



LINKING NUTRIENTS TO ALTERATIONS IN AQUATIC LIFE IN CALIFORNIA WADEABLE STREAMS



Linking Nutrients to Alterations in Aquatic Life in California Wadeable Streams

A. Elizabeth Fetscher, Martha Sutula, and Ashmita Sengupta
Southern California Coastal Water Research Project

Naomi Detenbeck
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection Agency
Atlantic Ecology Division
Narragansett, Rhode Island

U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Atlantic Ecology Division
Narragansett, Rhode Island 02882

Preface

This is contribution number ORD-007337 of the Atlantic Ecology Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency. The information in this document has been funded wholly by the U.S. Environmental Protection Agency under contract EP-12-D-000146 to the Southern California Coastal Water Research Program (SCCWRP). This document has been reviewed by the U.S. Environmental Protection Agency, Office of Research and Development, and approved for publication. This report was first reviewed through ORD's internal peer review process and approved for publication. In accordance with guidance in the US EPA's Peer Review Handbook for Influential Science Information, the document was subsequently sent out for an independent, external peer review to five subject matter experts with expertise in algal and macroinvertebrate ecology, nutrient effects on freshwater streams, and nutrient criteria development. The document was revised based on reviewer recommendations. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Table of Contents

List of Tables	v
List of Figures.....	vii
List of Acronyms and Abbreviations.....	xi
Acknowledgements.....	xiii
Executive Summary.....	xiv
1. Introduction.....	1
1.1 Introduction and Project Objectives	1
1.2 Document Organization	2
1.3 Literature Cited.....	3
2. Estimation of Reference and Ambient Concentrations of Stream Eutrophication Indicators	5
2.1 Introduction	5
2.2 Methods.....	5
2.2.1 Approach.....	5
2.2.2 Data Sources, Site Selection, and Stream Sampling Protocol.....	5
2.2.3 Distribution of Wadeable Stream Nutrients and Primary Producer Indicator Values.....	12
2.3 Results	13
2.3.1 Distribution of Nutrients and Primary Producer Indicators at Reference Sites.....	13
2.3.2 Ambient Distribution of Nutrients and Primary Producer Abundance Indicator	14
2.4 Discussion	19
2.5 Literature Cited.....	20
3. Thresholds of Adverse Effects of Primary Producer Abundance and Nutrients on Wadeable Stream	
Aquatic Life Indicators	21
3.1 Introduction	21
3.2 Methods.....	23
3.2.1 Conceptual Approach.....	23
3.2.2 Aquatic Life and Stressor Indicators	25
3.2.3 Detection of Ecological Thresholds.....	30
3.2.4 Data Sources	31
3.2.5 Data Analyses.....	33
3.3 Results	40
3.3.1 BMI and Diatom Responses to Biomass and Nutrient Gradients Based on Shifts in Community Composition	40
3.3.2 Biotic Responses to Biomass Gradients Based on Shifts in Integrative Measures of Community Composition (Metrics and Indices)	50
3.3.3 Examining Relative Influence of Biomass, Nutrients, and Other Factors on Integrative ALI Measures.....	63
3.3.4 Thresholds for Biomass and Nutrient Effects on Biotic Response.....	75

3.4 Discussion	77
3.4.1 Statistically Detected Thresholds of Adverse Effect in California Wadeable Streams	78
3.4.2 Variable Response of ALI Types to Biomass and Nutrients: The Biological Condition Gradient.....	84
3.4.3 Study Findings in Context of Policy Applications	85
3.5 Literature Cited.....	85
4. Evaluation of Nutrient Numeric Endpoint Benthic Biomass Spreadsheet Tool	91
4.1 Introduction	91
4.2 Methods.....	92
4.2.1 Conceptual Approach to Validation and Error Analysis.....	92
4.2.2 Background on Benthic Biomass Spreadsheet Tool Development and Testing (Objective 1).....	92
4.2.3 Data Sources	97
4.2.4 Validation of the NNE BBST Tool	98
4.2.5 Bias Analysis (Objective 2)	98
4.2.6 Exploratory Analyses Using Boosted Regression Trees (Objective 3) and Bayesian CART Analysis (Objective 4)	99
4.3 Results	102
4.3.1 Model Performance (Objective 1)	102
4.3.2 Random Forest Regression for Bias and Variance Analysis (Objective 2).....	105
4.3.3 Results of BRT Analyses (Objective 3).....	108
4.3.4 Results of Bayesian CART Analyses (Objective 4)	115
4.4 Discussion	122
4.4.1 Validation Exercise Shows Considerable Room for Improvement in BBST	122
4.4.2 Inclusion of Landscape and Site-Specific Factors Provide Avenue for Model Refinement	123
4.4.3 Summary of Validation and Recommendations for Refining Wadeable Stream Nutrient Algal Abundance Models	126
4.5 Literature Cited.....	126
Appendix A. Important Definitions	128
Appendix B. Graphics and Tables Supporting Analyses of Reference and Ambient Concentrations of Stream Eutrophication Indicator.....	131
Appendix C. Graphics and Tables Supporting Analyses of Thresholds of Adverse Effects of Primary Producer Biomass and Nutrient on Wadeable Stream Aquatic Life.....	141
Appendix D. Graphics and Tables Supporting Evaluation of Benthic Biomass Response Models.	
Appendix E. Quality Assurance/Quality Control Summary	193
Literature Cited.....	199

List of Tables

Table 2.1. Metric descriptions and codes for stream primary producer abundance indicators.	10
Table 2.2. Variables used for assigning sites to “site disturbance classes” per the state’s bio-objectives process (adapted from Ode et al., under review).	11
Table 2.3. Sources of data for landscape, meteorological, and geological explanatory variables used in predictive models.	12
Table 2.4. Median, 75 th , and 95 th percentiles of raw (unweighted) TN, TP benthic chlorophyll <i>a</i> , AFDM, and macroalgal percent cover (PCT_MAP), statewide and by region, at reference sites (both probability and targeted datasets included).	14
Table 2.5. Extent estimates for the site-evaluation categories based on reconnaissance information across the PSA and SMC probability surveys from 2008-2011.	15
Table 2.6. Statewide estimates for distributional properties of primary producer abundance indicator values in California perennial, wadeable streams.	15
Table 2.7. Estimated median values (with 95% confidence intervals) for selected ambient stream nutrient and primary producer abundance indicators statewide and by region.	17
Table 2.8. Percent of perennial wadeable stream kilometers exceeding 75 th and 95 th percentiles of statewide or regional reference values for nutrient and primary producer abundance gradients.	17
Table 3.1. Description of aquatic life indicators (ALIs) used in the analyses.	26
Table 3.2. Descriptions of the focal stressor gradients and other explanatory variables used in the analyses, listed in the table alphabetically according to the shorthand version of the name, within their respective categories.	28
Table 3.3. Sample sizes.	32
Table 3.4. Summary of analytical techniques used for threshold estimation.	34
Table 3.5. Results of CART analyses with NMS axis 1 scores for either the BMI or the diatom community as the response variable.	43
Table 3.6. TITAN and nCPA results for BMI and diatom community composition data.	45
Table 3.7. Summary of piecewise regression results for all ALI response types for which at least one version of the analysis (weighted or unweighted) fulfilled all four “strict” criteria, as described in the Methods.	51
Table 3.8. Summary of boosted regression tree models of ALI variables, and relative influence (and rank) of biomass and nutrient predictors used in each.	64
Table 3.9. Relative influence of predictors from BRT models.	65
Table 3.10. Partial Mantel coefficients (95% CIs) for correlation between biomass/nutrient predictors and ALI variables and p-values.	69
Table 3.11. Quantitatively determined thresholds of stream (or river) ALI responses to nutrient concentrations.	83
Table 4.1. Summary of types of models contained in BBST.	93

Table 4.2. Stream scouring cutoff values for the watersheds developed based on precipitation and watershed imperviousness.....	98
Table 4.3. Final set of classification variables used in B-CART analysis.	102
Table 4.4. Distribution of residuals of ≥ 0.5 standard deviation of mean for chlorophyll <i>a</i> by model type.	103
Table 4.5. Model performance (R^2 , slope and intercept) for all sites combined.....	103
Table 4.6. Model performance (R^2) for all, Reference, Intermediate, and Stressed sites (see Chapter 2) for predicted mean chlorophyll <i>a</i>	104
Table 4.7. Variables ranked according to importance for random forest regression for Chlorophyll <i>a</i> (Chl <i>a</i>) and AFDM by model type.	106
Table 4.8 Relative influence of nutrient species on abundance of stream biomass of six different types, from BRT models that included environmental co-factors.	109
Table 4.9. Partial Mantel coefficients (95% CIs) for correlation between nutrient predictors and biomass variables; <i>p</i> values.	111
Table 4.10. Results of B-CART analyses based on reduced sets of classification variables.	116
Table 4.11. Variables retained in regression analyses to predict benthic biomass (log10 chlorophyll <i>a</i>) based on Dodds-type models for nodes in B-CART models 1 and 3, and based on QUAL2K-type models for nodes in B-CART models 2 and 4.	120
Table A.1 Definition of beneficial uses applicable to freshwater habitat.....	130
Table B.1. Number of sites within each Level III ecoregion (Omernik, 1987) in the South Coast, by site disturbance class.....	139
Table C.1. TITAN change point values for BMI and diatom taxa (“pure” and “reliable”).....	144
Table C.2. Results of piecewise regressions for all analyses in which “relaxed” criteria were met	158
Table C.3. Shown are 1) a summary of thresholds, from all analyses, for the chlorophyll <i>a</i> , AFDM, TN, and, TP gradients, and 2) mean distributions of ALI values among sites that had gradient values below and above the indicated threshold	166
Table C.4. Summary of recommended numeric endpoints for stream NNE indicators, by beneficial use, from Tetra Tech (2006).....	178
Table D.1. Details on the models	187
Table D.2. Results of Bayesian CART analysis of full data set using all potential classification variables	189
Table D.3. Frequency of inclusion of classification variables in Bayesian CART TNTP models (Training sets 1-10).....	189
Table D.4. Results of Bayesian CART analysis of full data that includes PSA ecoregions	190
Table D.5. Results of Bayesian CART analysis of full data set that uses an empirical rather than the PSA ecoregions.....	190
Table E.1. Nutrient fractions for samples from 47 USGS NAWQA stream stations in California sampled biweekly over the year	194

List of Figures

Figure 2.1. All algae sampling sites (probability and targeted) included in this report, shown by the Perennial Stream Assessment (PSA) ecoregion in which they occur.	7
Figure 2.2. Statewide CDFs for biomass measures and macroalgal percent cover (attached and/or unattached combined) by site disturbance class	16
Figure 2.3. CDFs for biomass measures and macroalgal percent cover (attached and/or unattached combined), broken down by PSA6 ecoregion.....	18
Figure 3.1. Simplified conceptual model of eutrophication in wadeable streams depicting the relationship between nutrients, stream co-factors, ecological response and ecosystem services	22
Figure 3.2. Conceptual model depicting stages of change in biological conditions in response to an increasing stressor gradient. Reproduced from Davies and Jackson (2006)	24
Figure 3.3. Examples of two statistical approaches used to derive quantitative water quality goals	25
Figure 3.4. Examples of types of threshold relationships.....	30
Figure 3.5. ALI response to a stressor gradient showing the “reference envelope” along with “resistance” and “exhaustion” thresholds	31
Figure 3.6. Scatterplots and splines for non-metric multidimensional scaling (NMS) axis 1 values from the benthic macroinvertebrate (BMI) community against biomass and selected cover and nutrient gradients on log scale, using the statewide data set.....	41
Figure 3.7. Scatterplots and splines for NMS axis 1 values from the diatom community against biomass and selected cover and nutrient gradients on log scale, using the statewide data set	42
Figure 3.8. Cut points from CART analyses using NMS axis 1 scores from either the BMI or the diatom community as the response variable.....	44
Figure 3.9. Plots of “sum(z)” scores (depicted as dots) from TITAN analysis of BMI community data along chlorophyll a, AFDM, and TN gradients, and the cumulative threshold frequency graphs (depicted as lines) for the sum(z) scores.....	47
Figure 3.10. Plots of taxon-specific change from TITAN analysis of diatom community data along AFDM and TP gradients	48
Figure 3.11. Nonparametric change point analysis (nCPA) results	49
Figure 3.12. Summary of TITAN and nCPA change points along biomass/nutrient gradients, based on BMI and diatom community composition using the statewide dataset.....	50
Figure 3.13. Breakpoints, with 95% confidence intervals, for the chlorophyll a gradient, from piecewise regressions using all available ALI data types.....	53
Figure 3.14. Breakpoints, with 95% confidence intervals, for the AFDM gradient, from piecewise regressions using all available ALI data types.....	54
Figure 3.15. Breakpoints, with 95% confidence intervals, for the TN gradient, from piecewise regressions.....	55

Figure 3.16. Breakpoints, with 95% confidence intervals, for the TP gradient, from piecewise regressions.....	56
Figure 3.17. Piecewise regression plot of diatom ALI variable, RAWlowP on a TP gradient (top) and SiZer map from analysis of the same two variables (bottom).....	57
Figure 3.18. Piecewise regression plot of BMI ALI variable, Intolerant_PercentTaxa on a TN gradient (top) and SiZer map from analysis of the same two variables (bottom).....	58
Figure 3.19. Piecewise regression plot of BMI ALI variable, Taxonomic_Richness on a chlorophyll a gradient (top) and SiZer map from analysis of the same two variables (bottom)	59
Figure 3.20. Piecewise regression plot of BMI ALI variable, Intolerant_PercentTaxa on an AFDM gradient (top) and SiZer map from analysis of the same two variables (bottom).....	60
Figure 3.21. Closeup of piecewise regression plot of hybrid ALI variable, the IBI H20 on an AFDM gradient (top) and SiZer map from analysis of the same two variables (bottom).....	61
Figure 3.22. Distribution of ALI values among sites with stressor gradient (i.e., biomass or nutrient concentration) values below vs. above the threshold that had been determined based on piecewise regression	62
Figure 3.23. Heat map showing relative influence (%) of predictor variables (biomass, nutrients, and environmental co-factors) on ALI response variables, from 14 independent BRT models. Yellow = low influence, red = high.....	67
Figure 3.24. Summary of the relative influence of biomass and nutrient predictors on ALIs, from the BRT models	68
Figure 3.25. Partial dependence plots of chlorophyll a from BRT models predicting three ALI response types: CSCI, Taxonomic Richness, and RAWNhet	72
Figure 3.26. Partial dependence plots of AFDM from BRT models predicting four ALI response types: the metrics EPT_Percent and RAWDO100; and the IBIs H20 and D18.....	73
Figure 3.27. Partial dependence plots of TN from BRT models predicting three ALI response types: the soft algae IBI, S2; and the BMI ALIs, Taxonomic Richness and Intolerant Percent Taxa.....	74
Figure 3.28. Partial dependence plots of TP from BRT models predicting three ALI response types: the soft-algae ALI, RAWmeanZHR; and the IBIs, D18 and H20.....	75
Figure 3.29. Summary of results across analyses using the chlorophyll a, AFDM, TN, and TP gradients, stratified by assemblage type.....	76
Figure 3.30. Ranges of thresholds of ALI response by “ALI category” (as described in section 3.2.2) for two biomass and two nutrient gradients.....	77
Figure 4.1. Example of user interface for the BBST (example highlights output plot for the Dodds 1997 version of the model) with input and output panels.....	94
Figure 4.2. Example of user interface for BBST Standard QUAL2K model with input and output panels.	96

Figure 4.3. Sample plots of validation data showing measured versus predicted chlorophyll <i>a</i> by standard QUAL2K model, with 1:1 slope lines.....	104
Figure 4.4. Relative influence of variables for some selected models for chlorophyll <i>a</i> (top row: a, b, c) and AFDM (bottom row: d, e, f)	107
Figure 4.5. Heat map showing relative influence (%) of predictor variables (nutrients and environmental co-factors) on biomass response variables, from six independent BRT models.....	120
Figure 4.6. Three-dimensional plot (two views) of NO _x and mean monthly maximum ambient air temperature (the mean of the month the sample was collected and the two months prior) from BRT model for chlorophyll <i>a</i>	113
Figure 4.7. Partial dependence plots of percent fine substrata (left) and percent canopy cover (right) from the BRT model for AFDM	113
Figure 4.8. Three-dimensional plot of percent canopy cover (XCENMID) and conductivity from a BRT model for PCT_MAP (percent macroalgal cover).	114
Figure 4.9. Three-dimensional plot of percent sand + fine substrata and days of accrual from BRT model for PCT_MCP (percent macrophyte cover).....	115
Figure 4.10. Bayesian regression tree for Model 1.....	117
Figure 4.11. Bayesian regression tree for Model 2.....	117
Figure 4.12. Location of sampling sites corresponding to nodes in Bayesian CART Model 3	118
Figure 4.13. Sampling station locations corresponding to final nodes in Bayesian CART Model 4	119
Figure B.1. Histograms of biomass and algal/macrophyte cover data, all California probability data combined.	131
Figure B.2. Boxplots of biomass, ash-free dry mass, and macroalgal percent cover	134
Figure B.3. Cumulative distribution functions of biomass, ash-free dry mass, and macroalgal percent cover, by region, for all probability sites.....	137
Figure B.4. CDFs for benthic chlorophyll <i>a</i> , for the “xeric” and “mountain” Level III ecoregions (Omernik 1987) within the South Coast.....	139
Figure B.5. Within-ecoregion estimated percent of stream kilometers lower than the lowest proposed NNE endpoint for chlorophyll <i>a</i> (100 mg m ⁻²), by site disturbance class.	140
Figure C.1. Examples of plots of TITAN sum(z) scores for taxa that decrease in frequency along the gradient of interest (in black) and those that increase (in red).	141
Figure C.2 Examples of plots of TITAN change points for individual taxa.	142
Figure C.3. Examples of a SIZer map.....	143
Figure D.1. Comparison of the NNE stations (3053) to the daily precipitation station from NOAA (981, July 2012).....	179

Figure D.2. Comparison of the predicted precipitation between the PRISM and ARCGIS (predicted for this study)	180
Figure D.3. Ranking by variable importance for the Dodds 97 model for AFDM using random forests.....	181
Figure D.4. Ranking by variable importance for the Dodds 02 model for AFDM using random forests.....	182
Figure D.5. Ranking by variable importance for the QUAL2K model for AFDM using random forests.....	183
Figure D.6. Ranking by variable importance for the Dodds 97 for Chlorophyll <i>a</i> using random forests	184
Figure D.7. Ranking by variable importance for the Dodds 02 for Chlorophyll <i>a</i> using random forests	185
Figure D.8. Ranking by variable importance for the QUAL2K models for Chlorophyll <i>a</i> 186 using random forests.....	186
Figure D.9. Predicted versus observed normalized log ₁₀ chlorophyll <i>a</i> biomass (mg/m ²) for 1) TNTP training set, b) TNTP test set, c) DINDIP training set, and d) DINDIP test sets used in Bayesian CART analysis (Dodds-type model, all potential classifiers).....	191
Figure E.1. Frequency of a month of a) maximum annual total N and b) maximum annual total P in 47 California NAWQA streams.....	195
Figure E.2. Value for a) total N and b) total P by month relative to maximum monthly values in 47 California NAWQA streams.....	196
Figure E.3 Relationship between annual maximum and growing season average values for a) total N and b) total P in 47 California NAWQA streams	197
Figure E.4. Chlorophyll <i>a</i> levels (log-transformed) across sampling dates, by year, for South coast (blue) and all other sites (red) within the state	198
Figure E.5. Chlorophyll <i>a</i> levels (log-transformed) across sampling dates normalized by log total N, by year, for South Coast (blue) and all other sites (red) within the state	199

List of Acronyms and Abbreviations

Acronym or abbreviation	Definition
AFDM	Ash-free dry mass
AG_2000_1K	Percent agricultural land use in catchment within a 1-km radius from sampling site
AG_2000_5K	Percent agricultural land use in catchment within a 5-km radius from sampling site
AG_2000_WS	Percent agricultural land use in catchment
ALI	Aquatic life indicator
BCG	Biological condition gradient
BBST	Benthic biomass spreadsheet tool (Tetra Tech 2006)
BMI	Benthic macroinvertebrate
BRT	Boosted regression tree analysis
BURC	Beneficial use risk category
CART	Classification and regression tree analysis
CDF	Cumulative distribution function
CODE21_2000_1K	Percent "Code 21" land use in catchment within a 1-km radius from sampling site
CODE21_2000_5K	Percent "Code 21" land use in catchment within a 5-km radius from sampling site
CODE21_2000_WS	Percent "Code 21" land use in catchment
COLD	Cold freshwater habitat beneficial use
COMM	Commercial and sport fishing beneficial use
CPOM	Coarse particulate organic matter
CSCI	California Stream Condition Index (the BMI-based statewide index for stream bioassessment; Mazor et al., under review)
DO	Dissolved oxygen
D18	Diatom Index of Biotic Integrity (IBI); Fetscher et al. 2014
EMAP	Environmental Monitoring and Assessment Program
EPT_Percent	Percent BMI individuals that are Ephemeroptera, Plecoptera, or Trichoptera
EPT_PercentTaxa	Percent BMI taxa that are Ephemeroptera, Plecoptera, or Trichoptera
EPT_Taxa	Number of BMI taxa that are Ephemeroptera, Plecoptera, or Trichoptera
EQ	Equation
FPOM	Fine particulate organic matter
GHCND	Daily global historical climatology network
GRTS	Generalized random tessellation stratified method for creating a spatially balanced probability survey
H20	Diatom + soft algae ("hybrid") index of biotic integrity (IBI); Fetscher et al. 2014
H21	Diatom + soft algae ("hybrid") index of biotic integrity (IBI); Fetscher et al. 2014
H23	Diatom + soft algae ("hybrid") index of biotic integrity (IBI); Fetscher et al. 2014
IBI	Index of biotic integrity
Intolerant_Percent	Percent BMI individuals that are "intolerant"; Ode et al. 2005
Intolerant_PercentTaxa	Percent BMI taxa that are "intolerant"; Ode et al. 2005
Intolerant_Taxa	Number of BMI taxa that are "intolerant"; Ode et al. 2005
LRWQCB	Lahontan Regional Water Quality Control Board
MIGR	Migration of aquatic organisms beneficial use
MRDS	Mineral Resources Data System
MSE	Mean square error
MUN	Municipal (beneficial use)
N	Nitrogen
N:P	Nitrogen-to-phosphorus ratio
nCPA	Nonparametric change point analysis
NH ₄	Ammonium
NMS	Nonmetric multidimensional scaling analysis
NNE	Nutrient numeric endpoint
NOAA	National Oceanic and Atmospheric Agency
NO _x	Nitrate + nitrite
NPS	Non-point source
OBEM	Out of bag error method
O/E	Observed over expected taxa from RIVPACS models for BMI taxa; Mazor et al., under review
P	Phosphorus

PCT_CPOM	Percent cover of coarse particulate organic matter in streambed
PCT_FN	Percent cover of fine substrata in streambed
PCT_MAP	Macroalgal percent cover
PCT_MCP	Macrophyte percent cover
PCT_MIAT1	Percent presence of thick (1mm+) microalgae
PCT_SAFN	Percent sand + fines in streambed
PHab	Physical habitat (measures collected in a stream reach)
PRISM	Parameter-elevation relationships on independent slopes model
propAchMin	Proportion of diatom valves that are <i>Achnanthydium minutissimum</i>
propTaxaZHR	Proportion of total of total soft-algae taxa recorded that are in the Zygnemataceae, heterocystous cyanobacteria, or Rhodophyta
PSA	Perennial stream assessment
QUAL2K	A river and stream water quality model
RARE	Rare, threatened, or endangered species beneficial use
RAWDO100	Proportion diatoms requiring nearly 100% DO saturation; van Dam et al. 1994
RAWDO50	Proportion diatoms requiring at least 50% DO saturation; van Dam et al. 1994
RAWeutro	Proportion eutrophication indicator diatoms; van Dam et al. 1994
RAWlowN	Proportion low-N indicator diatoms; Potapova and Charles 2007
RAWlowP	Proportion low-P indicator diatoms; Potapova and Charles 2007
RAWlowTPsp	Proportion of soft algal taxa that are considered "low TP" indicators; Fetscher et al. 2014
RAWmeanZHR	Mean of the metrics propTaxaZHR and RAWpropBiovolZHR; Fetscher et al. 2014
RAWNhet	Proportion nitrogen-heterotroph diatoms; van Dam et al. 1994
RAWpropBiovolChlor	Proportion of total soft algae biovolume that is Chlorophyta
RAWpropBiovolZHR	Proportion of total soft algae biovolume that is in the Zygnemataceae, heterocystous cyanobacteria, or Rhodophyta
RAWpropGreenCRUS	Proportion of green algal biovolume belonging to <i>Cladophora glomerata</i> , <i>Rhizoclonium hieroglyphicum</i> , <i>Ulva flexuosa</i> , or <i>Stigeoclonium</i> species
RCMP	California's Reference Condition Management Program
REC-1	Contact water recreation beneficial use
REC-2	Non-contact water recreation beneficial use
RIVPACS	River Invertebrate Prediction and Classification System
S2	Soft algae index of biotic integrity (IBI); Fetscher et al. 2014
Shannon_Diversity	Shannon Diversity Index for BMI taxa
Simpson_Diversity	Simpson Diversity Index for BMI taxa
SiZer	Significant zero crossings analysis
SMC	Stormwater Monitoring Coalition
SPWN	Spawning, reproduction, and/or early development beneficial use
SRP	Soluble reactive phosphorus
SWAMP	California State Water Resources Control Board Surface Water Ambient Monitoring Program
SWRCB	State Water Resources Control Board
Taxonomic_Richness	Richness of BMI taxa
TITAN	Threshold indicator taxa analysis
TN	Total nitrogen
Tolerant_Percent	Percent BMI individuals that are "tolerant"; Ode et al. 2005
Tolerant_PercentTaxa	Percent BMI taxa that are "tolerant"; Ode et al. 2005
Tolerant_Taxa	Number of BMI taxa that are "tolerant"; Ode et al. 2005
TP	Total phosphorus
URBAN_2000_1K	Percent urban land use in catchment within a 1-km radius from sampling site
URBAN_2000_5K	Percent urban land use in catchment within a 5-km radius from sampling site
URBAN_2000_WS	Percent urban land use in catchment
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey
W1_HALL	A riparian disturbance index; Kaufmann et al. 1999
WARM	Warm freshwater habitat beneficial use
WILD	Wildlife habitat beneficial use
WQO	Water Quality Objective
XDENMID	Percent canopy cover

Acknowledgements

The California Surface Water Ambient Monitoring Program (SWAMP) funded Reference Condition Management Program and Perennial Stream Assessment, and the Southern California Stormwater Monitoring Coalition (SMC) provided data. Taxonomic identifications for biotic samples were carried out for the aforementioned monitoring programs by the following individuals: Joseph Slusark, Dan Pickard, and Austin Richards of the California Department of Fish and Wildlife Aquatic Bioassessment Laboratory (for benthic macroinvertebrate data); Rosalina Stancheva, Robert Sheath, and Christina Fuller of California State University San Marcos (for soft-algae data); and Patrick Kociolek, Evan Thomas, and Carrie Graeff of the University of Colorado at Boulder (for diatom data). The draft CSCI was developed by Raphael Mazor in collaboration with members of the California Biological-Objectives Technical Team and Science Advisory Panel. Mark Engeln, Abel Santana, Rebecca Schaffner, Marco Sigala, Carly Beck and Bryan White assisted with data management and literature review. The document benefited from review comments on an earlier version by EPA staff (Nathan Smucker, Karen Blocksom, Lester Yuan, Glen Thursby, and Wayne Munns), SCCW staff (Eric Stein), San Diego Regional Water Control Board (Lilian Busse), and Tetra Tech (Michael Paul, Jon Butcher), and valuable feedback on the revised version from Walter Dodds, Michelle Evans-White (University of Arkansas), R. Jan Stevenson (Michigan State University, SCCWRP (Stephen Weisberg, Kenneth Schiff), Los Angeles County Sanitation District (Anne Heil, George Gallis, Joshua West), and two anonymous reviewers.

Executive Summary

In wadeable streams, nutrient enrichment, in concert with other site-specific factors, can result in the overabundance of algal biomass, low dissolved oxygen and altered biotic communities. These changes can have adverse effects on stream ecosystem services. Scientifically-based water quality objectives (WQO) and tools that relate these objectives to nutrient management are needed in California to prevent eutrophication from occurring and to provide targets to restore waterbodies where adverse effects have already occurred.

The California State Water Resources Control Board (SWRCB) is developing nutrient water quality objectives for the State's surface waters. USEPA guidance on nutrient objective development generally recommends three means to set nutrient objectives (USEPA 2000): 1) a reference approach, based on a statistical percentile of nutrient or biotic response indicators in minimally-disturbed waterbodies; 2) an empirical stress-response approach, based on statistical analyses of field data on nutrients, algal abundance and indicators of aquatic life; or 3) a process-based approach, involving identification of ecological responses of concern and mechanistically modeling the linkage back to nutrient loads and other co-factors controlling response.

Among the approaches that the SWRCB staff is considering is the process-based approach, known as the Nutrient Numeric Endpoint (NNE) framework (Tetra Tech 2006). The NNE framework is intended to serve as numeric guidance to translate narrative WQO. It consists of two tenets: 1) assessment and recommended numeric (regulatory) endpoints based on the ecological response of an aquatic waterbody to eutrophication (e.g., algal abundance, dissolved oxygen [DO]) to assess waterbody condition and 2) scoping-level models that link the response indicator endpoints to nutrient inputs and other site-specific factors and management controls. These scoping models were intended to be used to establish default nutrient targets for point source discharge and municipal stormwater permits and total maximum daily loads (Tetra Tech 2006). Tetra Tech (2006) developed the benthic biomass spreadsheet tool (BBST) for use in streams. As the SWRCB prepares to propose nutrient objectives for wadeable streams, scientific analyses of improved data from California statewide stream probabilistic and targeted bioassessment surveys can strengthen the scientific basis for policy decisions. In the context of this study, "endpoints" refer to policy decisions on levels at which point management action should be taken; "thresholds" refer to the output of scientific analyses.

The objectives of this project are three-fold:

- Estimate the natural background and ambient concentrations of nutrients and candidate indicators of primary producer abundance in California wadeable streams;
- Explore relationships and identify thresholds of adverse effects of nutrient concentrations and primary producer abundance on aquatic life indicators in California wadeable streams;
- Evaluate the Benthic Biomass Spreadsheet Tool for California wadeable streams using existing data sets and recommend avenues for refinement.

The intended outcome of this study is research, NOT recommendations for regulatory endpoints for nutrient and response indicators for California wadeable streams. The findings of this research study, as well as other analyses, may be used as lines of evidence considered to support SWRCB policy decisions on nutrient objectives for wadeable streams.

Study Findings

The majority of the State’s Wadeable Streams sampled are below the 75th percentile of minimally disturbed “reference sites.” California’s perennial, wadeable streams, as assessed during the bioassessment index period of late spring through mid-summer, exhibited a skew toward the low end of the primary producer abundance gradient. Nearly 66% of perennial wadeable stream kilometers had estimated benthic chlorophyll *a* and 59% had estimated TN and TP values below **the 75th percentile of each variable at reference sites statewide**¹. Among the regions, a gradient in algal abundance and nutrient concentrations was observed from high in areas developed by urban and agricultural land uses (South Coast, Central Valley) to low in areas of the state with lower density development (e.g., North Coast and Sierra regions).

Statistically detectable thresholds were found for benthic chlorophyll *a*, ash-free dry mass (AFDM), and nutrients; benthic chlorophyll *a* thresholds were below those of TetraTech (2006). This study found statistically significant relationships and thresholds of adverse effects of benthic chlorophyll *a*, AFDM, and TN and TP concentrations on indicators of benthic macroinvertebrate (BMI) and algal community structure—employed in this study as indicators of aquatic life. Integrative aquatic life indicators (ALIs) such as indices of biotic integrity corresponded to higher thresholds whereas ALI measures specific to constrained groups of “sensitive” taxa generally corresponded to lower thresholds, illustrative of the paradigm of the biological condition gradient. Most of these thresholds of effect exceeded the 75th percentile of these indicators among reference stream reaches statewide, but they were often less than the 95th percentile. The range of benthic chlorophyll *a* thresholds in this study were generally substantially below the current NNE endpoints protective of beneficial uses recommended by TetraTech (2006; 100 and 150 mg/m² chlorophyll *a* for cold [salmonid] and warm water respectively). However, it should be noted that our results are based on instantaneous measurement at low-flow conditions, and as such, do not reflect year-long loads or storm flows. It is not clear to what degree the types of ALI-stressor relationships we observed would hold during rain events.

Validation exercise indicates that there is considerable room for improvement in BBST; inclusion of landscape and site-scale factors provide avenue for model refinement. The BBST models show poor fit, particularly among “stressed” sites (one-third of the data set), when validated against a statewide dataset, which contains benthic chlorophyll *a* data as currently measured in California ambient monitoring programs. The poor fit is understandable, given that the BBST was optimized for North American temperate streams and that the model predicts *maximum* algal abundance, a value not verifiably captured during the period in which sampling to generate the project data set occurred. Several landscape- and site-scale explanatory variables were high in their relative influence in the BBST model predicted-observed variance analysis and in preliminary nutrient-algal response models. Nutrient concentrations were important predictors in BBST model predicted-observed variance analysis and boosted regression tree (BRT) models, albeit occupying less prominent roles than other factors, such as temperature and stream substratum type. This finding validates the fundamental NNE approach: site-specific co-factors that vary across the California landscape can influence algal response to nutrients. It also suggests that model refinements are possible; inclusion of these site- and landscape-scale explanatory variables in preliminary nutrient-algal response models substantially improved model fit over existing BBST models.

¹ The analogous values, if considering the 95th percentile of Reference sites, are 90% of stream kilometers for Chlorophyll *a* and approximately 78% for nutrients.

Recommendations

- Statistical analyses for threshold detection were conducted on statewide scale; resources were insufficient to look at the question of whether there is scientific evidence for regionalization of thresholds. Additional analyses are recommended to look specifically at this question.
- An alternative approach to establish levels protective of ALI is to use predictive regression models to estimate concentrations of nutrients or algal abundance that are linked to a quantitative ALI target. We recommend such analyses based on the benthic invertebrate and stream algal IBI.
- A comprehensive effort to develop nutrient-algal abundance models for wadeable streams should be undertaken, considering a full range of predictive and probabilistic statistical models. The compiled dataset now includes a variety of explanatory variables that are available to begin a more thorough set of analyses. More than one model categorized by classes may be necessary in order to capture the range of nutrient-response relationships statewide. More complex mechanistic models could be considered over the long-term if the need to offer greater flexibility and applications to site-specific waterbody assessment are warranted.

1. Introduction

1.1 Introduction and Project Objectives

Eutrophication¹ of water resources is a major environmental issue in California, with demonstrated links among anthropogenic changes in watersheds, increased nutrient loading, harmful algal blooms, hypoxia, and impacts on aquatic food webs. In wadeable streams, elevated nutrient concentrations, in concert with other site-specific factors, can result in the overabundance of algal biomass, low dissolved oxygen and altered biotic communities, with a suite of adverse effects on stream ecosystem services and beneficial uses (Appendix A, Table A.1). High algal abundance can alter hydrology and interfere with spawning, foraging, and shelter (Biggs 2000, Quinn and Hickey 1990), limit the growth of benthic diatoms as food sources for scraper/grazers (Steinman 1996), and deteriorate water quality (Quinn and Gilliland 1989). Wadeable stream algal blooms can also negatively impact human health and other ecosystem services or beneficial uses, through toxin-forming harmful algal blooms, proliferation of pathogenic bacteria, taste/odor problems in municipal drinking water supplies and compromised aesthetics (Biggs 2000, Lembi 2003, Suplee et al. 2009, Fovet et al. 2012). In California, examples of eutrophication in wadeable streams have been well-documented (e.g., Southern California, Mazor et al. 2014). Scientifically-based water quality objectives and tools that relate these objectives to management controls are needed to prevent eutrophication from occurring and to provide targets to restore waterbodies where adverse effects have already occurred.

USEPA guidance on nutrient objective development generally recommends three means to set nutrient objectives (USEPA 2000): 1) a reference approach, 2) an empirical stress-response approach, and 3) a mechanistic, process-based approach. The reference waterbody approach involves characterization of the distributions of nutrients in “minimally disturbed” waterbodies. Nutrient concentrations are chosen at some statistical percentile of those reference waterbodies. The empirical stress-response approach involves establishing statistical relationships between the causal or stressor variable (in this case nutrient concentrations or loads) and the ecological response (changes in algal or aquatic plant biomass or community structure, changes in sediment or water chemistry such as dissolved oxygen, pH). The process-based approach involves identifying the ecological responses of concern and mechanistically modeling the linkage back to nutrient loads and other co-factors controlling response (e.g., hydrology, grazers, denitrification, etc.).

The California SWRCB is developing nutrient water quality objectives for the State’s surface waters. Among the approaches that SWRCB staff is considering is a process-based approach, known as the Nutrient Numeric Endpoint (NNE) framework (Tetra Tech 2006). The NNE framework, intended to serve as numeric guidance to translate narrative WQO, consists of two tenets: 1) numeric (regulatory) endpoints based on the ecological response of an aquatic waterbody to eutrophication (e.g., algal abundance, dissolved oxygen [DO]) to assess waterbody condition and 2) models that link the response indicator endpoints (e.g., algal abundance) to nutrient inputs and other site-specific factors and management controls. These models are intended to be used to establish nutrient targets for point source discharge and municipal stormwater permits and total maximum daily loads (Tetra Tech 2006). Tetra Tech (2006) developed the benthic biomass spreadsheet tool for use in establishing “scoping levels” nutrient targets in streams. As the SWRCB prepares to propose nutrient objectives for wadeable streams, analysis using newly available data from statewide stream

¹ See definition of eutrophication and other key terms in Appendix A.

bioassessment surveys can improve the scientific basis for policy decisions on nutrient objectives. In the context of this study, “endpoints” refer to regulatory decisions at which point management action should be taken, while “thresholds” refer to the output of scientific analyses.

The objectives of this research project are three-fold:

- Estimate the natural background and ambient concentrations of nutrients and candidate indicators of primary producer abundance in California wadeable streams;
- Explore relationships and identify thresholds of adverse effects of nutrient concentrations and primary producer abundance on indicators of aquatic life in California wadeable streams; and
- Evaluate the Benthic Biomass Spreadsheet Tool (BBST) for California wadeable streams using existing data sets and recommend avenues for refinement.

The intended outcome of this study is research, NOT final regulatory endpoints for nutrient and response indicators for California wadeable streams. In this context, this research can provide: 1) improved understanding of the corresponding quantitative thresholds at which eutrophication stressors (e.g., nutrient concentrations, algal abundance) begin to exert adverse effects on aquatic life measures, 2) context for these thresholds by summarizing available data on reference and ambient concentrations of stressors and 3) an improved understanding of what types of nutrient-response modeling may be appropriate, given existing data. The findings of this research study, as well as other analyses, may be used as lines of evidence considered to support SWRCB policy decisions on nutrient objectives for wadeable streams.

1.2 Document Organization

The document is organized as follows:

Chapter 1: Introduction, Objectives and Document Organization

Chapter 2: Estimation of Reference and Ambient Concentrations of Algal Biomass

Chapter 3: Investigating Nutrient and Primary Producer Abundance Thresholds for Aquatic Life Response

Chapter 4: Validation of NNE Benthic Biomass Spreadsheet Tool and Investigation of Stream Nutrient Relationships with Biomass

Appendices

1.3 Literature Cited

- Biggs, B.J.F. 2000. New Zealand Periphyton Guideline: Detecting, Monitoring and Managing the Enrichment of Streams. Ministry for Environment Publication, Wellington, New Zealand, 151 pp.
- Fovet, O., Belaud G., Litrico X., Charpentier S., Bertrand C., Dollet P., and C. Hugodot. 2012. A model for fixed algae management in open channels using flushing flows. *River Research and Applications* 28:960–972.
- Lembi, C.A. 2003. Control of nuisance algae. In: Wehr JD, Sheath RG (eds). *Freshwater Algae of North America Ecology and Classification*. Academic Press, New York, pp 805 - 834.
- Quinn, J.M. and B.W. Gilliland. 1989. The Manawatu River cleanup - Has it worked? *Transactions of the Institution of Professional Engineers, New Zealand*, 16:22-26.
- Quinn, J.M. and C.W. Hickey. 1990. Magnitude of effects of substratum particle size, recent flooding and catchment development on benthic invertebrate communities in 88 New Zealand rivers. *New Zealand Journal of Marine and Freshwater Research*, 24: 411-427.
- Steinman, A.D. 1996. Effects of grazers on freshwater benthic algae. In R. G. Stevenson et al. (eds), *Algal Ecology*. Academic Press, San Diego, CA: 341–365.
- Suplee, M.W., Watson V., Teply M. and H. McKee. 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.
- Tetra Tech 2006. Technical Approach to Develop Nutrient Numeric Endpoints for California. Tetra Tech, Inc. http://rd.tetratech.com/e pa/Documents/CA_NNE_July_Final.pdf
- US EPA 2000. Nutrient Criteria Technical Guidance Manual for Rivers and Streams. US. EPA Office of Science and Technology. EPA-822-B-00-002. July 2000. http://www2.epa.gov/sites/production/files/documents/guidance_rivers.pdf

2. Estimation of Reference and Ambient Concentrations of Stream Eutrophication Indicators

2.1 Introduction

As the SWRCB prepares to propose nutrient objectives for wadeable streams, newly available data from statewide and regional stream surveys can improve the scientific basis for policy decisions on regulatory endpoints. These policy decisions should be supported, in part, by the distribution of natural background concentrations from minimally disturbed reference sites and the distribution of ambient concentrations across the full population of wadeable streams. At the time in which the NNE framework was conceived (Tetra Tech 2006), this had not been summarized. Distribution of natural background concentrations and ambient levels are key considerations in the process of determining the scientific basis and the cost/benefits of policy decisions on regulatory endpoints. Here, natural background refers to the absence or near absence of anthropogenic effects and ambient levels, refers to all streams, including those affected by anthropogenic activities.

This section addresses two key questions:

- What is the distribution of the values of nutrient and algal abundance indicators at “Reference” sites that are subjected to minimal anthropogenic disturbance?
- What are the ambient distributions of these indicators in California perennial, wadeable streams statewide and by ecoregions of interest?

2.2 Methods

2.2.1 Approach

The California NNE framework proposes to establish regulatory endpoints for algal abundance, dissolved oxygen and pH in order to assess the beneficial use status of wadeable streams (Tetra Tech 2006). The Surface Water Ambient Monitoring Program (SWAMP) has since adopted a standardized algal monitoring protocol which includes alternate measures of algal abundance (e.g., ash-free dry mass [AFDM] and algal percent cover) (Fetscher et al. 2009). Currently, data are available on 938 sites using this standardized protocol, thus providing the opportunity to summarize nutrient concentrations and algal abundance indicators at the statewide and ecoregional scale.

2.2.2 Data Sources, Site Selection, and Stream Sampling Protocol

Data Sources

Survey data were compiled from the following wadeable stream monitoring programs:

- Statewide Perennial Stream Assessment (PSA),
- Statewide Reference Condition Management Program (RCMP), and
- Southern California Stormwater Monitoring Coalition (SMC)

The probabilistic survey design for the California ambient surveys (PSA, SMC) is based on the methods described in Stevens and Olsen (2004). The quality assurance parameters for the California datasets are

based on those established for the Surface Water Ambient Monitoring Program (SWAMP 2008). In some places (where noted), non-probability data (i.e., from sites subjectively selected for “targeted” sampling) are also included. In probability surveys, sites are selected using a combination of stratification and unequal probability weighting that yields a spatially balanced distribution of sites. Because of the objective way in which sites are selected, regional/statewide estimates of perennial wadeable stream condition with known confidence limits can be generated from the survey data. For more information on probability surveys, see Stevens and Olsen (2004). All references to statewide or regional stream percentiles in this report are based on this sampling framework and the operational definition of perennial wadeable streams.

The probability surveys reported on here are those of 1) the State of California Perennial Stream Assessment (PSA), and 2) the southern California Stormwater Monitoring Coalition (SMC). Results from these two programs were used to generate regional and statewide estimates of stream condition for nutrients and indicators of primary producer abundance. In addition to probability data, data from targeted sampling sites were also included in the analyses. These data from targeted sites come from the state’s Reference Condition Management Program (RCMP) and a recently completed project geared toward developing stream algal assemblage data for use in bioassessment of stream condition. Taken together, the available data represent 938 wadeable, perennial² stream reaches sampled from 2007 through 2011, including the sampling frames for probability surveys throughout the state (National Hydrography Data Set [NHD] v2, www.horizon-systems.com/nhdplus; Figure 2.1). Of these, 575 of the reaches were sampled as part of the probability surveys, and the remaining 363 were targeted. Sampling was largely conducted as one-time site visits (91% of samples) within the time frame spanning late spring to early fall, with the vast majority occurring in May through August. For sites with both benthic macroinvertebrate and algae data, the two assemblages were sampled during the same visit.

Site Selection and Evaluation for Probability Surveys

The `spsurvey` package (Kincaid and Olsen 2008) in R (R Core Team 2008) was used in establishing the list of “probability sites” for each year’s statewide (PSA) and regional (SMC) probability survey. This involved using a technique called Generalized Random Tessellation Stratified sampling site selection (GRTS; Stevens and Olsen 2004) to create spatially-balanced survey designs. As long as sites are sampled in the order in which they appear on the list, spatial balance among them is preserved, and the resulting dataset can be used to generate estimates of natural resource extent and condition with known confidence limits. The design of each survey was based on a “linear” resource sensu Kincaid and Olsen (2008). The reporting unit for this type of survey was in terms of length (e.g., stream kilometers). Once sampling sites were identified, they were inspected to determine whether they belonged to the target sampling population (perennial, wadeable streams in California), whether permission could be secured for sampling, whether they were safe to access, and whether they could be reached within a timeframe that would not compromise holding times for analytes.

² We used the PSA operational definition of “perennial”, i.e., those stream reaches with surface flow during the sampling period. A “wadeable” reach was defined as that which is <1m deep for at least 50% of its length.

Based on these factors, as well as whether a sample was successfully collected, sites were then classified into one of four “evaluation categories”:

- site is part of survey’s “target population”, and was sampled
- site is part of “target population”, but was not sampled
- site is not part of “target population”
- unknown

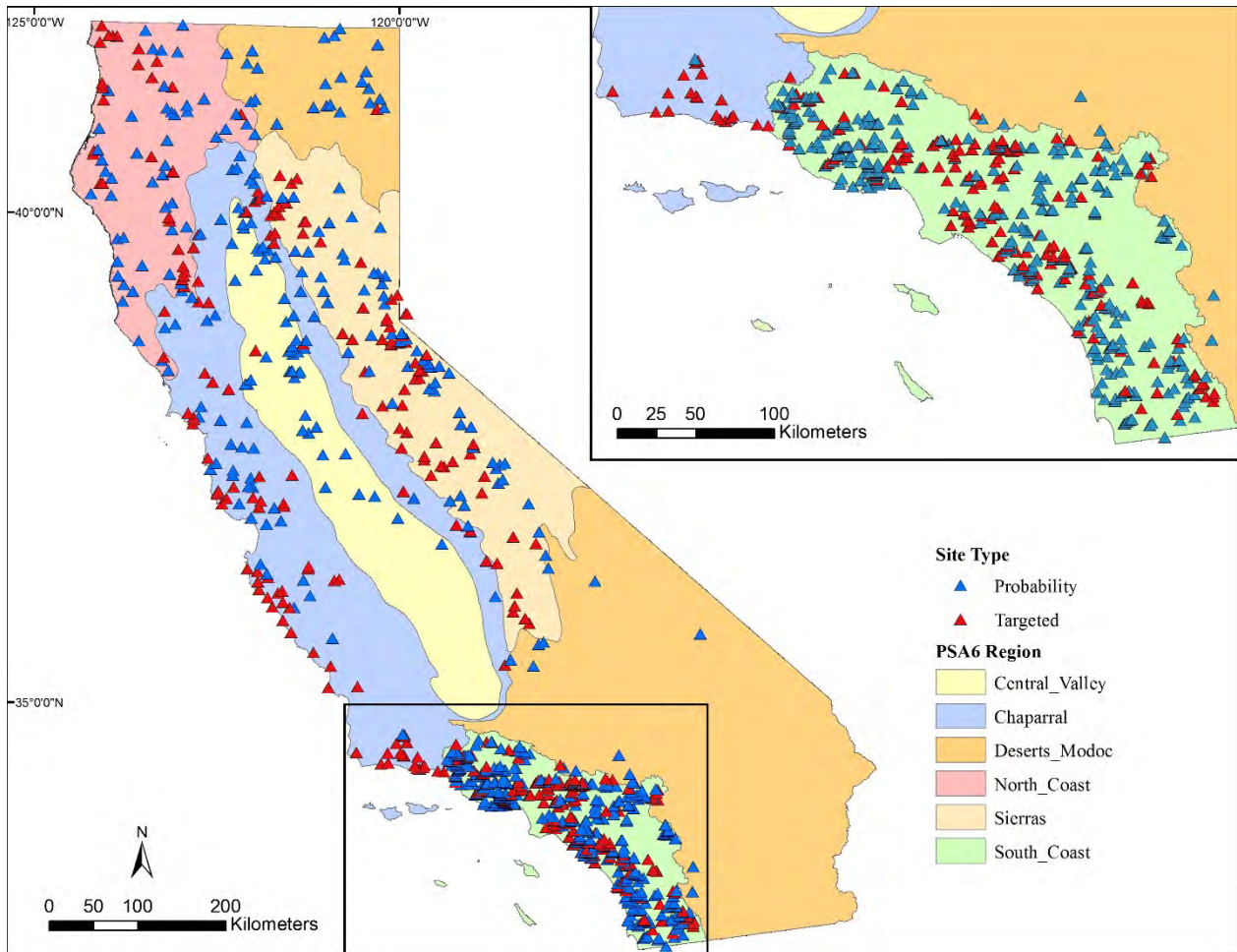


Figure 2.1. All algae sampling sites (probability and targeted) included in this report, shown by the Perennial Stream Assessment (PSA) ecoregion in which they occur. State bioassessment programs use a combination of Omernik (1995) ecoregions and Regional Water Quality Control Board boundaries to partition the state for assessment purposes. “PSA6” refers to the version of the classification scheme that encompasses six ecoregions.

Description of Stream Algal Field Sampling and Laboratory Analysis Protocols Utilized in Compiled Wadeable Stream Survey Data

The field sampling and laboratory analyses protocols used in the compiled stream survey data are briefly described in this section. The types and distribution of primary producer abundance across channel habitats can vary widely among stream types. For this reason, it is important to assess primary producer abundance within a stream in a number of different ways, because each individual indicator captures this distribution differently. For example, both benthic chlorophyll *a* and ash-free dry mass (AFDM) measure algal biomass, but chlorophyll *a* is a proxy for the measurement of live algal biomass, while AFDM measures both live and dead biomass, as well as organic matter imported into the survey site. Furthermore, algae and macrophytes can occupy different “compartments” within the stream (i.e., floating on the surface, attached to cobbles/boulders, interstitially distributed within the upper layer of gravel and fine sediments), all of which are included across the sample types upon which results are reported here. The ability to look at a combination of measures may provide a more robust overall assessment of algal/macrophyte abundance. Based on this rationale, the SWAMP standardized algal assessment protocol yields the following data types for indicators of stream primary producer abundance (Fetscher et al. 2009):

- Algal biomass:
 - benthic chlorophyll *a*
 - benthic ash-free dry mass (AFDM)
- Algal cover:
 - macroalgal percent cover
 - microalgal percent cover and thickness
- Macrophyte percent cover³

In addition to primary producer abundance indicators, total and dissolved inorganic nitrogen and phosphorus concentrations were also assessed. Chlorophyll *a* is under consideration for use within the current NNE framework. Other indicators (e.g., percent cover, AFDM, or other measures) may be considered for inclusion in the future.

A “multi-habitat” method was employed to quantitatively collect benthic algae at each sampling site⁴. This method, SWAMP’s Standard Operating Procedures (Fetscher et al. 2009), is based largely on the procedures of EPA’s Environmental Monitoring and Assessment Program (EMAP; Peck et al. 2006) and is analogous to SWAMP’s method for collecting benthic macroinvertebrates (Ode 2007). It involves objectively collecting from a known surface area specimens from a variety of stream substrata, in proportions aligning with relative abundances of substratum types in the stream. Specifically, eleven subsamples are collected at objectively determined locations, one from each of 11 transects that are spaced equidistantly from one another, across the 150-m long sampling reach. For systems with a mean wetted width >10, the sampling reach is 250 m long. The subsamples are then combined into a single “composite” sample for laboratory analyses. As such, a given composite sample may have been collected from any combination of cobbles, gravel, sand, and other substratum types. The goal is to achieve a representative sample of the benthic algae from each sampling reach, in terms of both community composition and biomass.

³ Macrophytes technically refer to both macroalgae and rooted aquatic vegetation. In this context, we define macrophytes as rooted aquatic vegetation.

⁴ BMIs and algae were collected in tandem at each of the 11 subsampling locations described at each study site; first BMIs, then algae, slightly offset so that sampling locations did not interfere with one another.

Various measures of algal and macrophyte cover were carried out using the methods outlined in Fetscher et al. (2009). This involved recording point-intercept presence/absence of microalgae, macroalgae, and macrophytes at each of 105 points objectively positioned (in a pre-determined grid) throughout each stream reach. Macroalgae that was attached to the stream bottom was recorded separately from that which was unattached and free-floating at the time of assessment. Microalgae was measured based on presence/absence of a biofilm on stream substrata. The thickness of the microalgal biofilm was also recorded using ordinal thickness codes.

For algal biomass, filtered aliquots of quantitatively sampled algal material were analyzed for chlorophyll *a* content using EPA 445.0, and for AFDM using WRS 73A.3. Chlorophyll *a* and AFDM concentrations measured in the laboratory were transformed into mass per area of stream bottom sampled (e.g., mg/m²).

Most algal/macrophyte field metrics were calculated as percent cover estimates based on the percentage of sampling points at which the type of algae/macrophyte was observed. The midpoint values of the ranges corresponding to each thickness code for mean microalgal thickness were averaged across all 105 sampling points per site (Fetscher et al. 2009). A “nuisance algae” metric combining information from both macroalgae and thick microalgae (>1 mm) was also calculated. A summary with descriptions of the metrics associated with algal/macrophyte cover is provided in Table 2.1.

Sites were grouped into “disturbance classes” throughout the following analyses. To assign sites to disturbance classes, we used the same set of screening criteria as that employed by the State of California’s Biological Objectives initiative (Ode et al., under review). Under this approach, sites are classified according to the degree of anthropogenic disturbance they are exposed to, based on surrounding land uses and local riparian disturbance measures. Table 2.2 provides a list of the factors that were used for classifying sites into one of the three disturbance classes: “Reference”, or those sites that are exposed to the lowest levels of anthropogenic disturbance based on the variables considered, “Stressed”, or those sites exposed to the highest levels, and “Intermediate”, or those sites falling between the “Reference” and “Stressed” groups.

Table 2.1. Metric descriptions and codes for stream primary producer abundance indicators.

Metric Code	Description
PCT_MAA	Percent Presence of Attached Macroalgae (defined as algal mats or filaments easily visible to the naked eye)
PCT_MAP	Percent Presence of Macroalgae (Attached and/or Unattached)
PCT_MAU	Percent Presence of Unattached Macroalgae
PCT_MIAT1	Percent Presence of Thick Microalgae (1mm+)
PCT_MIAT1P	Percent Presence of Thick Microalgae (1mm+), where Microalgae Present
PCT_MIATP	Percent Presence of Microalgae
PCT_NSA	Percent Presence of Nuisance Algae (Macroalgae and/or Thick Microalgae [1mm+] counts as “presence” at a given point)
XMIAT	Mean Microalgae Thickness (mm)
XMIATP	Mean Microalgae Thickness (mm) where Microalgae Present
PCT_MCP	Percent Presence of Macrophytes

Table 2.2. Variables used for assigning sites to “site disturbance classes” per the state’s bio-objectives process (adapted from Ode et al., under review). WS: Watershed. 5K: Watershed clipped⁵ to a 5-km buffer of the sample point. 1K: Watershed clipped to a 1-km buffer of the sample point. W1_HALL: proximity-weighted human activity index (Kaufmann et al. 1999). In order to be considered “Reference” condition, all criteria listed in the “Threshold” column for “Reference” must be met. If any of the criteria in the “Stressed” column apply, that site is considered “Stressed”. Sites not falling into either of these categories default to “Intermediate”. Data sources are as follows: A: National Landcover Data Set (2006, <http://www.epa.gov/mrlc/nlcd-2006.html>). B: Custom roads layer (P. Ode, pers. comm.). C: National Hydrography Dataset Plus (v2, <http://www.horizon-systems.com/nhdplus/>). D: National Inventory of Dams. E: Mineral Resource Data System (MRDS 2014). F: Field-measured variables (Fetscher et al. 2009).

Variable	Scale*	Threshold (Reference)	Threshold (Stressed)	Unit	Source
% Agriculture	1k, 5k, WS	<3	>50	%	A
% Urban	1k, 5k, WS	<3	>50	%	A
% Ag + % Urban	1k and 5k	<5	>50	%	A
% Code 21 ⁶	1k and 5k	<7	>50	%	A
	WS	<10	>50	%	A
Road density	1k, 5k, WS	<2	>5	km/km ²	B
Road crossings	1k	<5	-	crossings/ km ²	B, C
	5k	<10	-	crossings/ km ²	B, C
	WS	<50	-	crossings/ km ²	B, C
Dam distance	WS	>10	-	km	D
% Canals and pipelines	WS	<10	-	%	C
Instream gravel mines	5k	<0.1	-	mines/km	C, E
Producer mines	5k	0	-	mines	E
W1_HALL	reach	<1.5	>5	NA	F

*For variables in which multiple spatial scales are used for determining site classification, in the case of the “Reference” boundary, the value indicated must apply to all spatial scales listed, whereas for the “Stressed” boundary, the indicated value need only apply for one of the listed spatial scales.

Secondary data for watershed characterization were derived from the sources described below. Watershed and local habitat characteristics are required both as co-variates in periphyton and macroinvertebrate response models and as predictors of watershed disturbance regimes. Factors affecting instream periphyton growth and biomass accrual include nutrients (and their ratios), solar radiation, temperature, shading from riparian cover, incised stream channels, local topography, mean stream velocity, substratum type, abundance of grazers, and frequency, magnitude, and time since droughts or scouring flows. Field data were collected by PSA, RCMP, and SMC monitoring programs. Sources of landscape, meteorological, and geology data are listed in Table 2.3.

⁵ Only the land within the catchment contributing to the sampling site was included within the indicated radii (i.e., the area was clipped at the watershed boundaries).

⁶ “Code 21” encompasses a wide range of land uses primarily characterized by heavily managed vegetation (e.g., low-density residential development, parks, golf courses, highway medians)

Table 2.3. Sources of data for landscape, meteorological, and geological explanatory variables used in predictive models. DEM = digital elevation model.

Data Type/Variable	Data Source	Description or Download
Minimum and maximum air temperature per month (2007-2012)	PRISM	http://www.prism.oregonstate.edu/products/matrix.phtml , http://www.prism.oregonstate.edu/docs/index.phtml
Solar Radiation (for topographic shading)	ArcMap 10 tool Solar Radiation using DEM data from NHDPlus	http://www.horizon-systems.com/nhdplus/
Cloud cover, mean percent per month (2007-2012)	MODIS Cloud data from NASA	http://ladsweb.nascom.nasa.gov/data/
Land cover/land use	National Landcover Data Set, 2006	http://www.epa.gov/mrlc/nlcd-2006.html
Hydrology	National Inventory of Dams and NHD Plus	http://geo.usace.army.mil/pgis/f?p=397:1:0; http://www.horizon-systems.com/nhdplus/index.php
Elevation	National Elevation Dataset	http://ned.usgs.gov/
Drainage area (from DEM)	NHDPlus	http://www.horizon-systems.com/nhdplus/
Geology maps	USGS	http://mrddata.usgs.gov/geology/state/
Total precipitation per month (2007-2012)	PRISM	http://www.prism.oregonstate.edu/products/matrix.phtml , http://www.prism.oregonstate.edu/docs/index.phtml
Basin slope (from DEM)	NHDPlus	http://www.horizon-systems.com/nhdplus/

2.2.3 Distribution of Wadeable Stream Nutrients and Primary Producer Indicator Values

To provide an overview of the values for each of the indicators of primary producer abundance in the ambient surveys' target population (i.e., California perennial, wadeable streams), we generated descriptive statistics for estimated data distributions and cumulative distribution functions (CDFs) (Kincaid and Olsen 2009), using the *spsurvey* package in R on the probability subset of data. A CDF depicts the estimated probability distribution of values of a given indicator relative to the cumulative proportion of the geographic unit of interest, i.e., percent of stream length in the state.

Each site in the combined probability surveys for the different programs/years has an associated weight in units of stream length, which reflects how much of the state's stream network, within the stratum (e.g., landcover type, region, watershed) in which that site is found, is "represented by that site". The more sites in a given stratum, the less weight each site is assigned. Because data from multiple surveys with different stratification schemes were combined for this report, it was necessary to create mutually exclusive "cross-categories" corresponding to the intersection of the different strata from the various surveys. Once cross-categories were created, the weights of all sites had to be adjusted to reflect the combined numbers of sites within each new cross-category. Adjusted weights were calculated for each cross-category by dividing the total stream length within that cross-category by the number of sites evaluated during site reconnaissance. Once weights were adjusted, statewide extent and magnitude estimates for the various primary producer indicator values could be computed (see below).

It is not uncommon for some of the sites generated in a probability-based design to prove unsuitable for sampling for a variety of reasons that include: 1) the site being found, during reconnaissance, not to be part of the survey's designated "target population"; or 2) the site is within the target population, but for some logistical reason, it cannot be sampled (e.g., access denial, physical barriers or sheer distance of the site from nearest roads). Comprehensive documentation is required in order to classify sites into "evaluation categories" based on the results of site reconnaissance. If insufficient information regarding why samples were not collected is provided by field crews, the default classification for a site is "Unknown".

We chose to describe the ambient distribution of nutrients and primary producer abundance statewide and by ecoregion relative to the 75th percentile of reference sites. The percent of stream kilometers with indicator values below the 75th percentile of reference were calculated using the Horvitz-Thompson estimator (1952), which is a weighted average of sample values where weights are adjusted according to design implementation. Confidence intervals were based on local neighborhood variance estimators (Stevens and Olsen 2003), which assumes that samples located close together tend to be more alike than samples that are far apart. Graphical output for all analyses in the report was generated using the R package *ggplot2* (Wickham 2009). All graphics and statistical analyses in the report were carried out using R (version 2.15.1, R Core Team 2012), unless otherwise noted.

2.3 Results

2.3.1 Distribution of Nutrients and Primary Producer Indicators at Reference Sites

For the most part, quality of "Reference" sites, as identified by our standard set of screens, did not noticeably vary among regions in terms of the distribution of percent open (undeveloped) space within the contributing watersheds. The one notable exception was the South Coast (and particularly the xeric portion thereof), which did have a somewhat lower overall percentage of open space (96%) than other regions (which ranged from 98 to 99%).

Chlorophyll *a*, AFDM, macroalgal percent cover, and nutrients (TN and TP) exhibited a considerable degree of variability in values among Reference sites, but their distributions were highly skewed toward the low end of the stressor gradients (Table 2.4). At the 75th percentile, the ranges in nutrients and primary producer indicator values among ecoregions were fairly narrow (i.e., 0.10-0.31 mg/L TN, 0.02-0.04 mg/L TP, 8-27 mg/m² chlorophyll *a*, 6-27 g/m² AFDM, and 15-37% cover of macroalgae). For primary producer indicators and TN, North Coast and Sierra Nevada reference sites represented the lower end of that range, while South Coast, Desert-Modoc and Central Valley represented the upper end of the range.

Table 2.4. Median, 75th, and 95th percentiles of raw (unweighted) TN, TP benthic chlorophyll *a*, AFDM, and macroalgal percent cover (PCT_MAP), statewide and by region, at Reference sites (both probability and targeted datasets included).

Statistic by Primary Producer Indicator type		Statewide	Chaparral	Central Valley ¹	Deserts-Modoc	North Coast	South Coast	Sierra Nevada
		n=263	n=56	n=1	n=10	n=41	n=74	n=81
Chlorophyll <i>a</i> (mg/m ²)	Median	6.9	8.9		10.7	6.2	12.5	3.1
	75 th	14.6	16.4	23.0	26.5	9.2	24.4	7.9
	95 th	44.1	46.2		32.0	25.1	124.8	28.3
AFDM (g/m ²)	Median	5.4	6.2		13.4	4.0	16.3	3.7
	75 th	11.9	10.0	12.9	23.9	6.0	26.8	5.8
	95 th	34.0	19.7		36.7	14.8	130.6	12.2
Macroalgal percent cover (%)	Median	7.0	3.5		30.5	5.5	9.5	7.0
	75 th	22.9	15.9	41.0	36.8	15.0	26.0	23.0
	95 th	45.7	38.9		55.9	36.5	60.0	50.3
TN (mg/L)	Median	0.091	0.090		0.223	0.090	0.138	0.065
	75 th	0.161	0.144	0.155	0.281	0.117	0.308	0.100
	95 th	0.462	0.264		0.467	0.212	0.925	0.185
TP (mg/L)	Median	0.019	0.022		0.027	0.016	0.018	0.021
	75 th	0.032	0.042	0.027	0.041	0.020	0.035	0.032
	95 th	0.074	0.088		0.079	0.045	0.106	0.060

¹ The Central Valley ecoregion had only one site in the Reference site disturbance class; values in the table represent the results of this single site.

2.3.2 Ambient Distribution of Nutrients and Primary Producer Abundance Indicator

The proportions of sites falling into the four site “evaluation categories” are shown in Table 2.5. By far, the majority of stream kilometers in the state were estimated to fall outside of the surveys’ “target population”, either because they were non-perennial or non-wadeable stream reaches. The proportion of sites for which samples were collected represented about 10% of the state total stream kilometers.

Analysis of the statewide ambient wadeable stream data showed that algal biomass parameters, (chlorophyll *a*, AFDM) and nutrients exhibited broad ranges in concentrations, but their distributions were very highly skewed toward the low end (Table 2.6, Figure B.1). This was also generally true of primary producer percent cover metrics, with the exception of percent presence of microalgae (PCT_MIATP).

CDFs of site disturbance classes show a good amount of separation of reference, intermediate and stressed sites for chlorophyll *a* and AFDM, but not for macroalgal % cover (Figure 2.2). Boxplots of the distributions are provided in Appendix B, Figure B.2.

Table 2.5. Extent estimates for the site-evaluation categories based on reconnaissance information across the PSA and SMC probability surveys from 2008-2011.

Site Evaluation Category	Number of Sites Sampled*	Estimated Stream Kilometers (% of State Total)	Confidence Interval (95%)
Part of survey's "target population", and sampled	572	33,499 (10)	29,101 - 37,897
Part of "target population", but not sampled	400	43,438 (13)	37,973 - 48,903
Not part of "target population"	3362	238,195 (74)	231,300 - 245,089
Unknown	174	9,510 (3)	7,270 - 11,750

* Note that each sample for the input data used in the analysis represents either a one-time sampling event, or an average (for the small subset of stream reaches for which multiple samples over time were available).

Table 2.6. Statewide estimates for distributional properties of primary producer abundance indicator values in California perennial, wadeable streams. Data are from combined PSA and SMC probability surveys from 2008-2011. SE: standard error of the mean; CI: confidence interval (95%). Indicator acronyms are defined in Table 2.1.

Indicator	Range of Measured Values (N)	Estimated Mean (SE)	Estimated Median (CI)	Estimated 90th percentile (CI)
TN (mg/L)	0.01-26 (538)	0.533 (0.074)	0.131 (0.111-0.156)	1.035 (0.846-1.428)
TP (mg/L)	0.002-4.5 (536)	0.086 (0.008)	0.028 (0.024-0.031)	0.190 (0.150-0.280)
Chlorophyll <i>a</i> (mg m ⁻²)	0.22-1504 (536)	21 (2)	8 (6-12)	47 (39-64)
AFDM (g m ⁻²)	0.07-489 (525)	16 (2)	7 (6-8)	40 (23-50)
PCT_MAP (%)	0-98 (480)	16 (1)	6 (4-9)	51 (41-56)
PCT_MAA (%)	0-98 (480)	14 (1)	5 (3-7)	43 (36-52)
PCT_MAU (%)	0-87 (480)	2 (0.5)	0 (0-0)	3 (2-9)
PCT_MCP (%)	0-98 (480)	10 (1)	4 (2-5)	25 (20-39)
PCT_MIAT1 (%)	0-94 (478)	7 (1)	2 (0.5-2)	20 (13-32)
PCT_MIAT1P (%)	0-100 (464)	8 (1)	2 (1-3)	22 (16-41)
PCT_MIATP (%)	0-100 (478)	76 (2)	86 (83-93)	99 (99-100)
PCT_NSA (%)	0-100 (478)	20 (2)	11 (9-13)	52 (50-62)
XMIAT (mm)	0-6 (478)	0.5 (0.03)	0.3 (0.3-0.4)	1 (0.8-1.5)
XMIATP (mm)	0-20 (464)	0.6 (0.03)	0.4 (0.4-0.5)	1 (0.8-1.6)

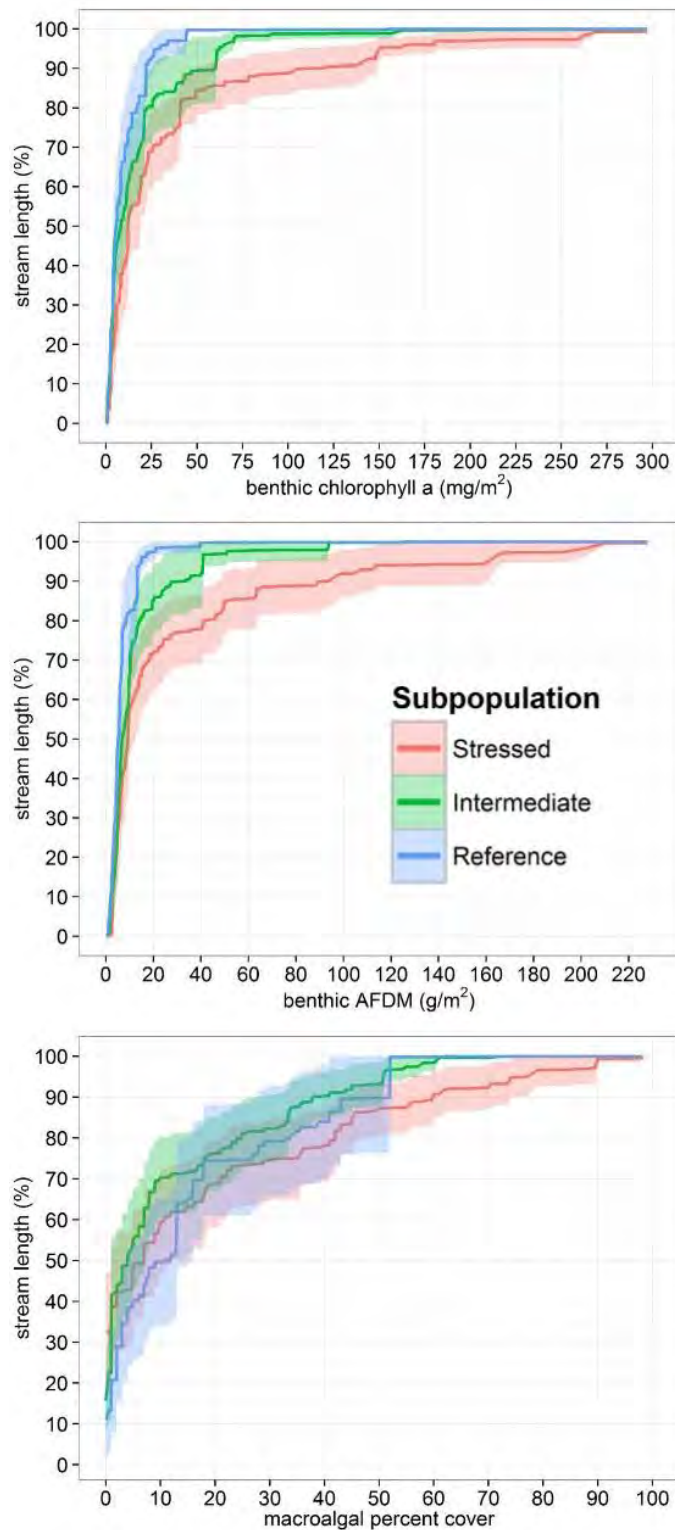


Figure 2.2. Statewide CDFs for biomass measures and macroalgal percent cover (attached and/or unattached combined) by site disturbance class. The graphs show the estimated probability distributions of the three types of primary producer abundance indicators relative to the cumulative proportion of stream length. Highlighted areas delineate the 95% confidence intervals for each estimate.

As with the reference sites, the ranges in median values of nutrients and primary producer indicator values among ecoregions were fairly narrow (i.e., 0.05-0.48 mg/L TN, 0.02-0.09 TP, 6-26 mg m⁻² chlorophyll *a*, 5-17 g m⁻² AFDM, and 1-20% cover of macroalgae, Table 2.7, Figure 2.3). North Coast and Sierra Nevada sites represented the lower end of that range, while South Coast and Central Valley consistently represented the upper end.

Table 2.7. Estimated median values (with 95% confidence intervals) for selected ambient stream nutrient and primary producer abundance indicators statewide and by region. Data are from combined PSA and SMC probability surveys from 2008-2011.

Indicator	Chaparral	Central Valley	Deserts-Modoc	North Coast	South Coast	Sierra Nevada
Chlorophyll <i>a</i> (mg m ⁻²)	13 (5.6-17.4)	12.6 (7.5-21.6)	8.9 (5.8-11)	5.7 (4-11.3)	25.7 (19.2-40.7)	5.7 (2.9-12)
AFDM (g m ⁻²)	6.6 (5.8-9.3)	13 (10.3-18.6)	10.2 (6.9-12.4)	5.5 (4.6-6.5)	17.2 (10.9-23.9)	4.8 (4.1-9.4)
Macroalgal percent cover (PCT_MAP, %)	5 (3-17.7)	16.9 (4.9-33.9)	11.9 (7-21.7)	7 (3-12.9)	20.1 (14.6-29.8)	1 (0.2-4)
TN (mg/L)	0.251 (0.135-0.365)	0.480 (0.332-0.890)	0.257 (0.192-0.364)	0.104 (0.080-0.130)	0.744 (0.540-0.989)	0.052 (0.042-0.081)
TP (mg/L)	0.034 (0.023-0.094)	0.095 (0.041-0.196)	0.041 (0.028-0.053)	0.028 (0.020-0.029)	0.050 (0.045-0.090)	0.020 (0.016-0.021)

Statewide, the percentage of stream kilometers that exceeded the 75th percentile of statewide reference values ranged from 27 % for macroalgal percent cover to a high of 41 % for TP and TN (Figure 2.3, Table 2.8). This range was generally greater among the regions and inconsistent by indicator group by region. For example, regions that were on the lower end of the absolute concentration range (North Coast and Sierra Nevada) had a higher percentage of miles exceeding their respective 75th percentile of eco-regional reference, putting them within range of South Coast, a region consistently at the upper edge of concentration range (Figure 2.2, Table 2.8). This is due to a proportionally lower Ecoregional reference value.

Table 2.8 Percent of perennial Wadeable stream kilometers exceeding 75th and 95th percentiles of statewide or regional Reference values for nutrient and primary producer abundance gradients. By this definition, 25% and 5% of Reference sites, respectively, exceed the indicated value as well. Data are from combined PSA and SMC probability surveys from 2008-2011.

Gradient	Statewide		Chaparral		Central Valley		Deserts-Modoc		North Coast		South Coast		Sierra Nevada	
	75th	95th	75th	95th	75th	95th	75th	95th	75th	95th	75th	95th	75th	95th
TN (mg/L)	41	22	58	44	85	-	48	30	40	12	75	2	27	6
TP (mg/L)	41	21	42	38	58	-	70	26	69	4	62	32	23	2
Chlorophyll <i>a</i> (mg/m ²)	34	10	39	8	34	-	14	7	43	10	54	15	43	10
AFDM (g/m ²)	27	11	32	12	55	-	19	10	48	2	34	6	41	18
Macroalgal percent cover (PCT_MAP; %)	26	11	35	17	33	-	16	4	22	9	42	16	16	9

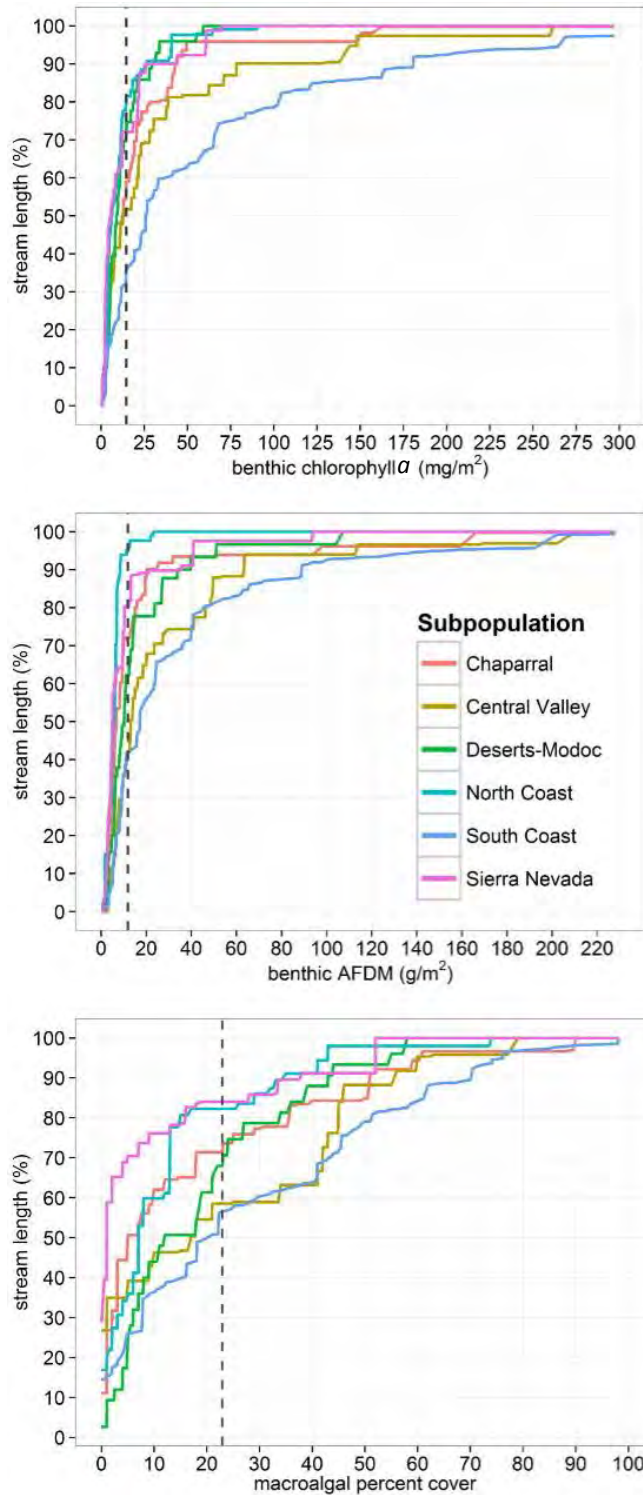


Figure 2.3. CDFs for biomass measures and macroalgal percent cover (attached and/or unattached combined), broken down by PSA6 ecoregion. The graphs show the estimated probability distributions of the three types of primary producer abundance indicators relative to the cumulative proportion of stream length. The vertical dashed line on each graph denotes the 75th percentile among Reference sites, statewide. Confidence intervals for each CDF can be viewed on the individual graphs for each ecoregion provided in Figure B.3. In addition, a further breakdown of the CDFs within the South Coast ecoregion (i.e., “xeric” and “mountain” subregions) is provided in Figure B.4.

2.4 Discussion

California's perennial, wadeable streams, as assessed during the PSA index period⁷, exhibited a skew toward the lower end of the nutrient and primary producer abundance gradient, although nutrients and primary producer abundance were understandably higher in intensively developed regions like South Coast and Central Valley. Statewide, only an estimated 34% of perennial stream kilometers had chlorophyll *a* values exceeding the 75th percentile of Reference sites⁸ statewide. This percentage was slightly higher for nutrients (41% of stream kilometers for both TN and TP).

Interpretation of the ambient distribution of nutrients and algae should be tempered by an understanding that the data may not represent the peak concentrations. Algal abundance and nutrient concentrations vary seasonally as a function of stream flow, temperature, available light, grazing pressure, nutrient source and other factors (Dodds et al. 2002). The data utilized in this survey represent a single time point taken during a late spring-summer index period. This index period was established to optimize condition assessment for benthic macroinvertebrates, not stream algal abundance, per se. The optimum period for stream algal assessments has not been established (see Appendix E for discussion). Atmospheric deposition is a significant component of N loading to California ecosystems, with a much more significant contribution of dry deposition to loading. Atmospheric deposition can have a more far-reaching effect than point sources and can affect otherwise pristine montane streams due to atmospheric transport. Impacts of dry deposition on stream water chemistry can be delayed from the dry summers until fall/winter when rains occur (Bytnerowicz and Fenn 1996).

The 75th percentiles for TN and TP estimated from the probability-based samples of reference streams are similar to those modeled for RF1 reaches in the corresponding nutrient ecoregions by Smith et al. (2003). Smith's values were based on models developed from estimated yields of USGS reference gaging stations (1976 – 1997), with corrections for wet atmospheric N deposition based on interpolated NADP values from 1980 – 1993. Smith et al. estimated an upper quartile of 0.21 (s.d. = 0.07) mg N/L with wet deposition and 0.18 (s.d. = 0.07) after correction for wet atmospheric deposition for annual flow-weighted instream nutrient concentrations in the Central Valley and Western Forested Mountain nutrient ecoregions, and 0.11 (s.d. = 0.04) with and 0.05 (s.d. = 0.07) without wet deposition for the Xeric West ecoregion. For total P concentrations, Smith et al. estimated 75th percentiles of 0.02 (s.d. = 0.005) and 0.03 (s.d. = 0.015) mg P/L for Central Valley and Western Forested Mountain ecoregions or Xeric West ecoregions, respectively. There are likely to be some differences between our estimates and those of Smith et al. due to differences in sampling dates as atmospheric N deposition has been declining, and because Smith et al. calculated estimated concentrations across the RF1 stream network which extends to larger systems than the perennial wadeable stream dataset. Corrections by Smith et al for atmospheric deposition are likely underestimates because of the dominance of dry deposition N sources in the arid west (Bytnerowicz and Fenn 1996).

⁷ The PSA index period for stream sampling starts in May for drier parts of the state and June or July in colder/wetter parts of the state (depending upon stream flow conditions), and lasts for two to three months.

⁸ In the case of the reference sites, values are given here for all available data combined (i.e., probability plus non-probability, or "targeted", sites)

2.5 Literature Cited

- Bytnerowicz, A. and M. E. Fenn. 1996. Nitrogen deposition in California forests: a review. *Environmental Pollution*, 92:127-146.
- Dodds, W.K., Smith V.H., and K. Lohman. 2002. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 865-874.
- Fetscher, A.E., Busse L.B., and P.R. Ode. 2009. Standard Operating Procedures for Collecting Stream Algae Samples and Associated Physical Habitat and Chemical Data for Ambient Bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 002. (updated May 2010)
- Horvitz, D.G. and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47:663-685.
- Kaufmann, P.R., Levine P., Robinson E.G., Seeliger C., and D.V. Peck. 1999. Surface waters: Quantifying physical habitat in wadeable streams. EPA/620/R-99/003. US EPA. Office of Research and Development. Washington, DC.
- Kincaid, T. and A.R. Olsen. 2008. SPSurvey package for R. Available from <http://www.epa.gov/nheerl/arm>
- Mineral Resources Data System (MRDS). 2014. <http://tin.er.usgs.gov/mrds/>. Accessed 18 January 2014.
- Ode, P. 2007. SWAMP Bioassessment Procedures: Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Available from http://mpsl.mlml.calstate.edu/phab_sopr6.pdf
- Ode, P.R., Rehn A.C., Mazor R.D., Schiff K. C., May J. T., Brown L. R., Gillett D., Herbst D., Lunde K. and C.P. Hawkins. In review. Evaluating the adequacy of a reference pool for use in environmentally complex regions.
- Peck, D.V., Herlihy A.T., Hill B.H., Hughes R.M., Kaufmann P.R., Klemm D., Lazorchak J.M., McCormick F.H., Peterson S.A., Ringold P.L., Magee R., and M. Cappaert. 2006. Environmental Monitoring and Assessment Program-Surface Waters Western Pilot Study: Field operations manual for wadeable streams. U.S. Environmental Protection Agency, Washington, D.C. EPA/620/R-06/003.
- R Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>
- R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>
- Smith, R. A., Alexander R.B., and G.E. Schwarz. 2003. Natural background concentrations of nutrients in streams and rivers of the conterminous United States. *Environmental Science and Technology*, 37:3039-3047.
- Stevens, D.L. and A.R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association*, 99:262-278.
- Stevens, D.L. and A.R. Olsen. 2003. Variance estimation for spatially balanced samples of environmental resources. *Environmetrics*, 14:593-610.
- Tetra Tech. 2006. Technical Approach to Develop Nutrient Numeric Endpoints for California. Tetra Tech, Inc. http://rd.tetratech.com/epa/Documents/CA_NNE_July_Final.pdf
- Wickham, H. 2009. ggplot2: elegant graphics for data analysis. Springer, New York.

