



ELSEVIER

Ecological Indicators 7 (2007) 807–826

ECOLOGICAL
INDICATORS

This article is also available online at:
www.elsevier.com/locate/ecolind

Development of a RIVPACS-type predictive model for bioassessment of wadeable streams in Wyoming

Eric G. Hargett^{a,*}, Jeremy R. ZumBerge^b, Charles P. Hawkins^c, John R. Olson^c

^a Wyoming Department of Environmental Quality-Water Quality Division, 122 W. 25th Street, Herschler Bldg. 4-W, Cheyenne, WY 82002, USA

^b Wyoming Department of Environmental Quality-Water Quality Division, 1866 S. Sheridan Avenue, Sheridan, WY 82801, USA

^c Western Center for Monitoring and Assessment of Freshwater Ecosystems, Department of Aquatic, Watershed, and Earth Resources, Utah State University, Logan, UT 84322, USA

Received 30 March 2006; received in revised form 2 October 2006; accepted 4 October 2006

Abstract

RIVPACS models produce a community-level measure of biological condition known as *O/E*, which is derived from a comparison of the observed (*O*) biota with those expected (*E*) to occur in the absence of anthropogenic stress. We used benthic macroinvertebrate and environmental data collected at 925 stream monitoring stations, from 1993 to 2001, to develop, validate, and apply a RIVPACS model to assess the biological condition of wadeable streams in Wyoming. From this dataset, 296 samples were identified as reference, 157 of which were used to calibrate the model, 46 to validate it, and 93 to examine temporal variability in reference site *O/E*-values. We used cluster analyses to group the model development reference sites into biologically similar classes of streams and multiple discriminant function analysis to determine which environmental variables best discriminated among reference groups. A suite of 14 categorical and continuous environmental variables best discriminated among 15 reference groups and explained a large proportion of the natural variability in biota within the reference dataset. Eleven of the predictor variables were derived from GIS. As expected, mean *O/E*-values for reference sites used in model development and validation were near unity and statistically similar. Temporal variability in *O/E*-values for reference sites was low. Test site values ranged from 0 to 1.45 (mean = 0.73). The model was accurate in both space and time and precise enough (S.D. of *O/E*-values for calibration data = 0.17) to detect modest alteration in biota associated with anthropogenic stressors. Our model was comparable in performance to other RIVPACS models developed in the United States and can produce effective assessments of biological condition over a broad, ecologically diverse region. We also provide convincing evidence that RIVPACS models can be developed primarily with GIS-based predictor variables. This framework not only simplifies the extraction of predictor variable information while potentially reducing expenditures of time and money in the collection of predictor variable information, but opens the door for development and/or application of RIVPACS models in regions where there is a paucity of local-scale, abiotic information.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Benthic macroinvertebrates; Bioassessment; Biological assessment of freshwater ecosystems; Biological indices; GIS; Modeling; *O/E*; Predictive model; RIVPACS; Streams; Wyoming

* Corresponding author. Tel.: +1 307 777 6701; fax: +1 307 777 5973.

E-mail address: eharge@state.wy.us (E.G. Hargett).

1. Introduction

RIVPACS (River InVertebrate Prediction And Classification System) is a multivariate predictive model that allows detection and interpretation of anthropogenic stress on invertebrate assemblages of streams and rivers (Clarke et al., 2003; Moss et al., 1987; Wright et al., 1993, 2000). Its derivative, AUSRIVAS (AUStralian RIVER Assessment System), is widely used to assess the biological condition of streams in Australia (Smith et al., 1999). Although these predictive models have been used in Great Britain and Australia for more than a decade, their potential has not been fully explored in the United States (Hawkins, 2006; Hawkins and Carlisle, 2001; Hawkins et al., 2000).

RIVPACS models make site-specific predictions of the benthic macroinvertebrate fauna expected in the absence of anthropogenic stressors. Those predictions are based on empirical relationships between individual taxon probabilities of capture and natural environmental features (e.g., latitude, substrate composition, alkalinity, elevation, etc.) that are derived from data collected from a reference site network. The deviation of the observed from the expected fauna, is usually (but not necessarily) measured by the ratio (O/E) of the observed (O) to expected (E)-values of one or more biotic index (Clarke et al., 1996). When the O/E index is expressed in units of taxa richness, it can be a measure of compositional similarity and thus a community-level measure of biological integrity (Hawkins, 2006).

There has been growing recognition among natural resource managers throughout the United States that biological indices such as O/E are useful and desirable tools in the evaluation of biological integrity of streams and to satisfy requirements under Sections 305(b) and 303(d) of the Clean Water Act (1972, amended in 1977). Use of RIVPACS models in the United States for bioassessment purposes is still a relatively new concept, but is gaining popularity. Several predictive models developed with datasets from the states of California, Maine, North Carolina, Ohio, Oregon, Washington, and the Mid-Atlantic Highlands region (e.g., Hawkins, 2006; Hawkins et al., 2000) show promise as effective tools in the evaluation of stream biological condition in the United States.

Advantages of a RIVPACS model compared to other bioassessment tools (e.g., multimetric and biotic indices) include intuitive output, ease of biological interpretation and its inherent standardization to site-specific conditions (Hawkins, 2006). One of the goals of the Wyoming Department of Environmental Quality-Water Quality Division (WDEQ/WQD), the primary entity responsible for protecting and managing biological integrity of streams in Wyoming, is the continued development of the most effective and applicable bioassessment tools to ascertain the condition of aquatic life in streams and rivers. The applicability of the RIVPACS approach to freshwater systems internationally as well as its promise as an effective bioindex tool in the United States presents an ideal opportunity to develop such a model for use in bioassessments in Wyoming. To our knowledge, a Wyoming RIVPACS model would represent the first attempt at developing such a regional framework for the United States intermountain west.

The objective of this study was to develop and evaluate a RIVPACS model that was applicable to wadeable streams in the State of Wyoming. To do so, we used an extensive statewide database that contained 9 years of benthic macroinvertebrate, physical, and chemical data collected at reference and non-reference sites by the WDEQ/WQD. We followed the latest techniques in both the development of RIVPACS models as well as assessing their accuracy and precision. We then applied the model to our entire dataset to assess the biological condition of individual sites.

2. Methods and materials

2.1. Study area

Wyoming is biologically diverse, with much of this diversity attributable to variability in geology, climate, topography, and other environmental features of the state (Knight, 1994). The State of Wyoming straddles the continental divide and encompasses 251,489 km² (97,100 mi²). Wyoming is characterized by abrupt topographic relief and numerous types of exposed granitic, volcanic, and sedimentary bedrock. Elevation ranges from 939 to 4207 m (3081–13,802 ft) with a mean of 2030 m (6660 ft). Average annual precipitation ranges from 15 to 150 cm (6–59 in.), which is mostly in

the form of rain in the plains regions and snow in the mountain and intermountain basins. Temperature in Wyoming varies widely due to the great topographic relief of the state. For example, mean daily maximum and minimum temperatures for July range from 32 to <24 °C (90 to <75 °F) and 13 to 0 °C (55–32 °F), respectively.

Omernik and Gallant (1987) divided Wyoming into five level III ecoregions: Middle Rockies, Southern Rockies, Northwestern Great Plains, Wyoming Basin, and the Western High Plains. The Middle Rockies consist of the Black Hills in northeastern Wyoming, the Bighorn Mountains in northcentral Wyoming, and the Teton, Absaroka, Gallatin, Wyoming, Salt River, Wind River, Beartooth, and other ranges of northwestern/western Wyoming. Because of differences in abiotic and biotic characteristics between the different mountain ranges, the Middle Rockies ecoregion is fairly heterogeneous and as a result, can be divided into three sub-regions: Middle Rockies East (Black Hills), Middle Rockies Central (Bighorn Mountains), and Middle Rockies West (mountain ranges of northwest and western Wyoming). The Laramie, Medicine Bow, and Sierra Madre ranges of south-central and southeast Wyoming comprise the Southern Rockies. The mountains of Wyoming are characterized by coniferous forest, aspen groves, sub-alpine meadows, and alpine tundra. The mixed-grass prairie of the Northwestern Great Plains makes up most of the eastern one-third of the state and the short-grass prairie Western High Plains are confined to the southeast corner of Wyoming. The remainder of the state is considered part of the Wyoming Basin, which is a high desert elevated plateau that consists of sagebrush, greasewood, and saltbush shrublands. Adding to the ecological diversity of Wyoming are escarpments of sedimentary and granitic rock scattered throughout the plains and basin regions of the state. Most streams in the mountains are classified as coldwater systems by the WDEQ/WQD based on the maximum temperature criteria of 20 °C (68 °F) (WDEQ, 2001). Streams in the plains and basin regions are a diverse mixture of coldwater and warmwater systems.

2.2. Reference and test samples

The WDEQ/WQD collected a total of 925 samples between 1993 and 2001 (Fig. 1). From this dataset,

296 samples were identified as reference based on whether water chemistry met numeric in-stream aquatic life criteria (WDEQ, 2001), sites possessed stable and diverse bed, bank, and in stream habitat conditions, and whether sites were minimally impacted by anthropogenic stressors. A subset of 157 reference sites sampled from 1993 to 1999 was used in model development (reference calibration). Another subset of 46 reference sites sampled in 2000 and 2001 was used in validation to evaluate whether the model could correctly assess sites of known reference condition (reference validation). An additional 93 samples that were collected at 32 previously sampled reference sites (repeat reference samples) were used to evaluate temporal variability in reference site *O/E*-values. The number of revisits to each of these reference sites ranged from one to eight, with most sites having at least three revisits. The remaining 629 samples were designated as test samples and evaluated with the model.

