APPLIED ISSUES

Implications of community concordance for assessing stream integrity at three nested spatial scales in Minnesota, U.S.A.

CHRISTINE L. DOLPH*, DAVID D. HUFF[†], CHRISTOPHER J. CHIZINSKI[‡] AND BRUCE VONDRACEK[§]

Water Resources Science Program, University of Minnesota, St. Paul, MN, U.S.A.
[†]N. O. A. A., National Marine Fisheries Service, University of California, Santa Cruz, Southwest Fisheries Science Center, Santa Cruz, CA, U.S.A.
[‡]U.S. Geological Survey, Nebraska Fish and Wildlife Cooperative Research Unit, School of Natural Resources, University of Nebraska, Lincoln, NE, U.S.A.
[§]U.S. Geological Survey, Minnesota Fish and Wildlife Cooperative Research Unit[1], University of Minnesota, St. Paul, MN, U.S.A.

SUMMARY

1. Fish and invertebrate assemblage data collected from 670 stream sites in Minnesota (U.S.A.) were used to calculate concordance across three nested spatial scales (statewide, ecoregion and catchment). Predictive taxa richness models, calibrated using the same data, were used to evaluate whether concordant communities exhibited similar trends in human-induced taxa loss across all three scales. Finally, we evaluated the strength of the relationship between selected environmental variables and the composition of both assemblages at all three spatial scales.

2. Significant concordance between fish and invertebrate communities occurred at the statewide scale as well as in six of seven ecoregions and 17 of the 21 major catchments. However, concordance was not consistently indicative of significant relationships between rates of fish and invertebrate taxa loss at those same scales.

3. Fish and invertebrate communities were largely associated with different environmental variables, although the composition of both communities was strongly correlated with stream size across all three scales.

4. Predictive taxa-loss models for fish assemblages were less sensitive and precise than models for invertebrate assemblages, likely because of the relatively low number of common fish taxa in our data set. Both models, however, distinguished reference from non-reference sites.

5. The importance of concordance, geographic context and scale are discussed in relation to the design and interpretation of stream integrity indicators. In particular, our findings suggest that community concordance should not be viewed as a substitute for an evaluation of how assemblages respond to environmental stressors.

Keywords: canonical correspondence analysis, fish, macroinvertebrates, procrustes, RIVPACS.

Introduction

'Ecological integrity' has become a key concept and management goal in natural resource disciplines, particularly among water resource scientists (Karr & Chu, 2000; Norris & Hawkins, 2000; Karr & Yoder, 2004). This term has been defined in a number of ways (Boulton, 1999; Karr, 1999; Norris & Thoms, 1999) but generally refers to whether aquatic systems sustain the levels of biodiversity and ecosystem function comparable to that which they provided prior to modern human civilisation (Karr & Dudley, 1981; Karr, 1999). Because funding for conservation planning is limited, water quality monitoring programmes have typically sought to provide cost-effective assessments of stream integrity by sampling a small number of aquatic assemblages (e.g. Barbour *et al.*, 1999; Paavola *et al.*, 2003; Bilton *et al.*, 2006). These assemblages are used as 'indicators' or 'surrogates' for the condition of the overall biodiversity and/or functionality of an ecosystem (Flather *et al.*, 1997; Lawler *et al.*, 2003; Lewandowski, Noss & Parsons, 2010).

Despite widespread use in stream monitoring, the degree to which indicator taxa actually reflect trends among other organisms is generally not well known (Paavola *et al.*, 2003; Bilton *et al.*, 2006). Recently, a growing body of literature has sought to address this question by quantifying the degree of concordance between various assemblages (Heino, 2010; Johnson & Hering, 2010 and references therein). Community concordance refers to the extent to which different assemblages exhibit similar spatial variation in community structure (*sensu* Jackson & Harvey, 1993). Because it is based on species identities (rather than just numbers of species), concordance can provide a more comprehensive picture of community similarity than richness measures alone (Su *et al.*, 2004; Pawar *et al.*, 2007).

Community concordance between assemblages can arise from a number of factors, including shared responses to the same environmental drivers (Kilgour & Barton, 1999; Heino *et al.*, 2005; Pawar *et al.*, 2007; Heino, 2010), co-losses of sensitive species in response to environmental stress (Mykrä *et al.*, 2008a; Yates & Bailey, 2010) and strong biotic interactions (Jackson & Harvey, 1993; Heino, 2002; Johnson & Hering, 2010). Concordance can also be affected by the life-history traits of stream assemblages, for example, the dispersal and reproductive capabilities of various taxa may govern their access to beneficial environments, thereby limiting the extent to which environmental gradients and biotic interactions might otherwise drive spatial patterns in community composition (Moritz *et al.*, 2001; Townsend *et al.*, 2003; Campbell Grant, Lowe & Fagan, 2007; Pawar *et al.*, 2007; Grenouillet *et al.*, 2008; Shurin, Cottenie & Hillebrand, 2009). Finally, community concordance among major taxonomic groups may be greater at broader spatial scales, because large-scale patterns in the composition of these groups are probably controlled by the same regional environmental gradients (Paavola *et al.*, 2006).

Because of the potentially complex nature of community concordance, it is unclear whether concordant communities might exhibit similar responses to environmental disturbance and thus constitute appropriate surrogates for one another in the context of stream assessment. If we are to use evidence of concordance to make generalisations about the relationships between different organism groups – especially for the purposes of designing biological monitoring programmes – it is necessary to establish (i) whether concordant communities indicate similar levels of stream integrity when evaluated using standard biological indices and (ii) whether concordance among multiple taxonomic groups arises from a shared response by these groups to certain environmental stressors. To our knowledge, however, only Mykrä *et al.* (2008a) have explicitly evaluated how patterns in community concordance relate to patterns in the biotic indices that are typically used to monitor aquatic systems. Efforts to resolve the implications of community concordance for the biological assessment of streams must also address the issue of scale (Johnson *et al.*, 2007). For example, the use of one taxonomic group as an indicator for the condition of others may not be appropriate if concordance does not occur at the same scale at which biotic indices are designed (Paavola *et al.*, 2006). Moreover, understanding the response of different communities to local-, catchment- and regional-scale environmental drivers is a critical prerequisite to the effective design of conservation programmes, which are likely to alleviate biological impairment only insofar as they address stressors on a scale to which organisms readily respond.

The goal of this study was to evaluate whether significantly concordant communities yielded equivalent indications of stream integrity at different spatial

scales. Specifically, we examined how concordance between fish and invertebrate assemblages varied among 670 stream sites that occurred across three nested spatial scales (statewide, ecoregion and catchment) in Minnesota, U.S.A.. We then developed predictive models (e.g. RIVPACS or taxa-loss models; Wright, 2000; Hawkins, 2006) for fish and invertebrates and evaluated whether concordant communities exhibited similar degrees of taxa loss at each of the three spatial scales. Finally, we evaluated whether the spatial patterns observed for fish and invertebrate assemblage composition could be explained by similar environmental variables. We hypothesised that (i) concordance between fish and invertebrate communities would be strongest at broader spatial scales and weaker at narrower scales, (ii) concordant communities would respond to similar environmental variables and (iii) fish and invertebrate communities that exhibited concordance at a given scale would exhibit similar responses to environmental stress and would therefore show similar degrees of departure from reference conditions as quantified by taxa-loss models.

Methods

Study area and data sources

The state of Minnesota is characterised by seven ecological regions (i.e. 'ecoregions'; Omernik, 1987) and 81 major catchments. The term 'major catchment' refers to a drainage area assigned an eight-digit Hydrologic Unit Code (HUC) according to the classification system developed by the U.S. Geological Survey (USGS, 1982). These catchments define drainage areas for the state's large river (i.e. generally fourth–sixth order) and lake systems and were established historically based on a size criterion, with each catchment constituting an area of roughly 1800 km^2. Some of the catchments exclude upstream contributing areas and thus are most accurately described as 'administrative' rather than 'true' catchments. These 81 catchments now constitute the primary spatial scale at which state agencies implement water quality assessment and conservation efforts.

The southern and western ecoregions of Minnesota (Driftless Area, Western Cornbelt Plains, Northern Glaciated Plains and Red River Valley) are predominantly agricultural, whereas the northern regions (Northern Lakes and Forests and Northern Minnesota
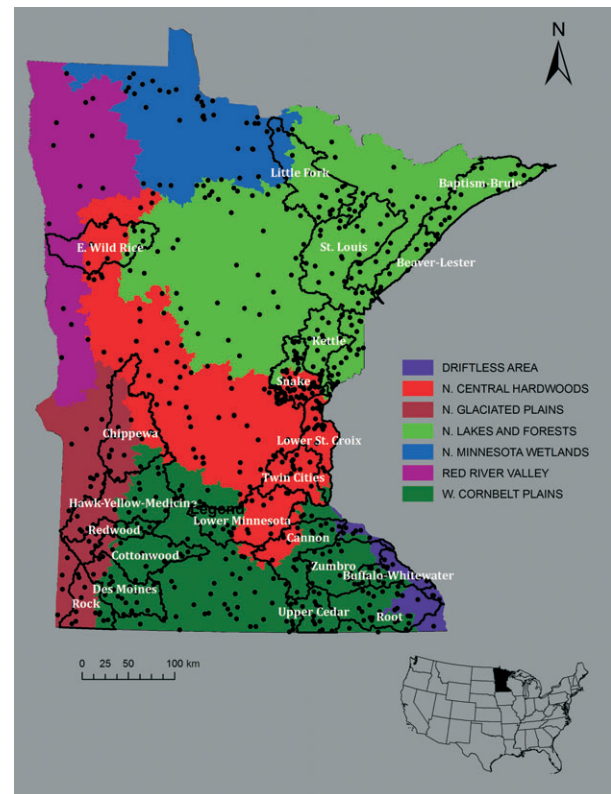

Fig. 1 Stream sites (closed circles, $n = 670$) from which fish and invertebrate samples were collected by the Minnesota Pollution Control Agency between 1996 and 2006. Black outlines indicate boundaries of major catchments (eight-digit hydrologic unit code); coloured areas are level III Omernik (1987) ecoregions.

