10-1-2016

Database Analysis to Support Nutrient Criteria Development (Phase I)

B. E. Haggard

J.T. Scott

M.A. Evans-White

ARKANSAS WATER RESOURCES CENTER | PUBLICATION MSC381
FUNDED BY THE TEXAS COMMISSION ON ENVIRONMENTAL QUALITY

DATABASE ANALYSIS TO SUPPORT NUTRIENT CRITERIA DEVELOPMENT (PHASE I)

2016 October

ARKANSAS WATER RESOURCES CENTER
DIVISION OF AGRICULTURE RESEARCH & EXTENSION
University of Arkansas System

The intent of this publication of the Arkansas Water Resources Center is to provide a location whereby a final report on water research to a funding agency can be archived.

The Texas Commission on Environmental Quality (TCEQ) contracted with University of Arkansas researchers for a multiple year project titled “Database Analysis to Support Nutrient Criteria Development”.

This publication covers the first of three phases of that project and has maintained the original format of the report as submitted to TCEQ. This report can be cited either as an AWRC publication (see below) or directly as the final report to TCEQ.

Haggard, B.E., J.T. Scott, and M.A. Evans-White. 2016. Database Analysis to Support Nutrient Criteria Development (Phase I). Arkansas Water Resources Center, Fayetteville, AR, MSC Publication 381: 183 pp.

DATABASE ANALYSIS TO SUPPORT NUTRIENT CRITERIA DEVELOPMENT

final report submitted in fulfillment of
contract number: 582-11-12791 to:

Laurie Eng
Project Manager, Water Quality Standards Group
Texas Commission on Environmental Quality

By

Principal Investigator:
Brian E. Haggard, Professor and Director, Arkansas Water Resources Center
Co-Principal Investigators:
J. Thad Scott, Assistant Professor, Crop, Soil and Environmental Sciences Department
Michelle Evans-White, Assistant Professor, Biological Sciences Department
University of Arkansas

August 31, 2011

Executive Summary

A large amount of literature on nutrient criteria has emerged, since the EPA first called for the development of numeric criteria by states and tribes in 1998. The review of the literature (see **Chapter 1**) showed that two main approaches existed, where one approach focused on the frequency distribution of median concentrations from streams or a select group representing reference conditions and the other examined statistical, predictive relations between nutrients and biological response variables. Predictive approaches have focused on establishing relations between nutrient concentrations and algae, macroinvertebrates and fish communities, and determining whether thresholds exist that can aide in nutrient criteria development. Most predictive approaches have occurred at the state or even watershed level, which may or may not be comparable to the nutrient criteria guidance provided by EPA at the aggregate eco-region level. Although not completely comparable, it is interesting to note that the ranges in numeric thresholds are similar when estimated via the frequency distribution method or via predicting biological response to nutrients.

The purpose of this project was to provide statistical support to the Texas Commission on Environmental Quality (TCEQ) that would aide in this agency's development of numeric nutrient criteria for streams, rivers and reservoirs. The first step in this process was the acquisition and compilation of geospatial, water quality (e.g., chemical concentrations), and bioassessment data from 2,482 stations spanning 23 basins across Texas. Following data reorganization and reduction, median values for each parameter were estimated at each, individual station and then compiled into an overall median database. The parameters of primary concern were total phosphorus (TP), ortho-phosphate (PO_4-P), total nitrogen (TN), nitrate plus nitrite N (NO_x-N), and sestonic chlorophyll-a (chl-a). Frequency distributions including the minimum, 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles, and maximum of these parameters were calculated at multiple spatial scales, such as by basin, eco-region levels III and IV, and basin by level III eco-region. These distributions are intended to provide guidance to TCEQ per EPA recommendations, and **Chapter 2** shows that variations in the percentiles exist between basins, eco-regions and other spatial-scales.

States across the US are moving forward with the development of nutrient criteria, but many states are concerned about the legitimacy of promulgating one numeric criterion across the whole state that represents multiple basins, various eco-regions, and a myriad of land uses. **Chapter 3** evaluated the potential to use categorical geographic factors to predict median TP concentrations from stations with available data, because TP represented the nutrient parameter with the largest number of medians (based on a minimum of 10 observations per station). Several of the categorical variables explained significant amounts of variation in median TP concentrations, and the order of the primary splits was: basin by level III eco-region, level IV eco-region, level III eco-region, and then land use (urban then forested). The stations were then separated into two groups, based on basin by level III ecoregions representing one group with a mean TP concentration of 0.19 mg/L based on 865 medians (i.e., the "low P" category) and the second group with a mean of 0.77 mg/L based on 169 medians ("high P" category).

The three databases used in the next Chapter were the overall median database, the "low P" median database, and the "high P" median database.

The frequency distribution approach should be used in conjunction with other statistically valid methods evaluating stressor-response relations in streams. **Chapter 4** evaluated potential biological thresholds, using categorical and regression tree (CART) analyses on the three median databases to identify changepoints which result in substantial ecological change. TP thresholds that resulted in ecological changes ranged from 0.06 to 0.09 mg/L in the overall and "low P" median database, when using Secchi depth, dissolved oxygen (DO) flux, and chl-a as biological response variables. There was only one meaningful, ecological changepoint in the "high P" median database (~0.25 mg/L). Similarly, TN thresholds were identified in the overall and "low P" median database, ranging from 0.8 to 1.6 mg/L; no TN thresholds were significant in the "high P' median database. These numeric thresholds should be used to provide guidance to TCEQ, when moving forward to develop numeric nutrient criteria.

A subset of streams has more intensive biological and habitat data available for select time periods, and these indices of biological integrity (i.e., fish IBI and RBIBI) and habitat scores (HQI) were integrated with median nutrient concentrations representing the index, non critical or critical period that the biological data was collected. In the complete median bioassessment database, TP thresholds that resulted in changes in the IBI scores for fish and macroinvertebrates were 0.059 and 0.065 mg/L, respectively; thresholds were much more variable in the "low P" and "high P" groupings within the bioassessment database. TN thresholds that resulted in ecological changes ranged from 1.3 to 2.5 mg/L and did not vary much between data sources (i.e., overall bioassessment data, "low P" or "high P"). Overall, habitat (HQI) was an important covariate that explained more variation in IBI scores than did nutrients, except for the 0.059 mg/L TP thresholds that explained 21% of the variation in RBIBI scores across the entire bioassessment database. **Chapter 5** details these statistical analysis and relations between nutrients, habitat and IBI scores.

The final chapter, **Chapter 6**, shifts from streams and rivers to reservoirs, and focuses on the classical relations between TP, Secchi depth and chl-a (in raw and median data) and how chl-a concentrations change over time; the same statistical procedure (CART) was used in this chapter to identify thresholds in TP and time. In these Texas reservoirs, the most consistent threshold in TP correlated with changes in Secchi depth and chl-a was approximately 0.04 mg/L (based on CART analysis of the median database). The raw data showed a TP threshold of 0.06 mg/L, but this model was much weaker than those developed with the median database. CART analysis on chl-a over time identified statistical significant temporal thresholds in chl-a concentrations at all reservoirs tested. However, the temporal thresholds were inconsistent in time across these reservoirs and only four of the temporal thresholds appeared to be related to changes in method detection limits or other methodological changes for chl-a.

These statistical analyses are intended to support TCEQ's efforts in the development of numeric nutrient criteria, and we are not advocating for specific numbers but simply providing literature review, database management, and statistical support to TCEQ through this contract.

Table of Contents

Executive Summary i

Table of Contents iii

List of Tables v

List of Figures vi

Chapter 1: Nutrient Criteria in Streams Literature Review 1-1
- Background 1-2
- Percentile Analysis of Nutrient or Algal Variables 1-3
- Predictive Relationships between Nutrients and Ecological Conditions 1-10
- Applicable Texas Studies 1-20
- Synthesis and Conclusions 1-23
- Literature Cited 1-24
- Appendix 1-1: Stream Study Report from the State of Texas Reviews 1-34

Chapter 2: Database Development, Median Calculations and Frequency Distributions 2-1
- Introduction 2-2
- Methods 2-3
- Results and Discussion 2-5
- References 2-12
- Appendix 2-1: Level II and Level IV Ecoregions in Texas 2-13
- Appendix 2-2: Texas River and Coastal Basins 2-14
- Appendix 2-3: Frequency Distribution of Median Nutrient and Chlorophyll-a Concentrations among Basin by Level III Ecoregions in Texas, 2000-2010. 2-15
- Appendix 2-4: Frequency Distribution of Median Nutrient and Chlorophyll-a Concentrations among Level IV Ecoregions in Texas, 2000-2010. 2-22

Chapter 3: Threshold Analysis to Identify Station Groupings Representing Similar Nutrient Conditions 3-1
- Introduction 3-3
- Methods 3-3
- Results and Discussion 3-3
- References 3-6
- Appendix 3-1: CART on Nutrients versus Geography and Watershed Attributes 3-7
- Appendix 3-2: Changepoint on Nutrients versus Basin by Ecoregion III 3-9
- Appendix 3-3: Changepoint on Nutrients versus Ecoregion IV 3-10
- Appendix 3-4: Changepoint on Nutrients versus Ecoregion III 3-11
- Appendix 3-5: Changepoint on Nutrients versus Developed Land 3-12

Chapter 4: Threshold Analysis on Streams and Rivers Median Database 4-1
- Introduction 4-2
- Methods 4-3
- Results and Discussion 43
- References 4-15
- Appendix 4-1: CART Models on Biologicals versus Nutrients—All Median Data 4-16

Appendix 4-2: CART Models on Biologicals versus Nutrients—Data from Low P Basin by Ecoregion III 4-22
Appendix 4-3: CART Models on Biologicals versus Nutrients—Data from High P Basin by Ecoregion III 4-28

Chapter 5: Threshold Analysis on Streams and Rivers Database with Bioassessment Data 5-1
Introduction 5-2
Methods 5-2
Results and Discussion 5-4
References 5-15
Appendix 5-1: CART Models on Biologicals versus Nutrients and Habitat 5-16
Appendix 5-2: CART Models on Biologicals versus Nutrients and Habitat from "Low P" Basin by Ecoregion 5-24
Appendix 5-3: CART Models on Biologicals versus Nutrients and Habitat from "High P" Basin by Ecoregion 5-26

Chapter 6: Threshold Analyses on Reservoir Data 6-1
Introduction 6-2
Methods 6-2
Results and Discussion 6-3
References 6-12
Appendix 6-1: Threshold Analyses on Median Reservoir Data 6-14
Appendix 6-2: Threshold Analyses on Reservoir Raw Data 6-16
Appendix 6-3: Thresholds in Reservoir Chlorophyll-a in Time Series 6-18

List of Tables

Number	Title	Page
1-1	Total nitrogen (mg/L) 25th and 75th percentiles across United States Environmental Protection Agency (EPA) Aggregate Nutrient Ecoregions taken from peer-reviewed literature and compared to EPA suggested criteria.	1-5
1-2	Total phosphorus (µg/L) 25th and 75th percentiles across United States Environmental Protection Agency (EPA) Aggregate Nutrient Ecoregions taken from peer-reviewed literature and compared to EPA suggested criteria.	1-5
1-3	Suspended chlorophyll-a (µg/L) and turbidity (NTU or FTU) 25th and 75th percentiles across US Environmental Protection Agency (EPA) Aggregate Ecoregions from peer-reviewed literature and compared to EPA suggested criteria.	1-10
1-4	Benthic algal total nitrogen (TN) and total phosphorus (TP) thresholds determined by various statistical analyses.	1-14
1-5	Benthic macroinvertebrate total nitrogen (TN) and total phosphorus (TP) criteria determined by various statistical analyses.	1-17
1-6	Stream fish total nitrogen (TN) and total phosphorus (TP) thresholds determined by various threshold analyses.	1-17
2-1	Frequency distribution of median nutrient and chlorophyll-a concentrations among basins in the State of Texas, 2000-2010.	2-6
2-2	Frequency distribution of median nutrient and chlorophyll-a concentrations among level III ecoregions in the State of Texas, 2000-2010.	2-9
2-3	EPA recommended nutrient criteria for total phosphorus, total nitrogen, and chlorophyll-a for the aggregate ecoregions that in the State of Texas (EPA, 2000).	2-11
3-1	Separation of individual stream and river stations by basin and level III ecoregions based on median total phosphorus (TP) concentrations; the groups represent similar TP conditions.	3-4
4-1	Nutrient thresholds identified in the primary split of CART analyses for the entire streams and rivers median data set and for data sets that were grouped according to the “Low P” and “High P” basin by ecoregion III classification.	4-3
5-1	Nutrient thresholds identified in the primary split of CART analyses for the entire bioassessment median dataset and for datasets that were grouped according to the “Low P” and “High P” basin by ecoregion III classification.	5-4
5-2	Nutrient thresholds identified in the primary split of CART analyses for the entire bioassessment median dataset and for datasets that were grouped according to the “Low P” and “High P” basin by ecoregion III classification.	5-11
6-1	Phosphorus thresholds identified by CART analyses for reservoir median and raw data databases.	6-3

List of Figures

Number	Title	Page
1-1	Boxplots of the ratio of the literature-estimated total nitrogen (TN) Criterion to the EPA-suggested TN criterion for each aggregate ecoregion. *Texas Aggregate Ecoregion.	1-9
1-2	Boxplots of the ratio of the literature-estimated total phosphorus (TP) thresholds to the EPA-suggested TP Criterion for each aggregate ecoregion. *Texas Aggregate Ecoregion.	1-9
2-1	Distribution of data collected from reference condition streams and the general stream population and the associated percentile distribution used to develop nutrient criteria.	2-2
4-1	CART model for median Secchi depth versus median nutrient concentrations (TP, TN, and NO_x-N) from the entire median dataset (model $r^2 = 0.25$).	4-4
4-2	CART model for median chlorophyll-a concentrations determined by spectrophotometry versus median nutrient concentrations (TP, TN, and NO_x-N) from the entire median dataset (model $r^2 = 0.37$).	4-5
4-3	CART model for median chlorophyll-a concentrations determined by fluorometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the entire median dataset (model $r^2 = 0.84$).	4-6
4-4	CART model for median merged chlorophyll-a data versus median nutrient concentrations (TP, TN, and NO_X-N) from the entire median dataset (model $r^2 = 0.54$).	4-7
4-5	CART model for Secchi depth versus median nutrient concentrations (TP, TN, and NO_X-N) from the “Low P” Basin x Ecoregion III dataset (model $r^2 = 0.24$).	4-8
4-6	CART model for median chlorophyll-a concentration determined by spectrophotometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the “Low P” basin by ecoregion III dataset (model $r^2 = 0.40$).	4-9
4-7	CART model for median chlorophyll-a concentration determined by fluorometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the “Low P” basin by ecoregion III dataset (model $r^2 = 0.74$).	4-10
4-8	CART model for median merged chlorophyll-a data versus median nutrient concentrations (TP, TN, and NO_X-N) from the “Low P” Basin x Ecoregion III dataset (model $r^2 = 0.74$).	4-11
4-9	CART model for median Secchi depth versus median nutrient concentrations (TP, TN, and NO_X-N) from the “High P” Basin x Ecoregion III dataset (model $r^2 = 0.52$).	4-13
4-10	CART model for median chlorophyll-a measured with fluorometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the “High P” Basin x Ecoregion III dataset (model $r^2 = 0.17$).	4-14
5-1	CART model for median Secchi depth versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model $r^2 = 0.12$).	5-5

5-2 CART model for median dissolved oxygen flux versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.43). 5-6

5-3 CART model for median chl-a concentrations determined by spectrophotometry versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.20). 5-7

5-4 CART model for median chl-a concentrations determined by fluorometry versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.33). 5-7

5-5 CART model for median merged chl-a data versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.23). 5-8

5-6 CART model for fish IBI versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.60). 5-9

5-7 CART model for RBIBI versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.21). 5-10

5-8 CART model for fish IBI versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment “Low P” Basin x Ecoregion median dataset (model r^2 = 0.48). 5-12

5-9 CART model for RBIBI vebnbrsus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment “Low P” Basin x Ecoregion median dataset (model r^2 = 0.29). 5-13

5-10 CART model for fish IBI versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment “High P” Basin x Ecoregion median dataset (model r^2 = 0.19). 5-14

6-1 CART model for median Secchi depth versus median TP from the reservoir median dataset (model r^2 = 0.34). 6-4

6-2 CART model for median chl-a concentrations versus median TP concentrations from the reservoir median dataset (model r^2 = 0.27). 6-4

6-3 CART model for Secchi depth versus TP from the reservoir raw dataset (model r^2 = 0.13). 6-5

6-4 CART model for chl-a concentrations versus TP concentrations from the reservoir raw dataset (model r^2 = 0.06). 6-6

6-5 CART model for chl-a concentrations through time at Lake O’ the Pines (model r^2 = 0.36). 6-7

6-6 CART model for chl-a concentrations through time at Lake Tawakoni (model r^2 = 0.27). 6-7

6-7 CART model for chl-a concentrations through time at Lake Fork (model r^2 = 0.05). 6-8

6-8 CART model for chl-a concentrations through time at Lake Houston (model r^2 = 0.36). 6-8

6-9 CART model for chl-a concentrations through time at Lake Conroe (model r^2 = 0.05). 6-9

6-10 CART model for chl-a concentrations through time at Lake Whitney (model r^2 = 0.36). 6-9

6-11 CART model for chl-a concentrations through time at Lake Belton (model r^2 = 0.20). 6-10

6-12 CART model for chl-a concentrations through time at Lake Travis (model r^2 = 0.05). 6-11

6-13 CART model for chl-a concentrations through time at Lake Buchanan (model r^2 = 0.10). 6-11

6-14 CART model for chl-a concentrations through time at Lake Amistad (model r^2 = 0.06). 6-12

Chapter 1: Nutrient Criteria in Streams Literature Review

Michelle A. Evans-White

Assistant Professor, Biological Sciences Department
University of Arkansas, Science and Engineering 601, Fayetteville, Arkansas 72701
Corresponding Author: mevanswh@uark.edu

EXECUTIVE SUMMARY: A large amount of literature on nutrient criteria has been generated since the EPA first called for the development of nutrient criteria by states and tribes in 1998. Our objective was to summarize and evaluate the use of the percentile analysis of nutrient or algal variables and the use of predictive statistical analyses to develop nutrient criteria for streams and rivers. The focus of our review was on peer-reviewed literature published after 1998. Two main approaches exist for criteria development. One approach identifies reference reaches based on percentile analysis of data frequency distributions. The other main approach examines predictive relationships between nutrient or algal criteria variables and response variables representing ecological condition to establish desired levels for criteria. The focus of literature utilizing the percentile approach to nutrient criteria development has been on comparing 25th and 75th percentile approaches, examining factors that cause variability within aggregate ecoregions, and comparing 25th and 75th percentile results obtained with more modern data or with data collected using a more rigorous statistical design compared to EPA-proposed criteria. Overall consensus of many studies suggests that aggregate ecoregions are too coarse for criteria development, but the majority of comparisons between EPA percentiles and literature percentiles are most abundant at the aggregate ecoregion level. We found that literature TN criteria for the Great Plains Grass and Shrublands (IV) and the Southeastern Temperate Forested Plains and Hills (IX), both aggregate ecoregions found in Texas, were highly variable and were up to approximately 2.5 times higher than EPA-suggested criteria. Therefore, future studies might focus on more appropriate factors describing variability in this Texas aggregate ecoregion. Literature TP criteria were more variable for the Great Plains Grass and Shrublands aggregate ecoregion than any other excluding the Eastern Coastal Plain. Finally, little TN and TP criteria data are available for the Texas-Louisiana Coastal and Mississippi Alluvial Plain compared to other Texas aggregate ecoregions. Predictive approaches have focused on establishing relationships between water quality and algae, macroinvertebrate, and fish communities, attributing causation, and determining whether threshold points exist that can aid in nutrient criteria development. Many studies have found linear and non-linear relationships between benthic and suspended algae, macroinvertebrate communities, and fishes. Most of the predictive approaches have occurred at the state or watershed level and may not be directly comparable to EPA aggregate ecoregions. Although not completely comparable, criteria estimated via the percentile method had similar ranges to biological threshold criteria.

BACKGROUND

Eutrophication occurs when ecosystem primary production is stimulated by the addition of nutrients needed for growth such as nitrogen (N) or phosphorus (P). Cultural eutrophication occurs when humans are the source of growth-limiting nutrients to ecosystems. Large primary producer populations can cause large fluctuations in dissolved oxygen (DO) concentrations between the day when photosynthesis predominates and the night when respiration by algae and bacteria are the dominant metabolic processes. Dissolved oxygen levels can become low enough in the dark hours that biota at higher trophic levels including insects, crustaceans, mollusks, and fishes become stressed or die from the lack of sufficient oxygen to respire. Chronic nutrient enrichment can also favor algal species that cause drinking water taste and odor problems or that produce toxins that can harm animal and human health (Carmichael 1986).

Elevated N concentrations can also have direct negative effects on aquatic biota due to toxicity (Camargo and Alonso 2006). The most common forms of dissolved N in freshwaters are nitrate (NO_3^-) and ammonium (NH_4^+). The concentration of NH_4^+ and ammonia (NH_3) are interrelated via chemical equilibrium; as temperature and pH increase, NH_3 becomes more abundant. Ammonia can be toxic to fishes at 0.8-0.35 mg/L [96 h LC50 (lethal concentration with 50% mortality of test organisms); Ball 1967, Rice and Bailey 1980] and to macroinvertebrates at 0.11-0.65 mg/L (96 h LC50; Mummert et al. 2003, Alonso and Camargo 2003). Ammonium (NH_4^+) is nontoxic or less toxic than NH_3 is to aquatic organisms (Russo 1985, Constable et al. 2003, Camargo and Alonso 2006). Nitrate levels can be directly toxic to sensitive invertebrates and fishes above 2 mg/L NO_3-N (Camargo et al. 2005) and can cause blue baby syndrome in infants when concentrations exceed 10 mg/L in drinking water (Knobelach et al. 2000). There is also some evidence linking ingested nitrates to gastric and other cancers in humans (Joossens et al. 1996; Vermeer et al. 1998).

National Water Quality Inventories consistently conclude that excess nutrients are impairing US waters, where impairment is defined as not meeting designated use criteria. A 1998 Water Quality Inventory by the US Environmental Protection Agency (EPA) found that 40% of US rivers were impaired and listed nutrients as the second leading cause of impairment (EPA 2002). Total N (TN) and total P (TP) in US streams and rivers have been estimated to exceed background levels by a factor of 6.4 and 2.0, respectively (Smith et al. 2003). In 2006, the EPA published their Wadeable Streams Assessment reporting that 42% of the nation's stream length was in poor, 25% in fair, and 28% in good biological condition. Nitrogen and P were listed as the stressors affecting the largest percentage of stream length and were ranked a close second behind sediments as posing a risk to biological condition (EPA 2006; Paulson et al. 2008). Wadeable streams are important because they account for approximately 90% of the total length of perennial streams and rivers. In addition, small, wadeable streams can have significant impacts on water quality downstream in larger rivers (Alexander et al. 2008; Dodds and Oakes 2006; 2008).

The Clean Water Action Plan was a presidential initiative released in 1998 that provides a blueprint for establishing nutrient criteria to protect and restore US waters. Water quality criteria should protect the

designated use for the water body and be part of an antidegradation policy to protect existing water quality. Criteria should be scientifically based and designated uses should involve economic, political, and social considerations including the protection of downstream waters. The EPA identified 9 key steps in the criteria development process for streams (EPA 2000). First, federal and state agencies, tribes, and stakeholders must identify water quality needs and goals. They must sort rivers and streams into groups having comparable characteristics including trophic status. Variables to evaluate or predict the degree of eutrophication must be chosen, and these need to include TN, TP, chlorophyll *a* (chl-a) to estimate algal biomass, and an estimate of turbidity. They must then design a monitoring strategy for collecting nutrient and algal biomass data, collect the data, and analyze the data with a focus on statistical analysis relating nutrient variables to algal biomass or to changes in ecological condition indicating eutrophication such as low DO. Criteria can then be developed based on reference conditions or on data analyses. Finally, government agencies and tribes must implement the nutrient control strategies and measure the effectiveness of those strategies to reassess the validity of the criteria.

Since the first call for development of nutrient criteria in 1998, a large amount of peer-reviewed literature has been produced focusing on statistical analysis approaches to developing nutrient criteria. These analyses can be sorted into two main types. One approach identifies reference reaches based on percentile analysis of data frequency distributions. The other main approach examines predictive relationships between nutrient or algal criteria variables and response variables representing ecological condition to establish desired levels for criteria. The objective of this report is to summarize and evaluate the use of the percentile analysis of nutrient or algal variables and the use of predictive statistical analyses to develop nutrient criteria for streams and rivers. The focus of our review is on peer-reviewed literature published after 1998.

PERCENTILE ANALYSIS OF NUTRIENT OR ALGAL VARIABLES

The EPA suggests that reference criteria be based on the 75th percentile of a distribution of a population of reference streams. They also suggested that the lower 25th percentile of a population of all streams within each region can be used as a surrogate if a reference stream population does not exist. In 2000, the EPA provided an initial aggregate Omernik Level III ecoregional TN, TP, suspended chl-a, and turbidity criteria for streams (Tables 1-1 through 1-3). Given the lack of reference streams in many ecoregions, they reported the lower 25th percentile of a population of all streams within each region excluding the Southern Florida Coastal Plain region for data collected from 1991-1995. Since this initial attempt by the EPA to provide nutrient criteria, several studies have used a similar approach to define criteria and have compared their analyses to the EPA's original proposed nutrient criteria. Results of these studies are reviewed in this section and reported in Tables 1-1 through 1-3. Special emphasis was placed on patterns observed in Texas aggregate ecoregions including the Xeric West (III), Great Plains Grass and Shrublands (IV), South Central Cultivated Great Plains (V), Southeastern Temperate Forested Plains and Hills (IX), and the Texas-Louisiana Coastal and Mississippi Alluvial Plains (X) aggregate ecoregions.

Rohm et al. (2002) found that the 14 EPA aggregated ecoregions described meaningful spatial patterns in TN and TP taken in a subset of streams sampled in the US EPA National Eutrophication Survey (NES). However, the authors suggested that their percentile analysis using NES data would not be appropriate for use in developing criteria because the absence of point sources to streams could not be confirmed and the data were dated. Therefore, percentiles were not included in Tables 1-1 and 1-2. Although aggregate ecoregions were useful in explaining spatial patterns in TN and TP data in this study, the authors suggested that these aggregate ecoregions were too coarse of a scale for setting nutrient criteria using survey data alone. Incorporation of land cover variables has been found to explain 3 and 6 times more variation in TP and TN data than ecoregions in NES data (Wickham et al. 2005). So, incorporating land cover may fine tune nutrient criteria estimates. Rohm et al. (2002) also used EPA Environmental Monitoring and Assessment Program (EMAP) data from the Central and Eastern Forested Plains ecoregion to examine 75th percentile TN and TP concentrations (Tables 1-1 and 1-2). They argued that EMAP data should be used because it was collected using a statistical survey design so that each site sampled was representative of a portion of the 'target population' of aquatic resources within a particular region. In addition, EMAP data should be critically examined to filter out sites that have known and quantified impacts related to nutrient enrichment before 75th percentiles are examined.

Several factors constrain the use of the 75th percentile reference stream approach. Pristine reference sites are virtually non-existent causing managers to use sites in moderately developed watersheds. Further, atmospheric N deposition can be an important factor altering stream chemistry in even primarily forested watersheds (Flum and Nodvin 1995). Finally, most of the US reference sites are located in small watersheds because too few large watersheds remain undeveloped. Smith et al. (2003) attempted to overcome these limitations by developing an empirical model incorporating regression models and the SPARROW transport model that would provide background TN and TP yield and concentrations. They used data from 63 minimally impacted small US Geological Survey (USGS) reference basins located in the 14 EPA nutrient ecoregions of the coterminous US. Explanatory variables in their model included annual runoff, basin size, atmospheric N deposition rate and region-specific factors (e.g., geology and vegetation type). When atmospheric N deposition was incorporated into the model, the upper quartile of TN background concentrations in each region were 15-100% higher than those estimated when deposition was not incorporated (Table 1-1). Texas aggregate regions IV, V, and X were at the high end of the range of TP concentration 75th percentiles reported (Table 1-2) while TN concentration 75th percentiles in most of the Texas aggregate ecoregions were some of the lowest reported in this study (Table 1-1). Regional background TN ($r = 0.60$) and TP ($r = 0.63$) concentrations correlated positively with EPA estimates. However, large local variation in runoff due to large variation in elevation and differences in cumulative in-stream loss at junctions of small tributaries and large rivers led to large local variation in background TN and TP concentrations within some ecoregions. This variation suggests that some of the background concentrations in streams in these regions exceed the proposed criteria. As suggested by Rohm et al. (2002), this study suggests that the 14 aggregate ecoregions used by the EPA were too coarse to be used for establishing nutrient criteria.

Table 1-1. Total nitrogen (mg/L) 25th and 75th percentiles across United States Environmental Protection Agency (EPA) Aggregate Nutrient Ecoregions taken from peer-reviewed literature and compared to EPA suggested criteria.

	Aggregate EPA Nutrient Ecoregions													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
25TH Percentile EPA (2000)	0.31	0.12	0.38	0.56	0.88	2.18	0.54	0.38	0.69	0.76	0.31	0.9	--	0.71
25th percentiles of a general population														
Palmstrom 2005	--	--	--	--	--	--	0.48	0.29	2.01	--	0.29	--	--	1.85
Suplee et al. 2007	--	0.08*	--	0.61	0.60	--	--	--	--	--	--	--	--	--
Herlihy and Sifneos 2008	--	0.07	0.78	0.44	0.99	1.86	0.58	0.27	0.33	0.92	0.16	--	--	0.62
Longing and Haggard 2010	--	--	--	0.61	0.86	--	--	--	0.53	--	0.21	--	--	--
Tran and Smith 2010	--	--	--	--	--	--	0.53	--	--	--	--	--	--	--
75th percentiles of a reference population														
Rohm et al. 2002	--	--	--	--	--	--	--	--	--	--	0.37	--	--	--
Smith et al. 2003	0.18	0.18	0.05	0.12	0.37	0.44	0.17	0.18	0.17	0.55	0.17	0.61	0.65	0.63
Smith et al. 2003 (with N deposition)	0.21	0.21	0.11	0.21	0.51	0.62	0.33	0.28	0.28	0.67	0.29	0.71	0.79	0.76
Suplee et al. 2007	--	0.13*	--	1.30	1.12	--	--	--	--	--	--	--	--	--
Herlihy and Sifneos 2008	--	0.15	0.29	0.93	1.19	2.5	--	0.39	0.68	--	0.294	--	--	--
Tran and Smith 2010	--	--	--	--	--	--	0.48	--	--	--	--	--	--	--

* Value represents a mean of medians for the Northern Rockies, Middle Rockies, and Canadian Rockies; -- = no data

Table 1-2. Total phosphorus (µg/L) 25th and 75th percentiles across United States Environmental Protection Agency (EPA) Aggregate Nutrient Ecoregions taken from peer-reviewed literature and compared to EPA suggested criteria.

	Aggregate EPA Nutrient Ecoregions													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
25TH Percentile EPA (2000)	47	10	21.9	23	67	76.2	33	10	36.6	128	10	40	--	3.8
25th percentiles of a general population														
Palmstrom 2005	--	--	--	--	--	--	20	16	42	--	12	--	--	82
Suplee et al. 2007	--	13*	--	20	20	--	--	--	--	--	--	--	--	--
Herlihy and Sifneos 2008	--	3	10.4	18.9	34.4	65.8	17	6.8	20.4	147	3.9	--	--	22.7
Longing and Haggard 2010	--	--	--	20	70	--	--	--	60	--	20	--	--	--
Tran and Smith 2010	--	--	--	--	--	--	29	--	--	--	--	--	--	--
75th percentiles of a reference population														
Rohm et al. 2002	--	--	--	--	--	--	--	--	--	--	13	--	--	--
Smith et al. 2003	20	20	30	70	70	60	30	20	50	60	20	30	40	20
Smith et al. 2003 (with N deposition)	--	9*	--	170	140	--	--	--	--	--	--	--	--	--
Suplee et al. 2007	--	19	40	86.8	107	181	--	10.2	60.1	--	17.7	--	--	--
Herlihy and Sifneos 2008	--	--	--	--	--	--	18	--	--	--	--	--	--	--
Tran and Smith 2010	--	--	--	--	--	--	--	--	--	--	13	--	--	--

* Value represents a mean of medians for the Northern Rockies, Middle Rockies, and Canadian Rockies; -- = no data

Dodds and Oakes (2004) described a multiple regression approach to estimate reference nutrient concentrations that incorporated land use. Initially, they quantified regional variation in TN and TP concentrations using region as a categorical predictor and land use as a covariate. They then developed multiple linear regressions where land use classifications were independent variables and nutrient concentrations were dependent variables. This allows calculation of the intercept, which represents the expected nutrient concentration in the absence of human activity. This approach offers a method of establishing a reference condition when reference sites are rare.

Another assumption associated with the 75th percentile estimates of reference streams and the 25th percentile of a general population of streams is that the two estimates should be similar. However, a study examining both approaches using data collected from Montana reference and a general population of streams found that reference 75th percentiles ranged from the 4th to the 97th percentile of the general population of streams (Suplee et al. 2007). Differences between the TN 75th percentile of all references sites and the EPA suggested TN criteria were most dramatic for the Great Plains Grass and Shrublands (IV) and the South Central Cultivated Great Plains (V), which are both aggregate ecoregions found in Texas (Table 1-1). Differences between TP 75th percentiles and EPA suggested criteria were less dramatic (Table 1-2). Differences in the outcomes of the two percentile approaches to criteria estimation suggest that the 25th percentile of the general population of streams may be either overly stringent or not protective enough. In addition, the researchers matched nutrient concentrations from 5 regional scientific studies to their corresponding reference population and found them to match on average to the 86th percentile with a coefficient of variation of 13% (Suplee et al. 2007). The consistency between the concentrations found to cause change in a "beneficial water use", as defined by Montana, in the scientific studies and the reference stream populations suggests that nutrient concentrations at high percentiles of reference stream distributions represent a meaningful threshold where changes in ecological condition can occur.

An analysis of reference streams at a broader regional scale also illustrated inconsistencies between the 25th and 75th percentile approaches. Herlihy and Sifneos (2008) analyzed data from 1392 wadeable streams across the coterminous US sampled after year 2000 from EPA's Wadeable Stream Assessment (WSA) database and used the 25th and 75th percentile approach to estimate nutrient criteria. TN ($r > 0.95$) and TP ($r > 0.90$) percentiles estimated by the EPA and the WSA data were highly correlated. However, criteria estimated using the 75th percentile WSA TN data were higher than those estimated from EPA and WSA 25th percentiles, which were similar in many ecoregions (Table 1-1). Criteria estimated using 75th percentile WSA TP data were higher than those estimated using the 25th percentile approach and the 25th percentile criteria using WSA data were higher than those for EPA (Table 2). Therefore, this study provided further evidence that 25th percentiles of general populations were not good approximates of 75th percentile estimates of reference streams. Approximately 39 and 47% of national stream length exceeded TP and TN criteria when they were based on WSA 75th percentiles.

A more focused data analysis of streams in the Western Coastal Plains was conducted to 1) examine how the choice of data screening method to determine least-disturbed reference sites altered criteria

results, 2) determine the amount of variation among level III ecoregions, and 3) determine the factors related to undisturbed nutrient concentrations in the Willamette and Central Valley, Western Forested Plains, and Xeric West ecoregions (Sifneos and Herlihy 2008). TP and TN reference percentiles ranged from 0.011-0.030 mg/L and 0.088-0.480 mg/L depending upon the screening method, respectively. The most conservative TP and TN 75th percentile estimates were associated with orthophotograph screening (see method description in Stoddard et al. 2005). The least conservative TP and TN 75th percentile estimates were associated with Oregon Department of Environmental Quality screening with the highest human disturbance index (HDI) classification A that classifies sites as ideal reference sites. Reference-site TN ($p < 0.0001$) and TP concentrations ($p < 0.0001$) varied significantly among level III ecoregions suggesting that the aggregate nutrient ecoregion is too coarse a spatial scale for criteria development. Regression tree models included runoff, elevation, acid neutralizing capacity, forest composition, substrate size, and Omernik level III ecoregion explained 46 to 48% of the variance in undisturbed site nutrient concentrations.

Inconsistencies were found between EPA suggested criteria and those estimated in a study synthesizing published data for over 300 streams in small, primarily forested watersheds (1-1000 ha) across the US (Ice et al. 2003; Binkley et al. 2004). Unfortunately this analysis provided no information on TN or on TP, but dissolved organic N (DON), NO_3-N, dissolved inorganic P (DIP), and dissolved organic P (DOP) were reported. Median NO_3-N concentrations in streams from forested watersheds across the US were 0.15 mg/L, with concentrations greater than 1 mg/L found in Northeastern forests and in Western forests with alder, which is well above any of the EPA proposed TN criteria for forested ecoregions. Fewer studies reported DON compared to nitrate N (68 versus 256, respectively). The DON median was 0.08 mg /L. Coniferous forests generally had a higher median DON (0.7 mg/L) than did hardwood forests (0.1 mg/L). Median DIP concentrations were 0.004 mg/L. Northeastern forested streams had a higher median (0.015 mg/L) than did Southeastern (0.007 mg/L) and Western streams (0.003 mg /L). The median DOP was 0.35 mg/L. However, only 26 studies reported DOP compared to 80 studies reporting DIP. While this study suggests that N concentrations in some primarily forested streams may exceed EPA criteria, comparisons between the two are limited because this study did not represent a sample of randomly selected streams sampled with common protocols and in a balanced manner across forest types, management practices, or regions.

A basin approach to setting nutrient criteria may be more appropriate than an ecoregion approach given the large spatial variation in nutrient concentrations found within some ecoregions. In addition, management at the basin level may be needed to successfully meet criteria goals in higher order lotic ecosystems and large basins can include multiple ecoregions and political boundaries (i.e., states, tribes). Longing and Haggard (2010) examined extensive water quality data from 589 streams and river stations from 1996-2006 in the Red River Basin, USA that includes the Central and Forested Uplands, Great Plains Grass and Shrublands, South Central Cultivated Great Plains, and the Southeastern Temperate Forested Plains. The 25th percentiles for TN were similar to or below EPA suggested criteria in each ecoregion (Table 1-1), but TP 25th percentiles were most often greater than those recommended by the EPA (Table 1-2). The 25th percentile TN (mg/L), TP (mg/L), and suspended chl-a (μg/L)

concentrations were reported for several Texas level III ecoregions including the Southwestern Tablelands (0.61, 0.02, and 1.52 , respectively), High Plains (--, 0.047, and 86.8, respectively), Central Great Plains (0.86, 0.07, and 6.95, respectively), Texas Blackland Prairies (--, 0.015, and 5.98, respectively), Cross Timbers (0.41, 0.04, and 10.0, respectively), East Central Texas Plains (0.64, 0.05, and 5.0, respectively), and the South Central Plains (0.59, 0.07, and 2.7, respectively). In contrast to some other studies, Longing and Haggard (2010) found no statistically significant difference among TN, TP, or chl-a criteria computed within level III ecoregions of the Southeastern Temperate Forested Plains and Hills aggregate ecoregion. Samples sizes were not large enough to test level III ecoregion differences in the other aggregate ecoregions sampled. Therefore, nutrient criteria development in the Red River basin could focus more on aggregate ecoregions than on level III ecoregions.

Overall consensus of many studies suggests that aggregate ecoregions are too coarse for criteria development (Rohm et al. 2002; Smith et al. 2003; Herlihy and Sifneos 2008). However, few criteria at a smaller spatial scale have been reported. Therefore, comparisons between EPA percentiles and literature percentiles are most abundant at the aggregate ecoregion level. We examined the ratio of the literature-estimated criteria to the EPA-suggested criteria for TN and TP in each ecoregion (Figures 1-1 and 1-2). Literature TN criteria for the Xeric West (III) aggregate ecoregion were generally lower than EPA-suggested criteria. Literature TN criteria for the Great Plains Grass and Shrublands (IV) and the Southeastern Temperate Forested Plains and Hills (IX) were highly variable and were up to approximately 2.5 times higher than EPA-suggested criteria (Figure 1-1). Since Longing and Haggard (2010) did not find a statistically significant amount of variation associated with level III ecoregions in the Southeastern Temperate Forested Plains and Hills, future studies might focus on more appropriate factors describing variability in this Texas aggregate ecoregion. Literature TP criteria were more variable for the Great Plains Grass and Shrublands aggregate ecoregion than any other excluding the Eastern Coastal Plain (Figure 1-2). Very little TN and TP criteria data are available for the Texas-Louisiana Coastal and Mississippi Alluvial Plain compared to other Texas aggregate ecoregions.

Percentile analysis of chl-a data has been used as the basis for establishing stream trophic state (Dodds et al. 1998). An analysis of published data for temperate stream sites proposed using the lower and upper third of sestonic or benthic chl-a distributions to establish the boundary between oligotrophic-mesotrophic systems and between mesotrophic-eutrophic systems (Dodds et al. 1998). Oligotrophic-mesotrophic boundaries fell at 20 mg/m^2, 60 mg/m^2, and 10 μg/L for mean benthic chl-a, maximum benthic chl-a , and suspended chl-a, respectively. Mesotrophic-eutrophic boundaries fell at 70 mg/m^2, 200 mg/m^2, and 30 μg/L for mean benthic chl-a, maximum benthic chl-a, and suspended chl-a, respectively. Less percentile data on suspended chl-a were available compared to TN and TP data in peer-reviewed literature. Of the studies examining suspended chl-a, most have estimated criteria at a higher concentration of chl-a than the EPA criteria (Table 1-3). Criteria estimates from 25th percentile of general stream populations for the South Central Cultivated Great Plains and the Southeastern Temperate Forested Plains and Hills were 2-4 times higher than EPA estimates. Upper and lower quantile distributions of turbidity data were also rare (Table 1-3). Most studies have found slightly

higher turbidity criteria than EPA suggested criteria. One estimate existed for the Southeastern Temperate Forested Plains and Hills, and it was lower than the EPA suggested criteria (Palmstrom 2005).

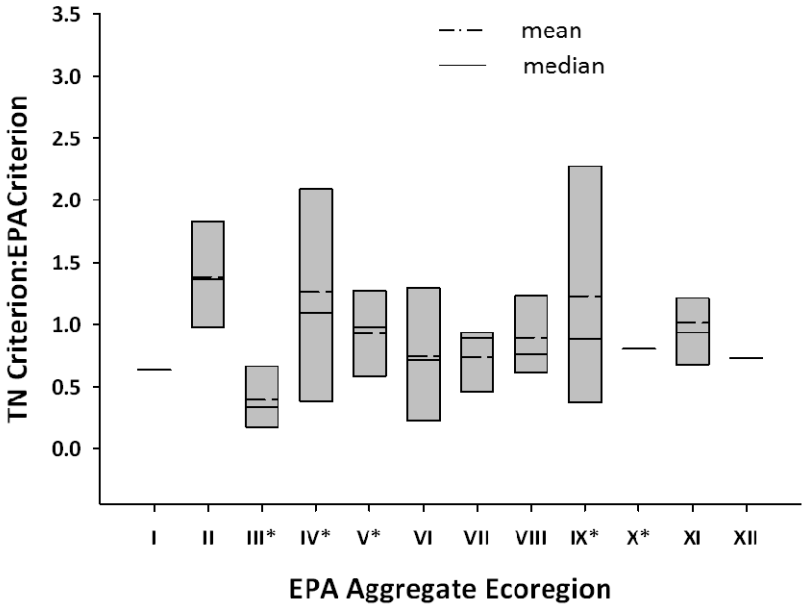


Figure 1-1. Boxplots of the ratio of the literature-estimated total nitrogen (TN) Criterion to the EPA-suggested TN criterion for each aggregate ecoregion. *Texas Aggregate Ecoregion

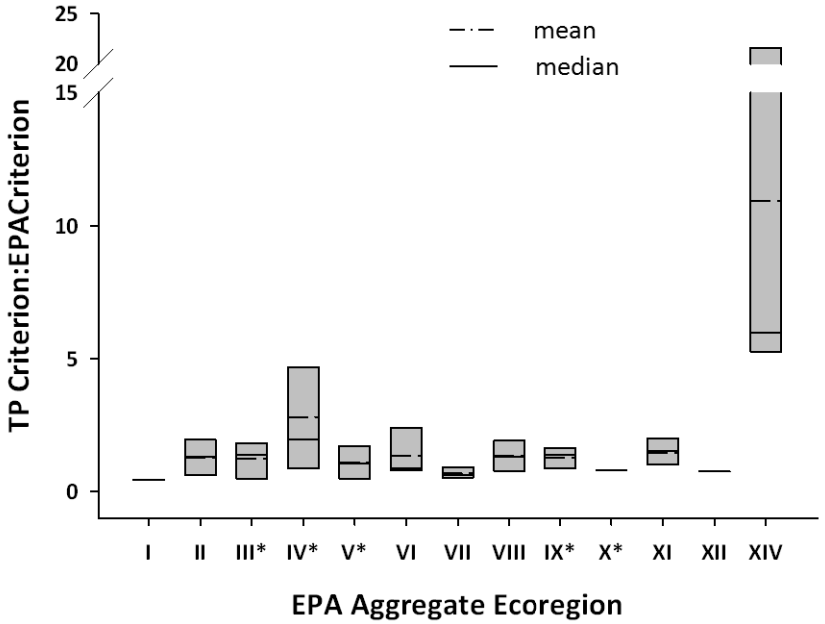


Figure 1-2. Boxplots of the ratio of the literature-estimated total phosphorus (TP) thresholds to the EPA-suggested TP Criterion for each aggregate ecoregion. *Texas Aggregate Ecoregion

Table 1-3. Suspended chlorophyll-a (µg/L) and turbidity (NTU or FTU) 25th and 75th percentiles across US Environmental Protection Agency (EPA) Aggregate Ecoregions from peer-reviewed literature and compared to EPA suggested criteria.

	Aggregate EPA Nutrient Ecoregions													
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
Chlorophyll a (µg/L)														
EPA 2000	1.80	1.08	1.78	2.40	3.00	2.70	1.50	0.63	0.93	2.10	1.61	0.40		3.75
Longing and Haggard 2010	--	--	--	1.52	6.78	--	--	--	3.76	--	0.75	--	--	--
Palmstrom 2005	--	--	--	--	--	--	--	--	3.47	--	4.35	--	--	4.00
Tran and Smith 2010	--	--	--	--	--	--	2.30	--	--	--	--	--	--	--
Turbidity (NTU or FTU)														
EPA 2000	4.25*	1.30	2.34*	4.21*	7.83*	6.36*	1.70	1.30*	5.70*	17.50*	2.30	1.90	--	3.04*
Palmstrom 2005	--	--	--	--	--	--	1.70	1.40	4.00	--	1.60	--	--	4.50
Tran and Smith 2010	--	--	--	--	--	--	2.70	--	--	--		--	--	--

* Value represents a mean of medians for the Northern Rockies, Middle Rockies, and Canadian Rockies; -- = no data

PREDICTIVE RELATIONSHIPS BETWEEN NUTRIENTS AND ECOLOGICAL CONDITIONS

Threshold Analyses

Predictive relationships between criteria variables and response variables can be linear or non-linear. Linear analyses are well described relative to non-linear responses even though many responses to eutrophication are non-linear (Dodds et al. 2010). A system can often respond rapidly with a relatively small change in a criteria variable (e.g., TN, TP, chl-a, turbidity) and the challenge becomes identifying the point or threshold where that rapid change occurs in an objective manner. Statistical threshold analyses have received considerable attention in water quality criteria literature since the EPA first requested criteria development. Much of the current literature cited in this review used these threshold approaches, but we will not do an exhaustive review of the benefits and limitations of each approach because Dodds et al. (2010) has already reviewed many of these methods. Methods included in the review were breakpoint or piecewise regression, cumulative frequency distributions (Paul and McDonald 2005; Utz et al. 2009; Hilderbrand et al. 2010), nonlinear curve fitting, nonparametric changepoint analysis (nCPA; Qian et al. 2003; King and Richardson 2003), quantile regression (Chaudhuri and Loh 2002; Cade and Noon 2003), recursive partitioning or regression tree (Breiman et al. 1984; De'ath and Fabricus 2000, De'ath 2002), regime shift detection (Rodinav 2004; Gal and Anderson 2010; Sonderegger et al. 2009), significant zero crossings (SiZer; Sonderegger et al. 2009), threshold indicator taxa analysis (TITAN; Baker and King 2010), and two-dimensional Kolmogorov-Smirnov test (2DKS; Garvey et al. 1998). Dodds et al. (2010) also examined thresholds in macroinvertebrate richness across stream TP concentrations using breakpoint regression, cumulative frequency, quantile regression tree, nCPA, 2DKS, regime shift, and SiZer to assess variability associated with statistical methodology. They found that threshold estimates varied 3-fold depending upon the type of analyses. Breakpoint regression, 2DKS, and SiZer yielded the highest (least conservative) threshold concentrations. The

authors suggested that a more conservative approach to estimating thresholds would be most prudent given that it is difficult to predict whether a system can return to its prior condition once a threshold has been crossed (Dodds et al. 2010).

Algal Community Responses

Suspended Algae

Relationships between nutrients and phytoplankton biomass in lentic systems have been useful for managing lake and reservoir water quality (Peters 1986) and relationships between suspended chl-a exist in lotic ecosystems as well. Van Nieuwenhuyse and Jones (1996) reported a significant positive curvilinear relationship between river suspended chl-a and TP data (R^2 = 0.67; n = 292) compiled from the literature. Catchment area was also positively related to chl-a and a regression model including both TP and catchment area explained 18% more variance than TP alone. The relationship between catchment area and suspended chl-a suggests that physical factors relating to catchment area, such as hydraulic flushing rate (Kilkus et al. 1975; Soballe and Kimmel 1987), may also regulate suspended chl-a.

Relationships between suspended chl-a and nutrients have also been found at a smaller spatial scale. Lohman and Jones (1999) examined factors explaining suspended chl-a at 23 sites on 13 Missouri Ozark streams and found a curvilinear relationship between TN (range = 0.220-8.435 mg/L; R^2 = 0.70) and TP (range = 0.006-3.093 mg/L; R^2 = 0.78). Catchment area explained an extra 14 and 12% of variance in TN- and TP-chl-a relationships. When the relationship between nutrients and suspended chl-a was limited to 17 sites without known point source pollution, relationships became linear, likely because the range of nutrient concentrations observed at these sites occurred in the linear range of TN (range = 0.172-0.765 mg/L) and TP (range = 0.006-0.119 mg/L) of the curvilinear relationship. Percent row crop and forested land use also were good predictors of suspended chl-a at sites without known point source pollution providing some support for a causal link between non-point sources and suspended chl-a. Lohman and Jones (1999) also examined whether time after a catastrophic flooding event altered the relationship between explanatory factors and suspended chl-a. They found that TP, TN, and catchment area remained the main factors explaining variation in suspended chl-a 0, 14, 28, and 42 days after a catastrophic flood (55-74% of variance explained), but models explained less variance than those based on long-term averages.

Relationships between suspended chl-a and nutrients may also depend on light availability. Royer et al. (2008) examined suspended chl-a collected in statewide surveys of >100 Illinois stream and river sites with 75% of the sites having a TP concentration $\geq$0.112 mg/L (range = 0.007-2.8 mg/L) and TN $\geq$1.0 mg/L (range = 0.21-18.7 mg/L) at base-flow discharge. Watershed area was the best predictor of suspended chl-a as has been found in other studies (R^2 range = 0.20-0.51; Van Nieuwenhuyse and Jones 1996; Lohman and Jones 1999). However, no relationships were found between nutrient concentrations and seston chl-a at high or at base flow. Shading may have played a role in the absence of nutrient relationships. A statistically significant correlation was found between TP and suspended chl-a in the

low flow data when sites with canopy cover ≤25% and TP ≤0.2 mg/L were examined (r = 0.62). A threshold was visually identified at approximately 0.07 mg/L.

Management options were analyzed for the River Ouse in northeastern England that drains a largely unpolluted network (Hutchins et al. 2010). A daily river quality model (QUESTOR) predicted a shift in the river from oligotrophic/mesotrophic conditions to mesotrophic/eutrophic in approximately 70 years due to climate change predictions. In addition, the model suggested that reducing nutrient pollution enough to suppress phytoplankton growth was a more costly option than establishing riparian shading. However, the model did not incorporate possible downstream costs and much literature exists linking nutrient concentrations in river networks to nutrient concentrations, algal blooms, and hypoxia in coastal waters (Dodds 2006; Alexander et al. 2008).

Water residence time may also be an important factor controlling suspended sediments on rivers with impoundments. PhosFate, a GIS based catchment model, was used to examine factors contributing to eutrophication of the Zala River catchment in western Hungary (Honti et al. 2010). Small reservoirs and impoundments were identified as one of the key factors, which also included P concentrations, determining algal biomass presumably because they increase water residence time (WRT). However, given the social importance of these impoundments, it was a more realistic management strategy to minimize WRT, incorporate more riparian buffers, and employ best agricultural practices in the watershed. Water residence time can also be an important factor controlling suspended chl-a in the Kalamazoo River basin in Michigan, USA (Reid and Hamilton 2007).

Benthic algae

Several early studies have found linear relationships between nutrients and benthic algal biomass (Biggs and Close 1989; Lohman et al. 1992; Biggs 1995; Biggs et al. 1998b; Dodds et al. 1997; Biggs et al. 1999; Chétalat et al. 1999) and were reviewed in Biggs (2000). Biggs (2000) examined nutrient and benthic algal relationships from 30 sites in New Zealand streams and rivers and found that nutrient concentrations explained 12-22.6% and 29.5-32.5% of the variation in mean and maximum benthic chl-a, respectively. Days of accrual explained more variation in mean and maximum benthic chl-a (39.7 and 61.8%, respectively) than did nutrient concentrations. Multiple regression models incorporating both variables explained greater than 40% and 70% of the variation in mean and maximum benthic chl-a highlighting the need to incorporate hydrologic disturbances to increase the predictive ability of models.

Dodds et al. (2002; 2006) compiled temperate stream periphyton biomass and nutrient concentration data from the literature (n = 300) and from a subset of 620 National Water Quality Monitoring Network (NAWQA) sites to determine if water column nutrients and non-nutrient factors (e.g., temperature, latitude, land use, substrate type) were linked to periphyton biomass. The greatest amount of variance in mean and maximum periphyton chl-a was always explained by TN or TP (~40%; positive relationships). Stream gradient and latitude were negatively correlated and temperature and substrate type (dummy variable where 0=artificial substrate and 1=natural substrate) were positively correlated with mean and maximum periphyton chl-a, but these variables explained less variation. Ecoregion and land use also

described significant variation in periphyton chl-a in NAWQA data. The large amount of variation in the benthic chl-a data may be due to hydrological disturbances (Lohman et al. 1992; Biggs 1995; Biggs 2000), differences in grazing pressure (Stevenson et al. 2006), or variation in light availability and current velocity (Stevenson et al. 1996) . Therefore, relationships describing a greater amount of variation in algal biomass may be more attainable at a smaller spatial scale where these variables can be carefully measured. Threshold criteria derived by breakpoint regression were generally higher than those estimated by 2DKS analyses (Table 1-4). Both of these methods were found to be less conservative than other statistical threshold methods available (Dodds et al. 2010).