2.3. Macroinvertebrates

Benthic macroinvertebrates were collected from riffles with a Surber sampler (0.09 m² = 1 ft²) and 500- μ mesh. This habitat was selected because it is considered to have the greatest diversity and density of organisms and provides the best measure of overall health of the macroinvertebrate assemblage (Barbour et al., 1999). In the absence of riffles, samples were collected in runs. Each sample consisted of a composite of eight Surber samples collected randomly along a 30.5 m (100 ft) maximum length of riffle or run. Samples were collected from downstream to upstream to avoid habitat disruption, placed in polyethylene bottles and preserved in either 10% formalin or 99% isopropanol. All samples were collected during wadeable baseflow conditions and in habitats that possessed sufficient depth to remain submerged during periods of low flow. Repeat samples for a given site were collected from the riffle or run identified during the initial monitoring of that site and within 1 week of the original sample date (day of year). Samples in the montane regions were sampled from 1 August to 31 October, whereas plains and basin streams were sampled from 15 July to 31 October. Repeat samples were generally collected at 1 year intervals following initial sampling.

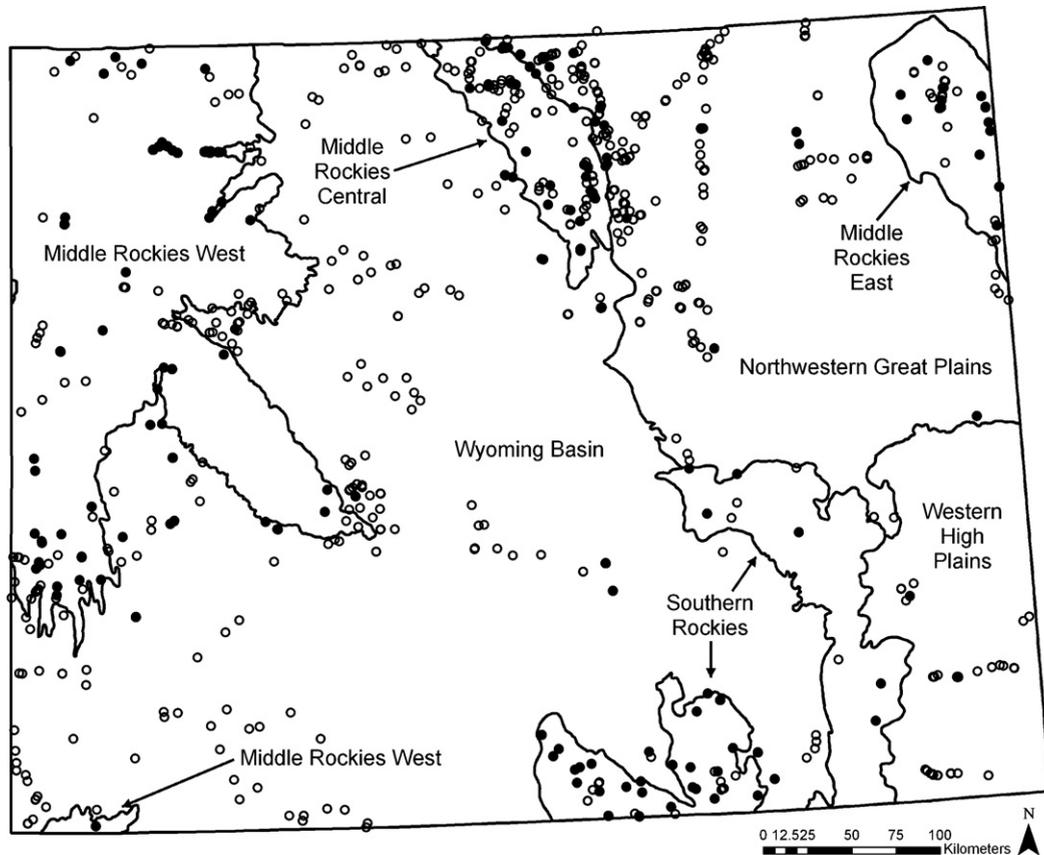


Fig. 1. Location of reference (solid circles) and test (open circles) sites in relation to ecoregions and sub-regions of Wyoming.

Macroinvertebrate data were reported as raw taxa counts and identified to the lowest taxonomic level possible (usually genus). To ensure taxonomic resolution was consistent among samples, we aggregated data to operational taxonomic units (OTUs) as described by Hawkins et al. (2000). A sub-sample of 300 randomly selected individuals was used in predictive model development and testing. Ostermiller and Hawkins (2004) concluded that the precision, accuracy and sensitivity of RIVPACS-type models increases with sample size and recommended sub-sample counts of at least 300–350 individuals.

2.4. Water chemistry, physical habitat and landscape characterization

Water quality data (pH, temperature, sulfate, total phosphorus, nitrate-nitrogen, alkalinity, hardness, total suspended solids, turbidity, chloride, dissolved

oxygen, and conductivity) associated with each sample were collected at one location directly below the base of the riffle or run sampled for macroinvertebrates and prior to macroinvertebrate sampling to minimize contamination from disturbance of the upstream sample area. At each site, stream current velocity (m/s), stream discharge (m³/s) and substrate composition of the macroinvertebrate sample riffle/run were gathered. Substrate composition was visually estimated as the percentage of cobble (64–254 mm), coarse gravel (25.4–63 mm), fine gravel (7.62–25.3 mm), sand (<7.62 mm, gritty), silt (<7.62 mm, soft and fine), and clay (<7.62 mm, solid and slick). Within each of the eight Surber samples for a riffle/run, visual estimates of substrate composition were conducted and averaged. We grouped values for each substrate category into two substrate composition variables used in subsequent analyses: coarse substrate (sum of cobble, coarse gravel, and fine gravel)

and fine substrate (sum of sand, silt, and clay). Water quality data, stream current velocity, stream discharge and substrate composition were measured at the time of each macroinvertebrate sample collection. A digital planimeter or a geographic information system (GIS) was used to calculate watershed area (m^2) and elevation (m) from 1:24,000 United States Geological Survey (USGS) DRG-E topographic maps. We obtained latitude and longitude coordinates with a handheld global positioning system (GPS) (WGS 1984 datum) and converted them to decimal degrees. Sites were assigned to one of seven ecoregion or sub-region classifications as described previously and classified as located in either a plains, foothills, or mountains landscape. Each site was identified by dominant stream origin: montane snowmelt, spring, mixture or other. Primary contributing geology for each site was identified from a 1:500,000 geological bedrock map of Wyoming (USGS, 1994). Chemical and physical weathering rates and rock nutrient content (either phosphorous, sulfate or nitrate) of geological formations within a site's watershed were developed based on the 1:500,000 geological bedrock map (J.R. Olson, Utah State University, unpublished data). Chemical activity of geological formations was assigned an ordinal ranking from low activity (1: granitics, gneiss) to high activity (5: limestone, dolomite). Likewise, physical activity of rock formations was assigned an ordinal ranking from low activity (1: granitics, gneiss, limestone) to high activity (5: siltstone, shale). Rock nutrient content was classified into one of three ranked (ordinal) categories that ranged from low (1) activity (granitics and gneiss with nearly no nutrient content) to high (3) activity (phosphate, gypsum). Mean, majority, and maximum index values for chemical, physical, and nutrient activity were calculated with GIS analysis of the distribution of geology classes in the watershed.

2.5. Predictive model construction

2.5.1. Data analysis

Chemical and physical variables were evaluated for normality and either square-root, log, or $\log(X + 1)$ transformed as necessary. PC-ORD (Version 4.0) (McCune and Medford, 1999) or STATISTICA (Version 6.0) (StatSoft, 2001) was used for all statistical analyses.

2.5.2. Classification of reference samples

The Sørensen (Bray–Curtis) similarity index was used to measure the compositional similarity between all pairs of reference site samples. Rare taxa, defined as those taxa that were collected at 10 or fewer sites within the reference dataset, were excluded from the classification analysis. A flexible hierarchical unweighted pair-group average (UPGMA) agglomerative clustering method with $\beta = -0.5$ was then used to cluster samples based on these similarities.

2.5.3. Prediction of class membership

Once we identified groups of reference sites, stepwise multiple discriminant function analysis (DFA) was used to determine which environmental variables were most strongly associated with group membership. Candidate predictor variables included: stream origin, geological chemical activity, geological physical activity, geological nutrient activity, primary contributing bedrock geology, date (day of year) of collection, water chemistry variables (alkalinity, conductivity, hardness, and sulfate), level III ecoregions and sub-regions, latitude, longitude, elevation, watershed area, landscape type, percent substrate type (cobble, coarse gravel, fine gravel, sand, silt/clay), percent coarse substrate, percent fine substrate, and velocity. Final selection of variables for inclusion in the discriminant model was based on the results of both forward and backward DFA analysis, ease of variable measurement, and ease of ecological interpretation. The final discriminant model was used to estimate the probability that a new site belonged to each of the biotically defined classes.

2.5.4. Estimating probabilities of capture, E , and O/E

We estimated site-specific probabilities of capture (pc) as the frequencies of occurrence of taxa observed within each reference site group weighted by the DFM-derived probabilities that a site was a member of each class. The number of pc -values ≥ 0.5 were summed to estimate the number of expected (E) taxa in each sample. Use of $pc \geq 0.5$ results in a more precise index than $pc \geq 0$ (Hawkins et al., 2000; Ostermiller and Hawkins, 2004; Van Sickle et al., 2005). O/E was estimated as the ratio of the observed number of predicted taxa (O) to E .

2.6. Model validation and responsiveness

Following methods described by Hawkins et al. (2000) and Van Sickle et al. (2005), we evaluated the accuracy and precision of the predictive model by applying it to both the reference calibration dataset in addition to the independent reference validation dataset. We evaluated model accuracy and precision by estimating the standard deviation of O/E -values in calibration and validation data sets and by determining how well O was correlated with E . The slope of the regression of O on E should be near 1 for RIVPACS models with good accuracy (Hawkins et al., 2000; Van Sickle et al., 2005). Likewise, the scatter of points along the regression line (evaluated as the coefficient of determination, r^2) should be small for a model with good precision. Hawkins (2006), Hawkins et al. (2000) and Van Sickle et al. (2005) found that the best performing models generally have reference site O/E standard deviations <0.20 and account for a significant amount of the variation in O among reference sites.