Wetlands) are predominantly a mix of forests and wetlands. The North Central Hardwoods region is characterised by widespread urban and suburban development associated with the Minneapolis-St. Paul metropolitan area, as well as by woodland and crops.

Fish and benthic macroinvertebrate data used in this study were collected from 670 stream sites in Minnesota between 1996 and 2006 (Fig. 1). This data set was aggregated from a number of independent studies conducted by the Minnesota Pollution Control Agency (MPCA) and the U.S. Environmental Protection Agency within Minnesota streams. While some of these studies adhered to a randomised study site selection, others targeted specific sites known to occur along a gradient of human disturbance and to represent a range of stream sizes and habitat types. In addition, the Snake River catchment was intensively sampled as a pilot project in the MPCA's recently established catchment-based biomonitoring programme. As a result, while the sampling sites included in

this study occur throughout the state of Minnesota, not all areas of the state are equally represented. Both fish and macroinvertebrates were collected once from each site during base-flow conditions from mid-June to mid-October. All samples were collected from sampling reaches equal to 35 times the mean wetted width of a stream or to a maximum distance of 500 m.

Macroinvertebrates were collected using a D-Frame dip-net with 500-μm mesh. Each sample represented a composite of 20 sampling efforts divided proportionately among five types of productive habitats present at the sampling reach (riffles, undercut banks, submerged/emergent vegetation, snags/woody debris and leaf packs). For each sampling effort, a 30-cm^2 area of substratum immediately upstream of the D-net was disturbed. Samples were collected in a downstream to upstream direction until all major habitat types contained within the reach were sampled and preserved in 100% denatured ethanol. In the laboratory, macroinvertebrate samples were subsampled using a 61 × 61 cm gridded screen tray divided into 144 5.1 cm^2 squares. Each sample was spread evenly across this grid, and organisms were picked from randomly selected grid squares until a minimum of 300 organisms were collected or until all individual specimens had been picked from the sample (if the sample contained <300 individuals). All specimens were identified by the MPCA to the lowest possible taxonomic resolution, typically genus. We eliminated ambiguous taxa in invertebrate samples (i.e. taxa identified to more than one taxonomic level) by aggregating or eliminating taxa as described by Yuan, Hawkins & Van Sickle (2008), resulting in a set of unambiguous operational taxonomic units (OTUs; Ostermiller & Hawkins, 2004). We sought to standardise abundance counts to similar values across all samples by randomly subsampling invertebrate samples with >300 individuals such that each sample contained exactly 300 individuals. Not every invertebrate sample in our data set contained as many as 300 individuals; however, all samples did contain at least 200 individuals. Ultimately, our data set consisted of 426 sites (64%) with 300 individuals and 244 (36%) with between 200 and 299 individuals.

Fish were captured using one of four different types of electrofishing gear, depending on stream size. Fish habitat was sampled in proportion to its availability within the stream reach, with greater electrofishing time spent sampling proportionally more abundant habitat types. In smaller streams (<3 m wide), virtually all fish habitat could be sampled, whereas in larger streams MPCA biologists were constrained to weave among habitat types. All captured fish were sorted into buckets and identified to species. The number of fish captured ranged from 1 to 4275 individuals. Fish samples containing 160 or more individuals have been shown to yield biological index scores that are resistant to the effects of sampling error; however, collections this large may not be achievable for highly degraded stream sites (Dolph *et al.*, 2010). Of the 670 sites in our data set, 435 (65%) contained 160 fish or more and 235 (35%) contained fewer than 160 fish. The mean number of fish per sample was 444. Because we wished to include both degraded and intact streams in our analysis, we did not exclude sites based on fish abundance. Fish abundance values were standardised to the length of stream reach sampled; across all stream sites, the number of fish per 100 m ranged from 0.7 to 90.

For concordance and pCCA analysis, the abundance of fish and macroinvertebrates was log(x + 1)-transformed to reduce the influence of abundant taxa (McCune & Grace, 2002). For the development of RIVPACS models, fish and macroinvertebrate abundance values were converted to presence/absence.

Community concordance

To analyse community concordance, we organised fish and invertebrate data at three nested spatial scales: (i) statewide, (ii) ecoregion and (iii) major catchment. We chose to evaluate concordance at the ecoregion and catchment scales because the composition of stream communities is often thought to respond to gradients operating at these scales (Allan, 2004; Pyne, Rader & Christensen, 2007). Although ecosystems are clearly not organised according to state boundaries, we selected the statewide scale as the broadest geographic unit for our concordance analysis because stream monitoring programmes in the United States are administered primarily at this scale. We limited our analysis of concordance to those spatial 'units' (i.e. individual catchments or ecoregions) where samples from at least 10 stream sites were available (Fig. 1).

We used non-metric multidimensional scaling (NMDS) based on Bray–Curtis (Sørensen) ranked distances between stream samples to reveal patterns in fish and invertebrate assemblages. For each spatial

unit, we determined the appropriate number of dimensions for each NMDS solution by computing stress values. Clarke (1993) suggested that stress values <0.2 correspond to reasonably good ordinations, but that stress values <0.1 minimise the risk of false inferences. We generated NMDS configurations from 10 random starts at each level of dimensionality from one to five and selected the lowest number of dimensions above which stress values did not decrease appreciably. In selecting a final NMDS solution for each assemblage in each spatial unit, we avoided solutions based on local optima by performing the analysis with a maximum of 100 random starts until a stable solution was reached. If no stable solution was reached after 100 starts, we selected the solution with the lowest stress value.

We used a Procrustean analysis (Peres-Neto & Jackson, 2001) to evaluate the degree and significance of concordance between fish and invertebrate NMDS configurations at each spatial scale. In Procrustes analysis, one (NMDS) configuration is kept fixed as reference, whereas the other is rotated, translated, reflected, dilated and scaled until the sum-of-squared residuals (m^2) between corresponding coordinates in both configurations is minimised. Lower m^2 values indicate higher degrees of concordance. A correlation-like statistic (R) can also be derived from m^2 values and is calculated as $R = \sqrt{(1 - m^2)}$. We evaluated the statistical significance of the Procrustean fit using Protest, which repeatedly ($n = 999$) randomises the configuration of one matrix and recalculates m^2. The percentage of m^2 values ≤ the observed m^2 provides the significance level (i.e. the P-value) of the test. Heino (2010) suggests that strong concordance between multiple organism groups is indicated by R values >0.7 (or m^2 values <0.5).

Partial canonical correspondence analysis

We determined whether the compositions of concordant fish and invertebrate assemblages were related to similar environmental gradients using partial canonical correspondence analysis (pCCA; Birks, Peglar & Austin, 1994; ter Braak & Verdonschot, 1995; Legendre & Legendre, 1998) for each assemblage type at each of the three spatial scales. Because we were primarily interested in evaluating whether concordant communities were related to similar environmental variables, we restricted the pCCA analysis to those

spatial units in which significant community concordance had been found. We performed pCCA for fish and invertebrate communities in each of the spatial units using a set of 23 candidate environmental variables (Table 1) while controlling for the effects of seasonal and annual variation by treating the day (i.e. calendar day, 1–365) and year a site was sampled as covariables. All environmental variables were either collected at each stream site by the MPCA or compiled from publicly available spatial data sets and transformed for normality when necessary.

The number of environmental variables was large relative to the number of stream sites in some spatial units, particularly at the catchment scale. Therefore, we used forward selection to determine the most significant environmental variables in each pCCA model (ter Braak & Verdonschot, 1995), including additional variables until either (i) the resulting additional canonical eigenvalues were no longer significant ($P > 0.05$) when the previous canonical axes were treated as covariables or (ii) a total of up to five variables had been included. Although the decision to limit the pCCA models to five variables was arbitrary, our objective was to identify a limited number of environmental parameters that were the most relevant to the composition of fish and invertebrate communities.

We tested the significance of each canonical eigenvalue and each overall pCCA model using Monte Carlo tests with 999 permutations (Legendre & Legendre, 1998). In pCCA, canonical axes with eigenvalues that are statistically significant and greater than 0.30 can generally be considered important environmental determinants of community composition (ter Braak & Verdonschot, 1995).

