Ann. Limnol. - Int. J. Lim. 47 (2011) 199–209
© EDP Sciences, 2011
DOI: 10.1051/limn/2011026

Available online at:
www.limnology-journal.org

The importance of land use/land cover data in fish and mussel conservation planning

Robert L. Hopkins II[1*] and Matt R. Whiles[2]

[1] School of Sciences, University of Rio Grande, Rio Grande, OH 45674, USA
[2] Department of Zoology, Southern Illinois University, Carbondale, IL 62901, USA

Received 21 December 2010; Accepted 16 April 2011

Abstract – Freshwater fish and mussel diversity is declining at an alarming rate across North America. Human land uses and disturbances within watersheds have been implicated as the primary cause of declines. In this paper, we demonstrate the utility of land use/land cover (LULC) variables in species distribution modeling and conservation planning using a straightforward multiscale approach for prioritizing freshwater fish and mussel conservation areas in the upper Green River catchment (Ohio River basin, USA). We developed distribution and species richness models for 10 uncommon fishes and 14 rare mussels using multiscale landscape data and boosted regression tree (BRT) analyses based on LULC composition and pattern, geology composition, and soil composition data. We then used probability of occurrence, endemicity, prevalence, trend and range of individual species to estimate the conservation value of each stream reach. Conservation areas were defined for three spatial scales nested within the catchment management zone (focal areas, riparian management buffer and subcatchment management zone) using a simple optimization technique. Priority conservation areas were located primarily in the eastern (upper Green River) and southern (upper Barren River) portions of the catchment. We found that focal species richness is explained most by soil composition in the subcatchment. However, nested within the subcatchment scale focal species richness responded positively to percent forest and negatively to patch density of developed/exposed land in the reach buffer. For both the reach and riparian buffers, retaining forested tracts of land and limiting the level of development and fragmentation would benefit the focal species.

Key words: Landscape / aquatic biodiversity / reserve selection / species distribution models

Introduction

There is increasing concern over precipitous population declines and rates of extinction of native freshwater fauna in North America (Rahel, 2000; Abell, 2002; Strayer, 2006). According to Ricciardi and Rasmussen (1999), approximately 50% of the mussel species and 25% of the fish species in North America are imperiled, with extinction rates approaching 6.4 and 2.4% per decade, respectively. The pattern of imperilment in the southern United States, a region rich in freshwater biological diversity, exceeds the continental patterns; nearly 30% of the fish fauna (Warren *et al.*, 1997) and 60% of the mussel fauna within this region are considered imperiled (Neves *et al.*, 1997). Concurrent with widespread declines in freshwater biodiversity is increasing recognition of the role(s) of native biodiversity and species composition in ecocatchment integrity. There is mounting evidence that ecosystem stability, functioning and resistance to invasion by non-indigenous species are contingent, at least to some degree, upon the composition and diversity of native species (Tilman, 1999; Naeem and Wright, 2003; Srivastava and Vellend, 2005).

Habitat degradation stemming from human land uses within the watershed has been cited as the principal cause of decline in freshwater biodiversity and shifts in faunal structure (Saunders *et al.*, 2002). Human activities within the watershed often result in hydrologic modifications, altered nutrient deliveries and increased sediment loading to receiving waters (Allan, 2004). In summary, species declines, the ongoing transformation of the landscape by humans, and the ecological and economic values of biodiversity warrant increased research in freshwater biodiversity conservation.

Species distribution models (SDMs), in conjunction with reserve-selection algorithms, have revolutionized

*Corresponding author: rhopkins@rio.edu

the conservation planning process (Guisan and Thuiller, 2005; Moilanen *et al.*, 2008). SDMs afford biologists the ability to identify key elements of the environmental niche of a species based on current distribution patterns and generate distribution probability maps. These spatially explicit SDMs are of paramount importance to resource managers designing conservation plans. However, resource managers are also concerned with developing practical management strategies, which hinge on modeled species–environment relationships. A recurring problem is that data used to model species occurrences often have limited implications for management or are incompatible with spatial units designated for conservation. For example, Moilanen *et al.* (2008) developed predictive models for several New Zealand fishes using environmental data captured at several spatial scales. The models performed very well and were used for predictive mapping and selection of conservation areas. A disconnect arises when one considers how to actively manage or create policy for those conservation areas: many of the environmental variables used (*e.g.*, reach slope, average temperature) offer limited information for developing management plans. We suggest that, in addition to natural environmental variables, metrics quantifying land use/land cover (LULC) – which relate to human influences – should also be integrated into SDMs. Such models may provide resource managers with sufficient information to develop practical landscape management strategies.

