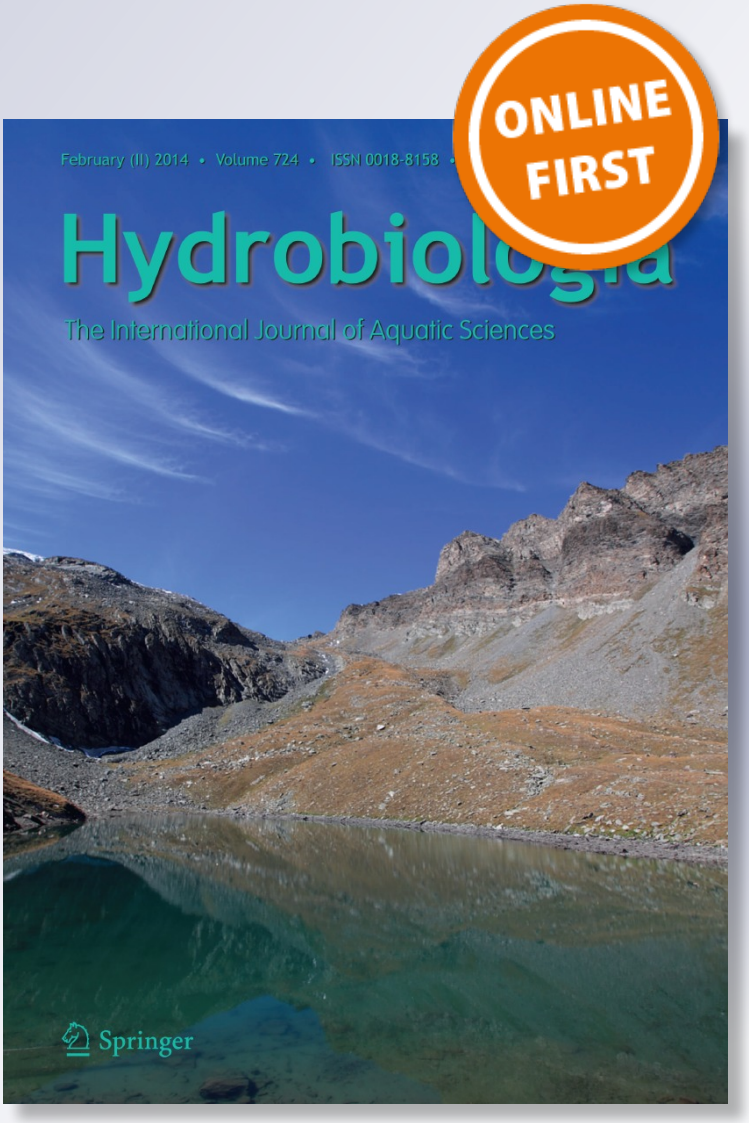
Agricultural disturbance response models for invertebrate and algal metrics from streams at two spatial scales within the U.S.

Ian R. Waite

Hydrobiologia
The International Journal of Aquatic Sciences

ISSN 0018-8158

Hydrobiologia
DOI 10.1007/s10750-013-1774-4



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media Dordrecht (outside the USA). This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com”.

Springer

Author's personal copy

Hydrobiologia
DOI 10.1007/s10750-013-1774-4

PRIMARY RESEARCH PAPER

Agricultural disturbance response models for invertebrate and algal metrics from streams at two spatial scales within the U.S.

Ian R. Waite

Received: 26 February 2013 / Revised: 25 November 2013 / Accepted: 27 November 2013
© Springer Science+Business Media Dordrecht (outside the USA) 2013

Abstract As part of the USGS study of nutrient enrichment of streams in agricultural regions throughout the United States, about 30 sites within each of eight study areas were selected to capture a gradient of nutrient conditions. The objective was to develop watershed disturbance predictive models for macroinvertebrate and algal metrics at national and three regional landscape scales to obtain a better understanding of important explanatory variables. Explanatory variables in models were generated from landscape data, habitat, and chemistry. Instream nutrient concentration and variables assessing the amount of disturbance to the riparian zone (e.g., percent row crops or percent agriculture) were selected as most important explanatory variable in almost all boosted regression tree models regardless of landscape scale or assemblage. Frequently, TN and TP concentration and riparian agricultural land use variables showed a threshold type response at relatively low values to biotic metrics modeled. Some measure of habitat condition was also commonly selected in the final invertebrate models, though the variable(s) varied across regions. Results suggest national models tended to account for more general landscape/climate differences, while regional models incorporated both broad landscape scale and more specific local-scale variables.

Keywords Scale · Stream bioassessment · Predictive models · Riparian · Invertebrates · Algae

Handling editor: Sonja Stendera

Electronic supplementary material The online version of this article (doi:10.1007/s10750-013-1774-4) contains supplementary material, which is available to authorized users.

I. R. Waite (✉)
Oregon Water Science Center, U.S. Geological Survey, 2130 SW 5th Ave, Portland, OR 97211, USA
e-mail: iwaite@usgs.gov

Introduction

Ecological modeling has increased tremendously in the last decade, and models have been developed individually for both stream invertebrate and algal indicators in select regions. Stream ecologists are trying to understand the spatial scales and processes associated with human and natural disturbances that are affecting the biota. Extensive research has concluded that nutrients are a major cause of impairment in streams, rivers, lakes, and estuaries and that agricultural land use practices at various scales are a leading contributor of nutrients to these receiving waters (Allan, 2004; Herlihy & Sifneos, 2008; Stevenson et al., 2008; Dubrovsky et al., 2010; Munn et al., 2010). The destruction of natural riparian areas along stream corridors and loss of wetlands in the watershed has further exacerbated nutrient loading from agricultural lands by reducing or eliminating nutrient uptake, denitrification, and sedimentation of

Published online: 18 December 2013

Author's personal copy

adsorbed phosphorus (Verhoeven et al., 2006). As a result, the development of nutrient criteria has become an important pursuit by the state and federal agencies. For example, the Environmental Protection Agency (EPA) divided the U.S. into 14 nutrient ecoregions to account for the inherently large spatial variability in natural and anthropogenic influences across the U.S. There have also been a large number of studies trying to understand the effects of nutrients on biological integrity in streams with a focus on macroinvertebrates and periphyton assemblages (Waite et al., 2004; Smith et al., 2007; Lücke & Johnson, 2009; Black et al., 2010; Larned, 2010; Yuan, 2010). Most of these studies have focused on a specific region or at one spatial scale; few studies have been conducted across various spatial scales to better understand the effects of different natural and agricultural practices on biological assemblages across these scales. As an exception to this, Riseng et al. (2011) developed structural equation models using the same dataset analyzed here and found that national models did not account for important differences in biotic responses to agricultural disturbance that were captured by regional models.

Macroinvertebrates are the most commonly used biological assemblage in bioassessments nationally and internationally, but they are not directly affected by nutrients and as a result may be less sensitive than algal assemblages to the effects of agricultural land use and nutrient enrichment in particular (Wang et al., 2007; Stevenson et al., 2008; Yuan, 2010). There is some concern that relying primarily on macroinvertebrates for nutrient-related bioassessments may not adequately identify lower trophic level impacts (Brown et al., 2009; Friberg et al., 2010; Munn et al., 2010). Waite (2013) found that invertebrate and algal metrics both had strong correlations to nutrients and agricultural disturbance but that the strength varied regionally and by biotic group, suggesting that no single biotic group works well in all regions. This concurs with what other researchers have suggested (Hughes & Peck, 2008; Brown et al., 2009), that difference in the response of the different biological assemblages suggests that when possible, it is important to evaluate more than one indicator in each bioassessment region.

The hierarchical connection of landscape disturbance such as agricultural practices to instream nutrient concentrations and then subsequent changes to biological communities is complex with many possible interactions among biotic and environmental variables (e.g., biological uptake, light, habitat, temperature, nutrient cycling and limitation, etc.). Riseng et al. (2004) found that algal biomass was influenced by three variables describing very different environmental components: high-flow disturbance, nutrient concentration, and grazer biomass with various interactions among these variables. This complexity and interaction among environmental factors can often result in poor correlations of land use with nutrient concentrations and nutrients with biological indicators. In addition, landscape spatial scale often is also important; Potapova & Charles (2007), for example, found that nutrient-diatom relationships were improved when regional factors were included. Waite (2013) and Riseng et al. (2011) also found that the strength of correlations of invertebrate and algal metrics to agricultural disturbance variables (agricultural intensity and nutrient concentrations) varied with spatial scale and geographic region.

Models provide a useful framework for testing hypotheses, determining potential direct and indirect linkages (Riseng et al., 2011), and directing where further research is needed. Turak et al. (2010) were able to use multiple linear regression (MLR) models to predict stream macroinvertebrate observed/expected (O/E) metrics using three local disturbance factors (point sources, housing, and flow disturbance) explaining approximately 76% of the variation. Waite et al. (2010) developed macroinvertebrate response models for three regions in the western United States, and the best MLR models from each individual region required only two or three explanatory variables (urban and/or agricultural land use plus a natural setting variable) to model macroinvertebrate metrics that explained 41–74% of the variation. On the other hand, Clapcott et al. (2012) found that land use variables only explained between 6 and 36% of the variation in the presence/absence of invertebrate metrics in New Zealand, and responses for fish metrics were even lower. Brown et al. (2009) and Cuffney et al. (2010) studying the effects of urbanization in nine metropolitan areas found a wide range in the strength of correlations of urban intensity to macroinvertebrate and algal indicators, from 0.2 to 0.9 for invertebrates and 0.05 to 0.8 for algae.