Predictive model development

We developed separate predictive models for fish and invertebrate assemblages using the 'best-subsets' method developed by Van Sickle, Huff & Hawkins (2006). We identified 'best available' reference sites (*sensu* Stoddard *et al.*, 2006) using an index designed by the MPCA to quantify the degree of human disturbance acting upon a stream site (Table 2). The highest index score a site could receive was 81, which indicated a minimal degree of alteration from human activity. Streams in each ecoregion were designated as reference sites if disturbance scores fell within the upper quartile of scores for that ecoregion. The total

Table 1 Environmental variables used in pCCA[1] and predictive model development[2]. Unless a specific data transformation is noted, no transformation was performed

Variable	Description / Unit	Source
Ecoregion[2]	Seven categories; six dummy variables	Omernik (1987)
Catchment area[1,2]	Land area (km^2) draining into sampling reach; log10-transformed	Digital Elevation Models
Stream gradient[1,2]	Change in stream elevation over the sampling reach (m km^{-1}); log10-transformed	1 : 24 k scale USGS topographic maps
Air temperature[1,2]	Annual mean, 1961–1990	NRCS; Daly *et al.* (2000)
Precipitation (average)[1,2]	Annual mean, 1961–1990	NRCS; Daly *et al.* (2000)
Precipitation (seasonal)[1,2]	Precipitation accumulated through 1st October of year in which assemblage data were collected; log10-transformed	Minnesota Climatology Working Group
Year[1,2]	Year in which assemblage data were collected	MPCA
Calendar Day[1,2]	Day of the year (1–365) on which assemblage data were collected	MPCA
Latitude[1,2]		MPCA
Longitude[1,2]		MPCA
Thalweg depth[1,2]	Mean thalweg depth of sampling reach	MPCA
Stream depth[1]	Mean depth of sampling reach	MPCA
Stream width[1]	Mean bankfull width of sampling reach; log10-transformed	MPCA
Flow[1]	m^3 s^{-1}; log10-transformed	MPCA
Water temperature[1,†]	°C	MPCA
Conductivity[1,†]	μmho cm^{-1} at 25 °C; square-root-transformed	MPCA
Dissolved oxygen[1,†]	mg L^{-1}	MPCA
pH[1,†]	Power-transformed (=4)	MPCA
% Fines[1]	% of stream substratum covered by fine sediments; arcsine-square-root-transformed	MPCA
Nitrogen[1,†]	mg L^{-1}; 1/square-root-transformed	MPCA
Phosphorus[1,†]	mg L^{-1}; log10-transformed	MPCA
% Riparian disturbance[1]	% of 100-m riparian area adjacent to sampling reach affected by human activities	MPCA
% Forest[1]	% of catchment draining into sampling reach characterised by forest; arcsine-square-root-transformed	2001 NLCD; Homer *et al.* (2007)
% Wetland[1]	% of land use in catchment draining into sampling reach characterised by wetlands; arcsine-square-root-transformed	2001 NLCD; Homer *et al.* (2007)
% Agriculture[1]	% of land use in catchment draining into sampling reach characterised by agriculture (row crops); arcsine-square-root-transformed	2001 NLCD; Homer *et al.* (2007)
% Urban[1]	% of land use in catchment draining into sampling reach characterised by urban land use; arcsine-square-root-transformed	2001 NLCD; Homer *et al.* (2007)

NLCD, National land cover database; NRCS, Natural resource conservation service, U.S. Department of Agriculture; pCCA, partial canonical correspondence analysis.
†Denotes one-time grab samples collected by the MPCA.

number of reference sits was 169, and disturbance scores among these sites ranged from 43 to 81 (mean = 64). We randomly selected approximately half of these reference sites (*n* = 88, hereafter referred to as calibration sites) to develop predictive taxa-loss models; the other half was used to test predictive model performance (*n* = 81, hereafter referred to as validation sites). The remaining 501 sites were treated as non-reference, or test sites. Total disturbance scores at test sites ranged from 5 to 78, with a mean of 47.

For each assemblage, we calculated Sørensen dissimilarities between all pairs of reference sites in the calibration data set, based on the presence/absence of

taxa found at one or more of these sites. A flexible *β*-algorithm (*β* = −0.6) was used to cluster calibration sites based on their dissimilarity (Legendre & Legendre, 1998; McCune & Grace, 2002), and the resulting cluster dendrogram was pruned to identify distinct biological groups of reference sites. Because we used different criteria for selecting reference sites in each ecoregion, we sought to determine whether disturbance levels were consistent within reference site groups by using boxplots to evaluate the range of total disturbance scores within each group.

A subset of 11 environmental variables (Table 1) was used to predict the probability that a stream site

Table 2 Metrics used to determine total disturbance index scores for each stream site. Primary metrics[1] receive a minimum and maximum of 0 and 10 points, respectively, with least altered values receiving higher scores. Secondary metrics[2] adjust the total score by one point each. All secondary metrics could only result in a downward adjustment of the score (i.e. points lost from the total), except that one additional point was added for sites that were characterised by a number of road crossings that fell within the lower quartile of all sites. Total index scores were calculated by combining primary and secondary metrics. 'Human activity' refers to any type of human land use or human-made structure (i.e. buildings, fences, removal of vegetation, agricultural land use)

Metrics
% agricultural land use [1, †]
% impervious surfaces [1, †]
Number of animal units per km² [1, †]
Number of point sources per km² [1, †]
% riparian habitat affected by human activity [1, †]
% of stream channel that is channelised [1, †]
% sampling reach that is channelised [1, ‡]
% of 30-m riparian corridor adjacent to sampling reach affected by human activity [1, ‡]
% riparian corridor characterised by agricultural land use [2, ‡]
% agricultural land use with >3% slope [2, †]
Number of road crossings per stream km [2, †]
Number of feedlots per km² [2, †]
Urban area immediately upstream of sampling reach [2, ‡]
Feedlot immediately upstream of sampling reach [2, ‡]
Continuous discharge present <8 stream km from sampling reach [2, ‡]

[†]Values pertain to catchment area draining into the sampling reach.
[‡]Values pertain to sampling reach.

belonged to each reference site group. These 11 variables were chosen from the larger group of 23 because they tend to be resistant to human activities. Probability of group membership was combined with taxa occurrence probabilities within reference site groups to estimate the occurrence probability for each taxon at each steam site. The total number of taxa expected to occur at each site (*E*) was calculated by summing the probabilities of all taxa predicted to occur with a probability ≥ 0.5. Restricting the pool of expected taxa to those with probabilities of capture ≥0.5 has been shown to eliminate variability associated with the presence or absence of rare taxa (Van Sickle, Larsen & Hawkins, 2007).

The aim of the predictive model is to make site-specific predictions about the number of expected taxa by explaining natural variability in species richness across the validation data set. Thus, predictive model performance can be evaluated by determining the

reduction of the root mean squared error (RMSE) of the O/E values at validation sites below the RMSE of a null model richness ratio (O_N/E_N; Van Sickle *et al.*, 2005). The null model predicts the probability of occurrence for each taxon as the proportion of all calibration sites at which that taxon was found; it does not therefore explain any of the natural variability in species richness across reference sites.

Predictive model performance can also be evaluated by how closely the standard deviation (SD) of O/E among calibration sites approaches a replicate-sampling SD (SD_R; Van Sickle *et al.*, 2005). The SD_R represents the maximum precision (i.e. minimum variability) in O/E scores that a model could produce given random sampling variability in biological samples; SD_R at calibration sites therefore provides a theoretical lower limit for SD(O/E) that could be achieved by any predictive model (Van Sickle *et al.*, 2005). RMSE(O/E) at validation sites, RMSE(O_N/E_N), SD(O/E) at calibration sites and SD_R are all included as outputs in the best-subsets approach (Van Sickle, Huff & Hawkins, 2006).

We tested the congruence between fish and invertebrate O/E scores at each spatial unit by conducting linear regression between the two score types and using the coefficient of determination (R^2 values) to evaluate the strength of the linear relationship. Finally, we gauged the sensitivity of each model (i.e. ability to detect changes in community composition resulting from human activity) by evaluating whether O/E scores at non-reference sites differed significantly from those at reference sites using an unpaired, one-sided *t*-test.

Statistical programming

Data analysis was performed in *R* (version 2.12.0). We used the *metaMDS*, *protest*, *cca* and *permutest* functions (vegan: community ecology package; Oksanen *et al.*, 2009) to perform NMDS, Procrustean analysis, pCCA and permutation tests for each assemblage at each scale, respectively. For predictive model development and selection, we used a set of *R* scripts modified from those originally developed by Van Sickle *et al.* (2006).

Results

Community concordance

NMDS ordinations with three dimensions had stress values <0.2, indicating a good fit for both taxonomic

groups at all three spatial scales (e.g. Fig. 2). Spatial units with more stream sites tended to have higher stress values, and invertebrate solutions had higher stress values than fish solutions. Concordance between fish and invertebrate assemblages was significant at the statewide scale, in six of seven ecoregions and in 17 of 21 major catchments (Table 3). In most cases, concordance was relatively weak (R values <0.7). However, concordance was relatively strong ($R \geq 0.7$) in four of 21 catchments and in the Red River Valley ecoregion. The highest concordance (i.e. lowest m^2 and highest R values) occurred in the Rock catchment ($R = 0.80$).