To further evaluate the overall precision of the predictive model, we constructed a null model and estimated the standard deviation expected for replicate reference site samples following the methods of Van Sickle et al. (2005). Together, the null model and replicate sample standard deviations estimate the minimum and maximum levels of precision, respectively, that is theoretically attainable by any RIVPACS model given the reference dataset (Van Sickle et al., 2005). Comparison of the predictive model's precision with these minimum and maximum levels of precision provided context regarding how well the model performed relative to its possible performance. We also examined model bias by testing (ANOVA, $\alpha = 0.05$) whether reference site O/E -values were associated with ecoregions/sub-regions and landscape setting. We used Tukey multiple comparison tests to identify sets of ecoregions and landscape types that differed in mean reference site O/E -values.

2.7. Application of the model to reference and test sites

Based on protocols described by Clarke et al. (1996), Hawkins et al. (2000), and Moss et al.

(1987), all repeat reference and test sites were evaluated as to whether they were within the experience of the model. These procedures determine if the Mahalanobis squared distances between a test site and each classification group in multivariate discriminant space are greater than expected as measured by a χ^2 -test. Repeat reference or test samples that failed this test were not considered in the evaluation of model performance. For those test samples evaluated with the model, we then used ANOVA ($\alpha = 0.05$) followed by Tukey multiple comparison tests to determine whether O/E -values differed among ecoregions/sub-regions and landscapes as described above. We used the 10th and 90th percentiles of reference calibration O/E -values to establish thresholds to infer whether a test site was significantly different from reference and thus biologically impaired (Clarke et al., 1996; Ostermiller and Hawkins, 2004). The parametric two-sample t -test was used to determine if significant differences existed between (1) mean calibration and validation O/E -values, (2) mean reference, repeat reference, and test site O/E -values, (3) mean O/E -values at calibration and known degraded test sites from the 1993 to 1999 dataset, and (4) mean values of continuous predictor variables for sites that fell either outside or within model experience. We used scatter-plots to visualize how those sites that fell outside model experience differed from the population of reference sites in predictor variable values. To examine inter-annual variability in O/E -values, we calculated the mean coefficient of variation (CV) for repeat reference samples.

3. Results

3.1. Operational taxonomic units

Two-hundred and nineteen operational taxonomic units (OTUs) were identified from the reference calibration dataset (Appendix A). Ninety-six of these taxa occurred at 10 or more calibration sites and were used to create the biotic classification. Among all 219 OTUs, 25%, 21%, 12%, 11%, 11%, 10%, and 6% of OTUs were allocated among the Chironomidae, Trichoptera, Diptera, Ephemeroptera, non-insects, Plecoptera, and Coleoptera, respectively.

Table 1
Mean and range of values for selected environmental variables at reference and test sites

	Reference ^a (n = 296)		Test (n = 629)	
	Mean	Range	Mean	Range
Julian sample data	266.7	217–305	267.5	92–305
Alkalinity (mg/L)	104.5	10–310	158.9	5–1,630
Conductivity ($\mu\text{S}/\text{cm}$)	312.5	23–2420	838.6	8–21,700
Hardness as CaCO_3 (mg/L)	157.2	8–1455	276.7	11–2,680
Sulfate (mg/L)	81.5	9–2037	257.7	10–8,678
Elevation (m)	2031	1091–2798	1834	1,024–2,930
pH	8	6–9	8	5–10.5
Chloride (mg/L)	6.5	<5–89	53.3	<5–6,685
Turbidity (NTU)	2.9	<0.1–58.2	7.6	0–425
Nitrate-nitrogen (mg/L)	0.2	<0.1–4.5	0.2	<0.1–12.8
Total phosphorous (mg/L)	<0.1	<0.1–0.3	0.1	<0.1–4.0
Water temperature ($^{\circ}\text{C}$)	9.6	2.1–19	10.5	0–37.5
Watershed area (km^2)	450.1	0.8–7620	771.6	0.8–11,103
Percent located in the foothills	26.1	–	23.9	–
Percent located in the mountains	54.8	–	35.1	–
Percent located in the plains	19.1	–	41	–
Percent coarse gravel	15.8	0–78	19.3	0–91
Percent cobble	58.4	0–95	45.1	0–95
Percent fine gravel	11.6	0–59	16.9	0–83
Percent sand	10.1	0–98	11.8	0–100
Percent silt	3.6	0–99	5.2	0–100
Percent clay	0.6	0–46	1.4	0–84
Percent coarse substrate	85.8	0–100	81.3	0–100
Percent fine substrate	14.3	0–100	18.4	0–100
Velocity (m/s)	0.5	0.1–1.0	0.5	0.5–3.3
Discharge (m^3/s)	1.2	<0.1–14.2	1.2	<0.1–531

^a Includes calibration, validation and repeat reference samples.

3.2. Reference and test sample characteristics

Considerable environmental variation occurred among both reference and test sites (Table 1). Reference and test sites generally occurred at similar elevations, were well buffered, alkaline, and sampled within comparable dates. However, we found differences in substrate composition between reference and test sites. These differences were minimal when individual substrate composition categories were compiled as either coarse or fine substrate. Overall, test sites had higher conductivity, greater watershed area, lower elevation, and elevated levels of hardness, alkalinity, sulfate, total phosphorous, nitrate-nitrogen, and turbidity relative to reference sites. Reference sites were distributed throughout most perennial lotic systems within the ecoregion/sub-regions of Wyoming (Fig. 1). The distribution of test sites among ecoregions/sub-regions in Wyoming was not similar

to the distribution of reference sites primarily because a greater proportion of test sites were located in the Northwestern Great Plains and Wyoming Basin ecoregions (Table 4 and Fig. 1). Reference and test sites were absent from large areas of the southern Northwestern Great Plains and central and southern portions of the Wyoming Basin because most streams in these regions are ephemeral to intermittent.

3.3. The predictive model

3.3.1. Reference site classification

We derived 15 reference groups from the cluster analysis of the calibration samples (Fig. 2). Percent chaining of the cluster analysis was low (0.52). All reference groups contained a minimum of five reference samples. Generally, samples from upper-elevation montane regions (groups 1–11) were distinct from those from the plains and the low-elevation

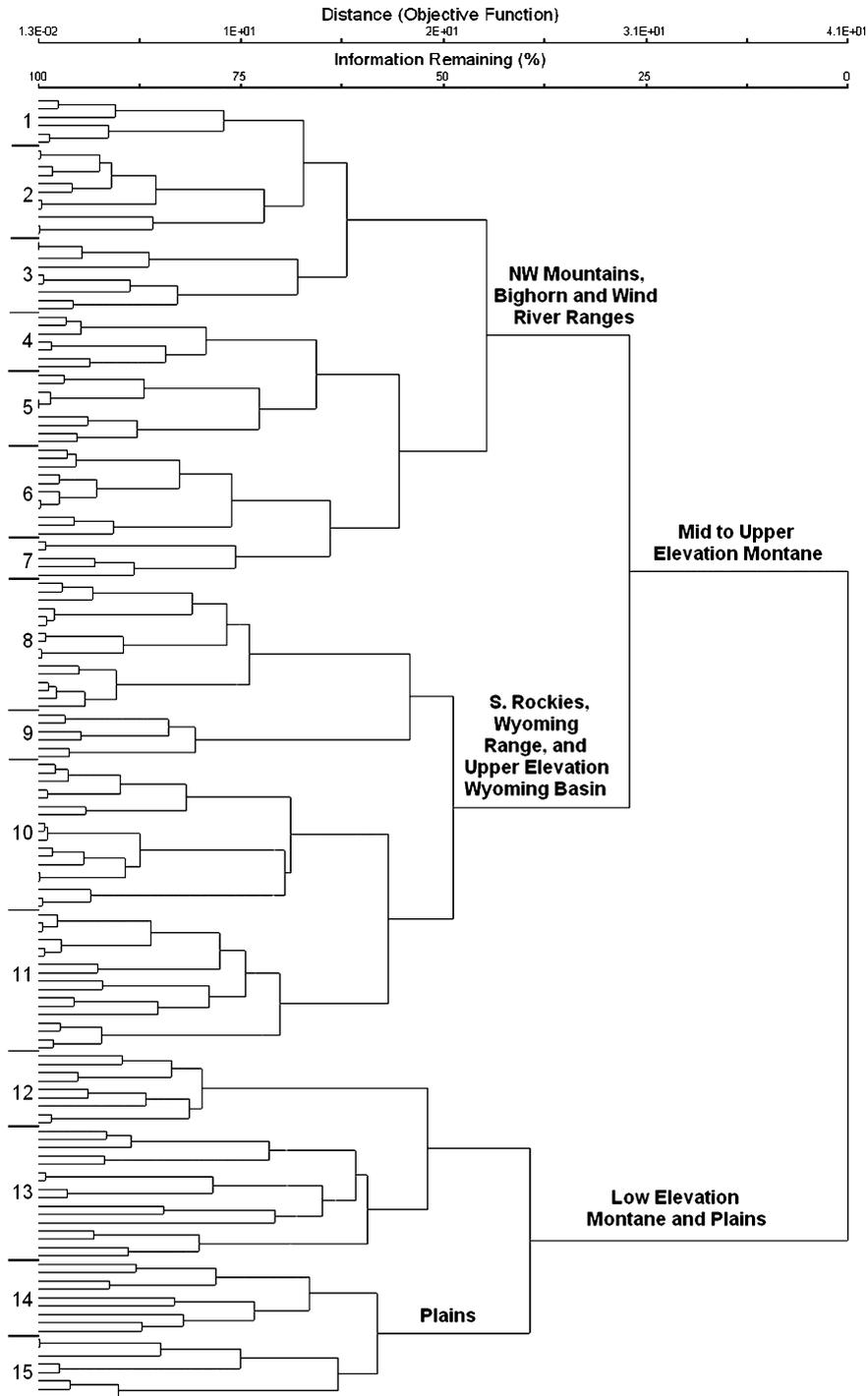


Fig. 2. UPGMA flexible beta cluster analysis of calibration samples showing the 15 biologically defined groups that were used to develop the RIVPACS model.

mountains of northeastern Wyoming (groups 12–15). Reference site samples from the Northwestern Great Plains and Western High Plains ecoregions tended to fall in either group 14 or 15. Samples from streams in the mid to upper-elevations of the Absaroka, Gallatin and Beartooth mountains of northwestern Wyoming clustered into groups 1, 2, 3, or 9. The majority of samples from streams in the sub-alpine regions of the Bighorn Mountains clustered together to form group 6. Regional structuring of macroinvertebrate communities was less pronounced among assemblage groups 5, 7, 8, 10, and 11. Samples from these groups were scattered among the mountain ranges of the Southern Rockies ecoregion, the Wyoming, Salt River, and Wind River ranges of western Wyoming, and foothills along the eastern slope of the Bighorn Mountains.

