SWAMP
Surface Water Ambient Monitoring Program

CALIFORNIA NPS PROGRAM

Technical Report | 2008

An Index of Biotic Integrity (IBI) for Perennial Streams in California's Central Valley

December 2008

CALIFORNIA Water Boards

www.waterboards.ca.gov/swamp

An Index of Biotic Integrity (IBI) for Perennial Streams in California's Central Valley

Andrew C. Rehn[1], Jason T. May[2] and Peter R. Ode[1]

[1]California Aquatic Bioassessment Lab
Dept. of Fish and Game
2005 Nimbus Road
Rancho Cordova, CA 95670

[2]U.S. Geological Survey
Water Resources Division
Placer Hall, 6000 J Street,
Sacramento, CA 95819

Executive Summary

Bioassessment is the science of using aquatic organisms as indicators of ecological condition in streams in rivers. Many types of organisms can be used as indicators, for example fish or algae, but bioassessment is most frequently based on benthic macroinvertebrates (BMIs), which are small but visible bottom-dwelling organisms such as insects. BMI data sets typically consist of long lists of species (or taxa) found in a sample and their relative abundances. These data can be simplified into measures of biological condition such as indices of biotic integrity (IBIs) that are designed to be sensitive to human-caused alterations to the landscape, to stream channels and riparian zones, and to water chemistry. IBIs function much like economic indicators: high IBI scores reflect good ecological conditions while low IBI scores reflect poor ecological conditions.

Bioassessment is increasingly used throughout California by water quality monitoring programs, but in the Central Valley bioassessment is more challenging than in other regions of the state because the entire landscape and most streams are highly altered by human activities such as urbanization, agriculture and water diversions. This makes it impossible to evaluate how BMIs respond across a complete gradient of human disturbance within the region, that is, from minimally disturbed reference sites where human activity is absent or minimal and which therefore set the benchmark for biological expectations, to the most altered sites with degraded biology. In the Central Valley, minimally disturbed reference sites are no longer available. Even the ‘least-disturbed’ sites, which represent the best-available chemical, physical and biological habitat conditions given the current state of the landscape, are markedly disturbed. Reference sites in other parts of California, such as the Sierra Nevada or the Sierra foothills, may be significantly less disturbed than Central Valley reference sites.

In this study, BMI data sets from 11 studies conducted at various intervals over the last 14 years were compiled to build an IBI for Central Valley streams. Data were not collected consistently by the different studies, and many gaps were present in associated physical habitat and water chemistry data sets. This could be corrected for BMI samples by standardizing to a consistent level of taxonomic effort. Gaps in other data could not be addressed. Criteria for defining ‘best-available’ reference sites were established as data allowed and were based on local urban and agricultural intensity, stream channel and riparian condition, and stream substrate composition. Eighty BMI metrics were evaluated for inclusion in the IBI based on 4 criteria: 1) sufficient range for scoring; 2) responsiveness to land use and/or local disturbance variables measured at the 150-meter sampling reach (as data allowed); 3) good discrimination between reference and test sites; 4) lack of correlation with other responsive metrics. Five final metrics were selected and scored for inclusion in the IBI: collector richness (number of taxa that are collector-feeders), predator richness (number of taxa that are predators), percent EPT taxa (percent of taxa that are mayflies, stoneflies, or caddisflies), percent clinger taxa (percent of taxa that cling to vegetation) and Shannon diversity (a composite measure of taxonomic richness and evenness of abundance). The final IBI showed good discrimination between reference and test sites, and was validated with an independent data set. BMI metrics and the final IBI were more strongly related to reach-scale physical habitat variables than to water chemistry or land use variables, but detailed water chemistry was lacking for many

sites, and some studies have shown that response to land use diminishes when more than 10% of a watershed is degraded by human activities.

Despite data gaps that were less than ideal for indicator development, this study is the first to set expectations for Central Valley BMI assemblages based on best-available reference sites. The Central Valley IBI can be used as a general interpretive framework for benthic samples collected from perennial streams on the valley floor and provides an objective means for rating biological condition in a region with high urban and agricultural intensity. The ability to rank sampling sites relative to explicitly defined biological expectations is essential to any biological monitoring program. Therefore, this index may prove useful in several monitoring applications, including California's non-point source CMAP program where sampling was stratified to assess and compare stream condition in urban, agricultural and forested watersheds, in stormwater monitoring programs, in point-source pollution investigations, and in stream restoration monitoring. Key recommendations include: 1) that all future bioassessment projects in the Central Valley should collect quantitative physical habitat and water chemistry with consistent protocols at all sites; *in situ* chemistry and rapid (qualitative) physical habitat are not sufficient for screening reference sites or evaluating BMI responses to stressor gradients; 2) that bioassessment should be added to NPS monitoring where programs are already collecting more intensive stressor data such as pesticides, nutrients and metals. This will provide California's monitoring programs with better datasets to support future analyses.

Introduction

California's Central Valley is a semi-arid tectonic basin of approximately 58,000 km^2 with climate, geology, physiography and land use sufficiently distinct to constitute a Level III ecoregion within the United States (Omernik 1987). The ecoregion comprises two major river drainages that converge to form the San Francisco Estuary: the Sacramento River and its tributaries in the north and the San Joaquin River and its tributaries in the south. The smaller Tulare Basin in the extreme south is hydrologically closed. With a Mediterranean climate of pronounced wet winters and dry summers, much of the Central Valley's annual water supply comes from snow melt in surrounding mountain ranges, primarily the Cascades to the north and Sierra Nevada to the east. Large-scale damming of nearly every river in the region to capture, store and divert spring snow melt has transformed the valley floor, with its rich soils and long growing season, into the most productive agricultural region in the United States. Other widespread hydrologic alterations associated with irrigation-subsidized agriculture include an extensive network of canals and aqueducts, channelization of natural stream drainages for flood control and use of natural or modified channels as effluent drains. In addition, the Central Valley's human population is expected to reach 7 million by 2010 (www.greatvalley.org), placing ever increasing demands on regional water resources.

Bioassessment is increasingly utilized in California as a tool for freshwater resource management (e.g., Ode et al. 2005a; Rehn 2008), but is especially challenging in the Central Valley because human activities in the region are so extensive. In bioassessment, water quality and biological conditions at sampling sites are often related to upstream land use, but water removal, subsidization, inter-basin transfer and extensive artificial channels often make the concept of a watershed difficult to apply on the valley floor. Transformation of the valley floor into farm fields, orchards and urban sprawl is nearly complete, thereby reducing the utility of conventional land use measures like "percent of watershed in agriculture" because complete gradients of disturbance no longer exist at the landscape scale. By contrast, gradients in chemical and physical conditions such as nutrient concentrations and riparian structure do still exist at a localized scale, making quantitative characterization of these variables especially critical in regional stream surveys.

Despite the challenges, several agencies have collected benthic macroinvertebrate (BMI) samples from streams, sloughs and canals on the valley floor during the last several years. However, an interpretive framework for resulting data sets is lacking because expectations for valley floor BMI assemblages have never been defined in the context of regional reference conditions (Hall et al. 2006). Ode et al. (2005) identified a pool of least disturbed reference sites in the Sacramento Valley but emphasized methods for reference site selection in highly disturbed regions rather than development of a biological index. Other studies have associated BMI metrics or multivariate ordination axes with local environmental stressor gradients (e.g., Brown and May 2000a, b; Griffith et al. 2003), but no attempt was made to set expectations for BMI assemblages or to include least disturbed sites. Such studies are invaluable for understanding how different human stressor gradients influence the composition and structure of BMI assemblages across sites, but they do not allow assessment of biological condition at a particular site with respect to least-disturbed or best-attainable sites for the region. In this study, existing BMI data sets were compiled to: 1) determine if they sufficiently capture the range of biological conditions across Central Valley streams to develop a regional index of biotic integrity (IBI); 2) if data were found to be sufficient, develop an IBI; 3) evaluate relationships between IBI scores and component metrics with human stressor gradients across a broader range of sites than was available in any previous single data set.

Methods

Data sets

Data were compiled from 11 different bioassessment projects conducted at various intervals over the last 14 years (Table 1). Some projects focused exclusively on Central Valley streams and/or canals, whereas other projects were of broader scope but included sampling sites on the valley floor (elevation < 85 m). BMI sampling protocols, physical habitat (PHAB) measurements and water chemistry measurements were not consistent among projects (Table 1; Appendix 1). All projects sampled BMIs with kick-net protocols, but the type of habitat sampled, the total area sampled, and the number of BMIs counted in the laboratory varied. To facilitate data comparability, BMI taxa lists from the various projects were standardized at "Level I" taxonomic effort as defined by the Southwestern Association of Freshwater Invertebrate Taxonomists (www.swrcb.ca.gov/swamp/docs/safit/ste_list.pdf). Nine-hundred count samples were standardized at 500-count using randomized subsampling without replacement. Samples with < 300 organisms were omitted from analyses.

Compiled data sets included much redundancy for several reasons: 1) many projects sampled sites repeatedly over multiple seasons and/or years; 2) different projects frequently sampled the same sites or sites very close together; 3) some projects sampled consecutive sites, often separated by only 1 km, along a single stream. From an initial total of 740 benthic samples from 314 sites, 168 samples from 141 sites were selected for construction and testing of the IBI (Fig. 1, Appendix 2). A development set comprising 112 sites was used for screening metrics (see below). Development sites were selected to minimize data redundancy; multiple (or at least adjacent) sites on a single stream and repeat samples over multiple seasons or years were avoided. Sites with the most comprehensive PHAB and water chemistry data were given preference for the development set. Preference also was given to late summer/fall samples since most projects sampled in the fall. A validation set comprising 56 samples not used in metric screening was used as an independent test of whether the final index discriminated between reference and test sites. Some validation samples were from the same sites used in the development set, but represented different sampling events. Ideally, a validation set would not contain development sites, but data redundancy and a paucity of reference sites (see below) precluded complete independence.