Benthic chl-a has also been found to be positively correlated with TN or TP at a smaller spatial scale (Busse et al. 2006; Stevenson et al. 2006; Stevenson et al. 2008). Busse et al. (2006) measured algal cover, algal biomass as chl-a, and physical and chemical variables at 14 sites with a range of land use types in a southern California watershed. They found that benthic chl-a was positively correlated ($r \geq 0.75$) with percent urbanization and the proportion of upstream land covered by impervious surfaces (range = 2-55%) as well as to TN (range = 0.395-4.490 mg/L) and TP (range = 0.037-0.398 mg/L). However, TP explained more variation (r = 0.86-0.88) in algal measures than did TN (r = 0.75-0.82). Correlations between urban measures and benthic chl-a were greater when land use in a 500 m cone upstream of the sampling point was delineated than when a 100 m cone or the whole catchment was delineated. Total P explained the most variation in total and benthic chl-a in multiple regression models including physical and chemical variables. A quadratic relationship between TP and benthic chl-a explained more variation than a linear one suggesting saturating conditions were reached between 100 - 200 μg TP/L. Percent full sun was also positively related to total, sestonic, and benthic chl-a in some seasons. Current speed was negatively related to sestonic algae and positively related to benthic algae. Algal responses to increasing TP suggested P-limitation. However, responses of chl-a on nutrient diffusing substrata often indicated N-limitation or no positive response to nutrients. In addition, water TN:TP suggested co-limitation or N-limitation and that these ratios were not predictive of results on nutrient diffusing substrata or relationships between nutrients and benthic chl-a.

Relationships between nutrients and benthic algae were also examined from 104 northwestern Kentucky and Michigan streams for a 2 month period (Stevenson et al. 2006). Diatom biomass indicators were higher in Kentucky streams where hydrology constrains invertebrate grazer biomass than in Michigan streams where hydrologic stability allows higher invertebrate grazer biomass to accrue. However, diatom indicators were not related to nutrient concentrations. Positive correlations were found between nutrients and benthic chl-a and percent area of substratum covered by *Cladophora* in both regions, but *Cladophora* responded more positively to nutrients in Michigan than in Kentucky. Total N and TP explained similar amounts of variance in benthic chl-a (r = 0.30-0.43) and *Cladophora* cover (r = 0.30-0.67) within each region. They found most benthic algal responses between 0.01 and 0.03 mg/L (max ~1.0 mg/L) and between 0.40 and 1.0 mg/L (max ~6.0 mg/L) of TP and TN, respectively. Mean periphyton N:P was 11 and 9 and water column N:P was 103 and 84 in Kentucky and Michigan streams, respectively, and were poorly related to algal biomass.

Table 1-4. Benthic algal total nitrogen (TN) and total phosphorus (TP) thresholds determined by various statistical analyses.

Dependent Variable	Criterion Estimation Method[a]	TN Estimated Criterion (mg/L)	TP Estimated Criterion (mg/L)	Citation
Mean chlorophyll a	Regression	0.537	0.043	Dodds et al. 2002;2006
Maximum chlorophyll a	Regression	0.602	0.062	Dodds et al. 2002;2006
Mean chlorophyll a	2DKS	0.515	0.027	Dodds et al. 2002;2006
Maximum chl-a	2DKS	0.367	0.027	Dodds et al. 2002;2006
Mean chlorophyll a	nCPA	NA	0.013	Stevenson et al. 2008
Mean Ash-Free Dry Mass	nCPA	NA	0.008	Stevenson et al. 2008
Acid phosphotase activity	nCPA	NA	0.007	Stevenson et al. 2008
Alkaline phosphotase activity	nCPA	NA	0.007	Stevenson et al. 2008
Number of diatom taxa	nCPA	NA	0.012	Stevenson et al. 2008
Diatom eveness	nCPA	NA	0.020	Stevenson et al. 2008
Proportion of native diatom taxa	nCPA	NA	0.012	Stevenson et al. 2008
Porportion of low-P native taxa	nCPA	NA	0.019	Stevenson et al. 2008
Diatom species similarity to reference	nCPA	NA	0.027	Stevenson et al. 2008
Percent low-P diatom individuals	nCPA	NA	0.019	Stevenson et al. 2008
% high-P diatom individuals	nCPA	NA	0.012	Stevenson et al. 2008
Mean chl-a	nCPA	0.435*	0.038	Miltner 2010
Fine -grained depositional substrate				
% abundance of pollution tolerant diatoms	Regression	0.860	0.280	Black et al. 2011
Alkalophilus diatom richness	Regression	NS	0.050	Black et al. 2011
% abundance of pollution senstive diatoms	Regression	NS	0.090	Black et al. 2011
% abundance high TN diatoms	Regression	0.610	0.060	Black et al. 2011
% abundance high TP diatoms	Regression	0.710	0.060	Black et al. 2011
% abundance N heterotrophs	Regression	1.500	0.100	Black et al. 2011
% abundance motile algae	Regression	0.270	0.060	Black et al. 2011
% richness motile algae	Regression	1.490	0.090	Black et al. 2011
Coarse-grained substrate (rock or wood)				
Alkalophilus diatom richness	Regression	1.250	0.030	Black et al. 2011
% abundance high TN diatoms	Regression	1.450	0.070	Black et al. 2011
% abundance high TP diatoms	Regression	1.300	0.080	Black et al. 2011
% abundance N heterotrophs	Regression	0.590	0.130	Black et al. 2011
% abundance motile algae	Regression	NS	0.200	Black et al. 2011
% richness motile algae	Regression	1.790	0.070	Black et al. 2011

*Dissolved Inorganic Nitrogen; [a] Regression = Breakpoint or Piecewise, nCPA = nonparametric changepoint analysis, 2DKS = 2-dimensional Kolmogrov Smirnov Test

Another way to examine the link between nutrients and algae is to determine nutrients limiting algal growth. Stevenson et al. (2008) measured the potential for P-limitation across a gradient of TP in Mid-Atlantic Highland streams (n=607). Acid and alkaline phosphatase production was measured along with periphyton chl-a, AFDM, and diatom taxonomic composition. Acid and alkaline phosphatase activity was negatively related to TP suggesting the increasing concentrations of TP relieved P-limitation. Chlorophyll *a* and AFDM were positively related to TP concentrations. Thresholds calculated by nCPA occurred between 0.01-0.02 mg/L (Table 1-4) and corresponded nicely with the 75th percentile concentration of reference sites, which was 0.012 mg/L in this region.

Many studies have found that algal community composition can also provide a strong relationship with nutrient concentrations (Stevenson et al. 2008; Porter et al. 2008; Justus et al. 2010; Black et al. 2011; but see Stevenson 2006). Stevenson et al. (2008) found that the number of diatom taxa, evenness, proportion of expected native taxa, and the number of high P taxa were positively related to TP. TP

generally explained more variation (R^2 = 0.04-0.33) for these community composition variables than for algal standing stock measures such as chl-a (R^2 = 0.07) and AFDM (R^2 = 0.07). Porter et al. (2008) analyzed benthic algal community metrics from 976 streams and rivers collected from 1993-2001 through the NAWQA Program and found that algal species indicators for trophic condition, organic enrichment, salinity, motility, and taxa richness had significant positive correlations with TN and TP concentrations (range r = 0.31-0.57). The proportional abundance of N-fixing diatoms decreased as TN concentrations increased (r = -0.30). In addition, diatom species associated with high dissolved oxygen (DO) were often negatively correlated with TN and TP r ranged from -0.34 to -0.40. Justus et al. (2010) also found positive correlations between a nutrient index that combined TN and TP and algal indices including the relative abundance of most tolerant diatoms (r = 0.80), the combined relative abundance of three *Cymbella* spp (r = -0.71), mesosaprobic algae percent taxonomic richness (r = 0.65), and the relative abundance of obligate N heterotrophic diatoms (r = 0.57). Black et al. (2011) examined algal communities on coarse- (i.e., rock and wood > 64mm) and fine-grained substrate from 73 stream sites sampled as part of USGS NAWQA. Sites from two agricultural regions, Washington (n = 23) and Nebraska (n = 23), and sites from across the western USA that were characterized by <10% agricultural or urban land use were used in the analysis. Piecewise regression identified thresholds in several diatom indices (Table 1-4). Thresholds for fine and coarse-grained substrate were reported, but they were not statistically different suggesting that taxonomic information from both substrate types may not be needed to establish algal thresholds (Black et al. 2011).

Relationships between algal community composition and nutrient concentrations may not be as strong in larger rivers. Snyder et al. (2002) examined diatom community composition data from 12 large rivers (> 5th order) in Idaho, USA across a DIN (range = 0.002-1.36 mg/L) and TP (range = 0.007-0.100 mg/L) gradient and found no community level response to increasing nutrients. Instead, principle component analysis groups were mainly determined by drainage basin. Few studies have examined relationships between benthic algae and nutrients in non-wadeable streams, which should be a target of future research.

Another area of future research would include linking nutrient concentrations with designated uses. Some early studies attempted to describe levels of benthic algae that may constitute nuisance levels that may impair a designated use (Horner et al. 1983; Welch et al. 1988), but algal levels that constitute a nuisance to the general public are not well established. Suplee et al. (2009) used an On-River survey, which consisted of 44 trips on rivers to interview recreators, and a By-Mail survey that was sent to 2000 randomly selected voting individuals. The surveys consisted of 8 randomly ordered photographs with varying levels of benthic algae (< 50-1276 mg/m^2) that were shown to responders. The responders had to rate each photo as desirable or undesirable. As benthic chl-a levels increased, the percent of desirable responses decreased. The overall average percentage of people rating a photograph as desirable was less than 50% for all photographs with chl-a levels $\geq$200 mg/m^2 in On-River and By-Mail surveys. Photographs showing chl-a levels <150 mg/m^2 were rated as desirable on average greater than 50% of the time. This finding supports previous work suggesting 150 mg/m^2 represents a nuisance threshold of chl-a (Horner et al. 1983; Welch et al. 1988).

Macroinvertebrate and Fish Community Responses

There is a rich history of linking macroinvertebrate and fish multimetric biological indices to water quality (Hilsenhoff 1977; Karr 1981; Washington 1984) and this method remains important in more current studies examining water quality impacts on macroinvertebrates and fishes (e.g., Waite and Carpentar 2000; Smith et al. 2007; Justus et al. 2010). After the US EPA's call for the development of scientifically defensible nutrient criteria, the focus in bioassessment shifted from not only determining biological traits and species that were associated with changes in water quality, but also to developing methods that can attribute causation in large spatial scale observational studies (King and Richardson 2003; Wang et al. 2006; DeZwart et al. 2007; Yuan 2010) and to defining particular nutrient concentrations or thresholds where traits or species shift (reviewed Dodds et al. 2010). One of the earliest studies using this approach used observational and experimental data to examine macroinvertebrate responses to changes in TP concentrations in wetlands (King and Richardson 2003). A nonparametric changepoint analysis (nCPA) was used to examine threshold shifts in experimental and in observational data allowing more effective estimates of impairment or risk associated TP than by examining observational data alone. This analysis was also important because it described a way to attach confidence levels to thresholds (King and Richardson 2003; Smith et al. 2005)

Wang et al. (2006) used regression tree analysis, which is similar to nCPA, and a 2DKS analysis to examine macroinvertebrate and fish water quality thresholds from 240 wadeable Wisconsin streams with seasonal median TP and TN ranging from 0.012-1.641 mg/L and 0.131-21.260 mg/L, respectively. They found that 66 and 69% of the fish and macroinvertebrate measures were correlated with at least one nutrient measure. Percentages and individuals of Ephemeroptera, Trichoptera, and Plecoptera (EPT), the Hilsenhoff biotic index, and richness were the macroinvertebrate measures with the greatest correlations with nutrient measures. Threshold estimates ranged from 0.61 to 1.68 mg/L and 0.04 to 0.09 mg/L of TN and TP, respectively (Table 1-5). Percentages of carnivorous, intolerant, and omnivorous fishes, the index of biotic integrity, and salmonid abundance were fish measures with the greatest correlations with nutrient measures and threshold estimates ranged from 0.54 to 1.83 mg/L and 0.06 to 0.09 mg/L of TN and TP, respectively (Table 1-6). Mean ($\pm$ stdev) macroinvertebrate TN thresholds estimated by regression tree analysis (1.248 ± 0.33 mg/L) tended to be higher than those estimated by 2DKS (0.888 ± 0.213). However, macroinvertebrate TP thresholds were similar for both methods (Regression tree = 0.075 $\pm$ 0.024; 2DKS = 0.078 $\pm$ 0.025). Mean ($\pm$ stdev) fish TN and TP thresholds estimated by regression tree analysis (1.26 $\pm$ 0.494 and 0.078 $\pm$ 0.015 mg/L, respectively) tended to be higher than 2DKS estimates (0.558 $\pm$ 0.035 and 0.063 $\pm$ 0.005 mg/L, respectively). This tendency for 2DKS to yield less conservative estimates was also found by Dodds et al. (2010). Redundancy analysis indicated that nutrients explained 22 and 15% of the variance in macroinvertebrate and fish assemblages. Catchment and instream habitat explained the most variation in macro-invertebrate (42%) and fish (46%) assemblages.

Table 1-5. Benthic macroinvertebrate total nitrogen (TN) and total phosphorus (TP) criteria determined by various statistical analyses.

Dependent Variable	Criterion Estimation Method[a]	TN Estimated Criterion (mg/L)	TP Estimated Criterion (mg/L)	Citation
Percentage of EPT individuals	Regression Tree	1.680	0.080	Wang et al. 2006
Percentage of EPT taxa	Regression Tree	1.30	0.090	Wang et al. 2006
Hilsenhoff Biotic Index	Regression Tree	1.140	0.090	Wang et al. 2006
Taxa richness	Regression Tree	0.870	0.040	Wang et al. 2006
Percentage of EPT individuals	2DKS	0.980	0.090	Wang et al. 2006
Percentage of EPT taxa	2DKS	1.110	0.090	Wang et al. 2006
Hilsenhoff Biotic Index	2DKS	0.610	0.090	Wang et al. 2006
Taxa richness	2DKS	0.850	0.040	Wang et al. 2006
Taxa richness	Regression Tree	1.925	0.150	Weigel and Robertson 2007
Mean pollution tolerance value	Regression Tree	0.634	0.064	Weigel and Robertson 2007
Taxa richness	nCPA	1.040	0.050	Evans-White et al. 2009
Primary consumer richness	nCPA	1.140	0.050	Evans-White et al. 2009
Gathering consumer richness	nCPA	0.930	0.060	Evans-White et al. 2009
Scraping consumer richness	nCPA	NS	0.050	Evans-White et al. 2009
Shredding consumer richness	nCPA	NS	0.050	Evans-White et al. 2009

Table 1-6. Stream fish total nitrogen (TN) and total phosphorus (TP) thresholds determined by various threshold analyses.

Dependent Variable	Criterion Estimation Method[a]	TN Estimated Criterion (mg/L)	TP Estimated Criterion (mg/L)	Citation
Percentage of carnivorous individuals	Regression Tree	1.220	0.090	Wang et al. 2006
Index of Biotic Integrity	Regression Tree	1.360	0.070	Wang et al. 2006
Salmonid individuals	Regression Tree	0.630	0.060	Wang et al. 2006
Percentage of intolerant individuals	Regression Tree	1.830	0.090	Wang et al. 2006
Percentage of carnivorous individuals	2DKS	0.540	0.060	Wang et al. 2006
Index of Biotic Integrity	2DKS	0.540	0.060	Wang et al. 2006
Salmonid individuals	2DKS	0.610	0.060	Wang et al. 2006
Percentage of intolerant individuals	2DKS	0.540	0.070	Wang et al. 2006
Index of Biotic Integrity	Regression Tree	0.634	0.139	Weigel and Robertson 2007
Percent Biomass of Round Suckers	Regression Tree	0.634	0.091	Weigel and Robertson 2007

Weigel and Robertson (2007) used a regression tree analysis for macroinverebrate and fish assemblages sampled from 41 sites on 34 nonwadeable Wisconsin rivers with TN ranging from 0.415-5.485 mg/L and TP ranging from 0.023-0.497 mg/L. Macroinvertebrate taxa richness and the mean pollution tolerance value were most consistently and highly correlated with nutrient variables. For fishes, the index of biotic integrity and the percentage of fish biomass composed of round suckers (*Cycleptus* spp., *Hypentelium* spp., *Minytrema* spp., and *Moxostoma* spp.) were the most highly correlated with nutrient variables. Mean thresholds of all metrics occurred at approximately 0.957 mg/L and 0.111 mg/L of TN and TP, respectively (Tables 1-5 and 1-6). Nutrients, suspended chl-a, water clarity, and land cover (forest or row-crop agriculture) explained 61% of the variation in macroinvertebrate variables, but they were correlated with each other to such an extent that redundancy analysis could not attribute variation to individual factors. The same variables explained 44% of the variation in fish assemblages with nutrients and other water chemistry variables explaining 25% and 13% of the variation, respectively.

Aquatic biodiversity can often have negative threshold relationships with water-quality variables at large spatial scales (Wang et al. 2006; Weigel and Robertson 2007), but the specific mechanism(s) driving these threshold relationships are not well established. Evans-White et al. (2000) hypothesized that

resource quality [(i.e., carbon:phosphorus (C:P)] might partly drive macroinvertebrate primary consumer (grazer and detritivore) richness thresholds by altering growth or competitive interactions among species with differing resource demands as has been found in some manipulative P dosing studies (Cross et al. 2006, 2007, Singer and Battin 2007). Evans-White et al. (2009) estimated TN (Table 1-4), TP (Table 1-5), and turbidity thresholds for macroinvertebrate richness across trophic levels and feeding groups in Central Plains (USA) streams. They also determined if mean taxon body C:P of feeding groups with diversity losses were negatively related to TP, a pattern that would suggest that communities from streams with high nutrient concentrations were dominated by a few species with high dietary P demands. More than 75% of the sampling events came from 4 level III ecoregions including the Central Great Plains, Central Irregular Plains, Flint Hills, and the Ozark Highlands. Primary consumers were more sensitive to TN and TP (threshold mean = 1.0 mg/L and 0.06 mg/L, respectively) than secondary consumers (TP threshold mean = 0.09 mg/L), a result supporting the resource quality hypothesis because predators are usually not limited by N or P. Turbidity reduced richness regardless of feeding mode [threshold mean = 4.7 NTU; range = 1.79 (scraper richness)-10.75 (collector-filterer richness)], a result suggesting that turbidity and nutrient macroinvertebrate thresholds were caused by different factors. The TP-richness threshold could be caused partially by changes in food quality because the mean body C:P of shredding and collector-gathering taxa declined as TP increased (threshold mean = 0.07 and 0.75 mg/L, respectively). Mean scraper C:P was not related to TP and, other factors may cause declines in scraper richness (Yuan 2010). Results from Evans-White et al. (2009) support the hypothesis that changes in detrital resource quality can contribute to large-scale losses in biodiversity in nutrient-enriched streams. Within macroinvertebrate detritivorous feeding groups, P-rich detritus might allow faster growing taxa with higher body P demands to out-compete slower growing taxa adapted to lower quality detritus. Thresholds in this study may be the most applicable to Texas because a majority of samples came from the Central Great Plains ecoregion. However, TP thresholds were higher than those observed in some applicable Texas studies (King and Winemiller 2009; King et al. 2009).

Threshold type analyses have also been used to make trophic level categories. Smith et al. (2007) examined nutrients and macroinvertebrates in 129 locations from 116 streams sampled across New York USA. They established a nutrient biotic index (NBI) estimated from species TP and nitrate (NO_3^-) optima that was linearly related to TP ($r = 0.65$) and to NO_3^- concentrations ($r = 0.57$). The trophic state of sample sites was then estimated by using additive tree clusters based on mean pair-wise Bray-Curtis similarities yielding oligotrophic and eutrophic boundaries at $\leq$0.0175 and $\geq$0.065 mg/L and $\leq$0.24 and $\geq$0.98 mg/L of TN and TP, respectively. They suggested using these trophic boundaries as thresholds to establish nutrient criteria.

Biotic assessments would be most useful if they contributed information on the condition or magnitude of alteration of an ecosystem and on the potential causes of impairment. Often scientist and managers are better at the latter than the former. Several previous studies have addressed potential causes via redundancy analysis (Wang et al. 2006; Weigel and Robertson 2007) and by linking experimental data with observational data (King and Richardson 2003; Evans-White et al. 2009). DeZwart et al. (2006) employed another approach that linked fish, habitat, and chemistry data collected from Ohio River

sampling events (n = 1552), assessed the biological condition at each site, and attributed impairment to multiple probable causes. Biological condition was estimated from the proportion of native species predicted to occur at a site that were actually observed. Predicted occurrences were estimated in a similar manner to the River Invertebrate Prediction and Classification System (RIVPACS) that provides species-specific probabilities of capture based on the geographic location and habitat characteristics of each site. Toxic exposure effects were estimated using species sensitivity distributions and toxicity mixture principles and generalized linear regression models described species abundance and habitat relationships. They found that average losses of fish species in Ohio rivers was 40% and water chemistry, toxicity, effluent, and habitat loss were estimated to explain 28, 3, 3, and 16% of losses. Site-specific causes were also shown in pie charts mapped onto river segments that allowed easier communication of the causes of impairment.

Yuan (2010) proposed yet another method that may be used to identify with more confidence causal effects of particular factors in large spatial scale observational data. He used propensity scores, which have frequently been used in epidemiological, sociological, and economic studies, but have not been used in ecology. Propensity functions can summarize multiple covariate contributions as one parameter. Essentially, nutrient concentrations can be related to covariate values with regression and then the predicted concentrations in each stream becomes the propensity score. Next, the scores can be stratified into groups with similar covariate distributions and then causal effects of nutrients on dependent variables can be estimated more confidently. Yaun (2010) uses propensity scores to estimate the effect of increasing TN on benthic macroinvertebrate grazers in small streams of the western US (n = 827 sampling sites). The response of grazer richness to increasing TN varied across groups with similar covariate distributions. In large, wadeable open-canopied streams, grazer richness was negatively related to TN, but in small closed-canopied streams, grazer richness responded positively to increasing TN. Benthic chl-a responded positively to TN in both stream types. Thus, Yuan (2010) proposed that increasing algal biomass stimulated grazer richness in small closed-canopied streams and shifting algal community composition may have caused declines in grazer richness in the larger streams. While this method can increase confidence in a causal relationship, it can only account for covariates that are measured.

Land use factors can also play a role in benthic macroinvertebrate threshold patterns. King et al. (2005) reported that a threshold estimated via nCPA in benthic macroinvertebrate biotic composition occurred with as little as 21% developed land in Coastal plain stream watersheds. Biotic composition changed rapidly between 21 and 32% developed land and nearly a 100% probability of a threshold beyond 32% developed land. King and Baker (2010) also found numerous macroinvertebrate species declined between 0.5%-2% impervious cover in Coastal Plain streams using TITAN. These studies indicate that land use should also be considered when assessing the biological quality of streams, especially stream systems.

APPLICABLE TEXAS STUDIES

Wastewater treatment plants (WWTP) are one of the nutrient sources that have led to elevated stream N and P concentrations in the Brushy Creek watershed (Alan Plumer Associates 1996), in streams located in the Edwards Plateau of Central Texas (Mabe 2007), and in the South Fork of the San Gabriel River (Brazos River Authority 1998). Alan Plumer Associates (1996) assessed the effects of WWTP effluent on benthic algal cover downstream. Their goal was to determine if improving P removal from the effluent of a WWTP managed by the City of Leander would reduce stream algal biomass. They sampled 9 study sites that represented a stream that receives wastewater discharges (n = 4), a stream that receives effluent that has been treated to remove P (n = 3), and a stream receiving no wastewater effluent (n = 2) in the Brushy Creek watershed. Algal cover was estimated with a qualitative scoring system from photographs taken at each site. Algal cover was lower at sites that had no effluent discharge compared to sites that had effluent discharge, although no statistics were reported. The P concentrations in tributaries that received discharges from WWTPs that provided P reduction were not different from those that received WWTP discharge without P reduction. The lack of an algal cover response to wastewater treatment processes may have been due to the photographic method utilized to measure stream algal communities.

Agrilife Research (2010) assessed four methods of periphyton sampling including periphyton scraping followed by chl-a and AFDM measurement in the laboratory (method 1), areal coverage in a bucket (method 2), picking up cobbles in a transect to estimate percent coverage on the cobble (method 3), and percent coverage estimated from photos of a 1 m^2 sampling frame (method 4). Results suggested that periphyton measured using methods 2 and 4 were limited in turbid waters that limit visibility of the benthos. Qualitative measures of periphyton coverage from method 3 did correlate with quantitative measures of chl-a and AFDM, while measures from method 2 only correlated with AFDM. Percent coverage estimated from photographs (method 4) was never correlated with quantitative measures and was not correlated with measures of water quality. Correlation analysis between water quality variables and measures of periphyton biomass usually showed no relationship or a weak relationship (i.e., $r < 0.55$). Several periphyton measures had negative correlations with TN and positive correlations with DO supporting a direct link between water quality, benthic algae, and DO.

Mabe (2007) sampled water quality, algae, benthic macroinvertebrates and fishes from 15 streams located in the Edwards Plateau of Central Texas in summer 2005. Streams were grouped into streams receiving wastewater effluent (WW), streams not receiving wastewater effluent (NWW), and least disturbed (LD) streams. Streams receiving wastewater effluent always had higher nutrient concentrations than NWW and LD streams, which were not statistically different. Benthic chl-a tended to be higher in WW streams than in NWW and LD streams, but it was not statistically different due to a large amount of variance potentially associated with hydrology; however, suspended chl-a was higher in WW streams than in NWW and LD streams. Benthic and suspended chl-a were positively correlated with TN or to NO_3-N and TP was positively related to macroalgal cover and diel DO ranges. In addition,

TP and macroalgal cover were positively related to minimum DO providing further empirical evidence for the link between water quality, algae, and DO concentrations.

Nonpoint nutrient sources such as poultry production and poultry litter application also contribute to reduced water quality in Texas streams. A study examining water quality in several small streams in the Brazos River basin, some associated with agricultural lands, found that several of the streams had exceedances of DO, chl-a, *E.coli*, and fecal coliform criteria (Contreras 2007). Keisling et al. (2006) examined 33 streams across Texas. Thirty percent of the stream study sites showed exceedances of the DO criteria for more than 50% of the observations. In addition, 30% of the sites had IBI scores that exceeded criteria more than 50% of the time. They found that suspended chl-a was negatively correlated to 24 h average (R^2 = -0.22), minimum (R^2 = -0.22), and maximum DO. According to Keising et al. (2006), the decline in DO minimum seemed most pronounced when suspended chl-a concentrations exceeded 10 μg/L. There was also a positive relationship between periphyton chl-a and diel DO maximum (R^2 = 0.12) and 24 h DO range (R^2 = 0.18). Finally, positive linear relationships between and benthic chl-a (R^2= 0.19) and NO_3^- and benthic chl-a (R^2= 0.13) were found. Benthic chl-a was negatively related to sestonic chl-a (R^2= 0.12) across these Texas streams.

A two-year survey was conducted to estimate periphyton and macrophyte diversity and abundance at 11 locations within the North Bosque River watershed (McFarland et al. 2008). Twenty-four macrophyte species were identified and *Justicia americana* (L.) Vahl was the dominant taxa ranked by AFDM. Periphyton chl-a suggested mesotrophic or eutrophic conditions were indicated at 36% of the sites. Suspended chl-a concentrations of phytoplankton indicated mesotrophic conditions at 73% of sites. Given the disparity between benthic and suspended chl-a measurements, it may be best to consider both when trying to estimate trophic status in these streams.

Several studies have used algal growth responses to experimentally added nutrients (Stanley et al. 1990; Dávalos-Lind and Lind 1999; Matlock. and Rodriguez 1999**).** Stanley et al. (1990) evaluated nutrient limitation of periphyton and phytoplankton in the Upper Guadalupe River using clay pots and glass bottles enriched with nutrients. Periphyton on pots enriched with P had significantly higher chl-a in 78% of the trials. Pots enriched with N alone did not result in significantly higher chl-a, but N was found to be secondarily limiting at potentially 22% of sites. Flow-through experiments found that periphyton responses to enrichment were greatest when ambient P was less than 0.010 mg/L. In contrast to periphyton, phytoplankton was either primarily or secondarily limited by N in 75% of the trials and variability in the limiting nutrient in reservoirs seemed to increase with flow variability. Dávalos-Lind and Lind (1999) examined whether the potential for water entering Lake Waco from 4 tributaries differ in their ability to promote algal growth and potentially reservoir eutrophication and whether their growth potential changes seasonally. Algal bioassays of *Selenastrum capricornutum* Printz and native phytoplankton collected from a Lake Waco North Bosque Arm site showed that tributary water algal growth potential (AGP) means ranged from 292-857% and had less growth potential than did reservoir water which ranged from 711-1285%. Of the tributaries, the South Bosque River had the highest AGP at 857%. AGP within sites varied greatly by season with the winter and spring having higher AGP than the

late spring and summer. Reservoir AGPs were less variable seasonally than tributaries. The growth limiting nutrient was P at most sampling sites on most dates. Nitrogen occasionally produced a greater algal response than P in the Lake Waco North Bosque Arm site. In addition, co-limitation of N and P did occur in the autumn months at some of the reservoir sites. The growth response to added P varied seasonally for reservoir sites with a lower algal growth response in the summer compared to in other seasons. A similar pattern was not seen in tributaries because of low summer flow.

Matlock and Rodriguez (1999) assessed nutrient limitation and trophic status of tributaries in the Lake Waco/Bosque River watershed in north-central Texas, where dairy farming and row crop agriculture have been associated with declining water quality. Periphytometers (Matlock et al. 1998) were placed at ten sites in five streams (North Bosque, Middle Bosque, Lower Bosque, Hog Creek, and Neil's Creek) with different nutrient characteristics. The ratio of chl-a in the no nutrients added periphytometer to any nutrient amended periphytometer was named a lotic ecosystem trophic status index (LETSI) and is useful for making comparisons of stream biotic response to nutrients (see Matlock et al. 1999). In July 1997 and May 1998 most sites were either P-limited, co-limited, or not limited by nutrients. In October 1998, one site was N-limited while other sites were not nutrient limited. Therefore, seasonality played a significant role in algal responses to nutrients.

King and Winemiller (2009) collected periphyton chl-a, nutrient ratio, and community composition and suspended chl-a data from Brazos and Trinity River wadeable streams (n=64) located in the Cross Timbers (29), Blackland Prairies (32), and East Central Texas (33) level III ecoregions from June-August 2008. Epilithic periphyton C:P ratios had a negative threshold relationship with TP at ~0.020 mg/L, while periphyton C:P ratios on muddy/sandy substrate did not respond to nutrients. Suspended chl-a had a positive threshold relationship to TN (~0.35 mg/L) and TP (~0.025 mg/L) in ecoregion 29, but was not related in ecoregions 32 and 33 potentially due to a lower sample size and nutrient concentration range. Ordinations suggested that algal metrics should be stratified by ecoregion, but not by basin. Ecoregion 29 algal species ordinations were related to TP, TN, pasture, outfalls, sediment, and chloride but ordinations in ecoregions 32 and 33 were not related to nutrient or nutrient-related variables. Threshold analyses indicated that 31 algal species declined between 0.015 and 0.025 mg/L, while 36 algal species increased between 0.020 and 0.050 mg/L. Most of these taxa also responded negatively to decreasing C:P periphyton ratios, which strengthens the link between changes in surface water TP and algal responses.

King et al. (2009) also did a more in depth study in the Brazos River basin within the Cross Timbers Level III ecoregion and linked field observations to an experimental stream study examining the effects of TP and other water quality parameters on stream dissolved oxygen concentrations and on benthic algae and macroinvertebrate communities. They found that periphyton C:P, N:P, and C:N had a significant negative threshold response to increasing TP with TP concentrations around 0.020 mg/L showing the most consistent threshold decline in periphyton stoichiometry. Dissolved inorganic N (DIN) had less control over periphyton stoichiometry than did TP. The thickness of microbial films growing on rocks also showed threshold declines around 0.02 mg/L. Filamentous green algal cover exhibited 2 distinct

increases around 0.020 and 0.200 mg/L and chl-a:AFDM shifted towards a greater fraction of chl-a at these same concentrations. Hydrology influenced patterns in DO, water temperature, and pH across streams. In 2006, which was a drought year, few of the streams were flowing and DO $\leq$2 mg/L was found in streams that had TP concentrations >0.027 mg/L. Minimum DO during this drought year was significantly predicted by algal variables suggesting that these metrics may be linked quantitatively to aquatic life use standards.

TITAN analysis of algal species data indicated that a TP concentration around 0.02 mg/L was the threshold of greatest overall decline in species (King et al. 2009). Numerous species including *Cladophora* showed threshold increases simultaneously. A few taxa showed chl-a threshold increases between 0.20-0.50 mg/L corresponding to filamentous green algal cover and chl-a: AFDM threshold increases. Macro-invertebrates showed more consistent relationships to TP at sites with no flow (2006) compared to sites with flow (2007), but there were fewer sites sampled with flow than without flow. TITAN thresholds were reported for individual taxa and more taxa showed negative compared to positive responses.

A controlled 28-day P-dosing experiment conducted in 12 artificial streams having TP concentrations of 0.019-0.021 (control), 0.038-0.040 (low), and 0.127-0.137 (high) mg/L based on patterns found in the Brazos River basin (King et al. 2009). They found that periphyton biomass on ceramic tiles was higher in the high P treatment relative to the low P treatment and control on day 14, but by day 28, the low and high P treatments had similar biomass that was higher than controls. In addition, *Cladophora* biomass was significantly higher in the low and high P treatments compared to the control on day 28. Periphyton C:P differed significantly among control (mean ~320), low (mean ~230), and high (mean ~150) streams. Algal species responses to P enrichment mirrored field responses. Control streams increasingly resembled algal communities from the P-enriched field sites over time until 28 days. Five of the 7 algal taxa showing responses to P-enrichment in the experiment also showed responses in the field. Macroinvertebrate taxonomic composition did not differ among treatments in the experiment as they did in the field, which may be attributed to the relatively short dosing period in the experiment relative to the life cycle of macroinvertebrate taxa. The appendix contains an expanded summary of each of these studies conducted in the State of Texas.

SYNTHESIS AND CONCLUSIONS

Two main approaches for nutrient criteria development were assessed in this review. The focus of literature utilizing the percentile approach to nutrient criteria development has been on comparing 25th and 75th percentile approaches, examining factors that cause variability within aggregate ecoregions, and comparing 25th and 75th percentile results obtained with more modern data or with data collected using a more rigorous statistical design to EPA-proposed criteria. Overall consensus of many studies suggests that aggregate ecoregions are too coarse for criteria development (Rohm et al. 2002; Smith et al. 2003; Sifneos and Herlihy 2008), but the majority of comparisons between EPA percentiles and literature percentiles are most abundant at the aggregate ecoregion level. We found that TN criteria for

the Xeric West (III) aggregate ecoregion were generally lower than EPA-suggested criteria. In addition, literature TN criteria for the Great Plains Grass and Shrublands (IV) and the Southeastern Temperate Forested Plains and Hills (IX) were highly variable and were up to approximately 2.5 times higher than EPA-suggested criteria (Figure 1-1). A previous study found no statistical difference among level III ecoregions in the Southeastern Temperate Forested Plains and Hills (Longing and Haggard 2010). Therefore, future studies might focus on more appropriate factors describing variability in this Texas aggregate ecoregion. Literature TP criteria were more variable for the Great Plains Grass and Shrublands aggregate ecoregion than any other excluding the Eastern Coastal Plain (Figure 1-2). Very little TN and TP criteria data are available for the Texas-Louisiana Coastal and Mississippi Alluvial Plain compared to other Texas aggregate ecoregions. Finally, very few studies reported percentile derived criteria for suspended chl-a and turbidity.

Predictive approaches have focused on establishing relationships between water quality and algae, macroinvertebrate, and fish communities, attributing causation, and determining whether threshold points exist that can aid in nutrient criteria development. Many studies have found linear and non-linear relationships between benthic and suspended algae, macroinvertebrate communities, and fishes. Most of the predictive approaches have occurred at the state level and may not be directly comparable to EPA aggregate ecoregions. However, benthic algal criteria threshold estimates ranged from 0.007 to 0.100 mg/L and 0.270 to 1.500 mg/L TP and TN, respectively (Table 1-4), which is within the range of criteria proposed by percentile analysis (Tables 1-1 and 1-2). It also encompasses the P-threshold estimates provided for Texas benthic algal communities by King et al. (2009) and King and Winemiller (2009). Benthic macroinvertebrate and fish TN thresholds ranged from 0.610 to 1.925 mg/L and 0.540 to 1.830 mg/L (Tables 1-5 and 1-6). Published benthic macroinvertebrate TP thresholds ranged from 0.040 to 0.150 mg/L, which is slightly higher than TP thresholds estimated for Texas streams that occur primarily around 0.015-0.020 mg/L (King and Winemiller 2009; King et al. 2009). Published fish TP thresholds ranged from 0.060 to 0.139 and were comparable to benthic macroinvertebrate thresholds. Fish TP thresholds estimated for Texas were generally lower than literature thresholds (~0.015-0.025 mg/L; King and Winemiller 2009). Although not completely comparable, criteria estimated via the percentile method had similar ranges to biological threshold criteria.

REFERENCES

Agrilife Research. 2010. Analysis and summary of data collection methods for nutrients in Texas streams. Contract #582-9-90439-08

Alan Plumer Associates. 1996. City of Leander study of algae cover in the upper Brushy Creek watershed. Report 501-0200.

Alexander RB, Smith RA, Schwarz GE, Boyer EW, Nolan JV, Brakebill JW. 2008. Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River basin. Environmental Science and Technology 42:822-830.

Alonso A, Camargo JA. 2003. Short term toxicity of ammonia, nitrite, and nitrate to the aquatic snail *Potamopyrgus antipodarum* (Hydrobiidae, Mollusca). Bulletin of Environmental Contamination and Toxicology 271:52-58.

Baker ME, King RS. 2010. A new method for detecting and interpreting biodiversity and ecological community thresholds. Methods in Ecology and Evolution 1:25-37.

Bayer CW, Davis JR, Twidwell SR, Kleinsasser R, Linam G, Mayes K, Hornig E. 1992. Texas aquatic ecoregion project: an assessment of least disturbed streams. Draft Report

Biggs BJF. 2000. Eutrophication of streams and rivers: dissolved nutrient chlorophyll relationships for benthic algae. Journal of the North American Benthological Society. 19:17-31.

Biggs BJF. 1995. The contribution of flood disturbance, catchment geology, and land use to the habitat template of periphyton in stream ecosystems. Freshwater Biology 33: 17-31

Biggs BJF, Close ME. 1989. Periphyton biomass dynamics in gravel bed rivers: the relative effects of flow and nutrients. Freshwater Biology 22:209-231.

Biggs BJF, Smith RA, Duncan MJ. 1999. Velocity and sediment disturbance of periphyton in headwater streams: biomass and metabolism. Journal of the North American Benthological Society 18:222-241.

Biggs BJF, Stevenson RJ, Lowe RL. 1998. A habitat matrix conceptual model for stream periphyton. New Zealand Journal Of Marine And Freshwater Research 143:25-56.

Binkley D, Ice GG, Kaye J., Williams CA. 2004. Nitrogen and phosphorus concentrations in forest streams of the United States. Journal of the American Water Resources Association 40:1277-1291.

Black RW, Moran PW, Frankforter JD 2011. Response of algal metrics to nutrients and physical factors and identification of nutrient thresholds in agricultural streams. Environmental Monitoring and Assessment 175:397-417.

Brazos River Authority. Environmental Services. 1998. South Fork of San Gabriel River Nutrient Study. NPDES Permit # WQ0014477001

Breimen L, Freeman JH, Olshen RA, Stone CJ. 1984. Classification and regression trees. Wadsworth International Group, Belmont, California.

Busse LB, Simpson JC, Cooper SD. 2006. Relationships among nutrients, algae, and land use in urbanized southern California streams. Canadian Journal of Fisheries and Aquatic Sciences 63:2621-2638.

Cade BS, Noon BR. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and Environment 1:412-420.

Carleton JN, Park RA, Clough JS. 2009. Ecosystem modeling applied to nutrient criteria development in rivers. Environmental Management 44:485-492.

Carlisle DM, Hawkins CP. 2008. Land use and the structure of western US stream invertebrate assemblages: predictive models and ecological traits. Journal of the North American Benthological Society 27:986-999.

Carmichael WW. 1986. Algal toxins. Advances in Botanical Research Incorporating Advances in Plant Pathology 12:47-101.

Camargo JA, Alonso A. 2006. Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment. Environmental International 32:831–849.

Camargo, J. A., A. Alonso, and A. Salamanca. 2005. Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. Chemosphere 58:1255-1267.

Chalar G, Arocena R, Pacheco JP, Fabian D. 2010. Trophic assessments of streams in Uruguay: a trophic state index for benthic invertebrates (TSI-BI). Ecological Indicators 11:362-369.

Chaudhuri P and Loh W. 2002. Nonparametric estimation of conditional quantiles using quantile regression trees. Bernoulli 8:561-576.

Chételat J, Pick FR, Morin A, Hamilton PB. 1999. Periphyton biomass and community compositionin rivers of different nutrient status. Canadian Journal of Fisheries and Aquatic Sciences 56:560-569.

Constable M, Charlton M, Jensen F, McDonald K, Craig G, Taylor KW. 2003. An ecological risk assessment of ammonia in the aquatic environment. Human and Ecological Risk Assessment 9:527-548.

Contreras, C. 2007. Nutrient effects in small Brazos basin streams historical data review. Texas Parks and Wildlife Department. Water Quality Technical Series WQTS-2007-01

Cross WF, Wallace JB, Rosemond AD. 2007. Nutrient enrichment reduces constraints on material flows in a detritus-based food web. Ecology 88:2563–2575.

Cross WF, Wallace JB, Rosemond AD, Eggert SL. 2006. Whole-system nutrient enrichment increases secondary production in a detritus-based ecosystem. Ecology 87:1556–1565.

Dávalos-Lind, L. and O.T. Lind. 1999. The algal growth potential of and growth-limiting nutrients in Lake Waco and its tributary waters, part 1 (narrative). Texas Institute of Applied Environmental Research Report.

De'ath G. 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. Ecology 81:3178-3192.

De'ath G, Fabricus KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

DeZwart D, Dyer SD, Posthuma L, Hawkins CP. 2006. Predictive models attribute effects on fish assemblages to toxicity and habitat alteration. Ecological Applications 16:1295-1310.

Dodds WK. 2008. Trophic state, eutrophication, and nutrient criteria in streams. Trends in Ecology and Evolution 22:669-676.

Dodds WK, Clements WH, Gido K, Hilderbrand RH, King RS. 2010. Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. Journal of the North American Benthological Society 29:988-997.

Dodds WK, Oakes RM. 2008. Headwater influences on downstream water quality. Environmental Management 41:367-377.

Dodds WK, Oakes RM. 2006. Controls on nutrients across a prairie stream watershed. Environmental Management 37:634-646.

Dodds, W. K., V. H. Smith, and K. Lohman. 2006. Erratum: nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Canadian Journal of Fisheries and Aquatic Sciences 63:1190-1191.

Dodds WK, Oakes RM. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. Limnology and Oceanography Methods 2:333-341.

Dodds WK. 2003. Misuse of inorganic N and soluble reactive P concentrations to indicate nutrient status of surface waters. Journal of the North American Benthological Society 22:171-381.

Dodds WK, Smith VH, Lohman K. 2002. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Canadian Journal of Fisheries and Aquatic Sciences 59:865-874.

Dodds WK, Jones JR, Welch EB. 1998. Suggested classification of stream trophic distributions of temperate stream types by chlorophyll, total nitrogen, and phosphorus. Water Research 32:1455-1462.

Dodds WK, Smith VH, Zander B. 1997. Developing nutrient targets to control benthic chlorophyll levels in streams: a case study of the Clark Fork River. Water Research 31:1738-1750.

Environmental Protection Agency (EPA). 2006. Wadeable Streams Assessment. EPA-841-B-06-002.

Environmental Protection Agency (EPA). 2002. National Water Quality Inventory: 2000 Report. EPA-841-F-02-001.

Environmental Protection Agency (EPA). 2000a. Nutrient criteria technical guidance manual: rivers and streams. EPA-822-B-00-002.

Evans-White MA, Dodds WK, Huggins DG, Baker DS. 2009. Thresholds in macroinvertebrate biodiversity and stoichiometry across water-quality gradients in Central Plains (USA) streams. Journal of the North American Benthological Society 28:855-868.

Flum T, Nodvin SC. 1995. Factors affecting stream water chemistry in the Great Smokey Mountains. Environmental Science and Technology 37:1575-1580.

Gal G, Anderson W. 2010. A novel approach to detecting a regime shift in a lake ecosystem. Methods in Ecology and Evolution 1:45-52.

Garvey JE, Marschell EA, Wright RA. 1998. From star charts to stoneflies: detecting relationships in continuous bivariate data. Ecology 79:442-447.

Hawkins CP, Paulsen SG, Van Sickle J, Yuan LL. 2008. Regional assessments of stream ecological condition: scientific challenges associated with the USA's national wadeable steam assessment. Journal of the North American Benthological Society 27:805-807.

Hilderbrand RJ, Utz RM, Stranko SA, Raesly RL. 2010. Applying thresholds to forecast potential biodiversity loss from human development. Journal of the North American Benthological Society 29:1009-1016.

Hilsenhoff WL. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin Wisconsin Department of Natural Resources 100.

Herhily AT, Paulsen SG, Van Sickle J, Stoddard JL, Hawkins CP, Yuan LL. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. Journal of the North American Benthological Society 27:860-877.

Herlihy AT, Sifneos JC. 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. Journal of the North American Benthological Society 27:932-948.

Honti M, Istvánovics V, Kovács AS. 2010. Balancing between retention and flushing in river networks—optimizing nutrient management to improve trophic state. Science of the Total Environment 408:4712-4721.

Horner RR, Welch EB, Veenstra. 1983. Development of nuisance periphytic algae in laboratory streams in relation to enrichment and velocity. Pages 121-134 in Periphyton of freshwater ecosystems. Wetzel RG (ed), Junk Publishers, The Hague.

Hutchins MG, Johnson AC, Deflandre-Vlandas A, Comber S, Posen P, Boorman D. 2010. Which offers more scope to suppress river phytoplankton blooms: reducing nutrient pollution or riparian shading? Science of the Total Environment 408: 5065-5077.

Ice G, Binkley D. 2003. Forest stormwater concentrations of nitrogen and phosphorus: a comparison with EPA's proposed water quality criteria. Journal of Forestry 101:21-28.

Joossens JV, Hill MJ, Elliot P,Stamler R, Stamler J, Lessafre E, Dyer A, Nichols R, Kesteloot H. 1996. Dietary salt, nitrate, and stomach cancer mortality in 24 countries. International Journal of Epidemiology 3:494-504.

Justus BG, Petersen JC, Femmer SR, Davis JV, Wallace JE. 2009. A comparison of algal, macroinvertebrate and fish assemblage indices for assessing low-level nutrient enrichment in wadeable Ozark streams. Ecological Indicators 10:627-638.

Karr JR. 1981. Assessment of biotic integrity using fish communities. Fisheries 6(6):21-27.

Keisling, RL, Canova MG, Aragon-Long SC, Hornig CE. 2006. 24 hour Dissolved Oxygen and Use Attainment Study-Texas. United States Geological Survey.

Kilkus SP, LaPerriere JD, Bachmann RW. 1975. Nutrients and algae in some central Iowa streams. Journal of Water Pollution Control Federation 47:77-82.

King RS, Baker ME. 2010. Considerations for analyzing ecological community thresholds in response to anthropogenic environmental gradients. Journal of the North American Benthological Society 29:998-1008.

King, RS, Winemiller KO. 2009. Development of biological indicators of nutrient enrichment for application in Texas streams. Water Quality Assessment Program Special Study #98665304

King, RS, Brooks BW, Back JA, Taylor JM, Fulton BA. 2009. Linking observational and experimental approaches for the development of regional nutrient criteria for wadeable streams. Final Report #CP-966137-01.

King RS, Baker ME, Whigham DF, Weller DE, Jordan TE, Kazyak PF, Hurd MK. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

King RS, Richardson CJ. 2003. Integrating bioassessment and ecological risk assessment: an approach to developing numerical water-quality criteria. Environmental Management 31:795-809.

Knobeloch L, Salna B, Hogan A, Postle J, Anderson H. 2000. Blue babies and nitrate-contaminated well water. Environmental Health Perspectives 108:675-678.

Lin LS, Markus M, Russell A. 2007. Stream classification system based on susceptibility to algal growth in response to nutrients. Journal of Environmental Engineering 133:692-697.

Lohman K and Jones JR. 1999. Nutrient-sestonic chlorophyll relationships in northern Ozark streams. Canadian Journal of Fisheries and Aquatic Sciences 56:124-130.

Lohman K, Jones JR, Perkins BD. 1992. Effects of nutrient enrichment and flood frequency on periphyton biomass in Northern Ozark streams. Canadian Journal of Fisheries and Aquatic Sciences 49:1148-1205.

Longing SD, Haggard BE. 2010. Distributions of median nutrient and chlorophyll concentrations across the Red River Basin, USA. Journal of Environmental Quality 39:1966-1974.

Mabe JA. 2007. Nutrient and biological conditions of selected small streams in the Edwards Plateau, Central Texas, 2005-06, and implications for development of nutrient criteria. United States Geological Survey. Scientific Investigations Report 2007-5195.

Matlock, M.D. and A.D. Rodriguez. 1999. Preliminary report of findings: periphytometer study on streams in the Lake Waco/Bosque River Watershed. Texas Institute for Applied Environmental Research

Matlock MD, Storm DE, Smolen MD, Matlock ME, McFarland AMS, Hauck LM. 1999. Development and application of a lotic ecosystem trophic status index. Transactions of the American Society of Agricultural Engineers 42:651-656.

Matlock MD, Matlock ME, Storm DE, Smolen MD, and Henley WJ. 1998. A quantitative passive diffusion periphytometer for lotic ecosystems. Journal of the American Water Resources Association 34:1141-1147.

McFarland AMS, Millican JS, Back JA, Brister JS. 2008. Trophic status related to periphyton and macrophyte distribution and abundance along the North Bosque River in north central Texas. Texas Journal of Science 60:3-22.

McMahon G, Alexander RB, Qian S. 2003. Support of total maximum daily load programs using spatially referenced regression models. Journal of Water Resources Planning and Management 129:315-329.

Miltner RJ. 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. Environmental Management 45:842-855.

Morgan AM, Royer TV, David MB, Gentry LE. 2006. Relationships among nutrients, chlorophyll-a, and dissolved oxygen in agricultural streams in Illinois. Journal of Environmental Quality 35:1110-1117.

Mulholland PJ, Webster JR. 2010. Nutrient dynamics in streams and the role of J-NABS. Journal of the North American Benthological Society 29:100-117.

Mummert AK, Neves RJ, Newcomb TJ, Cherry DS. 2003. Sensitivity of juvenile freshwater mussels (*Lampsilis fasciola*, and *Villosa iris*) to total and unionized ammonia. Environmental Toxicology and Chemistry 22:2545-2553.

Munn M., Frey J, Tesoriero A. 2010. The influence of nutrients and physical habitat in regulating algal biomass in agricultural streams. Environmental Management 45:603-615.

Ode PR, Hawkins CP, Mazor RD. 2008. Comparability of biological assessments derived from predictive models and multimetric indices of increasing geographic scope. Journal of the North American Benthological Society 27:967-985.

Palmstrom N. 2005. Development of regional nutrient criteria and implications for states-- and the regulated community. Lake and Reservoir Management 21:174-185.

Paul JF, McDonald ME. 2005. Development of empirical, geographically specific water quality criteria: a conditional probability analysis approach. Journal of the American Water Resources Association. 41: 1211-1223.

Peters RH. 1986. The role of prediction in limnology. Limnology and Oceanography 31:1143-1159.

Porter SD, Mueller DK, Spahr NE, Munn MD, Dubrovsky NM. 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Freshwater Biology 53:1036-1054.

Qian SS, King RS, Richardson CJ. 2003. Two methods for the detection of environmental thresholds. Ecological Modeling 166:87-97.

Radloff, PL, Contreras C, Whisenant A, Bronson JM. 2010. Nutrient effects in small Brazos Basin streams final report. Texas Parks and Wildlife Department. WQS-2010-02.

Reid NJ, Hamilton SK. 2007. Controls on algal abundance in a eutrophic river with varying degrees of impoundment (Kalamazoo River, Michigan, USA). Lake and Reservoir Management 23: 219-230.

Rodinov S. 2004. A sequential algorithm for testing climate regime shifts. Geophysical Research Letters 31:L090204.

Rohm CM, Omernik JM, Woods AJ, Stoddard JL. 2002. Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development. Journal of the American Water Resources Association 38:213-239.

Royer TV, David MB, Gentry LE, Mitchell CA, Starks KM, Heatherly T, Whiles MR. 2008. Assessment of chlorophyll-a as a criterion for establishing nutrient standards in streams and rivers of Illinois. Journal of Environmental Quality 37:437447.

Russo R C. 1985. Ammonia, nitrite and nitrate. Pages 455-471 *in* G. M. Rand & S. R. Petrocelli, editors. Fundamentals of aquatic toxicology. Hemisphere Publishing Corporation, Washington, D.C., U.S.A.

Singer GA, Battin TJ. 2007. Anthropogenic subsidies alter consumer-resource stoichiometry, biodiversity, and food chains. Ecological Applications 17:376–389.

Shapiro MH, Holdsworth SM, Paulsen SG. 2008. The need to assess the condition of aquatic resources in the US. Journal of the North American Benthological Society 27:808-811.

Shukla JB, Misra AK, Chandra P. 2008. Mathematical modeling and analysis of the depletion of dissolved oxygen in eutrophied water bodies affected by organic pollutants. Nonlinear Analysis 9:1851-1865.

Smith AJ, Tran CP. 2010. A weight-of-evidence approach to define nutrient criteria protective of aquatic life in large rivers. Journal of the North American Benthological Society 29:875-891.

Smith JG, Beauchamp JJ, Stewart AJ. 2003. Alternative approach for establishing acceptable thresholds on macroinvertebrate community metrics. Journal of the North American Benthological Society 24:428-440.

Smith RA, Alexander RB, Schwarz GE. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous Unites States. Environmental Science and Technology 37:3039-3047.

Snyder EB, Robinson CT, Minshall GW, Rushforth SR. 2002. Regional patterns in periphyton accrual and diatom assemblage structure in a heterogenous nutrient landscape. Canadian Journal of Fisheries and Aquatic Sciences 59:564-577.

Soballe DM, Kimmel BL. 1987. A large-scale comparison of factors influencing phytoplankton abundance in rivers, lakes, and streams. Ecology 68:1943-1954.

Sonderegger DL, Wang H, Clements WH, Noon BR. 2009. Using SiZer to detect thresholds in ecological data. Frontiers in Ecology and the Environment 7:190-195.

Stanley EH, Short RA, Harrison JW, Hall R, Wiedenfeld RC. 1990. Variation in nutrient limitation of lotic and lentic algal communities in a Texas (USA) river. Hydrobiologia 206:61-71.

Stevenson RJ, Hill BH, Herlihy AT, Yuan LL, Norton SB. 2008. Algae-P relationships, thresholds, and frequency distributions guide nutrient criterion development. Journal of the North American Benthological Society 27:783-799

Stevenson RJ, Pan Y, Manoylov KM, Parker CA, Larsen DP, Herlihy AT. 2008. Development of diatom indicators of ecological conditions for streams of the western US. Journal of the North American Benthological Society 27:1000-1016.

Stevenson RJ, Rier ST, Riseng CM, Schultz RE, Wiley MJ. 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 RJ, Bothwell M, Lowe RL (editors). 1996. Algal ecology: freshwater benthic ecosystems. Academic Press, San Diego, California.

Stoddard, JL, Peck DV, Olsen AR, Larsen P, Van Sickle J, Hawkins CP, Hughes RM, Whittier TR, Lomnicky G, Herlihy AT, Kaufmann PR, Peterson SA, Ringold PL, Paulsen SG, Blair R. 2005. Environmental Monitoring and Assessment Program (EMAP): western streams and rivers statistical summary. EPA

620/R-05/006. Office of Research and Development, US Environmental Protection Agency, Washington DC.

Suplee MW, Varghese A, Cleland J. 2007. Developing nutrient criteria for streams: an evaluation of the frequency distribution method. Journal of the American Water Resources Association 43(2):453-472.

Suplee MW, Watson V, Teply M, McKee H. 2009. How green is too green? Public opinion of what constitutes undesirable algae levels in streams. Journal of the American Water Resources Association 45:123-140.

