl-a were not identified because benthic chl-a was not significantly correlated to nutrient concentrations and did not produce valid change points. As discussed previously, a number of biotic and abiotic variables influence algal biomass accrual (Marks et al 2000).

The ranges of valid change-points were fairly narrow for each nutrient and site class (at most 1.24 mg/L for TN and 0.08 mg/L for TP). For TN, median candidate thresholds were greatest in the TN Flat site class and least in the TN Steep site class. Likewise for TP, TP Steep sites had lower median change-points and increasing change-point medians were in the TP Flat-Moderate and TP high-Volcanic classes. The EPT taxa, Plecoptera taxa, weighted average disturbance (wa_OptCat_L1DisTot), and weighted average nitrogen preference (wa_OptCat_LNtl) metrics had the most valid change-points associated with them.

Table 17. Change-points (CP) as candidate thresholds from selected benthic macroinvertebrate (BMI), diatom and dissolved oxygen (DO) metrics. Values in gray font did not pass the tests for valid change-points.

Metric	TN (mg/L) TN Flat	TN (mg/L) TN Moderate	TN (mg/L) TN Steep	TP (mg/L) TP High-Volcanic	TP (mg/L) TP Flat-Moderate	TP (mg/L) TP Steep
EPTTax	0.49	0.25	0.42	0.067	0.044	0.030
EphemTax	0.49	0.22	0.28	0.058	0.044	0.030
PlecoTax	0.56	0.33	0.25	0.063	0.041	0.027
IntolTax	0.48	0.29	0.39	0.061	0.051	0.029
Toler percent	0.66	0.40	0.26	0.083	0.052	0.041
EPT percent	0.97	0.36	0.22	0.047	0.014	0.029
Pleco percent	0.35	0.33	0.14	0.114	0.044	0.027
NonIn percent	0.72	1.26	0.23	0.083	0.014	0.018
ShredTax	0.48	0.25	0.23	0.047	0.151	0.017
Clngr percent	1.09	0.49	0.28	0.122	0.051	0.022
Median CP BMI	0.53	0.31	0.28	0.063	0.044	0.029
wa_OptCat_DisTotMMI	0.48	0.52	0.16	0.068	0.056	0.035
wa_OptCat_L1DisTot	0.50	0.38	0.26	0.068	0.066	0.034
wa_OptCat_L1Ptl	0.48	0.52	0.13	0.066	0.032	0.036
wa_OptCat_LNtl	0.47	0.39	0.19	0.068	0.078	0.035
wa_OptCat_NutMMI	0.47	0.52	0.15	0.066	0.056	0.035

Metric	TN (mg/L) TN Flat	TN Moderate	TN Steep	TP (mg/L) TP High-Volcanic	TP Flat-Moderate	TP Steep
pi_NAWQA_TN_1	0.66	0.67	0.13	0.084	0.028	0.019
pi_Ptpv_TP_all_Hi	0.52	0.71	0.21	0.094	0.032	0.029
x_Shan_e	0.70	0.51	0.25	0.071	0.034	0.027
Median CP diatoms	0.49	0.45	0.18	0.068	0.056	0.035
DO_min	0.63	0.34	0.30	0.066	0.039	0.035
Pmax4hr	0.70	0.37	0.36	0.059	0.099	0.035
Median valid CP BMI, diatoms, & DO	0.50	0.36	0.22	0.067	0.044	0.035
Reference 90th quantile	0.69	0.42	0.30	0.105	0.071	0.054

Change-points for Delta DO and Pmax4hr were calculated based on macroinvertebrate metrics and nutrient concentrations (Table 18). Using nutrient concentrations in the CPA as a response to DO statistics is somewhat circular and might not be an appropriate application of the technique. However, median values for the DO change-points derived from macroinvertebrate metrics were equal to the medians when the nutrient-derived change-points were also included. Change-points were derived for all sites and for the TP Flat-Moderate sites. In the TP High-Volcanic and TP Steep site classes, there were not enough samples for valid change-point analyses.

Table 18. Change-points (CP) as candidate DO thresholds from selected benthic macroinvertebrate (BMI) metrics and nutrient concentrations. Values in gray font did not pass the tests for valid change-points.

Metric	Delta DO All sites	Delta DO TP Flat-Moderate	Pmax4hr All sites	Pmax4hr TP Flat-Moderate
EPTTax	1.74	1.99	0.358	0.254
EphemTax	1.60	1.88	0.298	0.254
PlecoTax	2.34	2.09	0.275	0.254
IntolTax	2.37	2.42	0.254	0.254
Toler percent	2.46	2.44	0.474	0.439
EPT percent	1.72	2.44	0.290	0.338
Pleco percent	2.02	1.99	0.214	0.214
NonIn percent	2.41	2.42	0.298	0.322
ShredTax	1.56	2.44	0.145	0.214
Clngr percent	2.26	2.44	0.331	0.254
LogTN	5.77	5.73	0.679	0.679
LogTP	2.03	2.06	0.269	0.351
Median	2.30	2.42	0.290	0.254

6. Synthesis of Multiple Thresholds

The strength of an analysis with numerous approaches and response endpoints comes from the multiple lines of evidence. They can be used to show central tendencies and ranges in candidate thresholds. The central tendency of candidate thresholds shows corroborated evidence, which give greater confidence in a selected threshold. Threshold selection may be based on confidence in an individual analytical technique, corroboration from multiple lines of evidence, and/or on corroborating evidence from the scientific literature.

Both reference percentiles and the stressor-response results were considered in selection of proposed thresholds. All candidate nutrient thresholds were compiled and summarized for each variable and site class using cumulative distribution function (CDF). CDF curves show the proportion (referred to as the "cumulative proportion") of candidate thresholds that are less than a selected concentration (in this case the 90th quantile of reference sites). All of the valid candidate thresholds were shown in tables and CDF curves in Appendix A. Synthesis of the multiple thresholds also included review of the individual analytical techniques and corroboration from multiple lines of evidence and the scientific literature. Evidence from the stressor-response analyses and the scientific literature supports the selected quantile values as thresholds.

The CDF curves place the reference 90th quantile values within the ranges of stressor-response thresholds. If the proportion of the candidate thresholds below the 90th quantile was greater than 65%, a quantile between 90 and 74 was selected to move the proposed threshold closer to the central tendency of all candidate thresholds. TP in the High-Volcanic site class is the exception. In this site class a larger proportion (86%) of the candidate thresholds are below the 75th quantile as biological responses occur within the range of reference TP concentrations. Since this is a biological response to natural conditions, the 75th quantile was selected as the proposed threshold. This synthesis resulted in proposed thresholds shown in Tables 19.

Table 19. Proposed TN and TP thresholds by nutrient site classes with related statistical foundation

	TN (mg/L)			TP (mg/L)		
	TN Flat	TN Moderate	TN Steep	TP High-Volcanic	TP Flat-Moderate	TP Steep
Reference quantile[1]	85	80	90	75	90	90
Proposed threshold	**0.65**	**0.37**	**0.30**	**0.084**	**0.061**	**0.03**
Reference quantile 90% confidence interval[1]	0.59 – 0.84	0.35 – 0.41	0.26 – 0.34	0.072 – 0.09	0.051 – 0.069	0.016 – 0.053
Stressor-response candidate thresholds median[2]	0.52	0.33	0.26	0.067	0.066	0.029
Cumulative proportion[2]	57%	61%	63%	86%	48%	53%

[1] from Table 8
[2] proportion of candidate thresholds below the selected reference quantile

These thresholds are supported by the scientific literature. The values are in the range of thresholds found in peer reviewed literature and USEPA approved numeric nutrient criteria. Numeric nutrient standards for wadeable streams in different Montana ecoregions range from 0.275-1.3 for TN and 0.025-0.15 for TP (MTDEQ 2014). A Review of Stream Nutrient Criteria Development in the United States conducted by M. A. Evans-White, B. E. Haggard, and J. T. Scott and published in the Journal of Environmental Quality (Evans-White, et.al. 2013) found the following:

- percentile analysis of ecoregions found in NM produced TN thresholds ranging from 0.3 - 0.9 and TP thresholds of 0.01 - 0.1 (5 studies)
- benthic macroinvertebrate derived thresholds ranged from 0.6 - 1.7 for TN and 0.04 - 0.15 for TP (using 13 difference metrics in 3 studies)
- benthic algal derived thresholds ranged from 0.4 - 1.1 for TN and 0.01 - 0.07 for TP (using 19 difference metrics in 4 studies)
- of states with numeric nutrient standards (excluding Nevada's very high values) TN criteria were 0.2 - 2.0 (5 states) and TP criteria were 0.01 - 0.10 (9 states)

For Delta DO values, the reference distribution 90th quantile values were similar in the TP High-Volcanic and TP Flat-Moderate site classes. Due to small sample sizes in the other site classes, stressor-response analyses were only possible in the TP Flat-Moderate site class. The regression interpolation using nutrient thresholds were close and slightly lower than the 90th quantile value. Change-point analysis suggested lower thresholds for all macroinvertebrate metrics. Change-point values for all sites were similar to those derived in the TP Flat-Moderate site class and were generally in the range described by the reference distribution in the TP Steep sites. The 90th quantile of the Delta DO reference distribution is the proposed Diel DO thresholds. These thresholds will be applied by TP site class as it was used in calculating the threshold and Delta DO was significantly correlated with TP.