Environmental gradients affecting concordant communities

Of the 23 candidate environmental variables in our data set, fish and invertebrate assemblage composition was often strongly associated with some aspect of stream size (e.g. stream width or catchment area) across all three spatial scales. For fish communities, at least one of these parameters was among the first or second environmental variables selected in pCCA analyses at the statewide scale, in five of six ecoregions and in 10 of 17 catchments (see Appendix S1). Similarly, stream width or catchment area was among the top two variables most correlated with invertebrate assemblage composition in five of six ecoregions and in eight of 17 catchments. Catchment

area was also among the five environmental variables selected by pCCA for invertebrate communities statewide.

At the statewide scale, stream gradient and % agricultural land cover were identified as important variables for both fish and invertebrate communities (see Appendix S1). Fish communities were also associated with geographic gradients (i.e. latitude and longitude), whereas invertebrates were associated with average air temperature and stream flow. At ecoregion and catchment scales, fish and invertebrate pCCA models often shared no more than two variables (including catchment area or stream width); two exceptions were the Snake catchment, where fish and invertebrates were both associated with catchment area, stream gradient and % urban land use, and the Northern Glaciated Plains ecoregion, where both fish and invertebrate communities were associated with % fines, latitude and some aspect of precipitation. Fish and invertebrate communities were associated with longitude in the Northern Lakes and Forests ecoregion and with dissolved oxygen in the Northern Minnesota Wetlands ecoregion. In addition, both assemblages were associated with % forest in the Baptism-Brule catchment, with pH in the Root catchment and with % fines in the Rock catchment. Apart from stream size, fish assemblages across all scales were most commonly associated with stream gradient, % forest, % agriculture, longitude and nitrogen concentrations, whilst invertebrate assemblages were
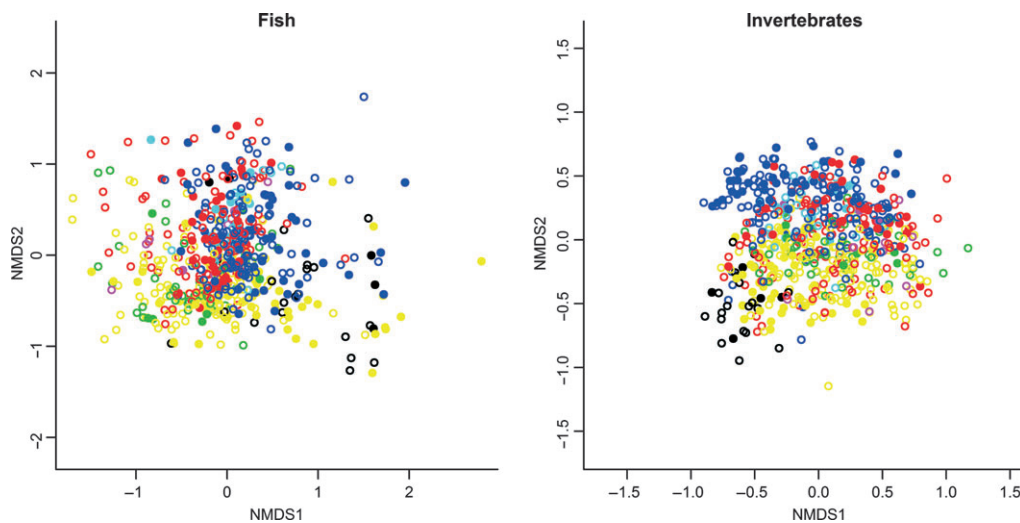

Fig. 2 NMDS ordination solutions for statewide invertebrate and fish data sets. Only two of three dimensions are shown. Stress values = 0.19 (invertebrates); 0.18 (fish). Solid circles are reference sites; open circles are test sites. Colours refer to the ecoregion within which each site is located: black, Driftless Area; red, Northern Central Hardwoods; blue, Northern Lakes and Forests; yellow, Western Cornbelt Plains; Turquoise, Northern Minnesota Wetlands; green, Northern Glaciated Plains; pink, Red River Valley.

Table 3 Concordance between fish and macroinvertebrate communities, and congruence between fish and invertebrate O/E scores, at three spatial scales. Concordance was performed using non-metric multidimensional scaling (NMDS) followed by Procrustes rotation and Protest significance test with 999 permutations. Significant concordance between fish and macroinvertebrate communities and significant relationships ($P < 0.5$) between fish and invertebrate O/E scores are indicated in bold. Note that lower m^2 values indicate higher concordance

Scale	Unit	*n*	Concordance			Congruence in O/E scores	
			m^2	R	P	R^2	P
Statewide	Minnesota	611	**0.6909**	**0.5560**	**<0.001**	**0.04**	**<0.001**
Ecoregion	Driftless Area	24	0.8635	0.3695	0.152	0.12	0.098
	Northern Central Hardwoods	142	**0.7609**	**0.4890**	**0.001**	**0.04**	**0.016**
	Northern Glaciated Plains	54	**0.7980**	**0.4494**	**0.001**	**0.09**	**0.026**
	Northern Lakes & Forests	165	**0.6586**	**0.5843**	**0.001**	0.00	0.993
	Northern Minnesota Wetlands	35	**0.7270**	**0.5225**	**<0.001**	**0.11**	**0.048**
	Red River Valley	11	**0.5305**	**0.6852**	**0.012**	0.18	0.199
	Western Cornbelt Plains	180	**0.7081**	**0.5403**	**0.001**	**0.09**	**<0.001**
Catchment	Baptism-Brule	18	**0.5493**	**0.6713**	**0.001**	0.13	0.138
	Beaver-Lester	10	0.7793	0.4698	0.542	0.25	0.139
	Buffalo-Whitewater	18	0.8864	0.3371	0.490	0.06	0.309
	Cannon	19	**0.5845**	**0.6446**	**0.001**	0.11	0.161
	Chippewa	16	**0.6283**	**0.6097**	**0.002**	0.01	0.664
	Cottonwood	14	**0.6953**	**0.5520**	**0.001**	0.18	0.149
	Des Moines Headwaters	23	**0.6035**	**0.6297**	**0.001**	0.07	0.208
	Eastern Wild Rice	10	**0.7795**	**0.4696**	**0.027**	0.06	0.502
	Hawk-Yellow-Medicine	16	**0.6237**	**0.6134**	**0.004**	0.18	0.100
	Kettle	14	**0.6488**	**0.5926**	**0.015**	0.00	0.999
	Little Fork	14	**0.6664**	**0.5776**	**0.011**	0.23	0.084
	Lower Minnesota	18	**0.5392**	**0.6788**	**0.001**	**0.27**	**0.027**
	Lower St. Croix	18	**0.6414**	**0.5988**	**0.001**	0.03	0.531
	Redwood	11	**0.5229**	**0.6907**	**0.005**	0.20	0.170
	Rock	13	**0.3544**	**0.8035**	**0.001**	0.08	0.344
	Root	27	**0.7566**	**0.4934**	**0.002**	0.03	0.374
	Snake	59	**0.6837**	**0.5624**	**0.001**	0.00	0.897
	St. Louis	23	**0.6430**	**0.5975**	**0.001**	0.10	0.132
	Twin Cities	12	0.8478	0.3901	0.633	0.00	0.852
	Upper Cedar	13	0.8189	0.4256	0.338	0.13	0.219
	Zumbro	18	**0.6303**	**0.6080**	**0.002**	0.04	0.448

frequently associated with % fines, stream gradient and stream flow across all scales.

Predictive model performance and congruence in O/E scores

We identified eleven and seven reference site groups from the cluster analysis of the invertebrate and fish assemblages, respectively (Fig. 3). The final set of reference site groups for invertebrates was generated by pruning the dendrogram at a height of *c.* 1.3 and for fish by pruning at a height of *c.* 2.9. Reference sites in invertebrate groups six and eight exhibited the highest variability in total disturbance scores; other groups were less variable (Fig. 4). For the fish classification, groups four, five and six exhibited high

variability in total disturbance score, whereas groups one, two, three and seven were less variable (Fig. 4).

Discriminant function models for fish and invertebrate assemblages used catchment area and stream gradient to differentiate reference site groups, in agreement with the pCCA results. Both models also included the year a stream was sampled. Other variables included in the invertebrate model were ecoregion and latitude, whereas the fish assemblage model included average annual air temperature, precipitation totals during the sampling year and mean thalweg depth.