Utz R, Hilderbrand RH, Boward DM. 2009. Identifying regional differences in threshold responses of aquatic invertebrates to landcover gradients. Ecological Indicators 9:556-567.

Van Nieuwenhuyse EE, Jones JR. 1996. Phosphorus-chlorophyll relationship in temperate streams and its variation with stream catchment area. Canadian Journal of Fisheries and Aquatic Sciences 53:99-105.

Vermeer ITM, Pachen D, Dellinga JW, Kleinjans JCS, van Maanen JMS. 1998. Volatile N-nitrosamine formation after intake of nitrate at the ADI level in combination with an amine-rich diet. Environmental Health Perspectives 106: 459-463.

Waite IR, Carpenter KD. 2000. Associations among fish assemblage structure and environmental variables in Willamette Basin Steams, Oregon. Transactions of the American Fisheries Society 129:754-770.

Wang LZ, Robertson DM, Garrison PJ. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implication to nutrient criteria development. Environmental Management 39:194-212.

Washington HG. 1984. Diversity, biotic and similarity indices: a review with special relevance to aquatic ecosystems. Water Resources 18: 653-694.

Weigel BM, Robertson DM. 2007. Identifying biotic integrity and water chemistry relations in nonwadeable rivers of Wisconsin: toward the development of nutrient criteria. Environmental Management 40:691-708.

Welch EB, Jacoby JM, Horner RR, Seeley MR. 1988. Nuisance biomass levels of periphytic algae in streams. Hydrobiologia 157:161-168.

Wickham JD, Riitters KH, Wade TG, Jones KB. 2005. Evaluating the relative roles of ecological regions and land-cover composition for guiding establishment of nutrient criteria. Landscape Ecology 20:791-798.

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

APPENDIX 1-1: Stream Study Report from the State of Texas Reviews

Agrilife Research. 2010. Analysis and summary of data collection methods for nutrients in Texas streams. Contract #582-9-90439-08

The objective of this study was to assess quantitative and qualitative measures of periphyton biomass and to assess their efficacy as an indicator of nutrient enrichment in Texas streams. The work focused on 30 wadeable stream and river stations and the data presented was collected at each station in June 2010. Sampling sites selected were located included the Cross Timbers (29), Texas Blackland Prairies (32), East Central Texas Plains (33), and Western Gulf Coastal Plains (34) level III ecoregions. Data collection activities included water chemistry sampling and analyses, routine water quality field measurements, four methods of periphyton sampling and assessment, flow measurement, and physical habitat assessment. The four methods used for periphyton sampling and assessment included periphyton scraping followed by chlorophyll a (chl-a) and ash free dry mass (AFDM) measurement in the laboratory (method 1), areal coverage in a bucket (method 2), picking up cobbles in a transect to estimate percent coverage on the cobble (method 3), and percent coverage estimated from photos of a $1m^2$ sampling frame (method 4). Methods 2 and 3 quantified macro- and micro-algae. Method 3 also noted the coverage of moss. Measurements of the maximum length of macroalgae and the thickness of microalgae were taken at each location using method 2. Method 4 classified the extent of 1) trace, thin film, or sparse scatter of filaments, 2) felt short filaments or scattered cushions, nodules or clumps, and 3) thick mats of benthic or floating periphyton. Method 4 also classified algal condition as good, fair, or poor.

The natural log transformed combined fraction of microalgae and macroalgae taken using method 2 had a significant correlation with AFDM ($r = 0.52$, $p = 0.006$). Macroalgal cover estimated by method 3 was positively correlated with chl-a ($r = 0.37$, $p = 0.044$) and AFDM ($r = 0.46$, $p = 0.013$). However, neither microalgal cover nor moss cover estimated in method 3 had statistically significant correlations with chl-a or AFDM. When all three coverage measures were added together, a significant relationship was found with chl-a ($r = 0.44$, $p = 0.016$) and AFDM ($r = 0.41$, $p = 0.029$). No significant correlations were found between the average cover values by category or combined categories taken using method 4 and chl-a or AFDM measures.

Results suggested that periphyton methods 2 and 4 are limited in turbid waters that limit visibility of the benthos. Qualitative measures of periphyton coverage from method 3 did correlate with quantitative measures of chl-a and AFDM, while measures from method 2 only correlated with AFDM. Method 4 measurements were never correlated with quantitative measures and were not correlated with measures of water quality. Correlation analysis between water quality variables and measures of periphyton biomass usually showed no relationship or a weak relationship (i.e., $r < 0.55$). Several periphyton measures had negative correlations with TN and positive correlations with DO. The nature of relationships between pH and periphyton measures varied with the method and the parameter measured. The authors cautioned readers about relationships between algal measures and pH since they were based on instantaneous measurements collected during daylight hours. Since large diel ranges in DO and to a lesser degree pH are expected in eutrophic stream systems, the time of day when instantaneous measurements are taken will affect the observed relationships with water chemistry variables and the interpretation of the data.

Alan Plumer Associates. 1996. City of Leander study of algae cover in the upper Brushy Creek watershed. Report 501-0200.

The objective of this study was to evaluate whether removing a higher level of phosphorus (P) from water discharged from a wastewater treatment plant (WWTP) in the city of Leander would reduce benthic algal cover downstream. The study design involved sampling 9 study sites that represented a stream that receives wastewater discharges (n = 4), a stream that receives effluent that has been treated to remove P (n = 3), and a stream receiving no wastewater effluent (n = 2) in the Brushy Creek watershed. Sites were sampled on 9 dates from late summer 1993 to spring 1995. Algal cover was estimated from photographs taken at each site and was categorized as having a trace, thin film, or sparse scatter of filaments (category 1), having short filaments, scattered small cushions, nodules, or clumps (category 2), or having thick mats, either benthic or floating (class 3). Ammonia, nitrate plus nitrite, total Kjeldahl nitrogen and orthophosphate and total P (TP) were also measured in water grab samples taken at each site. Flow data were taken from a USGS gaging station outside of the watershed.

No statistical analysis information exists for the algal cover data, which ranged from 0-86%, 0-39%, 0-100%, and 0-100% for classes 1, 2, 3, and for the adjusted combined algal coverage value, respectively. Algal cover was lower at sites that had no effluent discharge compared to sites that had effluent discharge although no statistics were reported.

In-stream ammonia nitrogen concentrations ranged from 0.01-0.92 mg/L. The highest concentration was taken in June 1994 from station 2 that was representative of urban development, but no WWTP discharge. Nitrate plus nitrite nitrogen ranged from 0.01 to 3.54 mg/L. The highest concentration was found at station 3 below Shipman dam and the WWTP in June 1994. Again, one of the highest nitrate concentrations (3.16 mg/L) was found at station 2, but in a different month (February 1994). Total Kjehldahl nitrogen ranged from 0.02 mg/L at station 2 in May to 2.92 mg/L at station 1, a stream without WWTP discharges or urbanization, in July 1993. In-stream inorganic nitrogen concentrations ranged from 0.02 at several sites to 12.0 mg/L at station 3 in July 1994.

Orthophosphate ranged from 0.001 mg/L at station 1 in February to 2.162 mg/L at station 3 in July 1994. Total phosphorus concentrations ranged from 0.01 mg/L found at several sites to 5.39 mg/L found at station 3 in June 1994. Spearman correlation analyses revealed significant correlations between orthophosphate and TP (values not reported), but no relationships between algal cover and water chemistry were found. The median ratio of total nitrogen to total phosphorus was 5.2 suggesting that the stream may be limited by nitrogen instead of phosphorus.

The report suggests that algal production in this stream might be limited by water depth, velocity, surface area, shading bottom characteristics, or some other physical factor. The authors also advise that orthophosphate should not necessarily be considered reflective of total phosphorus and that nonpoint sources provided phosphorus loads in sufficient quantities to support extensive algal growth. The phosphorus concentrations in tributaries that received discharges from WWTPs that provide phosphorus reduction were not different from those that received WWTP discharge without phosphorus reduction.

Bayer, C.W., Davis, J.R., Twidwell, S.R., Kleinsasser R., Linam G., Mayes K., and Hornig E. 1992. Texas aquatic ecoregion project: an assessment of least disturbed streams. Draft Report.

The objectives of this study were to aid in the classification of small streams for water quality criteria by 1) improving physicochemical and biological databases for small Texas streams, 2) determining if regional patterns in this data exist, 3) defining and verifying Texas ecoregions, and 4) developing procedures for assessing aquatic life use (ALU) in Texas streams. The study provides data from 72 wadeable streams in 11 of the 12 possible Texas ecoregions (Ecoregion 23 missing) collected from 1986 to 1990 during the "summer period", which generally is from June-Sept, but can extend into October. The "summer period" was sampled to provide observations during a critical low flow and elevated temperature period. A few streams were sampled in spring 1989 to identify conditions during a critical spawning period and to examine whether summer samples adequately represented fish assemblages in Texas streams. Each stream was sampled at one site during each sampling event. The number of streams per ecoregion ranged from 4-11.

This report contained excellent physical descriptions of each site including drainage basin size, soil types, floral, land use, and stream characteristic data. All physicochemical data among ecoregions were statistically different (overall ANOVA $p \leq 0.02$) except for DO ($p = 0.1$), daily minimum DO (mean range = 4.4-5.8), and fecal coliform numbers (all samples < 200/100mL). Total phosphorus (TP; mean range = 0.02-0.28 mg/L), orthophosphate (mean range = 0.01-0.25 mg/L), and ammonia nitrogen (N; mean range = 0.02-0.30) differed among ecoregions ($p < 0.01$). Kjeldahl-N (mean range = 0.30-1.2 mg/L) differed among ecoregions ($p = 0.05$), but nitrate-N (mean range = 0.01-1.96) did not differ. Total dissolved solids (mean range = 111-6471 mg/L; $p < 0.0001$), chloride (mean range = 15-3040 mg/L; $p < 0.0001$), and sulfate concentrations (mean range = 8-857 mg/L; $p = 0.05$) also differed among ecoregions. No significant differences were found among ecoregions for total suspended solids or volatile suspended solids. Carbonaceous five day biochemical oxygen demand concentrations were generally less than 3 mg/L except for in the Western Gulf Coastal Plain where 57% of the streams sampled were $\leq$3.0 mg/L. Suspended chlorophyll *a* and turbidity means ranged from 2-8 mg/L and 0.4-13.2 NTU, respectively.

Benthic macroinvertebrate scores based on TWC macrobenthic criteria indicated that 77.8% of the streams met the criteria for high or exceptional subcategory designations, but only 67.9% meth this designation if the Ohio Invertebrate Community Index was used. The average number of species in each ecoregion ranged from 23.67 in ecoregion 26 to 56.75 in ecoregion 24. The EPT index ranged from 6.33 in ecoregion 26 to 15.00 in ecoregion 30. A majority of the samples that fell below the high category were from intermittent streams. The greatest number of species was found in eastern Texas relative to the Panhandle and West Texas. Fewer species of darters existed as you move from east to west across Texas.

This study concluded that data on physical habitat, water quality, macroinvertebrates and fishes indicated that a presumptive use of high aquatic life use was justified for many perennial streams in the state. Adjustments for regional or site-specific characteristics should be made to the ALU classes and supporting criteria.

Brazos River Authority. Environmental Services. 1998. South Fork of San Gabriel River Nutrient Study. NPDES Permit # WQ0014477001

The objectives of this study were to conduct sampling and analysis to assess nutrient concentrations and algal growth in the South Fork San Gabriel River in preparation for the development of a wastewater treatment plant that will serve the Liberty Hill CCN, in western Williamson County. Sampling was planned to occur at least 12 months prior to discharge and at least two years after discharge began. Benthic algal biomass measured as chlorophyll a was assessed using periphytometers and by visual estimates of percent areal coverage using permanent transects and photographs. Suspended chlorophyll a was also measured. The following water chemistry parameters were also measured: total dissolved solids, sulfate, chloride, nitrate-nitrogen (N), nitrite-N, orthophosphate-P, total ammonia, TP, and TKN. They also measured pH, DO (mg/L and %sat), conductivity, salinity, temperature, secchi depth, depth of the stream at the sample point, weather (clear, partly cloudy, cloudy, rain), total water depth, water velocity, flow severity, and percent aquatic vegetation. Samples were taken from November 2005 to October 2008. Wastewater discharge began Nov. 1, 2006.

Average ammonia N ranged from 0.10 mg/L at Mankins Branch to 0.22 mg/L at Upper Middle and was below the screening level (1.0 mg/L) at all sites. Average nitrate plus nitrite N was below the screening level (1.0 mg/L) at 36% of sites and ranged from 0.12 mg/L at Lower Middle to 12.06 mg/L at Mankins Branch. Orthophosphate was below the screening level (0.1 mg/L) at approximately 73% of the sites and got as high as 1.60 mg/L at Mankins Branch. Total phosphorus was below the screening level (0.2 mg/L) at 73% of sites and got as high as 1.70 mg/L at Mankins Branch. Chlorophyll *a* was below the screening level (30 µg/L) in 82% of the streams and ranged from 4.49 µg/L at South Fork to 21.56 µg/L at Mankins Branch.

A total of 17 macrophyte species were found in this study. Macrophyte total percent coverage ranged from 3-40 and richness ranged from 1-10. Total algae percent coverage ranged from <5 to 70. The only bivariate analysis conducted in the report was between flow and total phosphorus and nitrate and no statistics were reported. Basically, higher nitrate plus nitrite and total phosphorus concentrations were found at lower flow.

The author's recommendations for nutrient reduction suggested that a majority of sites that exceeded screening levels for nitrate plus nitrite N and total phosphorus were taken from the San Gabriel at Berry Creek and at Mankins Branch at County Road 104 and 100 during non-high flow conditions. Samples taken from the San Gabriel at highway 95 at low flow also exceeded screening levels for nitrate plus nitrite N. Most of these sites were associated with wastewater treatment plant discharges that appeared to be affecting water quality. The authors recommended that the City of Georgetown consider nutrient removal at the wastewater treatment plants.

The data is available on an Excel file (BRA_SFSG_Final_13Nov08.xls). The filename for the quality assurance project plan and the water quality monitoring plan are BRA SFSG QAPP for San Gabriel Project updated Jun08.doc and BRA SFSGworkplanFeb06.doc. The final report is San Gab BRA Algae 1998.pdf.

Contreras, C. 2007. Nutrient effects in small Brazos basin streams historical data review. Texas Parks and Wildlife Department. Water Quality Technical Series WQTS-2007-01

The study objective was to add to the body of data of the effects of nutrient enrichment on small streams in Texas Brazos River basin and to increase knowledge about the distribution and status of freshwater mussels in this basin. Six small perennial stream sites were selected in the Brazos River basin. Sites were sampled for basic water quality, instantaneous flow, diel variables, organics and metals, fish, benthic invertebrates including mussels, periphyton, and habitat between March 15 – October 15 (index period) and July 1-September 30 (critical period) each year from 1989-2004.

Duck Creek, which contains a significant amount of poultry production and poultry litter application in its watershed and has 3 industrial wastewater discharges, was historically designated as a stream not supporting contact recreation due to the *E. coli* geometric mean and was listed as a concern for life use support based on DO grab samples. In the present study, instantaneous DO, sulfate, and fecal coliform grab samples exceeded sample criterion 2, 8, 47, and 22% of the time, respectively.

In previous studies conducted in the 1987-88, water quality standards and criteria in the Navasota River were being met. However, in 2006 one sample station was listed as not supporting contact recreation because of an *E. coli* geometric mean. Fecal coliform single sample measurements also caused it to be listed as a concern for near non-attainment. In the present study, around 2% of samples exceeded the sulfate criterion. The geometric mean of *E. coli* samples did exceed the criterion 42% of the time. Fecal coliform samples exceeded the grab sample criterion 23% of the time, but the geometric mean remained below the criterion. Chlorophyll *a* also exceeded the screening level 23% of the time. However, none of the water quality measurements exceeded any criteria.

There was no historical water quality data associated with Clear Creek or its tributaries. Little Elm Creek had few exceedances. The 2-h peak flow limit, the daily average for $CBOD_5$, and the daily maximum for $CBOD_5$ were each exceeded one time. In the present study, fecal coliform exceeded the criterion in 1 out of 2 samples and nitrate exceeded the screening level.

The minimum DO in Walnut Creek was below the criterion 6 times. Some of the wastewater treatment discharges violated several criteria including pH, ammonia nitrogen, fecal coliform, and $CBOD_5$. Sulfate exceeded the criterion 2% of the time. The grab sample mean for *E. coli* and for coliform was exceeded 45 and 21% of the time, respectively. There were no other exceedances.

Willis Creek had previous exceedances of the *E. coli* geometric mean and single-sample criterion. Fecal coliform was near non-attainment in single sample measurements. In the present study, *E. coli* exceeded the single sample criterion 43% of the time and the geometric mean was exceeded. Fecal coliforms exceeded the single sample criterion 36% of the time, but the geometric mean was not exceeded. Nitrate exceeded the screening level a majority of the time and nitrite exceeded the screening level once. In 2004, biological data was taken twice. The habitat quality index was high each time and the benthic IBI score scored within the exceptional or the high life use category. Diel DO measurements indicated that the site was exceptional and fish IBI was high and intermediate to high.

Dávalos-Lind, L. and O.T. Lind. 1999. The algal growth potential of and growth-limiting nutrients in Lake Waco and its tributary waters, part 1 (narrative). Texas Institute of Applied Environmental Research Report.

The objectives of this study were to 1) examine whether the potential for water entering Lake Waco from 4 tributaries differed in their ability to promote algal growth and reservoir eutrophication, 2) examine whether algal growth potential (AGP) differ between reservoir headwaters and near the dam, 3) examine whether there was seasonal variation in the AGP from the different sources, 4) determine whether nitrogen (N) or phosphorus (P) limited algal growth in the reservoir and its tributaries, and 5) determine whether native phytoplankton community structure changed in response to the addition of the growth-limiting nutrient. Algal bioassays of *Selenastrum capricornutum* Printz and native phytoplankton collected from the Lake Waco North Bosque Arm site were used to address these objectives 1-4 during 2 annual cycles and AGP was reported in the percentage increase in fluorescence on the date of maximum fluorescence compared to the initial fluorescence. *In situ* bioassays were conducted to examine objective 5.

Comparisons among years, seasons, and sites were done only by comparing means, standard deviations, and coefficients of variation. No analysis of variance or means comparison statistical test results were reported. AGP's estimated for *S. capricornutum* and for native phytoplankton were correlated (R^2 = 0.62). So most of the discussion focused on AGPs calculated for *S. capricornutum.* There were no major differences in results and conclusions between years, but AGP varied greatly among sampling sites. Tributary water *S. capricornutum* AGP means ranged from 292-857% and had less growth potential than did reservoir water which ranged from 711-1285%. Of the tributaries, the South Bosque River had the highest AGP at 857%. AGP within sites varied greatly by season with the winter and spring having higher AGP than the late spring and summer. Reservoir AGPs were less variable seasonally than tributary AGPs.

The growth limiting nutrient was P at most sampling sites on most dates. Nitrogen occasionally produced a greater algal response than P in the Lake Waco North Bosque Arm site. In addition, co-limitation of N and P did occur in the autumn months at some of the reservoir sites. However, N most often produced no AGP suggesting that ambient supplies were adequate for algal growth. The growth response to added P varied seasonally for reservoir sites with a lower algal growth response in the summer compared to in other seasons. A similar pattern was not seen in tributaries because of low summer flow.

Taxonomic analysis of reservoir algal communities in Lake Waco was typical of mesotrophic to eutrophic lakes. Cyanobacterial species were dominant much of the year, but increased in abundance during the summer months. Chlorophytes and diatoms reached maximum abundance in the spring. Centrate diatoms were abundant mostly during the winter and early spring. When P was added to experimental cultures, Cyanobacteria declined in relative abundance while Bacillariophyta increased.

Keisling, R.L, M.G. Canova, S.C. Aragon-Long, C.E. Hornig. 2006. 24 hour Dissolved Oxygen and Use Attainment Study-Texas. United States Geological Survey.

****Figure numbers cited in the text of this report do not match figure numbers in the separate figure file. This fact along with figure and table legends that were not very descriptive made it difficult to confirm specifics about some of the variables actually shown in figures.****

The objective of this study was to evaluate whether aquatic life use (ALU) based on dissolved oxygen (DO) criteria were attained and to determine whether ALU conditions and DO data were related. Water quality, macroinvertebrate, and fish data were taken for 33 streams across Texas from May 2003 to July 2005. Temperature, pH, conductivity, and DO data were taken during 24-72 hour deployments during the index period from March 15 to October 15 and during the critical assessment period of May to September for 2003 and 2004.

Significant monthly patterns in average diel DO were often found within years at each site, but few sites had significant differences between the first and second 24 hours of deployment (mean relative percent difference between 1st and 2nd 24h = 3.1%). Monthly variation was not necessarily consistent across years within sites. However, 24-h mean DO by sites did not differ as frequently across sites and had no statistically significant differences between years. Therefore, the authors suggest using 24-hour mean DO to compare across sites. Specific conductance and temperature averaged over 24 hours did not exhibit many differences between months and years within sites.

Thirty percent of the stream sites showed exceedances of the DO criteria for more than 50% of the observations. In addition, 30% of the sites had IBI biocriteria that exceeded criteria more than 50% of the time. However, there was no systematic relationship between macroinvertebrate and fish index of biological integrity (IBI) and the 24-hour average DO collected during biological sampling periods.

A positive and a negative linear statistical relationship was found between periphyton chlorophyll *a* (chl-a) and the IBI (R^2 = 0.27; $p < 0.01$) and the habitat quality index (HQI) (R^2 = 0.18, $p < 0.03$), respectively. The IBI was not correlated with sestonic chl-a, but it was positively correlated with drainage area (statistical output not reported).

Seston chl-a was weakly negatively correlated to 24 h average (R^2 = -0.22), minimum (R^2 = -0.22), and maximum DO (statistical data not reported). According to the authors, the decline in DO minimum seemed most pronounced when seston concentrations exceeded 10 µg/L. Seston chl-a also had a weak positive correlation with HQI and a possible threshold effect was found between periphyton chl-a and seston chl-a (R^2 = 0.12). There was also a positive relationship between periphyton chl-a and diel DO maximum (R^2 = 0.12) and 24 h DO range (R^2 = 0.18). Finally, positive linear relationships between phosphate phosphorous and benthic chl-a (R^2= 0.19) and nitrate and benthic chl-a (R^2= 0.13) were found. Benthic chl-a was negatively related to sestonic chl-a (R^2= 0.12).

King, R.S. and K.O. Winermiller. 2009. Development of biological indicators of nutrient enrichment for application in Texas streams. Water Quality Assessment Program Special Study #98665304

The objectives of this study were to 1) compare TN and TP concentration estimates from Baylor University (BU) and Texas Commission on Environmental Quality (TCEQ) laboratories, 2) evaluate the soft-substrate periphyton responses to nutrients, 3) estimate periphyton and other biological thresholds if present, 4) evaluate responses of fishes to enrichment and sedimentation, and 5) recommend responsive ecological indicators and nutrient concentrations that cause changes in those ecological indicators. Data were collected from Brazos and Trinity River wadeable streams (n=64) located in the Cross Timbers (29), Blackland Prairies (32), and East Central Texas (33) level III ecoregions from June-August 2008.

Total P measurements from TCEQ and BU were positively related (R^2=0.90) when points below the 50 µg/L TP TCEQ LOD, which represented 55% of the sites, were removed from the analysis. TN data between TCEQ representing TKN+nitrate-N+ammonia-N and BU representing a persulfate digest were similar (R^2=0.91), but variance in TCEQ data increased as concentrations decreased and TCEQ was consistently above the 1:1 line for the two methods. The TCEQ LOD may have contributed to this pattern. Periphyton nutrient ratio responses to nutrient concentrations differed between rocky (ecoregion 29) and sandy/muddy streams (ecoregions 32 and 33). Epilithic periphyton nutrient C:P ratios had a threshold decline with increasing TP, while periphyton on muddy/sandy substrate did not respond to nutrients. Suspended chlorophyll *a* had a positive threshold relationship to TN (~350 µg/L) and TP (~25 µg/L) in ecoregion 29, but was not related in ecoregions 32 and 33 potentially due to a lower sample size and nutrient concentration range.

Ordinations suggested that algal metrics should be stratified by ecoregion, but not by basin. Ecoregion 29 algal species ordinations were related to TP, TN, pasture, outfalls, sediment, and chloride but ordinations in ecoregions 32 and 33 were not related to nutrient or nutrient-related variables. Threshold analyses indicated that 31 algal species declined between 15 and 25 µg/L TP, while 36 algal species increased between 20 and 50 µg/L TP. Most of these taxa also responded negatively to decreasing C:P periphyton ratios, which strengthens the link between changes in surface water TP and algal responses.

Fish community structure at sites from ecoregion 32 with low TP, chloride, substrate embeddedness, mud-silt, chlorophyll *a*, filtrable and nonfiltrable residue, and high periphyton C:P, C:N, and N:P ratios were groups from sites that had opposite values for these same variables. Ecoregions 32 and 33 had weak relationships with outfalls, rowcrop, pasture and impervious cover, but they weren't strong enough relationships to be interpretable at this point. Four fish species declined between 15-25 µg/L TP and four species increased at 30 µg/L TP. Most of the same species responded to periphyton C:P, chloride, mud-silt cover, substrate embeddedness, outfalls, and pastures. Five potential fish metrics were suggested: fish community index (nMDS Axis 1), percent grazing herbivore abundance, percent abundance of darters, percent abundance of nutrient-tolerant cyprinids.

King, R.S., B.W. Brooks, J.A. Back, J.M. Taylor, B.A. Fulton. 2009. Linking observational and experimental approaches for the development of regional nutrient criteria for wadeable streams. Final Report #CP-966137-01.

This study linked a field and an experimental stream study examining the effects of TP and other water quality parameters on stream dissolved oxygen concentrations and on benthic algae and macroinvertebrate communities. The observational field study was conducted in the Brazos River basin within the Cross Timbers Level III ecoregion. Relationships between biota and TP were evaluated using ordinations and TITAN. Periphyton C:P, N:P, and C:N had a significant negative threshold response to increasing TP with TP concentrations around 19-20 µg/L showing the most consistent threshold decline in periphyton stoichiometry. DIN had less control over periphyton stoichiometry than did TP. The thickness of microbial films growing on rocks also showed threshold declines around 20 µg TP/L. Filamentous green algal cover exhibited 2 distinct increases around 20 and 200 µg TP/L and chl-a:AFDM shifted towards a greater fraction of chl-a at these same concentrations. Hydrology influenced patterns in DO, water temperature, and pH across streams. In 2006, which was a drought year, few of the streams were flowing and DO $\leq$2 mg/L was found in streams that had >27 µg TP/L. Minimum DO during this drought year was significantly predicted by algal variables suggesting that these metrics may be linked quantitatively to aquatic life use standards. TITAN analysis of algal species data indicated that 19.2-21.6 µg TP/L was the threshold of greatest overall decline in species. Numerous species including *Cladophora* showed threshold increases simultaneously. A few taxa showed threshold increases between 200-500 µg TP/L corresponding to filamentous green algal cover and chl-a: AFDM threshold increases. Macroinvertebrates showed more consistent relationships to TP at sites with no flow (2006) compared to sites with flow (2007), but there were fewer sites sampled with flow than without flow. TITAN thresholds were reported for individual taxa and more taxa showed negative compared to positive responses.

A controlled 28-day P-dosing experiment was conducted in 12 artificial streams having either 8 (controls), 20, or 100 µg PO_4-P/L. TP ranged from 19.1-20.5, 37.7-40.3, and 127.2-137.2 µg/L in the controls, low P, and high P streams, respectively. Periphyton biomass on ceramic tiles was higher in the high P treatment relative to the low P treatment and control on day 14, but by day 28, the low and high P treatments had similar biomass that was higher than controls. *Cladophora* biomass was significantly higher in the low and high P treatments compared to the control on day 28. Periphyton C:P differed significantly among control (mean ~320), low (mean~230), and high (mean ~150) streams. Algal species responses to P enrichment mirrored field responses. Control streams increasingly resembled algal communities from the P-enriched field sites over time until 28 days. Five of the 7 taxa showing responses to P-enrichment in the experiment also showed responses in the field. Macroinvertebrate taxonomic composition did not differ among treatments in the experiment as they did in the field, which may be attributed to the relatively short dosing period in the experiment relative to the life cycle of macroinvertebrate taxa.

The authors suggest that streams with 20 µg TP/L and possibly as low as 15 µg TP/L should experience a threshold decline in biological integrity. TP concentrations $\geq$200 µg/L may have more consistent nuisance algal growth and fewer taxa.

Mabe, J.A. 2007. Nutrient and biological conditions of selected small streams in the Edwards Plateau, Central Texas, 2005-06, and implications for development of nutrient criteria. United States Geological Survey. Scientific Investigations Report 2007-5195.

Water quality, algae, benthic macroinvertebrates and fishes were sampled from 15 streams located in the Edwards Plateau of Central Texas in summer 2005. A subset (n = 8) of these streams were re-sampled in summer 2006. Spearman rank correlations and linear regressions were used to analyze bivariate data and Kruskal Wallace tests were used to assess differences in univariate data among streams receiving wastewater effluent, streams not receiving wastewater effluent, and least disturbed streams. Total N ranged from 0.12 to 4.81 mg/L and nitrate-nitrite N (NO_3-NO_2-N) ranged from 0.004 to 4.67 mg/L. Total P ranged from 0.001 to 3.52 mg/L and 82% of the orthophosphate (PO_4) concentrations were less than 0.004 mg/L. Mean measured N and P concentrations from least-disturbed sites that could serve as criteria were 0.18, 0.068, 0.265, and 0.003 mg/L for TKN, NO_3-NO_2-N, TN, and TP, respectively. Streams receiving wastewater effluent always had higher nutrient concentrations than those not receiving effluent and the least disturbed streams, which were not statistically different.

Benthic algal chlorophyll a (chl-a; median = 40.8 mg/m^2; range = 11.2-148 mg/m^2) and ash free dry mass (AFDM; range=4.50 – 55.7 g/m^2) were measured as well as phytoplankton chl-a. Positive relationships existed between TN and benthic algae (R^2 = 0.26) and between NO_3-NO_2-N and benthic algae (R^2 = 0.26). A significant negative relationship existed between TP and AFDM (R^2 = 0.38). Phytoplankton chl-a ranged from 0.70 to 6.3 mg/L and was best explained by a multiple regression model including nitrate-nitrite concentrations (positive effect) and discharge (negative effect) (R^2 = 0.37). Macroalgal cover across sites was positively related to TP (R^2 = 0.36). Microalgal cover was negatively related to TP (R^2 = 0.28).

Mean diel DO concentrations ranged from 4.88 to 7.62 mg/L (median = 6.28 mg/L) and minimums ranged from 2.35 to 6.86 mg/L (median = 4.61 mg/L). Minimum diel pH ranged from 7.24-8.07 (median = 7.73). Maximum diel pH ranged from 7.49-8.99 (median = 8.06). Total P was negatively correlated with diel DO minimums (r = -0.52) and positively correlated with diel DO ranges (r = 0.43). Macroalgal cover was negatively related to diel DO minimums (r = -0.47) and positively correlated with diel DO ranges (r = 0.43). Diel pH range was related to macroalgal coverage (r = 0.44) and negatively related to AFDM (r = -0.54).

Total N was the only nutrient measure correlated with benthic invertebrate ALU scores (r = 0.50). Taxa richness, EPT taxa richness, and the ratio of intolerant to tolerant taxa were positively correlated with TN. The percent scrapers and the number of non-insect taxa were positively correlated with TP. Phytoplankton chl-a was significantly related to more macroinvertebrate variables than benthic chl-a or AFDM were. The fish ALU score was positively correlated to TN (r = 0.53) and TP (r = 0.49). Fish richness and intolerant species richness was positively correlated with TN, nitrate, and TP. The only algal biomass variable correlated with a fish metric was benthic AFDM, which was negatively correlated with the fish ALU score.

Matlock, M.D. and A.D. Rodriguez. 1999. Preliminary report of findings: periphytometer study on streams in the Lake Waco/Bosque River Watershed. Texas Institute for Applied Environmental Research

The objective of this study was to assess nutrient limitation and trophic status of tributaries in the Lake Waco/Bosque River watershed in north-central Texas where dairy farming and row crop agriculture have been associated with declining water quality. Ten sites in five streams (North Bosque, Middle Bosque, Lower Bosque, Hog Creek, and Neil's Creek) with different nutrient characteristics were sampled. These different nutrient characteristics were not described in the study. Periphytometers receiving no nutrients, amended N, amended P, or both were used to measure algal nutrient limitation over from July 1997 to December 1998. This period included a severe drought (50 year return period) and a severe flood (25 year event). Periphytometers receiving a gradient of phosphorus concentrations (75, 150, 300, 500, and 800 µg phosphate/L) were deployed twice to determine the concentration that elicited a significant biological response. Chlorophyll *a* (chl-a) associated with periphyton production was used as an indicator of baseline primary productivity (no nutrients added) and maximum primary productivity (MPP; assumed to occur in the N+P treatment) in response to nutrient amendment. The MPP served to represent the rate of periphyton growth when nutrients are not limiting. The ratio of chl-a in the no nutrients added periphytometer to any nutrient amended periphytometer was named a lotic ecosystem trophic status index (LETSI) and is useful for making comparisons of stream biotic response to nutrients.

Neil's Creek, the reference sub-watershed, was P-limited in July 1997, the Middle Bosque River below Crawford and the North Bosque River at Clifton were co-limited, and the North Bosque River above and below the Stephensville wastewater treatment plant (WWTP) outfall were not limited by nutrients. The site below the WWTP and had the highest base productivity (4.58 µg cm^{-2}), whereas all other sites ranged from 0.47 to 1.73 µg cm^{-2}. The LETSI-NP ranged from 0.08-0.90, with Niel's Creek MPP around 0.18. In April, Neil's Creek remained P-limited and the Middle Bosque below Crawford was not limited by nutrients. Base primary production was higher in Neil's Creek (0.71 µg cm^{-2}) than in the Middle Bosque below Crawford (0.56 µg cm^{-2}) and the LETSI-NP value indicated that Niel's Creek was reaching a higher percentage of its MPP (70%) compared to the MPP of the Middle Bosque site (56%). In May 1998, the site below the WWTP again demonstrated the highest MPP (9.02 µg cm^{-2}) and was not nutrient limited. All other sites were P-limited (North Bosque River at Clifton and Valley Mills), co-limited (Middle Bosque River), or not limited by nutrients. Phosphorus gradient responses at the North Bosque Valley Mills sited indicated a statistically significant growth response at 800 µg phosphate/L, with a critical concentration mean of 650 µg phosphate. The phosphorus gradient periphytometers in the North Bosque at Hico elicited no response greater than the control. The final periphytometer sampling event occurred in October 1998 and the Hico site was N-limited, while others were not nutrient limited. The North Bosque below Stephenville WWTP and at Valley Mills and the Middle Bosque below Crawford had the highest MPP. In conclusion, seasonality plays a significant role in the algal response to nutrients. The WWTP site was degraded by nutrient enrichment. Nutrient assimilative capacity was highest in the late spring to early summer compared to other seasons.

Porter, T. 2010. Letter addressing Docket ID No. EPA-HQ-OW-2009-0596. Association of State and Interstate Water Pollution Control Administrators.

The Association of State and Interstate Water Pollution Control Administrators (ASIWPCA) offered several comments on the USEPA's Proposed Rule on water quality standards for nitrogen (N) and phosphorus (P) for Florida's lakes and flowing waters. The ASIWPCA suggests that nutrients as a "natural" part of any ecosystem present a unique challenge because the criteria development approach suggested by the EPA is designed to address threshold pollutants. They argue that N and P are not threshold pollutants. They also suggest that nutrients have weak concentration-response relationships. They do not believe that the nutrient criteria development process in Florida should be used as a blueprint for other states. They believe an alternative approach based on best available technology/practices that optimize reductions of point and non-point loading sources should be used. They also suggest that nutrient standards should only be applied when there is biological confirmation of an impact related to human nutrient sources and there is confidence that nutrient control is the key to use attainment.

Radloff, P.L., Contreras C., Whisenant, A., and Bronson, J.M. 2010. Nutrient effects in small Brazos Basin streams final report. Texas Parks and Wildlife Department. WQS-2010-02.

The objectives of this study were to add to the body of data examining the effects of nutrient enrichment on small streams in Texas and to increase knowledge about the distribution and status of freshwater mussels in the Brazos River Basin. They sampled six sites in North Central Texas streams 4 times for water quality, periphyton, benthic macroinvertebrates including mussels, and fishes. Habitat and flow data were also collected.

Streams all had mean dissolved oxygen (DO) greater than 5.0 mg/L and minima were greater than 3.0 mg/L except when streams were not flowing (3 sampling events) and in June 2008 the minimum DO in fell below 3.0 mg/L in Duck Creek. In ecoregion 32, specific conductance tended to increase from May 07 to August 08, but the same trend was not observed in ecoregion 33. The Habitat Quality Index rated all sites as either intermediate (Little Elm and Tributary of Little Elm Creeks) or high (Willis, Clear, Duck and Walnut Creeks). Nitrate levels were higher in ecoregion 32 than in ecoregion 33 streams and consistently exceeded criterions. Total phosphorous (TP) levels were high and exceeding criteria in Tributary to Little Elm, which receives wastewater treatment plant effluent, and Willis Creek.

Suspended chlorophyll a (chl-a) only exceeded TCEQ screening levels on two occasions in Willis Creek. Periphyton chl-a and ash free dry mass values ranged from 8.4-39 mg m^{-2} and 0.72-1.6 mg m^{-2}, which are below nuisance levels. High values were associated with Tributary of Little Elm Creek. ANOSIM indicated that diatom communities differed between streams in ecoregion 32 and 33 and Tributary of Little Elm Creek had the highest percentage of tolerant and eutrophic taxa and the lowest percentage of sensitive taxa. Aquatic vegetation surveys suggested that cover and thickness were low. Macro- and micro-algal composite scores were a third of the maximum score.

Statewide IBI scores for macroinvertebrates collected using kick-net or snag/woody debris collections indicated that Willis and Little Elm Creeks had limited aquatic life use (ALU) in May 2007. Clear and Walnut Creeks received exceptional ALU scores in May 2007 and Willis Creek received an exceptional ALU score in July 2008. All other sites and events were rated intermediate or high. An average of the four sampling events at each site indicated that Little Elm and its tributary rated intermediate and Willis, Clear, Duck, and Walnut Creeks rated high. Mussels sampled using timed, random searches were all dead or recently dead. Willis Creek had the highest richness with 9 species, but most sites had 3 or fewer species.

Fish collected by seining or electroshocking were assigned regionalized IBI scores. Only one collection effort at Willis Creek in May 2007 resulted in a rating of limited ALU. All other collections received an intermediate or high ALU score. When all 4 sampling events were averaged within each stream, Little Elm and its tributary, Duck Creek, and Walnut Creek received intermediate scores and Willis and Clear Creeks received a high score.

Stream flow differed between the two ecoregions and influenced biological communities indicating that nutrient criteria for wadeable streams need to account for this difference.

Stoner, N.K. 2011. Memorandum: Working in partnership with states to address phosphorus and nitrogen pollution through use of a framework for state nutrient reductions. United States Environmental Protection Agency.

The US EPA reiterated its commitment to the partnership between states and stakeholders to make greater progress in reducing nutrient [nitrogen (N) and phosphorus (P)] loadings to our nation's waters. Nutrient enrichment of US waters has increased dramatically over the last 50 years. The EPA understands the need for flexibility in this collaborative effort but provides a framework to use as a planning tool and suggests that the timetable for nutrient criteria is flexible as long as the state is making near-term reductions in nutrient loadings to state waters while the criteria are being developed. The framework advises to 1) prioritize watersheds on a statewide basis for N and P loading reductions, 2) set watershed loading reduction goals based upon best available information, 3) ensure effectiveness of point source permits in targeted/priority sub-watersheds, 4) use tools to accelerate the adoption of agricultural conservation practices, 5) identify how the state will assure nutrient reductions from developed communities not covered by the Municipal Separate Storm Sewer Systems program, 6) identify accountability and verification measures for framework objectives 3-5, 7) report implementation activities annually and load reductions and environmental impacts associated with management activities in target watersheds biannually, and 8) develop a work plan and schedule for numeric criteria development.

Chapter 2: Database Development, Median Calculations and Frequency Distributions

Leslie B. Massey[1] and Brian E. Haggard[2]

[1]Project Manager, Arkansas Water Resources Center, UA Division of Agriculture
[2]Professor and Director, Arkansas Water Resources Center, UA Division of Agriculture
203 Engineering Hall, Fayetteville, Arkansas, 72701
Corresponding Author: lbartsc@uark.edu

EXECUTIVE SUMMARY: The acquisition and compilation of geospatial, water quality and bioassessment data from 2,482 stations spanning 23 watersheds in the State of Texas was conducted to support the development of nutrient criteria by the Texas Commission on Environmental Quality. Following the reorganization and reduction of the data, medians were calculated for each station from the data collected from 2000 to 2010 and compiled into an overall median database to be used in advanced statistical analysis. The primary parameters of concern were identified as total phosphorus (TP), total nitrogen (TN), ortho-phosphate (PO_4-P), nitrate plus nitrite-nitrogen (NO_x-N) and chlorophyll-a (chl-a). Frequency distributions (minimum, 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles, and maximum) were calculated at multiple spatial scales including basin, level III ecoregion, level IV ecoregion and basin by level III ecoregion. The data analyzed for this study represented the general nutrient population, and therefore the 25^{th} percentile distribution of the medians was used for simple, numeric comparisons. Between basins the range in the 25^{th} percentile median concentration for TP, TN and chl-a was 0.05-0.30 mg/L, 0.40-4.70 mg/L, and 3.0-30.5 µg/L, respectively. At the larger Level III ecoregion scale the range of the 25^{th} percentile median concentration for total phosphorus, total nitrogen and chlorophyll-a was 0.03-0.171 mg/L, 0.31-1.19 mg/L, and 3.0-11.4 µg/L, respectively. This study showed that variations in the 25^{th} percentile median concentrations exist between basins and ecoregions and at other spatial scales, and this frequency distribution method should only be used in conjunction with other statistically valid methods of evaluating stressor-response relationship in Texas streams.

INTRODUCTION

The Clean Water Action Plan released in 1998, established a national set a nutrient criteria for the 14 aggregate ecoregions across the United States, five of which lie partly within the State of Texas. These numerical values were set for both causative (e.g., nutrients) and response (e.g., chlorophyll and transparency) variables which are associated with the prevention and assessment of eutrophic conditions in streams and reservoirs. However, local and regional influences on water quality can contribute to median concentrations that are different than what the US Environmental Protection Agency (EPA) has recommended (e.g., Smith et al. 2003). Therefore, states, tribes and others have the option of adopting the criteria set by the EPA or establishing scientifically defensive nutrient criteria for lakes and reservoirs of various spatial scales (e.g., basins and ecoregions) specific to their area of concern. Two commonly accepted statistical approaches to develop nutrient criteria include using percentile analysis of data frequency distributions discussed here and stressor-response relationships discussed in subsequent chapters.

The frequency distribution method does not require prior knowledge of individual stream conditions to set nutrient criteria; the criteria are developed relative to the population of streams and reservoirs in a specific area (e.g., state, basin or ecoregion). The EPA (2000) has suggested two different statistical methods to identify nutrient criteria based on percentile analysis of data frequency distributions. The first method establishes the 75th percentile of a distribution of reference or minimally impacted stream conditions as a criterion; the second is based upon the 25th percentile of the general condition nutrient concentration. The EPA (2000) suggests that both approaches should result in similar criterion (Figure 2-1); however, studies have shown that a comparison of criterion between approaches can be highly variable (Suplee et al. 2007 and Herlihy and Sifeneos 2008). For example, Suplee et al., 2007 showed that the 75th percentile of reference condition data can range anywhere from the 4th to the 97th percentile of the general population data (Suplee et al., 2007). In addition, the 75th percentile approach is somewhat constrained due to the limited existence of true reference condition streams. Nonetheless, the frequency distribution method is a tool that can aid states when setting nutrient criteria.

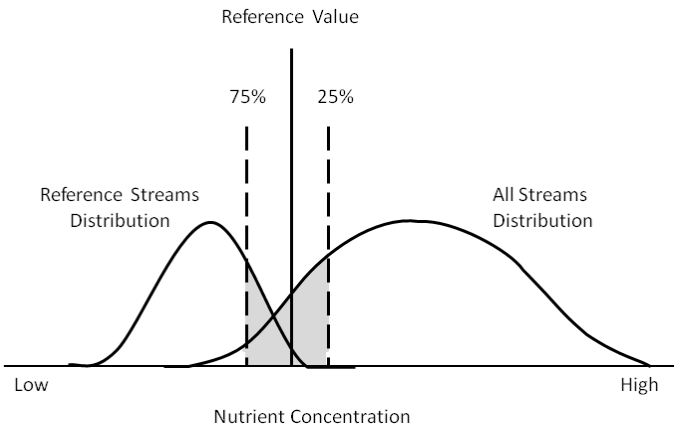


Figure 2-1. Distribution of data collected from reference condition streams and the general stream population and the associated percentile distribution used to develop nutrient criteria.

Several studies have shown inconsistencies between the EPA suggested criteria and criteria developed based on the 25th or 75th percentiles of data specific to a basin or ecoregion (Ice et al. 2003; Binkley 2004; Longing and Haggard 2010). This could be due to the fact that aggregate ecoregions are too coarse to be used for establishing nutrient criteria for all spatial scales, and the basin or smaller ecoregion level might be more appropriate for the development of nutrient criteria (Rohm et al. 2002). The objective of this chapter is to discuss the frequency distribution of median data acquired from the Texas Commission on Environmental Quality (TCEQ) at various spatial scales including individual basins, level III ecoregions, level IV ecoregions, and basin by level III ecoregion combinations.

METHODS

Geospatial Database

A geospatial database contained within a Microsoft Excel file was provided by TCEQ that identified land use and land cover data for 98 percent of the water quality stations and all but one of the bioassessment stations included in this study. The geospatial descriptors included percent open water, perennial ice or snow, low intensity residential, high intensity residential, commercial/ industry/transportation, bar rock/sand/clay, quarries/strip mines/gravel pits, transitional, deciduous forest, evergreen forest, mixed forest, shrubland, orchards/vineyards, grassland/herbaceous, pasture/hay, row crops, small grain, fallow, urban/ recreational grasses, woody wetlands, and emergent herbaceous wetlands, and drainage area and slope. These descriptors were reduced to ten categories including percent water, ice, urban (i.e., low intensity, high intensity, commercial/industrial/ transportation, and transitional), barren (i.e.,bare rock/sand/clay, and quarries/strip mines/gravel pits), forest (i.e., deciduous, evergreen, mixed and shrubland), row crop (i.e., row crop, small grains, fallow, and orchard/vineyards/other), pasture (i.e., grassland/herbaceous, pasture/hay, urban/recreational grasses), and wetland (i.e., woody and emergent herbaceous wetlands), and drainage area and slope. For the purpose of advanced statistical analysis, these categories were further reduced to percent developed (i.e., urban, barren, row crop, and pasture), percent forest, and drainage area. TCEQ also provided a separate file that identified level III and level IV ecoregion for each of the Station IDs present in the water quality database (Appendix 2-1).

Water Quality Database

Data Acquisition, Compilation and Reduction

TCEQ provided a database of water quality data collected from 1968 to 2010 from freshwater streams and rivers throughout the State of Texas. The collected data was from 2,482 stations spanning 23 watersheds (Appendix 2-2) and was divided among 38 Microsoft Excel worksheets within four Microsoft Excel workbooks. The data described 171 stream characteristics and water quality parameters including nutrients, sediments, transparency, physico-chemical parameters, as well as others.

For the purposes of advanced statistical analyses conducted during this project, only data collected from 2000 to 2010 was used. Therefore, the database was sorted and any data collected before calendar

year 2000 was removed. The data received from TCEQ were output to a single column format within the files, so the data were reorganized into a useable format. The data was sorted by Basin ID and a new Microsoft Excel worksheet was created for each individual basin. Each basin worksheet was then restructured using the pivot table function in Microsoft Excel so that each parameter and the associated data were unique to an individual column. Any estimated data points (i.e., those reported with a < or >) were flagged and used in the database without the associated qualifying sign.

Several additional parameters were calculated from the original data provided. Where species of nitrogen were not measured directly, a value for total nitrite-N (NO_2-N), total nitrate-N (NO_3-N), nitrite plus nitrate-N (NO_x-N), total Kjeldahl N (TKN), and total N (TN) were calculated if the necessary N species were measured instead. In addition, dissolved oxygen (DO) flux (i.e., 24 hour maximum minus 24 hour minimum) was calculated, and the two methods of chlorophyll analysis were merged into one parameter by averaging the values for the two methods (i.e., spectrophotometric and fluorometric). In most cases there was little to no overlap between the two methods, and the merged data was included in the database as a separate parameter in addition to the other two measures of chlorophyll-a (chl-a).

Due to the volume of data provided, several parameters were removed from the median database. Parameters were excluded due to lack of data, duplication of parameters, and or because TCEQ project staff identified the parameter as non-vital for the focus of this analysis (e.g., legacy historical parameters no longer used, parameters with unclear analysis method, etc.).

Median and Frequency Distribution Calculations

Median values of each parameter were calculated for each Station ID. Median values were calculated based on at least 10 data points, i.e. no medians were calculated if less than 10 data points were available for a given parameter at a given station. The calculated medians for each Station ID were then compiled into one database, and this database was merged with the GIS database according to the unique Station ID number.

Frequency distributions (minimum value, 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th} percentiles and maximum value) for water quality parameters TP (TCEQ parameter code 00665), TN (calculated parameter code 00600C; measured TN or TN calculated from other measured N species), NO_x-N (calculated parameter 00620C; average NO_x-N measured from multiple analytical methods), PO_4-P (TCEQ parameter code 00671), and chl-a (TCEQ parameter code 70953) were calculated using Microsoft Excel. Frequency distributions were calculated for multiple spatial scales including basin, level III ecoregion, level IV ecoregion and basin by level III ecoregion (i.e., unique combinations of basin and level III ecoregions combined).

Data Quality Assurance and Control

Data quality checks were employed frequently throughout the database reorganization and data calculation processes. The original source files were maintained in unmanipulated form, and subsequent changes to each database were saved under unique file names. Data transferred from one file to the next were always checked for accuracy by comparing first and last rows and the row count

between files. In addition, when calculations were preformed, including median and frequency distribution calculations, at least 10 percent of calculations were checked for accuracy following the secondary data quality assurance project plan (QAPP).

RESULTS AND DISCUSSION

Database Management

After the databases were reorganized and reduced, a total of 69 water quality, physico-chemical and stream characteristic parameters remained in the water quality and bioassessment databases. These reduced databases were used to calculate median values and or frequency distributions and to conduct initial statistical analysis. The main parameters of interest for this project were identified as the causal variables TP, TN, NO_x-N, PO_4-P and the response variable chl-a. The frequency distribution for these parameters across the spatial scales basin and level III ecoregion are provided in Tables 2-1 and 2-2; the distributions for level IV ecoregion and basin by level III ecoregion are provided in Appendices 2-3 and 2-4, respectively, but are not specifically discussed within the text. These data summaries include count (i.e., number of median values per spatial classification), minimum and maximum values and percentiles (10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th}). This database represents a general nutrient population, because streams were not classified according to the degree of impact.

Frequency Distribution by Basin

The State of Texas is divided into 23 river basins (Basins 1-23) and one coastal basin (Basin 24; Appendix 2-2). River basin waters are the surface inland waters comprising the major streams and their tributaries while coastal basin waters are surface inland waters that discharge or in some way interconnect with bays or the Gulf of Mexico. The 25^{th} percentile of the median TP concentrations was less than 0.10 mg/L at 63% of the basins in Texas, and the 25^{th} percentiles at these basins ranged from 0.05 to 0.81 mg/L. The basins with less data tended to have 25^{th} percentile median concentrations that were greater than 0.10 mg/L. For example, six of the seven basins where the 25^{th} percentile was greater than 0.10 mg/L had 33 or less contributing median data points, and the EPA recommends a minimum of 30 data points be used when analyzing frequency distributions to guide nutrient criteria development (EPA, 2000). The 25^{th} percentile of median concentrations of PO_4-P data followed a pattern similar to that observed for TP, and the 25^{th} percentiles of the medians of these parameters were positively correlated ($R^2 = 0.57$; $p = 0.0005$). Basin 22, The Nueces-Rio Grande Costal Basin, had a 25^{th} percentile median PO_4-P concentration that was greater than 0.10 mg/L, but only 6 median data points contributed to the frequency distribution at this basin. The 25^{th} percentile PO_4-P concentrations ranged from 0.04-0.10 mg/L for the other basins.

Less TN data was available for analysis compared to other measured parameters, and the 25^{th} percentile data distribution could only be calculated for 48% of the basins. Furthermore, only three basins (i.e., Trinity River Basin (8), Brazos River Basin (12) and Colorado River Basin (14)) had more than 12 median

data points contributing to the frequency distribution of the data; the 25th percentile of median TN concentrations ranged from 0.44 to 0.84 mg/L at these three basins. The range in the 25th percentile of median TN concentrations across all basins was from 0.40 to 4.70 mg/L. The 25th percentile of median NO_x-N concentrations ranged from 0.04 to 0.21 mg/L for most of the basins, while two basins (San Antonio River Basin (19) and Nueces-Rio Grande Coastal Basin (22)) with limited contributing data (i.e., 10 or less median data points) lied outside this range. The 25th percentile of the median concentrations of TN and NO_X-N were positively correlated (R^2 = 0.97; $p < 0.0001$); however, this relationship was primarily driven by two data points.

The 25th percentile chl-a data distribution was calculated for 70% of the basins and ranged from 1.5-5.2 µg/L, except for the Red River Basin (2) and the Nueces-Rio Grande Coastal Basin (22) which exhibited 25th percentile concentrations that were greater than that observed for the other basins. Similar to the patterns observed for the frequency distribution of the other parameters, these basins also had fewer than 10 median data points from which the distribution was calculated. The 25th percentile of the median concentrations of chl-a were positively correlated to nutrients (TP: R^2 = 0.48; p = 0.004; PO_4-P: R^2 = 0.29; p = 0.037), especially to nitrogen (TN: R^2 = 0.70; p = 0.003; NO_X-N: R^2 = 0.82; $p < 0.0001$), but these relationships were driven by 25th percentile of chl-a calculated for Basin 22.

Table 2-1. Frequency distribution of median nutrient and chlorophyll-a concentrations among Basins in the State of Texas, 2000-2010.