3.3.2. Predictor variables

Fourteen predictor variables were selected that best accounted for the natural variability in taxonomic composition among the 15 reference groups (Table 2). Those variables most important in discriminating between groups were log watershed area; log percent coarse substrate; dummy values (0/1) for Western High Plains, Middle Rockies West, Middle Rockies Central, and Northwestern Great Plains ecoregions; latitude; and longitude. Other factors were less important in distinguishing groups.

3.3.3. Predictive model validation

O/E -values for both the calibration ($\chi^2 = 11.82$, $P = 0.06$) and validation ($\chi^2 = 2.08$, $P = 0.05$) datasets

Table 2

Predictor variables (with corresponding F -values) used in the discriminant model

Log watershed area	8.85
Log percent coarse substrate	6.01
Western High Plains (WHP)	6.00
Middle Rockies West (MRW)	4.98
Latitude	4.80
Middle Rockies Central (MRC)	4.54
Northwestern Great Plains (NGP)	3.35
Longitude	3.03
Geological chemical activity (majority in watershed)	2.37
Early Archaen gneiss bedrock geology	2.17
Log alkalinity	2.10
Log elevation	1.92
Middle Rockies East (MRE)	1.79
Geological nutrient activity (majority in watershed)	1.39

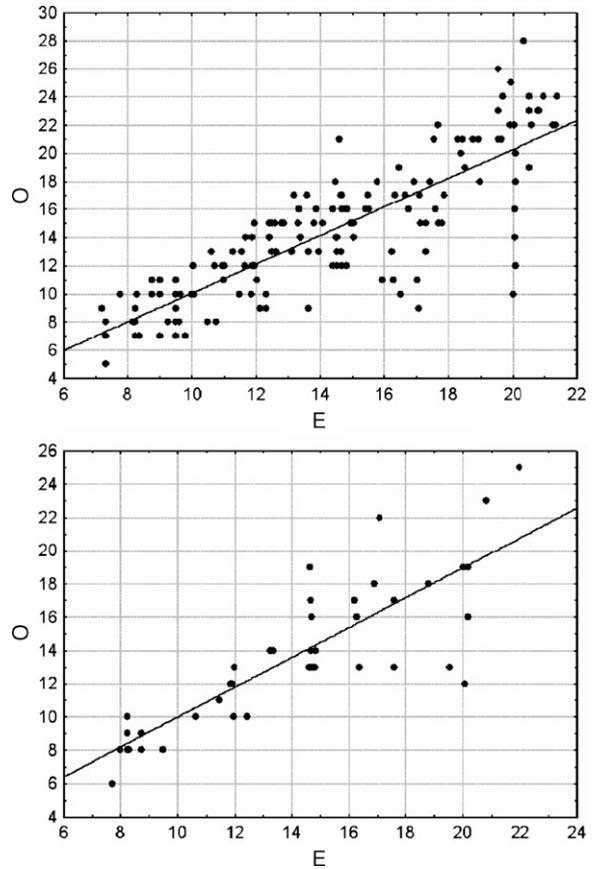


Fig. 3. Relationship between observed (O) and expected (E) OTU richness for calibration (top) and validation (bottom) samples.

did not show any statistically significant departure from normal distributions based on goodness of fit tests. Predicted values of E were similar to values of O for both calibration ($r^2 = 0.70$, $P < 0.001$) and validation ($r^2 = 0.69$, $P < 0.001$) datasets (Fig. 3). The mean and standard deviation of calibration O/E -values were 1.01 and 0.174, respectively (Table 3). The mean and standard deviation for validation O/E -values were similar to calibration values, 0.98 and 0.152, respectively (Table 3). Mean O/E -values for calibration and validation datasets were not significantly different from one another ($t = 0.789$, $P = 0.43$) (Fig. 4). There was no evidence that calibration O/E -values varied among classification groups ($F = 1.391$, $P = 0.165$), indicating the model was unbiased and predicted similarly among reference groups. In contrast, calibration O/E -values were significantly

Table 3

Means, standard deviations (S.D.), percentiles, and minimum and maximum *O/E*-values for reference and test samples

	Mean	S.D.	10th	25th	75th	90th	Minimum	Maximum
Reference calibration ($n = 157$)	1.01	0.174	0.75	0.90	1.13	1.21	0.50	1.44
Reference validation ($n = 46$)	0.98	0.152	0.79	0.89	1.06	1.16	0.60	1.30
Repeat reference samples ($n = 93$)	0.98	0.187	0.74	0.88	1.12	1.19	0.41	1.37
Test ($n = 509$)	0.72	0.281	0.30	0.54	0.93	1.08	0.00	1.45

Includes all repeat reference samples and only those test samples that were within the experience of the model.

different ($F = 2.527$, $P = 0.02$) among ecoregion/sub-regions (Table 4). The Tukey multiple comparison test showed that calibration *O/E*-values were significantly lower in the Western High Plains (mean = 0.90) relative to the other regions, though not significantly less than 1 ($t = -1.289$, $P = 0.25$). These results imply that the model may over-predict *E* in the WHP,

although the evidence is not conclusive. Mean calibration *O/E*-values among landscape types (i.e., mountains, foothills, and plains) were similar ($F = 0.269$, $P = 0.77$) (Table 4). The standard deviation of calibration *O/E*-values (0.174) was moderately larger than the replicate sample standard deviation of 0.140 (theoretical best model) but appreciably less than that for the null model (0.289). These results indicate that the predictive model accounted for a large proportion of the natural variability in *O* within the reference dataset.

3.3.4. *O/E*-values of repeat reference samples

Mean *O/E*-values for repeat reference samples (0.98) were no different from that for calibration samples (Fig. 4, $t = 1.31$, $P = 0.19$). The CV of *O/E*-values for repeat reference samples ranged from 0.05 to 45.65 with a mean of 11.9, implying that sampling error was somewhat less than that implied by the estimated replicate sample standard deviation (S.D. = CV/100 if mean is 1).

3.3.5. *O/E*-values of test sites

Mean test site *O/E*-values (0.73) were significantly lower than calibration *O/E*-values ($t = 11.95$, $P < 0.001$) (Fig. 4). Mean *O/E*-values (0.53) for known degraded test sites from the 1993 to 1999 record were substantially lower than mean values (1.01) for calibration sites ($t = 11.73$, $P < 0.001$) (Fig. 4). Mean test sample *O/E*-values varied by 0.19 *O/E* units among ecoregion/sub-regions ($F = 6.43$, $P < 0.001$, Table 4), with larger *O/E*-values occurring in streams in the mountainous sub-regions of the Middle Rockies West and Middle Rockies Central than in streams in the Middle Rockies East, Southern Rockies and basin and plains regions (Table 4). Mean test site *O/E*-values varied less markedly among landscape types (0.14 *O/E* units) than ecoregions, but mean values from streams in plains landscapes (0.72)

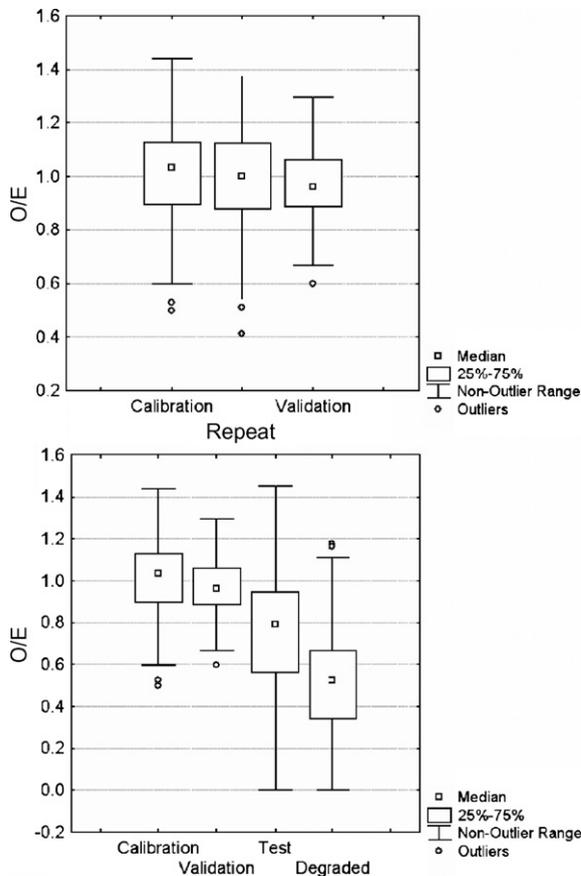


Fig. 4. Box and whisker plots of *O/E*-values for different sets of reference samples (top) and reference and non-reference samples (bottom).