Models of biological indicators related to measures of watershed disturbance, and natural factors are becoming more common, but still little has been done relating and comparing disturbance to multiple biotic

Author's personal copy

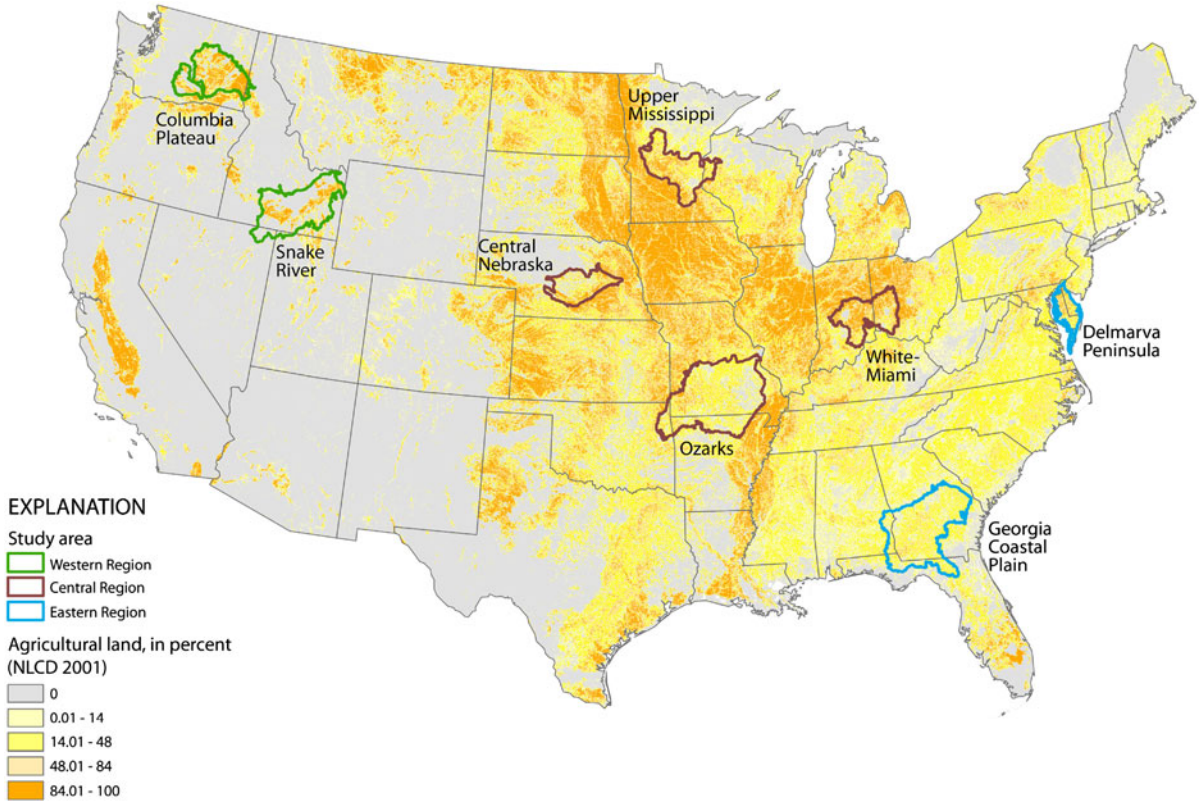


Fig. 1 Map showing agricultural land in percent and the eight study areas coded by region: west, central, and east

indicators at multiple spatial scales. The objectives of this research were to develop models of agricultural intensity relating the response of algal and macroinvertebrate metrics to variables describing landscape disturbance, instream habitat, and water quality (e.g., nutrient enrichment) at multiple landscape spatial scales to obtain a better understanding of important explanatory variables that predict the biotic condition of agriculturally influenced streams. Do watershed and riparian scaled variables predict algal and invertebrate metrics equally among various geographic regions that are different in agricultural intensity and natural settings? If not, how can model development and application of these biotic indicators across various agricultural settings be informed by these differences? Our general hypothesis is that algal and invertebrate indicators will respond to agricultural intensity and disturbances in different ways, and each will highlight unique watershed and riparian explanatory variables. All data was collected as part of the U.S. Geological Survey's National Water Quality Assessment program (NAWQA).

Methods

Study area

This study was conducted in eight study areas characterized by extensive agricultural land use and located throughout the United States (Fig. 1). To take into account regional differences due to both natural and anthropogenic influences, the study area datasets were aggregated into (1) a single national dataset and (2) into three regional datasets based on unit proximity and similarities in climate and physiography (Cushing & Allan, 2001; Riseng et al., 2011). The Columbia Plateau (CCYK) and the Upper Snake River study areas (USNK) are arid to xeric and were grouped as the Western region; all sites were located in the western US and within the Columbia River basin. Agriculture is heavily dependent on irrigation, while forests and wetlands are relatively rare. The Upper Mississippi (UMIS), Central Nebraska (CNBR), White-Miami (WHMI), and the Ozarks (OZRK) study areas are located in the semi-

Author's personal copy

Table 1 Description, variable code, and definition of explanatory environmental (landscape, riparian, habitat, and water chemistry) and response (invertebrate and algal metrics) variables used for model development

Explanatory variables: landscape		
Description	Variable Code	Definition
Watershed scale variables		
Percent agricultural landuse	WS_AG	Percent watershed area in agricultural landuse (NLCD 2001 category 81, 82)
Mean annual precipitation	WS_MnAnnPrecip	Mean annual precipitation (cm)
Soil infiltration rate	Soil_Mod-Infil	Hydrologic soil group B, moderate infiltration rate (min. infiltration rate 4–8 mm/h)
Total nitrogen input	TN Input	Input to the watershed of total nitrogen from fertilizer application, atmospheric deposition and animal manure (kg/km^2)
Total phosphorus input	TP Input	Input to the watershed of total phosphorus loading from fertilizer application, atmospheric deposition and animal manure (kg/km^2)
Mean annual temperature	Mean Temp.	Mean annual temperature for the watershed for years 1980–1997 (degrees Celsius) (Daymet)
Riparian scale variables		
Percent agricultural landuse	Rip_AG	Percent buffer area in agricultural landuse (NLCD 2001 category 81, 82) in riparian
Percent row crops in riparian	Rip RowCrop	Percent buffer area in row crops (NLCD 2001 category 81)
Percent agricultural landuse in 100 m riparian buffer (orthophoto)	AG Rip_Buf	Percent agricultural landuse (cropland + farmstead) within 100 m riparian buffer for the stream segment based on orthophotos
Habitat variables		
Average instream velocity	Ave. Velocity	Average water column velocity measured at 3 locations at 11 equidistant transects
Average silt	Ave. Silt	Average percent silt covering dominant substrates measured at 3 locations at 11 equidistant transects
Percent fines	% Fines	Percentage of fine substrates (sand and silt categories) measured at 3 locations at 11 equidistant transects
Percent habitat cover	Habitat Cover	Percent occurrence of all fish habitat cover among 5 measured locations at 11 equidistant transects
Percent canopy cover	Canopy Cover	Average percent canopy closure, mid-channel measurements
Percent run habitat	Runs	Percent run habitat type in the stream reach
Water chemistry variables		
Dissolved oxygen concentration	DO Conc.	Average DO concentration measured during two water quality samplings (mg/l)
Total nitrogen concentration	TN Conc.	Average TN concentration measured during two water quality samplings (mg/l)
Total phosphorus concentration	TP Conc.	Average TP concentration measured during two water quality samplings (mg/l)
Water temperature	Temp.	Average water temperature measured during two water quality samplings (Centigrade)
Response variables: invertebrate metrics		
Observed/expected	O/E	Ratio of number of observed taxa at a site over the expected taxa based on null model of reference sites from each region

Author's personal copy

Table 1 continued

Explanatory variables: landscape		
Description	Variable Code	Definition
Tolerant richness	RICHTOL	Average USEPA tolerance values for sample based on richness
Ephemeroptera, Plecoptera, and Trichoptera richness	EPTR	Richness composed of Mayflies, Stoneflies and Caddisflies for a sample
Algae metrics		
Observed/expected	O/E Algae	Ratio of number of observed taxa at a site over the expected taxa based on null model of reference sites from each region
Percent eutrophic diatoms plus soft algae	Eutrophic_D	Percentage of diatoms and soft algae classified as eutrophic
Percent low inorganic nitrogen diatoms	ON_AL	Percentage of diatoms classified as low inorganic nitrogen based on van Dam (1994)

humid plains and were grouped in the Central region; all sites were located in the central US, draining to the Mississippi River. Preliminary analysis also suggested that the Ozarks fit better in the central rather than the eastern region (Riseng et al., 2011). Wetlands were more abundant in the Central than Western region. The Georgia Coastal Plain (GCP) and the Delmarva Peninsula (DLMV) study areas were grouped in the Eastern region: all sites located in the eastern US along the coast and draining to the Atlantic Ocean. Coastal Plain streams typically had riparian zones with forested floodplains and abundant canopy cover and the highest amount of wetlands. Analyses were conducted at both the national (across all eight areas) and three regional landscape scales.

