

A Multimetric Assessment of Stream Condition in the Northern Lakes and Forests Ecoregion Using Spatially Explicit Statistical Modeling and Regional Normalization

EDWARD A. BAKER*

*Michigan Department of Natural Resources,
488 Cherry Creek Road, Marquette, Michigan 49855, USA*

KEVIN E. WEHRLY AND PAUL W. SEELBACH

*Michigan Department of Natural Resources, Institute for Fisheries Research,
212 Museums Annex, Ann Arbor, Michigan 48109, USA*

LI WANG¹

*Wisconsin Department of Natural Resources,
1350 Femrite Drive, Monona, Wisconsin 53716, USA*

MICHAEL J. WILEY

*School of Natural Resources and Environment,
University of Michigan, Ann Arbor, Michigan 48109, USA*

THOMAS SIMON

*U.S. Fish and Wildlife Service,
620 South Walker Street, Bloomington, Indiana 47403, USA*

Abstract.—We sampled fish communities, water temperature, water chemistry, physical habitat, and catchment characteristics for 94 stream sites selected randomly throughout the Northern Lakes and Forests ecoregion and used those data to explicitly model reference conditions and assess ecological stream condition at each site via a regional normalization framework. The streams we sampled were first order through fourth order, and the catchments ranged from 0.9 to 458 km². We developed multiple linear regression (MLR) models that predicted fish community metrics, water chemistry characteristics, and local physical habitat from catchment characteristics; we used these models to compare existing conditions with the conditions that would be expected based on the regression models. Our results indicated that the fish communities were relatively unimpaired because the catchment variables associated with human-induced land use change were important in only 1 of the 10 fish metric models. Agricultural land use was a significant variable in the MLR equation for species of *Lepomis* (sunfish). Agricultural land use and urban land use were both significant variables in all of the MLR models predicting water chemistry variables; urban land use was a significant variable in the MLR model predicting the percent coverage of all instream cover types. Regional normalization indicated that none of the sites were impaired based on fish community attributes. However, our analysis based on water chemistry metrics indicated that 22–35% of the sites were impaired and that, based on physical habitat, 6–14% of the sites were impaired. A comparison with other published studies of the ecoregion suggested that the regional normalization process correctly characterized stream condition.

Resource agencies increasingly rely on biological communities as indicators of stream water quality. Such biologically based approaches, or

bioassessments, are appealing because they incorporate attributes of aquatic assemblages that are sensitive to the multiple factors influencing streams and provide a cost-effective method to assess the ecological status of a large number of sites across broad geographic regions (Karr and Chu 1999). Establishment of reference conditions is fundamental to all bioassessment approaches (e.g., index of biotic integrity [IBI]: Karr 1981; River Invertebrate Prediction and Classification System

* Corresponding author: bakerea@michigan.gov

¹ Present address: Michigan Department of Natural Resources, Institute for Fisheries Research, 212 Museums Annex, Ann Arbor, Michigan 48109, USA.

Received November 18, 2003; accepted November 24, 2004
Published online May 11, 2005

[RIVPACS]: Wright et al. 1993). Reference conditions represent the biological characteristics expected in the anthropogenically unaffected condition and serve as the benchmark against which sites of unknown status (test sites) are judged (Reynoldson et al. 1997).

Traditionally, reference conditions have been defined by characterizing a group of sites that the investigator judges to be least impaired. Two common types of bioassessments (IBI and [RIVPACS]) rely on the use of reference sites. Test sites are then compared with reference sites to evaluate the degree of ecological impairment at the test sites. An important assumption inherent in this approach is that test and reference sites would have equivalent physical, chemical, and biotic characteristics in the absence of human impacts. In lotic systems, identifying appropriate reference sites for comparison with test sites is made difficult by the large variability that exists among sites. This variability must be accounted for to accurately assess ecological status. Several methods have been developed to ensure that test and reference sites are equivalent, all of which employ some form of classification (reviewed in Seelbach et al. 2002). The IBI-type assessments often use ecoregions (Omernik and Gallant 1988) as the basic classification unit. In this case, all streams within an ecoregion are assumed to be similar, and test sites are compared with reference sites based on membership in the ecoregion. Frequently, however, streams are not homogeneous at the ecoregion scale (Wang et al. 2003), and finer-scale classifications have been developed based on physical attributes that are important in structuring biological assemblages. In these cases, test sites are compared with reference sites having similar characteristics, such as stream size (Ohio Environmental Protection Agency 1987; Lyons et al. 2001) or water temperature (Lyons et al. 1996; Mundahl and Simon 1998). The RIVPACS-type approaches classify reference sites into groups based on similarity in taxonomic composition (Wright et al. 1993; Norris 1996; Hawkins et al. 2000). A modeling approach is then used to compare test sites with appropriate reference sites, based on similarity of physical characteristics.

An alternative method for establishing reference conditions is the development of site-specific expectations derived from landscape-context statistical modeling (reviewed by Seelbach et al. 2002). In this approach, statistical models that relate patterns in coarse-scale features (e.g., catchment geology or land use) to site-scale ecological response

variables (e.g., stream flow or species richness) are developed for a particular geographic region. When these models include anthropogenic factors as predictor variables, they can be used to generate site-specific expectations or reference conditions by setting these factors to zero. For example, for an urbanized catchment, if the urban land-use value is set to zero and the model is used to recalculate fish community metrics, the model output represents the expected fish community in that stream in the absence of urbanization. This approach is based on the view that large-scale features are important in controlling the local ecological character of streams (Frissell et al. 1986; Poff 1997). This approach eliminates the need for investigators to select and classify reference sites, eliminating a potential source of bias in stream bioassessment. Landscape-context statistical models are similar in intent to stream classification, but instead of using strata, variance is explicitly accounted for by modeling. Recent studies have shown that certain ecological characteristics of aquatic sites can be readily predicted from catchment-scale variables (Wiley et al. 1997; Wehrly et al. 1998; Soranno et al. 1999; Baker et al. 2001).