Table 20. Threshold ranges for Delta DO derived from reference distributions (Ref Dist 90th), the reference distribution 90% confidence interval (Ref Dist CI90), regression interpolation range (Reg Int range), change-point analysis (CPA) median, and CPA ranges associated with benthic macroinvertebrates (BMI) and nutrients.

	TP High-Volcanic	TP Flat-Moderate	TP Steep	All Classes
Ref Dist 90th	5.02	4.08	1.79	4.16
Ref Dist CI90	3.13 - 7.24	3.52 - 7.27	1.40 - 2.37	3.27-7.13
Reg Int range	NA	3.34 - 4.06	NA	NA
CPA median	NA	2.42	NA	2.30
CPA BMI range	NA	1.88 – 2.44	NA	1.56 – 2.46

References

Antweiler, R.C. and H.E. Taylor. 2008. Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets: I. Summary statistics. Environ. Sci. Technol. 42:3732–3738.

Barbour, M.T., J. Gerritsen, B.D. Snyder and J.B. Stribling. 1999. Chapter 6 in the EPA Rapid Bioassessment Protocol for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish, Second Edition, EPA 841-B-99-002.

Biggs, B.J.F., 2000, Eutrophication of streams and rivers: dissolved nutrient-chlorophyll relationships for benthic algae: Journal of American Benthological Society, v. 19, no. 1, p. 50-67.

Daily, J. P., N.P. Hitt, D.R. Smith, and C.D. Snyder. 2012. Experimental and environmental factors affect spurious detection of ecological thresholds. Ecology, 93(1):17-23.

Evans-White, M. A., B. E. Haggard, and J. T. Scott. 2015. A Review of Stream Nutrient Criteria Development in the United States. Journal of Environmental Quality. 07/2013; 42(4):1002-14.

Griffith, G.E., J.M. Omernik, M.M. McGraw, G.Z. Jacobi, C.M. Canavan, T.S. Schrader, D. Mercer, R. Hill, and B.C. Moran. 2006. Ecoregions of New Mexico: Reston, Virginia, U.S. Geological Survey (map scale 1:1,400,000).

Helsel, D. 2010. Much ado about next to nothing: Incorporating nondetects in science. Ann. Occup. Hyg., 54(3): 257–262.

Jacobi, G.Z., M.D. Jacobi, M.T. Barbour, and E.W. Leppo. 2006. Benthic macroinvertebrate stream condition indices for New Mexico wadeable streams. Prepared for the New Mexico Environment Department, Santa Fe.

Jessup, B.K., D. Eib, L. Guevara, J. Hogan, F. John, S. Joseph, P. Kaufmann, and A. Kosfiszer. 2015. New Mexico Nutrient Thresholds for Perennial Wadeable Streams. Prepared for the U.S. Enviromental Protection Agency, Region 6, Dallas, TX and the New Mexico Environment Department, Santa Fe, NM. Prepared by Tetra Tech, Inc., Montpelier, VT.

Jessup, B.K., P. Kaufmann, F. John, L.S. Guevara, S. Joseph. 2014. Bedded sediment conditions and macroinvertebrate responses in New Mexico streams: a first step in establishing sediment criteria. Journal of the American Water Resources Association. 50(6):1558-1574.

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(6):795-809.

Marks, J. C., Power, M. E. and Parker, M. S. 2000. Flood disturbance, algal productivity, and interannual variation in food chain length OIKOS 90: 20–27. Copenhagen

Montana Department of Environmental Quality. 2014. DEPARTMENT CIRCULAR DEQ-12A Montana Base Numeric Nutrient Standards. http://deq.mt.gov/Portals/112/Water/WQPB/Standards/NutrientWorkGroup/PDFs/NutrientRules/CircularDEQ12A_July2014_FINAL.pdf

NMAC (New Mexico Administrative Code). 2013. State of New Mexico Standards for Interstate and Intrastate Streams. 20.6.4. New Mexico Water Quality Control Commission. As amended through June 5, 2013.

New Mexico Environment Department Surface Water Quality Bureau (NMED/SWQB). 2008. State of New Mexico Nutrient Criteria Development Plan, Revision 4. Prepared by Surface Water Quality Bureau, New Mexico Environment Department. Accessed 01/04/2013 at: http://www.nmenv.state.nm.us/swqb/Nutrients/index.html

New Mexico Environment Department Surface Water Quality Bureau (NMED/SWQB). 2015. State of New Mexico Surface Water Quality Bureau Standard Operating Procedures for Sample Collection and Handling. Available at: http://www.nmenv.state.nm.us/swqb/.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchak, F.H. McCormick, S.A. Peterson, P.L. Ringold, T. Magee, and M. Cappaert, 2006. Environmental Monitoring and Assessment Program-Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA/620/R-06/003. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C.

Peterson, C.G., A.C. Weibel, N.B. Grimm, and S.G. Fisher. 1994. Mechanisms of benthic algal recovery following spates: comparison of simulated and natural events. Oecologia August 1994, Volume 98, Issue 3-4, pp 280-290

Porter, S. D. 2008. Algal attributes: an autecological classification of algal taxa collected by the National Water-Quality Assessment Program. US Geological Survey.

Potapova, M., and D.F. Charles, 2007, Diatom metrics for monitoring eutrophication in rivers of the United States: Ecological Indicators, v. 7, p. 48–70.

Qian, S.S., and T.F. Cuffney. 2012. To threshold or not to threshold? That's the question. Ecological Indicators 15:1-9.

Qian, S. S., R. S. King, and C.J. Richardson. 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modelling, 166(1):87-97.

Stevenson, R.J., S.T. Rier, C.M. Riseng, R.E. Schultz, and M.J. Wiley. 2006. Comparing effects of nutrients on algal biomass in streams in two regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia. 561:149-165.

Stevenson, R. J, Y. Pan, K. M. Manoylov, C. A. Parker, D. P. Larsen, and A.T. Herlihy. 2008. Development of diatom indicators of ecological conditions for streams of the western US. J. N. Am. Benthol. Soc., 27(4):1000–1016.

Stoddard, J.L., D.V. Peck, A.R. Olsen, D.P. Larsen, J. Van Sickle, C.P. Hawkins, R.M. Hughes, T.R. Whittier, G. Lomnicky, A.T. Herlihy, P.R. Kaufmann, S.A. Peterson, P.L. Ringold, S.G. Paulsen, and R. Blair. 2005. Western Streams and Rivers Statistical Summary. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC. EPA 620/R-05/006.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., and Norris, R.H. 2006. Setting Expectations for the Ecological Condition of Streams: The Concept of Reference Condition. Ecological Applications, 16(4), 2006, pp. 1267–1276.

Tetra Tech. 2014. Quality Assurance Project Plan for Nutrient-Scientific Technical Exchange Partnership System (N-STEPS) – Secondary Data Analysis and Model Development in Support of Numeric Nutrient Criteria. Revision 1. Prepared for U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Health and Ecological Criteria Division by Tetra Tech, Inc., Fairfax, VA.

________. 2012. Quality Assurance Project Plan for New Mexico Nutrient Framework Technical Support; QAPP 343. Prepared by Tetra Tech, Inc., Fairfax, VA, September 2012.

________. 2011a. Quality Assurance Project Plan for Nutrient-Scientific Technical Exchange Partnership System (N-STEPS) – Secondary Data Analysis. Revision 0. Prepared for U.S. Environmental Protection Agency, Office of Water, Office of Science and Technology, Health and Ecological Criteria Division by Tetra Tech, Inc., Fairfax, VA.

Therneau, T., B. Atkinson and B. Ripley. 2013. rpart: Recursive Partitioning. R package version 4.1-3. http://CRAN.R-project.org/package=rpart

USEPA (U.S. Environmental Protection Agency. 2010. Using Stressor-response Relationships to Derive Numeric Nutrient Criteria. EPA-820-S-10-001.

________. 2009. Empirical Approaches for Nutrient Criteria Derivation. Science Advisory Board Review Draft.

________. 2006. Estimation and Application of Macroinvertebrate Tolerance Values. Report No. EPA/600/P04/116F. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, D.C.

________. 2004. Wadeable Stream Assessment: Field Operations Manual. EPA841-B-04-004. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, DC.