Reference sites

Reference sites represent biological condition when human disturbance is absent or minimal (Hughes 1995; Bailey et al. 2005; Stoddard et al 2005). In regions like the Central Valley, where essentially all watersheds have been altered and the landscape has been greatly transformed, minimally disturbed reference sites are no longer available and even least-disturbed sites are markedly disturbed. It is important for readers of this report to know that by 'reference site' we mean 'best-available site' given the extent of regional landscape modification. Reference sites in other parts of California, such as the Sierra Nevada or the north coast, may be significantly less disturbed than Central Valley reference sites.

Ideally, candidate reference sites are required to pass a series of chemical and physical criteria to be considered least-disturbed. However, inconsistency in PHAB, water chemistry and land use data across projects precluded application of consistent reference site screening criteria. For the most part, sites were screened as data allowed. To facilitate consistency in land use screens across sites, a qualitative ranking of urban and agricultural intensity within 1km upstream (or within a 1km radius around canals with no clearly definable watershed) was carried out using visual assessment of National Landcover Database (NLCD; 2001) GIS layers. Sites were ranked as having either 'low', 'moderate' or 'high' intensity of urban and agriculture: sites ranked as 'high' for either land use category were omitted from the reference pool. EMAP and CMAP sites had the most inclusive set of chemical and physical data associated with benthic samples and were screened using thresholds outlined in Stoddard et al. (2005). Ode et al. (2005b) used a mix of qualitative and quantitative PHAB data to evaluate candidate Sacramento Valley reference sites but did not list explicit thresholds for inclusion in the reference pool. Sites from Ode et al. (2005b) were screened using the following criteria: total qualitative PHAB score ≥125, sand and fine substrate ≤ 50%, riparian disturbance index (W1_HALL from Kaufmann et al. [1999]) < 0.5. Most other data sets had only qualitative PHAB data available. For those projects, sites with total qualitative PHAB score ≥150 and visually estimated fine substrate < 50% were considered reference. Whenever possible, additional chemistry data, land use data or qualitative notes about site condition were used to evaluate sites for inclusion in the reference pool.

Metrics screening and IBI evaluation

Eighty BMI metrics (Appendix 3) were evaluated for use in a Central Valley biotic index based on four criteria: 1) sufficient range for scoring; 2) responsiveness to land use and reach-scale disturbance variables (as data allowed); 3) good discrimination between reference and test sites; 4) lack of correlation with other responsive metrics. Richness metrics with range < 10 and percentage metrics with a value of '0' at ≥ 70% of development sites were excluded. Responsiveness was assessed using visual inspection of biotic metric vs. stressor gradient scatter plots and linear regression coefficients. Urban and agricultural intensity were qualitatively ranked based on visual assessment of NLCD (2001) GIS layers at two spatial scales: within 1km upstream (or within a 1km radius around canals with no clearly definable watershed) and within the entire upstream watershed for sites with clearly definable watersheds. Rankings were on a 0 to 10 scale with 10 representing high intensity. Metrics were selected as responsive if they showed either a linear or a wedge-shaped relationship with stressor gradients (note: biological metrics often show a wedge-shaped relationship with single stressor gradients where the upper boundary represents a threshold of biological response; multiple limiting factors may result in lower metric values than expected if response were to the single gradient alone [Blackburn 1992]). Stressor variables reported at < 30 development sites were omitted from metric screening. Correlated stressor variables were not omitted (because of large gaps in the compiled chemical and physical data), and none were selected *a priori* as most appropriate for metric screening. Box-and-whisker plots were used to evaluate BMI metrics for discrimination between reference and test sites. Metrics with non-overlapping quartiles between reference and test sites were considered to show good discrimination. Metrics that passed the range, responsiveness and discrimination tests were tested for redundancy. Pairs of metrics with Pearson correlation coefficients $|r| \geq 0.7$ were considered redundant and the least responsive metric of the pair was eliminated.

Metrics were scored on a 0–10 scale using statistical properties of raw metric values from reference and test sites to define metric ceilings and floors. Any site with a metric value equal to or greater than the 80th percentile of reference sites received a score of 10; any site with a metric value equal to or less than the 20th percentile of test sites received a score of 0. The remaining range of intermediate metric values was divided equally and assigned scores of 1 through 9. An IBI score was calculated for each site by summing the constituent metric scores and adjusting the index to a 100-point scale. To test whether the distribution of IBI scores in reference and test sites might have resulted from chance, score distributions in the development set were compared to those in the validation set.

Seasonal and inter-annual variation

Sites that were sampled in both spring (March-June) and fall (September-November) were used to evaluate if final IBI metrics varied seasonally and would need adjusted scoring scales. Sites that were sampled 3 or more times over different seasons and/or years were used to evaluate variance in IBI scores over time. Coefficient of variation (CV) was plotted as a function of mean IBI score to determine if variability in IBI scores increased or decreased as a function of biological condition. Seasonality and inter-annual variance analyses were not restricted to the development data set, but drew from the initial combined pool of 314 sites.

Results

Most REMAP samples (collected with the reachwide method) were dropped from analyses due to low counts: 56% of REMAP samples had < 100 organisms, and 84% had < 300 organisms. Other projects included here that used the reachwide method did not produce such a high frequency of low count samples. However, a recent study by Mazor et al. (in prep) also found that the EPA's reachwide sampling method frequently produced samples with insufficient counts for calculation of IBI and O/E scores when used in low gradient, highly altered streams in California. Mazor et al. (in prep) found that targeted-habitat sampling methods usually yielded sufficient sample sizes for index calculation when collected from study reaches adjacent to those where low-count reachwide samples were taken.

Sixteen of the 141 sites used for construction and testing of the IBI passed all land use, PHAB and chemistry screens and were considered best-available reference sites for the Central Valley (Appendix 2). Three of the final 16 reference sites had either same-day duplicate samples or repeat visits in different years and were used in both development and validation of the IBI. In some cases (e.g., Butte Creek and Auburn Ravine), sites low in a watershed were considered test sites while sites higher in the same watershed were considered reference sites.

Five of the 80 evaluated metrics were selected as most-responsive to stressor gradients, best-discriminating between reference and test sites and least correlated with each other: collector richness, predator richness, percent EPT taxa, percent clinger taxa and Shannon diversity (Table 2). A relatively large proportion of metrics (19%), mostly involving intolerant taxa and shredder taxa, was rejected because they had a value of '0' at ≥ 70% of development sites (Appendix 3). By contrast, only 3% (2/70) of metrics failed this criterion in the North Coast IBI (Rehn et al. 2005), and only 8% (5/61) failed this criterion in the Southern California IBI (Ode et al. 2005). The lack of intolerant organisms on the Valley floor is expected given the extent of human influence on regional streams, and lack of shredders in lower watersheds is predicted by the river continuum concept (Vannote et al. 1980). Candidate metrics for inclusion in the final IBI responded more strongly to local PHAB variables than to chemistry or land use disturbance variables (Fig. 2, Table 3). However, large data gaps and lack of consistency in physical and chemical data among projects precluded thorough screening of metric response across all sites. In many cases, different subsets of sites were used to screen metrics for responsiveness against different stressors depending on data availability. The general lack of response to water chemistry variables may have been due to lack of chemistry data at most sites or inconsistency in collection methods and data reporting when chemistry data were collected.

The final IBI showed good discrimination between reference and test sites in both development and validation data sets (Fig. 3). Despite the noted limitations in screening reference sites and evaluating metric responsiveness, the characterization of reference conditions and subsequent IBI scoring presented here is repeatable and not likely due to chance. Like most metrics, the IBI showed poor dose-response relationships with urban intensity at both local and watershed scales and with agricultural intensity at the local scale. By contrast, IBI score decreased significantly as percent agriculture in the watershed increased (Fig. 4). The final 5 IBI metrics did not vary between spring and fall samples and did not require seasonal adjustments in scoring (Fig. 5). Variability in IBI score increased as mean IBI score decreased at sites with at least 3 repeat visits (Fig. 6), i.e. sites with higher IBI scores showed less inter-annual variability.

Discussion

Despite data gaps that were less than ideal for indicator development, this study is the first to set expectations for Central Valley BMI assemblages based on best-available reference sites. However, the interpretive framework developed here should be regarded as preliminary and additional questions remain. For example, most sampling sites (and all reference sites) were concentrated in the Sacramento Valley (Fig. 1). It is therefore unclear how accurately the IBI reflects reference conditions in the more arid San Joaquin Valley, although the current IBI probably sets expectations higher than would be possible from analysis of San Joaquin reference sites alone (if they even exist).

A number of studies have been published on Central Valley BMI assemblages and their relation to environmental stressors. For example, Brown and May (2000a) used canonical correspondence analysis (CCA) to evaluate relationships between BMI assemblages from snag samples and environmental gradients in the lower Sacramento and San Joaquin drainages. BMIs were most strongly related to specific conductance and percent of watershed in agricultural + urban land use, although their first CCA axis explained only 10% of taxonomic variance across sites. BMIs also were related to gradient, elevation, dominant substrate type and water temperature in Brown and May (2000a), the latter two variables being potential indicators of human activity in watersheds. Griffith et al. (2003) used redundancy analysis and CCA and found that BMI metrics and genera abundances were most related to aspects of channel and riparian condition, substrate composition and soluble salts, but those authors did not include land use gradients in their analyses. Hall et al. (2006) found that of the 10 qualitative PHAB metrics recommended by Barbour et al. (1999), BMI metrics from the Stanislaus, Tuolumne and Merced Rivers were most strongly correlated with channel flow status and bank stability.