Using landscape context models to estimate reference conditions, Wiley et al. (2003) developed an assessment approach called "regional normalization." The regional normalization process involves two steps. First, observed indicator data are re-expressed as deviations from their expected scores, based on explicitly modeled reference conditions for each site. Second, the deviations are then normalized (scaled) by dividing by a measure of the natural regional variation observed for that metric. This type of scaling is useful because the larger the regional variance a metric exhibits, the less any unit deviation from an expected value tells us about relative status. Furthermore, this scaling allows direct comparison of normalized scores computed from very different metrics (e.g., native fish taxa, macroinvertebrate families, water temperature, and conductivity). The normalized scores also can be used to construct ad hoc summary metrics (i.e., multimetrics) and to compare trends within and between classes of indicators (typical goals of assessment programs). The regional normalization approach has been used successfully to model reference conditions and to assess ecological stream conditions, by using data on fish assemblages across portions of five ecoregions in Michigan's Lower Peninsula, including the southeastern portion of the Northern Lakes and Forests (NLF) ecoregion (Wiley et al. 2003). Regional nor-

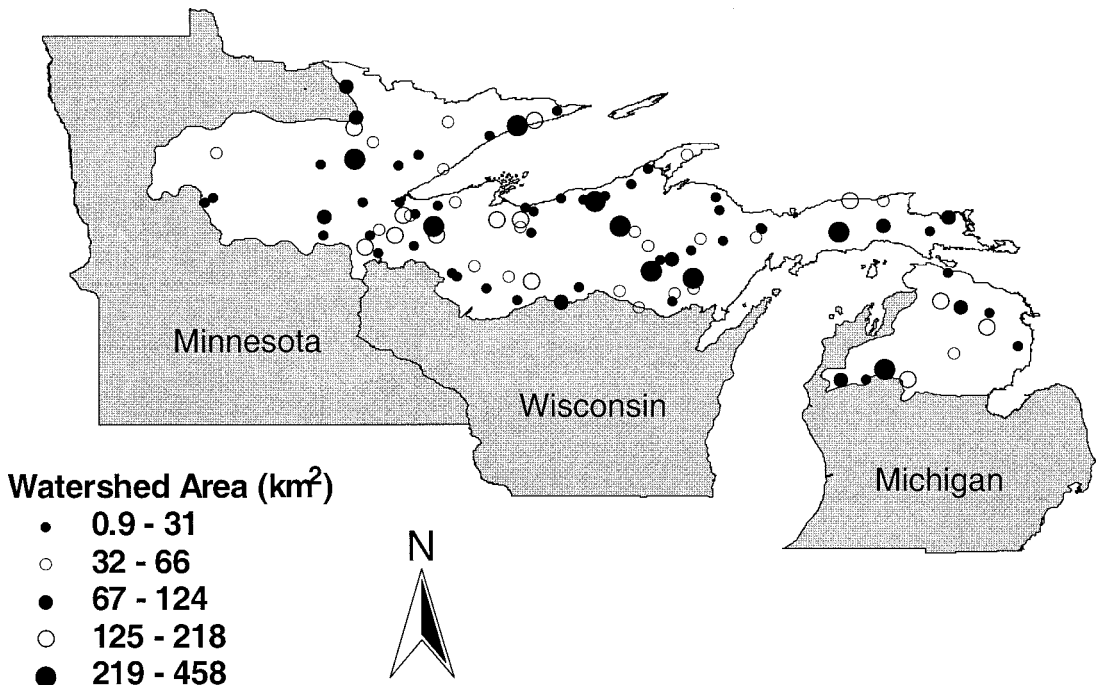


FIGURE 1.—Northern Lakes and Forests Ecoregion overlain on states of Minnesota, Wisconsin, and Michigan. Variation in catchment area (km^2) is shown for stream sites randomly sampled during 1998–1999.

malization was also used to assess conditions in the Huron River watershed in southeastern Michigan by using data on macroinvertebrates and water chemistry (Wiley and Martin 2000). Wiley et al. (2003) concluded that streams in the NLF were in better condition than streams in other ecoregions in Michigan. However, this conclusion was based on sampling from only a small portion of the NLF ecoregion in Michigan's Lower Peninsula.

Wang et al. (2003) used multivariate statistical techniques to evaluate the relative importance of watershed-scale and reach-scale habitat conditions in structuring stream fish assemblages at 79 sites in the NLF ecoregion. They concluded that reach-scale factors were the primary determinant of stream-fish assemblage in relatively undisturbed stream ecosystems but that watershed-scale factors became increasingly important as the degree of landscape modifications increased. Wang et al. (2003) also concluded that stream-fish assemblages in the NLF ecoregion could be grouped into one of three environment–fish associations that were largely explained by water temperature regime and stream size.

In this study we applied the regional normalization technique to 94 sites (the same sites used by Wang et al. plus 15 additional sites) to assess

stream fish, habitat, and water chemistry conditions in the NLF ecoregion. Our objective was to use the regional normalization technique to assess stream fish community, water chemistry, and habitat conditions for first-order to fourth-order streams located throughout the NLF ecoregion and, in so doing, provide a description of the methodology. In this study, we modeled reference condition for stream sites in the NLF ecoregion. We first developed statistical multiple linear regression (MLR) models using catchment-scale and reach-scale characteristics. These MLR equations were then used to model site-specific reference conditions for a suite of ecological (fish, chemistry, and physical habitat) indicator metrics typically used in IBI-type assessments. Finally, we summarized the ecological status of first-order to fourth-order streams across the NLF ecoregion.

Study Site

The NLF ecoregion lies within the formerly glaciated section of the upper midwestern USA (Figure 1) and encompasses $180,500 \text{ km}^2$ in northern Minnesota, Wisconsin, and Michigan (Omernik and Gallant 1988). The landscape of the NLF ecoregion is quite varied containing bedrock uplands and a mosaic of glacial deposits, including

end and ground moraines, ice-contact ridges, outwash sand plains, lacustrine sand and clay plains, and glacial fluvial valleys. Predominant landcover includes a mixture of conifer and deciduous forests, wetlands, marshes, and swales. Most streams in the ecoregion are perennial, commonly originating in lakes or wetlands, and have a low stream density of about 1.0 stream kilometer per square kilometer.

Anthropogenic activities have influenced the landscape through land-use practices such as timber harvest, residential development, agriculture, livestock grazing, and light industrial development. Factors associated with these activities, such as increased runoff and siltation, reduction of riparian corridors and stream habitat, channel and drainage modification, and dewatering and filling of lowlands, have potentially altered stream resources. Despite these changes, the NLF ecoregion contains a relatively undisturbed landscape compared with other ecoregions in the Midwest (e.g., Eastern Corn Belt Plain; Simon and Dufour 1997).

Methods

Survey design.—During 1998–1999 we sampled 94 first-order through third-order (as defined by Strahler 1957; 1:100,000 map scale) wadeable stream sites distributed throughout the NLF ecoregion. Sampling sites were selected using a random stratified probability-based design (Overton et al. 1991). The sample population of streams in the region was delineated from digitized U.S. Geological Survey (USGS) topographic maps (1:100,000 scale). The U.S. Environmental Protection Agency (USEPA; Pacific Ecology Division, Corvallis, Oregon) selected sample sites using the RF3 algorithm for equal weighting of stream segments (Overton et al. 1991). Sites selected for sampling were defined by a latitude, longitude point on the stream. This latitude, longitude point then defined the midpoint of the sampled reach. Sample probabilities were weighted according to density so that approximately equal numbers of first-order, second-order, and third-order streams were selected across the region; although equal numbers of each order were selected, not all selected sites were sampled. Reasons for omitting sites included site access difficulties due to land ownership, excess depth at the site, proximity to road crossings, etc. Streams were sampled from June to October during low flows (at least 3 d after rain events).