Site selection

In each study area, sites were selected to represent a nutrient concentration gradient ranging from "little impacted" to "highly impacted" by agricultural land use. The initial selection of sites relied partially on modeled estimates of nitrogen and phosphorus input to each of the independent watersheds derived from county-level fertilizer sales, atmospheric deposition (nitrate), and livestock data (Ruddy et al., 2006). Final selection of sites was based upon modeled nutrient loading to the watershed, existing USGS nutrient data, and similarity of stream habitat with other sites within the study area. This approach yielded 28–30 wadeable sites within each study area that spanned the greatest range in nutrient concentrations possible, given similar stream habitats within the study area. The sites were sampled between 2003 and 2007, and watershed area ranged from 0.22 to 6,378 km^2 with a median value of 137 km^2, though the range in watershed area for the fifth and ninety-fifth percentile was more restrictive in size (10–1,370 km^2). The largest watersheds were in the dry areas of central Nebraska that had large contributing area, yet the resultant stream sizes were small. Mean and ranges of other select descriptive and environmental variables for each of the eight study areas are provided in Table 2.

Watershed and riparian characterization

Geographical Information System-derived watershed land cover, soil characteristics, topographic features, runoff estimates, and riparian land cover were calculated for each watershed upstream of the study site. Calculations were made using the National Hydrography Dataset (NHD 100 K; US Geological Survey, 2003a), the National Land Cover Dataset (NLCD2001; US Geological Survey, 2003b), and digitized orthophoto quarter quadrangles to characterize riparian conditions as described in Johnson & Zelt (2005). Riparian variables were determined from GIS land cover (NLCD) at the reach scale 250 m from the stream centerline for a set distance upstream (base-10 logarithm of the upstream drainage area) using methods outlined in Johnson & Zelt (2005); this method was used so that riparian area was scaled proportionally to the size of the watershed and not based on a set

Author's personal copy

Table 2 Summary of important environmental variables by study area

Region	Study area	Study area code	Mean annual precipitation (cm)	Regional crops	Percent agriculture	TN loading in watershed	TP loading in watershed	Agricultural intensity index
Western	Columbia Plateau	CCYK	40 (22–91)	Orchards, vineyards, wheat/small grains, alfalfa, dried beans	36 (0–91)	4,100 (330–15,000)	510 (40–2,200)	49 (0–100)
	Upper Snake River	USNK	45 (22–75)	Potatoes, sugar beets, alfalfa wheat/small grains	19 (0–92)	3,200 (200–17,000)	670 (20–3,600)	35 (0–95)
	Central Nebraska	CNBR	68 (59–78)	Corn, soybeans, alfalfa	47 (14–91)	8,600 (3,600–15,000)	1,530 (640–2,600)	80 (45–100)
Central	Ozarks	OZRK	122 (116–130)	Alfalfa, wheat/small grains, grass hay, soybeans	38 (1–81)	4,600 (700–14,000)	1,100 (80–3,800)	49 (0–85)
	Upper Mississippi	UMIS	78 (69–88)	Corn, soybeans, alfalfa, wheat/small grains	56 (2–89)	4,700 (600–10,000)	870 (40–2,000)	68 (20–100)
	White–Miami	WHMI	108 (98–115)	Corn, soybeans	83 (67–92)	8,400 (6,000–13,000)	1,590 (1,100–3,200)	91 (80–100)
Eastern	Georgia Coastal Plains	GCP	123 (117–132)	Cotton, peanuts, soybeans, wheat/small grains	32 (5–65)	2,800 (600–6,500)	750 (80–1,800)	46 (10–80)
	Delmarva Peninsula	DLMV	114 (113–116)	Corn, soybeans, wheat/small grains	61 (23–92)	8,000 (2,300–16,000)	1,470 (190–3,600)	82 (40–100)

Numeric variables presented as average, range within parentheses (TN and TP loading in units of km/km^2)

Author's personal copy

distance. Additional measurements of riparian land cover were estimated using orthophotography calculated within 100-m buffers adjacent to the study reaches (Zelt & Munn, 2009). The specific watershed and riparian variables used in this study are identified in Table 1. Additional natural setting variables included mean elevation, soil permeability, soil type, base flow index, and minimum flow for 30 days prior to biological sampling. Mean annual temperature determined as the average from 1980 to 1997 from Daymet data (Thorton & Running, 1999, University of Montana, 2005).

Habitat characterization

A study reach of approximately 20–30 wetted channel widths in length was established at each site. Reach-scale data were collected along 11 transects distributed proportionally within the study reach. Habitat measures for each study reach included an assessment of instream and bank features such as water velocity, depth, width, fish habitat cover, substrate size, substrate embeddedness, channel bankfull width, bank vegetation cover, canopy closure made using a spherical densitometer, and measurement of instantaneous stream discharge (Rantz, 1982). Habitat characterizations were performed during low-flow conditions. A detailed description of the reach-scale habitat and riparian delineation methods is available in Fitzpatrick et al. (1998) and Biggs & Kilroy (1994). Study reach habitat variables that were collected or derived and used in the data analysis are shown in Table 1.

Water quality/chemistry

Water chemistry samples were collected twice at all sites: once 30 days prior to the habitat characterization and once during the habitat assessments. Samples were collected using isokinetic depth-integrated equal-width increment or multivertical grab samples depending on water depth and velocity, stored on ice, and maintained at 4°C prior to analysis that occurred within a month (Patton & Gilroy, 1998; US Geological Survey, 2006). Field water quality properties were measured during each sampling event and included water temperature, dissolved oxygen, specific conductance, and pH using laboratory meters that were calibrated daily prior to use. Water chemistry analysis included nutrients (TN and TP plus dissolved species), alkalinity, dissolved organic carbon, and suspended sediment. All laboratory analyses for chemical constituents were conducted at the USGS National Water Quality Laboratory in Denver, CO, using methods by Fishman (1993) and Patton and Kryskalla (2003). Table 1 includes the water quality variables that were collected or derived and used in the data analysis.

Algae and macroinvertebrate sampling

Algae and macroinvertebrate samples were collected from erosional (coarse) or depositional (fine-grained) substrate at five locations in the study reaches during the summers of 2003 and 2007. Coarse substrate was defined as either gravel or cobble typical of riffles or large woody debris in areas where sand substrates dominated, while fine-grained substrate was typically sand or silt. A detailed discussion of algal and macroinvertebrate sampling protocols can be found in Moulton et al. (2002) and Cuffney (2003). Where coarse substrates occurred they were sampled for algae by scrapping the biological material off rock or wood from a measured area then composited into one sample with a known quantity of water and preserved with 5% formalin for identification (Moulton et al., 2002). Depositional substrate was sampled for algae by inverting a 47-mm diameter plastic Petri dish on to the substrate, sliding a spatula under the dish, and rinsing the material into a container and preserving it with 5% formalin for identification. All algal community samples were shipped to the Academy of Natural Science in Philadelphia for algal identification, density, and biovolume calculations (Charles et al., 2002). Algae data for all sites were collected during stable low-flow periods, typically between July and September. Algae and diatoms were identified to the lowest taxonomic level possible, usually species or subspecies. Taxonomic resolution varied among sites and over time, so the data were modified before data analysis to establish a consistent level of taxonomy across all sites and time. For macroinvertebrates, five composited samples collected from coarse substrates or woody debris as defined with a 500-Um net were identified to the lowest taxonomic level possible, typically the genus level. For each site, a minimum of 300 organisms were counted and identified.

Over 150 metrics were calculated for each of the algal and macroinvertebrate assemblages using the

Author's personal copy

USGS developed programs Invertebrate and Algae Data Analysis System (ADAS and IDAS; Cuffney, 2003). Metrics were then screened to select three metrics for each assemblage that showed the best initial responses to various agricultural disturbance indicators based on initial Spearman correlations and previous experience (Waite, 2013). Ecological tolerance values indicate how well each taxon is expected to tolerate pollution and were obtained from the U.S. EPA (Wisseman, 1996; Barbour et al., 1999; Cuffney, 2003). Tolerance values ranged from 1 to 10, with low values indicating intolerant taxa (1–4) and high numbers indicating tolerant taxa (7–10). Average tolerance values were calculated using both national and region-specific tolerance estimates; national tolerances were an average of all the regional tolerances (Cuffney, 2003). Final algal and invertebrate metrics that were used in the data analysis are shown in Table 1. See Porter et al. (2008) for more detailed information on algal metrics and Cuffney (2003) for invertebrate metrics. The observed/expected ratio metric, O/E, was calculated following Van Sickle et al.'s (2005) null model method using the low AG-Index category as the minimally impacted expected sites. Only taxa that were identified to at least the genus level and occurred at greater than 0.65 and 0.45% of the sites in the low category for invertebrate and algal data, respectively, were used. Ostermiller and Hawkins (2004) discuss the statistical and biological reasons why use of an intermediate percent capture threshold such as 50% may have advantages over the inclusion of all the taxa; the general intent of this measure is to reduce variability associated with rare species. I adjusted the percentage slightly to better match the number of taxa found at the minimally impacted sites in this study. O/E values were generated for each region separately.