Metrics may have shown stronger responses to reach-scale PHAB variables than to land use in the present study because reach-scale stressor gradients cover a wide range from low to high (i.e., from good condition to poor condition), whereas most Central Valley sites have medium to high levels of watershed disturbance. Recent studies have found that stream biota respond to land use gradients most strongly in the range of 0-10% development in a watershed, but responses are muted when land use exceeds 10% (Hatt et al. 2004; Walsh et al. 2007). Mazor et al. (in prep) also found that two biotic indices (IBI and O/E) responded weakly to land use variables in low gradient, highly altered streams in California. Mazor et al. (in prep) did not evaluate the responses of the indices to reach-scale degradation. Given these results, land use criteria may be inappropriate as reference screens for Central Valley streams because of the high degree of alteration to essentially all watersheds and the generally poor response of BMI metrics to land use variables (Table 3).

Data were insufficient to evaluate whether sloughs and large rivers should be treated as separate waterbody types in the Central Valley, but several lines of evidence suggest this IBI can be used to score these waterbody types until the question can be addressed explicitly. Sloughs and large rivers were included in development of the Central Valley IBI, and none of the reference criteria established here explicitly excludes them from the reference pool. For example, most sloughs included in the Sacramento Valley Reference project (Ode et al. 2005) had < 50% fine sediment, and several large river sites from various projects had relatively high habitat condition scores (total PHAB scores > 150), but were excluded because they were downstream of dams. Also, sloughs and large river sites did not necessarily receive low IBI scores. For example, Union School Slough received a score of 68, and the Yuba River at Marysville received a score of 82, both of which fall within the distribution of IBI scores at reference sites (Fig. 3).

Conclusions and recommendations

1) This index can be used as a general interpretive framework for benthic samples collected at perennial streams (including sloughs and large rivers) on the Central Valley floor.

2) We recommend that all future bioassessment projects in the Central Valley collect quantitative PHAB and water chemistry with consistent protocols at all sites (e.g., the latest SWAMP Bioassessment Protocols, Ode 2007). *In situ* chemistry and rapid (qualitative) PHAB are not sufficient for thorough screening of reference sites and evaluation of metric responses to stressor gradients.

3) We recommend that non-point source (NPS) studies (e.g., grant projects, NPS monitoring) take advantage of opportunities to add biological sampling at monitoring sites where programs are already collecting more intensive stressor data (e.g., pesticides, nutrients, metals, etc.). This will provide California's monitoring programs with better datasets to support future analyses.

4) We recommend using reference conditions in the Sacramento Valley as a benchmark for conditions in the San Joaquin Valley unless future data sets suggest otherwise. The sparseness of data points (especially reference sites) in the southern San Joaquin and Tulare Basins may not be easily remedied given the more arid landscape of the southern Central Valley. However, no bioassessment projects have targeted the area south of Merced, and the only sampling points in the southernmost valley come from probability surveys (REMAP, EMAP and CMAP). Future reconnaissance and targeted sampling efforts will be required to fully assess the applicability of this index in the southern San Joaquin and Tulare Basins.

5) While we found no evidence of seasonal differences in IBI performance, the question of whether spring and fall samples need separate scoring scales could be more thoroughly answered if more reference sites had been sampled in the spring. The current comparisons (Fig 5), while fairly robust because of the large sample size, were based mostly on test sites.

6) Because of the high frequency of low-count samples collected with the EPA reachwide method in low gradient, sand bottom streams (see Results), we recommend the use of a modification of the reachwide sampling method in these stream types: sampling points at transects should alternate between 0%, 50% and 100% of stream width (i.e., margin-center-margin) instead of the standard 25%, 50% and 75% of stream width. This approach has been adopted for low gradient streams in California (Ode 2007) and also in national protocols (USEPA 2007).

7) The higher inter-annual variability in IBI score observed at sites with low IBI scores (Fig. 6) may indicate that sites in moderate to poor biological condition should be sampled multiple times before being listed as impaired or non-impaired in a regulatory context.

References

Barbour MT, Gerritsen J, Griffith GE, Frydenborg R, McCarron E, White JS, Bastian ML. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* **15**:185-211.

Barbour, M.T., J. Gerritsen, B.D. Snyder and J.B. Stribling. 1999. Revision to rapid bioassessment protocols for use in stream and rivers: periphyton, BMIs and fish. EPA 841-D-97-002. U.S. Environmental Protection Agency. Washington DC.

Blackburn, T.M., J.H. Lawton and J.N. Perry. 1992. A method for estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. Oikos 65: 107-112.

Brown, L.R. and J.T. May. 2000a. Macroinvertebrate assemblages on woody debris and their relations with environmental variables in the lower Sacramento and San Joaquin river drainages, California. Environmental Monitoring and Assessment 64:311-329.

Brown, L. R. and J.T. May 2000b. Benthic macroinvertebrate assemblages and their relations with environmental variables in Sacrament and San Joaquin river drainages, California, 1993-1997. U.S. Geological Survey Water-Resources Investigations Report 00-4125. Sacramento, CA 33 pgs.

Griffith, M.B., P. Husby, R.K. Hall, P.R. Kaufmann and B.H. Hill. 2003. Analysis of macroinvertebrate assemblages in relation to environmental gradients among lotic habitats of California's Central Valley. Environmental Monitoring and Assessment 82:281-309.

Hall, L.W., W.D. Killen and R.D. Anderson. 2006. Characterization of benthic communities and physical habitat in the Stanislaus, Tuolumne, and Merced Rivers, California. Environmental Monitoring and Assessment 115:223-264.

Harrington, J.M. 1999. California stream bioassessment procedures. California Department of Fish and Game, Water Pollution Control Laboratory. Rancho Cordova, CA.

Hatt, B.E., T.D. Fletcher, C.J. Walsh, and S.L. Taylor. 2004. The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams. Environmental Management 34: 112-124.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-47 in W.S. Davis and T.P. Simon (eds.) Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. CRC Press, Inc. Boca Raton.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger and D.V. Peck. 1999. Surface waters: quantifying physical habitat in wadeable streams. US EPA, Office of Research and Development. EPA/620/R-99/003.

Mazor, R.D., K. Schiff, K. Ritter, A.C. Rehn and P.R. Ode. (in prep). Bioassessment tools in novel habitats: an evaluation of indices and sampling methods in low-gradient streams in California.

Ode, P.R. 2007. Standard operating procedures for collecting macroinvertebrate samples and associated physical and chemical data for ambient bioassessments in California. California State Water Resources Control Board Surface Water Ambient Monitoring Program (SWAMP) Bioassessment SOP 001.

Ode PR, Rehn AC, May JT. 2005a. A quantitative tool for assessing the ecological condition of streams in southern coastal California. Environmental Management 35:493-504.

Ode, P.R., D.P. Pickard, J.P. Slusark and A.C. Rehn. 2005b. Adaptation of a reference site selection methodology to creeks and sloughs of California’s Sacramento Valley and alternative strategies for developing a regional bioassessment framework. California Department of Fish and Game, Water Pollution Control Laboratory. Rancho Cordova, CA.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the Association of American Geographers 77(1):118-125.

Peck DV, Herlihy AT, Hill BH, Hughes RM, Kaufmann PR, Klemm DJ, Lazorchak JM, McCormick FH, Peterson SA, Ringold PL, Magee T, Cappaert M. 2006. Environmental Monitoring and Assessment Program – Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA 620/R-06/003. U.S. Environmental Protection Agency, Office of Research and Development: Washington, D.C.

Rehn, A.C. 2008. Benthic macroinvertebrates as indicators of biological condition below hydropower dams on west slope Sierra Nevada streams, California, USA. River Research and Applications. DOI 10.1012/rra1121.

US EPA. 2007. National Rivers and Streams Assessment: Field Operations Manual. EPA-841-B-07-009. U.S. Environmental Protection Agency, Washington, D.C.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell and C.E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130-137.

Walsh, C.J., K.A. Waller, J. Gehling, and R. MacNally. 2007. Riverine invertebrate assemblages are degraded more by catchment urbanisation than by riparian deforestation. Freshwater Biology 52: 574-587.

Figure 1. California's Central Valley ecoregion with 141 BMI sampling locations used in IBI development and validation. Circled crosses indicate reference sites; plain circles indicate test sites.