Here, we present a multiscale landscape approach for the selection and management of areas for fish and mussel conservation using a case study of the upper Green River catchment (Ohio River basin), USA. The upper Green River catchment is the fourth most biologically diverse aquatic ecosystem in North America (Master *et al.*, 1998) and is particularly well known for the high number of endemic species. Six fishes, one mussel and one crayfish occur nowhere else in the world. Our chief objectives were to (1) develop SDMs for 10 fishes and 14 mussels of conservation concern, (2) use a simple ranking index and distribution probabilities of each species to map and identify conservation areas at three spatial scales and (3) demonstrate the utility of SDMs in making scale-specific recommendations for landscape management. The overarching purpose of our study was to illustrate *via* a case study the potential value of incorporating LULC variables as foundational elements in SDMs used for conservation planning.

Methods

Study area and species data

The upper Green River catchment (Ohio River basin) is located in the central-eastern United States (Fig. 1) and drains an area of approximately 13 000 km^2 with elevation ranging from 110 to 550 m above sea level. Three main subsystems comprise the upper Green River catchment – Barren River, Nolin River and Green River – and each is impounded by a reservoir. Physiography is varied, with dissected uplands in the southern, eastern and northern areas of the watershed. In contrast, the central and western portions of the drainage have extensive karsting and much lower-relief topography. Upland areas are predominantly forested with scattered pastureland, while low-gradient regions are heavily used for agriculture.

We extracted occurrence data for 24 focal species (10 fishes and 14 mussels) of conservation concern from a dataset consisting of museum-based information (Hopkins, 2009; Hopkins *et al.*, 2009). Data were available for 304 sites sampled between 1995 and 2008: many sites received repeated visits during the 13-year sample interval. The proportion of sample site occupied by each species ranged from 0.02 to 0.46, with a mean of 0.08. Species richness was estimated as the total number of focal species collected at each site.

To build each SDM, we coded sites as positive if the respective species was ever detected: sites at which a species was not collected were treated as probable absences (Elith *et al.*, 2006). The majority of survey efforts was completed by the Kentucky State Nature Preserves Commission, Kentucky Department of Fish and Wildlife Resources (KDFWR), Southern Illinois University, and the Kentucky Division of Water: each effort was intended as a quasi-comprehensive aquatic survey. However, these efforts were not standardized and a variety of sampling techniques were used. Therefore, to reduce the confounding effects of variable sampling efficiencies – and because absence records are never completely reliable – we employed a binomial deviance loss function in model fitting. This loss function is robust to noisy occurrence data containing false negative observations (Elith *et al.*, 2008). Furthermore, Elith *et al.* showed excellent performance of boosted regression trees (BRTs) (our method of analysis) even when species "absence" data consisted of randomly generated pseudo-absences.

Landscape variables

We analyzed three spatial scales for each sample site: (1) the subcatchment, (2) a 100 m subcatchment riparian buffer and (3) a 100 m riparian buffer for a 1000 m upstream reach. Each scale was delineated based on a 30 m digital elevation model and hydrographic map. The resulting polygons were used as masks for data extraction. For each SDM, we calculated 75 landscape metrics for each site – 25 for each spatial scale – comprising four broad types of landscape features: LULC pattern, LULC composition, soil composition and geology composition using FRAGSTATS (McGarigal and Marks, 1995) (Table 1). The potential problem with using a total of 75 predictor variables in each species model (*e.g.*, overfitting) is addressed by our statistical method of choice. Landscape data were derived from maps provided by the Kentucky Division of Geographic Information. All spatial