Agricultural intensity index (AG-Index)

An index of agricultural intensity was developed that combines four landscaped-based variables to attempt to assess the overall agricultural intensity within the watershed. The four variables selected were percent agricultural land use in the watershed based on NLCD 2001 data, percent agricultural land use in the riparian zone based on orthophotography interpreted data, and TN and TP individual input to the watershed (Zelt & Munn, 2009). Each variable was given the following scores based on these percentile ranges across all sites for that variable: 0–10% = 0, 11–30% = 2, 31–50% = 3, 51–75% = 4, >75% = 5. The scores were then summed, divided by 20 (the total possible score), then multiplied by 100 to shift the scores to a percent from 0 to 100, with a score of 100 being high agricultural intensity.

Model development

Environmental datasets were quality checked and assembled in MS Excel files then read into R (R Development Core Team, 2006, version 2.10.0). Evaluation and reduction of environmental variables followed a common step-wise approach; variables were assessed for general distribution and possible outliers via scatter plots and correlation statistics; redundant (>0.80 Spearman correlation among similar variables) or surrogate variables were removed. This process was also assessed through the use of Primer's BIO-ENV procedure (PRIMER, version 6: Clarke & Gorley, 2006) which calculates a hierarchical subset of variables that explain the greatest portion of the variance in the original full dataset. This was done separately for each variable grouping (i.e., habitat, water chemistry, and riparian and watershed land cover) before the final subset of variables was pulled together. Overall there were 20–100 variables per variable grouping that were evaluated and screened; habitat had the most initial variables, water chemistry the least. Once a subset of environmental variables and biological metrics were determined, the metrics were related to key indicators of agricultural land use (AG-Index, % agriculture in the watershed and riparian zone, TN and TP instream concentrations, etc.) via Spearman rank correlation statistics (Waite, 2013).

Boosted regression tree (BRT) models were developed for the national scale and each regional scale (Western, Central and Eastern regions). Models were developed for three invertebrate and algal metrics as response variables, each model starting with the same set of explanatory variables. Model performance was assessed using a variety of statistics, including adjusted mean sum of squares (R^2) and root mean squared error (RMSE). Model residuals, potential outliers, and interaction terms were evaluated. A description of variables used in model development is provided in Table 1.

Regression trees are one type of technique within the commonly used decision tree family, and their use

Author's personal copy

and technical details have been described extensively in the literature (e.g., Breiman et al., 1984; De'ath & Fabricius, 2000; Prasad et al., 2006); therefore, I will only provide a brief overview. Decision trees attempt to explain variation in one categorical (classification) or continuous (regression) response variable by one or more explanatory variables, the resultant output being a dendogram, or tree, with varying numbers of branches or nodes. Trees are developed following a hierarchical binary splitting procedure that attempts to find the best single explanatory variable that minimizes within group and maximizes among group dissimilarity in the response variable at each split. A fitted decision tree results in a set of the explanatory variables and their associated predictive power, classification, and regression trees (CART) are an example.

Random forests (RF) and BRT are among a family of techniques used to advance single CART trees by averaging the results for each binary split from numerous trees or forests, thus, reducing the predictive error and improving overall performance (De'ath, 2007; Elith et al., 2008). In BRT, after the initial tree has been generated, successive trees are grown on reweighted versions of the data giving more weight to those cases that are incorrectly classified than those that are correctly classified within each growth sequence. Thus, as more and more trees are grown in BRT, the large number of trees increases the chance that cases that are difficult to classify initially are correctly classified, thus, representing an improvement to the basic averaging algorithm used in RF (De'ath, 2007). Boosted trees and random forest models retain the positive aspects of single trees seen in CART models, yet have improved predictive performance, nonlinearities, and interactions are catered to or easily assessed, and they can provide an ordered list of the importance of the explanatory variables (Cutler et al., 2007; De'ath, 2007). Though RF and BRT offers improved modeling performance over CART, the simple single tree obtained from CART is lost, making it more difficult to visualize the results. Partial dependency plots are a way to visualize the effect of a specific explanatory variable on the response variable after accounting for the average effects of all other explanatory variables (De'ath, 2007; Elith et al., 2008); these are presented in this paper for select models as examples. BRT models were run using the gbm library in R using specific code from Elith et al. (2008). I used untransformed data (Table 1), a bag fraction of 0.75 and a learning rate of 0.001 for developing our models; a bag fraction of 0.75 means that each tree is developed using a random selection of 75% of the data. The learning rate influences the total number of trees evaluated for the model. I used R^2 values to compare MLR and BRT models because it is a common and well understood measure that is common to both approaches (Aertsena et al., 2010). Explanatory variables in the final BRT models were pruned by using a combination of variable importance scores, evaluation of interactions, and partial dependency responses and gbm simplify scores following rules outlined by Elith et al. (2008). I eliminated explanatory variables that did not explain at least 0.07 percent in variable importance or showed a response pattern that did not follow the expected response form. I also took into account whether there were interaction affects among explanatory variables when selecting variables to reduce down to the most parsimonious final set of variables.

Results

Table 2 provides a comparison of important descriptor and environmental variables (mean and range) for each of the eight study areas. To smooth possible single year anomalies, the TN and TP watershed input data presented in Table 2 are averages for the 3-year period 2001–2004 (pers. comm. Zelt). The two study areas in the western region had the lowest average precipitation, unique crop types (i.e., orchards, vineyards, potatoes, etc.), relatively low nutrient input values, and lower average agricultural intensity index values (AG-Index) than the other six study areas. The Delmarva Peninsula, White-Miami, Upper Mississippi, and Central Nebraska areas had higher average % agriculture and AG-Index values plus smaller ranges in these values than the other four study areas (Table 2). On the other hand, Columbia Plateau, Upper Snake, Ozarks, and Georgia Coastal Plan study areas had lower average agriculture intensity but a larger gradient or total range in nutrients and agriculture intensity than the other study areas.

Invertebrate models were developed for EPT richness, richness of tolerant taxa (RICHTOL), and O/E metrics at the national scale and for each of the three regions (Table 3). Models were successfully developed

Author's personal copy

Table 3 Comparison of model evaluation statistics for boosted regression tree (BRT) models for three macroinvertebrate metrics at two landscape spatial scales (national and regional), number of variables in final model in parentheses

Macroinvertebrates					
	Model statistic	National $n = 229$	Western $n = 59$	Central $n = 116$	Eastern $n = 54$
EPT richness					
BRT	R^2	0.69 (5)	0.73 (5)	0.84 (4)	0.63 (5)
	RMSE	2.56	2.54	2.12	1.88
Richness of tolerant taxa (RICHTOL)					
BRT	R^2	0.68 (4)	0.81 (4)	0.88 (5)	0.70 (4)
	RMSE	0.45	0.44	0.27	0.24
Observed/expected taxa (O/E)					
BRT	R^2	0.71 (5)	0.52 (4)	0.85 (3)	0.65 (5)
	RMSE	0.14	0.18	0.11	0.14

EPT Richness—Total taxa richness of mayflies, stoneflies and caddisflies

Table 4 Comparison of explanatory variables for boosted regression trees (BRT) models for three macroinvertebrate metrics at two landscape spatial scales (national and regional)

Macroinvertebrates metrics	National		Western		Central		Eastern	
EPT richness	Rip RowCrop	40	Rip RowCrop	30	TP Conc.	45	AG Rip_Buf	30
	Mean Temp.	25	TN Input	30	Rip RowCrop	23	TP Conc.	25
	% Fines	18	Ave. Velocity	15	TN Conc.	17	Ave. Silt	22
	TP Conc.	9	AG-Index	15	% Fines	15	Ave. Velocity	14
	TN Conc.	8	TP Conc.	11			TN Conc.	9
Richness of tolerant taxa (RICHTOL)	Rip RowCrop	38	Ave. Velocity	26	TP Conc.	45	DO Conc.	40
	Ave. Silt	30	TN Input	22	% Fines	17	TN Conc.	26
	Mean Temp.	19	DO Conc.	18	Rip RowCrop	13	Rip RowCrop	19
	TP Conc.	13	AG Rip_Buf	17	Ave. Silt	13	Ave. Velocity	15
			Ave. Silt	16	Ave. Velocity	12		
Observed/expected taxa (O/E)	Rip RowCrop	48	TN Input	34	% Fines	50	Rip RowCrop	28
	TP Conc.	24	AG-Index	25	Rip RowCrop	26	Habitat Cover	27
	TN Conc.	13	TP Conc.	23	TP Conc.	24	TN Conc.	27
	% Fines	8	Rip RowCrop	17			DO Conc.	9
	AG-Index	7					Ave. Silt	9

EPT richness—Total taxa richness of mayflies, stoneflies and caddisflies. Variables are listed in order of importance in each model with associated model importance values. See Table 1 for variable definitions

at the national scale that had R^2 values similar to the values at some of the individual regions, except for RICHTOL where national models were lower than all the regional models. BRT R^2 values for EPT Richness and O/E were 0.69 and 0.71 nationally, while regionally values showed more variability, ranging from a low of 0.52 (Western O/E) to a high of 0.88 (Central RICHTOL) (Table 3). RICHTOL had the highest R^2 for any of the three invertebrate metrics across the three regions, but O/E had the highest value for the national models. The Central region consistently had the highest BRT model R^2 values, all >0.84, the Western region metrics ranged from 0.52 to 0.81, and the Eastern region had values between 0.63 and 0.70. The explanatory variables in order of importance value in the BRT models for each invertebrate model are provided in Table 4, all final BRT models had between 3 and 5 explanatory variables. There was a high degree of similarity across invertebrate response metrics and spatial scales in the final explanatory variables that were retained in each