Total Phosphorus (TP); mg/L)

Basin	Count	MIN	10th	25th	Median	75th	90th	MAX
1	18	0.060	0.067	0.081	0.095	0.143	0.231	0.542
2	66	0.020	0.050	0.060	0.115	0.219	0.845	4.200
3	15	0.060	0.074	0.110	0.202	0.464	0.900	1.680
4	32	0.023	0.060	0.080	0.103	0.150	0.276	7.150
5	28	0.060	0.060	0.064	0.105	0.133	0.175	0.245
6	68	0.060	--	0.063	0.115	0.201	--	3.260
7	6	0.110	--	0.163	0.178	0.219	--	0.345
8	109	0.029	0.060	0.060	0.080	0.195	0.972	2.880
9	2	0.160	--	--	0.580	--	--	1.000
10	156	0.050	0.103	0.150	0.828	1.271	1.830	3.280
11	33	0.035	0.092	0.150	0.220	0.540	0.734	0.930
12	133	0.040	0.060	0.060	0.100	0.290	0.940	7.430
13	10	0.065	0.125	0.173	0.193	0.217	0.241	0.340
14	118	0.010	0.020	0.060	0.060	0.079	0.287	2.235
15	1	--	--	--	0.370	--	--	--
16	9	0.090	0.154	0.190	0.210	0.245	0.305	0.305
18	75	0.007	0.021	0.050	0.050	0.065	0.278	0.880
19	68	0.019	0.057	0.064	0.150	0.842	1.416	3.195
20	2	0.071	--	--	0.660	--	--	1.240
21	38	0.002	0.050	0.060	0.065	0.134	0.160	0.490
22	13	0.090	0.182	0.300	0.730	0.920	1.068	1.420
23	63	0.004	0.052	0.060	0.095	0.214	0.392	1.420

Total Nitrogen (TN; mg/L)

Basin	Count	MIN	10th	25th	Median	75th	90th	MAX
1	1	--	--	--	0.71	--	--	--
2	12	0.29	0.33	0.40	0.56	0.86	1.66	1.72
3	0	--	--	--	--	--	--	--
4	11	0.59	0.69	0.84	0.95	1.06	1.12	1.21
5	3	0.79	--	--	1.10	--	--	1.10
6	0	--	--	--	--	--	--	--
7	3	1.02	--	--	1.38	--	--	1.60
8	50	0.48	0.74	0.84	1.15	1.58	7.83	10.04
9	0	--	--	--	--	--	--	--
10	3	1.20	--	--	1.45	--	--	1.80
11	1	--	--	--	1.90	--	--	--
12	58	0.26	0.37	0.62	1.16	2.69	4.07	15.24
13	8	0.90	1.15	1.28	1.41	1.62	1.66	1.75
14	55	0.20	0.30	0.44	0.87	1.56	2.45	7.15
15	0	--	--	--	--	--	--	--
16	0	--	--	--	--	--	--	--
18	7	1.01	1.01	1.07	1.19	1.40	1.45	1.50
19	12	1.04	1.18	1.63	2.83	3.67	6.76	9.57
20	0	--	--	--	--	--	--	--
21	7	0.77	0.91	1.02	1.03	1.58	2.65	3.75
22	9	1.52	1.78	4.70	5.32	6.70	7.80	8.00
23	9	0.49	0.59	0.63	1.26	1.41	1.60	7.75

Nitrate plus Nitrite-Nitrogen (NO_X-N; mg/L)

Basin	Count	MIN	10th	25th	Median	75th	90th	MAX
1	12	0.02	0.04	0.08	0.20	2.25	6.63	11.10
2	39	0.04	0.04	0.06	0.09	0.32	1.79	9.34
3	2	0.38	--	--	4.49	--	--	8.60
4	13	0.05	0.05	0.07	0.21	0.41	0.55	7.79
5	31	0.04	0.05	0.08	0.13	0.17	0.23	0.74
6	49	0.04	0.05	0.06	0.12	0.44	0.99	9.01
7	4	0.06	--	0.08	0.08	0.11	--	0.17
8	62	0.04	0.09	0.21	0.40	0.87	6.03	8.72
9	2	0.04	--	--	0.52	--	--	1.00
10	27	0.04	0.04	0.04	0.04	0.33	1.18	12.30
11	39	0.02	0.04	0.07	0.18	1.04	1.87	3.92
12	51	0.03	0.04	0.05	0.19	0.44	1.22	14.90
13	8	0.04	0.11	0.19	0.27	0.45	0.47	0.49
14	134	0.02	0.02	0.07	0.26	1.11	2.02	12.70
15	1	--	--	--	1.07	--	--	--
16	7	0.08	0.13	0.19	0.23	0.28	0.50	0.81
18	47	0.03	0.13	0.21	0.43	0.76	1.34	11.70
19	8	0.18	0.30	0.69	3.14	4.15	5.61	8.57
20	1	--	--	--	2.34	--	--	-
21	23	0.02	0.04	0.10	0.17	0.61	1.92	2.75
22	10	0.23	0.51	3.13	3.82	4.32	4.71	5.20
23	35	0.04	0.07	0.13	0.26	0.39	0.68	4.40

Ortho-Phosphate (PO_4-P; mg/L)

Basin	Count	MIN	10th	25th	Median	75th	90th	MAX
1	18	0.020	0.040	0.040	0.040	0.065	0.156	0.460
2	54	0.020	0.020	0.040	0.040	0.080	0.663	3.860
3	10	0.040	0.040	0.040	0.045	0.120	0.396	0.990
4	14	0.020	0.026	0.040	0.040	0.049	0.060	1.285
5	19	0.040	0.040	0.040	0.040	0.040	0.060	0.130
6	38	0.020	0.040	0.040	0.053	0.119	0.213	2.870
7	3	0.040	--	--	0.117	--	--	0.309
8	87	0.010	0.020	0.040	0.040	0.060	0.731	1.675
9	2	0.100	--	--	0.550	--	-	1.000
10	157	0.010	0.040	0.080	0.695	1.195	1.860	3.200
11	24	0.040	0.040	0.058	0.095	0.370	0.855	1.310
12	187	0.003	0.032	0.040	0.040	0.245	1.901	8.000
13	10	0.040	0.067	0.074	0.097	0.111	0.908	8.000
14	91	0.010	0.020	0.040	0.040	0.040	0.290	4.865
15	1	--	--	--	0.265	--	--	--
16	0	--	--	--	--	--	--	--
18	25	0.020	0.030	0.040	0.040	0.040	0.172	0.965
19	65	0.010	0.020	0.101	0.951	2.540	3.306	5.140
20	1	--	--	--	1.190	--	--	--
21	22	0.013	0.022	0.040	0.040	0.040	0.151	0.340
22	6	0.040	--	0.115	0.358	0.379	--	0.384
23	30	0.006	0.007	0.040	0.040	0.110	0.351	0.340

Chlorophyll-a (Chl-a; µg/L)

Basin	Count	MIN	10th	25th	Median	75th	90th	MAX
1	6	3.00	3.23	3.94	6.50	21.4	29.7	33.5
2	9	3.00	4.82	12.6	21.5	38.2	40.7	44.3
3	13	3.00	3.00	3.36	4.28	7.26	21.0	28.4
4	17	2.00	3.00	3.00	3.30	5.00	5.12	5.54
5	0	--	--	--	--	--	--	--
6	12	3.00	3.00	3.02	7.24	11.7	20.1	22.1
7	2	15.5	--	--	26.6	--	--	37.7
8	15	3.00	3.03	5.15	10.2	13.3	17.5	20.4
9	1	--	--	--	3.00	--	--	--
10	14	3.00	3.00	3.62	5.51	8.36	9.36	10.1
11	5	3.00	--	3.00	3.00	3.39	--	15.8
12	97	3.00	3.19	3.30	5.57	13.7	26.1	72.2
13	6	0.73	--	1.47	3.01	3.44	--	5.71
14	54	3.00	3.00	5.00	5.00	9.16	28.1	76.8
15	2	0.14	1.08	2.48	4.82	7.15	8.56	9.49
16	0	--	--	--	--	--	--	--
18	12	3.00	--	--	3.00	--	--	8.63
19	7	3.00	3.00	3.00	3.00	3.00	3.12	3.30
20	1	--	--	--	5.00	--	--	--
21	23	2.00	3.00	3.00	5.00	8.35	10.3	27.8
22	5	17.3	--	30.5	37.6	77.8	--	100
23	26	3.00	3.00	3.02	7.89	18.6	30.8	100

Frequency Distributions by Level III Ecoregion

Texas is divided into 11 level III ecoregions comprised of deserts (9%), tablelands (9%), timbers (9%), plateaus (9%), prairies (9%), and plains (55%). The 25th percentiles of median TP concentrations were calculated for all of the level III ecoregions. Comparisons of these 25th percentiles were similar to the range observed by basin where most (81%) of the 25th percentile of median TP concentrations at the level III ecoregions were 0.10 mg/L or less (i.e, 0.03-0.10 mg/L). However, two ecoregions in the Texas plains, the High Plains and South Central Plains, had 25th percentiles of 0.12 mg/L and 0.17 mg/L, respectively. The 25th percentiles of the median PO_4-P concentrations were less varied and ranged from 0.02 to 0.04 mg/L across all level III ecoregions. Interestingly, the 25th percentiles of median TP and PO_4-P concentrations were not correlated across the level III ecoregion (R^2=0.19; p=0.175).

The ecoregions in the Texas Plains had the highest 25th percentile of median TN concentrations which, ranged from 0.97 mg/L to 1.19 mg/L, except for the High Plains ecoregion where data was not available for analysis. The 25th percentile of median TN concentrations were less than 0.52 mg/L at the other level III ecoregions with the lowest value observed at Edwards Plateau (0.31 mg/L). The 25th percentile distribution of NO_X-N data was calculated for the same level III ecoregions (i.e., all but High Plains where only three median data points were available). The 25th percentile of median NO_X-N concentrations ranged from 0.03 mg/L at the Southwest Table lands to 0.28 mg/L at the Texas Blackland Prairies, which is similar to the range observed at the basin level. The 25th percentiles of the medians of TN and NO_X-N calculated for level III ecoregions were not correlated (R^2=0.12; p=0.332).

The 25th percentiles of the median chlorophyll-a concentrations ranged from 3.00 to 3.30 µg/L at 64% of the level III ecoregions across Texas. Three level III ecoregions had 25th percentile median chlorophyll-a concentrations that were greater than 3.30 µg/L with the highest 25th percentile of 11.4 µg/L observed in Chihuahuan Deserts. No 25th percentile was calculated for the High Plains ecoregion, because only two median data points were available in this ecoregion. No significant correlations existed between the 25th percentile of median chl-a and nutrients at the level III ecoregion scale (TP: R^2 = 0.05; p = 0.513; PO_4-P: R^2 = 0.03; p = 0.608, TN: R^2 = 0.04; p = 0.563; NO_X-N: R^2 = 0.12; p = 0.332).

Table 2-2. Frequency distribution of median nutrient and chlorophyll-a concentrations among level III ecoregions in the State of Texas, 2000-2010.

Total Phosphorus (TP; mg/L)

Level III Ecoregion	Count	MIN	10th	25th	Median	75th	90th	MAX
24-Chihuahuan Deserts	30	0.005	0.050	0.060	0.098	0.344	0.613	0.790
25-High Plains	4	0.145	--	0.171	0.205	0.731	--	2.235
26-Southwestern Tablelands	43	0.020	0.042	0.060	0.080	0.128	0.316	1.130
27-Central Great Plains	57	0.050	0.060	0.060	0.080	0.200	0.464	1.980
29-Cross Timbers	112	0.029	0.060	0.060	0.080	0.213	0.537	1.980
30-Edwards Plateau	94	0.007	0.016	0.030	0.050	0.060	0.060	1.895
31-Southern Texas Plains	48	0.002	0.057	0.060	0.091	0.137	0.167	0.322
32-Texas Blackland Prairies	208	0.013	0.050	0.060	0.070	0.195	0.922	4.200
33-East Central Texas Plains	72	0.050	0.061	0.100	0.255	0.858	1.638	7.430
34-Western Gulf Coastal Plain	219	0.050	0.064	0.101	0.255	0.849	1.174	7.430
35-South Central Plains	176	0.050	0.073	0.118	0.348	0.897	1.202	7.430

Total Nitrogen (TN; mg/L)

Level III Ecoregion	Count	MIN	10th	25th	Median	75th	90th	MAX
24-Chihuahuan Deserts	5	0.49	--	0.63	1.20	1.26	--	1.77
25-High Plains	0	--	--	--	--	--	--	--
26-Southwestern Tablelands	10	0.33	0.36	0.42	0.56	0.70	0.85	1.72
27-Central Great Plains	4	0.29	--	0.97	1.39	1.61	--	1.71
29-Cross Timbers	51	0.26	0.35	0.52	0.99	1.51	3.48	15.24
30-Edwards Plateau	32	0.20	0.27	0.31	0.50	0.94	1.13	7.04
31-Southern Texas Plains	7	0.61	0.86	1.02	1.24	1.48	1.70	1.92
32-Texas Blackland Prairies	70	0.45	0.72	0.88	1.29	2.90	7.00	10.04
33-East Central Texas Plains	18	0.57	0.99	1.19	2.30	3.78	7.85	9.57
34-Western Gulf Coastal Plain	38	0.57	1.01	1.19	1.93	3.75	7.76	9.57
35-South Central Plains	14	0.57	1.01	1.19	1.93	3.75	7.76	9.57

Nitrate plus Nitrite-Nitrogen (NO_X-N; mg/L)

Level III Ecoregion	Count	MIN	10th	25th	Median	75th	90th	MAX
24-Chihuahuan Deserts	15	0.04	0.048	0.13	0.34	0.48	0.68	1.17
25-High Plains	3	0.10	--	--	0.47	--	--	12.70
26-Southwestern Tablelands	40	0.02	0.02	0.03	0.06	0.28	1.67	11.10
27-Central Great Plains	34	0.02	0.04	0.06	0.19	1.52	3.09	6.60
29-Cross Timbers	58	0.02	0.04	0.05	0.21	0.43	1.17	14.90
30-Edwards Plateau	87	0.02	0.03	0.08	0.21	0.47	1.26	6.77
31-Southern Texas Plains	30	0.02	0.07	0.11	0.21	0.39	0.99	2.54
32-Texas Blackland Prairies	92	0.03	0.15	0.28	0.55	1.27	4.65	11.70
33-East Central Texas Plains	34	0.02	0.05	0.17	0.39	1.65	4.35	8.57
34-Western Gulf Coastal Plain	96	0.00	0.04	0.08	0.25	1.09	3.23	5.20
35-South Central Plains	116	0.02	0.05	0.11	0.30	1.36	3.21	8.57

Phosphate (PO_4-P; mg/L)

Level III Ecoregion	Count	MIN	10th	25th	Median	75th	90th	MAX
24-Chihuahuan Deserts	19	0.006	0.007	0.040	0.040	0.140	0.412	0.580
25-High Plains	4	0.040	--	0.040	0.040	0.481	--	1.805
26-Southwestern Tablelands	44	0.020	0.020	0.038	0.040	0.046	0.198	0.750
27-Central Great Plains	43	0.020	0.040	0.040	0.040	0.078	0.763	2.720
29-Cross Timbers	120	0.003	0.017	0.040	0.040	0.068	0.384	7.250
30-Edwards Plateau	63	0.010	0.020	0.020	0.040	0.040	1.068	2.785
31-Southern Texas Plains	22	0.006	0.040	0.040	0.040	0.058	0.160	2.000
32-Texas Blackland Prairies	169	0.010	0.020	0.040	0.040	0.631	2.000	8.000
33-East Central Texas Plains	75	0.020	0.040	0.040	0.130	1.027	3.306	7.815
34-Western Gulf Coastal Plain	187	0.020	0.040	0.040	0.145	1.000	3.270	7.815
35-South Central Plains	121	0.020	0.040	0.040	0.280	1.026	3.245	7.815

Chlorophyll-a (Chl-a; mg/L)

Level III Ecoregion	Count	MIN	10th	25th	Median	75th	90th	MAX
24-Chihuahuan Deserts	16	3.00	5.02	11.4	14.8	23.8	33.0	52.6
25-High Plains	2	54.8	--	--	55.3	--	--	55.9
26-Southwestern Tablelands	11	3.00	3.45	6.50	26.0	35.2	38.2	44.3
27-Central Great Plains	19	3.00	5.28	9.81	14.9	32.5	44.7	72.2
29-Cross Timbers	67	3.00	3.30	3.30	7.20	12.6	16.8	39.9
30-Edwards Plateau	37	2.00	3.00	3.00	3.00	5.00	7.18	46.1
31-Southern Texas Plains	20	3.00	3.00	3.00	3.93	6.82	9.91	11.4
32-Texas Blackland Prairies	39	3.00	3.00	3.00	4.00	7.76	14.1	28.3
33-East Central Texas Plains	29	3.00	3.00	3.25	4.28	7.26	9.86	76.8
34-Western Gulf Coastal Plain	45	3.00	3.00	3.26	4.49	7.50	14.1	76.8
35-South Central Plains	42	3.00	3.00	3.30	5.00	8.30	12.6	76.8

The EPA has suggested nutrient criteria for TP, TN and chl-a for the 14 aggreagate ecoregions in the United States, and the five located within Texas are presented in Table 2-3. The EPA suggested nutrient criteria range from 0.01 to 0.07 mg/L of TP (EPA 2000), and most (64%) of the 25th percentiles calculated during this study at the level III ecoregions within Texas fell within the upper end of this recommended range. However, 36 percent of the 25th percentiles were greater than 0.10 mg/L, and the only aggregate ecoregion with recommended criteria above 0.10 mg/L was Aggregate Ecoregion X, which lies in the Gulf Coast portion of Texas (EPA 2000). The range of EPA suggested criteria for TN among the aggregate ecoregions in Texas was 0.12 to 0.88 mg/L (EPA 2000) compared to the range of level III ecoregions calculated in this study which ranged from 0.31 to 1.19 mg/L, where 45% of the level III ecoregion's 25th percentiles were greater than 0.88 mg/L. The 25th percentile of median chl-a concentrations were also typically greater than the range suggested for the aggregate ecoregions in Texas. The 25th percentile of median chlorophyll-a concentrations from this study ranged from 3.00 to 11.4 µg/L while the range in EPA suggested criteria was 0.93 to 3.00 µg/L. These differences highlight the fact that local and regional impacts can influence the distribution of data, and that criteria specific to an area (i.e., basin or ecoregion) should be developed to take into account variations that can occur at spatial scales smaller than the aggregate ecoregion.

Table 2-3. EPA recommended nutrient criteria for total phosphorus, total nitrogen, and chlorophyll-a for the aggregate ecoregions that in the State of Texas (EPA, 2000).

Aggregate Ecoregion	Total Phosphorus (mg/L)	Total Nitrogen (mg/L)	Chlorophyll-a (mg/L)
Aggregate Ecoregion II	0.01	0.12	1.08
Aggregate Ecoregion III	0.02	0.38	1.78
Aggregate Ecoregion IV	0.02	0.56	2.40
Aggregate Ecoregion V	0.07	0.88	3.00
Aggregate Ecoregion IX	0.04	0.69	0.93 S[1]
Aggregate Ecoregion X	0.13	0.76	2.10 S[1]

[1]Chlorophyll-a measured by spectrophotometric method with acid correction

The development of frequency distributions from median parameter concentrations is an important first step in the development of nutrient criteria, and the 25th percentile method recommended by the EPA (2000) *should* be used as a guide when setting criteria for specific basins and or ecoregions. The frequency distribution is also a good method to estimate the number of sites within a spatial scale (e.g., basins, ecoregions) that could exceed the developed criteria. However, this study as well as others (Ice et al. 2003; Binkley 2004; Longing and Haggard 2010) have shown that the 25th frequency distribution can vary from one basin or ecoregion to another and at different spatial scales. These studies have shown that 25th percentiles based on regional data often significantly differ from that developed for the aggregate ecoregions. The frequency distribution method should only be one of many tools used to support the development of numeric nutrient criteria. The Science Advisory Board (SAB) has advised the EPA that the stressor-response approach is a legitimate, scientifically based method for developing nutrient criteria when correctly applied, and this approach is the focus of the following chapters.

REFERENCES

Environmental Protection Agency (EPA). 2000. Nutrient criteria technical guidance manual: rivers and streams. EPA-822-B-00-002.

Herlihy AT, Sifneos JC. 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. Journal of the North American Benthological Society 27:932-948.

Ice G, Binkley D. 2003. Forest stormwater concentrations of nitrogen and phosphorus: a comparison with EPA's proposed water quality criteria. Journal of Forestry 101:21-28.

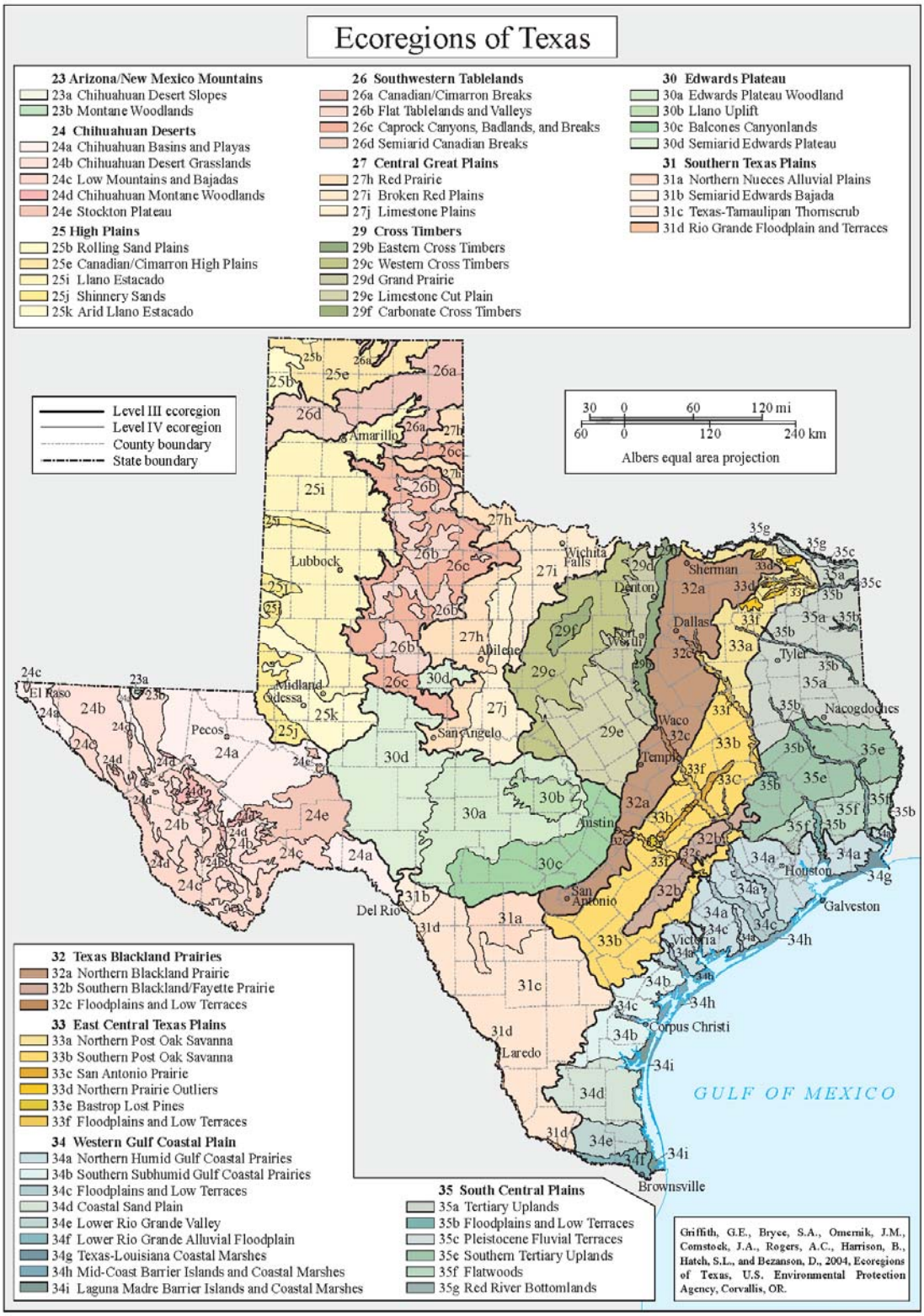
Longing SD, Haggard BE. 2010. Distributions of median nutrient and chlorophyll concentrations across the Red River Basin, USA. Journal of Environmental Quality 39:1966-1974.

Rohm CM, Omernik JM, Woods AJ, Stoddard JL. 2002. Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development. Journal of the American Water Resources Association 38:213-239.

Smith RA, Alexander RB, Schwarz GE. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous Unites States. Environmental Science and Technology 37:3039-3047.

Suplee MW, Varghese A, Cleland J. 2007. Developing nutrient criteria for streams: an evaluation of the frequency distribution method. Journal of the American Water Resources Association 43(2):453-472.

APPENDIX 2-1: Level II and Level IV Ecoregions in Texas (ftp.epa.gov/wed/ecoregions/tx/tx_eco_pg.pdf).



APPENDIX 2-2: Texas River and Coastal Basins (http://www.tceq.texas.gov/publications/gi/gi-316/gi-316_intro.html/at_download/file).

APPENDIX 2-3: Frequency Distribution of Median Nutrient and Chlorophyll-a Concentrations among Basin by Level III Ecoregions in Texas, 2000-2010.

Total Phosphorus (TP; mg/L)

Basin-Level III	Count	MIN	10th	25th	Median	75th	90th	MAX
1-25-High Plains	1	--	--	--	0.180	--	--	--
1-26-Southwestern Tablelands	16	0.060	0.073	0.084	0.095	0.139	0.258	0.542
1-35-South Central Plains	1	--	--	--	0.060	--	--	--
2-26-Southwestern Tablelands	18	0.020	0.040	0.043	0.060	0.060	0.375	1.010
2-27-Central Great Plains	26	0.050	0.060	0.071	0.138	0.235	0.506	1.155
2-29-Cross Timbers	2	0.210	--	--	0.220	--	--	0.230
2-32-Texas Blackland Prairies	8	0.060	0.074	0.133	0.238	2.608	3.710	4.200
2-33-East Central Texas Plains	5	0.105	--	0.120	0.147	0.193	--	0.930
2-35-South Central Plains	7	0.108	0.109	0.110	0.120	0.135	0.223	0.333
3-32-Texas Blackland Prairies	4	0.060	--	0.060	0.078	0.285	--	0.855
3-33-East Central Texas Plains	6	0.100	--	0.281	0.460	0.834	--	1.680
3-35-South Central Plains	5	0.120	--	0.140	0.190	0.202	--	0.382
4-33-East Central Texas Plains	2	0.100	--	--	3.625	--	--	7.150
4-35-South Central Plains	30	0.023	0.060	0.079	0.103	0.140	0.226	1.460
5-32-Texas Blackland Prairies	2	0.19			0.216			0.245
5-33-East Central Texas Plains	1	--	--	--	0.160	--	--	--
5-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
5-35-South Central Plains	25	0.060	0.060	0.060	0.100	0.125	0.140	0.20
6-33-East Central Texas Plains	0	--	--	--	--	--	--	--
6-34-Western Gulf Coastal Plain	1	--	--	--	0.130	--	--	--
6-35-South Central Plains	67	0.060	0.060	0.062	0.110	0.203	0.284	3.260
7-34-Western Gulf Coastal Plain	6	0.110	--	0.163	0.178	0.219	--	0.345
8-29-Cross Timbers	20	0.029	0.060	0.060	0.080	0.303	0.791	1.980
8-32-Texas Blackland Prairies	76	0.050	0.060	0.060	0.068	0.131	0.905	2.880
8-33-East Central Texas Plains	7	0.060	0.444	0.760	0.920	1.040	1.127	1.138
8-35-South Central Plains	6	0.060	--	0.140	0.150	0.168	--	0.870
9-34-Western Gulf Coastal Plain	2	0.160	--	--	0.580	--	--	1.000
10-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
10-34-Western Gulf Coastal Plain	121	0.060	0.130	0.270	0.920	1.325	1.860	3.280
10-35-South Central Plains	35	0.050	0.062	0.085	0.145	0.365	1.720	3.280
11-34-Western Gulf Coastal Plain	33	0.035	0.092	0.150	0.220	0.540	0.734	0.930
12-25-High Plains	2	0.145	--	--	0.188	--	--	0.230
12-26-Southwestern Tablelands	3	0.060	--	--	0.080	--	--	1.130
12-27-Central Great Plains	9	0.060	0.076	0.095	0.340	0.470	1.396	1.980
12-29-Cross Timbers	83	0.040	0.060	0.060	0.080	0.180	0.504	1.960
12-30-Edwards Plateau	2	0.060	--	--	0.060	--	--	0.060
12-32-Texas Blackland Prairies	15	0.050	0.054	0.085	0.190	0.330	1.108	1.525
12-33-East Central Texas Plains	14	0.065	0.080	0.090	0.191	1.835	3.042	7.430
12-34-Western Gulf Coastal Plain	5	0.150	--	0.272	0.290	0.780	--	1.710
13-34-Western Gulf Coastal Plain	10	0.065	0.125	0.173	0.193	0.217	0.241	0.340
14-25-High Plains	1	--	--	--	2.235	--	--	--
14-26-Southwestern Tablelands	6	0.060	--	0.063	0.095	0.139	--	0.282
14-27-Central Great Plains	22	0.060	0.060	0.060	0.060	0.069	0.088	0.131
14-29-Cross Timbers	7	0.060	0.060	0.060	0.060	0.120	0.615	1.300
14-30-Edwards Plateau	50	0.010	0.020	0.050	0.060	0.060	0.060	1.895
14-32-Texas Blackland Prairies	21	0.013	0.020	0.020	0.050	0.160	0.194	0.482
14-33-East Central Texas Plains	6	0.060	--	0.211	0.360	0.370	--	0.385
14-34-Western Gulf Coastal Plain	5	0.125	--	0.274	0.300	0.328	--	0.360
15-34-Western Gulf Coastal Plain	1	--	--	--	0.370	--	--	--
15-35-South Central Plains	0	--	--	--	--	--	--	--
16-32-Texas Blackland Prairies	2	0.090	--	--	0.168	--	--	0.245

16-33-East Central Texas Plains	2	0.200	--	--	0.215	--	--	0.230
16-34-Western Gulf Coastal Plain	5	0.170	--	0.190	0.210	0.305	--	0.305
18-30-Edwards Plateau	25	0.007	0.011	0.015	0.050	0.050	0.050	0.060
18-32-Texas Blackland Prairies	31	0.050	0.050	0.050	0.060	0.063	0.410	0.880
18-33-East Central Texas Plains	15	0.050	0.050	0.055	0.090	0.248	0.329	0.370
18-34-Western Gulf Coastal Plain	4	0.060	--	--	0.155	--	--	0.310
19-30-Edwards Plateau	6	0.019	--	0.021	0.037	0.328		1.780
19-31-Southern Texas Plains	0	--	--	--	--	--	--	--
19-32-Texas Blackland Prairies	49	0.020	0.060	0.063	0.112	0.573	1.310	3.195
19-33-East Central Texas Plains	11	0.204	0.474	0.796	0.852	0.993	1.260	2.610
19-34-Western Gulf Coastal Plain	2	0.602	--	--	0.626	--	--	0.650
20-33-East Central Texas Plains	0	--	--	--	--	--	--	--
20-34-Western Gulf Coastal Plain	2	0.071	--	--	0.656	--	--	1.240
21-30-Edwards Plateau	7	0.010	0.022	0.040	0.050	0.055	0.060	0.060
21-31-Southern Texas Plains	26	0.002	0.060	0.060	0.076	0.134	0.160	0.322
21-33-East Central Texas Plains	3	0.105	--	--	0.130	--	--	0.490
21-34-Western Gulf Coastal Plain	2	0.138	--	--	0.138	--	--	0.139
22-34-Western Gulf Coastal Plain	13	0.090	0.182	0.300	0.730	0.920	1.068	1.420
23-24-Chihuahuan Deserts	30	0.005	0.050	0.060	0.098	0.344	0.613	0.790
23-30-Edwards Plateau	4	0.060	--	--	0.060	--	--	0.060
23-31-Southern Texas Plains	22	0.040	0.051	0.063	0.100	0.140	0.167	0.248
23-34-Western Gulf Coastal Plain	7	0.074	0.078	0.085	0.228	0.250	0.264	0.270
24-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--

Total Nitrogen (TN; mg/L)

Basin-Level III	Count	MIN	10th	25th	Median	75th	90th	MAX
1-25-High Plains	0	--	--	--	--	--	--	--
1-26-Southwestern Tablelands	1	--	--	--	0.71	--	--	--
1-35-South Central Plains	0	--	--	--	--	--	--	--
2-26-Southwestern Tablelands	9	0.33	0.35	0.42	0.52	0.68	0.94	1.72
2-27-Central Great Plains	3	0.29	--	--	1.20	--	--	1.71
2-29-Cross Timbers	0	--	--	--	--	--	--	--
2-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
2-33-East Central Texas Plains	0	--	--	--	--	--	--	--
2-35-South Central Plains	0	--	--	--	--	--	--	--
3-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
3-33-East Central Texas Plains	0	--	--	--	--	--	--	--
3-35-South Central Plains	0	--	--	--	--	--	--	--
4-33-East Central Texas Plains	0	--	--	--	--	--	--	--
4-35-South Central Plains	11	0.59	0.69	0.84	0.95	1.06	1.12	1.21
5-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
5-33-East Central Texas Plains	0	--	--	--	--	--	--	--
5-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
5-35-South Central Plains	3	0.79	--	--	1.10	--	--	1.10
6-33-East Central Texas Plains	0	--	--	--	--	--	--	--
6-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
6-35-South Central Plains	0	--	--	--	--	--	--	--
7-34-Western Gulf Coastal Plain	3	1.02	--	--	1.38	--	--	1.60
8-29-Cross Timbers	6	0.48	--	1.01	1.06	1.11	--	1.14
8-32-Texas Blackland Prairies	40	0.66	0.74	0.81	1.16	1.44	7.83	10.04
8-33-East Central Texas Plains	4	3.79	--	--	7.08	--	--	8.07
8-35-South Central Plains	0	--	--	--	--	--	--	--
9-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
10-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
10-34-Western Gulf Coastal Plain	3	1.20	--	--	1.45	--	--	1.80
10-35-South Central Plains	0	--	--	--	--	--	--	--

11-34-Western Gulf Coastal Plain	1	--	--	--	1.90	--	--	--
12-25-High Plains	0	--	--	--	--	--	--	--
12-26-Southwestern Tablelands	0	--	--	--	--	--	--	--
12-27-Central Great Plains	0	--	--	--	--	--	--	--
12-29-Cross Timbers	42	0.26	0.34	0.50	0.90	1.58	3.52	15.24
12-30-Edwards Plateau	0	--	--	--	--	--	--	--
12-32-Texas Blackland Prairies	5	2.70	--	2.93	4.05	5.53	--	7.69
12-33-East Central Texas Plains	2	0.94	--	--	1.07	--	--	1.20
12-34-Western Gulf Coastal Plain	9	0.70	1.02	1.10	1.70	2.70	5.28	13.80
13-34-Western Gulf Coastal Plain	8	0.90	1.15	1.28	1.41	1.62	1.66	1.75
14-25-High Plains	0	--	--	--	--	--	--	--
14-26-Southwestern Tablelands	0	--	--	--	--	--	--	--
14-27-Central Great Plains	1	--	--	--	1.57	--	--	--
14-29-Cross Timbers	3	0.85	--	--	0.91	--	--	1.28
14-30-Edwards Plateau	27	0.20	0.26	0.30	0.43	0.80	1.02	7.04
14-32-Texas Blackland Prairies	14	0.45	0.58	0.71	1.19	1.51	4.99	7.15
14-33-East Central Texas Plains	6	0.57	--	2.02	2.30	2.48	--	2.83
14-34-Western Gulf Coastal Plain	4	1.71	--	1.76	1.82	1.86	--	1.87
15-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
15-35-South Central Plains	0	--	--	--	--	--	--	--
16-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
16-33-East Central Texas Plains	0	--	--	--	--	--	--	--
16-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
18-30-Edwards Plateau	2	1.01	--	--	1.07	--	--	1.13
18-32-Texas Blackland Prairies	2	1.39	--	--	1.45	--	--	1.50
18-33-East Central Texas Plains	3	1.02	1.05	1.10	1.19	1.30	1.37	1.41
18-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
19-30-Edwards Plateau	2	1.16	--	--	1.25	--	--	1.33
19-31-Southern Texas Plains	0	--	--	--	--	--	--	--
19-32-Texas Blackland Prairies	9	1.04	1.60	2.35	2.95	3.32	5.18	6.98
19-33-East Central Texas Plains	1	--	--	--	9.57	--	--	--
19-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
20-33-East Central Texas Plains	0	--	--	--	--	--	--	--
20-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
21-30-Edwards Plateau	1	--	--	--	0.77	--	--	--
21-31-Southern Texas Plains	4	1.02	--	1.02	1.13	1.41	--	1.92
21-33-East Central Texas Plains	2	1.01	--	--	2.38	--	--	3.75
21-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
22-34-Western Gulf Coastal Plain	9	1.52	1.78	4.70	5.32	6.70	7.80	8.00
23-24-Chihuahuan Deserts	5	0.49	--	0.63	1.20	1.26	--	1.77
23-30-Edwards Plateau	0	--	--	--	--	--	--	--
23-31-Southern Texas Plains	3	0.61	--	--	1.41	--	--	1.56
23-34-Western Gulf Coastal Plain	1	--	--	--	1.40	--	--	--
24-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--

Nitrite Plus Nitrate-Nitrogen (NO_x-N; mg/L)

Basin-Level III	Count	MIN	10th	25th	Median	75th	90th	MAX
1-25-High Plains	1	--	--	--	0.10	--	--	--
1-26-Southwestern Tablelands	11	0.02	0.04	0.07	0.25	3.13	6.82	11.10
1-35-South Central Plains	0	--	--	--	--	--	--	--
2-26-Southwestern Tablelands	15	0.04	0.05	0.06	0.14	0.32	1.16	3.47
2-27-Central Great Plains	14	0.04	0.04	0.06	0.073	0.37	1.83	5.88
2-29-Cross Timbers	0	--	--	--	--	--	--	--
2-32-Texas Blackland Prairies	3	0.07	--	--	0.45	--	--	9.34
2-33-East Central Texas Plains	4	0.05	--	0.12	0.19	0.24	--	0.26
2-35-South Central Plains	3	0.04	0.04	0.04	0.04	0.13	0.18	0.22

3-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
3-33-East Central Texas Plains	1	--	--	--	0.38	--	--	--
3-35-South Central Plains	1	--	--	--	8.60	--	--	--
4-33-East Central Texas Plains	0	--	--	--	--	--	--	--
4-35-South Central Plains	13	0.05	0.05	0.07	0.21	0.41	0.55	7.79
5-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
5-33-East Central Texas Plains	1	--	--	--	0.17	--	--	--
5-34-Western Gulf Coastal Plain	1	--	--	--	0.06	--	--	--
5-35-South Central Plains	29	0.04	0.05	0.08	0.13	0.17	0.25	0.74
6-33-East Central Texas Plains	0	--	--	--	--	--	--	--
6-34-Western Gulf Coastal Plain	1	--	--	--	0.23	--	--	--
6-35-South Central Plains	48	0.04	0.05	0.06	0.12	0.46	1.03	9.01
7-34-Western Gulf Coastal Plain	4	0.06		0.08	0.08	0.11	--	0.17
8-29-Cross Timbers	9	0.04	0.06	0.21	0.23	0.39	0.41	0.45
8-32-Texas Blackland Prairies	46	0.06	0.13	0.26	0.41	0.87	4.81	8.72
8-33-East Central Texas Plains	5	0.06	--	2.14	5.04	6.22	--	6.39
8-35-South Central Plains	2	0.09	--	--	0.14	--	--	0.19
9-34-Western Gulf Coastal Plain	2	0.04	0.136	0.28	0.52	0.76	0.904	1
10-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
10-34-Western Gulf Coastal Plain	8	0.04	0.04	0.04	0.05	0.47	0.95	1.34
10-35-South Central Plains	19	0.04	0.04	0.04	0.04	0.17	1.85	12.30
11-34-Western Gulf Coastal Plain	39	0.02	0.04	0.07	0.18	1.04	1.87	3.92
12-25-High Plains	0	--	--	--	--	--	--	--
12-26-Southwestern Tablelands	1	--	--	--	0.05	--	--	--
12-27-Central Great Plains	0	--	--	--	--	--	--	--
12-29-Cross Timbers	46	0.03	0.04	0.05	0.20	0.51	1.35	14.90
12-30-Edwards Plateau	0	--	--	--	--	--	--	--
12-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
12-33-East Central Texas Plains	2	0.14	--	--	0.15	--	--	0.17
12-34-Western Gulf Coastal Plain	2	0.19	--	--	0.21	--	--	0.22
13-34-Western Gulf Coastal Plain	8	0.04	0.11	0.19	0.27	0.45	0.47	0.49
14-25-High Plains	2	0.47	--	--	6.59	--	--	12.70
14-26-Southwestern Tablelands	13	0.02	0.02	0.02	0.02	0.03	0.06	0.08
14-27-Central Great Plains	20	0.02	0.03	0.13	0.34	1.97	3.57	6.60
14-29-Cross Timbers	3	0.02	0.05	0.09	0.16	0.29	0.36	0.41
14-30-Edwards Plateau	59	0.02	0.03	0.07	0.14	0.45	1.80	6.77
14-32-Texas Blackland Prairies	28	0.03	0.18	0.31	0.51	1.24	2.02	6.34
14-33-East Central Texas Plains	5	0.02	--	1.38	1.73	1.78	--	2.05
14-34-Western Gulf Coastal Plain	4	1.03	--	1.11	1.20	1.32	--	1.44
15-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
15-35-South Central Plains	1	--	--	--	1.07	--	--	--
16-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
16-33-East Central Texas Plains	2	0.26	--	--	0.28	--	--	0.30
16-34-Western Gulf Coastal Plain	5	0.08	--	0.16	0.21	0.23	--	0.81
18-30-Edwards Plateau	22	0.05	0.12	0.17	0.25	0.46	0.59	0.66
18-32-Texas Blackland Prairies	11	0.33	0.43	0.62	0.83	1.47	10.75	11.70
18-33-East Central Texas Plains	11	0.03	0.05	0.21	0.40	0.85	0.97	1.34
18-34-Western Gulf Coastal Plain	3	0.66	--	--	0.77	--	--	1.34
19-30-Edwards Plateau	2	0.18	--	--	0.27	--	--	0.35
19-31-Southern Texas Plains	0	--	--	--	--	--	--	--
19-32-Texas Blackland Prairies	4	0.8		2.00	3.24	4.15		4.34
19-33-East Central Texas Plains	1	--	--	--	8.57	--	--	--
19-34-Western Gulf Coastal Plain	1	--	--	--	3.88	--	--	--
20-33-East Central Texas Plains	0	--	--	--	--	--	--	--
20-34-Western Gulf Coastal Plain	1	--	--	--	2.34	--	--	--
21-30-Edwards Plateau	4	0.37	--	0.40	0.50	0.61	--	0.63
21-31-Southern Texas Plains	15	0.02	0.05	0.11	0.12	0.74	1.85	2.55

21-33-East Central Texas Plains	2	0.06	--	--	1.41	--	--	2.75
21-34-Western Gulf Coastal Plain	2	0.03	--	--	0.05	--	--	0.08
22-34-Western Gulf Coastal Plain	10	0.23	0.51	3.13	3.82	4.32	4.71	5.20
23-24-Chihuahuan Deserts	15	0.04	0.05	0.13	0.34	0.48	0.68	1.17
23-30-Edwards Plateau	0	--	--	--	--	--	--	--
23-31-Southern Texas Plains	15	0.05	0.07	0.13	0.26	0.35	0.57	0.72
23-34-Western Gulf Coastal Plain	5	0.15	--	0.15	0.18	0.50	--	0.52
24-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--

Ortho-Phosphate (PO_4-P; mg/L)

Basin-Level III	Count	MIN	10th	25th	Median	75th	90th	MAX
1-25-High Plains	1	--	--	--	0.040	--	--	--
1-26-Southwestern Tablelands	16	0.020	0.040	0.040	0.040	0.080	0.177	0.460
1-35-South Central Plains	1	--	--	--	0.040	--	--	--
2-26-Southwestern Tablelands	17	0.020	0.020	0.020	0.020	0.040	0.192	0.700
2-27-Central Great Plains	20	0.020	0.038	0.040	0.040	0.076	0.599	1.035
2-29-Cross Timbers	3	0.040	--	--	0.040	--	--	0.705
2-32-Texas Blackland Prairies	3	0.060	--	--	0.210	--	--	3.860
2-33-East Central Texas Plains	5	0.040	--	0.050	0.080	0.100	--	0.840
2-35-South Central Plains	6	0.040	--	0.040	0.040	0.070	--	0.130
3-32-Texas Blackland Prairies	3	0.040	--	--	0.040	--	--	0.990
3-33-East Central Texas Plains	4	0.040	--	0.048	0.095	0.188	--	0.330
3-35-South Central Plains	3	0.040	--	--	0.040	--	--	0.060
4-33-East Central Texas Plains	0	--	--	--	--	--	--	--
4-35-South Central Plains	14	0.020	0.026	0.040	0.040	0.049	0.060	1.285
5-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
5-33-East Central Texas Plains	1	--	--	--	0.060	--	--	--
5-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
5-35-South Central Plains	18	0.040	0.040	0.040	0.040	0.040	0.060	0.130
6-33-East Central Texas Plains	0	--	--	--	--	--	--	--
6-34-Western Gulf Coastal Plain	1	--	--	--	0.050	--	--	--
6-35-South Central Plains	37	0.020	0.040	0.040	0.056	0.120	0.217	2.870
7-34-Western Gulf Coastal Plain	3	0.040	--	--	0.117	--	--	0.309
8-29-Cross Timbers	13	0.020	0.020	0.020	0.040	0.040	0.056	0.725
8-32-Texas Blackland Prairies	64	0.010	0.030	0.040	0.040	0.040	0.788	1.675
8-33-East Central Texas Plains	6	0.020	0.030	0.115	0.393	0.610	0.708	0.750
8-35-South Central Plains	4	0.050	0.050	--	0.060	0.070	--	0.070
9-34-Western Gulf Coastal Plain	2	0.100	--	--	0.550	--	--	1.000
10-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
10-34-Western Gulf Coastal Plain	119	0.040	0.060	0.190	0.845	1.370	1.890	3.200
10-35-South Central Plains	38	0.010	0.014	0.040	0.055	0.145	0.926	2.745
11-34-Western Gulf Coastal Plain	24	0.040	0.040	0.058	0.095	0.370	0.855	1.310
12-25-High Plains	2	0.040	--	--	0.040	--	--	0.040
12-26-Southwestern Tablelands	5	0.040	--	0.040	0.040	0.040	--	0.750
12-27-Central Great Plains	11	0.040	0.040	0.040	0.100	0.825	2.705	2.720
12-29-Cross Timbers	98	0.003	0.011	0.040	0.040	0.097	0.361	7.250
12-30-Edwards Plateau	6	0.040	--	0.040	1.020	2.000	--	2.000
12-32-Texas Blackland Prairies	28	0.040	0.040	0.040	0.090	1.609	2.000	8.000
12-33-East Central Texas Plains	28	0.040	0.040	0.040	0.040	0.181	2.352	7.815
12-34-Western Gulf Coastal Plain	9	0.040	0.040	0.095	0.235	1.260	2.628	5.000
13-34-Western Gulf Coastal Plain	10	0.040	0.067	0.074	0.097	0.111	0.908	8.000
14-25-High Plains	1	--	--	--	1.805	--	--	--
14-26-Southwestern Tablelands	6	0.031	--	0.040	0.040	0.040	--	0.083
14-27-Central Great Plains	12	0.020	0.040	0.040	0.040	0.040	0.040	0.040
14-29-Cross Timbers	6	0.040	--	0.040	0.040	0.048	--	2.260

14-30-Edwards Plateau	42	0.010	0.020	0.020	0.040	0.040	0.040	2.785
14-32-Texas Blackland Prairies	11	0.010	0.010	0.020	0.040	0.107	0.208	0.440
14-33-East Central Texas Plains	8	0.040	0.051	0.216	0.309	1.478	4.820	4.865
14-34-Western Gulf Coastal Plain	5	0.040	--	0.179	0.223	0.226	--	0.274
15-34-Western Gulf Coastal Plain	1	--	--	--	0.265	--	--	--
15-35-South Central Plains	0	--	--	--	--	--	--	--
16-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
16-33-East Central Texas Plains	0	--	--	--	--	--	--	--
16-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
18-30-Edwards Plateau	2	0.040	--	--	0.040	--	--	0.040
18-32-Texas Blackland Prairies	17	0.020	0.026	0.040	0.040	0.040	0.202	0.965
18-33-East Central Texas Plains	5	0.040	--	0.040	0.062	0.120	--	0.130
18-34-Western Gulf Coastal Plain	1	--	--	--	0.040	--	--	--
19-30-Edwards Plateau	6	0.010	0.015	0.020	0.360	1.525	1.900	2.000
19-31-Southern Texas Plains	1	--	--	--	2.000	--	--	--
19-32-Texas Blackland Prairies	43	0.020	0.020	0.040	0.791	2.248	3.024	5.140
19-33-East Central Texas Plains	15	0.189	0.768	1.028	3.000	3.300	3.432	3.660
19-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
20-33-East Central Texas Plains	0	--	--	--	--	--	--	--
20-34-Western Gulf Coastal Plain	1	--	--	--	1.190	--	--	--
21-30-Edwards Plateau	4	0.013	--	0.018	0.030	0.040	--	0.040
21-31-Southern Texas Plains	15	0.020	0.040	0.040	0.040	0.040	0.120	0.215
21-33-East Central Texas Plains	3	0.040	--	--	0.070	--	--	0.340
21-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
22-34-Western Gulf Coastal Plain	6	0.04	--	0.115	0.358	0.379	--	0.384
23-24-Chihuahuan Deserts	19	0.006	0.007	0.040	0.040	0.140	0.412	0.580
23-30-Edwards Plateau	3	0.040	--	--	0.040	--	--	0.040
23-31-Southern Texas Plains	6	0.006	--	0.040	0.045	0.058	--	0.160
23-34-Western Gulf Coastal Plain	2	0.120	--	--	0.140	--	--	0.160
24-34-Western Gulf Coastal Plain	3	4.600	--	--	5.100	--	--	5.100

Chlorophyll-a (Chl-a; mg/L)

Basin-Level III	Count	MIN	10th	25th	Median	75th	90th	MAX
1-25-High Plains	0	--	--	--	--	--	--	--
1-26-Southwestern Tablelands	5	3.45	--	5.42	7.58	26.0	--	33.5
1-35-South Central Plains	1	--	--	--	3.00	--	--	--
2-26-Southwestern Tablelands	2	38.2	--	--	41.3	--	--	44.3
2-27-Central Great Plains	5	3	3.912	5.28	12.6	19.95	20.85	21.45
2-29-Cross Timbers	1	--	--	--	39.9	--	--	--
2-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
2-33-East Central Texas Plains	0	--	--	--	--	--	--	--
2-35-South Central Plains	1	--	--	--	37.9	--	--	--
3-32-Texas Blackland Prairies	2	7.02	--	--	7.88	--	--	8.73
3-33-East Central Texas Plains	6	3.00	--	3.00	3.18	4.05	--	7.23
3-35-South Central Plains	5	3.65	--	4.10	5.83	24.1	--	28.4
4-33-East Central Texas Plains	0	--	--	--	--	--	--	--
4-35-South Central Plains	17	2.00	3.00	3.00	3.30	5.00	5.12	5.54
5-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
5-33-East Central Texas Plains	0	--	--	--	--	--	--	--
5-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
5-35-South Central Plains	0	--	--	--	--	--	--	--
6-33-East Central Texas Plains	0	--	--	--	--	--	--	--
6-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
6-35-South Central Plains	12	3.00	3.00	3.02	7.24	11.7	20.1	22.1
7-34-Western Gulf Coastal Plain	2	15.5	--	--	26.6	--	--	37.7
8-29-Cross Timbers	6	3.00	--	6.42	7.89	10.9	--	15.1

8-32-Texas Blackland Prairies	6	3.08	--	5.56	10.5	11.2	--	14.9
8-33-East Central Texas Plains	1	--	--	--	3.00	--	--	--
8-35-South Central Plains	2	19.2	--	--	19.8	--	--	20.4
9-34-Western Gulf Coastal Plain	1	--	--	--	3.00	--	--	--
10-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
10-34-Western Gulf Coastal Plain	11	3.00	3.56	4.69	6.94	8.70	9.52	10.1
10-35-South Central Plains	3	3.00	--	--	3.00	--	--	3.80
11-34-Western Gulf Coastal Plain	5	3.00	--	3.00	3.00	3.39	--	15.8
12-25-High Plains	2	54.8	--	--	55.3	--	--	55.9
12-26-Southwestern Tablelands	3	3.00	--	--	11.5	--	--	36.9
12-27-Central Great Plains	8	5.27	6.97	10.4	14.4	37.8	48.8	72.2
12-29-Cross Timbers	54	3.00	3.30	3.30	5.55	12.5	15.3	31.1
12-30-Edwards Plateau	2	3.00	--	--	3.00	--	--	3.00
12-32-Texas Blackland Prairies	14	3.00	3.09	3.31	4.75	12.6	25.8	28.3
12-33-East Central Texas Plains	11	3.10	3.25	3.30	3.80	6.33	9.17	12.6
12-34-Western Gulf Coastal Plain	3	5.00	--	--	7.40	--	--	25.2
13-34-Western Gulf Coastal Plain	6	0.73	--	1.47	3.01	3.44	--	5.71
14-25-High Plains	0	--	--	--	--	--	--	--
14-26-Southwestern Tablelands	1	--	--	--	31.5	--	--	--
14-27-Central Great Plains	6	8.72	--	13.4	26.0	33.2	--	68.8
14-29-Cross Timbers	6	5.00	--	7.73	9.50	11.9	--	20.1
14-30-Edwards Plateau	24	3.00	3.00	3.00	5.00	5.00	7.44	46.1
14-32-Texas Blackland Prairies	6	3.00	--	3.88	5.00	5.00	--	5.00
14-33-East Central Texas Plains	6	5.00	--	5.00	5.00	5.08	41.0	76.8
14-34-Western Gulf Coastal Plain	5	5.00	--	5.00	5.00	8.40	--	9.62
15-34-Western Gulf Coastal Plain	1	--	--	--	9.49	--	--	--
15-35-South Central Plains	1	--	--	--	0.14	--	--	--
16-32-Texas Blackland Prairies	0	--	--	--	--	--	--	--
16-33-East Central Texas Plains	0	--	--	--	--	--	--	--
16-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
18-30-Edwards Plateau	2	3.00	--	--	3.00	--	--	3.00
18-32-Texas Blackland Prairies	6	3.00	--	3.00	3.00	3.00	3.32	3.64
18-33-East Central Texas Plains	3	3.00	--	--	7.58	--	--	8.63
18-34-Western Gulf Coastal Plain	1	--	--	--	4.50	--	--	--
19-30-Edwards Plateau	1	--	--	--	3.00	--	--	--
19-31-Southern Texas Plains	0	--	--	--	--	--	--	--
19-32-Texas Blackland Prairies	5	3.00	--	3.00	3.00	3.00	--	3.30
19-33-East Central Texas Plains	1	--	--	--	3.00	--	--	--
19-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--
20-33-East Central Texas Plains	0	--	--	--	--	--	--	--
20-34-Western Gulf Coastal Plain	1	--	--	--	5.00	--	--	--
21-30-Edwards Plateau	4	2.00	--	2.75	3.00	3.00	--	3.00
21-31-Southern Texas Plains	16	3.00	3.00	3.00	5.00	7.99	10.1	11.4
21-33-East Central Texas Plains	1	--	--	--	27.8	--	--	--
21-34-Western Gulf Coastal Plain	2	5.00	--	--	7.05	--	--	9.10
22-34-Western Gulf Coastal Plain	5	17.3	--	30.5	37.6	77.8	--	100
23-24-Chihuahuan Deserts	16	3.00	5.02	11.4	14.8	23.8	33.0	52.6
23-30-Edwards Plateau	4	3.00	--	3.00	3.00	4.27	--	8.09
23-31-Southern Texas Plains	4	3.00	--	3.00	3.00	3.02	--	3.10
23-34-Western Gulf Coastal Plain	2	4.54	--	--	4.81	--	--	5.07
24-34-Western Gulf Coastal Plain	0	--	--	--	--	--	--	--

APPENDIX 2-4. Frequency Distribution of Median Nutrient and Chlorophyll-a Concentrations among Level IV Ecoregions in Texas, 2000-2010.