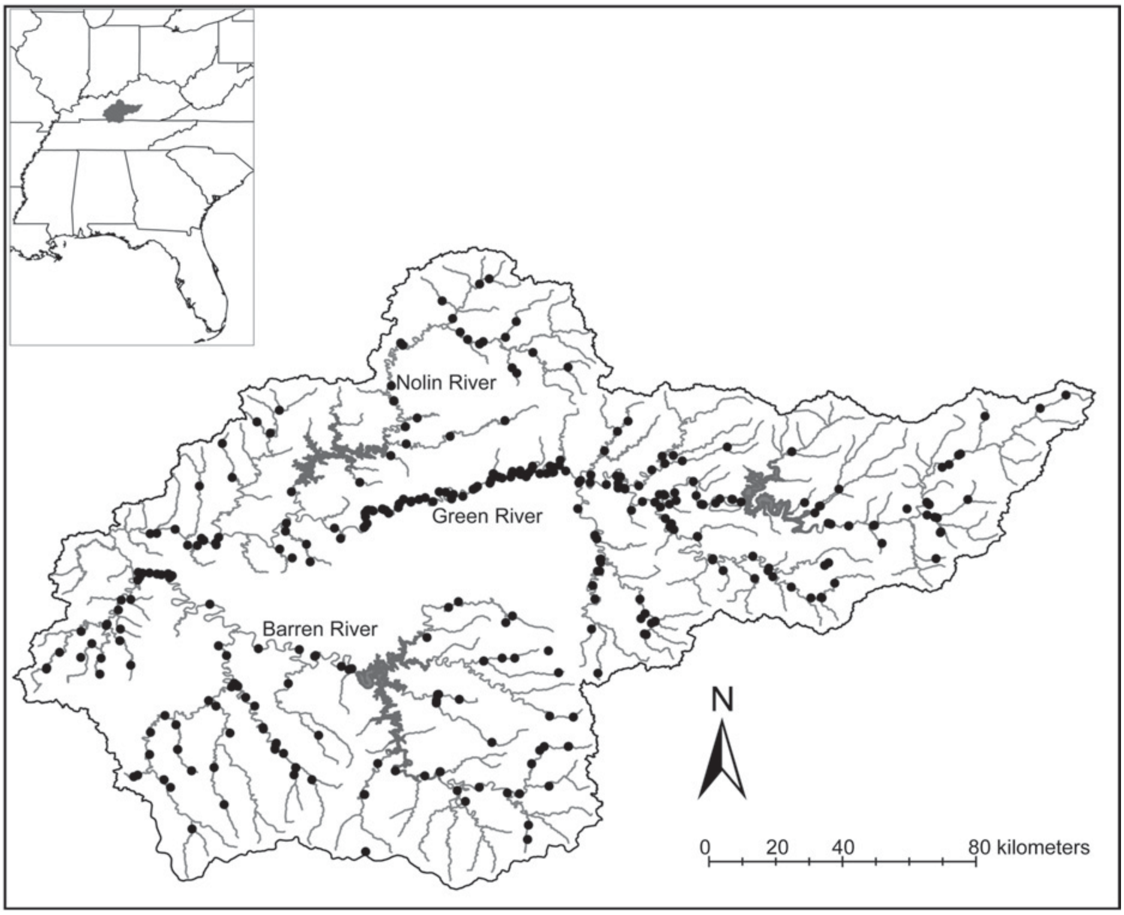

Fig. 1. Map of upper Green River catchment showing locality of sample sites. All streams of approximately second order and higher are shown. Inset map shows position of watershed in the eastern United States.

processing and mapping was accomplished using ArcGIS 9.2 (Environmental System Research Institute, 2006).