3. Thresholds of Adverse Effects of Primary Producer Abundance and Nutrients on Wadeable Stream Aquatic Life Indicators

3.1 Introduction

Nutrient overenrichment, in concert with other site-specific factors, can result in the overabundance of organic matter in a process known as eutrophication (Nixon 1995). The adverse effects of eutrophication on stream ecosystem services such as biodiversity, water quality and aesthetics have been well vetted in the literature (Figure 3.1). Nutrients, together with a complex suite of reach scale environmental factors, can have direct and indirect effects on biotic communities (Wotton et al. 1996, Stevenson et al. 1997, Risang et al. 2004). Nutrients stimulate autotrophic production. This high algal abundance can alter hydrology and interfere with spawning, foraging, and shelter (Biggs 2000, Quinn and Hickey 1990). Filamentous algae that proliferates in high nutrient conditions can limit the growth of benthic diatoms as a food source for primary consumers such as scraper/grazers (Steinman 1996). Excessive organic matter accumulation can cause declines in dissolved oxygen, leading to deteriorated habitat quality (Quinn and Gilliland, 1989). Nutrients also increase heterotrophic production, a pathway much less well studied than autotrophic pathways (Evans-White 2014, Dodds 2007). Most studies to date have demonstrated short-term stimulation of bacterial and fungal growth with nutrient additions, improved nutritional quality of leaf litter (decreased C:N ratios) and concomitant increases in detritivore biomass (Greenwood et al. 2007, Connolly and Pearson 2013, Tant et al. 2013). Longer-term enrichment of headwater streams can lead to a decreased efficiency of trophic transfers, with an increased loss of carbon downstream and reduced productivity of top predators if nutrients have differential effects on primary consumers with different degrees of resistance to predators (Davis et al. 2010, Suberkropp et al. 2010). Algal blooms can also negatively impact ecosystem services by causing taste/odor problems, cyanobacterial toxin production (Chorus and Bartram 1999; Aboal et al. 2002; Douterelo et al. 2004), blocked of filtration systems, and compromised aesthetics (Biggs 2000, Lembi 2003, Suplee et al. 2009, Fovet et al. 2012).

While the conceptual models for adverse effects are generally well accepted, research is needed to better quantify relationships among nutrients, stream landscape- and site-scale environmental co-factors and ecological responses (Stevenson et al. 2011). In particular, thresholds in the ecological responses along these environmental gradients help develop stakeholder consensus for management action and provide a basis for evaluating the cost vs. benefits of different options (Stevenson and Sabater 2010, Muradian 2010). Used in this context, an ecological threshold refers to a marked change in a dependent variable (ecological response) within a small range of the independent variable (stressor).

Over the past decade, there has been a tremendous increase in peer-reviewed science examining levels or thresholds of nutrients adversely affecting aquatic life indicators. Most examples to date have used benthic macroinvertebrates, algal, and/or fish community composition as AL measures in empirical stress-response relationships with nutrients. In addition, most of these studies have focused on streams in temperate climates. It is unclear how applicable these thresholds are in California's Mediterranean climate, when many of the major co-factors controlling response to nutrients or algal abundance are fundamentally different (e.g., rainfall frequency, flow, temperature, and light availability). In some cases, benthic algal abundance

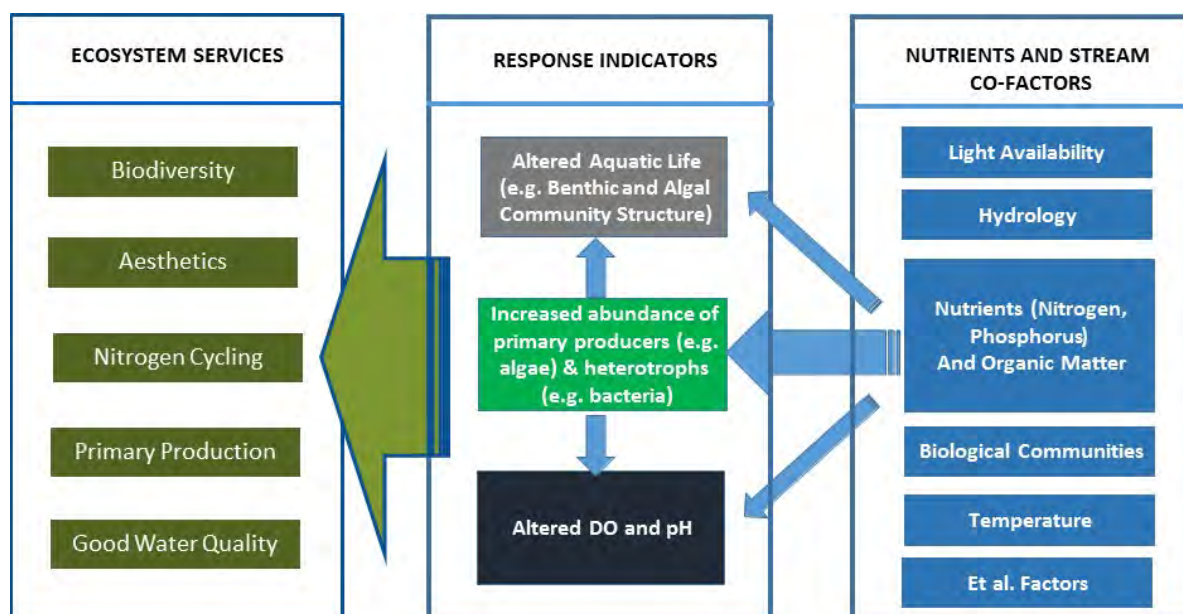


Figure 3.1. Simplified conceptual model of eutrophication in wadeable streams depicting the relationship between nutrients, stream co-factors, ecological response and ecosystem services.

and/or dissolved oxygen measures have served as intermediate response variables for inferring impacts thereto (Wang et al. 2007, Weigel and Robertson 2007, Stevenson et al. 2008, Miltner 2010, Smith and Tran 2010, and Suplee and Watson 2013).

Measures of algal abundance such as benthic chlorophyll *a*, algal percent cover, or ash-free dry mass are of interest to California water quality managers for use in assessment of eutrophication. This is because such measures have a strong mechanistic linkage with nutrients, but are more robust measures of the impacts of nutrient enrichment on the ecosystem services because they integrate stream co-factors (Figure 3.1). However, few studies have identified thresholds in relationships among algal abundance measures and aquatic life (e.g. benthic macroinvertebrate [BMI] or algal community composition). Most studies are based on watershed- or reach-scale mechanistic models that link algal abundance to dissolved oxygen. Very few empirical stress-response studies have been published looking specifically at thresholds of algal abundance (benthic chlorophyll *a*, ash-free dry mass, or macroalgal percent cover) that adversely affect BMI or algal communities in wadeable streams. Of these, the majority have been conducted in New Zealand and are, therefore, of uncertain applicability to California's wadeable streams. As California state water quality managers are interested in using biological response to assess status of stream beneficial uses vis-à-vis nutrients, this study focused on investigations of thresholds of algal abundance measures as well as nutrients on aquatic life measures.

Over the past 10 years, California's investment in a stream bioassessment program has produced a large data set that can be used to investigate the linkage of nutrient and algal abundance to ALI, employing both reference and empirical stress-response approaches. The previous section of this report summarized the reference distribution of nutrients and algal abundance indicators. The objectives of this section are to 1) investigate relationships of nutrients and primary producer abundance indicators with BMI and algal community measures of ALI and 2) determine levels of nutrients and primary producer abundance indicators associated with adverse effects to these AL measures.

3.2 Methods

3.2.1 Conceptual Approach

Three basic approaches have been used in establishing levels of stream nutrients and algal abundance that are protective of aquatic life (U.S. Environmental Protection Agency 2000): 1) reference approach, 2) empirical stress-response approach, and 3) process-based approach. Of these approaches, the reference and empirical stress-response versions are among the most commonly used quantitative approaches to establish WQOs across large geographic areas, such as California wadeable streams. The latter approach involves quantifying the relationships among stressor gradients (e.g. nutrients, algal abundance) and aquatic life measures that are representative of ecosystem services. The Biological Condition Gradient (BCG) is a useful conceptual model for empirical stress-response studies. The BCG model describes the changes in aquatic communities, measured by aquatic life indicators, as a function of stress (Davies and Jackson 2006; Figure 3.2). It predicts the transition of biotic communities, as measured by ALI indicators, as a function of increasing stress, from pristine to slightly modified ecological condition, then moderate, and finally, very low ecological condition. These relationships can be linear or non-linear in nature.

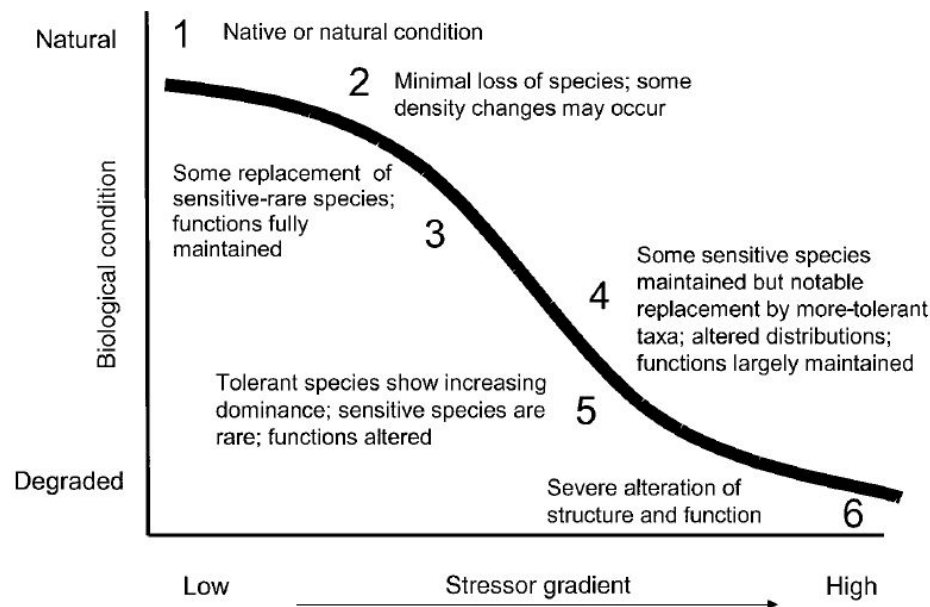


Figure 3.2. Conceptual model depicting stages of change in biological conditions in response to an increasing stressor gradient. Reproduced from Davies and Jackson (2006).

In this study, we investigated the relationships between eutrophication stressors (e.g., nutrient and primary producer abundance indicators) and BMI and algal community structure as measures of ALI (Figure 3.1). BMI and algal community structure were selected as ALI measures because: 1) they are the assemblages of choice for bioassessment in California statewide and regional programs, 2) BMI and/or diatom community composition have been used as the basis for WQO development in various other states and countries, and 3) a large and geographically broad set of ambient survey data is available for both assemblages, using standardized field and laboratory protocols (Ode 2007 and Fetscher et al. 2009). Other potential indicators of ALI attainment, such as dissolved oxygen or pH, could be used for setting biomass/nutrient WQOs, but data on diel ranges and fluctuations are not available statewide. A large number of BMI and algal IBI metrics were

evaluated in order to better understand the complexity of the community response to the chosen stressor gradients.

A wide variety of statistical methods have been used to model the fundamental relationships among stressors, community responses, and environmental co-factors that mediate response to stress. With respect to setting quantitative water quality goals, two approaches are commonly used (Figure 3.3): 1) statistical change point detection and 2) regression methods to relate stressors to quantitative ecosystem service targets (e.g. percentile of index of biological integrity corresponding to a percentile of reference sites) (EPA 2010). In this study, we investigated where along the gradients of nutrients and primary producer abundance these ALI measures exhibited break points, as evidence of thresholds of adverse effects. We focused on breakpoints or thresholds rather than regressions that extrapolate to quantitative ecosystem service targets because: 1) California has not yet officially adopted into policy stream BMI or algal IBI targets, and 2) break points and other types of thresholds are valuable to describe the response surface of ecosystem change to stressors over the BCG (Stevenson et al. 2011).

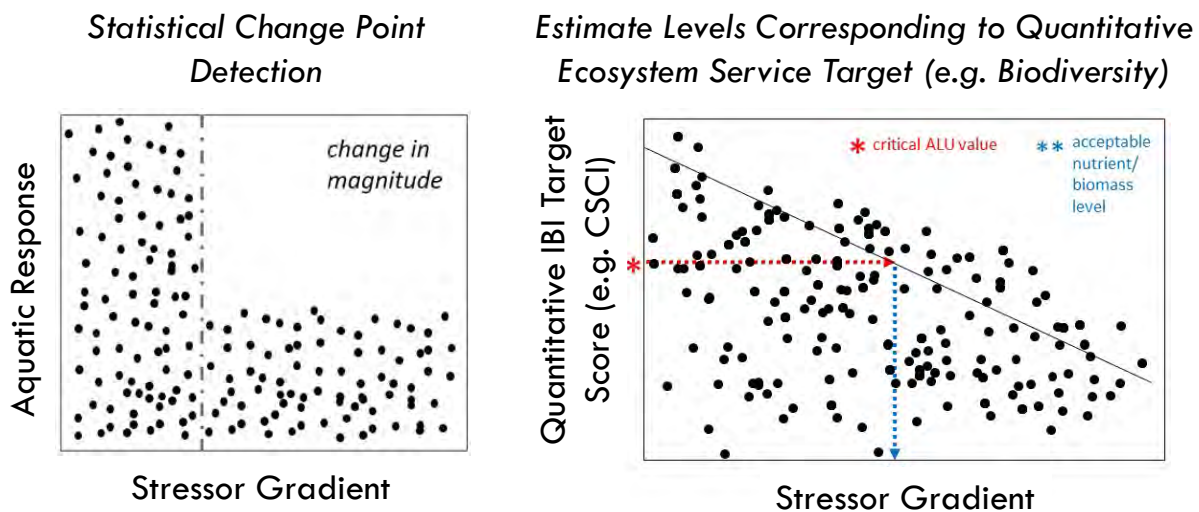


Figure 3.3. Examples of two statistical approaches used to derive quantitative water quality goals. (EPA 2010)

3.2.2 Aquatic Life and Stressor Indicators

We utilized several biotic assemblages in order to examine multiple lines of evidence for effects of algal biomass and nutrients on stream communities. Some measures from these assemblages were selected because they have explicit connections to eutrophication as known indicators of dissolved oxygen, organic matter, or stream nutrient levels, allowing us to remove to some degree the effect of other confounding stressors. Other ALIs were selected because they were developed to serve as indicators of overall stream condition. Thus they may be responsive to changes in the stream environment resulting from nutrient enrichment, even if they were not developed specifically for assessing nutrient (and excessive biomass) impacts. Furthermore, within assemblages, we used several types of metric and index that describe different aspects of the communities or summarize community composition as a whole. This facilitated an evaluation of how different levels of stress may have different degrees of effect on stream communities. An understanding of the magnitude and extent of effects of stress on biotic communities across such a “response surface” conveys information relevant for risk classification (Tetra Tech 2006). Within the BMI and

algal assemblages, we used three basic types of ALI measures: 1) “raw” community composition, as summarized in the form of axes from non-metric multidimensional scaling ordinations (NMS; see Section 3.2.5), 2) calculated metrics that describe specific aspects of a biotic community according to taxon-specific attributes/ecological preferences, and 3) calculated multimetric indices that provide more holistic ways of summarizing community composition. Unlike the NMS axes, metrics and indices have pre-established polarity of scoring that is indicative of “good” vs. “bad” water-body condition.

To facilitate interpretation of the results of our analyses within the context of the BCG conceptual model, we grouped the metric- and index-based ALI measures into four categories (Table 3.1):

- **Sensitive:** metrics based on “sensitive” taxa, i.e., those that are known, based on the literature, to be highly responsive to relatively low levels of generalized stress. Also included in this group are “tolerant” taxa because of the loosely inverse relationship between metrics describing proportion of sensitive taxa and proportion of tolerant.
- **Low-nutrients:** metrics based on taxa that have been associated with low-nutrient conditions by previous studies in the literature
- **Eutrophication:** metrics based on taxa that are tolerant to various aspects of eutrophication, according to the literature
- **Integrative:** indices that provide an integrative measure of community composition to provide inference into overall water-body condition

Along the BCG gradient, sensitive metrics would be expected to respond at level 2, with functional changes (eutrophication metrics) occurring at level 4 and integrative indices showing significant impacts at levels 4 and 5.

Table 3.1. Description of aquatic life indicators (ALIs) used in the analyses. Indicators are listed in the table alphabetically according to the shorthand version of the name, grouped by assemblage type. The ALI categories used, INT= integrative, SEN= sensitive, EUTRO= eutrophication and NUT = low nutrient, are defined above. IBI =index of biotic integrity.

Abbreviated Name	ALI Category	Description of Variable	ALI Response to Stress
ALI - Benthic Macroinvertebrate			
CSCI	INT	California Stream Condition Index (the BMI-based statewide multimetric index for stream bioassessment; Mazor et al., under review)	decrease
EPT_Percent ⁹	SEN	percent BMI individuals that are Ephemeroptera, Plecoptera, or Trichoptera	decrease
EPT_PercentTaxa	SEN	percent BMI taxa that are Ephemeroptera, Plecoptera, or Trichoptera	decrease
EPT_Taxa	SEN	number of BMI taxa that are Ephemeroptera, Plecoptera, or Trichoptera	decrease

⁹ Not considered “integrative”, because EPT account for only a subset of the whole community.

Table 3.1 (continued)

Abbreviated Name	ALI Category	Description of Variable	ALI Response to Stress
Intolerant_Percent	SEN	percent BMI individuals that are "intolerant"; Ode et al. 2005	decrease
Intolerant_PercentTaxa	SEN	percent BMI taxa that are "intolerant"; Ode et al. 2005	decrease
Intolerant_Taxa	SEN	number of BMI taxa that are "intolerant"; Ode et al. 2005	decrease
O/E	INT	observed over expected taxa from RIVPACS models for BMI taxa; Mazor et al., under review	decrease
Shannon_Diversity	INT	Shannon diversity index for BMI taxa	decrease
Simpson_Diversity	INT	Simpson diversity index for BMI taxa	decrease
Taxonomic_Richness	INT	richness of BMI taxa	decrease
Tolerant_Percent	SEN	percent BMI individuals that are "tolerant"; Ode et al. 2005	increase
Tolerant_PercentTaxa	SEN	percent BMI taxa that are "tolerant"; Ode et al. 2005	increase
Tolerant_Taxa	SEN	number of BMI taxa that are "tolerant"; Ode et al. 2005	increase
BMI community	INT	NMS axis 1 score (from ordination of BMI community composition data)	no expectation
ALI - Diatom			
D18	INT	diatom IBI; Fetscher et al. 2014	decrease
propAchMin	SEN	proportion of diatom valves that are Achnantheidium minutissimum	decrease
RAWDO100	EUTRO	proportion diatoms requiring nearly 100% DO saturation; van Dam et al. 1994	decrease
RAWDO50	EUTRO	proportion diatoms requiring at least 50% DO saturation; van Dam et al. 1994	decrease
RAWeutro	EUTRO	proportion eutrophication indicator diatoms; van Dam et al. 1994	increase
RAWlowN	NUT	proportion low-N indicator diatoms; Potapova and Charles 2007	decrease
RAWlowP	NUT	proportion low-P indicator diatoms; Potapova and Charles 2007	decrease
RAWNhet	EUTRO	proportion nitrogen-heterotroph diatoms; van Dam et al. 1994	increase
diatom community	INT	NMS axis 1 score (from ordination of diatom community composition data)	no expectation
ALI – Hybrid of Diatoms and Soft-bodied Algae			
H20 ¹⁰	INT	diatom + soft algae ("hybrid") IBI; Fetscher et al. 2014	decrease
H21	INT	diatom + soft algae ("hybrid") IBI; Fetscher et al. 2014	decrease

¹⁰ H20, H21, and H23 differ in terms of the type of soft-algal information they include. H20 includes only species presence/absence, H21 includes only species biovolumes, and H23 includes both types of data. For more details, see Fetscher et al. (2014).

Table 3.1 (continued)

Abbreviated Name	ALI Category	Description of Variable	ALI Response to Stress
H23	INT	diatom + soft algae ("hybrid") IBI; Fetscher et al. 2014	decrease
ALI – Soft Algae			
propTaxaZHR	NUT	proportion of total of total soft-algae taxa recorded that are in the Zygnemataceae, heterocystous cyanobacteria, or Rhodophyta	decrease
RAWlowTPsp	NUT	proportion of soft algal taxa that are considered "low TP" indicators; Fetscher et al. 2014	decrease
RAWmeanZHR	NUT	mean of the metrics propTaxaZHR and RAWpropBiovolZHR; Fetscher et al. 2014	decrease
RAWpropBiovolChlor	EUTRO	proportion of total soft algae biovolume that is Chlorophyta	increase
RAWpropBiovolZHR	NUT	proportion of total soft algae biovolume that is in the Zygnemataceae, heterocystous cyanobacteria, or Rhodophyta	decrease
RAWpropGreenCRUS	EUTRO	proportion of green algal biovolume belonging to Cladophora glomerata, Rhizoclonium hieroglyphicum, Ulva flexuosa, or Stigeoclonium spp.	increase
S2	INT	soft algae IBI; Fetscher et al. 2014	decrease

The focal stressor gradients used in the analyses were of two categories: 1) primary producer abundance, including measures of benthic chlorophyll *a*, AFDM and algal and macrophyte percent cover and 2) concentrations of total and dissolved inorganic nitrogen and phosphorus. In addition to these, other environmental gradients that could influence relationships between the focal stressor gradients and ALIs were included in analyses, where possible. These are listed in Table 3.2.

Table 3.2 Descriptions of the focal stressor gradients and other explanatory variables used in the analyses, listed in the table alphabetically according to the shorthand version of the name, within their respective categories.

Abbreviated Name	Description of Variable
PRIMARY PRODUCER ABUNDANCE	
AFDM	benthic ash-free dry mass
PCT_MAP	macroalgal percent cover
PCT_MCP	macrophyte percent cover
PCT_MIAT1	percent presence of thick (1mm+) microalgae
<i>none</i>	benthic chlorophyll <i>a</i>
<i>none</i>	soft algal total biovolume
NUTRIENT	
NH ₄	ammonium
NO _x	nitrate + nitrite
SRP	soluble reactive P

Table 3.2 (continued)

Abbreviated Name	Description of Variable
TN	total nitrogen
TP	total phosphorus
LANDSCAPE- Development	
AG_2000_1K	percent agricultural land use within a 1-km radius from sampling site
AG_2000_5K	percent agricultural land use within a 5-km radius from sampling site
AG_2000_WS	percent agricultural land use in catchment
CODE21_2000_1K	percent "Code 21" ¹¹ land use within a 1-km radius from sampling site
CODE21_2000_5K	percent "Code 21" land use within a 5-km radius from sampling site
CODE21_2000_WS	percent "Code 21" land use in catchment
URBAN_2000_1K	percent urban land use in catchment within a 1-km radius from sampling site
URBAN_2000_5K	percent urban land use in catchment within a 5-km radius from sampling site
URBAN_2000_WS	percent urban land use in catchment
<i>none</i>	site disturbance class
LANDSCAPE-Geographic	
<i>none</i>	ecoregion
<i>none</i>	latitude
<i>none</i>	longitude
<i>none</i>	site elevation
<i>none</i>	watershed area
<i>none</i>	percent sedimentary geology in the catchment
LANDSCAPE – Meteorological	
<i>none</i>	mean monthly % cloud cover (3-mo antecedent mean)
<i>none</i>	mean monthly max temperature (3-mo antecedent mean)
<i>none</i>	mean monthly solar radiation (3-mo antecedent mean)
<i>none</i>	total precipitation (3-mo antecedent total)
LOCAL PHYSICAL HABITAT (PHab)	
PCT_CPOM	percent cover of coarse particulate organic matter in streambed
PCT_FN	percent cover of fine substrata in streambed
PCT_SAFN	percent sand + fines in streambed
W1_HALL	a riparian disturbance index; Kaufmann et al. 1999
XDENMID	percent canopy cover
<i>none</i>	days of accrual (i.e., estimated number of days since last scour event)
<i>none</i>	mean stream depth
<i>none</i>	mean stream width

¹¹ "Code 21" encompasses a wide range of land uses primarily characterized by heavily managed vegetation (e.g., low-density residential development, parks, golf courses, highway medians)

Table 3.2 (continued)

Abbreviated Name	Description of Variable
<i>none</i>	slope, reach-level
<i>none</i>	stream discharge
<i>none</i>	stream temperature
WATER CHEMISTRY (GENERAL)	
<i>none</i>	alkalinity
<i>none</i>	conductivity
<i>none</i>	turbidity

3.2.3 Detection of Ecological Thresholds

Different types of ecological threshold exist (Figure 3.4). The relationship between gradient and response can involve a change in the slope relating the response variable to the stressor gradient (as in the graph on the left of Figure 3.4), or a change in magnitude of the response variable's value (as depicted in the "step" response model on the right side; Brendan et al. 2008). The dashed line in each figure represents a response threshold. It is important to note by the time the threshold is reached in Figure 3.2, an ecologically significant change may have occurred in the value of that ALI.

Some stressor-response relationships may involve a more complex type of change-in-slope than that described in Figure 3.4. For instance, low values of the stressor gradient may be tolerated within a certain range at the low end of the stressor gradient (i.e., the "reference envelope") without a concomitant decline in ALI response. In addition, Cuffney et al. (2010) distinguished between "resistance thresholds" (marked by a sharp decline in ecosystem condition following an initial no-effect zone) and "exhaustion thresholds" (marked by a sharp transition to zero slope at the end of a stressor gradient at which point the ALI response is essentially saturated; Figure 3.5). Furthermore, different ALI measures, within or between biotic assemblages, can exhibit different thresholds of response to a given stressor gradient depending upon their varying levels of susceptibility. In aggregate, this array represents the response surface of the BCG (Davies and Jackson 2006; Figure 3.2).

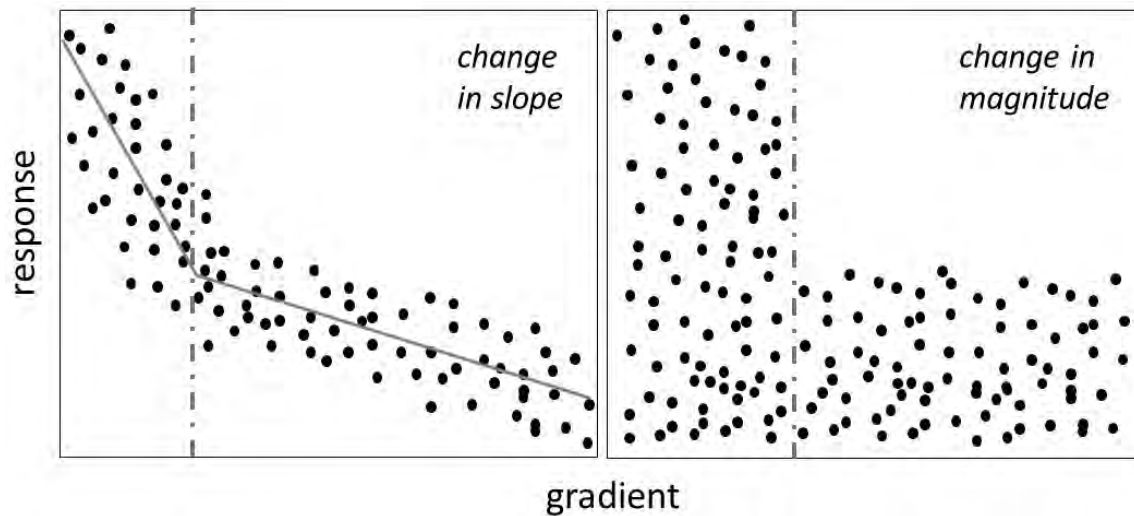


Figure 3.4. Examples of types of threshold relationships.

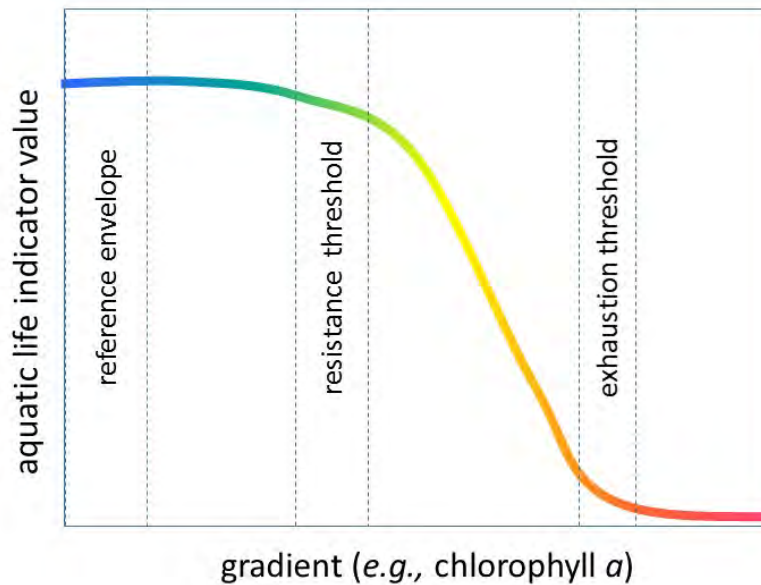


Figure 3.5. ALI response to a stressor gradient showing the “reference envelope” along with “resistance” and “exhaustion” thresholds.

Statistical techniques vary in terms of what types of threshold they are most appropriate for detecting (Brenden et al. 2008). We employ multiple statistical analyses in examining thresholds because each method has a unique set of attendant advantages and limitations and no one technique is universally accepted among scientists. Furthermore, when results of different tests converge on similar values, there is a greater level of confidence and less likelihood that a given result was merely an artifact of the statistical method used. We began the study by conducting a small set of initial analyses across a broad swath of the available ALI measures and biomass/nutrient gradients, then conducted additional analyses on the subset of ALI-biomass/nutrient combinations that yielded the strongest relationships, in order to look for support for preliminary thresholds identified. Section 3.2.2 and 3.2.4 summarize the techniques used in this section and what threshold types they can detect. Once we identified thresholds across analyses, we summarized results by grouping them according to the four ALI categories listed above, in order to provide a snapshot of the study’s findings within the context of the BCG concept.

For the purposes of this study, we define “endpoints” or “objectives” to refer to policy decisions to regulate levels that are deemed an unacceptable risk, while “thresholds” refer to the output of statistical analyses. The results of this study may be among those that the SWRCB considers in its synthesis of the science that will support policy decisions. However, the thresholds produced in the course of this study should not be construed as policy.

3.2.4 Data Sources

The dataset used for the analyses in this chapter is described in detail in Chapter 2, Section 2.2.1. Included are sites from both the probability and targeted surveys. Table 3.1 lists the response variables (ALI indicators), while Table 3.2 provides a list of the stressor variables, as well as site-specific and landscape-level co-factors. Sample sizes for analyses varied according to the variables used (Table 3.3).

Table 3.3. Sample sizes.

Gradient	ALI Type			
	BMI	diatom	hybrid	soft
Chlorophyll <i>a</i>	545	850	784	804
AFDM	521	819	756	776
PCT_MAP	452	745	679	700
PCT_MCP	452	745	679	700
PCT_MIAT1	452	744	679	700
TN	574	873	775	796
TP	582	892	767	788
NO _x	593	931	769	790
SRP	581	935	769	790

In looking for evidence of biomass thresholds for impacts to ALIs, our primary focus was on algal biomass as measured by benthic chlorophyll *a* concentrations. Chlorophyll *a* is a primary line of evidence in the recommended NNE endpoints for Wadeable streams (TetraTech 2006), and in general, it is the typical means of quantifying eutrophication of Wadeable streams (US Environmental Protection Agency 2000). However, alternative indicators of stream primary producer abundance were explored 1) to which ALIs may be more directly responsive and/or 2) which may be more directly tied to nutrient impacts. These included AFDM, soft algal total biovolume, macroalgal percent cover (PCT_MAP), macrophyte cover (PCT_MCP), and percent presence of thick (1mm+) microalgae (PCT_MIAT1), for certain analyses.

For the BMIs, ALIs included the California Stream Condition Index (CSCI), a draft statewide multimetric index for stream assessment recently developed by Mazor et al. (under review), as well as a statewide “Observed/Expected” RIVPACS-type (Wright et al., 1993) predictive model based on BMI taxa, also developed by Mazor et al. (under review). In addition, several classical metrics based on the BMI community were used. Finally, community composition data were ordinated using Nonmetric Multidimensional Scaling (NMS) to reduce the dimensionality of the dataset, allowing the use of NMS axis scores as response variables that summarize information about the BMI community in each sampling site. We made the assumption that values of the ALI response to the right of an identified threshold along an increasing stressor represent adverse effects relative to the values on the left. Similarly, King and Richardson (2008) used an NMS-based approach to assess biological impairment in the Everglades resulting from experimental P additions. For benthic stream algae, some ALIs were based on indices developed by Fetscher et al. (2014), which use community composition of diatoms and/or non-diatom (“soft”) algae. Although developed for use in southern California streams, they have some applicability in other parts of the state (Fetscher et al. 2013). Selected metrics that comprise the IBIs, and NMS scores based on diatom community composition, also were included. Lists of the ALI variables, primary producer biomass variables, and other variables (landscape, meteorological, local physical habitat, and water chemistry) used in the analyses are provided in Tables 3.1 and 3.2.

3.2.5 Data Analyses

There are a number of challenges to determining the existence of ecological thresholds. First, the noisiness of the dataset may interfere with threshold detection by making it difficult to discern whether or where there is a clear, abrupt change in the response variable along the stressor gradient. Second, different taxa within any given biotic assemblage may respond somewhat differently to any given stressor (Baker and King 2010). Finally, multiple factors potentially influence the value of any given ALI response measure, making it difficult to ascribe ALI response solely to the stressor of interest.

Analytical techniques differ in terms of whether and how confounding factors such as other sources of stress can be taken into account, and also differ in their susceptibility to outliers. As such, we used a variety of techniques to attempt to mitigate these challenges and seek consensus in results among different techniques (Dodds et al. 2010; Smucker et al. 2013a,b). We also looked at different measures of ALI within and across biotic assemblages. In some analyses, we were able to control for potential confounding factors that could influence ALI response variables. This was made possible by the large size of our dataset and the fact that sites throughout California, and across varying levels and types of anthropogenic disturbance, were sampled. In addition, a large number of local physical habitat (PHab), landscape-level geographic, meteorological variables, and water chemistry measures were available for most of the sampling sites.

The analytical techniques used for exploring potential biomass thresholds for ALI response can be grouped into two broad categories. Table 3.4 provides a summary of the key assumptions, strengths and limitations of the different approaches:

- Analyses that use basic species data for evaluating shifts in “raw” community composition
- Nonmetric multidimensional scaling (followed by classification and regression tree analysis; NMS/CART)
- Threshold indicator taxa analysis (TITAN)
- nonparametric change point analysis (nCPA)
- Analyses that include higher-order variables, such as biotic metrics and indices, as integrative measures reflecting aspects of community composition
- Piecewise regression
- Significant zero crossings (SiZer)
- Boosted regression trees (BRT; including partial Mantel tests to pre- and post-screen predictor variables. Boosted regression tree and partial Mantel tests were also used for examination of nutrient and other environmental co-factor effects on biomass—see Chapter 4).

The following section provides a brief introduction to each analytical technique.

Table 3.4. Summary of analytical techniques used for threshold estimation.

Analytical Technique	Strengths	Limitations	Type of Threshold (refer to Figure 3.4)
CART	Number of thresholds does not have to be established a priori but can be manually limited by user. Least absolute deviation method can be used to reduce sensitivity to outliers. Can handle multiple potential predictors of thresholds.	This technique can overfit classification and regression trees. Bootstrapping is desirable to determine robustness and level of confidence associated with solutions. Will find a break-point whether one exists or not.	magnitude
TITAN	Provides separate change points for taxa to allow user to assess a community-level change point (if it exists); multiple assessment measures are available for determining confidence in change points	Some degree of interpretation is involved in determining what constitutes a “community-level change point”	magnitude
Piecewise Regression	Intuitive, conceptually easy for non-experts to grasp; provides several measures of uncertainty for determining confidence in the breakpoint	User must specify number of breakpoints <i>a priori</i> ; this technique will “find a breakpoint” whether a true threshold exists or not; sensitive to outliers	slope
SiZer	No requirement for a priori determination of the number of break points	SiZer maps can be difficult to interpret; output does not include a numeric threshold (only visual, subject to interpretation); no measure of uncertainty	slope
BRT	Insensitive to data distributions as well as the presence of outliers, can fit both linear and nonlinear relationships, and automatically handles interaction effects between pairs of predictors	Partial effects plots are created using the mean of other predictor variables so care must be taken in interpretation if interactions exist.	slope (thresholds identified from partial dependence plots); magnitude thresholds can be determined through subsequent CART analysis

Nonmetric multidimensional scaling (NMS)

Nonmetric multidimensional scaling is an ordination technique that reduces the dimensionality of information in a dataset in order to summarize its major gradients. The product of an NMS ordination conducted on community composition data is a series of scores. The plots produced provide insight into similarity in species composition among samples. The closer two sample points are to one another within NMS ordination space, the more similar they are in terms of the types and proportions of species they contain.

NMS analyses were performed separately for BMI and diatom community composition. Proportion data were used for both the BMI and diatom-based ordinations. Only sites with at least 450 BMI (or diatom) individuals were used in the analyses. Furthermore, only taxa that represented at least 1% relative abundance for at least two sites in the dataset were included. NMS was run using PC-ORD software (version 6; McCune and Grace, 2002) with the Bray-Curtis distance measure and “slow and thorough” autopilot mode. This measure and mode runs initial ordinations to determine the best dimensionality (stability criterion of 0.00001, maximum of six axes, 40 runs with real data, and 50 randomized runs). A second round of ordinations is run using the selected dimensionality (stability criterion of 0.00001, one run with real data, up to 400 iterations). From each NMS ordination, we selected the axis that was most strongly associated with biomass and nutrients for use as a response variable in subsequent analyses (e.g., CART, see below).

For determining Pearson correlation coefficients that incorporate sample weights to describe the relationship between NMS scores and biomass/nutrient gradients, we used the R package “weights” (Pasek and Tahk 2012). Sample weights were calculated as the number of stream kilometers represented by each sampling site (weights account for differences in the number of sites in each stratum and stream kilometers in the stratum; see Chapter 2). To facilitate use of the “weights” package, which provides only Pearson correlation coefficients, the non-normal data were first rank-transformed.

Classification and regression tree (CART) analysis

Classification and regression trees (De’ath and Fabricius 2000) is an analytical method that “builds trees” via a recursive series of binary splits of the data set into successively smaller groups of observations. Splitting occurs along one or more explanatory variables, which can be categorical, or continuous. For classification trees, during each recursion, the split chosen maximizes homogeneity of response values within the resulting two groups. We used CART to identify “cut points” (i.e., locations of the split) in explanatory variables, such as biomass and nutrients, as an indication of thresholds of their effects on ALI response variables. The NMS 1 axis scores for the BMI and diatom communities were used as response variables in the CART analyses.

Depending upon the version of CART analysis run, explanatory variables included either all of the following: Chlorophyll *a*, AFDM, the different nutrient types, ecoregion, and site disturbance class (in which case, the output of the analysis is referred to as “ALL”), or only a single explanatory variable was used (either AFDM alone or only chlorophyll *a* [“CHLA”] alone). For the latter two versions of the analysis, the number of splits used in the tree building was restricted to two. Because community composition could vary geographically and, therefore, might influence the outcome of the analysis, CART analyses were run both statewide and within the South Coast ecoregion (Figure 2.1), where the highest density of data were available. These latter groupings of data facilitated an assessment of the possible effect of biogeographic variation on cut point values.

CART analysis was carried out using SYSTAT v. 13 software (Systat Software, Inc., Chicago, IL) using the following options: least absolute deviation (which renders the analysis less sensitive to outliers), a maximum number of 2 splits, p-values of 0.05 for total and incremental variance explained and minimum of five objects in final classes. One thousand bootstrap replicates were run to generate confidence intervals for split values. In this and subsequent analyses, the number of bootstrap replicates run is chosen to ensure robust and reasonably precise results within practical limits.

Threshold indicator taxa analysis (TITAN)

Threshold indicator taxa analysis (TITAN) is an analytical technique that represents a combination of indicator analysis and change point analysis. TITAN identifies if synchronous declines occur in multiple species along an environmental gradient¹² of interest. To conduct TITAN, we used the R package “mvpart” (version 1.6-0, Therneau and Atkinson 2009) with scripts modified by Baker and King (2010). We used TITAN to look at BMI and diatom species responses to biomass and to nutrient levels in order to identify potential thresholds. TITAN is still being debated as to the validity of change point values identified and for that reason may be considered more exploratory. TITAN uses indicator value scores from indicator species analysis (Dufrêne and Legendre 1997) to integrate occurrence, relative abundance, and directionality of taxa responses. It identifies the optimum value (i.e., “change point”) of a continuous variable, x , that partitions sample units while maximizing taxon-specific scores. Indicator z scores standardize original scores relative to the mean and standard deviation of permuted samples along x , thereby emphasizing the relative magnitude of change and increasing the contributions of taxa with low occurrence frequencies but high sensitivity to the gradient. TITAN distinguishes negative (z^-) and positive (z^+) taxa responses to the gradient. It tracks cumulative responses of declining $\text{sum}(z^-)$, which we refer to as “decreasers”, and increasing $\text{sum}(z^+)$ taxa, which we refer to as “increasers”, in the community. Narrow peaks in $\text{sum}(z)$ scores along the environmental gradient of interest (x -axis) and the presence of many taxa with change points at similar levels of that gradient indicate a community threshold.

Bootstrapping is used to estimate indicator taxon “reliability” and “purity” as well as uncertainty around the location of individual taxa and community change points. Indicator “purity” as defined by Baker and King (2010) is the proportion of change-point response directions (positive or negative) among bootstrap replicates that agree with the observed response. As such, “pure indicators” are those that are consistently assigned the same response direction, regardless of abundance and frequency distributions generated by resampling the original data. For the purposes of this report, “pure taxa” are defined as those for which $\text{purity} \geq 0.95$. Indicator “reliability” is defined by Baker and King (2010) as the proportion of bootstrap change points whose indicator value scores consistently result in P-values below one or more user-determined probability levels. For the purposes of this report, “reliable indicators” are those with repeatable and consistently large indicator value maxima (specifically, ≥ 0.95 of the bootstrap replicates achieving $P \leq 0.05$). Examples of TITAN output and its interpretation are provided in Figures C.1 and C.2. We used 500 bootstrap replicates in order to identify pure and reliable indicator taxa, and to establish uncertainty around taxa change-points (i.e., 5 and 95% quantiles; Baker and King 2010). In order to downweight the influence on

¹² A multivariate version of this package, that would allow multiple stressor gradients to be used in determining taxon-specific z -scores, is currently under development (M. Baker, personal communication 2014), and not available for use in the present version of the report.

indicator values of taxa with high relative abundances, both BMI and diatom data were analyzed as $\log_{10}(x+1)$ -transformed species relative abundances. Taxa with fewer than five occurrences on either side of a partition (during the TITAN binary partitioning process) were eliminated.

Nonparametric change point analysis (nCPA)

Nonparametric change point analysis (Qian et al. 2003) is an analytical technique used when a step function (i.e., change in magnitude, as described in Figure 3.4) is assumed. It seeks the point along a gradient at which the sum of the deviance in the response variable, to the left of the point, plus the deviance to the right of the point, is maximally lower than that across the data set as a whole. The sum of the deviance values is calculated iteratively along the gradient, and the point at which maximal deviance reduction is realized reflects a community-level change-point (or threshold) in the relationship.

To conduct nCPA on BMI and diatom community data, we used the R package “mvpart” (version 1.6-0, Therneau and Atkinson 2009) with scripts modified by Baker and King (2010). ALI response variables for the nCPA analyses were based on Bray-Curtis and Euclidean distances as the dissimilarity metrics for the community data. In addition, 5th and 95th quantiles were determined using 500 bootstrap replicates. Data were prepared for analysis in the same way as described above for TITAN.

Piecewise regression

We used piecewise linear regression to detect change in slope in the relationship between ALI response variables and biomass/nutrient gradients to search for possible “breakpoints” in the response of each available ALI variable to biomass and nutrient gradients (Muggeo 2003). Before running piecewise regression analysis, scatterplots for all ALIs against the various biomass and nutrient gradients were visualized (Muggeo 2008), and there were no cases in which it was clear, based on the plots, that >1 breakpoint was present. Therefore, as a conservative default, all analyses were run coercing a single breakpoint. Piecewise regression was one of the few analyses for which sample weights could be incorporated. For each ALI/gradient combination, the analyses were run both with and without incorporating sample weights as described above.

Because piecewise regression will always “find a break point” whether or not one truly exists, we created a set of four criteria against which to evaluate the output of each analysis, in order to distill the full list of ALI/gradient combinations into a subset for which high confidence could be ascribed to the breakpoints identified. Two levels of criteria were employed: “strict” and “relaxed”. In order for the piecewise regression output for a given ALI/gradient combination to be assigned to one of these levels: 1) it had to result in a significant Davies' (1987)¹³ test (indicating that the slopes on either side of the break point were significantly different from one another at $\alpha = 0.05$); 2) at least one of the two slopes had to be significantly different from zero (as assessed by ensuring that the 95% CI [confidence interval] around at least one of the slopes did not straddle zero); 3) the CI around the break point had to be sufficiently narrow (i.e., the CI width divided by the breakpoint value had to be <0.5 for the “relaxed” level, and <0.3 for the “strict”); and 4) the adjusted R^2 for the regression had to be sufficiently high (i.e., at least 0.1 for the “relaxed” level, and at least 0.25 for the “strict”). To conduct piecewise regressions, we used the R package “segmented” (Muggeo 2008).

Significant Zero crossings (SiZer)

¹³ Using a standard “k” value of 10, *per* Muggeo (2008).

The “significant zero crossings” analytical technique fits multiple smoothing curves through a scatterplot using locally weighted polynomials, with the goal of assessing the nature of the first derivative of each curve (indicating the direction of slope) at intervals along a gradient (Chaudhuri and Marron 1999). SiZer attempts to distill the significant, “real” features of a curve in a dataset by “looking past” any noise that may be present. To accomplish this, curve smoothing of the relationship of a response variable (e.g., an ALI measure) to a gradient (e.g., a biomass or nutrient gradient) is conducted at various bandwidths. The bandwidth is the ranges along the x-axis over which the polynomial smoothing is conducted. Unlike the approach for piecewise regression using the “segmented” package, SiZer does not require the user to propose a priori the number of breakpoints upon which to base output. Thus the data can freely “speak for themselves” as to how many thresholds may be present.

The output of a SiZer analysis is a “SiZer map”, which uses color coding to allow users to visualize where along the gradient (the x-axis) the first derivative is significantly positive (depicted by blue) and where it is significantly negative (depicted by red) for different curve-smoothing bandwidths (which are represented on the y-axis as “h” (\log_{10} -transformed)). Areas in which the derivative is neither increasing nor decreasing significantly are indicated by purple, and grey means that data are insufficient to make a determination for that gradient-bandwidth combination. SiZer does not explicitly provide threshold values, but the SiZer map supplies the user with the means to make inferences. Specifically, a point along the stressor gradient where a narrow band of transition from purple to red, purple to blue, or vice versa is consistent across many bandwidths is a compelling indication of a significant and robust change in slope (and corresponding threshold of response). Examples of SiZer maps and their interpretation are provided in Figure C.3. To produce SiZer maps, we used the R package “SiZer” (Sonderegger 2011).

Boosted regression trees (BRT)

With the exception of the CART analysis, the analyses described above all involved looking at response of some form of ALI to a single independent variable (i.e., chlorophyll *a* concentration, or some other biomass or nutrient measure). In order to look at the effect of biomass/nutrients within the context of other potential predictors, which could confound ALI responses, and to facilitate an evaluation of the relative importance of biomass/nutrients as compared to other potential determining factors, we employed boosted regression tree analysis. BRT analysis was used for two purposes in this report to assess:

- Biomass and nutrient relationships with ALIs, as well as to look for evidence of thresholds of response to biomass/nutrients, while holding other predictors constant
- Nutrient and other environmental co-factor relationships with biomass of various types

One of the daunting aspects of determining nutrient effects on stream primary producer biomass is the fact that nutrients do not act in isolation. Rather, their influence on biomass is mediated by any of a number of environmental co-factors, which can limit the potential for biomass accrual even when nutrient levels are high. As such, determining the influence of nutrients on biomass requires accounting for the effects of the co-factors. To this end, we used BRT analysis to investigate nutrient effects on biomass levels in conjunction with other environmental co-factors.

BRT combines the strengths of regression trees with a machine-learning algorithm called “boosting”, which is an adaptive method for combining many simple models to give improved predictive performance. The final BRT model is essentially an additive regression model in which individual terms are simple trees, fitted in a

forward, stage-wise fashion. BRTs randomly add predictor variables and identify “relative influence” of each predictor based on how often it is selected and whether it improves the model. Advantages of BRT analysis are that it is insensitive to data distributions (thus requiring no transformation) as well as the presence of outliers, can fit both linear and nonlinear relationships, and automatically handles interaction effects between pairs of predictors. BRTs can also be used to plot the “partial dependence” of the response variable on an individual predictor, which is a way of looking at the relationship of the response variable to the predictor when all other predictors are held constant, at their mean values in the dataset (Elith et al. 2008).

Rather than conducting BRT analyses on all possible ALI response variables, we selected a cross-section of variable types from each assemblage and sought to reduce redundancy. For example, for the BMI assemblage, only one type of diversity index was used, as well as only one metric each from the BMI "trio" metric groups (i.e., Taxa, Percent, and PercentTaxa).

BRTs were run with tree complexity = 5, learning rate = 0.001, and bag fraction = 0.5, and all final models were built with >1000 trees, the number of which was optimized per model (for most models, except where noted) to maximize model performance while reducing overfitting. A 10-fold cross-validation procedure without replacement (90% training, 10% validation) was employed that used all data for training and validation steps (Elith et al. 2008). We utilized the model-simplification procedure described in Elith et al. (2008) to reduce the number of predictor variables in the final model for each ALI. For BRT analysis, we used the R package “dismo” (Hijmans et al. 2013).

Partial Mantel tests

Because some of the predictors available for use in the BRT analyses (i.e., the landscape variables at nested spatial scales; Table 3.2) had a high likelihood of being correlated with one another, we used partial Mantel tests in a “prescreening” step to determine if effects of the landscape variables at each scale could be detected after accounting for other scales. Any non-significant land-use variables were then excluded from the BRT analyses for that ALI type. We also used partial Mantel tests for “post-screening” the suite of predictor variables remaining in each final BRT model after having completed the BRT model-simplification procedure (see above) in order to determine which variables had significant partial Mantel correlation coefficients when the other predictor variables from the final BRT models were taken into account. This was accomplished by including all top-ranked (i.e., those with the highest relative influence) biomass and/or nutrient predictor variables, as well as any non-nutrient/non-biomass predictors that ranked above them. A geographic-distance variable was also incorporated, in order to evaluate the potential for spatial autocorrelation. Note that the Mantel test can return erroneously low p-values in the presence of spatial autocorrelation, thus it is important to rule out.

To prepare data for the partial Mantel tests, we first transformed all non-normal data, using arcsine-square-root for proportion data (such as land use) and \log_{10} for other data types (such as chlorophyll a , AFDM, and nutrient concentrations). We also included information on geographic distance among sites in order to test for potential spatial autocorrelation (hereafter referred to as “space”) in the ALI relationship with the nutrients/biomass and other variables (King et al. 2005). To accomplish this, we first transformed latitude and longitude into Universal Transverse Mercator (UTM) coordinates. Euclidean distance matrices were then calculated for each variable based on the transformed values. For partial Mantel tests, we used the R package “ecodist” (Goslee and Urban 2007).

3.3 Results

This section provides results from each statistical technique we employed, organized analysis-by-analysis. The final part of the results section provides a summary of all thresholds identified per ALI category, to facilitate visualization and interpretation of study results within the context of the BCG concept.

3.3.1 BMI and Diatom Responses to Biomass and Nutrient Gradients Based on Shifts in Community Composition

NMS and CART

NMS was used on BMI and diatom data to depict sampling site relationships to one another based on community composition. Proximity of sites to one another along an NMS axis is an indication that those sites share a more similar community composition than sites that are further away in ordination space. The same NMS ordination axes from this analysis were used as response variables in CART models (see Methods). The spline fits to NMS and subsequent CART analyses with NMS scores as the response variable revealed similar qualitative (NMS) and quantitative (CART) thresholds for the biomass and nutrient stressor gradients examined. CART-derived thresholds were slightly higher than perceived resistance thresholds on NMS spline plots and well below perceived exhaustion thresholds. CART analyses carried out with all explanatory variables (“ALL”) generally included AFDM, TN, or TP as splitting variables, but chlorophyll *a* was rarely included.

All four versions of the NMS analyses resulted in 3-axis solutions. In each case, NMS axis 1 had the strongest relationship with biomass and nutrients and was therefore selected. For the BMI analyses statewide, final stress was 17.7 and percent variance explained by NMS 1 was 28.7 while the results were 17.2 and 29.3, respectively, for the South Coast. For the diatom analyses, final stress was 19.2 and percent variance explained by NMS 1 was 27.3 for the statewide data set, and 19.7 and 20.6, respectively, for the South Coast. Scatterplots of statewide NMS axis 1 against biomass and nutrient gradients are provided in Figures 3.6 and 3.7.

All graphs show consistent and significant relationships between biomass/nutrient gradients and NMS scores, indicating that sites that share similar biomass/nutrient concentrations are also similar in species composition. Pearson correlation coefficients indicate that, for both assemblages, the relationships are strongest for nutrients, chlorophyll *a*, and AFDM, whereas the weakest relationships are between NMS scores and the percent cover metrics. Particularly strong are the relationships between the diatom community composition and nutrients (especially TP). Furthermore, the scatterplots show that the most pronounced relationship between the diatom community (NMS axis 1) and TP occurs between a lower qualitative threshold of approximately 0.01 and a higher one at 0.1 mg/L, whereas for TN, the most pronounced relationship for both diatom and BMI communities occurs between a lower threshold of approximately 0.1 and a higher one at 1 mg/L. These observations are corroborated by the results of the CART analyses of diatom and BMI NMS axis 1 scores (Table 3.5, Figure 3.8), in which median cut point values for TP and TN were consistently <0.1 and <1 mg/L, respectively, and closer to visually perceived resistance thresholds in spline fits. All median cut points for chlorophyll *a* were <31 mg/m², and for AFDM were <42 g/m². Note that CART analyses carried out with all explanatory variables (“ALL”) generally included AFDM, TN, or TP as splitting variables, but chlorophyll *a* was rarely included. Ecoregion and site disturbance class were never included in final trees.

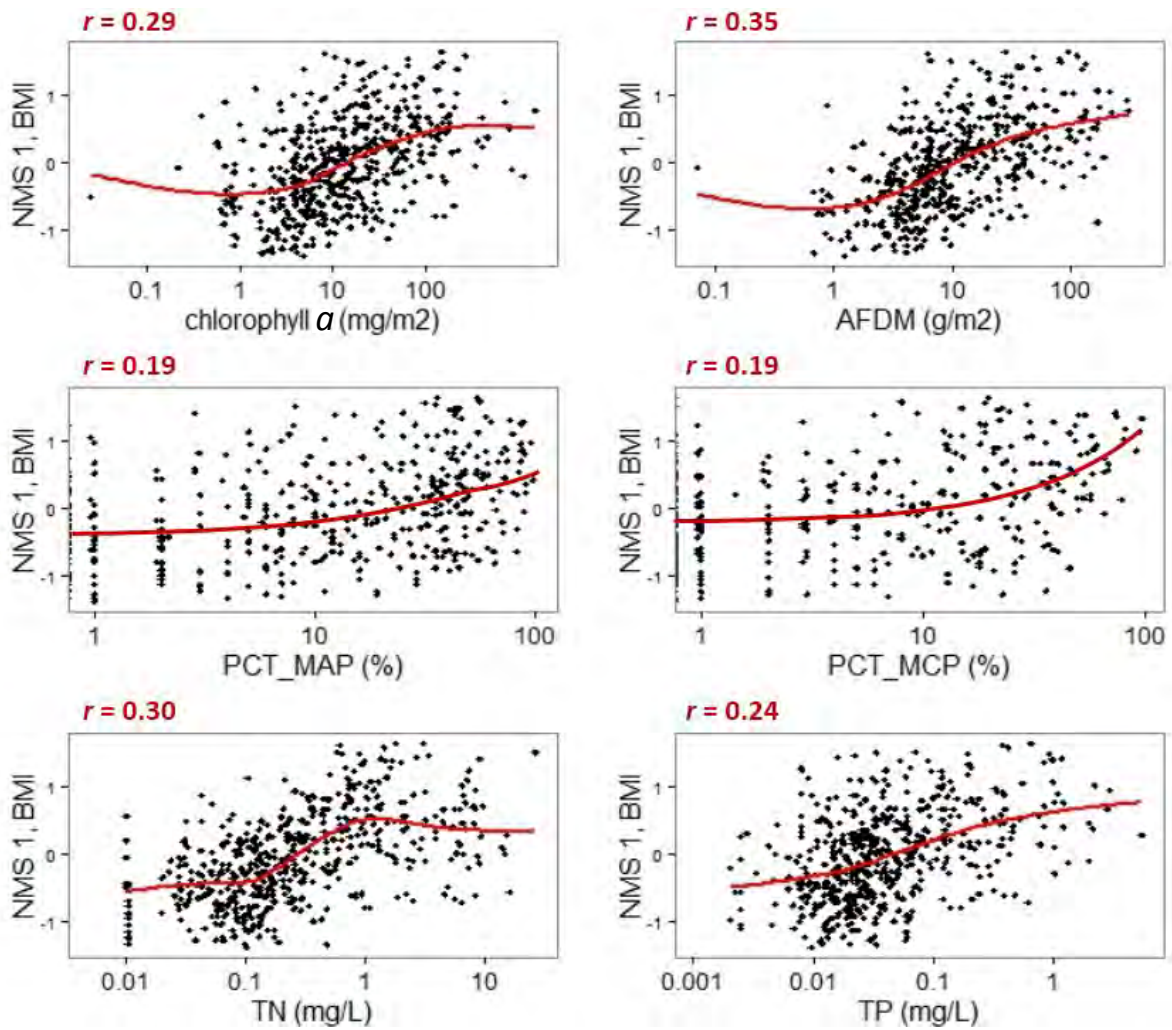


Figure 3.6. Scatterplots and splines for non-metric multidimensional scaling (NMS) axis 1 values from the benthic macroinvertebrate (BMI) community against biomass and selected cover and nutrient gradients on log scale, using the statewide data set. The Pearson correlation coefficient for each relationship is provided to the upper left of each graph. Correlation analyses were performed on rank-transformed data, and sample weights were used in the analyses. All relationships were highly statistically significant ($p < 0.0001$). PCT_MAP and PCT_MCP are percent cover of macroalgae and macrophytes, respectively.

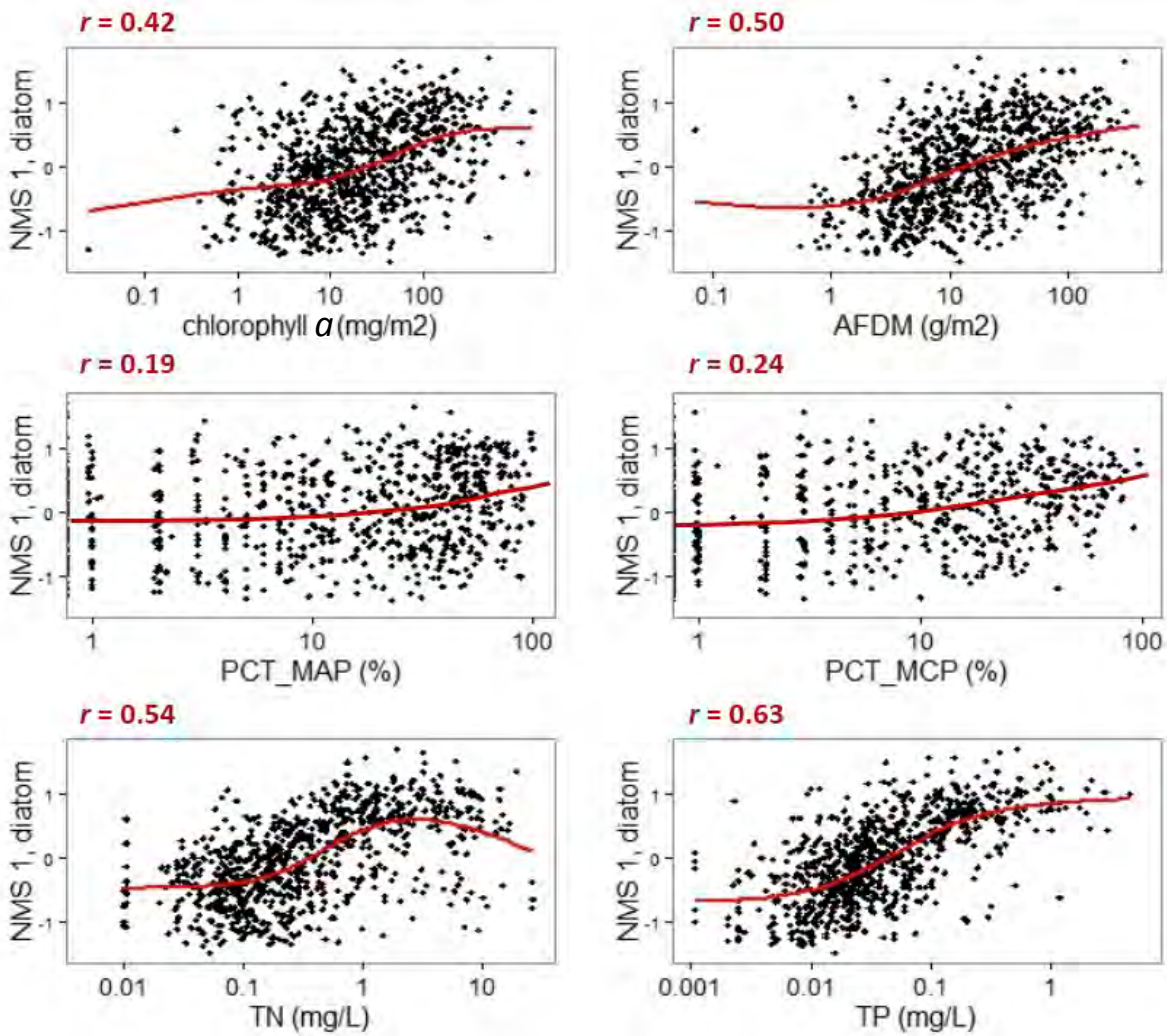


Figure 3.7. Scatterplots and splines for NMS axis 1 values from the diatom community against biomass and selected cover and nutrient gradients on log scale, using the statewide data set. The Pearson correlation coefficient for each relationship is provided to the upper left of each graph. Correlation analyses were performed on rank-transformed data, and sample weights were used in the analyses. All relationships were highly statistically significant ($p < 0.0001$). PCT_MAP and PCT_MCP are percent cover of macroalgae and macrophytes, respectively.

Table 3.5. Results of CART analyses with NMS axis 1 scores for either the BMI or the diatom community as the response variable. Separate analysis were run for the statewide dataset and for the South Coast ecoregion. Model runs included either the full set of explanatory variables (“ALL”, see Methods), or chlorophyll *a* or AFDM alone. Cut points are the median values, from 1,000 bootstrap runs, at which the first split in the indicated splitting variable was made during tree building. “Frequency” refers to the number of bootstrap replicates in which the variable in question was the splitting variable at the first node.

Splitting Variable	Cut Point (95% CI)	Assemblage	Region	Explanatory Variables in Model	Frequency	Model Fit
Chlorophyll <i>a</i> (mg/m ²)	6.2 (6.2 - 6.2)	BMI	statewide	ALL	1	0.21
	23.6 (4.1 - 61.9)	BMI	statewide	chlorophyll <i>a</i>	810	0.14
	21.9 (11.4 - 30.1)	BMI	South Coast	ALL	3	0.15
	30.8 (2.7 - 86.0)	BMI	South Coast	chlorophyll <i>a</i>	173	0.09
	23.6 (4.1 - 61.9)	diatom	statewide	chlorophyll <i>a</i>	810	0.14
	23.6 (4.1 - 61.9)	diatom	South Coast	chlorophyll <i>a</i>	810	0.14
AFDM (g/m ²)	12.6 (4.8 - 35.8)	BMI	statewide	ALL	634	0.22
	30.8 (4.1 - 88.7)	BMI	statewide	AFDM	239	0.10
	41.8 (6.0 - 159.3)	BMI	South Coast	ALL	59	0.18
	25.2 (4.0 - 75.0)	BMI	South Coast	AFDM	988	0.13
	25.9 (3.2 - 103.3)	diatom	statewide	ALL	50	0.24
	18.5 (3.6 - 54.1)	diatom	statewide	AFDM	840	0.15
	18.5 (3.6 - 54.1)	diatom	South Coast	AFDM	840	0.15
TN (mg/L)	0.29 (0.09 - 0.75)	BMI	statewide	ALL	179	0.22
	0.65 (0.22 - 1.8)	BMI	South Coast	ALL	42	0.19
	0.61 (0.12 - 2.2)	diatom	statewide	ALL	554	0.25
	0.60 (0.18 - 1.7)	diatom	South Coast	ALL	82	0.24
TP (mg/L)	0.058 (0.021 - 0.12)	BMI	statewide	ALL	9	0.19
	0.055 (0.017 - 0.12)	BMI	South Coast	ALL	435	0.21
	0.080 (0.012 - 0.25)	diatom	statewide	ALL	221	0.25
	0.070 (0.01 - 0.19)	diatom	South Coast	ALL	523	0.25
NH ₄ (mg/L)	0.013 (0.008 - 0.018)	BMI	statewide	ALL	2	0.19
	0.045 (0.005 - 0.18)	BMI	South Coast	ALL	11	0.18
SRP (mg/L)	0.080 (0.055 - 0.12)	BMI	South Coast	ALL	43	0.20
	0.074 (0.016 - 0.20)	diatom	statewide	ALL	21	0.24
	0.078 (0.012 - 0.15)	diatom	South Coast	ALL	304	0.25

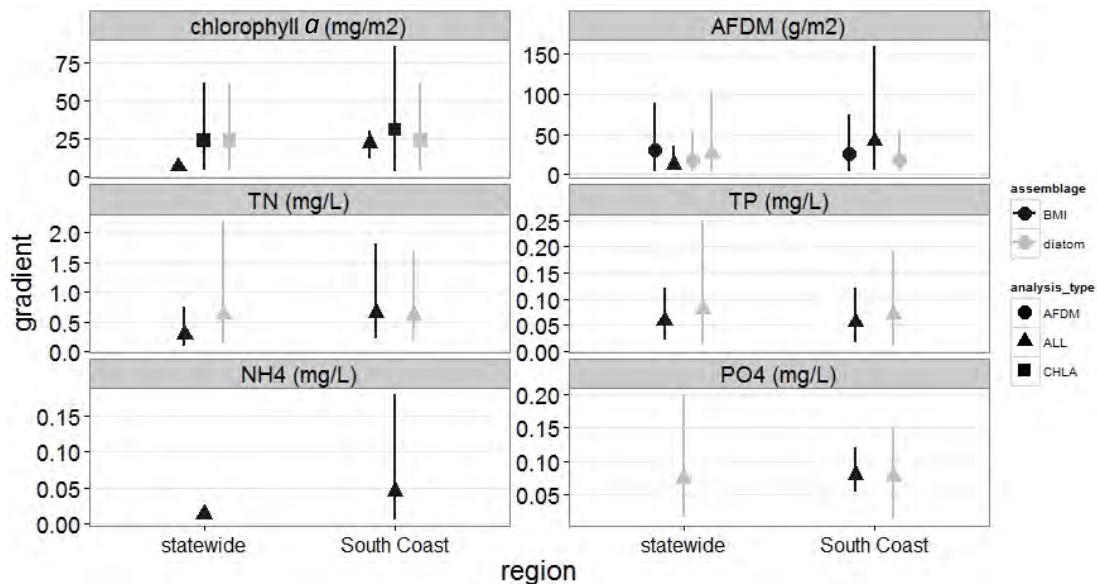


Figure 3.8. Cut points from CART analyses using NMS axis 1 scores from either the BMI or the diatom community as the response variable. Cut points are the median values, from 1,000 bootstrap runs, at which the first split in the indicated splitting variable was made during tree building. Error bars correspond to cut point 95% confidence intervals. Separate analyses were run for the statewide dataset and for the South Coast ecoregion. Model runs included either the full set of explanatory variables (“ALL”, see Methods), or chlorophyll *a* (“CHLA”) or AFDM alone. Y-axes correspond to the stressor gradients, which are labeled in the upper strip of each panel.

TITAN and nCPA

TITAN and nCPA were used on BMI and diatom community composition data in order to detect change points in biotic response along biomass and nutrient stressor gradients. Based on the nCPA results and results for the TITAN “decreaser” taxa: Chlorophyll *a* change points were always <27 mg/m², AFDM change points were always <13 g/m², TN change points were always <0.5 mg/L, and TP change points were always <0.09 mg/L.

Table 3.6 provides the mean change points (i.e., points along biomass/nutrient gradients where taxa show the greatest change in frequency and relative abundance, and which, therefore, can be interpreted as thresholds) derived from the nCPA analyses. Only results for the pure and reliable taxa from the TITAN analyses are included. Table C.1 provides TITAN change point values for individual taxa. TITAN change points for “increaser” taxa were invariably higher, sometimes substantially so, than those for “decreaser” taxa. Numbers of pure and reliable taxa were low for the percent cover ALLs relative to the other biomass/nutrient gradient types. For macroalgal percent cover (PCT_MAP), change points from nCPA and TITAN “increasers” were all <36%, and for macrophyte percent cover (PCT_MCP), change points were all <19%. Figures 3.9-3.11 show examples of TITAN and nCPA change points for BMI and diatom communities along several biomass and nutrient gradients, and Figure 3.12 provides a graphical summary of change points from all TITAN and nCPA analyses. TITAN analyses show a narrow range of response to AFDM and TP for sensitive (decreaser) taxa with relatively narrow confidence intervals. Appearance of tolerant (increasing) taxa was more gradual with much wider confidence intervals (Figure 3.10). Overall, BMI community composition showed a very sharp threshold of response along a gradient of TN, while responses along gradients of macroalgal and macrophyte cover were more diffuse (Figure 3.11).

Table 3.6. TITAN and nCPA results for BMI and diatom community composition data. Included are change points from TITAN (sum[z+] and sum[z-]) analyses and nCPA analyses based on Euclidean and Bray-Curtis distance measures (nCPA.euc and nCPA.bc). Also provided are quantiles (tau = 0.05 through 0.95) of each estimated change point distribution. The values provided for the TITAN analysis are mean values among only the "pure" and "reliable" taxa (see Methods for more details). "Tau = 0.95, max." is the tau = 0.95 value for the taxon (among the pure and reliable taxa) that had the highest tau = 0.95 value for the analysis in question.

Gradient	Analysis Type	Assemblage	# Taxa*	Change Point	tau =					
					0.05	0.1	0.5	0.9	0.95	(max)0.95
Chlorophyll <i>a</i> (mg/m ²)	TITAN.decreasers	BMI	92	16.51	4.45	6.02	13.28	22.83	27.74	95.17
		Diatom	72	16.06	4.54	6.19	14.13	26.92	32.26	108.96
	TITAN.increasers	BMI	23	43.40	7.90	10.01	37.75	84.58	108.32	580.45
		Diatom	61	83.98	18.78	27.31	69.01	188.87	233.19	747.25
	nCPA.euc	BMI	217	15.98	5.63	6.15	17.37	49.28	57.22	-
		Diatom	409	26.73	12.13	12.98	26.65	49.55	55.87	-
	nCPA.bc	BMI	217	17.15	10.28	11.03	17.15	20.99	22.87	-
		Diatom	409	26.73	13.88	15.21	28.39	50.73	56.84	-
AFDM (g/m ²)	TITAN.decreasers	BMI	90	7.05	3.27	3.80	6.67	10.85	13.04	52.91
		Diatom	65	10.19	4.26	5.10	9.42	16.48	18.72	66.52
	TITAN.increasers	BMI	34	23.06	6.67	7.45	20.78	55.71	71.87	185.37
		Diatom	100	73.01	14.74	20.96	60.52	130.20	153.92	304.33
	nCPA.euc	BMI	217	11.42	6.13	6.61	10.92	16.10	18.75	-
		Diatom	408	7.80	5.34	5.59	8.23	23.08	26.81	-
	nCPA.bc	BMI	217	10.86	6.25	6.61	10.91	12.52	12.93	-
		Diatom	408	12.71	5.98	6.80	11.61	21.87	24.70	-
PCT_MAP (%)	TITAN.decreasers	BMI	74	13.53	2.00	3.60	12.66	25.82	29.99	82.95
		diatom	47	11.63	1.99	2.62	10.80	23.55	28.34	73.78
	TITAN.increasers	BMI	24	25.96	8.16	11.35	27.57	56.10	60.86	89.26
		diatom	35	37.44	15.10	19.10	37.81	58.62	65.32	93.00
	nCPA.euc	BMI	203	35.00	12.00	15.00	33.00	60.05	70.74	-
		diatom	387	14.00	7.81	9.69	18.10	36.84	37.07	-
	nCPA.bc	BMI	203	19.78	11.00	16.00	23.00	39.00	41.37	-
		diatom	387	14.00	12.00	12.92	18.00	35.65	36.97	-
PCT_MAP (%)	TITAN.decreasers	BMI	52	5.37	0.29	0.66	5.20	15.72	18.97	59.02
		diatom	23	3.46	0.17	0.55	4.21	13.81	17.19	41.68
	TITAN.increasers	BMI	23	34.92	12.54	16.02	32.99	52.27	56.66	80.02
		diatom	80	29.89	6.80	10.02	28.83	54.54	60.60	82.00
	nCPA.euc	BMI	203	18.05	6.00	8.00	14.00	20.98	23.09	-
		diatom	387	7.00	2.00	2.86	6.26	18.07	26.13	-
	nCPA.bc	BMI	203	18.05	6.00	8.00	15.00	21.00	25.53	-
		diatom	387	7.00	2.00	2.86	6.92	14.57	26.68	-

* Number of taxa, for TITAN, is the number of pure and reliable taxa, not the total number of taxa evaluated in the analysis (the latter of which is the same number as that provided in the corresponding nCPA analysis).

Table 3.6 Continued.

Gradient	Analysis Type	Assemblage	# Taxa*	Change Point	tau =						
					0.05	0.1	0.5	0.9	0.95	(max) 0.95	
TN (mg/L)	TITAN.decreasers	BMI	117	0.20	0.08	0.09	0.18	0.34	0.39	1.33	
		diatom	96	0.29	0.07	0.09	0.25	0.53	0.63	4.12	
	TITAN.increasers	BMI	31	1.49	0.27	0.38	1.16	3.57	4.52	13.96	
		diatom	103	1.75	0.43	0.56	1.48	4.63	5.91	16.42	
	nCPA.euc	BMI	220	0.32	0.29	0.30	0.37	0.54	0.59	-	
		diatom	407	0.37	0.32	0.34	0.45	0.60	0.61	-	
	nCPA.bc	BMI	220	0.32	0.25	0.26	0.31	0.45	0.48	-	
		diatom	407	0.48	0.33	0.35	0.47	0.60	0.63	-	
	TP (mg/L)	TITAN.decreasers	BMI	103	0.04	0.01	0.01	0.03	0.06	0.08	0.46
			diatom	68	0.04	0.02	0.02	0.04	0.06	0.07	0.35
TITAN.increasers		BMI	21	0.11	0.04	0.05	0.09	0.41	0.61	2.07	
		diatom	98	0.18	0.05	0.06	0.15	0.45	0.56	2.07	
nCPA.euc		BMI	220	0.08	0.04	0.05	0.08	0.10	0.12	-	
		diatom	406	0.03	0.02	0.02	0.03	0.06	0.08	-	
nCPA.bc		BMI	220	0.08	0.04	0.05	0.08	0.10	0.11	-	
		diatom	406	0.05	0.02	0.03	0.05	0.09	0.10	-	

* Number of taxa, for TITAN, is the number of pure and reliable taxa, not the total number of taxa evaluated in the analysis.

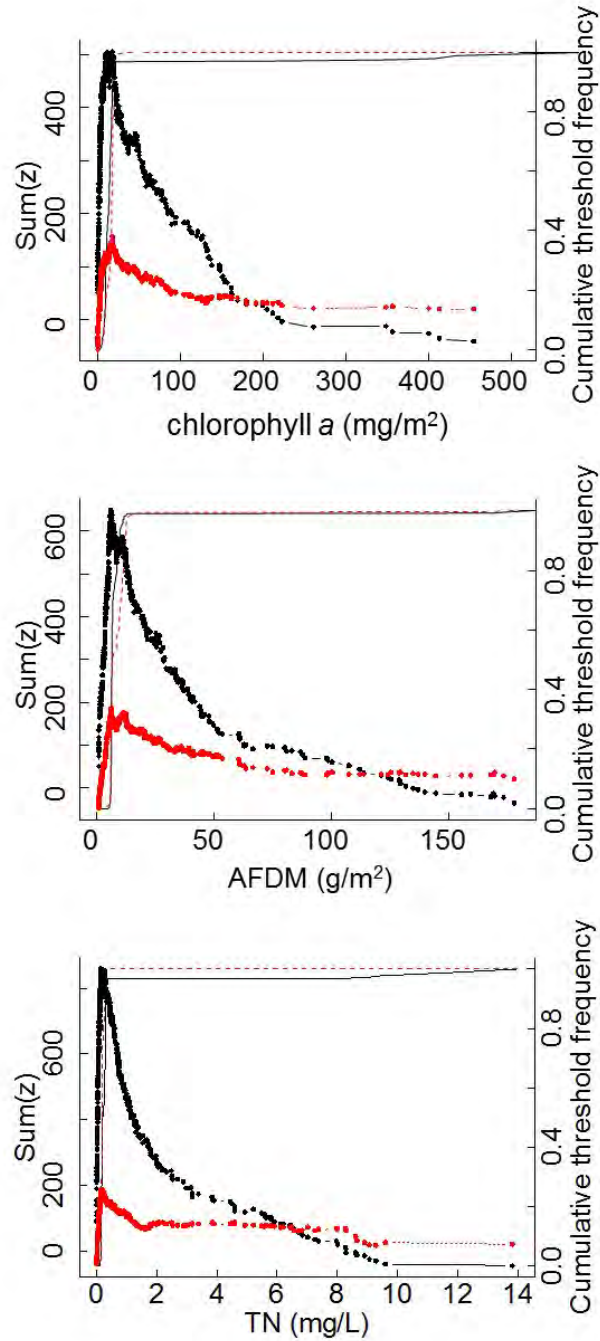


Figure 3.9. Plots of “sum(z)” scores (depicted as dots) from TITAN analysis of BMI community data along chlorophyll a, AFDM, and TN gradients, and the cumulative threshold frequency graphs (depicted as lines) for the sum(z) scores. Black dots (and solid lines) correspond to “decreaser” taxa, and red dots (and dotted lines) correspond to “increaser” taxa. Highest sum(z) scores across a gradient correspond to where the greatest change in component species’ relative abundances occurred. In contrast to Table 3.6, sum(z) across all taxa (not just “pure” and “reliable”) are represented in these plots.

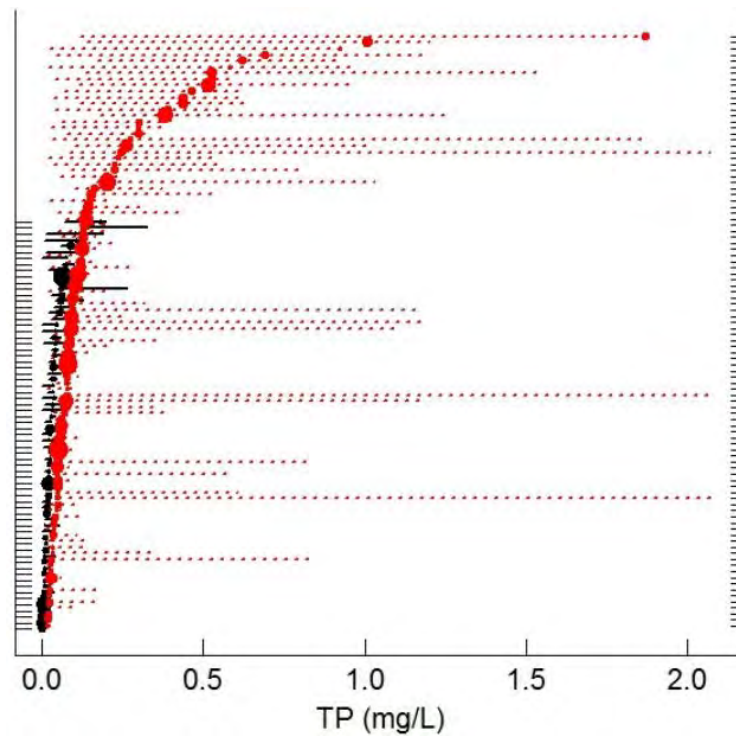
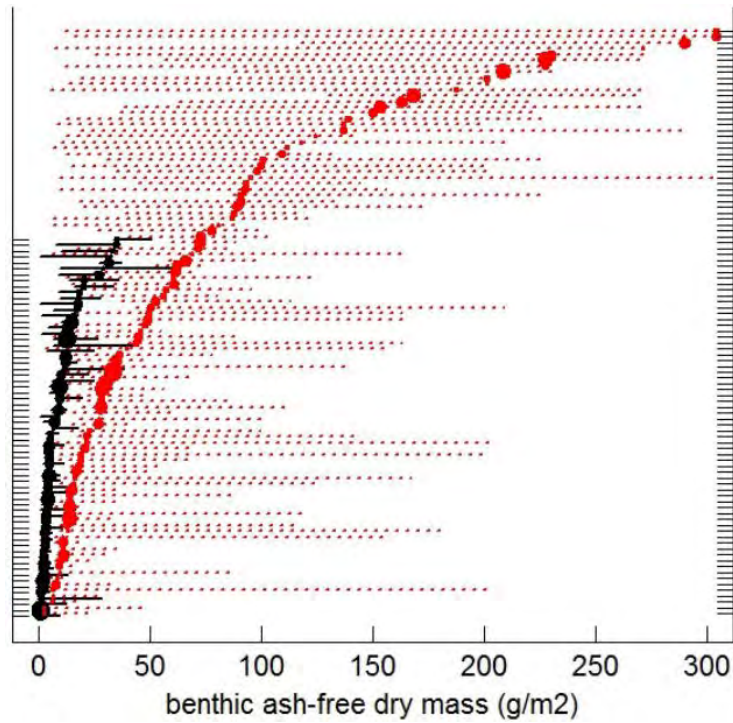


Figure 3.10. Plots of taxon-specific change from TITAN analysis of diatom community data along AFDM and TP gradients. Black plots refer to sum(z) scores for “decreaser” taxa, and red plots correspond to “increaser” taxa. Horizontal lines overlapping each symbol represent 5th and 95th percentiles from 500 bootstrap replicates. See Table C.1 for lists of the decreaser and increaser taxa (too numerous to label legibly here), their individual change point values, and the ranks thereof.

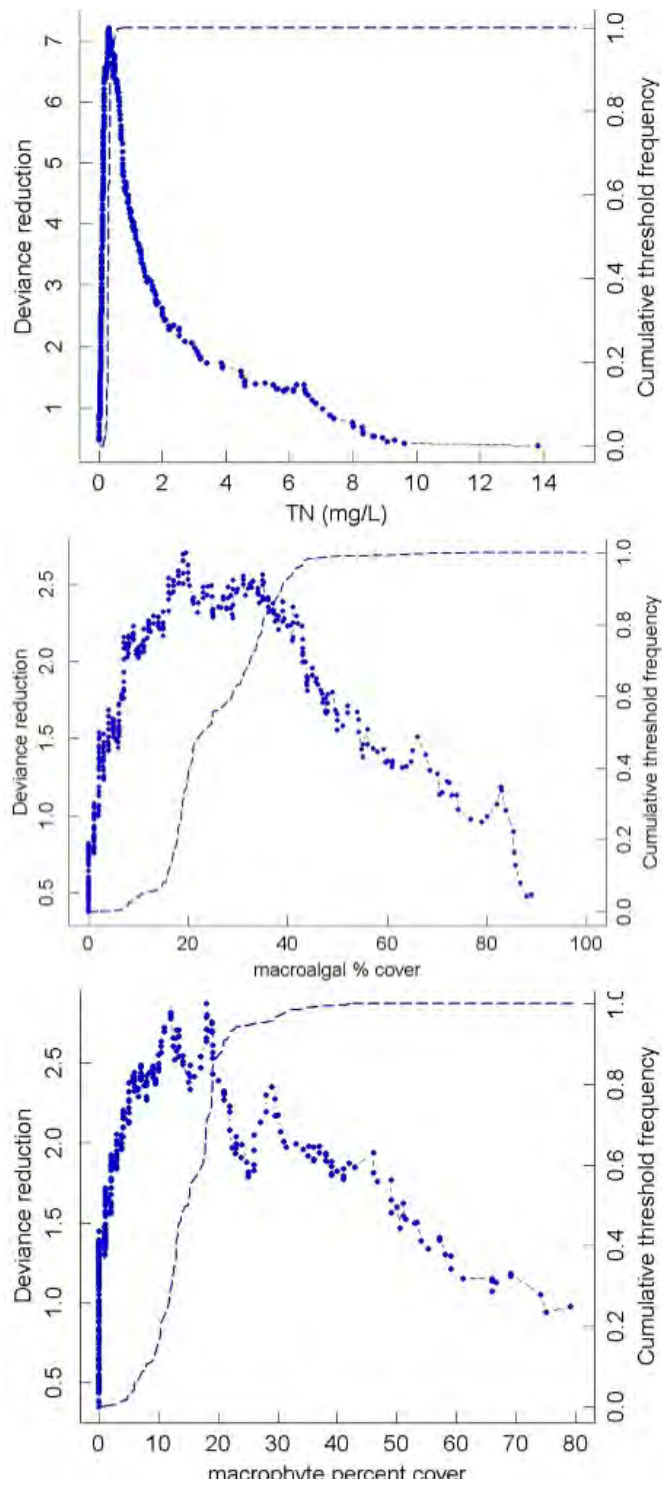


Figure 3.11. Nonparametric change point analysis (nCPA) results. Shown are deviance reduction values across TN, macroalgal percent cover (PCT_MAP), and macrophyte percent cover (PCT_MCP) for the BMI community (distance measure = Bray-Curtis).