Chemistry.—We measured alkalinity, hardness, dissolved oxygen, pH, turbidity, and conductivity on the same day that fish communities and physical

habitat were sampled. Alkalinity, hardness, and dissolved oxygen were determined by titration; pH, turbidity, and conductivity were measured with electronic meters. Water samples were also collected, preserved, and shipped to the USEPA laboratory in Duluth, Minnesota, for analysis of total suspended solids and nutrient (nitrogen and phosphorous) concentrations.

Temperature.—In all sampled streams, we placed electronic temperature loggers (Optic Stowaway loggers, Onset Computer Corporation) that recorded temperature once per hour during deployment periods beginning in April or May and ending in September or October. Mean maximum and minimum temperatures were calculated (average of daily average) for each stream for 24 June to 21 July. This period was selected because it is when streams in the ecoregion are typically warmest (Michigan Department of Natural Resources [MDNR], unpublished data).

Reach-scale variables.—We collected reach-scale habitat and riparian disturbance data using methods derived from both the USEPA Environmental Monitoring and Assessment Program protocols (Lazorchak et al. 1998) and the Wisconsin habitat rating system (Simonson et al. 1994). Reach-scale habitat variables measured included instream and riparian features. We also measured discharge and air temperature when the fish community was sampled. Stream order, sinuosity, and stream linkage number were determined from 1:24,000 scale topographic maps. Although stream order was determined from 1:100,000 scale maps for the purpose of selecting sample sites, we used stream order determined from 1:24,000 scale maps in the analysis.

Catchment-scale variables.—Digital maps from the three states were compiled into a comprehensive Geographic Information System (GIS). Individual upstream catchment boundaries were delineated for each site based upon subwatershed divides mapped by MDNR from USGS 1:24,000 scale topographic maps. Watershed boundaries were then locally modified for each sample site by using a 3-arc-second digital elevation model (at a scale of 1:250,000). Land use, surficial geology, bedrock geology, slope, precipitation, and growing degree-days were summarized for each catchment using Arc/Info software (ESRI, Inc.). Groundwater flux data were summarized for each catchment by using a spatial model (map layer) predicting maximum and minimum potential groundwater velocities (based on Darcy's Law; Wiley et al. 1997; Baker et al. 2003) via minimum and maximum

conductivity values for each surficial and bedrock geology class.

Because long-term flow data were not available for sampled sites we predicted base-flow water yield (90% exceedence flow; $\text{m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$) at each stream site from an unpublished MLR equation developed from USGS stream gauging data and catchment characteristics for gauged sites in the NLF ecoregion. Independent variables in the MLR equation included catchment area, average catchment slope, percent surface water bodies in catchment, percent sand and gravel (from surficial geology), and percent forested wetland in the catchment. The MLR equation explained 94.5% of the variability in the streamflow data and was significant at $P < 0.001$.

Fish assemblages.—Fish assemblages were sampled via electrofishing with either a backpack unit or gasoline-powered generator mounted on a tow-barge. Backpack units were used in streams too small to float the tow-barge (streams generally < 10 m wide). Lyons (1992b) noted that sampling 35 times the average stream width was necessary for accurately characterizing stream-fish species richness. We sampled 35 times the average wetted channel width of the selected site (the level Lyons 1992b noted was necessary for accurately characterizing stream-fish species richness) but never less than 150 m or more than 500 m. Electrofishing started at the downstream end of the reach and proceeded upstream for 1 pass. Fish were identified and counted by species in the field. Three voucher specimens of each species were preserved in 10% formalin for laboratory confirmation of field identifications.

We used literature references to classify fish into structural and functional categories for metric computation (Lyons 1992a; Lyons et al. 1996). Because the fish fauna in the NLF ecoregion is naturally species poor compared with other ecoregions in the Midwest, we modified certain metrics according to Lyons (1992a). For this analysis, the count of darter species included the number of *Cottus* (i.e., sculpin) and *Noturus* (i.e., madtoms) species, and the number of sunfish species metric included yellow perch *Perca flavescens*. We found no differences in richness metrics among major drainage basins (i.e., Great Lakes, Mississippi, and Hudson Bay) in the NLF (analysis of variance [ANOVA], $P > 0.05$), and therefore, combined data from all basins in further analyses. We also estimated fish density (number/ m^2) at each site by dividing the total number of fish captured by the

surface area sampled (mean stream width \times reach length).

Data analysis.—We constructed MLR models to predict fish, water chemistry, reach-scale physical habitat, and riparian disturbance metrics from both catchment-scale and reach-scale physical variables. Before constructing models, we calculated Pearson's product-moment correlation coefficients among variables in the data set. We attempted MLR model construction only when there were significant correlations between dependent and independent variables and only independent variables that were significantly correlated ($\alpha = 0.05$) with the dependent variables were considered for inclusion in the models. We then constructed models using a stepwise progression. Independent variables were included in models if they were significant ($\alpha = 0.05$) and improved the model fit (r^2 value). Models and model coefficients were considered significant at $\alpha = 0.05$. When necessary, data were log-transformed to meet assumptions of normality.

The MLR models were then used to calculate expected (reference) conditions at sampled sites. When measures of human impacts were significant variables in a MLR model, coefficients for those variables (included agricultural or urban land use) were set to zero and the model was recalculated. Resultant model output represented an expectation for conditions in the absence of agricultural or urban land use. When models did not include measures of agricultural or urban land use as significant variables, the models were not adjusted, and resultant predictions were assumed to directly describe reference conditions. Status was determined for each metric at each site by first calculating a deviation score. For most metrics deviation scores were determined by subtracting the modeled expectation from the observed value. Deviation scores for chemistry metrics, channel width : depth ratio, the number of native species metric for cold-water streams, and the percent omnivore metric were calculated by subtracting the observed value from the modeled expectation. In this way, observed values that exceed expected values result in a negative deviation score. Normalized scores for a particular metric were then calculated by dividing the deviation score for that metric by the standard deviation of the modeled expectation for that metric.

To assess site condition, we constructed box plots of the normalized metric scores. These plots are simple to view, and provide a powerful suite of information on the status of individual metrics

(Wiley et al. 2003). For example, sites that have normalized metric scores near zero are considered unimpacted by anthropogenic disturbance because the observed conditions are similar to the expected conditions predicted by the MLR models. Sites that have normalized metric scores that are far from zero are considered impacted because the observed conditions deviate from expected. The farther a site score is from zero, the greater the impact. We arbitrarily interpreted ± 2 standard deviations from expected as describing the statistically expected range. Deviations exceeding $+2$ SD signified exceptional conditions, and those less than 2 SD signified impacted conditions (Wiley et al. 2003). In addition, we calculated summary multimetrics for the fish community, water chemistry, and reach-scale physical habitat by averaging the individual metrics for each indicator (Wiley et al. 2003). These summary multimetrics were used to represent the overall status of the sampled streams in the NLF Ecoregion.