The invertebrate data set contained 15 taxa that occurred in ≥0.5 of all calibration reference sites (i.e. 15 null model taxa), and the fish data set had five taxa. Both the fish and the invertebrate predictive
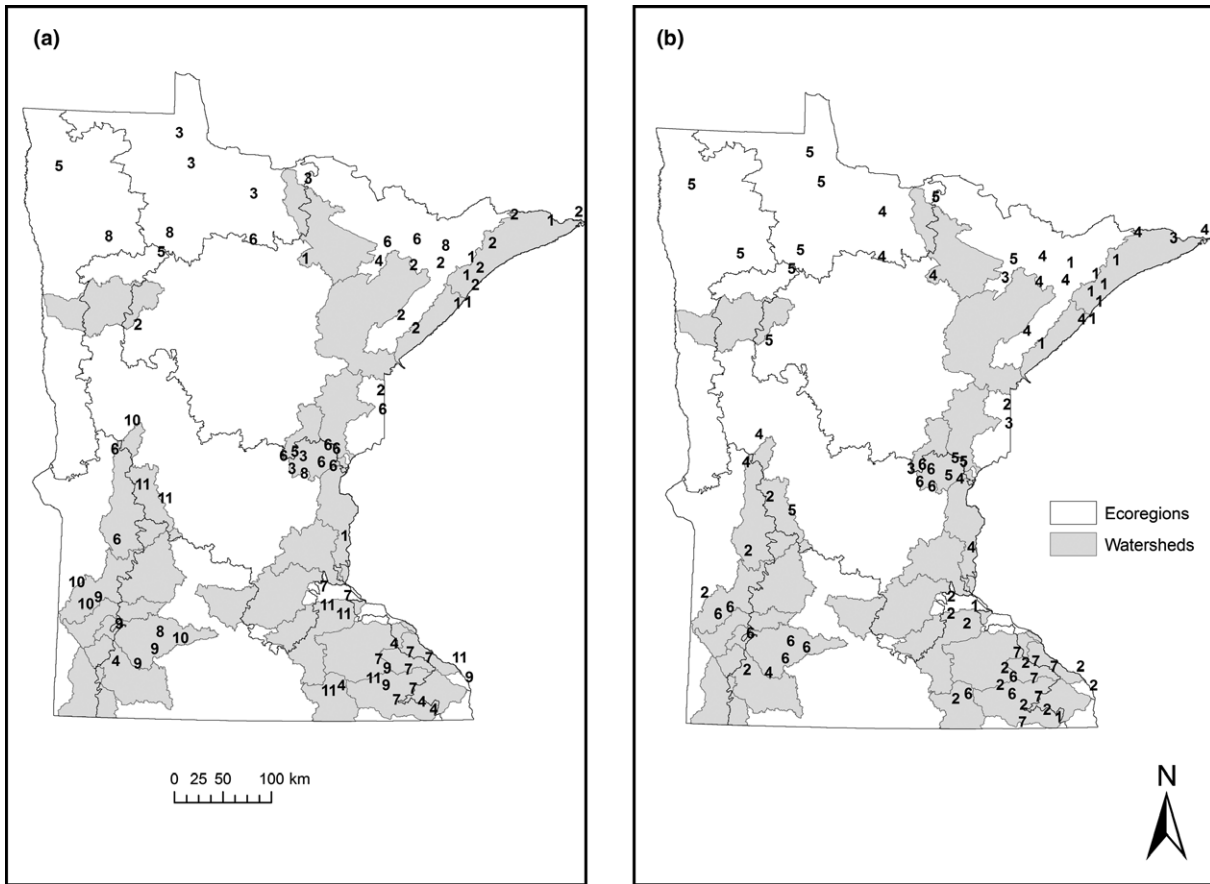
Fig. 3 Reference site groups identified during predictive model development for (a) invertebrates and (b) fish.
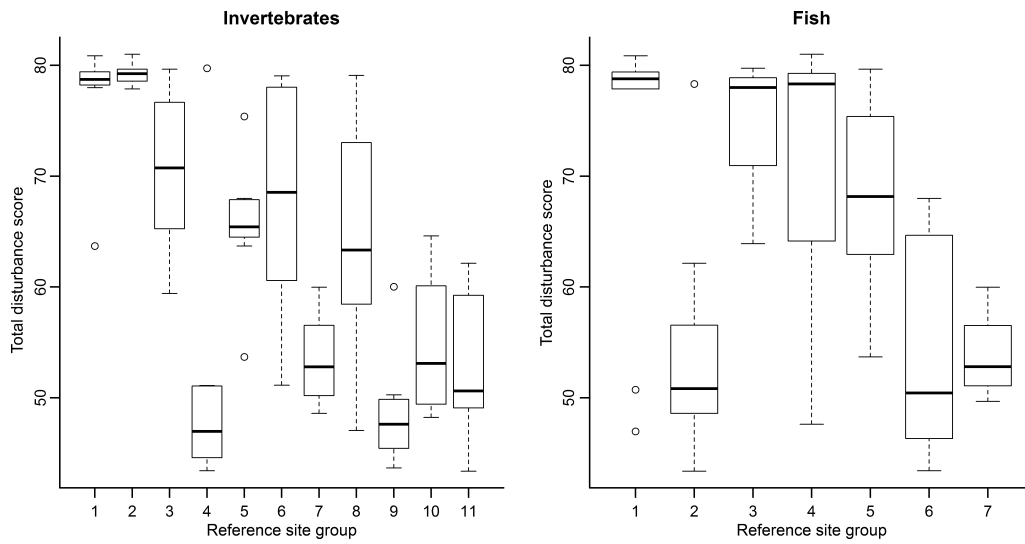

Fig. 4 Distribution of total disturbance scores among calibration reference site groups identified for invertebrate and fish data sets during predictive model development. Boxes represent first and third quartiles, black lines are medians, whiskers are 1.5 × the interquartile range, and circles are outliers.

Statistic	Invertebrates	Fish
Mean O/E (calibration sites)	1.032	1.044
Mean O/E (validation sites)	0.914	0.956
RMSE(O_N/E_N) for null model (validation sites)	0.303	0.548
RMSE(O/E) for predictive model (validation sites)	0.199	0.403
Predictive improvement over the null model	0.104	0.145
SD(O/E) for calibration sites	0.191	0.382
SD_R(O/E) for calibration site pseudoreplicates (theoretical lower bound of variability)	0.140	0.310

O/E, observed to expected taxa ratio; RMSE, root mean squared error.

models indicated improvement in RMSE(O/E) relative to their respective null models (Table 4). The fish assemblage model indicated a greater absolute decrease in RMSE(O/E) relative to the null model, although the fish assemblage model was also less precise in absolute terms [i.e. higher RMSE(O/E)] than the invertebrate model. Both models also indicated little ability for predictive improvement; in both cases, the SD(O/E) for the calibration sites was only slightly greater than the SD_R (Table 4). Finally, O/E scores derived from both the invertebrate and the fish assemblage models were significantly lower for test sites than for reference sites (one-sided unpaired *t*-tests; fish: $t = -1.78$, d.f. = 308.32, $P = 0.04$; invertebrates: $t = -6.14$, d.f. = 395.32, $P < 0.001$). For invertebrates, the mean score among reference sites was 0.98, whereas the mean score among test sites was 0.86. For fish, the mean score among reference sites was 1.00 and the mean score among test sites was 0.94.

Relationships between fish and invertebrate O/E scores were significant at the statewide scale, in four of seven ecoregions and in one of 21 catchments (Table 3). However, relationships were weak (R^2 values <0.3); the strongest relationship occurred in the Lower Minnesota catchment ($R^2 = 0.27$). Spatial units characterised by significantly correlated fish and invertebrate O/E scores also did not necessarily reflect units in which significant concordance between fish and invertebrate assemblages had been found. For example, although concordance between fish and invertebrate assemblages was significant and strong in the Rock basin, no significant relationship was found between fish and

invertebrate O/E scores. Indeed, we found a significant relationship between fish and invertebrate O/E scores in only one of the four catchments for which concordance values were relatively strong ($R \geq 0.7$). This phenomenon occurred at larger scales as well; communities in the Northern Lakes and Forests and Red River Valley ecoregions exhibited concordance at the regional scale but did not have significantly correlated O/E scores. Conversely, no significant relationships between fish and invertebrate O/E scores occurred in spatial units where fish and invertebrate communities were not significantly concordant.

Discussion

A number of recent studies have sought to evaluate the degree of concordance among stream assemblages (reviewed in Heino, 2010). Practitioners who take this approach have sometimes argued that communities exhibiting similar spatial patterns may respond to the same environmental drivers and therefore exhibit similar responses to disturbance (Kilgour & Barton, 1999; Paszkowski & Tonn, 2000; Mykrä *et al.*, 2008a). The appeal of using concordance measures to predict whether two communities will respond similarly to environmental change may be related to the relative straightforwardness of this approach, especially when compared with the relatively complex process of developing biological indices for each taxonomic group. Measures of assemblage concordance require only species abundance data, whereas development of biological indices often entails a series of analyses – namely, selection of index type, reference sites and predictor variables. If assemblage concordance was in fact indicative of a parallel response to disturbance by multiple taxonomic groups, managers could use concordance measures to justify a monitoring approach that uses one group as a surrogate indicator for the condition of others rather than dedicate the additional time and effort required to develop indices for all taxonomic groups of interest and compare their performance.

From a practical standpoint, we were interested primarily in whether concordant communities would yield congruent O/E scores. In contrast with our initial hypothesis, our analyses indicated that significant concordance between fish and invertebrates did not necessarily indicate whether these two communities would respond similarly to environmental change, i.e. concordant communities did not always exhibit similar

trends in O∕E scores. However, although we found significant concordance between fish and invertebrate assemblage composition at the statewide scale, and in most major catchments and ecoregions, concordance was generally weak. As Heino (2010) pointed out, significant concordance should not be confused with strong concordance, especially given that the randomisation tests typically used to evaluate concordance tend to yield statistical significance even if the strength of correlations is low. Heino (2010) further argues that only strong correlations ($m^2 < 0.5$; $R > 0.7$) can provide sufficient cause to predict the composition of one taxonomic group based on the other. Thus, it is perhaps not surprising that the weakly concordant communities we documented in this study failed to exhibit tightly linked rates of taxa loss in response to disturbance.