Total Phosphorus (TP; mg/L)

Level IV	Count	MIN	10th	25th	Median	75th	90th	MAX
24a-Chihuahuan Basins & Playas	1	--	--	--	0.020	--	--	--
24b-Chihuahuan Desert Grssland	1	--	--	--	0.188	--	--	--
24c-Low Mountains and Bajadas	6	0.095	--	0.120	0.201	0.329	--	0.730
24e-Stockton Plateau	2	0.050	--	--	0.050	--	--	0.050
25i-Llano Estacado	4	0.145	--	0.171	0.205	0.731	--	2.235
26a-Canadian/Cimarron Breaks	14	0.060	0.063	0.080	0.088	0.115	0.269	0.542
26b-Flat Tablelands & Valleys	14	0.060	0.060	0.060	0.080	0.145	0.227	0.282
26c-Caprock Canyon/Bdlnd/Brk	20	0.020	0.040	0.048	0.060	0.083	0.785	1.130
26d-Semiarid Canadian Breaks	4	0.075	--	0.109	0.133	0.156	--	0.190
27h-Red Prairie	29	0.050	0.060	0.060	0.060	0.090	0.110	1.980
27i-Broken Red Plains	18	0.060	0.100	0.136	0.208	0.298	0.793	1.155
27j-Limestone Plains	10	0.060	0.060	0.060	0.063	0.378	0.548	1.250
29b-Eastern Cross Timbers	9	0.060	0.060	0.080	0.295	0.980	1.208	1.960
29c-Western Cross Timbers	31	0.060	0.060	0.060	0.120	0.256	0.540	1.300
29d-Grand Prairie	13	0.029	0.056	0.060	0.070	0.080	0.314	1.980
29e-Limestone Cut Plain	59	0.040	0.060	0.060	0.080	0.119	0.373	1.075
29f-Carbonate Cross Timbers	0	--	--	--	--	--	--	--
30a-Edwards Plateau Woodland	16	0.050	0.060	0.060	0.060	0.060	0.070	1.895
30b-Llano Uplift	6	0.050	--	0.060	0.060	0.060	--	0.060
30c-Balcones Canyonlands	59	0.007	0.014	0.020	0.050	0.050	0.060	1.780
30d-Semiarid Edwards Plateau	13	0.060	0.060	0.060	0.060	0.060	0.060	0.075
31a-Northern Nueces Allv Plns	12	0.002	0.051	0.060	0.060	0.060	0.060	0.070
31c-Texas-Tamaulipan Thrnsrcb	19	0.050	0.058	0.066	0.110	0.146	0.177	0.322
31d-Rio Grande Fldpln/Terrace	17	0.040	0.066	0.085	0.102	0.150	0.198	0.248
32a-Northern Blackland Prairie	193	0.013	0.050	0.060	0.065	0.180	0.878	4.200
32b-S Blackland/Fayette Prair	5	0.090	--	0.090	0.245	0.410	--	0.520
32c-Floodplains & Low Terrace	10	0.170	0.187	0.369	0.763	1.158	1.818	2.880
33a-Northern Post Oak Savanna	10	0.060	0.096	0.109	0.170	0.834	2.227	7.150
33b-Southern Post Oak Savanna	42	0.050	0.060	0.090	0.225	0.836	1.737	7.43
33c-San Antonio Prairie	4	0.08	--	0.089	0.151	0.98925	--	3.33
33d-Northern Prairie Outliers	0	--	--	--	--	--	--	--
33f-Floodplains & Low Terrace	16	0.060	0.130	0.325	0.380	0.923	1.040	1.138
34a-N Humid Gulf Cstal Prair	175	0.035	0.110	0.180	0.620	1.043	1.606	3.280
34b-S Subhumid Glf Cstl Prair	5	0.071	--	0.09	0.1385	0.18	--	1.24
34c-Floodplains & Low Terrace	18	0.060	0.136	0.183	0.282	0.355	0.616	0.780
34f-Lower Rio Grnd Allv Fldpl	18	0.074	0.087	0.231	0.500	0.763	0.988	1.42
34g-Texas-Louisiana Cstl Marsh	2	0.160	--	--	0.165	--	--	0.170
34h-Mid-Coast Barr Isl C Marsh	1	--	--	--	0.310	--	--	--
35a-Tertiary Uplands	53	0.023	0.060	0.079	0.120	0.196	0.280	1.460
35b-Floodplains & Low Terrace	44	0.060	0.060	0.078	0.108	0.140	0.167	0.382
35c-Pleistocene Flvl Terraces	2	0.202	--	--	0.268	--	--	0.333
35e-Southern Tertiary Uplands	28	0.050	0.060	0.118	0.163	0.243	0.945	3.260
35f-Flatwoods	44	0.060	0.060	0.060	0.085	0.340	1.522	3.280
35g-Red River Bottomlands	5	0.108	--	0.110	0.110	0.120	--	0.150

Total Nitrogen (TN; mg/L)

Level IV	Count	MIN	10th	25th	Median	75th	90th	MAX
24a-Chihuahuan Basins & Playas	20	0.22	0.30	0.71	0.96	1.41	1.88	2.40
24b-Chihuahuan Desert Grssland	0	--	--	--	--	--	--	--
24c-Low Mountains and Bajadas	2	1.26	--	--	1.52	--	--	1.77
24e-Stockton Plateau	0	--	--	--	--	--	--	--
25i-Llano Estacado	0	--	--	--	--	--	--	--
26a-Canadian/Cimarron Breaks	0	--	--	--	--	--	--	--
26b-Flat Tablelands & Valleys	0	--	--	--	--	--	--	--
26c-Caprock Canyon/Bdlnd/Brk	9	0.33	0.35	0.42	0.52	0.68	0.94	1.72
26d-Semiarid Canadian Breaks	1	--	--	--	0.71	--	--	--
27h-Red Prairie	1	--	--	--	0.29	--	--	--
27i-Broken Red Plains	2	1.20	--	--	1.46	--	--	1.71
27j-Limestone Plains	1	--	--	--	1.57	--	--	--
29b-Eastern Cross Timbers	2	1.02	--	--	1.06	--	--	1.11
29c-Western Cross Timbers	9	0.85	0.90	1.00	1.20	1.51	2.25	2.69
29d-Grand Prairie	3	0.48			1.09			1.14
29e-Limestone Cut Plain	37	0.26	0.34	0.48	0.81	1.54	3.66	15.24
29f-Carbonate Cross Timbers	0	--	--	--	--	--	--	--
30a-Edwards Plateau Woodland	8	0.36	0.37	0.45	0.70	0.90	0.95	1.03
30b-Llano Uplift	5	0.24	--	0.26	0.30	0.30	--	0.52
30c-Balcones Canyonlands	19	0.20	0.27	0.35	0.64	1.05	1.19	7.04
30d-Semiarid Edwards Plateau	0	--	--	--	--	--	--	--
31a-Northern Nueces Allv Plns	0	--	--	--	--	--	--	--
31c-Texas-Tamaulipan Thrnsrcb	4	1.02	--	--	1.13	--	--	1.92
31d-Rio Grande Fldpln/Terrace	3	0.61	--	--	1.41	--	--	1.56
32a-Northern Blackland Prairie	62	0.45	0.72	0.84	1.21	2.65	5.45	10.04
32b-S Blackland/Fayette Prair	2	1.39	--	--	1.45	--	--	1.50
32c-Floodplains & Low Terrace	6	1.43	--	2.95	4.66	7.51	--	8.80
33a-Northern Post Oak Savanna	0	--	--	--	--	--	--	--
33b-Southern Post Oak Savanna	8	0.94	0.99	1.01	1.30	3.06	5.50	9.57
33c-San Antonio Prairie	1	--	--	--	1.2	--	--	--
33d-Northern Prairie Outliers	0	--	--	--	--	--	--	--
33f-Floodplains & Low Terrace	9	0.57	1.66	2.28	2.53	6.39	7.82	8.07
34a-N Humid Gulf Cstal Prair	12	1.02	1.11	1.24	1.41	1.60	1.74	1.80
34b-S Subhumid Glf Cstl Prair	0	--	--	--	--	--	--	--
34c-Floodplains & Low Terrace	16	0.70	1.00	1.54	1.74	1.98	2.93	13.80
34f-Lower Rio Grnd Allv Fldpl	10	1.40	1.50	2.56	5.16	6.36	7.78	8.00
34g-Texas-Louisiana Cstl Marsh	0	--	--	--	--	--	--	--
34h-Mid-Coast Barr Isl C Marsh	0	--	--	--	--	--	--	--
35a-Tertiary Uplands	12	0.59	0.80	0.90	1.01	1.10	1.12	1.21
35b-Floodplains & Low Terrace	2	0.69	--	0.72	0.75	0.79	--	0.82
35c-Pleistocene Flvl Terraces	0	--	--	--	--	--	--	--
35e-Southern Tertiary Uplands	0	--	--	--	--	--	--	--
35f-Flatwoods	0	--	--	--	--	--	--	--
35g-Red River Bottomlands	0	--	--	--	--	--	--	--

Nitrate plus Nitrite (NO_X-N; mg/L)

Level IV	Count	MIN	10th	25th	Median	75th	90th	MAX
24a-Chihuahuan Basins & Playas	19	0.03	0.03	0.04	0.45	0.81	1.12	1.39
24b-Chihuahuan Desert Grssland	1	--	--	--	0.14	--	--	--
24c-Low Mountains and Bajadas	6	0.06	0.20	0.33	0.35	0.61	0.93	1.17
24e-Stockton Plateau	0	--	--	--	--	--	--	--
25i-Llano Estacado	3	0.10	--	--	0.47	--	--	12.70
26a-Canadian/Cimarron Breaks	10	0.02	0.04	0.04	0.12	1.08	7.25	11.1
26b-Flat Tablelands & Valleys	10	0.02	0.02	0.02	0.05	0.17	1.55	3.47
26c-Caprock Canyon/Bdlnd/Brk	20	0.02	0.02	0.02	0.06	0.10	0.36	1.48
26d-Semiarid Canadian Breaks	3	0.14	--	--	0.31	--	--	4.90
27h-Red Prairie	18	0.02	0.05	0.12	0.34	1.93	3.91	6.60
27i-Broken Red Plains	11	0.04	0.05	0.06	0.08	0.32	1.77	5.88
27j-Limestone Plains	5	0.02	--	0.03	0.14	0.77	--	2.36
29b-Eastern Cross Timbers	4	0.21	--	0.22	0.31	0.59	--	1.15
29c-Western Cross Timbers	9	0.02	0.04	0.15	0.20	0.27	0.48	0.75
29d-Grand Prairie	9	0.06	0.06	0.06	0.22	0.31	0.402	0.45
29e-Limestone Cut Plain	36	0.03	0.04	0.05	0.16	0.54	1.89	14.9
29f-Carbonate Cross Timbers	0	--	--	--	--	--	--	--
30a-Edwards Plateau Woodland	8	0.03	0.03	0.09	0.31	0.45	0.59	0.60
30b-Llano Uplift	5	0.02	--	0.02	0.02	0.03	--	0.04
30c-Balcones Canyonlands	66	0.03	0.07	0.11	0.24	0.46	0.67	6.77
30d-Semiarid Edwards Plateau	8	0.02	0.02	0.07	0.71	1.80	1.86	1.97
31a-Northern Nueces Allv Plns	2	0.92	--	--	1.73	--	--	2.55
31c-Texas-Tamaulipan Thrnsrcb	15	0.02	0.05	0.11	0.12	0.29	1.19	2.00
31d-Rio Grande Fldpln/Terrace	13	0.05	0.07	0.12	0.26	0.33	0.63	0.72
32a-Northern Blackland Prairie	82	0.03	0.15	0.27	0.49	1.23	4.02	11.7
32b-S Blackland/Fayette Prair	3	0.33	--	--	0.43	--	--	0.75
32c-Floodplains & Low Terrace	7	0.21	0.69	1.61	2.43	5.41	6.46	6.93
33a-Northern Post Oak Savanna	4	0.05	--	0.12	0.19	0.24	--	0.26
33b-Southern Post Oak Savanna	18	0.03	0.06	0.15	0.35	0.88	1.80	8.57
33c-San Antonio Prairie	1	--	--	--	0.17	--	--	--
33d-Northern Prairie Outliers	0	--	--	--	--	--	--	--
33f-Floodplains & Low Terrace	11	0.02	0.17	0.88	1.78	3.59	6.22	6.39
34a-N Humid Gulf Cstal Prair	65	0.02	0.04	0.06	0.18	0.79	1.30	3.92
34b-S Subhumid Glf Cstl Prair	2	0.03	0.26	0.61	1.19	1.76	2.11	2.34
34c-Floodplains & Low Terrace	13	0.08	0.14	0.22	0.66	1.13	1.40	3.88
34f-Lower Rio Grnd Allv Fldpl	15	0.15	0.16	0.36	2.95	3.96	4.55	5.20
34g-Texas-Louisiana Cstl Marsh	0	--	--	--	--	--	--	--
34h-Mid-Coast Barr Isl C Marsh	1	--	--	--	1.34	--	--	--
35a-Tertiary Uplands	35	0.04	0.13	0.16	0.31	0.53	0.89	7.79
35b-Floodplains & Low Terrace	27	0.05	0.05	0.05	0.08	0.13	0.24	1.07
35c-Pleistocene Flvl Terraces	2	0.04	--	--	4.32	--	--	8.60
35e-Southern Tertiary Uplands	24	0.04	0.04	0.04	0.10	0.33	1.21	9.01
35f-Flatwoods	27	0.04	0.04	0.05	0.06	0.12	0.70	12.3
35g-Red River Bottomlands	1	--	--	--	0.22	--	--	--

Ortho-Phosphate (PO4-P; mg/L)

Level IV	Count	MIN	10th	25th	Median	75th	90th	MAX
24a-Chihuahuan Basins & Playas	3	0.004	--	--	0.004	--	--	0.005
24b-Chihuahuan Desert Grssland	1	--	--	--	0.040	--	--	--
24c-Low Mountains and Bajadas	5	0.020		0.040	0.040	0.040		0.580
24e-Stockton Plateau	1	--	--	--	0.040	--	--	--
25i-Llano Estacado	4	0.040	--	0.040	0.040	0.481	--	1.805
26a-Canadian/Cimarron Breaks	14	0.040	0.040	0.040	0.040	0.065	0.193	0.460
26b-Flat Tablelands & Valleys	14	0.040	0.040	0.040	0.040	0.043	0.060	0.0832
26c-Caprock Canyon/Bdlnd/Brk	19	0.020	0.020	0.020	0.031	0.040	0.404	0.750
26d-Semiarid Canadian Breaks	4	0.020	--	0.035	0.040	0.061	--	0.125
27h-Red Prairie	19	0.020	0.040	0.040	0.040	0.040	0.613	1.980
27i-Broken Red Plains	16	0.020	0.040	0.040	0.040	0.085	0.693	1.035
27j-Limestone Plains	8	0.020	0.034	0.040	0.040	0.263	0.615	1.350
29b-Eastern Cross Timbers	11	0.020	0.040	0.040	0.350	0.760	1.515	7.250
29c-Western Cross Timbers	33	0.040	0.040	0.040	0.040	0.124	0.329	2.260
29d-Grand Prairie	21	0.009	0.020	0.020	0.040	0.040	0.040	0.170
29e-Limestone Cut Plain	53	0.003	0.006	0.018	0.040	0.090	0.313	2.870
29f-Carbonate Cross Timbers	2	0.040	--	--	0.040	--	--	0.040
30a-Edwards Plateau Woodland	17	0.040	0.040	0.040	0.040	0.040	0.488	2.785
30b-Llano Uplift	6	0.040		0.040	0.040	0.040		0.040
30c-Balcones Canyonlands	37	0.010	0.017	0.020	0.040	0.040	1.880	2.000
30d-Semiarid Edwards Plateau	3	0.040	--	--	0.040	--	--	0.040
31a-Northern Nueces Allv Plns	9	0.040	0.040	0.040	0.040	0.040	0.432	2.000
31c-Texas-Tamaulipan Thrnsrcb	9	0.020	0.036	0.040	0.040	0.060	0.171	0.215
31d-Rio Grande Fldpln/Terrace	4	0.006	--	0.039	0.055	0.085	--	0.160
32a-Northern Blackland Prairie	153	0.010	0.020	0.040	0.040	0.631	2.000	8.000
32b-S Blackland/Fayette Prair	2	0.200	--	--	0.203	--	--	0.205
32c-Floodplains & Low Terrace	14	0.040	0.040	0.048	0.224	0.746	1.344	1.675
33a-Northern Post Oak Savanna	6	0.040	0.040	0.043	0.065	0.095	0.468	0.835
33b-Southern Post Oak Savanna	44	0.020	0.040	0.040	0.265	2.640	3.449	7.815
33c-San Antonio Prairie	4	0.040	--	0.040	0.041	1.048	--	4.065
33d-Northern Prairie Outliers	0	--	--	--	--	--	--	--
33f-Floodplains & Low Terrace	21	0.040	0.040	0.004	0.140	0.330	0.445	0.750
34a-N Humid Gulf Cstal Prair	161	0.040	0.060	0.100	0.700	1.230	1.885	12.600
34b-S Subhumid Glf Cstl Prair	3	0.040			0.040			1.190
34c-Floodplains & Low Terrace	16	0.040	0.040	0.115	0.211	0.298	3.130	8.00
34f-Lower Rio Grnd Allv Fldpl	6	0.120	0.140	0.205	0.358	0.379	0.382	0.384
34g-Texas-Louisiana Cstl Marsh	1	--	--	--	0.040	--	--	--
34h-Mid-Coast Barr Isl C Marsh	0	--	--	--	--	--	--	--
35a-Tertiary Uplands	31	0.020	0.040	0.040	0.050	0.068	0.145	1.285
35b-Floodplains & Low Terrace	36	0.020	0.040	0.040	0.040	0.040	0.065	0.130
35c-Pleistocene Flvl Terraces	1	--	--	--	0.130	--	--	--
35e-Southern Tertiary Uplands	22	0.010	0.031	0.043	0.063	0.111	0.714	2.870
35f-Flatwoods	27	0.010	0.021	0.040	0.065	0.615	1.123	2.745
35g-Red River Bottomlands	4	0.040	--	0.040	0.040	0.049	--	0.075

Chlorophyll-a (Chl-a; mg/L)

Level IV	Count	MIN	10th	25th	Median	75th	90th	MAX
24a-Chihuahuan Basins & Playas	16	0.04	0.05	0.06	0.07	0.19	0.46	0.69
24b-Chihuahuan Desert Grssland	1	--	--	--	20.5	--	--	--
24c-Low Mountains and Bajadas	4	13.8	--	27.5	33.0	38.6	--	52.6
24e-Stockton Plateau	2	3.00	--	--	3.46	--	--	3.92
25i-Llano Estacado	2	54.8	--	--	55.3	--	--	55.9
26a-Canadian/Cimarron Breaks	4	5.41	--	7.04	16.8	27.9	--	33.5
26b-Flat Tablelands & Valleys	4	3.00	--	7.27	11.5	21.5	--	31.5
26c-Caprock Canyon/Bdlnd/Brk	3	36.9	--	--	38.2	--	--	44.3
26d-Semiarid Canadian Breaks	1	--	--	--	3.45	--	--	--
27h-Red Prairie	11	3.00	5.27	6.49	13.9	32.5	68.8	72.2
27i-Broken Red Plains	4	12.6	--	14.3	17.4	20.3	--	21.5
27j-Limestone Plains	4	8.72	--	10.7	24.4	37.8	--	38.7
29b-Eastern Cross Timbers	5	3.28	--	6.81	13.3	14.8	--	39.9
29c-Western Cross Timbers	24	3.30	4.32	6.76	9.13	13.0	20.3	31.1
29d-Grand Prairie	5	3.00	--	9.85	11.6	14.5	--	15.1
29e-Limestone Cut Plain	33	3.00	3.06	3.30	3.30	9.05	14.3	19.7
29f-Carbonate Cross Timbers	0	--	--	--	--	--	--	--
30a-Edwards Plateau Woodland	14	3.00	3.00	3.00	5.00	6.38	9.54	46.1
30b-Llano Uplift	5	3.00	--	5.00	5.00	5.00	--	5.00
30c-Balcones Canyonlands	14	2.00	3.00	3.00	3.00	3.00	3.00	5.00
30d-Semiarid Edwards Plateau	4	3.00	--	3.00	3.00	4.27	--	8.09
31a-Northern Nueces Allv Plns	7	3.00	3.00	3.00	3.00	3.93	5.62	7.64
31c-Texas-Tamaulipan Thrnsrcb	11	3.00	3.00	4.00	6.40	9.45	10.4	11.4
31d-Rio Grande Fldpln/Terrace	2	3.00	--	--	3.05	--	--	3.10
32a-Northern Blackland Prairie	32	3.00	3.00	3.00	3.33	5.14	10.5	21.7
32b-S Blackland/Fayette Prair	0	--	--	--	--	--	--	--
32c-Floodplains & Low Terrace	7	3.50	4.40	5.00	10.2	19.5	27.8	28.3
33a-Northern Post Oak Savanna	3	3.00	--	--	3.00	--	--	3.00
33b-Southern Post Oak Savanna	14	3.00	3.03	3.26	5.44	9.03	23.2	76.8
33c-San Antonio Prairie	3	3.80	3.83	3.87	3.94	4.32	4.54	4.69
33d-Northern Prairie Outliers	0	--	--	--	--	--	--	--
33f-Floodplains & Low Terrace	9	3.00	3.29	4.28	5.00	5.00	5.53	7.26
34a-N Humid Gulf Cstal Prair	24	0.73	3.00	3.00	5.08	8.54	9.59	37.7
34b-S Subhumid Glf Cstl Prair	4	5.00	--	5.00	41.4	83.4	--	100
34c-Floodplains & Low Terrace	11	2.64	3.46	4.75	5.00	8.75	15.8	25.2
34f-Lower Rio Grnd Allv Fldpl	5	4.54	--	5.07	17.3	30.9	--	37.6
34g-Texas-Louisiana Cstl Marsh	1	--	--	--	15.5	--	--	--
34h-Mid-Coast Barr Isl C Marsh	0	--	--	--	--	--	--	--
35a-Tertiary Uplands	19	2.00	3.00	3.00	3.06	4.98	8.63	22.1
35b-Floodplains & Low Terrace	17	0.14	3.17	3.65	6.14	12.0	21.9	28.4
35c-Pleistocene Flvl Terraces	1	--	--	--	4.1	--	--	--
35e-Southern Tertiary Uplands	1	--	--	--	3.80	--	--	--
35f-Flatwoods	3	3.00	--	--	3.00	--	--	3.00
35g-Red River Bottomlands	1	--	--	--	37.9	--	--	--

Chapter 3: Threshold Analysis to Identify Station Groupings Representing Similar Nutrient Conditions

Brian E. Haggard[1] and J. Thad Scott[2]

[1]Professor and Director, Arkansas Water Resources Center, UA Division of Agriculture,
203 Engineering Hall, Fayetteville, AR 72701
[2]Assistant Professor, Crop Soil and Environmental Sciences Department, UA Division of Agriculture,
115 Plant Science Building, Fayetteville, AR 72701
Corresponding Author: haggard@uark.edu

EXECUTIVE SUMMARY: States across the US are moving forward with the development of numeric nutrient criteria, but many states are concerned about the legitimacy of promulgating one numeric criterion across the whole state that represents multiple basins, various level III and IV ecoregions, and different land uses (e.g., forest, pasture, row crop and urban). Therefore, the objective of this chapter was to predict median total phosphorus (TP) concentrations as a function of categorical geographic factors (basin and ecoregions) and watershed attributes (catchment area, percent developed and percent forested). We conducted Categorical and Regression Tree (CART) analyses on the median database for streams and rivers (described in Chapter 2) to group stream stations by these attributes. The focus (dependent variable) of these analyses was on median TP concentrations (TCEQ parameter code 00665) across all stations, because TP represented the nutrient with the largest number of medians (based on a minimal of 10 data points per station). CART analyses were performed using the MVPART library in R 2.8.1 (http://www.r-project.org/). We required a minimum of 20 observations to be used in any single split in the CART model and that each terminal node in the model had a minimum of ten observations. The order of the primary, significant splits was: basin by level III ecoregion (R^2 = 0.29), level IV ecoregions (R^2 = 0.22), level III ecoregions (R^2 = 0.18), percent developed (R^2 = 0.11) and then percent forested. The stations were separated into two groups based on basin by level III ecoregion combinations, representing one group with a mean TP concentration of 0.191 mg/L based on 865 individual station medians ("low P" group) and the second group with a mean TP concentration of 0.771 mg/L based on 169 medians ("high P" group). The station groupings were then used to prepare median databases, which were then used in subsequent threshold analysis between nutrients and biological response variables. The three databases used in the following chapters were: (1) overall median database including all stations, (2) median database representing stations in the "low P" category, and (3) median database representing stations in the "high P" category.

INTRODUCTION

States across the US are moving forward with the development of numeric nutrient criteria, although the pace varies by state and the political, legal and environmental pressures each state is facing. The US Environmental Protection Agency (EPA) provides guidance to state agencies using the frequency distribution approach for aggregate ecoregions (EPA 2000). Rohm et al. (2002) showed that the 14 EPA aggregate ecoregions described patterns in total nitrogen (TN) and phosphorus (TP) at a large spatial scale, based on data from the National Eutrophication Survey (NES). This is a coarse spatial scale relative to jurisdictional boundaries, and many studies have shown that frequency distributions within ecoregions of individual watersheds (e.g., Longing and Haggard 2010) and states (e.g., Suplee et al. 2007; Sifneos and Herlihy 2008) are quite different from the EPA recommendations.

The development of nutrient criteria can be a costly process from the efforts needed to evaluate the physical, chemical and biological conditions of streams to that required to push the numeric criteria through promulgation. Many states are concerned about the legitimacy of promulgating one numeric criterion across the whole state that represents multiple basins, various level III and IV ecoregions, and different land uses (e.g., forest, pasture, row crop and urban); for these reasons, it might not feasible to develop numeric criteria for individual watersheds or ecoregions. Thus, states need to explore defensible approaches to aggregating stream stations into categories to assist in this process.

Regression tree models have been used to explain the variance (i.e., deviation) in nutrient concentrations of streams as a function of land use, ecoregion, and other watershed attributes. For example, Sifneos and Herlihy (2008) showed that select physico-chemical properties, runoff, elevation, ecoregion, and forest composition explained almost half of the variance in nutrient concentrations of relatively undisturbed streams in Oregon. The observation that undisturbed watershed conditions may not exist because of the effects of even minimal development (King and Baker 2010), atmospheric deposition (Flum and Nodvin 1995) and small catchment areas (Smith et al. 2003) suggests that statistical techniques need to be explored which group all streams by similar nutrient conditions. The objective of this chapter was to explore the relations between median nutrient concentrations (focusing on TP) at Texas streams and watershed attributes (both numeric and categorical), providing a defensible approach from which Texas streams could be grouped by watershed attributes.

METHODS

We conducted Categorical and Regression Tree (CART) analyses on the median database for streams and rivers (described in Chapter 2) to group stream stations by watershed attributes into similar nutrient conditions. The focus (dependent variable) of these analyses was on median TP concentrations (TCEQ parameter code 00665) across all stations, because TP represented the nutrient with the largest number of medians (based on a minimal of 10 data points per station). The watershed attributes considered included catchment area, percent developed land, percent forested, level III and IV ecoregions, basin

numbers, and basin by level III ecoregion. These factors represented the independent variables used to group the median station concentrations into similar TP conditions.

CART analysis is a means to reduce data, based on quantifying thresholds in independent variables that are correlated with shifts in the magnitude and/or variability of dependent variables. This statistical procedure can also provide hierarchical structure in independent variables, showing multiple thresholds from the same or different independent variables. CART analysis is very useful for resolving nonlinear, hierarchical, and high-order interactions among predictor variables (De'Ath and Fabricius 2000) and for detecting numerical values that lead to ecological changes (Qian and others 2003). CART models use recursive partitioning to separate data into subsets that are increasingly homogeneous; for example, subsets of data representing similar nutrient conditions. This iterative process invokes a tree-like classification that can reveal relationships that are often difficult to reconcile with conventional linear models (Urban 2002). We "pruned" CART models to generate final models that balanced accuracy within the available dataset with robustness to novel data (Urban 2002). CART models were cross-validated to determine "pruning size" (i.e., the number of predictor variables included in the model). Model cross-validations were conducted using 10 random and similarly sized subsets of our data according to the method detailed by De'ath and Fabricius (2000). The optimum tree size for each model was selected using the minimum cross-validated error rule (De'ath and Fabricius 2000).

CART analyses were performed using the MVPART library in R 2.8.1 (http://www.r-project.org/). We required a minimum of 20 observations to be used in any single split in the CART model and that each terminal node in the model has a minimum of ten observations. CART analysis is insensitive to missing data. Therefore, we did not remove observations from the data set due to missing values. However, we did require that all calculated medians have a minimum of ten observations used in calculating the median value (see Chapter 2). The appendix shows the CART models, statistics and a graphical representation of the station groupings with mean median TP concentrations for each group.

RESULTS AND DISCUSSION

First, we analyzed the median database for obvious outliers and then looked at the relation between median TP concentrations and percent developed land across all stations. This relation was significant ($P < 0.05$) showing a threshold with developed land in the order of 64%, explaining approximately 8% of the variation in median TP concentrations. Next, we evaluated a database where we excluded all stations with median TP concentrations greater than 2 mg/L and observed a similar, significant threshold in percent developed land (±1%); this resulted in a 1.5 times increase in model strength without changing the threshold. Further, we then excluded all stations within median TP concentrations exceeding 1 mg/L which resulted in a similar threshold in percent developed land, but it did not increase model strength. This strategy allowed us to reduce the database to stations with median TP concentration less than 2 mg/L and yielded the strongest predictive power in CART analyses, while also maximizing the number of medians included in the analyses. Over 1000 stations had available TP data, and these were used to identify station groupings by geographic and watershed attributes.

Based on this database (median TP concentration less than 2 mg/L), we investigated thresholds using CART and various independent variables like geographic factors (i.e., basin and ecoregions) and watershed characteristics (i.e., catchment area, percent developed and percent forested). The analysis showed that the strongest predictor of median TP concentrations was basin by level III ecoregion, explaining over 29% of the variation. CART broke the stations into two groups (Table 3-1), representing one group with a mean TP concentration of 0.191 mg/L based on 865 individual station medians (hereafter, referred to as "low P" group) and the second group with a mean TP concentration of 0.771 mg/L based on 169 medians (hereafter, "high P" group). There were 73 unique combinations of basin by level III ecoregion across Texas based on the available stations and data, and 63 were in the "low P" category and 10 were in the "high P" category. Further, analysis of these grouping should focus on the locations across Texas and potential similarities in watershed attributes.

The change point analysis showed that basin by level III ecoregion was the best, although several other attributes also resulted in the separation of this data into two groups. The order of the primary, significant splits was: basin by level III ecoregion (R^2 = 0.29), level IV ecoregions (R^2 = 0.22), level III ecoregions (R^2 = 0.18), percent developed (R^2 = 0.11) and then percent forested. As several studies have shown that catchment land use influences stream nutrient concentrations (e.g., see Buck et al. 2004; Popova et al. 2006; Haggard et al. 2007), this analysis showed that the stations could be separated into groups based on less than or greater than 65% developed land; the threshold for forested land was 29%. While land use was significant, these station median TP concentrations were separated by basin identification combined with level III ecoregions, because these categorical variables explained the largest amount of median variations in TP concentration.

The stations grouping were then used to prepare median databases, which were then used in subsequent threshold analysis between nutrients and biological response variables. The three databases used in the following chapters were: (1) overall median database including all stations, (2) median database representing stations in the "low P" category, and (3) median database representing stations in the "high P" category. Future analysis should evaluate how median N concentrations and biological response variables might be separated by CART using basins, ecoregions and other watershed attributes.

Table 3-1. Separation of individual stream and river stations by basin and level III ecoregions based on median total phosphorus (TP) concentrations; the groups represent similar TP conditions.

Mean TP = 0.191 mg/L		Mean TP = 0.771 mg/L	
Basin	Ecoregion III	Basin	Ecoregion III
1	25-High Plains	3	33-East Central Texas Plains
1	35-South Central Plains	8	33-East Central Texas Plains
1	26-Southwestern Tablelands	9	34-Western Gulf Coastal Plain
2	32-Texas Blackland Prairies	10	34-Western Gulf Coastal Plain
2	27-Central Great Plains	12	27-Central Great Plains
2	33-East Central Texas Plains	12	34-Western Gulf Coastal Plain
2	35-South Central Plains	19	33-East Central Texas Plains
2	26-Southwestern Tablelands	19	34-Western Gulf Coastal Plain
2	29-Cross Timbers	20	34-Western Gulf Coastal Plain
3	32-Texas Blackland Prairies	22	34-Western Gulf Coastal Plain

3	35-South Central Plains
4	33-East Central Texas Plains
4	35-South Central Plains
5	32-Texas Blackland Prairies
5	33-East Central Texas Plains
5	35-South Central Plains
5	34-Western Gulf Coastal Plain
6	33-East Central Texas Plains
6	35-South Central Plains
6	34-Western Gulf Coastal Plain
7	34-Western Gulf Coastal Plain
8	32-Texas Blackland Prairies
8	35-South Central Plains
8	29-Cross Timbers
10	32-Texas Blackland Prairies
10	35-South Central Plains
11	34-Western Gulf Coastal Plain
12	32-Texas Blackland Prairies
12	30-Edwards Plateau
12	33-East Central Texas Plains
12	25-High Plains
12	26-Southwestern Tablelands
12	29-Cross Timbers
13	34-Western Gulf Coastal Plain
14	32-Texas Blackland Prairies
14	27-Central Great Plains
14	30-Edwards Plateau
14	33-East Central Texas Plains
14	25-High Plains
14	26-Southwestern Tablelands
14	34-Western Gulf Coastal Plain
14	29-Cross Timbers
15	35-South Central Plains
15	34-Western Gulf Coastal Plain
16	32-Texas Blackland Prairies
16	33-East Central Texas Plains
16	34-Western Gulf Coastal Plain
18	32-Texas Blackland Prairies
18	30-Edwards Plateau
18	33-East Central Texas Plains
18	34-Western Gulf Coastal Plain
19	32-Texas Blackland Prairies
19	30-Edwards Plateau
19	31-Southern Texas Plains
20	33-East Central Texas Plains
21	30-Edwards Plateau
21	33-East Central Texas Plains
21	31-Southern Texas Plains
21	34-Western Gulf Coastal Plain
23	24-Chihuahuan Deserts
23	30-Edwards Plateau
23	31-Southern Texas Plains
23	34-Western Gulf Coastal Plain
24	34-Western Gulf Coastal Plain

REFERENCES

Buck, O., D.K. Niyogi, and C.R. Townsend. 2004. Scale–dependence of land use affects on water quality of streams in agricultural catchments. Environ. Pollut. 130:287–299.

De'ath G, Fabricus KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

Environmental Protection Agency (EPA). 2000. Nutrient criteria technical guidance manual: rivers and streams. EPA-822-B-00-002.

Flum T, Nodvin SC. 1995. Factors affecting stream water chemistry in the Great Smokey Mountains. Environmental Science and Technology 37:1575-1580.

Haggard, B.E., D.R. Smith, and K.R. Brye. 2007. Variations in stream water and sediment phosphorus among select Ozark catchments. J. Environ. Quality 36:1725–1734.

King RS, Baker ME. 2010. Considerations for analyzing ecological community thresholds in response to anthropogenic environmental gradients. Journal of the North American Benthological Society 29:998-1008.

Longing SD, Haggard BE. 2010. Distributions of median nutrient and chlorophyll concentrations across the Red River Basin, USA. Journal of Environmental Quality 39:1966-1974.

Popova, Y.A., V.G. Keyworth, B.E. Haggard, D.E. Storm, R.A. Lynch, and M.E. Payton. 2006. Stream nutrient limitation and sediment interactions in the Eucha–Spavinaw Basin. J. Soil Water Conserv. 61:105–115.

Qian SS, King RS, Richardson CJ. 2003. Two methods for the detection of environmental thresholds. Ecological Modeling 166:87-97.

Rohm CM, Omernik JM, Woods AJ, Stoddard JL. 2002. Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development. Journal of the American Water Resources Association 38:213-239.

Smith RA, Alexander RB, Schwarz GE. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous Unites States. Environmental Science and Technology 37:3039-3047.

Suplee MW, Varghese A, Cleland J. 2007. Developing nutrient criteria for streams: an evaluation of the frequency distribution method. Journal of the American Water Resources Association 43(2):453-472.

Urban, D.L. (2002) "Classification and regression trees" In: McCune, B., Grace, J.B. eds. , Analysis of Ecological Communities, MjM Software Design, Gleneden Beach, Oregon, USA, pp 221-231.

APPENDIX 3-1: CART on Nutrients versus Geography and Watershed Attributes

TP versus GEOGRAPHY (Ecoregion III, Basin by Ecoregion III, Ecoregion IV) and WATERSHED ATTRIBUTES (watershed area, %developed, %forest)

Call:
mvpart(form = tp_00665 ~ III_ecoregion_code + basin_ecoregionIII +
IV_ecoregion + area + developed + forest, data = streamslt2,
xval = 10, method = "anova", minsplit = 20, minbucket = 10)
n=1034 (1417 observations deleted due to missingness)

	CP	nsplit	rel error	xerror	xstd
1	0.29181396	0	1.000000	1.0026534	0.07859875
2	0.06347487	1	0.708186	0.7710723	0.06854989

Node number 1: 1034 observations, complexity param=0.291814
mean=0.2857219, MSE=0.1579689
left son=2 (865 obs) right son=3 (169 obs)
Primary splits:
basin_ecoregionIII splits as -LRLLRLLLLRLLLLL-LLL-LLLLLLLLLLR-RLLL-RLLLLRLLLL-LLLLLLLRLLLLLL--LLLLRLLR, improve=0.29181400, (0 missing)

IV_ecoregion splits as LLLLLLLLLLLLRLLL-LLLLLLLLLRRRL-RRLLRLLLLLLLL, improve=0.21955670, (0 missing)
III_ecoregion_code splits as LLLLRLLLLRL, improve=0.18165570, (0 missing)
developed < 64.89586 to the left, improve=0.10833560, (33 missing)
forest < 29.29489 to the right, improve=0.09850967, (33 missing)

Node number 2: 865 observations
mean=0.1908202, MSE=0.08035611

Node number 3: 169 observations
mean=0.7714615, MSE=0.2731773

APPENDIX 3-2: Changepoint on Nutrients versus Basin by Ecoregion III

Total P versus [Basin by Ecoregion III]

Call:
mvpart(form = tp_00665 ~ basin_ecoregionIII, data = streamslt2,
xval = 10, method = "anova", minsplit = 20, minbucket = 10)
n=1034 (1417 observations deleted due to missingness)

	CP	nsplit	rel error	xerror	xstd
1	0.29181396	0	1.000000	1.001174	0.07862712
2	0.02548322	1	0.708186	0.789926	0.06993250

Node number 1: 1034 observations, complexity param=0.291814
mean=0.2857219, MSE=0.1579689
left son=2 (865 obs) right son=3 (169 obs)
Primary splits:
basin_ecoregionIII splits as -LRLLRLLLLRLLLLL-LLL-LLLLLLLLLLR-RLLL-RLLLLRLLLL-LLLLLLLRLLLLLL--LLLLRLLR, improve=0.291814, (0 missing)
Node number 2: 865 observations
mean=0.1908202, MSE=0.08035611

Node number 3: 169 observations
mean=0.7714615, MSE=0.2731773

APPENDIX 3-3: Changepoint on Nutrients versus Ecoregion IV

Total P versus Ecoregion IV

Call:
mvpart(form = tp_00665 ~ IV_ecoregion, data = streamslt2, xval = 10,
method = "anova", minsplit = 20, minbucket = 10)
n=1034 (1417 observations deleted due to missingness)

	CP	nsplit	rel error	xerror	xstd
1	0.21955670	0	1.0000000	1.003109	0.07878892
2	0.01349732	1	0.7804433	0.810264	0.06792526

Node number 1: 1034 observations, complexity param=0.2195567
mean=0.2857219, MSE=0.1579689
left son=2 (767 obs) right son=3 (267 obs)
Primary splits:
IV_ecoregion splits as LLLLLLLLLLLLRLLL-LLLLLLLLLRRRL-RRLLRLLLLLLLL,
improve=0.2195567, (0 missing)

Node number 2: 767 observations
mean=0.1758423, MSE=0.07413421
Node number 3: 267 observations
mean=0.6013686, MSE=0.2644815

APPENDIX 3-4: Changepoint on Nutrients versus Ecoregion III

Total P versus Ecoregion III

Call:
mvpart(form = tp_00665 ~ III_ecoregion_code, data = streamslt2,
xval = 10, method = "anova", minsplit = 20, minbucket = 10)
n=1034 (1417 observations deleted due to missingness)

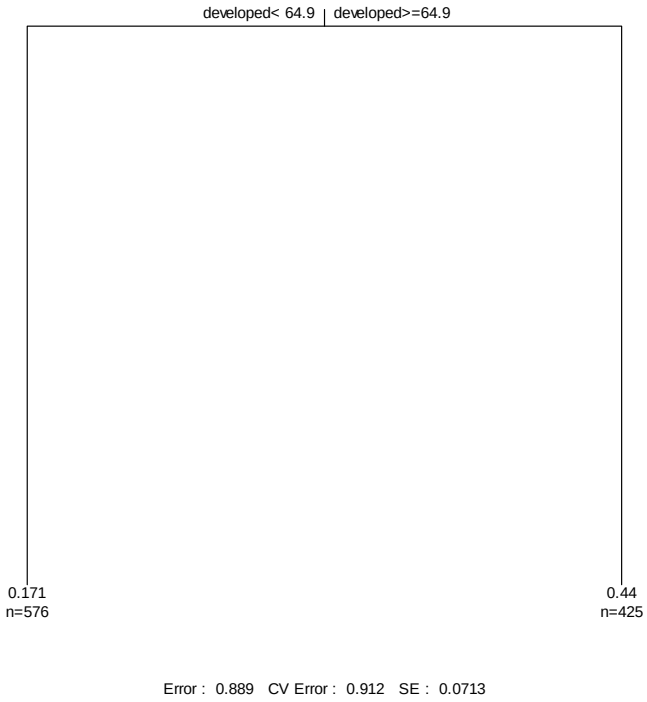
	CP	nsplit	rel error	xerror	xstd
1	0.1816557	0	1.0000000	1.005192	0.07891902
2	0.0100000	1	0.8183443	0.828467	0.06949043

Node number 1: 1034 observations, complexity param=0.1816557
mean=0.2857219, MSE=0.1579689
left son=2 (756 obs) right son=3 (278 obs)
Primary splits:
III_ecoregion_code splits as LLLLRLLLLRL, improve=0.1816557, (0 missing)

Node number 2: 756 observations
mean=0.1829979, MSE=0.08511085
Node number 3: 278 observations
mean=0.565072, MSE=0.2493685

APPENDIX 3-5: Changepoint on Nutrients versus Developed Land

Total P versus Developed Land



Call:
mvpart(form = tp_00665 ~ developed, data = streamslt2, xval = 10,
method = "anova", minsplit = 20, minbucket = 10)
n=1001 (1450 observations deleted due to missingness)

	CP	nsplit	rel error	xerror	xstd
1	0.11067635	0	1.0000000	1.0027280	0.07984827
2	0.03256173	1	0.8893237	0.9121781	0.07131606

Node number 1: 1001 observations, complexity param=0.1106763
mean=0.2850214, MSE=0.1597255
left son=2 (576 obs) right son=3 (425 obs)
Primary splits:
developed < 64.89586 to the left, improve=0.1106763, (0 missing)

Node number 2: 576 observations
mean=0.1708131, MSE=0.06781099
Node number 3: 425 observations
mean=0.4398073, MSE=0.2426602

Chapter 4: Threshold Analysis on Streams and Rivers Median Database

J. Thad Scott[1], Brian E. Haggard[2] and Leslie B. Massey[3]

[1]Assistant Professor, Crop Soil and Environmental Sciences Department, UA Division of Agriculture,
115 Plant Science Building, Fayetteville, AR 72701
[2]Professor and Director, Arkansas Water Resources Center, UA Division of Agriculture
[3]Project Manager, Arkansas Water Resources Center, UA Division of Agriculture
203 Engineering Hall, Fayetteville, Arkansas, 72701
Corresponding Author: jts004@uark.edu

EXECUTIVE SUMMARY: We used categorical and regression tree analyses on median data from Texas rivers and streams to identify changepoints which result in substantial ecological change. Identical analyses were conducted on three data sets: 1) the entire rivers and streams median data set where a minimum of ten observations, and total phosphorous (TP) concentrations less than 2 mg/L were required in calculating medians, 2) a reduced data set that included from only "Low P" basin by ecoregion combinations as defined in Chapter 3, and 3) a reduced data set that included from only "high P" basin by ecoregion combinations as defined in Chapter 3. Commonly measured biological variables were used in the analysis. Thresholds for TP, total nitrogen (TN), and nitrite plus nitrate nitrogen (NO_X-N) were derived in all analyses. TP thresholds that resulted in ecological changes ranged from 0.060 to 0.094 mg/L in the whole data set and in the "Low P" basin by ecoregion dataset. Only one TP threshold, 0.25 mg/L P, was found to be ecologically meaningful from the "high P" basin by ecoregion dataset. Similarly, TN thresholds that resulted in ecological changes ranged from 0.8 to 1.6 mg/L in the whole data set and in the "Low P" basin by ecoregion dataset. No TN thresholds were found to be useful from the "high P" basin by ecoregion dataset. Thresholds in NO_X-N generally predicted an increase in algal biomass (decreased Secchi depth, increased chlorophyll-a concentration) as inorganic nitrogen concentrations decreased. This pattern supports ecological theory, but supports the notion of using caution in assuming that increased dissolved inorganic nutrient concentrations should be related to a greater trophic state in aquatic ecosystems.

INTRODUCTION

The objective of this chapter was to identify nutrient concentrations that were correlated with a change in the magnitude or variability of commonly measured biological variables in Texas rivers and streams. The analyses conducted in this chapter focused on median water quality values in order to capture thresholds that were important at broad temporal and geographic scales.

Classification and regression tree (CART) analysis is an empirical modeling technique that is useful for identifying ecological thresholds and hierarchical structure in predictor variables (De'ath and Fabricius 2000). CART uses recursive partitioning to divide data into subsets that are increasingly homogeneous, invoking a tree-like classification that can explain relationships that may be difficult to reconcile with conventional linear models (Urban 2002). Categorical variables (e.g., station location, basin, ecoregion or land-use classifications) may also be used as independent variables in CART analysis, which provides another advantage to using CART rather than traditional regression techniques. CART and other similar methods have been used to identify thresholds and hierarchical structure in environmental correlates of various biological processes in aquatic ecosystems (King et al. 2005, East and Sharfstein 2006). King et al. (2005) used CART to specifically identify thresholds in nutrient concentrations which resulted in shifts in ecological structure and function. These thresholds were used to recommend specific water quality nutrient criteria for the Florida Everglades ecosystem.

METHODS

We conducted CART analyses on the median database for streams and rivers (described in Chapter 2) to identify thresholds in nutrient concentrations that resulted in measurable changes in common biological responses. The biological (dependent) variables included in the analyses were: median Secchi depth (m), median 24 hour dissolved oxygen (DO) flux, median chlorophyll-a (chl-a) measured with spectrophotometry, median chl-a measured with fluorometry, and median merged chl-a data (combined spectrophotometric and fluormetric measurements). The nutrient (independent) variables included in the analysis were total phosphorus (TP; TCEQ parameter code 00665), total nitrogen (TN; calculated parameter code 00600C), and nitrite plus nitrate-nitrogen (NO_X-N; calculated parameter code 00631C).

CART analysis is a form of data reduction that aims to: 1) quantify thresholds in independent variables that are correlated with shifts in the magnitude and/or variability of dependent variables, and 2) identify hierarchical structure in independent variables. CART analysis is very useful for resolving nonlinear, hierarchical, and high-order interactions among predictor variables (De'Ath and Fabricius 2000) and for detecting numerical values that lead to ecological changes (Qian et al. 2003). CART models use recursive partitioning to separate data into subsets that are increasingly homogeneous. This iterative process invokes a tree-like classification that can reveal relationships that are often difficult to reconcile with conventional linear models (Urban 2002).

CART analyses were performed using the MVPART library in R 2.8.1 (http://www.r-project.org/). We required a minimum of 20 observations to be used in any single split in the CART model and that each terminal node in the model had a minimum of ten observations. CART analysis is insensitive to missing data. Therefore, we did not remove observations from the data set due to missing values. However, we did require that all calculated medians have a minimum of ten observations used in calculating the median value. We also analyzed the median database for obvious outliers and omitted stations where median TP concentrations exceeded 2.0 mg/L. This strategy yielded the strongest predictive power in CART analyses while also maximizing the number of medians included in the analyses. We first ran CART models using all median data from the streams and rivers median database. Secondarily, we reran CART models after limiting data to stations from "low P" or "high P" basin by ecoregion III combinations (see Chapter 3 for details on this classification). Because CART analysis involves recursive partitioning, models may sometimes be over-fit (i.e. too many independent variables that decrease the statistical rigor of final model). We "pruned" CART models to generate final models that balanced accuracy within the available dataset with robustness to novel data (Urban 2002). CART models were cross-validated to determine "pruning size" (i.e., the number of predictor variables included in the model). Model cross-validations were conducted using 10 random and similarly sized subsets of our data according to the method detailed by De'ath and Fabricius (2000). The optimum tree size for each model was selected using the minimum cross-validated error rule (De'ath and Fabricius 2000).

RESULTS AND DISCUSSION

Analyses using all three datasets resulted in statistically-valid thresholds in nutrient concentrations that were correlated with biological variation (Table 4.1). TP and TN thresholds that were correlated with biological responses were similar in the CART analyses conducted on all median data and on the "Low P" basin by ecoregion III data. However, NO_X-N thresholds differed between the two data sets. NO_X-N thresholds were much less in the analysis where we used the "Low P" basin by ecoregion III dataset. The "high P" basin by ecoregion III dataset yielded few statistically valid primary splits in nutrient concentrations. However, the few thresholds that were identified were much greater nutrient concentrations than the thresholds identified using the entire dataset or the "Low P" basin by ecoregion III dataset.

Table 4-1. Nutrient thresholds identified in the primary split of CART analyses for the entire streams and rivers median data set and for data sets that were grouped according to the "Low P" and "high P" basin by ecoregion III classification.

	ALL DATA			"Low P" basin x Ecoregion III			"high P" basin x Ecoregion III		
BIOLOGICAL VARIABLE	TP (mg/L)	TN (mg/L)	NO_X-N (mg/L)	TP (mg/L)	TN (mg/L)	NO_X-N (mg/L)	TP (mg/L)	TN (mg/L)	NO_X-N (mg/L)
Secchi	0.075[1]	1.5[2]	2.9[3]	0.075[1]	1.2[2]	0.03[3]	0.25[1]	na	2.7[2]
DO Flux	0.060[1]	0.82[2]	0.10[3]	0.060[1]	0.81[2]	0.09[3]	na	na	na
Chl-a spec	0.094[1]	1.6[2]	0.02[3]	0.094[1]	1.6[3]	0.02[2]	0.77[1]	na	na
Chl-a fluor	0.068[1]	1.4[2]	0.36[3]	0.094[1]	1.4[3]	0.36[2]	0.22[1]	na	na
Chl-a merge	0.094[1]	1.6[2]	2.9[3]	0.084[1]	1.2[2]	0.02[3]	0.19[1]	na	2.7[2]

[1]Strongest predictor of biological response between the three nutrient variables
[2]Second Strongest predictor of biological response between the three nutrient variables
[3]Weakest predictor of biological response between the three nutrient variables

Although all of the thresholds reported in Table 4-1 were statistically valid according to our defined methods in CART analysis, some of the thresholds should be taken with caution because they may not represent justified ecological relationships. Rather, it is useful to show these thresholds relative to the relationship(s) between each of the independent and dependent variables. Including all these graphs in this report was not feasible because so many variables were included in the analyses. Therefore, in the following sections we graphically present CART models for each of the strongest primary (and secondary and tertiary, when applicable) predictors of biological responses in the entire dataset and the "Low P" and "high P" basin by ecoregion III datasets. Graphical presentations were also limited to CART models where the primary split explained a minimum of 15% of the variance (i.e. partial $r^2 \geq 0.15$) in the biological (dependent) variable. The complete results of all possible CART models using these datasets are available in Appendices 4-1 through 4-3.

CART on Complete Median Dataset

Several CART models were developed from the complete median dataset on Texas streams and rivers. The models presented in this section met the minimum requirements for inclusion in the mainbody of this document.

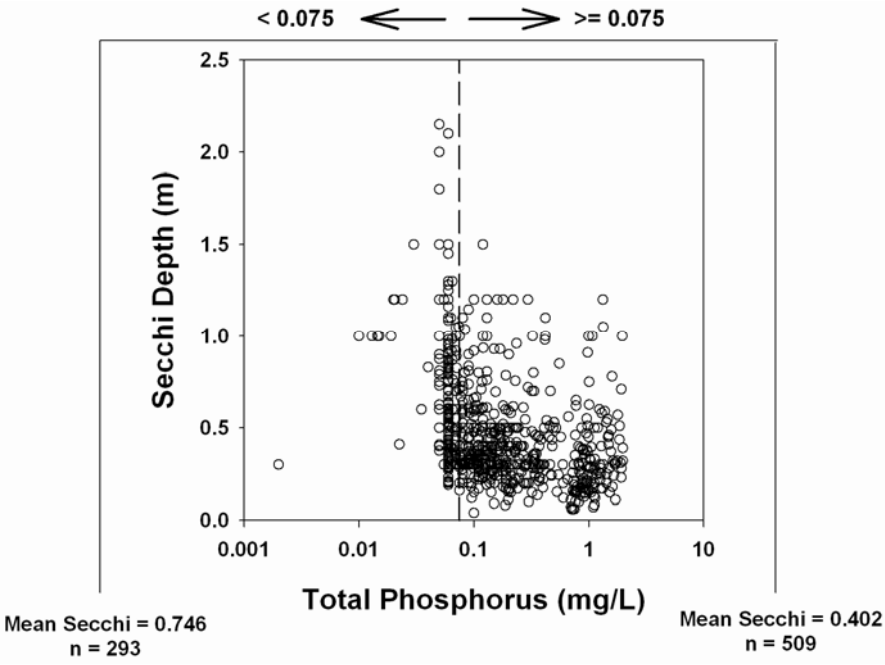


Figure 4-1.CART model for median Secchi depth versus median nutrient concentrations (TP, TN, and NO_x-N) from the entire median dataset (model $r^2 = 0.25$).

Median TP was the strongest predictor of median Secchi depth (Figure 4-1). Median Secchi depth was approximately 0.8 m when the median TP in the water column was less than 0.075 mg/L. When TP was equal to or exceeded the 0.075 mg/L threshold, median Secchi depth decreased to 0.4 m. There were no

statistically valid secondary splits in this CART model, but the threshold in median TP concentrations explained 25% of the variation in median Secchi depth across all river and stream stations included in the dataset.

Nutrient concentrations only explained 11% of the variation in the primary split on DO flux in the overall median database. This fell below our prescribed minimum for showing graphical results.