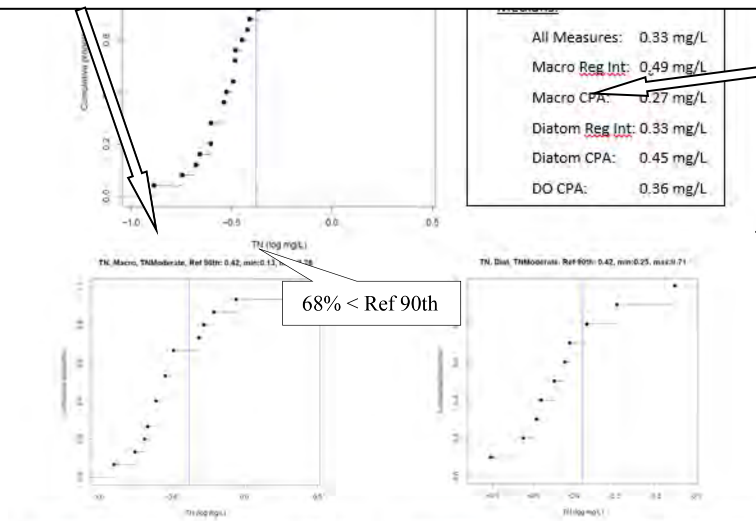
________. 2000a. Nutrient Criteria Technical Guidance Manual: Rivers and Streams. EPA-822-B-00-002. U.S. Environmental Protection Agency, Washington, DC. http://www.epa.gov/ost/criteria/nutrient/guidance/rivers/index.html

Woodruff, L., W.F. Cannon, D.B. Smith, and F. Solano, 2015. The distribution of selected elements and minerals in soil of the conterminous United States, J. Geochem. Explor., http://dx.doi.org/10.1016/j.gexplo.2015.01.006)

Appendix A Candidate Threshold Summary

Candidate Threshold Summary Legend

CDF diagram with all thresholds including 90th quantile of the reference distributions (vertical line with dashed vertical confidence interval) and change-point analysis and regression interpolation from both assemblages. X-axis is log mg/L. Y-axis is proportion of candidate thresholds less than X. The percentage of valid biological thresholds less than the reference 90th quantile is emphasized.



Title, with nutrient and site class

Summary statistics, with candidate threshold, confidence interval and medians of candidate thresholds by assemblage and method.

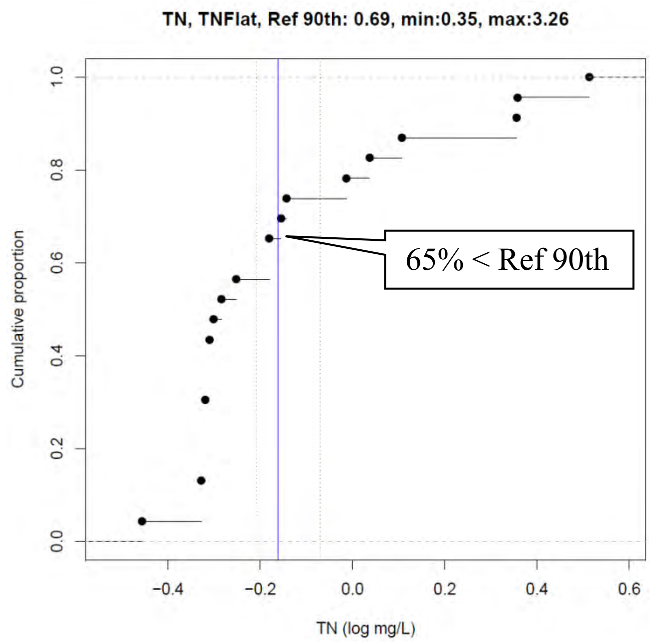
CDF diagrams for diatoms and macroinvertebrates, including both methods. Nutrients are shown on a log scale on the axes and are back-transformed in the text.

Macro Metric	Reg. Int.	CPA
EPTTax	0.18	0.25
EphemTax	0.62	0.22
PlecoTax	3.26	0.33
IntolTax	0.53	0.29
Toler percent	0.29	0.40
EPT percent	0.13	0.36
Pleco percent	0.49	0.33
NonIn percent	0.21	1.26
ShredTax	2.28	0.25
Clngr percent	0.88	0.49

Diatom Metric	Reg. Int.	CPA
wa_OptCat_DisTotMMI	0.36	0.52
wa_OptCat_L1DisTot	0.30	0.38
wa_OptCat_L1Ptl	0.43	0.52
wa_OptCat_LNtl	0.33	0.39
wa_OptCat_NutMMI	0.32	0.52
pi_NAWQA_TN_1	[illegible]	0.67
pi_Ptpv_TP_all_Hi	0.25	0.71
x_Shan_e	[illegible]	0.51

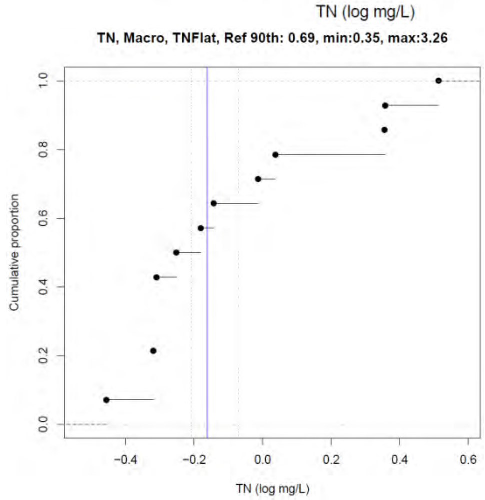
Threshold values from macroinvertebrate and diatom assemblages and from regression interpolation (Reg. Int.) and change-point analysis (CPA). Gray-shaded values are not valid and were not used in the CDF or summary statistics.

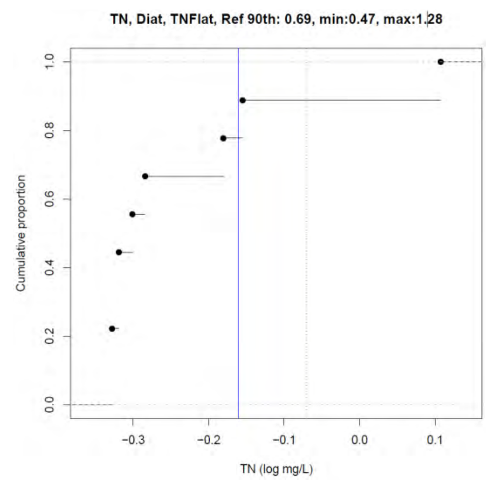
7.1 TN Threshold Synthesis – Flat Site Class



90th Quantile of Ref: **0.69 mg/L**
90% CI: 0.62 – 0.85 mg/L
Medians:

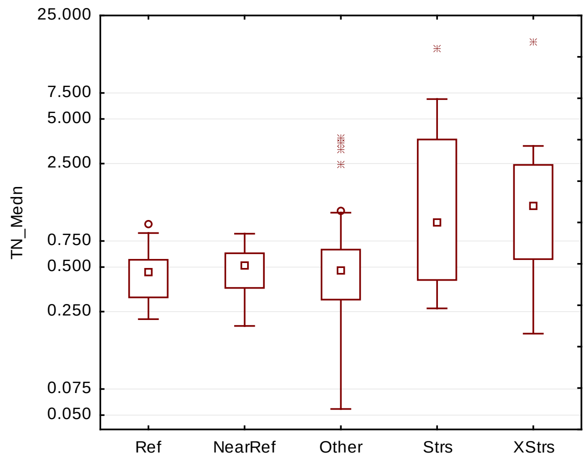
All Measures:	0.52 mg/L
Macro Reg Int:	2.28 mg/L
Macro CPA:	0.53 mg/L
Diatom Reg Int:	1.28 mg/L
Diatom CPA:	0.49 mg/L
DO CPA:	NA





Macro Metric	Reg.Int.	CPA
EPTTax	3.70	0.49
EphemTax	2.27	0.49
PlecoTax	3.26	0.56
IntolTax	3.48	0.48
Toler percent	500.8	0.66
EPT percent	398.7	0.97
Pleco percent	0.49	0.35
NonIn percent	28.3	0.72
ShredTax	2.28	0.48
Clngr percent	108.6	1.09

Diatom Metric	Reg.Int.	CPA
wa_OptCat_DisTotMMI	10.4	0.48
wa_OptCat_L1DisTot	18.3	0.50
wa_OptCat_L1Ptl	7.45	0.48
wa_OptCat_LNtl	10.5	0.47
wa_OptCat_NutMMI	9.26	0.47
pi_NAWQA_TN_1	1.28	0.66
pi_Ptpv_TP_all_Hi	7.98	0.52
x_Shan_e	2.26	0.70



TN values were lowest in reference sites in the TN Flat site class and increased most in the stressed and extremely stressed sites (Figure 45). More than half of the stressed and extremely stressed sites were greater than the 90th quantile value (0.69 mg/L TN). Also see Section 4.2.

Figure 4. **Site median TN value distributions along the disturbance gradient for sites in the TN Flat site class.**

Several change-points were identified near 0.50 mg/L for both diatoms and macroinvertebrate metrics (e.g., Figure 46). Also see Section 4.5 and Appendix L.