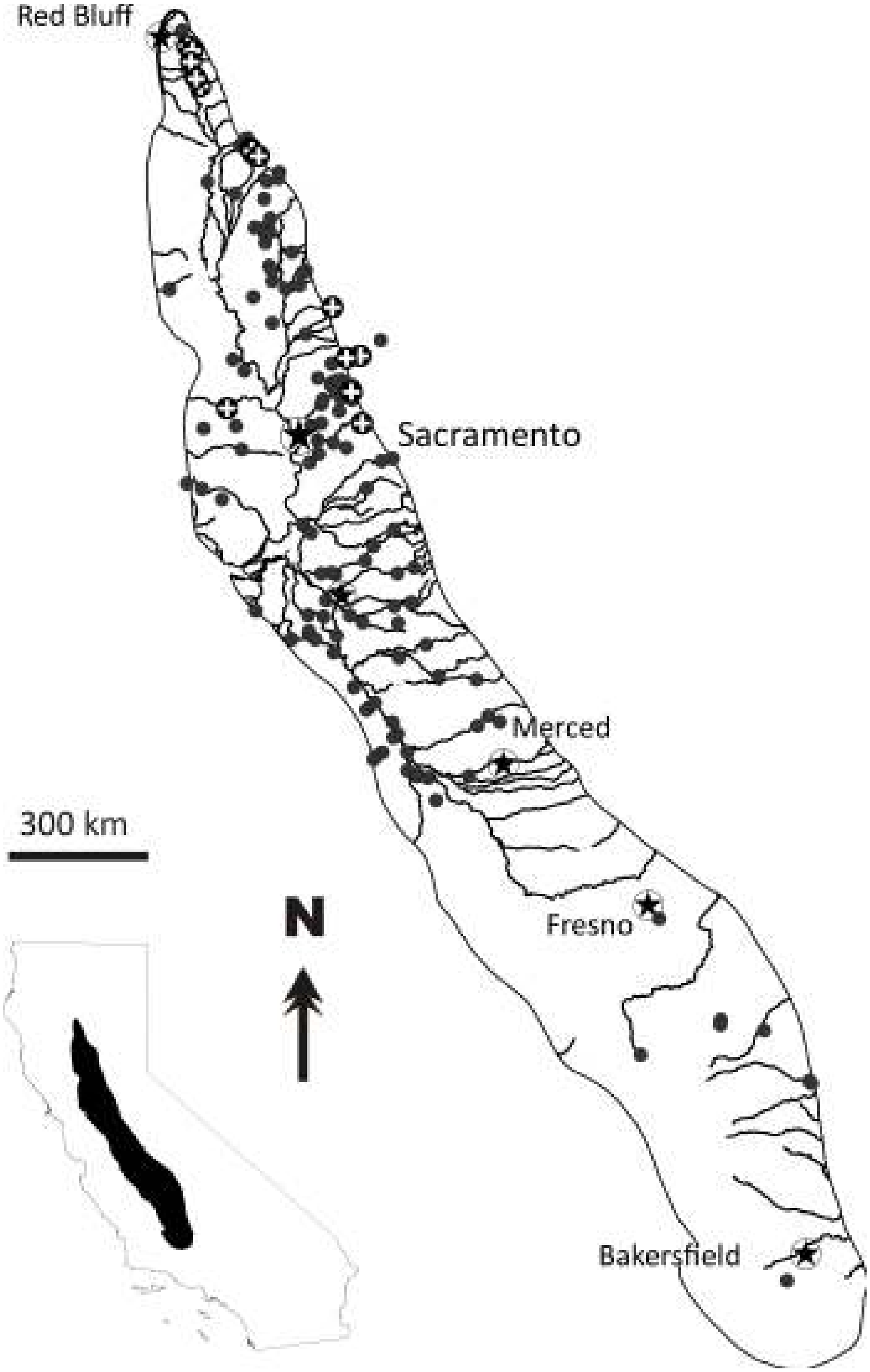

Figure 2. Example scatterplots showing stronger response of metrics to in-stream and riparian variables (a-c) than to land use (d) or water chemistry (e) variables.

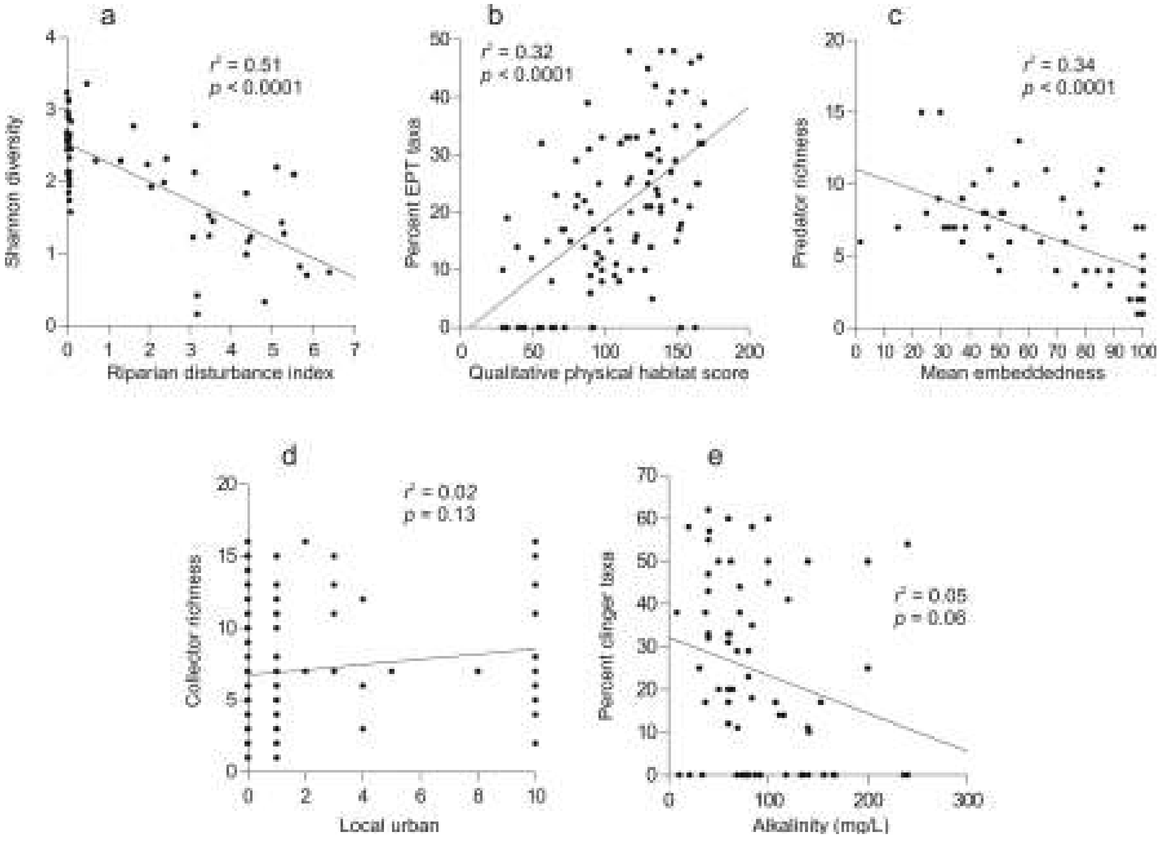

Figure 3. Box-and-whisker plots of IBI scores at Central Valley reference and test sites. Samples in the validation set were not used in metric screening or scoring. The hypothetical least-disturbed reference distribution is to emphasize that existing Central Valley reference sites represent best-available conditions given the extent of regional stream and landscape alteration. Boxes indicate median values and interquartile ranges, whiskers indicate 95th percentiles, outliers are indicated by an x.

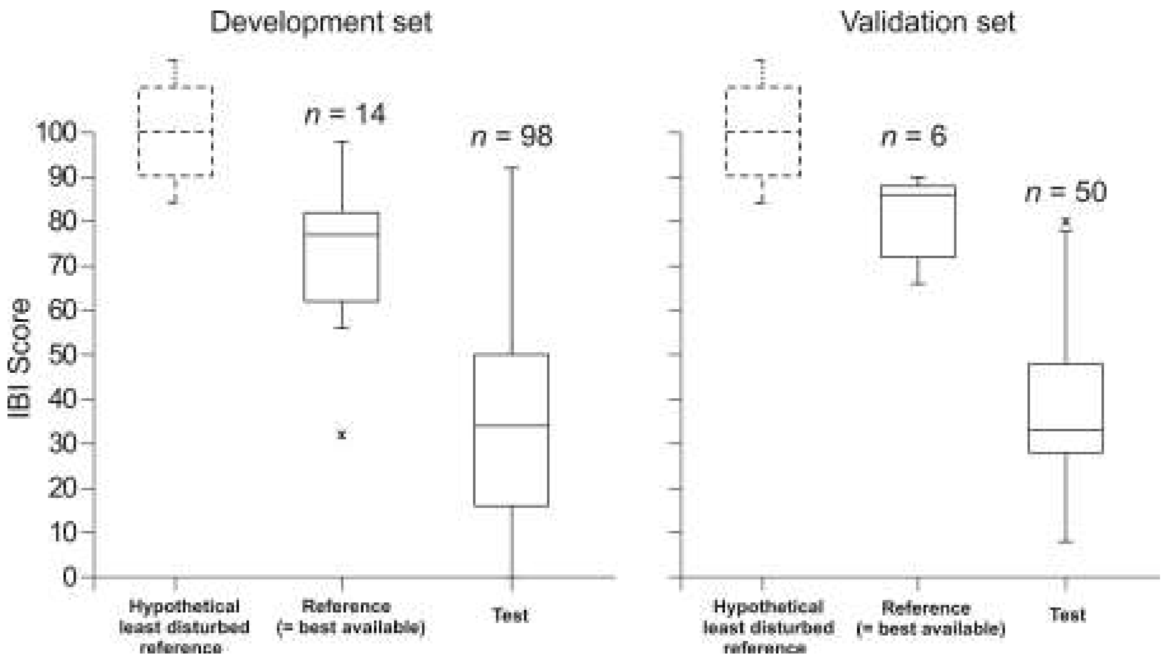

Figure 4. Scatterplots of IBI scores vs. qualitative land use intensity at local and watershed scales. Lines are best-fits from least-squares regressions.