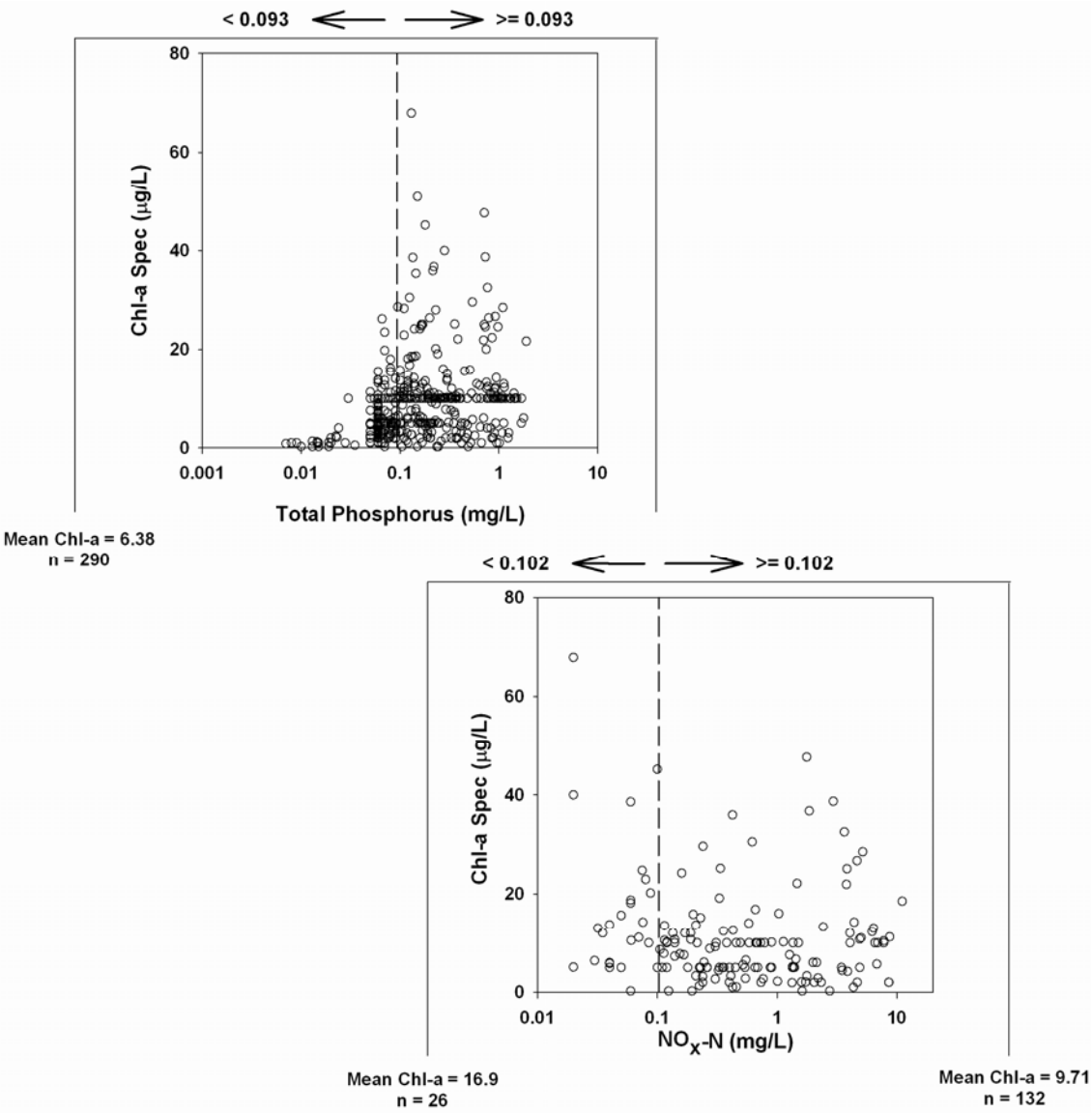


Figure 4-2. CART model for median chlorophyll-a concentrations determined by spectrophotometry versus median nutrient concentrations (TP, TN, and NO_x-N) from the entire median dataset (model r^2 = 0.37).

Median TP was the strongest predictor of median chl-a concentrations measured using spectrophotometry (Figure 4-2). Median chl-a was approximately 6.4 µg/L when the median TP in the water column was less than 0.093 mg/L. NO_X-N was an important secondary predictor when TP concentrations equaled or exceeded 0.093 mg/L. Median chl-a concentrations were greatest (~17 µg/L) when TP

equaled or exceeded the 0.093 mg/L threshold and NO_X-N was less than 0.1 mg/L. However, when TP equaled or exceeded the 0.093 mg/L threshold and NO_X-N was greater than 0.1 mg/L the median chl-a concentration was only 9.1 µg/L. This pattern makes sense ecologically in that algal biomass expressed as chl-a was greatest when TP was high, but dissolved inorganic N (NO_X-N) was low.

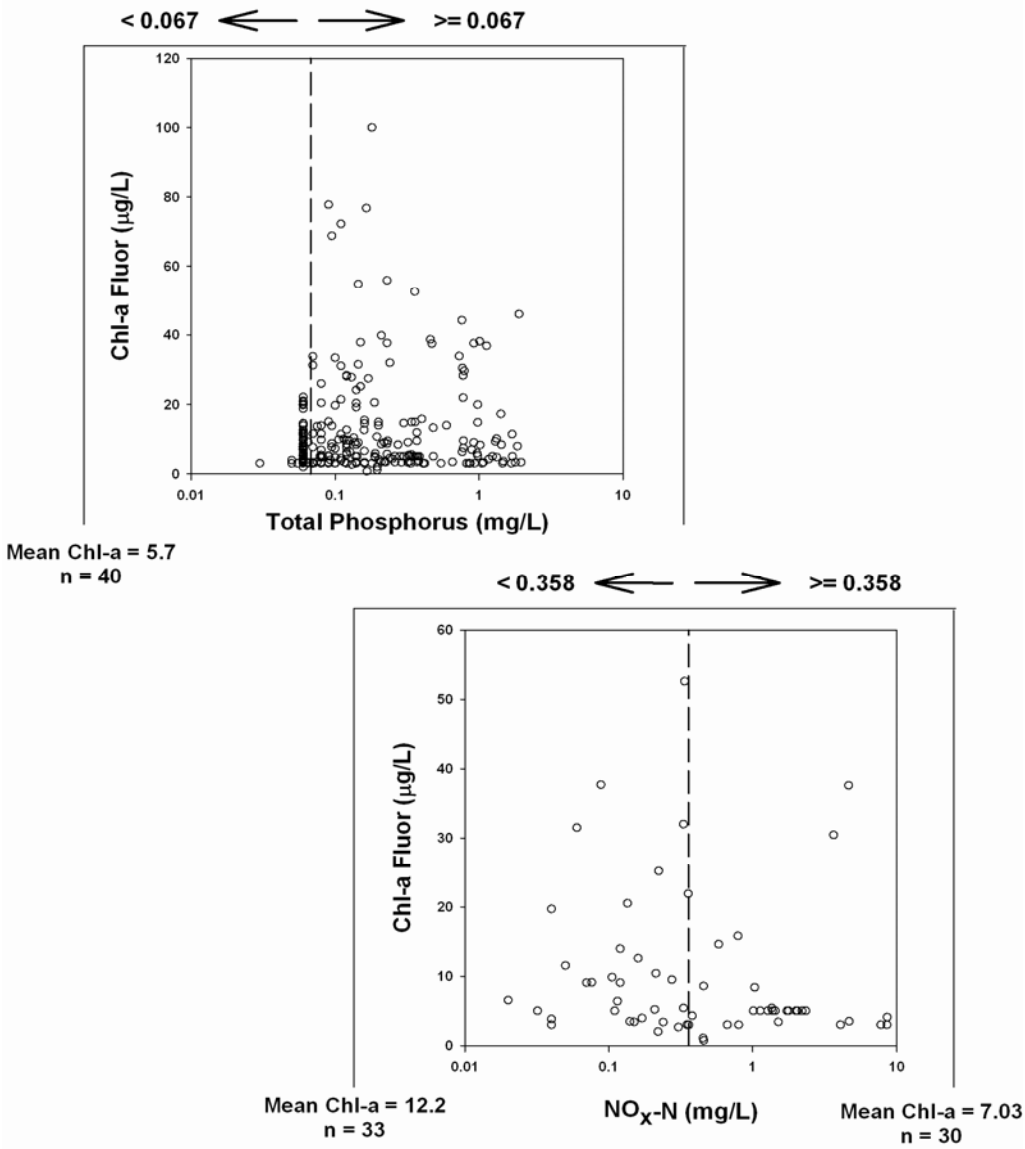


Figure 4-3. CART model for median chlorophyll-a concentrations determined by fluorometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the entire median dataset (model r^2 = 0.84).

Model structure for the CART model used to predict median chl-a concentrations from fluometric measurements (Figure 4-3) was very similar to the structure for the CART model predicting chl-a from spectrophotometric measurements (Figure 4-2). However, these models differed in their threshold values and in the median values of the reduced data. Median TP was the strongest predictor of median chl-a concentrations measured using fluorometry (Figure 4-3). Median chl-a was approximately 5.7 µg/L

when the median TP in the water column was less than 0.067 mg/L. NO_X-N was an important secondary predictor when TP concentrations equaled or exceeded 0.067 mg/L. Median chl-a concentrations were greatest (12.2 µg/L) when TP equaled or exceeded the 0.067 mg/L threshold and NO_X-N was less than 0.36 mg/L. However, when TP equaled or exceeded the 0.067 mg/L threshold and NO_X-N was greater than 0.36 mg/L the median chl-a concentration was only 7.0 µg/L. This pattern also makes sense ecologically in that algal biomass expressed as chlorophyll-a was greatest when TP was high, but inorganic N (NO_X-N) was low.

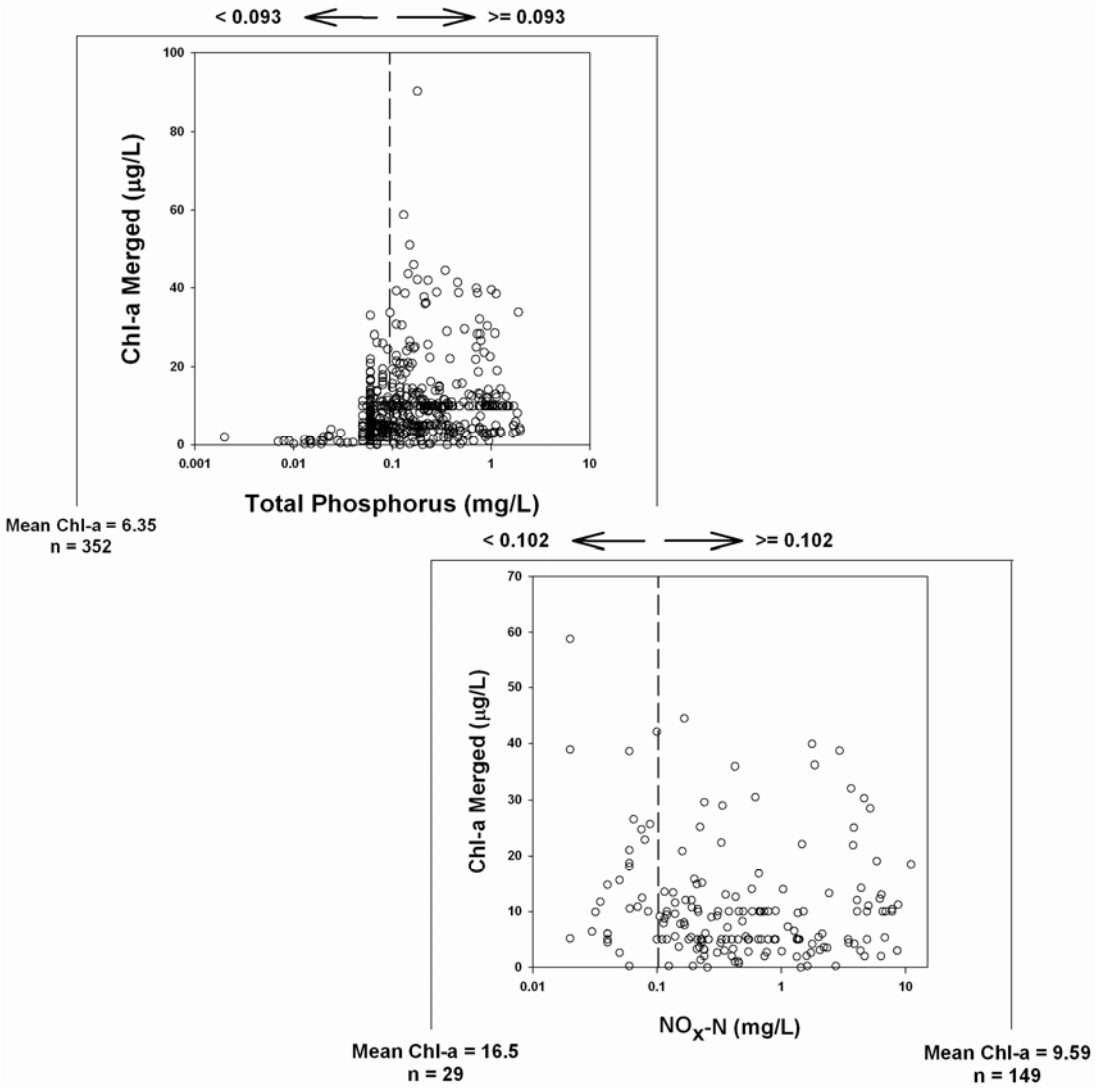


Figure 4-4. CART model for median merged chlorophyll-a data versus median nutrient concentrations (TP, TN, and NO_X-N) from the entire median dataset (model r^2 = 0.54).

When data from the spectrophotometric and fluorometric methods were merged into a single variable, the model structure for the CART model used to predict median merged chl-a data (Figure 4-4) was most similar to the structure for the CART model predicting chl-a from spectrophotometric measurements (Figure 4-2). The two models did not differ in their threshold values for nutrient concentrations but did

differ slightly in the median values of the reduced data. Median TP was the strongest predictor of median merged chl-a concentrations (Figure 4.4). Median chl-a was approximately 6.4 µg/L when the median TP in the water column was less than 0.093 mg/L. NO_X-N was an important secondary predictor when TP concentrations equaled or exceeded 0.093 mg/L. Median chl-a concentrations were greatest (16.5 µg/L) when TP equaled or exceeded the 0.093 mg/L threshold and NO_X-N was less than 0.1 mg/L. However, when TP equaled or exceeded the 0.093 mg/L threshold and NO_X-N was greater than 0.1 mg/L the median chl-a concentration was only 9.6 µg/L. This pattern also makes sense ecologically in that algal biomass expressed as chl-a was greatest when TP was high, but dissolved inorganic N (NO_X-N) was low.

The results of CART modeling on the entire median database in this study indicated that nutrient thresholds can explain substantial variability in some measurements of biological variability in Texas streams and rivers. Secchi depth and chl-a concentrations were generally related to nutrient concentrations in predictable patterns. It was interesting to note that different measurements of chlorophyll-a concentrations resulted in very similar model structure but slightly different threshold concentrations and median response values. Total nitrogen was generally the second-strongest predictor of biological variables in the primary split of all CART models, but TP was always the strongest predictor (Table 4-1).

CART on "Low P" Basin by Ecoregion III Dataset

Several CART models were developed from the "Low P" basin by ecoregion III dataset on Texas streams and rivers. The models presented in this section met the minimum requirements for inclusion in the mainbody of this document.

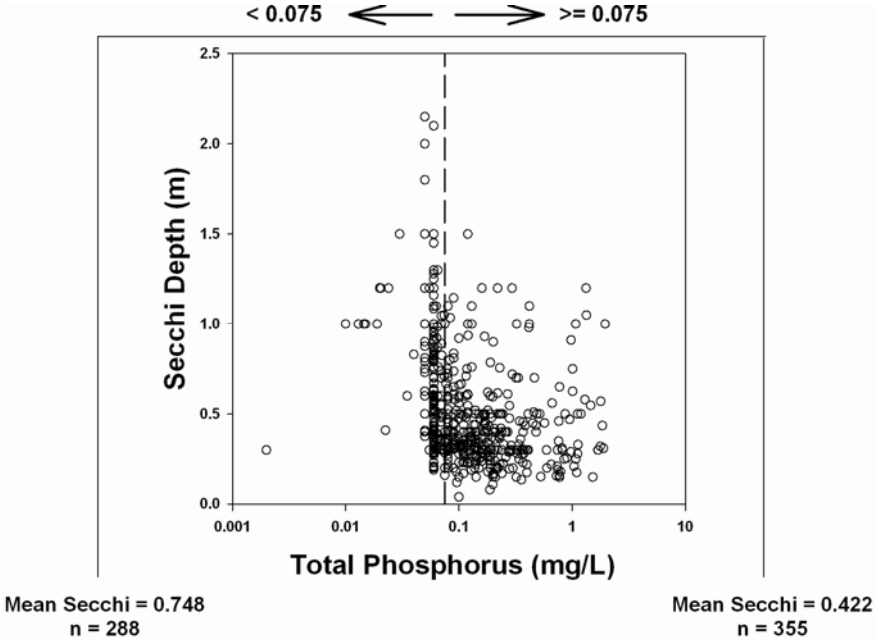


Figure 4-5. CART model for Secchi depth versus median nutrient concentrations (TP, TN, and NO_X-N) from the "Low P" basin by ecoregion III dataset (model r^2 = 0.24).

Median TP was also the strongest predictor of median Secchi depth in the dataset that was limited to "Low P" basin by ecoregion medians (Figure 4-5). Similar to the results from the entire dataset shown above, median Secchi depth was approximately 0.8 m when the median TP in the water column was less than 0.075 mg/L. When TP equaled or exceeded the 0.075 mg/L threshold, median Secchi depth decreased to 0.4 m. There were no statistically valid secondary splits in this CART model, but the threshold in median TP concentrations explained 24% of the variation in median Secchi depth across all river and stream stations included in the dataset.

Nutrient concentrations only explained 11% of the variation in the primary split on dissolved oxygen flux in the "Low P" basin by ecoregion III database. This fell below our prescribed minimum for showing graphical results.

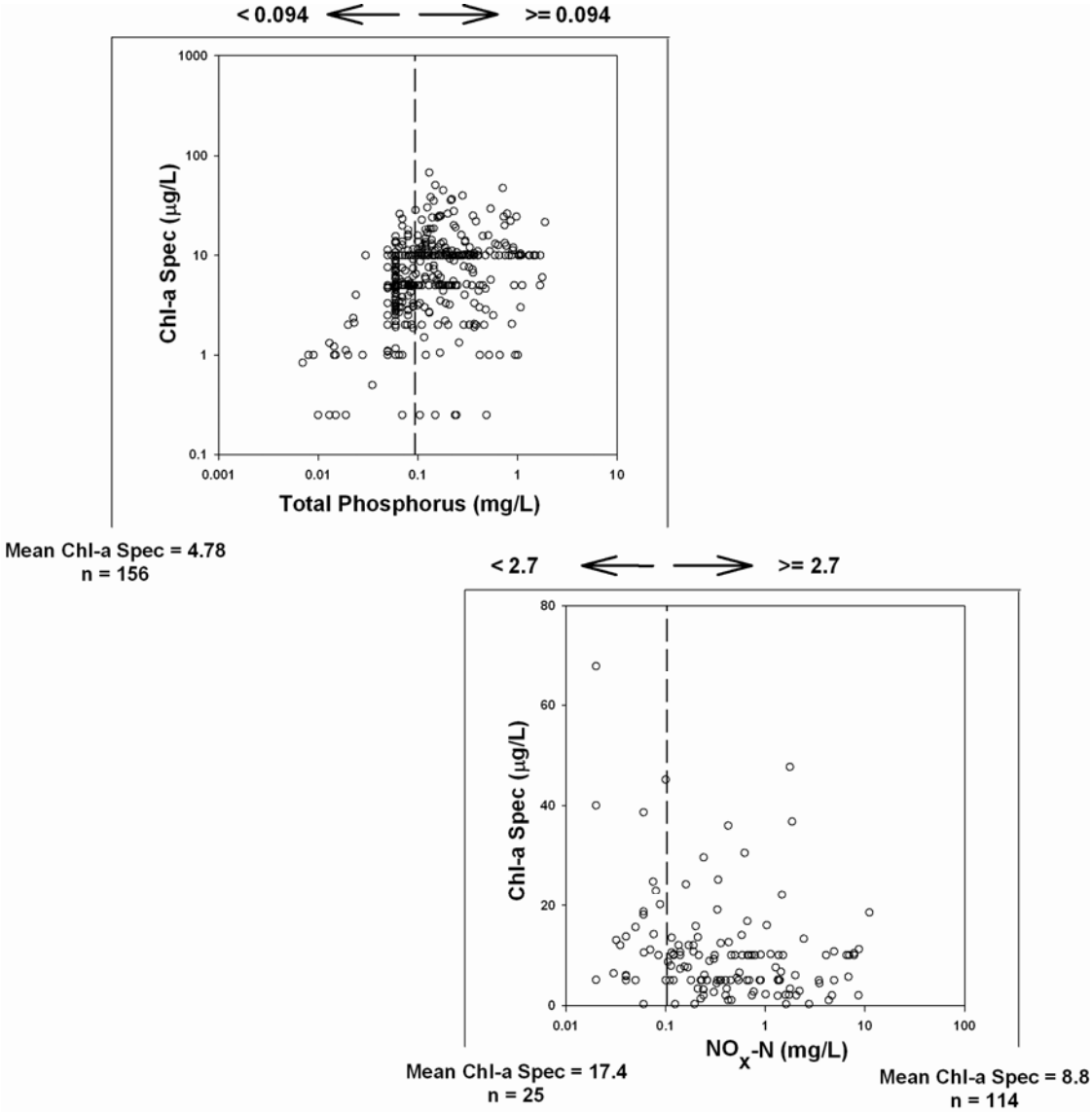


Figure 4-6. CART model for median chlorophyll-a concentration determined by spectrophotometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the "Low P" basin by ecoregion III dataset (model r^2 = 0.40).

Median TP was the strongest predictor of median chl-a concentrations measured using spectrophotometry (Figure 4-6). Median chl-a was approximately 4.8 µg/L when the median TP in the water column was less than 0.094 mg/L. NO_X-N was an important secondary predictor when TP concentrations equaled or exceeded 0.094 mg/L. Median chl-a concentrations were greatest (17.4 µg/L) when TP equaled or exceeded the 0.094 mg/L threshold and NO_X-N was less than 2.7 mg/L. However, when TP equaled or exceeded the 0.094 mg/L threshold and NO_X-N was greater than 2.7 mg/L the median chl-a concentration was only 8.8 µg/L. This pattern makes sense ecologically in that algal biomass expressed as chl-a was greatest when TP was high, but dissolved inorganic N (NO_X-N) was low.

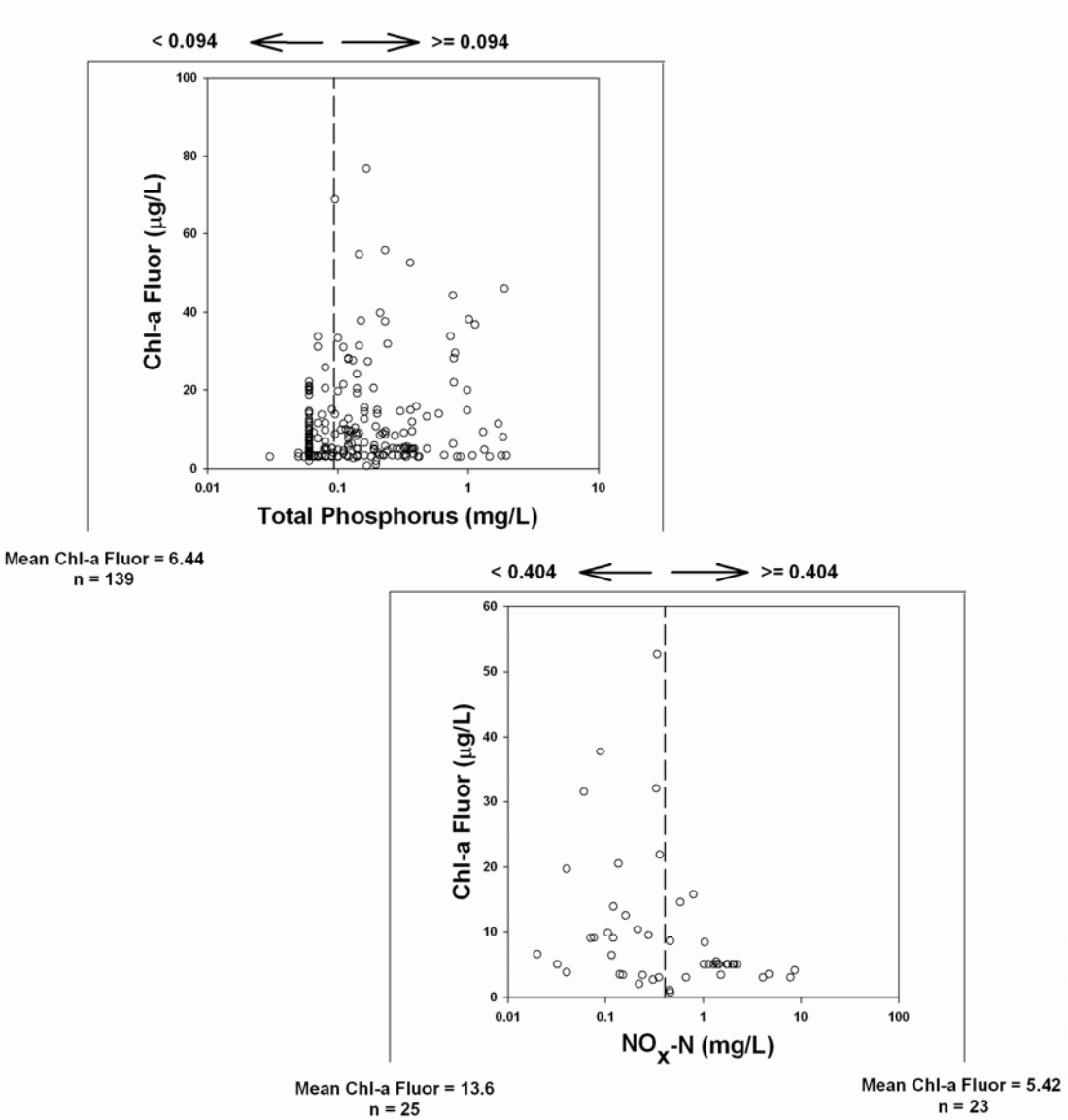


Figure 4-7. CART model for median chlorophyll-a concentration determined by fluorometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the "Low P" basin by ecoregion III dataset (model r^2 = 0.74).

When using the "Low P" basin by ecoregion dataset, model structure for the CART model used to predict median chl-a concentrations from fluometric measurements (Figure 4-7) was very similar to the structure for the CART model predicting chl-a from spectrophotometric measurements (Figure 4-6). However, these models differed in their threshold values and in the median values of the reduced data. Median TP was the strongest predictor of median chl-a concentrations measured using fluorometry (Figure 4-7). Median chl-a was approximately 6.4 µg/L when the median TP in the water column was less than 0.094 mg/L. NO_X-N was an important secondary predictor when TP concentrations equaled or exceeded 0.094 mg/L. Median chl-a concentrations were greatest (13.6 µg/L) when TP equaled or exceeded the 0.094 mg/L threshold and NO_X-N was less than 0.40 mg/L. However, when TP equaled or exceeded the 0.094 mg/L threshold and NO_X-N was greater than 0.40 mg/L the median chl-a concentration was only 5.4 µg/L. This pattern also makes sense ecologically in that algal biomass expressed as chl-a was greatest when TP was high, but dissolved inorganic N (NO_X-N) was low.

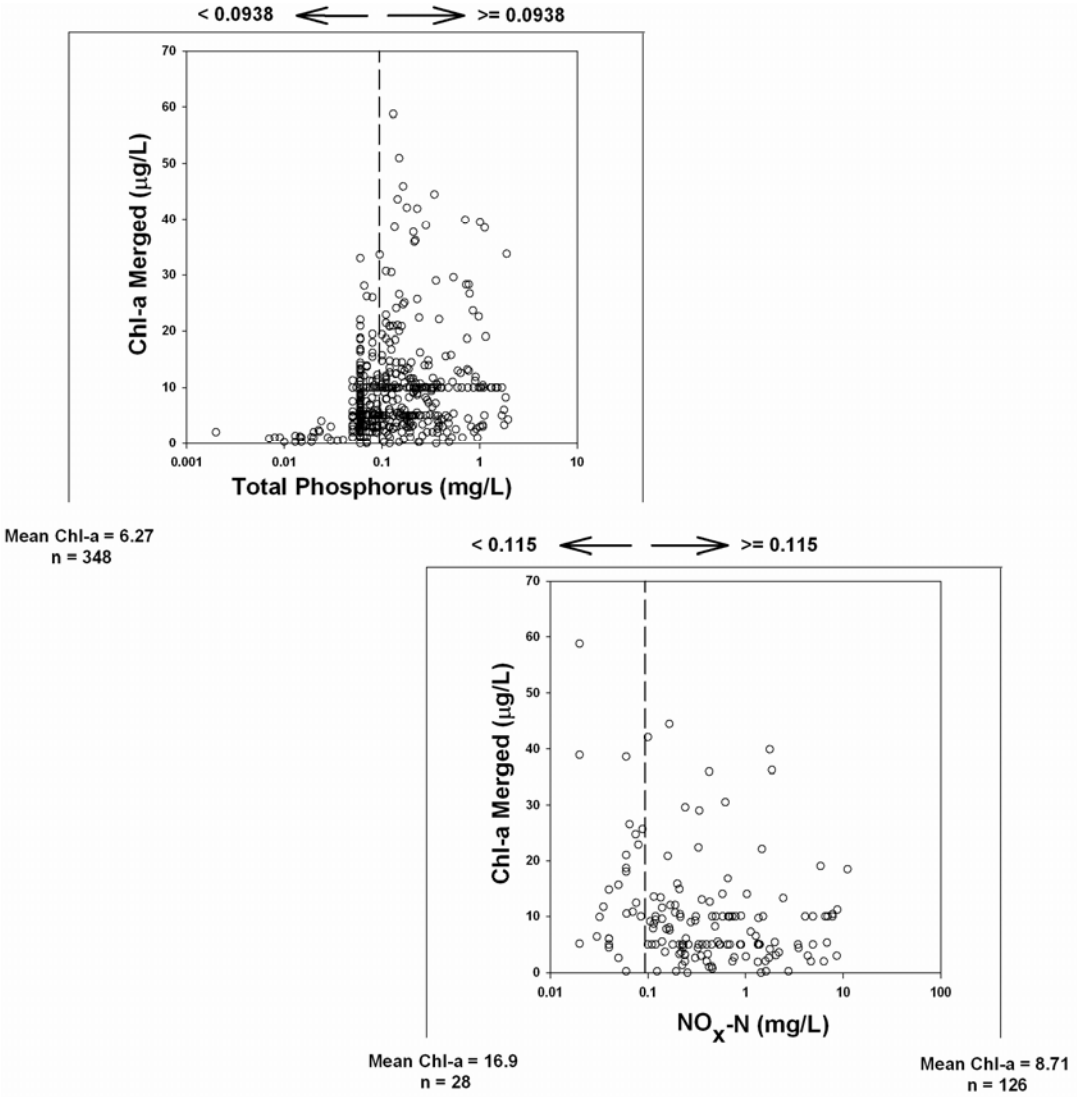


Figure 4-8. CART model for median merged chlorophyll-a data versus median nutrient concentrations (TP, TN, and NO_X-N) from the "Low P" basin by ecoregion III dataset (model r^2 = 0.74).

Merged chl-a data in the "Low P" basin by ecoregion III dataset resulted in a similar model structure to other CART models on chl-a data (Figure 4-8). Median TP was the strongest predictor of median merged chl-a concentrations in these data (Figure 4-8). Median chl-a was approximately 6.3 µg/L when the median TP in the water column was less than 0.094 mg/L. NO_X-N was an important secondary predictor when TP concentrations equaled or exceeded 0.094 mg/L. Median chl-a concentrations were greatest (16.9 µg/L) when TP equaled or exceeded the 0.094 mg/L threshold and NO_X-N was less than 0.12 mg/L. However, when TP equaled or exceeded the 0.094 mg/L threshold and NO_X-N was greater than 0.12 mg/L the median chl-a concentration was only 8.7 µg/L. This pattern also makes sense ecologically in that algal biomass expressed as chl-a was greatest when TP was high, but dissolved inorganic N (NO_X-N) was low.

The results of CART modeling on the "Low P" basin by ecoregion database in this study provided further evidence that nutrient thresholds can explain substantial variability in some measurements of biological variability in Texas streams and rivers. The relationships observed in these data were very similar to those observed when the entire database was used to construct CART models. As seen previously, Secchi depth and chl-a concentrations were generally related to nutrient concentrations in predictable patterns. As with the full database, TN was generally the second-strongest predictor of biological variables in the primary split of all CART models, but TP was always the strongest predictor (Table 4-1). Also, the threshold values for all nutrients in this reduced dataset were similar to those observed for the complete median dataset (Table 4-1).

CART on "High P" Basin by Ecoregion III Dataset

Only two CART models could be developed from the "high P" basin by ecoregion III dataset on Texas streams and rivers that met the minimum requirements for inclusion in the mainbody of this document.

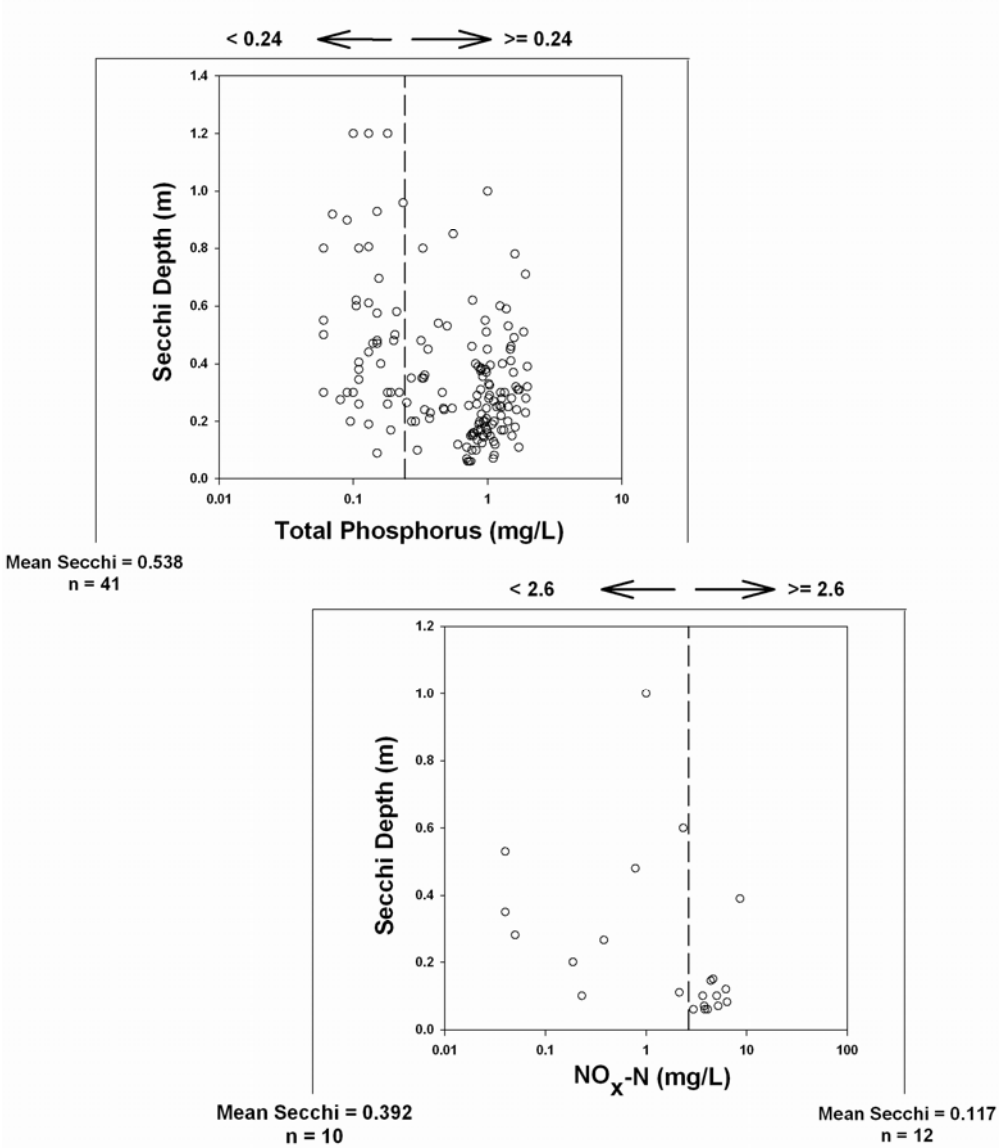


Figure 4-9. CART model for median Secchi depth versus median nutrient concentrations (TP, TN, and NO_X-N) from the "high P" basin by ecoregion III dataset (model r^2 = 0.52).

Median TP was also the strongest predictor of median Secchi depth in the dataset that was limited to "high P" basin by ecoregion medians (Figure 4-9). Median Secchi depth in this database was approximately 0.5 m when the median TP in the water column was less than 0.24 mg/L. NO_X-N was an important secondary predictor when TP values exceeded or equaled 0.24 mg/L. When TP equaled or exceeded the 0.24 mg/L threshold and NO_X-N was greater than 2.6 mg/L, median Secchi depth was 0.12

m. However, when TP equaled or exceeded 0.24 mg/L and NO_X-N was less than 2.6 mg/L, median Secchi depth was 0.4 m. This pattern did not make sense ecologically. Although we expected the Secchi depths to be least at high TP concentrations, the decreased transparency at both high P and high NO_X-N concentrations does not fit a predictable pattern. Rather, we would expect to see less transparency (i.e. greater Secchi depth) when NO_X-N concentrations were highest because that would be indicative of efficient inorganic nutrient use by the algae, which was consistently observed above in the chl-a CART models.

No statistically-valid CART model could be produced for the DO flux in the "high P" basin by ecoregion III database.

Nutrient concentrations only explained 9% of the variation in the primary split in median chl-a measured using spectrophotometry in the "high P" basin by ecoregion III database. This fell below our prescribed minimum for showing graphical results.

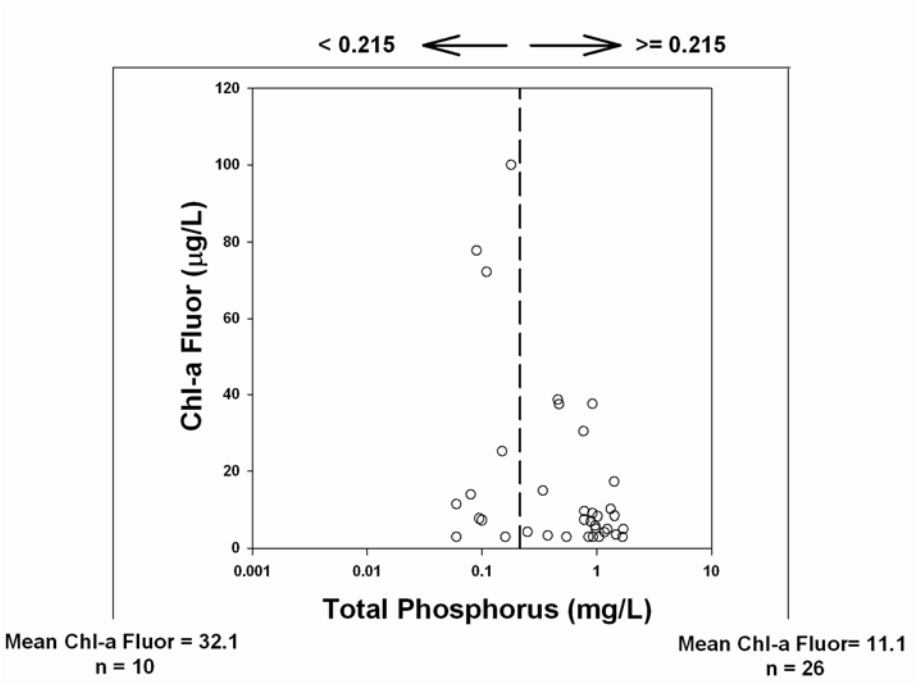


Figure 4-10. CART model for median chlorophyll-a measured with fluorometry versus median nutrient concentrations (TP, TN, and NO_X-N) from the "high P" basin by ecoregion III dataset (model r^2 = 0.17).

Median TP was the strongest predictor of median chl-a concentrations measured using fluorometry (Figure 4-10). Median chl-a was approximately 32 µg/L when the median TP in the water column was less than 0.22 mg/L. Median chl-a concentrations were greatest (13.6 µg/L) when TP equaled or exceeded the 0.22 mg/L threshold. This pattern did not make sense ecologically and appears to be driven primarily by three medians that all exceeded 75 µg/L chl-a.

Nutrient concentrations only explained 7% of the variation in the primary split in median chl-a measured using spectrophotometry in the "high P" basin by ecoregion III database. This fell below our prescribed minimum for showing graphical results.

The results of CART modeling on the "high P" basin by ecoregion database did not provide much useful information. Only two CART models were statistically valid and both yielded results that did not support well established ecological theory. One exception was the primary split of TP concentrations in predicting Secchi depth in this dataset. A threshold of 0.24 mg/L TP separated relatively high and low Secchi depths. This threshold was substantially greater than the thresholds observed in the analyses on the entire dataset and on the "Low P" basin by ecoregion III dataset (Table 4-1). However, that difference was expected because the separation of data specifically targeted relatively P-rich locations in this analysis.

REFERENCES

De'ath G, Fabricus KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

East T. L., Sharfstein B. 2006. Development of a decision tree model for the prediction of the limitation potential of phytoplankton in Lake Okeechobee, Florida, USA. Arch. Hydrobiol. 2006;165:127-144.

King RS, Baker ME, Whigham DF, Weller DE, Jordan TE, Kazyak PF, Hurd MK. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

Qian SS, King RS, Richardson CJ. 2003. Two methods for the detection of environmental thresholds. Ecological Modeling 166:87-97.

Urban, D.L. (2002) "Classification and regression trees" In: McCune, B., Grace, J.B. eds. , Analysis of Ecological Communities, MjM Software Design, Gleneden Beach, Oregon, USA, pp 221-231.

APPENDIX 4-1: Models on Biologicals versus Nutrients—All Median Data

Secchi versus TP, TN, and NO_x-N

```
Call:
mvpart(form = Secchi_00078 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=826 (1625 observations deleted due to missingness)

       CP nsplit rel error    xerror      xstd
1 0.2541120      0  1.000000 1.0016417 0.06949203
2 0.1253307      1  0.745888 0.8031284 0.05211658

Node number 1: 826 observations,    complexity param=0.254112
  mean=0.5242149, MSE=0.1136746
  left son=2 (509 obs) right son=3 (293 obs), 24 observations remain
  Primary splits:
     tp_00665   < 0.07475  to the right, improve=0.23467150, (24 missing)
     tn_00600C  < 1.504964 to the right, improve=0.02542678, (663 missing)
     nox_00631C < 2.85     to the right, improve=0.01321244, (449 missing)

Node number 2: 509 observations
  mean=0.4019484, MSE=0.05338546

Node number 3: 293 observations
  mean=0.7461766, MSE=0.146287
```

DO flux versus TP, TN, and NO_X-N

```
Call:
mvpart(form = doflux_89856C ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=102 (2349 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.11145714      0 1.0000000 1.027205 0.2642652
2 0.02999099      2 0.7770857 1.122797 0.2207620

Node number 1: 102 observations,    complexity param=0.1114571
  mean=2.122843, MSE=2.859015
  left son=2 (68 obs) right son=3 (32 obs), 2 observations remain
  Primary splits:
     tp_00665   < 0.06025 to the right, improve=0.06616325, (2 missing)
     tn_00600C  < 0.816   to the right, improve=0.03794745, (51 missing)
     nox_00631C < 0.09945 to the right, improve=0.02136768, (29 missing)

Node number 2: 68 observations,    complexity param=0.1114571
  mean=1.837574, MSE=2.011308
  left son=4 (36 obs) right son=5 (10 obs), 22 observations remain
  Primary splits:
     nox_00631C < 0.115   to the right, improve=0.05174488, (22 missing)
     tp_00665   < 0.338   to the left,  improve=0.04955176, (0 missing)
     tn_00600C  < 1.195   to the left,  improve=0.03001522, (36 missing)

Node number 3: 32 observations
  mean=2.779219, MSE=4.166981

Node number 4: 36 observations
  mean=1.718056, MSE=2.000484

Node number 5: 10 observations
  mean=2.669, MSE=2.125259
```

Chl-a Spectrophotometry versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlaspec_32211 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=590 (1861 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.18573416      0 1.0000000 1.001453 0.1533704
2 0.05044318      2 0.6285317 1.102125 0.1228702

Node number 1: 590 observations,    complexity param=0.1857342
  mean=8.775006, MSE=56.47014
  left son=2 (290 obs) right son=3 (294 obs), 6 observations remain
  Primary splits:
     tp_00665   < 0.0935   to the left,  improve=0.09570816, (6 missing)
     tn_00600C  < 1.562725 to the left,  improve=0.04986554, (430 missing)
     nox_00631C < 0.0245   to the right, improve=0.02017316, (273 missing)

Node number 2: 290 observations
  mean=6.376181, MSE=18.22351

Node number 3: 294 observations,    complexity param=0.1857342
  mean=11.0497, MSE=78.93059
  left son=6 (132 obs) right son=7 (26 obs), 136 observations remain
  Primary splits:
     nox_00631C < 0.10275  to the right, improve=0.048686560, (136 missing)
     tn_00600C  < 1.562725 to the left,  improve=0.029156710, (220 missing)
     tp_00665   < 0.70025  to the left,  improve=0.007724358, (0 missing)

Node number 6: 132 observations
  mean=9.705499, MSE=72.90104

Node number 7: 26 observations
  mean=16.9175, MSE=232.0491
```

Chl-a Flourometry versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlafluor_70953 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=309 (2142 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.42066688      0 1.0000000 1.0020031 0.2164213
2 0.01555814      2 0.1586662 0.9508442 0.1967900

Node number 1: 309 observations,    complexity param=0.4206669
  mean=10.5613, MSE=178.2886
  left son=2 (111 obs) right son=3 (197 obs), 1 observation remains
  Primary splits:
     tp_00665   < 0.0675  to the left,  improve=0.070857510, (1 missing)
     tn_00600C  < 1.3525  to the left,  improve=0.007233137, (251 missing)
     nox_00631C < 0.3585  to the right, improve=0.006858616, (216 missing)

Node number 2: 111 observations,    complexity param=0.01555814
  mean=5.852387, MSE=20.61608
  left son=4 (13 obs) right son=5 (98 obs)
  Primary splits:
     tp_00665   < 0.0575  to the left,  improve=0.049785880, (0 missing)
     nox_00631C < 0.195   to the right, improve=0.048926240, (82 missing)
     tn_00600C  < 0.616   to the left,  improve=0.003382745, (89 missing)

Node number 3: 197 observations,    complexity param=0.4206669
  mean=13.26745, MSE=247.6661
  left son=6 (30 obs) right son=7 (33 obs), 134 observations remain
  Primary splits:
     nox_00631C < 0.3585  to the right, improve=0.008475183, (134 missing)
     tp_00665   < 0.18825 to the right, improve=0.007794559, (0 missing)
     tn_00600C  < 1.3525  to the left,  improve=0.004524351, (161 missing)

Node number 4: 13 observations
  mean=3.070769, MSE=0.06009941

Node number 5: 98 observations,    complexity param=0.01555814
  mean=6.221378, MSE=22.18035
  left son=10 (12 obs) right son=11 (15 obs), 71 observations remain
  Primary splits:
     nox_00631C < 0.195   to the right, improve=0.053844620, (71 missing)
     tn_00600C  < 0.616   to the left,  improve=0.005242796, (77 missing)
```

Node number 6: 30 observations
mean=7.030583, MSE=62.79806

Node number 7: 33 observations
mean=12.1603, MSE=138.4481

Node number 10: 12 observations
mean=4.583333, MSE=1.841389

Node number 11: 15 observations
mean=8.773333, MSE=36.75162

Chl-a Merged versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlamerge_70953C ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=715 (1736 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.27024120      0 1.0000000 1.0015840 0.1532872
2 0.04383012      2 0.4595176 0.9986235 0.1204772

Node number 1: 715 observations,    complexity param=0.2702412
  mean=8.97765, MSE=77.2065
  left son=2 (352 obs) right son=3 (355 obs), 8 observations remain
  Primary splits:
     tp_00665   < 0.09375  to the left,  improve=0.086252920, (8 missing)
     tn_00600C  < 1.562725 to the left,  improve=0.031737340, (532 missing)
     nox_00631C < 2.85     to the left,  improve=0.009602496, (360 missing)

Node number 2: 352 observations
  mean=6.354687, MSE=24.65693

Node number 3: 355 observations,    complexity param=0.2702412
  mean=11.54497, MSE=112.7686
  left son=6 (149 obs) right son=7 (29 obs), 177 observations remain
  Primary splits:
     nox_00631C < 0.10275  to the right, improve=0.028756760, (177 missing)
     tn_00600C  < 1.562725 to the left,  improve=0.019167900, (266 missing)
     tp_00665   < 0.184    to the right, improve=0.005079723, (0 missing)

Node number 6: 149 observations
  mean=9.585433, MSE=75.72699

Node number 7: 29 observations
  mean=16.4719, MSE=186.3458
```

APPENDIX 4-2: CART Models on Biologicals versus Nutrients—Data from Low P Basin by Ecoregion III

Secchi versus TP, TN, and NO_x-N

```
Call:
mvpart(form = Secchi_00078 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=667 (1450 observations deleted due to missingness)

       CP nsplit rel error    xerror      xstd
1 0.2379961      0  1.000000 1.0017910 0.07420215
2 0.1915764      1  0.762004 0.8415712 0.05827608

Node number 1: 667 observations,    complexity param=0.2379961
  mean=0.562503, MSE=0.1198315
  left son=2 (355 obs) right son=3 (288 obs), 24 observations remain
  Primary splits:
     tp_00665   < 0.07475 to the right, improve=0.211887500, (24 missing)
     tn_00600C  < 1.175   to the right, improve=0.015813910, (522 missing)
     nox_00631C < 0.0296  to the left,  improve=0.002772206, (317 missing)
Node number 2: 355 observations
  mean=0.422112, MSE=0.05176277
Node number 3: 288 observations
  mean=0.7484714, MSE=0.1476714
```

DO flux versus TP, TN, and NO_X-N

```
Call:
mvpart(form = doflux_89856C ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=93 (2024 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.12784308      0 1.0000000 1.039064 0.2777063
2 0.04121899      2 0.7443138 1.108387 0.2698922

Node number 1: 93 observations,    complexity param=0.1278431
  mean=2.121022, MSE=2.985469
  left son=2 (60 obs) right son=3 (31 obs), 2 observations remain
  Primary splits:
     tp_00665   < 0.06025 to the right, improve=0.08713811, (2 missing)
     tn_00600C  < 0.816   to the right, improve=0.04756599, (48 missing)
     nox_00631C < 0.09945 to the right, improve=0.02927733, (27 missing)

Node number 2: 60 observations,    complexity param=0.1278431
  mean=1.768, MSE=2.062424
  left son=4 (29 obs) right son=5 (10 obs), 21 observations remain
  Primary splits:
     nox_00631C < 0.115   to the right, improve=0.08127144, (21 missing)
     tp_00665   < 0.338   to the left,  improve=0.07255784, (0 missing)
     tn_00600C  < 1.265   to the left,  improve=0.02327779, (34 missing)

Node number 3: 31 observations
  mean=2.855968, MSE=4.112907

Node number 4: 29 observations
  mean=1.506034, MSE=1.996723

Node number 5: 10 observations
  mean=2.669, MSE=2.125259
```

Chl-a Spectrophotometry versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlaspec_32211 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=543 (1574 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.19844134      0 1.0000000 1.002511 0.1693815
2 0.05523163      2 0.6031173 1.047536 0.1380916

Node number 1: 543 observations,    complexity param=0.1984413
  mean=8.512299, MSE=54.83031
  left son=2 (288 obs) right son=3 (249 obs), 6 observations remain
  Primary splits:
     tp_00665   < 0.0935   to the left,  improve=0.09396533, (6 missing)
     nox_00631C < 0.0245   to the right, improve=0.02541811, (245 missing)
     tn_00600C  < 1.562725 to the left,  improve=0.01896113, (397 missing)

Node number 2: 288 observations,    complexity param=0.05523163
  mean=6.336953, MSE=18.09846
  left son=4 (20 obs) right son=5 (268 obs)
  Primary splits:
     tp_00665   < 0.029    to the left,  improve=0.112170700, (0 missing)
     nox_00631C < 0.778    to the right, improve=0.020679500, (134 missing)
     tn_00600C  < 0.495    to the left,  improve=0.008342541, (205 missing)

Node number 3: 249 observations,    complexity param=0.1984413
  mean=10.914, MSE=80.76241
  left son=6 (114 obs) right son=7 (25 obs), 110 observations remain
  Primary splits:
     nox_00631C < 0.10275  to the right, improve=0.075298670, (110 missing)
     tn_00600C  < 1.509964 to the left,  improve=0.011972190, (189 missing)
     tp_00665   < 0.23075  to the right, improve=0.009515208, (0 missing)

Node number 4: 20 observations
  mean=1.12125, MSE=0.8149122

Node number 5: 268 observations,    complexity param=0.05523163
  mean=6.726185, MSE=17.20666
  left son=10 (19 obs) right son=11 (117 obs), 132 observations remain
  Primary splits:
     nox_00631C < 0.778    to the right, improve=0.03337636, (132 missing)
     tp_00665   < 0.0525   to the left,  improve=0.02216367, (0 missing)
     tn_00600C  < 0.495    to the left,  improve=0.01430103, (193 missing)
```

Node number 6: 114 observations
mean=8.800446, MSE=60.16318

Node number 7: 25 observations
mean=17.3942, MSE=235.4227

Node number 10: 19 observations
mean=2.675263, MSE=5.020225

Node number 11: 117 observations
mean=5.743825, MSE=15.48604

Chl-a Fluorometry versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlafluor_70953 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=273 (1844 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.37357535      0 1.0000000 1.009035 0.2029476
2 0.01780068      2 0.2528493 0.919790 0.1791377

Node number 1: 273 observations,    complexity param=0.3735753
  mean=9.723764, MSE=127.8256
  left son=2 (139 obs) right son=3 (133 obs), 1 observation remains
  Primary splits:
     tp_00665   < 0.09375   to the left,  improve=0.089854460, (1 missing)
     nox_00631C < 0.3585    to the right, improve=0.017050950, (187 missing)
     tn_00600C  < 1.3525    to the left,  improve=0.003886938, (218 missing)

Node number 2: 139 observations
  mean=6.437806, MSE=32.90759

Node number 3: 133 observations,    complexity param=0.3735753
  mean=13.23002, MSE=203.7177
  left son=6 (23 obs) right son=7 (25 obs), 85 observations remain
  Primary splits:
     nox_00631C < 0.4044062 to the right, improve=0.029317800, (85 missing)
     tp_00665   < 0.6925    to the left,  improve=0.018742450, (0 missing)
     tn_00600C  < 1.81525   to the right, improve=0.003039967, (104 missing)

Node number 6: 23 observations
  mean=5.416848, MSE=12.20271

Node number 7: 25 observations
  mean=13.56, MSE=158.7484
```

Chl-a Merged versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlamerge_70953C ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=651 (1466 observations deleted due to missingness)

       CP nsplit rel error   xerror      xstd
1 0.2411303      0 1.0000000 1.001650 0.12341664
2 0.0540229      2 0.5177394 1.037098 0.09926795

Node number 1: 651 observations,    complexity param=0.2411303
  mean=8.531635, MSE=63.86653
  left son=2 (348 obs) right son=3 (295 obs), 8 observations remain
  Primary splits:
     tp_00665   < 0.09375   to the left,  improve=0.09102999, (8 missing)
     tn_00600C  < 1.195     to the left,  improve=0.01675841, (486 missing)
     nox_00631C < 0.0245    to the right, improve=0.01431673, (320 missing)

Node number 2: 348 observations
  mean=6.273851, MSE=23.44532

Node number 3: 295 observations,    complexity param=0.2411303
  mean=11.14268, MSE=94.54869
  left son=6 (126 obs) right son=7 (28 obs), 141 observations remain
  Primary splits:
     nox_00631C < 0.10275   to the right, improve=0.054897770, (141 missing)
     tn_00600C  < 1.509964  to the left,  improve=0.009393214, (224 missing)
     tp_00665   < 0.6896665 to the left,  improve=0.007526353, (0 missing)

Node number 6: 126 observations
  mean=8.706147, MSE=64.28106

Node number 7: 28 observations
  mean=16.88161, MSE=188.133
```

APPENDIX 4-3: CART Models on Biologicals versus Nutrients—Data From High P Basin by Ecoregion III

Secchi versus TP, TN, and NO_x-N

```
Call:
mvpart(form = Secchi_00078 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=159 (175 observations deleted due to missingness)

       CP nsplit rel error   xerror      xstd
1 0.2635922      0 1.0000000 1.021706 0.1621587
2 0.0100000      2 0.4728157 0.912742 0.1300673

Node number 1: 159 observations,    complexity param=0.2635922
  mean=0.3635975, MSE=0.05589922
  left son=2 (118 obs) right son=3 (41 obs)
  Primary splits:
     tp_00665   < 0.2425  to the right, improve=0.18916660, (0 missing)
     nox_00631C < 2.6455  to the right, improve=0.04542242, (132 missing)

Node number 2: 118 observations,    complexity param=0.2635922
  mean=0.3029831, MSE=0.03193622
  left son=4 (12 obs) right son=5 (10 obs), 96 observations remain
  Primary splits:
     nox_00631C < 2.6455  to the right, improve=0.10886450, (96 missing)
     tp_00665   < 0.57825 to the right, improve=0.04131705, (0 missing)

Node number 3: 41 observations
  mean=0.5380488, MSE=0.08385839

Node number 4: 12 observations
  mean=0.11725, MSE=0.007698188

Node number 5: 10 observations
  mean=0.3915, MSE=0.06718025
```

DO Flux versus TP, TN, and NO_X-N

No splits possible -- try decreasing cp

Chl-a Spectrophotometry versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlaspec_32211 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=47 (287 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.04423492      0 1.0000000 1.081922 0.3160349
2 0.01000000      2 0.9115302 1.210664 0.3205365

Node number 1: 47 observations,    complexity param=0.04423492
  mean=11.81011, MSE=65.40629
  left son=2 (26 obs) right son=3 (21 obs)
  Primary splits:
     tp_00665 < 0.77375 to the right, improve=0.03220433, (0 missing)

Node number 2: 26 observations
  mean=10.50577, MSE=35.19699

Node number 3: 21 observations,    complexity param=0.04423492
  mean=13.425, MSE=98.09405
  left son=6 (11 obs) right son=7 (10 obs)
  Primary splits:
     tp_00665 < 0.3575  to the left,  improve=0.08396488, (0 missing)

Node number 6: 11 observations
  mean=10.68864, MSE=46.67356

Node number 7: 10 observations
  mean=16.435, MSE=137.36
```

Chl-a Fluorometry versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlafluor_70953 ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=36 (298 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.17307636      0 1.0000000 1.065072 0.4533584
2 0.03180164      1 0.8269236 0.992791 0.3483801

Node number 1: 36 observations,    complexity param=0.1730764
  mean=16.91264, MSE=515.3078
  left son=2 (26 obs) right son=3 (10 obs)
  Primary splits:
     tp_00665 < 0.215 to the right, improve=0.1730764, (0 missing)

Node number 2: 26 observations
  mean=11.05577, MSE=127.6064

Node number 3: 10 observations
  mean=32.1405, MSE=1202.256
```

Chl-a merged versus TP, TN, and NO_X-N

```
Call:
mvpart(form = chlamerge_70953C ~ tp_00665 + tn_00600C + nox_00631C,
   data = streamstplt2, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=64 (270 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.07223685      0 1.0000000 1.068098 0.5046990
2 0.03080409      1 0.9277632 1.013424 0.4145686

Node number 1: 64 observations,    complexity param=0.07223685
  mean=13.51445, MSE=190.2929
  left son=2 (53 obs) right son=3 (11 obs)
  Primary splits:
     tp_00665   < 0.185  to the right, improve=0.07223685, (0 missing)
     nox_00631C < 2.6455 to the left,  improve=0.05646686, (40 missing)

Node number 2: 53 observations
  mean=11.82538, MSE=91.8884

Node number 3: 11 observations
  mean=21.65273, MSE=584.4462
```

Chapter 5: Threshold Analysis on Streams and Rivers Database with Bioassessment Data

J. Thad Scott[1], Brian E. Haggard[2] and Leslie B. Massey[3]

[1]Assistant Professor, Crop Soil and Environmental Sciences Department, UA Division of Agriculture,
115 Plant Science Building, Fayetteville, AR 72701
[2]Professor and Director, Arkansas Water Resources Center, UA Division of Agriculture
[3]Project Manager, Arkansas Water Resources Center, UA Division of Agriculture
203 Engineering Hall, Fayetteville, Arkansas, 72701
Corresponding Author: jts004@uark.edu

EXECUTIVE SUMMARY: We used categorical and regression tree analyses on median bioassessment data from Texas rivers and streams to identify changepoints which result in substantial ecological change. Identical analyses were conducted on three data sets: 1) the entire rivers and streams median bioassessment dataset where a minimum of ten observations were required in calculating medians, 2) a reduced data set that included data from only "Low Phosphorus" (P) basin by ecoregion combinations as defined in Chapter 3, and 3) a reduced data set that included data from only "High P" basin by ecoregion combinations as defined in Chapter 3. Commonly measured biological variables were used in the analysis, and thresholds for total P (TP), total nitrogen (TN), and nitrite plus nitrate-nitrogen (NO_X-N) were derived in all analyses. In the complete bioassessment median database, TP thresholds that resulted in changes in Fish IBI and RBIBI were 0.065 and 0.059 mg/L, respectively. TP thresholds for these two parameters were much more variable when data were limited to either "Low P" or "High P" basin by ecoregion III geographic areas. TN thresholds that resulted in ecological changes ranged from 1.3 to 2.5 mg/L and did not differ drastically between data source (complete data versus "Low P" and "High P" basin by ecoregion III data). Habitat quality index was an important covariate that explained more variation than nutrients in Fish IBI and RBIBI in most analyses. The most prominent exception was a TP threshold of 0.059 mg/L that explained 21% of the variation in RBIBI across the entire bioassessment median database.