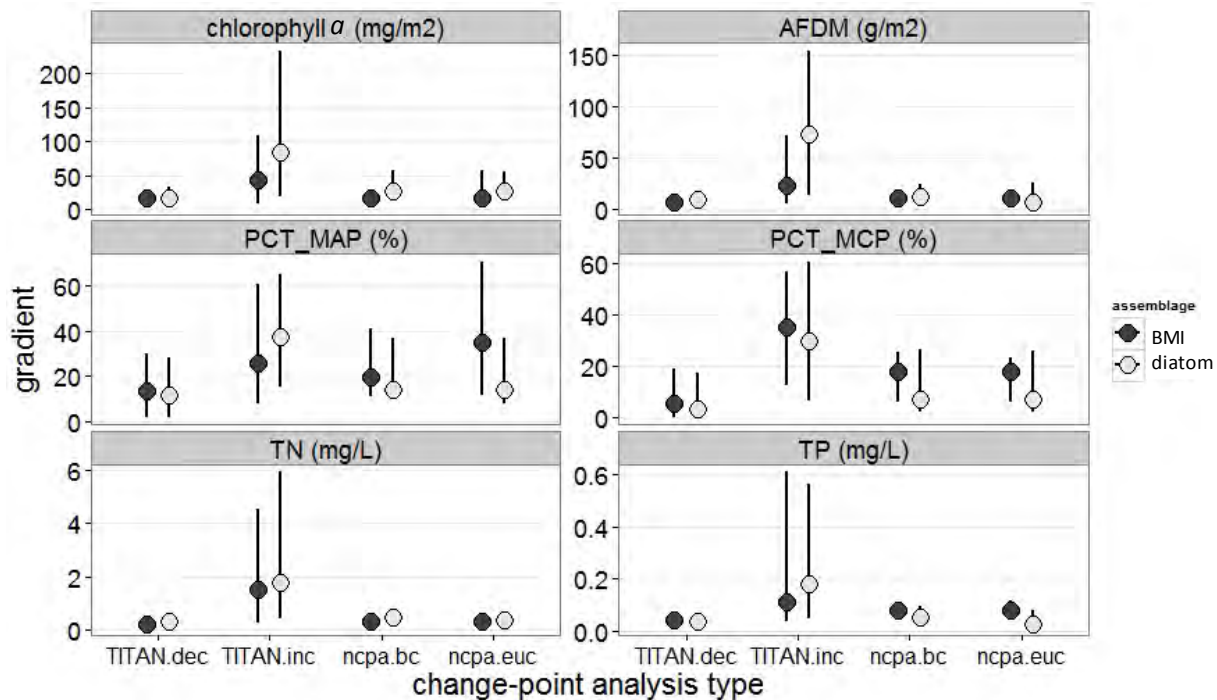


Figure 3.12. Summary of TITAN and nCPA change points along biomass/nutrient gradients, based on BMI and diatom community composition using the statewide dataset. Vertical lines associated with each change point represent 5th - 95th percentiles from 500 bootstrap replicates. TITAN values represent means among pure and reliable taxa. Y-axes correspond to the stressor gradients, which are labeled in the upper strip of each panel.

3.3.2 Biotic Responses to Biomass Gradients Based on Shifts in Integrative Measures of Community Composition (Metrics and Indices)

Piecewise regression and SiZer

Piecewise regression and SiZer are different approaches to evaluating relationships (and identifying potential thresholds or response) between biomass and nutrient stressor gradients and ALIs. We used a variety of ALIs that included both metrics and more integrative indices, such as IBIs, as opposed to the previous analyses that focused on “raw” community data. Chlorophyll *a* breakpoints, as estimated via piecewise regression, ranged from approximately 25 to 150 mg/m². For AFDM, over half of the ALI breakpoints were estimated (in the unweighted analyses) to be <20 g/m². Estimated TN breakpoints from unweighted piecewise regressions ranged from approximately 0.5 to 1.1 mg/L, while those for TP ranged from 0.075 to 0.12 mg/L. There was generally a high degree of correspondence between the piecewise regression output and the SiZer map for the various ALI/gradient combinations.

Table 3.7 provides a summary of the piecewise regression analysis output for all ALI/gradient combinations for which at least one member of each analysis pair (weighted/unweighted) passed all four “strict” evaluation criteria; Table C.2 is an extended version, providing the output for all ALI/gradient combinations. Breakpoint estimates arising from analyses not incorporating sample weights were almost invariably lower than those with weights (and CIs for the latter tended to be broader). Furthermore, analyses including weights were less likely than those without weights to result in output that successfully met all four screening criteria (even for the “relaxed” version).

For chlorophyll *a* breakpoints, over half of the ALIs have values of <100 mg/m² (among the analyses that did not incorporate sample weights). However, there was little agreement among ALI variables, and CIs were generally broad, especially for the ALIs with the higher breakpoints (Figure 3.13). Furthermore, breakpoints were generally not well supported, as few of them (N=10; Table C.2, Table 3.7) passed the “relaxed”, and none passed the “strict”, screening criteria. Nonetheless, of those that did achieve the “relaxed” criteria, all four assemblages were represented. Breakpoint values were still highly variable, ranging from 23 to 113 mg/m². In general, there was a high degree of interdigitation of breakpoints among assemblages (e.g., BMI ALIs were represented across the full range of values generated; Figure 3.13). Exhaustion thresholds occurred at lowest levels for SENS indicators, followed by EUTR indicators (decreased DO, increasing saprobicity, then increasing green algal biovolume and finally nuisance green algae), NUTR indicator taxa, and finally INT indicators. Sensitive and eutrophication indicators tended to have the smallest CI for chlorophyll *a* breakpoints as compared to nutrient and integrative indicators.

Table 3.7. Summary of piecewise regression results for all ALI response types for which at least one version of the analysis (weighted or unweighted) fulfilled all four “strict” criteria, as described in the Methods.

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
TN (mg/L)	D18	unweighted	0.88 (0.07), 0.26	-45.66 (-53.16 – -38.16)	0.38 (-0.22 – 0.97)	0.37	yes	yes
		weighted	1.29 (0.13), 0.50	-34.89 (-39.88 – -29.91)	0.25 (-1.54 – 2.04)	0.31	yes	no
	EPT_PercentTaxa	unweighted	0.68 (0.04), 0.16	-0.56 (-0.63 – -0.48)	0.00 (0.00 – 0.01)	0.59	yes	yes
		weighted	0.72 (0.06), 0.22	-0.55 (-0.62 – -0.47)	0.00 (-0.01 – 0.01)	0.46	yes	no
	EPT_Taxa	unweighted	0.63 (0.04), 0.14	-27.25 (-30.69 – -23.81)	0.01 (-0.22 – 0.23)	0.60	yes	yes
		weighted	0.62 (0.05), 0.21	-31.09 (-35.72 – -26.46)	-0.07 (-0.81 – 0.67)	0.41	yes	no
	H20	unweighted	1.06 (0.06), 0.25	-40.18 (-45.02 – -35.33)	0.29 (-0.22 – 0.80)	0.53	yes	yes
		weighted	1.29 (0.12), 0.46	-32.65 (-36.85 – -28.46)	-0.14 (-1.64 – 1.37)	0.39	yes	no
	H21	unweighted	0.68 (0.05), 0.18	-58.63 (-67.41 – -49.85)	-0.19 (-0.72 – 0.34)	0.46	yes	yes
		weighted	1.19 (0.12), 0.47	-35.13 (-40.14 – -30.12)	-0.33 (-2.02 – 1.37)	0.34	yes	no
	H23	unweighted	0.77 (0.04), 0.18	-56.32 (-63.40 – -49.25)	-0.18 (-0.69 – 0.33)	0.53	yes	yes
		weighted	1.21 (0.11), 0.45	-34.86 (-39.56 – -30.16)	-0.31 (-1.90 – 1.27)	0.36	yes	no
TN (mg/L)	Intolerant_Percent Taxa	unweighted	0.62 (0.04), 0.15	-0.57 (-0.65 – -0.50)	0.00 (-0.01 – 0.00)	0.57	yes	yes
		weighted	0.58 (0.05), 0.19	-0.65 (-0.75 – -0.55)	0.00 (-0.02 – 0.01)	0.41	yes	no

Table 3.7 Continued

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?	
	Intolerant_Taxa	unweighted	0.52 (0.03), 0.13	-25.78 (-29.45 – -22.10)	-0.06 (-0.25 – 0.14)	0.53	yes	yes	
		weighted	0.51 (0.04), 0.18	-31.35 (-36.28 – -26.43)	-0.13 (-0.81 – 0.55)	0.37	yes	no	
S2		unweighted	0.83 (0.06), 0.24	-52.74 (-60.75 – -44.72)	-0.80 (-1.43 – -0.17)	0.46	yes	yes	
		weighted	0.93 (0.14), 0.53	-34.70 (-43.39 – -26.01)	-1.38 (-3.23 – 0.46)	0.20	no	no	
	Taxonomic_Richness	unweighted	0.71 (0.05), 0.19	-31.26 (-35.72 – -26.80)	-0.07 (-0.40 – 0.27)	0.54	yes	yes	
		weighted	0.71 (0.07), 0.27	-33.33 (-38.93 – -27.72)	-0.11 (-1.08 – 0.86)	0.37	yes	no	
	Tolerant_Percent Taxa	unweighted	0.67 (0.05), 0.18	0.42 (0.35 – 0.48)	0.00 (-0.01 – 0.00)	0.50	yes	yes	
		weighted	0.73 (0.06), 0.23	0.41 (0.35 – 0.46)	0.00 (-0.01 – 0.01)	0.45	yes	no	
TP (mg/L)	D18	unweighted	0.12 (0.01), 0.03	-352.10 (-406.00 – -298.20)	-4.04 (-9.01 – 0.93)	0.41	yes	yes	
		weighted	0.14 (0.01), 0.05	-290.00 (-335.00 – -245.00)	-9.16 (-19.92 – 1.60)	0.35	yes	no	
	H20	unweighted	0.11 (0.01), 0.03	-369.00 (-420.80 – -317.20)	-3.79 (-7.96 – 0.38)	0.50	yes	yes	
		weighted	0.13 (0.01), 0.05	-275.10 (-315.90 – -234.30)	-6.31 (-15.21 – 2.59)	0.38	yes	no	
	H23	unweighted	0.11 (0.01), 0.03	-371.90 (-426.40 – -317.30)	-2.34 (-6.92 – 2.24)	0.46	yes	yes	
		weighted	0.14 (0.01), 0.05	-266.10 (-309.70 – -222.60)	-5.41 (-15.22 – 4.40)	0.34	yes	no	
	RAWlowP	unweighted	0.08 (0.01), 0.02	-6.78 (-7.97 – -5.60)	-0.03 (-0.10 – 0.03)	0.35	yes	yes	
		weighted	0.08 (0.01), 0.03	-6.02 (-7.24 – -4.81)	-0.03 (-0.20 – 0.13)	0.22	yes	no	
	SRP (mg/L)	H20	unweighted	0.13 (0.01), 0.04	-315.70 (-360.00 – -271.40)	-1.12 (-7.26 – 5.02)	0.37	yes	yes
			weighted	0.14 (0.01), 0.05	-284.30 (-330.80 – -237.80)	4.09 (-8.30 – 16.48)	0.29	yes	no

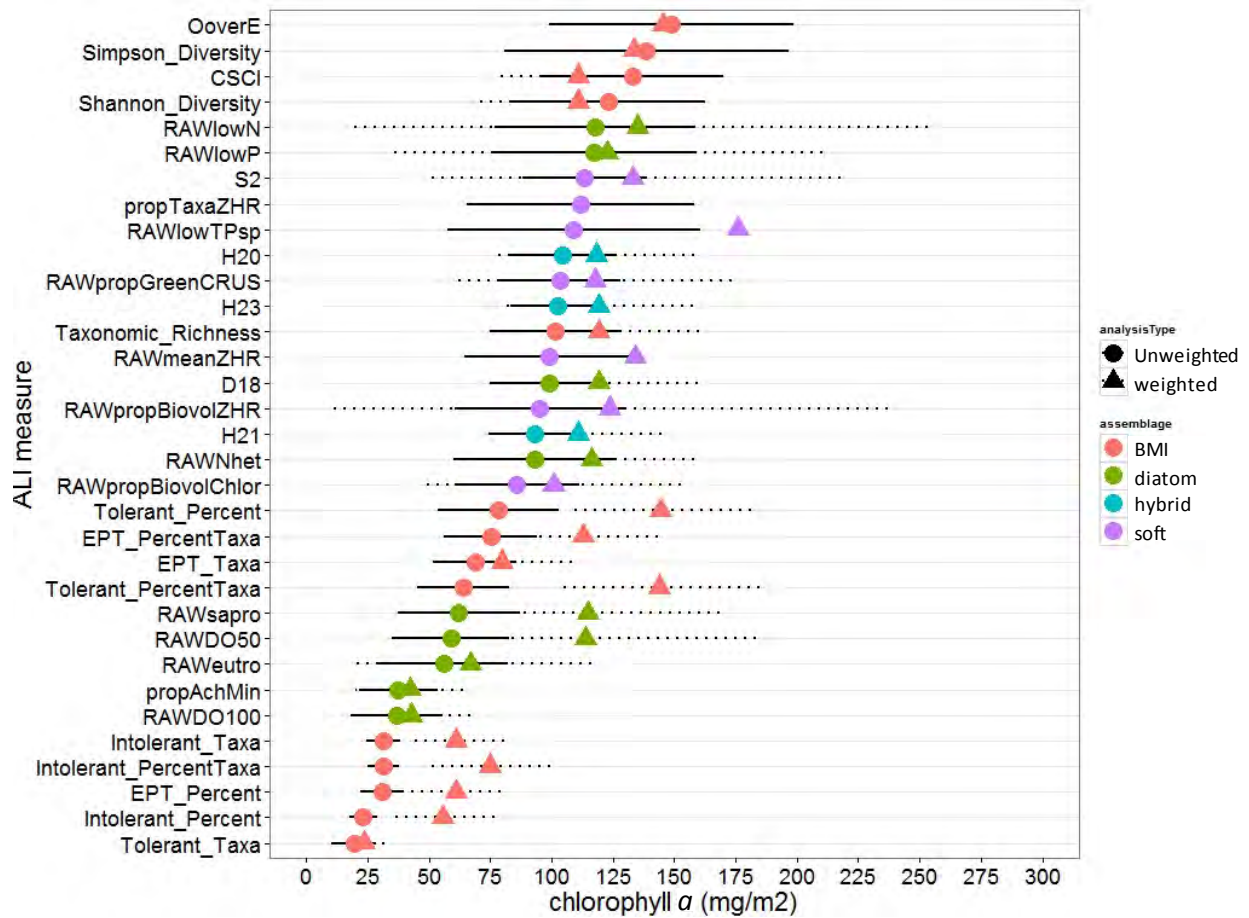


Figure 3.13. Breakpoints, with 95% confidence intervals, for the chlorophyll *a* gradient, from piecewise regressions using all available ALI data types. Triangles correspond to analyses using sample weights and circles correspond to unweighted. BMI ALI measures are in pink, diatom = green, hybrid = blue, and soft = purple. Solid lines are the 95% CI for unweighted analyses, and dashed are for weighted. Note that fewer than half of the ALI measures' piecewise regressions met the "relaxed" criteria for confidence in the breakpoint, as described in Methods, and none met the "strict" criteria. Details on analysis results are provided in Table 3.7 and Table C.2.

Breakpoints for AFDM (Figure 3.14) exhibited a higher degree of consensus among ALIs than was observed for chlorophyll *a*. In addition, a higher number of ALIs achieved the “relaxed” criteria (but still none achieved the “strict”). Among those ALIs achieving the “relaxed” criteria, all four assemblages were represented, and estimated AFDM breakpoint values occupied the relatively narrow range of 7 to 39 g/m²; Table C.3, Table 3.7). Again, exhaustion thresholds tended to be lower for SENS and EUTR indicators and greatest for INTI and NUTR indicators although ranges of mean breakpoints tended to be narrower than for chlorophyll *a*. Highest weighted mean breakpoints were associated with proportional biovolume in green algae, the soft algal IBI and proportion nuisance green algae. For the AFDM gradient, BMI breakpoints tended to occur at lower values than diatom breakpoints. Again, indicators of initial DO depletion and increasing saprobicity had exhaustion thresholds lower than those for filamentous greens and in the same range as those for SENS and INTI BMI indicators.

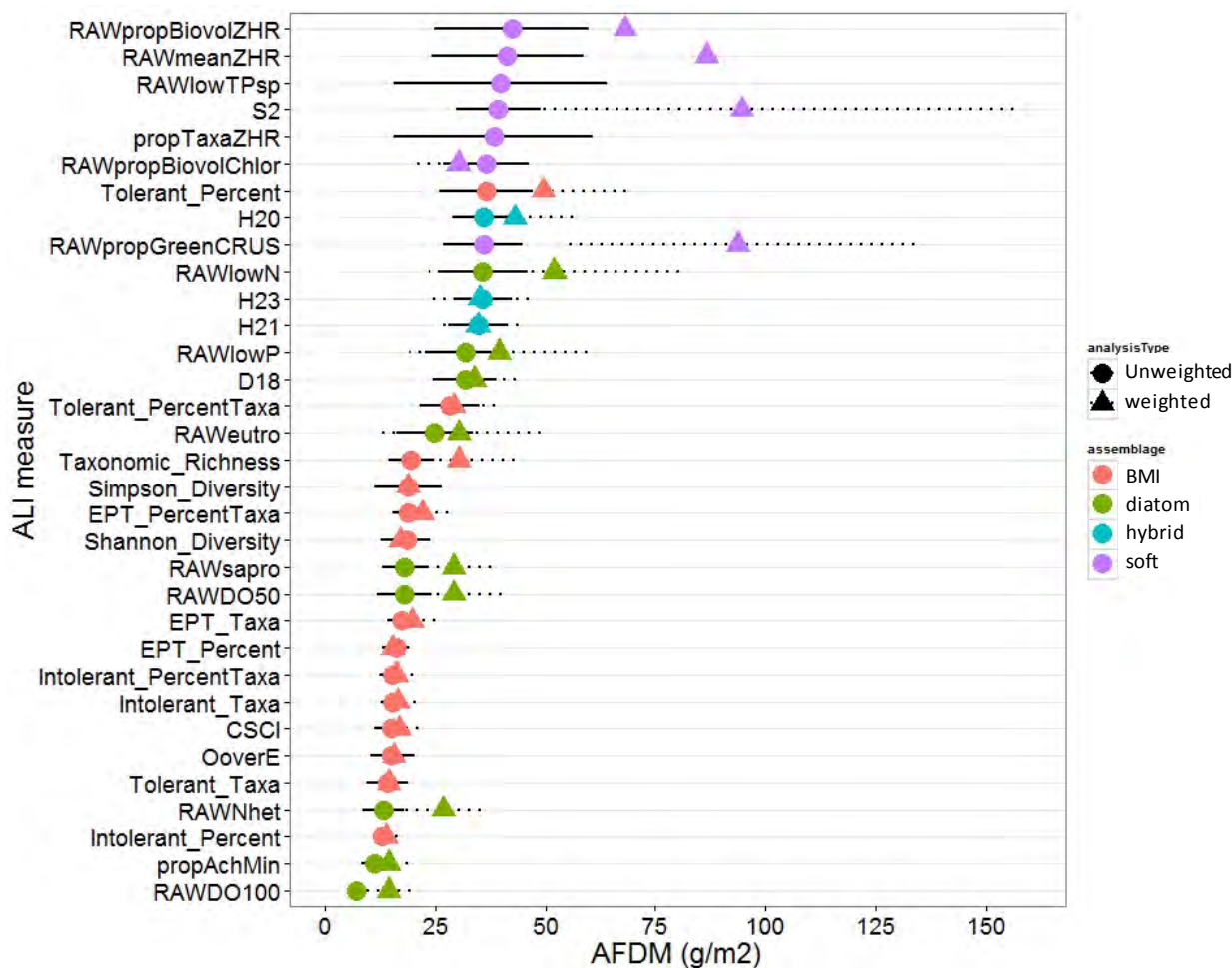


Figure 3.14. Breakpoints, with 95% confidence intervals, for the AFDM gradient, from piecewise regressions using all available ALI data types. Triangles correspond to analyses using sample weights and circles correspond to unweighted. BMI ALI measures are in pink, diatom = green, diatom+soft hybrid = blue, and soft algae = purple. Solid lines are the 95% CI for unweighted analyses, and dashed are for weighted. Note that fewer than half of the ALI measures’ piecewise regressions met the “relaxed” criteria for confidence in the breakpoint, as described in Methods, and none met the “strict” criteria. Details on analysis results are provided in Table 3.7 and Table C.2..

Results of unweighted piecewise regression analyses against the TN gradient achieved the “strict” criteria for eleven ALIs (Table 3.7). The ALIs represented all four assemblages, and their estimated TN breakpoints ranged from approximately 0.5 to 1.1 mg/L, lending support, via multiple lines of evidence, that a variety of instream ecological changes occur below 1.1 mg/L TN. Breakpoints were generally lower for the BMI assemblage relative to the algal assemblages (Figure 3.15), and for this assemblage, break points were very similar for the weighted and unweighted versions of the analysis.

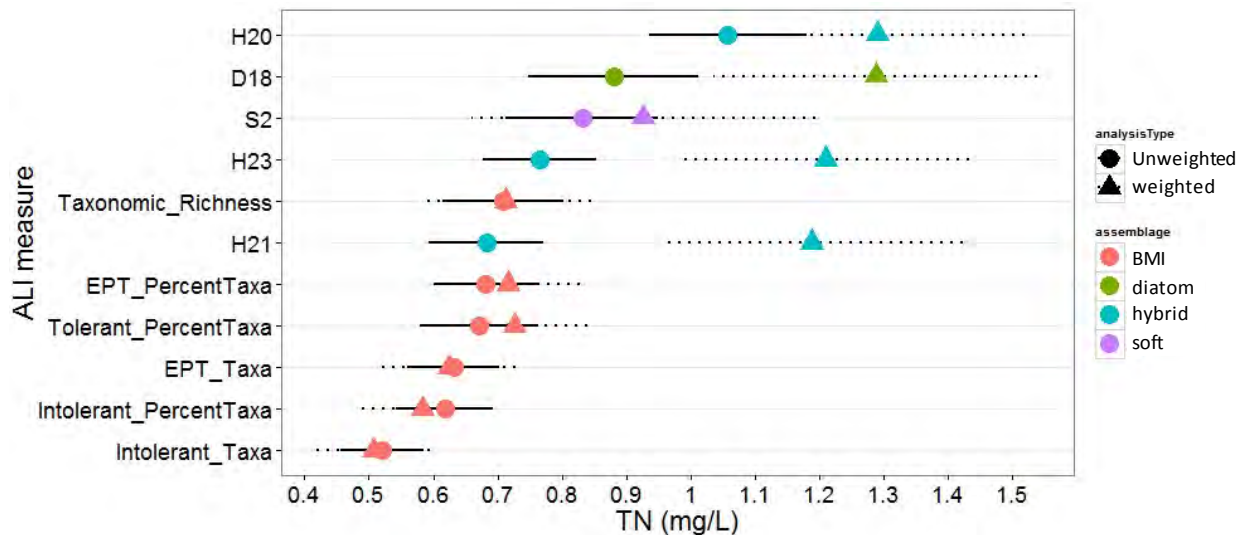


Figure 3.15. Breakpoints, with 95% confidence intervals, for the TN gradient, from piecewise regressions. This graph include only the ALI measures for which at least one analysis type (usually the unweighted version) fulfilled all four of the "strict" criteria for confidence in the breakpoint, as described in Methods. Triangles correspond to analyses using sample weights and circles correspond to unweighted. BMI ALI measures are in pink, diatom = green, hybrid = blue, and soft = purple. Solid lines are the 95% CI for unweighted analyses, and dashed are for weighted.

Results of unweighted piecewise regression analysis against the TP gradient achieved the “strict” criteria for four ALIs (Table 3.7). The ALIs represented only two of the assemblages (the diatoms and the hybrids, the latter of which include information about the diatom community), and their estimated TP breakpoints ranged from approximately 0.075 to 0.12 mg/L (Figure 3.16).

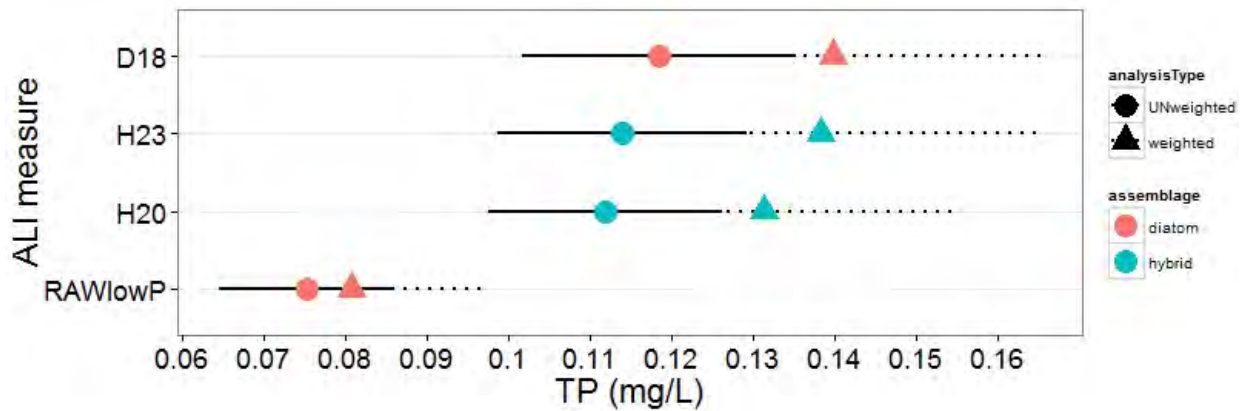


Figure 3.16. Breakpoints, with 95% confidence intervals, for the TP gradient, from piecewise regressions. This graph includes only the ALI measures for which at least one analysis type (usually the unweighted version) fulfilled all four of the "strict" criteria for confidence in the breakpoint, as described in Methods. Triangles correspond to analyses using sample weights and circles correspond to unweighted. Diatom ALI measures are in pink and hybrid = blue. Solid lines are the 95% CI for Unweighted analyses, and dashed are for weighted.

Whereas piecewise regression is focused on identifying the breakpoints (and associated uncertainty levels around them) in ALI response along a gradient, SiZer plays the complementary role of establishing if, where, and to what level of resolution, one or more slopes in the relationship between response variable and gradient are "real" and significant. Examples across a diverse array of ALI and gradient types are provided in Figures 3.17 through 3.21. In each of these cases, the mean ALI value decreased from the lowest to the highest biomass (or nutrient gradient) value, and the portion of the gradient where a downward slope in ALI value was most strongly supported by SiZer immediately preceded where the estimated breakpoint in slope occurred, as identified by piecewise regression. Thus the two methods, which are based on different approaches, were always in close agreement. As such, different lines of evidence supported essentially the same location for each threshold, thereby reducing the possibility that that location of any given estimated threshold was a mere artifact of the analytical method employed.

In general, for ALI/gradient relationships in which the first slope was particularly steep and the estimated breakpoint based on piecewise regression had a narrow CI, the SiZer map exhibited a correspondingly sharp/narrow downward red "peak", indicating high confidence, at fine as well as coarse resolution (i.e., bandwidth) leading up to that breakpoint. This is exemplified by relationships between ALIs and nutrient gradients, as shown in Figures 3.17 and 3.18. More gradual initial slopes in the piecewise regressions, and softer antecedent, downward "peaks" in the red portion of the SiZer maps were characteristic of the chlorophyll *a* and AFDM gradients (Figures 3.19 and 3.20). Figure 3.21 provides an example of a more weakly supported threshold for the ALI response variable, in this case, for the ALI H20 (a "hybrid" algae IBI) against the AFDM gradient. Here, the CI around the piecewise regression break point is broad, and the downward "peaks" in the red portion of the SiZer map are broad rather than sharp and localized over a narrow range in the gradient.

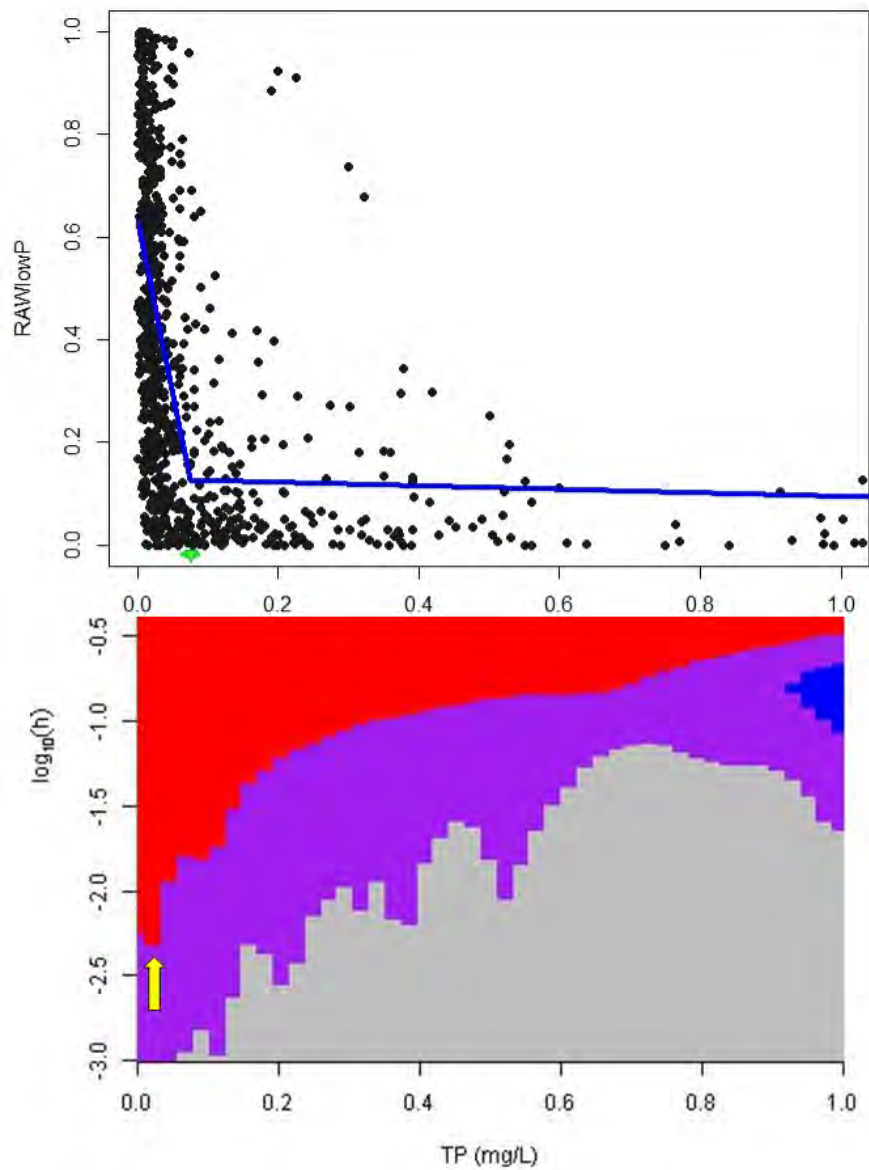


Figure 3.17. Piecewise regression plot of diatom ALI variable, RAWlowP on a TP gradient (top) and SiZer map from analysis of the same two variables (bottom). Note that x-axis was truncated to focus on the area of the break point. On the piecewise regression, the green line under data points along the x-axis is the 95% confidence interval¹⁴ for the breakpoint, and the green triangle is the breakpoint. On the SiZer map, the downward-extended (“downward peak”) red portion of the graph (yellow arrow at ~0.02 mg/L TP) indicates an area of well-supported, significantly negative slope that holds even at narrow bandwidths (see Methods for interpretation of color coding on a SiZer map, as well as definition of the y-axis).

¹⁴ Note that the CI on this graph is so narrow that it is barely discernible.

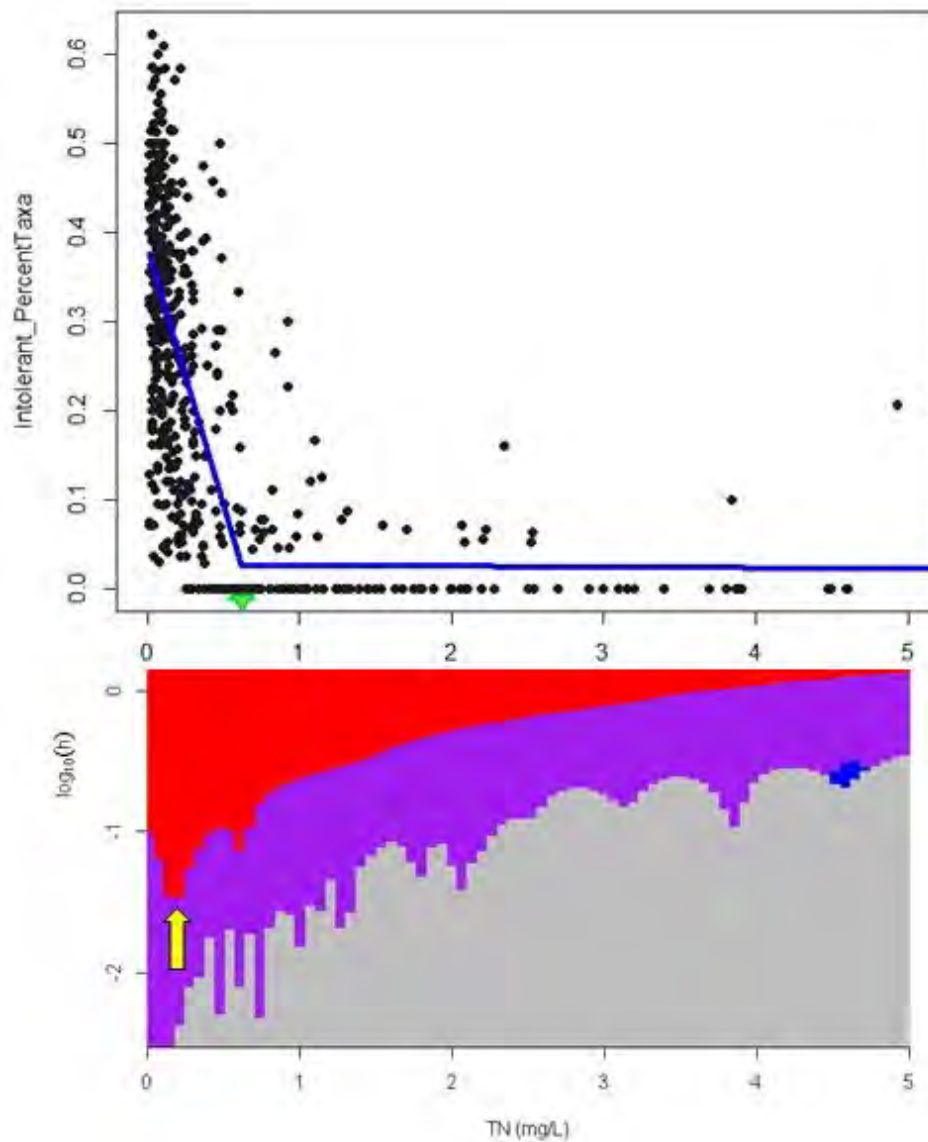


Figure 3.18. Piecewise regression plot of BMI ALI variable, Intolerant_PercentTaxa on a TN gradient (top) and SiZer map from analysis of the same two variables (bottom). Note that x-axis was truncated to focus on the area of the break point. On the piecewise regression, the green line under data points along the x-axis is the 95% confidence interval for the breakpoint, and the green triangle is the breakpoint. On the SiZer map, the downward-extended (“downward peak”) red portions of the graph indicate areas of well-supported, significantly negative slopes that hold even at narrow bandwidths (see Methods for interpretation of color coding on a SiZer map, as well as definition of the y-axis). The slope that holds for the broadest range of bandwidths is marked with a yellow arrow at ~0.14 mg/L TN, however note that there is a second slope further down the gradient, near 0.8.

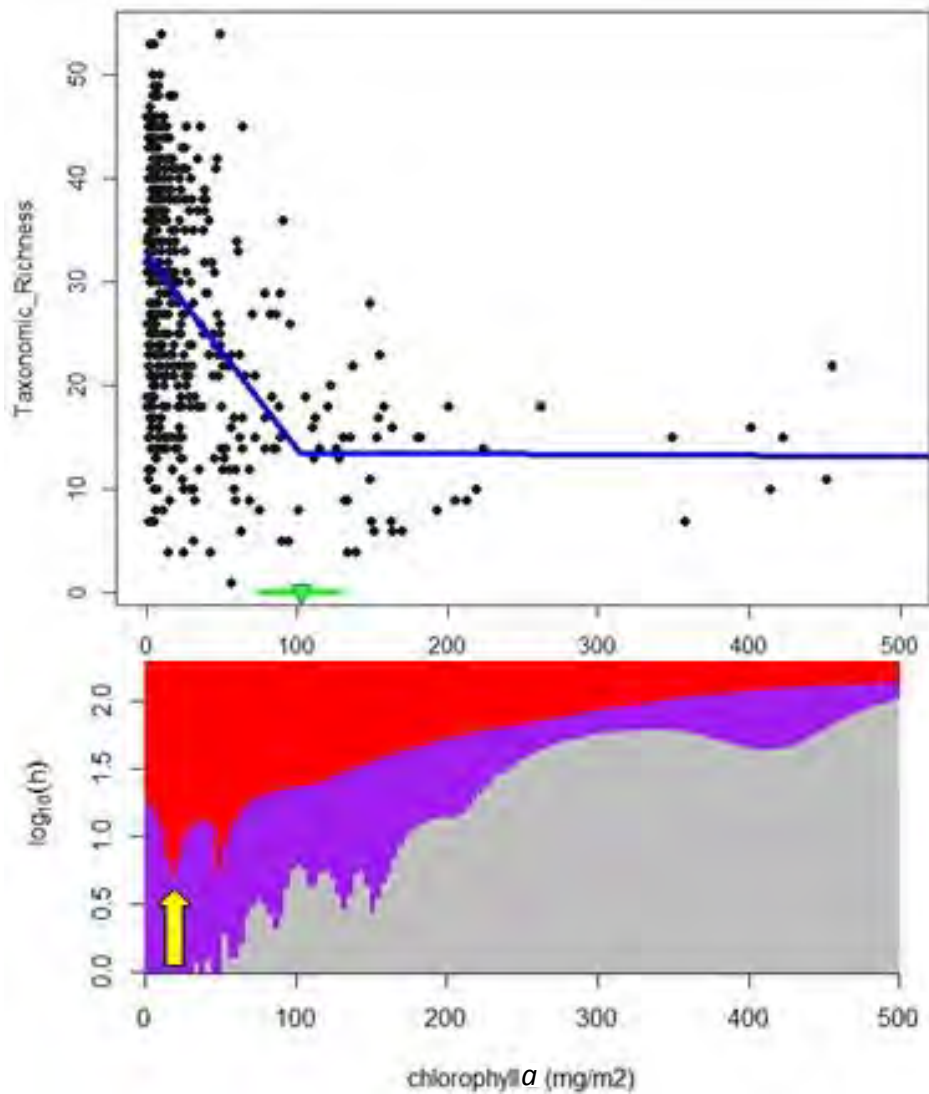


Figure 3.19. Piecewise regression plot of BMI ALI variable, Taxonomic_Richness on a chlorophyll a gradient (top) and SiZer map from analysis of the same two variables (bottom). Note that x-axis was truncated to focus on the area of the break point. On the piecewise regression, the green line under data points along the x-axis is the 95% confidence interval for the breakpoint, and the green triangle is the breakpoint. On the SiZer map, the downward-extended (“downward peak”) red portions of the graph indicate areas of well-supported, significantly negative slopes that hold even at narrow bandwidths (see Methods for interpretation of color coding on a SiZer map, as well as definition of the y-axis). The slope that holds for the broadest range of bandwidths is marked with a yellow arrow at ~15 mg/m² chlorophyll a, however note that there is a second slope further down the gradient, near 55.

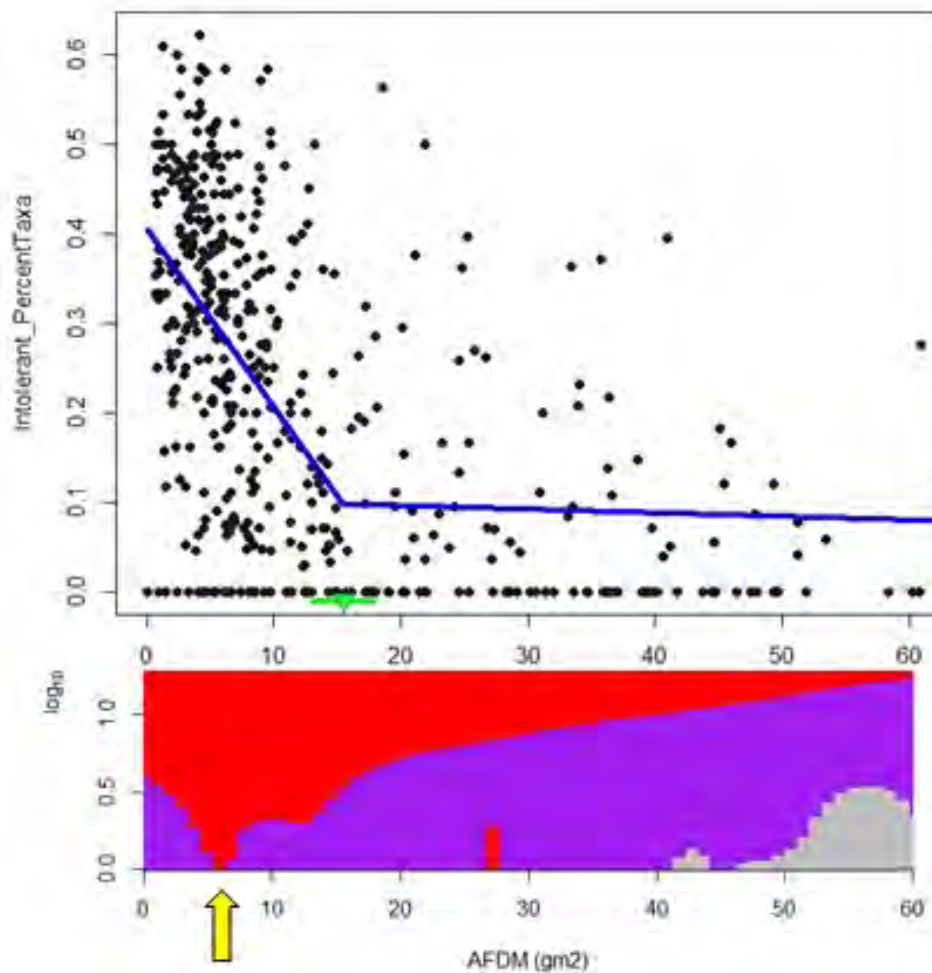


Figure 3.20. Piecewise regression plot of BMI ALI variable, Intolerant_PercentTaxa on an AFDM gradient (top) and SiZer map from analysis of the same two variables (bottom). Note that x-axis was truncated to focus on the area of the break point. On the piecewise regression, the green line under data points along the x-axis is the 95% confidence interval for the breakpoint, and the green triangle is the breakpoint. On the SiZer map, the downward-extended (“downward peak”) red portions of the graph indicate areas of well-supported, significantly negative slopes that hold even at narrow bandwidths (see Methods for interpretation of color coding on a SiZer map, as well as definition of the y-axis). The slope that holds for the broadest range of bandwidths is marked with a yellow arrow at ~6 g/m² AFDM, however note that there is a second slope further down the gradient, near 13.

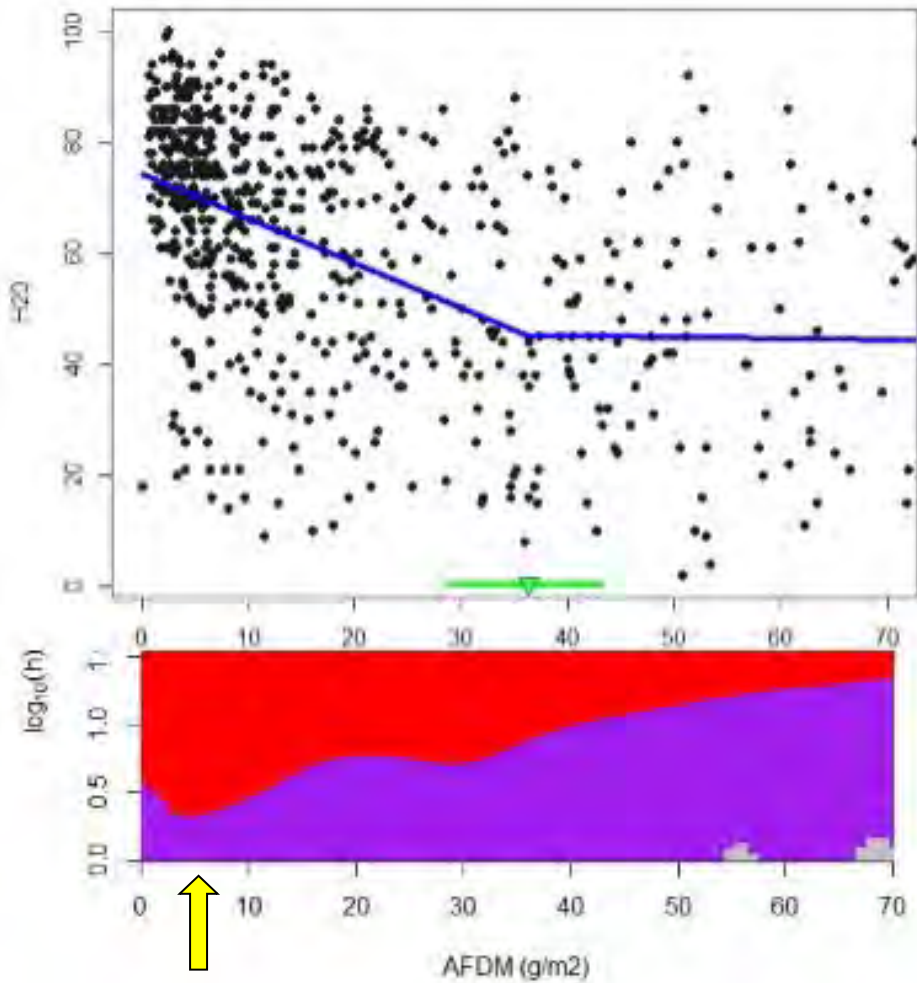


Figure 3.21. Closeup of piecewise regression plot of hybrid ALI variable, the IBI H20 on an AFDM gradient (top) and SiZer map from analysis of the same two variables (bottom). Note that x-axis was truncated to focus on the area of the breakpoint. On the piecewise regression, the green line under data points along the x-axis is the 95% confidence interval for the breakpoint, and the green triangle is the breakpoint. On the SiZer map, the downward-extended (“downward peak”) red portions of the graph indicate areas of significantly negative slopes that hold even at narrow bandwidths (see Methods for interpretation of color coding on a SiZer map, as well as definition of the y-axis). The slope that holds for the broadest range of bandwidths is marked with a yellow arrow at $\sim 4 \text{ g/m}^2$ AFDM, however note that there is a second slope further down the gradient, near 32.

The thresholds identified in Figures 3.17-3.21 all exhibited hallmarks of “exhaustion” thresholds (see Section 3.2.3). To illustrate this, we generated boxplots showing distributions of the ALI values for sites binned by gradient values (specifically, those falling below the identified threshold vs. those falling above; Figure 3.22). All five ALIs in this example were expected to decrease in value with increasing stress, and in each case the mean ALI value below the threshold exceeded that above the threshold. The reason the thresholds were interpreted as being of the “exhaustion” variety is that the distribution of values above the threshold exhibited a substantially narrower interquartile range (IQR) than those below the threshold. The response of H20 ALI along the AFDM gradient was an exception, in which IQRs were similar above and below the

threshold. However, even in this latter case, the range in AFDM values below the threshold was only 1/10 the range in values above the threshold within the project dataset (which extends to 450 g/m² — note that the x-axis in Figure 3.21 is truncated to allow easier viewing of the break point). As such, the ratio of IQR to range in gradient was substantially higher below the threshold than above it. In summary, the higher variability in ALI values below the threshold strongly supports the threshold as being “exhaustion” rather than “resistance”. This same pattern is evident across by far the majority of ALI-gradient combinations we examined (Table C.3). This result and others (see below) provides support that, except in the rare cases where noted, the thresholds we identified were exhaustion thresholds.

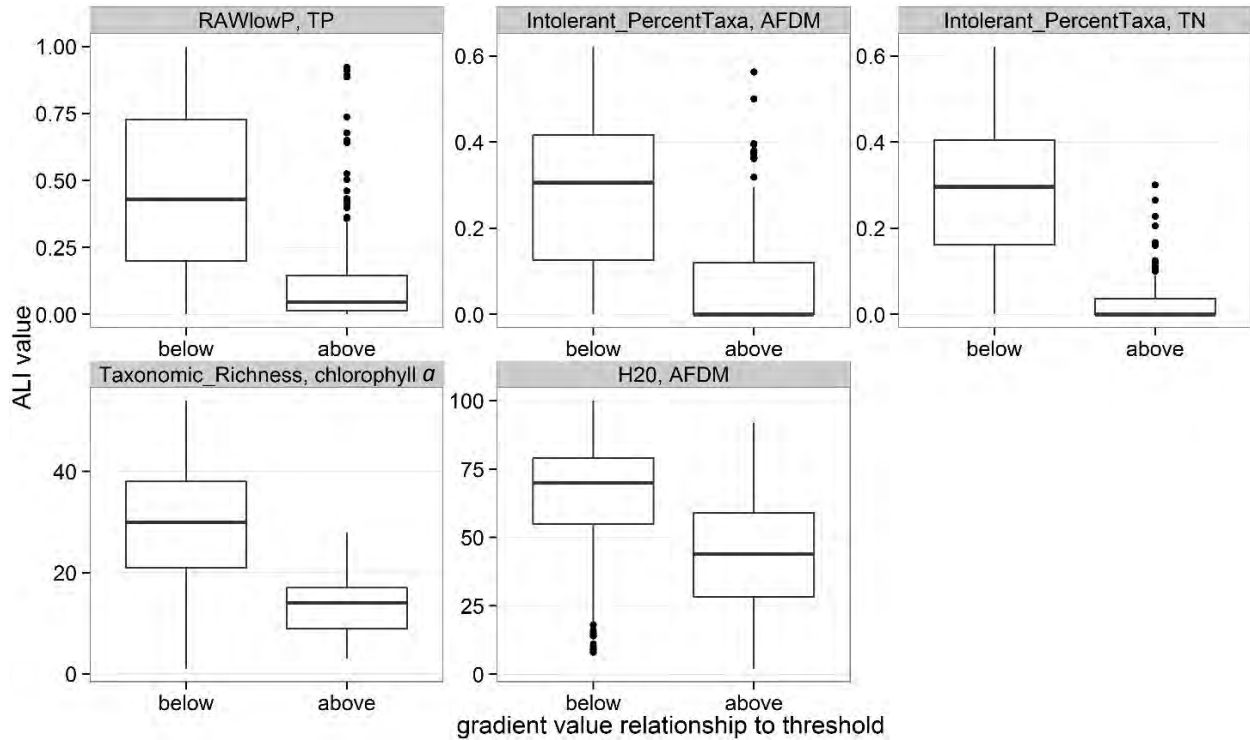


Figure 3.22. Distribution of ALI values among sites with stressor gradient (i.e., biomass or nutrient concentration) values below vs. above the threshold that had been determined based on piecewise regression. The strip above each panel in the plot indicates the type of ALI followed by the type of gradient in question. The ALI values are indicated by the y-axis.

3.3.3 Examining Relative Influence of Biomass, Nutrients, and Other Factors on Integrative ALI Measures

BRT and partial Mantel tests

BRT, a modeling approach, is the one type of analysis used in this study that allowed us to incorporate effects of other potential confounding factors (other stressor types as well as natural gradients) on the relationship between biomass/nutrients and ALI responses. Partial Mantel tests were used to determine whether important predictors of ALI response, based on BRT models, were statistically significant when other factors were controlled for.

For most ALIs, nutrients outranked biomass variables in terms of their relative influence in BRT models (Tables 3.8 and 3.9)¹⁵. The exceptions were for the soft-algae ALIs (in which for two of the four ALI types tested, biomass in the form of soft algal total biovolume was the biomass/nutrient predictor with the highest relative influence) and for the BMI metric EPT_Percent (for which AFDM ranked higher than any of the nutrients). In general, among biomass types, AFDM was the highest-ranked predictor for the greatest number of ALIs, followed by soft algal total biovolume and chlorophyll *a*. PCT_MAP and PCT_MCP were not top-ranked predictors for any of the ALIs examined.

The overall top-ranked predictor for most of the BMI ALIs was TN (Tables 3.8 and 3.9), and for diatoms, it was phosphorus (either as TP or SRP). TN was also a top predictor for one of the soft-algae ALIs (the index, S2), and NO_x was the top-ranked predictor for another soft-algae ALI (RAWmeanZHR). Summaries of the relative influence of all predictor variables, and specifically among the biomass and nutrient variables, are depicted graphically in Figures 3.23 and 3.24, respectively.

Results of the partial Mantel tests on the top-ranked predictors from the final BRT models are provided in Table 3.10. Most of the top-ranked predictors were found to be significantly correlated with their respective ALI response variables when the effects of the other top-ranked predictors, as well as spatial autocorrelation—in terms of geographic distance between sites, were controlled for. The latter generally did not have a significant effect on the ALI response variables (or the effect was relatively small, if significant), suggesting that predictor-response relationships observed in the BRT analyses were not merely artifacts of spatial autocorrelation.

¹⁵ Note that we re-ran two test cases for BRT with the input data transformed to improve normality and found that the results were nearly identical to those we provide in Table 3.8 (which was based on untransformed data), thus confirming that data transformation is not necessary for BRT analysis.

Table 3.8. Summary of boosted regression tree models of ALI variables, and relative influence (and rank) of biomass and nutrient predictors used in each. Boldface type corresponds to biomass or nutrient predictors that ranked highest within each model. Each model contained only one type of biomass predictor (as indicated by the column, “biomass type included in model”). Biomass type selected for each model was based on what biomass type ranked highest in an analogous model containing all five predictors (data not shown). “Model cv correlation (se)” refers to the cross-validation correlation coefficient (with standard error), indicating reliability of each model (Elith et al. 2008). Dashes indicate that the predictor in question was not part of the final BRT model for that ALI variable.

Assemblage	ALI type	Biomass type(s) included in model	Highest ranked predictor (relative influence)	Model cv correlation (se)	# Trees	# Predictors in final model	Relative influence of (rank)					
							biomass	TN	NO _x	NH ₄	TP	SRP
BMI (N = 611)	Intolerant_PercentTaxa	AFDM	TN (27.12)	0.932 (0.005)	5000	16	2.63 (9)	27.12 (1)	-	-	-	-
	Taxonomic_Richness	chlorophyll <i>a</i>	TN (30.96)	0.847 (0.009)	5050	36	1.59 (13)	30.96 (1)	0.27 (35)	0.47 (31)	3.07 (6)	2.24 (10)
	CSCI ¹⁶	chlorophyll <i>a</i>	URBAN_2000_5K (16.2)	0.829 (0.012)	5850	31	2.35 (12)	5.38 (6)	0.43 (31)	-	2.66 (11)	1.06 (22)
	Shannon_Diversity	soft algal total biovolume	TN (15.23)	0.727 (0.016)	4900	35	2.31 (12)	15.23 (1)	0.61 (35)	1.21 (27)	2.12 (13)	1.72 (19)
	EPT_Percent	AFDM	ecoregion (10.75)	0.717 (0.021)	5750	20	10.70 (2)	8.61 (3)	3.47 (15)	-	-	-
diatom (N = 888)	D18	AFDM	TP (20.53)	0.773 (0.015)	5400	31	3.71 (7)	5.38 (5)	0.80 (30)	0.94 (26)	20.53 (1)	4.89 (6)
	RAWeutro	AFDM	ecoregion (9.64)	0.664 (0.021)	6850	29	2.85 (10)	1.30 (28)	1.33 (27)	1.38 (26)	9.26 (3)	9.26 (2)
	RAWDO100	AFDM	SRP (11.31)	0.648 (0.025)	7050	27	3.64 (10)	1.73 (25)	2.61 (16)	-	8.42 (2)	11.31 (1)
	RAWNhet	chlorophyll <i>a</i>	conductivity (11.89)	0.641 (0.035)	5900	36	3.01 (11)	5.34 (6)	1.03 (32)	1.86 (21)	7.10 (3)	5.52 (5)
Hybrid algae (N = 809)	H20	AFDM	URBAN_2000_WS (24.12)	0.847 (0.009)	5950	35	2.57 (9)	18.47 (2)	1.52 (10)	0.77 (23)	12.37 (3)	2.92 (7)
soft algae (N = 845)	S2	soft algal total biovolume	TN (25.99)	0.781 (0.022)	5950	15	6.90 (5)	25.99 (1)	11.82 (3)	-	6.62 (6)	-
	RAWprop GreenCRUS	soft algal total biovolume	soft algal total biovolume (35.01)	0.727 (0.015)	5500	11	35.01 (1)	7.36 (5)	9.28 (3)	-	-	-
	RAWprop BiovolChlor	soft algal total biovolume	soft algal total biovolume (31.07)	0.658 (0.021)	4500	18	31.07 (1)	5.13 (6)	7.33 (3)	-	3.29 (10)	2.22 (16)
	RAWmean ZHR	soft algal total biovolume	NO _x (18.37)	0.624 (0.015)	4650	22	8.96 (3)	10.90 (2)	18.37 (1)	1.42 (20)	8.68 (4)	2.21 (14)

¹⁶ Note that the CSCI scoring tool was in draft form at the time of preparation of this report and is subject to change before being finalized.

Table 3.9. Relative influence of predictors from BRT models. The top-ranked predictor in each model is in bold. Dashes indicate that the predictor in question was not a part of the final model for the ALI measure in question.

Predictor Type	Predictor	ALI Measure													
		CSCI	EPT_Percent	Intolerant_PercentTaxa	Shannon_Diversity	Taxonomic_Richness	D18	RAWDO100	RAWeutro	RAWNhet	H2O	RAWmeanZHR	RAWprop BiovolChlor	RAWprop GreenCRUS	SZ
Biomass	soft algal biovolume	-	-	-	2.3	-	-	-	-	-	-	9.0	31.1	35.0	6.9
	AFDM	-	10.7	2.6	-	-	3.7	3.6	2.9	-	2.6	-	-	-	-
	Chlorophyll <i>a</i>	2.4	-	-	-	1.6	-	-	-	3.0	-	-	-	-	-
Nutrient	TN	5.4	8.6	27.1	15.2	31.0	5.4	1.7	1.3	5.3	18.5	10.9	5.1	7.4	26.0
	TP	2.7	-	-	2.1	3.1	20.5	8.4	9.3	7.1	12.4	8.7	3.3	-	6.6
	NO _x	0.4	3.5	-	0.6	0.3	0.8	2.6	1.3	1.0	1.5	18.4	7.3	9.3	11.8
	SRP	1.1	-	-	1.7	2.2	4.9	11.3	9.3	5.5	2.9	2.2	2.2	-	-
	NH ₄	-	-	-	1.2	0.5	0.9	-	1.4	1.9	0.8	1.4	-	-	-
Other	conductivity	7.3	4.5	3.6	9.8	5.9	14.4	4.1	6.2	11.9	7.0	2.9	3.8	7.2	2.3
	ecoregion	2.9	10.8	22.4	7.3	9.0	2.5	5.8	9.6	0.7	1.3	6.1	2.6	4.0	4.0
	URBAN_2000_5K	16.2	5.4	9.5	11.2	13.6	-	-	-	-	-	-	-	-	-
	slope, reach	12.1	8.4	5.4	2.4	1.2	2.0	2.5	2.3	2.2	1.1	3.5	5.0	-	3.9
	elevation	1.4	-	4.0	0.7	0.4	1.9	3.3	4.8	1.5	3.8	3.0	9.3	10.4	7.3
	CODE_21_2000_5K	-	-	1.1	0.9	0.6	2.5	2.1	2.1	1.1	3.2	7.1	6.0	9.0	15.0
	canopy cover (%)	2.7	-	2.2	2.0	1.9	1.5	7.0	7.5	5.9	0.8	2.3	2.5	4.9	3.3
	URBAN_2000_WS	-	-	-	-	-	9.1	-	-	5.2	24.1	-	-	-	-
	discharge	3.6	6.7	1.2	1.9	1.0	0.7	2.5	1.8	1.1	0.5	4.0	5.4	5.3	2.4
	stream temperature	2.0	3.5	12.0	2.5	0.8	1.1	3.9	4.8	2.2	0.7	2.1	-	-	-
	URBAN_2000_1K	6.6	4.3	1.5	1.6	2.7	2.0	2.6	2.1	7.6	1.5	-	-	-	-
	alkalinity	2.3	2.7	2.9	1.9	0.8	1.2	7.3	6.4	1.9	1.2	2.2	-	-	-
	sands & fines (%)	0.6	3.0	1.4	2.7	2.3	6.6	3.4	2.6	3.4	2.7	-	-	-	-
pH	1.0	3.4	-	1.2	0.8	1.1	1.8	1.7	2.5	0.7	4.4	3.2	3.2	2.5	

Table 3.9 Continued.

Predictor Type	Predictor	ALI Measure													
		CSCI	EPT_Percent	Intolerant_PercentTaxa	Shannon_Diversity	Taxonomic_Richness	D18	RAWDO100	RAWeutro	RAWNhet	H2O	RAWmean-ZHR	RAWprop BiovolChlor	RAWprop-GreenCRUS	S2
	longitude	1.3	-	-	3.5	5.0	1.3	1.6	1.5	1.2	0.4	3.7	3.1	-	2.8
	mean monthly max temp (3-mo span)	1.0	-	-	2.4	1.2	0.9	2.7	2.4	1.0	0.6	1.3	2.3	4.3	2.9
	watershed area	2.2	3.5	-	0.9	0.6	1.0	2.2	2.2	1.1	1.1	2.1	3.6	-	2.1
	CODE_21_2000_WS	-	3.4	1.8	1.3	1.4	2.1	2.9	2.3	1.2	1.1	-	-	-	-
	stream depth	1.1	-	-	5.6	2.0	1.8	-	2.0	3.5	0.7	-	-	-	-
	coarse particulate organic matter (%)	0.7	3.9	-	0.7	0.5	0.9	4.8	2.6	1.9	0.9	-	-	-	-
	site disturbance class	11.3	-	-	-	0.2	1.4	-	-	2.2	1.3	-	-	-	-
	turbidity	1.2	-	-	1.7	0.8	2.2	2.0	2.2	1.8	1.4	1.8	-	-	-
	sedimentary geology (%)	0.9	3.6	-	0.9	0.6	2.1	1.8	1.3	2.4	1.4	-	-	-	-
Other	latitude	1.4	-	-	4.3	2.9	-	2.1	1.5	0.5	0.5	1.6	-	-	-
	total precipitation (3-mo span)	0.8	4.0	-	1.2	0.6	1.0	-	-	4.0	0.5	-	2.2	-	-
	finest (%)	3.6	3.7	1.2	0.6	0.8	0.9	-	-	0.4	-	-	2.1	-	-
	stream width	-	-	-	0.7	0.4	-	4.3	3.0	1.3	0.4	-	-	-	-
	mean monthly solar radiation (3-mo span)	0.9	-	-	1.8	0.8	-	1.7	1.6	0.9	0.6	1.4	-	-	-
	W1_HALL (riparian disturbance index)	2.1	-	-	1.6	0.8	1.8	-	-	2.1	1.0	-	-	-	-
	mean monthly % cloud cover (3-mo span)	-	2.6	-	1.8	0.4	-	-	-	1.1	0.5	-	-	-	-
	Ag_2000_WS	1.0	-	-	1.7	1.3	-	-	-	-	-	-	-	-	-
	Ag_2000_1K	-	-	-	-	-	-	-	-	2.3	0.6	-	-	-	-

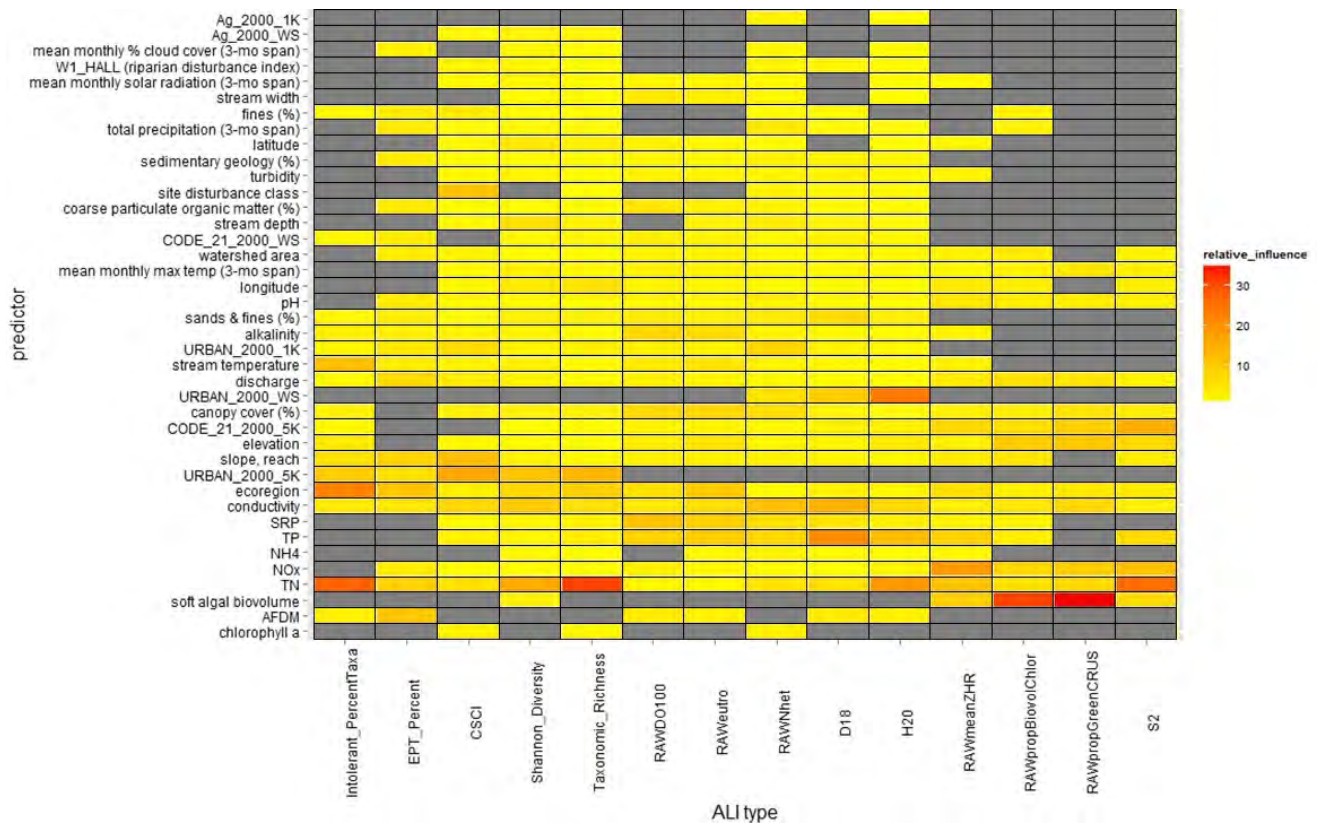


Figure 3.23. Heat map showing relative influence (%) of predictor variables (biomass, nutrients, and environmental co-factors) on ALI response variables, from 14 independent BRT models. Yellow = low influence, red = high. The biomass and nutrient predictor variables are grouped at the bottom of the graph. All climate variables are based on data for the month in which the sample in question was collected, averaged with the prior two months. Grey boxes indicate that the corresponding predictor type was not included in the final BRT model for that ALI type.

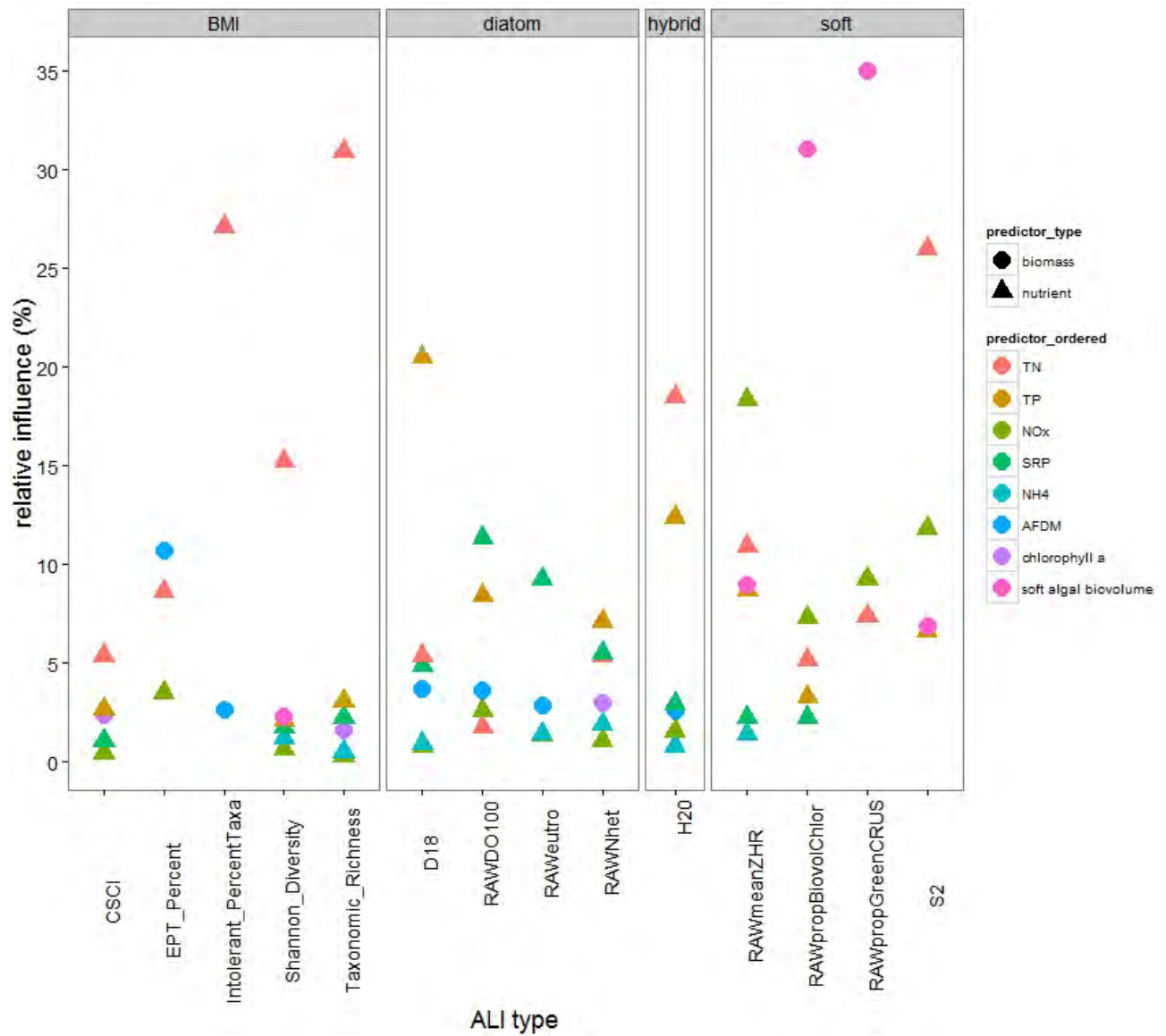


Figure 3.24. Summary of the relative influence of biomass and nutrient predictors on ALIs, from the BRT models. This is the same information as that presented in Table 3.9, but in graphical form, and focusing only on biomass and nutrients, to facilitate comparison of magnitudes of influence for the two types of predictors.

Table 3.10. Partial Mantel coefficients (95% CIs) for correlation between biomass/nutrient predictors and ALI variables and p-values. Grey boxes correspond to explanatory variables that were not included in the partial Mantel test for the ALI variable in question. “Space” refers to the geographic distance between sites (for testing the significance of spatial autocorrelation). “NS” = not significant; dashes correspond to predictors that were included as explanatory variables in the partial Mantel tests for the indicated ALI variables, but (because they did not fall under the categories of biomass, nutrients, or “space”) were not the focal variable in the tests. Values in bold correspond to significant partial Mantel tests.

Predictor	CSCI	EPT_Percent	Intolerant_Percent-Taxa	Shannon_Diversity	Taxonomic_Richness	D18	RAWDO-100	RAWeutro	RAWNhet	RAW/mean Z HR	RAWprop-Biovol-Chlor	RAWprop Green-Biovol-CRUS	S2	H20
chlorophyll <i>a</i>	0.06 (0.04–0.08) 0.005				0.04 (0.02–0.06) 0.022				0.07 (0.04–0.09) 0.002					
AFDM		0.18 (0.16–0.20) 0.001	0.15 (0.13–0.17) 0.001			0.12 (0.11–0.14) 0.001	0.01 (0.01–0.03) NS	0.01 (-0.001–0.022); NS						0.11 (0.09–0.13) 0.001
soft algal bio-volume				0.00 (-0.01–0.01) NS						0.01 (0.00–0.01); NS	0.210 (0.194 – 0.224); 0.001	0.18 (0.17–0.20) 0.001	0.05 (0.04–0.061) 0.001	
TN	0.17 (0.14–0.18) 0.001	0.24 (0.22–0.26) 0.001	0.34 (0.32–0.36) 0.001	0.15 (0.12–0.17) 0.001	0.26 (0.24–0.28) 0.001	0.07 (0.05–0.09) 0.001			0.06 (0.03–0.08) 0.006	0.06 (0.04–0.07) <0.001		0.07 (0.06–0.08) 0.001	0.10 (0.09–0.12) 0.001	0.17 (0.15–0.19) 0.001
NO _x										0.12 (0.11– 0.14) <0.001	0.14 (0.12–0.15) 0.001	0.11 (0.09–0.12) 0.001	0.19 (0.18–0.21) 0.001	
TP	0.07 (0.04–0.09) 0.003			-0.05 (-0.07– -0.03) NS	0.04 (0.03–0.06) 0.004	0.16 (0.14–0.18) 0.001	0.08 (0.06–0.09) 0.001	0.09 (0.07–0.10) 0.001	0.10 (0.08–0.12) 0.001	0.06 (0.05–0.07) <0.001			0.08 (0.07–0.09) 0.001	0.16 (0.14–0.18) 0.001
SRP					-0.05 (-0.06– -0.03) NS	0.01 (-0.01–0.03) NS	0.04 (0.03–0.05) 0.023	0.037 (0.021 – 0.051); 0.013	-0.03 (-0.05– -0.01) NS					0.03 (0.01–0.05) 0.049
space	-0.04 (-0.05– 0.03) NS	0.03 (0.01–0.04) 0.012	0.07 (0.06–0.09) 0.001	-0.02 (-0.03– -0.01) NS	-0.02 (-0.03– -0.01) NS	0.02 (0.01–0.03) 0.026	-0.002 (-0.02– -0.01) NS	0.004 (-0.01–0.01) NS	-0.03 (-0.04– -0.02) NS	0.04 (0.03–0.05); <0.001	0.01 (0.0–0.02) NS	-0.03 (-0.04– 0.02) NS	-0.01 (-0.02–0.01) NS	-0.03 (-0.04– 0.02) NS
conductivity	-													

Table 3.10. (continued)

Predictor	CSCI	EPT_Percent	Intolerant_Percent-Taxa	Shannon_Diversity	Taxonomic_Richness	D18	RAWDO-100	RAWeutro	RAWNhet	RAWmean Z HR	RAWprop-Biovol-Chlor	RAWprop Green-Biovol-CRUS	S2	H2O
canopy cover (%)	-				-		-	-	-					
stream temperature			-	-			-	-						
alkalinity			-				-	-						
stream width							-	-						
CPOM (%)							-							
elevation			-					-			-	-	-	-
URBAN_2000_1K	-				-				-					
sand & fines (%)				-	-	-			-					-
URBAN_2000_WS						-			-					-
stream depth				-	-				-					
total precipitation (3-mo span)									-					
URBAN_2000_5K	-		-	-	-									
slope, reach	-		-	-										
discharge	-													
fines (%)	-													
CODE_21_2000_5K												-	-	-
latitude				-	-									
longitude				-	-									
mean monthly max temp (3-mo span)				-										

Chlorophyll *a* had low relative influence in all three BRT models in which it was included as a predictor (Tables 3.8 and 3.9), and changes in slope in the corresponding partial dependence plots (Figure 3.25) were gradual along the chlorophyll *a* gradient, making it difficult to discern clear thresholds. However, at least for the BMI ALIs, a possible threshold was weakly evident in the vicinity of 60 mg/m², and for all three ALIs, the partial dependence plots leveled off by 100 mg/m², suggesting that any thresholds of effect of chlorophyll *a* on those ALIs is <100 mg/m².

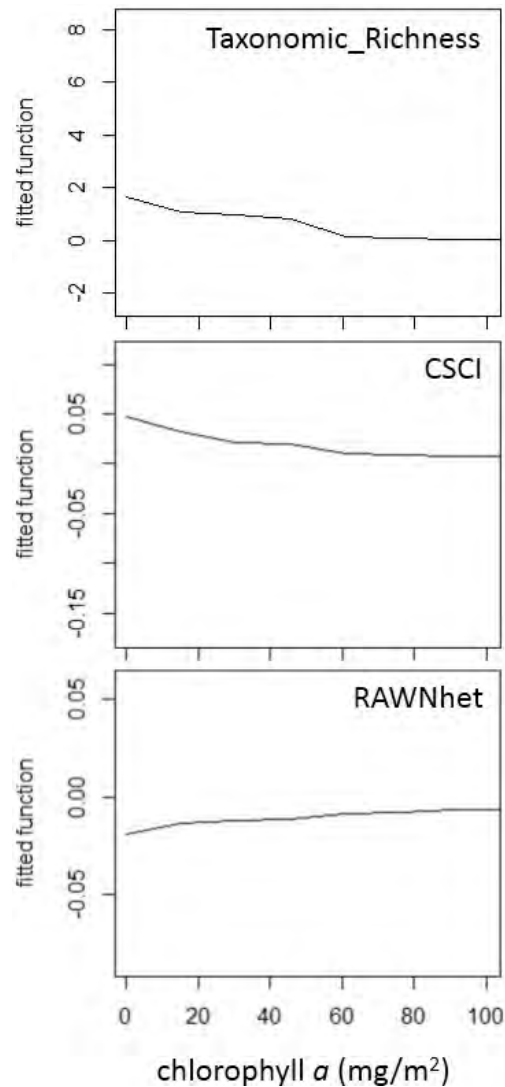


Figure 3.25. Partial dependence plots of chlorophyll *a* from BRT models predicting three ALI response types: CSCI, Taxonomic Richness, and RAWNhet. Y-axes correspond to the standardized, fitted ALI variables. Graphs do not show entire gradient length, but are cut off at the point beyond which there are no further changes in slope.

AFDM was included as a predictor in the BRT models for six ALIs spanning the BMI and diatom assemblages as well as the diatom/soft algae (“hybrid”) IBI, H20. Examples of partial dependence plots from these models are provided in Figure 3.26. AFDM was the predictor with the second highest relative influence on the BMI ALI, EPT_Percent, among a total of 20 predictors in the final model (Tables 3.8 and 3.9). There was a precipitous drop in the fitted value for EPT_Percent along the AFDM gradient until around 25 g/m², beyond which no further decline was evident. A roughly similar pattern, albeit less pronounced, was realized for the BRT model with the diatom ALI, RAWDO100, as the response variable (however, AFDM was not a significant explanatory variable for this ALI in the partial Mantel test; Table 3.10). The IBIs H20 and D18 exhibited similar initial breakpoints of approximately 35 g/m² in their partial dependence plots; beyond that point, the fitted response variables continued to decline (although much more shallowly; Figure 3.26). Only at approximately 180 g/m² AFDM was no further decline evident for any of the ALIs.

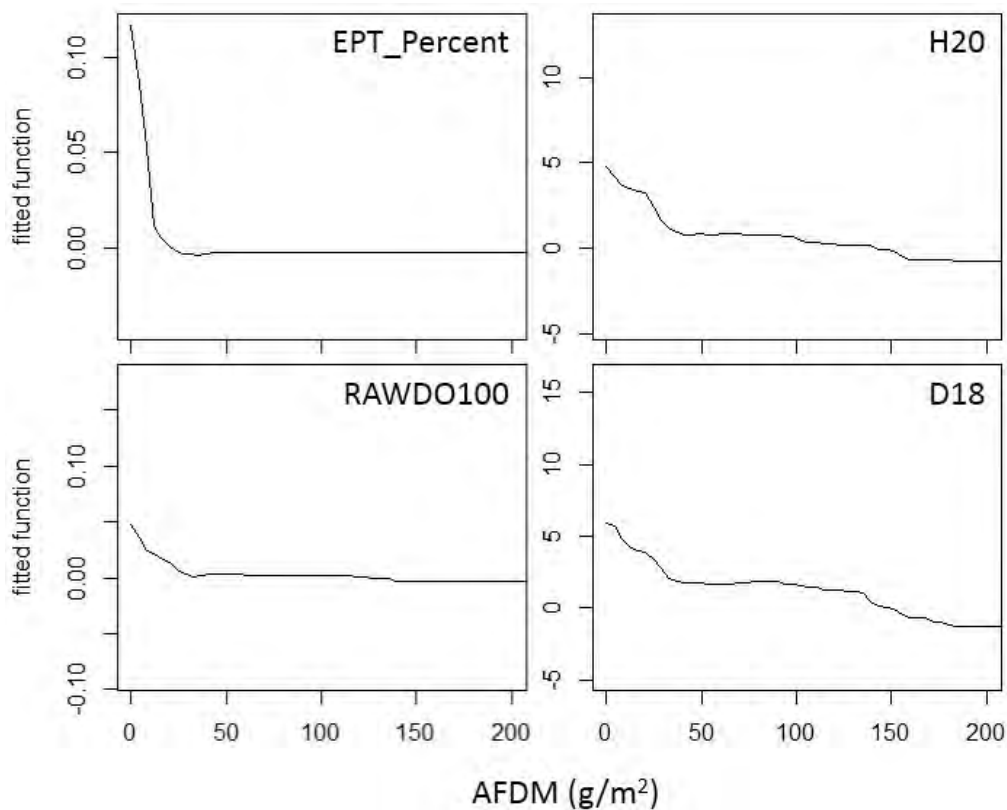


Figure 3.26. Partial dependence plots of AFDM from BRT models predicting four ALI response types: the metrics EPT_Percent and RAWDO100; and the IBIs H20 and D18. Y-axes correspond to the standardized, fitted ALI variables. Graphs do not show entire gradient length, but are cut off at the point beyond which there are no further changes in slope.

Whereas biomass was rarely a top-ranked predictor among the BRT models for the fourteen ALIs, the opposite was true for nutrients (Tables 3.8 and 3.9). Nitrogen in one form or another was the top-ranked predictor for three BMI ALIs and two soft-algae ALIs, and phosphorus in one form or another was the top-ranked predictor for two diatom ALIs. Partial dependence plots of TN from BRT models for three ALIs are provided in Figure 3.28. Three breakpoints were observed in the plots, depending upon the ALI in question. For two of the ALIs (S2 and Taxonomic Richness), there was an initial breakpoint at 0.3 mg/L TN, where the curve transitioned from more-or-less flat to a strong negative slope, and a final breakpoint at 0.8 mg/L TN,

after the curve became essentially flat again. The former may be considered a “resistance” threshold, and the latter an “exhaustion” threshold (see Introduction to this Section for definitions). In the case of Taxonomic Richness, there was also a gentler break in slope of the partial dependence plot around 0.55 mg/L TN, which was also the location of the break in slope for the partial dependence plot for Intolerant Percent Taxa (Figure 3.27). Thus, for the three ALIs, which collectively represent two different assemblages, similar patterns of response to TN were observed, providing weight of evidence for threshold locations along this gradient.

Partial dependence plots for the TP gradient (Figure 3.28) were reasonably congruent across ALI response variables. There was a precipitous drop in the fitted values for RAWmeanZHR, D18, and H20 along the TP gradient until around 0.05 – 0.1 mg/L TP, beyond which no further decline was evident for RAWmeanZHR, and minor fluctuations in slope were observed for D18 and H20.

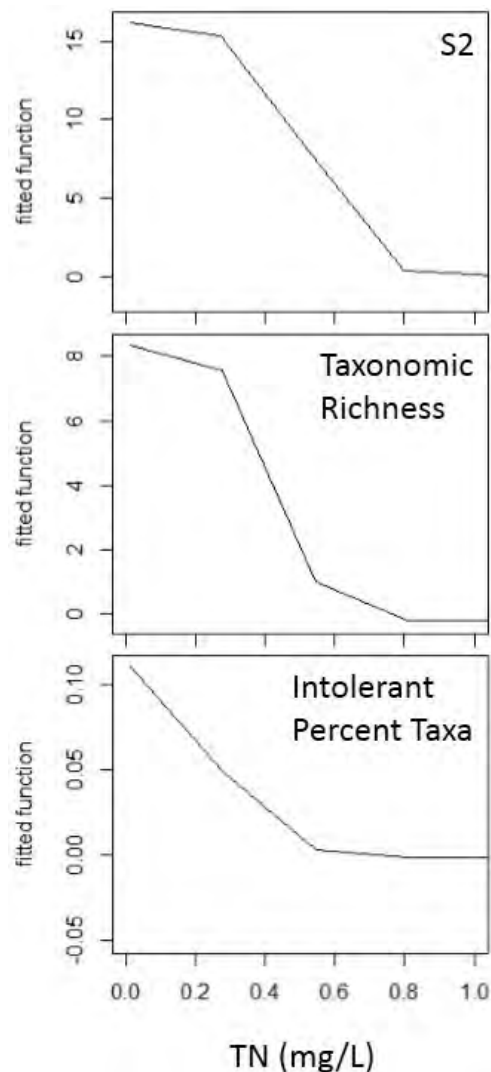


Figure 3.27. Partial dependence plots of TN from BRT models predicting three ALI response types: the soft algae IBI, S2; and the BMI ALIs, Taxonomic Richness and Intolerant Percent Taxa. Y-axes correspond to the standardized, fitted ALI variables.