Table 4
Mean *O/E*-values, *P*-values, and *F*-values for comparisons between ecoregion/sub-region and landscape types for both reference calibration and test samples (significant values are italicized)

	Landscape			Ecoregion/sub-region							<i>F</i>	<i>P</i>		
	Mountains	Foothills	Plains	<i>P</i>	MRC	MRE	MRW	SR	WB	NGP			WHP	
Reference calibration	1.02 (86)	1.00 (41)	1.00 (30)	0.269	0.770	0.93 (33)	1.00 (9)	1.05 (57)	1.04 (28)	0.99 (14)	1.08 (10)	0.90 (6)	2.527	0.020
	A	A	A	-	-	A	A	A	A	A	A	B	-	-
Test	0.86 (176)	0.81 (130)	0.72 (203)	15.22	<0.001	0.83 (125)	0.77 (16)	0.89 (136)	0.70 (22)	0.70 (32)	0.76 (146)	0.72 (32)	6.43	<0.001
	A	A	B	-	-	A	B	A	B	B	B	B	-	-

Numbers of samples for each landscape type and bioregion are noted in parentheses. Capital letters denote homogeneous groups as determined from the Tukey multiple comparison tests.

MRC: Middle Rockies Central, MRE: Middle Rockies East, MRW: Middle Rockies West, SR: Southern Rockies, NGP: Northwestern Great Plains, WB: Wyoming Basin and WHP: Western High Plains.

were significantly lower ($F = 15.22, P < 0.001$) than those from mountain (0.86) and foothill (0.81) regions (Table 4). Of the 509 test site samples within the model’s experience, 62% ($n = 317$) fell outside the 10th (0.75) and 90th (1.21) percentiles of reference calibration *O/E*-values. Of these 317 impaired test site samples, 309 had *O/E*-values below the 10th percentile and only 8 samples were above the 90th percentile. A sizeable proportion of test site *O/E*-values in the plains (59%, $n = 203$) and foothills (46%, $n = 130$) were outside the central 80% of reference calibration *O/E*-values. In the mountains, only 28% ($n = 176$) of test site *O/E*-values were considered significantly different from reference.

3.3.6. Sites outside the experience of the model

Sites associated with 120 (16.6%) of the 722 combined repeat reference and test samples fell outside the experience of the model and could not be assessed. All 120 samples that fell outside model experience were test samples. The majority of these test sites occurred in four regions of the state: the Bear River basin, lower Big Horn River basin, eastern foothills of the Bighorn Mountains, and upper Powder River basin (Fig. 5). On average, sites that fell outside of the experience of the model differed in elevation (mean = 1446 m, $t = 8.51, P < 0.001$), percent coarse substrate (mean = 37.2, $t = 10.36, P < 0.001$), and alkalinity (mean = 229 mg/L, $t = 8.78, P < 0.001$). Several of these sites that fell outside the cluster of reference sites defined in predictor variable space were largely low-elevation streams with higher alkalinities and higher percentages of fine sediment (Fig. 6).

4. Discussion

Benthic macroinvertebrates are a particularly effective indicator of the condition of lotic systems and are thus commonly used in bioassessments of water quality conditions (Barbour et al., 1999; Plafkin et al., 1989). Effective assessments, however, must account for natural heterogeneity in assemblage structure, something that can be strongly influenced by both abiotic and biotic factors that are manifest at both large (e.g., regional) and local-scales (Allan, 1995). In Wyoming, the diversity of assemblage structure was evident in the 15 biotically defined

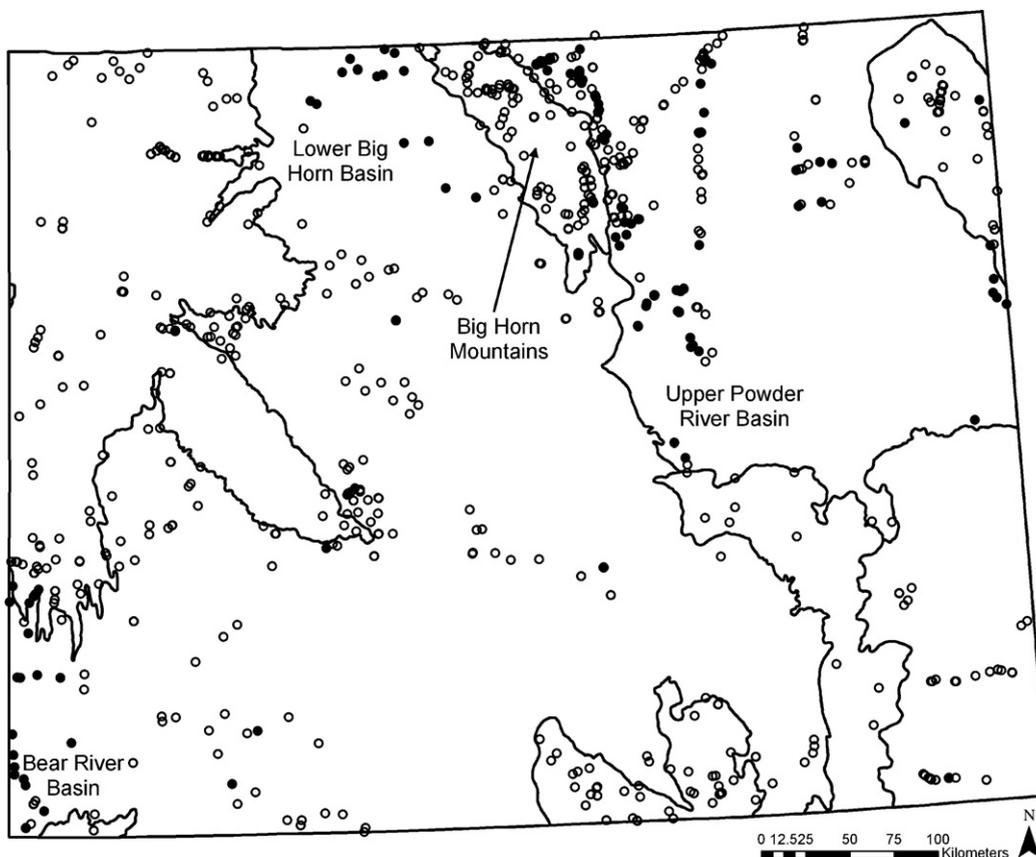


Fig. 5. Location of samples that were outside (solid circles) and within (open circles) model experience.

groups derived from the reference dataset. This high diversity in macroinvertebrate assemblages in Wyoming appears to be structured, in large part, by the high environmental heterogeneity within the state which our model largely captured through the use of 14 predictor variables.

In our model, the strongest predictor variables were those that described the substrate composition of the streambed (percent coarse substrate) and the area of the drainage basin upstream from a sample location. These results are consistent with a large literature that shows substrate type is a major determinant of the distribution and abundance of aquatic insects (review by Minshall, 1984) and that stream biota exhibit strong longitudinal ‘zonation’ along river networks in response to changes in physical and chemical factors with distance downstream (review by Hynes, 1970). The significance of the ecoregion/sub-region predic-

tors suggest that differences in stream macroinvertebrate assemblages across Wyoming are tied to some extent to broad-scale abiotic or biotic factors acting upon streams, which may include stream geomorphology, vegetation, climate, soils, hydrology, and biogeography. We observed that several of the 15 reference groups were at least partially associated with particular ecoregions/sub-regions and inclusion of these predictors increased our model precision. Use of these predictors allowed us to better characterize factors that influence assemblage distribution and perhaps capture region-specific expected values into a single predictive model. The use of latitude, longitude, and elevation by the model implies that temperature has a large role in structuring aquatic assemblages, an inference consistent with the ideas and observations of Hawkins et al. (1997, 2000), Vannote and Sweeney (1980) and Vannote et al. (1980) among others.

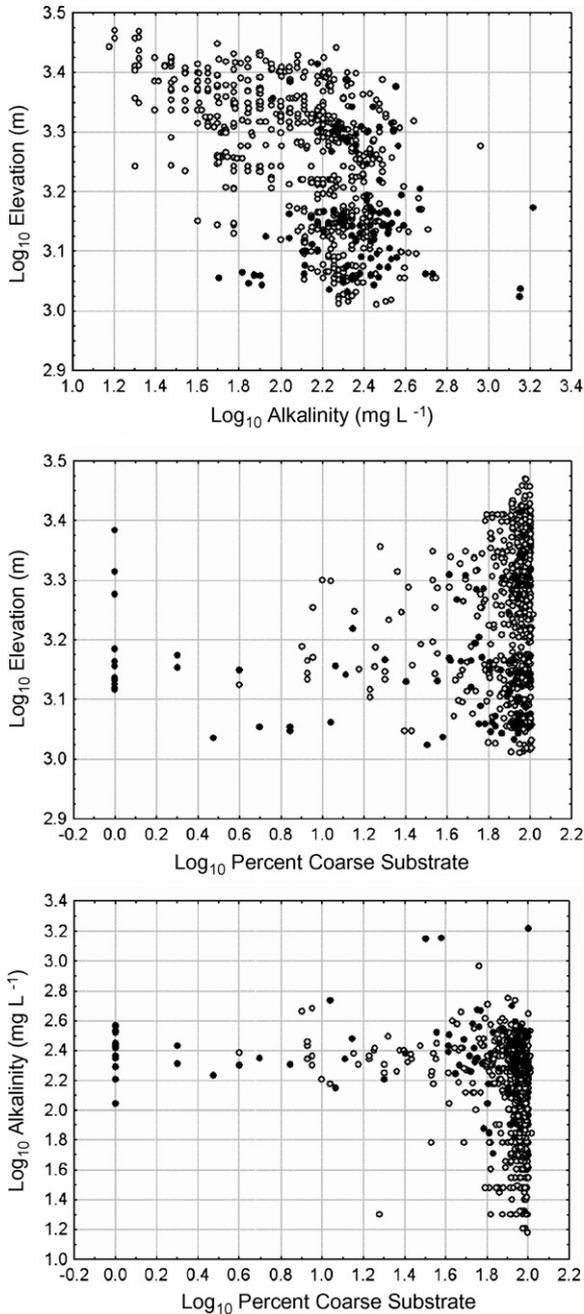


Fig. 6. Relationships between alkalinity, percent coarse substrate and elevation for sites within (open circles) and outside (solid circles) model experience.