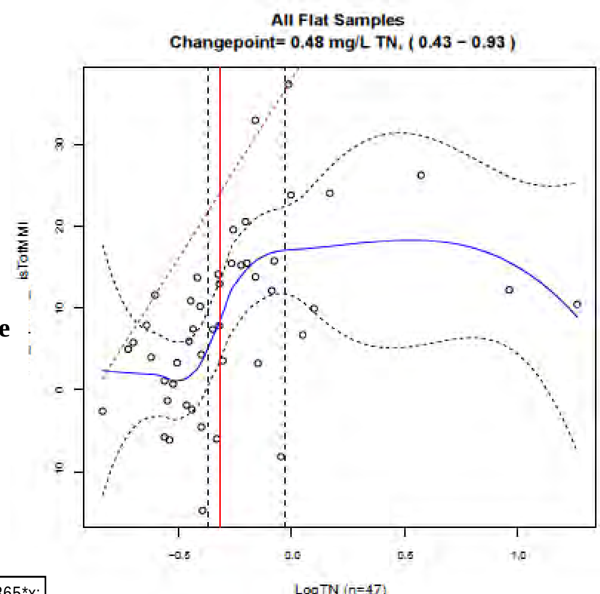
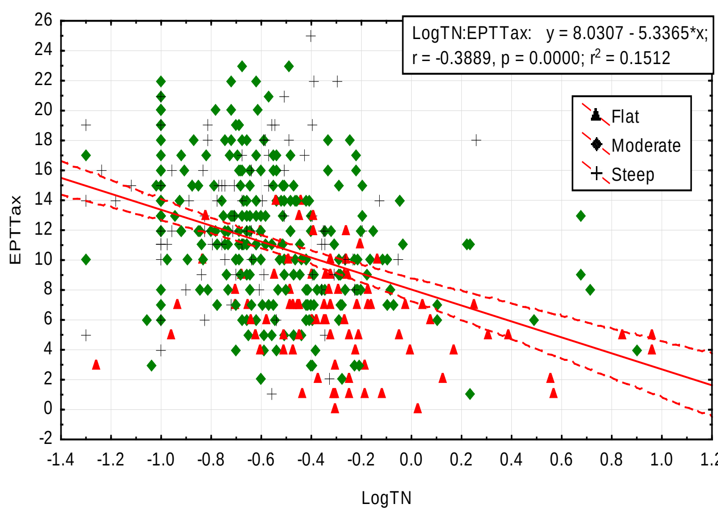


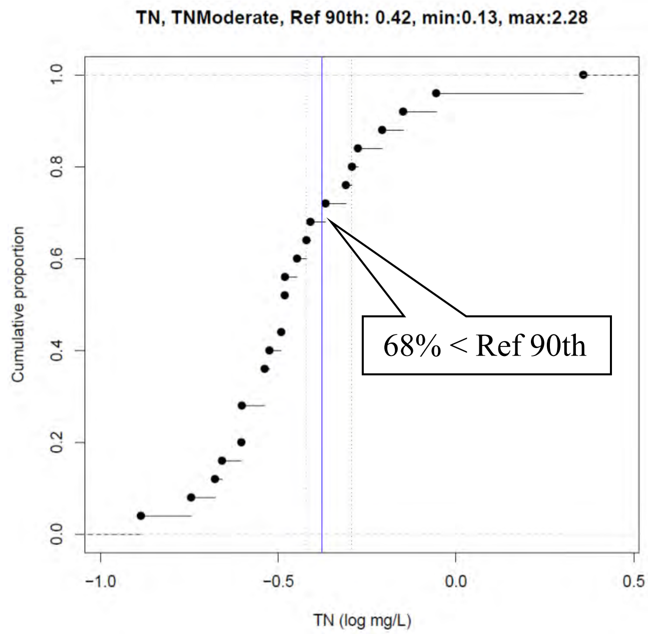
Figure 5. **Change-point plot for TN and the weighted average disturbance index diatom metric, showing the change-point with confidence intervals (vertical solid and dashed lines), the 95th quantile regression line (slanting dashed line), and the LOWESS regression (blue fitted line).**



While all but one metric had a significant overall regression equation (e.g., Figure 47), the reference quartiles of the metrics in the TN Flat sites were substantially different than those in the other site classes. This resulted in high interpolated values for TN. Also see Section 4.4 and Appendix K.

Figure 6. **Regression plot for TN and EPT taxa. In the TN Flat site class, the reference quartile for EPT taxa was 5 taxa, which translates to 3.7 mg/L TN.**

7.2 TN Threshold Synthesis – Moderate Site Class



90th Quantile of Ref: **0.42 mg/L**
90% CI: 0.38 – 0.51 mg/L
Medians:

All Measures: 0.33 mg/L

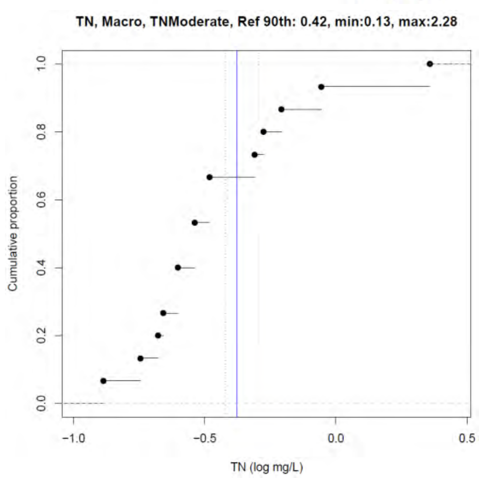
Macro Reg Int: 0.49 mg/L

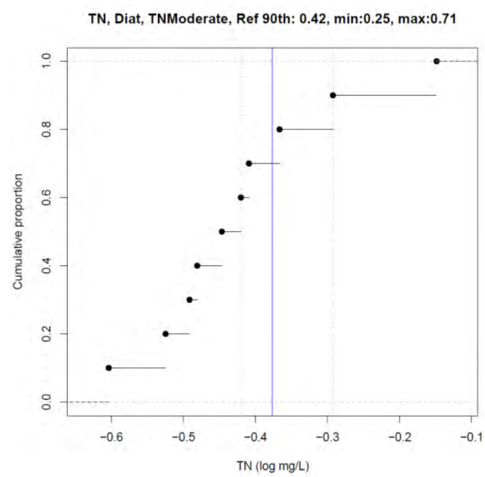
Macro CPA: 0.27 mg/L

Diatom Reg Int: 0.33 mg/L

Diatom CPA: 0.45 mg/L

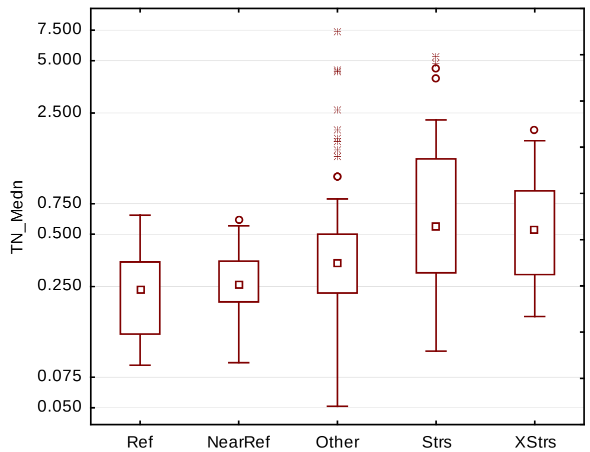
DO CPA: 0.36 mg/L





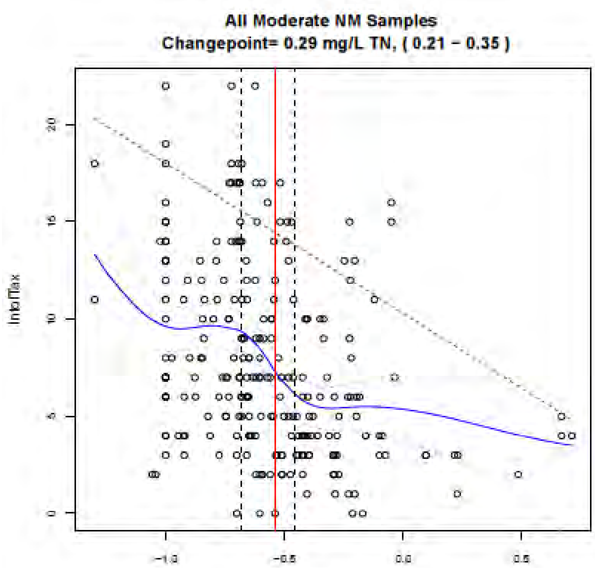
Macro Metric	Reg. Int.	CPA
EPTTax	0.18	0.25
EphemTax	0.62	0.22
PlecoTax	3.26	0.33
IntolTax	0.53	0.29
Toler percent	0.29	0.40
EPT percent	0.13	0.36
Pleco percent	0.49	0.33
NonIn percent	0.21	1.26
ShredTax	2.28	0.25
Clngr percent	0.88	0.49

Diatom Metric	Reg. Int.	CPA
wa_OptCat_DisTotMMI	0.36	0.52
wa_OptCat_L1DisTot	0.30	0.38
wa_OptCat_L1Ptl	0.43	0.52
wa_OptCat_LNtl	0.33	0.39
wa_OptCat_NutMMI	0.32	0.52
pi_NAWQA_TN_1	4.36	0.67
pi_Ptpv_TP_all_Hi	0.25	0.71
x_Shan_e	16.6	0.51



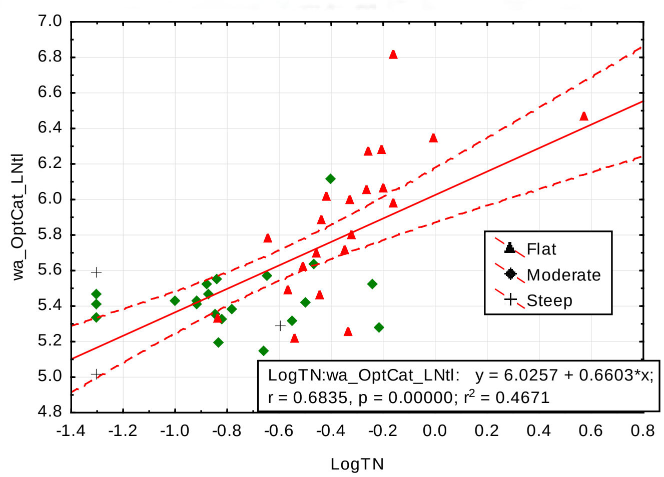
TN values were lowest in reference sites in the TN Moderate site class and increased most in the stressed and extremely stressed sites (Figure 48). More than half of the stressed and extremely stressed sites were greater than the 90th quantile value (0.42 mg/L TN). Also see Section 4.2.