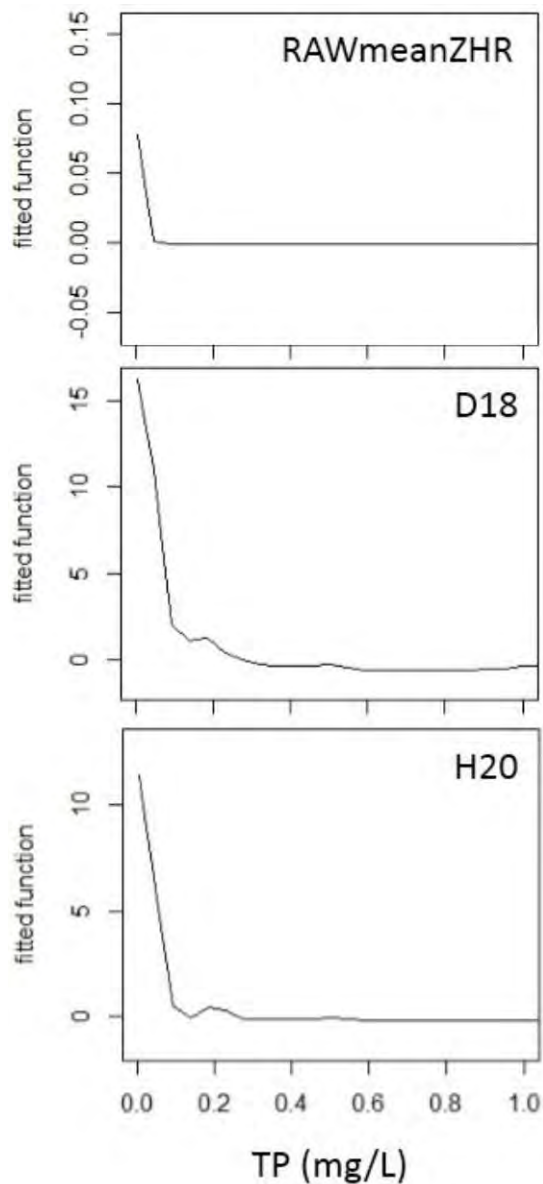


Figure 3.28. Partial dependence plots of TP from BRT models predicting three ALI response types: the soft-algae ALI, RAWmeanZHR; and the IBIs, D18 and H20. Y-axes correspond to the standardized, fitted ALI variables.

3.3.4 Thresholds for Biomass and Nutrient Effects on Biotic Response

We employed a wide variety of analytical methods, ALI response variables from different biotic communities, and primary producer abundance measures to evaluate potential thresholds of effect of biomass on stream ALIs. The sheer volume of output from this effort practically guaranteed that the results would not all point to a single biomass or nutrient threshold. However, for many of the gradients examined, there was a reasonable degree of consensus among analytical techniques and ALI response types within each of the four “ALI categories” (Figures 3.29- 3.30), even between biotic assemblages, thus providing a weight of evidence for fairly narrow ranges of threshold values. Most of the thresholds we observed could be classified as “exhaustion” thresholds (as defined in Section 3.2.3). In other words, ALI responses, as inferred through our approach, were generally saturated at the point along the stressor gradient at which we observed most

thresholds. Thus most of our thresholds would best be considered "backstops", and this would be important to keep in mind when considering these results in any policy decisions.

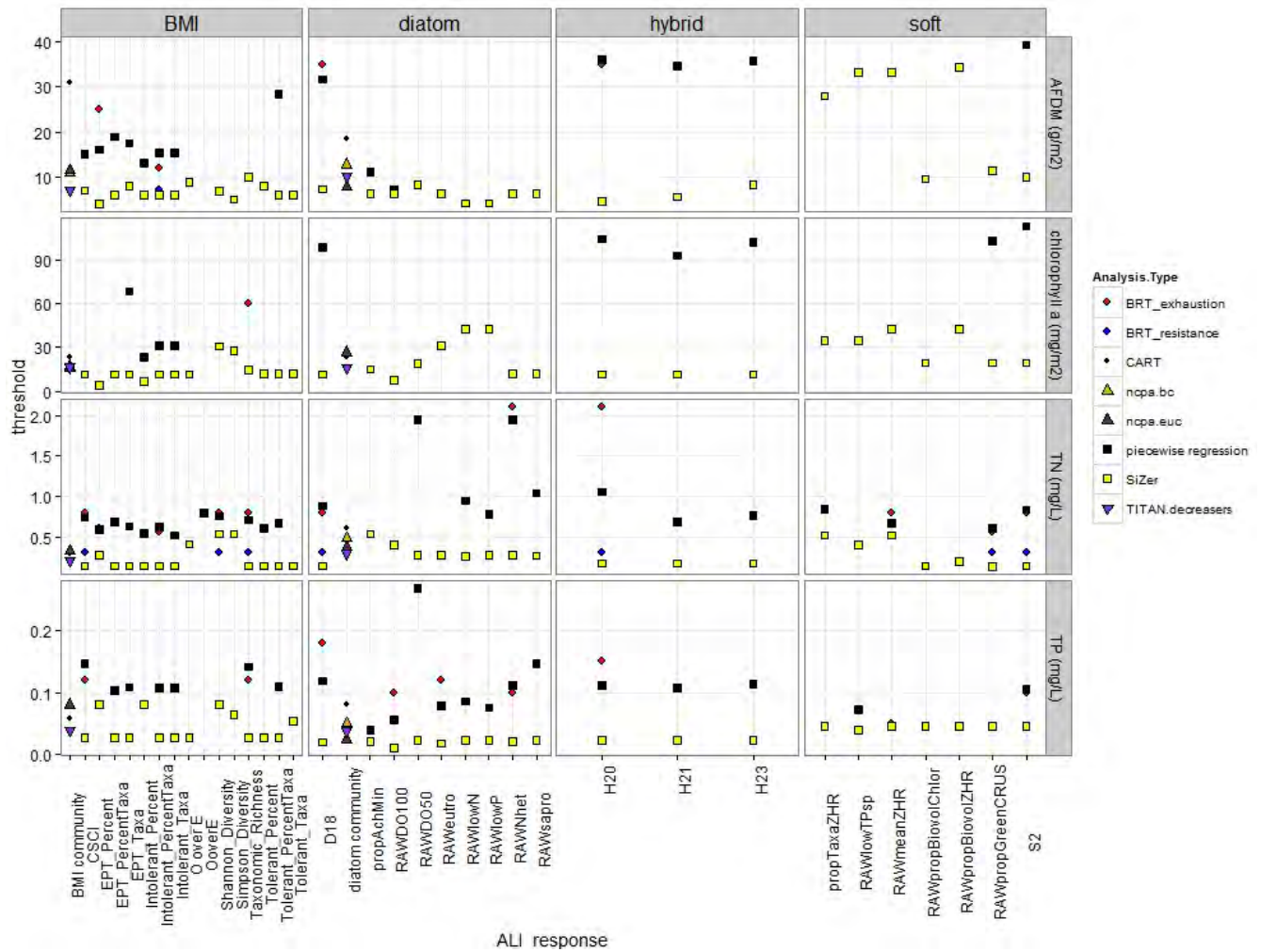


Figure 3.29. Summary of results across analyses using the chlorophyll a, AFDM, TN, and TP gradients, stratified by assemblage type. The y-axis corresponds to the threshold that was identified by the analysis in question for each of the gradients. For the CART results, data are based on the models using the full statewide dataset. For the piecewise regression analyses, break points (thresholds) are given only for those that passed at least the “relaxed” evaluation criteria. For the BRT thresholds, only those for which the predictor was significantly correlated with the ALI response in the partial Mantel test are included. Note that the values corresponding to SiZer analyses can more appropriately be viewed as indicative of a significant, dramatic slope change that precedes a threshold, rather than as a threshold in itself. Analysis-specific confidence limits, where applicable, are provided in the figures and tables presented previously. See Table 3.1 for definitions of the ALI variables. A tabular version of this information is provided in Table C.3.

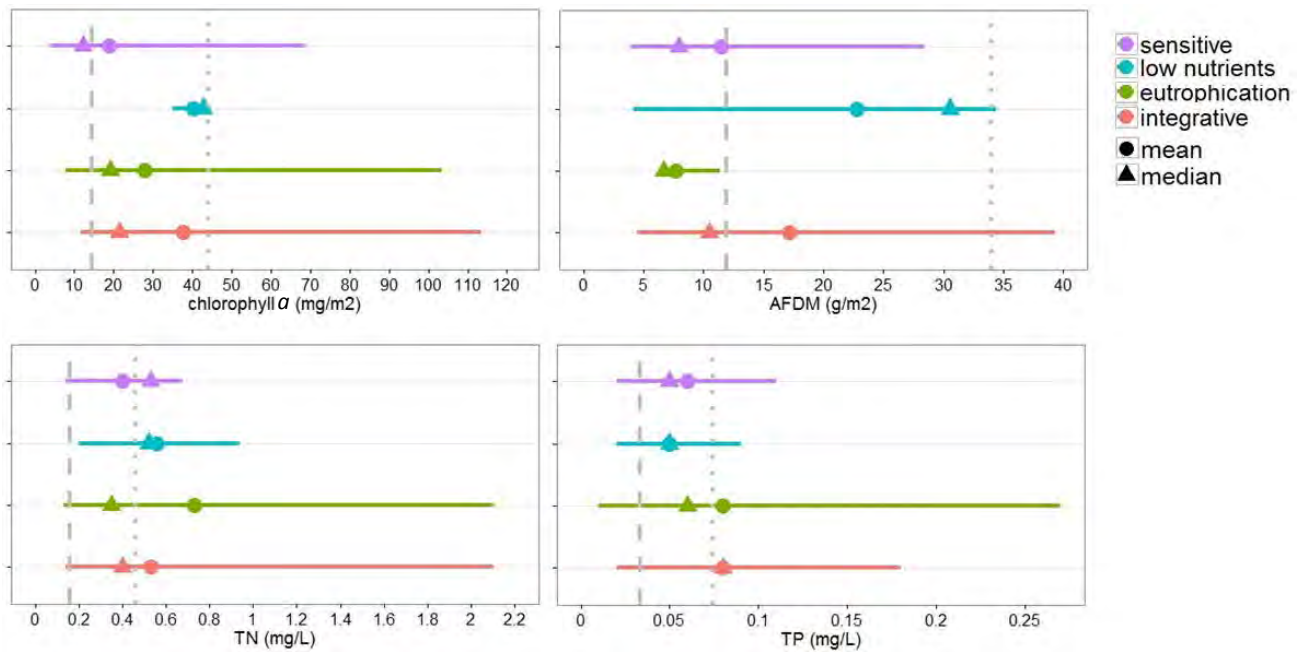


Figure 3.30. Ranges of thresholds of ALI response by “ALI category” (as described in section 3.2.2) for two biomass and two nutrient gradients. The same data that are shown in Figure 3.29, all assemblages and analyses combined, were used to make these graphs. Circles correspond to the mean of thresholds within each category, and triangles are the medians. Dashed lines indicate the 75th percentile of the indicator in question among Reference sites statewide, and dotted lines indicate the 95th.

3.4 Discussion

This study found evidence for a range of thresholds of effect for benthic chlorophyll *a*, AFDM, and TN and TP concentrations on BMI and algal community structure. Most of the thresholds observed could be classified as "exhaustion" thresholds¹⁷-- a sharp transition in the stressor gradient at which point the response variable reaches a natural limit (Cuffney et al. 2010). Thus we have generally characterized these thresholds as indicative of “adverse” effects on the ALI responses used. Integrative ALIs (such as IBIs) corresponded to higher thresholds whereas ALI measures specific to constrained groups of “sensitive” taxa generally corresponded to lower thresholds, illustrative of the paradigm of the biological condition gradient (Davies and Jackson 2006; Figure 3.2). In this discussion, we employ the median range within the ALI categories (sensitive, low nutrient, eutrophication, integrative) in order to summarize and compare with the literature. These ranges do not imply value judgments with respect to rigor of analytical approach nor importance of AL indicator type and thus should not be construed as policy recommendations.

Most of these thresholds of effect exceeded the 75th percentile of these indicators among Reference stream reaches statewide, but they were often less than the 95th percentile (Figure 3.30). Statistically significant relationships between stressors (benthic chlorophyll *a* concentrations, AFDM, nutrients) and a variety of ALIs were observed. However, change points in the response to AFDM and nutrient concentrations were more discernible than that for chlorophyll *a* (as currently measured in California ambient monitoring programs). These conclusions are based on analytical criteria for assessing the level of confidence in thresholds and the

¹⁷ Ecologically meaningful resistance thresholds may not always exist (or may be so low as to be undetectable with available methods/data), and few were apparent based on our analyses.

degree of consistency of thresholds across ALI indicators as multiple lines of evidence, both within and among assemblages.

3.4.1 Statistically Detected Thresholds of Adverse Effect in California Wadeable Streams

Benthic Chlorophyll *a*

Benthic chlorophyll *a* had a statistically-significant relationship with many California stream ALIs. Overall, thresholds of adverse effects ranged from 4 to 113 mg/m² chlorophyll *a*, and median thresholds within ALI categories ranged from 12 to 43 mg/m². Most of our analysis- and ALI-specific chlorophyll *a* thresholds exceeded the 75th percentile of chlorophyll *a* values among Reference stream reaches statewide (14.6 mg/m²; Chapter 2). A review of literature revealed only one study that used statistical methods to detect thresholds of adverse effects of benthic chlorophyll *a* on ALI indicators in wadeable streams. Miltner (2010) found a change point at 107 mg/m² related to changes in the abundance of Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa in Ohio streams. EPT are benthic macroinvertebrate taxa typically associated with “clean water” streams.

The range of thresholds found in this study were substantially lower than the NNE endpoints recommended for streams by a working group of international experts, regulatory agencies and stakeholders (Tetra Tech 2006, Appendix A). Two factors should be considered in comparing thresholds in literature that had been cited in support of the NNE thresholds proposed by Tetra Tech (2006; e.g., Biggs 2000; Quinn and Hickey 1990) with those of our study: 1) the temporal breadth of sampling and the temporal statistic that is the basis for the threshold, and 2) the range of ALI indicators considered (benthic invertebrates vs. salmonid fisheries).

Based on a study of 31 reaches in 21 New Zealand streams, Biggs (2000) observed that chlorophyll *a* concentrations exceeding ~13-20 mg/m² were associated with a 50% reduction in the percentage of EPT taxa. These chlorophyll numbers fall within the range of thresholds found in this study, which is based on a one-time sample in a spring – summer index period. In contrast, Biggs’ (2000) values are based on mean monthly samples. Biggs (2000) goes on to note that mean monthly sampling over the course of a year in 16 oligotrophic streams (defined as those with catchments having < 1% developed land use) yielded a 90th percentile of 20 mg/m² and a mean peak biomass of 47 mg/m². It is on the basis of this work that Biggs (2000, p. 97) stated, “I recommend that the mean monthly biomass not exceed 15 mg/m² and the peak biomass not exceed 50 mg/m² for the protection of benthic biodiversity in streams”. He goes on to add that the two measures imply that ALIs can continue to thrive when benthic algal abundance is elevated for a short duration, but that more substantial adverse effects would occur with chronic algal blooms. Unfortunately, repeat sampling that would be helpful to relate the one-time sample taken during the PSA spring-summer index period to mean monthly or maximum statistics has not been conducted for California. Thus it is important that any application of thresholds from this study to policy development consider the temporal statistic and the monitoring frequency with which regulatory decisions would be made.

Thresholds arising from the present study were derived based on changes to algal and benthic macroinvertebrate community composition, while NNE endpoints are also supported by literature linked to salmonid beneficial uses. Biggs (2000) asserted that protection of salmonids affords a slightly higher algal biomass than is protective of benthic invertebrate “clean water species”. Quinn and Hickey (1990) demonstrated that trout biomass increased from oligotrophic (< 20 mg/m²) to mesotrophic (20-100 mg/m²) streams, but then fell three-fold in eutrophic streams (> 100 mg/m²). Biggs (2000) demonstrated that mean

monthly benthic algal biomass in New Zealand streams that are “renowned for their trout fisheries” was 23 mg/m², with average maximum biomass of 171 mg/m².

Further modeling studies by Quinn and McFarlane (1989) link abundance of macroalgae at 21 °C in excess of 120 mg/m² chlorophyll *a* to depressed dissolved oxygen (DO) (i.e., < 5 mg/L). Similarly, Miltner (2010) found a change point in the 24-h DO concentration range to occur at a benthic chlorophyll *a* concentration of 182 mg/m² and suggested that this biomass level not be exceeded in order to maintain DO levels > 4 mg/L, and protect “existing high-quality waters”. For the California streams, algal indicators of oxygen-saturated waters (RAWDO100) and oxygen-depleted waters (RAWDO50) showed exhaustion thresholds of 45 and 115 mg/m² chlorophyll *a*, respectively. Though temperature and other site-specific factors play a role in determining the amount of algal biomass that would result in depression of stream DO, the scientific basis for establishing separate biomass endpoints for COLD and WARM Wadeable streams remains unclear, and our study does not further inform this debate.

Statistical confidence in benthic chlorophyll *a* thresholds found in this study was not as strong as for AFDM and nutrient concentration thresholds, based on analytical criteria for assessing the level of uncertainty in thresholds. None of the piecewise regression analyses for chlorophyll *a* fulfilled the “strict” criteria for determining confidence in the breakpoint, however some fulfilled the “relaxed” criteria. While a reasonable degree of consensus in thresholds was found among ALIs within assemblages, relatively poor agreement was found between assemblages—indicating a variable biotic response to chlorophyll *a*. Furthermore, partial dependence plots from the BRT analyses exhibited roughly linear relationships between chlorophyll *a* and predicted ALI response and suggested only weak thresholds. Thus we recommend use of predictive regression models to estimate benthic chlorophyll *a* concentrations that are quantitatively linked to an ALI target (such as CSCI, once the index is finalized and a quantitative target is established).

Our limited ability to detect benthic chlorophyll *a* thresholds may be due to: 1) heterogeneity of the streams across more than 100 miles of latitude and 2) low precision of the rapid stream assessment protocol employed in ambient surveys. BRT analyses revealed that chlorophyll *a* had a relatively weak influence on ALI response variables within the context of other predictors such as nutrient concentrations, stream physical habitat measures, meteorological variables, and land-use. Fetscher et al. (2009) found relatively poor precision in streams with chlorophyll *a* values exceeding approximately 50 mg/m². This is likely due to the high degree of patchiness of macroalgae, which is often the primary contributor to high values of algal biomass (Sheath et al. 1986, Wehr and Sheath 2003)¹⁸. This has led to the suggestion that a higher density of sampling may be needed in order to overcome some of the sampling error contributed by the patchiness.

¹⁸ Meaning that soft algae are more likely to proliferate to nuisance conditions than diatoms, as measured by Chlorophyll *a*.

AFDM

AFDM is an alternative measure of biomass, incorporating live as well as dead autochthonous and allochthonous organic matter. As with benthic chlorophyll *a*, peer-reviewed literature provided little in the way of examples of wadeable stream studies using quantitative methods to detect AFDM thresholds of effect on ALIs. The only work we found that suggests AFDM thresholds linked to ALIs was Biggs (2000) in which a 50% reduction in the number of EPT taxa was found to correspond to AFDM levels $> 5 \text{ g/m}^2$, based on a study of 31 sites across 21 New Zealand streams. This value aligns with the lower range of thresholds found in the present study (4 to 39 g/m^2 overall, with a median values within ALI categories ranging from 7 to 31 g/m^2), with similar caveats as those stated above regarding mismatch between our two studies in terms of temporal sampling.

In the present study, AFDM was, overall, the biomass variable with the strongest influence on BMIs and diatoms in the BRT analyses, and was the second-highest-ranked predictor for the ALI measure, EPT_Percent. AFDM exhibited similar thresholds of effect across ALIs from different biotic assemblages based on piecewise regressions, although thresholds tended to be lower for BMI indicators. Furthermore, CIs for these thresholds were generally narrow, despite the fact that a mix of organic matter sources (labile and refractory) are found in streams across California, and their modes of action on both algal and BMI communities differ. AFDM may, in general, be a more suitable predictor of ALI responses than chlorophyll *a*, an unsurprising result given that AFDM is the most integrative and quantitative measure of biomass that we have available. AFDM is more quantitative than the percent cover metrics, which either ignore thickness or estimate it into bins of varying width, and it is the most integrative biomass indicator because it includes all forms of stream organic matter (microbial biomass and live and dead algal and vascular plants—in terms of allochthonous inputs and autochthonous production). This is due to the fact that AFDM captures live and dead algal biomass as well as fungal and bacteria biomass, which are also stimulated by nutrient overenrichment (Gulis and Suberkropp 2004, Carr et al. 2005). In fact, in their recent review of stream nutrient criteria development approaches, Evans-White et al. (2013) asserted that “heterotrophic bases for criteria establishment should be considered in conjunction with the more traditional autotrophic bases for criteria establishment.” AFDM has the added advantage that it is less susceptible to degradation than chlorophyll *a*, or to variability in the algal C:chlorophyll *a* ratio, as noted above.

As an indicator, AFDM is not without challenges, however. The 75th percentile value of Reference sites (11.9 g/m^2) lies squarely within mid-range of thresholds detected, suggesting that some wadeable streams are naturally carbon-enriched (e.g., forests with terrestrial carbon inputs). This would render AFDM an indicator prone to false positives, without controlling for exogenous factors. It is worth noting that Biggs (2000a) does not recommend specific criteria for AFDM, because “AFDM is more prone to large measurement error with low biomass accrual.” It may be advisable to move California’s PSA program toward piloting a carbon-enrichment measure that provides information on carbon source as well as biomass. For example, benthic C:N ratio can be used to indicate algal (labile) versus terrestrial (refractory) sources of carbon to sediments (e.g., Ruttenberg and Goñi 1997). More work may also be needed on detrital-based headwater streams. In other regions of the country, when nutrients have a disproportionate impact on predator-resistant consumers, headwater streams have shown long-term declines in organic matter as detritivore activity increases in response to moderate nutrient enrichment.

Other Biomass Gradients

In addition to chlorophyll *a* and AFDM, we looked at several other stream primary producer abundance indicators, including several types of algal and macrophyte percent cover, and soft-algal total biovolume. Macroalgal percent cover (PCT_MAP) has been suggested as an efficient and informative means of estimating stream algal biomass (Fetscher and McLaughlin 2008), because it can be assessed rapidly at a much higher spatial density than traditional benthic chlorophyll *a* biomass samples can be quantitatively collected. Results of our analyses failed to find macroalgal percent cover as a strong predictor of ALI responses, and no well-supported thresholds along this gradient were apparent in the analyses we conducted. However, other types of analytical approaches (e.g., those discussed in the Introduction to this chapter) that are not based on thresholds may be useful to incorporate in future work. Also, it is worth noting that nuisance algal mats are of great concern from an aesthetic standpoint (Biggs 2000, Lembi 2003, Suplee et al. 2009), and there are percent cover thresholds in the literature relating to aesthetic (REC-2) beneficial uses. For example, Welch et al. (1988) and Biggs (2000) have suggested that macroalgal percent cover in the range of 20-30% and above is unacceptable from the standpoints of aesthetics and recreation. Thus macroalgal percent cover may merit numeric endpoints on the basis of REC-2 beneficial uses regardless of whether strong relationships between macroalgal percent cover and ALIs can be discerned.

The lack of thresholds of effect of macroalgal percent cover detected in this study may be a consequence of the way this biomass type is currently measured. The rapid, point-intercept procedure that assesses macroalgal presence/absence along a predetermined grid of 105 points (Fetscher et al. 2009) takes into account only two-dimensional (areal) cover, ignoring thickness, which is potentially an important determinant of biomass. Thus implementing some form of area-weighted biomass, that quantifies algal biomass at specific points in the stream in addition to recording cover at a high density of observation points, may be a means of obtaining higher precision information about stream algal biomass. However, such an effort would likely add considerably to field time during sampling. Nonetheless, percent cover information, as currently collected, may be useful as a screening variable to place a “ceiling” on the amount of benthic chlorophyll *a* and/or AFDM likely present in a stream (Fetscher et al. 2013). This would require establishing a relationship (e.g., via quantile regression) to determine an upper bound for the “maximum” amount of chlorophyll *a* (or AFDM) possible, given a specific percent cover value. Such knowledge would allow the user to rule out a chlorophyll *a* or AFDM-based biomass exceedance when percent cover outcomes are below a pre-determined value.

Nutrients

TN and TP concentration had strong, statistically-significant relationships with stream ALI indicators; thresholds detected in this study ranged from 0.13 to 2.1 mg/L for TN and 0.01 to 0.27 mg/L for TP¹⁹ (medians within ALI categories ranged from 0.35 to 0.53 mg/L TN and 0.05 to 0.08 mg/L TP among the ALI categories). These ranges largely fell within the collective ranges of values from the literature (0.41 to 1.79 mg/L for TN; 0.0082 to 0.28 mg/L for TP; Table 3.11), thus lending additional support for the numbers we derived. Most of our analysis- and ALI-specific TN thresholds exceeded the 75th percentile of TN values among California Reference stream reaches (0.162 mg TN/L), while the 75th percentile of TP

¹⁹ Note, however, that many “sensitive” taxa had even lower thresholds, based on TITAN analysis (see Appendix C.2)

Table 3.11. Quantitatively determined thresholds of stream (or river) ALI responses to nutrient concentrations. “Min.” refers to the minimum value from each publication, across all ALI types and analytical methods employed. “Max.” is the maximum for this value.

Citation	Region	ALI measure(s)	gradient(s)	threshold detection method	max.			
					min. TP (mg/L)	TP (mg/L)	min. TN (mg/L)	max. TN (mg/L)
the present study	California	BMI, algae	biomass, nutrients	TITAN, nCPA, CART, piecewise regression, BRT	0.011	0.267	0.13	2.1
Baker et al. 2010	Everglades	BMI	TP	TITAN and nCPA	0.015	0.019	-	-
Black et al. 2011	western United States	diatoms	TN,TP	piecewise regression	0.03	0.28	0.59	1.79
Evans-White et al. 2009	Kansas, Nebraska, Missouri	BMI	TN,TP	nCPA	0.05	0.05	1.04	1.04
Paul et al. 2007	southeastern Pennsylvania	BMI, diatoms	TP	nCPA	0.038	0.064	-	-
Qian et al. 2003	Everglades	BMI	TP	change point estimated using the nonparametric & the Bayesian methods	0.011	0.014	-	-
Richardson et al. 2007	Everglades	algal, macrophyte and BMI	TP	Bayesian change point analysis	0.008	0.024	-	-
Smith et al. 2010	New York State	BMI, diatom	TN,TP	nCPA	0.009	0.07	0.41	1.2
Smith et al. 2007	New York State	BMI	TP, NO ₃	Hodges-Lehmann estimation	0.065	0.065	0.98 (NO ₃)	0.98 (NO ₃)
Smucker et al. 2013a	Connecticut	diatoms	TP	boosted regression trees	0.019	0.082	-	-
Stevenson et al. 2008	Mid-Atlantic Highlands	diatoms	TP	lowess regression and regression tree analysis	0.012	0.027	-	-
Wang et al. 2007	Wisconsin	fish, BMI	TN,TP	regression tree analysis & 2-dimensional Kolmogorov-Smirnov techniques	0.06	0.09	0.54	0.61
Weigel and Robertson 2007	Wisconsin	fish, BMI	TN,TP	regression tree analysis	0.06	0.06	0.64	0.64

concentrations at Reference sites (0.033 mg TP/L) was within the lower end of the range of TP thresholds we observed. The agreement in nutrient concentration thresholds between those identified in our study and what is presented in the literature is somewhat surprising, given that all but one of the studies were conducted in different biogeographic provinces (i.e., east of the Rocky Mountains) and across a diverse array of stream types. In particular, several studies were conducted in regions with cooler climates and/or those with higher levels of precipitation year-round than that which represents the bulk of our study region, and some were conducted in rivers rather than wadeable streams. Black et al. (2011) is the only study from the

western United States. Their ranges of thresholds of effects on diatom communities in agriculturally-dominated to low-impact Wadeable streams in the western U.S were 0.03-0.28 mg/L for TP and 0.59-1.79 mg/L for TN.

A recent review by Evans-White et al. (2013) summarized common approaches to stream nutrient criteria development and thresholds of effect that have been reported for nutrients on BMI and fish ALIs. All studies covered by the review are included in Table 3.11 of this report. Included in the review were the numeric criteria for 12 states that were established between 2010 and 2012. Criteria for TP range from 0.01 to 0.49 mg/L and criteria for TN range from 0.13 to 5 mg/L.

Our findings, along with those in recent studies, suggest that nutrients may be exerting direct effects on ALIs via means not mediated through pathways typically cited in eutrophication literature (e.g., via increases in primary production and concomitant reduction in dissolved oxygen levels; Dodds and Welch 2000). Direct effects of stream nutrients can occur through nutrient toxicity (Camargo and Alonso 2006). Nutrient enrichment can also precipitate changes in instream food quality. Under this latter scenario primary consumers with a high nutrient demand are disproportionately affected by low-quality food relative to those with lower nutrient demands (Sterner and Elser 2002). This results in altered competitive interactions among species (Evans-White et al. 2009), which, in turn, decrease diversity and cause shifts in benthic community structure (Gafner and Robinson 2007, Singer and Battin 2007). More recent studies have demonstrated effects of moderate nutrient loading on headwater streams as the result of effects on heterotrophic production and food web shifts (Davis et al. 2010, Suberkropp et al. 2010).

Statistically, confidence in the nutrient concentration thresholds is high. In BRT models, nitrogen in one form or another was the top-ranked predictor for several BMI and soft algae ALIs and phosphorus was the top-ranked predictor for two diatom ALIs. This was despite the fact that a wide variety of land use, geographic, meteorological, geological, and local stream physical habitat variables (as well as algal biomass) were included as predictors in the models. The piecewise regression analyses for which confidence in the breakpoint was highest (i.e., those that passed the “strict” criteria) were based on ALI responses to nutrient gradients. TITAN analyses indicated well-supported, community-level change points along nutrient gradients. BRT partial dependence plots revealed easily-discernible breaks in slope across nutrient gradients, not only for tolerance/sensitivity type metrics, but also for several of the more integrative measures (e.g., IBIs), and a relatively high level of consensus in nutrient thresholds from partial dependence plots was observed across biotic assemblages. Thus, based on the output of widely different analytical techniques for multiple biotic assemblages, narrow ranges of thresholds with high confidence were realized for both TN and TP.

3.4.2 Variable Response of ALI Types to Biomass and Nutrients: The Biological Condition Gradient

The gradient of thresholds of ALI response to algal abundance indicators and nutrients illustrates the paradigm of the biological condition gradient (BCG, Davies and Jackson 2006, Figure 3.2). Integrative ALIs (such as IBIs) tended to correspond to higher thresholds whereas ALI measures specific to constrained groups of “sensitive” taxa generally corresponded to lower thresholds. At the same time, ALI variables that were based on highly integrative indices (e.g., CSCI, and the algae IBIs) tended to exhibit threshold responses to biomass that were not as well-supported as those for individual metrics based on sensitive/intolerant taxa (e.g., EPT_Percent and Intolerant_PercentTaxa). This finding conforms to the observation of Baker and King (2010) that integrative indices may blur taxon-specific change points, relative to information about individual taxa or small groups of taxa that share similar autecological characteristics. That notwithstanding, some of

the index-based ALI variables (e.g., the algae IBI, H2O and the soft algae IBI, S2) were highly responsive directly to nutrients, for which clear thresholds were observed. Many of the ALI metrics also exhibited high nutrient responsiveness, such as RAWmeanZHR, for which nitrogen was a strong predictor, and for which marked thresholds were observed. This is not surprising, because this metric incorporates information about the relative abundance of heterocystous cyanobacteria in the algal community. These organisms are capable of fixing N₂ and are therefore excellent indicators of stream nitrogen limitation (Stancheva et al. 2013). In general, ALI variables showed strong responsiveness to nutrients, but in different ways. Diatom-based ALIs were more influenced by phosphorus (in accordance with the findings of Ponader et al. 2008), whereas BMIs and soft algae were more influenced by nitrogen. Thus assessing multiple assemblages concurrently may provide a broader perspective on stream nutrient status.

Our results did diverge from the traditional BCG gradient paradigm in one respect. The BCG paradigm suggests that little change will occur in functional level parameters until systems have degraded to levels 4 and 5. However, diatom indicators suggested changes in DO regime may be occurring at lower levels, coincident with the loss of sensitive macroinvertebrate taxa.

3.4.3 Study Findings in Context of Policy Applications

The thresholds for algal abundance and nutrients that were derived from this study are based on a data set that represents an index period of late spring-early summer. Since nutrient management occurs year-round, it is important to consider the extent to which our analyses can be applied outside the index period. We acknowledge that thresholds may differ for other times of the year and other stream types. For example, our results are based on instantaneous measurement at low-flow conditions, and as such, do not reflect year-long loads or storm flows. It is not clear to what degree the types of ALI-stressor relationships we observed would hold during rain events. Similarly, although the target population for the surveys that generated the data was perennial, wadeable streams, in reality, some of these streams are actually intermittent. It is not always possible to distinguish between perennial and intermittent hydrology unless the site is visited in the late-summer or fall, prior to the onset of the rainy season (which is outside of the index period for sampling).

Finally, differences in the type of biotic community that can be supported by different wadeable stream types (e.g., low-order, high-gradient mountain streams vs. concrete-lined low-gradient streams in developed areas) may affect the nature of response thresholds. Our statewide data set included a mix of stream types spanning multiple regions and broad natural and anthropogenic gradients, including channelized systems. With the exception of CART analyses of NMS axis breakpoints, we did not explicitly test for differences in response across regions or stream classes (e.g., natural versus modified channels). Analyses presented in Chapter 4 demonstrate that multiple factors (many correlated with urban development) may be influencing and modifying the response of benthic algal biomass to nutrients. Thus, it is possible that relationships and thresholds presented here could be further refined through stratification of the current data set.

3.5 Literature Cited

- Abol, M., Puig M.A., Ríos H., and E. López-Jiménez. 2000. Relationship between macroinvertebrate diversity and toxicity of (Cyanophyceae Cyanobacteria) in some streams from eastern Spain. *Verh. Int. Ver. Limnol.* 27:555–559.
- Baker, M.E. and R.S. King. 2010. A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods in Ecology and Evolution*, 1:25–37.
- Biggs, B.J.F. 2000. New Zealand Periphyton Guideline: Detecting, Monitoring and Managing the Enrichment of Streams. Ministry for Environment Publication, Wellington, NZ. 151 pp.
- Black, R.W., Moran P.W., and J.D. Frankforter. 2011. Response of algal metrics to nutrients and physical factors and identification of nutrient thresholds in agricultural streams. *Environmental Monitoring and Assessment*, 175(1-4): 397-417.
- Brenden, T.O., Wang L., and Z. Su 2008. Quantitative identification of disturbance thresholds in support of aquatic resource management. *Environmental Management*, 42:821–832.
- Camargo, J.A., and A. Alonso. 2006. Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment. *Environmental International*, 32:831–849.
- Carr, G. M., Morin A., and P.A. Chambers. 2005. Bacteria and algae in stream periphyton along a nutrient gradient. *Freshwater Biology*, 50:1337-1350.
- Chaudhuri, P. and J.S. Marron. 1999. SiZer for exploration of structures in curves. *Journal of the American Statistical Association*, 94:807–823.
- Chorus, J., and J. Bartram. (eds.) 1999. Toxic Cyanobacteria in Water. A Guide to their Public Health Consequences, Monitoring and Management. World Health Organization, E & F N Spon., 416pp.
- Connolly, N.M. and R.G. Pearson. 2013. Nutrient enrichment of a heterotrophic stream alters leaf litter nutritional quality and shredder physiological condition via the microbial pathway. *Hydrobiologia*, 718:85-92.
- Cuffney, T.F., Brightbill R.A., May J.T., and I.R. Waite. 2010. Responses of benthic macroinvertebrates to environmental changes associated with urbanization in nine metropolitan areas. *Ecological Applications*, 20:1384–1401.
- Davies, S.P. and S.K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. *Ecological Applications*, 16:121-1266.
- Davies, R.B. 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 74:33–43.
- Davis, J.M., Rosemond A.D., Eggert S.L., Cross W.F., and J. Bruce Wallace. 2010. Long-term nutrient enrichment decouples predator and prey production. *PNAS*, 107(1):121-6.
- De'ath, G. and K.E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81: 3178-3192.
- Dodds, W.K. 2007. Trophic state, eutrophication and nutrient criteria in streams. *Trends in Ecology and Evolution*, 22(12):669-76.
- Dodds, W. K., and E. B. Welch. 2000. Establishing nutrient criteria in streams. *Journal of the North American Benthological Society*, 19:186–196.

- Dodds, W.K., Clements W.H., Gido K., Hilderbrand R.H., and R.S. King. 2010. Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. *Journal of the North American Benthological Society*, 29:988–997.
- Douterelo, I., Perona E., and P. Mateo. 2004. Use of cyanobacteria to assess water quality in running waters. *Environmental Pollution*, 127:377–384.
- Dufrêne, M. and P. Legendre. 1997. Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67:345–366.
- Elith, J., Leathwick J.R., and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77:802–813.
- Evans-White, M., Dodds W.K., Huggins D.G., and D.S. Baker. 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.
- Evans-White, M.A., Haggard B.E., and J.T. Scott. 2014. A review of stream nutrient criteria development in the United States. *Journal of Environmental Quality*, 42:1002-1014.
- Fetscher, A.E., Stancheva R., Kociolek J.P., Sheath R.G., Stein E.D., Mazor R.D., Ode P.R, and L.B. Busse. 2014. Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination. *Journal of Applied Phycology*, 26:433-450.
- Fetscher, A.E., Sutula M.A., Busse L.B., and E.D. Stein. 2013. Condition of California perennial, wadeable streams based on algal indicators. Technical Report for the California Surface Water Ambient Monitoring Program (SWAMP), Sacramento, CA.
- Fetscher, A.E., Busse L.B., and P.R. Ode 2009. Standard Operating Procedures for Collecting Stream Algae Samples and Associated Physical Habitat and Chemical Data for Ambient Bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 002. (updated May 2010)
- Fetscher, A.E. and K. McLaughlin. 2008. Incorporating bioassessment using freshwater algae into California's Surface Water Ambient Monitoring Program (SWAMP). Technical Report 563. Southern California Coastal Water Research Project. Costa Mesa, CA.
- Fovet, O., Belaud G., Litrico X., Charpentier S., Bertrand C., Dollet P., and C. Hugodot. 2012. A model for fixed algae management in open channels using flushing flows. *River Research and Applications*, 28:960–972.
- Gafner, K., and C. T. Robinson. 2007. Nutrient enrichment influences the responses of stream macroinvertebrates to disturbance. *Journal of the North American Benthological Society*, 26:92–102.
- Goslee, S.C. and D.L.Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, 22(7):1-19.
- Greenwood, J.L., Rosemond A.D., Wallace J.B., Cross W.F. and H.S. Weyers. 2007. Nutrients stimulate leaf breakdown rates and detritivore biomass: Bottom-up effects via heterotrophic pathways. *Oecologia*, 151(4): 637-649.
- Gulis, V. and K. Suberkropp. 2004. Effects of whole-stream nutrient enrichment on the concentration and abundance of aquatic hyphomycete conidia in transport. *Mycologia*, 96(1): 57-65.
- Hijmans, R.J., Phillips S., Leathwick J., and J. Elith. 2013. Dismo: Species distribution modeling. R package version 0.8-11. <http://CRAN.R-project.org/package=dismo>

- Kaufmann, P. R., Levine P., Robinson E. G., Seeliger C., and D. V. Peck. 1999. Quantifying Physical Habitat in Wadeable Streams. EPA/620/R-99/003, U.S. Environmental Protection Agency, Washington, D.C.
- King, R. S., Baker M.E., Whigham D.F., Weller D.E., Jordan T.E., Kazyak P.F., and M.K. Hurd. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. *Ecological Applications*, 15:137–153.
- King, R.S. and C.J. Richardson. 2008. Macroinvertebrate and fish responses to experimental P additions in Everglades sloughs. Pp 477-504 In: *The Everglades Experiments: Lessons for Ecosystem Restoration (Ecological Studies 201)*, C. J. Richardson. Springer-Verlag, New York.
- Lembi, C.A. 2003. Control of nuisance algae. In: Wehr JD, Sheath RG (eds). *Freshwater Algae of North America Ecology and Classification*. Academic Press, New York pp. 805 - 834.
- Mazor, R.D., Rehn A., Ode P.R., Engeln M., Schiff K., and C.P. Hawkins. Maximizing the applicability of a bioassessment index in environmentally complex regions. *Freshwater Science*. Under review.
- McCune, B. and J.B. Grace, 2002. *Analysis of Ecological Communities*. MjM Software, Gleneden Beach, Oregon, 300 pp., ISBN-13: 978-0972129008.
- Miltner, R.J. 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. *Environmental Management*, 45:842-855.
- Muggeo, V.M.R. 2003. Estimating regression models with unknown break-points. *Statistics in Medicine*, 22:3055–3071.
- Muggeo, V.M.R. 2008. Segmented: An R Package to Fit Regression Models with Broken-Line Relationships. *R News*, 8/1, 20-25. URL <http://cran.r-project.org/doc/Rnews/>.
- Nixon, S. 1995. Coastal eutrophication: A definition, social causes, and future concerns. *Ophelia*, 41: 199-220.
- Ode, P. 2007. SWAMP Bioassessment Procedures: Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Available from http://mpsl.mlml.calstate.edu/phab_sopr6.pdf
- Ode, P.R., Rehn A.C., and J.T. May. 2005. A quantitative tool for assessing the integrity of southern coastal California streams. *Environmental Management*, 35:493-504.
- Pasek, J. and Tahk A. 2012. Weights: Weighting and Weighted Statistics. R package version 0.75. <http://CRAN.R-project.org/package=weights>
- Paul, M. J., and L. Zheng. 2007. Development of Nutrient Endpoints for the Northern Piedmont Ecoregion of Pennsylvania: TMDL Application. Tetra Tech, Owing Mills, MD. Prepared for the U.S. Environmental Protection Agency Region 3. This report is available on the EPA Region III website at: http://www.epa.gov/reg3wapd/tmdl/pa_tmdl/IndianCreekNutrient/NEPStudySEPA.pdf.
- Ponader, K.C., Charles D.F., Belton T.J., and D.M. Winter. 2008. Total phosphorus inference models and indices for coastal plain streams based on benthic diatom assemblages from artificial substrata. *Hydrobiologia*, 610:139–152.
- Potapova, M. and D.F. Charles. 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. *Ecological Indicators*, 7: 48-70.
- Qian, S.S., King, R.S., and C.J. Richardson. 2003. Two statistical methods for the detection of environmental thresholds. *Ecological Modelling*, 166:87–97.
- Quinn, J.M., and B.W. Gilliland. 1989. The Manawatu River cleanup - Has it worked? *Transactions of the Institution of Professional Engineers, New Zealand*, 16:22-26.

- Quinn, J.M. and C.W. Hickey. 1990. Magnitude of effects of substratum particle size, recent flooding and catchment development on benthic invertebrate communities in 88 New Zealand rivers. *New Zealand Journal of Marine and Freshwater Research*, 24: 411-427.
- Quinn, J.M. and P.N. McFarlane. 1989. Effects of slaughterhouse and dairy factory wastewaters on epilithon: A comparison in laboratory streams. *Water Research*, 23: 1267-1273.
- Richardson, C.J., King R.S., Qian S.S., Vaithyanathan P., Qualls R.G., and C.A. Stow. 2007 Estimating ecological thresholds for phosphorus in the Everglades. *Environmental Science and Technology*, 41(23): 8084-8091.
- Riseng, C.M., Wiley M.J., and R.J. Stevenson. 2004. Hydrologic disturbance and nutrient effects on benthic community structure in midwestern US streams: A covariance structure analysis. *Journal of the North American Benthological Society*, 23:309-326.
- Ruttenberg, K.C. and M.A. Goñi. 1997. Phosphorous distributions, C:N:P ratios, and $\delta^{13}C_{org}$ in arctic, temperate, and tropical coastal sediments: tools for characterizing bulk sedimentary organic matter. *Marine Geology*, 139:123-145.
- Sheath, R.G., Morison M.O., Korch J.E., Kaczmarczyk D., and K.M. Cole. 1986. Distribution of stream macroalgae in south-central Alaska. *Hydrobiologia*, 135:259–269.
- Singer, G.A., and T.J. Battin. 2007. Anthropogenic subsidies alter consumer-resource stoichiometry, biodiversity, and food chains. *Ecological Applications*, 17: 376–389.
- Smith, A.J., and C.P. Tran. 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, A. J., Bode R.W., and G.S. Kleppel. 2007. A nutrient biotic index (NBI) for use with benthic macroinvertebrate communities. *Ecological Indicators*, 7(2): 371-386.
- Smucker, N. J., Becker M., Detenbeck N.E., and A.C. Morrison. 2013a. Using algal metrics and biomass to evaluate multiple ways of defining concentration-based nutrient criteria in streams and their ecological relevance. *Ecological Indicators*, 32:51-61.
- Smucker, N.J., Detenbeck N.E., and A.C. Morrison. 2013b. Diatom responses to watershed development and potential moderating effects of near-stream forest and wetland cover. *Freshwater Science*, 32:230–249.
- Sonderegger, D. 2011. SiZer: Significant Zero Crossings. R package version 0.1-4. <http://CRAN.R-project.org/package=SiZer>
- Stancheva, R., Sheath R.G., Read B.A., McArthur K.D., Schroepfer C., Kociolek J.P., and A.E. Fetscher. 2013. Nitrogen-fixing cyanobacteria (free-living and diatom endosymbionts): Their use in southern California stream bioassessment. *Hydrobiologia*, 720:111-127.
- Steinman, A.D., 1996. Effects of grazers on freshwater benthic algae. In R. G. Stevenson et al. (eds), *Algal Ecology*. Academic Press, San Diego, CA. pp 341–365.
- Sterner, R.W. and J. J. Elser. 2002. *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere*. Princeton University Press. Princeton, New Jersey.
- Stevenson, R.J., Hill B.H., Herlihy A.T., Yuan L.L., and S.B. Norton. 2008. Algae–P relationships, thresholds, and frequency distributions guide nutrient criterion development. *Journal of the North American Benthological Society*, 27:783-799.
- Stevenson, R.J., Bennett B.J., Jordan D.N., and R.D. French. 2012. Phosphorus regulates stream injury by filamentous green algae, thresholds, DO, and pH. *Hydrobiologia*, 695:25-42.

- Suberkropp, K., Gulis V., Rosemond A.D., and J.P. Bernstead. 2010. Ecosystem and physiological scales of microbial responses to nutrients in a detritus-based stream: Results of a 5-year continuous enrichment. *Limnology and Oceanography*, 55(1):149-160.
- Suplee, M.W., Watson V., Teply M., and H. McKee. 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.
- Suplee, M.W., and V. Watson, 2013. Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana’s Wadeable Streams and Rivers—Update 1. Helena, MT: Montana Dept. of Environmental Quality.
- Tant, C.J., Rosemond A.D., and M.R. First. 2013. Stream nutrient enrichment has a greater effect on coarse than on fine benthic organic matter. *Freshwater Science*, 32(4):1111-1121.
- Tetra Tech. 2006. Technical Approach to Develop Nutrient Numeric Endpoints for California. Tetra Tech, Inc. http://rd.tetratech.com/e pa/Documents/CA_NNE_July_Final.pdf
- Therneau, T. M. and B. Atkinson. 2009. mvpart: Multivariate regression trees. R package version 3.1 - 41, R port by Brian Ripley, extensions and adaptations of rpart to mvpart by G. De'ath. <http://CRAN.R-project.org/package=mvpart>
- U.S. Environmental Protection Agency. 2000. Nutrient Criteria Technical Guidance Manual, Rivers and Streams. U.S. Environmental Protection Agency, Office of Water, Washington, D.C. EPA-822-B-00-002.
- van Dam, H., Mertens A., and J. Sinkeldam. 1994. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. *Netherlands Journal of Aquatic Ecology*, 28:117–133.
- Wang, L., Robertson D.M., and P.J. Garrison. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: Implication to nutrient criteria development. *Environmental Management*, 39:194–212.
- Wehr, J.D. and R.G. Sheath. 2003. Freshwater habitats of algae. In: Wehr, J.D. and R.G. Sheath (eds) *Freshwater Algae of North America: Ecology and Classification*. Academic Press, San Diego, CA. pp 11–57.
- Weigel, B.M., and D.M. Robertson. 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, E.B., Jacoby J.M., Horner R.R., and M.R. Seeley. 1988. Nuisance biomass levels of periphytic algae in streams. *Hydrobiologia*, 157:161-168.
- Wootton J.T., Parker M.S., and M.E. Power. 1996. Effects of disturbance on river food webs. *Science*, 273:1558-1561.
- Wright, J.F., Furse M.T., and P.D. Armitage. 1993. RIVPACS: A technique for evaluating the biological quality of rivers in the UK. *European Water Pollution Control*, 3: 15-25.

4. Evaluation of Nutrient Numeric Endpoint Benthic Biomass Spreadsheet Tool

4.1 Introduction

The California State Water Resources Control Board is developing nutrient water quality objectives for the State's surface waters. Among the approaches that the SWRCB staff is considering is an ecological response approach, known as the Nutrient Numeric Endpoint framework (Tetra Tech 2006). The NNE framework, intended to serve as numeric guidance to translate narrative WQO, consists of two tenets: 1) use of response indicators to assess the status of waterbody condition with respect to eutrophication and other adverse effects of nutrient-overenrichment and 2) use of models to link regulatory endpoints for response indicators back to nutrients and other management controls. To facilitate the translation of NNE response indicators to nutrients, SWRCB supported the development of scoping-level nutrient-algal abundance models. These scoping models (e.g., the benthic biomass spreadsheet tool [BBST] for streams) were intended to be used as a starting point for setting site-specific numeric nutrient targets (Tetra Tech 2006). The intent was that the BBST helps users determine what the appropriate nutrient concentration targets should be, given other environmental co-factors at play at a site, based on the proposed algal abundance endpoints.

The NNE spreadsheet tools were developed during a period when relatively little California wadeable stream data was available to optimize the models. Thus the existing BBST was considered provisional, pending availability of larger datasets to aid in its refinement. Since that time, the Surface Water Ambient Monitoring Program (SWAMP) has developed a bioassessment program for perennial, wadeable streams, focused on the use of benthic macroinvertebrate community composition, water chemistry, measures of physical habitat, and toxicity to assess ecological condition. SWAMP has supported the development of standardized protocols for the collection of stream algal data (Fetscher et al. 2009) with the intent of adding algal community measures to its suite of indices of biological integrity (Fetscher et al. 2014). Since 2007, data from 1032 sites have been collected throughout California, permitting an evaluation of the suitability of the BBST for use in regulatory application in wadeable streams, which is a goal of the present analysis, and whether additional refinements are needed (e.g., regionally specific model coefficients to improve performance).

The objectives of this component of the study are to:

- Evaluate performance of the BBST for California perennial, wadeable streams,
- Explore sources of bias and error in BBST model predictions, in order to recommend potential refinements for wadeable stream nutrient-algal abundance models,
- Understand the relative influence of nutrients and environmental co-factors on stream primary producer abundance using Boosted Regression Trees (BRT), and
- Explore potential regional variation in predictive model coefficients using linearized versions of the Dodds and QUAL2K models through a Bayesian Classification and Regression Tree (B-CART) analysis.

4.2 Methods

4.2.1 Conceptual Approach to Validation and Error Analysis

We used existing data to validate the BBST in California wadeable streams and identify aspects of its underlying models that may require refinement for particular stream types or regions. BBST validation entailed assessment of the accuracy and bias of its underlying models relating nutrients to stream benthic biomass. This was accomplished by comparing observed biomass values (chlorophyll *a* and AFDM) with predicted values generated via the BBST models (Objective 1; TetraTech 2006). Furthermore, to facilitate an exploration of potential sources of model error and inform recommendations for future refinements to the BBST, the magnitude of deviation of BBST model predictions from observed biomass was used as the response variable in random forest models with site-specific and landscape-level factors as explanatory variables (Objective 2). To build upon this effort, we also used two exploratory approaches to begin investigating other ways to model biomass response to nutrients:

- BRT (Objective 3) and
- B-CART analysis (Objective 4).

4.2.2 Background on Benthic Biomass Spreadsheet Tool Development and Testing (Objective 1)

The BBST estimates algal density as AFDM (g/m²) and benthic chlorophyll *a* (mg/m²) using five methods: two versions of models by Dodds et al. (1997 and 2002), and three versions of the QUAL2K model: standard, revised, and revised with accrual (Chapra and Pelletier, 2003). Model set-up and initial testing are described in detail Tetra Tech (2006) and summarized here. Table 4.1 summarizes the models by type, input parameters, and major differences. All five BBST models predict chlorophyll *a* and AFDM. Total nitrogen (TN) and phosphorus (TP) are the base input variables for both Dodds and QUAL2K models, though additional variables, such as canopy closure, water temperature, and stream depth, are included as secondary input variables in the QUAL2K models.

Dodds 1997 and 2002 Models

The Dodds models (1997, 2002) are statistical log-log regression models of the mean and maximum²⁰ values of chlorophyll *a* as a function of stream TN and TP concentrations from field monitoring data (Eqs. 1 and 2). The Dodds (1997) model was developed for wadeable streams in temperate climates, using a compilation of data from the Clark Fork River, Montana, and 205 sites throughout North America and New Zealand. In the BBST, AFDM is calculated by dividing the chlorophyll *a* values by a constant (2.5; Tetra Tech 2006). The coefficient of determination (R^2) was 0.43 for the mean seasonal chlorophyll *a*, and 0.35 for maximum seasonal chlorophyll *a* (Chl *a*).

$$\log(\text{mean Chl } a) = -3.223 + 2.826 \log(\text{TN}) - 0.431(\log [\text{TN}])^2 + 0.254 \log(\text{TP}) \quad \text{Eq(1)}$$

$$\log(\text{max Chl } a) = -2.702 + 2.785 \log(\text{TN}) - 0.433(\log [\text{TN}])^2 + 0.305 \log(\text{TP}) \quad \text{Eq(2)}$$

²⁰ In the work of Dodds et al. (2002), “maximum” appears to be intended to represent the spatially-averaged, temporal maximum algal growth potential (in response to nutrient and light availability) in the absence of temporary reductions in biomass density due to grazing, scour, and other factors. It is thus intended to be a temporal maximum, identified via multiple samples taken over the growing season.

Table 4.1. Summary of types of models contained in BBST.

Model	Type	Input Parameters	Comment
Dodds 1997	Log-log polynomial Regression	TN, TP	Dodds 1997 is a second order, log-log regression relationship between TN, TP and chlorophyll <i>a</i> . It was developed for wadeable streams in <i>temperate</i> climates, using a compilation of data from the Clark Fork River, Montana, and 205 sites throughout North America and New Zealand (for a total of 300 sites).
Dodds 2002	Log-log linear Regression	TN, TP	Dodds 2002 is a first order log-log regression relationship between TN, TP and chlorophyll <i>a</i> . In addition to the 1997 dataset, the 2002 version included additional data from the USGS National Stream Water Quality Monitoring Network stream data (972 sites from two datasets).
QUAL2K (standard)	Simulation Model	Inorganic nutrients, Stream Depth, Stream Velocity, Canopy Closure, Unshaded Solar Radiation	River and Stream Water Quality Model (QUAL2K) standard version is a parametric representation of the inorganic nutrient constituents, and physical parameters such as light, temperature and uses default model parameters.
QUAL2K (revised)	Simulation Model	TN, TP, Stream Depth, Stream Velocity, Canopy Closure, Unshaded Solar Radiation	In the QUAL2K revised version, the default kinetic parameters for benthic algae were adjusted to the Dodds (2002) results for a better fit for application in California. A nutrient availability fraction was also added.
QUAL2K (accrual)	Simulation Model	TN, TP, Stream Depth, Stream Velocity, Canopy Closure, Unshaded Solar Radiation, Days of Accrual	Days of accrual, which accounts for the scouring effect of rain events on algal biomass, was incorporated using Biggs (2000) regression coefficients into the revised QUAL2K model.

In the revised model based on Dodds et al. (2002) the regression equation was changed **to a first order log-log linear relationship** and included additional data from the USGS National Stream Water Quality Monitoring Network stream data. In addition to the nutrient concentrations, the effect of stream gradient, water temperature, and latitude was also examined, but not included in the linear regression equation (Eqs. 3- 4)

$$\log(\text{mean Chl } a) = 0.155 + 0.236 \log(\text{TN}) + 0.443\log(\text{TP}) \quad \text{Eq (3)}$$

$$\log(\text{max Chl } a) = 0.714 + 0.372 \log(\text{TN}) + 0.223\log(\text{TP}) \quad \text{Eq (4)}$$

USER INPUTS

Site:
Analyst:
Date:

Nutrient Concentrations (mg/L)

	Average	Minimum	Maximum
Ammonia-N	0.03	0.02	0.05
Nitrite-N	0.001	0.001	0.001
Nitrate-N	0.14	0.05	0.2
Organic N	0.32	0.1	1
Total N (calc)	0.491	0.171	1.251
Inorganic P	0.025	0.003	0.05
Organic P	0.0036	0.002	0.01
Total P (calc)	0.0286	0.005	0.06

Unshaded Solar Radiation (cal/cm²/d)

	Average	Minimum	Maximum
Enter manually	572	445	649

Estimate Latitude: 37.00 Month Range: Jun - Sep

Stream Inputs

Stream Depth (m)	1
Stream Velocity (m/s)	0.3
Water Temperature (°C)	20.0
Days of Accrual (optional)	120
Canopy Closure	<input type="checkbox"/>
f	0.9
Closure (%)	30
Light Extinction Coeff. (1/m)	0.5

Calculate

Method & Target Selection

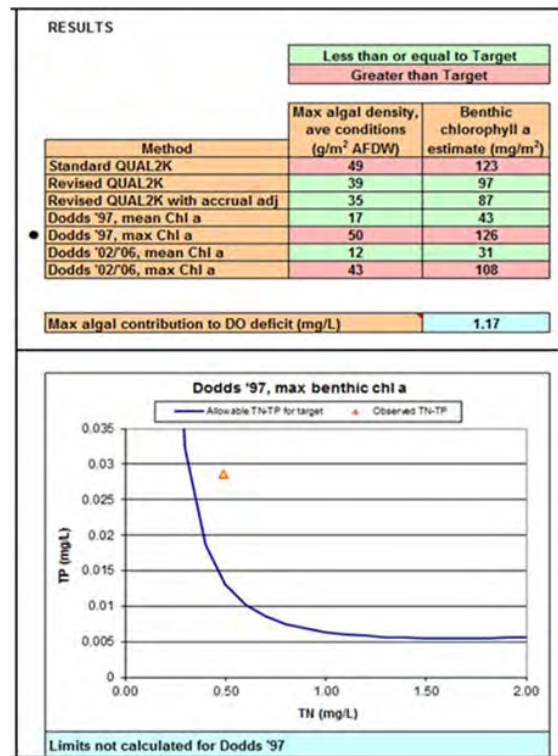
Select Method: Dodds '97, max benthic chl a

Target Max Benthic Chl a (mg/m²): 100

Corresponding Algal Density (g/m² AFDW): 40

California Benthic Biomass Tool, v14a (July 2012)

a) Input Panel of Dodds 1997 model



b) Output Panel of Dodds 1997 model

Figure 4.1. Example of user interface for the BBST (example highlights output plot for the Dodds 1997 version of the model) with input and output panels. The user inputs nutrient and stream data on the left panel, and the max algal density and benthic chlorophyll *a* values are estimated on the output panel on the right side. The output panel also shows the allowable TN and TP plot for a given site for a user-selected model. Note that only the TN/TP inputs are required for estimates based on Dodds' models, whereas the other nutrient types and the environmental data are required for the QUAL2k estimates.

Figure 4.1 shows the user interface screen for the Dodds et al. (1997 and 2002) models in the BBST. The user enters ammonia, nitrate, nitrite, organic nitrogen, phosphate, and organic phosphorus concentrations in the input panel (Figure 4.1a), and the maximum (and/or mean) algal density and chlorophyll *a* are predicted on the output panel (Figure 4.1b) along with an allowable TN, TP plot. The plot shows a threshold above which the combination of TN and TP is estimated to result in exceedance of a user-stated biomass target. The observed TN and TP values are plotted on the graph as a triangle to allow the user to visualize whether, and to what extent, existing nutrient conditions could lead to an exceedance of the biomass target. Additional entries on the input panel, such as canopy closure, do not come into play for the Dodds versions of the model.

QUAL2K Models

Versions of the River and Stream Water Quality Model (QUAL2K) in the BBST are a parametric representation of the benthic algal component of the mechanistic steady state model developed by Chapra and Pelletier (2003). This simple parametric representation was adapted to provide initial estimates of benthic algal responses to availability of light and nutrients, and can be adjusted to achieve general agreement with the empirical relationships developed by Dodds et al. (1997, 2002).

The model calculates the steady state algal growth as

$$B = \frac{K_{pmax} \cdot \Phi_{Nb} \cdot \Phi_{Lb}}{K_{rb} + K_{db}} \quad \text{Eq (5)}$$

Where, K_{pmax} is the maximum photosynthetic rate at a reference temperature of 20°C, (Φ_{Nb}) is the benthic algae nutrient attenuation factor represented by the Michaelis-Menten nutrient limitation equation for inorganic nitrogen and phosphorus. Φ_{Lb} represents a light limitation factor with a benthic algae light parameter, K_{rb} is the temperature-dependent benthic algae respiration rate, and K_{db} is the temperature-dependent benthic algae respiration rate. The prediction of biomass uses only the sum of respiration and death as a combined loss term, and the model is unable to distinguish the processes independently. Equations to estimate individual components of equation 5 are not provided in this report but can be found in the TetraTech 2006 report.

The standard QUAL2K model uses the default model parameters. The user provides ammonia, nitrite, nitrate as N, total Kjeldahl Nitrogen, phosphate as P, and total phosphorus in addition to hydrology information such as stream depth and velocity, and site-specific information such as solar radiation (Figure 4.2a). The maximum algal density and chlorophyll *a* is predicted on the output panel (Figure 4.2b) along with an allowable TN and TP plot.

In the revised QUAL2K model, the default parameters were optimized to achieve a better agreement between the Dodds 2002 equation for maximum chlorophyll *a* and steady-state QUAL2K predictions based on the Ecoregion 6 dataset (Tetrattech 2006). Thus the standard QUAL2K was optimized for wadeable streams in temperate climates.

The revised QUAL2K + ACCRUAL model accounts for the scouring effect of rain events on algal biomass, where the days of accrual is defined as the average time between flood events greater than 3 times the median flow in a stream (Biggs 2000). Flow volume is a useful surrogate for velocity, as changes in flow volume correlate with changes in velocity. Sudden increases in velocity (e.g., by a factor of two to three) can result in the scour of algae adapted to a constant velocity. A simple statistical representation for the effects of the hydraulic regime on the biomass was created based on analysis of the mean number of days available for biomass accrual (Biggs 2000). The best fit regression for maximum monthly density of benthic algal biomass (mg/m^2 chlorophyll *a*) included days of accrual and soluble inorganic nitrogen (SIN) concentration, although the coefficients for days of accrual are similar for regressions using accrual only and using accrual and soluble reactive phosphorus.

USER INPUTS

Site:

Analyst:

Date:

Nutrient Concentrations (mg/L)

	Average	Minimum	Maximum
Ammonia-N	0.03	0.02	0.05
Nitrite-N	0.001	0.001	0.001
Nitrate-N	0.14	0.05	0.2
Organic N	0.32	0.1	1
Total N (calc)	0.491	0.171	1.251
Inorganic P	0.025	0.003	0.05
Organic P	0.0036	0.002	0.01
Total P (calc)	0.0286	0.005	0.06

Unshaded Solar Radiation (cal/cm²/d)

	Average	Minimum	Maximum
Enter manually	572	445	649

Estimate Latitude: Month Range: -

Stream Inputs

Stream Depth (m)	1
Stream Velocity (m/s)	0.3
Water Temperature (°C)	20.0
Days of Accrual (optional)	120
Canopy Closure	<input type="text" value="0.9"/>
Closure (%)	30
Light Extinction Coeff. (1/m)	0.5

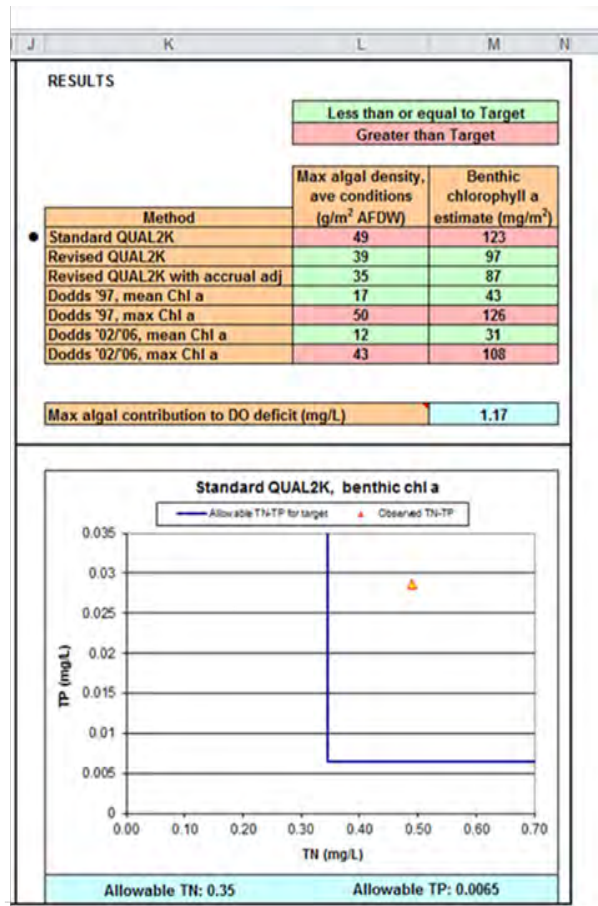
Calculate

Method & Target Selection

Select Method:

Target Benthic Chl a (mg/m ²)	100
Corresponding Algal Density (g/m ² AFDW)	40

California Benthic Biomass Tool, v14a (July 2012)



a) Input Panel of Standard QUAL2K model

b) Output Panel of Standard QUAL2K model

Figure 4.2. Example of user interface for BBST Standard QUAL2K model with input and output panels. The user inputs nutrient and stream data on the left panel, and the maximum algal density and benthic chlorophyll a values are estimated on the output panel on the right side. The output panel also shows the allowable TN and TP plot for a given site. Note that only the TN/TP inputs are required for estimates based on Dodds' models, whereas the other nutrient types and the environmental data are required for the QUAL2k estimates

Initial Model Testing

Through the phases of model development, the BBST was tested on two limited datasets from California: 1) data from 35 sites collected over 2000-2002 by the Lahontan Regional Water Quality Control Board (LRWQCB, N= 93) and 2) provisional data from the Western Stream Environmental Monitoring and Assessment Program (EMAP) dataset from 2000-2002 (n=103). Tetra Tech (2006) states that both datasets lacked critical information such as stream hydrology, light availability, and days of accrual at the time of their availability. Based on comparison of the LRWQCB 6 and EMAP data to Dodds, the report concluded that the equations are qualitatively reasonable for predicting mean and maximum potential growth of benthic algae in California streams in the absence of severe light or scour limitation. However, it was noted that the Dodds (2002) statistical relationships are quite weak, with R^2 values uniformly <0.5 . This was attributed to the fact that light and scour limitation play important roles in observed chlorophyll *a*. The report concludes that inclusion of average days of accrual and canopy closure might improve the results.

4.2.3 Data Sources

The California wadeable stream data, as described in Section 3.2.4, were used in the analyses presented in this chapter. These data represent one-time sampling events (as opposed to seasonal mean/max data) of stream algal response, water chemistry and other *in situ* variables taken during the time period of April through October throughout California. These instantaneous snap shots have an unquantified relationship with maximum or mean biomass. This variability is compounded by a large latitudinal variability in climate, rainfall, elevation, hydrology, geology, land use, and vegetation cover of wadeable streams throughout the state.

The dataset was used to explain the bias and variance in the predictions. We selected 52 predictor variables ranging from catchment geological composition to streambed attributes to meteorological data derived from existing data sources to examine potential factors contributing to model error (Table 3.2).

Estimating Days of Accrual

Biggs (2000) demonstrated that the predictive ability of regression equations could be improved (from an $R^2 <0.4$ to an $R^2 >0.7$) by inclusion of a measure of average days of accrual. The revised QUAL2K + ACCRUAL attempts to account for accrual based on the average time between flood events greater than 3 times the median flow in a stream. However, no guidance is given for how to estimate the days of accrual from ambient monitoring data consisting of one-time site visits. Given that velocity and flow are one-time measures at the sites in the wadeable stream data set, we developed a methodology to estimate the average days of accrual using the size of storm events as measured by daily precipitation data, and readily available parameters.

A daily precipitation database for all the study sites was developed by matching site data with daily precipitation from the Daily Global Historical Climatology Network (GHCND; NOAA National Climatic Data Center) and was spatially interpolated, using the nearest-neighbor method, during the period from 2007-2012 (Figure D.1). Raster files of daily precipitation data by site were generated for the period of 2007-2011.

We established cutoff values for scouring events as a function of storm size (daily precipitation) and extent of urbanization using best professional judgment (Table 4.2). The days of accrual were estimated for each site using a recursive algorithm that counts the number of days between the sampling date and an antecedent “scouring” event with a value equal to or higher than the cutoff. A nominal default of 120 days accrual was

used for any sites with no precipitation data, a time period roughly corresponding to the duration of the Mediterranean dry season.

Table 4.2. Stream scouring cutoff values for the watersheds developed based on precipitation and watershed imperviousness

Watershed Imperviousness (%)	Cutoff for Scouring Event, Precipitation (inches)
<5	0.5
5-25	0.4
25-50	0.3
>50	0.2

4.2.4 Validation of the NNE BBST Tool

The five underlying models of the NNE BBST tool: 1) Dodds 1997, 2) Dodds 2002, 3) QUAL2K 4) QUAL2k Revised, and 5) QUAL2k Revised + Accrual were validated against the observed benthic chlorophyll *a* data for 1031 sites.

Currently, the spreadsheet models are set up to conduct site-specific assessments. In order to conduct model runs in a more efficient manner, the BBST was recoded using R scripts, enabling batch runs for data from all sites. Model output from single runs using the original interface was checked against output of the R script to ensure accuracy of model translation. The models were validated by comparing predicted versus observed values, using a linear regression, and the performance was measured in terms of coefficient of determination (R^2) and slope.

4.2.5 Bias Analysis (Objective 2)

We conducted analysis of factors affecting bias using the randomForest package in R (Liaw and Wiener 2002). A dataset with predicted-minus-observed values for chlorophyll *a* and AFDM and the 52 selected explanatory variables was constructed. For the purpose of selecting the top predictor variables, any missing values of the explanatory variables were populated in the dataset using nearest-neighbor interpolation. After the preliminary predictor variable selection process, no interpolated values were used in the subsequent random forest regression analysis. Bias is the deviation of predicted values from the observed values (chlorophyll *a* for this study) resulting from usage of poor explanatory variables in the model or just incorrect choice of models. In the bias analysis process we try to examine the impact of a single or a group of explanatory variables on the prediction abilities of a given model.

Nonlinear multiple regression techniques in the randomForest package were used to determine the importance of the predictor variables. The strength and correlation of the predictor variables were estimated using the out-of-bag error method (OBEM). The error is estimated internally during the run. Each tree is constructed using a different bootstrap sample from the original data. In order to cross-validate, about one-third of the cases are left out of the bootstrap sample used for tree construction and are used to estimate the OBEM error. Each regression forest produces a variable importance plot based on the percent increase mean square error (MSE) for a given explanatory variable, and the total variance explained by the multivariate regression.

4.2.6 Exploratory Analyses Using Boosted Regression Trees (Objective 3) and Bayesian CART Analysis (Objective 4)

Two statistical methods were used to conduct exploratory analyses in order to suggest model refinements: BRT and B-CART analyses. Boosted regression trees allow nonlinear relationships and variable interactions to be represented in model predictions. Bayesian CART analyses provide a simplified set of regression models to predict algal biomass by site class, along with a set of classification rules to define groups.

BRT Analyses

BRT and partial Mantel tests were used to conduct exploratory analysis to investigate the relationship of nutrients to biomass response variables within a context of a large suite of environmental co-factors. See Chapter 3 for a detailed description of BRT and partial Mantel tests.

Bayesian CART

To further explore potential refinement of models explaining benthic algal response to nutrients, we applied a Bayesian Classification and Regression Tree (B-CART) analysis to the data matrix containing nutrients and potential co-factors influencing the response of benthic chlorophyll *a*. Bayesian CART is an approach to the development of Regression Trees that is informed by the analyst's prior knowledge of tree form and distribution of potential model coefficients. Unlike the Classification Tree analysis applied in Chapter 3, Regression Tree analysis is designed to optimize the fit of regression models within each final group rather than the difference in mean values among groups.

Regional variation in nutrient – Chlorophyll *a* relationships for lakes across the United States has been successfully explored using Bayesian Classification and Regression Tree (B-CART) analysis (Freeman et al. 2009). Regression Trees are designed to simultaneously classify observations and optimize the fit of regression models within each final class. Thus they provide a tool for assessing whether regionally specific model coefficients are appropriate.

We applied the CGMlidCART program for Bayesian CART analysis developed by Chipman et al. (available at: <http://www.rob-mcculloch.org/code/CART/index.html>) to data matrices containing \log_{10} chlorophyll *a* mg/m^2 biomass as the dependent variable and multiple explanatory variables. Because the CGMlidCART program cannot function with missing values, we substituted medians for missing values in the matrices²¹. Based on the protocol suggested by Chipman et al. (2002), we evaluated regression trees using a range of model parameters, chose the “most visited tree” among model iterations for each model, and used Aikake's Information Criterion (Burnham and Anderson 1998) to compare the fit of alternative models. See Appendix D for a detailed discussion of Chipman's fitting protocol.

Bayesian CART allows the user to choose both potential classifier variables and variables to include in regression models for end nodes. We tested alternative regression tree models based on selection of classification variable sets and selection of regression variable sets. Two types of regression trees were fitted to the data based on regression variables included, one set analogous to the modified Dodds model (“Dodds-type”), with $\log_{10}\text{TP}$, $(\log_{10}\text{TP})^2$, $\log_{10}\text{TN}$, $(\log_{10}\text{TN})^2$, \log_{10} (days accrual), and \log_{10} (days accrual)² as independent variables in the final regression models. The second-order term for total N allows the model to incorporate nutrient saturation effects at high levels. Although Dodds only included a second-order term for

²¹ This corresponded to 5 percent of values for variables retained in the final models.

total N, we added a second-order term for total P as well to reflect the possibility of P saturation at high nutrient levels. The second set of regression tree models were analogous to the QUAL2K model (“QUAL2K-type”), in which the steady state value for algal biomass, B, is predicted as:

$$B = (K_{pmax} \cdot \Theta_{Nb} \cdot \Theta_{Lb}) / (k_{rb} + k_{db})$$

Where B = steady state value for biomass

K_{pmax} = maximum photosynthetic rate in the absence of limiting factors

Θ_{Nb} = nutrient limitation term

Θ_{Lb} = light limitation term

k_{rb} = loss rate due to respiration

and k_{db} = loss rate due to death (e.g., grazing).

The equation can be linearized by applying a log transformation to both sides of this expression:

$$\log_{10} B = \log_{10} K_{pmax} + \log_{10} \Theta_{Nb} + \log_{10} \Theta_{Lb} - \log_{10} (k_{rb} + k_{db})$$

Temperature dependence is incorporated into QUAL2K with the Arrhenius relationship:

$$K_{pmax} = K_{pmax,20} \Theta^{(T-20)}$$

Log transformation of this term yields:

$$\log_{10} K_{pmax} + (T-20)\log_{10}\Theta$$

so temperature was used as a predictor without log-transformation.

The nutrient limitation term in QUAL2K is nonlinear because it is based on the Michaelis-Menten equation:

$$\Theta_{Nb} = \min\left(\left(\frac{n_a + n_n}{k_{sNb} + n_a + n_n}\right), \left(\frac{p_i}{k_{sPb} + p_i}\right)\right),$$

Where k_s = half-saturation constant.

To facilitate the application of the B-CART approach, which requires a linear model, we substituted the Dodds form of the nutrient limitation term, which allows for nutrient saturation at high N or P, and used the TN:TP or DIN:DIP ratios as potential classifiers to account for switching between N and P limitation.

The light limitation term in QUAL2KE is:

$$\Theta_{Lb} = (I_0 e^{-k_e H}) / (K_{Lb} + (I_0 e^{-k_e H}))$$

where I_0 = light incident on the water surface

K_e = light attenuation coefficient

H = water depth, and

K_{lb} = half-saturation coefficient

The term K_e should be proportional to turbidity, so an interaction term, Turbidity x Depth, was added in the regression tree. The incident light term was calculated as solar radiation x fraction cloud-free sky x (1 - fraction canopy). To account for the effects of potential light saturation, a second-order available light term (incident light²) also was incorporated into the regression tree model. Potential algal loss due to scouring was incorporated into the regression tree with the days of accrual and days of accrual² terms. Finally, each type of regression tree model was estimated with two alternate forms, one using total nutrients (TN, TP) and one using the dissolved inorganic forms (DIN = NH₄-N + NO_x).

We identified a reduced set of potential classification variables from the full potential set described in Table 3.2 supplemented by a few new interaction terms. In addition to the classification variables in Table 3.2, we considered interaction terms for turbidity x depth, stream power (watershed area x slope), stream power x antecedent precipitation, and stream power x antecedent precipitation x % sands and fines (an index of potential substrate disturbance). Details of the classification variable reduction process are provided in Appendix D.

The final reduced set of classification variables used to develop Regression Trees is listed in Table 4.3. Because of the redundancy of PSA ecoregion and geographical coordinates as classifiers, we constructed two final sets of alternative regression trees, one set using PSA ecoregions but not geographical coordinates as classifiers, and a second set excluding PSA ecoregions as classifiers but using geographical coordinates as classifiers to define “empirical nutrient regions”. We also compared trees with and without geographical coordinates as predictor variables in regressions to allow them to reflect the effect of smooth climatic gradients.

Final model selections were based on maximization of log-likelihood values when comparing models with equal numbers of predictor variables (regression coefficients and classifiers) or minimization of Aikake’s Information Criteria (AIC) values for comparison of disparate models (Burnham and Anderson 1998). Values of AIC can only be compared for model runs using equivalent training sets. We also calculated an r^2 value for regressions relating predicted to observed values to summarize the percent variation explained in each regression tree.

Table 4.3. Final set of classification variables used in B-CART analysis.

Classification Variable	Definition
Longitude	Degrees longitude
Latitude	Degrees latitude
NH ₄	Instream NH ₄ value (mg NH ₄ -N/L)
URB21_5K	Percent NLCD "Code 21" land use within a 5-km radius from sampling site
JulianDay	Day of year (1-365)
PSAc	Perennial Stream Assessment ecoregion (1-6)
Year	Year of sample
Conductivity	Instream conductivity
REFSITESTAT	Site disturbance status (Reference, Intermediate, Stressed) as defined in Chapter 2

Because B-CART forces the computation of regression coefficients for each final node of the tree regardless of whether or not they are significant, we ran separate regression analyses on each final node (or combined nodes where final node size was insufficient for model fitting). We conducted regression analyses in SAS using PROC GLMSELECT with the default stepwise selection method (SAS version 9.3, Copyright [©] 2002-2010, SAS Institute, Cary, NC). Final regressions were tested for model assumptions, i.e., normality of residuals (Wilk-Shapiro test) and homogeneity of variance (via plots of residuals versus predicted values), although these assumptions are not required for the original nonparametric B-CART analyses.

4.3 Results

4.3.1 Model Performance (Objective 1)

The validation data illustrate that the BBST has very poor model fits for all model types, although QUAL2K (standard, revised, and revised with accrual) performs marginally better than the Dodd's models for chlorophyll *a* (Tables 4.4-4.5; Figure 4.3). The models overpredict lower values and underpredict higher values for both chlorophyll *a* and AFDM, with slopes ranging from 0.1-0.55, and positive intercepts ranging from 30-122. Results of linear regressions are shown to illustrate the poor match between observed values and model predictions. However, because model fits were so poor, model assumptions for linear regressions were not met even after multiple standard transformations (log, square root, inverse, power) and higher order equations were applied. In general, both variance and residuals tended to decrease with the mean, and residuals were not normally distributed.

Table 4.4. Distribution of residuals of ≥ 0.5 standard deviation of mean for chlorophyll a by model type. Slopes were significantly different from 1 at p -value <0.05 .

0.5 Standard Deviation			
Chlorophyll a	Number of sites	Over Predict	Under Predict
Dodds 97 Mean	806	88	96
Dodds 97 Max	806	301	35
Dodds 02 Mean	806	91	83
Dodds 02 Max	806	284	26
QUAL2K Standard	868	234	62
QUAL2K Revised	846	264	27
QUAL2K Revised + ACCRUAL	846	213	38

Overall the model performance is comparable for all sites combined as well as for the Reference and Intermediate sites (Table 4.6), and it declines for Stressed sites. The model performance is comparable for the three QUAL2K models, even though the QUAL2K revised with accrual was expected to perform better than the other two versions due to customized accrual information. All models tend to overpredict the chlorophyll a and AFDM concentrations.

Table 4.5. Model performance (R^2 , slope and intercept) for all sites combined.

	Chlorophyll a						AFDM					
	Mean			Max			Mean			Max		
	R^2	Slope	Intercept	R^2	Slope	Intercept	R^2	Slope	Intercept	R^2	Slope	Intercept
Dodd 97	0.16	0.33	0.93	0.16	0.31	1.43	0.21	0.44	0.43	0.21	0.42	0.93
Dodd 02	0.20	0.32	0.99	0.18	0.22	1.72	0.21	0.40	0.53	0.20	0.27	1.28
QUAL2k Standard	0.15	0.33	1.23	NA	NA	NA	0.13	0.36	0.82	NA	NA	NA
QUAL2K revised	0.26	0.34	1.48	NA	NA	NA	0.25	0.40	1.04	NA	NA	NA
QUAL2K revised +Accrual	0.20	0.43	1.15	NA	NA	NA	0.21	0.54	0.64	NA	NA	NA

Table 4.6. Model performance (R^2) for all, Reference, Intermediate, and Stressed sites (see Chapter 2) for predicted mean chlorophyll a .

	R^2 (All Sites)	R^2 (Ref + Inter)	R^2 (Stressed)
Dodd's 97	0.16	0.11	0.04
Dodd's 02	0.20	0.15	0.07
QUAL2k Standard	0.15	0.11	0.03
QUAL2K revised	0.26	0.20	0.11
QUAL2K revised with accrual	0.20	0.13	0.10

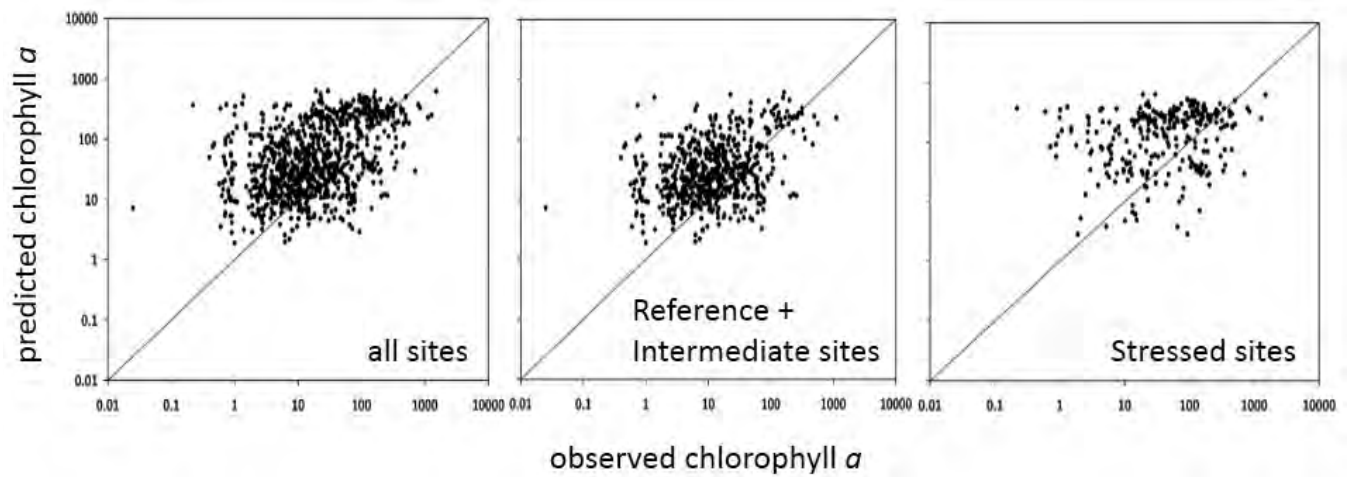


Figure 4.3. Sample plots of validation data showing measured versus predicted chlorophyll a by standard QUAL2K model, with 1:1 slope lines.

4.3.2 Random Forest Regression for Bias and Variance Analysis (Objective 2)

- The random forest regression ranks the explanatory variables that account for the variance and bias. These explanatory variables can be divided into three major categories: 1) water chemistry variables, 2) other site-specific parameters, such as physical habitat and 3) land use (Table 4.7). Recurring key explanatory variables from these categories are observed for both chlorophyll *a* and AFDM. The variance explained by the models ranges from 0-53%, signaling that model fit could be improved if refinements are made for how these variables are currently used in the models.
- Water chemistry variables, such as nutrients (TN, TP, NO_x, NH₄, and SRP), chloride, and conductivity were highly ranked in all our random forest models. Nutrients are consistently the most important explanatory variables, except for the cases in which Dodd's 97 mean AFDM models were being evaluated. Chloride, a reasonable surrogate for urbanization, was a key indicator for the chlorophyll *a* predictions. Conductivity is also a strong indicator of water quality associated with urbanization. A number of site-specific parameters, such as air and water temperature, canopy cover, solar radiation, reach slope, mean width of wetted channel, substratum composition (fines and particle size less than sand) are ranked high for all the models. Coarse particulate organic matter is a critical cofactor mostly for the AFDM regression models.
- Indicators of urbanization ranked highest as predictors in the Random Forest regression analyses for both chlorophyll *a* and AFDM. For some of the variables, such as road density and urban land use, the value measured within the 1 km radius of the site was important, rather than the watershed level value. In contrast, the watershed-level values for urban land use and Code 21²², as well as values of W1_Hall (an indicator of human disturbance that is local to the sampling reach) were important co-factors for the Dodds models.
- Note that explanatory variable rankings are qualitative rather than quantitative. For example, in the Dodds 97 mean chlorophyll *a* regression (Figure 4.4a), the reach slope was ranked as the most important variable, with 17% MSE, and water temperature also ranked in the top 15. However, the MSE of water temperature was negative, implying zero influence. Some of the regressions performed poorly for all variables. For example, for the QUAL2K with ACCRUAL model (Figure 4.4c), regression analysis showed no significant relationship between the predicted-minus-observed chlorophyll *a* and the explanatory variables, and the variable importance ranks also showed insignificant MSE change. The influence of the explanatory variables was stronger for the max Dodds models compared to the mean Dodds predictions for both chlorophyll *a* (variance for mean = 0.18, and max = 0.53, for Dodds 02) and AFDM (variance for mean = 0, and max = 0.46, for Dodds 02). The QUAL2K with accrual had the lowest variance and the weakest relationships between the predicted–observed biomass and the explanatory variables. It should be noted that the % Increase MSE values are not comparable between the models and only have meaning when making comparisons between variables within a given model.