Statistical modeling and predictive mapping

Relationships between species responses and environmental predictors were analyzed using BRTs, a machine-learning form of logistical regression using decision trees and a boosting algorithm (Elith *et al*., 2008). In contrast to traditional classification and regression tree analysis which models data using a single tree, BRT employ boosting which combines numerous trees into a single model. Each successive tree focuses on unexplained portions of the response variance; thus predictive deviance is reduced in a stagewise fashion. There are two primary parameters for BRT models that are controlled by the modeler, tree complexity (TC) and learning rate (LR). TC represents the number of nodes in each individual tree, and controls the interaction depth of the input variables. The LR controls the effects of successive tree addition on model performance: low LRs reduce the effects of successive trees and high LRs increase the relative influence of successive trees. Consequently, low LRs result in more complex models (*e.g*., larger number of trees), but may be suitable when data are particularly noisy. The BRT approach has been shown to exhibit relatively high predictive accuracy compared to other SDM techniques (Elith *et al*., 2006). Furthermore, BRT only incorporate important variables into the model, provide robust information regarding variable importance, can model non-linear relationships, and are relatively easy to interpret (Elith *et al*., 2008).

We fitted each model in R (R Development Core Team, 2006) using the gbm package (Ridgeway, 2006), code provided by Elith *et al*. (2008) as well as some novel code. Parameters for each species' BRT model were optimized by systematically altering the TC and LR of the model until (1) the initial estimate of predictive deviance of the model was minimized, (2) predictive accuracy was maximized measured using the area under the curve (AUC) score, (3) the number of trees in the model was not excessively high or low (*i.e*., between approximately 1000 and 3500) (Elith *et al*., 2008) and (4) the model did not overfit the data, indicated by a steep rise in predictive deviance after the minimum is reached (Elith *et al*., 2008). A bag fraction of 0.5 was used in each model to ensure a sufficient amount of stochasticity and avoid overfitting the model to the data (Elith *et al*., 2008). The bag fraction

Table 1. Description of landscape environmental variables used in species distribution models (SDMs).

Variable type	Description
Land use/land cover (LULC) composition	
Pct_agr	Percent proportion of agricultural land use
Pct_pas	Percent proportion of pasture land use
Pct_for	Percent proportion of forest cover
Pct_dex	Percent proportion of developed/exposed land
LULC pattern	
Pd_agr	Number of agricultural patches per 100 ha
Pd_pas	Number of pasture patches per 100 ha
Pd_for	Number of forest patches per 100 ha
Pd_dex	Number of developed/exposed patches per 100 ha
Ps_agr	Mean agricultural patch size in hectares
Ps_pas	Mean pasture patch size in hectares
Ps_for	Mean forest patch size in hectares
Ps_dex	Mean developed/exposed patch size in hectares
Pd	Total patch density per 100 ha
Contagion	Probability of adjacent cells being the same LULC type
Soil composition	
Pct_fl	Percent proportion of fine-loamy soils
Pct_f	Percent proportion of fine soils
Pct_fs	Percent proportion of fine-silty soils
Pct_c	Percent proportion of clayey soils
Pct_cs	Percent proportion of coarse-silty soils
Pct_ls	Percent proportion of loamy-skeletal soils
Bedrock geology composition	
Pct_ac	Percent proportion of alluvium/coal measure
Pct_fly	Percent proportion of flysch
Pct_mol	Percent proportion of molasse
Pct_shl	Percent proportion of shale carbonate
Pct_she	Percent proportion of shelf carbonate

is the proportion of training data randomly selected at each stage to build the next tree in the model. Selecting a value less than 1.0 introduces randomness into the model and results in more robust performance.

The predictive and explanatory performances of the BRT models were evaluated using a ten-fold cross validation (CV) procedure. Each focal species dataset was randomly divided into 70% of the presence/absence data to develop BRT models, while the remaining 30% of the data were used to test the model. For each CV trial, two measures were calculated to assess the model's predictive and explanatory performance: the percent response variance explained and area under the curve (AUC) calculated for each species model. The AUC measures the ability of a model to accurately discriminate between presence and absence sites. The CV procedure was repeated ten times with resubstitution for each study species and the mean and standard error of these two performance measures were calculated.