Results

Stream Characteristics

Of the 94 stream sites we sampled, 17 were first order, 30 second order, 39 third order, and 8 fourth order. Sampled streams were distributed across the NLF ecoregion (Figure 1). Catchment characteristics for the sampled sites were quite variable. For example, median catchment area 42.5 km^2 and ranged from 0.9 to 458.0 km^2 . Sampled streams were largely undisturbed by agricultural or urban development. The median agricultural land use was 0.6% of catchment areas and ranged from 0% to 72% . The median urban land use was 0% of catchment areas and ranged from 0% to 50% . Land cover in the sampled catchments was predominated by deciduous, coniferous, or mixed forests (median, 68.1% of the land cover) and wetlands (16.3%).

Water quality metrics varied by at least an order of magnitude among streams sampled in the NLF ecoregion. Nutrient concentrations were generally low, although some streams did have exceptionally high values for total phosphorous and total nitrogen.

Reach-scale habitat was also extremely variable among sampled sites and indicated a range of channel conditions from small, shallow, intermittent streams to large, deep perennial streams. Many of the sites were low gradient streams crossing sand plains, till plains, and lake plains and were predominated by soft substrates. However, some

sites were on slopes of bedrock formations or moraines and had high local gradients and substrates composed of gravel, cobble, and boulder.

We captured 42,032 fish representing 61 different species from 13 families. Cyprinids were the most commonly caught taxon among streams (93.6% occurrence) and were also the most abundant taxon among all fish, representing 54.5% of all individuals captured. Based on percent occurrence, catostomids (75.5%) and umbrids (71.3%) were also common, and percids and sticklebacks (Gasterodeidae) occurred in over half of the streams. Salmonids were collected at 50% of the sites sampled.

The number of species captured at a site varied across the ecoregion (range = 1 – 26 , median = 11), as did the number of fish captured per site (range = 4 – $2,890$, median = 272). Fish density among sites ranged from 0.01 to 2.96 fish/m^2 and median fish density was 0.16 fish/m^2 .

Application of Regional Normalization to NLF Streams

Statistically significant MLR models relating fish community attributes to catchment-scale and reach-scale habitat features were constructed for 10 of the fish community metrics analyzed (Table 1). No statistically significant MLR models were constructed for the number of darter species, percent tolerant species, or percent insectivorous species metrics. The constructed models included both catchment-scale and reach-scale variables, and explained 23 – 57% of the variation in fish community metrics. The number of sunfish species at a site was positively correlated with the percent agricultural land use in the catchment and was the only fish community metric that was related to our measures of human disturbance in the landscape.

Statistically significant MLR models were developed for total phosphorous, conductivity, total nitrogen, and total suspended solids concentrations (Table 2). The models accounted for 19 – 44% of the variability in the water chemistry metrics, and all four models included urban or agricultural land use or both as predictors. All four water chemistry metrics were positively related to agricultural land use. Conductivity and total nitrogen were positively related to urban land use, and total suspended solids concentration was negatively related to urban land use.

Significant MLR models were developed for five of the reach-scale habitat variables (Table 3) and explained 18 – 57% of the variation in the data set. Of the six models, the model for percent of all

TABLE 1.—Multiple linear regression models used for streams in the Northern Lakes and Forests ecoregion relating fish community attributes to reach-scale and catchment-scale habitat attributes. Models are of the general form $\log_e(Y) = \text{constant} + B_1\log_e X_1 + B_2\log_e X_2 + B_3\log_e X_3, \dots, B_n\log_e X_n$.

Dependent variable	Independent variables	Coefficient	Coefficient standard error	Coefficient P-value	Model R^2	Model P-value
Number of native species (NATSP)	Minimum basin groundwater index value	-0.112	0.037	0.003	0.567	<0.005
	Area sampled	0.122	0.054	0.026		
	Average maximum temperature	1.719	0.330	<0.001		
	Bankfull depth/catchment area	-2.583	0.703	<0.001		
Number of sucker species (SCKR)	Average maximum temperature	0.042	0.013	0.002	0.324	<0.005
	Catchment slope	-0.127	0.034	<0.001		
Fish density (number/m ²) (DNSTY)	Channel width : depth ratio	0.156	0.065	0.019	0.322	<0.005
	Average channel width	-0.146	0.025	<0.001		
	Average maximum temperature	0.120	0.037	0.002		
Percent lithophilous spawners (LITHO)	Total nitrogen	0.000042	0.0	0.014	0.363	<0.005
	Average maximum temperature	0.499	0.108	<0.001		
	Catchment area	0.021	0.012	0.027		
Number of sunfish species (SUNF)	Percent all cover types	-0.002	0.001	0.002	0.346	<0.005
	Percent total agricultural land use	1.205	0.401	0.003		
	Average minimum temperature	0.077	0.025	0.003		
	Catchment area	0.080	0.037	0.036		
Percent carnivorous fish (CARNV)	Percent total water land cover	4.353	1.328	0.001	0.388	<0.005
	Average catchment slope	0.029	0.013	0.031		
	Average maximum temperature	-0.401	0.092	<0.001		
	Dissolved oxygen	0.020	0.006	0.002		
Percent coolwater and coldwater individuals (CLCLD)	Total nitrogen	0.00005	0.0	0.16	0.538	<0.005
	Water yield	1.630	0.634	0.012		
	Average minimum temperature	-0.062	0.008	<0.001		
	Turbidity	-0.059	0.018	0.002		
Number of intolerant species (INTOL)	Channel width : depth ratio	-0.063	0.031	0.044	0.547	<0.005
	Dissolved oxygen	0.114	0.019	<0.001		
	Catchment area	0.109	0.030	<0.001		
	Longitude	0.044	0.014	0.002		
Percent brook trout (BKT)	Average maximum temperature	0.045	0.014	0.002	0.287	<0.005
	Average minimum temperature	-0.431	0.093	<0.001		
	Total phosphorous	-0.050	0.015	0.002		
Percent omnivorous individuals (OMNIV)	Water yield	-0.049	0.016	0.002	0.233	<0.005
	Percent peat and muck (from surficial geology)	0.213	0.084	0.013		
	Percent bank erosion	0.018	0.006	0.007		

cover types was the only one that included either agricultural or urban land use as a predictor. The percent of instream cover present in a reach was positively related to the percent of urban land use in the catchments.