²² “Code 21” encompasses a wide range of land uses primarily characterized by heavily managed vegetation (e.g., low-density residential development, parks, golf courses, highway medians).

Table 4.7. Variables ranked according to importance for random forest regression for Chlorophyll *a* (Chl *a*) and AFDM by model type.

Standard Qual 2k		Revised Qual 2k		Qual2k with accrual		Dodds 97 mean		Dodds 02 mean		Dodds 97 max		Dodds 02 max	
Chl <i>a</i>	AFDM	Chl <i>a</i>	AFDM	Chl <i>a</i>	AFDM	Chl <i>a</i>	AFDM	Chl <i>a</i>	AFDM	Chl <i>a</i>	AFDM	Chl <i>a</i>	AFDM
Urban LU (1 km)	NO _x	TN	Substratum (fine)	Code 21 (1 km)	Substratum (fine)	Reach Slope	Precip (3 mo)	Ag LU (1 km)	Urban LU (WS)	Ag LU (1 km)	TP	TP	TP
Road Dens (1 km)	Substratum (fine)	Urban LU (1 km)	TN	Chloride	TN	NO _x	Code 21	Air Temp (3 mo)	Precip (3 mo)	CPOM	SRP	TN	TN
TN	TN	TP	Water Temp	Substratum (>sand)	Air Temp (same mo)	TN	Sub-stratum (fine)	Canopy Cover	Substratum (fine)	Air Temp (3 mo)	Substratum (fine)	SRP	NO _x
TKN	CPOM	Road Den (1 km)	Urban LU (1 km)	Alkalinity	Geology Cenoz	Code 21 (WS)	Urban LU (WS)	CPOM	Code 21	Wet Chnl Width	TN	Alkalinity	Substratum (fine)
TP	Cond	NO _x	Air Temp	TKN	Wet Chnl Width	TP	Turbidity	Alkalinity	Turbidity	Canopy Cover	Cond	Substratum (<sand)	SRP
Chloride	Chloride	Substratum (<sand)	TP	Wet Chnl Width	Geology (Quat)	Catchment Slope (1 km)	Solar Rad	Wet Chnl Width	Catchment Slope (1 km)	Alkalinty	Solar Rad	Alkalinity	Reach Slope
NO _x	SRP	Urban LU (WS)	Cond	TP	Road Dens (WS)	Chloride	Road Dens (WS)	Substratum (<sand)	Road Dens (WS)	Urban LU (1 km)	Reach Slope	Cond	Cond
Alkalinity	TP	SRP	SRP	CPOM	Water Depth	Alkalinity	Catchment Slope (1 km)	Flow	Light Extinction	Geology (Quat)	Site Elev		Urban LU (1Km)
SRP	Water Temp	Chloride	NO _x	Catchment Slope (1 km)	Turbidity	Solar Radiation	Light Extinction	Ag LU (WS)	Turbidity	Urban LU (WS)	Road Dens (1 km)		CPOM
NH ₄	Air Temp	Alkalinity	Road Dens (1 km)	Urban LU (1 km)	Precip (3 mo)	SRP	Ag LU (WS)	Geology (Quat)	Air Temp	NH ₄	NO _x		Chloride
Cond	Canopy Cover	Cond	Solar Rad	Cond	Area	Substratum (<sand)	Turbidity	Chloride	Water Depth	Road Dens (1 km)	Substratum (<sand)		Solar Rad
Substratum (<sand)	Wet Chnl Width	Air Temp	Canopy Cover	Road Dens (1 km)	Solar Rad	Cond	Sample Site Elev	Area	Solar Rad	Substratum (<sand)	CPOM		Substratum (>sand)
Reach Slope	Area	Water Temp	Wet Chnl Width	Area	Geology (ign/met)	W1 HALL	TKN	NH ₄	Geology (Quat)	Code 21 (1 km)	Code 21		Wet Chnl Width
Air Temp	Flow	Area	Flow	Wet Chnl Width	Urban LU (WS)	Water Temp	Ag LU (WS)	Precip (3 mo)	Accrual	Solar Rad	Canopy Cover		Canopy Cover

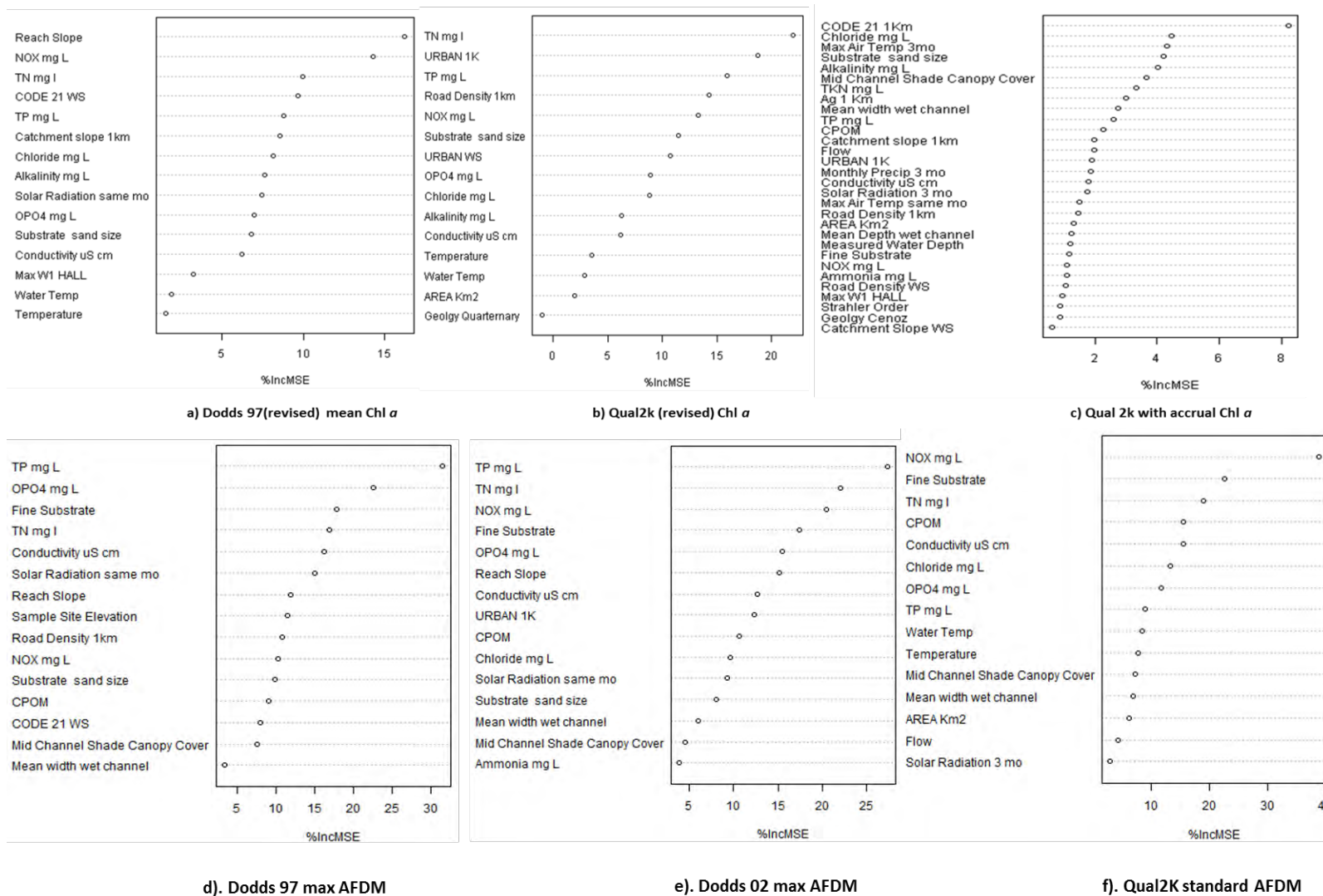


Figure 4.4. Relative influence of variables for some selected models for chlorophyll a (top row: a, b, c) and AFDM (bottom row: d, e, f). All other model output is provided in Appendices D3 - D8. Predictor variables ranked on the Y-axis, and the mean squared error values are listed on the x-axis. Variable names are given in Table 3.2.

4.3.3 Results of BRT Analyses (Objective 3)

Results from the BRT analyses examining nutrient and other environmental co-factor effects on the six biomass response variables are summarized in Table 4.8 and in the heat map in Figure 4.5. The relative influence of all predictor variables for the six models are provided in Figures D.3 – D.8. Wide differences were observed among biomass types in terms of what environmental co-factors most strongly predicted biomass levels.

The final BRT models had numbers of predictors ranging from 27 to 33. In no case was a nutrient the top predictor for any given biomass type, however for AFDM, NH_4 was the second highest-ranked predictor, with a relative influence of nearly 8%, and for chlorophyll *a*, NO_x was the fifth highest-ranked predictor, with a relative influence of nearly 6%. Partial Mantel tests (Table 4.9) indicated that nitrogen correlated significantly with both of these biomass variables when other high-ranking predictors from the BRT models, as well as spatial autocorrelation, were accounted for. In the case of chlorophyll *a*, SRP was also significant, although the Mantel partial correlation coefficient was very low. Nutrient predictors collectively had a low relative influence on the percent-cover biomass types (i.e., PCT_MAP, PCT_MCP, and PCT_MIAT1), and no nutrients were significantly correlated with percent cover metrics based on the partial Mantel tests.

Of the six biomass types tested in this study, chlorophyll *a* was the most directly responsive to nutrients, based on BRT analysis, with a total of >16% of the relative influence on chlorophyll *a* attributable to nutrient concentrations, amid the 20 other environmental co-factors (physical habitat, meteorological, landscape, water chemistry; see Table 4.7) included in the models (Figure 4.5). The biomass variable that was least responsive to nutrients was macrophyte percent cover (PCT_MCP), for which <6% relative influence was attributed to nutrients (Table 4.8). Among the nutrients, nitrogen species were invariably associated with a higher degree of influence on biomass than phosphorus species, but the overall relative influence of nitrogen vs. phosphorus varied by biomass type. The difference between the two was most dramatic for chlorophyll *a*, with nitrogen species collectively accounting for 3 times the relative influence realized for phosphorus species.

Of the non-nutrient predictors, stream temperature was the most likely to influence biomass; it was the top-ranked predictor for both chlorophyll *a* and percent presence of thick (1mm+) microalgae (PCT_MIAT1; Appendix D1), exhibiting almost 10% relative influence in both cases. The top-ranked predictors for the other biomass response variables included the substratum-specific percent fines and percent sand + fines, as well as percent canopy cover and conductivity.

Table 4.8 Relative influence of nutrient species on abundance of stream biomass of six different types, from BRT models that included environmental co-factors (see Figure 4.5 for a full list of predictors in each final model). “Model cv correlation (se)” refers to the cross-validation correlation coefficient (with standard error), indicating reliability of each model (Elith et al. 2008). PCT_MAP is macroalgal percent cover; PCT_MCP is macrophyte percent cover; PCT_MIAT1 is percent presence of thick (1mm+) microalgae. Dashes indicate that the predictor in question was not included in the final model for that biomass type. Bold values correspond to the highest ranked nutrient predictor for that biomass type.

Biomass Type	Highest Ranked Predictor (relative influence)	Model CV Correlation (se)	# Trees	# Predictors in Final Model	Relative Influence of (rank)				
					TN	NO _x	NH ₄	TP	SRP
AFDM (N = 847)	finest (%) (11.91)	0.628 (0.027)	6000*	30	1.38 (24)	-	7.70 (2)	-	3.83 (11)
Chlorophyll <i>a</i> (N = 878)	stream temperature (9.85)	0.503 (0.051)	6000*	25	3.78 (10)	5.97 (5)	2.83 (14)	-	4.05 (9)
PCT_MAP (N = 771)	canopy cover (%) (13.4)	0.643 (0.022)	8250	33	2.73 (11)	3.17 (7)	2.25 (16)	1.49 (23)	1.46 (24)
PCT_MCP (N = 771)	sand & fines (%) (17.43)	0.680 (0.025)	7200	33	1.80 (17)	1.16 (25)	1.45 (21)	0.72 (33)	0.77 (32)
PCT_MIAT1 (N = 770)	stream temperature (9.23)	0.458 (0.03)	5550	32	2.66 (17)	2.85 (15)	3.07 (14)	3.40 (12)	3.55 (11)
soft algal total biovolume (N = 914)	conductivity (14.81)	0.599 (0.024)	5000	27	1.51 (26)	3.21 (9)	2.00 (22)	3.31 (8)	2.08 (21)

*For these models, tree number was not optimized. A fixed number of 6,000 trees was used.

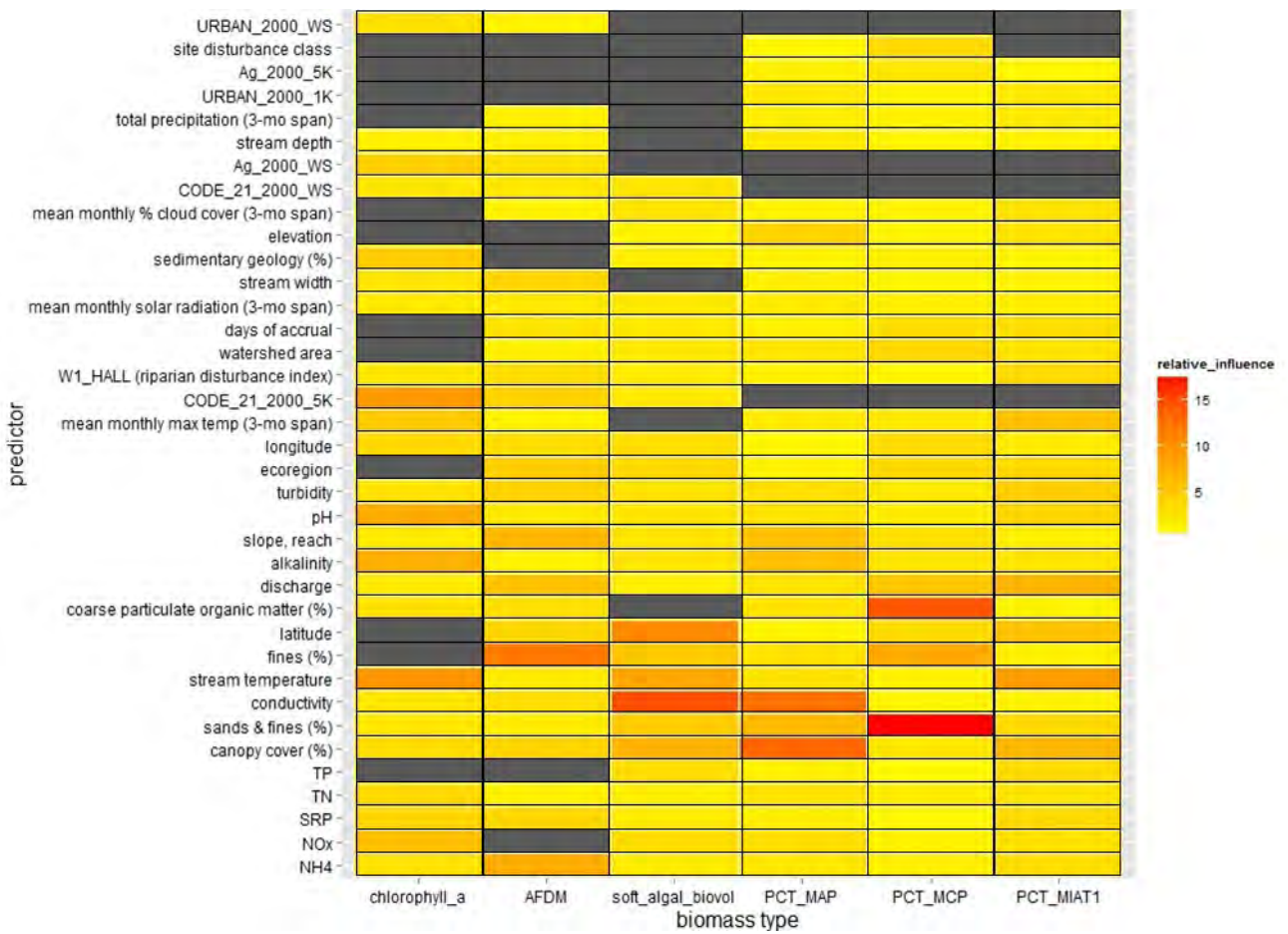


Figure 4.5. Heat map showing relative influence (%) of predictor variables (nutrients and environmental co-factors) on biomass response variables, from six independent BRT models. Yellow = low influence, red = high. The five nutrient-based predictor variables are grouped at the bottom of the graph. All climate variables are based on data for the month in which the sample in question was collected, averaged with the prior two months. Grey boxes indicate that the corresponding predictor type was not included in the final BRT model for that biomass type.

Table 4.9. Partial Mantel coefficients (95% CIs) for correlation between nutrient predictors and biomass variables; p values. Grey boxes correspond to explanatory variables that were not included in the partial Mantel test for the ALI variables in question. “Space” refers to the geographic distance between sites (for testing the significance of spatial autocorrelation). “NS” = not significant; dashes correspond to predictors that were included as explanatory variables in the partial Mantel tests for the indicated biomass variables, but (because they did not fall under the categories of nutrients or “space”) were not the focal variable in the tests. Values in bold correspond to significant partial Mantel tests.

Explanatory Variable	Chlorophyll a	AFDM	Soft Algal Biovolume	PCT_MAP	PCT_MCP	PCT_MIAT1
TN					-0.036 (-0.057 – -0.011); NS	-0.041 (-0.062 – -0.020); NS
NO _x	0.156 (0.138 – 0.175); 0.001		0.014 (0.006 – 0.022); NS	0.020 (0.008 – 0.030); NS		0.041 (0.019 – 0.055); NS
NH ₄		0.083 (0.060 – 0.103); 0.001				0.045 (0.001 – 0.071); NS
TP			0.000 (-0.009 – 0.009); NS			-0.052 (-0.077 – -0.023); NS
SRP	0.041 (0.023 – 0.058); 0.019	0.005 (-0.008 – 0.018); NS				0.048 (0.025 – 0.074); NS
space	0.072 (0.054 – 0.090); 0.001	-0.009 (-0.027 – 0.006); NS	0.040 (0.029 – 0.050); 0.001	-0.004 (-0.015 – 0.004); NS	-0.019 (-0.037 – -0.006); NS	-0.064 (-0.082 – -0.046); NS
CODE_21_2000_5K	-	-				
stream temperature	-		-			-
alkalinity	-			-	-	
mean monthly max temp (3-mo span)	-				-	-
pH	-					-
Ag_2000_WS	-					
sedimentary geology (%)	-					
canopy cover (%)		-	-	-	-	-
latitude		-	-		-	-
finer (%)		-	-		-	
slope, reach		-		-	-	
discharge		-			-	-
turbidity		-			-	-
stream width		-				
sand & fines (%)			-	-	-	-
conductivity			-	-		

Table 4.9. (continued)

Explanatory Variable	Chlorophyll <i>a</i>	AFDM	Soft Algal Biovolume	PCT_MAP	PCT_MCP	PCT_MIAT1
elevation				-		
days of accrual					-	-
Ag_2000_5K					-	
watershed area					-	
longitude					-	
coarse particulate organic matter (%)					-	
W1_HALL (riparian disturbance index)						-
mean monthly % cloud cover (3-mo span)						-

Figure 4.6 shows the joint influence of NO_x and a critical environmental co-factor, temperature, on chlorophyll *a* levels. The latter had a particularly strong influence, corresponding to an abrupt rise in biomass response within the range of approximately 26-28°C. Both stream water temperature on the day of sampling and antecedent ambient air temperature were important determinants of chlorophyll *a* concentrations based on BRT models (Figure 4.5 and Appendix D2). Together, these two temperature measures accounted for 15% of the relative influence in predicting chlorophyll *a* (Appendix D1).

Of all predictors in the BRT model for AFDM, percent fine substrata had the highest relative influence (Table 4.8, Appendix D1). Percent canopy cover also exhibited a fairly high relative influence, however the relationship between this co-factor and AFDM was not monotonic; rather, very high (in particular) and very low, canopy cover values were the two states that corresponded to predictions of higher AFDM values (Figure 4.7).

Percent canopy cover was the most important predictor of percent cover of macroalgae (PCT_MAP; Table 4.8), accounting for over 13% relative influence. Unlike the case with AFDM, percent canopy cover had gradual, monotonic relationship with macroalgal cover, in which macroalgal percent cover decreased steadily with increasing canopy cover (except for levels >80% canopy cover, at which point PCT_MAP dropped off precipitously; Figure 4.8). The interaction between canopy cover and conductivity was significant in predicting macroalgal percent cover: high conductivity (>500 μS, and especially >2500 μS) combined with low canopy cover were conditions favoring high macroalgal percent cover.

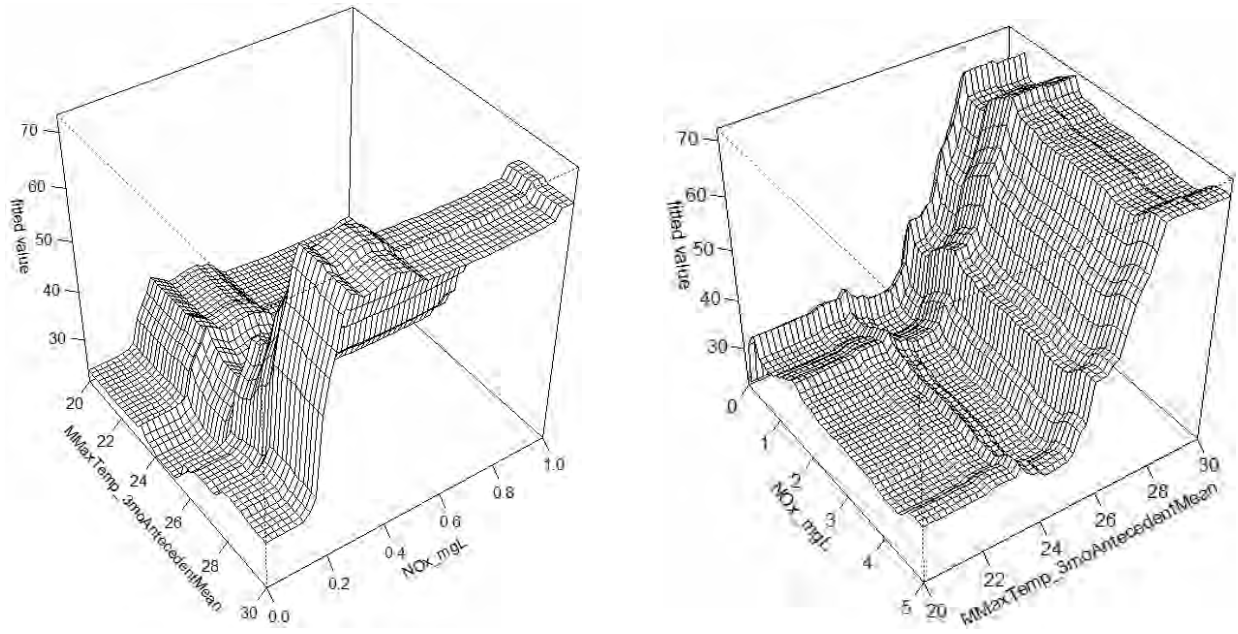


Figure 4.6. Three-dimensional plot (two views) of NO_x and mean monthly maximum ambient air temperature (the mean of the month the sample was collected and the two months prior) from BRT model for chlorophyll a.

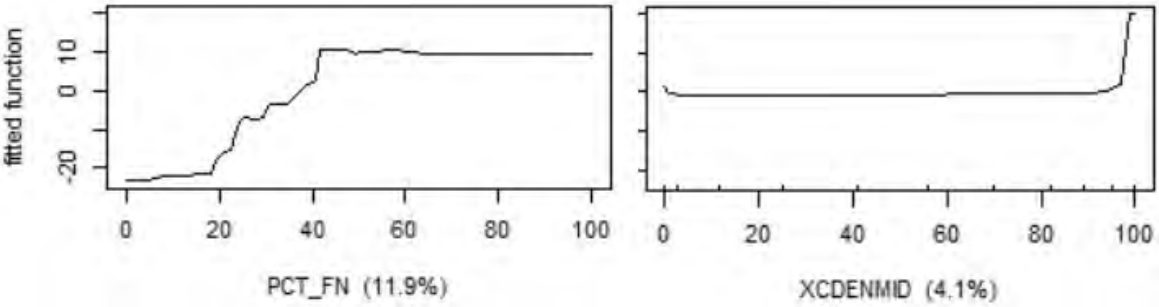


Figure 4.7. Partial dependence plots of percent fine substrata (left) and percent canopy cover (right) from the BRT model for AFDM. The y-axes correspond to the fitted AFDM variable. The values in parentheses are the relative influence of the variable indicated on the x-axis on the response variable. PCT_FN is percent fines; XC DENMID is percent canopy cover.

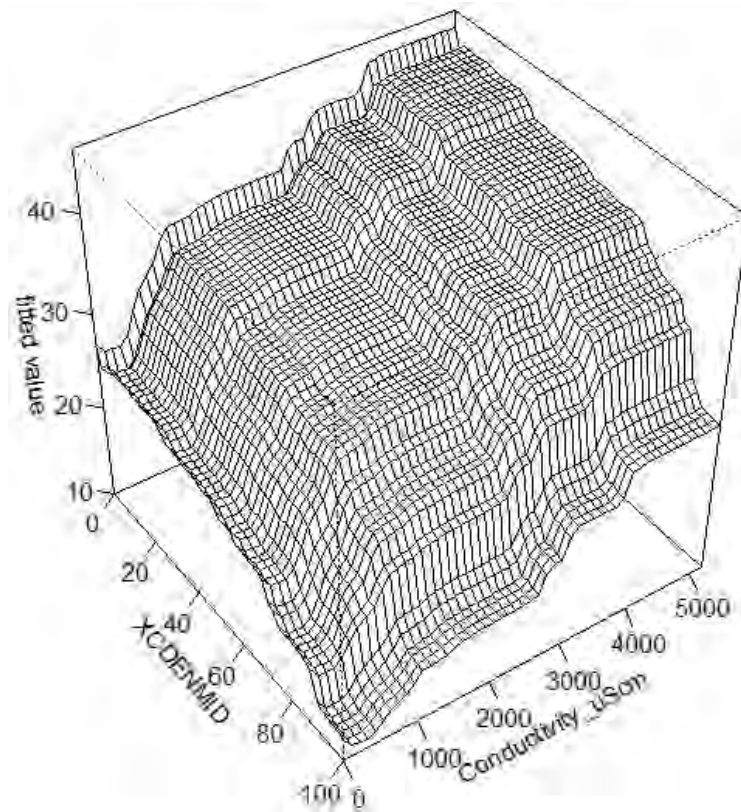


Figure 4.8. Three-dimensional plot of percent canopy cover (XC DENMID) and conductivity from a BRT model for PCT_MAP (percent macroalgal cover). The two predictors exhibited a significant interaction in their relationship to PCT_MAP.

A substantially different set of environmental co-factors came into play as key predictors for macrophyte percent cover (PCT_MCP) relative to what was observed for the other biomass variables (Figure 4.5). The highest-ranked predictor of PCT_MCP (with a relative influence of >17%) was percent sand + fines (Table 4.7), which had a significant interaction with several other predictors, such as days of accrual (Figure 4.9). Sites with percent sand + fines >40% (and especially >80%), with days of accrual exceeding approximately 100 days, had particularly high macrophyte percent cover.

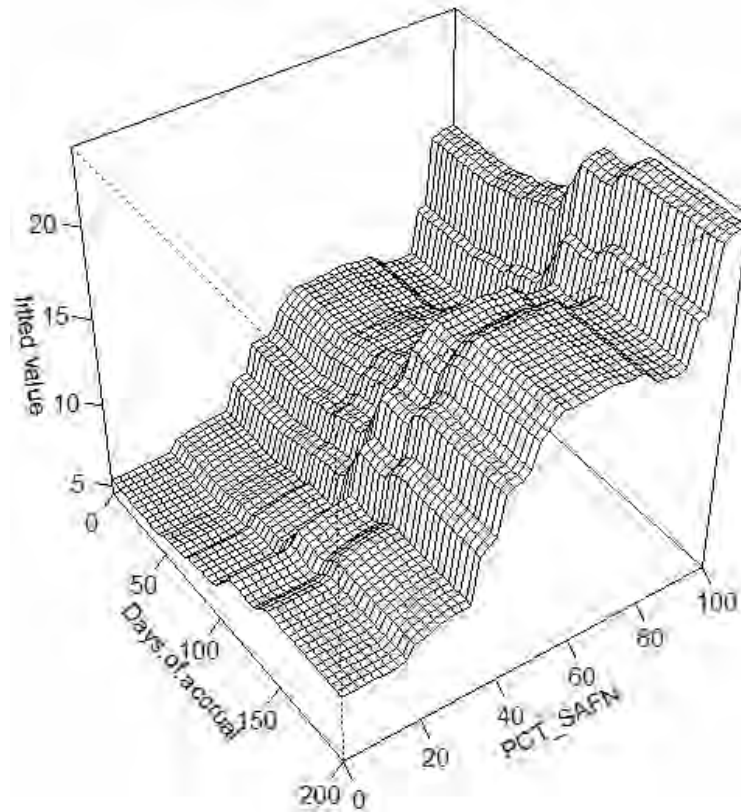


Figure 4.9. Three-dimensional plot of percent sand + fine substrata and days of accrual from BRT model for PCT_MCP (percent macrophyte cover).

4.3.4 Results of Bayesian CART Analyses (Objective 4)

Bayesian CART Trees

Bayesian CART analyses were run with a training set to fit the model, with independent model validation using a test set (10% of observations). The Bayesian CART analysis with a full set of classification variables yielded a relatively high explanatory power for the training set ($r^2 = 0.84$), with only a slightly lower value for the validation test set ($r^2 = 0.80$). Model fit based on AIC was even better for the Dodds-type DINDIP model (training $r^2 = 0.91$, test $r^2 = 0.88$; Table D.2). For the Bayesian CART analyses with reduced classification sets, and within the PSA ecoregion regression trees, the Dodds-type TNTP models outperformed the Dodds DINDIP models. In both cases, model fit improved significantly and substantially (from test set r^2 values of 0.23 or 0.44 to 0.63 or 0.81, respectively) when geographic coordinate predictors were included as continuous variables. For PSA ecoregion DINDIP regression trees, QUAL2K-type models performed slightly worse than the simpler Dodds-type models (Table D.4). For the best Dodds-type TNTP model, Julian Day, NH_4 , and local urbanization were included as final classifiers; PSA ecoregions were not. For the best QUAL2K-type DINDIP model, Julian Day and local urbanization were included as final classifiers.

Overall, based on AIC values, the empirical nutrient region regression trees performed better than the PSA ecoregion regression trees (Tables D.4, D.5). Again, for Dodds-type models, TNTP models outperformed DINDIP models, and model fits were significantly improved by the addition of geographical coordinate covariates. Some reversals of these trends are apparent in the test set r^2 values due to the presence of a few

outliers (Table D.5, Figure D.9). Again, the simpler Dodds-type models yielded a better fit than the QUAL2K-type models and explained about 10% more variation, although even the QUAL2K-type models performed much better than the original BBST models with the California data set. For the best Dodds-type TNTP model, latitude, longitude, Julian Day, and NH₄ were included as final classifiers, while for the best QUAL2K-type DINDIP model, latitude, longitude, Julian Day and local urbanization were included as final classifiers (Table 4.10).

Table 4.10. Results of B-CART analyses based on reduced sets of classification variables. Models used reduced set of four or five potential classification variables (PSA ecoregion (PSAcI) OR Latitude and Longitude, Julian Day, NH₄, and URB21_5K). Training set used is 1. Predictor variables for Dodds-type models included TN, TP, TN2, TP2, days accrual, and days accrual2 (TNTP models) or the same variables with DIN and DIP substituted for TN and TP (DINDIP models). Predictor variables for the QUAL2K-type models also included stream temperature, incident light, and a turbidity x water depth interaction). Models were also run with or without latitude and longitude as predictors. Model numbers are provided for a subset for ease of reference in the text. Model fits are assessed based on the Aikake’s Information Criteria (AIC) values and by r^2 values associated with plots of predicted versus observed values for log₁₀ chlorophyll *a* (mg/m²) for training and test sets.

Model No.	Model Type	Independent Variable Type	Lat/ Long included	Number independent regression variables	Log Likelihood	Most Visited Tree Size	AIC	Predicted vs. Observed r^2		Final Classification Variables		
								Training	Test			
Model 1	Dodds	TNTP	Yes	8	736.17	7	-1346.33	0.79	0.81	JulDay	NH ₄	URB21_5K
Model 2	QUAL2K	DINDIP	Yes	11	635.90	4	-1175.79	0.69	0.66	JulDay	URB21_5K	
Model 3	Dodds	TNTP	Yes	8	1249.05	23	-2084.1	0.92	0.57	Lat, Long	JulDay	NH ₄
Model 4	QUAL2K	DINDIP	Yes	11	1033.65	18	-1635.3	0.9	0.72	Lat, Long	JulDay	URB21_5K

The Regression Trees generated by the final B-CART analyses are illustrated in Figures 4.10-4.13. Figures 4.10 and 4.11 show the sequential splits of the original training set into different nodes, along with the classification variable and rule associated with each split. For example, Tree 1 generates five classes of sites after two are collapsed due to insufficient size. The first split separates off a small early spring class (n = 8) with Julian Day of sample less than 124. Nodes B and C have very low NH₄, with Node B representing spring values and C representing summer values. Node D and E represent sites with higher NH₄ and with low versus higher localized URB21_5K values (Figure 4.10). Tree 2 describes four classes of sites, a late summer set, a spring/midsummer non-Reference set, and two Reference sets (low localized URB21_5K development), one from the spring sampling period and the second from the summer sampling period (Figure 4.11).

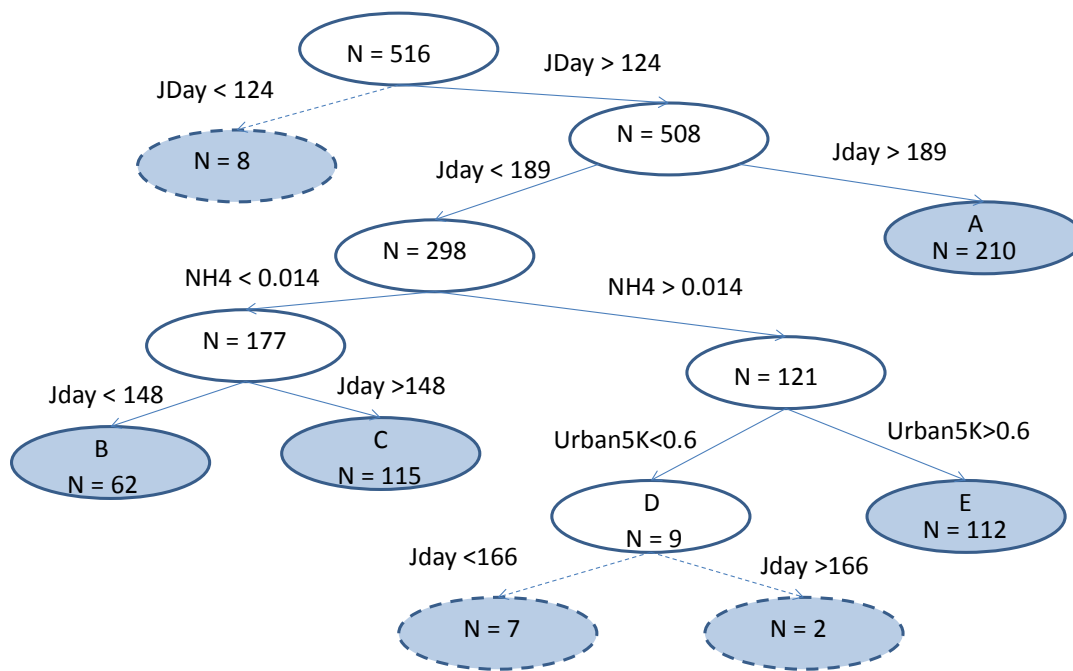


Figure 4.10. Bayesian regression tree for Model 1. Ovals represent nodes in tree and arrows correspond to classification rules. Filled-in ovals represent final nodes in tree. Dashed lines and borders are added to indicate nodes that should be collapsed due to inadequate final node size.

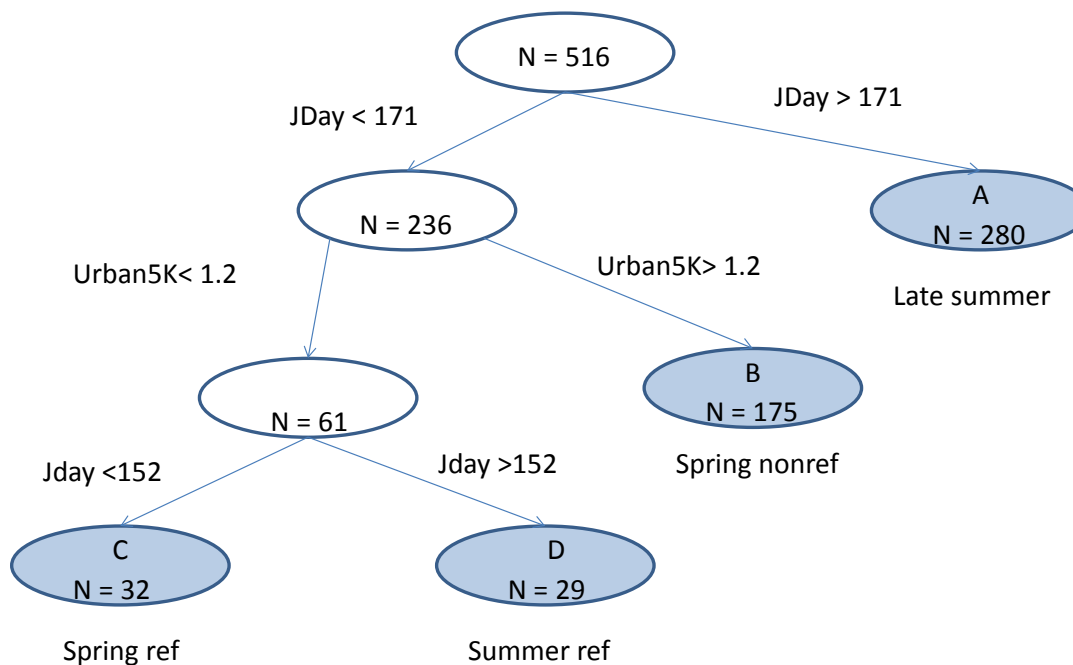


Figure 4.11. Bayesian regression tree for Model 2. Ovals represent nodes in tree and arrows correspond to classification rules. Filled-in ovals represent final nodes in tree. Dashed lines and borders are added to indicate nodes that should be collapsed due to inadequate final node size.

Structures for trees 3 and 4 were more complex but similar (Figures 4.12, 4.13). Because of the complex tree structure with a large number of end nodes defined by geographic coordinates, the nodes for Trees 3 and 4 are illustrated in map form, with the legend indicating nodes with classification rules based on nongeographic variables. The initial split in both trees represented separation of northern from southern sites. In Tree 3, southern sites were then distinguished on the basis of NH_4 levels, and then by a combination of season and geographic region. Northern nodes were further classified only by geographic region. In Tree 4, the southern nodes were differentiated by season (spring versus summer, spatial region, and then level of localized urban development (URB21_5K), while northern nodes were classified by geographic region.

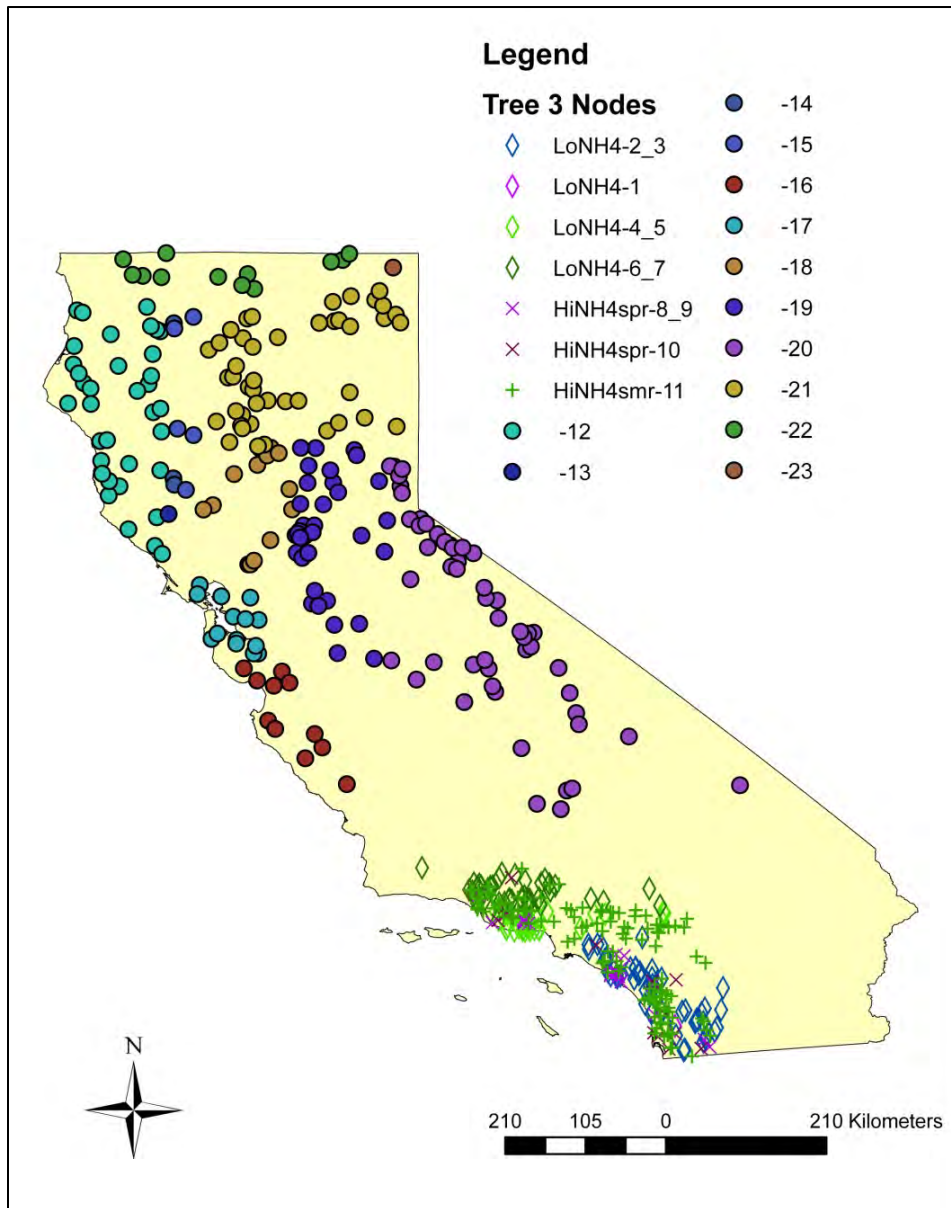


Figure 4.12. Location of sampling sites corresponding to nodes in Bayesian CART Model 3. Nodes not classified according to NH_4 level or season (solid circles) were classified solely on the basis of latitude and longitude. HiNH4 = High NH_4 , LoNH4 = Low NH_4 , spr = spring, smr = summer.

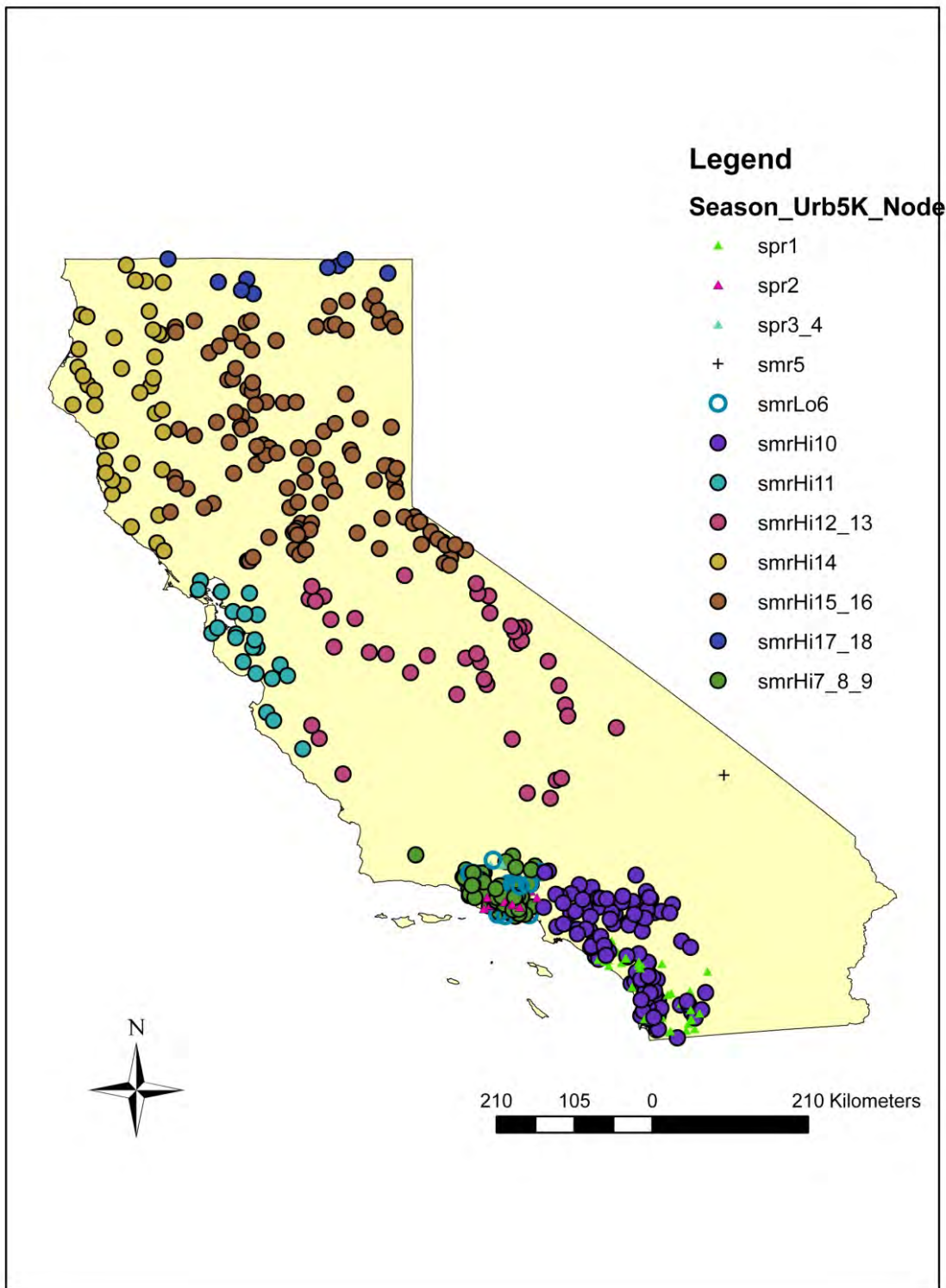


Figure 4.13. Sampling station locations corresponding to final nodes in Bayesian CART Model 4. Nodes were classified based on lat/long coordinates, Julian day (season), and level of NLCD Code 21 urbanization within 5K radius. spr = spring, smr = summer, Lo = Low URB21_5K, Hi = High URB21_5K.

Regression Analysis of Bayesian CART Nodes

Table 4.11 presents the results of the stepwise regressions performed on observations from the end nodes of the Regression Trees (after collapsing nodes of inadequate size). Final regressions were performed using both training and test set observations combined. Model assumptions for linear regressions (i.e., normality of residuals, homogeneity of variance) were not always met. These assumptions are not required for the original Bayesian CART analyses, but final regression models will need to be refined. For both Trees 1 and 2 (the Dodds-type and QUAL2K model types without lat/long predictors), days of accrual was retained as the best predictor of benthic algal biomass for samples collected in early spring, spring with low NH₄ values, or spring with low urban21_5k values (Table 4.11). However, the sign of the regression coefficient was negative, opposite of that predicted by Dodds models in which biomass is expected to accrue over time following a spate. For the two smaller nodes this effect is probably due to outliers. For nodes representing samples collected in mid to late summer either total or dissolved inorganic N were retained in regression models.

For both Trees 3 and 4 (the Dodds-type and QUAL2K model types with lat/long predictors), either latitude or longitude was retained as the best predictor in regression models for most end nodes. Longitude was always associated with a positive effect, while latitude effects varied within region. Total N or total P were selected as the primary explanatory variable in regressions for only a few regions.

Table 4.11. Variables retained in regression analyses to predict benthic biomass (log₁₀ chlorophyll a) based on Dodds-type models for nodes in B-CART models 1 and 3, and based on QUAL2K-type models for nodes in B-CART models 2 and 4. Nodes are numbered from left to right in B-CART trees in corresponding Figures 4.12 and 4.13. Regressions were fit using stepwise regression with a y-intercept (Int). Sign of regression coefficients is given inside parentheses following parameter. Node characteristics describe the classification rules producing each final group. – No regression results due to low class size.

Tree	Node	Size	Regression variables	adj r ²	Node Characteristics	
1	1 ²³	9	Int	log10days (-)	0.59	early spring
1	2 ^{24,24}	70	Int	log10days (-)	0.19	low NH ₄ spring
1	3	125	Int			low NH ₄ summer
1	4_5	10	Int			high NH ₄ late spring low urban21_5K
1	6	123	Int			high NH ₄ late spring high urban21_5K
1	7 ²⁵	236	Int	log10TN (+)	0.11	mid to late summer
2	1 ²⁵	34	Int	log10days (-)	0.58	low urban21_5K, early spring
2	2 ²⁵	35	Int			low urban21_5K, late spring
2	3 ^{24, 25}	191	Int	log10shadsolr (+)	0.06	spring, hi urban21_5K
2	4 ^{24, 25}	313	Int	log10DIN (+)	0.14	summer

²³ Residuals demonstrated heterogeneity of variance

²⁴ Residuals not distributed normally according to Wilk-Shapiro test (p <0.01)

Table 4.11 (continued)

Tree	Node	Size	Regression variables		adj r ²	Node Characteristics
3	1	9	Int			low NH ₄
3	2_3	9	Int			low NH ₄
3	4_5	43	Int	longitude (+)	0.7	low NH ₄
3	6_7	48	Int	longitude (+)	6	low NH ₄
3	8_9	21	Int	log10TP(-)	0.24	high NH ₄ spring
3	10	13	Int	logTN(+)	0.34	high NH ₄ spring
3	11	129	Int			high NH ₄ summer
3	12	38	Int	latitude (-)	0.72	
3	13	4	Int	log10TN (+)	0.74	
3	14	3	Int	log10days(+)	1	
3	15	5	Int	latitude (-)	0.76	
3	16	13	Int	latitude (-)	0.86	
3	17	16	Int			
3	18	15	Int	latitude (+)	0.79	
3	19	45	Int	latitude (-)	0.71	
3	20	55	Int	latitude (+)	0.48	
3	21 ²⁵	55	Int	longitude (+)	0.65	
3	22	15	Int	log10TP(+)	0.23	
3	23	1	--			
4	1 ^{24,25}	58	Int	latitude (-)	0.27	spring
4	2	30	Int	longitude (+)	0.49	spring
4	3_4	60	Int	longitude (+)	0.47	spring
4	5	2	--			summer
4	6	17	Int	latitude(-)	0.45	summer low Urb21_5K
4	7_8_9 ²⁴	69		longitude (-)	0.11	summer high Urb21_5K
4	10	99	Int	temperature (+)	0.07	summer high Urb21_5K
4	11	22	Int	latitude (-)	0.52	summer high Urb21_5K
4	12_13 ²⁴	40	Int	longitude (+)	0.26	summer high Urb21_5K
4	14	36	Int	latitude (-)	0.82	summer high Urb21_5K
4	15_16 ^{24,25}	129	Int	longitude (+)	0.52	summer high Urb21_5K
4	17_18	11	Int	longitude (+)	0.44	summer high Urb21_5K

4.4 Discussion

4.4.1 Validation Exercise Shows Considerable Room for Improvement in BBST

The BBST Dodds and QUAL2K models showed a very poor fit when validated against a statewide dataset of 1031 Wadeable Stream sites in California over 2007-2011. The perceived poor performance of the underlying BBST models is understandable for a variety of reasons, including difference in climate and hydrology between the data set supporting Dodds model development and the California Wadeable Stream dataset and differences in the spatial and temporal representativeness of modeled output versus the validation data set. In addition, algal primary production may be affected by other factors in the California Wadeable Stream dataset than those considered in the models. Finally, precision of observed benthic algal biomass data, as currently measured in California ambient monitoring programs, is uncertain. Data used to develop the original Dodds models were compiled from various sources reflecting different temporal intensity of sampling and collection or analytical methods, which are not necessarily comparable to the California stream data.

At the time of BBST development, California Wadeable Streams data were scarce and thus models were optimized for available national or international datasets. Fundamental differences in the factors controlling primary production between these national and California Wadeable Streams is an obvious reason for poor model performance. The California Wadeable Stream dataset is comprised largely of sites from a Mediterranean climate and perennial to intermittent flow regimes, while Dodds et al. (1997, 2002) models are derived from largely temperate, Wadeable Stream data. The BBST QUAL2K models (Tetra Tech 2006) were optimized to Dodds et al. (1997, 2002). In the application of their empirical model to a USGS data set, Dodds et al. (2002) report best fit at R^2 of 0.18, comparable to our findings. Though the Tetra Tech report (2006) suggests that the equations proposed by Dodds et al. (2002) are qualitatively reasonable for predicting mean and maximum potential growth of benthic algae in California streams in the absence of severe light or scour limitation, they also report low R^2 (~0.20) values for model validation of the RWQCB 6 data.

Another reason for poor model performance is a fundamental inequality in predicted biomass versus what is measured on both temporal and spatial scales. Conceptually, the BBST models are predicting algal abundance as spatially and seasonally averaged means or maximum values. A true validation of this model output is difficult, as large, geographically expansive Wadeable Stream datasets rarely have both good temporal and spatial (within streams) resolution. The California Wadeable Stream dataset reflects a one-time sampling of both explanatory variables and biomass responses, integrated over 150-m stream reaches, sampled over the growing season²⁵, across a 1000-mile range of latitude. Thus, these data are not likely to be representative of a spatially and temporally averaged “mean”, nor maximum, values. Furthermore, it is likely that site-specific factors acting on the expression of primary producer biomass include ones beyond those which are typically considered in eutrophication models. As such, recommended future work includes time-course sampling in streams to understand seasonal means and maxima, as well work toward better understanding the potential role of a wider-array of site-specific factors.

In addition, factors associated with urbanization had strong explanatory power for models’ lack of fit and we observed the poorest model performance in the Stressed sites. Approximately one-third of our data are from

²⁵ The index period for stream sampling for the validated data used here starts in May for drier parts of the state and June or July in colder/wetter parts of the state (depending upon stream flow conditions), and lasts for two to three months.

Stressed sites. Urbanization impacts a stream in numerous ways (Walsh et al. 2005, Booth et al. 2004), including increased scouring incidents. Stream channelization can also lead to increased water temperatures. Urban-derived toxicants have the potential to lower the biomass and might also explain the discrepancy in the model predictions. Application of the BBST may improve if applied on smaller spatial scales, where site-specific and landscape factors controlling eutrophication are more homogeneous (see Chapter 5). Biggs (1995) observed significant variation in the benthic algae population in streams based on land use and underlying geology. Other studies report stronger relationships (based on R^2 values) between the nutrients and biomass, along with secondary co-factors (such as days of accrual and watershed area) when applied to a homogeneous set of sites at a watershed scale (Biggs 2000, Van Nieuwenhuysse and Jones, 1996, Dodds 2006). The NNE-BBST performance improved when applied to the 270 Reference sites in our study ($R^2 \sim 0.2$) compared to aggregation of 1031 more heterogeneous sites. Finally, previous work by Fetscher (unpublished) suggests that, at least in streams supporting macroalgae and relatively high benthic algal biomass, precision of biomass estimates, based on data as currently collected for California ambient monitoring programs, is uncertain.

4.4.2 Inclusion of Landscape and Site-Specific Factors Provide Avenue for Model Refinement

Preliminary BRT and Bayesian CART analyses indicate that inclusion of landscape and site-specific factors into statistical stress-response models appeared to improve model fit over existing BBST Dodds et al. (1997, 2002) and QUAL2K models. Several landscape- and site-scale explanatory variables were high (relative to nutrients) in their relative influence in the variance analysis of the difference between observed and BBST-predicted biomass, and in the preliminary BRT models. This finding validates the fundamental NNE approach (i.e., that site-specific co-factors that vary across the California landscape can control algal response to nutrients, thus overriding a simple nutrient limitation on algal abundance). Some of these variables, such as a water temperature, canopy cover, and solar radiation are already included in the BBST QUAL2K, providing validation that the fundamental factors considered in the TetraTech (2006) modeling approach are relevant. Other explanatory variables not previously available for BBST modeling, such as ambient air temperature, "CODE 21" land use, alkalinity, sedimentary geology, solar radiation, and sediment percent fines, had a high level of relative influence in preliminary BRT models, though the importance of these variables varied among models predicting benthic chlorophyll a , AFDM and macroalgal percent cover (PCT_MAP).

Bayesian CART, a modeling approach that directly incorporates nutrients and/or mechanistic relationships into the model, also found improved fits with inclusion of variables representing geographic position (latitude and longitude). This suggests that model fit could be substantially improved by regionalizing coefficients. Both modeling approaches suggest that the strong influence of environmental gradients associated with latitude and longitude are not well-represented by the PSA ecoregions. Ecoregion was also among the predictors that exhibited somewhat surprisingly low relative influence on the biomass measures in BRT models (<5%). Bayesian CART models incorporating PSA ecoregion underperformed relative to those incorporating geographic position. This makes intuitive sense, as latitude and longitude can capture multi-factor gradients in temperature, precipitation regime, slope, and cloudiness/fog. A single ecoregion could contain both east- and west-facing slopes, such that east-west direction would be a poor predictor of monotonic environmental conditions within a single PSA ecoregion. For example, elevation (and temperature) could both increase and decrease with distance east.

Inclusion of explanatory variables that are integrative over time and space, in addition to instantaneous “snap-shots” taken in the field, may help to improve model performance. Examples of this are proxies for air temperature and available light, developed through GIS analysis. Elevation and distance from the coast (e.g., cover and fog) produce strong gradients in air temperature and available light that are only imperfectly captured in the QUAL2K model (which uses only latitude and sampling month range to calculate available solar radiation). Water temperature at time of sampling is typically employed in the model, yet water (and air) temperatures show a strong diurnal variation, so the time of field sampling could confound model application for regulatory purposes. In preliminary BRT models, both antecedent ambient air temperature and stream temperature on the day of sampling exhibited strong correlations with chlorophyll *a*, the levels of which increased dramatically over a relatively narrow range of antecedent ambient air temperature (approximately 26-28°C). Thus inclusion of landscape-scale monthly averaged air temperature variables may help to improve the prediction of temperature influence on primary production over use of a one-time measure of stream water temperature alone.

Other explanatory variables, such as “Code 21” land use and W1_Hall, indicators of development and riparian disturbance, respectively, are not typically included in mechanistic models of eutrophication (e.g., QUAL2k), yet they were identified as having a high relative influence in both BRT and B-CART models. These represent indirect effects and may actually represent a suite of stressors. A number of recent studies have reported a positive correlation between increase in urbanization and benthic biomass (Catford et al. 2007, Walsh et al. 2005). Cuffney et al. (2005) observed that basin-scale land use changes were the most important variables influencing the benthic response to urbanization. Rather than a single metric of urbanization, it is often the interaction among multiple impacts of urbanization that has the most significant influence on the benthic algal biomass (Taylor et al. 2004). A number of stressors such as hydromodification effects on hydroperiod and stream channel morphology and habitat type, as well as chemical contaminants, such as herbicides and heavy and trace metals, can affect algal abundance. These factors, and their interactions related to “Code 21” land use and W1_Hall, are difficult to model mechanistically.

Nutrient concentrations were important predictors in variance analysis of the difference between observed and BBST-predicted biomass, and in BRT models, albeit occupying less prominent roles than other factors. However, Bayesian CART models illustrated that inclusion of season when modeling the role of nutrients is important. Model results showed that total or dissolved inorganic N (and occasionally total P) was a better predictor of benthic algal biomass measured in the summer than in the spring. This could be because in the summer, when nutrient levels are associated with baseflow, grab samples are more likely to be representative of available nutrients than in the spring.

The Bayesian CART results suggest there are seasonal shifts in controlling factors, with days of accrual being a better predictor of benthic algal biomass for spring samples, and total or dissolved inorganic N (and occasionally total P) being a better predictor of benthic algal biomass measured in the summer. Biggs (2000) reported an improvement in the regression model from a R^2 of 0.40 to 0.74 with the inclusion of accrual information. The negative sign of regression coefficients we observed to be associated with days of accrual could have resulted either from an inaccurate specification of threshold discharges associated with scouring events or from the effect of scouring events on macroinvertebrate grazer populations. Low or intermediate-level events in spring could be associated with pulses of nutrient inputs but without sufficient power to remove existing algal biomass. Predicting site-specific scour based on land-use and historic meteorology data is challenging, and it is possible that we cannot currently estimate it using the available PSA data with

sufficient accuracy. Better information is needed on levels of discharge associated with scouring of algal biomass in California systems. In addition, numerous other factors could be at play that are unrelated to precipitation. For instance, anthropogenically originating controlled releases could be at play, but are difficult or impossible to model on a large scale. Other factors that are difficult to account for based on the existing dataset can (like scour) lead to the removal of algae and macrophytes from streams, thus reducing biomass. These include herbicides and algaecides, but we do not currently have these data on a large scale in order to include them in our models.

Among biomass types, chlorophyll *a* was the most responsive to nutrients in the BRT models. The higher responsiveness of chlorophyll *a* than AFDM to stream nutrient concentrations could be due to the fact that algal chlorophyll *a* production is by necessity stimulated by stream water-column nutrients, whereas stream AFDM can be subsidized by allochthonous material, thus weakening its connection to stream nutrient levels. However, as noted in Chapter 3, several findings suggested that AFDM may, in general, be a more meaningful predictor of ALI responses than chlorophyll *a*. Thus there may be value in assessing chlorophyll *a* and AFDM jointly in order to determine nutrient impacts to ALIs, as they represent two important components in the linkage of nutrient concentrations to ALIs: the former being more directly responsive to nutrients, and the latter apparently having more direct influence on ALIs.

In determining what factors belong in predictive models for biomass levels in response to nutrients and other environmental co-factors, it is important to consider the possible mechanisms behind observed relationships, such as those presented earlier in this section. For instance, percent fine substrata was the top-ranked “predictor” for AFDM biomass, and yet, it is possible that the strong relationship between this predictor and the response variable is not causative in nature, but rather the result of the fact that the organic component of fine bed material in a stream is, in itself, AFDM in the form of fine particulate organic matter (FPOM). Another observation that will be useful in further model development is the non-monotonic relationship between canopy cover and AFDM. At the low end of the canopy cover gradient, the somewhat elevated predicted AFDM is likely the result of increased sun exposure supporting instream primary production. But the high AFDM predicted at the high end of the canopy cover gradient cannot be due to the same phenomenon; rather it is most likely the result of allochthonous organic matter input to the stream from the canopy itself, which, following breakdown by shredders and weathering, would be included in the FPOM pool analyzed as part of AFDM. In both of these cases, anthropogenic nutrient loading would not be responsible for some fractions of the AFDM, and any predictive models must take this into account. It is important to keep in mind, however, that, while the results of the preliminary BRT modeling presented here offer insights into nutrient-biomass relationships, a more thorough approach to nutrient-algal abundance modeling is still needed in order to better refine and optimize predictive models for wadeable streams.

4.4.3 Summary of Validation and Recommendations for Refining Wadeable Stream Nutrient Algal Abundance Models

Our analyses indicate that the existing BBST models could be improved substantially, and existing data can be used to pursue refinements. The compiled dataset now includes a variety of explanatory variables that are available to begin a more thorough set of analyses. If algal abundance is among numeric endpoints utilized, then we recommend revising scoping models for wadeable streams, considering a full range of predictive statistical models.

US Environmental Protection Agency (2009) recommends using a variety of modeling approaches including regression (e.g., linear, logistic, quantile, and piecewise), change point analysis, and structural equation modeling to explore relationships between nutrients and algal abundance. More than one model categorized by classes, such as regions within the state, may be necessary in order to capture the range of nutrient-response relationships statewide. More complex mechanistic models could be considered over the long-term if the need to offer greater flexibility and applications to site-specific waterbody assessment are warranted.

Although the Bayesian CART trees incorporating latitude and longitude as both classification and predictor variables were the most accurate predictors of benthic algal biomass, in practice they may be too complex to be useful to managers. It is likely that these could be simplified by incorporating degree days (cumulative temperature effect) and distance from the coast as classification variables in place of latitude/longitude. The significance of temperature in predicting potential peak algal biomass is apparent in both BRT and Bayesian CART results. Temperature effects appear to be captured in classifier variables (e.g., season, latitude and longitude) but not as a continuous variable in final linear regressions. This could reflect the temperature optimum of filamentous green algal taxa that are responsible for the larger biomass accruals, as illustrated by the nonlinear interaction plots for NO_x and 3-month antecedent air temperature (Figure 4.6).

4.4 Literature Cited

- Biggs, B.J.F. 1995. The contribution of disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. *Freshwater Biology*, 33:419-438.
- Biggs, B.J.F. 2000. Eutrophication of streams and rivers: Dissolved nutrient-chlorophyll relationships for benthic algae. *Journal of the North American Benthological Society*, 19(1): 17-31.
- Booth, D.B., Karr J.R., Schauman S., Konrad C.P., Morley S.A., Larson M.G., and S.J. Burges. 2004. Reviving urban streams: Land use, hydrology, biology, and human behavior. *Journal of the American Water Resources Association*, 40:1351–1364.
- Burnham, K.P. and D.R. Anderson. 1998. *Model Selection and MultiModel Inference* (2nd ed.) Springer, New York, NY.
- Catford, J.A., Walsh C.J., and J. Beardall. 2007. Catchment urbanization increases benthic microalgal biomass in streams under controlled light conditions. *Aquatic Sciences*, 69(4), 511-522.
- Chapra, S. and G. Pelletier. 2003. *QUAL2K: A Modeling Framework for Simulation River and Stream Water Quality: Documentation and User's Manual*. Civil and Environmental Engineering Dept., Tufts University, Medford, MA.
- Chipman, H.A., George E.I., and R.E. McCulloch. 1998. Bayesian CART model search. *Journal of the American Statistical Association*, 93:935-948.
- Chipman, H., George E.I., and R.E. McCulloch. 2002. Bayesian treed models. *Machine Learning*, 48:299-320.
- Cuffney, T.F., Zappia H., Giddings E.M.P., and J.F. Coles. 2005. Effects of urbanization on benthic macroinvertebrate assemblages in contrasting environmental settings: Boston, Massachusetts; Birmingham, Alabama; and Salt Lake City, Utah. Pages 361–407 in L. R. Brown, R. H. Gray, R. M. Hughes, and M. R. Meador, editors. *Effects of urbanization on stream ecosystems*. American Fisheries Society, Symposium 47, Bethesda, Maryland.
- Dodds, W.K., Smith V.H., and K. Lohman. 2002. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 865-874.
- Dodds, W.K., Smith V.H. and K. Lohman. 2006. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams (Erratum). *Canadian Journal of Fisheries and Aquatic Sciences*, 63:1190-1191.
- Dodds, W.K., Smith V.H., and B. Zander. 1997. Developing nutrient targets to control benthic chlorophyll levels in streams: A case study of the Clark Fork River. *Water Research*, 31(7): 1738-1750.
- Dodds, W.K. 2006. Nutrients and the "Dead Zone": Ecological stoichiometry and depressed dissolved oxygen in the northern Gulf of Mexico. *Frontiers in Ecology and the Environment*, 4:211-217.
- Elith, J., Leathwick J.R. and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77: 802–813.
- Fetscher A.E., Busse L.B., and P.R. Ode. 2009. *Standard Operating Procedures for Collecting Stream Algae Samples and Associated Physical Habitat and Chemical Data for Ambient Bioassessments in California*. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 002. (updated May 2010)
- Fetscher, A.E., Stancheva R., Kociolek J.P., Sheath R.G., Mazor R.D., Stein E.D., Ode P.R., and L.B. Busse 2013. Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs. a combination. *Journal of Applied Phycology*, 26:433-450.

- Freeman, A.M., Lamon III E.C., and C.A. Stow. 2009. Nutrient criteria for lakes, ponds, and reservoirs: A Bayesian TREED model approach. *Ecological Modelling*, 220:630-639.
- Liaw, A and M. Wiener. 2002. Classification and Regression by Random Forest. *R News* 2(3): 18--22.
- Taylor, S. L., Roberts S.C., Walsh C.J., and B.E. Hatt. 2004. Catchment urbanization and increased benthic algal biomass in streams: Linking mechanisms to management. *Freshwater Biology*, 49: 835--851.
- Tetra Tech. 2006. Technical Approach to Develop Nutrient Numeric Endpoints for California. Tetra Tech, Inc. http://rd.tetrattech.com/epa/Documents/CA_NNE_July_Final.pdf
- U.S. Environmental Protection Agency. 2009. Guidance on the Development, Evaluation, and Application of Environmental Models. Office of the Science Advisor, Council for Regulatory Environmental Modeling. Washington, DC. EPA/100/K-09/003.
- Van Nieuwenhuysse, E.E. and J.R. Jones. 1996, Phosphorus-chlorophyll relationship in temperate streams and its variation with stream catchment size, *Canadian Journal of Fisheries and Aquatic Science* 53:99-105.
- Walsh C.J, Roy A.H., Feminella J.W., Cottingham P.D., Groffman P.M., and R.P. Morgan II. 2005. The urban stream syndrome: Current knowledge and the search for a cure. *Journal of the North American Benthological Society*, 24(3): 706-723.

Appendix A. Important Definitions

Important Definitions

For those outside the regulatory world, distinction between terms like “criteria,” “standards,” “objectives,” and “endpoints” can be confusing. The purpose of this section is to provide definitions of the terms that are used in this document within the context of California water quality regulations.

Eutrophication: Eutrophication is defined as the acceleration of the delivery, *in situ* production of organic matter, and accumulation of organic matter (Nixon 1995). One main cause of eutrophication in estuaries is nutrient overenrichment (nitrogen, phosphorus and silica). However, other factors influence primary producer growth and the build-up of nutrient concentrations, and hence modify (or buffer) the response of a system to increased nutrient loads (hereto referred to as **co-factors**). These **co-factors** include hydrologic residence times, mixing characteristics, water temperature, light climate, and grazing pressure.

Indicator: A characteristic of an ecosystem that is related to, or derived from, a measure of biotic or abiotic variable, that can provide quantitative information on ecological condition, structure and/or function. With respect to the water quality objectives, indicators are the ecological parameters for which narrative or numeric objectives are developed.

Numeric Endpoint: Within the context of the NNE framework, numeric endpoints are thresholds that define the magnitude of a response indicator that is considered protective of ecological health. These numeric endpoints serve as guidance to Regional Boards in translating narrative nutrient or biostimulatory substance water quality objectives. They are called “numeric endpoints” rather than “numeric objectives” to distinguish the difference with respect to SWRCB policy. Objectives are promulgated through a public process and incorporated into basin plans. Numeric endpoints are guidance that can evolve over time without the need to go through a formal standards development process.

Water Quality Criteria: Section 303 of the Clean Water Act gives the States and authorized Tribes power to adopt water quality criteria with sufficient coverage of parameters and of adequate stringency to protect designated uses. In adopting criteria, States and Tribes may:

- Adopt the criteria that US EPA publishes under §304(a) of the Clean Water Act;
- Modify the §304(a) criteria to reflect site-specific conditions; or
- Adopt criteria based on other scientifically-defensible methods.

The State of California’s water criteria are implemented as “water quality objectives,” as defined in the Water Code (of the Porter Cologne Act; for further explanation, see below).

States and Tribes typically adopt both **numeric** and **narrative** criteria. **Numeric** criteria are quantitative. **Narrative** criteria lack specific numeric targets but define a targeted condition that must be achieved.

Section 303(c)(2)(B) of the Clean Water Act requires States and authorized Tribes to adopt numeric criteria for priority toxic pollutants for which the Agency has published §304(a) criteria. In addition to narrative and numeric (chemical-specific) criteria, other types of water quality criteria include:

- Biological criteria: a description of the desired biological condition of the aquatic community, for example, based on the numbers and kinds of organisms expected to be present in a water body.
- Nutrient criteria: a means to protect against nutrient over-enrichment and cultural eutrophication.
- Sediment criteria: a description of conditions that will avoid adverse effects of contaminated and uncontaminated sediments.

Water Quality Objectives: The Water Code (Porter-Cologne Act) provides that each Regional Water Quality Control Board shall establish water quality objectives for the waters of the state (i.e., ground and surface waters) which, in the Regional Board's judgment, are necessary for the reasonable protection of beneficial uses and for the prevention of nuisance. The State of California typically adopts both **numeric** and **narrative** objectives. **Numeric** objectives are quantitative. **Narrative** objectives present general descriptions of water quality that must be attained through pollutant control measures. Narrative objectives are also often a basis for the development of numerical objectives.

Water Quality Standards: Water quality standards are the foundation of the water quality-based control program mandated by the Clean Water Act. Water quality standards define the goals for a waterbody by designating its uses, setting criteria to protect those uses, and establishing provisions to protect water quality from pollutants. A water quality standard consists of three basic elements:

1. **Designated uses** of the water body (e.g., recreation, water supply, aquatic life, agriculture; Table A.1),
2. **Water quality criteria** to protect designated uses (numeric pollutant concentrations and narrative requirements), and
3. **Antidegradation policy** to maintain and protect existing uses and high quality waters.

Table A.1. Definition of beneficial uses applicable to freshwater habitat.

Cold Freshwater Habitat (COLD) - Uses of water that support cold water ecosystems including, but not limited to, preservation or enhancement of aquatic habitats, vegetation, fish or wildlife, including invertebrates.

Commercial and Sport Fishing (COMM) - Uses of water for commercial or recreational collection of fish, shellfish, or other organisms including, but not limited to, uses involving organisms intended for human consumption or bait purposes.

Contact Water Recreation (REC-1) - Uses of water for recreational activities involving body contact with water, where ingestion of water is reasonably possible. These uses include, but are not limited to, swimming, wading, water-skiing, skin and SCUBA diving, surfing, white water activities, fishing, or use of natural hot springs.

Migration of Aquatic Organisms (MIGR) - Uses of water that support habitats necessary for migration, acclimatization between fresh and salt water, or other temporary activities by aquatic organisms, such as anadromous fish

Non-contact Water Recreation (REC-2) – Uses of water for recreational activities involving proximity to water, but not normally involving body contact with water, where ingestion of water is reasonably possible. These uses include, but are not limited to, picnicking, sunbathing, hiking, beachcombing, camping, boating, tide pool and marine life study, hunting, sightseeing, or aesthetic enjoyment in conjunction with the above activities.

Rare, Threatened, or Endangered Species (RARE) - Uses of water that support habitats necessary, at least in part, for the survival and successful maintenance of plant or animal species established under state or federal law as rare, threatened or endangered.

Spawning, Reproduction, and/or Early Development (SPWN) - Uses of water that support high quality aquatic habitats suitable for reproduction and early development of fish. This use is applicable only for the protection of anadromous fish.

Warm Freshwater Habitat (WARM) - Uses of water that support warm water ecosystems including, but not limited to, preservation or enhancement of aquatic habitats, vegetation, fish or wildlife, including invertebrates.

Wildlife Habitat (WILD) - Uses of water that support terrestrial ecosystems including, but not limited to, preservation and enhancement of terrestrial habitats, vegetation, wildlife (e.g., mammals, birds, reptiles, amphibians, invertebrates), or wildlife water and food sources.

Appendix B. Graphics and Tables Supporting Analyses of Reference and Ambient Concentrations of Stream Eutrophication Indicator

B.1. Histograms of Biomass and Algal/Macrophyte Cover Data

Figure B.1 Histograms of biomass and algal/macrophyte cover data, all California probability data combined. Y-axes indicate number of sites (N).

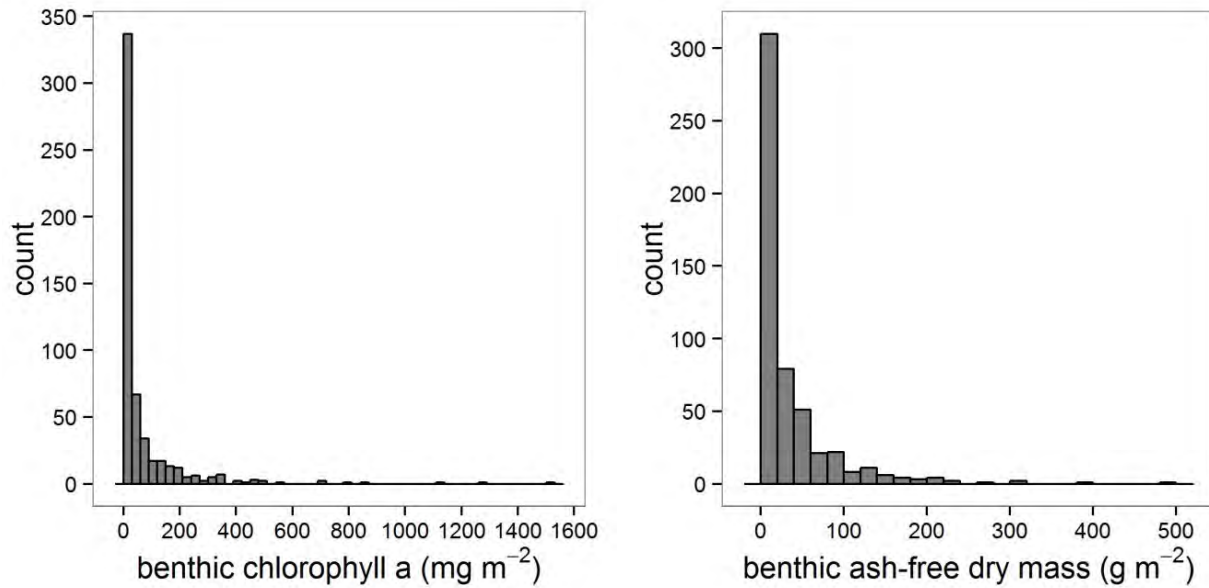


Figure B.1 (continued)

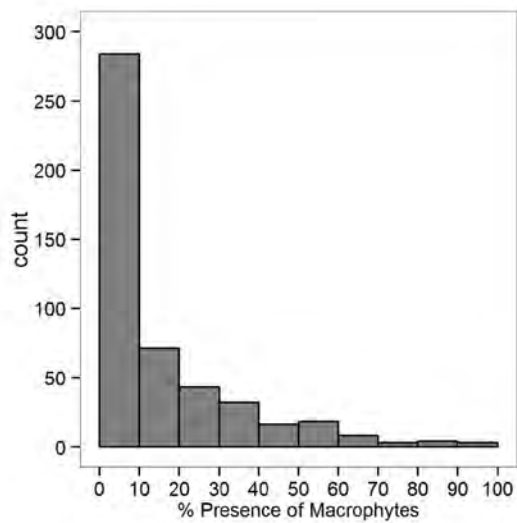
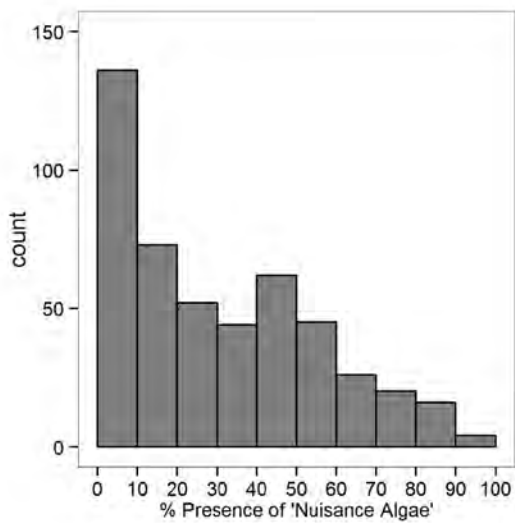
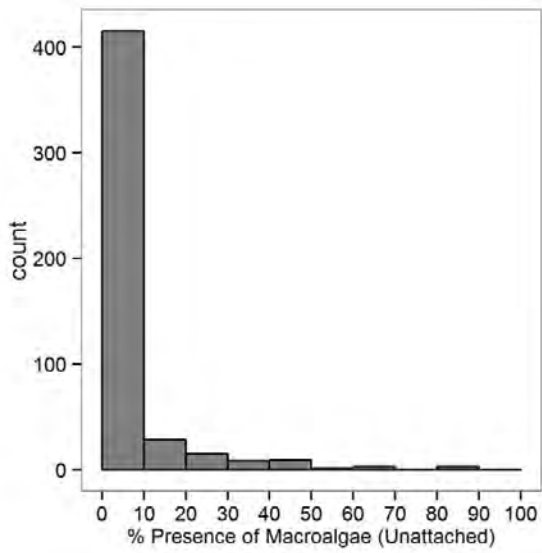
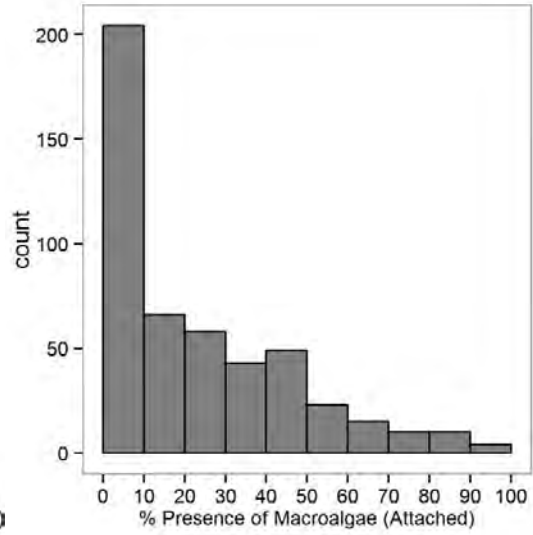
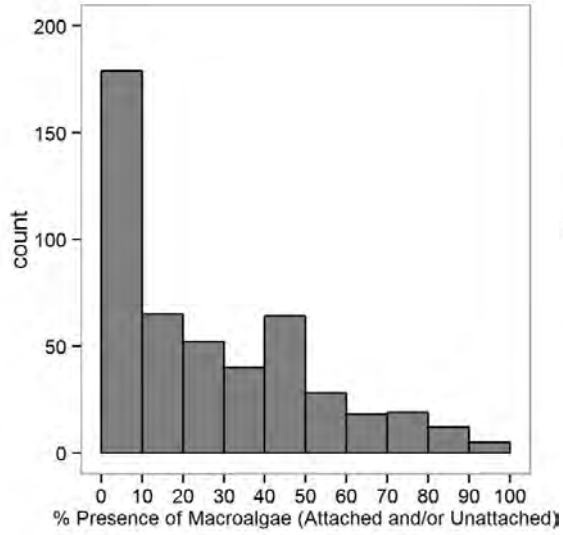


Figure B.1 (continued)

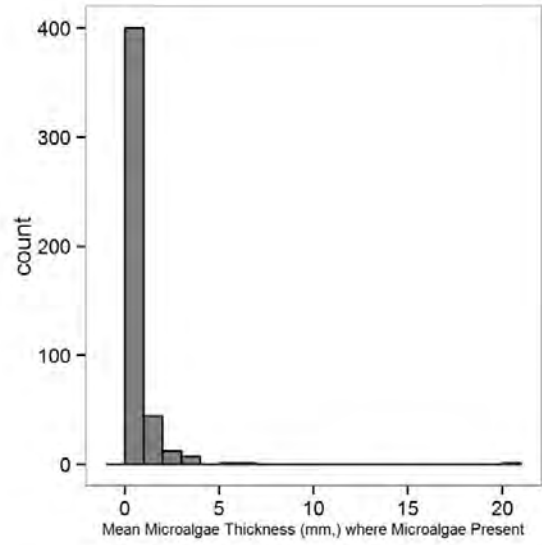
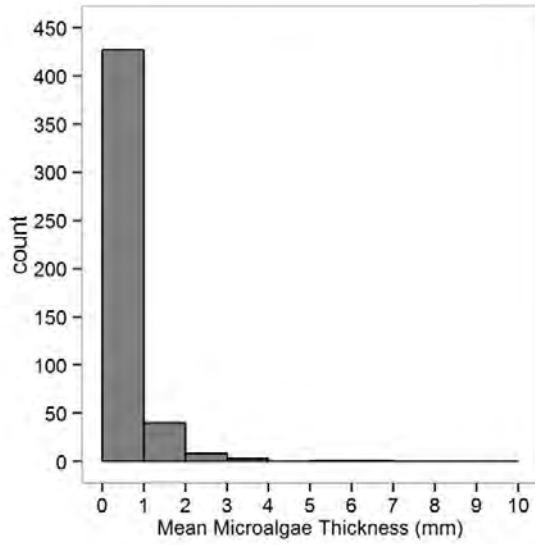
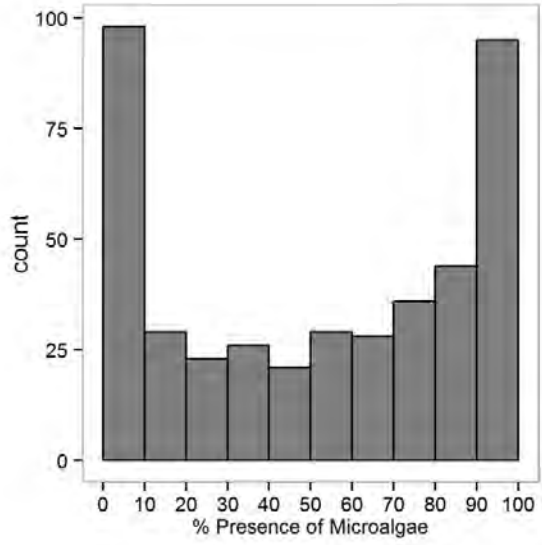
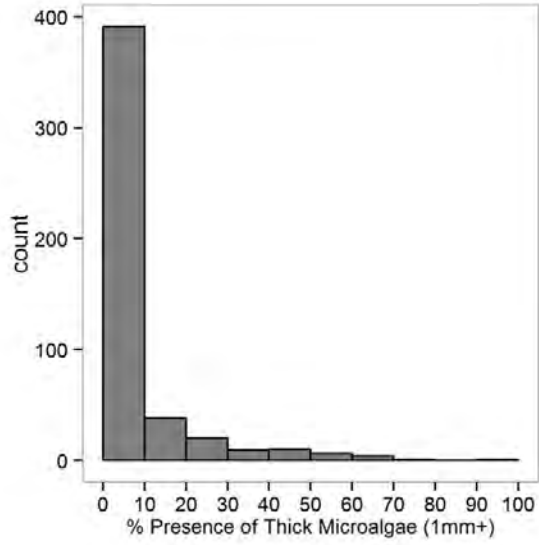


Figure B.2 Boxplots of Biomass, Ash-Free Dry Mass, and Macroalgal Percent Cover.