Spatial mapping was performed by dividing all streams into 3000 m reaches and quantifying environmental predictors for the reach mid-point. In total, 1108 reaches were delineated. Probability of occurrence for each species was predicting using the optimal BRT model. We applied an occurrence probability threshold of 0.5 to develop a presence/absence map for each species (scores ≥ 0.5 equal presence), a typical approach for delimiting species distributions (Hirzel and Guisan, 2002; Brotons *et al.*, 2004). These distribution maps were used to estimate parameters for the priority index equation discussed below.

Relationships of focal species richness and environmental variables were analyzed following a similar procedure. Using code provided by Elith *et al.* (2008), the relative influence of each environmental variable at each spatial scale was quantified and relationships were examined using partial dependence plots (*e.g.*, fitted functions), which show the general effect of each variable on the response by integrating out the effects of all other variables. Pairwise interactions of partial dependence plots of the most influential LULC variables were used to visualize generalized effects of LULC on species richness. Note that these models are not intended to predict focal species richness, but instead are used to identify general global responses to select variables and provide insight for management strategies.

Selecting areas for conservation

To prioritize areas for conservation we employed a simple method which ranks stream reaches according to (1) probability of occurrence, (2) connectivity, (3) degree of endemism, (4) stability and (5) prevalence for each focal species. Each measure is described in detail below. Our methodology was based on a technique developed by Filipe *et al.* (2004). However, in the current study we added a measure of connectivity, added a measure of species trend and quantified endemism more explicitly. First, we calculated a priority index for each species k for each reach j (PI_{kj}):

$$\mathrm{PI}_{kj} = \frac{E_k \times T_k}{O_k}\left(\frac{P_{kj(\mathrm{down})} + P_{kj} + P_{kj(\mathrm{up})}}{3}\right)$$

where E_k is the ratio of the study area (km^2) to the total range (km^2) of species k, T_k is the score of distributional trend for species k, O_k is the proportion of stream reaches occupied by species k, and P_{kj} is the probability of occurrence of species k at reach j, one reach downstream of reach j ($P_{kj(\mathrm{down})}$) and one reach upstream of reach j ($P_{kj(\mathrm{up})}$). The total range of species k was estimated using distribution maps provided by the KDFWR (2005). Point-locality maps (*e.g.*, dot maps) were available for the state of Kentucky and generalized polygon maps were available for the southeastern US. We set the minimum allowable value of $E_k = 1$; thus, species with $E_k > 1$ are endemic to the study area. Two scores were used for species trend (T_k): these were 1, stable; 2, declining. These numbers were selected to produce higher scores for declining species and were based on status categories provided by the KDFWR (2005). The parameter O_k and all P_{kj} were derived from predictive maps developed using the BRT models. A relative measure of population connectivity was calculated by averaging probability scores for three contiguous reaches (see the parenthetical portion of the equation). The priority index equation was designed so that a species highly endemic to the study area, sporadically distributed, and showing evidence of a range contraction was assigned a high value.

A conservation index for each reach j ($j = 1, \ldots, N$) (CI_j) within the study area was then calculated as the sum of the priority indices:

$$\mathrm{CI}_j = \sum_{k=1}^{S} \mathrm{PI}_{kj}$$

where S is the number of focal species and PI_{kj} is the priority index of species k in reach j. Once calculated, we standardized the CI to range between 0 and 100. Following the methods of Filipe *et al.* (2004), a cumulative-frequency curve of increasing CI values was used to define the threshold for selection of reaches as focal areas for conservation. This technique is designed to optimize the relationship between CI and the number of reaches selected for conservation. A cumulative-frequency curve of CIs was used to determine the threshold value for selecting focal areas: this approach reveals the relationship between the number of reaches conserved and the total conservation value. The inflection point of the curve was first identified, and then the cut-off point was lowered (if necessary) to ensure selection of at least 15% of stream reaches.