Author's personal copy

Table 5 Comparison of model evaluation statistics for boosted regression tree (BRT) models for three algal metrics at two landscape spatial scales (national and regional), number of variables in final model in parentheses

Algae	Model statistic	National $n = 229$	Western $n = 59$	Central $n = 116$	Eastern $n = 54$
% Low inorganic N diatoms (BRT)	R^2	0.77 (5)	0.74 (5)	0.79 (3)	0.84 (4)
	RMSE	11.0	13.1	14.2	10.1
% Eutrophic diatoms (BRT)	R^2	0.78 (4)	0.70 (5)	0.83 (4)	0.73 (5)
	RMSE	12.2	10.9	9.7	13.3
Observed/expected taxa (O/E) (BRT)	R^2	0.72 (5)	0.76 (6)	0.89 (4)	0.78 (3)
	RMSE	0.16	0.12	0.10	0.15

model. TN and/or TP concentration and some measure of riparian land use (either Rip_RowCrop or AG Rip_Buf) were typically one of the most important variables in every model (Table 4). The Central region model for O/E explained 85% of the variation with only three explanatory variables, percent fines, explained 50% of the model variation. Additional important explanatory variables found in the various models included some measure of substrate condition (% Fines or Ave. Silt) or habitat characteristics (Ave. Velocity or Habitat Cover). Mean annual air temperature within the watersheds was in two national models but not in any of the regional models, and average DO concentration and the Agricultural Intensity Index (AG-Index) were each retained in a few models.

Algal models were developed for the metrics % low inorganic N diatoms, % eutrophic diatoms, and O/E at the national landscape scale and for each of the three regions (Table 5). Models were successfully developed at the national scale that had R^2 values similar to the values at the individual regions. BRT R^2 values were between 0.72 and 0.78 nationally, while regional values showed slightly more variability, ranging from a low of 0.70 (Western % Eutrophic) to a high of 0.89 (Central O/E) (Table 5). Comparing among algal metrics, O/E had the highest R^2 value in the West and Central regions and percent low inorganic N diatoms the highest in the Eastern region. Across regions, the Central region had the highest BRT R^2 values for two of the three algal metrics, the Eastern region for the other metric. Explanatory variables in order of importance in the models for each BRT algal model are provided in Table 6, all final BRT models had between 3 and 6 explanatory variables. There was a high degree of similarity across algal response metrics and scales in the final explanatory variables retained in each model. Similar to the invertebrate models, TN, TP, or some measure of riparian land use were selected as the most important variable in almost all models. Either one or both of TN, TP concentration were retained in every model except one, while all but three models had either Rip_RowCrop or AG Rip_Buf (Table 6). There was such a dramatic gradient in nutrient enrichment in the Central region that for two of the models one variable, TP concentration, explained more than 80% of the model variation. Some measure of substrate condition (Ave. Silt and % Fines) and habitat characteristic (% Runs, Habitat Cover and Ave. Velocity) were also common explanatory variables selected in the final BRT models. Other measures of water quality other than TN and TP concentration, DO concentration, pH and temperature, were also retained in a few models. Unlike the invertebrate models, mean annual air temperature was retained in only one out of the three national models and not in any of the regional models. AG-Index was retained in two models and % agriculture in the watershed in only one model.

The partial dependency plots show the response form of individual explanatory variables with the effect of other variables in the model accounted for; though these plots could be produced for each BRT model, for brevity only two examples are shown, one each for an invertebrate and algal metric (Figs. 2, 3, respectively) (partial dependency plots for all of the other invertebrate and algal metrics are available as supplemental material, see link on title page). Partial dependency plots are shown for the four variables retained in the Central BRT model for EPT richness (R^2 of 0.84): TP concentration, % riparian row crops, TN concentration and % fines. These variables explain 37, 23, 20, and 19% of the variation in the full model,

Author's personal copy

Table 6 Comparison of explanatory variables for boosted regression trees (BRT) models for three algal metrics at two landscape spatial scales (national and regional)

Algae metrics	National		Western		Central		Eastern	
% Low inorganic N diatoms	Rip RowCrop	37	AG-index	39	TP conc.	83	Rip RowCrop	64
	TP Conc.	25	TP conc.	27	TN conc.	10	pH	14
	MeanTemp.	16	TN conc.	15	DO conc.	7	% AG	13
	pH	12	% Runs	11			Habitat Cover	10
	TN Conc.	10	AG Rip_Buf	8				
% Eutrophic diatoms	Ave. Silt	47	TP conc.	40	TP conc.	87	AG Rip_Buf	46
	TP Conc.	30	pH	18	% Fines	5	pH	24
	AG Rip_Buf	13	Habitat cover	17	TN conc.	4	TN conc.	13
	% Fines	10	Ave. velocity	13	AG Rip_Buf	4	Habitat cover	9
			TN conc.	12			TP conc.	7
Observed/expected taxa (O/E)	Rip RowCrop	37	TP conc.	29	Rip RowCrop	53	AG Rip_Buf	38
	AG Rip_Buf	27	DO conc.	17	TP conc.	22	Rip RowCrop	34
	TP Conc.	9	Ave. silt	15	AG Rip_Buf	19	TN conc.	28
	Ave. Velocity	11	TN Input	14	% Fines	7		
	TN Conc.	13	Habitat cover	13				
			TN conc.	11				

Variables are listed in order of importance in each model with associated model importance values. See Table 1 for variable definitions

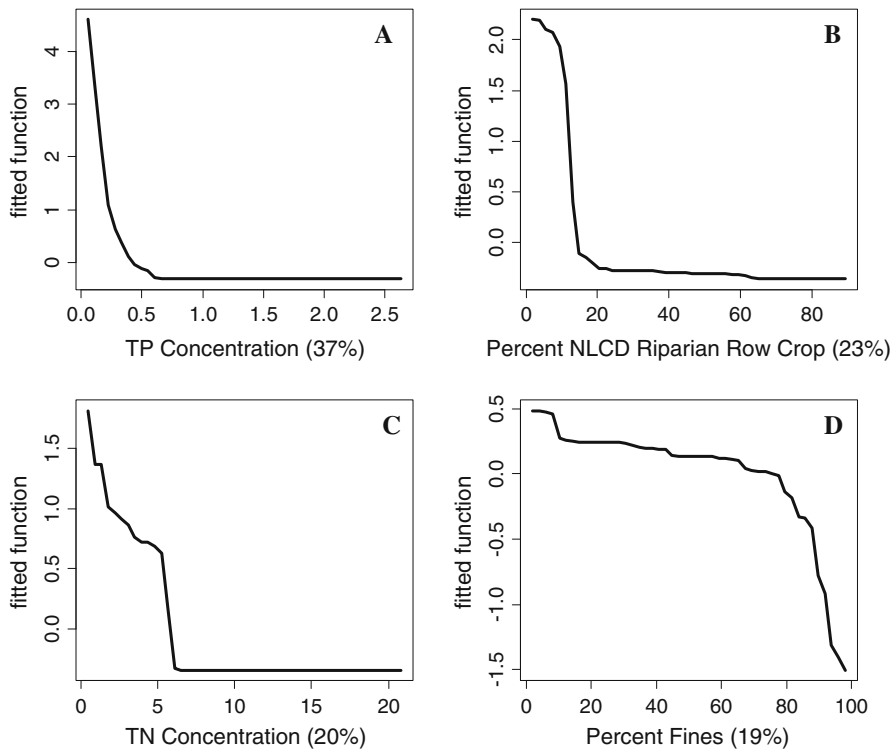


Fig. 2 Partial dependency plots for TP concentration (mg/l) (**A**), percent NLCD row crops in riparian (**B**), TN concentration (mg/l) (**C**), and percent fines in substrate (**D**) in the boosted regression model developed for EPT richness in central region. The *y*-axis fitted function represents the effect of the selected variable on the response variable EPT richness; the relative contribution of each explanatory variable is reported in parentheses. Refer to Table 1 for variable definitions