Boxplots (with “jitter” data points) of biomass, ash-free dry mass, and macroalgal percent cover, for all statewide data combined (i.e., probability plus target sites), stratified by site disturbance class.

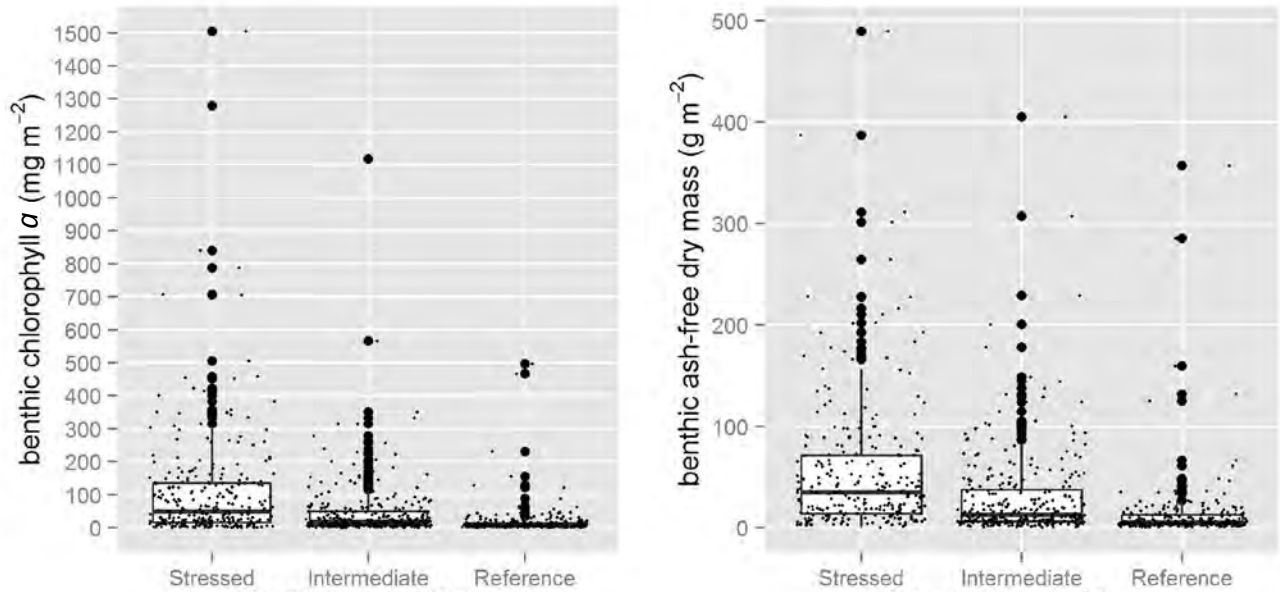


Figure B.2 (continued)

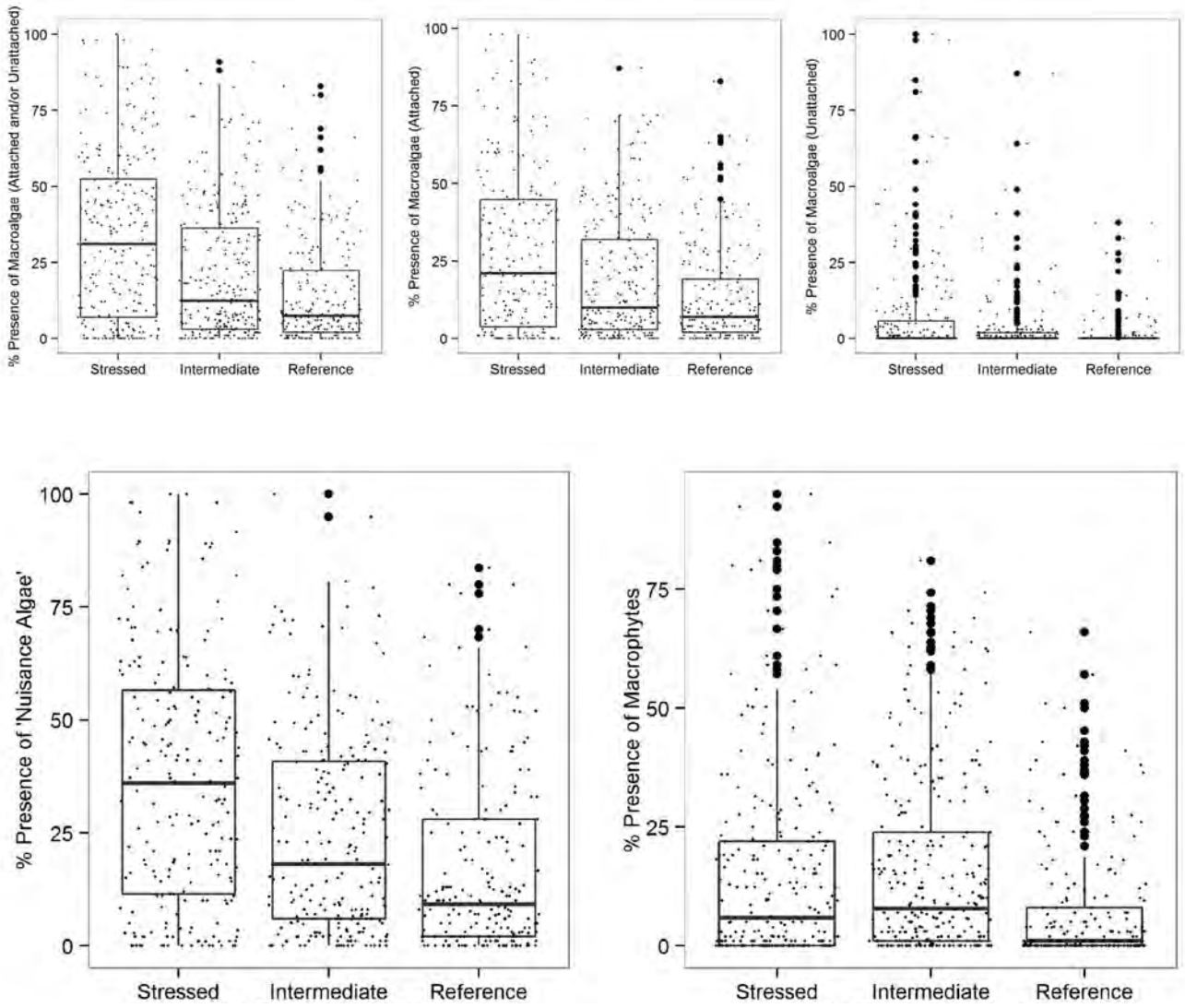


Figure B.2 (continued)

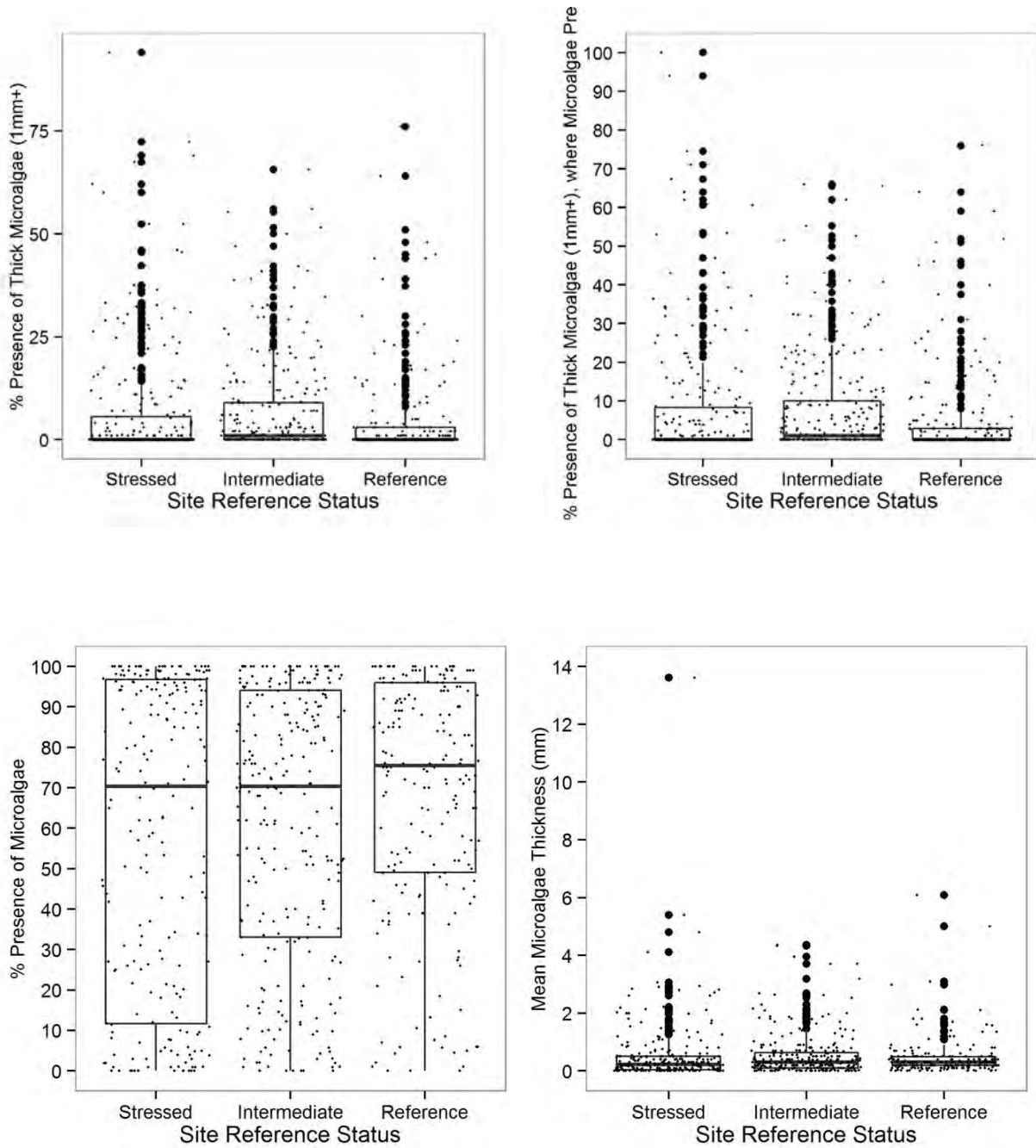


Figure B.3. Cumulative distribution functions of biomass, ash-free dry mass, and macroalgal percent cover, by region, for all probability sites. Shaded areas delineate 95% confidence intervals. Dashed lines correspond to the 75th percentile of values for the indicator in question among Reference sites statewide.

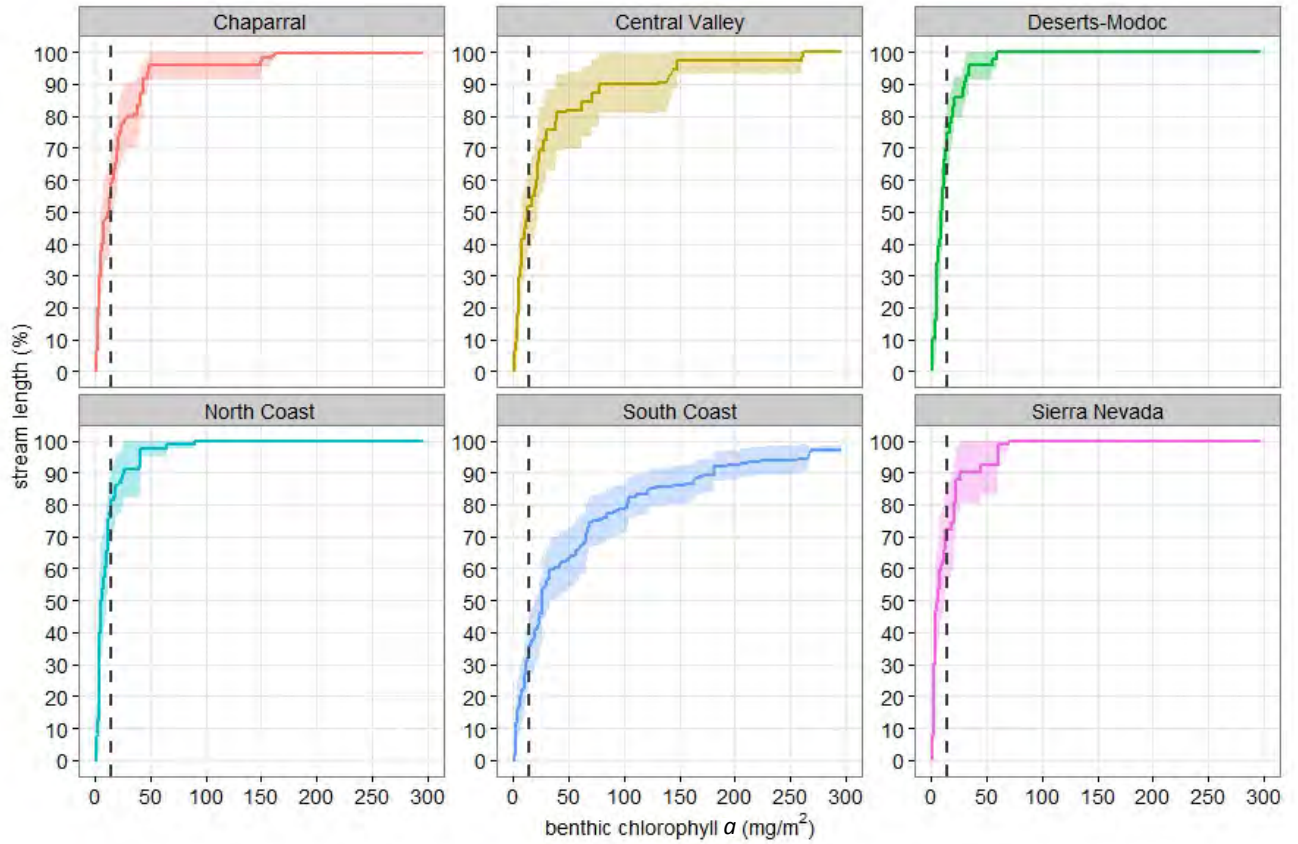
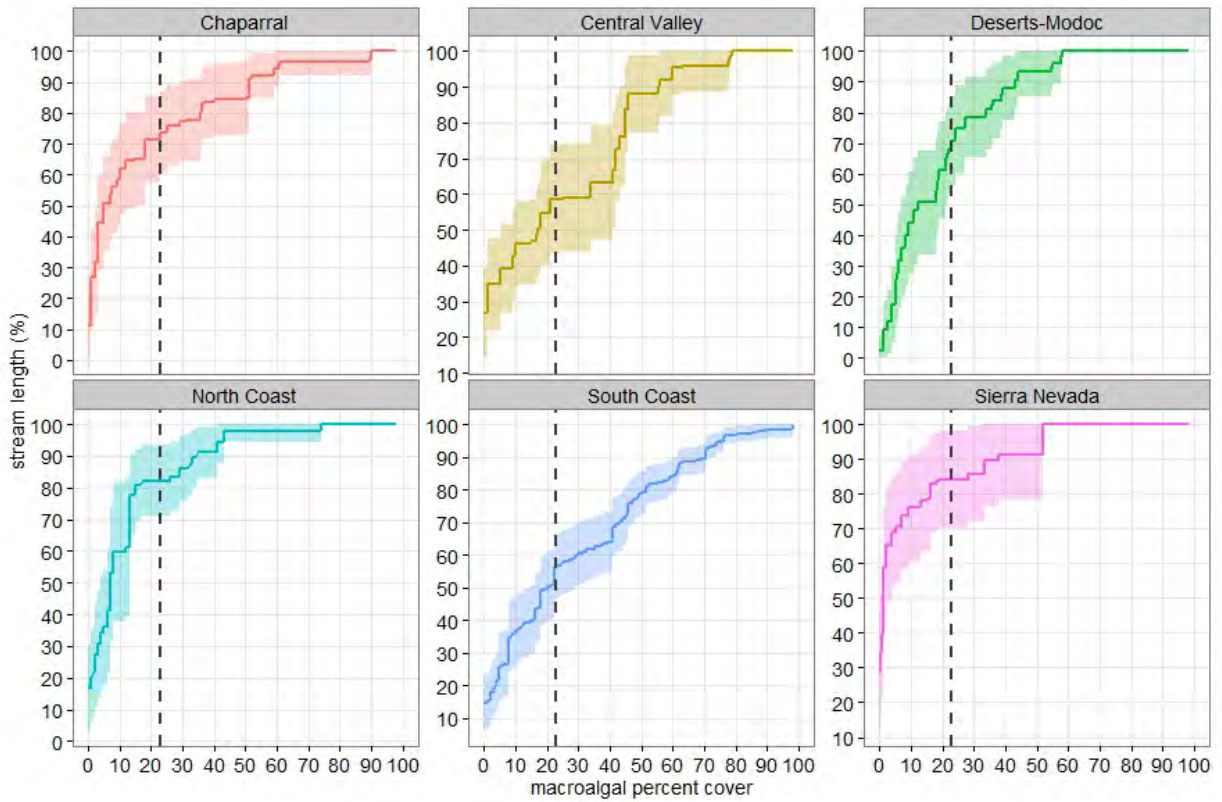
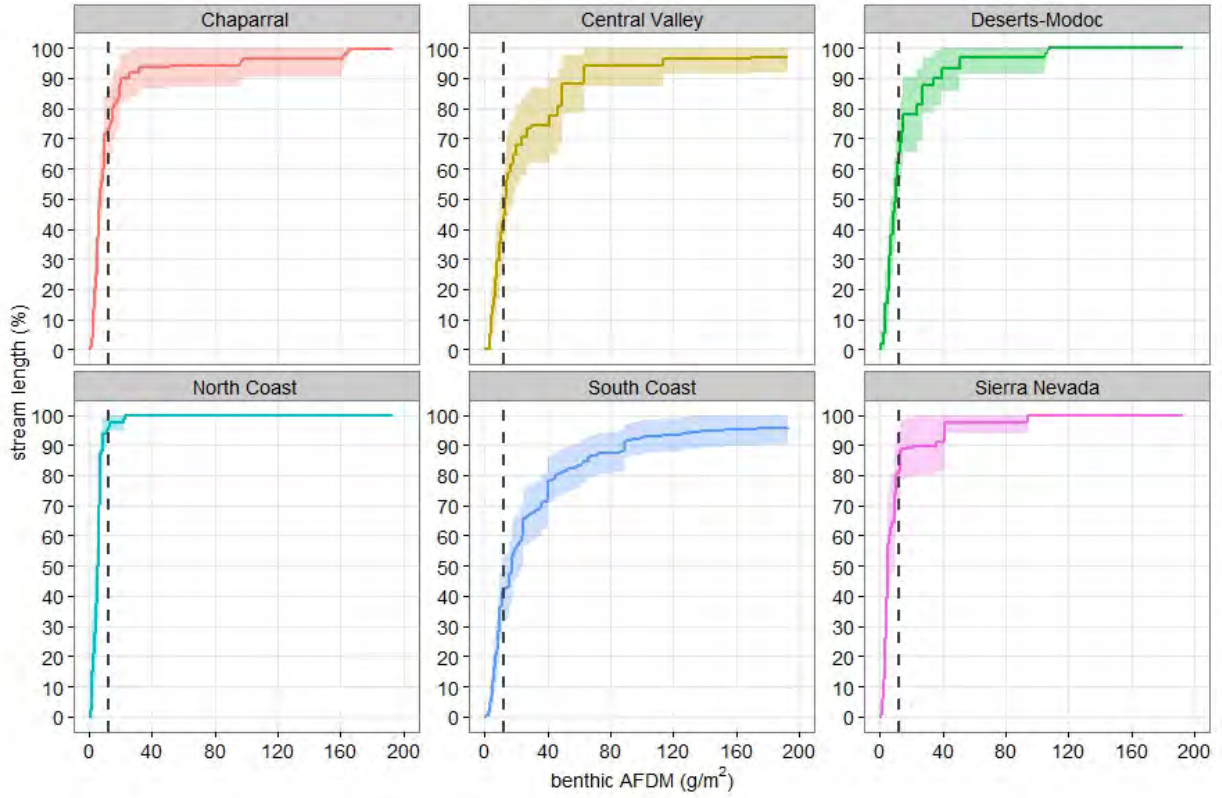


Figure B.3 (continued)



B.2. Chlorophyll *a* Distributions within South Coast

To investigate possible differences in chlorophyll *a* distributions within the PSA6 South Coast ecoregion, we conducted a set of analyses complementary to that which is presented in the main body of the report, in which this ecoregion was further divided into “xeric” and “mountain” zones. This subdivision was based on the Level III classification scheme of Omernik (1987). Multiple “reference” sites were sampled for chlorophyll *a* within both regions (Table B.1), however they were nearly three times as abundant in the mountain zone as in the xeric zone.

Table B.1. Number of sites within each Level III ecoregion (Omernik, 1987) in the South Coast, by site disturbance class.

Ecoregion	Reference	Intermediate	Stressed
South Coast Mountain	27	33	1
South Coast Xeric	11	80	144

For each of the 3 NNE endpoints for chlorophyll *a*, higher proportions of stream length exceeded endpoints within the xeric ecoregion than in the mountain ecoregion (Figures B.4). The same tendency was observed within each site disturbance class (where data were available; Figure B.5).

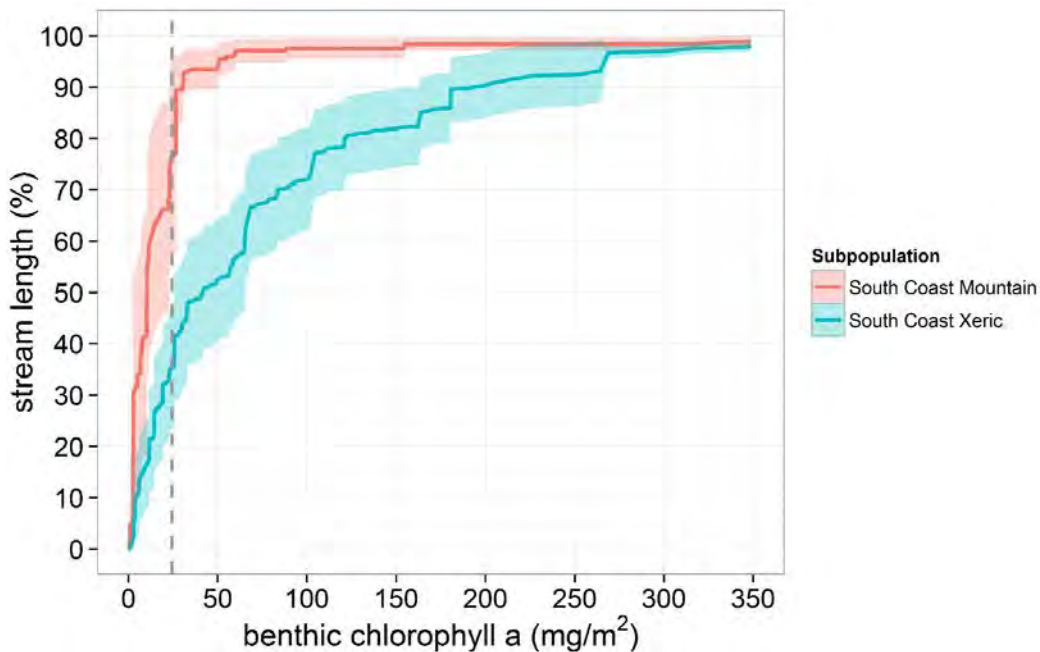


Figure B.4. CDFs for benthic chlorophyll *a*, for the “xeric” and “mountain” Level III ecoregions (Omernik 1987) within the South Coast. The graph shows the estimated probability distributions of chlorophyll *a* relative to the cumulative proportion of stream length. The dashed grey line on the graph denotes the 75th percentile of chlorophyll *a* values among Reference sites in the South Coast (24.4 mg/m²). Highlighted areas delineate the 95% confidence intervals for each estimate.

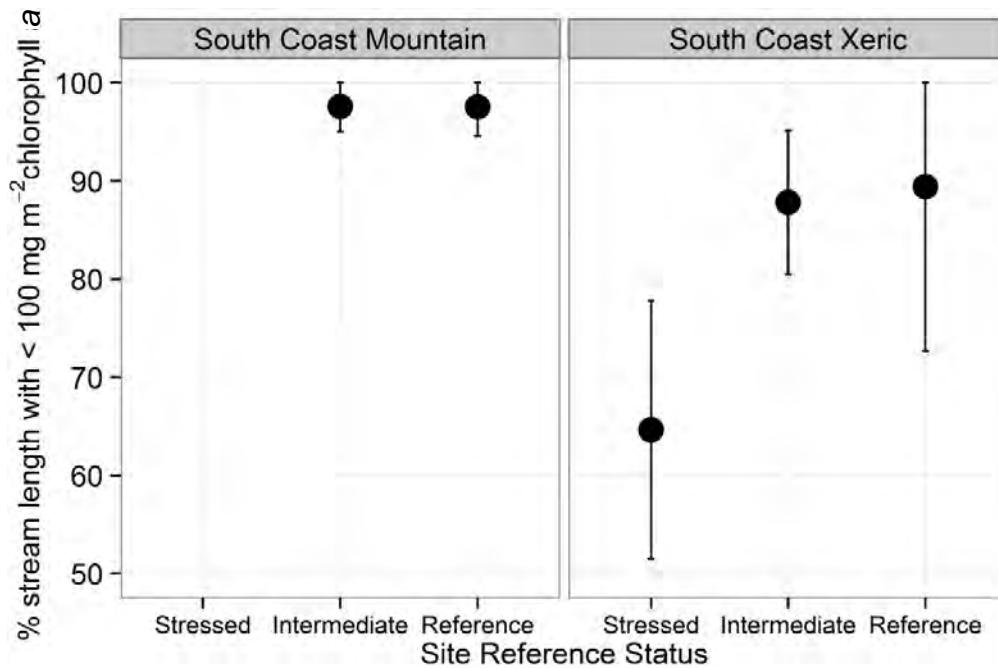


Figure B.5. Within-ecoregion estimated percent of stream kilometers lower than the lowest proposed NNE endpoint for chlorophyll a (100 mg m^{-2}), by site disturbance class. Bars indicate 95% confidence intervals. Note that y-axis scale begins at 50% mark. Due to insufficient sample size, no estimate is available for the “Stressed” site disturbance class within the South Coast Mountain ecoregion (Omernik JM. 1987. Ecoregions of the conterminous United States. Map [scale 1:7,500,000]. *Annals of the Association of American Geographers* 77:118-125).

Appendix C. Graphics and Tables Supporting Analyses of Thresholds of Adverse Effects of Primary Producer Biomass and Nutrient on Wadeable Stream Aquatic Life

C.1. Sample output from TITAN and SiZer analyses.

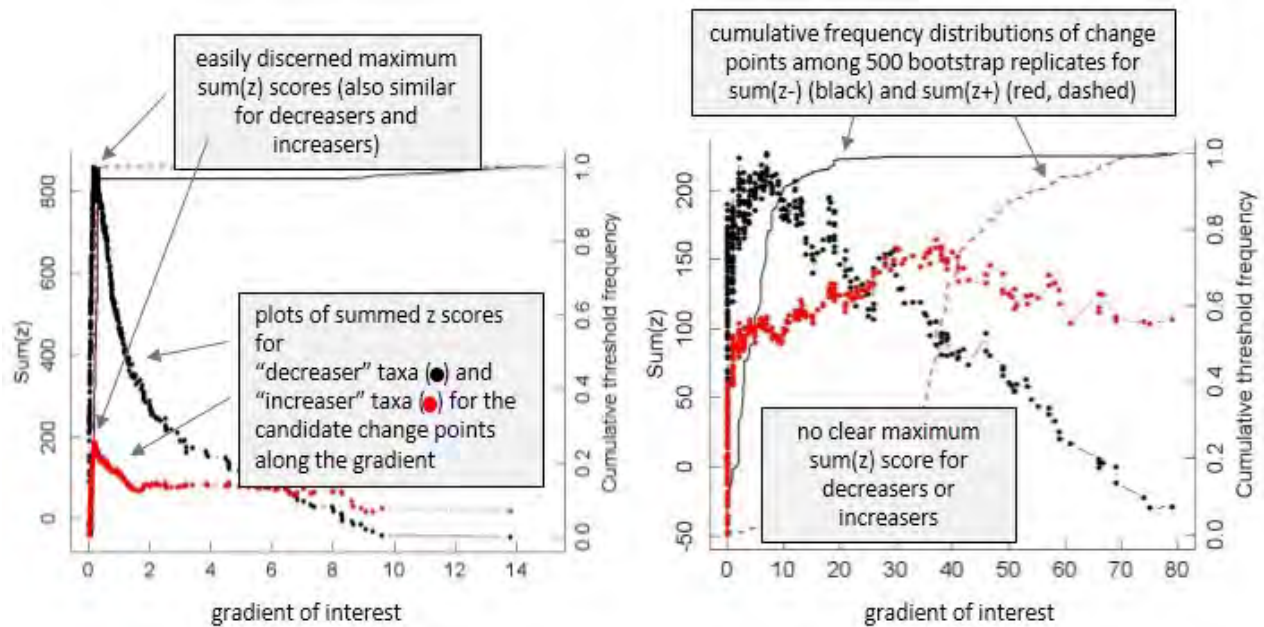


Figure C.1. Examples of plots of TITAN sum(z) scores for taxa that decrease in frequency along the gradient of interest (in black) and those that increase (in red). The graph on the left shows an example of a clear community-level change point, or threshold. The graph on the right indicates no clear community-level change point associated with the gradient in question.

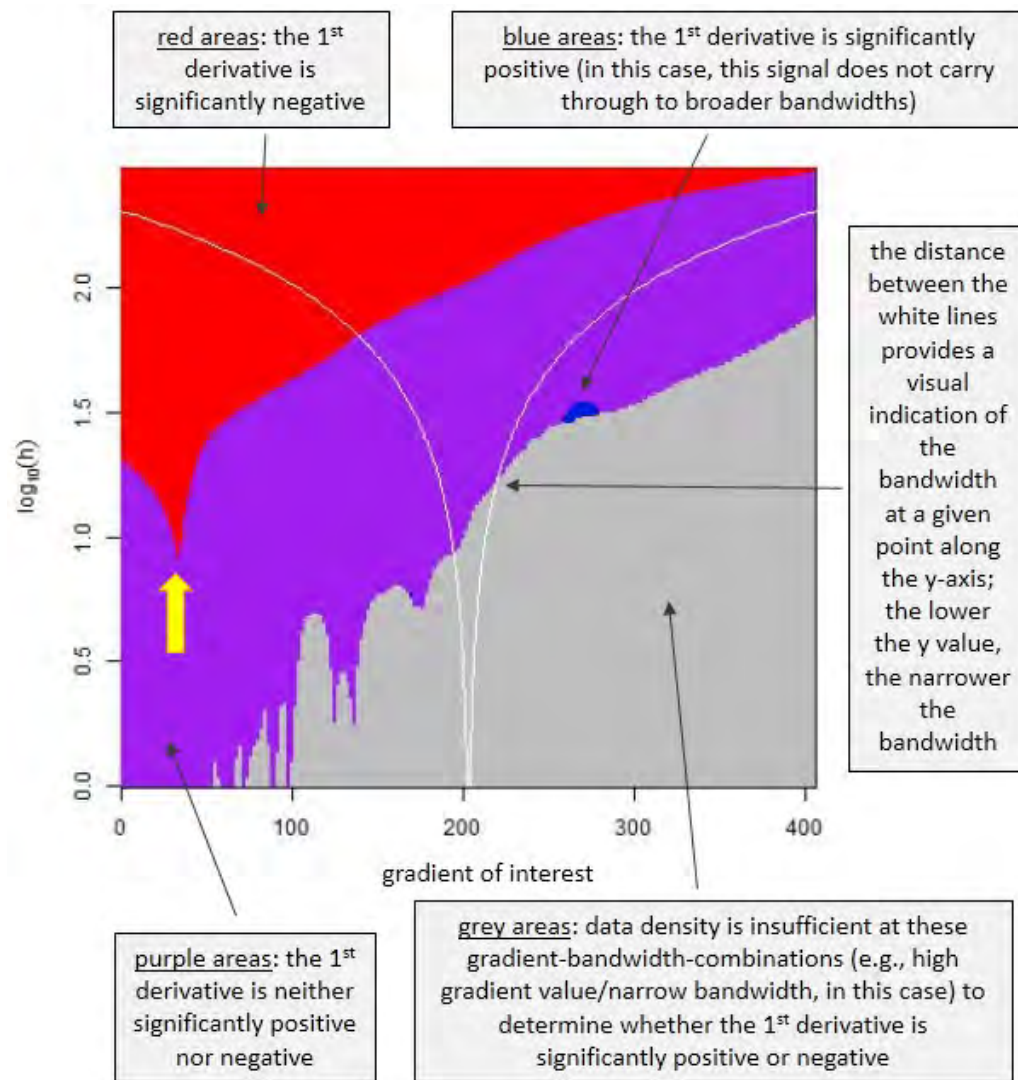


Figure C.3 Example of a SiZer map. This map indicates that there is an overall downward trend in the relationship between the response variable and the gradient, as evidenced by the red color spanning the entire length of the gradient at the top portion of the map (where the bandwidth is broadest). Moving downward along the y-axis, which represents the bandwidths (i.e., as the bandwidths for assessing the first derivative of the locally fitted polynomial relating response to gradient become narrower), red transitions to purple, for most of the gradient, indicating that as bandwidth narrows, the local derivative (i.e., slope) associated with most points along the gradient is no longer significantly positive (or negative). However, at one point along the gradient (i.e., at ~40, as indicated at the yellow arrow), a significant negative first derivative continues to be evident by virtue of the red coloration, even at relatively narrow bandwidths (i.e., extending far downward along the y-axis), providing compelling evidence for a well-supported, steep, negative slope in that narrow region of the gradient. Such a signature characteristically immediately precedes a threshold for this type of gradient-response relationship.

Table C.1. TITAN change point values for BMI and diatom taxa (“pure” and “reliable”)

Also shown is the order (increasing) of diatom taxa in terms of their change points, within the “decreaser” and “increaser” groups, for the AFDM and TP gradients. This information is supplemental to Figure 3.10, where the taxon codes are not legible due to the number of taxa involved. Number of “decreaser” diatom taxa is 65 for AFDM and 68 for TP. Number of “increaser” diatom taxa is 100 for AFDM and 98 for TP.

Direction Of Response		Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
decrease	BMI	<i>Acari</i>	48.59	14.42	-	41.37	0.00	0.51	0.02	-	
		<i>Acentrella</i>	48.70	18.30	-	-	-	0.22	0.04	-	
		<i>Agapetus</i>	14.05	5.97	-	19.00	27.50	0.16	0.06	-	
		<i>Ambrysus</i>	-	-	-	-	-	-	0.00	-	
		<i>Ameletus</i>	6.27	6.20	-	5.00	0.00	0.22	0.02	-	
		<i>Amiocentrus</i>	22.04	9.44	-	-	9.00	0.23	-	-	
		<i>Ampumixis</i>	18.31	6.03	-	7.00	26.50	0.07	0.06	-	
		<i>Anagapetus</i>	4.09	-	-	-	4.00	0.12	-	-	
		<i>Antocha</i>	46.94	8.86	-	45.32	27.50	0.17	0.05	-	
		<i>Apatania</i>	12.98	3.97	-	3.40	0.00	0.16	0.02	-	
		<i>Arctopsyche</i>	4.20	6.01	-	45.32	2.00	0.14	0.03	-	
		<i>Atherix</i>	2.47	4.54	-	-	-	0.21	0.02	-	
		<i>Atrichopogon</i>	5.35	-	-	36.60	-	0.50	-	-	
		<i>Attenella</i>	10.50	6.01	-	8.07	-	0.02	0.06	-	
		<i>Baetis</i>	11.12	6.33	-	16.17	19.02	0.19	0.03	-	
		<i>Berosus</i>	-	-	-	-	-	-	0.01	-	
		<i>Bezzia, Palpomyia</i>	84.70	-	-	-	-	0.50	0.07	-	
		<i>Blephariceridae</i>	-	3.23	-	6.00	-	0.01	0.02	-	
		<i>Brachycentrus</i>	5.39	1.87	-	-	-	0.02	-	-	
		<i>Caenis</i>	-	-	-	-	-	-	0.13	-	
decrease	BMI	<i>Calineuria</i>	20.55	6.50	-	8.00	1.00	0.20	0.02	-	
		<i>Caloparyphus, Euparyphus</i>	-	-	-	-	-	-	0.01	-	
		<i>Capniidae</i>	15.11	9.77	-	-	-	0.17	0.04	-	
		<i>Caudatella</i>	4.23	6.01	-	20.57	4.00	0.13	0.03	-	
		<i>Centroptilum</i>	77.87	18.72	-	-	-	0.28	0.04	-	
		<i>Ceratopsyche, Hydropsyche</i>	49.89	11.77	-	37.07	-	0.54	0.14	-	
		<i>Chelifera, Metachela</i>	3.88	12.25	-	4.38	0.00	0.27	-	-	
		<i>Cheumatopsyche</i>	-	-	-	-	-	-	0.13	-	
		<i>Chimarra</i>	-	-	-	-	-	0.01	-	-	
		<i>Cinygma</i>	8.00	9.73	-	3.96	-	0.10	0.08	-	
		<i>Cinygmula</i>	13.37	6.47	-	3.87	0.48	0.15	0.08	-	

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
decrease	BMI	<i>Cleptelmis</i>	8.09	1.07	-	-	-	0.17	0.02	-
		<i>Clinocera</i>	-	-	-	-	-	0.11	-	-
		<i>Cloeodes</i>	6.35	8.62	-	-	-	0.27	-	-
		<i>Cordulegaster</i>	12.81	-	-	6.33	-	-	-	-
		<i>Cultus</i>	-	14.83	-	-	-	0.06	0.02	-
		<i>Despaxia</i>	9.60	5.79	-	-	-	0.12	-	-
		<i>Deuterophlebia</i>	4.99	3.17	-	-	-	0.10	0.02	-
		<i>Diamesinae</i>	3.86	18.72	-	19.28	2.00	0.17	0.06	-
		<i>Dicosmoecus</i>	6.23	4.60	-	17.00	-	0.03	0.03	-
		<i>Dicranota</i>	17.40	9.04	-	7.31	0.00	0.22	0.08	-
		<i>Diphedor</i>	19.08	9.73	-	31.00	0.00	0.24	0.06	-
		<i>Dixa</i>	17.37	-	-	6.00	-	-	-	-
		<i>Dolophilodes</i>	4.33	4.32	-	-	-	0.17	0.00	-
		<i>Doroneuria</i>	15.11	6.01	-	0.00	4.00	0.10	0.06	-
		<i>Drunella</i>	9.27	6.01	-	16.17	2.00	0.21	0.02	-
		<i>Ecclisomyia</i>	12.96	7.42	-	0.00	-	0.13	0.00	-
		<i>Ecdyonurus</i>	21.15	4.11	-	0.00	0.00	0.29	0.02	-
		<i>Epeorus</i>	10.80	6.30	-	5.03	2.00	0.17	0.02	-
		<i>Ephemerella</i>	9.78	3.78	-	20.10	1.00	0.30	0.01	-
		<i>Eubrianax</i>	26.07	9.49	-	8.00	5.71	0.30	0.08	-
		<i>Frisonia</i>	6.35	4.76	-	2.00	1.00	0.30	0.01	-
		<i>Glossosoma</i>	11.57	6.17	-	8.44	27.50	0.21	0.02	-
		<i>Glutops</i>	17.47	6.50	-	4.76	-	0.16	0.04	-
		<i>Gumaga</i>	18.36	9.70	-	41.00	4.00	0.31	0.02	-
		<i>Helicopsyche</i>	-	-	-	-	-	0.29	0.09	-
		<i>Helodon, Prosimulium</i>	4.93	5.24	-	1.00	2.00	0.13	-	-
		<i>Hesperoconopa</i>	-	-	-	-	-	-	0.02	-
		<i>Hesperoperla</i>	46.60	10.24	-	33.50	6.67	0.06	0.02	-
<i>Heterlimnius</i>	17.15	-	-	4.00	-	0.30	0.07	-		
<i>Heteroplectron</i>	12.98	5.86	-	2.01	5.36	0.17	0.02	-		
<i>Hexatoma</i>	21.93	6.45	-	2.09	0.00	0.28	0.02	-		
<i>Hydraena</i>	18.73	-	-	0.00	-	0.60	0.03	-		
<i>Ironodes</i>	6.06	5.34	-	2.01	13.00	0.11	0.02	-		
decrease	BMI	<i>Isoperla</i>	-	-	-	-	-	0.31	0.04	-
		<i>Juga</i>	48.37	-	-	-	-	0.29	-	-
		<i>Lara</i>	6.27	7.36	-	-	-	0.15	-	-
		<i>Lepidostoma</i>	20.79	9.49	-	18.55	0.00	0.32	0.05	-
		<i>Limnophila</i>	-	1.80	-	-	0.00	0.17	0.04	-
		<i>Malenka</i>	18.36	6.02	-	10.00	12.00	0.60	0.03	-
		<i>Marilia</i>	-	-	-	-	-	0.25	0.02	-
		<i>Maruina</i>	-	4.91	-	-	14.00	0.17	0.01	-

Table C.1

Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)	
decrease	BMI	<i>Matriella, Serratella</i>	10.13	4.25	-	41.00	-	0.25	0.04	-	
		<i>Micrasema</i>	46.22	7.29	-	2.00	-	0.30	0.26	-	
		<i>Microcylloepus</i>	-	-	-	-	-	-	-	0.04	-
		<i>Moselia</i>	4.15	7.36	-	2.00	-	0.12	0.03	-	
		<i>Mystacides</i>	-	-	-	-	-	-	0.29	0.09	-
		<i>Narpus</i>	20.99	12.50	-	18.05	7.31	0.30	0.07	-	
		<i>Neohermes</i>	-	-	-	-	-	-	0.08	-	-
		<i>Neophylax</i>	12.96	5.97	-	12.00	-	0.20	0.02	-	
		<i>Neoplasta</i>	8.09	5.19	-	18.00	-	-	-	-	
		<i>Neotrichia</i>	12.96	3.10	-	-	-	0.02	0.02	-	
		<i>Nothotrichia</i>	-	-	-	-	-	0.11	0.01	-	
		<i>Ochrotrichia</i>	-	36.48	-	-	-	0.03	0.01	-	
		<i>Octogomphus, Specularis</i>	39.16	1.37	-	7.62	-	0.13	0.09	-	
		<i>Oecetis</i>	-	-	-	-	-	0.49	0.08	-	
		<i>Ophiogomphus</i>	-	-	-	-	-	0.09	0.00	-	
		<i>Optioservus</i>	20.55	12.46	-	38.55	7.31	0.31	0.09	-	
		<i>Ordobrevia</i>	17.37	5.97	-	3.00	0.00	0.30	0.02	-	
		<i>Oreodytes</i>	12.96	7.59	-	15.50	-	0.10	0.02	-	
		<i>Oreogeton</i>	3.04	1.31	-	-	0.00	0.09	0.00	-	
		<i>Orohermes</i>	15.68	4.50	-	0.00	-	0.13	0.00	-	
<i>Oroperla</i>	-	4.09	-	-	-	0.12	0.02	-			
<i>Orthoclaadiinae</i>	-	-	-	-	2.00	-	-	-			
<i>Paraleptophlebia</i>	19.96	8.95	-	19.00	0.00	0.36	0.09	-			
<i>Paraleuctra</i>	-	-	-	-	-	0.08	-	-			
<i>Parapsyche</i>	9.34	5.55	-	7.00	7.00	0.29	0.03	-			
<i>Parthina</i>	17.62	7.14	-	7.31	-	0.26	-	-			
<i>Pedomoecus</i>	2.36	1.50	-	13.50	-	0.10	0.03	-			
<i>Pericoma, Telmatoscopus</i>	-	-	-	30.50	-	-	-	-			
<i>Perlinodes</i>	17.31	6.97	-	8.00	0.00	0.21	0.02	-			
<i>Polycentropus</i>	17.37	6.33	-	16.17	0.00	0.23	0.01	-			
<i>Probezzia</i>	-	-	-	6.00	-	0.03	-	-			
<i>Procloeon</i>	-	-	-	-	-	0.44	-	-			
<i>Prodiamesinae</i>	-	-	-	-	-	0.05	-	-			
<i>Protoptila</i>	-	-	-	-	-	0.22	0.04	-			
<i>Psephenus</i>	-	-	-	-	-	0.29	0.02	-			
<i>Psychoglypha</i>	-	-	-	7.00	-	0.30	-	-			
<i>Pteronarcys</i>	15.98	6.63	-	26.33	0.00	0.31	0.01	-			
<i>Ptychopteridae</i>	-	-	-	7.31	-	0.05	-	-			
<i>Rhithrogena</i>	6.93	7.07	-	5.00	0.00	0.11	0.07	-			
<i>Rhizelmis</i>	-	-	-	-	-	0.09	0.00	-			

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
decrease	BMI	<i>Rhyacophila</i>	19.96	6.03	-	19.28	6.00	0.17	0.02	-
		<i>Rickera</i>	3.88	-	-	-	-	-	-	-
		<i>Sialis</i>	23.14	8.95	-	35.27	-	0.26	-	-
		<i>Sierraperla</i>	-	-	-	-	-	0.05	-	-
		<i>Simulium</i>	7.46	-	-	-	-	-	-	-
		<i>Skwala</i>	12.58	4.23	-	-	-	0.12	0.00	-
		<i>Sphaeriidae</i>	-	-	-	3.00	-	-	-	-
		<i>Stictotarsus</i>	-	-	-	-	-	0.62	0.03	-
		<i>Stilobezzia</i>	10.55	2.21	-	-	-	0.20	0.03	-
		<i>Suwallia</i>	4.86	1.33	-	-	1.00	0.01	0.01	-
		<i>Sweltsa</i>	19.96	7.25	-	5.00	0.00	0.15	0.02	-
		<i>Tabanus, Atylotus</i>	-	5.03	-	-	-	-	0.02	-
		<i>Timpanoga</i>	10.28	5.88	-	-	-	0.02	0.02	-
		<i>Tinodes</i>	-	-	-	-	-	0.50	0.01	-
		<i>Visoka</i>	5.79	2.31	-	0.00	-	0.09	0.06	-
		<i>Wiedemannia</i>	-	1.37	-	-	-	0.06	0.05	-
		<i>Wormaldia</i>	-	-	-	-	-	0.22	0.03	-
		<i>Yoraperla</i>	5.80	4.97	-	4.00	7.00	0.19	0.02	-
		<i>Yphria</i>	4.05	2.90	-	-	-	0.08	0.02	-
		<i>Zaitzevia</i>	20.55	12.39	-	45.00	0.00	0.15	0.09	-
<i>Zapada</i>	9.20	5.19	-	2.00	18.18	0.10	0.06	-		
decrease	Diatom	<i>Achnanthydium bisolettianum</i>	-	-	-	-	-	-	0.01	16
		<i>Achnanthydium deflexum</i>	18.72	4.28	22	16.00	-	0.67	0.02	23
		<i>Achnanthydium minutissimum</i>	26.65	12.83	48	9.00	8.57	0.56	0.07	59
		<i>Achnanthydium minutissimum var gracillima</i>	2.64	1.28	3	-	2.93	0.02	-	-
		<i>Achnanthydium sp 1 SWAMP KB</i>	3.85	2.58	11	-	-	0.05	0.01	9
		<i>Adlafia bryophila</i>	-	-	-	0.00	-	0.05	0.04	41
		<i>Adlafia minuscula</i>	12.41	18.52	56	-	-	0.04	-	-
		<i>Amphipleura pellucida</i>	-	-	-	-	5.71	0.45	0.04	48
		<i>Amphipleura sp 1 SCCWRP JPK</i>	-	-	-	-	-	0.13	0.01	17
		<i>Amphipleura sp A SWAMP JPK</i>	-	9.24	37	-	-	-	-	-
<i>Aulacoseira alpigena</i>	3.83	3.40	16	-	-	0.12	-	-		
<i>Aulacoseira ambigua</i>	4.02	-	-	-	-	-	-	-		
<i>Aulacoseira crenulata</i>	8.89	-	-	16.00	-	0.04	-	-		

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
decrease	Diatom	<i>Aulacoseira distans</i>	6.08	9.69	39	-	-	0.03	-	-
		<i>Aulacoseira italica</i>	30.51	-	-	-	-	1.33	-	-
		<i>Aulacoseira subarctica</i>	1.87	3.95	20	-	-	-	-	-
		<i>Brachysira vitrea</i>	-	-	-	-	-	0.03	0.01	12
		<i>Caloneis bacillum</i>	76.75	-	-	-	-	1.03	0.05	51
		<i>Cocconeis disculus</i>	-	-	-	4.00	-	0.37	-	-
		<i>Cocconeis placentula</i> <i>var euglypta</i>	7.61	-	-	25.00	-	0.09	0.14	67
		<i>Cocconeis placentula</i> <i>var lineata</i>	-	2.65	12	-	-	-	-	-
		<i>Cymbella affinis</i>	0.69	1.28	2	-	0.00	0.67	0.01	13
		<i>Cymbella cistula</i>	37.64	12.25	43	-	-	0.28	0.03	37
		<i>Cymbella tumida</i>	-	12.93	49	-	-	0.01	-	-
		<i>Cymbella turgidula</i>	-	-	-	-	-	0.22	0.02	21
		<i>Cymbopleura</i> <i>naviculiformis</i>	13.91	-	-	-	-	-	-	-
		<i>Diatoma hiemale</i>	-	-	-	-	-	0.02	0.02	24
		<i>Diatoma mesodon</i>	3.82	13.15	50	2.00	-	0.09	0.06	54
	<i>Diatoma moniliforme</i>	-	-	-	-	-	0.74	0.02	22	
	<i>Diatoma tenue</i>	35.73	-	-	-	-	-	0.01	4	
	<i>Diatoma vulgare</i>	-	-	-	-	-	0.02	0.04	40	
	<i>Diatoma vulgare var</i> <i>linearis</i>	-	-	-	-	0.00	0.21	0.00	2	
	<i>Didymosphenia</i> <i>geminata</i>	-	9.10	35	-	-	0.20	0.01	5	
	<i>Diploneis oblongella</i>	-	-	-	-	-	0.18	0.00	1	
	<i>Encyonema elginense</i>	2.60	3.42	17	-	-	0.22	0.01	15	
	<i>Encyonema minutum</i>	5.29	17.96	54	2.00	19.00	0.19	-	-	
<i>Encyonema muelleri</i>	-	-	-	-	-	0.08	0.01	18		
<i>Encyonema silesiacum</i>	0.90	-	-	-	-	0.21	0.02	31		
<i>Encyonopsis</i> <i>falaisensis</i>	-	-	-	-	-	-	0.01	14		
<i>Encyonopsis</i> <i>microcephala</i>	-	-	-	-	-	0.45	0.02	25		
<i>Epithemia adnata</i>	-	34.78	64	-	-	0.44	0.03	34		
<i>Epithemia sorex</i>	-	31.74	61	-	-	0.48	0.04	44		
<i>Epithemia turgida</i>	-	17.15	53	-	-	0.05	0.04	47		
<i>Epithemia turgida var</i> <i>westermanni</i>	49.40	18.30	55	-	-	0.04	-	-		
<i>Eunotia bilunaris</i>	-	-	-	-	-	0.09	-	-		
<i>Eunotia incisa</i>	2.71	-	-	-	-	-	-	-		
<i>Fragilaria capucina</i>	-	-	-	-	-	0.01	0.01	10		
decrease	Diatom									

Table C.1												
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)		
decrease	Diatom	<i>Fragilaria capucina</i> <i>var gracilis</i>	11.29	-	-	-	-	0.24	0.03	36		
		<i>Fragilaria capucina</i> <i>var rumpens</i>	0.67	2.45	9	-	-	0.25	-	-		
		<i>Fragilaria vaucheriae</i>	3.86	4.37	23	34.64	-	0.31	0.06	55		
		<i>Fragilaria vaucheriae</i> <i>var capitellata</i>	-	-	-	-	-	0.02	-	-		
		<i>Frustulia</i> <i>amphipleuroides</i>	-	-	-	-	-	0.04	-	-		
		<i>Frustulia krammeri</i>	-	-	-	1.90	-	-	-	-		
		<i>Frustulia vulgaris</i>	-	-	-	21.67	-	-	-	-		
		<i>Geissleria acceptata</i>	6.23	12.46	47	7.00	-	0.06	-	-		
		<i>Geissleria ignota</i>	-	4.50	26	-	-	0.10	-	-		
		<i>Gomphoneis geitleri</i>	-	1.31	5	-	-	0.03	-	-		
		<i>Gomphoneis minuta</i>	49.36	9.58	38	-	-	0.15	0.05	52		
		<i>Gomphoneis</i> <i>olivaceoides</i>	4.62	-	-	-	-	0.21	-	-		
		<i>Gomphoneis</i> <i>olivaceum</i>	48.94	1.30	4	-	-	0.03	0.04	42		
		<i>Gomphoneis</i> <i>pseudokunoi</i>	-	1.82	6	-	-	-	0.02	19		
		decrease	Diatom	<i>Gomphoneis rhombica</i>	-	-	-	-	-	0.04	-	-
				<i>Gomphonema</i> <i>acuminatum</i>	-	-	-	-	-	0.47	-	-
<i>Gomphonema</i> <i>angustatum</i>	-			-	-	-	-	-	0.01	8		
<i>Gomphonema</i> <i>bohemicum</i>	5.00			2.22	8	-	15.24	0.02	0.03	33		
<i>Gomphonema</i> <i>clavatum</i>	-			2.89	13	-	-	-	-	-		
<i>Gomphonema clevei</i>	22.29			4.50	27	0.00	-	0.21	-	-		
<i>Gomphonema</i> <i>kobayasii</i>	10.90			4.47	25	9.00	8.57	0.22	0.06	53		
<i>Gomphonema</i> <i>micropus</i>	-			-	-	-	-	-	0.09	63		
<i>Gomphonema</i> <i>minutum</i>	7.41			6.80	34	0.00	1.00	0.01	0.02	28		
<i>Gomphonema</i> <i>montanum</i>	-			5.30	32	-	-	0.05	0.00	3		
		<i>Gomphonema</i> <i>pumilum</i>	13.86	27.03	59	13.50	0.00	0.30	0.11	65		
		<i>Gomphonema sp B</i> <i>SWAMP JPK</i>	13.14	9.14	36	7.62	4.00	0.06	0.07	60		
		<i>Gomphonema sp C</i> <i>SWAMP JPK</i>	63.87	12.32	45	27.72	0.00	0.07	0.08	61		

Table C.1											
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)	
decrease	Diatom	<i>Gomphonema stoermeri</i>	-	-	-	-	-	-	0.02	26	
		<i>Gomphonema subclavatum</i>	-	-	-	-	-	-	0.03	32	
		<i>Gomphonema truncatum</i>	-	-	-	-	-	-	0.74	0.04	43
		<i>Gomphosphenia sp A SWAMP EWT</i>	2.13	5.03	31	-	-	-	0.05	-	-
		<i>Gomphosphenia sp A SWAMP JPK</i>	2.70	1.88	7	-	-	-	-	-	-
		<i>Gomphosphenia sp B SWAMP EWT</i>	-	-	-	-	-	-	0.05	-	-
		<i>Halamphora normanii</i>	-	-	-	0.00	-	-	-	-	-
		<i>Hannaea arcus</i>	2.51	4.77	30	0.00	0.00	0.06	0.03	38	
		<i>Hantzschia amphioxys</i>	5.80	-	-	11.59	-	-	-	-	-
		<i>Karayevia clevei</i>	21.90	33.78	63	39.00	-	0.89	-	-	
		<i>Karayevia oblongella</i>	17.22	-	-	3.00	0.00	0.13	-	-	
		<i>Karayevia suchlandtii</i>	7.09	4.10	21	4.00	1.00	0.18	0.02	30	
		<i>Kolbesia suchlandtii</i>	16.18	9.77	41	-	0.00	0.17	0.05	50	
		<i>Luticola mutica</i>	-	-	-	24.76	-	-	-	-	
		<i>Mastogloia smithii</i>	-	-	-	-	-	-	0.01	6	
		<i>Meridion circulare</i>	3.73	0.81	1	-	-	0.39	0.04	46	
		<i>Meridion circulare var constrictum</i>	1.20	-	-	2.09	-	0.48	-	-	
		<i>Navicula angusta</i>	5.35	32.83	62	-	-	-	-	-	
		<i>Navicula capitatoradiata</i>	63.93	-	-	-	-	1.14	-	-	
decrease	Diatom	<i>Navicula cryptocephala</i>	9.79	28.84	60	-	-	1.16	-	-	
		<i>Navicula cryptotenella</i>	15.28	10.01	42	40.60	0.00	0.24	0.06	58	
		<i>Navicula cryptotenelloides</i>	5.78	2.58	10	-	-	0.03	0.01	7	
		<i>Navicula digitoradiata</i>	-	-	-	-	-	-	0.01	11	
		<i>Navicula globulifera</i>	-	-	-	-	-	0.04	-	-	
		<i>Navicula margalithii</i>	-	-	-	-	0.00	-	-	-	
		<i>Navicula menisculus</i>	22.84	20.57	58	24.00	5.71	0.30	0.11	66	
		<i>Navicula radiosa</i>	-	-	-	-	-	0.74	0.04	39	
		<i>Navicula rhyngocephala</i>	-	-	-	1.98	-	-	-	-	
		<i>Navicula tenelloides</i>	-	-	-	0.00	-	-	-	-	
		<i>Nitzschia dissipata</i>	62.68	20.12	57	32.38	-	0.48	0.09	64	
<i>Nitzschia dissipata var media</i>	23.30	3.65	18	21.21	2.86	0.25	-	-			
<i>Nitzschia dubia</i>	-	-	-	3.00	-	-	-	-			

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
decrease	Diatom	<i>Nitzschia innominata</i>	-	4.50	28	-	-	0.11	-	-
		<i>Nitzschia linearis</i>	-	-	-	21.67	-	-	-	-
		<i>Nitzschia nana</i>	-	-	-	1.00	-	-	-	-
		<i>Nitzschia paleacea</i>	-	-	-	-	-	0.05	-	-
		<i>Nitzschia perminuta</i>	-	-	-	-	-	0.04	-	-
		<i>Nitzschia recta</i>	9.67	12.32	46	-	-	-	-	-
		<i>Nitzschia tenuirostris</i>	3.82	-	-	-	-	-	-	-
		<i>Pinnularia borealis</i>	2.44	-	-	-	-	-	-	-
		<i>Pinnularia microstauron</i>	12.13	-	-	-	-	-	-	-
		<i>Placoneis elginensis</i>	10.26	-	-	-	-	-	-	-
		<i>Planothidium dubium</i>	34.08	16.10	52	3.00	-	-	-	-
		<i>Planothidium haynaldii</i>	-	-	-	-	-	0.60	-	-
		<i>Planothidium lanceolatum</i>	-	-	-	15.67	-	-	-	-
		<i>Planothidium rostratum</i>	4.93	6.40	33	-	-	-	-	-
		<i>Psammothidium bioretii</i>	1.80	4.41	24	-	-	0.25	0.04	45
		<i>Psammothidium marginulatum</i>	8.18	3.17	15	7.00	-	-	-	-
		<i>Psammothidium subatomoides</i>	-	-	-	-	-	0.16	0.02	27
		<i>Reimeria sinuata</i>	18.25	14.90	51	18.07	3.00	1.14	0.18	68
		<i>Reimeria uniseriata</i>	10.04	3.95	19	-	-	-	0.02	20
		<i>Rhoicosphenia abbreviata</i>	-	-	-	27.29	-	1.69	-	-
		<i>Rhoicosphenia sp B SWAMP EWT</i>	16.18	12.25	44	-	-	0.09	-	-
		<i>Rhoicosphenia sp C SWAMP EWT</i>	48.94	-	-	-	-	-	-	-
		<i>Rhopalodia gibba</i>	-	35.13	65	-	-	0.37	0.06	56
		<i>Rossithidium nodosum</i>	-	-	-	-	-	0.10	-	-
		<i>Rossithidium pusillum</i>	12.63	9.70	40	19.42	2.00	0.29	0.04	49
		<i>Sellaphora bacillum</i>	-	-	-	-	-	0.29	0.08	62
		<i>Sellaphora hustedtii</i>	-	-	-	1.98	-	-	-	-
		<i>Sellaphora seminulum</i>	48.94	-	-	12.38	-	-	-	-
		<i>Sellaphora stroemii</i>	-	-	-	3.94	-	-	-	-
		<i>Stauroneis smithii</i>	-	-	-	9.71	-	-	-	-
<i>Staurosirella leptostauron</i>	-	-	-	-	-	0.80	-	-		
<i>Surirella angusta</i>	-	-	-	1.00	-	-	-	-		
<i>Synedra delicatissima</i>	-	-	-	-	-	-	0.03	35		

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
decrease	Diatom	<i>Synedra mazamaensis</i>	-	-	-	-	-	0.04	-	-
		<i>Synedra ulna</i>	-	-	-	-	-	1.30	0.06	57
		<i>Tabellaria fenestrata</i>	4.01	4.50	29	-	-	0.08	-	-
		<i>Tabellaria flocculosa</i>	2.58	3.17	14	0.00	0.00	0.11	0.02	29
increase	BMI	<i>Agabus</i>	-	-	-	-	-	-	0.02	-
		<i>Ambrysus</i>	3.58	-	-	17.50	-	-	-	-
		<i>Argia</i>	-	12.32	-	-	31.21	0.19	-	-
		<i>Caenis</i>	-	6.01	-	-	-	-	-	-
		<i>Callibaetis</i>	8.18	6.25	-	7.00	39.52	-	-	-
		<i>Caloparyphus, Euparyphus</i>	22.29	14.42	-	35.00	-	0.21	-	-
		<i>Calopterygidae</i>	18.41	6.02	-	60.98	12.38	-	-	-
		<i>Ceratopogon</i>	-	-	-	31.21	13.00	-	-	-
		<i>Chironominae</i>	5.63	2.89	-	-	-	-	-	-
		<i>Crangonyx</i>	-	-	-	-	-	0.30	-	-
		<i>Culicidae</i>	-	-	-	-	-	0.11	0.12	-
		<i>Culicoides</i>	-	60.87	-	-	-	0.52	0.08	-
		<i>Dasyhelea</i>	18.25	11.42	-	18.00	-	8.14	-	-
		<i>Dixella</i>	-	-	-	-	-	-	0.03	-
		<i>Dolichopodidae</i>	155.68	171.39	-	-	-	0.31	0.07	-
		<i>Dubiraphia</i>	-	7.97	-	-	36.00	-	-	-
		<i>Enochrus</i>	-	12.15	-	45.00	-	-	-	-
		<i>Ephydriidae</i>	-	169.11	-	2.00	67.83	0.94	0.39	-
		<i>Fallceon</i>	22.29	6.51	-	38.05	-	0.48	0.12	-
		<i>Ferrissia</i>	-	-	-	-	-	0.17	0.04	-
		<i>Gammarus</i>	-	-	-	-	-	0.55	0.14	-
		<i>Glossiphoniidae</i>	-	12.35	-	42.43	77.02	0.48	0.58	-
		<i>Gyraulus</i>	15.37	6.17	-	-	60.02	0.24	0.10	-
		<i>Hedriodiscus, Odontomyia</i>	-	-	-	-	4.00	-	-	-
		<i>Helisoma</i>	5.02	45.68	-	-	28.50	0.17	0.07	-
		increase	BMI	<i>Hemerodromia</i>	24.12	6.25	-	35.00	-	-
<i>Hyaella</i>	18.41			10.29	-	20.00	12.00	0.32	0.05	-
<i>Hydra</i>	-			6.30	-	-	18.00	6.46	-	-
<i>Hydrobiidae</i>	-			26.26	-	-	-	-	0.04	-
<i>Hydroptila</i>	22.86			12.32	-	19.28	-	0.21	-	-
<i>Ischnura</i>	-			17.17	-	11.00	50.24	0.24	-	-
<i>Laccobius</i>	-			-	-	-	41.50	-	-	-
<i>Lymnaea</i>	7.60			12.25	-	37.50	80.02	-	-	-
<i>Menetus</i>	-			-	-	-	-	0.22	-	-
<i>Muscidae</i>	34.08			19.00	-	-	-	0.61	0.03	-
<i>Nectopsyche</i>	12.00			-	-	10.00	13.17	4.59	-	-

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
increase	BMI	<i>Ochthebius</i>	-	-	-	-	26.00	0.28	-	-
		<i>Oligochaeta</i>	17.37	12.46	-	0.95	5.36	0.29	0.05	-
		<i>Orthocladinae</i>	159.69	-	-	-	-	1.80	-	-
		<i>Ostracoda</i>	29.67	12.35	-	26.83	9.31	0.41	0.08	-
		<i>Oxyethira</i>	-	5.32	-	50.00	-	-	-	-
		<i>Peltodytes</i>	-	-	-	28.71	57.62	-	-	-
		<i>Pericoma, Telmatoscopus</i>	-	-	-	-	-	8.00	-	-
		<i>Petrophila</i>	352.84	7.29	-	9.71	-	-	-	-
		<i>Physa, Physella</i>	17.15	20.29	-	39.00	18.00	0.21	0.03	-
		<i>Prostoma</i>	-	7.87	-	-	-	-	-	-
		<i>Psychoda</i>	-	-	-	-	-	-	0.19	-
		<i>Sphaeriidae</i>	-	-	-	-	36.50	-	-	-
		<i>Tanypodinae</i>	-	3.17	-	18.00	-	-	-	-
		<i>Tipula</i>	-	20.61	-	-	-	8.46	-	-
		<i>Trichocorixa</i>	-	26.75	-	-	-	0.60	0.06	-
		<i>Tricorythodes</i>	9.34	6.47	-	19.78	-	-	-	-
		<i>Tropisternus</i>	-	-	-	-	66.00	0.23	-	-
<i>Turbellaria</i>	18.41	-	-	-	-	0.46	0.03	-		
increase	Diatom	<i>Achnanthydium exiguum</i>	-	-	-	-	-	7.02	0.03	12
		<i>Achnanthydium exiguum var heterovalvum</i>	-	289.93	98	-	-	-	-	-
		<i>Amphora copulata</i>	-	71.95	63	-	-	0.74	0.05	21
		<i>Amphora libyca</i>	-	-	-	-	-	0.66	0.04	16
		<i>Amphora ovalis</i>	37.82	95.37	75	-	57.07	1.05	-	-
		<i>Amphora pediculus</i>	12.88	2.60	2	-	-	0.18	-	-
		<i>Amphora perpusilla</i>	10.35	28.55	37	67.63	4.76	0.21	0.02	5
		<i>Amphora sp 1 SCCWRP BSL</i>	20.55	-	-	-	-	-	-	-
		<i>Amphora sp 1 SWAMP JPK</i>	203.60	227.52	94	42.34	35.62	-	0.13	64
		<i>Amphora sp 5 SWAMP JPK</i>	747.25	153.88	87	-	-	1.91	0.52	90
		<i>Amphora stoermerii</i>	-	34.68	43	-	-	1.01	0.30	82
		<i>Aulacoseira granulata</i>	-	-	-	-	26.11	1.89	0.25	79
		<i>Bacillaria paradoxa</i>	37.82	28.27	36	-	20.99	0.54	0.08	44
		<i>Biremis sp 1 SCCWRP JPK</i>	-	168.60	89	-	-	-	-	-
		<i>Caloneis amphisbaena</i>	46.94	55.91	55	-	-	3.18	-	-
increase	Diatom	<i>Caloneis silicula</i>	-	9.24	8	-	17.07	-	-	-
		<i>Cocconeis pediculus</i>	22.67	7.80	6	7.77	-	4.28	-	-

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
increase	Diatom	<i>Craticula accomoda</i>	-	-	-	54.14	-	0.40	-	-
		<i>Craticula cuspidata</i>	-	-	-	-	-	0.89	0.17	73
		<i>Craticula halophila</i>	48.94	18.02	26	-	37.88	1.28	0.10	56
		<i>Cyclostephanos invisitatus</i>	-	-	-	-	-	-	0.23	76
		<i>Cyclostephanos tholiformis</i>	-	-	-	-	-	0.24	-	-
		<i>Cyclotella meneghiniana</i>	28.89	33.40	42	36.00	19.05	0.44	0.06	30
		<i>Cymatopleura solea</i>	-	68.86	62	-	25.36	-	-	-
		<i>Denticula kuetzingii</i>	20.85	18.66	27	35.27	-	2.11	-	-
		<i>Diademsis confervacea</i>	103.14	137.14	84	89.00	-	0.39	0.08	37
		<i>Diatoma vulgaris</i>	-	-	-	15.00	-	-	-	-
		<i>Diploneis elliptica</i>	-	-	-	-	70.48	-	-	-
		<i>Diploneis oblongella</i>	-	304.33	99	-	12.10	-	-	-
		<i>Diploneis smithii</i>	-	-	-	-	16.51	-	-	-
		<i>Discostella pseudostelligera</i>	-	-	-	-	-	0.24	0.08	42
		<i>Entomoneis alata</i>	63.93	-	-	-	39.00	0.94	0.05	28
		<i>Entomoneis paludosa</i>	-	49.58	52	-	-	6.64	0.08	39
		<i>Eolimna subadnata</i>	-	201.11	91	-	-	2.42	-	-
		<i>Eolimna subminuscula</i>	148.16	45.78	49	-	-	1.91	0.20	74
		<i>Eolimna tantula</i>	-	304.33	100	-	39.00	-	-	-
		<i>Epithemia sorex</i>	-	-	-	8.00	-	-	-	-
		<i>Epithemia turgida</i>	-	-	-	5.00	-	-	-	-
		<i>Fallacia monoculata</i>	-	6.20	3	-	28.29	0.57	0.05	24
		<i>Fallacia pygmaea</i>	-	209.19	93	-	-	0.57	0.14	69
		<i>Fallacia tenera</i>	-	36.48	45	-	-	0.70	0.14	68
		<i>Fragilaria capucina</i>	-	-	-	-	0.95	-	-	-
		<i>Frustulia creuzburgensis</i>	-	-	-	-	-	1.14	0.20	75
		<i>Frustulia vulgaris</i>	-	150.31	86	-	8.72	-	0.02	8
		<i>Gomphonema acuminatum</i>	8.09	-	-	-	25.86	-	-	-
		<i>Gomphonema affine</i>	10.55	-	-	-	70.48	-	-	-
		<i>Gomphonema augur</i>	121.56	60.75	57	47.82	15.24	3.86	-	-
<i>Gomphonema lagenula</i>	137.52	66.22	61	-	-	2.48	-	-		
<i>Gomphonema mexicanum</i>	-	-	-	25.00	70.95	0.17	0.62	94		
<i>Gomphonema parvulum</i>	46.86	27.37	33	45.86	15.24	0.62	0.11	58		

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
increase	Diatom	<i>Gomphonema pseudoaugur</i>	22.29	80.51	67	42.00	-	1.54	-	-
		<i>Gomphonema truncatum</i>	-	-	-	31.00	-	-	-	-
		<i>Gyrosigma acuminatum</i>	-	77.65	66	-	-	0.66	-	-
		<i>Gyrosigma nodiferum</i>	28.73	87.18	68	-	30.48	0.61	0.05	25
		<i>Halamphora coffeaeformis</i>	130.30	14.30	21	-	-	9.93	-	-
		<i>Halamphora normanii</i>	-	-	-	-	-	-	0.08	41
		<i>Halamphora veneta</i>	47.23	14.05	19	41.75	-	0.76	0.09	49
		<i>Hantzschia amphioxys</i>	-	-	-	-	46.33	0.31	0.08	40
		<i>Hippodonta capitata</i>	-	19.01	28	-	57.07	0.62	0.06	32
		<i>Hippodonta hungarica</i>	-	27.83	35	-	79.52	0.98	0.10	55
		<i>Hippodonta pumila</i>	-	-	-	-	-	0.26	-	-
		<i>Karayevia ploenensis</i>	-	-	-	-	-	0.46	0.04	19
		<i>Lemnicola hungarica</i>	-	-	-	-	51.21	-	0.13	66
		<i>Luticola cohnii</i>	-	87.70	69	-	-	-	-	-
		<i>Luticola goeppertiana</i>	-	-	-	-	-	-	0.39	86
		<i>Luticola mutica</i>	-	72.61	64	-	64.76	0.81	0.07	36
		<i>Mayamaea agrestis</i>	31.65	100.83	78	-	5.84	-	-	-
		<i>Mayamaea atomus</i>	-	11.42	14	6.00	-	-	-	-
		<i>Melosira varians</i>	-	-	-	-	2.96	-	0.02	2
<i>Navicula angusta</i>	-	-	-	-	-	-	0.04	15		
<i>Navicula arenaria</i>	-	62.50	60	-	-	-	-	-		
<i>Navicula cari</i>	-	-	-	-	50.24	1.33	-	-		
<i>Navicula cincta</i>	-	6.48	5	-	31.43	0.60	0.13	65		
<i>Navicula cryptocephala</i>	-	-	-	-	80.99	-	-	-		
<i>Navicula digitoradiata</i>	-	-	-	-	61.45	-	-	-		
<i>Navicula erifuga</i>	-	-	-	-	3.00	0.94	1.87	98		
<i>Navicula graciloides</i>	-	109.30	79	-	73.73	0.36	0.24	78		
<i>Navicula gregaria</i>	26.73	13.90	17	-	7.00	0.57	0.08	45		
<i>Navicula libonensis</i>	-	23.18	32	-	13.00	-	-	-		
<i>Navicula normaloides</i>	171.70	72.93	65	-	-	0.67	-	-		
<i>Navicula peregrina</i>	-	124.41	82	-	-	0.58	-	-		
<i>Navicula phyllepta</i>	28.04	47.96	50	-	3.81	-	-	-		
<i>Navicula radiosa</i>	-	-	-	-	25.86	-	-	-		
<i>Navicula radiosa var tenella</i>	-	27.37	34	-	-	-	-	-		
<i>Navicula recens</i>	-	13.97	18	-	-	9.74	-	-		
<i>Navicula rhyngocephala</i>	-	-	-	-	7.00	-	-	-		

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
increase	Diatom	<i>Navicula rostellata</i>	-	-	-	-	-	-	0.30	83
		<i>Navicula salinarum</i>	65.28	111.63	80	-	15.24	0.97	-	-
		<i>Navicula schroeteri</i>	-	90.46	72	-	-	1.12	0.09	50
		<i>Navicula sp 3 SWAMP JPK</i>	747.25	-	-	-	-	1.91	0.38	85
		<i>Navicula subrotundata</i>	-	-	-	-	-	0.62	0.46	89
		<i>Navicula tenelloides</i>	-	100.08	77	-	-	-	-	-
		<i>Navicula trivialis</i>	-	12.71	15	-	72.45	0.21	0.12	61
		<i>Navicula veneta</i>	13.40	11.40	13	-	8.00	0.16	-	-
		<i>Navicula viridula</i>	-	-	-	-	2.00	0.24	0.09	51
		<i>Navicula viridula var linearis</i>	-	-	-	-	6.92	-	-	-
		<i>Nitzschia acicularis</i>	-	-	-	-	2.00	-	-	-
		<i>Nitzschia amphibia</i>	26.43	17.20	25	31.38	59.04	0.76	0.06	33
		<i>Nitzschia amphiboides</i>	94.19	201.11	92	-	-	3.12	-	-
		<i>Nitzschia angustatula</i>	-	271.41	97	-	-	0.57	0.05	29
		<i>Nitzschia aurariae</i>	50.71	14.95	22	16.00	-	1.33	1.01	97
		<i>Nitzschia bacillum</i>	117.41	-	-	-	-	-	-	-
		<i>Nitzschia bryophila</i>	-	-	-	-	-	1.81	0.44	87
		<i>Nitzschia bulnheimiana</i>	25.97	11.27	12	90.50	-	1.47	0.12	60
		<i>Nitzschia capitellata</i>	126.93	-	-	-	-	8.63	0.09	48
		<i>Nitzschia communis</i>	-	93.23	73	-	37.07	2.03	0.08	46
		<i>Nitzschia commutata</i>	57.63	48.78	51	-	10.48	-	-	-
		<i>Nitzschia compressa var vexans</i>	-	50.88	53	-	63.33	-	0.02	7
		<i>Nitzschia desertorum</i>	77.89	14.13	20	29.00	-	1.05	0.69	95
		<i>Nitzschia dubia</i>	-	-	-	-	-	0.54	0.05	27
		<i>Nitzschia elegantula</i>	13.70	60.87	58	-	-	-	-	-
		<i>Nitzschia filiformis</i>	23.95	-	-	-	-	16.42	-	-
		<i>Nitzschia fonticola</i>	-	-	-	7.69	-	-	0.03	10
		<i>Nitzschia frustulum</i>	314.91	13.09	16	-	53.67	-	0.10	54
		<i>Nitzschia inconspicua</i>	63.87	2.26	1	70.74	-	0.22	0.03	9
		<i>Nitzschia intermedia</i>	-	-	-	-	-	-	0.92	96
		<i>Nitzschia lacuum</i>	9.38	8.38	7	-	-	-	-	-
		<i>Nitzschia liebethruthii</i>	10.30	32.83	41	-	-	0.66	-	-
<i>Nitzschia linearis</i>	-	16.63	24	-	29.26	-	-	-		
<i>Nitzschia microcephala</i>	15.28	29.77	40	18.02	12.00	0.29	0.08	47		
<i>Nitzschia minuta</i>	-	-	-	-	-	4.45	-	-		
<i>Nitzschia palea</i>	49.81	44.72	48	-	6.81	0.73	0.09	52		
<i>Nitzschia paleaeformis</i>	-	28.57	39	-	-	-	-	-		
<i>Nitzschia perminuta</i>	-	-	-	-	39.02	-	-	-		

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
increase	Diatom	<i>Nitzschia perspicua</i>	-	-	-	-	-	4.58	0.28	81
		<i>Nitzschia rosenstockii</i>	-	98.26	76	-	58.55	0.98	0.13	67
		<i>Nitzschia sigma</i>	-	-	-	-	-	0.62	0.44	88
		<i>Nitzschia solita</i>	28.04	-	-	-	-	-	0.03	14
		<i>Nitzschia umbonata</i>	-	-	-	-	-	3.75	0.15	70
		<i>Nitzschia valdecostata</i>	-	90.46	71	-	-	-	0.03	13
		<i>Nitzschia vitrea</i>	-	162.77	88	-	-	-	-	-
		<i>Parlibellus protracta</i>	-	6.26	4	-	-	0.38	0.02	6
		<i>Placoneis elginensis</i>	-	-	-	-	-	-	0.23	77
		<i>Planothidium delicatulum</i>	20.78	10.92	11	36.00	17.00	0.61	0.10	57
		<i>Planothidium engelbrechtii</i>	-	137.14	83	-	53.67	0.26	0.02	4
<i>Planothidium frequentissimum</i>	43.48	10.27	10	88.50	-	0.41	0.06	35		
increase	Diatom	<i>Planothidium lanceolatum</i>	-	39.47	46	-	42.00	-	0.02	3
		<i>Pleurosira laevis</i>	58.35	20.57	29	83.33	31.21	6.46	0.08	38
		<i>Psammodictyon constrictum</i>	121.56	61.55	59	-	21.00	0.94	-	-
		<i>Pseudostaurosira elliptica</i>	74.23	34.69	44	24.22	5.86	0.54	0.13	63
		<i>Pseudostaurosira parasitica</i>	5.29	117.98	81	-	39.00	-	0.05	26
		<i>Pseudostaurosira subsalina</i>	29.48	-	-	90.50	-	4.15	-	-
		<i>Rhoicosphenia sp 1 SCCWRP JPK</i>	-	-	-	-	20.95	-	-	-
		<i>Rhoicosphenia sp B SWAMP EWT</i>	-	-	-	-	-	-	0.04	17
		<i>Rhopalodia gibba</i>	-	-	-	3.00	-	-	-	-
		<i>Rhopalodia musculus</i>	-	-	-	-	10.00	-	-	-
		<i>Rhopalodia operculata</i>	-	138.66	85	-	-	0.20	-	-
		<i>Sellaphora hustedtii</i>	-	229.99	96	-	-	-	-	-
		<i>Sellaphora laevissima</i>	-	-	-	-	-	-	0.30	84
		<i>Sellaphora nyassensis</i>	-	-	-	-	39.00	-	-	-
<i>Sellaphora pupula</i>	-	21.87	30	-	2.00	0.46	0.26	80		
increase	Diatom	<i>Sellaphora seminulum</i>	-	-	-	-	-	-	0.02	1
		<i>Sellaphora sp 2 SWAMP JPK</i>	-	-	-	-	6.03	-	-	-
		<i>Simonsenia delognei</i>	-	187.78	90	-	-	-	-	-
		<i>Stauroneis smithii</i>	-	227.52	95	-	-	-	-	-
		<i>Staurosira construens</i>	-	-	-	3.98	62.86	-	0.58	93

Table C.1										
Direction Of Response	Assemblage	Taxon	Chloro phyll <i>a</i> (mg/m ²)	AFDM (g/m ²)	AFDM Change Point Order (diatoms)	PCT_MAP	PCT_MCP	TN (mg/L)	TP (mg/L)	TP Change Point Order (diatoms)
		<i>Staurosira construens</i> <i>var binodis</i>	-	-	-	16.27	-	-	-	-
		<i>Staurosira construens</i> <i>var venter</i>	-	-	-	32.00	-	0.17	0.05	20
		<i>Staurosirella pinnata</i>	-	-	-	-	-	-	0.03	11
		<i>Stephanodiscus</i> <i>medius</i>	-	-	-	-	-	-	0.05	22
		<i>Surirella angusta</i>	-	-	-	-	15.74	0.35	0.12	62
		<i>Surirella brebissonii</i>	28.73	22.00	31	-	57.62	1.07	0.15	71
		<i>Surirella brebissonii</i> <i>var kuetzingii</i>	-	15.98	23	-	10.48	0.58	0.04	18
		<i>Surirella brightwellii</i>	-	-	-	-	10.00	9.93	-	-
		<i>Surirella ovalis</i>	-	52.31	54	-	20.95	0.49	0.09	53
		<i>Surirella ovata</i>	-	-	-	-	-	5.59	0.53	92
		<i>Synedra ulna</i>	-	-	-	2.86	0.00	-	-	-
		<i>Tabularia fasciculata</i>	28.79	28.57	38	66.81	12.38	0.36	0.08	43
		<i>Tabularia tabulata</i>	-	-	-	-	-	0.38	0.06	31
increase	Diatom	<i>Thalassionema</i> <i>nitzschioides</i>	185.30	43.73	47	-	39.05	-	-	-
		<i>Thalassiosira</i> <i>weissflogii</i>	28.73	93.23	74	-	-	0.76	0.15	72
		<i>Tryblionella calida</i>	-	57.08	56	-	-	1.26	-	-
		<i>Tryblionella constricta</i>	63.87	9.49	9	-	20.80	0.97	0.06	34
		<i>Tryblionella hungarica</i>	136.63	89.37	70	-	25.00	0.74	0.11	59
		<i>Tryblionella levidensis</i>	-	-	-	-	-	0.69	0.05	23
		<i>Tryblionella littoralis</i>	-	-	-	-	-	0.70	0.53	91

Table C.2. Results of piecewise regressions for all analyses in which “relaxed” criteria were met.