Geographic location, watershed area, and elevation have been shown to be important predictor variables for a number of RIVPACS models developed in the United States (Hawkins, 2006; Hawkins et al., 2000). Such commonality suggests that these variables should be considered for development of any RIVPACS model. The use of several geology related predictors and alkalinity demonstrates that geologic setting probably has a strong effect on invertebrate distributions by affecting both physical and chemical aspects of the stream environment. Early Archaean gneiss bedrock geology appears to be a particularly important predictor, a formation that is common in some mountain ranges of northwestern and northcentral Wyoming.

4.1. Model performance and validation

In general, the Wyoming predictive model was accurate and precise enough to detect modest degradation. However, a small amount of systematic bias occurred associated with ecoregion setting (Table 4). For such a systematic bias to occur, the model had to over-predict for some streams and under-predict for others. We cannot fully explain this bias at this time, but because the model is statistical in nature, the average *O/E*-value for calibration samples must be approximately one, and any errors of under-prediction will be balanced by errors of over-prediction. The apparent over-prediction that occurred in the Western High Plains may be due to two factors: (1) incomplete representation of the macroinvertebrate assemblage in the Western High Plains from only six reference sites and (2) classification of these six sites into more than one reference group. Several of our reference groups were associated with a particular ecoregion/sub-region, including the Western High Plains. Reference sites in the Western High Plains were equivalent in site quality and generally associated with one reference group, however, two of the six reference sites were classified into a neighboring reference group. We suspect that the combined effects of these two factors limited the discriminatory power of the Western High Plains predictor since it did not entirely account for macroinvertebrate assemblage variation within the ecoregion and thus resulted in the systematic bias for sites within this ecoregion.

The fact that the mean and standard deviation of *O/E*-values for samples in our calibration and validation

datasets were similar implies that our model should accurately predict expected biological condition for new datasets. These results also indicate that there is no evidence to suggest that the model is overfit. In other words, the model did not spuriously associate random variation in biological structure with variation in predictor variables, something that could result in either imprecise or biased predictions.

4.2. Temporal variability in reference condition

Ideally, a RIVPACS model based on one-time sampling should be able to effectively account for both spatial and temporal variation in biotic structure (Hawkins, 2006; Hawkins et al., 2000). Temporal variability in macroinvertebrate assemblage structure is associated not only with life-history traits such as timing of emergence and dispersal, but also temporal variation in both environmental variables and measurement error. Such temporal variation in either biotic processes or environmental setting could influence the prediction of E and hence the estimate of O/E -values. However, temporal variability in biological condition, as measured by O/E , was relatively constant through time as indicated by the low mean CV of 12%. We can infer from these results that the Wyoming model was reasonably temporally robust in spite of inter-annual changes in macroinvertebrate assemblages, environmental factors, and crew sampling efficiencies.

4.3. Detection of anthropogenic disturbance

Because the model was able to predict invertebrate composition, and hence richness, across a broad range of naturally differing stream conditions, any deviation between observed and predicted assemblages beyond model error should be interpreted as biological degradation of a site associated with some stress, i.e., habitat alteration or direct toxicity by pollutants. Test sites with high O/E -values were primarily distributed among Wyoming's mountainous regions where the potential for human disturbance is minimal due to rough terrain and the limited number of population centers. Areas with high densities of test sites with appreciably low O/E -values were confined to the plains and basin regions, particularly in southwest, central, and northeast Wyoming. Predominant human activities that have the

potential to either directly or indirectly alter the biota of stream ecosystems in these areas include hydrologic modifications, point and non-point source pollutants, and stream alterations such as channelization, excessive aggradation or degradation of sediment in the stream channel, and urbanization.

Test site O/E -values that fell outside the 10th and 90th percentiles of the reference calibration O/E -values imply some degree of biological impairment caused by anthropogenic disturbances. The vast majority of impaired test site O/E -values were below the 10th percentile (0.75) of reference calibration O/E scores suggesting that at least 25% of expected taxa were absent from samples taken at these sites. The fact that impairment was most prevalent (for the period of record) for streams in the plains (59% of test sites) and foothills (46% of test sites) comes as no surprise given that the majority of human development (and their influences on aquatic systems) and the major population centers of Wyoming occur in low-land grassland/sagebrush steppe and foothill environments. Conversely, streams in the less altered mountainous (28% of test sites) landscapes of Wyoming were less degraded. Collectively, these findings suggest that different levels of biological degradation occur within Wyoming. Implementation of best management practices, pollutant removal, and/or broad-scale watershed improvement projects in Wyoming since the period of record analyzed here may have improved biological condition in some streams, but we have not yet been able to address that question.

4.4. Model limits

RIVPACS assessment software is typically programmed to flag sites that fall outside the experience of the model and thus prevent extrapolation of predictions to environmental settings beyond those used in model development. In general, poor or incomplete representation by reference sites of the range of naturally occurring conditions in the region of interest will result in data gaps that could lead to unreliable estimates of E at some sites. Flagging these problematic sites is therefore critical to help eliminate biases in site bioassessments. However, the use of dummy variables for categorical predictors (e.g., ecoregion) may cause problems with the use of the Mahalanobis squared distance. Unless reference sites

in each ecoregion used in predictions have a similar range in the values of the continuous predictors used (latitude, longitude, watershed area, elevation, etc.), a test site could be flagged as having a combination of conditions not observed in the reference data. This issue appears to be of particular concern in the WY model, which uses dummy variables for six categorical variables. Further analysis of the use and effects of categorical predictors in RIVPACS-type models is needed. These findings imply that when a site is flagged as being outside model experience, it is important to thoroughly evaluate whether predictor variable values are truly outside the expected range of reference conditions before excluding the accuracy of the result. Of the 120 samples that were flagged as having predictor variable values outside the experience of the model, only 27 sites were flagged because of clear extrapolations (Fig. 6). Conditions at these 27 sites are probably most indicative of general limitations in the model.

In general, unsuccessful model predictions at these 27 test sites appeared to be largely the result of percent coarse substrate and/or log alkalinity values that fell outside the range of conditions found at reference sites (Fig. 6). Many of these sites were located in low- to moderate-elevation regions of the state where natural substrate composition is likely characterized by moderate to high percentages of sand, silt, or organic material and alkaline water chemistry associated with sedimentary watershed geology. A handful of sites that fell outside model experience were characterized by coarse substrates and alkaline waters. Considering that very few reference sites used in the development of this model were collected from sites with these combinations of characteristics, it is understandable that the model could not predict assemblage composition at many sites with these general characteristics. This limitation in the model can only be remedied by inclusion of additional reference sites that expand the range of coverage for specific predictor variables or by the substitution of appropriate surrogate predictors for the ones currently in use.

Although most of the flagged sites appeared to fall outside model experience because of naturally high fine sediment or alkalinities, the alkalinities and fine sediment levels observed at a handful of sites were probably altered by human activity, and thus would confound assessments. For example, although alka-

linity in streams of Wyoming is predominantly a function of watershed geology (and hence natural), there were a few instances where alkalinities were probably elevated above natural ambient conditions by point or non-point source pollution. In such cases the model would either flag the sample if the alkalinity value was outside the range observed at reference sites or predict a fauna more associated with altered conditions than natural ones. A similar situation could occur if sediment composition were altered. At such sites, additional data should be used to evaluate the validity of any *O/E*-values obtained. In general, future revisions of the model should avoid the use of direct measures of alkalinity and substrate composition.

Although we thoroughly examined the suite of environmental variables in our dataset to ultimately choose the 14 predictor variables that most effectively discriminated among all reference groups, those variables did not account for 100% of the natural variability within the reference dataset. Clarke et al. (1996) suggested that errors in the predictive capability of a RIVPACS model may be the result of not including all necessary environmental variables to effectively predict the correct biological classes. Aspect, channel morphological parameters and climatic data may prove to be useful predictors and enhance the accuracy, precision, and spatial applicability in future revisions of our predictive model.

4.5. Model limitations associated with effluent dependent systems

Over a century of mineral development has altered the lotic ecology of streams in several watersheds of Wyoming, transforming what were once ephemeral or intermittent drainages into perennial systems fed almost exclusively by production water. In many of these cases, there can be a net environmental benefit from adding flow to an otherwise ephemeral or intermittent stream through increases in macroinvertebrate assemblage diversity and structure. Nonetheless, achievable biological condition may be less than would be expected for naturally perennial streams in the region due to concomitant alterations of channel and water quality conditions from effluent flows. Because of their unique characteristics, the sensitivity and general applicability of our model to effluent dependent systems is contingent on the

appropriateness of the model's predicted biological condition. Our model is unable to detect whether an effluent dependent stream was formerly ephemeral or intermittent and therefore makes a prediction based on the assumption that the stream is naturally perennial. Several of our test sites in the Upper Powder River Basin that fell outside model experience, were effluent dependent and likely possessed perennial stream characteristics that were unlike those expected by the model. For such streams, case-by-case decisions must be made on whether the network of reference sites used to build a model provides an adequate expectation of biological condition.

4.6. Comparison with other models and geographic predictor variables

Several RIVPACS models have been developed in the United States for the states of California, Colorado, Maine, North Carolina, Ohio, Oregon, Washington and the Mid-Atlantic Highlands region (e.g., Hawkins, 2006; Hawkins et al., 2000). The precision (standard deviation) of these models generally ranged from about 0.10 to 0.30. Those models that were most precise and most effective in detecting anthropogenic disturbance were developed from high resolution taxonomic data (genus and/or species) collected from natural substrates at reference sites of consistent quality. Considering that the Wyoming model was developed with genus and species level data collected from both minimally disturbed and least-disturbed reference sites, the overall precision of our model appears to be clearly comparable with other successful predictive models.