INTRODUCTION

The objective of this chapter was to identify nutrient concentrations that were correlated with a change in the magnitude or variability of bioassessment variables. Bioassessment can involve a number of different biological metrics but usually incorporate aspects of species diversity across environmental gradients. The metrics used in this study were Texas Fish Index of Biotic Integrity (Fish IBI) and the Texas Rapid Benthic Macroinvertebrate Index of Biotic Integrity (RBIBI). A limited number of stations had sufficient data for IBI estimates and this database was much smaller than the overall median water quality database presented in the previous chapter. However, some of the common biological measurements (Secchi depth, dissolved oxygen (DO) flux, and sestonic chlorophyll-a (chl-a)) were included in this database. We used that opportunity to compare thresholds that were derived from the overall median database for these common biological measurements to thresholds that were derived from the more limited bioassessment database.

Classification and regression tree (CART) analysis is an empirical modeling technique that is useful for identifying ecological thresholds and hierarchical structure in predictor variables (De'ath and Fabricius 2000). CART uses recursive partitioning to divide data into subsets that are increasingly homogeneous, invoking a tree-like classification that can explain relationships that may be difficult to reconcile with conventional linear models (Urban 2002). Categorical variables (e.g., station location, basin, ecoregion or land-use classifications) may also be used as independent variables in CART analysis, which provides another advantage to using CART rather than traditional regression techniques. CART and other similar methods have been used to identify thresholds and hierarchical structure in environmental correlates of various biological processes in aquatic ecosystems (King et al. 2005, East and Sharfstein 2006). King et al. (2005) used CART to specifically identify thresholds in nutrient concentrations which resulted in shifts in ecological structure and function. These thresholds were used to recommend specific water quality nutrient criteria for the Florida Everglades ecosystem.

METHODS

The Texas Commission on Environmental Quality (TCEQ) provided bioassessment data including fish IBI, RBIBI, ALU scores, HQI and others from 173 stations that spanned 16 basins. This database was provided in a useable format where Station ID and parameter labels and associated data were in a unique column. Water quality data were paired with bioassessment data based on Station ID. There were 11 stations in the bioassessment database that did not match Station IDs in the water quality database, so water quality stations located on the same or a nearby reach were paired with nine of these bioassessment data points. However, two stations did not have a similar station nearby, and the stations were removed from the bioassessment database.

Median values of each parameter were calculated for various time periods including IBI period (time frame during which all bioassessment data were collected at a give station), year, index period (March 15-October 15), non critical period (March 15-June 30; October 1-15), critical period (July 1-September

30), and monthly or matching date. The calculated medians for each Station ID were then compiled into a database for each of the identified time periods.

We conducted CART analyses on the bioassessment median database for streams and rivers to identify thresholds in nutrient concentrations that resulted in measurable changes in Fish IBI and RBIBI. The Habitat Quality Index (HQI) was also included as a potential predictor in these CART models because these biological indices are designed to capture the effect of degraded habitat on biological health. We also constructed CART models to identify nutrient thresholds in some common biological variables that were the primary focus of the previous chapter: median Secchi depth (m), median 24 hour dissolved oxygen (DO) flux, median chl-a measured with spectrophotometry, median chl-a measured with fluorometry, and median merged chl-a data (combined spectrophotometric and fluorometric measurements). The independent variables included in the analysis were total phosphorus (TP; TCEQ parameter code 00665), total nitrogen (TN; calculated parameter code 00600C), nitrite plus nitrate-nitrogen (NO_X-N; calculated parameter code 00631C), and the HQI.

CART analysis is a form of data reduction that aims to: 1) quantify thresholds in independent variables that are correlated with shifts in the magnitude and/or variability of dependent variables, and 2) identify hierarchical structure in independent variables. CART analysis is very useful for resolving nonlinear, hierarchical, and high-order interactions among predictor variables (De'Ath and Fabricius 2000) and for detecting numerical values that lead to ecological changes (Qian and others 2003). CART models use recursive partitioning to separate data into subsets that are increasingly homogeneous. This iterative process invokes a tree-like classification that can reveal relationships that are often difficult to reconcile with conventional linear models (Urban 2002).

CART analyses were performed using the MVPART library in R 2.8.1 (http://www.r-project.org/). We required a minimum of 20 observations to be used in any single split in the CART model and that each terminal node in the model had a minimum of ten observations. CART analysis is insensitive to missing data. Therefore, we did not remove observations from the data set due to missing values. However, we did require that all calculated medians have a minimum of ten observations used in calculating the median value. We first ran CART models using all median bioassessment data from the streams and rivers median database. Secondarily, we reran CART models after limiting data to stations from "low P" or "high P" basin by ecoregion III combinations (see Chapter 3 for details on this classification). CART models on the complete dataset were also developed for median Secchi depth (m), median 24 hour DO flux, median chl-a measured with spectrophotometry, median chl-a measured with fluorometry, and median merged chl-a data (combined spectrophotometric and fluorometric measurements) in order to compare these results to similar models built using the complete stream and river median database (Chapter 4). CART models for these variables were not constructed from the "Low P" or "High P" reduced bioassessment data. Because CART analysis involves recursive partitioning, models may sometimes be over-fit (i.e. too many independent variables that decrease the statistical rigor of final model). We "pruned" CART models to generate final models that balanced accuracy within the available dataset with robustness to novel data (Urban 2002). CART models were cross-validated to determine

"pruning size" (i.e., the number of predictor variables included in the model). Model cross-validations were conducted using 10 random and similarly sized subsets of our data according to the method detailed by De'ath and Fabricius (2000). The optimum tree size for each model was selected using the minimum cross-validated error rule (De'ath and Fabricius 2000).

RESULTS AND DISCUSSION

Analyses using all three bioassessment datasets resulted in statistically-valid thresholds in nutrient concentrations that were correlated with variation in bioassessment metrics (Table 5-1). TP thresholds were lowest in the analyses that included all bioassessment data. TP thresholds in the "Low P" and "High P" basin by ecoregion III bioassessment data were much more variable. Thresholds for TN and NO_X-N were reasonably similar between the complete bioassessment dataset and the "Low P" basin by ecoregion III bioassessment dataset. The "High P" basin by ecoregion III dataset provided no statistically valid thresholds for NO_X-N. Nutrients were very seldom the strongest predictor of bioassessment variables. Rather, HQI was generally the strongest primary predictor.

All of the thresholds reported in Table 5-1 were statistically valid according to our defined methods in CART analysis, and most represented predictable ecological relationships. However, it is useful to show these thresholds relative to the relationship(s) between each of the independent and dependent variables. Including all these graphs in this report was not feasible because so many variables were included in the analyses. Therefore, in the following sections we graphically present CART models for each of the strongest primary (and secondary and tertiary, when applicable) predictors of biological responses in the entire dataset and the "Low P" and "High P" basin by ecoregion III datasets. Graphical presentations were also limited to CART models where the primary split explained a minimum of 10% of the variance (i.e. partial $r^2 \geq 0.10$) in the bioassessment or biological (dependent) variable. The complete results of all possible CART models using these datasets are available in Appendices 5-1 through 5-3.

Table 5-1. Nutrient thresholds identified in the primary split of CART analyses for the entire bioassessment median dataset and for datasets that were grouped according to the "Low P" and "High P" basin by ecoregion III classification.

	ALL DATA			"Low P" Basin x Ecoregion III			"High P" Basin x Ecoregion III		
BIOLOGICAL VARIABLE	TP (mg/L)	TN (mg/L)	NO_x-N (mg/L)	TP (mg/L)	TN (mg/L)	NO_x-N (mg/L)	TP (mg/L)	TN (mg/L)	NO_x-N (mg/L)
Fish IBI	0.065[3]	1.3[2]	0.87[4]	0.088[2]	1.3[3]	1.0[4]	0.14[2]	2.5[3]	1.8[4]
RBIBI	0.059[1]	2.5[3]	0.028[4]	0.24[2]	1.1[4]	0.16[3]	0.073[2]	na	na

[1]Strongest predictor of biological response between the three nutrient variables and habitat quality index
[2]Second Strongest predictor of biological response between the three nutrient variables and habitat quality index
[3]Second weakest predictor of biological response between the three nutrient variables and habitat quality index
[4]Weakest predictor of biological response between the three nutrient variables and habitat quality index

CART Analysis of Bioassessment and Other Biological Responses to Nutrients and Habitat

Several CART models were developed from the complete bioassessment median dataset on Texas streams and rivers. The models presented in this section met the minimum requirements for inclusion in the mainbody of this document.

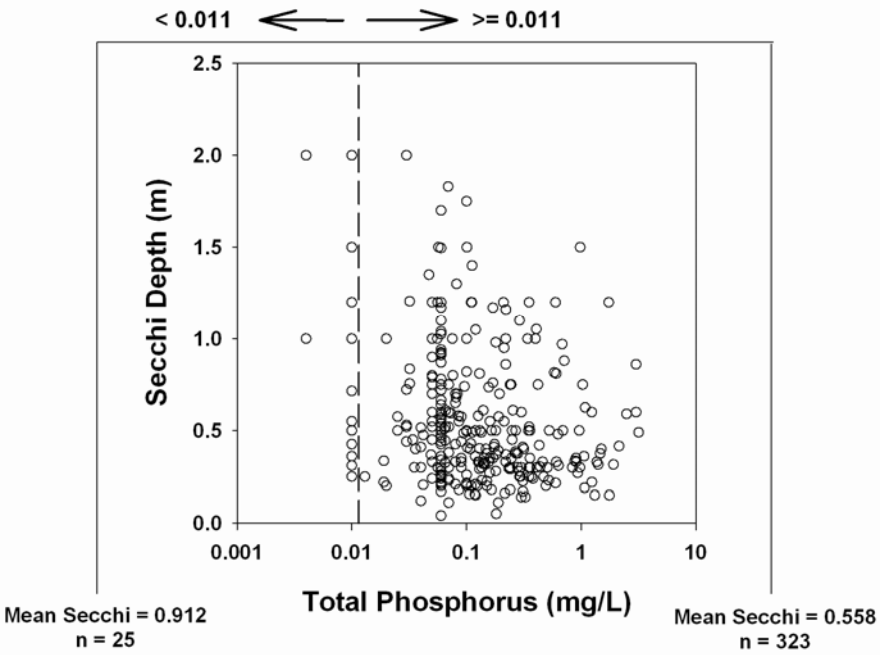


Figure 5-1. CART model for median Secchi depth versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model $r^2 = 0.12$).

Median TP was the strongest predictor of median Secchi depth in the bioassessment dataset (Figure 5-1). Median Secchi depth was approximately 0.9 m when the median TP in the water column was less than 0.011 mg/L. When TP equaled or exceeded the 0.011 mg/L threshold, median Secchi depth decreased to 0.6 m. There were no statistically valid secondary splits in this CART model, and the single threshold in median TP concentrations explained 12% of the variation in median Secchi depth across all river and stream stations included in the bioassessment dataset. However, visual inspection of the data (TP vs. Secchi depth) suggested that the censored values may have influenced the identified threshold identified in this analysis. This would be an important facet to investigate in future analysis.

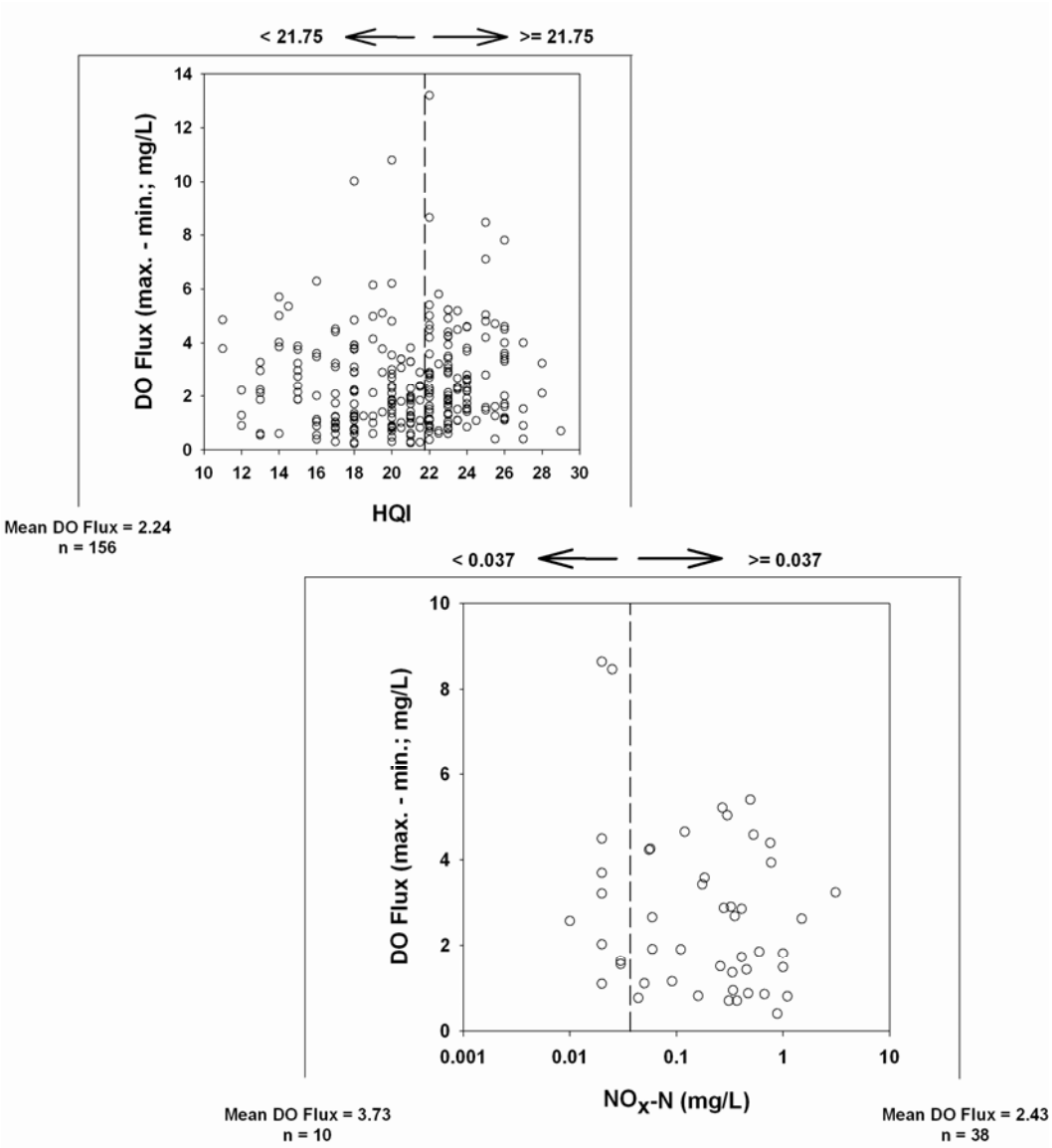


Figure 5-2. CART model for median dissolved oxygen flux versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.43).

HQI was the strongest predictor of median 24-hour dissolved oxygen flux in the bioassessment data (Figure 5-2). Median DO flux was approximately 2.2 mg/L when HQI was less than 21.8. NO_X-N was an important secondary predictor when HQI equaled or exceeded 21.8. DO flux was greatest (3.7 mg/L) when HQI equaled or exceeded the 21.8 index threshold and NO_X-N was less than 0.04 mg/L. However, when HQI equaled or exceeded the 21.8 index value threshold and NO_X-N was greater than 0.04 mg/L, the median DO flux was only 2.4 mg/L. This pattern may make sense ecologically, depending on how HQI might predict algal biomass and productivity in streams and rivers. However, greater DO flux associated with lower inorganic N (NO_X-N) does fit a predictable ecological pattern.

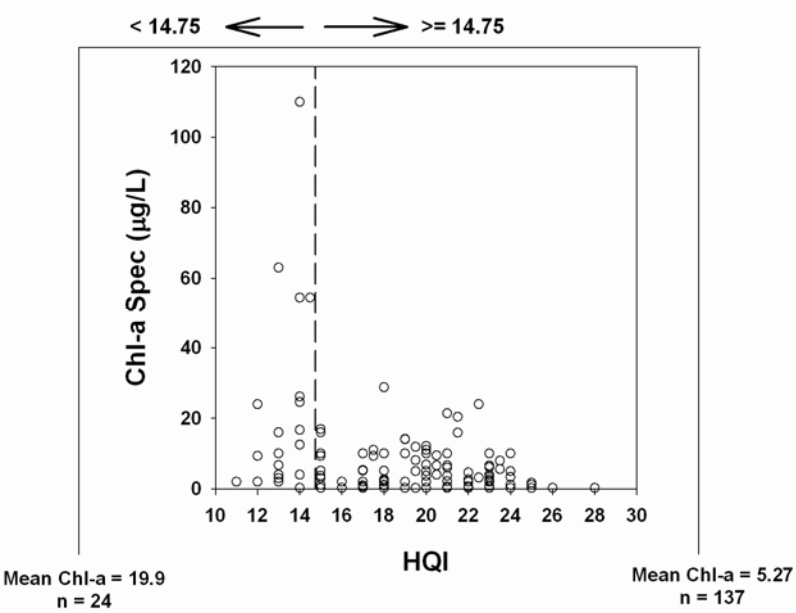


Figure 5-3. CART model for median chl-a concentrations determined by spectrophotometry versus median nutrient concentrations (TP, TN, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.20).

HQI was the strongest predictor of median chl-a concentrations measured with spectrophotometry in the bioassessment dataset (Figure 5-3). Median chl-a was approximately 19.9 µg/L when the HQI values was less than 14.8. When HQI equaled or exceeded the 14.8 threshold, median chl-a decreased to 5.3 µg/L. There were no statistically valid secondary splits in this CART model, and the single threshold in HQI explained 20% of the variation in median chl-a concentrations measured with spectrophotometry across all stations included in the bioassessment dataset.

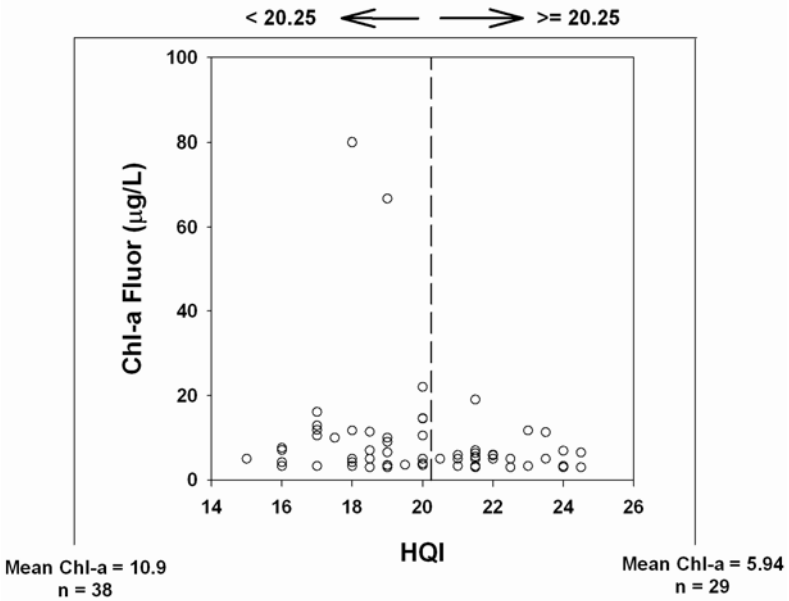


Figure 5-4. CART model for median chl-a concentrations determined by fluorometry versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.33).

HQI was the strongest predictor of median chl-a concentrations measured with fluorometry in the bioassessment dataset (Figure 5-4). Median chl-a was approximately 10.9 µg/L when the HQI values was less than 20.3. When HQI equaled or exceeded the 20.3 threshold, median chl-a decreased to 5.9 µg/L. There were no statistically valid secondary splits in this CART model, and the single threshold in HQI explained 33% of the variation in median chl-a concentrations measured with fluorometry across all stations included in the bioassessment dataset.

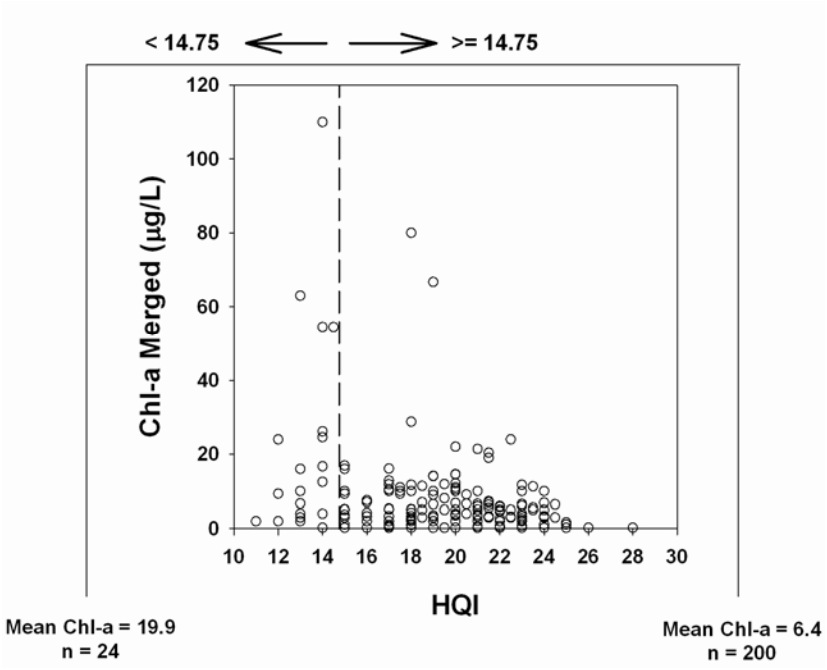


Figure 5-5. CART model for median merged chl-a data versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.23).

HQI was the strongest predictor of median merged chl-a concentrations in the bioassessment dataset (Figure 5-5). This model was most similar to the CART model constructed for chl-a measured with spectrophotometry from the bioassessment database. Median chl-a was approximately 19.9 µg/L when the HQI values was less than 14.8. When HQI equaled or exceeded the 14.8 threshold, median chl-a decreased to 6.4 µg/L. There were no statistically valid secondary splits in this CART model, and the single threshold in HQI explained 23% of the variation in median chl-a concentrations measured with fluorometry across all stations included in the bioassessment dataset.

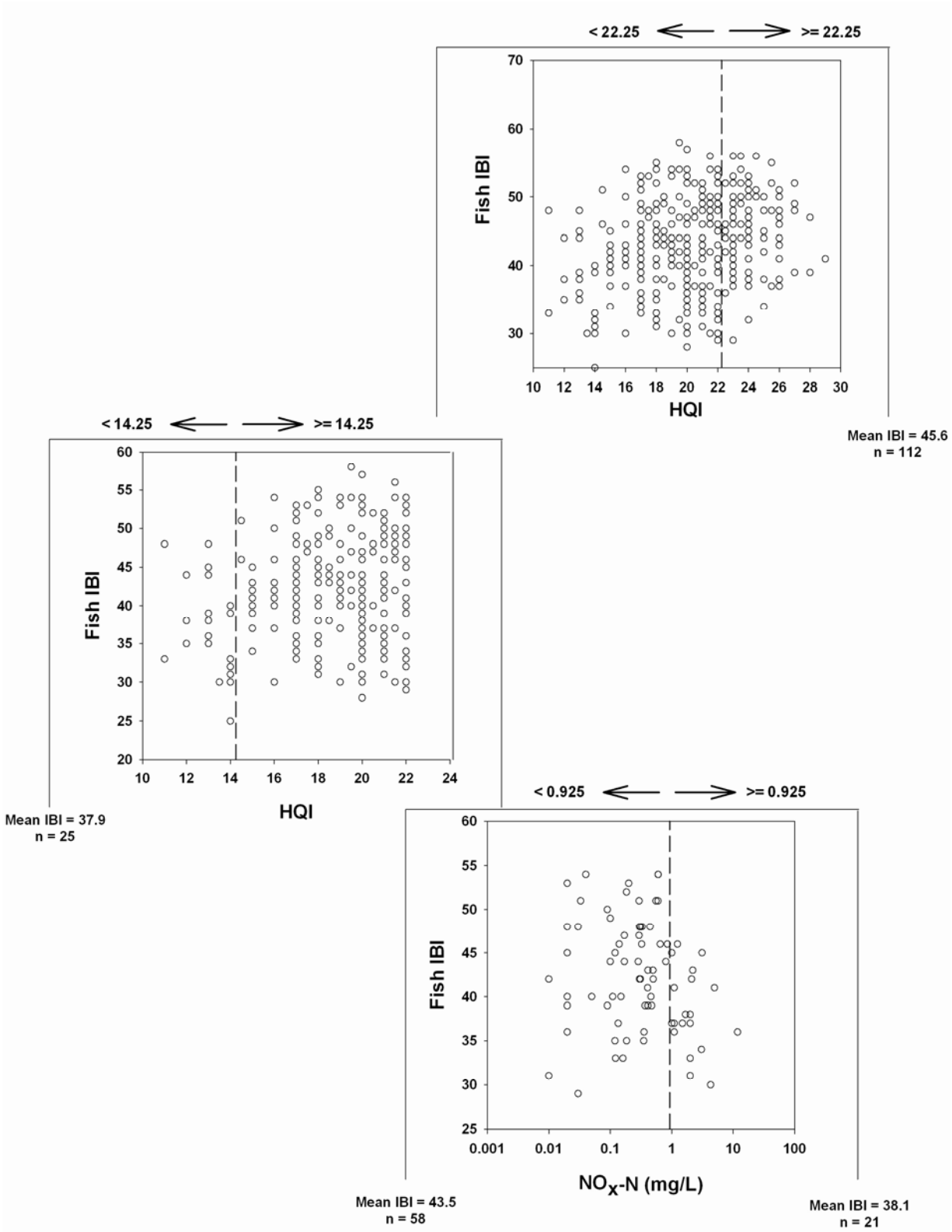


Figure 5-6. CART model for fish IBI versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.60).

HQI was the strongest predictor of Fish IBI in the bioassessment data (Figure 5.6). Fish IBI was approximately 45.6 when HQI was greater than 22.3. HQI was also an important secondary predictor following the initial HQI threshold. Fish IBI was least (37.9) when HQI was less than 14.3. However, when HQI was less than 22.5 but greater than 14.3, NO_X-N was an important tertiary predictor of Fish IBI. At these intermediate HQI values (14.3 – 22.5), Fish IBI was less (38.1) when NO_X-N was greater than 0.9 mg/L. The greatest Fish IBI values were observed at intermediate HQI values (14.3 – 22.5) and NO_X-N concentrations greater than 0.9 mg/L.

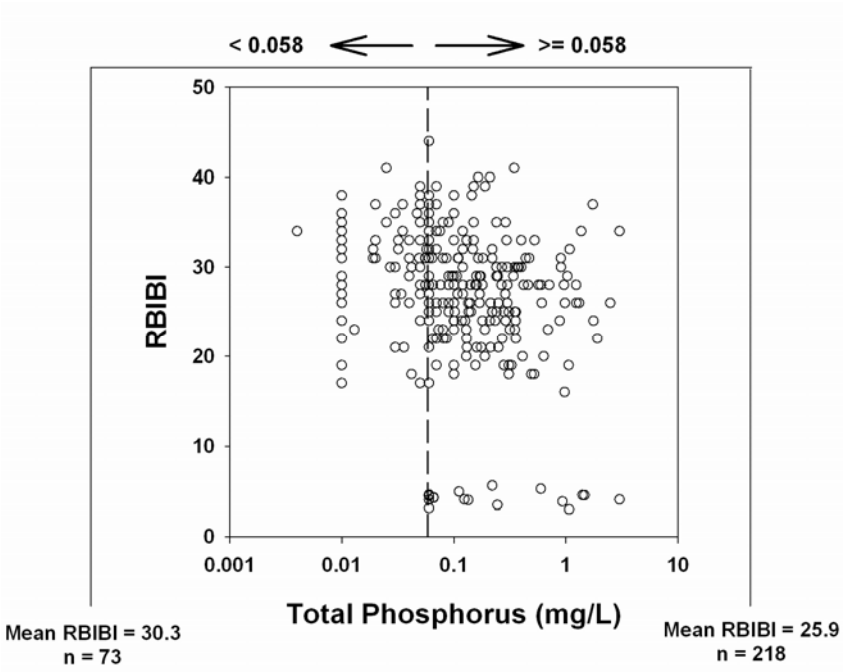


Figure 5-7. CART model for RBIBI versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment median dataset (model r^2 = 0.21).

Median TP was the strongest predictor of RBIBI in the bioassessment dataset (Figure 5-7). RBIBI was approximately 30.3 when the median TP in the water column was less than 0.058 mg/L. When TP equaled or exceeded the 0.058 mg/L threshold, RBIBI decreased to 25.9. There were no statistically valid secondary splits in this CART model, and the single threshold in median TP concentrations explained 21% of the variation in RBIBI across all river and stream stations included in the bioassessment dataset.

The results of CART modeling on the entire bioassessment median database in this study indicated that nutrient thresholds can explain some variability in bioassessment metrics for Texas streams and rivers. However, HQI was a very important covariate. The nutrient thresholds derived from the entire median bioassessment database (Table 5-1) were slightly lower than the nutrient thresholds derived for the common biological measurements in the previous chapter (Table 5-1). However, most of the thresholds observed for the common biological measurements in the bioassessment data were very similar to the thresholds observed for these same measurements in the larger water quality median database (Table 5-2).

Table 5-2. Nutrient thresholds identified in the primary split of CART analyses for the entire bioassessment median dataset and for datasets that were grouped according to the "Low P" and "High P" basin by ecoregion III classification.

	ALL MEDIAN DATA			BIOASSESSMENT MEDIAN DATA		
BIOLOGICAL VARIABLE	TP (mg/L)	TN (mg/L)	NO_x-N (mg/L)	TP (mg/L)	TN (mg/L)	NO_x-N (mg/L)
Secchi	0.075[1]	1.5[2]	2.9[3]	0.012[1]	1.8[3]	0.32[4]
DO flux	0.060[1]	0.82[2]	0.10[3]	0.068[4]	0.29[3]	0.04[2]
Chla-spec	0.094[1]	1.6[2]	0.02[3]	0.046[2]	1.6[4]	0.02[3]
Chla-fluor	0.068[1]	1.4[2]	0.36[3]	0.32[2]	na	na
Chla-merge	0.094[1]	1.6[2]	2.9[3]	0.038[2]	1.6[4]	0.11[3]

[1]Strongest predictor of biological response between the three nutrient variables and habitat quality index
[2]Second Strongest predictor of biological response between the three nutrient variables and habitat quality index
[3]Second weakest predictor of biological response between the three nutrient variables and habitat quality index
[4]Weakest predictor of biological response between the three nutrient variables and habitat quality index

CART Analysis of Bioassessment Response to Nutrients in "Low P" Basin by Ecoregion III

CART models were developed for Fish IBI and RBIBI from the "Low P" basin by ecoregion III bioassessment median dataset on Texas streams and rivers. The models presented in this section met the minimum requirements for inclusion in the main body of this document.

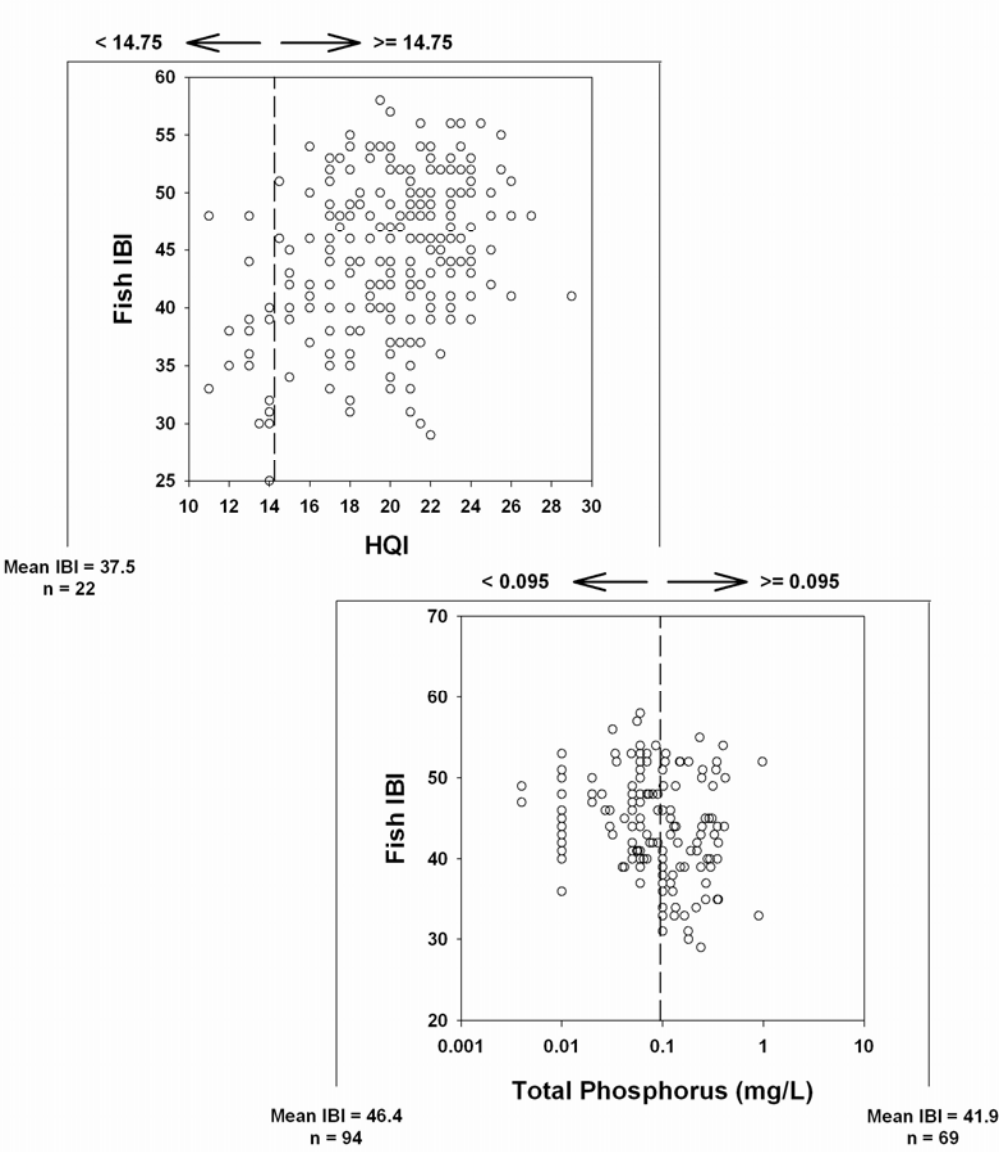


Figure 5-8. CART model for fish IBI versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment "Low P" basin by ecoregion median dataset (model r^2 = 0.48).

HQI was the strongest predictor of Fish IBI in the "Low P" basin by ecoregion III bioassessment data (Figure 5-8). Fish IBI was approximately 37.5 when HQI was less than 14.8. TP concentration was an important secondary predictor when HQI equaled or exceeded the 14.8 threshold. Fish IBI was greatest (46.4) when HQI equaled or exceeded the 14.8 index threshold and TP was less than 0.095 mg/L.

However, when HQI equaled or exceeded the 14.8 index value threshold and TP was greater than 0.095 mg/L, the Fish IBI was only 41.9.

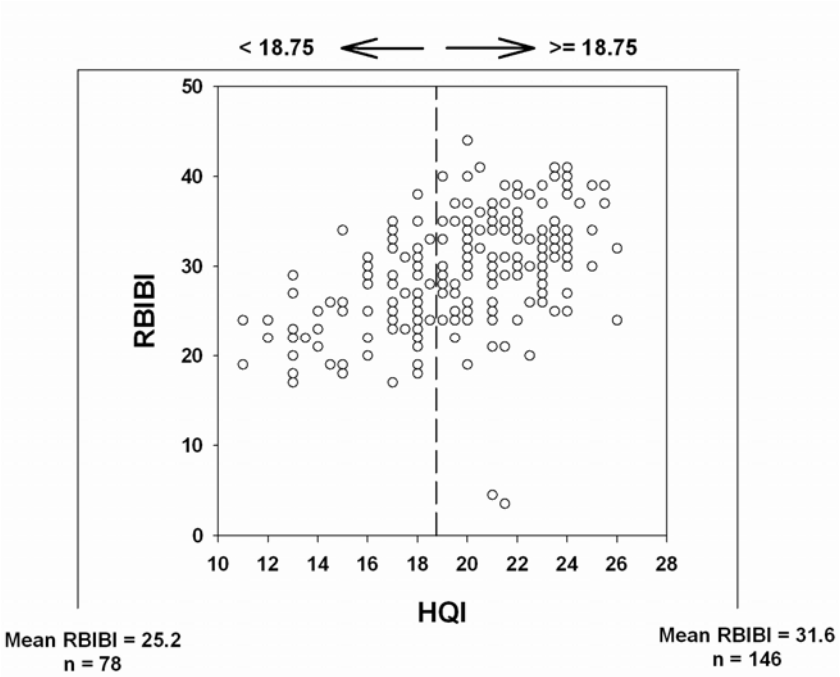


Figure 5-9. CART model for RBIBI vebnbrsus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment "Low P" basin by ecoregion median dataset (model r^2 = 0.29).

HQI was the strongest predictor of RBIBI in the "Low P" basin by ecoregion III bioassessment data (Figure 5-9). RBIBI was approximately 25.2 when HQI was less than 18.8. When HQI equaled or exceeded the 18.8 threshold, RBIBI increased to 31.6. There were no statistically valid secondary splits in this CART model, and the single threshold in HQI explained 29% of the variation in median RBIBI across all stations included in the bioassessment dataset.

The results of CART modeling on the "Low P" basin by ecoregion database in this study suggested that habitat quality was a much more important predictor of bioassessment data for "Low P" basin by ecoregion III geographic areas. One exception was the prediction of Fish IBI from these data. At stations where HQI equaled or exceeded 14.8, Fish IBI was greater when the TP concentration was less than 0.095 mg/L.

CART Analysis of Bioassessment Response to Nutrients in "High P" Basin by Ecoregion III

CART models were developed for Fish IBI and RBIBI from the "High P" basin by ecoregion III bioassessment median dataset on Texas streams and rivers. The models presented in this section met the minimum requirements for inclusion in the main body of this document.

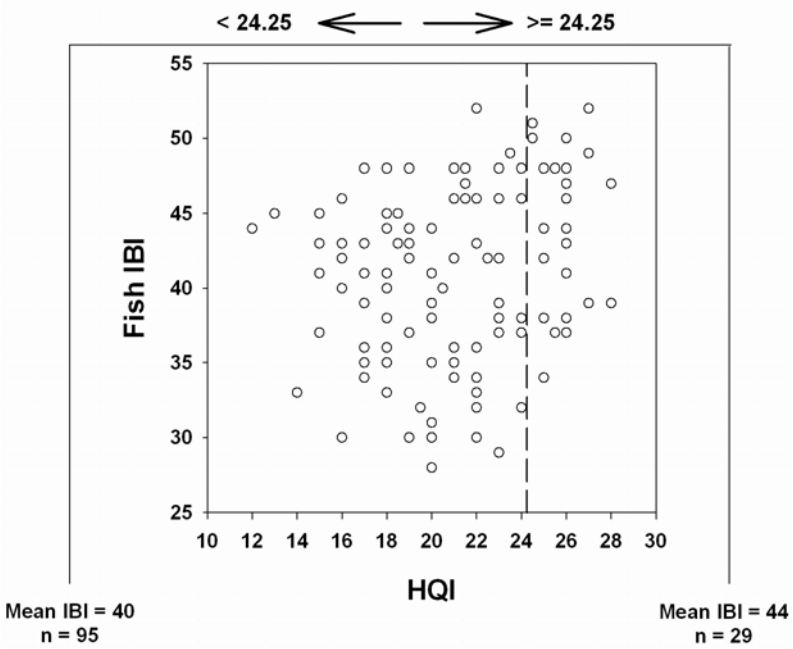


Figure 5-10. CART model for fish IBI versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment "High P" basin by ecoregion median dataset (model $r^2 = 0.19$).

HQI was the strongest predictor of Fish IBI in the "High P" basin by ecoregion III bioassessment data (Figure 5-10). Fish IBI was approximately 40 when HQI was less than 24.3. When HQI equaled or exceeded the 24.3 threshold, Fish IBI increased to 44. There were no statistically valid secondary splits in this CART model, and the single threshold in HQI explained 19% of the variation in median RBIBI across all stations included in the bioassessment dataset.

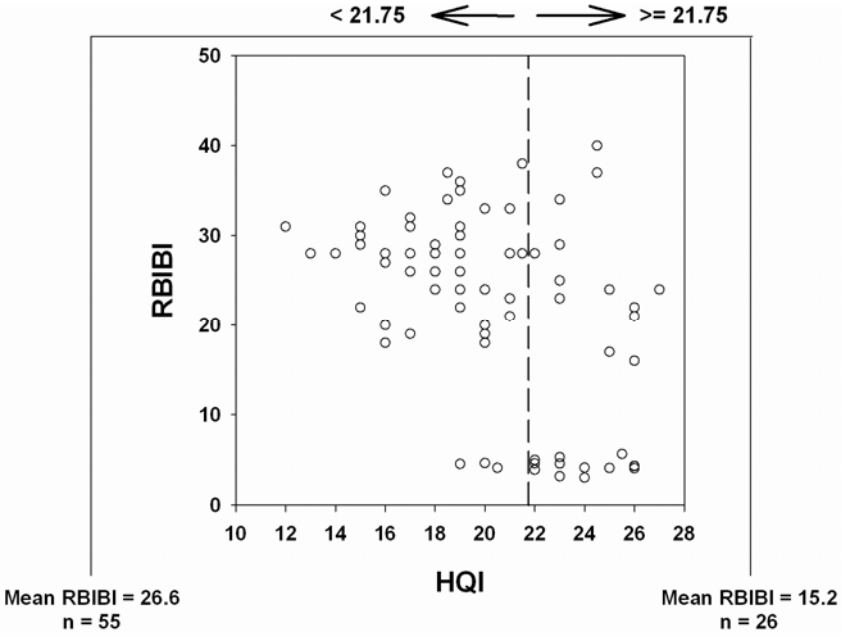


Figure 5-11. CART model for RBIBI versus median nutrient concentrations (TP, T N, and NO_X-N) and habitat quality from the bioassessment "High P" basin by ecoregion median dataset (model $r^2 = 0.28$).

HQI was the strongest predictor of RBIBI in the "High P" basin by ecoregion III bioassessment data (Figure 5-11). RBIBI was approximately 26.6 when HQI was less than 21.8. When HQI equaled or exceeded the 21.8 threshold, RBIBI decreased to 26. There were no statistically valid secondary splits in this CART model, and the single threshold in HQI explained 28% of the variation in median RBIBI across all stations included in the bioassessment dataset.

The results of CART modeling on the "High P" basin by ecoregion database in this study suggested that habitat quality was a much more important predictor of bioassessment data for "High P" basin by ecoregion III geographic areas. Nutrient concentrations were never the strongest primary split in CART models and did not provide any secondary splits using these data.

REFERENCES

De'ath G, Fabricus KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

East T. L., Sharfstein B. 2006. Development of a decision tree model for the prediction of the limitation potential of phytoplankton in Lake Okeechobee, Florida, USA. Arch. Hydrobiol. 2006;165:127-144.

King RS, Baker ME, Whigham DF, Weller DE, Jordan TE, Kazyak PF, Hurd MK. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

Qian SS, King RS, Richardson CJ. 2003. Two methods for the detection of environmental thresholds. Ecological Modeling 166:87-97.

Urban, D.L. (2002) "Classification and regression trees" In: McCune, B., Grace, J.B. eds. , Analysis of Ecological Communities, MjM Software Design, Gleneden Beach, Oregon, USA, pp 221-231.