Figure 7. **Site median TN value distributions along the disturbance gradient for sites in the TN Moderate site class.**



Most change-points were identified at TN values slightly less than the 90th quantile of reference sites (e.g., Figure 49). Also see Section 4.5 and Appendix L.

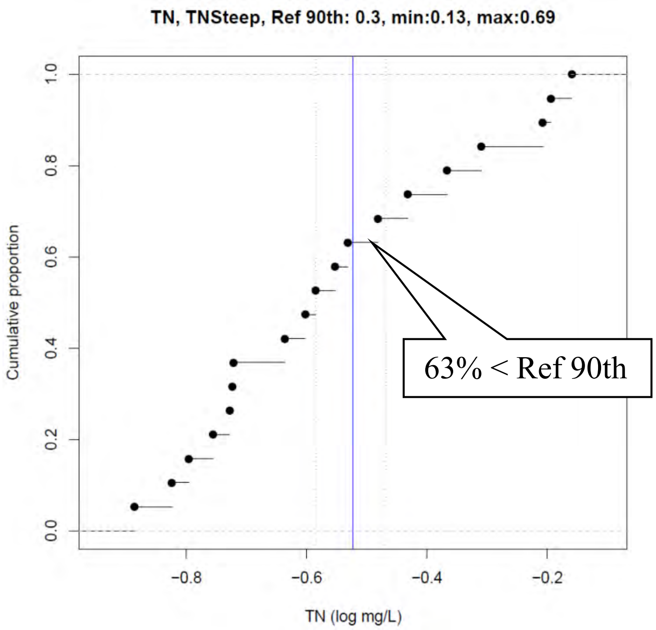
Figure 8. **Change-point plot for TN and the intolerant taxa macroinvertebrate metric, showing the change-point with confidence intervals (vertical solid and dashed lines), the 95th quantile regression line (slanting dashed line), and the LOWESS regression (blue fitted line).**



All but one metric had a significant overall regression equation (e.g., Figure 50). The reference quartiles of metrics in the TN Moderate sites were similar to those in the Steep site class. Also see Section 4.4 and Appendix K.

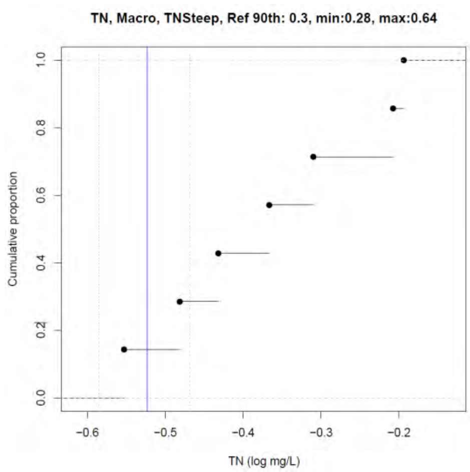
Figure 9. **Regression plot for TN and weighted average diatom nitrogen sensitivity. In the TN Moderate site class, the reference quartile for the metric was 6.0, which translates to 0.33 mg/L TN.**

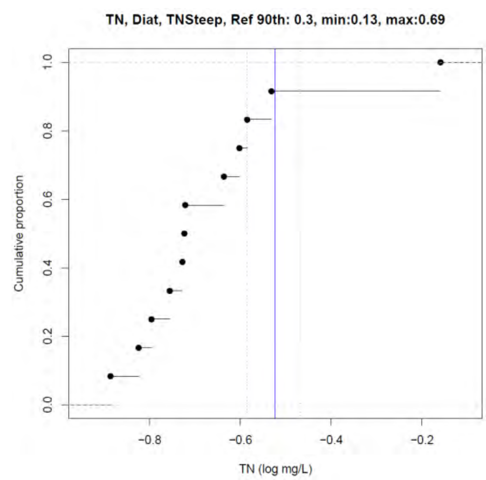
7.3 TN Threshold Synthesis – Steep Site Class



90th Quantile of Ref: **0.30 mg/L**
90% CI: 0.26 – 0.34 mg/L
Medians:

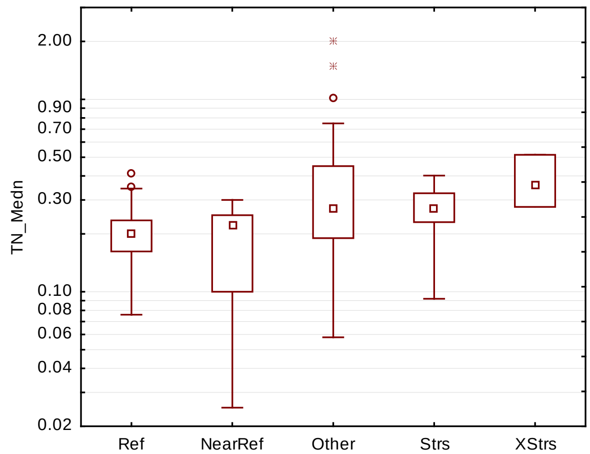
All Measures: 0.26 mg/L
Macro Reg Int: 0.46 mg/L
Macro CPA: 0.21 mg/L
Diatom Reg Int: 0.23 mg/L
Diatom CPA: 0.18 mg/L
DO CPA: 0.30 mg/L





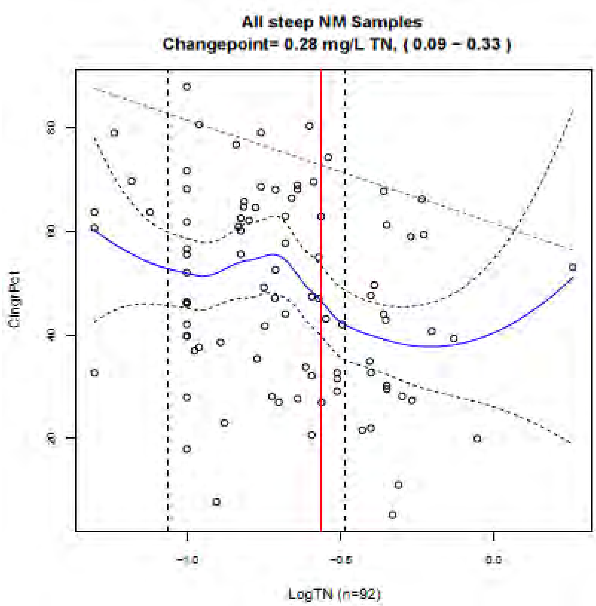
Macro Metric	Reg. Int.	CPA
EPTTax	0.43	0.42
EphemTax	0.62	0.28
PlecoTax	3.26	0.25
IntolTax	0.85	0.39
Toler percent	0.33	0.26
EPT percent	21.5	0.22
Pleco percent	0.49	0.14
NonIn percent	0.37	0.23
ShredTax	0.64	0.23
Clngr percent	2.44	0.28

Diatom Metric	Reg. Int.	CPA
wa_OptCat_DisTotMMI	0.19	0.16
wa_OptCat_L1DisTot	0.19	0.26
wa_OptCat_L1Ptl	0.29	0.13
wa_OptCat_LNtl	0.18	0.19
wa_OptCat_NutMMI	0.23	0.15
pi_NAWQA_TN_1	5.32	0.13
pi_Ptpv_TP_all_Hi	0.69	0.21
x_Shan_e	>10,000	0.25



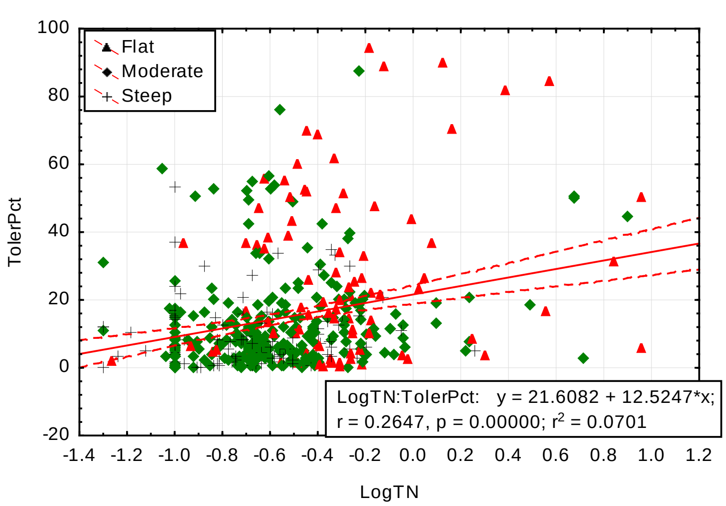
TN values were lowest in reference sites in the TN Steep site class and increased gradually with increasing stress (Figure 51). Approximately half of the stressed and extremely stressed sites were greater than the 90th quantile value (0.30 mg/L TN). Also see Section 4.2.