Most of these other predictive models used fewer predictor variables than in the Wyoming model. One reason that the Wyoming model may have needed more predictor variables is that Wyoming is an extraordinarily heterogeneous state both in terms of its physical environment (topography, geology and climate) and its invertebrate biota (elements of western and eastern faunas). This heterogeneity notwithstanding, most other models used many of the same predictor variables used in the Wyoming model, most notably elevation, latitude, longitude, and watershed area. Not only do these factors provide insight into attributes that structure regional stream faunas, but their common use across regions underscores their likely universal

importance. Such variables should probably be considered when developing RIVPACS models anywhere in the United States or elsewhere.

4.7. Utility of map/GIS variables in RIVPACS models

Our results present strong evidence that RIVPACS models can be developed based largely on map-level predictor variables. Out of the 14 predictor variables used in the Wyoming model, twelve were GIS-derived. For a variety of reasons, the model development team at Utah State University is now routinely developing RIVPACS models based on only GIS-derived predictor variables. Use of map-level predictors not only make RIVPACS model development a more inviting endeavor for managers in terms of cost and time for data collection, but GIS provides a readily available source of candidate predictor variables. GIS data is easily obtained from both government and non-profit spatial data clearing houses. These data can be used as surrogates for field collected variables that may be important in influencing macroinvertebrate assemblages but which are influenced by human disturbance. For example, sulfate bearing geological formations and their effects on stream chemistry and stream biota has been well documented in the literature. However, human activities such as permitted discharges from gas and oil production facilities hampered the use of this variable in the Wyoming model. Instead, nutrient content and the chemical weathering rate of geologic material in the watershed (both derived with GIS) served as adequate surrogates for sulfate. Percent coarse substrate was an important predictor variable in the Wyoming model, though human activities may influence site values of this variable in some streams, thereby affecting the predictive accuracy of our model. Predictions of potential substrate based on our understanding of the geomorphic controls of sediment size may allow us to avoid use of direct measures of substrate composition in many models. Buffington et al. (2004) were able to make predictions of sediment size based on DEM-derived channel slopes and catchment area/depth relationships. Refining these models to include other controls on substrate, such as valley width (R. Hill, Utah State University, unpublished data), should further improve our predictions of substrate, and hence biota.

In conclusion, the Wyoming RIVPACS model provides a clear, simple and defensible measure of biological condition of streams in a diverse landscape. As such, the model should greatly assist in helping Wyoming water quality managers establish priorities for stream management and develop ecologically realistic and meaningful restoration efforts and watershed protection plans.

Acknowledgments

We thank Dr. John Van Sickle for recommendations and assistance in the development of the model and Dr. Ralph Clarke and an anonymous referee for their constructive reviews. We also thank the numerous staff of the Wyoming Department of Environmental Quality-Water Quality Division who provided useful information in development of the model or conducted the field work, data entry, and/or quality assurance and quality control. This study was funded in part by the USEPA Region 8 (contract X988766-01). CPH and JRO were supported by EPA Science To Achieve Results (STAR) grant R-82863701 and Cooperative agreement X7-97830101 with Region 8 USEPA.

Appendix A

Operational taxonomic units (OTUs) derived from the reference calibration dataset. Asterisks denote the 96 OTUs that occurred at 10 or more sites.

Taxa group	Operational taxonomic unit
Acari	<i>Acari</i> *
Ephemeroptera	<i>Acentrella insignificans</i> *
Ephemeroptera	<i>Acentrella turbida</i> *
Ephemeroptera	<i>Ameletus</i> *
Diptera	<i>Antocha</i> *
Trichoptera	<i>Arctopsyche grandis</i> *
Diptera	<i>Atherix</i> *
Ephemeroptera	<i>Baetis</i> *
Trichoptera	<i>Brachycentrus americanus</i> *
Trichoptera	<i>Brachycentrus occidentalis</i> *
Chironomidae	<i>Brillia</i> *
Plecoptera	Capniidae*
Diptera	Ceratopogoninae*
Chironomidae	<i>Chaetocladius</i> *
Diptera	<i>Chelifera</i> *
Trichoptera	<i>Cheumatopsyche</i> *
Plecoptera	Chloroperlidae*
Ephemeroptera	<i>Cinygmula</i> *

Appendix A (Continued)

Taxa group	Operational taxonomic unit
Plecoptera	<i>Classenia sabulosa</i> *
Coleoptera	<i>Cleptelmis</i> *
Chironomidae	<i>Cricotopus</i> *
Chironomidae	<i>Diamesa</i> *
Diptera	<i>Dicranota</i> *
Ephemeroptera	<i>Diphetero hageni</i> *
Trichoptera	<i>Dolophilodes</i> *
Plecoptera	<i>Doroneuria</i> *
Ephemeroptera	<i>Drunella coloradoensis/flavilinea</i> *
Ephemeroptera	<i>Drunella doddsi</i> *
Ephemeroptera	<i>Drunella grandis/spinifera</i> *
Coleoptera	<i>Dubiraphia</i> *
Coleoptera	Dytiscidae*
Trichoptera	<i>Ecclisomyia</i> *
Oligochaeta	Enchytraeidae*
Ephemeroptera	<i>Epeorus</i> *
Ephemeroptera	<i>Ephemerella</i> *
Chironomidae	<i>Eukiefferiella</i> *
Trichoptera	<i>Glossosoma</i> *
Trichoptera	<i>Helicopsyche</i> *
Diptera	<i>Hemerodromia</i> *
Plecoptera	<i>Hesperoperla</i> *
Coleoptera	<i>Heterlimnius</i> *
Ephemeroptera	<i>Hexatoma</i> *
Amphipoda	<i>Hyalella azteca</i> *
Trichoptera	<i>Hydropsyche</i> *
Trichoptera	<i>Hydroptila</i> *
Plecoptera	<i>Isoperla</i> *
Trichoptera	<i>Lepidostoma</i> *
Trichoptera	<i>Limnephilus</i> *
Oligochaeta	Lumbriculidae*
Plecoptera	<i>Megarcys</i> *
Trichoptera	<i>Micrasema</i> *
Coleoptera	<i>Microcylloepus</i> *
Chironomidae	<i>Micropsectra</i> *
Nematoda	Nematoda*
Trichoptera	<i>Neothremma</i> *
Trichoptera	<i>Oecetis</i> *
Trichoptera	<i>Oligophlebodes</i> *
Oligochaeta	<i>Ophidonais serpentina</i> *
Coleoptera	<i>Optioservus</i> *
Chironomidae	<i>Orthocladius</i> *
Chironomidae	<i>Pagastia</i> *
Ephemeroptera	<i>Paraleptophlebia</i> *
Trichoptera	<i>Parapsyche elsis</i> *
Diptera	<i>Pericoma</i> *
Gastropoda	<i>Physella</i> *
Chironomidae	<i>Polypedilum</i> *
Chironomidae	<i>Potthastia</i> *
Trichoptera	<i>Protophila</i> *
Chironomidae	<i>Pseudosmittia</i> *
Plecoptera	<i>Pteronarcella</i> *
Plecoptera	<i>Pteronarcys</i> *
Chironomidae	<i>Rheocricotopus</i> *
Chironomidae	<i>Rheotanytarsus</i> *

Appendix A (Continued)

Taxa group	Operational taxonomic unit
Ephemeroptera	<i>Rhithrogena</i> *
Trichoptera	<i>Rhyacophila betteni</i> Gr.*
Trichoptera	<i>Rhyacophila brunnea</i> Gr.*
Trichoptera	<i>Rhyacophila coloradensis</i> Gr.*
Trichoptera	<i>Rhyacophila hyalinata</i> Gr.*
Trichoptera	<i>Rhyacophila pellisa</i> *
Ephemeroptera	<i>Serratella tibialis</i> *
Diptera	Simuliidae*
Plecoptera	<i>Skwala</i> *
Bivalvia	Sphaeriidae*
Chironomidae	<i>Stempellina</i> *
Plecoptera	Taeniopterygidae*
Chironomidae	<i>Tanytarsus</i> *
Chironomidae	<i>Thienemanniella</i> *
Diptera	<i>Tipula</i> *
Ephemeroptera	<i>Tricorythodes</i> *
Oligochaeta	Tubificidae*
Turbellaria	Turbellaria*
Diptera	<i>Tvetenia</i> *
Coleoptera	<i>Zaitzevia</i> *
Plecoptera	<i>Zapada cinctipes</i> *
Plecoptera	<i>Zapada columbiana</i> *
Plecoptera	<i>Zapada oregonensis</i> Gr.*
Trichoptera	<i>Agapetus</i>
Trichoptera	<i>Agraylea</i>
Hemiptera	<i>Ambrysus</i>
Trichoptera	<i>Amiocentrus aspilus</i>
Plecoptera	<i>Amphinemura</i>
Odonata	<i>Argia</i>
Diptera	Blephariceridae
Chironomidae	<i>Boreoheptagyia</i>
Coleoptera	<i>Brychius</i>
Ephemeroptera	<i>Caenis</i>
Ephemeroptera	<i>Callibaetis</i>
Chironomidae	<i>Cardiocladius</i>
Ephemeroptera	<i>Caudatella</i>
Diptera	Cecidomyiidae
Ephemeroptera	<i>Centropilum</i>
Trichoptera	<i>Ceraclea</i>
Trichoptera	<i>Chimarra</i>
Chironomidae	<i>Chironomus</i>
Ephemeroptera	<i>Choroterpes</i>
Chironomidae	<i>Cladotanytarsus</i>
Diptera	<i>Clinocera</i>
Chironomidae	<i>Conchapelopia</i>
Chironomidae	<i>Cryptochironomus</i>
Diptera	<i>Cryptolabis</i>
Plecoptera	<i>Cultus</i>
Chironomidae	<i>Demicryptochironomus</i>
Diptera	<i>Deuterothlebia</i>
Trichoptera	<i>Dicosmoecus atripes</i>
Trichoptera	<i>Dicosmoecus gilvipes</i>
Chironomidae	<i>Dicrotendipes</i>
Diptera	<i>Dixa</i>
Diptera	<i>Dixella</i>