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
Chlorophyll <i>a</i> (mg/m ²)	D18	unweighted	98.82 (12.11), 47.55	-0.32 (-0.39 – -0.24)	0.00 (-0.02 – 0.02)	0.19	yes	no
Chlorophyll <i>a</i> (mg/m ²)	D18	weighted	119.29 (20.48), 80.40	-0.32 (-0.38 – -0.25)	-0.01 (-0.07 – 0.06)	0.14	no	no
Chlorophyll <i>a</i> (mg/m ²)	EPT_Taxa	unweighted	68.77 (8.72), 34.25	-0.18 (-0.23 – -0.14)	0.00 (-0.01 – 0.01)	0.25	yes	no
Chlorophyll <i>a</i> (mg/m ²)	EPT_Taxa	weighted	79.77 (14.57), 57.24	-0.21 (-0.26 – -0.16)	0.00 (-0.03 – 0.02)	0.18	no	no
Chlorophyll <i>a</i> (mg/m ²)	H20	unweighted	104.54 (10.91), 42.88	-0.31 (-0.37 – -0.25)	-0.01 (-0.02 – 0.01)	0.28	yes	no
Chlorophyll <i>a</i> (mg/m ²)	H20	weighted	118.31 (20.40), 80.14	-0.27 (-0.33 – -0.21)	-0.01 (-0.07 – 0.05)	0.16	no	no
Chlorophyll <i>a</i> (mg/m ²)	H21	unweighted	93.13 (9.46), 37.14	-0.36 (-0.43 – -0.29)	-0.01 (-0.02 – 0.01)	0.28	yes	no
Chlorophyll <i>a</i> (mg/m ²)	H21	weighted	110.66 (17.42), 68.44	-0.34 (-0.40 – -0.27)	-0.01 (-0.07 – 0.05)	0.19	no	no
Chlorophyll <i>a</i> (mg/m ²)	H23	unweighted	102.55 (9.86), 38.71	-0.35 (-0.41 – -0.29)	-0.01 ((-0.02 – 0.01)	0.31	yes	no
Chlorophyll <i>a</i> (mg/m ²)	H23	weighted	119.02 (19.66), 77.18	-0.30 (-0.36 – -0.24)	-0.01 (-0.07 – 0.05)	0.17	no	no
Chlorophyll <i>a</i> (mg/m ²)	Intolerant_ Percent	unweighted	23.29 (2.91), 11.41	-0.01 (-0.01 – -0.01)	0.00 (0.00 – 0.00)	0.21	yes	no
Chlorophyll <i>a</i> (mg/m ²)	Intolerant_ Percent	weighted	55.73 (10.64), 41.79	0.00 (-0.01 – 0.00)	0.00 (0.00 – 0.00)	0.14	no	no
Chlorophyll <i>a</i> (mg/m ²)	Intolerant_ PercentTaxa	unweighted	31.40 (3.34), 13.11	-0.01 (-0.01 – -0.01)	0.00 (0.00 – 0.00)	0.29	yes	no
Chlorophyll <i>a</i> (mg/m ²)	Intolerant_ PercentTaxa	weighted	74.91 (12.58), 49.42	0.00 (-0.01 – 0.00)	0.00 (0.00 – 0.00)	0.20	no	no
Chlorophyll <i>a</i> (mg/m ²)	Intolerant_ Taxa	unweighted	31.43 (3.56), 13.99	-0.30 (-0.37 – -0.23)	0.00 (-0.01 – 0.00)	0.26	yes	no
Chlorophyll <i>a</i> (mg/m ²)	Intolerant_ Taxa	weighted	60.93 (10.03), 39.39	-0.22 (-0.28 – -0.17)	0.00 (-0.03 – 0.02)	0.18	no	no
Chlorophyll <i>a</i> (mg/m ²)	RAWpropGree nCRUS	unweighted	103.37 (12.60), 49.47	0.01 (0.00 – 0.01)	0.00 (0.00 – 0.00)	0.22	yes	no
Chlorophyll <i>a</i> (mg/m ²)	RAWpropGree nCRUS	weighted	117.59 (28.46), 111.78	0.00 (0.00 – 0.01)	0.00 (0.00 – 0.00)	0.09	no	no
Chlorophyll <i>a</i> (mg/m ²)	S2	unweighted	113.46 (13.00), 51.05	-0.32 (-0.39 – -0.26)	-0.01 (-0.03 – 0.01)	0.25	yes	no
Chlorophyll <i>a</i> (mg/m ²)	S2	weighted	133.22 (43.18), 169.53	-0.19 (-0.26 – -0.12)	-0.01 (-0.09 – 0.07)	0.06	no	no
AFDM (g/m ²)	CSCI	unweighted	15.02 (1.87), 7.35	-0.02 (-0.03 – -0.01)	0.00 (0.00 – 0.00)	0.23	yes	no
AFDM (g/m ²)	CSCI	weighted	16.75 (2.33), 9.14	-0.02 (-0.03 – -0.01)	0.00 (0.00 – 0.00)	0.19	no	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
AFDM (g/m ²)	D18	unweighted	31.72 (3.70), 14.55	-0.93 (-1.17 – -0.70)	-0.05 (-0.09 – -0.01)	0.24	yes	no
AFDM (g/m ²)	D18	weighted	33.91 (4.73), 18.59	-0.84 (-1.04 – -0.64)	-0.06 (-0.12 – 0.00)	0.20	no	no
AFDM (g/m ²)	EPT_Percent	unweighted	16.10 (1.62), 6.38	-0.02 (-0.03 – -0.02)	0.00 (0.00 – 0.00)	0.29	yes	no
AFDM (g/m ²)	EPT_Percent	weighted	15.42 (1.41), 5.52	-0.03 (-0.03 – -0.02)	0.00 (0.00 – 0.00)	0.30	yes	no
AFDM (g/m ²)	EPT_Percent Taxa	unweighted	18.93 (1.84), 7.23	-0.02 (-0.02 – -0.01)	0.00 (0.00 – 0.00)	0.35	yes	no
AFDM (g/m ²)	EPT_Percent Taxa	weighted	22.19 (2.96), 11.61	-0.01 (-0.01 – -0.01)	0.00 (0.00 – 0.00)	0.23	no	no
AFDM (g/m ²)	EPT_Taxa	unweighted	17.47 (1.64), 6.43	-0.78 (-0.95 – -0.61)	-0.02 (-0.04 – 0.00)	0.35	yes	no
AFDM (g/m ²)	EPT_Taxa	weighted	19.71 (2.72), 10.67	-0.65 (-0.83 – -0.48)	-0.02 (-0.06 – 0.03)	0.20	no	no
AFDM (g/m ²)	H20	unweighted	36.06 (3.68), 14.44	-0.81 (-0.98 – -0.63)	-0.03 (-0.06 – 0.01)	0.27	yes	no
AFDM (g/m ²)	H20	weighted	43.02 (6.79), 26.66	-0.56 (-0.69 – -0.43)	-0.04 (-0.09 – 0.02)	0.18	no	no
AFDM (g/m ²)	H21	unweighted	34.75 (3.40), 13.36	-0.91 (-1.10 – -0.72)	-0.02 (-0.06 – 0.02)	0.28	yes	no
AFDM (g/m ²)	H21	weighted	34.65 (4.74), 18.62	-0.78 (-0.96 – -0.60)	-0.03 (-0.09 – 0.02)	0.20	no	no
AFDM (g/m ²)	H23	unweighted	35.79 (3.40), 13.35	-0.91 (-1.09 – -0.73)	-0.04 (-0.08 – 0.00)	0.31	yes	no
AFDM (g/m ²)	H23	weighted	35.05 (5.79), 22.74	-0.66 (-0.83 – -0.49)	-0.06 (-0.12 – -0.01)	0.18	no	no
AFDM (g/m ²)	Intolerant_ Percent	unweighted	12.99 (1.10), 4.34	-0.02 (-0.02 – -0.02)	0.00 (0.00 – 0.00)	0.31	yes	no
AFDM (g/m ²)	Intolerant_ Percent	weighted	13.74 (1.45), 5.69	-0.02 (-0.02 – -0.01)	0.00 (0.00 – 0.00)	0.22	yes	no
AFDM (g/m ²)	Intolerant_ PercentTaxa	unweighted	15.41 (1.26), 4.94	-0.02 (-0.02 – -0.02)	0.00 (0.00 – 0.00)	0.38	yes	no
AFDM (g/m ²)	Intolerant_ PercentTaxa	weighted	16.20 (2.01), 7.90	-0.02 (-0.02 – -0.01)	0.00 (0.00 – 0.00)	0.21	yes	no
AFDM (g/m ²)	Intolerant_ Taxa	unweighted	15.29 (1.27), 4.98	-0.78 (-0.94 – -0.62)	-0.01 (-0.03 – 0.00)	0.36	yes	no
AFDM (g/m ²)	Intolerant_ Taxa	weighted	16.52 (2.04), 8.01	-0.71 (-0.89 – -0.53)	-0.01 (-0.05 – 0.03)	0.20	yes	no
AFDM (g/m ²)	propAchMin	unweighted	11.17 (1.41), 5.52	-0.02 (-0.03 – -0.01)	0.00 (0.00 – 0.00)	0.13	yes	no
AFDM (g/m ²)	propAchMin	weighted	14.52 (2.17), 8.52	-0.01 (-0.02 – -0.01)	0.00 (0.00 – 0.00)	0.10	no	no
AFDM (g/m ²)	RAWDO100	unweighted	7.10 (0.88), 3.45	-0.04 (-0.06 – -0.02)	0.00 (0.00 – 0.00)	0.10	yes	no
AFDM (g/m ²)	RAWDO100	weighted	14.39 (2.65), 10.40	-0.01 (-0.02 – -0.01)	0.00 (0.00 – 0.00)	0.06	no	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
AFDM (g/m ²)	S2	unweighted	39.33 (4.90), 19.24	-0.76 (-0.96 – -0.56)	-0.01 (-0.06 – 0.04)	0.18	yes	no
AFDM (g/m ²)	S2	weighted	94.63 (31.41), 123.32	-0.25 (-0.36 – -0.15)	0.01 (-0.11 – 0.12)	0.05	no	no
AFDM (g/m ²)	Tolerant_PercentTaxa	unweighted	28.40 (3.46), 13.61	0.01 (0.01 – 0.01)	0.00 (0.00 – 0.00)	0.30	yes	no
AFDM (g/m ²)	Tolerant_PercentTaxa	weighted	29.18 (4.86), 19.09	0.01 (0.00 – 0.01)	0.00 (0.00 – 0.00)	0.19	no	no
NO _x (mg/L)	H20	unweighted	0.61 (0.07), 0.28	-52.67 (-64.23 – -41.12)	-0.10 (-0.81 – 0.62)	0.33	yes	no
NO _x (mg/L)	H20	weighted	0.63 (0.09), 0.34	-60.65 (-73.24 – -48.06)	-0.70 (-2.53 – 1.14)	0.26	no	no
NO _x (mg/L)	H23	unweighted	0.38 (0.05), 0.18	-82.75 (-102.20 – -63.34)	-0.46 (-1.17 – 0.24)	0.33	yes	no
NO _x (mg/L)	H23	weighted	0.60 (0.09), 0.34	-63.72 (-77.25 – -50.19)	-0.64 (-2.60 – 1.33)	0.24	no	no
NO _x (mg/L)	S2	unweighted	0.29 (0.03), 0.13	-129.30 (-158.40 – -100.30)	-0.92 (-1.67 – -0.18)	0.39	yes	no
NO _x (mg/L)	S2	weighted	0.29 (0.05), 0.18	-114.80 (-147.20 – -82.37)	-0.98 (-2.88 – 0.93)	0.21	no	no
SRP (mg/L)	D18	unweighted	0.13 (0.01), 0.04	-329.50 (-378.60 – -280.40)	1.78 (-5.03 – 8.60)	0.33	yes	no
SRP (mg/L)	D18	weighted	0.13 (0.01), 0.04	-347.40 (-401.10 – -293.80)	7.36 (-6.93 – 21.64)	0.28	yes	no
SRP (mg/L)	EPT_PercentTaxa	unweighted	0.13 (0.01), 0.05	-2.37 (-2.84 – -1.89)	-0.01 (-0.06 – 0.04)	0.26	yes	no
SRP (mg/L)	EPT_PercentTaxa	weighted	0.14 (0.02), 0.07	-2.21 (-2.75 – -1.68)	-0.04 (-0.14 – 0.06)	0.22	no	no
SRP (mg/L)	EPT_Taxa	unweighted	0.13 (0.01), 0.06	-103.80 (-125.30 – -82.41)	-0.32 (-2.52 – 1.89)	0.25	yes	no
SRP (mg/L)	EPT_Taxa	weighted	0.13 (0.02), 0.07	-115.70 (-145.00 – -86.37)	-2.49 (-7.91 – 2.94)	0.20	no	no
SRP (mg/L)	H20	unweighted	0.13 (0.01), 0.04	-315.70 (-360.00 – -271.40)	-1.12 (-7.26 – 5.02)	0.37	yes	yes
SRP (mg/L)	H20	weighted	0.14 (0.01), 0.05	-284.30 (-330.80 – -237.80)	4.09 (-8.30 – 16.48)	0.29	yes	no
SRP (mg/L)	H21	unweighted	0.14 (0.01), 0.04	-299.60 (-347.30 – -251.80)	1.04 (-5.71 – 7.79)	0.31	yes	no
SRP (mg/L)	H21	weighted	0.14 (0.01), 0.05	-300.20 (-352.40 – -248.00)	5.03 (-8.87 – 18.93)	0.26	yes	no
SRP (mg/L)	H23	unweighted	0.14 (0.01), 0.04	-312.40 (-359.60 – -265.20)	-0.05 (-6.72 – 6.61)	0.34	yes	no
SRP (mg/L)	H23	weighted	0.14 (0.01), 0.05	-284.60 (-334.80 – -234.40)	2.78 (-10.60 – 16.15)	0.25	yes	no
SRP (mg/L)	Intolerant_PercentTaxa	unweighted	0.12 (0.01), 0.05	-2.36 (-2.87 – -1.85)	-0.01 (-0.06 – 0.03)	0.24	yes	no
SRP (mg/L)	Intolerant_PercentTaxa	weighted	0.12 (0.02), 0.07	-2.40 (-2.98 – -1.81)	-0.05 (-0.16 – 0.06)	0.20	no	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
SRP (mg/L)	Intolerant_Taxa	unweighted	0.12 (0.01), 0.06	-85.69 (-105.50 – -65.85)	-0.38 (-2.22 – 1.46)	0.21	yes	no
SRP (mg/L)	Intolerant_Taxa	weighted	0.12 (0.02), 0.07	-104.10 (-131.40 – -76.87)	-1.59 (-6.60 – 3.41)	0.17	no	no
SRP (mg/L)	propAchMin	unweighted	0.09 (0.01), 0.04	-2.92 (-3.59 – -2.24)	-0.01 (-0.07 – 0.05)	0.18	yes	no
SRP (mg/L)	propAchMin	weighted	0.09 (0.01), 0.05	-3.27 (-4.09 – -2.45)	-0.01 (-0.17 – 0.15)	0.13	no	no
SRP (mg/L)	RAWDO100	unweighted	0.08 (0.01), 0.03	-4.36 (-5.45 – -3.27)	0.04 (-0.04 – 0.12)	0.16	yes	no
SRP (mg/L)	RAWDO100	weighted	0.05 (0.00), 0.02	-6.94 (-8.58 – -5.31)	0.16 (-0.02 – 0.34)	0.13	yes	no
SRP (mg/L)	RAWDO50	unweighted	0.23 (0.02), 0.09	-1.08 (-1.26 – -0.90)	0.08 (0.03 – 0.13)	0.22	yes	no
SRP (mg/L)	RAWDO50	weighted	0.17 (0.02), 0.07	-1.36 (-1.60 – -1.13)	0.09 (0.01 – 0.17)	0.22	yes	no
SRP (mg/L)	RAWeutro	unweighted	0.09 (0.01), 0.04	3.99 (3.13 – 4.86)	-0.05 (-0.13 – 0.03)	0.19	yes	no
SRP (mg/L)	RAWeutro	weighted	0.08 (0.01), 0.04	4.14 (3.09 – 5.20)	-0.20 (-0.41 – 0.00)	0.11	yes	no
SRP (mg/L)	RAWlowN	unweighted	0.10 (0.01), 0.03	-4.74 (-5.57 – -3.92)	-0.02 (-0.11 – 0.06)	0.28	yes	no
SRP (mg/L)	RAWlowN	weighted	0.10 (0.01), 0.05	-4.59 (-5.56 – -3.61)	0.03 (-0.19 – 0.26)	0.16	yes	no
SRP (mg/L)	RAWlowP	unweighted	0.09 (0.01), 0.03	-5.28 (-6.19 – -4.38)	-0.04 (-0.12 – 0.05)	0.29	yes	no
SRP (mg/L)	RAWlowP	weighted	0.09 (0.01), 0.04	-5.59 (-6.66 – -4.52)	-0.02 (-0.25 – 0.20)	0.19	yes	no
SRP (mg/L)	RAWNhet	unweighted	0.14 (0.02), 0.07	1.65 (1.26 – 2.05)	0.01 (-0.04 – 0.07)	0.18	yes	no
SRP (mg/L)	RAWNhet	weighted	0.14 (0.02), 0.07	1.95 (1.53 – 2.37)	-0.02 (-0.13 – 0.09)	0.18	yes	no
SRP (mg/L)	RAWsapro	unweighted	0.14 (0.01), 0.06	-2.73 (-3.28 – -2.17)	0.06 (-0.02 – 0.14)	0.21	yes	no
SRP (mg/L)	RAWsapro	weighted	0.13 (0.02), 0.06	-2.85 (-3.48 – -2.22)	0.17 (0.00 – 0.33)	0.15	yes	no
SRP (mg/L)	S2	unweighted	0.13 (0.01), 0.06	-277.70 (-338.10 – -217.40)	-4.84 (-13.02 – 3.34)	0.20	yes	no
SRP (mg/L)	S2	weighted	0.14 (0.03), 0.13	-164.40 (-228.90 – -99.92)	-9.55 (-26.76 – 7.66)	0.07	no	no
SRP (mg/L)	Taxonomic_Richness	unweighted	0.14 (0.02), 0.07	-128.10 (-157.80 – -98.42)	-0.10 (-3.20 – 3.01)	0.22	yes	no
SRP (mg/L)	Taxonomic_Richness	weighted	0.14 (0.02), 0.08	-135.40 (-172.50 – -98.38)	-3.08 (-9.96 – 3.79)	0.19	no	no
SRP (mg/L)	Tolerant_PercentTaxa	unweighted	0.13 (0.02), 0.06	1.65 (1.24 – 2.05)	-0.01 (-0.05 – 0.03)	0.19	yes	no
SRP (mg/L)	Tolerant_PercentTaxa	weighted	0.15 (0.02), 0.08	1.55 (1.15 – 1.95)	0.03 (-0.05 – 0.10)	0.21	no	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
RAWpropGreenCRUS	propTaxaZHR	unweighted	0.86 (0.08), 0.32	-0.07 (-0.13 – -0.02)	-0.57 (-1.14 – 0.00)	0.13	yes	no
RAWpropGreenCRUS	propTaxaZHR	weighted	0.92 (0.10), 0.41	-0.01 (-0.07 – 0.05)	-0.72 (-2.77 – 1.32)	0.01	no	no
RAWpropGreenCRUS	S2	unweighted	0.99 (0.01), 0.02	-42.71 (-46.16 – -39.27)	1290 (3768 – 1188)	0.58	yes	yes
RAWpropGreenCRUS	S2	weighted	0.97 (0.03), 0.11	-33.86 (-39.20 – -28.52)	329.8 (896.6 – 237)	0.36	yes	yes
TN (mg/L)	CSCI	unweighted	0.74 (0.06), 0.25	-0.58 (-0.68 – -0.47)	0.01 (0.00 – 0.01)	0.40	yes	no
TN (mg/L)	CSCI	weighted	0.72 (0.07), 0.27	-0.73 (-0.85 – -0.60)	0.00 (-0.02 – 0.02)	0.37	yes	no
TN (mg/L)	D18	unweighted	0.88 (0.07), 0.26	-45.66 (-53.16 – -38.16)	0.38 (-0.22 – 0.97)	0.37	yes	yes
TN (mg/L)	D18	weighted	1.29 (0.13), 0.50	-34.89 (-39.88 – -29.91)	0.25 (-1.54 – 2.04)	0.31	yes	no
TN (mg/L)	EPT_Percent	unweighted	0.59 (0.07), 0.26	-0.60 (-0.75 – -0.45)	0.00 (0.00 – 0.01)	0.27	yes	no
TN (mg/L)	EPT_Percent	weighted	0.64 (0.08), 0.32	-0.60 (-0.74 – -0.47)	0.00 (-0.02 – 0.03)	0.23	yes	no
TN (mg/L)	EPT_Percent Taxa	unweighted	0.68 (0.04), 0.16	-0.56 (-0.63 – -0.48)	0.00 (0.00 – 0.01)	0.59	yes	yes
TN (mg/L)	EPT_Percent Taxa	weighted	0.72 (0.06), 0.22	-0.55 (-0.62 – -0.47)	0.00 (-0.01 – 0.01)	0.46	yes	no
TN (mg/L)	EPT_Taxa	unweighted	0.63 (0.04), 0.14	-27.25 (-30.69 – -23.81)	0.01 (-0.22 – 0.23)	0.60	yes	yes
TN (mg/L)	EPT_Taxa	weighted	0.62 (0.05), 0.21	-31.09 (-35.72 – -26.46)	-0.07 (-0.81 – 0.67)	0.41	yes	no
TN (mg/L)	H20	unweighted	1.06 (0.06), 0.25	-40.18 (-45.02 – -35.33)	0.29 (-0.22 – 0.80)	0.53	yes	yes
TN (mg/L)	H20	weighted	1.29 (0.12), 0.46	-32.65 (-36.85 – -28.46)	-0.14 (-1.64 – 1.37)	0.39	yes	no
TN (mg/L)	H21	unweighted	0.68 (0.05), 0.18	-58.63 (-67.41 – -49.85)	-0.19 (-0.72 – 0.34)	0.46	yes	yes
TN (mg/L)	H21	weighted	1.19 (0.12), 0.47	-35.13 (-40.14 – -30.12)	-0.33 (-2.02 – 1.37)	0.34	yes	no
TN (mg/L)	H23	unweighted	0.77 (0.04), 0.18	-56.32 (-63.40 – -49.25)	-0.18 (-0.69 – 0.33)	0.53	yes	yes
TN (mg/L)	H23	weighted	1.21 (0.11), 0.45	-34.86 (-39.56 – -30.16)	-0.31 (-1.90 – 1.27)	0.36	yes	no
TN (mg/L)	Intolerant_Percent	unweighted	0.54 (0.05), 0.19	-0.46 (-0.56 – -0.37)	0.00 (-0.01 – 0.00)	0.37	yes	no
TN (mg/L)	Intolerant_Percent	weighted	0.55 (0.06), 0.25	-0.50 (-0.61 – -0.40)	0.00 (-0.02 – 0.01)	0.26	yes	no
TN (mg/L)	Intolerant_PercentTaxa	unweighted	0.62 (0.04), 0.15	-0.57 (-0.65 – -0.50)	0.00 (-0.01 – 0.00)	0.57	yes	yes
TN (mg/L)	Intolerant_PercentTaxa	weighted	0.58 (0.05), 0.19	-0.65 (-0.75 – -0.55)	0.00 (-0.02 – 0.01)	0.41	yes	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
TN (mg/L)	Intolerant_Taxa	unweighted	0.52 (0.03), 0.13	-25.78 (-29.45 – -22.10)	-0.06 (-0.25 – 0.14)	0.53	yes	yes
TN (mg/L)	Intolerant_Taxa	weighted	0.51 (0.04), 0.18	-31.35 (-36.28 – -26.43)	-0.13 (-0.81 – 0.55)	0.37	yes	no
TN (mg/L)	O/E	unweighted	0.79 (0.09), 0.36	-0.46 (-0.57 – -0.35)	0.01 (0.00 – 0.02)	0.25	yes	no
TN (mg/L)	O/E	weighted	0.78 (0.09), 0.35	-0.64 (-0.77 – -0.51)	0.00 (-0.02 – 0.03)	0.29	yes	no
TN (mg/L)	propTaxaZHR	unweighted	0.84 (0.10), 0.37	-0.24 (-0.29 – -0.18)	0.00 (-0.01 – 0.00)	0.25	yes	no
TN (mg/L)	propTaxaZHR	weighted	0.70 (0.20), 0.77	-0.18 (-0.26 – -0.09)	-0.01 (-0.02 – 0.01)	0.06	no	no
TN (mg/L)	RAWDO50	unweighted	1.95 (0.19), 0.76	-0.11 (-0.13 – -0.09)	0.01 (0.00 – 0.01)	0.23	yes	no
TN (mg/L)	RAWDO50	weighted	1.89 (0.16), 0.62	-0.16 (-0.18 – -0.13)	0.02 (0.01 – 0.03)	0.33	yes	no
TN (mg/L)	RAWlowN	unweighted	0.94 (0.09), 0.37	-0.44 (-0.54 – -0.35)	0.01 (0.00 – 0.02)	0.24	yes	no
TN (mg/L)	RAWlowN	weighted	1.14 (0.23), 0.91	-0.32 (-0.41 – -0.22)	0.00 (-0.03 – 0.03)	0.09	no	no
TN (mg/L)	RAWlowP	unweighted	0.78 (0.08), 0.30	-0.53 (-0.65 – -0.42)	0.01 (0.00 – 0.01)	0.24	yes	no
TN (mg/L)	RAWlowP	weighted	0.85 (0.15), .59	-0.43 (-0.57 – -0.29)	0.00 (-0.02 – 0.03)	0.10	no	no
TN (mg/L)	RAWmeanZHR	unweighted	0.67 (0.07), 0.29	-0.44 (-0.55 – -0.34)	0.00 (-0.01 – 0.00)	0.25	yes	no
TN (mg/L)	RAWmeanZHR	weighted	0.90 (0.25), 0.97	-0.25 (-0.37 – -0.13)	-0.01 (-0.03 – 0.02)	0.06	no	no
TN (mg/L)	RAWNhet	unweighted	1.95 (0.22), 0.85	0.13 (0.11 – 0.16)	0.00 (-0.01 – 0.00)	0.23	yes	no
TN (mg/L)	RAWNhet	weighted	1.58 (0.20), 0.80	0.19 (0.15 – 0.22)	0.00 (-0.02 – 0.01)	0.21	no	no
TN (mg/L)	RAWprop GreenCRUS	unweighted	0.60 (0.07), 0.29	0.85 (0.61 – 1.08)	0.02 (0.01 – 0.03)	0.26	yes	no
TN (mg/L)	RAWprop GreenCRUS	weighted	0.58 (0.12), 0.46	0.54 (0.35 – 0.74)	0.02 (0.00 – 0.05)	0.11	no	no
TN (mg/L)	RAWsapro	unweighted	1.04 (0.11), 0.44	-0.33 (-0.40 – - 0.25)	0.01 (0.00 – 0.01)	0.22	yes	no
TN (mg/L)	RAWsapro	weighted	1.29 (0.19), 0.73	-0.28 (-0.34 – -0.22)	0.01 (-0.01 – 0.03)	0.17	no	no
TN (mg/L)	S2	unweighted	0.83 (0.06), 0.24	-52.74 (-60.75 – -44.72)	-0.80 (-1.43 – -0.17)	0.46	yes	yes
TN (mg/L)	S2	weighted	0.93 (0.14), 0.53	-34.70 (-43.39 – -26.01)	-1.38 (-3.23 – 0.46)	0.20	no	no
TN (mg/L)	Shannon_ Diversity	unweighted	0.76 (0.08), 0.30	-1.29 (-1.56 – -1.03)	-0.02 (-0.04 – 0.01)	0.36	yes	no
TN (mg/L)	Shannon_ Diversity	weighted	0.75 (0.11), 0.43	-1.35 (-1.71 – -1.00)	-0.01 (-0.07 – 0.05)	0.21	no	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
N (mg/L)	Taxonomic_Richness	unweighted	0.71 (0.05), 0.19	-31.26 (-35.72 – -26.80)	-0.07 (-0.40 – 0.27)	0.54	yes	yes
TN (mg/L)	Taxonomic_Richness	weighted	0.71 (0.07), 0.27	-33.33 (-38.93 – -27.72)	-0.11 (-1.08 – 0.86)	0.37	yes	no
TN (mg/L)	Tolerant_Percent	unweighted	0.61 (0.07), 0.28	0.48 (0.36 – 0.61)	-0.01 (-0.01 – 0.00)	0.24	yes	no
TN (mg/L)	Tolerant_Percent	weighted	0.74 (0.09), 0.37	0.38 (0.29 – 0.47)	0.00 (-0.02 – 0.01)	0.25	no	no
TN (mg/L)	Tolerant_PercentTaxa	unweighted	0.67 (0.05), 0.18	0.42 (0.35 – 0.48)	0.00 (-0.01 – 0.00)	0.50	yes	yes
TN (mg/L)	Tolerant_PercentTaxa	weighted	0.73 (0.06), 0.23	0.41 (0.35 – 0.46)	0.00 (-0.01 – 0.01)	0.45	yes	no
TP (mg/L)	CSCI	unweighted	0.15 (0.02), 0.06	-2.70 (-3.29 – -2.11)	0.00 (-0.06 – 0.06)	0.28	yes	no
TP (mg/L)	CSCI	weighted	0.15 (0.03), 0.10	-2.49 (-3.21 – -1.76)	-0.12 (-0.24 – 0.01)	0.23	no	no
TP (mg/L)	D18	unweighted	0.12 (0.01), 0.03	-352.10 (-406.00 – -298.20)	-4.04 (-9.01 – 0.93)	0.41	yes	yes
TP (mg/L)	D18	weighted	0.14 (0.01), 0.05	-290.00 (-335.00 – -245.00)	-9.16 (-19.92 – 1.60)	0.35	yes	no
TP (mg/L)	EPT_PercentTaxa	unweighted	0.10 (0.01), 0.04	-3.10 (-3.73 – -2.48)	-0.02 (-0.06 – 0.02)	0.37	yes	no
TP (mg/L)	EPT_PercentTaxa	weighted	0.12 (0.01), 0.06	-2.43 (-3.03 – -1.83)	-0.10 (-0.18 – 0.02)	0.30	no	no
TP (mg/L)	EPT_Taxa	unweighted	0.11 (0.01), 0.04	-129.00 (-157.40 – -100.50)	-0.74 (-2.49 – 1.02)	0.34	yes	no
TP (mg/L)	EPT_Taxa	weighted	0.14 (0.02), 0.08	-105.90 (-132.40 – -79.39)	-3.30 (-7.84 – 1.25)	0.25	no	no
TP (mg/L)	H20	unweighted	0.11 (0.01), 0.03	-369.00 (-420.80 – -317.20)	-3.79 (-7.96 – 0.38)	0.50	yes	yes
TP (mg/L)	H20	weighted	0.13 (0.01), 0.05	-275.10 (-315.90 – -234.30)	-6.31 (-15.21 – 2.59)	0.38	yes	no
TP (mg/L)	H21	unweighted	0.11 (0.01), 0.03	-372.10 (-434.90 – -309.30)	-2.74 (-7.41 – 1.93)	0.42	yes	no
TP (mg/L)	H21	weighted	0.14 (0.01), 0.05	-272.00 (-318.00 – -226.10)	-7.26 (-17.59 – 3.08)	0.32	yes	no
TP (mg/L)	H23	unweighted	0.11 (0.01), 0.03	-371.90 (-426.40 – -317.30)	-2.34 (-6.92 – 2.24)	0.46	yes	yes
TP (mg/L)	H23	weighted	0.14 (0.01), 0.05	-266.10 (-309.70 – -222.60)	-5.41 (-15.22 – 4.40)	0.34	yes	no
TP (mg/L)	Intolerant_PercentTaxa	unweighted	0.11 (0.01), 0.05	-2.60 (-3.22 – -1.97)	-0.02 (-0.06 – 0.02)	0.30	yes	no
TP (mg/L)	Intolerant_PercentTaxa	weighted	0.12 (0.02), 0.07	-2.24 (-2.92 – -1.57)	-0.08 (-0.17 – 0.01)	0.22	no	no
TP (mg/L)	Intolerant_Taxa	unweighted	0.11 (0.01), 0.05	-96.90 (-121.90 – -71.91)	-0.57 (-2.11 – 0.98)	0.27	yes	no
TP (mg/L)	Intolerant_Taxa	weighted	0.12 (0.02), 0.08	-93.62 (-123.20 – -64.03)	-2.62 (-6.70 – 1.46)	0.19	no	no

Table C.2 (continued)

Gradient	Response	Analysis Type	Breakpoint (SE), 95% Confidence Interval Width	Slope 1 (95% Confidence Interval)	Slope 2 (95% Confidence Interval)	Adjusted R ²	all 4 criteria (relaxed) fulfilled?	all 4 criteria (strict) fulfilled?
TP (mg/L)	propAchMin	unweighted	0.04 (0.00), 0.01	-7.42 (-9.19 – -5.66)	-0.05 (-0.10 – -0.01)	0.21	yes	no
TP (mg/L)	propAchMin	weighted	0.08 (0.01), 0.04	-3.83 (-4.81 – -2.85)	-0.01 (-0.13 – 0.11)	0.16	yes	no
TP (mg/L)	RAWDO100	unweighted	0.06 (0.01), 0.02	-5.75 (-7.43 – -4.08)	0.00 (-0.06 – 0.07)	0.15	yes	no
TP (mg/L)	RAWDO100	weighted	0.03 (0.00), 0.01	-10.88 (-13.43 – -8.33)	0.01 (-0.12 – 0.13)	0.15	yes	no
TP (mg/L)	RAWDO50	unweighted	0.27 (0.03), 0.10	-0.86 (-1.01 – -0.72)	0.04 (0.00 – 0.07)	0.27	yes	no
TP (mg/L)	RAWDO50	weighted	0.29 (0.03), 0.11	-0.89 (-1.03 – -0.75)	0.04 (-0.03 – 0.11)	0.32	yes	no
TP (mg/L)	RAWeutro	unweighted	0.08 (0.01), 0.03	4.47 (3.33 – 5.60)	-0.02 (-0.09 – 0.04)	0.19	yes	no
TP (mg/L)	RAWeutro	weighted	0.08 (0.01), 0.04	3.96 (2.70 – 5.22)	-0.08 (-0.24 – 0.07)	0.09	no	no
TP (mg/L)	RAWlowN	unweighted	0.09 (0.01), 0.03	-5.53 (-6.58 – -4.48)	-0.02 (-0.09 – 0.04)	0.32	yes	no
TP (mg/L)	RAWlowN	weighted	0.09 (0.01), 0.04	-4.73 (-5.90 – -3.55)	-0.05 (-0.21 – 0.12)	0.17	yes	no
TP (mg/L)	RAWlowP	unweighted	0.08 (0.01), 0.02	-6.78 (-7.97 – -5.60)	-0.03 (-0.10 – 0.03)	0.35	yes	yes
TP (mg/L)	RAWlowP	weighted	0.08 (0.01), 0.03	-6.02 (-7.24 – -4.81)	-0.03 (-0.20 – 0.13)	0.22	yes	no
TP (mg/L)	RAWlowTPsp	unweighted	0.07 (0.01), 0.03	-1.99 (-2.49 – -1.49)	0.00 (-0.03 – 0.02)	0.20	yes	no
TP (mg/L)	RAWlowTPsp	weighted	0.06 (0.01), 0.03	-2.01 (-2.71 – -1.32)	0.01 (-0.05 – 0.07)	0.09	no	no
TP (mg/L)	RAWsapro	unweighted	0.15 (0.01), 0.06	-2.56 (-3.05 – -2.06)	0.02 (-0.05 – 0.08)	0.25	yes	no
TP (mg/L)	RAWsapro	weighted	0.15 (0.02), 0.07	-2.46 (-2.98 – -1.94)	-0.01 (-0.14 – 0.12)	0.21	yes	no
TP (mg/L)	S2	unweighted	0.11 (0.01), 0.04	-374.80 (-453.90 – -295.70)	-3.11 (-8.62 – 2.40)	0.30	yes	no
TP (mg/L)	S2	weighted	0.11 (0.02), 0.07	-249.10 (-319.60 – -178.60)	-0.72 (-12.58 – 11.15)	0.14	no	no
TP (mg/L)	Taxonomic_Richness	unweighted	0.14 (0.02), 0.06	-129.10 (-157.00 – -101.20)	-0.73 (-3.34 – 1.87)	0.29	yes	no
TP (mg/L)	Taxonomic_Richness	weighted	0.15 (0.03), 0.10	-109.90 (-143.40 – -76.43)	-6.32 (-12.15 – -0.49)	0.22	no	no
TP (mg/L)	Tolerant_PercentTaxa	unweighted	0.11 (0.01), 0.04	2.17 (1.67 – 2.67)	0.00 (-0.03 – 0.03)	0.30	yes	no
TP (mg/L)	Tolerant_PercentTaxa	weighted	0.16 (0.02), 0.08	1.43 (1.08 – 1.78)	0.06 (0.00 – 0.12)	0.29	no	no

Table C.3. Shown are 1) a summary of thresholds, from all analyses, for the chlorophyll a , AFDM, TN, and TP gradients, and 2) mean and distributions of ALI values among sites that had gradient values below and above the indicated threshold. In addition, interquartile ranges (IQRs) of the ALI distributions are provided (see Figure 3.21 for example graphical depictions based on piecewise-regression-derived thresholds). To provide perspective on where each threshold lies relative to the full range of corresponding gradient values across the data set as a whole, maximum values are as follows: Chlorophyll a = 1504 mg/m², AFDM = 405 g/m², TN = 26.4 mg/L, TP = 5.4 mg/L. Note that in most cases, the threshold is far below the maximum for the gradient in question. As such, when normalized for the proportion of the range of gradient values represented by each bin (i.e., “below” vs. “above” the threshold), the IQR is, relatively speaking, greater for the sites with gradient values below the threshold than for those above.

Table C.3.		Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile			IQR
Gradient	25 th								50 th (median)	75 th	max	
AFDM (g/m ²)	BMI	BMIcommunity	CART	30.84	below	-0.19	-1.39	-0.65	-0.25	0.21	1.64	0.86
		BMIcommunity	CART	30.84	above	0.47	-0.88	0.05	0.47	0.82	1.63	0.76
		BMIcommunity	ncpa.bc	10.86	below	-0.35	-1.39	-0.74	-0.39	-0.04	1.62	0.70
		BMIcommunity	ncpa.bc	10.86	above	0.35	-0.96	-0.01	0.32	0.70	1.64	0.70
		BMIcommunity	ncpa.euc	11.42	below	-0.34	-1.39	-0.74	-0.39	-0.01	1.62	0.72
		BMIcommunity	ncpa.euc	11.42	above	0.36	-0.96	0.00	0.32	0.71	1.64	0.71
		BMIcommunity	TITAN.decreasers	7.05	below	-0.41	-1.39	-0.80	-0.44	-0.12	1.62	0.68
		BMIcommunity	TITAN.decreasers	7.05	above	0.22	-1.21	-0.17	0.20	0.65	1.64	0.82
		CSCI	pieciseregression	15.02	below	0.92	0.21	0.80	0.98	1.09	1.27	0.29
		CSCI	pieciseregression	15.02	above	0.70	0.10	0.51	0.68	0.93	1.22	0.41
		CSCI	SiZer	6.96	below	0.97	0.21	0.91	1.02	1.11	1.27	0.20
		CSCI	SiZer	6.96	above	0.76	0.10	0.58	0.77	0.98	1.23	0.40
		EPT_Percent	BRT_exhaustion	25.00	below	0.41	0.00	0.22	0.42	0.60	0.92	0.38
		EPT_Percent	BRT_exhaustion	25.00	above	0.18	0.00	0.01	0.11	0.34	0.77	0.33
		EPT_Percent	pieciseregression	16.10	below	0.43	0.00	0.26	0.45	0.62	0.90	0.36
		EPT_Percent	pieciseregression	16.10	above	0.19	0.00	0.01	0.12	0.33	0.92	0.32
		EPT_Percent	SiZer	3.94	below	0.56	0.00	0.44	0.59	0.71	0.86	0.27
		EPT_Percent	SiZer	3.94	above	0.31	0.00	0.09	0.29	0.47	0.92	0.38
		EPT_PercentTaxa	pieciseregression	18.93	below	0.44	0.00	0.33	0.49	0.57	0.75	0.23
		EPT_PercentTaxa	pieciseregression	18.93	above	0.23	0.00	0.10	0.22	0.35	0.67	0.24
EPT_PercentTaxa	SiZer	5.92	below	0.50	0.00	0.46	0.54	0.59	0.75	0.13		
EPT_PercentTaxa	SiZer	5.92	above	0.32	0.00	0.17	0.33	0.47	0.72	0.30		
EPT_Taxa	pieciseregression	17.47	below	15.17	0.00	9.00	17.00	21.00	34.00	12.00		
EPT_Taxa	pieciseregression	17.47	above	5.67	0.00	1.75	4.00	8.00	29.00	6.25		
EPT_Taxa	SiZer	7.92	below	17.13	0.00	12.00	19.00	23.00	34.00	11.00		
EPT_Taxa	SiZer	7.92	above	8.17	0.00	2.00	6.00	13.00	29.00	11.00		
AFDM (g/m ²)	BMI	Intolerant_Percent	pieciseregression	12.99	below	0.20	0.00	0.05	0.19	0.30	0.72	0.26
		Intolerant_Percent	pieciseregression	12.99	above	0.05	0.00	0.00	0.00	0.04	0.61	0.04
		Intolerant_Percent	SiZer	5.94	below	0.25	0.00	0.12	0.25	0.37	0.72	0.25
		Intolerant_Percent	SiZer	5.94	above	0.08	0.00	0.00	0.01	0.13	0.61	0.13
		Intolerant_PercentTaxa	BRT_exhaustion	12.00	below	0.30	0.00	0.18	0.32	0.43	0.62	0.25

Table C.3.								percentile					
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min						
								25 th	50 th (median)	75 th	max	IQR	
			Intolerant_PercentTaxa	BRT_exhaustion	12.00	above	0.09	0.00	0.00	0.04	0.14	0.56	0.14
			Intolerant_PercentTaxa	BRT_resistance	7.00	below	0.33	0.00	0.24	0.35	0.45	0.62	0.21
			Intolerant_PercentTaxa	BRT_resistance	7.00	above	0.13	0.00	0.00	0.07	0.23	0.58	0.23
			Intolerant_PercentTaxa	piecisewiregression	15.41	below	0.28	0.00	0.13	0.31	0.42	0.62	0.29
			Intolerant_PercentTaxa	piecisewiregression	15.41	above	0.08	0.00	0.00	0.00	0.12	0.56	0.12
			Intolerant_PercentTaxa	SiZer	5.94	below	0.35	0.00	0.26	0.38	0.46	0.62	0.20
			Intolerant_PercentTaxa	SiZer	5.94	above	0.15	0.00	0.00	0.09	0.26	0.58	0.26
			Intolerant_Taxa	piecisewiregression	15.29	below	9.99	0.00	3.00	10.00	16.00	25.00	13.00
			Intolerant_Taxa	piecisewiregression	15.29	above	2.38	0.00	0.00	0.00	3.00	24.00	3.00
			Intolerant_Taxa	SiZer	5.94	below	12.60	0.00	8.00	14.00	17.00	25.00	9.00
			Intolerant_Taxa	SiZer	5.94	above	4.82	0.00	0.00	2.00	8.00	24.00	8.00
			OoverE	SiZer	8.91	below	1.01	0.22	0.88	1.04	1.16	1.42	0.28
			OoverE	SiZer	8.91	above	0.82	0.15	0.64	0.80	1.02	1.39	0.38
			Shannon_Diversity	SiZer	6.93	below	2.38	0.08	1.99	2.52	2.85	3.50	0.86
			Shannon_Diversity	SiZer	6.93	above	1.92	0.03	1.53	1.92	2.36	3.38	0.83
			Simpson_Diversity	SiZer	4.95	below	0.81	0.32	0.77	0.86	0.90	0.95	0.12
			Simpson_Diversity	SiZer	4.95	above	0.73	0.01	0.66	0.77	0.86	0.95	0.20
			Taxonomic_Richness	SiZer	9.90	below	32.96	1.00	27.00	35.00	40.00	54.00	13.00
			Taxonomic_Richness	SiZer	9.90	above	21.81	3.00	14.00	20.00	28.00	53.00	14.00
			Tolerant_Percent	SiZer	7.92	below	0.06	0.00	0.01	0.02	0.05	0.93	0.04
			Tolerant_Percent	SiZer	7.92	above	0.23	0.00	0.04	0.12	0.36	0.97	0.32
			Tolerant_PercentTaxa	piecisewiregression	28.40	below	0.18	0.00	0.08	0.14	0.25	0.73	0.17
			Tolerant_PercentTaxa	piecisewiregression	28.40	above	0.35	0.06	0.25	0.33	0.46	0.77	0.21
			Tolerant_PercentTaxa	SiZer	5.94	below	0.13	0.00	0.06	0.10	0.16	0.70	0.10
			Tolerant_PercentTaxa	SiZer	5.94	above	0.26	0.00	0.14	0.24	0.37	0.77	0.23
			Tolerant_Taxa	SiZer	5.94	below	3.78	0.00	2.00	4.00	5.00	11.00	3.00
			Tolerant_Taxa	SiZer	5.94	above	5.44	0.00	4.00	5.00	7.00	12.00	3.00
			D18	BRT_exhaustion	35.00	below	70.55	2.00	60.00	76.00	88.00	100.00	28.00
			D18	BRT_exhaustion	35.00	above	48.71	0.00	30.00	52.00	66.50	96.00	36.50
			D18	piecisewiregression	31.72	below	71.49	2.00	60.00	76.00	88.00	100.00	28.00
	diatom		D18	piecisewiregression	31.72	above	48.69	0.00	30.00	52.00	68.00	96.00	38.00
			D18	SiZer	7.30	below	78.46	10.00	72.00	82.00	92.00	100.00	20.00
			D18	SiZer	7.30	above	57.23	0.00	42.00	62.00	76.00	100.00	34.00
			diatomCommunity	CART	18.51	below	-0.24	-1.48	-0.68	-0.32	0.16	1.71	0.84
			diatomCommunity	CART	18.51	above	0.32	-1.26	-0.09	0.40	0.76	1.66	0.85
			diatomCommunity	ncpa.bc	12.71	below	-0.30	-1.48	-0.74	-0.37	0.08	1.56	0.81
			diatomCommunity	ncpa.bc	12.71	above	0.29	-1.26	-0.12	0.33	0.74	1.71	0.86
			diatomCommunity	ncpa.euc	7.80	below	-0.38	-1.38	-0.78	-0.45	-0.07	1.36	0.71
			diatomCommunity	ncpa.euc	7.80	above	0.21	-1.48	-0.24	0.25	0.66	1.71	0.91
			diatomCommunity	TITAN.decreasers	10.19	below	-0.33	-1.38	-0.74	-0.41	0.01	1.36	0.75
AFDM (g/m ²)	diatom		diatomCommunity	TITAN.decreasers	10.19	above	0.25	-1.48	-0.18	0.30	0.71	1.71	0.90
			propAchMin	piecisewiregression	11.17	below	0.24	0.00	0.05	0.16	0.37	0.97	0.31
			propAchMin	piecisewiregression	11.17	above	0.10	0.00	0.00	0.02	0.13	0.85	0.13
			propAchMin	SiZer	6.22	below	0.27	0.00	0.07	0.20	0.42	0.97	0.35
			propAchMin	SiZer	6.22	above	0.12	0.00	0.00	0.04	0.16	0.87	0.15
			RAWDO100	piecisewiregression	7.10	below	0.43	0.00	0.19	0.40	0.63	1.00	0.44

Table C.3.												
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile				
								25 th	50 th (median)	75 th	max	IQR
		RAWDO100	pieciseregression	7.10	above	0.27	0.00	0.06	0.19	0.40	1.00	0.34
		RAWDO100	SiZer	6.35	below	0.44	0.00	0.21	0.41	0.64	1.00	0.44
		RAWDO100	SiZer	6.35	above	0.27	0.00	0.06	0.19	0.40	1.00	0.34
		RAWDO50	SiZer	8.30	below	0.96	0.45	0.96	0.99	1.00	1.00	0.04
		RAWDO50	SiZer	8.30	above	0.87	0.11	0.82	0.93	0.98	1.00	0.16
		RAWeutro	SiZer	6.25	below	0.48	0.01	0.28	0.48	0.70	1.00	0.42
		RAWeutro	SiZer	6.25	above	0.64	0.00	0.48	0.68	0.87	1.00	0.40
		RAWlowN	SiZer	4.15	below	0.52	0.00	0.28	0.53	0.75	1.00	0.48
		RAWlowN	SiZer	4.15	above	0.32	0.00	0.06	0.24	0.51	1.00	0.46
		RAWlowP	SiZer	4.15	below	0.53	0.00	0.28	0.54	0.81	1.00	0.52
		RAWlowP	SiZer	4.15	above	0.32	0.00	0.05	0.23	0.52	1.00	0.46
		RAWNhet	SiZer	6.25	below	0.08	0.00	0.00	0.03	0.09	0.73	0.08
		RAWNhet	SiZer	6.25	above	0.19	0.00	0.04	0.12	0.26	0.99	0.22
		H20	BRT_exhaustion	35.00	below	65.47	9.00	55.00	70.00	79.00	100.00	24.00
		H20	BRT_exhaustion	35.00	above	43.96	2.00	28.00	44.00	59.50	92.00	31.50
		H20	pieciseregression	36.06	below	65.30	8.00	55.00	70.00	79.00	100.00	24.00
		H20	pieciseregression	36.06	above	43.89	2.00	28.25	44.00	59.00	92.00	30.75
		H20	SiZer	4.50	below	74.00	18.00	69.00	76.00	85.00	100.00	16.00
		H20	SiZer	4.50	above	56.00	2.00	40.00	59.00	74.00	95.00	34.00
		H21	pieciseregression	34.75	below	65.25	4.00	53.00	69.00	81.00	100.00	28.00
	hybrid	H21	pieciseregression	34.75	above	42.36	0.00	28.50	43.50	56.00	93.00	27.50
		H21	SiZer	5.60	below	74.00	11.00	67.00	79.00	87.00	100.00	20.00
		H21	SiZer	5.60	above	53.00	0.00	37.00	54.00	71.00	100.00	34.00
		H23	pieciseregression	35.79	below	68.60	10.00	56.00	74.00	85.00	100.00	29.00
		H23	pieciseregression	35.79	above	44.40	5.00	29.00	42.00	59.50	90.00	30.50
		H23	SiZer	8.29	below	75.41	16.00	68.75	80.00	88.25	100.00	19.50
		H23	SiZer	8.29	above	54.24	5.00	39.00	55.00	72.00	99.00	33.00
	soft	propTaxaZHR	SiZer	27.98	below	0.22	0.00	0.10	0.21	0.33	0.71	0.23
		propTaxaZHR	SiZer	27.98	above	0.14	0.00	0.00	0.11	0.23	0.67	0.23
		RAWlowTPsp	SiZer	33.16	below	0.11	0.00	0.00	0.09	0.18	0.57	0.18
		RAWlowTPsp	SiZer	33.16	above	0.07	0.00	0.00	0.00	0.11	0.50	0.11
		RAWmeanZHR	SiZer	33.16	below	0.29	0.00	0.06	0.22	0.52	0.86	0.46
		RAWmeanZHR	SiZer	33.16	above	0.14	0.00	0.00	0.06	0.19	0.79	0.19
		RAWpropBiovolChlor	SiZer	9.50	below	0.30	0.00	0.00	0.07	0.61	1.00	0.61
		RAWpropBiovolChlor	SiZer	9.50	above	0.56	0.00	0.06	0.66	0.98	1.00	0.92
AFDM (g/m ²)	soft	RAWpropBiovolZHR	SiZer	34.38	below	0.36	0.00	0.00	0.19	0.74	1.00	0.74
		RAWpropBiovolZHR	SiZer	34.38	above	0.14	0.00	0.00	0.00	0.10	1.00	0.10
		RAWpropGreenCRUS	SiZer	11.40	below	0.15	0.00	0.00	0.00	0.00	1.00	0.00
		RAWpropGreenCRUS	SiZer	11.40	above	0.47	0.00	0.00	0.45	0.99	1.00	0.99
		S2	pieciseregression	39.33	below	58.56	0.00	42.00	63.00	78.00	100.00	36.00
		S2	pieciseregression	39.33	above	37.33	0.00	17.00	33.00	57.00	97.00	40.00
		S2	SiZer	10.00	below	64.23	0.00	55.00	68.00	78.00	98.00	23.00
		S2	SiZer	10.00	above	45.04	0.00	22.00	42.00	67.75	100.00	45.75
		BMIcommunity	CART	23.63	below	-0.20	-1.39	-0.65	-0.27	0.20	1.62	0.86
		BMIcommunity	CART	23.63	above	0.34	-1.12	0.00	0.32	0.68	1.64	0.68

Table C.3.								percentile				
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min					
								25 th	50 th (median)	75 th	max	IQR
Chlorophyll <i>a</i> (mg/m ²)	BMI	BMIcommunity	ncpa.bc	17.15	below	-0.27	-1.39	-0.68	-0.33	0.10	1.62	0.78
		BMIcommunity	ncpa.bc	17.15	above	0.32	-1.12	-0.05	0.32	0.68	1.64	0.72
		BMIcommunity	ncpa.euc	15.98	below	-0.28	-1.39	-0.70	-0.34	0.11	1.62	0.81
		BMIcommunity	ncpa.euc	15.98	above	0.32	-1.12	-0.07	0.30	0.67	1.64	0.74
		BMIcommunity	TITAN.decreasers	16.51	below	-0.27	-1.39	-0.68	-0.32	0.11	1.62	0.79
		BMIcommunity	TITAN.decreasers	16.51	above	0.32	-1.12	-0.06	0.31	0.67	1.64	0.74
		CSCI	SiZer	11.65	below	0.94	0.21	0.83	1.01	1.10	1.24	0.27
		CSCI	SiZer	11.65	above	0.76	0.10	0.59	0.76	0.97	1.27	0.38
		EPT_Percent	SiZer	3.97	below	0.49	0.00	0.35	0.55	0.69	0.88	0.34
		EPT_Percent	SiZer	3.97	above	0.32	0.00	0.11	0.30	0.50	0.92	0.39
		EPT_PercentTaxa	SiZer	11.41	below	0.46	0.00	0.36	0.50	0.58	0.75	0.21
		EPT_PercentTaxa	SiZer	11.41	above	0.31	0.00	0.17	0.29	0.45	0.72	0.28
		EPT_Taxa	pieciseregression	68.77	below	13.51	0.00	6.00	13.00	20.00	34.00	14.00
		EPT_Taxa	pieciseregression	68.77	above	3.49	0.00	1.00	3.00	5.00	15.00	4.00
		EPT_Taxa	SiZer	11.41	below	15.56	0.00	10.00	17.00	21.00	34.00	11.00
	EPT_Taxa	SiZer	11.41	above	8.91	0.00	2.00	7.00	14.00	29.00	12.00	
	Intolerant_Percent	pieciseregression	23.29	below	0.18	0.00	0.02	0.16	0.29	0.72	0.27	
	Intolerant_Percent	pieciseregression	23.29	above	0.05	0.00	0.00	0.00	0.04	0.61	0.04	
	Intolerant_Percent	SiZer	6.56	below	0.23	0.00	0.06	0.22	0.34	0.72	0.28	
	Intolerant_Percent	SiZer	6.56	above	0.10	0.00	0.00	0.02	0.16	0.61	0.16	
	Intolerant_PercentTaxa	pieciseregression	31.40	below	0.26	0.00	0.08	0.28	0.40	0.62	0.32	
	Intolerant_PercentTaxa	pieciseregression	31.40	above	0.07	0.00	0.00	0.00	0.10	0.51	0.10	
	Intolerant_PercentTaxa	SiZer	11.65	below	0.29	0.00	0.18	0.32	0.43	0.62	0.25	
	Intolerant_PercentTaxa	SiZer	11.65	above	0.13	0.00	0.00	0.06	0.23	0.58	0.23	
	Intolerant_Taxa	pieciseregression	31.43	below	9.11	0.00	2.00	9.00	15.00	25.00	13.00	
	Intolerant_Taxa	pieciseregression	31.43	above	2.19	0.00	0.00	0.00	3.00	19.00	3.00	
	Intolerant_Taxa	SiZer	11.65	below	10.41	0.00	5.00	10.00	16.00	25.00	11.00	
	Intolerant_Taxa	SiZer	11.65	above	4.29	0.00	0.00	1.00	6.50	24.00	6.50	
	OoverE	SiZer	11.65	below	0.99	0.31	0.86	1.03	1.16	1.39	0.31	
	OoverE	SiZer	11.65	above	0.84	0.15	0.65	0.86	1.03	1.42	0.38	
Chlorophyll <i>a</i> (mg/m ²)	BMI	Shannon_Diversity	SiZer	30.90	below	2.24	0.08	1.84	2.33	2.71	3.50	0.87
		Shannon_Diversity	SiZer	30.90	above	1.70	0.03	1.28	1.72	2.15	3.14	0.87
		Simpson_Diversity	SiZer	27.67	below	0.78	0.03	0.72	0.82	0.88	0.95	0.17
		Simpson_Diversity	SiZer	27.67	above	0.69	0.01	0.62	0.73	0.81	0.93	0.19
		Taxonomic_Richness	BRT_exhaustion	60.00	below	30.00	1.00	22.00	31.00	38.00	54.00	16.00
		Taxonomic_Richness	BRT_exhaustion	60.00	above	15.72	3.00	9.00	15.00	19.00	45.00	10.00
		Taxonomic_Richness	SiZer	14.89	below	31.93	4.00	25.00	34.00	40.00	54.00	15.00
		Taxonomic_Richness	SiZer	14.89	above	22.28	1.00	14.00	21.00	30.00	54.00	16.00
		Tolerant_Percent	SiZer	12.14	below	0.08	0.00	0.01	0.03	0.08	0.90	0.07
		Tolerant_Percent	SiZer	12.14	above	0.22	0.00	0.04	0.11	0.34	0.97	0.31
		Tolerant_PercentTaxa	SiZer	12.14	below	0.16	0.00	0.08	0.12	0.22	0.70	0.14
		Tolerant_PercentTaxa	SiZer	12.14	above	0.27	0.00	0.16	0.25	0.38	0.77	0.22
		Tolerant_Taxa	SiZer	12.14	below	4.34	0.00	3.00	4.00	6.00	12.00	3.00
		Tolerant_Taxa	SiZer	12.14	above	5.36	0.00	3.00	5.00	7.00	12.00	4.00
		diatom	D18	pieciseregression	98.82	below	68.21	2.00	54.00	74.00	86.00	100.00

Table C.3.												
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile				
								25 th	50 th (median)	75 th	max	IQR
		D18	pieciseregression	98.82	above	43.76	0.00	24.00	48.00	64.00	96.00	40.00
		D18	SiZer	11.65	below	74.61	2.00	66.00	80.00	90.00	100.00	24.00
		D18	SiZer	11.65	above	57.15	0.00	40.00	61.00	76.00	100.00	36.00
		diatomCommunity	CART	23.63	below	-0.21	-1.38	-0.66	-0.26	0.22	1.50	0.88
		diatomCommunity	CART	23.63	above	0.29	-1.48	-0.17	0.38	0.77	1.71	0.95
		diatomCommunity	ncpa.bc	26.73	below	-0.22	-1.38	-0.67	-0.27	0.22	1.50	0.89
		diatomCommunity	ncpa.bc	26.73	above	0.34	-1.48	-0.10	0.42	0.78	1.71	0.88
		diatomCommunity	ncpa.euc	26.73	below	-0.22	-1.38	-0.67	-0.27	0.22	1.50	0.89
		diatomCommunity	ncpa.euc	26.73	above	0.34	-1.48	-0.10	0.42	0.78	1.71	0.88
		diatomCommunity	TITAN.decreasers	16.06	below	-0.25	-1.38	-0.68	-0.30	0.14	1.50	0.82
		diatomCommunity	TITAN.decreasers	16.06	above	0.24	-1.48	-0.26	0.30	0.74	1.71	1.01
		propAchMin	SiZer	15.05	below	0.22	0.00	0.03	0.14	0.34	0.97	0.31
		propAchMin	SiZer	15.05	above	0.12	0.00	0.00	0.03	0.15	0.88	0.15
		RAWDO100	SiZer	7.77	below	0.41	0.00	0.16	0.35	0.60	1.00	0.44
		RAWDO100	SiZer	7.77	above	0.29	0.00	0.06	0.21	0.44	1.00	0.38
		RAWDO50	SiZer	18.77	below	0.94	0.35	0.93	0.98	1.00	1.00	0.06
		RAWDO50	SiZer	18.77	above	0.86	0.11	0.80	0.92	0.98	1.00	0.17
		RAWeutro	SiZer	31.07	below	0.54	0.01	0.34	0.56	0.75	1.00	0.42
	diatom	RAWeutro	SiZer	31.07	above	0.69	0.00	0.56	0.76	0.90	1.00	0.35
		RAWlowN	SiZer	42.72	below	0.42	0.00	0.16	0.38	0.66	1.00	0.50
		RAWlowN	SiZer	42.72	above	0.22	0.00	0.02	0.09	0.39	1.00	0.37
		RAWlowP	SiZer	42.72	below	0.42	0.00	0.15	0.38	0.68	1.00	0.52
		RAWlowP	SiZer	42.72	above	0.23	0.00	0.02	0.09	0.37	1.00	0.35
		RAWNhet	SiZer	11.97	below	0.10	0.00	0.01	0.04	0.13	0.69	0.12
	hybrid	RAWNhet	SiZer	11.97	above	0.20	0.00	0.05	0.13	0.29	0.99	0.24
		H20	pieciseregression	104.54	below	63.44	9.00	51.00	69.00	79.00	100.00	28.00
		H20	pieciseregression	104.54	above	37.57	2.00	21.00	38.00	49.75	88.00	28.75
		H20	SiZer	11.65	below	70.37	10.00	64.00	74.00	82.00	100.00	18.00
		H20	SiZer	11.65	above	52.55	2.00	36.00	55.00	71.00	95.00	35.00
		H21	pieciseregression	93.13	below	63.35	1.00	50.00	67.00	80.00	100.00	30.00
		H21	pieciseregression	93.13	above	36.69	0.00	20.50	39.00	49.00	93.00	28.50
		H21	SiZer	11.65	below	70.75	6.00	60.50	74.00	84.00	100.00	23.50
		H21	SiZer	11.65	above	51.24	0.00	34.00	51.00	69.00	100.00	35.00
		H23	pieciseregression	102.55	below	66.49	9.00	52.00	71.00	84.00	100.00	32.00
	soft	H23	pieciseregression	102.55	above	37.77	5.00	22.00	39.00	49.00	90.00	27.00
		H23	SiZer	11.65	below	74.01	16.00	66.00	78.00	88.00	98.00	22.00
		H23	SiZer	11.65	above	54.39	5.00	38.00	56.00	72.00	100.00	34.00
		propTaxaZHR	SiZer	34.95	below	0.22	0.00	0.10	0.22	0.33	0.71	0.23
		propTaxaZHR	SiZer	34.95	above	0.13	0.00	0.00	0.10	0.22	0.62	0.22
		RAWlowTPsp	SiZer	34.95	below	0.12	0.00	0.00	0.10	0.18	0.57	0.18
		RAWlowTPsp	SiZer	34.95	above	0.07	0.00	0.00	0.04	0.10	0.43	0.10
	BMI	RAWmeanZHR	SiZer	42.72	below	0.30	0.00	0.07	0.24	0.53	0.86	0.47
		RAWmeanZHR	SiZer	42.72	above	0.12	0.00	0.00	0.06	0.16	0.76	0.16
		RAWpropBiovolChlor	SiZer	19.42	below	0.32	0.00	0.00	0.09	0.64	1.00	0.64
		RAWpropBiovolChlor	SiZer	19.42	above	0.60	0.00	0.11	0.77	0.99	1.00	0.88

Table C.3.									percentile				
Gradient	Assemblage	ALI	analysis type	threshold	gradient value to threshold	mean	min						
								25 th	50 th (median)	75 th	max	IQR	
TN (mg/L)	BMI	RAWpropBiovolZHR	SiZer	42.72	below	0.38	0.00	0.00	0.22	0.80	1.00	0.80	
		RAWpropBiovolZHR	SiZer	42.72	above	0.12	0.00	0.00	0.00	0.09	1.00	0.09	
		RAWpropGreenCRUS	pieciseregression	103.37	below	0.25	0.00	0.00	0.00	0.51	1.00	0.51	
		RAWpropGreenCRUS	pieciseregression	103.37	above	0.71	0.00	0.40	0.95	1.00	1.00	0.60	
		RAWpropGreenCRUS	SiZer	19.42	below	0.16	0.00	0.00	0.00	0.00	1.00	0.00	
		RAWpropGreenCRUS	SiZer	19.42	above	0.50	0.00	0.00	0.57	0.99	1.00	0.99	
		S2	pieciseregression	113.46	below	57.80	0.00	40.00	62.00	78.00	100.00	38.00	
		S2	pieciseregression	113.46	above	27.38	2.00	13.00	20.00	35.00	87.00	22.00	
		S2	SiZer	19.42	below	63.46	0.00	52.00	67.00	78.00	100.00	26.00	
		S2	SiZer	19.42	above	42.55	0.00	20.00	37.00	65.00	97.00	45.00	
	BMI	BMIcommunity	CART	0.29	below	-0.36	-1.39	-0.74	-0.39	-0.02	1.12	0.72	
		BMIcommunity	CART	0.29	above	0.36	-1.34	0.00	0.35	0.77	1.64	0.77	
		BMIcommunity	ncpa.bc	0.32	below	-0.34	-1.39	-0.71	-0.35	0.04	1.12	0.75	
		BMIcommunity	ncpa.bc	0.32	above	0.40	-1.34	0.02	0.43	0.79	1.64	0.77	
		BMIcommunity	ncpa.euc	0.32	below	-0.34	-1.39	-0.71	-0.35	0.04	1.12	0.75	
		BMIcommunity	ncpa.euc	0.32	above	0.40	-1.34	0.02	0.43	0.79	1.64	0.77	
		BMIcommunity	TITAN.decreasers	0.20	below	-0.42	-1.39	-0.78	-0.45	0.08	1.12	0.86	
		BMIcommunity	TITAN.decreasers	0.20	above	0.30	-1.34	0.09	0.27	0.67	1.64	0.58	
		CSCI	BRT_exhaustion	0.80	below	0.92	0.21	0.80	0.98	1.08	1.27	0.28	
		CSCI	BRT_exhaustion	0.80	above	0.64	0.10	0.49	0.63	0.77	1.20	0.28	
TN (mg/L)	BMI	CSCI	BRT_resistance	0.30	below	0.98	0.36	0.89	1.02	1.10	1.27	0.20	
		CSCI	BRT_resistance	0.30	above	0.69	0.10	0.52	0.66	0.85	1.26	0.34	
		CSCI	pieciseregression	0.74	below	0.93	0.21	0.81	0.99	1.08	1.27	0.27	
		CSCI	pieciseregression	0.74	above	0.63	0.10	0.48	0.63	0.76	1.20	0.28	
		CSCI	SiZer	0.14	below	0.99	0.36	0.92	1.03	1.11	1.27	0.18	
		CSCI	SiZer	0.14	above	0.77	0.10	0.60	0.77	0.97	1.26	0.37	
		EPT_Percent	BRT_exhaustion	0.60	below	0.42	0.00	0.25	0.42	0.59	0.92	0.34	
		EPT_Percent	BRT_exhaustion	0.60	above	0.18	0.00	0.00	0.07	0.29	0.90	0.29	
		EPT_Percent	pieciseregression	0.59	below	0.42	0.00	0.25	0.42	0.59	0.92	0.34	
		EPT_Percent	pieciseregression	0.59	above	0.18	0.00	0.00	0.07	0.29	0.90	0.29	
	BMI	EPT_Percent	SiZer	0.27	below	0.46	0.00	0.30	0.47	0.63	0.92	0.33	
		EPT_Percent	SiZer	0.27	above	0.21	0.00	0.02	0.13	0.36	0.90	0.34	
		EPT_PercentTaxa	pieciseregression	0.68	below	0.45	0.00	0.35	0.48	0.56	0.75	0.21	
		EPT_PercentTaxa	pieciseregression	0.68	above	0.18	0.00	0.08	0.17	0.25	0.55	0.17	
		EPT_PercentTaxa	SiZer	0.14	below	0.51	0.15	0.46	0.53	0.59	0.75	0.13	
		EPT_PercentTaxa	SiZer	0.14	above	0.29	0.00	0.16	0.28	0.41	0.67	0.25	
		EPT_Taxa	pieciseregression	0.63	below	15.38	0.00	9.00	16.00	21.00	34.00	12.00	
		EPT_Taxa	pieciseregression	0.63	above	3.00	0.00	1.00	2.00	4.00	17.00	3.00	
		EPT_Taxa	SiZer	0.14	below	18.46	4.00	14.00	19.00	23.00	34.00	9.00	
		EPT_Taxa	SiZer	0.14	above	7.77	0.00	2.00	6.00	12.00	27.00	10.00	
Intolerant_Percent	Intolerant_Percent	pieciseregression	0.54	below	0.19	0.00	0.05	0.17	0.30	0.72	0.24		
	Intolerant_Percent	pieciseregression	0.54	above	0.01	0.00	0.00	0.00	0.00	0.53	0.00		
	Intolerant_Percent	SiZer	0.14	below	0.24	0.00	0.12	0.24	0.35	0.72	0.23		
	Intolerant_Percent	SiZer	0.14	above	0.07	0.00	0.00	0.00	0.11	0.61	0.11		
	Intolerant_PercentTaxa	BRT_exhaustion	0.55	below	0.29	0.00	0.17	0.30	0.41	0.62	0.25		
	Intolerant_PercentTaxa	BRT_exhaustion	0.55	above	0.03	0.00	0.00	0.00	0.05	0.33	0.05		
Intolerant_PercentTaxa	pieciseregression	0.62	below	0.28	0.00	0.16	0.30	0.41	0.62	0.24			

Table C.3.

Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile				
								25 th	50 th (median)	75 th	max	IQR
TN (mg/L)	diatom	Intolerant_PercentTaxa	pieciseregression	0.62	above	0.02	0.00	0.00	0.00	0.04	0.30	0.04
		Intolerant_PercentTaxa	SiZer	0.14	below	0.35	0.03	0.26	0.36	0.45	0.62	0.19
		Intolerant_PercentTaxa	SiZer	0.14	above	0.12	0.00	0.00	0.06	0.22	0.58	0.22
		Intolerant_Taxa	pieciseregression	0.52	below	10.12	0.00	4.25	10.00	15.75	25.00	11.50
		Intolerant_Taxa	pieciseregression	0.52	above	0.70	0.00	0.00	0.00	1.00	11.00	1.00
		Intolerant_Taxa	SiZer	0.14	below	12.74	1.00	8.00	13.00	17.00	25.00	9.00
		Intolerant_Taxa	SiZer	0.14	above	3.74	0.00	0.00	1.00	6.00	21.00	6.00
		OoverE	pieciseregression	0.79	below	0.98	0.29	0.84	1.01	1.15	1.42	0.31
		OoverE	pieciseregression	0.79	above	0.74	0.15	0.52	0.73	0.90	1.36	0.38
		OoverE	SiZer	0.40	below	1.01	0.38	0.88	1.04	1.17	1.42	0.29
		OoverE	SiZer	0.40	above	0.76	0.15	0.58	0.74	0.92	1.36	0.34
		Shannon_Diversity	BRT_exhaustion	0.80	below	2.28	0.46	1.87	2.36	2.71	3.50	0.84
		Shannon_Diversity	BRT_exhaustion	0.80	above	1.52	0.03	1.20	1.55	1.87	2.65	0.67
		Shannon_Diversity	BRT_resistance	0.30	below	2.40	0.46	2.01	2.50	2.78	3.50	0.77
		Shannon_Diversity	BRT_resistance	0.30	above	1.67	0.03	1.32	1.72	2.03	3.14	0.71
		Shannon_Diversity	pieciseregression	0.76	below	2.28	0.46	1.88	2.37	2.71	3.50	0.83
		Shannon_Diversity	pieciseregression	0.76	above	1.52	0.03	1.20	1.55	1.87	2.65	0.67
		Shannon_Diversity	SiZer	0.53	below	2.34	0.46	1.95	2.43	2.73	3.50	0.78
		Shannon_Diversity	SiZer	0.53	above	1.54	0.03	1.21	1.56	1.89	3.14	0.68
		Simpson_Diversity	SiZer	0.53	below	0.80	0.25	0.74	0.84	0.89	0.95	0.15
		Simpson_Diversity	SiZer	0.53	above	0.65	0.01	0.58	0.69	0.78	0.93	0.20
		Taxonomic_Richness	BRT_exhaustion	0.80	below	31.36	5.00	24.00	33.00	39.00	54.00	15.00
		Taxonomic_Richness	BRT_exhaustion	0.80	above	15.49	1.00	12.00	15.00	19.00	34.00	7.00
		Taxonomic_Richness	BRT_resistance	0.30	below	34.28	5.00	28.00	35.00	40.50	54.00	12.50
		Taxonomic_Richness	BRT_resistance	0.30	above	18.09	1.00	13.00	17.00	23.00	42.00	10.00
		Taxonomic_Richness	pieciseregression	0.71	below	31.84	5.00	25.00	33.00	39.00	54.00	14.00
		Taxonomic_Richness	pieciseregression	0.71	above	15.38	1.00	12.00	15.00	18.75	34.00	6.75
		Taxonomic_Richness	SiZer	0.14	below	35.47	15.00	30.50	36.00	41.00	54.00	10.50
		Taxonomic_Richness	SiZer	0.14	above	22.54	1.00	14.00	21.00	29.00	48.00	15.00
		Tolerant_Percent	pieciseregression	0.61	below	0.10	0.00	0.01	0.03	0.10	0.96	0.08
		Tolerant_Percent	pieciseregression	0.61	above	0.29	0.00	0.07	0.18	0.48	0.97	0.42
		Tolerant_Percent	SiZer	0.14	below	0.05	0.00	0.01	0.02	0.05	0.74	0.04
		Tolerant_Percent	SiZer	0.14	above	0.22	0.00	0.04	0.10	0.34	0.97	0.30
	Tolerant_PercentTaxa	pieciseregression	0.67	below	0.17	0.00	0.09	0.14	0.22	0.73	0.13	
	Tolerant_PercentTaxa	pieciseregression	0.67	above	0.36	0.00	0.27	0.36	0.45	0.77	0.18	
	Tolerant_PercentTaxa	SiZer	0.14	below	0.12	0.00	0.07	0.11	0.15	0.50	0.08	
	Tolerant_PercentTaxa	SiZer	0.14	above	0.28	0.00	0.17	0.26	0.38	0.77	0.21	
	Tolerant_Taxa	SiZer	0.14	below	4.04	0.00	3.00	4.00	5.00	12.00	2.00	
	Tolerant_Taxa	SiZer	0.14	above	5.43	0.00	4.00	5.00	7.00	12.00	3.00	
	D18	BRT_exhaustion	0.80	below	72.88	2.00	62.00	76.00	88.00	100.00	26.00	
D18	BRT_exhaustion	0.80	above	44.70	2.00	26.00	46.00	64.00	100.00	38.00		
D18	BRT_resistance	0.30	below	77.57	16.00	68.00	80.00	90.00	100.00	22.00		
D18	BRT_resistance	0.30	above	50.18	2.00	30.00	52.00	68.50	100.00	38.50		
D18	pieciseregression	0.88	below	72.62	2.00	62.00	76.00	88.00	100.00	26.00		
D18	pieciseregression	0.88	above	44.13	2.00	25.50	44.00	62.00	100.00	36.50		
D18	SiZer	0.14	below	80.11	16.00	73.50	82.00	92.00	100.00	18.50		

Table C.3.												
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile				
								25 th	50 th (median)	75 th	max	IQR
TN (mg/L)	diatom	D18	SiZer	0.14	above	57.25	2.00	40.00	60.00	76.00	100.00	36.00
		diatomCommunity	CART	0.61	below	-0.26	-1.48	-0.66	-0.27	0.11	1.19	0.77
		diatomCommunity	CART	0.61	above	0.50	-1.30	0.24	0.64	0.90	1.71	0.66
		diatomCommunity	ncpa.bc	0.48	below	-0.29	-1.48	-0.67	-0.30	0.07	1.09	0.74
		diatomCommunity	ncpa.bc	0.48	above	0.46	-1.30	0.20	0.62	0.88	1.71	0.67
		diatomCommunity	ncpa.euc	0.37	below	-0.32	-1.48	-0.68	-0.33	0.04	1.07	0.72
	hybrid	diatomCommunity	ncpa.euc	0.37	above	0.42	-1.30	0.07	0.58	0.85	1.71	0.78
		diatomCommunity	TITAN.decreasers	0.29	below	-0.34	-1.48	-0.69	-0.36	0.01	1.07	0.70
		diatomCommunity	TITAN.decreasers	0.29	above	0.37	-1.30	-0.08	0.52	0.81	1.71	0.89
		propAchMin	SiZer	0.53	below	0.22	0.00	0.04	0.14	0.32	0.97	0.28
		propAchMin	SiZer	0.53	above	0.07	0.00	0.00	0.01	0.07	0.82	0.07
		RAWDO100	SiZer	0.40	below	0.39	0.00	0.15	0.32	0.59	1.00	0.44
		RAWDO100	SiZer	0.40	above	0.24	0.00	0.05	0.14	0.34	0.99	0.29
		RAWDO50	pieciswiseregression	1.95	below	0.93	0.30	0.92	0.97	1.00	1.00	0.08
		RAWDO50	pieciswiseregression	1.95	above	0.79	0.11	0.70	0.86	0.95	1.00	0.25
		RAWDO50	SiZer	0.27	below	0.96	0.55	0.95	0.99	1.00	1.00	0.05
		RAWDO50	SiZer	0.27	above	0.85	0.11	0.78	0.91	0.97	1.00	0.19
		RAWeutro	SiZer	0.27	below	0.51	0.00	0.32	0.53	0.71	1.00	0.40
		RAWeutro	SiZer	0.27	above	0.68	0.01	0.52	0.76	0.90	1.00	0.38
		RAWlowN	pieciswiseregression	0.94	below	0.43	0.00	0.18	0.40	0.67	1.00	0.49
		RAWlowN	pieciswiseregression	0.94	above	0.15	0.00	0.02	0.05	0.17	0.98	0.15
		RAWlowN	SiZer	0.26	below	0.49	0.00	0.25	0.47	0.72	1.00	0.47
		RAWlowN	SiZer	0.26	above	0.22	0.00	0.03	0.10	0.34	0.99	0.31
		RAWlowP	pieciswiseregression	0.78	below	0.44	0.00	0.18	0.40	0.69	1.00	0.51
	RAWlowP	pieciswiseregression	0.78	above	0.16	0.00	0.02	0.05	0.20	0.99	0.18	
	RAWlowP	SiZer	0.27	below	0.50	0.00	0.25	0.47	0.75	1.00	0.49	
	RAWlowP	SiZer	0.27	above	0.22	0.00	0.03	0.10	0.32	0.99	0.29	
	hybrid	RAWNhet	BRT_exhaustion	2.10	below	0.12	0.00	0.02	0.07	0.17	0.96	0.15
		RAWNhet	BRT_exhaustion	2.10	above	0.31	0.01	0.12	0.25	0.48	0.87	0.37
		RAWNhet	pieciswiseregression	1.95	below	0.12	0.00	0.02	0.06	0.16	0.90	0.14
		RAWNhet	pieciswiseregression	1.95	above	0.32	0.01	0.12	0.26	0.48	0.96	0.37
		RAWNhet	SiZer	0.27	below	0.08	0.00	0.01	0.04	0.10	0.64	0.10
		RAWNhet	SiZer	0.27	above	0.23	0.00	0.07	0.16	0.33	0.96	0.26
		H20	BRT_exhaustion	2.10	below	64.18	4.00	52.00	69.00	79.00	100.00	27.00
		H20	BRT_exhaustion	2.10	above	35.13	2.00	22.00	35.00	48.50	70.00	26.50
		H20	BRT_resistance	0.30	below	72.73	26.00	66.00	74.00	81.00	100.00	15.00
		H20	BRT_resistance	0.30	above	44.01	2.00	28.00	42.00	59.00	99.00	31.00
	soft	H20	pieciswiseregression	1.06	below	67.19	4.00	58.00	71.00	80.00	100.00	22.00
		H20	pieciswiseregression	1.06	above	35.98	2.00	24.00	35.00	48.00	70.00	24.00
		H20	SiZer	0.17	below	74.47	28.00	69.00	75.00	82.00	100.00	13.00
		H20	SiZer	0.17	above	50.29	2.00	34.50	51.00	68.00	99.00	33.50
		H21	pieciswiseregression	0.68	below	68.02	1.00	57.00	70.00	81.00	100.00	24.00
H21		pieciswiseregression	0.68	above	38.88	3.00	24.00	38.00	51.00	96.00	27.00	
H21		SiZer	0.17	below	74.50	31.00	66.00	74.00	86.00	100.00	20.00	
H21		SiZer	0.17	above	49.10	1.00	33.00	49.00	66.50	100.00	33.50	

Table C.3.									percentile				
Gradient	Assemblage	ALI	analysis type	threshold	gradient value to threshold	mean	min						
								25 th	50 th (median)	75 th	max	IQR	
soft BMI	H23		pieciseregression	0.77	below	71.21	5.00	61.00	75.00	85.00	100.00	24.00	
	H23		pieciseregression	0.77	above	39.45	6.00	25.00	40.00	51.00	96.00	26.00	
	H23		SiZer	0.17	below	78.39	31.00	70.00	80.00	89.00	100.00	19.00	
	H23		SiZer	0.17	above	51.76	5.00	35.50	50.00	70.00	98.00	34.50	
	propTaxaZHR		pieciseregression	0.84	below	0.23	0.00	0.13	0.23	0.33	0.71	0.21	
	propTaxaZHR		pieciseregression	0.84	above	0.08	0.00	0.00	0.03	0.13	0.62	0.13	
	propTaxaZHR		SiZer	0.52	below	0.25	0.00	0.14	0.25	0.35	0.71	0.20	
	propTaxaZHR		SiZer	0.52	above	0.09	0.00	0.00	0.05	0.15	0.62	0.15	
	RAWlowTPsp		SiZer	0.40	below	0.14	0.00	0.04	0.13	0.20	0.57	0.16	
	RAWlowTPsp		SiZer	0.40	above	0.05	0.00	0.00	0.00	0.08	0.40	0.08	
	RAWmeanZHR		BRT_exhaustion	0.80	below	0.31	0.00	0.09	0.24	0.54	0.86	0.45	
	RAWmeanZHR		BRT_exhaustion	0.80	above	0.08	0.00	0.00	0.02	0.08	0.80	0.08	
	RAWmeanZHR		pieciseregression	0.67	below	0.31	0.00	0.10	0.26	0.54	0.86	0.44	
	RAWmeanZHR		pieciseregression	0.67	above	0.09	0.00	0.00	0.02	0.09	0.80	0.09	
	RAWmeanZHR		SiZer	0.52	below	0.33	0.00	0.11	0.28	0.55	0.86	0.44	
	RAWmeanZHR		SiZer	0.52	above	0.09	0.00	0.00	0.03	0.09	0.80	0.09	
	RAWpropBiovolChlor		SiZer	0.14	below	0.25	0.00	0.00	0.05	0.47	1.00	0.47	
	RAWpropBiovolChlor		SiZer	0.14	above	0.56	0.00	0.05	0.66	0.99	1.00	0.94	
	RAWpropBiovolZHR		SiZer	0.20	below	0.45	0.00	0.04	0.40	0.89	1.00	0.85	
	RAWpropBiovolZHR		SiZer	0.20	above	0.17	0.00	0.00	0.00	0.20	1.00	0.20	
	RAWpropGreenCRUS		BRT_exhaustion	0.55	below	0.17	0.00	0.00	0.00	0.00	1.00	0.00	
	RAWpropGreenCRUS		BRT_exhaustion	0.55	above	0.60	0.00	0.00	0.89	1.00	1.00	1.00	
	RAWpropGreenCRUS		BRT_resistance	0.30	below	0.14	0.00	0.00	0.00	0.00	1.00	0.00	
	RAWpropGreenCRUS		BRT_resistance	0.30	above	0.55	0.00	0.00	0.78	1.00	1.00	1.00	
	RAWpropGreenCRUS		pieciseregression	0.60	below	0.18	0.00	0.00	0.00	0.01	1.00	0.01	
	RAWpropGreenCRUS		pieciseregression	0.60	above	0.60	0.00	0.00	0.89	1.00	1.00	1.00	
	RAWpropGreenCRUS		SiZer	0.13	below	0.09	0.00	0.00	0.00	0.00	1.00	0.00	
	RAWpropGreenCRUS		SiZer	0.13	above	0.44	0.00	0.00	0.10	0.98	1.00	0.98	
	TN (mg/L)	S2		BRT_exhaustion	0.80	below	63.36	0.00	52.00	67.00	80.00	100.00	28.00
		S2		BRT_exhaustion	0.80	above	28.56	0.00	15.00	25.00	40.00	100.00	25.00
S2			BRT_resistance	0.30	below	68.05	3.00	58.00	70.00	82.00	100.00	24.00	
S2			BRT_resistance	0.30	above	35.87	0.00	17.00	32.50	53.50	100.00	36.50	
S2			pieciseregression	0.83	below	63.18	0.00	51.00	67.00	80.00	100.00	29.00	
S2			pieciseregression	0.83	above	28.24	0.00	15.00	25.00	38.00	100.00	23.00	
S2			SiZer	0.14	below	70.42	3.00	62.00	73.00	83.00	100.00	21.00	
S2			SiZer	0.14	above	44.56	0.00	23.00	42.00	65.50	100.00	42.50	
BMIcommunity			CART	0.06	below	-0.21	-1.39	-0.65	-0.27	0.20	1.50	0.85	
BMIcommunity			CART	0.06	above	0.33	-1.23	-0.07	0.32	0.77	1.64	0.84	
BMIcommunity			ncpa.bc	0.08	below	-0.19	-1.39	-0.64	-0.25	0.21	1.51	0.85	
BMIcommunity			ncpa.bc	0.08	above	0.45	-0.79	0.12	0.45	0.84	1.64	0.72	
BMIcommunity			ncpa.euc	0.08	below	-0.19	-1.39	-0.64	-0.25	0.21	1.51	0.85	
BMIcommunity			ncpa.euc	0.08	above	0.45	-0.79	0.12	0.45	0.84	1.64	0.72	
TN_mgL		BMIcommunity		TITAN.decreasers	0.04	below	-0.24	-1.39	-0.68	-0.29	0.19	1.49	0.87
		BMIcommunity		TITAN.decreasers	0.04	above	0.26	-1.27	-0.15	0.23	0.67	1.64	0.82
		CSCI		BRT_exhaustion	0.12	below	0.91	0.21	0.77	0.97	1.08	1.27	0.30