Figure 10. **Site median TN value distributions along the disturbance gradient for sites in the TN Steep site class.**



Most change-points were identified at TN values slightly less than the 90th quantile of reference sites. Only one macroinvertebrate metric gave acceptable change-point results (Figure 52). Also see Section 4.5 and Appendix L.

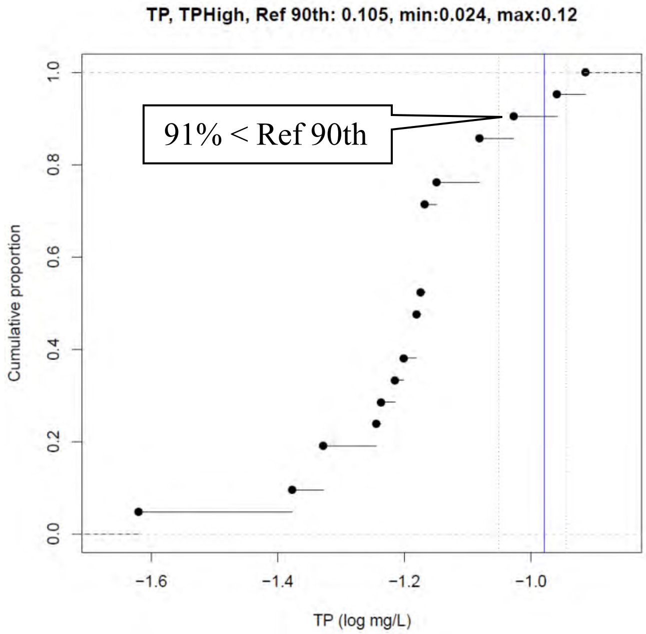
Figure 11. **Change-point plot for TN and the percent clinger macroinvertebrate metric, showing the change-point with confidence intervals (vertical solid and dashed lines), the 95th quantile regression line (slanting dashed line), and the LOWESS regression (blue fitted line).**



All but one metric had a significant overall regression equation (e.g., Figure 53). Interpolation results closest to the reference quantile value were for the macroinvertebrate percent tolerance metric. Also see Section 4.4 and Appendix K.

Figure 12. **Regression plot for TN and macroinvertebrate percent tolerance. In the TN Steep site class, the reference quartile for the metric was 15.5%, which translates to 0.33 mg/L TN.**

7.4 TP Threshold Synthesis – High-Volcanic Site Class



90th Quantile of Ref: **0.105 mg/L**
90% CI: 0.89 – 0.114 mg/L
Medians:

All Measures: 0.067 mg/L

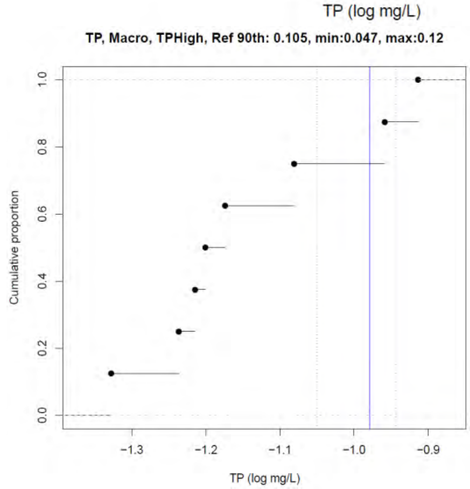
Macro Reg Int: 0.110 mg/L

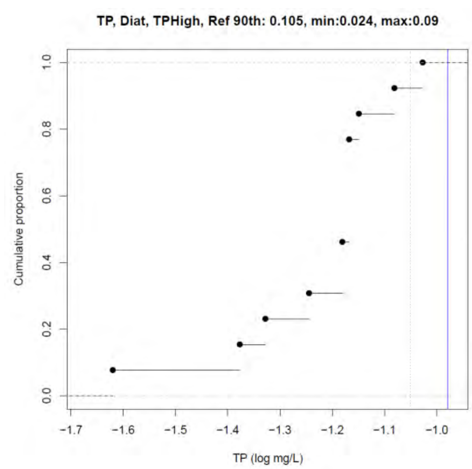
Macro CPA: 0.063 mg/L

Diatom Reg Int: 0.052 mg/L

Diatom CPA: 0.068 mg/L

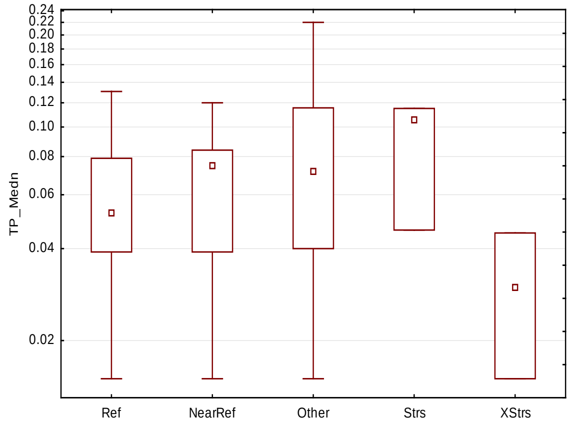
DO CPA: 0.059 mg/L





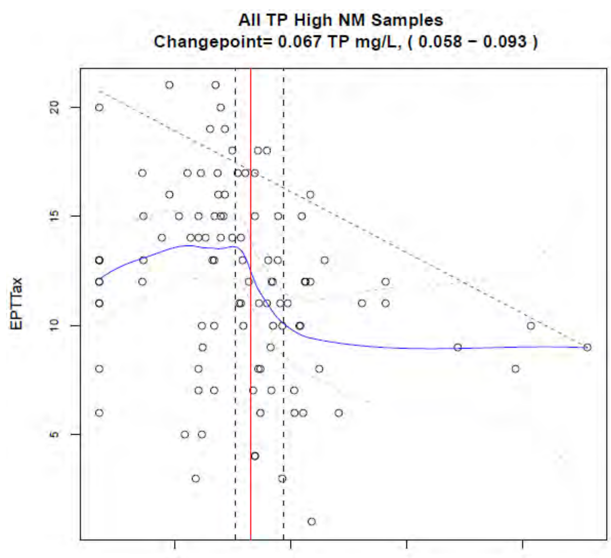
Macro Metric	Reg. Int.	CPA
EPTTax	0.11	0.067
EphemTax	1.00	0.058
PlecoTax	1.61	0.063
IntolTax	0.88	0.061
Toler percent	0.22	0.083
EPT percent	59,102	0.047
Pleco percent	1.74	0.114
NonIn percent	1.46	0.083
ShredTax	56.47	0.047
Clngr percent	13.22	0.122

Diatom Metric	Reg. Int.	CPA
wa_OptCat_DisTotMMI	0.042	0.068
wa_OptCat_L1DisTot	0.024	0.068
wa_OptCat_L1Ptl	0.068	0.066
wa_OptCat_LNtl	0.057	0.068
wa_OptCat_NutMMI	0.047	0.066
pi_NAWQA_TN_1	0.457	0.084
pi_Ptpv_TP_all_Hi	0.083	0.094
x_Shan_e	9.272	0.071



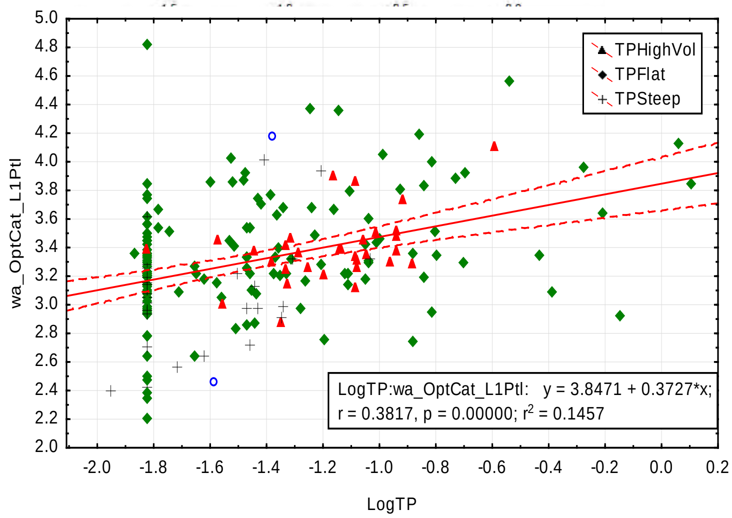
Median TP values in reference sites of the TP High-Volcanic site class were lower than those in the stressed sites and higher than those in the few extremely stressed sites (Figure 54). Most of the sites that were greater than the 90th quantile value (1.05 mg/L TP) were in the Other category. Also see Section 4.2.