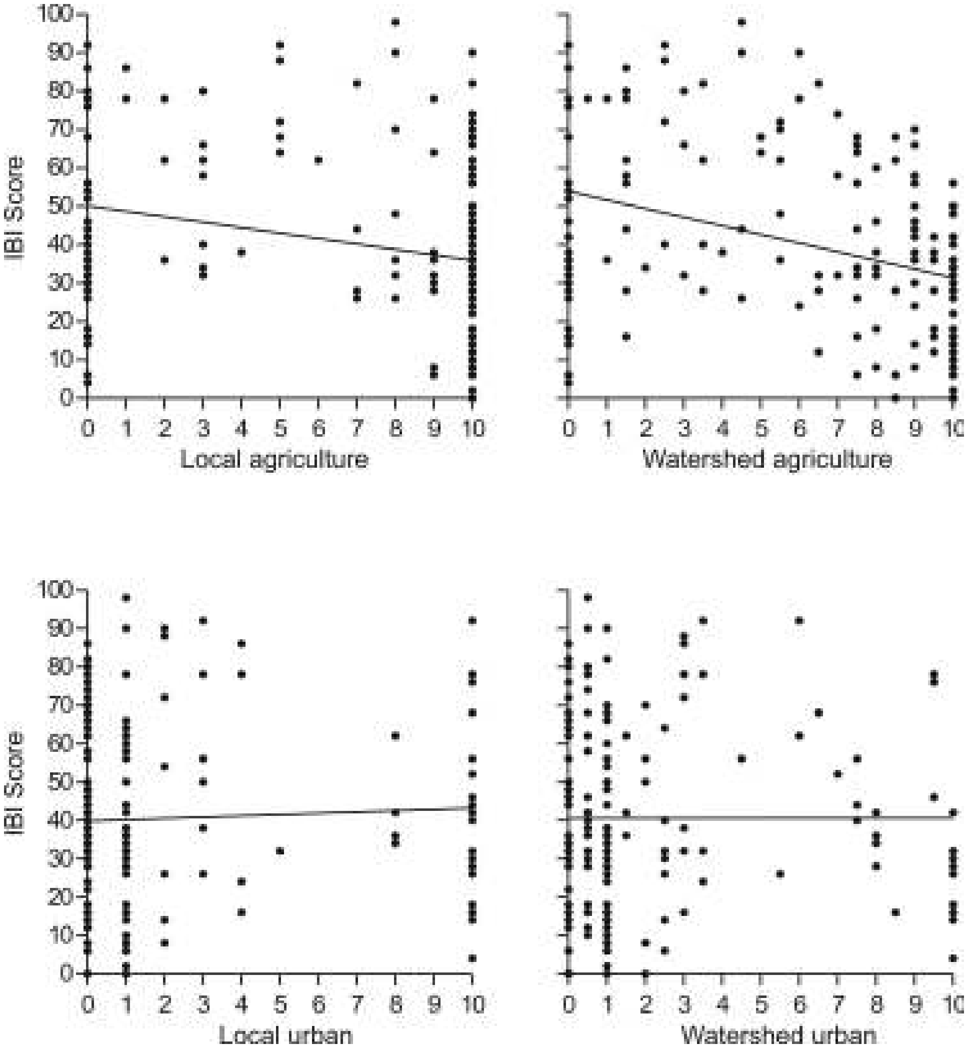

Figure 5. Box-and-whisker plots of raw values for the 5 final IBI metrics at sites with spring and fall samples. Most sites were sampled more than twice over years. Boxes indicate median values and interquartile ranges, whiskers indicate 95th percentiles, outliers are indicated by an x.

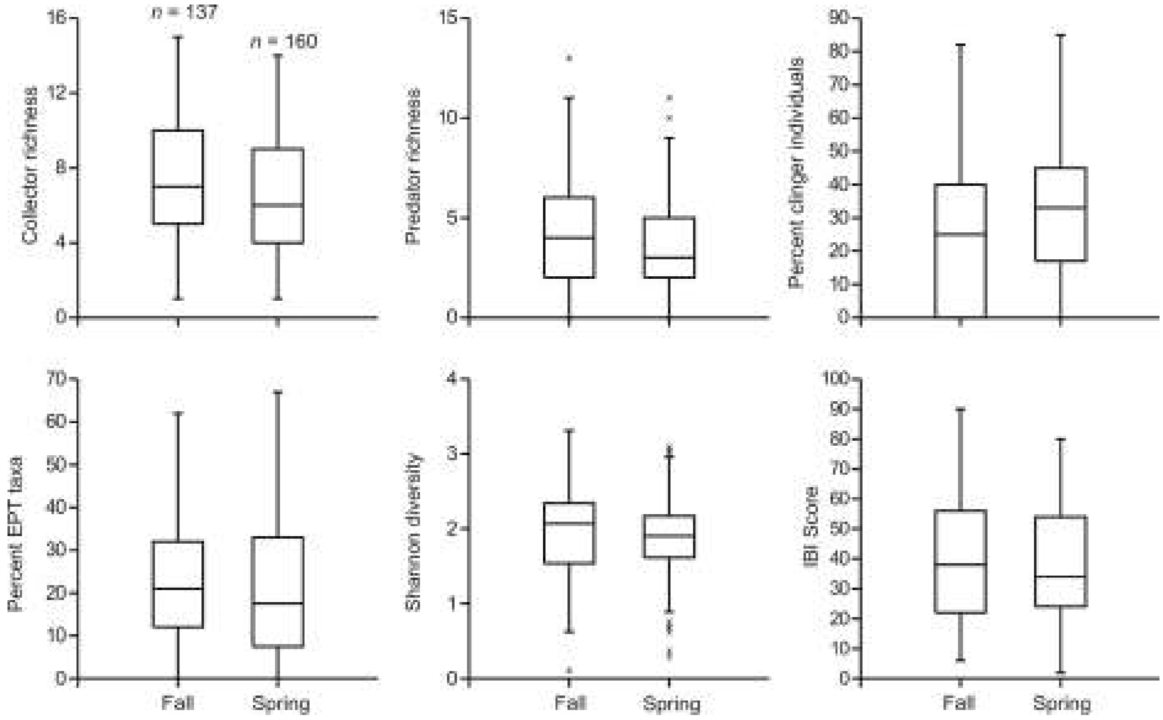

Figure 6. Scatterplot of mean IBI score vs. coefficient of variation at sites with at least 3 visits over time. Line is best-fit from least-squares regression.

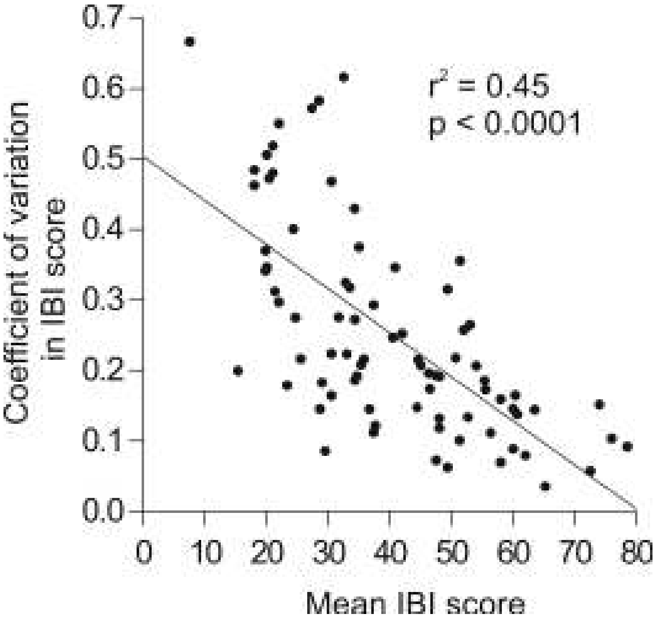

Table 1. Projects from which data were compiled for Central Valley IBI development. California Stream Bioassessment Procedures (CSBP) are from Harrington (1999); EPA multihabitat sampling method and qualitative PHAB method are from Barbour et al. (1999); EPA reachwide sampling method and quantitative PHAB method are from Peck et al. (2006). *In situ* water chemistry = pH, dissolved oxygen, conductivity, salinity and temperature measured with a hand held meter; lab analysis indicates that a water sample was collected for more detailed analysis of nutrients, metals and other analytes.

Project/sample processing lab	**# Streams**	**# Sites**	**Benthic sampling method**	**Specimen count**	**PHAB**	**Water chemistry**	**SEASON**
Sac River Basin- UC Davis ATL	16	45	CSBP	900	qualitative	lab analysis	Spring and Fall; 2000-2002
San Joaquin Ag Drains- UC Davis ATL	10	11	CSBP	900	qualitative	lab analysis	Spring and Fall; 2001
San Joaquin Basin TMDL- UC Davis ATL	16	22	EPA multihabitat	500	qualitative	lab analysis	Spring and Fall; 2002
Central Valley REMAP-ABL	82	82	EPA reachwide	500	quantitative	lab analysis	Late Summer-Fall; 1994-1995
University of Maryland-ABL	9	69	CSBP	900	qualitative	*in situ*	Spring; 2001-2007
Sac Valley Reference-ABL	30	30	EPA reachwide	500	quantitative	*in situ*	Fall 2002
Sac River Watershed-ABL	4	4	CSBP	900	qualitative	*in situ*	Fall; 2000-2002
Central Valley UAA-ABL	3	11	EPA reachwide	500	quantitative	*in situ*	Fall 2004
DPR San Joaquin-ABL	9	11	EPA reachwide	500	qualitative	*in situ* +pesticides	Spring 2005
Central Valley Bioassessment Project- UC Davis AEAL	24	34	EPA reachwide & multihabitat	500	qualitative	*in situ*	Spring and Fall; 2003-2005
EMAP/CMAP- ABL	22	22	EPA reachwide	500	quantitative	lab analysis	Summer and Fall; 2002-2006

Table 2. Scoring ranges for 5 component metrics in the Central Valley IBI.