Author's personal copy

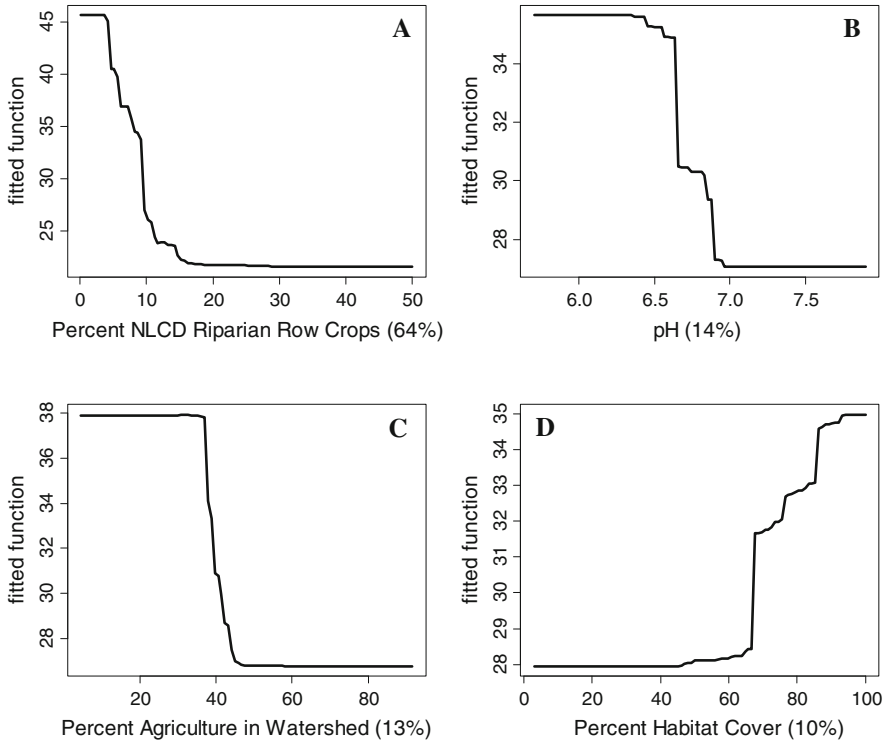


Fig. 3 Partial dependency plots for percent NLCD riparian row crops (**A**), stream pH (**B**), percent agriculture in watershed (NLCD) (**C**), and percent instream habitat cover (**D**) in the boosted regression model developed for the algal metric % Low inorganic N diatoms in east region. The *y*-axis fitted function represents the effect of the selected variable on the response variable % euthrophic; the relative contribution of each explanatory variable is reported in parentheses. Refer to Table 1 for variable definitions

respectively. The response form of the first three variables shows a threshold-type response, in general EPT richness starts high then drops off dramatically at low increases in these variables then the response of EPT flattens out for further increases in each variable (Fig. 2A–C). The response for percent fines shows a decrease in EPT richness from 0 to 10% then only slow average change until approximately 80% fines followed by a sharp decline (Fig. 2D).

The partial dependency plots for the Eastern BRT model for % low inorganic N diatoms (R^2 of 0.84) show a similar threshold type response for % riparian row crops as seen in the invertebrates, values decline rapidly at about 5% row crops in the riparian buffer (Fig. 3A). The % riparian row crop was the dominant variable in the model, explaining 64% of the variation; pH, percent agriculture in the watershed and percent habitat cover explained 14, 13, and 10% of the variation, respectively. The water quality variable pH shows a split step function response, suggesting that there may be two groups of sites in this region for pH (Fig. 3B). Percent agriculture in the watershed shows a rapid decline at about 40% agriculture (Fig. 3C). The last variable in the model, % instream habitat cover, shows almost no response until habitat cover increases above 65% where there is a rapid increase in low inorganic N diatoms with some variability until 100% cover (Fig. 3D).

Discussion

Comparison of invertebrate and algal indicators

In general, we would expect algal metrics to be more directly affected by nutrients than invertebrates or fish; however, the literature reveals a wide variation in the strength of the association of algae to nutrients and agricultural disturbance. Waite (2013) found a selected group of algae and macroinvertebrates metrics that were good indicators of agricultural disturbance but the strength of the correlation varied regionally and by biological assemblage. Individual biological metrics related to agricultural indicators of disturbance showed data analyzed at the regional scale had much higher correlations than at the national scale for both assemblages (Waite, 2013). Danielson et al. (2011) showed that locally developed algal metrics (Maine) were better correlated with disturbance than metrics from surrounding regions. In this study,

Author's personal copy

correlations above 0.77 to TP and 0.67 to TN were found for both algal and invertebrate metrics in select regions (Waite, 2013). Nutrient biotic indices for New York State (Smith et al., 2007) that were developed by summing nutrient tolerance scores for each invertebrate taxa for each sample only had correlations of 0.68 and 0.57 to TP and nitrate (NO3−), respectively, even though the original tolerance values were based on these nutrients. Porter et al. (2008) evaluated a large number of algal metrics and found the highest correlation to TN or TP was 0.57 and 0.37 for % watershed agriculture for % heterotrophic diatoms. Like this study, they found that total nutrients had higher correlations than dissolved nutrients (data not shown) and found higher correlations of algal metrics to nutrients in the western and central regions than eastern regions. One potential reason is that other studies have shown that pH is often a deciding factor for algal assemblages in the eastern regions and not so in the western and central regions (Stevenson et al., 2006). Black et al. (2010) also found poor correlation of TN to algal metrics across the western region; however, they did not assess the central and eastern regions. On the other hand, Pan et al. (2004) found low correlations of algal assemblages to % agriculture evaluated at the reach, network, and catchment scales. The variation in the response of the different biological assemblages to disturbance and across spatial scales suggests that when possible, it is important to evaluate more than one indicator in each bioassessment region (Hughes & Peck, 2008; Brown et al., 2009; Waite, 2013).

Comparison of models

Though BRT models have only recently begun to be applied in ecology, a number of researchers have shown the strength and promise of this modeling technique compared to other methods (De'ath, 2007; Elith et al., 2008; Aertsena et al., 2010; Leclere et al., 2011; Clapcott et al., 2012; Waite et al., 2012). Assessing a site index of three species of trees, Aertsena et al. (2010) compared five modeling techniques and found that generalized additive models (GAM) outperformed MLR and other techniques when user-friendliness and other qualitative measures were added as part of the evaluation process, with BRT models a close second. In their study, artificial neural network (ANN) models had good performance but were penalized for being complex, non-transparent, and having a large training effort. Modeling juvenile fish in a large river system, Leclere et al. (2011) found that BRT outperformed CART and generalized linear models (GLM) in prediction of the presence/absence of the fish species. Waite et al. (2012) compared MLR, CART, RF, and BRT models for predicting invertebrate metrics in small streams as an indicator of watershed disturbance; they showed that BRT models outperformed the other modeling techniques in three separate regions. Though the data are not shown, BRT models from this study explained higher amount of variation in every case compared to random forest or MLR models.

Comparison of models across scale

Ecologists generally believe that spatial scale is important in understanding ecosystem processes and species distributions and that scale should be taken into consideration in ecological research (Levin, 1992; Cooper et al., 1998; Allan, 2004; Waite et al., 2010). In contrast, managers generally desire models and criteria for as large a region as possible, and therefore, these disparate goals may be potentially in conflict with each other related to bioassessment modeling development and application. Cuffney et al. (2011) showed that responses of algae and invertebrate metrics to urbanization varied by biota, geographic region, and the antecedent land use in each region. Stevenson et al. (2006) modeling algal biomass found positive correlations with nutrients in two distinct regions, yet the amount of variation explained varied with nutrient concentration and region. Potapova & Charles (2007) suggested that algal indicators and optima were improved if developed regionally rather than nationally. Results reported here show that the strength of models varied across spatial landscape scale and biological assemblage. With some exception, models developed at the national scale were (high R^2 and low RMSE) comparable to regional models for both invertebrate and algal metrics. However, the national models tended to select more coarse scale explanatory variables (i.e., mean annual air temperature) whereas regional models selected more specific water quality, habitat, and riparian variables for both algae and invertebrates. The national models show that the effects of land use, nutrients, fine sediments, and riparian disturbance on biological indicators can be

Author's personal copy

detected and modeled; however, natural environmental setting variables are also usually important. For example, the variable "mean annual air temperature" was an important variable in many of the national models likely accounting for natural regional variability, suggesting the importance of accounting for natural regional variation when conducting bioassessments. These natural environmental setting variables are commonly captured in biome and ecoregion landscape classifications. Similar to the results from Waite (2013) that showed the correlation strength of indicator metrics to agricultural disturbance varied by biota, metric, and across regions so did the performance of models developed in this study. Strong models were developed for the three invertebrate metrics in the Central region and two of the three metrics in the Western region (R^2 for these BRT models ranged from 0.73 to 0.88). Though still relatively strong, the BRT invertebrate models in the Eastern region were a bit lower in performance (R^2 for Eastern BRT models 0.63–0.70) than most models from the other regions. One reason for this may be the lack of good reference sites in the eastern region, thus, decreasing the disturbance gradient and therefore, the statistical signal, this combined with small sample size. BRT models developed for the three algal metrics had R^2 from 0.70 to 0.89 for all models across all regions and nationally. In general, these results suggest that regional models allow for insight and interpretation at a relatively broad scale yet still with a high degree of specificity.