Normalized scores were calculated for each of the fish metric, water chemistry, and reach-scale habitat variables from the MLR models (Figure 2). For the number of sunfish species, water chemistry variables, and percent of all cover types, we set agricultural and urban land-use values to zero and

calculated deviation scores by subtracting the expected value from the observed value. Normalized water chemistry scores were positively correlated with agricultural land use, and conductivity and total nitrogen scores were positively correlated with percent urban land use.

Assessment of Stream Condition Across NLF

Our results indicated that the majority of streams we sampled were relatively unimpacted (Figure 2). Most of the normalized fish metric scores were

TABLE 2.—Multiple linear regression models used for streams in the Northern Lakes and Forests ecoregion relating water chemistry attributes to reach-scale and catchment-scale habitat attributes. Models are of the general form $\log_e(Y) = \text{constant} + B_1\log_eX_1 + B_2\log_eX_2 + B_3\log_eX_3, \dots, B_n\log_eX_n$.

Dependent variable	Independent variables	Coefficient	Coefficient standard error	Coefficient P-value	Model R ²	Model P-value
Total phosphorus (TP)	Percent agricultural land	3.636	0.531	<0.001	0.437	<0.005
	Longitude	-0.115	0.023	<0.001		
	Percent riparian woodlot	-0.074	0.036	0.042		
Conductivity (COND)	Percent agricultural land use	294.057	78.082	<0.001	0.272	<0.005
	Percent urban land use	849.493	207.016	<0.001		
	Percent peat and muck (from surficial geology)	-188.852	78.831	0.019		
Total nitrogen (TN)	Percent agricultural land use	2.044	0.449	<0.001	0.354	<0.005
	Percent urban land use	2.556	1.205	0.037		
	Percent riparian woodlot	-0.091	0.031	0.004		
	Percent wetland in catchment	1.714	0.426	<0.001		
Total suspended solids (TSS)	Percent agricultural land use	3.348	0.817	<0.001	0.189	<0.005
	Percent urban land use	-4.593	2.182	0.038		
	Percent clay substrate in stream	0.160	0.076	0.038		

distributed within ±1 SD of expected (zero), and only a few scores fell below -2 SD, the threshold denoting impacted conditions. Nearly all fish metrics had some normalized scores that exceeded +2 SD, indicating exceptional conditions. Normalized water chemistry metric scores were more widely distributed, but the majority of streams we sampled appeared relatively unimpacted. However, a num-

ber of scores fell below the threshold for concern suggesting that 22–35% of streams sampled had some impairment due to extreme water chemistry values (Figure 2). The distributions of normalized physical habitat metrics scores were also centered at or near zero and indicated that sampled streams were relatively unimpacted. The physical habitat of 6–14% of the sample streams deviated substan-

TABLE 3.—Multiple linear regression models used for streams in the Northern Lakes and Forests ecoregion relating reach-scale physical habitat attributes to reach-scale and catchment-scale habitat attributes. Models are of the general form $\log_e(Y) = \text{constant} + B_1\log_eX_1 + B_2\log_eX_2 + B_3\log_eX_3, \dots, B_n\log_eX_n$.

Dependent variable	Independent variables	Coefficient	Coefficient standard error	Coefficient P-value	Model R ²	Model P-value
Thalweg depth standard deviation (SDTHAL)	Catchment area	0.116	0.042	0.007	0.184	<0.005
	Catchment slope	-0.118	0.054	0.031		
Percent areal coverage of all cover types (%ALLCOV)	Percent urban land use	12.917	4.896	0.010	0.254	<0.005
	Percent bank erosion	-0.194	0.059	0.001		
	Channel width : depth ratio	-0.554	0.166	0.001		
Percent riparian woodlot (%WOODLOT)	Catchment slope	10.384	3.443	0.003	0.212	<0.005
	Longitude	2.647	1.309	0.046		
	Channel width : depth ratio	20.662	6.609	0.002		
Percent areal coverage of in-stream woody cover (%WDYCOV)	Catchment area	0.151	0.079	0.058	0.270	<0.005
	Percent riparian woodlot	0.327	0.061	<0.001		
Channel width : depth ratio (W : D ratio)	Channel width : depth ratio	-1.034	0.209	<0.001	0.568	<0.005
	Catchment area	0.147	0.028	<0.001		
	Percent riparian woodlot	0.112	0.025	<0.001		
	Percent instream large substrate	0.133	0.028	<0.001		
	Percent woody cover	-0.157	0.039	<0.001		

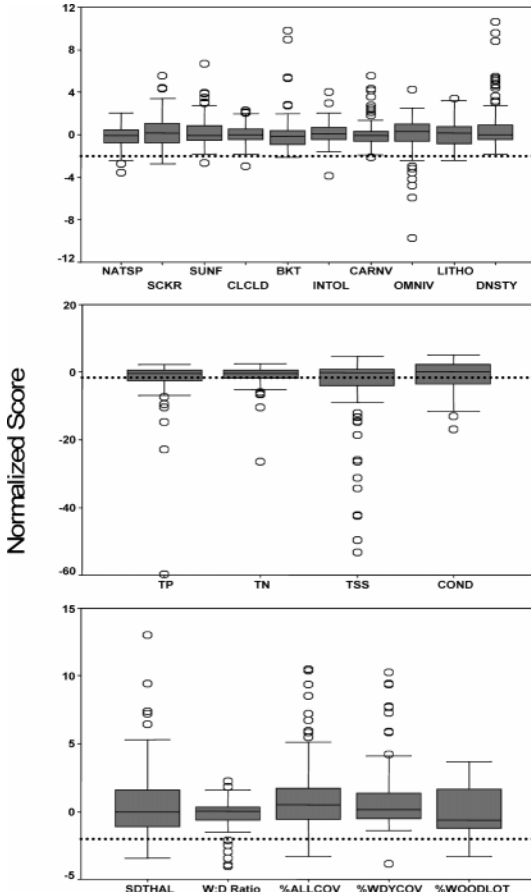


FIGURE 2.—Box plots of normalized fish community (upper panel), water chemistry (middle panel), and physical habitat (lower panel) metrics, scaled in terms of standard deviations, for streams in the Northern Lakes and Forests Ecoregion of Minnesota, Wisconsin, and Michigan. Dashed line indicates threshold of concern (-2 SD); circles represent individual site scores outside of interquartile range. Labels for x-axis defined in Tables 1–3.

tially from expected, indicating some impairment due to extreme physical habitat values (Figure 2).