Score	Collector richness	Predator richness	% EPT taxa	% Clinger taxa	Shannon diversity
0	0–3	0–1	0	0	≤ 1.28
1	4	2	1-4	1-6	1.29-1.47
2	5-6	3	5-9	7-12	1.48-1.66
3	7	4-5	10-13	13-19	1.67-1.84
4	8	6	14-18	20-25	1.85-2.03
5	9-10	7	19-22	26-31	2.04-2.22
6	11	8	23-27	32-37	2.23-2.40
7	12	9-10	28-31	38-44	2.41-2.59
8	13-14	11	32-36	45-50	2.60-2.78
9	15	12	37-40	51-57	2.79-2.96
10	≥ 16	≥ 13	41-100	58-100	≥ 2.97

Table 3. Coefficients of determination (r^2) for least-squares linear regressions between final candidate IBI metrics and stressor variables to which they best responded. Relationships between metrics and stressors often are not linear, but r^2 values still indicate relative strength of response. r^2 values < 0.1 are not shown. Values significant at $p < 0.0001$ are in bold. *N* for each stressor variable is listed in Appendix 1. Several final candidate metrics responded to individual components of the total qualitative PHAB score (from Barbour et al. 1999), but responses to the composite score were strongest.

Metric	watershed agriculture	local agriculture	watershed urban	local urban	riparian disturbance index	total qualitative PHAB score	channel alteration	percent sand and fines	mean embeddedness	in-stream habitat diversity	mean mid-channel canopy density	Kjeldhal N (mg/L)
Collector-filterer + collector-gatherer richness	**0.17**	--	--	--	**0.48**	**0.26**	**0.3**	0.14	**0.26**	0.2	--	0.15
Collector-gatherer richness	**0.14**	--	--	--	**0.37**	**0.2**	**0.22**	0.1	0.2	0.16	--	--
EPT richness	0.12	--	--	--	**0.33**	**0.36**	**0.33**	**0.23**	**0.31**	--	0.14	0.15
Non-insect taxa richness	--	--	--	--	**0.37**	**0.21**	**0.18**	--	0.22	0.17	--	0.13
Predator richness	**0.26**	**0.17**	--	--	**0.36**	**0.22**	**0.21**	--	**0.34**	--	0.14	0.17
Trichoptera richness	0.12	--	--	--	**0.27**	**0.36**	**0.34**	**0.24**	**0.27**	--	0.12	0.11
Percent burrower individuals	--	--	--	--	0.17	**0.24**	0.12	**0.19**	0.24	--	--	0.12
Percent Chironomidae individuals	--	--	--	--	0.22	0.12	--	--	0.21	--	--	0.32
Percent clinger taxa	--	--	--	--	**0.42**	**0.27**	**0.32**	**0.26**	**0.28**	0.18	--	0.19
Percent collector-filterer + collector-gatherer individuals	--	--	--	--	0.16	0.07	--	--	0.2	--	--	--
Percent collector-filterer individuals	--	--	--	--	0.2	**0.15**	0.17	--	--	--	--	0.1
Percent collector-filterer taxa	--	--	--	--	**0.4**	--	--	--	0.12	0.16	--	0.16
Percent collector-gatherer individuals	--	--	--	--	**0.33**	**0.23**	**0.24**	--	0.2	--	--	0.12
Percent Diptera individuals	--	--	--	--	0.17	--	--	--	0.18	--	--	0.31
Percent dominant taxon	--	--	--	--	**0.5**	**0.28**	**0.27**	--	0.18	0.15	--	0.28
Percent EPT taxa	--	--	--	--	**0.31**	**0.32**	**0.3**	**0.24**	**0.3**	--	--	0.22
Percent non-insect taxa	--	--	--	--	--	--	--	0.1	--	--	--	--
Percent predator individuals	--	--	--	--	--	--	--	--	--	0.11	--	--
Percent scraper individuals	--	--	--	--	--	--	0.11	--	0.11	--	--	--
Percent Trichoptera individuals	--	--	--	--	0.23	**0.17**	**0.18**	0.13	0.16	--	--	--
Percent Trichoptera taxa	--	--	--	--	**0.35**	**0.36**	**0.32**	**0.26**	**0.29**	0.13	--	0.16
Shannon Diversity	--	--	--	--	**0.51**	**0.40**	**0.38**	--	**0.26**	0.14	0.1	0.31
Taxonomic richness	**0.21**	0.12	--	--	**0.44**	**0.42**	**0.41**	**0.19**	**0.37**	0.12	0.17	0.18

Appendix 1. Summary of physical and chemical variables associated with benthic samples used in the Central Valley IBI development data set.

Variable	Number of sites	Mean	Min	Max
width (m)	61	6.67	0.13	32
depth (cm)	83	35.37	0.37	139.6
velocity (m/s)	20	0.76	0	2.93
sinuosity	23	1.10	1	1.89
gradient (% slope)	62	0.89	0	11.66
alkalinity (mg/L)	67	88.8	8.4	240
specific conductance (µS/cm^2)	101	228.9	0.04	3500
DO (mg/L)	84	10.85	1.2	263
pH	61	7.47	6.25	8.6
temperature (°C)	86	17.53	10.4	26.1
turbidity (NTU)	47	18.21	0.25	355
Ammonia (mg/L)	49	0.36	0	10.42
Arsenic (µg/L)	11	4.28	1.4	8.6
Boron (µg/L)	15	0.48	0.01	3.4
Cadmium (µg/L)	11	1.4	0	3.7
Chloride (mg/L)	30	50.95	0.3	905
Copper (µg/L)	10	3.45	1.8	8.43
Chromium (µg/L)	11	1.84	1.15	3.3
Kjeldhal N (mg/L)	30	1.15	0	5
Lead (µg/L)	11	5.06	0	29
Nickel (µg/L)	11	8.84	0	9.8
Nitrate (mg/L)	49	0.50	0	7
Phosphorous (mg/L)	50	0.77	0	9
Potassium (mg/L)	31	5.70	0.37	56.5
Sodium(mg/L)	31	63.60	1.1	673
Sulfate (mg/L)	31	39.06	0.11	530
Zinc (µg/L)	12	7.93	0.2	15
total dissolved solids (mg/L)	14	411.07	30	2820
total organic carbon (mg/L)	15	5.25	1	15.26
total suspended solids (mg/L)	29	194.51	0	2820
hardness (mg/L)	26	78.31	17	192
% concrete	98	2.26	0	97
% bedrock	82	0.80	0	25
% boulder	90	3.89	0	53.33
% small boulder	47	1.30	0	29
% large boulder	45	0.24	0	7
% cobble	90	12.06	0	67
% gravel	66	24.81	0	85
% course gravel	60	14.84	0	58
% fine gravel	63	8.83	0	38
% hardpan	79	5.14	0	66.66
% mud	20	37.75	3.33	90
% sand	85	24.88	0	82.85

Appendix 1 continued.

Variable	Number of sites	Mean	Min	Max
% fines	66	27.86	0	100
% sand and fines	105	50.01	0	100
mean embeddedness (quantitative)	54	65.96	1.72	100
riparian disturbance index (W1_HALL)	54	1.87	0	6.38
mean mid-channel canopy density	60	36.23	0	97.59
instream habitat diversity	54	1.33	0	6.28
qualitative embeddedness	31	16.25	2	93
qualitative epifaunal substrate	103	10.43	0	20
qualitative pool substrate	45	12.06	4	19
qualitative pool variablity	46	10.84	1	18
qualitative sediment deposition	103	10.06	0	20
qualitative channel flow status	103	13.04	0	20
qualitative channel alteration	103	10.96	1	20
qualitative channel sinuosity	46	7.76	0	20
qualitative bank stability	103	11.10	1	20
qualitative vegetative protection	103	10.85	0	20
qualitative riparian width	103	7.69	0	18
total qualitative habitat score	101	117.53	29	169
% fast-water habitat	32	7.65	0	100
% slow-water habitat	30	89.07	6	100
% pool	32	15.71	0	100
qualitative agricultural land use	112	31.94	0	100
qualitative urban land use	112	4.72	0	50
percent native land cover	46	26.88	0	100

Appendix 2. Sites used in development and validation of Central Valley IBI. For sites used in both development and validation sets, the collection date for development samples is listed first.