Comparison of explanatory variables across scales

A long list of explanatory variables was initially entered in the beginning of BRT model development. This included watershed scaled variables such as mean annual air temperature, mean annual precipitation, soil infiltration rate, percent agriculture, and TN and TP inputs to the watershed. The list also included riparian and reach-scale variables such as percent agriculture or row crops in the riparian, average silt, percent fines, average instream velocity, percent habitat cover, percent runs, percent canopy cover, pH, water temperature, DO, and TN and TP instream concentration. Given that each final model had a diverse array of potential explanatory variables covering a variety of scales and processes, it is noteworthy which explanatory variables were selected and not selected in the final models. With minor exception, watershed scaled variables commonly seen as important in previous research (i.e., percent agricultural land use and nutrient inputs in the watershed) were not important in most models for either algae or invertebrates. One exception was nutrient input to the watershed; TN input was selected in three regional models out of a possible 24 models (national and regional). On the other hand, instream nutrient concentration (either TN or TP or both) and variables assessing the amount of disturbance to the riparian zone (percent row crops based on NLCD or percent agriculture based on orthophotos) were selected as the most important explanatory variable in most of the BRT models regardless of landscape scale or assemblage. Some measure of habitat condition (habitat cover, % runs, velocity, % fines, and average silt) was also commonly selected in the final invertebrate models, though the specific habitat related variable(s) varied across regions. For example, percent fines and silt were not commonly selected in the Western region, but were common in the Central and Eastern regions. On the other hand, velocity and habitat cover were commonly selected in the Western and Eastern regions, yet rarely selected in the Central region. These selections seem to following the general geology and geomorphology differences across regions. The Western region generally has more high gradient streams with more riffles and higher velocity and less silts and fine substrate, while the Central region has lower gradient systems with lower velocity and higher amounts of silts and fine substrate. Agricultural land use is likely to add sediment to these systems, yet the naturally higher gradient systems may be able to flush sediments easier than the low gradient systems. The surprise related to this pattern was the Eastern region, which is also a low gradient system, yet in two out of the three invertebrate models, average velocity was selected. Riseng et al. (2011) looking at the causal pathway of effects of agricultural land use found that the variation explained was equally split between effects of habitat and water quality on invertebrate metrics in the central region. Recently, researchers have shown that human disturbance variables (i.e., land use and water quality) in combination with environmental setting variables such as climate, stream type, or elevation help explain the greatest amount of variation among biotic indicators (Maloney et al., 2009; Waite et al., 2010; Clapcott et al., 2012; Leclere et al., 2011). Thus, to increase our understanding and applicability to large

Author's personal copy

spatial areas, it may be important to take into account factors at multiple scales and a combination of processes including landscape and riparian disturbance, water quality, habitat, and natural variation.

Another benefit of RF and BRT models is the ability to see the response form of the individual explanatory variables retained in the models via partial dependency plots (De'ath, 2007; Elith et al., 2008; Aertsen et al., 2010; Clapcott et al., 2012). Though it wasn't possible to show all the partial dependency plots for all the algal and invertebrate models developed in this study in this manuscript, a general response pattern was seen throughout the models as shown in Figs. 2 and 3 (for the remaining dependency plots see link to supplemental information). TN and TP instream concentration and riparian disturbance variables frequently followed a threshold type response, where dramatic changes occurred at relatively low values followed by little further response. Though further analysis is needed to determine whether these are actual thresholds (Qian & Cuffney, 2012), the response form shown in the partial dependency plots from this study is suggestive of such a response. The thresholds seem to occur at values of approximately 0.01–0.02 mg/l and 0.5–1.0 mg/l for TP and TN concentration, respectively, though declines of biological condition start immediately with no apparent initial resistance. Stevenson et al. (2006) reported that benthic algae had a TP threshold of ca. 0.03 mg/l and a TN threshold of ca. 0.5 mg/l which is relatively close to the thresholds seen in this study. Thresholds for riparian disturbance (% row crop and % riparian agriculture orthophotos) occurred at approximately 5–10% disturbance with perhaps only a minor amount of initial resistance from 0 to 5% riparian land use. Clapcott et al. (2012) using BRT models for invertebrate metrics in New Zealand found no apparent threshold in the amount of indigenous vegetation removal in the catchment which followed a linear decreasing response.

Conclusions

For all biotic metrics and regions modeled, BRT models outperformed MLR models, frequently by a large amount. With minor exceptions, watershed scaled variables commonly seen as important in previous research (i.e., percent agricultural land use and nutrient inputs in the watershed), were not important in most models for either algae or invertebrates. On the other hand, instream nutrient concentration (either TN or TP or both) and variables assessing the amount of disturbance to the riparian zone (percent row crops based on NLCD or percent agriculture based on orthophotos) were selected as the most important explanatory variable in most models regardless of landscape scale or assemblage. Some measure of habitat condition (habitat cover, % runs, velocity, % fines, and average silt) was also commonly selected in the final invertebrate models, though the specific habitat related variable(s) varied across regions. Overall, the national models show that the effects of land use, nutrients, fine sediments, and riparian disturbance on biological indicators can be successfully detected and modeled; however, natural environmental setting variables are also usually important at these large scales. These large scale natural environmental settings variables are usually captured in biome and ecoregion landscape classifications. To increase our understanding and applicability to large spatial areas, it may be important to take into account factors at multiple scales and processes including landscape and riparian disturbance, water quality, habitat, and natural variation. In general, these results suggest national models tend to emphasize more general landscape differences and that regional models allow for insight and interpretation at a relatively broad scale yet still with a high degree of specificity. Frequently, TN and TP concentration and riparian agricultural land use variables showed a threshold type response to the biotic metrics modeled at relatively low values of the explanatory variables.

Acknowledgments I thank each of the many biologists and hydrologists with the NAWQA program who collected all the data used in this report and provided insights to the particularities of the data. I thank Mark Munn for giving me the funding to work on this project and who provided valuable advice and review. Dr. Yangdong Pan and Chris Parker (Portland State University) for advice on R programming and statistics. Finally, I thank one USGS reviewer and one outside reviewer for their review of the manuscript prior to submission to the journal.

References

Aertsena, W., V. Kinta, J. Van Orshovena, K. Özkanb & B. Muysa, 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. Ecological Modelling 221: 1119–1130.

Author's personal copy

Allan, J. D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annals and Review of Ecology, Evolution and Systematics 35: 257–284.

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

Biggs, B. J. F. & C. Kilroy, 1994. Stream Periphyton Monitoring Manual. The New Zealand Ministry for the Environment, Christchurch.

Black, R. W., P. W. Moran & J. D. Frankforter, 2010. Response of algal metrics to nutrients and physical factors and identification of nutrient thresholds in agricultural streams. Environmental Monitoring and Assessment 175: 397–417.

Breiman, L., J. H. Friedman, R. A. Olshen & C. J. Stone, 1984. Classification and Regression Trees. Wadsworth and Brooks/Cole, Monterey, California, USA.

Brown, L. R., T. F. Cuffney, J. F. Coles, F. Fitzpatrick, G. McMahon, J. Steuer, A. H. Bell & J. T. May, 2009. Urban streams across the USA: lessons learned from studies in 9 metropolitan areas. Journal of North American Benthological Society 28: 1051–1069.

Charles, D. F., C. Knowles & R. S. Davis, 2002. Protocols for the Analysis of Algal Samples Collected as Part of the U.S. Geological Survey National Water-Quality Assessment Program. Report No. 02-06. Patrick Center for Environmental Research, The Academy of Natural Sciences, Philadelphia.

Clapcott, J. E., K. J. Collier, R. G. Death, E. O. Goodwin, J. S. Harding, D. Kelly, J. R. Leathwick & R. G. Young, 2012. Quantifying relationships between land-use gradients and structural and functional indicators of stream ecological integrity. Freshwater Biology 57: 74–90.

Clarke, K. R. & R. N. Gorley, 2006. PRIMER v6: Users Manual/Tutorial PRIMER-E: Plymouth, England.

Cooper, S. D., S. Diehl, K. Kratz & O. Sarnelle, 1998. Implications of scale for patterns and processes in stream ecology. Australian Journal of Ecology 23: 27–40.

Cuffney, T. F., 2003. User's Manual for the National Water-Quality Assessment Program Invertebrate Data Analysis System (IDAS) Software: Version 3. U.S. Geological Survey Open-File Report 03-172, Raleigh, NC. Available at: ftp://ftpext.usgs.gov/pub/er/nc/raleigh/tfc/IDAS/.

Cuffney, T. F., R. A. Brightbill, J. T. May & I. R. Waite, 2010. Responses of benthic macroinvertebrates to environmental changes associated with urbanization in nine metropolitan areas of the conterminous United States. Ecological Applications 20: 1384–1401.

Cuffney, T. F., R. Kashuba, S. S. Qian, I. Alameddine, Y. Cha & B. Lee, 2011. Multilevel regression models describing regional patterns of invertebrate and algal responses to urbanization across the USA. Journal of North American Benthological Society 30: 797–819.

Cushing, C. E. & J. D. Allan, 2001. Streams: Their Ecology and Life. Academic Press, San Diego, California, USA.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson & J. J. Lawler, 2007. Random forests for classification in ecology. Ecology 88: 2783–2792.

Danielson, T. J., C. S. Loftin, L. Tsomides, J. L. DiFranco & B. Connors, 2011. Algal bioassessment metrics for wadeable streams and rivers of Maine, USA. Journal of North American Benthological Society 30: 1033–1048.

De'ath, G., 2007. Boosted trees for ecological modeling and prediction. Ecology 88: 243–251.

De'ath, G. & K. E. Fabricius, 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81: 3178–3192.