Figure 13. **Site median TP value distributions along the disturbance gradient for sites in the TP High-Volcanic site class.**



Most change-points were identified at TP values less than the 90th quantile of reference sites, like the EPT taxa response (Figure 55). Only the percent clinger macroinvertebrate metric had a higher change-point. Also see Section 4.5 and Appendix L.

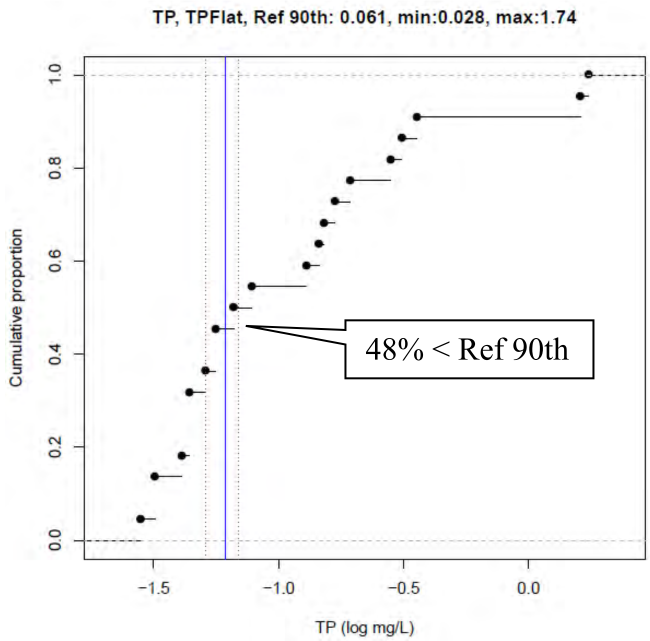
Figure 14. **Change-point plot for TP and the EPT taxa macroinvertebrate metric, showing the change-point with confidence intervals (vertical solid and dashed lines), the 95th quantile regression line (slanting dashed line), and the LOWESS regression (blue fitted line).**



All but one diatom metric had interpolated TP values less than the reference 90th quantile, such as the weighted average phosphorus diatom metric (Figure 56). Also see Section 4.4 and Appendix K.

Figure 15. **Regression plot for TP and weighted average diatom phosphorus sensitivity. In the TP High Volcanic sites, the metric upper reference quartile was 3.4, which translates to 0.068 mg/L TP.**

7.5 TP Threshold Synthesis – Flat-Moderate Site Class



90th Quantile of Ref: **0.061 mg/L**
90% CI: 0.051 – 0.069 mg/L
Medians:

All Measures: 0.066 mg/L

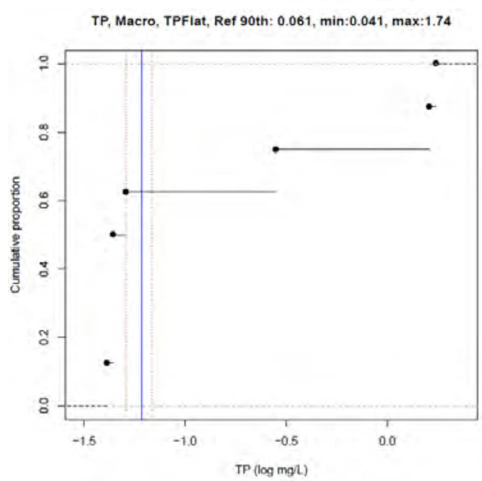
Macro Reg Int: 1.675 mg/L

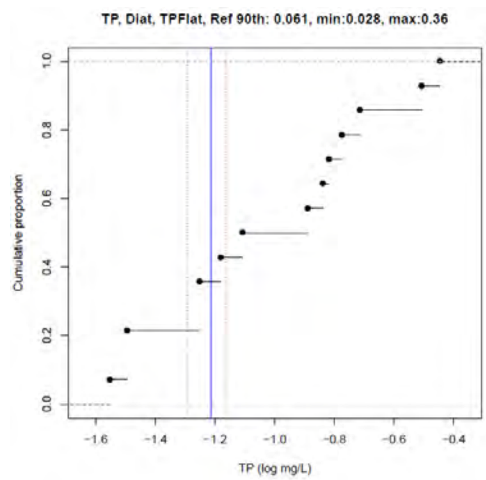
Macro CPA: 0.044 mg/L

Diatom Reg Int: 0.168 mg/L

Diatom CPA: 0.056 mg/L

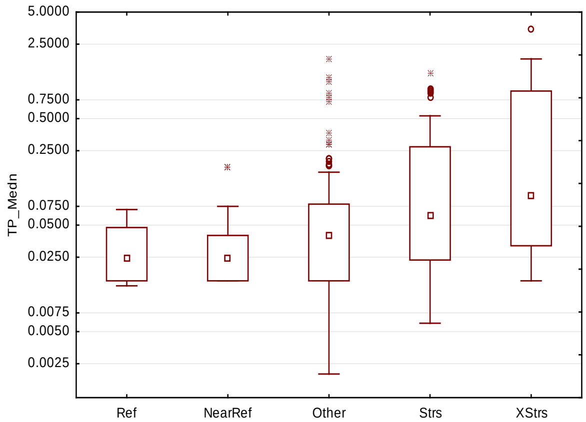
DO CPA: 0.099 mg/L





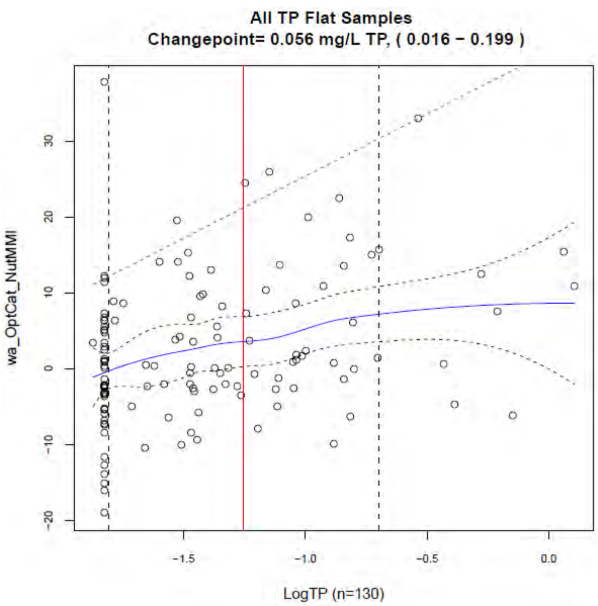
Macro Metric	Reg. Int.	CPA
EPTTax	3.39	0.044
EphemTax	211	0.044
PlecoTax	1.61	0.041
IntolTax	6.22	0.051
Toler percent	3.11	0.052
EPT percent	31.91	0.014
Pleco percent	1.74	0.044
NonIn percent	0.281	0.014
ShredTax	56.47	0.151
Clngr percent	5.01	0.051

Diatom Metric	Reg. Int.	CPA
wa_OptCat_DisTotMMI	0.168	0.056
wa_OptCat_L1DisTot	0.358	0.066
wa_OptCat_L1Ptl	0.145	0.032
wa_OptCat_LNtl	0.311	0.078
wa_OptCat_NutMMI	0.193	0.056
pi_NAWQA_TN_1	0.129	0.028
pi_Ptpv_TP_all_Hi	0.152	0.032
x_Shan_e	7.272	0.034



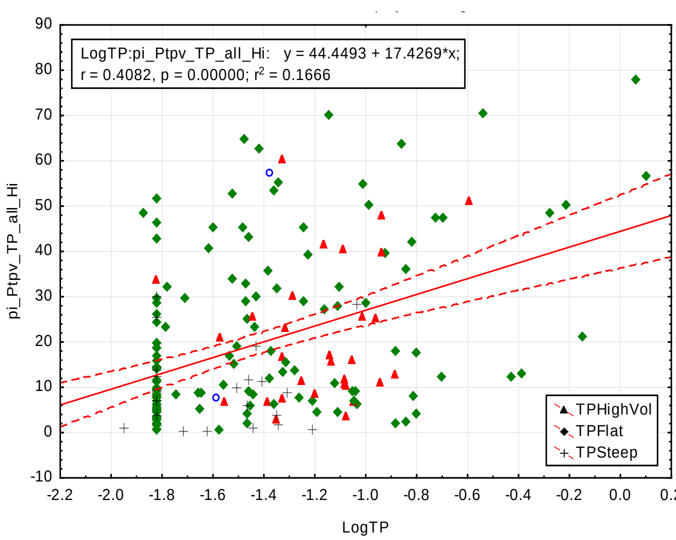
TP values were lowest in reference sites in the TP Flat-Moderate site class and increased gradually with increasing stress (Figure 57). More than half of the stressed and extremely stressed sites were greater than the 90th quantile (0.071 mg/L TP). Also see Section 4.2.

Figure 16. **Site median TN value distributions along the disturbance gradient for sites in the TP Flat-Moderate site class.**



Change-points for macroinvertebrates were lower than the 90th quantile of reference sites. For diatoms, valid change-points bracketed the 90th quantile. The weighted average nutrient index diatom metric had a change-point close to the reference 90th quantile (Figure 58). Also see Section 4.5 and Appendix L.