Project Name	Waterbody Name	Status	Collection Date*	Latitude	Longitude	Development	Validation
ATL Sac River	Auburn Ravine	Test	Fall-01	38.8700	-121.3566	X	
ATL Sac River	Auburn Ravine	Reference	Fall-01; Spring-02	38.9011	-121.2125	X	X
ATL Sac River	Butte Creek	Test	Fall-01; Spring-02	39.5301	-121.8584	X	X
ATL Sac River	Butte Creek	Reference	Fall-00; Spring-02	39.6994	-121.7771		X
ATL Sac River	Linda Creek	Test	Fall-00	38.7300	-121.2493	X	
ATL Sac River	Dry Creek	Test	Fall-00	38.7343	-121.3087	X	
ATL Sac River	Antelope Creek	Test	Fall-02; Spring-02	38.7876	-121.2489	X	X
ATL Sac River	Gilsizer Slough	Test	Fall-00; Spring-01	39.0260	-121.6592	X	X
ATL Sac River	Jack Slough	Test	Fall-00; Fall-01	39.1623	-121.5959	X	X
ATL Sac River	Main Canal	Test	Fall-01	39.3996	-121.7562	X	
ATL Sac River	Main Canal	Test	Fall-00; Spring-02	39.3924	-121.6840	X	X
ATL Sac River	Main Canal	Test	Fall-01; Spring-02	39.3952	-121.7160	X	X
ATL Sac River	Main Canal	Test	Fall-00	39.3804	-121.6787		X
ATL Sac River	Main Canal	Test	Fall-01	39.3779	-121.7062	X	
ATL Sac River	Main Canal	Test	Fall-01; Spring-02	39.4359	-121.6789	X	X
ATL Sac River	Pleasant Grove Creek	Test	Fall-00	38.8124	-121.4245	X	
ATL Sac River	Pleasant Grove Creek	Test	Fall-00	38.7959	-121.3555		X
ATL Sac River	Pleasant Grove Creek	Test	Fall-00; Fall-01	38.8055	-121.3087	X	X
ATL Sac River	South Branch Pleasant Grove Creek	Test	Fall-00	38.7711	-121.3159	X	
ATL Sac River	Wadsworth Canal	Test	Fall-01	39.2498	-121.6789	X	
ATL Sac River	Wadsworth Canal	Test	Fall-00	39.1897	-121.6620		X
ATL Sac River	Live Oak Slough	Test	Fall-00	39.2331	-121.6653	X	
ATL Sac River	Wadsworth Canal	Test	Fall-01	39.1273	-121.7566	X	
ATL Sac River	Rock Creek	Test	Fall-01	38.9643	-121.1101		X
ATL SJR AgDrains	Mtn House Creek	Test	5-Sep-01; 23-May-02	37.7856	-121.5356	X	X
ATL SJR TMDL	Los Banos Creek	Test	8-Oct-02	37.2764	-120.9539	X	
ATL SJR TMDL	Merced River	Test	1-Oct-02	37.4540	-120.6092		X
ATL SJR TMDL	Cosumnes River	Test	23-Oct-02	38.4904	-121.0978	X	
ATL SJR TMDL	Harding Drain	Test	9-Oct-02	37.4644	-121.0303	X	

*Exact collection dates were not known for the ATL Sac River project.

Appendix 2 continued.

Project Name	Waterbody Name	Status	Collection Date	Latitude	Longitude	Development	Validation
Central Valley UAA	New Alamo Creek	Test	04-Nov-04	38.3300	-121.8958	X	
Central Valley UAA	Morrison Creek	Test	05-Oct-04	38.5414	-121.2764	X	
Central Valley UAA	Morrison Creek	Test	11-Oct-04	38.4808	-121.4583	X	
Central Valley UAA	Alamo Creek	Test	15-Oct-04	38.3884	-122.0711	X	
CMAP	Big Chico Creek	Test	18-Jul-06	39.7489	-121.8008	X	
CMAP	New Creek	Test	28-Sep-05	40.1692	-122.1464	X	
CMAP	Cripple Creek	Test	19-Jul-04	38.6839	-121.3156	X	
CMAP	Arcade Creek	Test	16-Aug-06	38.6357	-121.4037	X	
CMAP	Morrison Creek	Test	10-Jul-06	38.5147	-121.4150	X	
CMAP	Dry Creek Tributary	Test	21-Jul-05	38.7086	-121.4056	X	
CMAP	Cherokee Canal	Test	25-Jul-06	39.5131	-121.7111	X	
CMAP	Little Chico Creek	Test	23-Jun-05	39.7208	-121.8383	X	
CMAP	Bear Creek	Test	07-Jul-05	38.0987	-121.1760	X	
CMAP	Pixley Slough	Test	26-Jun-06	38.0494	-121.3408	X	
CMAP	Dry Creek	Test	02-Jun-04	37.6475	-120.8056	X	
CMAP	Washington Colony Canal	Test	28-Aug-06	36.6956	-119.7164	X	
CMAP	Tulare Lake Canal	Test	29-Aug-06	36.1608	-119.8094	X	
CMAP	Stine Canal	Test	28-Jun-06	35.2711	-119.1078	X	
DPR 209	Bear Creek	Test	08-Jun-05	38.1549	-121.1336	X	
DPR 209	Laguna Creek	Test	20-Apr-05	38.3828	-121.1724		X
DPR 209	Little John Creek	Test	19-Apr-05	37.9201	-121.0269		X
DPR 209	Mormon Slough	Test	06-Jun-05	38.0494	-121.0134		X
DPR 209	Orestimba	Test	28-Jun-05	37.3294	-121.1096	X	
DPR 209	Marsh Creek	Test	25-Apr-05	37.8960	-121.7165		X
EMAP	Sacramento River	Test	14-May-02	39.5753	-122.0007		X
EMAP	Tule River	Test	20-Jun-02	36.0541	-118.9936	X	
REMAP	Cordua Canal	Test	31-Aug-95	39.2323	-121.4924	X	
REMAP	Unnamed Canal	Test	19-Jul-95	38.5586	-121.3427	X	
REMAP	Unnamed Canal	Test	26-Jul-95	38.8800	-121.8520	X	
REMAP	Unnamed Canal	Test	27-Jul-95	38.8369	-121.7942	X	
REMAP	Morrison Slough	Test	24-Aug-94	39.3404	-121.7024	X	
REMAP	Duck Creek	Test	13-Jul-95	37.9365	-120.9442	X	

Appendix 2 continued.

Project Name	Waterbody Name	Status	Collection Date	Latitude	Longitude	Development	Validation
REMAP	Unnamed Canal	Test	08-Sep-95	37.8550	-121.0067	X	
REMAP	Modesto Main Canal	Test	25-Aug-95	37.7177	-121.0002	X	
REMAP	Unnamed Canal	Test	12-Aug-94	37.7356	-121.3256	X	
REMAP	S.F. Persian Ditch	Test	03-Aug-94	36.2973	-119.4192	X	
REMAP	Davis Ditch	Test	03-Aug-94	36.2546	-119.2087	X	
REMAP	Evans Ditch	Test	27-Sep-95	36.2783	-119.4207	X	
REMAP	Campbell-Moreland Ditch	Test	23-Aug-95	36.0452	-118.9814	X	
Sac River Watershed Program	Dye Creek	Reference	16-Oct-02	40.1058	-122.1160	X	
Sac River Watershed Program	Yuba River	Test	27-Oct-00	39.1758	-121.5239	X	
Sac River Watershed Program	American River	Test	09-Nov-00	38.5681	-121.4222	X	
Sac River Watershed Program	Dry Creek	Reference	10-Oct-02	39.6798	-121.7423		X
Sac Valley Reference	Big Chico Creek	Test	12-Oct-04	39.7440	-121.8165		X
Sac Valley Reference	Deer Creek	Reference	22-Sep-04	39.9492	-122.0464	X	
Sac Valley Reference	Dye Creek	Reference	22-Sep-04	40.0883	-122.0903	X	
Sac Valley Reference	Mill Creek	Reference	22-Sep-04	40.0439	-122.0986	X	
Sac Valley Reference	Toomes Creek	Reference	22-Oct-04	39.9797	-122.0681	X	
Sac Valley Reference	Cache Creek	Reference	29-Sep-04	38.6870	-121.8765	X	
Sac Valley Reference	Putah Creek	Test	18-Oct-04	38.5272	-121.8017	X	
Sac Valley Reference	Union School Slough	Test	21-Oct-04	38.6070	-121.9920	X	
Sac Valley Reference	Ulatis Creek	Test	27-Sep-04	38.3694	-121.9947	X	
Sac Valley Reference	Willow Slough	Test	21-Oct-04	38.6198	-121.8327	X	
Sac Valley Reference	Bear River	Test	08-Oct-04	38.9849	-121.4868	X	
Sac Valley Reference	Dry Creek	Reference	04-Oct-04	39.0896	-121.3564	X	
Sac Valley Reference	South Honcut Creek,	Test	30-Sep-04	39.3042	-121.5650	X	
Sac Valley Reference	Jack Slough	Test	23-Oct-04	39.2250	-121.5104	X	
Sac Valley Reference	Alder Creek	Reference	25-Oct-04	38.6375	-121.1985	X	
Sac Valley Reference	Dry Creek	Test	15-Oct-04	38.7299	-121.3981	X	
Sac Valley Reference	Miners Ravine	Reference	25-Oct-04	38.7590	-121.2559	X	
Sac Valley Reference	Secret Ravine	Test	25-Oct-04	38.7600	-121.2574	X	

Appendix 2 continued.