Dubrovsky, N. M., K. R. Burow, G. M. Clark, J. M. Gronberg, P. A. Hamilton, K. J. Hitt, D. K. Mueller, M. D. Munn, B. T. Nolan, L. J. Puckett, M. G. Rupert, T. M. Short, N. E. Spahr, L. A. Sprague & W. G. Wilber, 2010. The Quality of Our Nation's Waters – Nutrients in the Nation's Streams and Groundwater, 1992–2004. U.S. Geological Survey Circular 1350.

Elith, J., J. R. Leathwick & T. Hastie, 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802–813.

Fishman, M. J., 1993. Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory; Determination of Inorganic and Organic Constituents in Water and Fluvial Sediments. Open File Report 93-125. USGS, Denver, CO, USA.

Fitzpatrick, F. A., I. R. Waite, P. J. D'Arconte, M. R. Meador, M. A. Maupin & M. E. Gurtz, 1998. Revised Methods for Characterizing Stream Habitat in the National Water Quality Assessment Program. Water-Resources Investigation Report 98-4052. USGS, Raleigh, NC, USA.

Friberg, N., J. Skriver, S. E. Larsen, M. L. Pedersen & A. Buffagni, 2010. Stream macroinvertebrate occurrence along gradients in organic pollution and eutrophication. Journal of Freshwater Biology 55: 1405–1419.

Herlihy, A. T. & J. C. Sifneos, 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. Journal of North American Benthological Society 27: 932–948.

Hughes, R. M. & D. V. Peck, 2008. Acquiring data for large aquatic resource surveys: the art of compromise among science, logistics, and reality. Journal of North American Benthological Society 27: 837–859.

Johnson, M. R. & R. B. Zelt, 2005. Protocols for Mapping and Characterizing Land Use/Land Cover in Riparian Zones. Open-File Report 2005-1302. USGS, Reston, VA, USA.

Larned, S. T., 2010. A prospectus for periphyton: recent and future ecological research. Journal of North American Benthological Society 29: 182–206.

Leclere, J., T. Oberdorff, J. Belliard & F. Leprieur, 2011. A comparison of modeling techniques to predict juvenile 0+ fish species occurrences in a large river system. Ecological Informatics 6: 276–285.

Levin, S., 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur Award Lecture. Ecology 73: 1943–1967.

Lücke, J. D. & R. K. Johnson, 2009. Detection of ecological change in stream macroinvertebrate assemblages using single metric, mulitmetric or multivariate approaches. Ecological Indicators 9: 659–669.

Maloney, K. O., D. E. Weller, M. J. Russell & T. Hothorn, 2009. Classifying the biological condition of small streams: an

Author's personal copy

example using benthic macroinvertebrates. Journal of North American Benthological Society 28: 869–884.

Moulton, S. R., J. Kennen, R. M. Goldstein & J. A. Hambrook, 2002. Revised Protocols for Sampling Algal, Invertebrate, and Fish Communities as Part of the National Water-Quality Assessment Program. Open-File Report 02-150. USGS, Reston, VA, USA.

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

Ostermiller, J. D. & C. P. Hawkins, 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of North American Benthological Society 23: 363–382.

Pan, Y., A. Herlihy, P. Kaufmann, J. Wigington, J. van Sickle & T. Moser. 2004. Linkages among land-use, water quality, physical habitat conditions and lotic diatom assemblages: A multi-spatial scale assessment. Hydrobiologia 515: 59–73.

Patton, C. J. & E. J. Gilroy, 1998. U.S. Geological Survey Nutrient Preservation Experiment – Experimental Design, Statistical Analysis, and Interpretation of Analytical Results. US Geological Survey Water-Resources Investigations Report 98-4118.

Patton, C. J. & J. R. Kryskalla, 2003. Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory: Evaluation of Alkaline Persulfate Digestion as an Alternative to Kjeldahl Digestion for determination of Total and Dissolved Nitrogen and Phosphorus in Water. Water-Resources Investigations Report 2003-4174. USGS, Denver, CO, USA.

Porter, S. D., D. K. Mueller, N. E. Spahr, M. D. Munn & N. M. Dubrovsky, 2008. Efficacy of algal metrics for assessing nutrient and organic enrichment in flowing waters. Journal of Freshwater Biology 53: 1036–1054.

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

Prasad, A. M., L. R. Iverson & A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9: 181–199.

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

R Development Core Team, 2006. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN3-900051-07-0. Retrieved from http://www.R-project.org.

Rantz, S. E., 1982. Measurement and Computation of Streamflow: Volume 1. Measurement of Stage and Discharge. US Geological Survey Water-Supply Paper, 2175.

Riseng, C. M., M. J. Wiley & R. J. Stevenson, 2004. Hydrologic disturbance and nutrient effects on benthic community structure in Midwestern US streams: a covariance structure analysis. Journal of North American Benthological Society 23: 309–326.

Riseng, C. M., M. J. Wiley, R. W. Black & M. D. Munn, 2011. Impacts of agricultural land use on biological integrity: a causal analysis. Ecological Applications 21: 3128–3146.

Ruddy, B. C., D. L. Lorenz & D. K. Mueller, 2006. County Level Estimates of Nutrient Inputs to the Land Surface of the Conterminous United States, 1982–2001. Scientific Investigations Report 2006-5012. USGS, Reston, VA, USA.

Smith, A. J., R. W. Bode & G. S. Kleppel, 2007. A nutrient biotic index (NBI) for use with benthic macroinvertebrate communities. Ecological Indicators 7: 371–386.

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

Stevenson, R. J., Y. Pan, K. M. Manoylov, C. A. Parker, D. P. Larsen & A. T. Herlihy, 2008. Development of diatom indicators of ecological conditions for streams of the western US. Journal of North American Benthological Society 27: 1000–1016.

Thorton, P. E. & S. W. Running, 1999. An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. Agricultural and Forest Meteorology 93: 211–228.

Turak, E., S. Ferrier, T. Barrett, E. Mesley, M. Drielsma, G. Manion, G. Doyle, J. Stein & G. Gordon, 2010. Planning for the persistence of river biodiversity: exploring alternative futures using process-based models. Journal of Freshwater Biology 56: 39–56.

University of Montana, 2005. Daymet, Daily Surface Weather and Climatological Summaries: Numerical Terradynamic Simulation Group, Daymet database, accessed August 2005, at http://www.daymet.org

US Geological Survey, 2003a. Coverages for 1:100 k stream tiles, assembled from National Hydrography Dataset. Reston: US Geological Survey (Retrieved 9 September 2009, from http://nhd.usgs.gov/index.html).

US Geological Survey, 2003b. National Land Cover Database (NLCD2001) Metadata for Mapping Zones by Multilayer Downloadable File. US Geological Survey (Retrieved 9 September 2009, from http://www.mrlc.gov/nlcd_multizone_map.php) [NLCD01 accessed October 2013 at http://www.mrlc.gov/nlcd01_data.php].

US Geological Survey, 2006. National Field Manual for the Collection of Water-Quality Data. Handbooks for Water-Resource Investigations, US Geological Survey Techniques of Water-Resources Investigations.

van Dam, H., A. Mertens & J. Sinkeldam, 1994. A coded check list and ecological indicator values of freshwater diatoms from the Netherlands. Netherlands Journal of Aquatic Ecology 28: 117–133.

Van Sickle, J., C. P. Hawkins, D. P. Larsen & A. T. Herlihy, 2005. A null model for the expected macroinvertebrate assemblage in streams. Journal of North American Benthological Society 24: 178–191.

Verhoeven, J. T. A., B. Arheimer, C. Yin & M. M. Hefting, 2006. Regional and global concerns over wetlands and water quality. Trends in Ecology and Evolution 21: 96–203.

Waite, I. R., 2013. Applications of an agricultural intensity index to invertebrate and algal metrics from streams at two scales within the U.S. Journal of the American Water Resources Association April 2013.

Waite, I. R., A. T. Herlihy, D. P. Larsen, N. S. Urquhart & D. J. Klemm, 2004. The effects of macroinvertebrate

Author's personal copy

taxonomic resolution in large landscape bioassessments: an example from the Mid-Atlantic Highlands, U.S.A. Journal of Freshwater Biology 49: 474–489.

Waite, I. R., L. R. Brown, J. G. Kennen, J. T. May, T. F. Cuffney, J. L. Orlando & K. A. Jones, 2010. Comparison of watershed disturbance predictive models for stream benthic macroinvertebrates for three distinct ecoregions in western US. Ecological Indicators 10: 1125–1136.

Waite, I. R., J. G. Kennen, J. T. May, L. R. Brown, T. F. Cuffney, K. A. Jones & J. L. Orlando, 2012. Comparison of stream invertebrate response models for bioassessment metrics. Journal of the American Water Resources Association 48: 570–583.

Wang, L., D. M. Robertson & P. J. Garrison, 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implication to nutrient criteria development. Journal of Environmental Management 39: 194–212.

Wisseman, B., 1996. Version 1.0. Unpublished Ecological Coding Attributes for Freshwater Invertebrate Taxa in Western North America. Available from Aquatic Biology Associates, Inc., Corvallis, OR.

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

Zelt, R. B. & M. D. Munn, 2009. Riparian and Associate Habitat Characteristics Related to Nutrient Concentrations and Biological Responses of Small Streams in Selected Agricultural Areas, United States, 2003-04. Scientific Investigations Report 2009-5224.