We constructed summary multimetric scores for each of the three groups of variables (fish, water chemistry, physical habitat) to evaluate the overall condition of the sampled streams (Figure 3). Multimetric scores were calculated using all of the variables that could be modeled with the MLR equations (Tables 1–3). Nearly all of the summarized fish metric site scores were nearly zero, indicating fish communities in sampled streams were very near the expectations predicted by the MLR models. The summarized water chemistry scores indicated overall potential impairment at 32% of

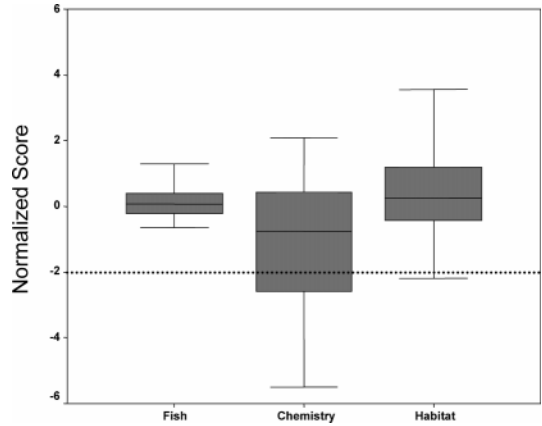


FIGURE 3.—Box plots of normalized summary multimetrics for fish, chemistry, and physical habitat, scaled in terms of standard deviations, for streams in the Northern Lakes and Forests Ecoregion of Minnesota, Wisconsin, and Michigan. Dashed line indicates threshold of concern (-2 SD).

the sampled sites (Figure 3). Agricultural and urban land use in the catchments of these impaired sites were both significantly greater (Mann–Whitney U -test, $Z_{\text{agricultural}} = -4.9$ and $Z_{\text{urban}} = -4.7$, $P < 0.001$) than in catchments of unimpaired sites (Figure 4). The summarized physical habitat metric scores indicated overall potential impairment at 2% of the sampled sites (Figure 3).

Discussion

Our assessment, based on the regional normalization approach, suggested that most streams in the Northern Lakes and Forests ecoregion are relatively unimpacted by agricultural and urban land-use changes. This result is probably due to the fact that the NLF has not experienced the widespread landcover changes (agriculture and urban development) typical of other Midwestern ecoregions (Omernik and Gallant 1988). Of the 10 MLR models for fish community metrics we constructed, only 1 incorporated agricultural or urban land use, suggesting that these land uses have had little affect on stream fish communities in the NLF. In a parallel study to ours, Butcher et al. (2001) found that benthic macroinvertebrate communities in the same NLF streams were relatively unimpacted. Only 12 of the 94 sites sampled were judged to be impaired in the development of a benthic community IBI for the NLF. Using multivariate statistical analysis, Wang et al. (2003) concluded that NLF stream fish community structure was related to watershed and local physical habitat conditions

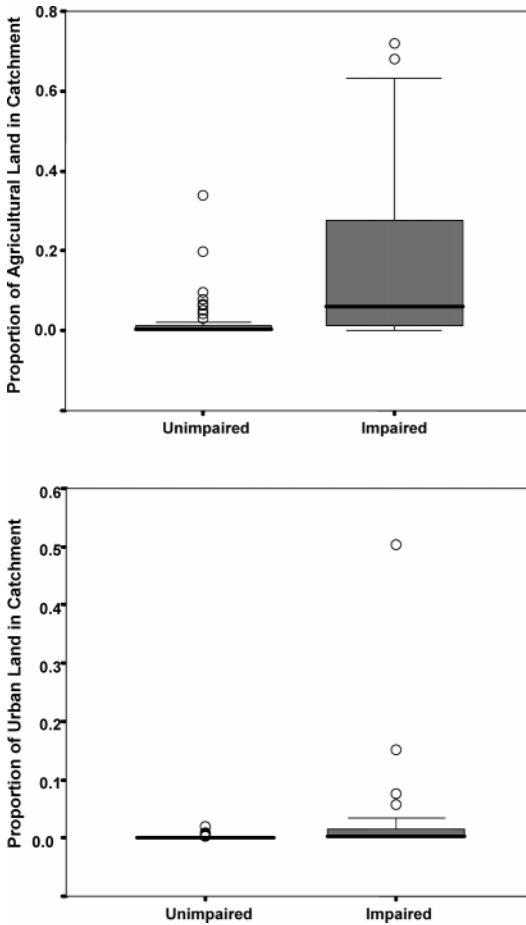


FIGURE 4.—Box plots showing differences in agriculture (upper panel) and urban land use (lower panel) in the catchments of stream sites that are unimpaired and impaired, based on normalized chemistry scores, for the Northern Lakes and Forests Ecoregion of Minnesota, Wisconsin, and Michigan. Circles represent individual sites outside of interquartile range.

unrelated to human land-use changes. Wang et al. (2003) also calculated IBI scores for 79 of the 94 sites we analyzed in this study and concluded that IBI scores were not an accurate indicator of stream condition in the NLF because they were lower than expected (58% of sites had IBI scores <50). The low IBI scores calculated by Wang et al. (2003) were attributed to the lack of an IBI specifically designed for cool-water or small warmwater streams that predominate in the NLF, a problem that is overcome by our use of spatially explicit MLR models and regional normalization.

Wiley et al. (2003) used regional normalization of bioassessment (fish) data to assess stream condition in portions of five ecoregions in lower Mich-

igan, including the southeast portion of the NLF. They also found little evidence of impairment in the NLF, in contrast to several other ecoregions with heavy agricultural and urban land usage (see figures in Wiley et al. 2003). However, the fish community summary metric used by Wiley et al. (2003) was constructed from only two fish community metrics: the intolerant fish species metric and the total fish species metric. Our analysis extends the conclusions of Wiley et al. (2003) to the entire NLF ecoregion and provides a more robust assessment of stream condition because we were able to use a larger set of fish community metrics in our analysis. In addition, our assessment integrated water chemistry and physical habitat metrics, in addition to the fish community metrics, to provide a more complete picture of stream status in the NLF ecoregion.

Our MLR models indicated water chemistry and some physical habitat attributes have been impacted in some streams, but these changes have apparently not translated to significant fish community alterations. All of our MLR models for water chemistry included percent agricultural land use and three out of the four models included percent urban land use. The negative influence of human land-use changes on water quality in streams is well documented (reviewed in Baker et al. 2001). The limited amount of agricultural and urban land use observed in the NLF ecoregion appears to explain the excessive chemistry values observed at a third of our sampling sites (Figure 4). That water chemistry and habitat indicators showed some impairment but fish communities did not underscores the need to consider multiple measures when assessing stream condition (Lyons 1992a; Karr and Chu 1999).

Although study sites included in our analysis appeared relatively unimpacted, NLF streams probably do suffer the effects of human disturbance. Nearly the entire ecoregion was clear-cut near the turn of the 20th century, and much of the land is still managed for timber production. In addition, some streams in this region are probably impacted by localized urban development in the region (Wang et al. 2000, 2001) and by point-source discharges associated with mining and other industries. The general lack of impacted sites in our data set may result from the probability-based site selection approach used in this study. Karr and Chu (1999) argued that by randomly selecting sites, the investigator runs the risk of excluding those that are severely degraded, especially if they are rare. Additional samples targeted

at impacted sites should be added to existing data sets and used to develop new regional normalization models. Such an effort would result in a more robust assessment of stream conditions in the NLF ecoregion.