Project Name	Waterbody Name	Status	Collection Date	Latitude	Longitude	Development	Validation
Sac Valley Reference	Butte Creek	Reference	12-Oct-04	39.6964	-121.7767	X	
Sac Valley Reference	Clear Creek	Test	15-Oct-04	39.5817	-121.6988	X	
Sac Valley Reference	Auburn Ravine**	Reference	13-Oct-04	38.8911	-121.2828	X	X
Sac Valley Reference	Comanche Creek**	Reference	22-Oct-04	39.7073	-121.7853	X	X
Sac Valley Reference	Dry Creek	Test	30-Sep-04	39.6197	-121.6365	X	
Sac Valley Reference	Gold Run Creek	Test	29-Oct-04	39.5934	-121.6396	X	
Sac Valley Reference	Hamlin Creek Trib.	Reference	20-Oct-04	39.6898	-121.7720	X	
Sac Valley Reference	Salt Creek	Test	11-Oct-04	39.1508	-122.1812	X	
Sac Valley Reference	Spring Creek	Test	11-Oct-04	39.1473	-122.1795	X	
UCD CVBP	Bear Creek	Test	2-Nov-03; 1-Oct-02	37.2556	-120.6519	X	X
UCD CVBP	Salt Slough	Test	2-Nov-03	37.2486	-120.8511	X	
UCD CVBP	Mud Slough	Test	17-Jun-04; 10-Jun-05	37.2542	-120.9069	X	X
UCD CVBP	Mud Slough	Test	6-Nov-03	37.2639	-120.9061		X
UCD CVBP	Merced River	Test	28-Oct-04	37.3497	-120.9578	X	
UCD CVBP	Ingalsby Slough	Test	3-Jun-05; 29-Oct-04	37.4918	-120.5578	X	X
UCD CVBP	Merced River	Test	29-Oct-04; 5-May-03	37.4702	-120.5005	X	X
UCD CVBP	Cosumnes River	Test	3-Nov-04; 26-May-04	38.5006	-121.0450	X	X
UCD CVBP	Mokelumne River	Test	21-Jul-03	38.2353	-121.4869	X	
UCD CVBP	Lone Tree Creek	Test	21-Oct-04	37.8556	-121.1847	X	
UCD CVBP	French Camp Slough	Test	22-Oct-04; 16-Oct-02	37.8817	-121.2492	X	X
UCD CVBP	Old River	Test	5-Nov-03	37.8047	-121.4494	X	
UCD CVBP	Mokelumne River	Test	22-Oct-04; 9-Oct-03	38.2225	-121.0344	X	X
UCD CVBP	Calaveras River	Test	21-Oct-04; 2-Oct-02	38.0727	-120.9310	X	X
UCD CVBP	Bear Creek	Test	3-Jun-05; 3-Oct-03	38.0432	-121.3224	X	X
UCD CVBP	Roberts Island, unnamed intake channel	Test	21-Jun-04	37.8767	-121.3766	X	
UCD CVBP	Roberts Island, Main Drain	Test	21-Jun-04	37.9415	-121.3692	X	
UCD CVBP	Middle River	Test	9-Nov-03	37.8816	-121.4544	X	
UCD CVBP	Grantline/Fabian Bell Canal	Test	6-Nov-03	37.8192	-121.4499	X	
UCD CVBP	Sugar Cut	Test	14-Nov-03	37.7859	-121.4194	X	
UCD CVBP	Disappointment and Pixley Slough	Test	11-Jun-03	38.0451	-121.3918	X	
UCD CVBP	San Joaquin River	Test	15-Nov-03; 16-Jun-03	37.8014	-121.3134	X	X

** Same-day duplicates from the Sac Valley Reference project's Auburn Ravine and Comanche Creek sites were used in development and validation sets.

Appendix 2 continued.

Project Name	Waterbody Name	Status	Collection Date	Latitude	Longitude	Development	Validation
UCD CVBP	Beaver Slough	Test	13-Nov-03	38.2042	-121.4463	X	
UCD CVBP	Orestimba Creek	Test	13-Oct-04; 28-May-03	37.4139	-121.0142	X	X
UCD CVBP	Ingram Creek	Test	2-Jun-05	37.6003	-121.2242	X	
UCD CVBP	Del Puerto Creek	Test	28-Oct-04	37.5214	-121.1486	X	
UCD CVBP	Orestimba Creek	Test	25-Oct-03; 18-Jun-01	37.3458	-121.0792	X	X
Univ of Maryland	Arcade Creek	Test	24-May-00	38.6250	-121.4556		X
Univ of Maryland	Kaseberg Creek	Test	17-May-06	38.7838	-121.3575		X
Univ of Maryland	Kaseberg Creek Trib 2	Test	18-May-06	38.7591	-121.3325		X
Univ of Maryland	Pleasant Grove Creek	Test	17-May-06	38.8058	-121.3064		X
Univ of Maryland	Pleasant Grove Creek Trib 2	Test	17-May-06	38.8041	-121.3283		X
Univ of Maryland	SF Pleasant Grove Creek	Test	19-May-06	38.7662	-121.2835		X
Univ of Maryland	Stanislaus River	Test	27-May-03	37.7711	-120.8683		X
Univ of Maryland	Tuolumne River	Test	25-May-03	37.6350	-120.6178		X
Univ of Maryland	Del Puerto Creek	Test	18-May-04	37.5403	-121.1186		X
Univ of Maryland	Del Puerto Creek	Test	23-May-06	37.5119	-121.1606		X
Univ of Maryland	Orestimba Creek	Test	20-May-04	37.4194	-121.0025		X
Univ of Maryland	Orestimba Creek	Test	15-May-02	37.4053	-121.0244		X
Univ of Maryland	Orestimba Creek	Test	20-May-03; 20-May-04	37.3189	-121.1217		X
Univ of Maryland	Salt Slough	Test	20-May-04	37.2483	-120.8533		X
Univ of Maryland	Salt Slough	Test	12-May-02	37.1592	-120.8122		X

Appendix 3. Eighty metrics screened for use in the Central Valley IBI. Discrimination between reference and non-reference sites is listed as 'good' (quartiles of reference and first downstream distributions do not overlap in box-and-whisker plots), 'fair' (quartiles overlap but at least one median is outside the other distribution's quartiles in box-and-whisker plots) or 'poor' (quartiles overlap and each median is within the other distribution's quartiles in box-and-whisker plots). See Barbour *et al*. (1996) for more detail on scoring discrimination in box-and-whisker plots. Metrics that failed the range test are marked with an asterisk (*); metrics that had a value of '0' at ≥ 70% of development sites are in italics. Metrics selected for inclusion in the IBI are in bold.

Metric	Response/ Discrimination	Notes
Coleoptera richness*		
Collector-filterer richness*		
Collector-gatherer richness	Fair/Good	
Collector-filterer + collector-gatherer richness	Good/Good	
Diptera richness*		
Elmidae richness*		
Ephemerellidae richness*		
Ephemeroptera richness*		
EPT Richness	Good/Good	correlated w/ all final metrics
Hydropsychidae richness*		
Intolerant EPT richness		
Intolerant richness		
Mollusca richness*		
Non-insect taxa richness	Good/Fair	
Plecoptera richness*		
Predator richness	Good/Good	
Scraper richness*		
Shredder richness*		
Trichoptera richness	Fair	
Percent Amphipoda individuals	Poor	
Percent Baetidae individuals	Poor	
Percent burrower individuals	Fair	
Percent collector-filterer + collector gatherer individuals	Fair	
Percent collector-filterer + collector gatherer taxa	Poor	
Percent collector-filterer taxa	Fair	
Percent collector-gatherer taxa	Poor	
Percent Chironomidae individuals	Fair /Poor	
Percent clinger taxa	Good/Good	
Percent collector-filterer individuals	Fair	
Percent collector-gatherer individuals	Fair	
Percent Corbicula individuals	Poor	
Percent Crustacea individuals	Poor	

Appendix 3 continued.

Metric	Response/ Discrimination	Notes
Percent Diptera individuals	Fair	
Percent Diptera taxa	Poor	
Percent dominant taxon	Good/Fair	correlated w/ Shannon Diversity
Percent Elmidae individuals		
Percent Ephemeroptera individuals	Poor	
Percent Ephemeroptera taxa	Poor	
Percent EPT individuals	Fair	
Percent EPT taxa	Good/Good	
Percent Gastropoda individuals	Poor	
Percent Glossosomatidae individuals		
Percent Hydropsychidae individuals	Poor	
Percent Hydroptilidae individuals		
Percent intolerant individuals		
Percent intolerant Diptera individuals *		
Percent intolerant Ephemeroptera individuals*		
Percent intolerant scraper individuals		
Percent intolerant taxa		
Percent intolerant Trichoptera individuals		
Percent Mollusca individuals	Poor	
Percent non-*Baetis Fallceon* Ephemeroptera individuals	Poor	
Percent non *Hydropsyche-Cheumatopsyche* Trichoptera individuals	Poor	
Percent non-Gastropoda scraper individuals		
Percent non-*Hydropsyche* Hydropsychidae individuals*		
Percent non-Insecta taxa	Fair	
Percent of Ephemeroptera individuals that are intolerant		
Percent of Trichoptera individuals that are intolerant		
Percent Oligochaeta individuals	Poor	
Percent Omnivore taxa		
Percent Perlodidae individuals *		
Percent Philopotamidae individuals *		
Percent Plecoptera individuals *		
Percent Plecoptera taxa*		
Percent predator taxa	Poor	
Percent predator individuals	Fair	
Percent Rhyacophilidae individuals *		
Percent scraper taxa	Poor	
Percent scraper individuals	Fair	
Percent sensitive EPT individuals	Poor	
Percent shredder taxa		
Percent shredder individuals		

Appendix 3 continued.

Metric	Response/ Discrimination	Notes
Percent Simuliidae individuals	Poor	
Percent tolerant individuals	Poor	
Percent tolerant taxa	Poor	
Percent Trichoptera individuals	Poor	
Percent Trichoptera taxa	Good/Fair	
Shannon Diversity	Good/Good	
Taxonomic richness	Good/Good	correlated w/ predator richness & Shannon Diversity
Weighted average tolerance value	Poor	