However, even for spatial units where concordance was relatively strong (i.e. R values approaching or exceeding 0.7), O∕E scores generally did not appear more strongly congruent than in other units. This observation questions whether strong concordance among communities might constitute sufficient justification for stream monitoring approaches that use one taxonomic group as a surrogate indicator for others. Concordance measures spatial patterns in the composition of two assemblages and is based on the entire composition of both groups. O∕E scores, in contrast, measure the losses of taxa relative to reference conditions and are based on the presence of a limited number of individual expected taxa. The loss of one or two taxa predicted to occur under reference conditions must, by definition, decrease an O∕E score. On the other hand, a loss of one or two species from one organismal group may not strongly affect the overall level of concordance between two assemblages as long as the key structural aspects that determine concordance are maintained. Thus, particularly at low or intermediate levels of disturbance, O∕E scores for two assemblages may diverge, whereas concordance measures may remain more or less constant. Further examination of how the loss of various taxa from one or both groups might affect the level of concordance between them could provide additional insight into the statistical properties of concordance measures relative to biological index scores.

Environmental drivers of stream communities

An examination of the environmental variables most strongly associated with the composition of fish and invertebrate communities provided insight into the nature of the concordance between them and into their respective responses to environmental stress. Namely, the consistent appearance of either catchment area or stream width in the pCCA models for fish and invertebrates suggests that the concordance we observed at all three spatial scales may have been driven in large part by parallel differences in both groups associated with stream size. The importance of stream size to community composition is consistent with a central tenet in the River Continuum Concept (Vannote *et al.*, 1980), which postulates that stream order, together with associated organic matter inputs, is the major factor driving the composition of stream communities. Indeed, a large body of literature has shown that aquatic communities change in a predictable fashion across longitudinal gradients in river systems (Ward, 1998; Grenouillet *et al.*, 2008 and references therein). In a recent example from Minnesota streams, Wan *et al.* (2010) demonstrated that the number of fish taxa in Minnesota streams increases uniformly with increasing stream catchment area and is highest in the largest river systems; invertebrate taxa richness appears to demonstrate a similar pattern over the same set of study sites (C.L. Dolph, unpublished data). Although we did not measure stream order *per se*, progressively larger catchment areas and stream widths generally indicate increasing stream order within a given study region (Allan & Castillo, 2008).

Besides the relatively consistent association with stream size, fish and invertebrate communities appeared to correlate largely with different sets of environmental variables. Thus, although concordance between fish and invertebrates may arise from parallel differences in the composition of each group associated with stream size, the level of concordance between them may be limited by their distinct sensitivities to other aspects of their environment. Catchment-scale land-use variables (e.g. % forest, % agriculture) were more often associated with fish than with invertebrate assemblage composition, whereas reach-scale variables (% fines, stream flow) were more often associated with invertebrate assemblage composition. However, both fish and invertebrate assemblages could be associated with either reach- or catchment-scale variables. Thus, as a number of other studies have indicated, the particular variables associated with the composition of stream communities probably depends on specific geographic context

(Johnson & Goedkoop, 2002; Magalhães, Batalha & Collares-Pereira, 2002; Townsend *et al.*, 2003; Burcher, Valett & Benfield, 2007; Johnson *et al.*, 2007; Infante *et al.*, 2009). It is also not surprising that % agriculture appeared strongly correlated with both fish and invertebrate communities at the statewide scale. The conversion of natural landscapes to agriculture results in changes to habitat structure, hydrology and the availability of nutrients and organic matter and has frequently been shown to diminish the integrity of stream assemblages (e.g. Roth, Allan & Erickson, 1996; Wang *et al.*, 1997; Allan, 2004; Lecerf *et al.*, 2006; Lyons, 2006; Piscart *et al.*, 2009).

Fish and invertebrate taxa-loss models

The differing sensitivities of fish and invertebrates to various environmental drivers may partly explain why relationships between fish and invertebrate O/E scores were weak. If the composition of each taxonomic group is structured by different environmental gradients, a disturbance that alters some aspect of the environment may result in taxa loss from one group but not the other. For example, stream flow was associated in a number of spatial units with invertebrate but not with fish assemblage composition. Presumably then, a change in flow conditions could more readily lead to the loss of invertebrate taxa than to the loss of fish taxa. The apparent discrepancy between fish and invertebrate O/E scores for Minnesota is consistent with findings by many previous studies, showing that different taxonomic groups fail to yield consistent indications of stream condition (Griffith *et al.*, 2005; Hering *et al.*, 2006; Carlisle *et al.*, 2008; Flinders, Horwitz & Belton, 2008; Mykrä *et al.*, 2008a; Beck & Hatch, 2009).

When two biotic indices convey divergent information about a stream's condition relative to reference sites, we must also consider whether these indices operate with uniform precision and accuracy. Indeed, unless the statistical performance of an index is explicitly evaluated, the sensitivity of a stream community to stressors cannot be disentangled from 'noise' (Fore, Karr & Conquest, 1994; Carlisle & Clements, 1999; Dolph *et al.*, 2010). In this study, both fish and invertebrate predictive models exhibited improved precision relative to their respective null models. Moreover, both models could detect departures from reference condition. The invertebrate assemblage model also exhibited a high degree of

precision (SD of O/E = 0.19) comparable with other RIVPACS models implemented nationwide (Hawkins, 2006; Hargett *et al.*, 2007; Carlisle *et al.*, 2008). Conversely, the fish model was relatively imprecise (SD of O/E = 0.38). The lack of precision associated with the fish model was likely due in part to the low number of common fish taxa in our data set – only five fish species were found in >50% of reference sites statewide. With few common fish taxa, a failure by the model to accurately predict the presence of only one or two species at a site could substantially alter the O/E score for a site.

The relative imprecision of our fish model can also be explained by a propensity for random sampling errors in the fish data set, as indicated by a high SD_R (0.31). The value of SD_R increases when the probability of capture Pi for a given taxon tends towards 0.5 (Van Sickle *et al.*, 2005). Because the true occurrence probabilities for taxa at reference sites are unknown, SD_R values are estimated using Pi values produced by the predictive models (Van Sickle *et al.*, 2005), i.e. Pi values for taxa at reference sites are calculated by determining the proportion of sites in a reference site group at which each taxon occurs. A high SD_R indicates a failure to consistently capture fish taxa across more than half of the sites within a reference site group. Given this limitation, an alternative DF model (i.e. a model consisting of additional or alternative predictor variables) probably would not substantially improve model performance. Instead, high heterogeneity in fish samples may limit the development of a precise predictive model for fish assemblages in Minnesota at a statewide scale. Differences in precision associated with the two predictive models may have contributed to the weak relationship between fish and invertebrate O/E scores, in addition to any real differences in the rates of taxa loss exhibited by these two groups that stem from their contrasting environmental sensitivities.

Predictive modelling, scale and reference condition

During predictive model development, we sought to account for the potential influence of geography on stream assemblages by selecting reference sites from every region of the state. However, reference sites in regions dominated by human land use had poor disturbance scores compared with sites in forested regions. When reference sites were clustered, within-

group disturbance levels could thus vary substantially for groups that were comprised of broadly distributed sites. Using variably disturbed reference sites to generate expectations for community composition could result in artificially low expected taxa estimates in less disturbed regions and artificially high estimates in more disturbed regions (because the inclusion of more disturbed reference sites will probably bias expectations towards less diverse, less abundant communities and *vice versa*). Moreover, because reference sites clustered differently for fish and invertebrate assemblages, expectations for each assemblage at a given site may have been based on two different estimates of reference condition. Such differences could have contributed to the general lack of congruence between fish and invertebrate O∕E scores. Yuan *et al.* (2008) argued that RIVPACS-type models should be built at spatial scales that match the scale at which assessment results will be reported to control for the spatial biases such as these. This type of analysis will soon be increasingly feasible in places where new and intensive sampling regimes implemented by management agencies should result in the availability of biological data for more stream sites at smaller (i.e. ecoregion and catchment) spatial scales. Finally, another consequence of using a statewide model to generate O∕E scores is that the assemblage–environment relationships used to predict community composition may rely primarily on large-scale environmental gradients and thus fail to accurately predict community composition at smaller scales (Marchant *et al.*, 1999; Mykrä, Heino & Muotka, 2007; Yuan *et al.*, 2008). However, although we did not have a sufficient number of study sites available to construct predictive models for each region and catchment to test this possibility, Mykrä *et al.* (2008b) found that taxa-loss models developed at larger and smaller scales performed with similar precision at small scales as long as rare taxa were excluded (as they were in our study).

Concordance and scale

Paavola *et al.* (2006) argued that the importance of regional-scale variables in shaping community composition may result in stronger concordance at broader scales. Our analyses appear to contradict this hypothesis. Although concordance between fish and invertebrate communities was significant at the state-wide scale and in most ecoregions, we also found significant assemblage concordance in most catchments (the smallest spatial scale we considered). Moreover, the degree of concordance at the catchment scale was as strong or stronger compared with other units and scales included in our study. Thus, concordance may occur at both small and large scales, depending on the geographic context.