Explicit Statistical Modeling in Stream Assessment

The regional normalization modeling approach used in this study provided a useful means to account for significant covariates of fish assemblage and stream characteristics. This approach may be useful in probability-based assessments where reference sites are not selected a priori. This approach may also be useful in areas that are heavily impacted, where reference sites are sparse. Our modeling approach is similar to that used in RIVPACS-type assessments (Hawkins et al. 2000), the major difference is that RIVPACS uses reference sites selected a priori, but the regional normalization approach models site-specific reference conditions. Reynoldson et al. (1997) argued that modeling in general would improve indicator accuracy and precision by accounting for important site-specific drivers. It would be useful to compare the performance of RIVPACS versus regional normalization approaches.

The MLR models developed in this study suggested that a mix of catchment, riparian, and site factors are important in structuring the ecological character of NLF streams. Our findings add to a growing list of studies that show that streams are structured by processes operating across multiple scales (Wiley et al. 1997; Wehrly 1999; Isaak and Hubert 2001). In our MLR models of fish metrics stream temperature was a significant predictor variable in 9 of the 10 models. The importance of temperature on the distribution and abundance patterns of fishes and invertebrates is well known and is an important variable to account for when developing biologically based indicators of stream condition (Hawkins et al. 1997; Lyons et al. 1996; Wehrly et al. 2003).

The relative fit of the MLR models developed in this study was somewhat low ($R^2 = 0.18\text{--}0.57$) but within the range of those developed by Wiley et al. (2003) for streams in Michigan's Lower Peninsula. High R^2 values are not a prerequisite for this methodology (Wiley et al. 2003) because the predictive models are used to develop site expectations but do not affect observed metric values. Specification errors in those models could, however, lead to biases in the assessment. Unexplainable random error, on the other hand, probably

only affects sensitivity of the normalized score because deviations are scaled by the variance in predicted values. A feature of the regional normalization approach is that metrics with high intrinsic variability (i.e., that are hard to predict) require larger absolute deviations to generate poor normalized scores.

Low normalizing model fit may be attributed to several factors. First, Wiley et al. (2003) showed that fish data collected using one-pass sampling, as used in this study, yielded poorer reference models than fish data collected with more intensive (multiple pass) sampling methods. Additionally, the use of backpack shockers in small streams and generator-powered shockers in larger streams may have introduced additional variance in the fish data because of different efficiencies of the two gears. Second, certain variables (e.g., water temperature, base flow yield, and groundwater potential) identified as important in structuring stream fish assemblages (Wiley et al. 1997; Zorn et al. 2002; Wehrly et al. 2003) were unavailable for some sites and had to be estimated. Poor estimates of these variables may have contributed to the low R^2 in the reference models. Third, single-point samples of variables like water chemistry may be inadequate for correctly characterizing streams. Water chemistry varies seasonally and in response to precipitation events (Allan 1995), so point measures spread across sites over the sampling season may have reduced the predictive ability of the MLR models. In spite of these potential sources of error and data limitations, we successfully constructed statistically significant predictive models for a suite of important fish, water chemistry, and habitat variables.

Stream Assessment Using Regional Normalization

In this study, we arbitrarily set the threshold of impairment for all metrics at -2 SD from the regional average predicted by the normalization model (this average is represented as the zero on the y-axis on data plots; e.g., Figures 2, 3). We also interpreted normalized scores that were greater than $+2$ SD from the regional average as exceptional. Wiley et al. (2003) noted that the criterion used to judge whether sites are impaired is clearly arbitrary and suggested that ± 2 SD for evaluating stream sites could be thought of as associated with 95% confidence intervals, particularly if index values are relatively normally distributed. However, identification of appropriate thresholds necessary to protect streams in this region will require additional information on how

individual metrics respond to water quality impairment (this will require additional data targeted from impaired sites). The fact that agricultural and urban land use was significantly higher in catchments we judged as impaired, based on water chemistry metrics, supports the criterion of -2 SD for judging impairment.

Our analysis indicated that fish communities have not been greatly affected by land-use changes associated with urbanization and agricultural development. However, we recognize that current fish communities, physical habitat, and water chemistry in NLF streams have been altered from pre-European settlement. For example, Arctic grayling *Thymallus arcticus* was a native species in several streams in the eastern NLF but have been extirpated. In addition, rainbow trout, brown trout, and other salmonines are exotic species that were intentionally introduced to many streams in the NLF. In this context, our conclusions concerning the degree of impact from agricultural and urban land use are relative to present-day communities typical of this region rather than pre-European settlement condition.

Regional normalization is a useful tool that allows for a more accurate and integrative assessment of stream conditions. By incorporating important covariates of fish, chemistry, and habitat data in the reference models, we were able to account for sources of variability and correct for systematic biases in the indicator metrics. In addition, the reference models provided an explicit and testable hypothesis for reference condition at a particular site. Finally, unit standardization in terms of regional variance enabled us to compare scores from different metrics (fish, water chemistry, and physical habitat), which facilitated a more integrative assessment of stream conditions in the NLF ecoregion.

Acknowledgments

This project was funded through the USEPA, Regional Environmental Monitoring and Assessment Program (REMAP) and USEPA, Region 5 as assistance identification number 826206-01-0 and request number #9728GH0023. Special thanks to Ron Carlson, Arthur Lubin, Kevin Pierard, and Linda Holst, all from USEPA, for their assistance with administration of this project. Field assistance was provided by partners on this project: Paul Kanehl of WDNR, Bureau of Research; Scott Niemela, Mike Feist, Joel Chirhart, and Konrad Schmidt of Minnesota Pollution Control Agency; Margaret Watkins and John Johnson of the Grand

Portage Chippewa Nation; and Paul Michael Stewart and Jason Butcher of USGS, Biological Resources Discipline, Porter, Indiana. Water chemistry was analyzed at the USEPA, National Health and Environmental Effects Research Laboratory, Duluth, Minnesota. John Lyons provided constructive comments on the manuscript. The opinions expressed herein do not necessarily reflect the opinions of the U.S. Fish and Wildlife Service or USEPA, although portions of this study may have been funded wholly or in part by those agencies.