Appendix A (Continued)

Taxa group	Operational taxonomic unit
Odonata	<i>Enallagma/Ischnura</i>
Chironomidae	<i>Endochironomus</i>
Chironomidae	<i>Euorthocladius</i>
Ephemeroptera	<i>Fallceon quelleri</i>
Gastropoda	<i>Ferrissia</i>
Gastropoda	<i>Fluminicola</i>
Amphipoda	Gammarus
Diptera	<i>Glutops</i>
Odonata	Gomphidae
Coleoptera	<i>Haliplus</i>
Oligochaeta	<i>Haplotalaxis</i>
Chironomidae	<i>Heleniella</i>
Coleoptera	<i>Helichus</i>
Hirudinea	<i>Helobdella stagnalis</i>
Ephemeroptera	<i>Heptagenia/Nixe</i>
Diptera	<i>Hesperoconopa</i>
Odonata	<i>Hetaerina</i>
Chironomidae	<i>Heterotrissocladius</i>
Chironomidae	<i>Hydrobaenus</i>
Gastropoda	Hydrobiidae
Coleoptera	Hydrophilidae
Plecoptera	<i>Isogenoides</i>
Coleoptera	<i>Lara avara</i>
Ephemeroptera	<i>Leptophlebia</i>
Trichoptera	<i>Leucotrichia</i>
Plecoptera	Leuctridae
Diptera	<i>Limmophora</i>
Chironomidae	<i>Limmophyes</i>
Chironomidae	<i>Lopescladius</i>
Oligochaeta	Lumbricidae
Gastropoda	Lymnaeidae
Chironomidae	<i>Macropelopia</i>
Plecoptera	<i>Malenka</i>
Trichoptera	<i>Mariliaa</i>
Diptera	<i>Maruina</i>
Trichoptera	<i>Mayatrichia</i>
Chironomidae	<i>Microtendipes</i>
Oligochaeta	<i>Nais communis</i>
Oligochaeta	<i>Nais elinguis</i>
Oligochaeta	<i>Nais variabilis</i>
Chironomidae	<i>Nanocladius</i>
Coleoptera	<i>Narpus</i>
Trichoptera	<i>Nectopsyche</i>
Trichoptera	<i>Neophylax</i>
Trichoptera	<i>Neotrichia</i>
Chironomidae	<i>Nilotanypus</i>
Trichoptera	<i>Ochrotrichia</i>
Chironomidae	<i>Odontomesa</i>
Coleoptera	<i>Ordobrevia</i>
Diptera	<i>Ormosia</i>
Trichoptera	<i>Oxyethira</i>
Chironomidae	<i>Parachaetocladius</i>
Chironomidae	<i>Parakiefferiella</i>
Chironomidae	<i>Paramerina</i>
Chironomidae	<i>Parametriocnemus</i>

Appendix A (Continued)

Taxa group	Operational taxonomic unit
Chironomidae	<i>Paraphaenocladius</i>
Chironomidae	<i>Paratanytarsus</i>
Chironomidae	<i>Paraorthocladius</i>
Chironomidae	<i>Pentaneura</i>
Lepidoptera	<i>Petrophila</i>
Chironomidae	<i>Phaenopsectra</i>
Gastropoda	Planorbidae
Oligochaeta	<i>Pristina</i>
Chironomidae	<i>Procladius</i>
Chironomidae	<i>Prodiamesa</i>
Plecoptera	<i>Prostoia</i>
Chironomidae	<i>Pseudochironomus</i>
Chironomidae	<i>Pseudodiamesa</i>
Chironomidae	<i>Pseudorthocladius</i>
Chironomidae	<i>Psilometriocnemus</i>
Trichoptera	<i>Psychomyia</i>
Diptera	<i>Ptychoptera</i>
Trichoptera	<i>Rhyacophila angelita</i> Gr.
Trichoptera	<i>Rhyacophila cyalinata</i> Gr.
Trichoptera	<i>Rhyacophila iranda</i> Gr.
Trichoptera	<i>Rhyacophila narvae</i>
Trichoptera	<i>Rhyacophila vagrita</i> Gr.
Trichoptera	<i>Rhyacophila verrula</i>
Megaloptera	<i>Sialis</i>
Oligochaeta	<i>Specaria</i>
Ephemeroptera	<i>Stenonema</i>
Chironomidae	<i>Stictochironomus</i>
Diptera	<i>Stratiomyiidae</i>
Chironomidae	<i>Subletta</i>

References

- Allan, J.D., 1995. Stream Ecology: Structure and Function of Running Waters. Chapman and Hall, London, UK, 400 pp.
- Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. EPA 841-B-99-002, 2nd ed. US Environmental Protection Agency, Office of Water, Washington, DC.
- Buffington, J.M., Montgomery, D.R., Greenberg, H.M., 2004. Basin-scale availability of salmonid spawning gravel as influenced by channel type and hydraulic roughness in mountain catchments. *Can. J. Fish. Aquat. Sci.* 61, 2085–2096.
- Clarke, R.T., Furse, M.T., Wright, J.F., Moss, D., 1996. Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. *J. Appl. Stat.* 23, 311–332.
- Clarke, R.T., Wright, J.F., Furse, M.T., 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. *Ecol. Model.* 160, 219–233.
- Hawkins, C.P., 2006. Quantifying biological integrity by taxonomic completeness: evaluation of a potential indicator for use in regional and global-scale assessments. *Ecol. Appl.* 16, 1277–1294.
- Hawkins, C.P., Hogue, J.N., Decker, L.A., Feminella, J.W., 1997. Channel morphology, water temperature, and assemblage structure of stream insects. *J. N. Am. Benthol. Soc.* 16, 728–749.
- Hawkins, C.P., Norris, R.H., Hogue, J.N., Feminella, J.W., 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecol. Appl.* 10, 1456–1477.
- Hawkins, C.P., Carlisle, D.M., 2001. Use of predictive models for assessing the biological integrity of wetlands and other aquatic habitats. In: Rader, R.B., Batzer, D.P., Wissinger, S.A. (Eds.), *Bioassessment and Management of North American Freshwater Wetlands*. John Wiley & Sons, New York, pp. 59–83.
- Hynes, H.B.N., 1970. *The Ecology of Running Waters*. University of Toronto Press, 555 pp.
- Knight, D.H., 1994. *Mountains and Plains: The Ecology of Wyoming Landscapes*. Yale University Press, New Haven and London, 338 pp.
- McCune, B., Medford, M.J., 1999. PC-ORD. Multivariate Analysis of Ecological Data, Version 4.0. MjM Software Design, Glenden Beach, OR.
- Minshall, G.W., 1984. Aquatic insect–substratum relationships. In: Resh, V.H., Rosenberg, D.M. (Eds.), *The Ecology of Aquatic Insects*. Praeger, New York, pp. 358–400.
- Moss, D., Furse, M.T., Wright, J.F., Armitage, P.D., 1987. The prediction of the macroinvertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biol.* 17, 41–52.
- Omerik, J.M., Gallant, A.L., 1987. Ecoregions of the West-central United States (Map). United States Environmental Protection Agency, Corvallis, OR.
- Ostermiller, J.D., Hawkins, C.P., 2004. Effects of sampling error on bioassessments of stream ecosystems: application of RIVPACS-type models. *J. N. Am. Benthol. Soc.* 23, 363–382.
- Plafkin, J.L., Barbour, M.T., Porter, K.D., Gross, S.K., Huges, R.M., 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers—Benthic Invertebrates and Fish. EPA/444/4-89-001. US Environmental Protection Agency, Office of Water, Washington, DC.
- Smith, M.J., Kay, W.R., Edward, D.H.D., Papas, P.J., Richardson, K.S.T.-J., Simpson, J.C., Pinder, A.M., Cale, D.J., Horwitz, P.H., Davies, J.A., Yung, F.H., Norris, R.H., Halse, S.A., 1999. AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. *Freshwater Biol.* 41, 269–282.
- StatSoft, 2001. STATISTICA, Version 6.0. StatSoft, Inc., Tulsa, OK.
- USGS, 1994. U.S. Geological Survey Bedrock Geology of Wyoming (Map) US Geological Survey, Denver, CO.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. The river continuum concept. *Can. J. Fish. Aquat. Sci.* 37, 130–137.
- Vannote, R.L., Sweeney, B.W., 1980. Geographic analysis of thermal equilibria: a conceptual model for evaluating the effects of

- natural and modified thermal regimes on aquatic insect communities. *Am. Nat.* 115, 667–695.
- Van Sickle, J., Hawkins, C.P., Larsen, D.P., Herlihy, A.T., 2005. A null model for the expected macroinvertebrate assemblage in streams. *J. N. Am. Benthol. Soc.* 24, 178–191.
- WDEQ, 2001. Wyoming Water Quality Rules and Regulations. Chapter 1. Wyoming Surface Water Quality Standards. Water Quality Division, Cheyenne, WY.
- Wright, J.F., Furse, M.T., Armitage, P.D., 1993. RIVPACS—a technique for evaluating the biological quality of rivers in the U.K. *Eur. Water Pollut. Contr.* 3, 15–25.
- Wright, J.F., Sutcliffe, D.W., Furse, M.T., 2000. Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. Freshwater Biological Association, Ambleside Cumbria, UK, 400 pp.