Strongly concordant communities at the catchment scale could be explained by a strong environmental gradient operating within certain catchments (Ormerod *et al.*, 1994; Paavola *et al.*, 2006). The widespread occurrence of catchment area or stream width in the pCCA models of fish and invertebrate assemblage composition indicates that stream size may constitute such a gradient. However, concordant fish and invertebrate communities at the catchment scale were not always associated with stream size or with any of the other environmental gradients we considered. In the Redwood catchment, for example, concordance was relatively strong but fish and invertebrate communities were associated with entirely different sets of environmental variables. Moreover, catchments characterised by concordant communities occurred both in the regions of intensively managed row crop (i.e. corn and soybean) production and in the regions of the state that drain into Lake Superior and are dominated (>80%) by forested areas and wetlands. If concordance at the catchment scale is driven by some environmental gradient that we failed to measure, the variables may not be the same in all catchments, given the very different geology, land use and climate patterns associated with the two types of catchments exhibiting concordance in this study. Alternatively, the concordance we observed at the catchment scale may also have been driven by biotic interactions or spatial autocorrelation rather than environmental sensitivities (Bahn & McGill, 2007; Currie, 2007; Grenouillet *et al.*, 2008; Johnson & Hering, 2010),

Implications for biological monitoring and conservation

The results of this study have several implications for the biological assessment of streams. First, given that concordant assemblages do not appear to yield tightly linked estimates of stream integrity, we suggest that a finding of community concordance should not be substituted for an evaluation of how different assemblages respond to disturbance. Instead, an

analysis of whether various assemblages respond to environmental degradation in similar ways (i.e. via similar rates of taxa loss or some other measure) would constitute the most direct and therefore strongest evidence in favour of a surrogate taxa approach to biological assessment.

Secondly, the distinct relationships fish and invertebrate assemblages appear to have with their environment suggest that a comprehensive assessment of stream integrity will be best achieved by monitoring for multiple taxonomic groups. Such an approach, however, poses the challenge of integrating potentially disparate responses of different assemblages (Carlisle *et al.*, 2008; Walters, Roy & Leigh, 2009). Some authors have argued that the interpretation of biological data could be simplified by employing a single multi-assemblage index to represent the integrity of the entire stream community (e.g. Griffith *et al.*, 2005), whereas others contend that combining multiple assemblages into a single score may obscure assemblage-specific alterations, as well as obfuscate the various unique pathways by which each assemblage is affected by human disturbance (Carlisle *et al.*, 2008; Walters *et al.*, 2009). In practice, however, the performance of a single index based on a multi-assemblage data set has rarely been tested. Additionally, research is needed to understand how patterns in fish and invertebrates assemblages correspond to other less studied groups such as macrophytes, algae, zooplankton, birds, mammals or molluscs.

Our findings also relate to the important concept of stressor identification (Cormier, Norton & Suter, 2003). Addressing the sources of environmental stress affecting stream assemblages requires a detailed understanding of how different groups respond to their environment and at which scale. Our observations indicate that fish and invertebrate communities may respond to distinct local- and catchment-scale variables depending on the geographic context. Thus, various conservation and restoration strategies – i.e. improving riparian or in-stream habitat or water quality, or changes to catchment land use – may affect fish or invertebrate populations more or less strongly depending on the study region. At the same time, however, environmental variables that correlated with both assemblages in a given spatial unit represent good candidate targets for conservation efforts in that particular geographic area. For example, the appearance of % fines in both fish and invertebrate pCCA models for the Rock catchment region suggests that efforts designed to mitigate excessive deposition of sediments might be an appropriate priority for conservation programmes aimed at broadly improving the integrity of stream communities in this catchment.

References

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

Allan J.D. & Castillo M.M. (2008) *Stream Ecology: Structure and Function of Running Waters*. Springer, Dordrecht, The Netherlands, p. 35.

Bahn V. & McGill B.J. (2007) Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography*, **16**, 733–742.

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*, 2nd edn. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water, Washington, DC.

Beck M.W. & Hatch L.K. (2009) A review of research on the development of lake indices of biotic integrity. *Environmental Reviews*, **17**, 21–44.

Bilton D.T., Mcabendroth L., Bedford A. & Ramsay P.M. (2006) How wide to cast the net? Cross-taxon congru-

ence of species richness, community similarity and indicator taxa in ponds *Freshwater Biology*, **51**, 578–590.

Birks H.J.B., Peglar S.M. & Austin H.A. (1994) *An Annotated Bibliography of Canonical Correspondence Analysis and Related Constrained Ordination Methods 1986–1993*. Botanical Institute, Bergen, Norway, 58pp.

Boulton A.J. (1999) An overview of river health assessment: philosophies, practice, problems and prognosis. *Freshwater Biology*, **41**, 469–479.

ter Braak C.J.F. & Verdonschot P.F.M. (1995) Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquatic Sciences*, **57**, 255–289.

Burcher C.L., Valett H.M. & Benfield E.F. (2007) The land-cover cascade: relationships coupling land and water. *Ecology*, **88**, 228–242.

Campbell Grant E.H., Lowe W.H. & Fagan W.F. (2007) Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters*, **10**, 165–175.

Carlisle D.M. & Clements W.H. (1999) Sensitivity and variability of metrics used in biological assessments of running waters. *Environmental Toxicology and Chemistry*, **18**, 285–291.

Carlisle D.M., Hawkins C.P., Meador M.R., Potapova M. & Falcone J. (2008) Biological assessments of Appalachian streams based on predictive models for fish, macroinvertebrate, and diatom assemblages. *Journal of the North American Benthological Society*, **27**, 16–37.

Clarke K.W. (1993) Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology*, **18**, 117–143.

Cormier S.M., Norton S.B. & Suter G.W. (2003) The U.S. Environmental Protection Agency's stressor identification Guidance: A process for determining the probable causes of biological impairments. *Human and Ecological Risk Assessment*, **9**, 1431–1443.

Currie D.J. (2007) Disentangling the roles of environment and space in ecology. *Journal of Biogeography*, **34**, 2009–2011.

Daly C., Taylor G.H., Gibson W.P., Parzybok T.W., Johnson G.L. & Pasteris P.A. (2000) High-quality spatial climate data sets for the United States and beyond. *Transactions of the American Society of Agricultural Engineers*, **43**, 1957–1962.

Dolph C.L., Sheshukov A.Y., Chizinski C.J., Vondracek B. & Wilson B. (2010) The Index of Biological Integrity and the bootstrap: can random sampling error affect stream impairment decisions? *Ecological Indicators*, **10**, 527–537.

Flather C.H., Wilson K.R., Dean D.J. & McComb W.C. (1997) Identifying gaps in conservation networks: of indicators and uncertainty in geographic-based analyses. *Ecological Applications*, **7**, 531–542.

Flinders C.A., Horwitz R.J. & Belton T. (2008) Relationship of fish and macroinvertebrate communities in the mid-Atlantic uplands: implications for integrated assessments. *Ecological Indicators*, **8**, 588–598.

Fore L.S., Karr J.R. & Conquest L.L. (1994) Statistical properties of an index of biological integrity used to evaluate water resources. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 1077–1087.

Grenouillet G., Brosse S., Tudesque L., Lek S., Baraille Y. & Loot G. (2008) Concordance among stream assemblages and spatial autocorrelation along a fragmented gradient. *Diversity and Distributions*, **14**, 592–603.

Griffith M.B., Hill B.H., McCormick F.H., Kaufmann P.R., Herlihy A.T. & Selle A.R. (2005) Comparative application of indices of biotic integrity based on periphyton, macroinvertebrates, and fish to southern Rocky Mountain streams. *Ecological Indicators*, **5**, 117–136.

Hargett E.G., ZumBerge J.R., Hawkins C.P. & Olson J.R. (2007) Development of a RIVPACS-type predictive model for bioassessment of wadeable streams in Wyoming. *Ecological Indicators*, **7**, 807–826.

Hawkins C.P. (2006) Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessment. *Ecological Applications*, **16**, 1277–1294.

Heino J. (2002) Concordance of species richness patterns among multiple freshwater taxa: a regional perspective. *Biodiversity and Conservation*, **11**, 137–147.

Heino J. (2010) Are indicator groups and cross-taxon congruence useful for predicting biodiversity in aquatic ecosystems? *Ecological Indicators*, **10**, 112–117.

Heino J., Paavola R., Virtanen R. & Muotka T. (2005) Searching for biodiversity indicators in running waters: do bryophytes, macroinvertebrates, and fish show congruent diversity patterns? *Biodiversity and Conservation*, **14**, 415–428.

Hering D., Johnson R.K., Kramm S., Schmutz S., Szoszkiewicz K. & Verndonschot P.F. (2006) Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. *Freshwater Biology*, **51**, 1757–1785.

Homer C., Dewitz J., Fry J., Coan M., Hossain N., Larson C., Herold N., McKerrow A., VanDriel J.N. & Wickham J. (2007) Completion of the 2001 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering & Remote Sensing*, **73**, 337–341.

Infante D.M., Allan J.D., Linke S. & Norris R.H. (2009) Relationship of fish and macroinvertebrate assemblages to environmental factors: implications for community concordance. *Hydrobiologia*, **623**, 87–103.

Jackson D.A. & Harvey H.H. (1993) Fish and benthic invertebrates: community concordance and commu-

nity-environment relationships. *Canadian Journal of Fisheries and Aquatic Science*, **50**, 2641–2651.