References

- Allan, J. D. 1995. Stream ecology: structure and function of running waters. Chapman and Hall, New York.
- Baker, M. E., M. J. Wiley, and P. W. Seelbach. 2001. GIS-based hydrological modeling of riparian areas: implications for stream water quality. *Journal of the American Water Resources Association* 37:1615–1628.
- Baker, M. E., M. J. Wiley, and P. W. Seelbach. 2003. A GIS model of subsurface water potential for aquatic resource inventory, assessment, and environmental management. *Environmental Management* 32:706–719.
- Butcher, J. T., P. M. Stewart, and T. P. Simon. 2001. Development of a benthic community index for the Northern Lakes and Forests ecoregion. U.S. Environmental Protection Agency, R-EMAP, Final Report for Grant 826206-01-0, Duluth, Minnesota.
- Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental Management* 10:199–214.
- Hawkins, C. P., J. N. Hogue, L. A. Decker, and J. W. Feminella. 1997. Channel morphology, water temperature, and assemblage structure of stream insects. *Journal of the North American Benthological Society* 16:728–749.
- Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10:1456–1477.
- Isaak, D. J., and W. A. Hubert. 2001. Production of stream habitat gradients by montane watersheds: hypothesis tests based on spatially explicit path analysis. *Canadian Journal of Fisheries and Aquatic Sciences* 58:1089–1103.
- Karr, J. R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6:21–27.
- Karr, J. R., and E. W. Chu. 1999. Restoring life in running waters: better biological monitoring. Island Press, Covelo, California.
- Lazorchak, J. M., D. L. Klemm, and D. V. Peck. 1998. Environmental Monitoring and Assessment Program Surface Waters: field operations and methods for measuring the ecological condition of Wadeable

- streams. U.S. Environmental Protection Agency, Report EPA/620/R-94/004F, Washington, D.C.
- Lyons, J. 1992a. Using the index of biotic integrity (IBI) to measure environmental quality in warmwater streams of Wisconsin. U.S. Forest Service, General Technical Report NC-149, Rhinelander, Wisconsin.
- Lyons, J. 1992b. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. *North American Journal of Fisheries Management* 12:198–203.
- Lyons, J., L. Wang, and T. D. Simonson. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. *North American Journal of Fisheries Management* 16:241–256.
- Lyons, J., R. R. Piette, and K. W. Niermeyer. 2001. Development, validation, and application of a fish-based index of biotic integrity for Wisconsin's large warmwater rivers. *Transactions of the American Fisheries Society* 130:1077–1094.
- Mundahl, N. D., and T. P. Simon. 1998. Development and application of an index of biotic integrity for coldwater streams of the Upper Midwestern United States. Pages 383–411 *in* T. P. Simon, editor. *Assessing the sustainability and biological integrity of water resources using fish communities*. CRC Press, Boca Raton, Florida.
- Norris, R. H. 1996. Predicting water quality using reference conditions and associated communities. Pages 32–52 *in* R. C. Bailey, R. H. Norris, and T. B. Reynoldson, editors. *Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach*. North American Benthological Society, Kalispell, Montana.
- Ohio Environmental Protection Agency. 1987. Biological criteria for the protection of aquatic life: Volume 1: The role of biological data in water quality assessment. Ohio Environmental Protection Agency, Division of Water Quality and Assessment Report, Columbus.
- Omernik, J. M., and A. L. Gallant. 1988. Ecoregions of the upper Midwest states. U.S. Environmental Protection Agency, EPA/600/3-88/037, Corvallis, Oregon.
- Overton, W. S., D. L. Stevens, and D. White. 1991. Design report for EMAP, Environmental Monitoring and Assessment Program. U.S. Environmental Protection Agency, EPA/600/3-91/053, Corvallis, Oregon.
- Poff, N. L. 1997. Landscape filters and species traits: toward mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society* 16:391–409.
- Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* 16:833–852.
- Seelbach, P. W., M. J. Wiley, P. A. Soranno, and M. T. Bremigan. 2002. Aquatic conservation planning: using landscape maps to predict ecological reference conditions for specific waters. Chapter 25 *in* K. Gutzwiller, editor. *Applying landscape ecology in biological conservation*. Springer-Verlag, New York.
- Simon, T. P., and R. Dufour. 1997. Development of index of biotic integrity expectations for the ecoregions of Indiana, V. Eastern corn belt plain. U.S. Environmental Protection Agency, EPA 905/R-96/002, Chicago, Illinois.
- Simonson, T. D., J. Lyons, and P. D. Kanehl. 1994. Guidelines for evaluating fish habitat in Wisconsin streams. U.S. Forest Service, General Technical Report NC-164, St. Paul, Minnesota.
- Soranno, P. A., K. E. Webster, J. L. Riera, T. K. Kratz, J. S. Baron, P. A. Bukavesckas, G. W. Kling, D. S. White, N. Caine, R. C. Lathrop, and P. R. Leavitt. 1999. Spatial variation among lakes within landscapes: ecological organization among lake chains. *Ecosystems* 2:395–410.
- Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. *Transactions of the American Geophysical Union* 38:913–920.
- Wang, L., J. Lyons, P. Kanehl, R. Bannerman, and E. Emmons. 2000. Watershed urbanization and changes in fish communities in southeastern Wisconsin streams. *Journal of the American Water Resources Association* 36:1173–1189.
- Wang, L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. *Environmental Management* 28:255–266.
- Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. M. Stewart. 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forests ecoregion, USA. *Canadian Journal of Fisheries and Aquatic Sciences* 60:491–505.
- Wehrly, K. E. 1999. The influence of thermal regime on the distribution and abundance of stream fishes in Michigan. Doctoral dissertation. University of Michigan, Ann Arbor.
- Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 1998. Landscape-based models that predict July thermal characteristics of lower Michigan rivers. Michigan Department of Natural Resources, Fisheries Research Report 2037, Ann Arbor.
- Wehrly, K. E., M. J. Wiley, and P. W. Seelbach. 2003. Classifying regional variation in thermal regime based on stream fish community patterns. *Transactions of the American Fisheries Society* 132:18–38.
- Wiley, M. J., and J. Martin. 2000. Current conditions, recent changes, and major threats to the Huron River: a report on eight years of citizen monitoring. Huron River Watershed Council, Report, Ann Arbor, Michigan. Available: <http://rivers.snre.umich.edu/mri/huron/hrwcrpt/index.htm>. (October 2003)
- Wiley, M. J., S. L. Kohler, and P. W. Seelbach. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. *Freshwater Biology* 37:133–148.

- Wiley, M. J., P. W. Seelbach, K. E. Wehrly, and J. S. Martin. 2003. Regional ecological normalization using linear models: a meta-method for scaling stream assessment indicators. Chapter 12 *in* T. P. Simon, editor. *Biological response signatures: indicator patterns using aquatic communities*. CRC Press, Boca Raton, Florida.
- Wright, J. F., M. T. Furse, and P. D. Armitage. 1993. RIVPACS: a technique for evaluating the biological water quality of rivers in the UK. *European Water Pollution Control* 3:15–25.
- Zorn, T. G., P. W. Seelbach, and M. J. Wiley. 2002. Distributions of stream fishes and their relationship to stream size and hydrology in Michigan's Lower Peninsula. *Transactions of the American Fisheries Society* 131:70–85.