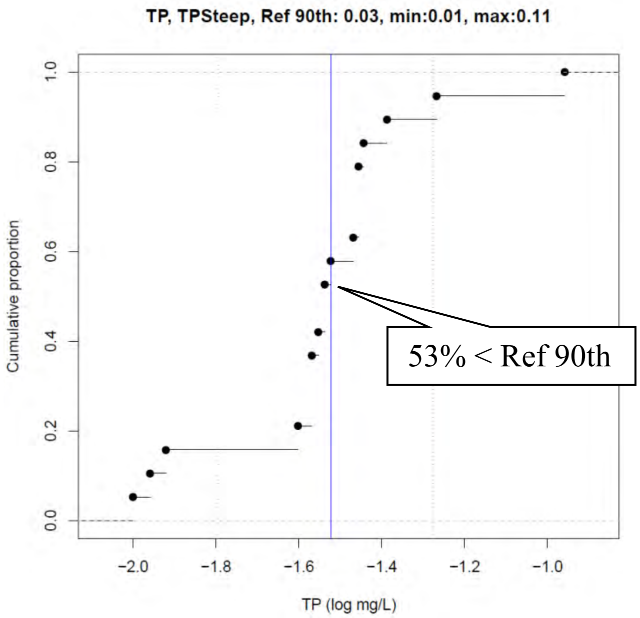
Figure 17. **Change-point plot for TP and the weighted average nutrient index diatom metric, showing the change-point with confidence intervals (vertical solid and dashed lines), the 95th quantile regression line (slanting dashed line), and the LOWESS regression (blue fitted line).**



Most regression interpolations for macroinvertebrate metrics were unreasonably high. Diatom results were also higher than the 90th quantile of reference sites (e.g., Figure 59). Also see Section 4.4 and Appendix K.

Figure 18. **Regression plot for TP and percent TP tolerant diatom metric. In the TP Flat-Moderate site class, the upper reference quartile for the metric was 30.2, which translates to 0.152 mg/L TP.**

7.6 TP Threshold Synthesis – Steep Site Class



90th Quantile of Ref: **0.030 mg/L**
90% CI: 0.016 – 0.053 mg/L
<u>Medians:</u>

All Measures: 0.029 mg/L

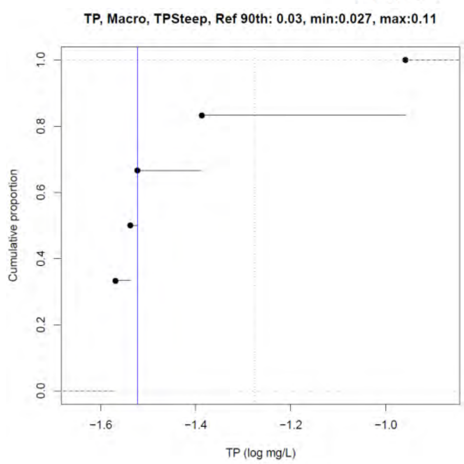
Macro Reg Int: 0.110 mg/L

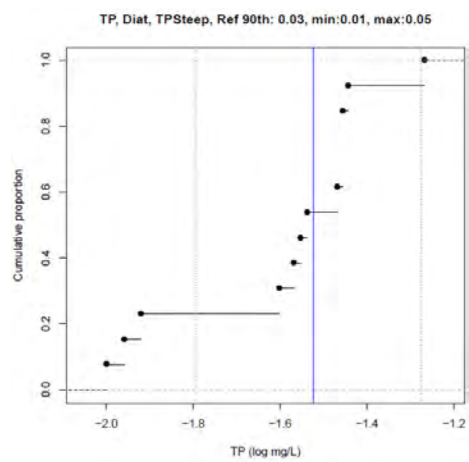
Macro CPA: 0.029 mg/L

Diatom Reg Int: 0.026 mg/L

Diatom CPA: 0.035 mg/L

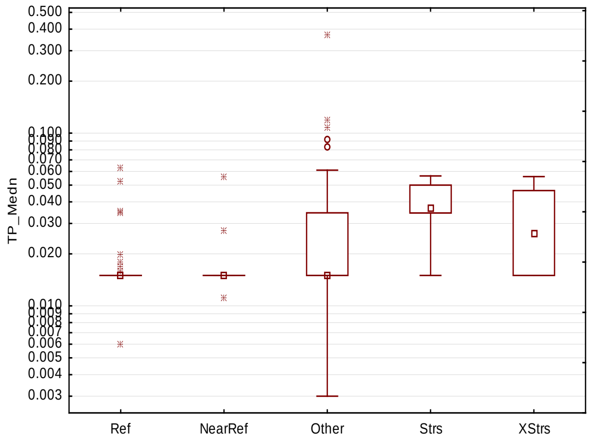
DO CPA: 0.035 mg/L





Macro Metric	Reg. Int.	CPA
EPTTax	0.11	0.030
EphemTax	1.00	0.030
PlecoTax	0.15	0.027
IntolTax	0.33	0.029
Toler percent	0.017	0.041
EPT percent	491758	0.029
Pleco percent	0.80	0.027
NonIn percent	0.003	0.018
ShredTax	0.60	0.017
Clngr percent	0.50	0.022

Diatom Metric	Reg. Int.	CPA
wa_OptCat_DisTotMMI	0.028	0.035
wa_OptCat_L1DisTot	0.027	0.034
wa_OptCat_L1Ptl	0.029	0.036
wa_OptCat_LNtl	0.054	0.035
wa_OptCat_NutMMI	0.025	0.035
pi_NAWQA_TN_1	0.010	0.019
pi_Ptpv_TP_all_Hi	0.011	0.029
x_Shan_e	0.012	0.027

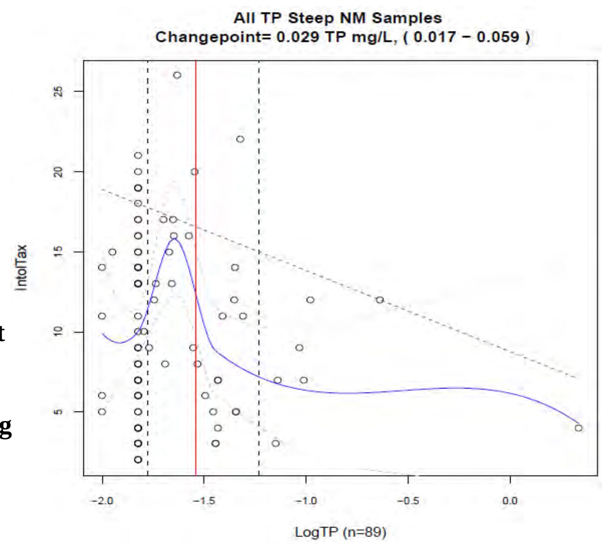


TP values were lowest in reference sites in the TP Steep site class and median values were highest in the stressed and extremely stressed sites (Figure 60). The 90th quantile value (0.054 mg/L TP) is exceeded most often in sites in the Other category. Also see Section 4.2.

Figure 19. **Site median TP value distributions along the disturbance gradient for sites in the TP Steep site class.**

Relatively few change-points were valid, but those that were fell near the 90th quantile of reference sites (e.g., Figure 61). Also see Section 4.5 and Appendix L.

Figure 20. **Change-point plot for TP and the intolerant taxa macroinvertebrate metric, showing the change-point with confidence intervals (vertical solid and dashed lines), the 95th quantile regression line (slanting dashed line), and the LOWESS regression (blue fitted line).**



Most regression interpolations for macroinvertebrate metrics were high. Diatom results were closer to the 90th quantile (e.g., Figure 62). Also see Section 4.4 and Appendix K.

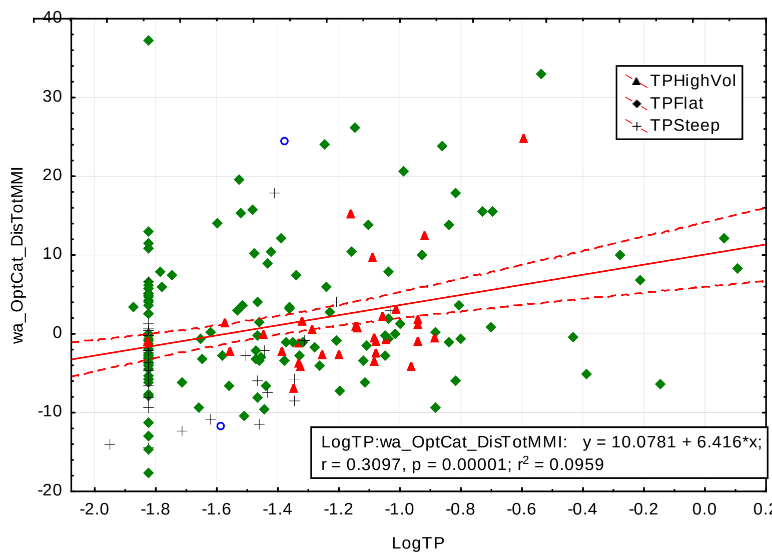


Figure 21. **Regression plot for TP and weighted average disturbance index metric. In the TP Steep site class, the reference quartile for the metric was 1.1, which translates to 0.028 mg/L TP.**