Johnson R.K. & Goedkoop W. (2002) Littoral macroinvertebrate communities: spatial scale and ecological relationships. *Freshwater Biology*, **47**, 1840–1854.

Johnson R.K. & Hering D. (2010) Spatial congruency of benthic diatom, invertebrate, macrophyte, and fish assemblages in European streams. *Ecological Applications*, **20**, 978–992.

Johnson R.K., Furse M.T., Hering D.T. & Sandin L. (2007) Ecological relationships between stream communities and spatial scale: implications for designing catchment-level monitoring programs. *Freshwater Biology*, **52**, 939–958.

Karr J.R. (1999) Defining and measuring river health. *Freshwater Biology*, **41**, 221–234.

Karr J.R. & Chu E.W. (2000) Sustaining living rivers. *Hydrobiologia*, **422/423**, 1–14.

Karr J.R. & Dudley D.R. (1981) Ecological perspective on water quality goals. *Environmental Management*, **5**, 55–68.

Karr J.R. & Yoder C.O. (2004) Biological assessment and criteria improve total maximum daily load decision making. *Journal of Environmental Engineering*, **130**, 594–604.

Kilgour B.W. & Barton D.R. (1999) Associations between stream fish and benthos across environmental gradients in southern Ontario, Canada. *Freshwater Biology*, **41**, 553–566.

Lawler J.J., White D., Sifneos J.C. & Master L.L. (2003) Rare species and the use of indicator groups for conservation planning. *Conservation Biology*, **17**, 875–882.

Lecerf A., Usseglio-Polatera P., Charcosset J., Lambrigot D., Bracht B. & Chauvet E. (2006) Assessment of functional integrity of eutrophic streams using litter breakdown and benthic macroinvertebrates. *Archiv für Hydrobiologie*, **165**, 105–126.

Legendre P. & Legendre L. (1998) *Numerical Ecology*, 2nd English edn. Elsevier Science B.V., Amsterdam, The Netherlands.

Lewandowski A.S., Noss R.F. & Parsons D.R. (2010) The effectiveness of surrogate taxa for the representation of biodiversity. *Conservation Biology*, **5**, 1367–1377.

Lyons J. (2006) A fish-based index of biotic integrity to assess intermittent headwater streams in Wisconsin, USA. *Environmental Monitoring and Assessment*, **122**, 239–258.

Magalhães M.F., Batalha D.C. & Collares-Pereira M.J. (2002) Gradients in stream fish assemblages across a Mediterranean landscape: contributions of environmental factors and spatial structure. *Freshwater Biology*, **47**, 1015–1031.

Marchant R., Hirst A., Norris R. & Metzeling L. (1999) Classification of macroinvertebrate communities across drainage basins in Victoria, Australia: conse-

quences of sampling on a broad spatial scale for predictive modelling. *Freshwater Biology*, **41**, 253–268.

McCune B. & Grace J.B. (2002) *Analysis of Ecological Communities*. MjM Software Design, Gleneden Beach, OR.

Moritz C., Richardson K., Ferrier S., Monteith G.B., Stanisic J., Williams S.E. *et al.* (2001) Biogeographical concordance and efficiency of taxon indicators for establishing conservation priority in a tropical rainforest biota. *Proceedings of the Royal Society of London Biological Sciences*, **268**, 1875–1881.

Mykrä H., Heino J. & Muotka T. (2007) Scale-related patterns in the spatial and environmental components of stream macroinvertebrate assemblage variation. *Global Ecology & Biogeography*, **16**, 149–159.

Mykrä H., Aroviita J., Hämäläinen H., Kotanen J., Vuori K.M. & Muotka T. (2008a) Assessing stream condition using macroinvertebrates and macrophytes: concordance of community responses to human impact. *Archiv für Hydrobiologia*, **172**, 191–203.

Mykrä H., Aroviita J., Kotanen J., Hämäläinen H. & Muotka T. (2008b) Predicting the stream macroinvertebrate fauna across regional scales: influence of geographical extent on model performance. *Journal of the North American Benthological Society*, **27**, 705–716.

Norris R.H. & Hawkins C.P. (2000) Monitoring river health. *Hydrobiologia*, **435**, 5–17.

Norris R.H. & Thoms M.C. (1999) What is river health? *Freshwater Biology* **41**, 197–209.

Oksanen J., Kindt R., Legendre P., O'Hara B., Simpson G.L., Solymos P. *et al.* (2009) Vegan: Community Ecology Package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists. R package version 1.15-3.

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

Ormerod S.J., Rundle S.D., Wilkinson S.M., Daly G.P., Dale K.M. & Juttner I. (1994) Altitudinal trends in the diatoms, bryophytes, macroinvertebrates and fish of a Nepalese river system. *Freshwater Biology*, **32**, 309–322.

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

Paavola R., Muotka T., Virtanen R., Heino J. & Kreivi P. (2003) Are biological classifications of headwater streams concordant across multiple taxonomic groups? *Freshwater Biology*, **48**, 1912–1923.

Paavola R., Muotka T., Virtanen R., Heino J., Jackson D. & Maki-Petays M. (2006) Spatial scale affects community concordance among fishes, benthic macroinvertebrates, and bryophytes in streams. *Ecological Applications*, **16**, 368–379.

Paszkowski C.A. & Tonn W.M. (2000) Community concordance between the fish and aquatic birds of lakes in northern Alberta, Canada: the relative importance of environmental and biotic factors. *Freshwater Biology*, **43**, 421–437.

Pawar S.S., Birand A.C., Ahmed M.F., Sengupta S. & Raman T.R.S. (2007) Conservation biogeography in North-east India: hierarchical analysis of cross-taxon distributional congruence. *Diversity and Distributions*, **13**, 53–65.

Peres-Neto P.R. & Jackson D.A. (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test *Oecologia*, **129**, 169–178.

Piscart C., Genoel R., Dolédec S., Chauvet E. & Marmonier P. (2009) Effects of intense agricultural practices on heterotrophic processes in streams. *Environmental Pollution*, **157**, 1011–1018.

Pyne M.I., Rader R.B. & Christensen W.F. (2007) Predicting local biological characteristics in streams: a comparison of landscape classifications. *Freshwater Biology*, **52**, 1302–1321.

Roth N.E., Allan J.D. & Erickson D.L. (1996) Landscape influences on stream biotic integrity assessed at multiple spatial scales. *Landscape Ecology*, **11**, 141–156.

Shurin J.B., Cottenie K.. & Hillebrand H. (2009) Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia*, **159**, 151–159.

Stoddard J.L., Larsen D.P., Hawkins C.P., Johnson R.K. & Norris R.H. (2006) Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecological Applications*, **16**, 1267–1276.

Su J.C., Debinski D.M., Jakubauskas M.E. & Kindscher K. (2004) Beyond species richness: community similarity as a measure of cross-taxon congruence for coarse-filter conservation. *Conservation Biology*, **18**, 167–173.

Townsend C.R., Dolédec S., Norris R., Peacock K. & Arbuckle C. (2003) The influence of scale and geography on relationships between stream community composition and landscape variables: description and prediction. *Freshwater Biology*, **48**, 768–785.

[USGS] United States Geological Survey (1982) *Codes for the Identification of Hydrological Units in the United States and the Caribbean Outlying Areas*. U.S. Geological Survey, Water Resources Division, Office of Water Data Coordination, Reston, VA.

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

Van Sickle J., Huff D.D. & Hawkins C.P. (2006) Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. *Freshwater Biology*, **51**, 359–372.

Van Sickle J., Larsen D.P. & Hawkins C.P. (2007) Exclusion of rare taxa affects performance of the O/E index in bioassessments. *Journal of the North American Benthological Society*, **26**, 319–331.

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

Walters D.M., Roy A.H. & Leigh D.S. (2009) Environmental indicators of macroinvertebrate and fish assemblage integrity in urbanizing watersheds. *Ecological Indicators*, **9**, 1222–1233.

Wan H., Chizinski C.J., Dolph C.L., Vondracek B. & Wilson B. (2010) The impact of rare taxa on a fish index of biotic integrity. *Ecological Indicators*, **10**, 781–788.

Wang L., Lyons J., Kanehl P. & Gatti R. (1997) Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. *Fisheries*, **22**, 6–12.

Ward J.V. (1998) Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. *Biological Conservation*, **83**, 269–278.

Wright J.F. (2000) An introduction to RIVPCAS. In: *Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques* (Eds J.F. Wright, D.W. Sutcliffe & M.T. Furse), pp. 1–24. Freshwater Biological Association, Ambleside, Cumbria, UK.

Yates A.G. & Bailey R.C. (2010) Covarying patterns of macroinvertebrate and fish assemblages along natural and human activity gradients: implications for bioassessment. *Hydrobiologia*, **637**, 87–100.

Yuan L.L., Hawkins C.P. & Van Sickle J. (2008) Effects of regionalization decisions on an O/E index for the US national assessment. *Journal of the North American Benthological Society*, **27**, 892–905.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Results of partial Canonical Correspondence Analysis (pCCA) using forward selection of variables for statewide[1], ecoregion[2] and catchment[3] datasets.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organised for online delivery, but are not copyedited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

(Manuscript accepted 1 February 2011)