Table C.3.												
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile				
								25 th	50 th (median)	75 th	max	IQR
TP_mgL	diatom	CSCI	BRT_exhaustion	0.12	above	0.63	0.10	0.48	0.65	0.76	1.12	0.28
		CSCI	pieciseregression	0.15	below	0.90	0.21	0.75	0.96	1.07	1.27	0.32
		CSCI	pieciseregression	0.15	above	0.61	0.10	0.42	0.63	0.75	1.11	0.33
		CSCI	SiZer	0.03	below	0.97	0.29	0.89	1.02	1.11	1.24	0.22
		CSCI	SiZer	0.03	above	0.77	0.10	0.60	0.77	0.99	1.27	0.38
		EPT_Percent	SiZer	0.08	below	0.41	0.00	0.21	0.41	0.58	0.92	0.37
		EPT_Percent	SiZer	0.08	above	0.18	0.00	0.00	0.07	0.30	0.90	0.29
		EPT_PercentTaxa	pieciseregression	0.10	below	0.43	0.00	0.32	0.47	0.55	0.75	0.24
		EPT_PercentTaxa	pieciseregression	0.10	above	0.19	0.00	0.08	0.18	0.27	0.64	0.19
		EPT_PercentTaxa	SiZer	0.03	below	0.48	0.09	0.39	0.50	0.58	0.75	0.19
		EPT_PercentTaxa	SiZer	0.03	above	0.30	0.00	0.16	0.29	0.46	0.72	0.30
		EPT_Taxa	pieciseregression	0.11	below	14.11	0.00	7.25	14.00	20.00	34.00	12.75
		EPT_Taxa	pieciseregression	0.11	above	3.83	0.00	1.00	3.00	4.00	23.00	3.00
		EPT_Taxa	SiZer	0.03	below	16.43	1.00	10.50	18.00	22.00	34.00	11.50
		EPT_Taxa	SiZer	0.03	above	8.63	0.00	2.00	6.00	14.00	31.00	12.00
		Intolerant_Percent	SiZer	0.08	below	0.17	0.00	0.02	0.14	0.27	0.72	0.25
		Intolerant_Percent	SiZer	0.08	above	0.02	0.00	0.00	0.00	0.00	0.53	0.00
		Intolerant_PercentTaxa	pieciseregression	0.11	below	0.25	0.00	0.09	0.26	0.39	0.62	0.30
		Intolerant_PercentTaxa	pieciseregression	0.11	above	0.04	0.00	0.00	0.00	0.06	0.44	0.06
		Intolerant_PercentTaxa	SiZer	0.03	below	0.30	0.00	0.18	0.32	0.42	0.62	0.24
		Intolerant_PercentTaxa	SiZer	0.03	above	0.14	0.00	0.00	0.07	0.28	0.58	0.28
		Intolerant_Taxa	pieciseregression	0.11	below	8.84	0.00	2.00	8.00	15.00	25.00	13.00
		Intolerant_Taxa	pieciseregression	0.11	above	1.20	0.00	0.00	0.00	1.00	16.00	1.00
		Intolerant_Taxa	SiZer	0.03	below	10.63	0.00	5.00	10.00	16.00	25.00	11.00
		Intolerant_Taxa	SiZer	0.03	above	4.71	0.00	0.00	1.00	9.00	25.00	9.00
		OoverE	SiZer	0.03	below	1.02	0.38	0.90	1.04	1.18	1.39	0.28
		OoverE	SiZer	0.03	above	0.85	0.15	0.67	0.87	1.05	1.42	0.38
		Shannon_Diversity	SiZer	0.08	below	2.26	0.08	1.83	2.35	2.70	3.50	0.87
		Shannon_Diversity	SiZer	0.08	above	1.60	0.03	1.25	1.71	1.95	2.74	0.70
		Simpson_Diversity	SiZer	0.06	below	0.79	0.03	0.72	0.83	0.88	0.95	0.17
		Simpson_Diversity	SiZer	0.06	above	0.68	0.01	0.63	0.73	0.80	0.93	0.18
		Taxonomic_Richness	BRT_exhaustion	0.12	below	30.17	1.00	22.00	31.00	38.00	54.00	16.00
		Taxonomic_Richness	BRT_exhaustion	0.12	above	17.05	3.00	12.25	16.00	20.00	41.00	7.75
		Taxonomic_Richness	pieciseregression	0.14	below	29.86	1.00	22.00	31.00	38.00	54.00	16.00
		Taxonomic_Richness	pieciseregression	0.14	above	15.97	3.00	12.00	15.00	18.00	41.00	6.00
		Taxonomic_Richness	SiZer	0.03	below	33.13	7.00	27.00	34.00	40.00	54.00	13.00
		Taxonomic_Richness	SiZer	0.03	above	23.56	1.00	14.00	22.00	33.00	54.00	19.00
		Tolerant_Percent	SiZer	0.03	below	0.09	0.00	0.01	0.03	0.08	0.90	0.07
		Tolerant_Percent	SiZer	0.03	above	0.21	0.00	0.03	0.09	0.33	0.97	0.30
		Tolerant_PercentTaxa	pieciseregression	0.11	below	0.19	0.00	0.09	0.15	0.25	0.73	0.16
Tolerant_PercentTaxa	pieciseregression	0.11	above	0.35	0.03	0.25	0.35	0.44	0.77	0.19		
Tolerant_PercentTaxa	SiZer	0.03	below	0.15	0.00	0.08	0.12	0.20	0.60	0.13		
Tolerant_PercentTaxa	SiZer	0.03	above	0.27	0.00	0.15	0.25	0.38	0.77	0.23		
Tolerant_Taxa	SiZer	0.05	below	4.59	0.00	3.00	4.00	6.00	12.00	3.00		
Tolerant_Taxa	SiZer	0.05	above	5.58	1.00	4.00	6.00	7.00	12.00	3.00		
D18	BRT_exhaustion	0.18	below	70.26	2.00	58.00	74.00	86.00	100.00	28.00		

Table C.3.														
Gradient	Assemblage	ALI	analysis type	threshold	gradient value relationship to threshold	mean	min	percentile						
								25 th	50 th (median)	75 th	max	IQR		
TP (mg/L)	diatom	D18	BRT_exhaustion	0.18	above	39.03	2.00	21.50	37.00	56.00	100.00	34.50		
		D18	pieciseregression	0.12	below	72.30	2.00	62.00	76.00	88.00	100.00	26.00		
		D18	pieciseregression	0.12	above	40.54	2.00	22.00	42.00	57.00	100.00	35.00		
		D18	SiZer	0.02	below	80.35	4.00	72.00	82.00	94.00	100.00	22.00		
		D18	SiZer	0.02	above	57.91	2.00	42.00	62.00	76.00	100.00	34.00		
		diatomCommunity	CART	0.08	below	-0.26	-1.48	-0.66	-0.28	0.11	1.50	0.77		
		diatomCommunity	CART	0.08	above	0.63	-0.98	0.44	0.69	0.92	1.71	0.48		
		diatomCommunity	ncpa.bc	0.05	below	-0.31	-1.48	-0.69	-0.34	0.06	1.50	0.75		
		diatomCommunity	ncpa.bc	0.05	above	0.51	-1.17	0.27	0.60	0.87	1.71	0.60		
		diatomCommunity	ncpa.euc	0.03	below	-0.44	-1.48	-0.82	-0.46	-0.10	1.09	0.72		
		diatomCommunity	ncpa.euc	0.03	above	0.30	-1.25	-0.15	0.39	0.74	1.71	0.88		
		diatomCommunity	TITAN.decreasers	0.04	below	-0.35	-1.48	-0.71	-0.37	0.00	1.16	0.71		
		diatomCommunity	TITAN.decreasers	0.04	above	0.42	-1.25	0.09	0.54	0.83	1.71	0.74		
		propAchMin	pieciseregression	0.04	below	0.24	0.00	0.06	0.16	0.36	0.97	0.30		
		propAchMin	pieciseregression	0.04	above	0.07	0.00	0.00	0.01	0.06	0.89	0.06		
		propAchMin	SiZer	0.02	below	0.28	0.00	0.08	0.21	0.45	0.97	0.38		
		propAchMin	SiZer	0.02	above	0.10	0.00	0.00	0.03	0.13	0.89	0.12		
		TP (mg/L)	hybrid	RAWDO100	BRT_exhaustion	0.10	below	0.37	0.00	0.14	0.31	0.58	1.00	0.44
				RAWDO100	BRT_exhaustion	0.10	above	0.19	0.00	0.03	0.10	0.21	0.96	0.18
				RAWDO100	pieciseregression	0.06	below	0.40	0.00	0.17	0.33	0.61	1.00	0.44
RAWDO100	pieciseregression			0.06	above	0.20	0.00	0.04	0.11	0.28	0.96	0.24		
RAWDO100	SiZer			0.01	below	0.52	0.00	0.22	0.52	0.79	1.00	0.57		
RAWDO100	SiZer			0.01	above	0.29	0.00	0.07	0.21	0.45	0.99	0.38		
RAWDO50	pieciseregression			0.27	below	0.93	0.30	0.92	0.97	1.00	1.00	0.08		
RAWDO50	pieciseregression			0.27	above	0.76	0.11	0.62	0.82	0.90	1.00	0.28		
RAWDO50	SiZer			0.02	below	0.96	0.55	0.96	0.99	1.00	1.00	0.04		
RAWDO50	SiZer			0.02	above	0.88	0.11	0.82	0.94	0.98	1.00	0.16		
RAWeutro	BRT_exhaustion			0.12	below	0.54	0.00	0.34	0.58	0.76	1.00	0.42		
RAWeutro	BRT_exhaustion			0.12	above	0.75	0.04	0.67	0.84	0.93	1.00	0.25		
RAWeutro	pieciseregression			0.08	below	0.52	0.00	0.32	0.55	0.73	1.00	0.41		
RAWeutro	pieciseregression			0.08	above	0.75	0.04	0.67	0.84	0.93	1.00	0.26		
RAWeutro	SiZer			0.02	below	0.45	0.00	0.22	0.46	0.63	0.97	0.41		
RAWeutro	SiZer			0.02	above	0.65	0.01	0.46	0.70	0.88	1.00	0.42		
RAWlowN	pieciseregression			0.09	below	0.45	0.00	0.20	0.43	0.69	1.00	0.49		
RAWlowN	pieciseregression			0.09	above	0.12	0.00	0.01	0.05	0.14	0.90	0.13		
RAWlowN	SiZer			0.02	below	0.54	0.00	0.32	0.53	0.78	1.00	0.46		
RAWlowN	SiZer			0.02	above	0.25	0.00	0.03	0.14	0.41	0.99	0.37		
RAWlowP	pieciseregression	0.08	below	0.47	0.00	0.20	0.43	0.73	1.00	0.53				
RAWlowP	pieciseregression	0.08	above	0.12	0.00	0.01	0.05	0.15	0.92	0.14				
RAWlowP	SiZer	0.02	below	0.56	0.00	0.34	0.57	0.81	1.00	0.47				
RAWlowP	SiZer	0.02	above	0.24	0.00	0.04	0.14	0.36	0.99	0.33				
TP (mg/L)	soft	RAWNhet	BRT_exhaustion	0.10	below	0.11	0.00	0.01	0.06	0.14	0.88	0.13		
		RAWNhet	BRT_exhaustion	0.10	above	0.28	0.00	0.10	0.22	0.40	0.99	0.30		
		RAWNhet	pieciseregression	0.11	below	0.11	0.00	0.01	0.06	0.14	0.90	0.13		
		RAWNhet	pieciseregression	0.11	above	0.28	0.00	0.10	0.22	0.40	0.99	0.29		
		RAWNhet	SiZer	0.02	below	0.07	0.00	0.01	0.03	0.08	0.69	0.08		
		RAWNhet	SiZer	0.02	above	0.20	0.00	0.05	0.13	0.27	0.99	0.22		

Table C.3.		Assemblage	ALI	analysis type	threshold	gradient value relationship		percentile				
Gradient	mean					min	25 th	50 th (median)	75 th	max	IQR	
												to threshold
soft	H20	BRT_exhaustion	0.15	below	65.01	4.00	54.00	69.00	79.00	100.00	25.00	
	H20	BRT_exhaustion	0.15	above	34.95	9.00	21.00	32.00	47.00	80.00	26.00	
	H20	piecewiseregression	0.11	below	67.07	4.00	58.00	71.00	80.00	100.00	22.00	
	H20	piecewiseregression	0.11	above	35.33	8.00	21.00	34.50	48.00	80.00	27.00	
	H20	SiZer	0.02	below	73.58	11.00	68.00	75.00	82.00	100.00	14.00	
	H20	SiZer	0.02	above	50.82	4.00	35.00	51.00	69.00	95.00	34.00	
	H21	piecewiseregression	0.11	below	66.47	1.00	54.00	69.00	80.25	100.00	26.25	
	H21	piecewiseregression	0.11	above	36.15	1.00	21.00	34.00	49.00	83.00	28.00	
	H21	SiZer	0.02	below	72.79	3.00	63.00	74.00	86.00	100.00	23.00	
	H21	SiZer	0.02	above	50.40	1.00	34.00	51.00	67.00	100.00	33.00	
	H23	piecewiseregression	0.11	below	69.47	5.00	58.00	74.00	85.00	100.00	27.00	
	H23	piecewiseregression	0.11	above	37.45	9.00	24.00	36.00	48.00	85.00	24.00	
	H23	SiZer	0.02	below	76.27	6.00	69.00	80.00	89.00	100.00	20.00	
	H23	SiZer	0.02	above	53.19	5.00	36.00	52.00	72.00	98.00	36.00	
	propTaxaZHR	SiZer	0.05	below	0.24	0.00	0.13	0.23	0.35	0.71	0.23	
	propTaxaZHR	SiZer	0.05	above	0.11	0.00	0.00	0.07	0.18	0.62	0.18	
	RAWlowTPsp	piecewiseregression	0.07	below	0.13	0.00	0.00	0.11	0.20	0.57	0.20	
	RAWlowTPsp	piecewiseregression	0.07	above	0.03	0.00	0.00	0.00	0.06	0.40	0.06	
	RAWlowTPsp	SiZer	0.04	below	0.14	0.00	0.05	0.13	0.21	0.57	0.16	
	RAWlowTPsp	SiZer	0.04	above	0.05	0.00	0.00	0.00	0.08	0.50	0.08	
	RAWmeanZHR	BRT_exhaustion	0.05	below	0.32	0.00	0.09	0.27	0.55	0.86	0.46	
	RAWmeanZHR	BRT_exhaustion	0.05	above	0.12	0.00	0.00	0.04	0.14	0.80	0.14	
	RAWmeanZHR	SiZer	0.05	below	0.32	0.00	0.09	0.28	0.56	0.86	0.46	
	RAWmeanZHR	SiZer	0.05	above	0.12	0.00	0.00	0.05	0.14	0.80	0.14	
	RAWpropBiovolChlor	SiZer	0.05	below	0.35	0.00	0.00	0.14	0.75	1.00	0.75	
	RAWpropBiovolChlor	SiZer	0.05	above	0.59	0.00	0.08	0.75	1.00	1.00	0.92	
	RAWpropBiovolZHR	SiZer	0.05	below	0.41	0.00	0.02	0.28	0.84	1.00	0.82	
	RAWpropBiovolZHR	SiZer	0.05	above	0.13	0.00	0.00	0.00	0.08	1.00	0.08	
	RAWpropGreenCRUS	SiZer	0.05	below	0.21	0.00	0.00	0.00	0.12	1.00	0.12	
	RAWpropGreenCRUS	SiZer	0.05	above	0.49	0.00	0.00	0.50	0.99	1.00	0.99	
TP (mg/L)	S2	BRT_exhaustion	0.10	below	61.06	0.00	45.00	67.00	80.00	100.00	35.00	
	S2	BRT_exhaustion	0.10	above	31.18	0.00	15.00	27.00	43.00	100.00	28.00	
	S2	piecewiseregression	0.11	below	60.82	0.00	45.00	66.00	79.50	100.00	34.50	
	S2	piecewiseregression	0.11	above	31.16	0.00	15.00	27.00	43.00	100.00	28.00	
	S2	SiZer	0.05	below	64.61	0.00	53.00	68.00	82.00	100.00	29.00	
	S2	SiZer	0.05	above	36.37	0.00	17.00	33.00	55.00	100.00	38.00	

Table C.4. Summary of recommended numeric endpoints for stream NNE indicators, by beneficial use, from Tetra Tech (2006).

Beneficial Use Risk Category I. Presumptive unimpaired (i.e., the beneficial use is supported)

Beneficial Use Risk Category II. Potentially impaired (i.e., the site may require an impairment assessment)

Beneficial Use Risk Category III. Presumptive impaired (i.e., the beneficial use is not supported or is highly threatened)

Response Variable	BURC							
	Boundary	COLD	WARM	REC-1	REC-2	MUN	SPWN	MIGR
Benthic algal biomass – max (mg chlorophyll <i>a</i> m ⁻²)	I/II	100	150	C	C	100	100	B
	II/II	150	200	C	C	150	150	B
Dissolved oxygen – mean of 7 daily min. (mg L ⁻¹)	I/II	9.5	6.0	A	A	A	8.0	C
	II/III	5.0	4.0	A	A	A	5.0	C
pH maximum – photosynthesis-driven	I/II	9.0	9.0	A	A	A	C	C
	II/III	9.5	9.5	A	A	A	A	A

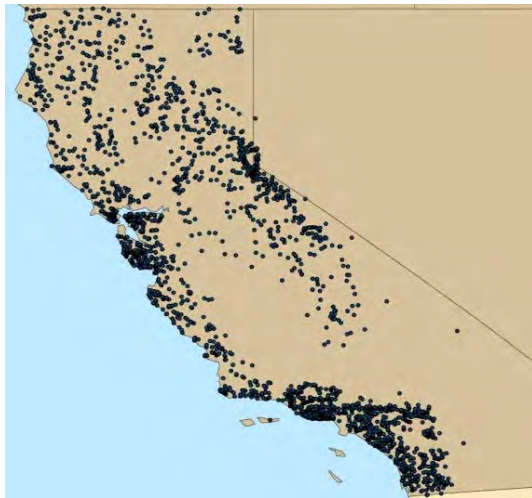
A – No direct linkage to the beneficial use

B – More research needed to quantify linkage

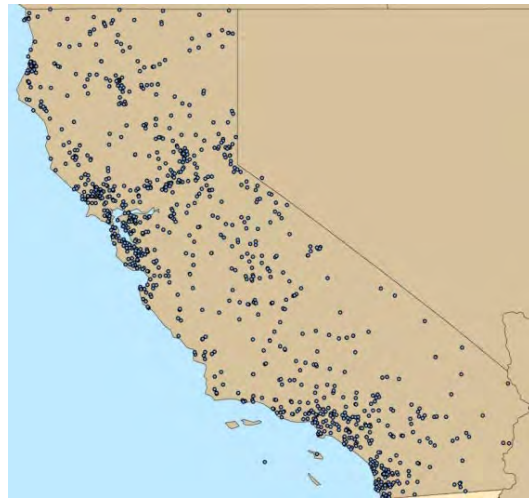
C – Addressed by Aquatic Life Criteria

Appendix D. Graphics and Tables Supporting Evaluation of Benthic Biomass Response Models the NNE Benthic Biomass Spreadsheet Tool

Figure D.1. Comparison of the NNE stations (3053) to the daily precipitation station from NOAA (981, July 2012).

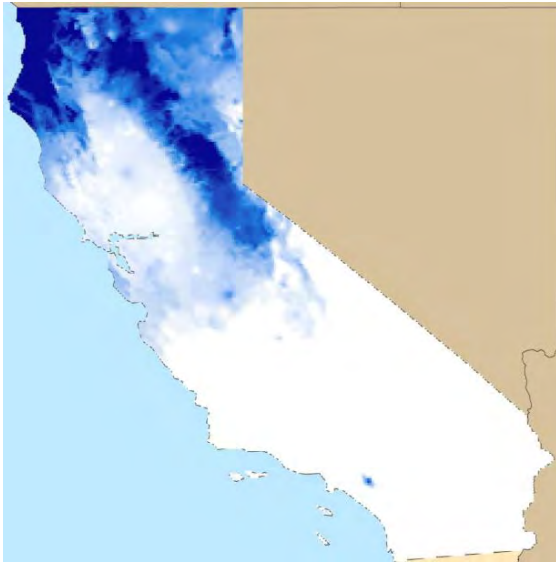


• NNE Stations

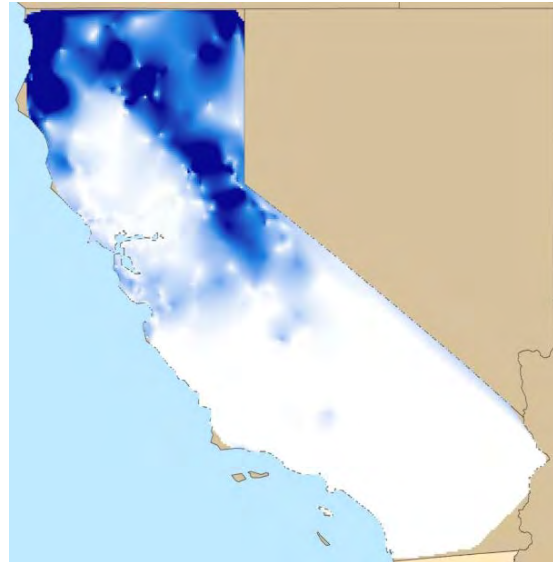


• Daily precipitation stations (NOAA)

Figure D.2. Comparison of the predicted precipitation between the PRISM and ARCGIS (predicted for this study). A good fit was observed for the predicted precipitation between the PRISM model and the ARCGIS model.



- Precipitation predicted by PRISM Model



- Precipitation predicted by ARCGIS model created for this study

Figure D.3. Ranking by variable importance for the Dodds 97 model for AFDM using random forests.

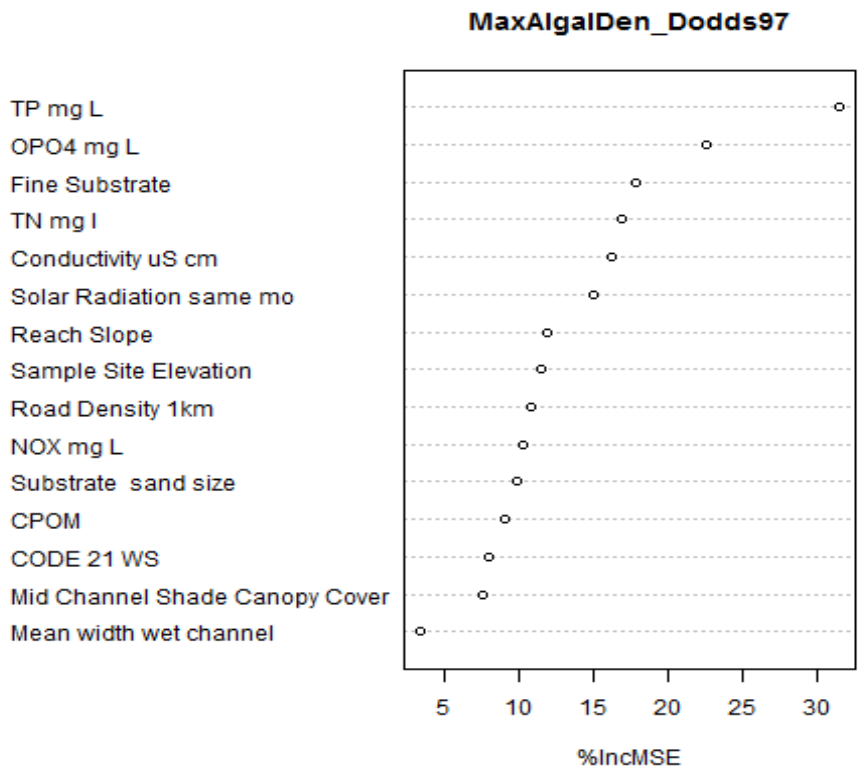
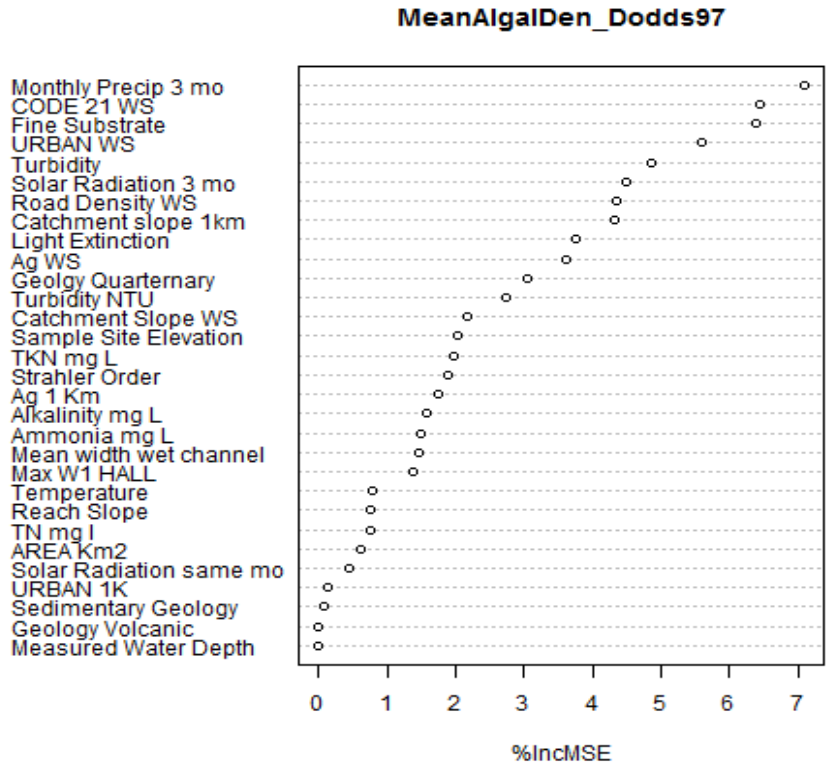


Figure D.4. Ranking by variable importance for the Dodds O2 model for AFDM using random forests.

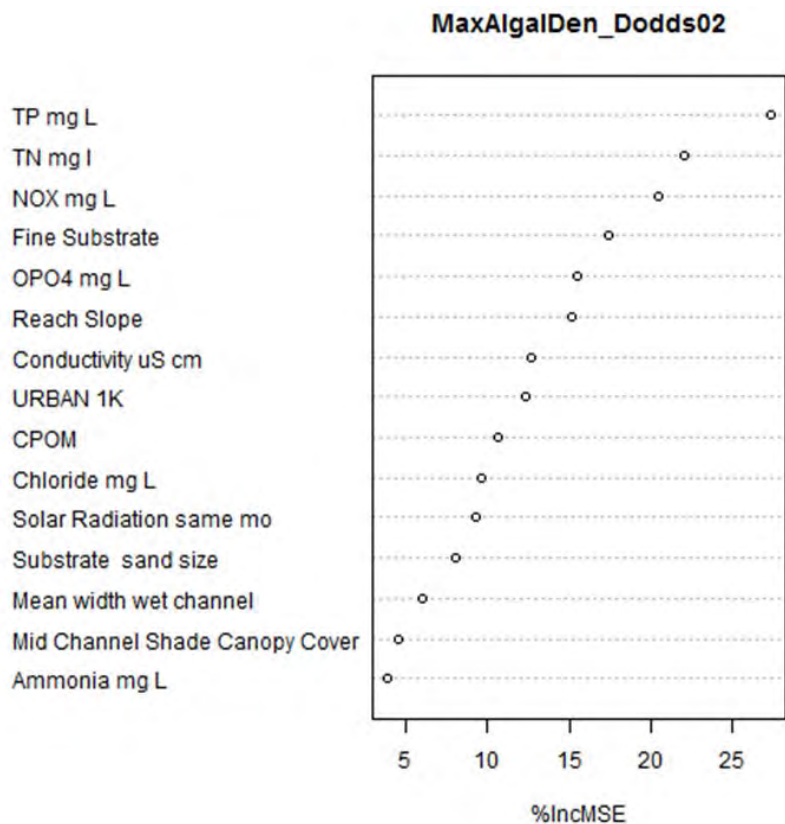
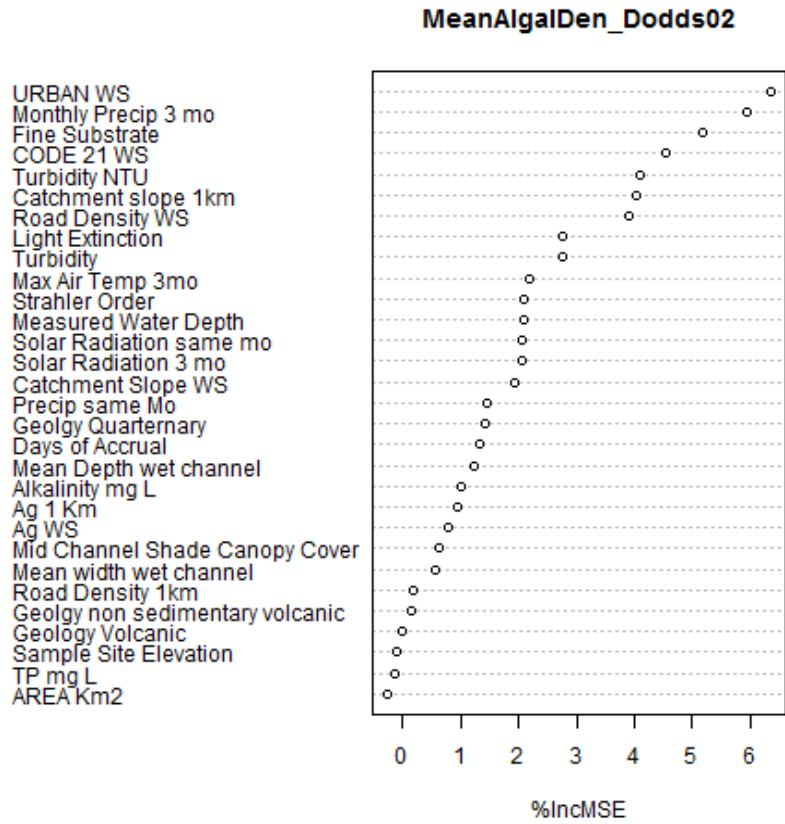


Figure D.5 Ranking by variable importance for the QUAL2K model for AFDM using random forests.

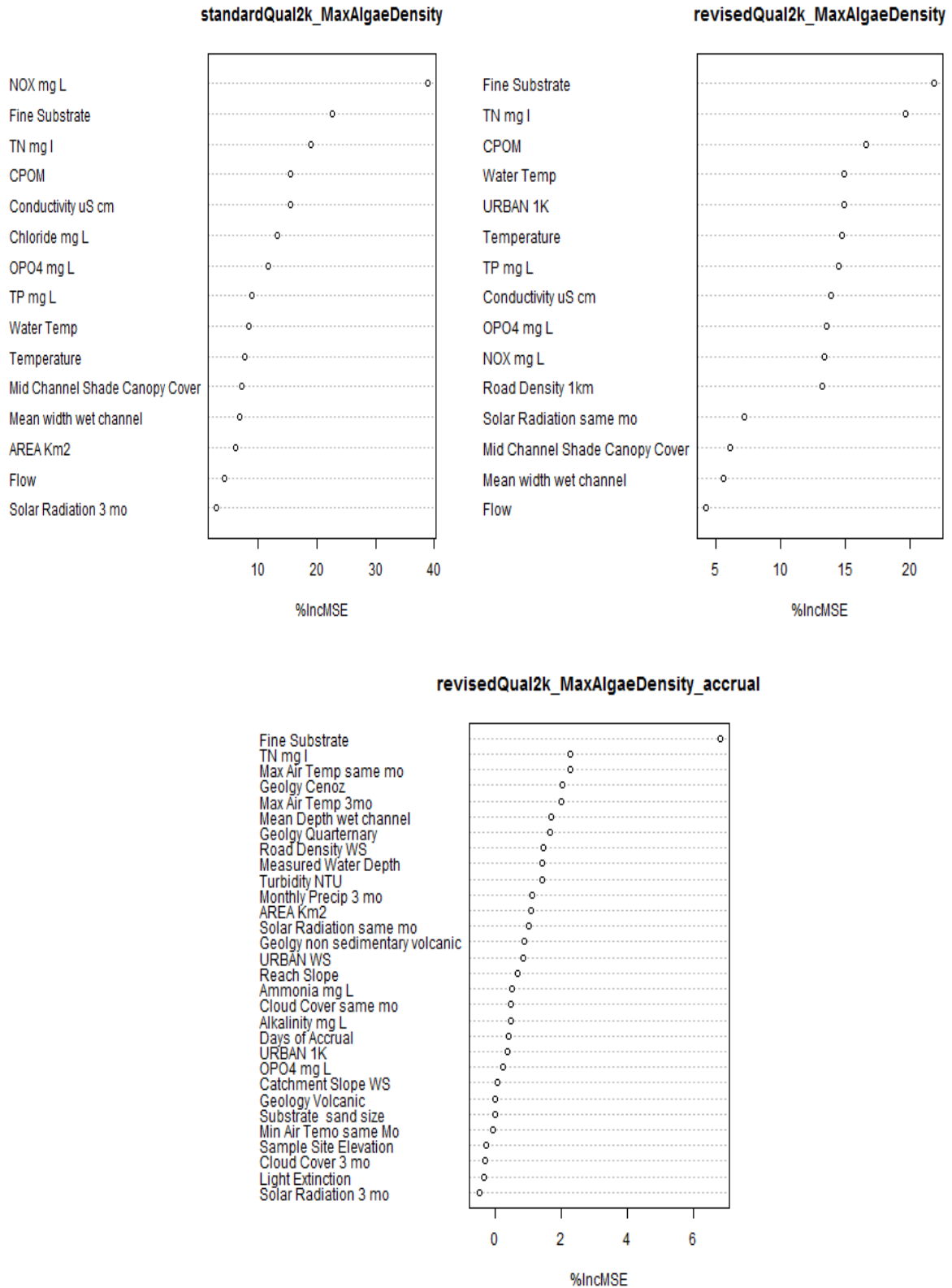


Figure D.6 Ranking by variable importance for the Dodds 97 for chlorophyll *a* using random forests.

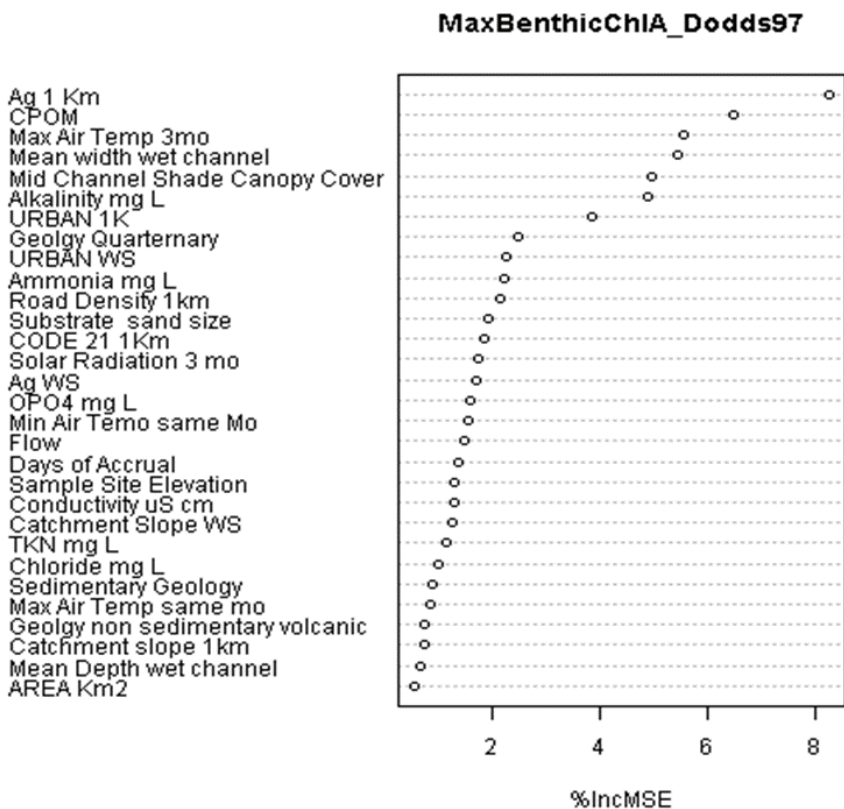
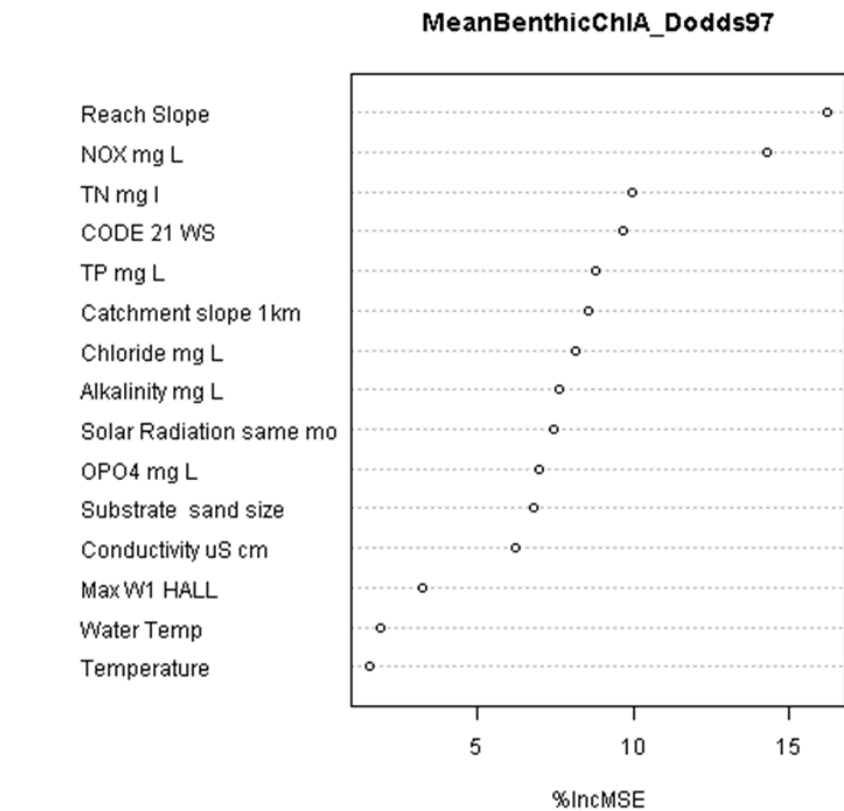


Figure D.7. Ranking by variable importance for the Dodds 02 for chlorophyll *a* using random forests.

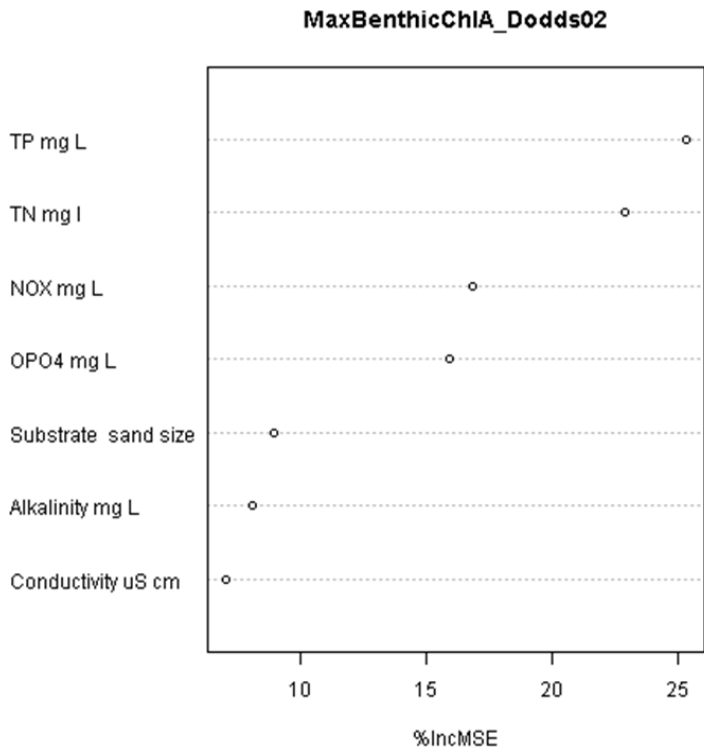
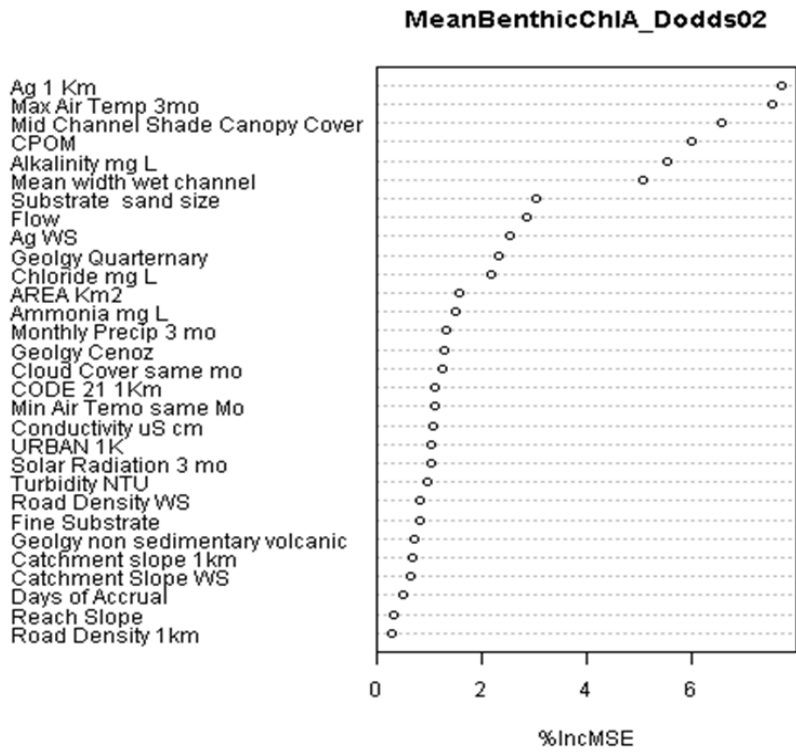


Figure D.8. Ranking by variable importance for the QUAL2K models for chlorophyll *a* using random forests.

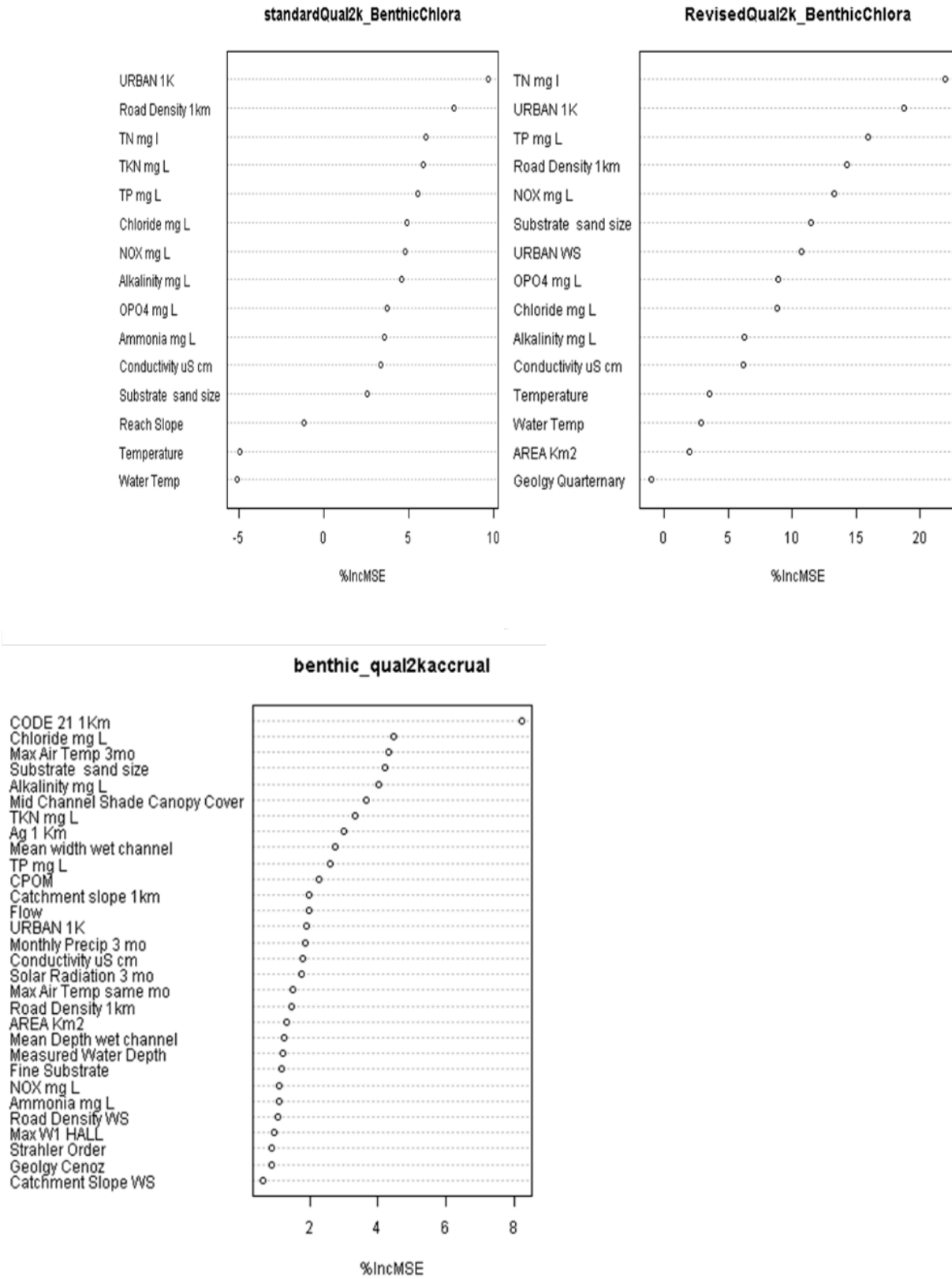


Table D.1 Details on the models. PCT_MAP is macroalgal percent cover, PCT_MCP is macrophyte percent cover, and PCT_MIAT1 is percent presence of thick (1mm+) microalgae. Dashes correspond to predictors that were not included in the final model for the biomass response variable in question.

Predictor	Chlorophyll a	AFDM	Soft Algal Total Biovolume	PCT_MAP	PCT_MCP	PCT_MIAT1	Mean Relative Influence
canopy cover (%)	3.3	4.07	7.13	13.4	2.23	6.8	6.16
sand & fines (%)	2.4	1.95	4.84	6.63	17.4	3.6	6.13
conductivity	2.4	3.09	14.81	12.7	1.01	1.2	5.86
fines (%)	-	11.9	5.15	2.5	7.8	1.3	5.73
stream	9.9	2.03	8.27	2.73	1.39	9.2	5.58
CODE_21_2000_5K	9.5	4.21	2.17	-	-	-	5.28
latitude	-	3.86	10.68	1.11	3.39	5.7	4.94
coarse particulate	2.9	3.11	-	2.68	14.5	1.2	4.87
discharge	2	5.63	1.56	2.59	5.34	6.9	4.00
Ag_2000_WS	4.5	2.86	-	-	-	-	3.66
alkalinity	7.3	1.21	2.45	5.96	2.42	2.5	3.65
slope, reach	1.9	6.78	2.54	5.91	2.86	1.5	3.59
pH	7.7	2.2	2.51	2.84	1.75	4	3.51
ecoregion	-	4.82	3.72	1.36	3.55	3.6	3.41
NOx	6	-	3.21	3.17	1.16	2.9	3.27
NH4	2.8	7.7	2	2.25	1.45	3.1	3.22
mean monthly max	5	1.23	-	1.88	2.06	5.8	3.19
turbidity	2.8	4.29	2.44	2.93	1.94	4.6	3.16
longitude	3.7	3.15	3.08	0.76	3.44	1.9	2.67
SRP	4.1	3.83	2.08	1.46	0.77	3.6	2.62
watershed area	-	1.37	2.17	2.44	3.8	2.6	2.48
days of accrual	-	2.55	2.16	1.38	3.16	3.1	2.47
CODE_21_2000_WS	2.4	2.26	2.68	-	-	-	2.44
TN	3.8	1.38	1.51	2.73	1.8	2.7	2.31
W1_HALL (riparian)	2.2	3.28	1.98	1.39	0.88	3.7	2.24
TP	-	-	3.31	1.49	0.72	3.4	2.23
elevation	-	-	1.24	4.05	0.96	2.6	2.21
site disturbance	-	-	-	0.73	3.64	-	2.19
stream width	2.5	3.99	-	1.63	0.82	1.2	2.03
sedimentary	4.8	-	1.67	1.13	1.47	1	2.00
URBAN_2000_WS	2.8	1.17	-	-	-	-	1.97
mean monthly solar	2.1	2.1	2.12	2.01	1.49	1.7	1.91
URBAN_2000_1K	-	-	-	2.1	1.24	2.2	1.83
mean monthly %	-	1.18	2.52	1.26	1.09	2.8	1.76
total precipitation	-	1.34	-	1.5	1.2	1.9	1.49
Ag_2000_5K	-	-	-	1.34	2.35	0.8	1.48
stream depth	1.6	1.47	-	1.97	0.9	1.2	1.43

Detailed Explanations of Bayesian CART Analysis

Bayesian CART Analysis Approach

Chipman et al (2002) have developed a Bayesian approach to CART which allows the user to specify a prior probability distribution $p(\theta, T)$, where θ represents the regression model parameters, T represents the tree structure and $p(\theta, T) = p(\theta | T)p(T)$. The BCART program uses the prior probability distributions to determine the chance of splitting each node and selects from a distribution of model parameters (β, σ) to produce a set of candidate trees. Four parameters are used to characterize prior probability distributions:

$$\Pr(\text{node splits} | \text{depth} = d) = \alpha(1+d)^{-\beta}$$

where α determines the number of final nodes, depth represents the order of splits or “level” in the tree, and β determines the shape of the tree (“bushiness”).

Chipman et al. (2002) suggest the following default values: $\alpha = 0.5$ and $\beta = 2$. The user also specifies a prior distribution for model coefficients and the residual variance. Chipman et al. suggest trying two bracketing values for normalized regression coefficients ($c = 1$ and 3), where smaller c values result in estimated coefficients that are shrunk towards 0 and smaller trees. Chipman et al. also suggest bracketing values for the fourth model parameter describing variability, of $0.404 s^2$ and $0.1173 s^2$, where s^2 is the residual variation. In practice, each regression tree model is run four times to cover the span of suggested a priori tree coefficients, and the tree with the largest log-likelihood value (minimum Aikake criterion) is selected. To avoid overfitting and facilitate selection of a robust solution, Chipman et al. suggest choosing the “most visited” tree among the multiple iterations rather than the overall “best” fitting tree.

The CGMlidCART program allows the user to specify a training data set to fit the regression tree models and a test data set to provide an independent validation of the models. We selected a random subset of values representing approximately 10% of the full data set ($n=57$) as a test set. In the first round of analyses, we included a full suite of potential classifier variables. In addition to the classification variables in Table 3.2, we added interaction terms for turbidity x depth, stream power (watershed area x slope), stream power x antecedent precipitation, and stream power x antecedent precipitation x % sands and fines (an index of potential substrate disturbance). To test the robustness of the original regression tree results (Dodds-type, TNTP), we repeated the analysis an additional nine times with different random subsets for the training and test sets and then chose a subset of classifiers (reduced set) based on their frequency of selection in the full models.

The ten runs of the Dodds-type TNTP model with full set of classifiers yielded trees of various sizes, i.e., from 9 to 21 final nodes (Table D.2). The first run yielded a relatively high explanatory power for the training test set ($r^2 = 0.84$), with only a slightly lower value for the validation test set ($r^2 = 0.80$). Model fit based on AIC was even better for the Dodds-type DINDIP model (training $r^2 = 0.91$, test $r^2 = 0.88$; Table D.2).

Table D.2. Results of Bayesian CART analysis of full data set using all potential classification variables (see Table .2 and Section 4). Predictor variables included TN, TP, TN2, TP2, days accrual, and days accrual2 (TNTP models) or the same variables with DIN and DIP substituted for TN and TP (DINDIP models).

Independent Variable Type	# Independent regression variables	CGM tree fitting parameters		Training Set	Log Likelihood	Most Visited Tree Size	AIC	Predicted vs. Observed r2	
		C	Variance					Training	Test
DINDIP	6	1	.117s2	1	955.081708	17	-1672.16	0.91	0.88
TNTP	6	1	.404s2	1	887.325251	21	-1480.65	0.84	0.8
TNTP	6	1	.404s2	2	942.584144	21	-1591.17		
TNTP	6	1	.404s2	3	1049.597041	20	-1819.19		
TNTP	6	1	.404s2	4	1082.228125	27	-1786.46		
TNTP	6	1	.404s2	5	912.540841	21	-1531.08		
TNTP	6	3	.404s2	6	840.74299	9	-1555.49		
TNTP	6	1	.117s2	7	973.827137	20	-1667.65		
TNTP	6	1	.117s2	8	947.849706	19	-1629.7		
TNTP	6	1	.404s2	9	886.811485	12	-1605.62		
TNTP	6	1	.404s2	10	1032.605069	20	-1785.21		

We chose the reduced set of classification variables based on the set of classifiers that were selected in at least half of the full runs (Table D.3). Even though PSA ecoregion only occurred in half of the full runs, it was retained for testing because it was potentially redundant with the latitude and longitude classifiers. Other classifiers in the reduced set included NH₄ (mg N/L), CODE_21_2000_5K (a measure of localized urbanization), and Julian day. To avoid redundancy, NH₄ was only included as a classifier in the TNTP models but was dropped from the DINDIP models.

Table D.3. Frequency of inclusion of classification variables in Bayesian CART TNTP models (Training sets 1-10). Only class variables occurring more than twice are listed.

Frequency	Classification Variable	Definition
46	Longitude	Degrees longitude
44	Latitude	Degrees latitude
20	NH4	Instream NH ₄ value (mg NH ₄ -N/L)
8	CODE_21_2000_5K	Percent NLCD "Code 21" land use within a 5-km radius from sampling site
7	JulianDay	Day of year (1-365)
5	PSAc	Perennial Stream Assessment ecoregion (1-6)
4	Year	Year of sample
4	Conductivity	Instream conductivity
3	REFSITESTAT	Site disturbance status (Reference, Intermediate, Stressed) as defined in Section 2

Table D.4. Results of Bayesian CART analysis of full data set that includes PSA ecoregion. Models used reduced set of four potential classification variables (PSA ecoregion (PSAci), Julian Day, NH₄, and Urban5K). Training set used is 1. Predictor variables for Dodds-type models included TN, TP, TN2, TP2, days accrual, and days accrual2 (TNTP models) or the same variables with DIN and DIP substituted for TN and TP (DINDIP models). Predictor variables for the QUAL2K-type models also included temperature, incident light, turbxdepth (see Table D.3 for definitions). Models were also run with or without latitude and longitude as predictors. Model numbers are provided for a subset for ease of reference in the text.

Model No.	Model Type	Independent Variable Type	Lat/Long included	Number independent regression variables	Log Likelihood	Most Visited Tree Size	AIC	Predicted vs. Observed r2		Final Classification Variables	
								Training	Test		
Model 1	Dodds	TNTP	No	6	583.02	5	-1096.05	0.51	0.44	JulDay NH4	Urban
	Dodds	TNTP	Yes	8	736.17	7	-1346.33	0.79	0.81	JulDay NH4	Urban
	Dodds	DINDIP	No	6	473.53	2	-919.064	0.32	0.23	Urban	
	Dodds	DINDIP	Yes	8	659.40	4	-1246.8	0.7	0.63	JulDay PSAci	
	QUAL2K	DINDIP	No	9	462.76	2	-885.526	0.34	0.29	Urban	
Model 2	QUAL2K	DINDIP	Yes	11	635.90	4	-1175.79	0.69	0.66	JulDay	Urban

Table D.5. Results of Bayesian CART analysis of full data set that uses an empirical rather than PSA ecoregion. Models used include a reduced set of five potential classification variables (Latitude, Longitude, Julian Day, NH₄, and Urban5K). Training set used is 1. Predictor variables for Dodds-type models included TN, TP, TN2, TP2, days accrual, and days accrual2 (TNTP models) or the same variables with DIN and DIP substituted for TN and TP (DINDIP models). Predictor variables for the QUAL2K-type models also included temperature, incident light, turbidity x depth. Models were also run with or without latitude and longitude as predictors. Model numbers are provided for a subset for ease of reference in the text.

Model No.	Model Type	Independent Variable Type	Lat/ Long included	Number independent regression variables	Log Likelihood	Most Visited Tree Size	AIC	Predicted vs. Observed r2		Final Classification Variables	
								Training	Test		
Model 3	Dodds	TNTP	No	6	1162.23	21	-2030.45	0.91	0.79	Lat Long	JulDay NH4
	Dodds	TNTP	Yes	8	1249.05	23	-2084.1	0.92	0.57	Lat Long	JulDay NH4
	Dodds	DINDIP	No	6	1030.95	21	-1767.9	0.84	0.71	Lat Long	JulDay Urban
	Dodds	DINDIP	Yes	8	1120.95	19	-1899.91	0.93	0.81	Lat Long	JulDay Urban
	QUAL2K	DINDIP	No	9	952.50	16	-1585	0.85	0.66	Lat Long	JulDay Urban
Model 4	QUAL2K	DINDIP	Yes	11	1033.65	18	-1635.3	0.9	0.72	Lat Long	JulDay Urban

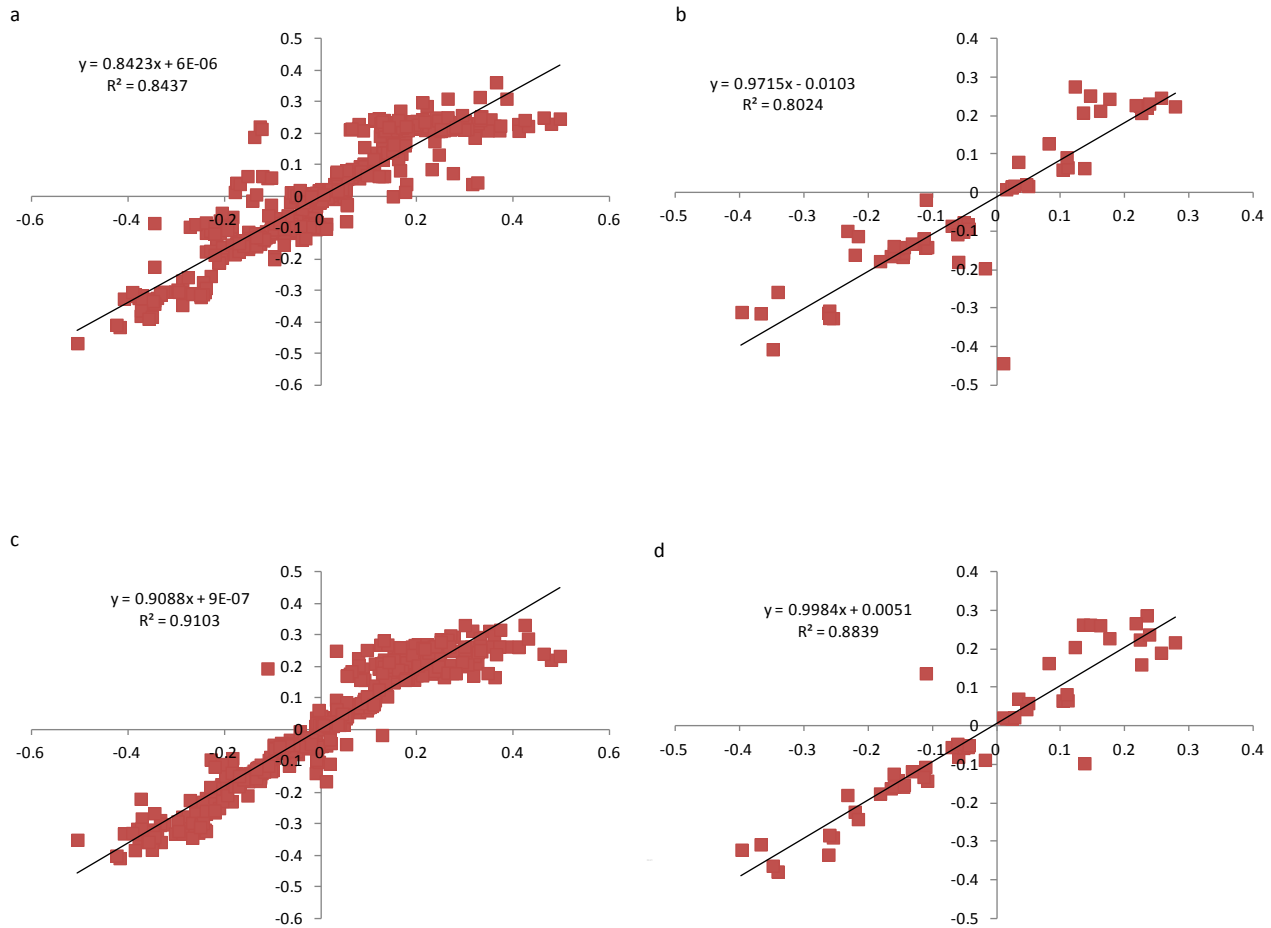


Figure D.9. Predicted versus observed normalized \log_{10} chlorophyll a biomass (mg/m²) for a) TNTP training set, b) TNTP test set, c) DINDIP training set, and d) DINDIP test sets used in Bayesian CART analysis (Dodds-type model, all potential classifiers). The line represents the fit of a linear regression predicting “predicted benthic algal biomass (chl a)” as a function of observed benthic biomass (\log_{10} chlorophyll a).

Appendix E. Quality Assurance/Quality Control Summary

Research conducted to produce this report followed guidelines specified in an approved Quality Assurance Plan (QAPP-AED-WDB-ND-2010-r2-01). The original water quality, biotic, and habitat data used in these analyses were not collected by this project but as part of existing California state or regional monitoring programs with existing approved QA plans (see Sections 2.2.2, 3.2.4, and 4.2.3). The quality assurance parameters for the California datasets used are based on those established for the Surface Water Ambient Monitoring Program (SWAMP 2008). General Quality Objectives for the state monitoring programs are described in Element A7 of SWAMP (2008):

SWAMP seeks to meet the following four objectives:

- *Create an ambient monitoring program that addresses all of California's hydrologic units using consistent and objective monitoring, sampling, and analytical methods; consistent data quality assurance (QA) protocols; and centralized data management.*
- *Document ambient water quality conditions in potentially clean and polluted areas. The scale for these assessments ranges from site-specific to statewide.*
- *Identify specific water quality problems preventing the State Board, the Regional Boards, and the public from realizing beneficial uses of water in targeted watersheds.*
- *Provide data to evaluate the overall effectiveness of regulatory water quality programs in protecting beneficial uses of California's waters.*

By definition, if the general Quality Objectives above are met, then the data collected under these monitoring programs should be of sufficient quality to meet the needs of the current project.

Measurement Quality Objectives for the SWAMP monitoring programs are available at http://www.swrcb.ca.gov/water_issues/programs/swamp/mqo.shtml. The changes in MQOs between 2008 and 2013 alluded to on that web site do not apply to any of the parameters included in the present study. Standard Operating Procedures used in the collection and processing of the samples under the established monitoring plans include Ode (2007) for stream benthic macroinvertebrates and habitat parameters and Fetscher et al. (2009) for benthic stream algae. QAQC protocols for bioassessment methods were supplemented with guidance found in QA Memos 1 and 2 (see http://www.swrcb.ca.gov/water_issues/programs/swamp/mqo.shtml)

Probability-based sampling frameworks used by the State of California Perennial Stream Assessment (PSA) and the southern California Stormwater Monitoring Coalition (SMC) monitoring programs are described in Section 2.2.2 of this report. Although each monitoring program was designed separately to assess the condition of perennial wadeable streams in California, the geographic scope of each differed so new sample weights had to be assigned when these data sets were combined. The calculation of adjusted sample weights used in creating composite cumulative probability distributions is also described in that section.

Objectives for targeted sampling frameworks differed from probability samples. These data from targeted sites come from the state's Reference Condition Management Program (RCMP) and a recently completely

project geared toward developing stream algal assemblage data for use in bioassessment of stream condition. Selection criteria for these monitoring programs are described in Ode and Schiff (2009)

In some cases, subsets of data from the monitoring programs were used based on the degree of anthropogenic disturbance associated with sites. Selection criteria for these disturbance classes are also described in Section 2.2.2 of this report.

In general, the data collected under the four monitoring programs described above were used for the intended purpose, i.e., to describe the ambient and/or reference condition of perennial wadeable streams in California or regions thereof and to determine whether designated uses were being met. As discussed in Section 2.2, the sampling window for California’s bioassessment programs has been chosen to maximize the chances of yielding complete samples across a range of wet and dry years. It was not chosen to assess the temporal variability in benthic algal biomass or to necessarily capture the maximum values of benthic algal parameters. However, given the need of the current study to assess the relationship between benthic algal biomass and community composition, it was appropriate to use samples collected concomitantly.

With the exception of a few field duplicates, the data available for the CA state monitoring programs do not include time series of stream nutrient concentrations or measurements of the full suite of nutrient forms. To assess the representativeness of instantaneous samples of nutrients collected during the growing season for state bioassessment monitoring programs, we analyzed data for 60 California NAWQA stations sampled at a fixed frequency over the year. We downloaded data from the USGS NAWQA Data Warehouse (U.S. Geological Survey 2001).

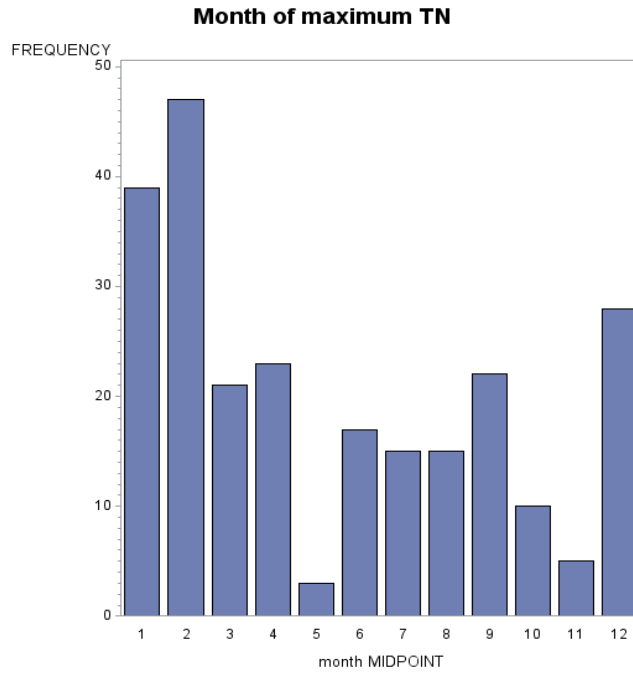
Table E.1. Nutrient fractions for samples from 47 USGS NAWQA stream stations in California sampled biweekly over the year.

	Average	Minima	Maxima	Std
Fraction dissolved inorganic N	0.53	0.05	0.97	0.24
Fraction dissolved N	0.89	0.57	1.20	0.13
Fraction particulate N	0.15	0.01	0.33	0.10
Fraction soluble reactive P	0.75	0.25	1.55	0.27
Fraction dissolved P	0.69	0.25	2.00	0.28
Fraction particulate P	0.30	0.00	0.75	0.22

Most of the total N and P in the CA NAWQA streams (Table E.1 above) are in dissolved form, as they are in the CA bioassessment streams. Although DIN and soluble reactive P are the most bioavailable forms, algae can utilize dissolved organic N and P as well. For the NAWQA streams, maximum total N and P tend to occur in January – February (Figure E.1a,b), followed by declines, then an increase during the growing season (Figure E.2a,b). TP tends to peak in July, while TN slowly increases from May through September. There is a fair amount of variability among biweekly TN and TP values over the growing season, which contributes to the noise in the relationships we have assessed. The predictive power of relationships could probably be improved by averaging values for multiple water quality samples over

a)

a)



b)

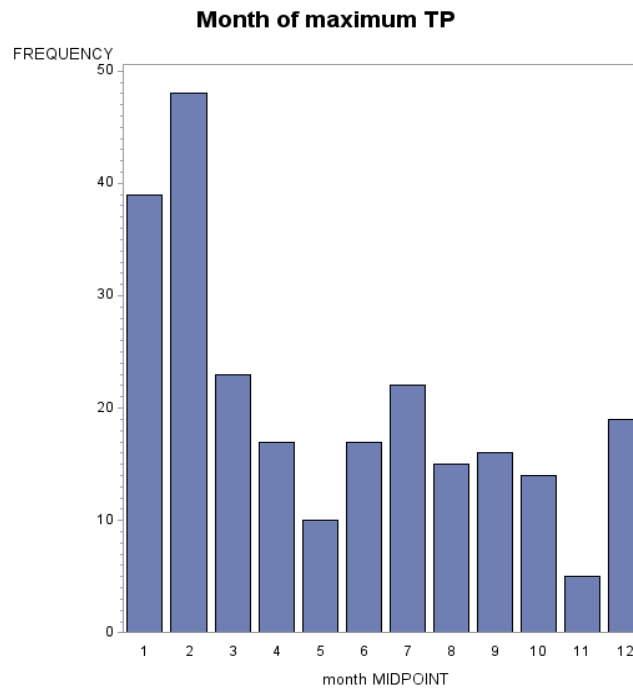
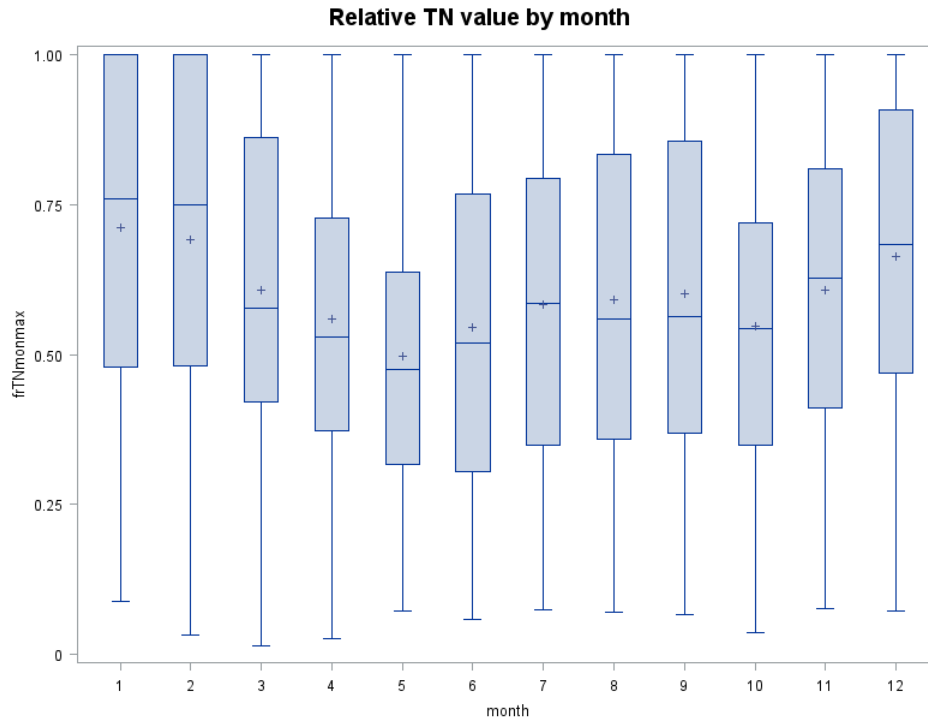


Figure E.1. Frequency of month of a) maximum annual total N and b) maximum annual total P in 47 California NAWQA streams.

a)



b)

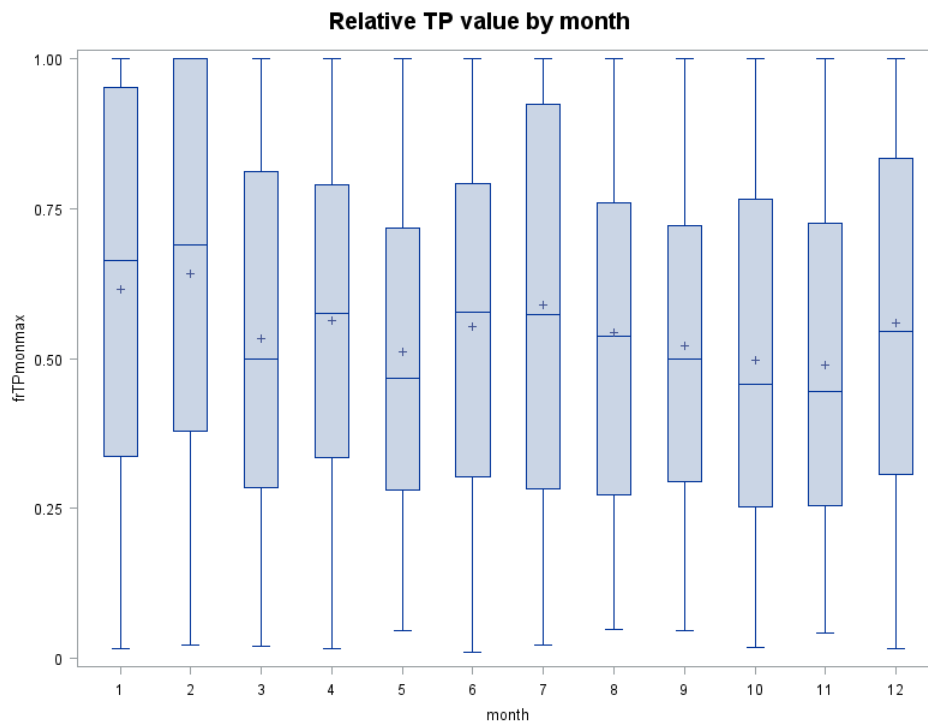
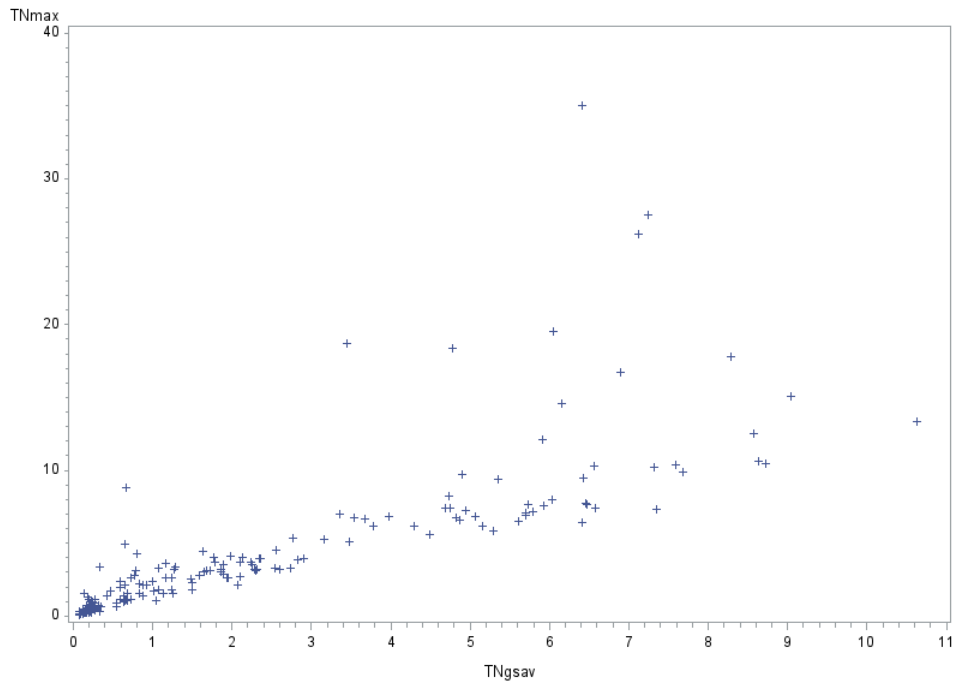


Figure E.2. Value for a) total N and b) total P by month relative to maximum monthly value in 47 California NAWQA streams.

a)

Relationship between annual maximum and average growing season TN



b)

Relationship between annual maximum and average growing season TP

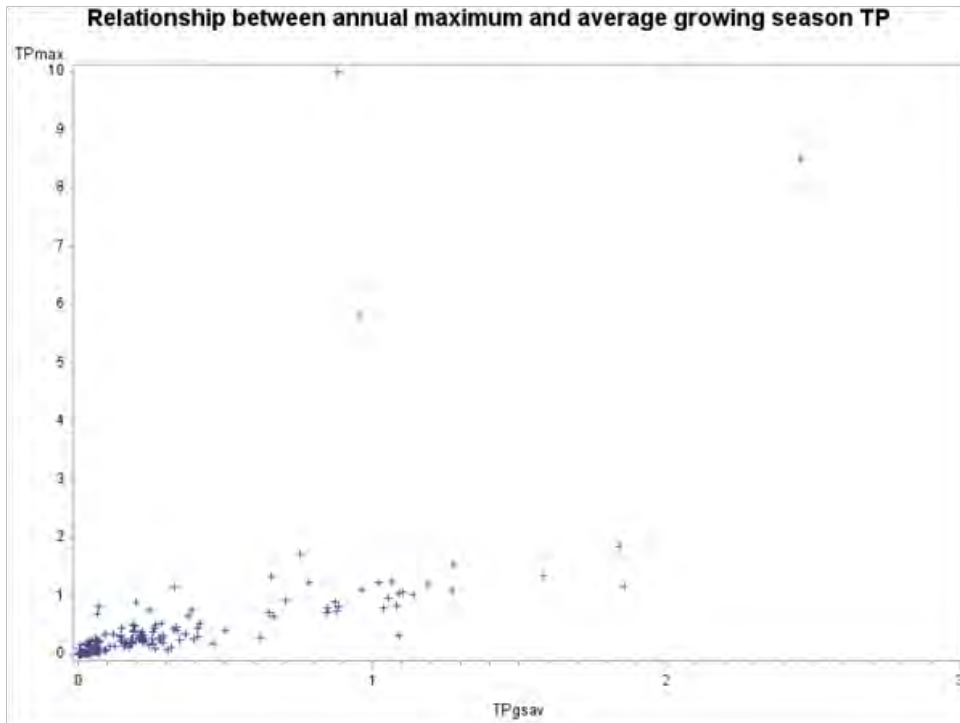


Figure E.3. Relationship between annual maximum and growing season average values for a) total N and b) total P in 47 California NAWQA streams.

the months preceding collection of benthic algal samples, or by using nutrient values inferred by diatom taxa composition (Pan and Stevenson 1996). Lohman and Priscu (1992) have documented evidence of luxury consumption of phosphate by *Cladophora* (in an N-limited portion of the Columbia River) such that ambient SRP was inversely related to cellular P content. In that same river segment, however, cellular N content of *Cladophora* did track ambient dissolved inorganic N levels, the limiting nutrient. In most cases represented by the CA NAWQA dataset, annual maxima recorded for TN and TP increase linearly in proportion to growing season averages and thus growing season values should represent the relative trophic condition of streams and facilitate cross-comparisons among systems (Figure E.3a, b). The few outliers in this relationship coincided with enriched systems.

The optimum period for stream algal assessments has not been established. A plot of stream survey data suggest that South Coast sites may be exhibiting a peak in June-July, although this trend may be confounded by disturbance class of sites sampled (e.g., urban streams sampled in late season). No seasonal peak is obvious for the remaining sites in the State (Figure E.4). When biomass values are normalized for TN, there is no consistent pattern of increasing biomass over the growing season in either South Coast or other sites, which would be captured by the “accrual” term in prediction equations (Figure E. 5).

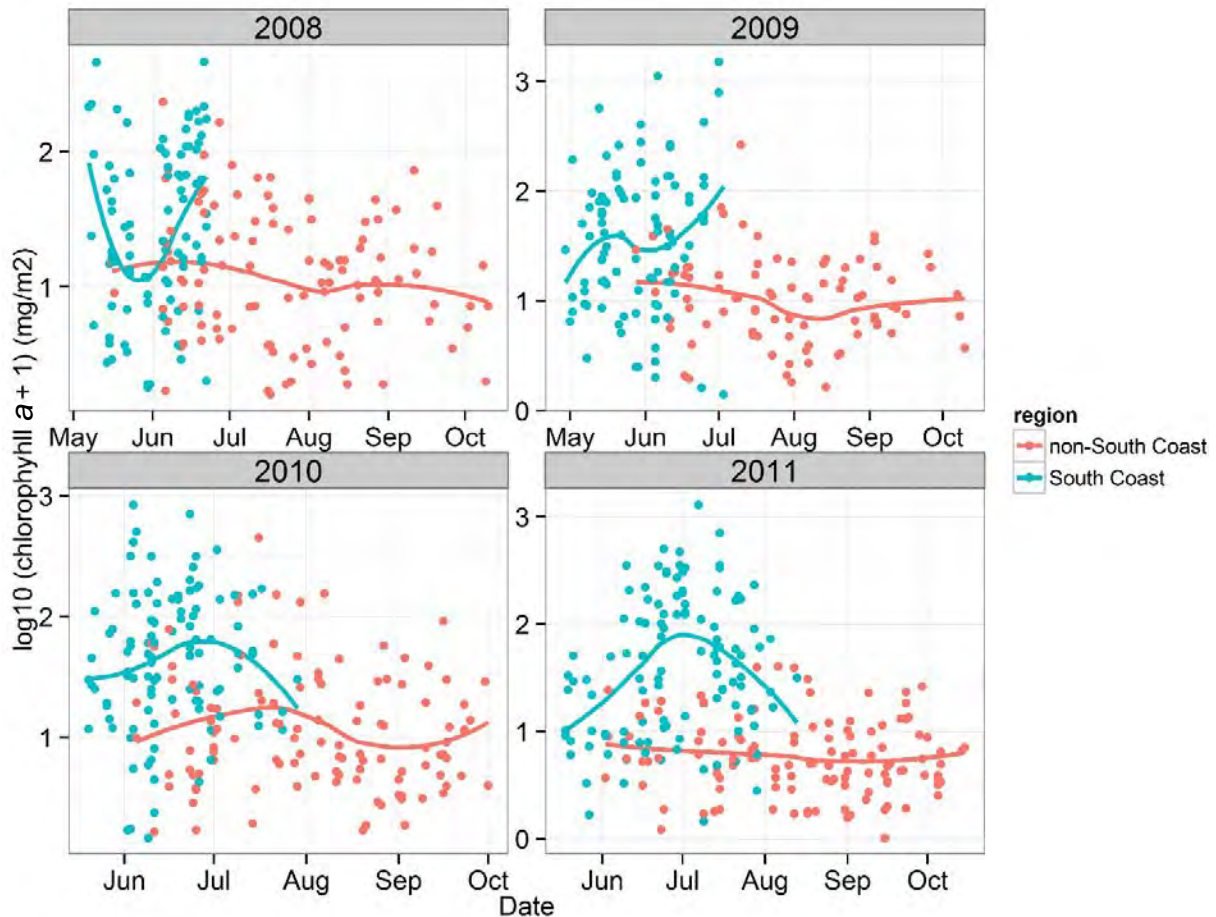


Figure E.4. Chlorophyll α levels (log-transformed) across sampling dates, by year, for South Coast (blue) and all other sites (red) within the state. Curves show time-averaged trends.

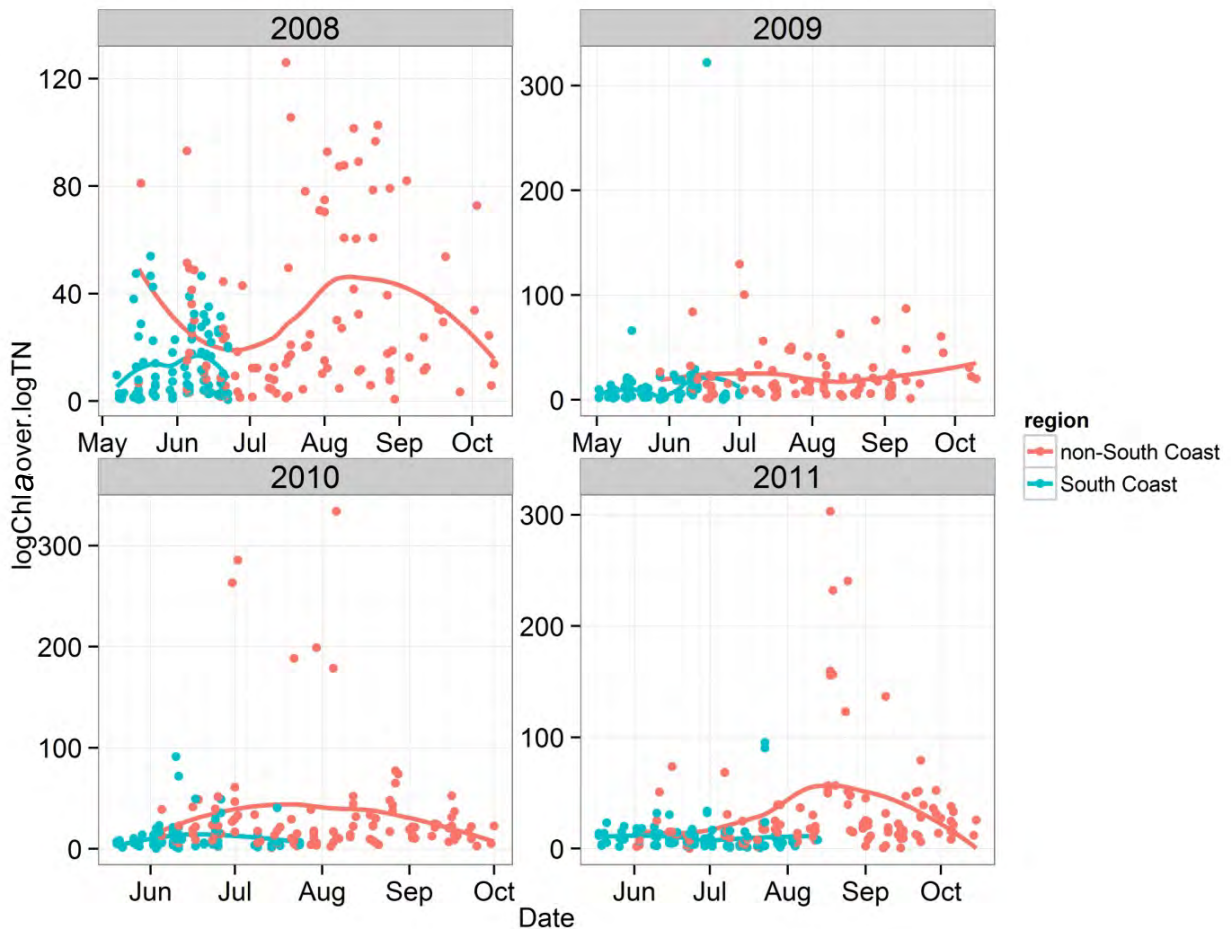


Figure E.5. Chlorophyll α levels (log-transformed) across sampling dates normalized by log total N, by year, for South Coast (blue) and all other sites (red) within the state. Curves show time-averaged trends.

Literature Cited

- Fetscher A.E., Busse L.B., and P.R. Ode. 2009. Standard Operating Procedures for Collecting Stream Algae Samples and Associated Physical Habitat and Chemical Data for Ambient Bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 002. (updated May 2010)
- Lohman, K. and J.C. Priscu. 1992. Physiological indicators of nutrient deficiency in Cladophora (Chlorophyta) in the Clark Fork of the Columbia River, Montana. *Journal of Phycology*, 28:443-448.
- Ode P. 2007. SWAMP Bioassessment Procedures: Standard operating procedures for collecting benthic macroinvertebrate samples and associated physical and chemical data for ambient bioassessment in California. Available from http://swamp.mpsl.mlml.calstate.edu/wp-content/uploads/2009/04/swamp_sop_bioassessment_collection_020107.pdf
- Ode P and K. Schiff. 2009. Recommendations for the Development and Maintenance of a Reference Condition Management Program (RCMP) to Support Biological Assessment of California's Wadeable Streams. SCCWRP Technical Report #581. www.sccwrp.org

Pan, Y. and R.J. Stevenson. 1996. Gradient analysis of diatom assemblages in Western Kentucky wetlands. *Journal of Phycology* 32:222-232.

SWAMP. 2008. Quality Assurance Program Plan of the California Surface Water Ambient Monitoring Program.

http://www.waterboards.ca.gov/water_issues/programs/swamp/docs/qapp/qapprp082209.pdf

U.S. Geological Survey 2001. National Water Information System (NWISWeb) [Surface Water/Bed Sediment]: U.S. Geological Survey database, accessed July 28, 2014, at <http://water.usgs.gov/nawqa/data>.



United States
Environmental Protection
Agency

Office of Research and Development
National Health and Environmental
Effects Research Laboratory
Atlantic Ecology Division
Narragansett, RI 02882

Official Business
Penalty for Private use
\$300