APPENDIX 5-1: CART Models on Biologicals Versus Nutrients and Habitat

Secchi versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = Secchi_00078 ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioassessment, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=385 (68 observations deleted due to missingness)

       CP nsplit rel error   xerror      xstd
1 0.1177987      0 1.0000000 1.006422 0.09465372
2 0.1100048      1 0.8822013 1.194966 0.09539202

Node number 1: 385 observations,    complexity param=0.1177987
  mean=0.5844039, MSE=0.1410318
  left son=2 (323 obs) right son=3 (25 obs), 37 observations remain
  Primary splits:
     tp_00665   < 0.0115 to the right, improve=0.053558410, (37 missing)
     hqi        < 22.75  to the left,  improve=0.047187090, (55 missing)
     tn_00600C  < 1.777  to the right, improve=0.011205630, (244 missing)
     nox_00631C < 0.3175 to the left,  improve=0.009048388, (243 missing)

Node number 2: 323 observations
  mean=0.5583854, MSE=0.1272923

Node number 3: 25 observations
  mean=0.9124, MSE=0.2714282
```

DO flux versus HQI, TP, T N, and NO_x-N

```
Call:
mvpart(form = doflux_89856C ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioassessment, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=318 (135 observations deleted due to missingness)

       CP nsplit rel error   xerror     xstd
1 0.2139459      0 1.0000000 1.012406 0.1561419
2 0.1252262      2 0.5721082 1.170550 0.1501700

Node number 1: 318 observations,    complexity param=0.2139459
  mean=2.493197, MSE=3.255269
  left son=2 (156 obs) right son=3 (120 obs), 42 observations remain
  Primary splits:
     hqi       < 21.75  to the left,  improve=0.01877992, (42 missing)
     nox_00631C < 0.042  to the right, improve=0.01619666, (172 missing)
     tn_00600C  < 0.2925 to the right, improve=0.01305175, (175 missing)
     tp_00665   < 0.0675 to the right, improve=0.01074530, (39 missing)

Node number 2: 156 observations
  mean=2.243878, MSE=2.825065

Node number 3: 120 observations,    complexity param=0.2139459
  mean=2.77925, MSE=3.851316
  left son=6 (38 obs) right son=7 (10 obs), 72 observations remain
  Primary splits:
     nox_00631C < 0.037  to the right, improve=0.028929420, (72 missing)
     tp_00665   < 0.247  to the left,  improve=0.009807842, (23 missing)
     hqi       < 24.75  to the left,  improve=0.008390555, (0 missing)
     tn_00600C  < 1.115  to the left,  improve=0.007414219, (74 missing)

Node number 6: 38 observations
  mean=2.433947, MSE=2.20332

Node number 7: 10 observations
  mean=3.7335, MSE=6.77961
```

Chl-a Spectrophotometry versus HQI, TP, TN, and NO_x-N

```
Call:
mvpart(form = chlaspec_32211 ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioassessment, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=184 (269 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.19649455      0 1.0000000 1.009224 0.4658518
2 0.06062243      1 0.8035054 1.114808 0.3849169

Node number 1: 184 observations,    complexity param=0.1964946
  mean=7.266149, MSE=133.2949
  left son=2 (137 obs) right son=3 (24 obs), 23 observations remain
  Primary splits:
     hqi       < 14.75  to the right, improve=0.177704800, (23 missing)
     tp_00665  < 0.0455 to the left,  improve=0.061607220, (2 missing)
     nox_00631C < 0.0225 to the right, improve=0.017266850, (67 missing)
     tn_00600C  < 1.595  to the right, improve=0.008740794, (68 missing)

Node number 2: 137 observations
  mean=5.274609, MSE=29.49813

Node number 3: 24 observations
  mean=19.88333, MSE=652.7391
```

Chl-a Fluorometry versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = chlafluor_70953 ~ hqi + tp_00665 + tn_00600C +
   nox_00631C, data = bioassessment, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
  n=92 (361 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.33509995      0   1.0000 1.022839 0.4847134
2 0.06830546      1   0.6649 1.101241 0.5026653

Node number 1: 92 observations,    complexity param=0.3351
  mean=8.795217, MSE=153.7959
  left son=2 (29 obs) right son=3 (38 obs), 25 observations remain
  Primary splits:
     hqi     < 20.25  to the right, improve=0.028904810, (25 missing)
     tp_00665 < 0.3215 to the right, improve=0.008437307, (10 missing)

Node number 2: 29 observations
  mean=5.936897, MSE=11.89253

Node number 3: 38 observations
  mean=10.92342, MSE=238.4983
```

Chl-a merge versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = chlamerge_70953C ~ hqi + tp_00665 + tn_00600C +
   nox_00631C, data = bioassessment, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=269 (184 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.22452032      0 1.0000000 1.004778 0.3429067
2 0.06245334      1 0.7754797 1.138603 0.2975405

Node number 1: 269 observations,    complexity param=0.2245203
  mean=7.762384, MSE=143.6523
  left son=2 (200 obs) right son=3 (24 obs), 45 observations remain
  Primary splits:
     hqi       < 14.75   to the right, improve=0.100860700, (45 missing)
     tp_00665  < 0.038   to the left,  improve=0.042966470, (12 missing)
     nox_00631C < 0.11475 to the right, improve=0.011512330, (140 missing)
     tn_00600C  < 1.595   to the right, improve=0.007277724, (141 missing)

Node number 2: 200 observations
  mean=6.396907, MSE=71.50354

Node number 3: 24 observations
  mean=19.88333, MSE=652.7391
```

Fish IBI versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = fish_ibi ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioassessment, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=422 (31 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.20011262      0 1.0000000 1.007439 0.05957296
2 0.07560369      3 0.3996621 1.016993 0.06366499

Node number 1: 422 observations,    complexity param=0.2001126
  mean=43.29384, MSE=42.29754
  left son=2 (255 obs) right son=3 (112 obs), 55 observations remain
  Primary splits:
     hqi       < 22.25  to the left,  improve=0.04607002, (55 missing)
     tn_00600C  < 1.2725 to the right, improve=0.03926833, (279 missing)
     tp_00665   < 0.0645 to the right, improve=0.03620055, (71 missing)
     nox_00631C < 0.8675 to the right, improve=0.03496700, (277 missing)

Node number 2: 255 observations,    complexity param=0.2001126
  mean=42.39216, MSE=45.04621
  left son=4 (25 obs) right son=5 (230 obs)
  Primary splits:
     hqi       < 14.25  to the left,  improve=0.04826007, (0 missing)
     nox_00631C < 0.9955 to the right, improve=0.03090909, (165 missing)
     tp_00665   < 0.078  to the right, improve=0.02587301, (49 missing)
     tn_00600C  < 1.2725 to the right, improve=0.02545707, (165 missing)

Node number 3: 112 observations
  mean=45.64286, MSE=32.19388

Node number 4: 25 observations
  mean=37.92, MSE=37.1936

Node number 5: 230 observations,    complexity param=0.2001126
  mean=42.87826, MSE=43.48953
  left son=10 (21 obs) right son=11 (58 obs), 151 observations remain
  Primary splits:
```

nox_00631C < 0.925 to the right, improve=0.045602100, (151 missing)
tn_00600C < 2.485 to the right, improve=0.036723510, (151 missing)
tp_00665 < 0.095 to the right, improve=0.028091890, (47 missing)
hqi < 20.25 to the left, improve=0.004952729, (0 missing)

Node number 10: 21 observations
mean=38.09524, MSE=18.6576

Node number 11: 58 observations
mean=43.53448, MSE=38.04191

RBIBI versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = rbibi ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioassessment, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=358 (95 observations deleted due to missingness)

       CP nsplit rel error   xerror      xstd
1 0.2099943      0 1.0000000 1.004490 0.10278899
2 0.1048387      1 0.7900057 0.980418 0.09810986

Node number 1: 358 observations,    complexity param=0.2099943
  mean=27.73338, MSE=62.00805
  left son=2 (218 obs) right son=3 (73 obs), 67 observations remain
  Primary splits:
     tp_00665   < 0.0585 to the right, improve=0.04788490, (67 missing)
     hqi        < 25.25  to the right, improve=0.03809872, (53 missing)
     tn_00600C  < 2.54   to the right, improve=0.01590284, (217 missing)
     nox_00631C < 0.0279 to the left,  improve=0.01080655, (216 missing)

Node number 2: 218 observations
  mean=25.91995, MSE=69.69328

Node number 3: 73 observations
  mean=30.32877, MSE=32.11109
```

APPENDIX 5-2: CART Models on Biologicals versus Nutrients and Habitat from "Low P" Basin by Ecoregion

Fish IBI versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = fish_ibi ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioasseslowp, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=274 (29 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.24173586      0 1.0000000 1.0091878 0.07543198
2 0.08442855      2 0.5165283 0.9039968 0.06448554

Node number 1: 274 observations,    complexity param=0.2417359
  mean=44.49635, MSE=43.92882
  left son=2 (22 obs) right son=3 (221 obs), 31 observations remain
  Primary splits:
     hqi       < 14.25   to the left,  improve=0.10156340, (31 missing)
     tp_00665  < 0.088   to the right, improve=0.09145679, (60 missing)
     tn_00600C  < 1.2725  to the right, improve=0.04729034, (151 missing)
     nox_00631C < 0.9955  to the right, improve=0.03421301, (150 missing)

Node number 2: 22 observations
  mean=37.54545, MSE=37.06612

Node number 3: 221 observations,    complexity param=0.2417359
  mean=45.36199, MSE=38.73774
  left son=6 (69 obs) right son=7 (94 obs), 58 observations remain
  Primary splits:
     tp_00665  < 0.095   to the right, improve=0.09122866, (58 missing)
     hqi       < 22.75   to the left,  improve=0.07321208, (0 missing)
     nox_00631C < 0.925   to the right, improve=0.06314320, (123 missing)
     tn_00600C  < 1.00345 to the right, improve=0.04593535, (124 missing)

Node number 6: 69 observations
  mean=41.94203, MSE=45.4749
Node number 7: 94 observations
  mean=46.37234, MSE=24.08477
```

RBIBI versus HQI, TP, TN, and NO_X-N

```
 Call:
mvpart(form = rbibi ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioasseslowp, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=253 (50 observations deleted due to missingness)

       CP nsplit rel error    xerror      xstd
1 0.2928113      0 1.0000000 1.0076307 0.1056727
2 0.1671904      1 0.7071887 0.8156807 0.1149335

Node number 1: 253 observations,    complexity param=0.2928113
  mean=29.46605, MSE=40.28475
  left son=2 (78 obs) right son=3 (146 obs), 29 observations remain
  Primary splits:
     hqi       < 18.75  to the left,  improve=0.20620090, (29 missing)
     tp_00665  < 0.2425 to the right, improve=0.04101562, (56 missing)
     nox_00631C < 0.155  to the left,  improve=0.02868461, (130 missing)
     tn_00600C  < 1.07   to the right, improve=0.00964115, (131 missing)

Node number 2: 78 observations
  mean=25.17949, MSE=24.81394

Node number 3: 146 observations
  mean=31.60897, MSE=36.11103
```

APPENDIX 5-3: CART Models on Biologicals versus Nutrients and Habitat from "High P" Basin by Ecoregion

Fish IBI versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = fish_ibi ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioasseshighp, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=148 (6 observations deleted due to missingness)

       CP nsplit rel error   xerror      xstd
1 0.1888674      0 1.0000000 1.019322 0.09326711
2 0.1540781      1 0.8111326 1.087264 0.09855365

Node number 1: 148 observations,    complexity param=0.1888674
  mean=41.06757, MSE=31.64408
  left son=2 (95 obs) right son=3 (29 obs), 24 observations remain
  Primary splits:
     hqi       < 24.25  to the left,  improve=0.076193740, (24 missing)
     tp_00665  < 0.1395 to the left,  improve=0.043806270, (11 missing)
     tn_00600C  < 2.485  to the right, improve=0.012341660, (128 missing)
     nox_00631C < 1.84   to the right, improve=0.007773736, (127 missing)
Node number 2: 95 observations
  mean=39.95789, MSE=32.92454
Node number 3: 29 observations
  mean=43.96552, MSE=23.13674
```

RBIBI versus HQI, TP, TN, and NO_X-N

```
Call:
mvpart(form = rbibi ~ hqi + tp_00665 + tn_00600C + nox_00631C,
   data = bioasseshighp, xval = 10, method = "anova", minsplit = 20,
   minbucket = 10)
 n=105 (49 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.27748386      0 1.0000000 1.0389534 0.1475469
2 0.05181039      1 0.7225161 0.8330359 0.1431550

Node number 1: 105 observations,    complexity param=0.2774839
  mean=23.55848, MSE=89.68732
  left son=2 (26 obs) right son=3 (55 obs), 24 observations remain
  Primary splits:
     hqi      < 21.75 to the right, improve=0.24172470, (24 missing)
     tp_00665 < 0.073 to the left,  improve=0.02323527, (11 missing)

Node number 2: 26 observations
  mean=15.21115, MSE=144.373

Node number 3: 55 observations
  mean=26.56636, MSE=55.46108
```

Chapter 6: Threshold Analyses on Reservoir Data

J. Thad Scott

Assistant Professor, Crop, Soil, and Environmental Sciences Department, UA Division of Agriculture
115 Plant Science Building, Fayetteville, AR 72701
Corresponding Author: jts004@uark.edu

EXECUTIVE SUMMARY: We used categorical and regression tree analyses on median and raw data from Texas reservoirs to identify thresholds which resulted in substantial ecological change. Identical Classification and regression tree (CART) models were built for total phosphorus (TP) versus Secchi depth and chlorophyll-a (chl-a) concentrations for both the median and raw data. We also used CART to identify temporal thresholds in chl-a concentrations in a select number of reservoirs to determine if methodological changes in chl-a determination that occurred in the early 2000's may be affecting temporal trends. The most consistent threshold in TP that were correlated with changes in Secchi depth and chl-a concentrations in the reservoir datasets was approximately 0.04 mg/L. A TP threshold of 0.060 mg/L was the strongest predictor of chl-a concentrations in the raw data, but this model was much weaker than all the others. CART analyses identified statistically valid temporal threshold in chl-a concentrations at all lakes tested. However, these results were inconsistent and only four of the temporal thresholds appeared to be related to changes in method detection limits for chl-a.

INTRODUCTION

The objective of this chapter was to identify total phosphorus (TP) concentrations that were correlated with a change in the magnitude or variability of commonly measured biological variables in Texas reservoirs. The analyses conducted in this chapter focused on both median water quality values in order to capture thresholds that were important at broad temporal and geographic scales. Furthermore, we were also interested in identifying temporal thresholds at which chlorophyll-a (chl-a) concentrations may have shifted, in order to quantify the importance of a methodological changes that occurred across the period of data.

Classification and regression tree (CART) analysis is an empirical modeling technique that is useful for identifying ecological thresholds and hierarchical structure in predictor variables (De'ath and Fabricius 2000). CART uses recursive partitioning to divide data into subsets that are increasingly homogeneous, invoking a tree-like classification that can explain relationships that may be difficult to reconcile with conventional linear models (Urban 2002). Categorical variables (e.g., station location, basin, ecoregion or land-use classifications) may also be used as independent variables in CART analysis, which provides another advantage to using CART rather than traditional regression techniques. CART and other similar methods have been used to identify thresholds and hierarchical structure in environmental correlates of various biological processes in aquatic ecosystems (King et al. 2005, East and Sharfstein 2006). King et al. (2005) used CART to specifically identify thresholds in nutrient concentrations which resulted in shifts in ecological structure and function. These thresholds were used to recommend specific water quality nutrient criteria for the Florida Everglades ecosystem.

METHODS

We conducted CART analyses on the median and raw data databases for reservoirs in order to identify thresholds in TP concentrations that resulted in measurable changes in Secchi depth and chl-a concentrations. We also used CART to identify temporal thresholds in chl-a concentration from ten reservoirs (Lake O' the Pines, Lake Tawakoni, Lake Fork, Lake Houston, Lake Conroe, Lake Whitney, Lake Belton, Lake Travis, Lake Buchanan, Lake Amistad) to determine if a methodological change in the early 2000's affected reported chl-a concentrations.

CART analysis is a form of data reduction that aims to: 1) quantify thresholds in independent variables that are correlated with shifts in the magnitude and/or variability of dependent variables, and 2) identify hierarchical structure in independent variables. CART analysis is very useful for resolving nonlinear, hierarchical, and high-order interactions among predictor variables (De'Ath and Fabricius 2000) and for detecting numerical values that lead to ecological changes (Qian and others 2003). CART models use recursive partitioning to separate data into subsets that are increasingly homogeneous. This iterative process invokes a tree-like classification that can reveal relationships that are often difficult to reconcile with conventional linear models (Urban 2002).

CART analyses were performed using the MVPART library in R 2.8.1 (http://www.r-project.org/). We required a minimum of 20 observations to be used in any single split in the CART model and that each terminal node in the model has a minimum of ten observations. CART analysis is insensitive to missing data. Therefore, we did not remove observations from the data set due to missing values. We first ran CART models using median data from reservoirs and then repeated the same analyses on the raw data from reservoirs. Because CART analysis involves recursive partitioning, models may sometimes be over-fit (i.e. too many independent variables that decrease the statistical rigor of final model). We "pruned" CART models to generate final models that balanced accuracy within the available dataset with robustness to novel data (Urban 2002). CART models were cross-validated to determine "pruning size" (i.e., the number of predictor variables included in the model). Model cross-validations were conducted using 10 random and similarly sized subsets of our data according to the method detailed by De'ath and Fabricius (2000). The optimum tree size for each model was selected using the minimum cross-validated error rule (De'ath and Fabricius 2000).

RESULTS AND DISCUSSION

Analyses using both the median and raw data datasets resulted in statistically-valid thresholds in TP concentrations that were correlated with biological variation (Table 6-1). TP thresholds were generally similar (~0.040 mg/L) across variables and across the median and raw data databases. However, a 0.060 mg/L TP threshold was identified as the best predictor of chl-a concentrations in the raw data database.

Table 6-1. Phosphorus thresholds identified by CART analyses for reservoir median and raw data databases.

BIOLOGICAL VARIABLE	Median Data TP (mg/L)	Raw Data TP (mg/L)
Secchi	0.032	0.039
Chlorophyll-a	0.043	0.060

All of the thresholds reported in Table 6-1 were statistically valid according to our defined methods in CART analysis. But it is still useful to show these thresholds relative to the relationship(s) between each of the independent and dependent variables. In the following sections we graphically present CART models for each of the thresholds identified in Table 6-1. The complete results of all possible CART models using these datasets are available in Appendices 6-1 through 6-3.

Threshold Analysis on Median Reservoir Data

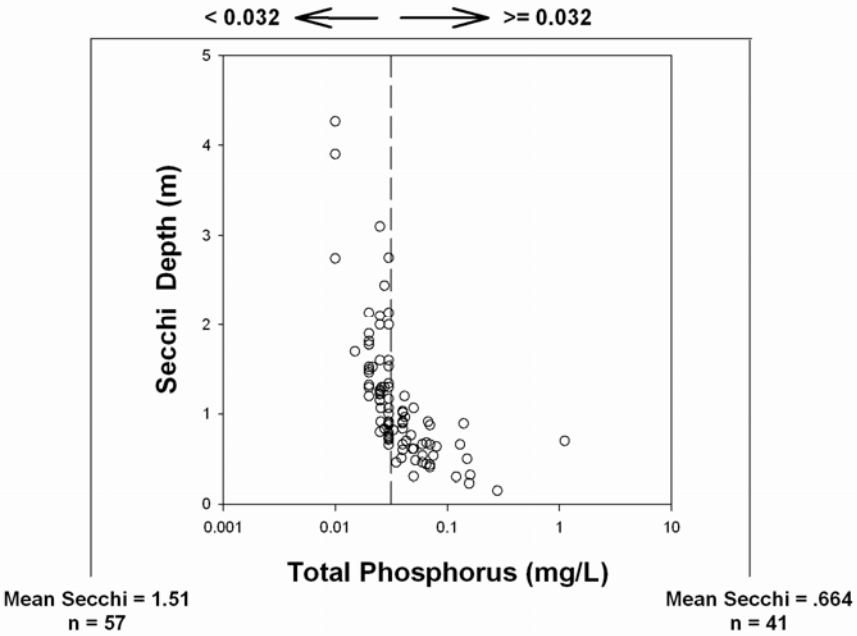


Figure 6-1.CART model for median Secchi depth versus median TP from the reservoir median dataset (model r^2 = 0.34).

Median TP was a strong predictor of median Secchi depth for Texas reservoirs (Figure 6-1). Median Secchi depth was approximately 1.5 m when the median TP in the water column was less than 0.032 mg/L. When TP equaled or exceeded the 0.032 mg/L threshold, median Secchi depth decreased to 0.7 m. The threshold in median TP concentrations explained 34% of the variation in median Secchi depth across all reservoirs.

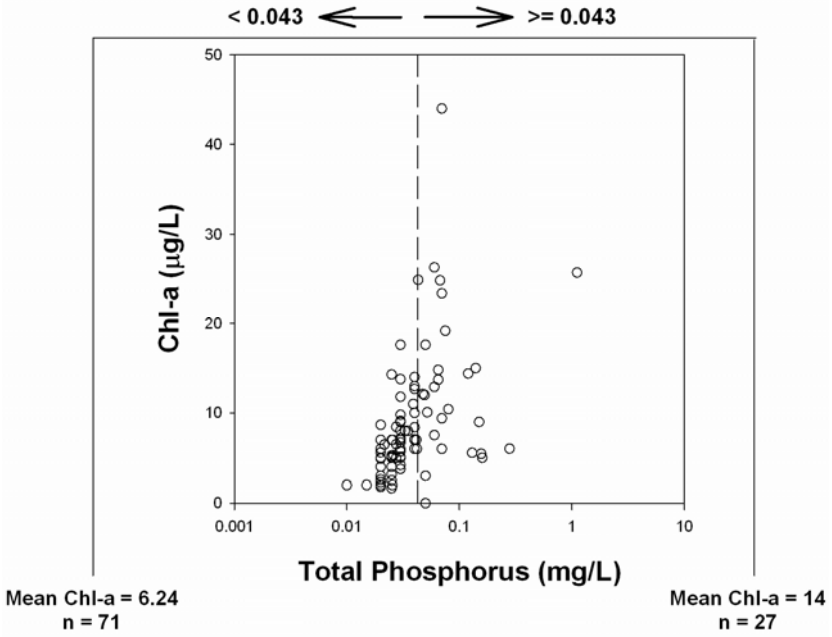


Figure 6-2. CART model for median chl-a concentrations versus median TP concentrations from the reservoir median dataset (model r^2 = 0.27).

Median TP was also a strong predictor of median chl-a concentrations for Texas reservoirs (Figure 6-2). Median chl-a concentrations were approximately 6.2 µg/L when the median TP in the water column was less than 0.043 mg/L. When TP equaled or exceeded the 0.043 mg/L threshold, median chl-a concentrations increased to 14 µg/L. The threshold in median TP concentrations explained 27% of the variation in median chl-a concentration across all reservoirs.

Threshold Analysis on Reservoir Raw Data

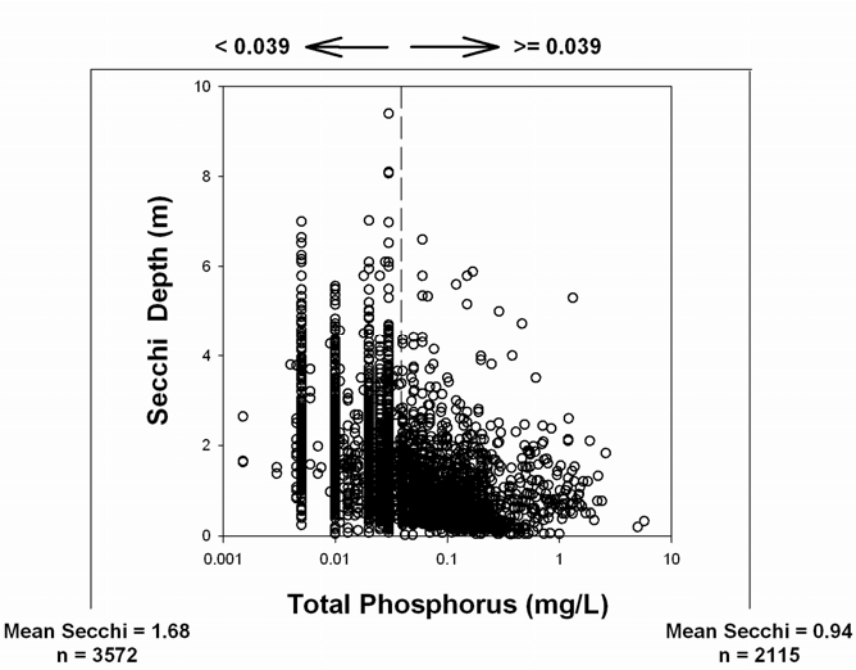


Figure 6-3. CART model for Secchi depth versus TP from the reservoir raw dataset (model r^2 = 0.13).

TP concentration was a reasonable predictor of Secchi depth in the raw data database for Texas reservoirs (Figure 6-3). Secchi depth was approximately 1.7 m when TP in the water column was less than 0.039 mg/L. When TP equaled or exceeded the 0.039 mg/L threshold, Secchi depth decreased to 0.9 m. The threshold in TP concentrations explained 13% of the variation in Secchi depth across all raw data from all the reservoirs.

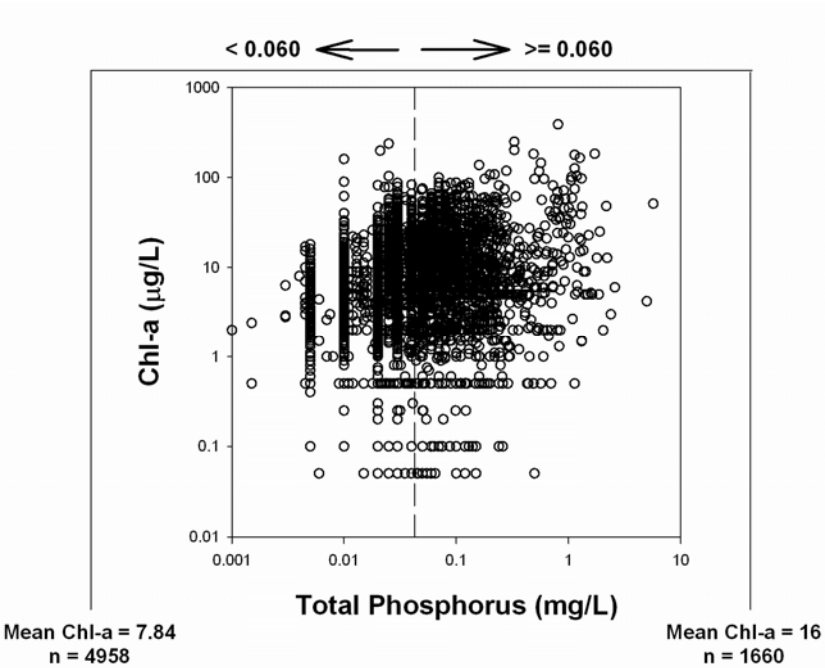


Figure 6-4. CART model for chl-a concentrations versus TP concentrations from the reservoir raw dataset (model r^2 = 0.06).

TP was a weak predictor of chl-a concentrations in the raw data database for Texas reservoirs (Figure 6-4). Chl-a concentrations were approximately 7.8 µg/L when the TP in the water column was less than 0.060 mg/L. When TP equaled or exceeded the 0.060 mg/L threshold, chl-a concentrations increased to 16 µg/L. The threshold in TP concentrations explained only 6% of the variation in chl-a across all reservoirs.

Temporal Changepoints in Reservoir Chlorophyll-a Concentrations

Temporal thresholds in chl-a concentrations were identified in data for all reservoirs examined. The models presented in this section met the minimum requirements of CART analysis, although some of the models have poor predictive power.

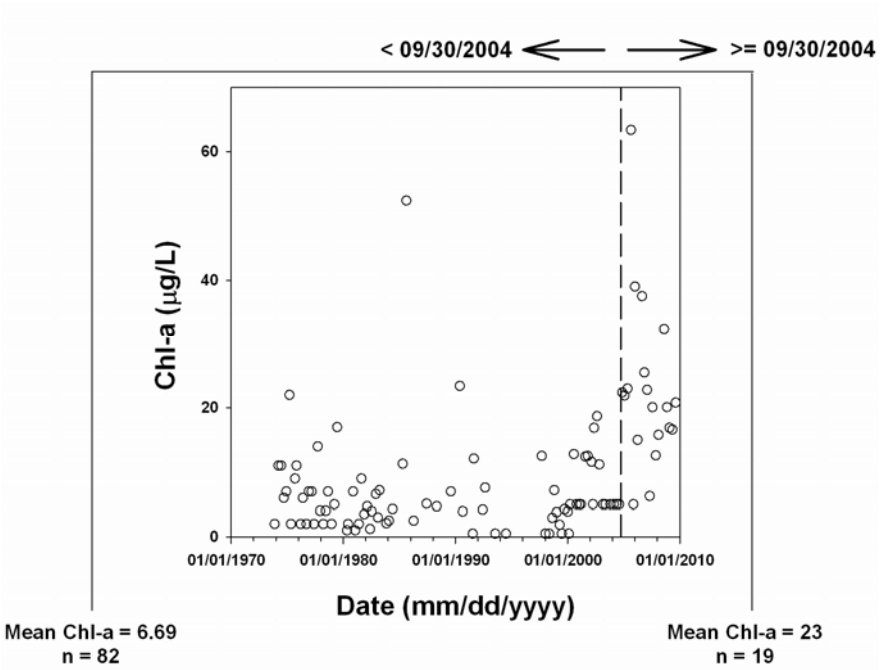


Figure 6-5. CART model for chl-a concentrations through time at Lake O' the Pines (model r^2 = 0.36).

Time was a good predictor of chl-a concentrations in Lake O' the Pines (Figure 6-5). Mean chl-a concentrations prior to September 2004 was 6.7 µg/L and mean chl-a after this data was 19 µg/L. Time explained 36% of the variation in chl-a data over the period of record in the reservoir database.

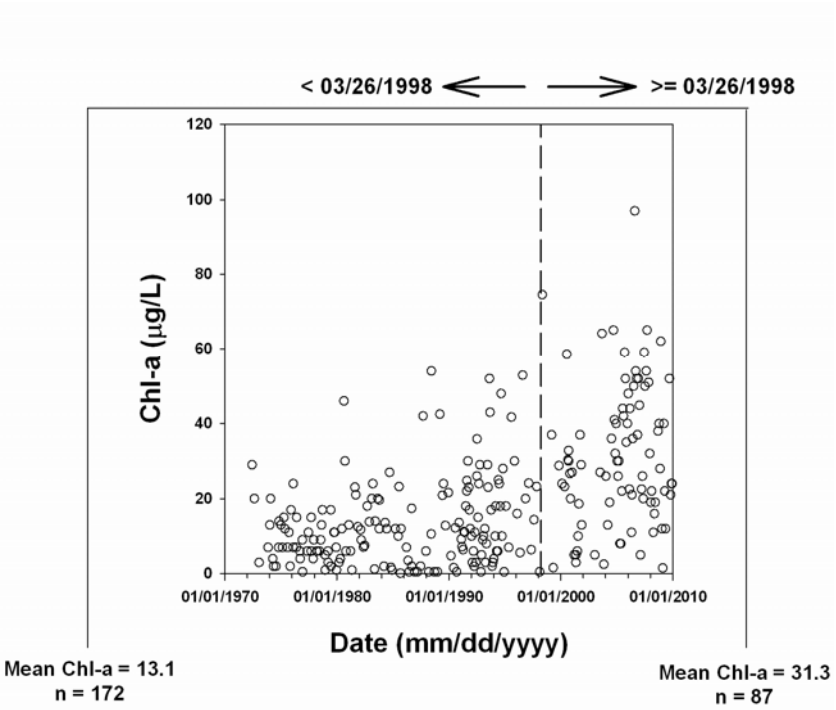


Figure 6.6. CART model for chl-a concentrations through time at Lake Tawakoni (model r^2 = 0.27).

Time was a good predictor of chl-a concentrations in Lake Tawakoni (Figure 6-6). Mean chl-a concentrations prior to March 1998 was 13.1 µg/L and mean chl-a after this data was 31.3 µg/L. Time explained 27% of the variation in chl-a data over the period of record in the reservoir database.

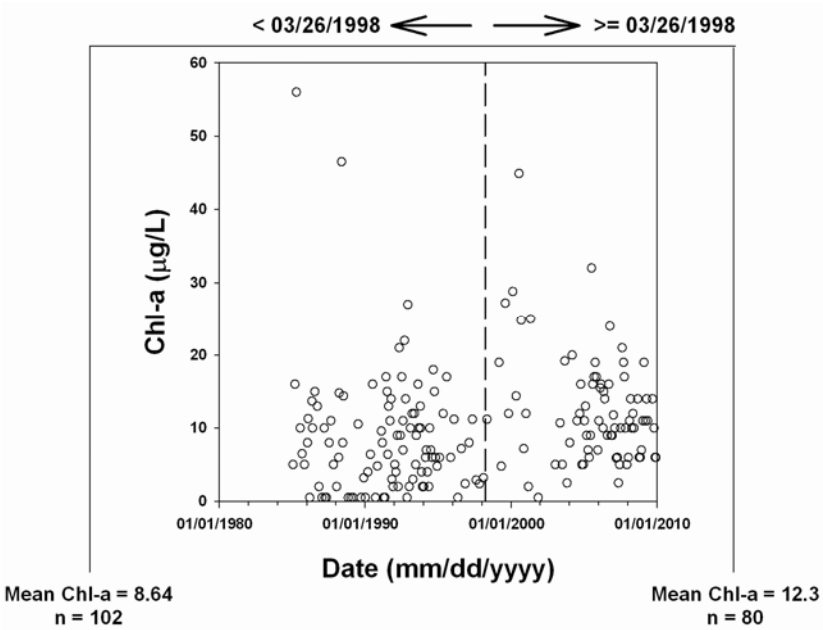


Figure 6-7. CART model for chl-a concentrations through time at Lake Fork (model r^2 = 0.05).

Time was a weak predictor of chl-a concentrations in Lake Fork (Figure 6-7). Mean chl-a concentrations prior to March 1998 was 8.6 µg/L and mean chl-a after this data was 12.3 µg/L. Time explained only 5% of the variation in chl-a data over the period of record in the reservoir database.

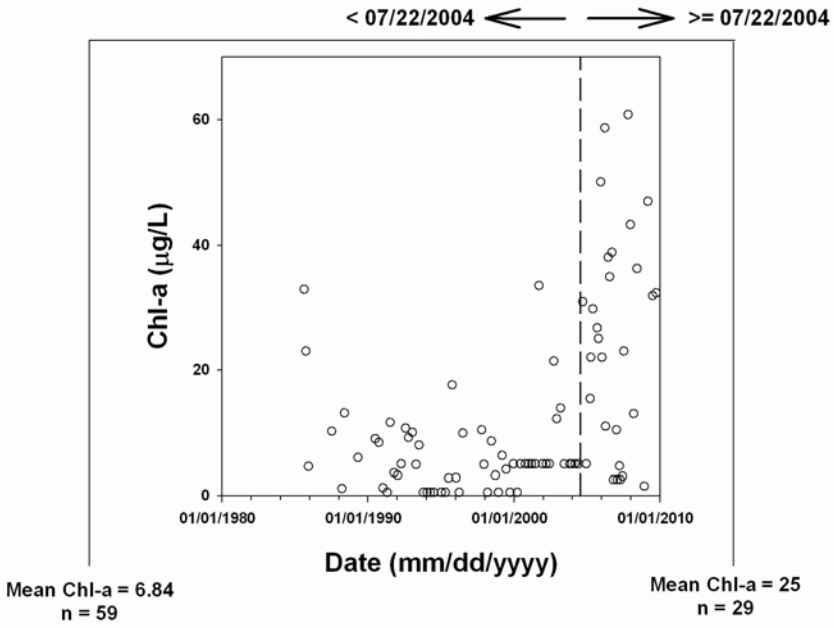


Figure 6-8. CART model for chl-a concentrations through time at Lake Houston (model r^2 = 0.36).

Time was a good predictor of chl-a concentrations in Lake Houston (Figure 6.8). Mean chl-a concentrations prior to July 2004 was 6.8 µg/L and mean chl-a after this data was 25 µg/L. Time explained 36% of the variation in chl-a data over the period of record in the reservoir database.

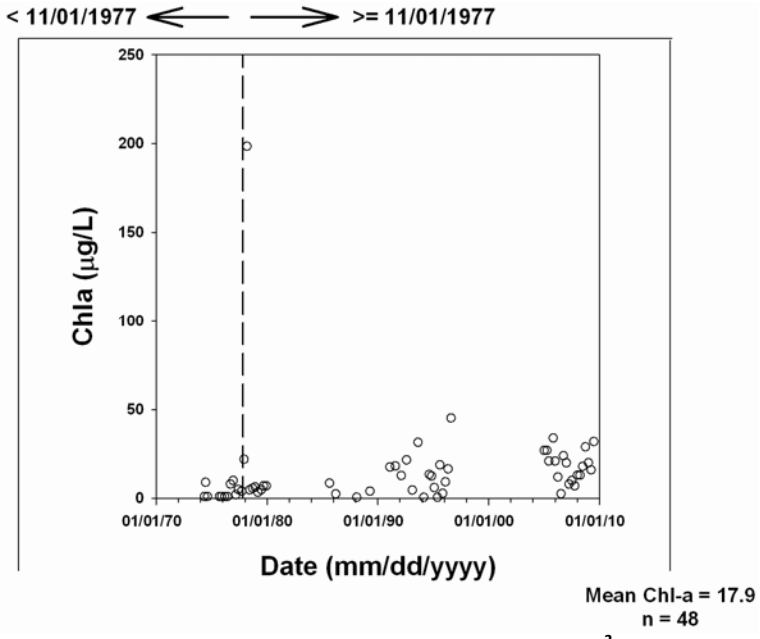


Figure 6-9. CART model for chl-a concentrations through time at Lake Conroe (model r^2 = 0.05).

Time was a weak predictor of chl-a concentrations in Lake Conroe (Figure 6-9). Mean chl-a concentrations prior to November 1977 was 3.7 µg/L and mean chl-a after this data was 17.9 µg/L. Time explained only 5% of the variation in chl-a data over the period of record in the reservoir database.

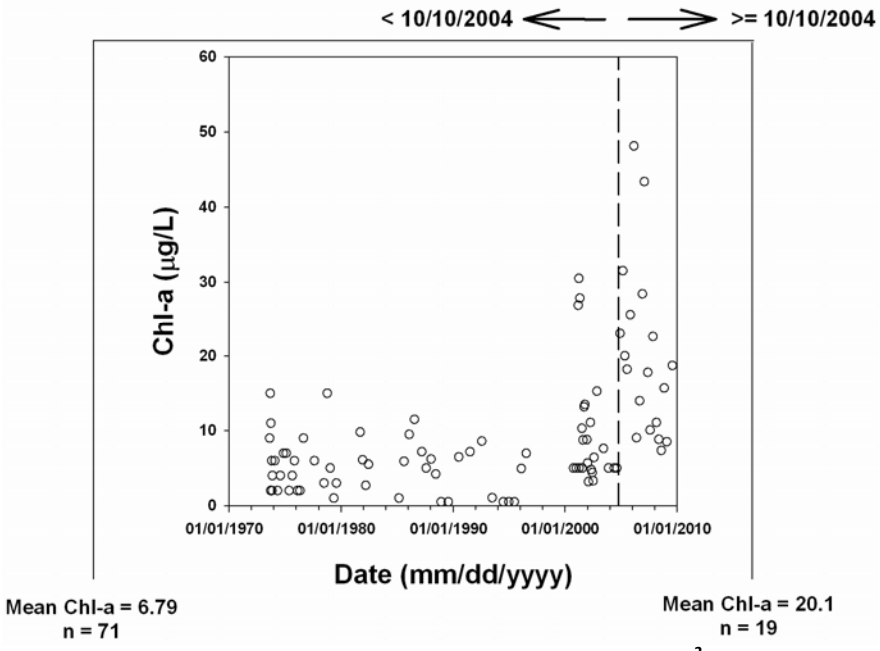


Figure 6-10. CART model for chl-a concentrations through time at Lake Whitney (model r^2 = 0.36).

Time was a good predictor of chl-a concentrations in Lake Whitney (Figure 6-10). Mean chl-a concentrations prior to October 2004 was 6.8 µg/L and mean chl-a after this data was 20.1 µg/L. Time explained 36% of the variation in chl-a data over the period of record in the reservoir database.

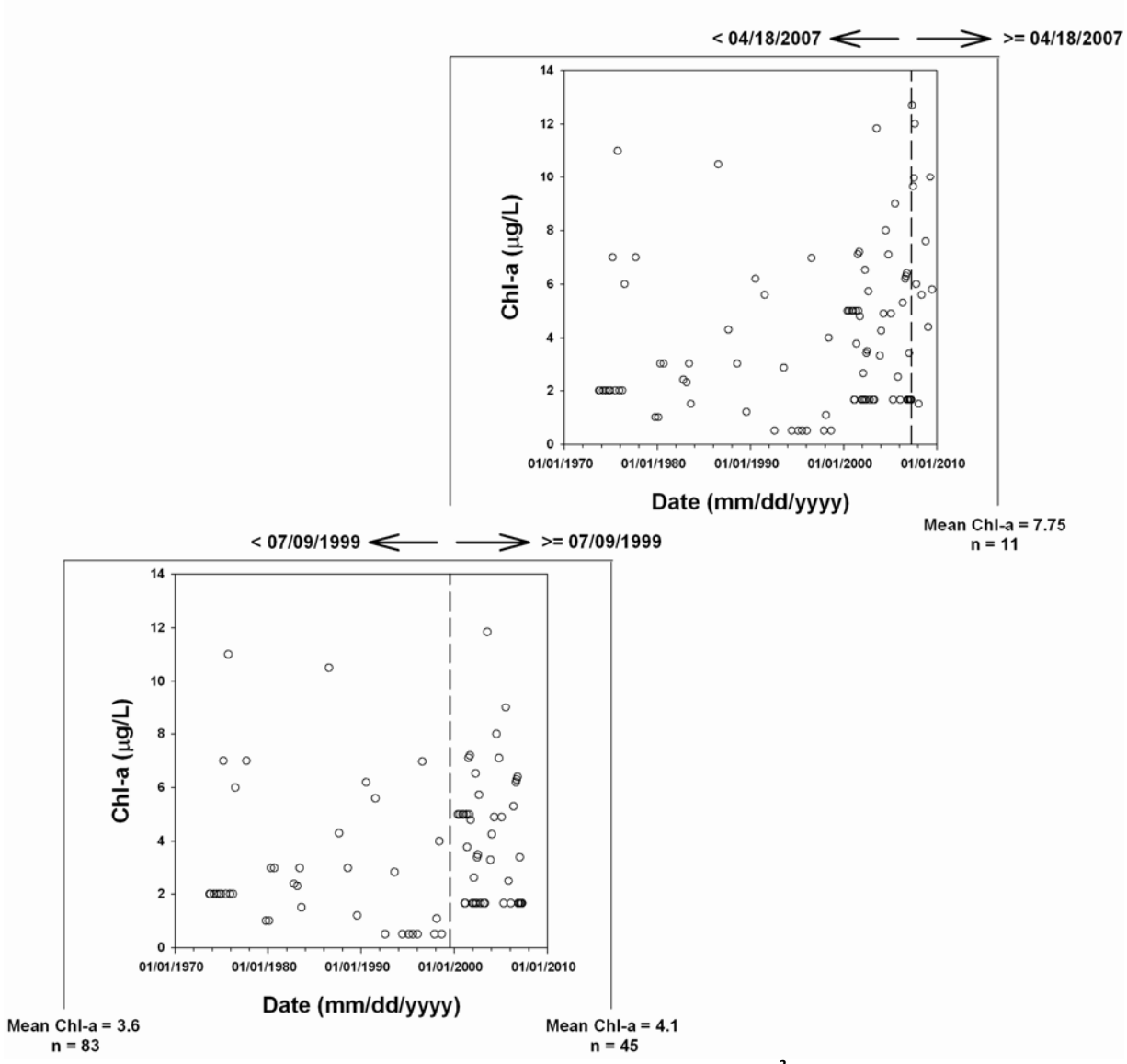


Figure 6-11. CART model for chl-a concentrations through time at Lake Belton (model r^2 = 0.20).

Time was a good predictor of chl-a concentrations in Lake Belton (Figure 6-11). Two temporal thresholds were identified. Mean chl-a concentrations after April 2007 was 7.8 µg/L. Mean chl-a concentrations prior to July 1999 was 3.6 µg/L. A mean chl-a concentration of 4.1 µg/L was observed between July 1999 and April 2007. Time explained a total of 20% of the variation in chl-a data over the period of record in the reservoir database.

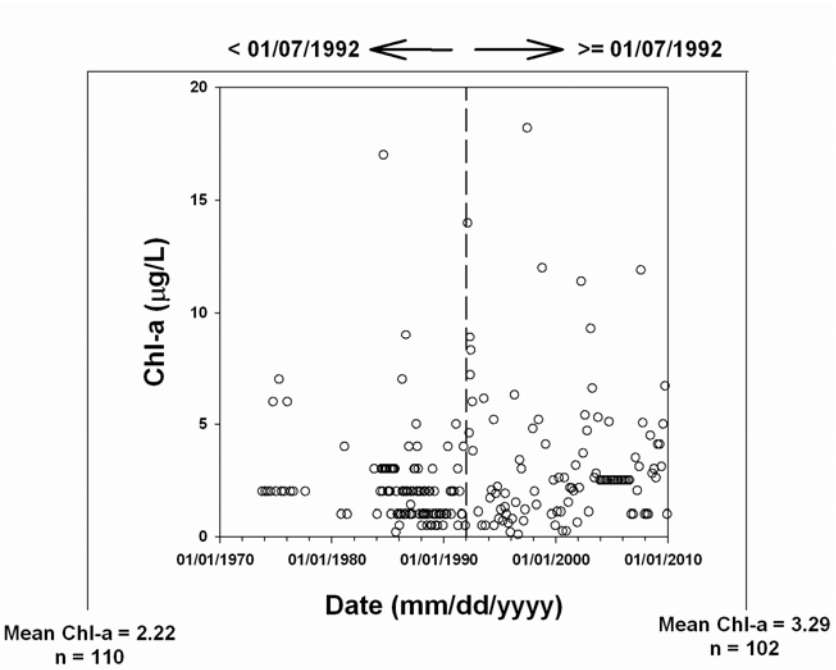


Figure 6-12. CART model for chl-a concentrations through time at Lake Travis (model r^2 = 0.05).

Time was a weak predictor of chl-a concentrations in Lake Travis (Figure 6-12). Mean chl-a concentrations prior to January 1992 was 2.2 µg/L and mean chl-a after this data was 3.3 µg/L. Time explained only 5% of the variation in chl-a data over the period of record in the reservoir database.

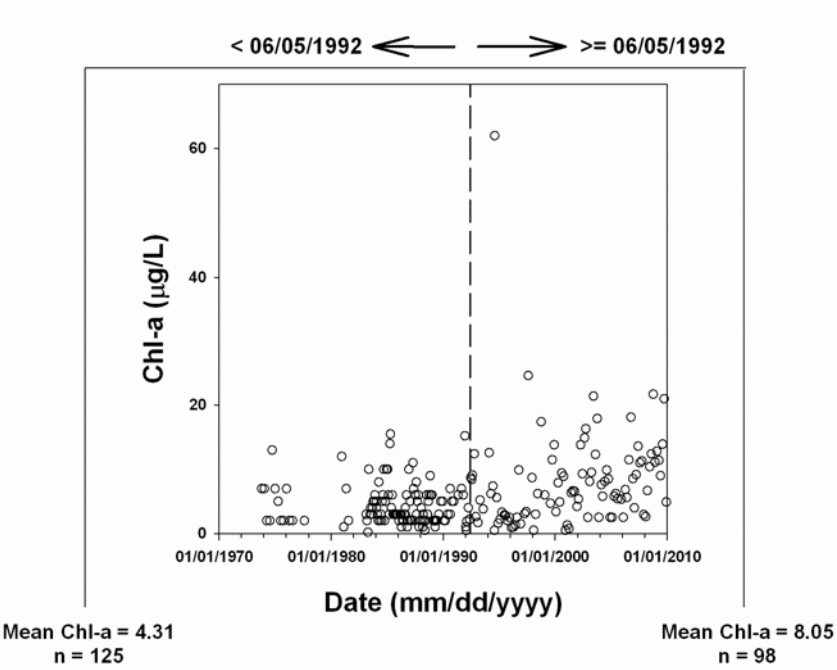


Figure 6-13. CART model for chl-a concentrations through time at Lake Buchanan (model r^2 = 0.10).

Time was a modest predictor of chl-a concentrations in Lake Buchanan (Figure 6-13). Mean chl-a concentrations prior to June 1992 was 4.3 µg/L and mean chl-a after this data was 8.1 µg/L. Time explained 10% of the variation in chl-a data over the period of record in the reservoir database.

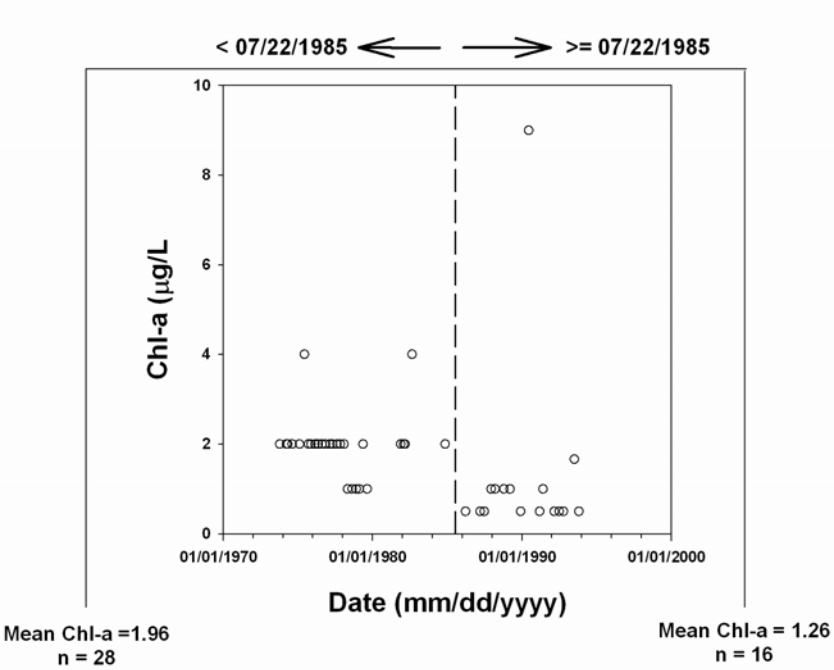


Figure 6.14. CART model for chl-a concentrations through time at Lake Amistad (model r^2 = 0.06).

Time was a weak predictor of chl-a concentrations in Lake Amistad (Figure 6-14). Mean chl-a concentrations prior to July 1985 was 2.0 µg/L and mean chl-a after this data was 1.3 µg/L. Time explained 6% of the variation in chl-a data over the period of record in the reservoir database.

The findings of the temporal trend CART models were somewhat inconclusive. Temporal thresholds were identified for all reservoirs tested, but these times did not always correspond to patterns that were obviously due to methodological change. However, temporal thresholds in some reservoirs (Lake O' the Pines, Lake Whitney, Lake Belton, Lake Amistad) did appear to be related to changes in detection levels for chl-a.

REFERENCES

De'ath G, Fabricus KE. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

East T. L., Sharfstein B. 2006. Development of a decision tree model for the prediction of the limitation potential of phytoplankton in Lake Okeechobee, Florida, USA. Arch. Hydrobiol. 2006;165:127-144.

King RS, Baker ME, Whigham DF, Weller DE, Jordan TE, Kazyak PF, Hurd MK. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

Qian SS, King RS, Richardson CJ. 2003. Two methods for the detection of environmental thresholds. Ecological Modeling 166:87-97.

Urban, D.L. (2002) "Classification and regression trees" In: McCune, B., Grace, J.B. eds. , Analysis of Ecological Communities, MjM Software Design, Gleneden Beach, Oregon, USA, pp 221-231.

APPENDIX 6-1: Threshold Analyses on Median Reservoir Data

Secchi Depth versus TP

```
Call:
mvpart(form = Secchi ~ tp, data = lakes, xval = 10, method = "anova",
   minsplit = 10, minbucket = 5)
 n= 98

       CP nsplit rel error   xerror      xstd
1 0.3424610      0  1.000000 1.036979 0.2645201
2 0.1511482      1  0.657539 0.833815 0.2088858

Node number 1: 98 observations,    complexity param=0.342461
  mean=1.158748, MSE=0.513442
  left son=2 (41 obs) right son=3 (57 obs)
  Primary splits:
     tp < 0.0315 to the right, improve=0.342461, (0 missing)

Node number 2: 41 observations
  mean=0.6643268, MSE=0.06042169

Node number 3: 57 observations
  mean=1.514384, MSE=0.5369879
```

Chl-a versus TP

```
Call:
mvpart(form = chla ~ tp, data = lakes, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n= 98

        CP nsplit rel error    xerror      xstd
1 0.27084635      0 1.0000000 1.0410373 0.3275534
2 0.05269619      1 0.7291537 0.8687543 0.2529734

Node number 1: 98 observations,    complexity param=0.2708463
  mean=8.383367, MSE=44.42739
  left son=2 (71 obs) right son=3 (27 obs)
  Primary splits:
     tp < 0.0425 to the left,  improve=0.2708463, (0 missing)

Node number 2: 71 observations
  mean=6.244225, MSE=11.60531

Node number 3: 27 observations
  mean=14.00852, MSE=87.062
```

APPENDIX 6-2: Threshold Analyses on Reservoir Raw Data

TP vs Chla

Mvpart (chla ~ tp, data = tx_res, xval = 10, method = "anova", minsplit = 15, minbucket = 10)

n=6618 (2069 observations deleted due to missingness)

CP	nsplit	rel error	xerror	xstd
1	0.05677135	0	1.0000000	1.0004735	0.130674
2	0.03098160	1	0.9432286	0.9507254	0.126769

Node number 1: 6618 observations, complexity param=0.05677135
mean=9.878475, MSE=219.1558
left son=2 (4958 obs) right son=3 (1660 obs)
Primary splits:
tp < 0.0595 to the left, improve=0.05677135, (0 missing)

Node number 2: 4958 observations
mean=7.837481, MSE=102.6913

Node number 3: 1660 observations
mean=15.97441, MSE=517.4034

TP vs Secchi Depth

```
mvpart(Secchi ~ tp, data = tx_res, xval = 10, method = "anova", minsplit = 15, minbucket = 10)

n=5687 (3000 observations deleted due to missingness)

CP              nsplit  rel error       xerror          xstd
1       0.1344066       0       1.0000000       1.0004083       0.03631524
2       0.0521979       1       0.8655934       0.8694248       0.03421188
3       0.0100000       2       0.8133955       0.8178509       0.03277695

Node number 1: 5687 observations,    complexity param=0.1344066
  mean=1.404842, MSE=0.9701435
  left son=2 (2115 obs) right son=3 (3572 obs)
  Primary splits:
     tp < 0.0385 to the right, improve=0.1344066, (0 missing)

Node number 2: 2115 observations
  mean=0.935566, MSE=0.5312636

Node number 3: 3572 observations,    complexity param=0.0521979
  mean=1.682704, MSE=1.022406
  left son=6 (2658 obs) right son=7 (914 obs)
  Primary splits:
     tp < 0.0115 to the right, improve=0.07885646, (0 missing)

Node number 6: 2658 observations
  mean=1.516199, MSE=0.8233658

Node number 7: 914 observations
  mean=2.166915, MSE=1.286151
```

APPENDIX 6-3: Thresholds in Reservoir Chlorophyll-a in Time Series

Chl-a versus time at the Lake of the Pines

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
  n=101 (8 observations deleted due to missingness)

         CP nsplit rel error   xerror      xstd
1 0.36254604      0  1.000000 1.019464 0.3142216
2 0.01435344      1  0.637454 0.661181 0.2428943

Node number 1: 101 observations,    complexity param=0.362546
  mean=9.756139, MSE=112.1713
  left son=2 (82 obs) right son=3 (19 obs)
  Primary splits:
     date_val < 38257.5 to the left,  improve=0.362546, (0 missing)

Node number 2: 82 observations
  mean=6.686463, MSE=49.99869

Node number 3: 19 observations
  mean=23.00421, MSE=164.3167
```

Chl-a versus time at Lake Tawakoni

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=259 (119 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.26526674      0 1.0000000 1.0119198 0.1218184
2 0.03173786      1 0.7347333 0.8528346 0.1056212

Node number 1: 259 observations,    complexity param=0.2652667
  mean=19.21622, MSE=279.436
  left son=2 (172 obs) right son=3 (87 obs)
  Primary splits:
     date_val < 35881 to the left,  improve=0.2652667, (0 missing)

Node number 2: 172 observations
  mean=13.09302, MSE=130.129

Node number 3: 87 observations
  mean=31.32184, MSE=353.9463
```

Chl-a versus time at Lake Fork

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=182 (87 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.05113782      0 1.0000000 1.018701 0.2368829
2 0.04855481      1 0.9488622 1.086465 0.2534968

Node number 1: 182 observations,    complexity param=0.05113782
  mean=10.26022, MSE=65.30843
  left son=2 (102 obs) right son=3 (80 obs)
  Primary splits:
     date_val < 35881 to the left,  improve=0.05113782, (0 missing)

Node number 2: 102 observations
  mean=8.641765, MSE=68.53038

Node number 3: 80 observations
  mean=12.32375, MSE=53.60256
```

Chl-a versus time at Lake Houston

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
  n=88 (139 observations deleted due to missingness)

         CP nsplit rel error    xerror      xstd
1 0.35574335      0 1.0000000 1.0324307 0.2087986
2 0.02719938      1 0.6442566 0.7198286 0.1418916

Node number 1: 88 observations,    complexity param=0.3557434
  mean=12.81011, MSE=203.7904
  left son=2 (59 obs) right son=3 (29 obs)
  Primary splits:
     date_val < 38192 to the left,  improve=0.3557434, (0 missing)

Node number 2: 59 observations
  mean=6.840678, MSE=49.04671

Node number 3: 29 observations
  mean=24.95483, MSE=298.6227
```

Chl-a versus time at Lake Conroe

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=60 (116 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.04846738      0 1.0000000 1.043380 0.8574350
2 0.02450914      1 0.9515326 1.072915 0.8950923

Node number 1: 60 observations,    complexity param=0.04846738
  mean=15.07717, MSE=671.5831
  left son=2 (12 obs) right son=3 (48 obs)
  Primary splits:
     date_val < 28428.5 to the left,  improve=0.04846738, (0 missing)

Node number 2: 12 observations
  mean=3.666667, MSE=11.22222

Node number 3: 48 observations
  mean=17.92979, MSE=795.986
```

Chl-a versus time at Lake Whitney

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=90 (40 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.35845917      0 1.0000000 1.0095620 0.2689785
2 0.05351437      1 0.6415408 0.6983135 0.1782392

Node number 1: 90 observations,    complexity param=0.3584592
  mean=9.598667, MSE=82.41821
  left son=2 (71 obs) right son=3 (19 obs)
  Primary splits:
     date_val < 38269 to the left,  improve=0.3584592, (0 missing)

Node number 2: 71 observations
  mean=6.786901, MSE=33.73567

Node number 3: 19 observations
  mean=20.10579, MSE=124.394
```

Chl-a versus time at Lake Belton

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
  n=94 (28 observations deleted due to missingness)

         CP nsplit rel error    xerror      xstd
1 0.19873621      0 1.0000000 1.0235902 0.1596227
2 0.03116437      1 0.8012638 0.9543573 0.1643070

Node number 1: 94 observations,    complexity param=0.1987362
  mean=4.097021, MSE=8.889953
  left son=2 (83 obs) right son=3 (11 obs)
  Primary splits:
     date_val < 39193.5 to the left,  improve=0.1987362, (0 missing)

Node number 2: 83 observations,    complexity param=0.03116437
  mean=3.613133, MSE=6.644995
  left son=4 (38 obs) right son=5 (45 obs)
  Primary splits:
     date_val < 36352   to the left,  improve=0.04220455, (0 missing)

Node number 3: 11 observations
  mean=7.748182, MSE=10.73145

Node number 4: 38 observations,    complexity param=0.03116437
  mean=3.036842, MSE=7.038111
  left son=8 (11 obs) right son=9 (27 obs)
  Primary splits:
     date_val < 33620.5 to the right, improve=0.1077146, (0 missing)

Node number 5: 45 observations
  mean=4.099778, MSE=5.795758

Node number 8: 11 observations
  mean=1.672727, MSE=4.076511

Node number 9: 27 observations
  mean=3.592593, MSE=7.177723
```

Chl-a versus time at Lake Travis

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=212 (29 observations deleted due to missingness)

        CP nsplit rel error   xerror      xstd
1 0.04858569      0 1.0000000 1.007434 0.2471018
2 0.03505264      2 0.9028286 1.074488 0.2564650

Node number 1: 212 observations,    complexity param=0.04858569
  mean=2.737075, MSE=7.035013
  left son=2 (110 obs) right son=3 (102 obs)
  Primary splits:
     date_val < 33614   to the left,  improve=0.04041155, (0 missing)

Node number 2: 110 observations
  mean=2.223636, MSE=4.223623

Node number 3: 102 observations,    complexity param=0.04858569
  mean=3.290784, MSE=9.476015
  left son=6 (92 obs) right son=7 (10 obs)
  Primary splits:
     date_val < 34210.5 to the right, improve=0.08758223, (0 missing)

Node number 6: 92 observations
  mean=2.990435, MSE=8.062067

Node number 7: 10 observations
  mean=6.054, MSE=14.01904
```

Chl-a versus time at Lake Buchanan

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
 n=223 (21 observations deleted due to missingness)

        CP nsplit rel error    xerror      xstd
1 0.10067610      0 1.0000000 1.0100087 0.4207207
2 0.03382898      1 0.8993239 0.9968433 0.4411556

Node number 1: 223 observations,    complexity param=0.1006761
  mean=5.954888, MSE=34.16256
  left son=2 (125 obs) right son=3 (98 obs)
  Primary splits:
     date_val < 33760.5 to the left,  improve=0.1006761, (0 missing)

Node number 2: 125 observations
  mean=4.3128, MSE=9.413276

Node number 3: 98 observations
  mean=8.049388, MSE=57.90423
```

Chl-a versus time at Lake Amistad

```
Call:
mvpart(form = chla ~ date_val, data = laketime, xval = 10, method = "anova",
   minsplit = 20, minbucket = 10)
  n=44 (6 observations deleted due to missingness)

         CP nsplit rel error   xerror      xstd
1 0.06039763      0 1.0000000 1.039916 0.6514814
2 0.01217147      1 0.9396024 1.139825 0.8116604

Node number 1: 44 observations,    complexity param=0.06039763
  mean=1.708182, MSE=1.900424
  left son=2 (16 obs) right son=3 (28 obs)
  Primary splits:
     date_val < 31247.5 to the right, improve=0.06039763, (0 missing)

Node number 2: 16 observations
  mean=1.26, MSE=4.10025

Node number 3: 28 observations
  mean=1.964286, MSE=0.4630102
```