Results

Statistical modeling

The predictive model for each species provided an excellent level of discrimination between species presence and absence with AUCs (estimated from the cross-validation analyses) ranging from 0.87 to 0.98 (Table 2). Tree complexities ranged from 3 to 7, and LRs varied from 0.05 to 0.001. For each of the SDMs only 7–12 landscape variables (of 75 possible) exhibited a relative influence of 5% or higher – exemplifying the ability of BRT to select only important variables for explaining response variance. For examples of individual species–environment relationships, see Hopkins (2009) and Hopkins and Burr (2009). The models also explained a high proportion of the response variance, ranging from 71.8 to 96.4%. Generally speaking, model performance was negatively correlated with the geographical range and prevalence of a species: models of species with narrow ranges (*e.g., Thoburnia atripinnis*, *Quadrula cylindrica*) or those that were very rare (*e.g., Percina macrocephala*, *Obovaria retusa*) typically had higher AUCs and a higher proportion of the response variance was explained.

Species attributes and predictive maps

The conservation attributes of species varied markedly (Table 2). Values for endemism (E) ranged from 1 to 13.4 – the majority of species had overall ranges exceeding the study area. However, seven species (six fishes and one mussel) were endemic. *Etheostoma barbouri* is restricted to extreme eastern portions of the study area and received the highest endemic value of 13.4. Sixteen species, including all the mussel species, showed evidence of decreasing ranges ($T = 2$). Rarity was common, with 18 species predicted to occur in 5% or fewer of the stream reaches. Only four species (*e.g., Etheostoma bellum*, *Etheostoma barrenense*, *Etheostoma rafinesquei* and *Villosa ortmanni*) were predicted to occupy more than 10% of the stream reaches. Conservation indices for stream reaches ranged from 0.85 to 100, with a median CI equal to 10.5 (Fig. 2). The highest scoring reaches were located in the Green River and associated headwater areas in the eastern region of the watershed. The upper Barren River also had a number of high scoring reaches. In contrast, only a very few scattered, isolated headwater tributaries in the Nolin River received a high score. For examples of individual distribution probability maps for the focal species, see Hopkins and Burr (2009).

Conservation areas

The inflection point on the CI curve occurred at a value of 37, and less than 7% of stream reaches scored at or above this value. The cut-off value for selecting a reach as a focal area was reduced to ≥ 25; using this value

Table 2. Summary of model performance and conservation attributes for each focal species.

	AUC ± SE	Percent variance explained	Endemism (*E*)	Trend score (*T*)	Proportion of reaches occupied (*O*)
Fish species					
Etheostoma barbouri	0.94 ± 0.05	71.8	13.4	2	0.04
Etheostoma barrenense	0.98 ± 0.02	92.4	4.9	1	0.19
Etheostoma bellum	0.91 ± 0.03	75.9	1.6	1	0.36
Etheostoma kantuckeense	0.87 ± 0.05	83.4	5.3	1	0.05
Etheostoma maculatum	0.92 ± 0.4	86.6	1	2	0.08
Etheostoma rafinesquei	0.89 ± 0.05	73.7	1.8	1	0.24
Percina macrocephala	0.98 ± 0.01	96.4	1	1	0.01
Percina stictogaster	0.96 ± 0.01	93.2	1	1	0.02
Phenacobius uranops	0.91 ± 0.03	85.6	1	2	0.06
Thoburnia atripinnis	0.95 ± 0.03	91.2	11.4	1	0.09
Mussel species					
Alasmidonta marginata	0.92 ± 0.02	92.3	1	2	0.01
Cumberlandia monodonta	0.92 ± 0.03	89.7	1	2	0.01
Cyprogenia stegaria	0.90 ± 0.03	86.3	1	2	0.05
Fusconaia subrotunda	0.89 ± 0.03	87.1	1	2	0.05
Lampsilis abrutpta	0.93 ± 0.04	84.5	1	2	0.01
Lampsilis ovata	0.95 ± 0.03	88.6	1	2	0.03
Obovaria retusa	0.95 ± 0.03	87.8	1	2	0.01
Plethobasus cyphus	0.92 ± 0.04	87.9	1	2	0.02
Pleurobema clava	0.96 ± 0.06	77.4	1	2	0.01
Pleurobema plenum	0.91 ± 0.05	87.2	1	2	0.03
Pleurobema rubrum	0.92 ± 0.05	86.9	1	2	0.02
Quadrula cylindrica	0.97 ± 0.03	92.0	1	2	0.03
Villosa lienosa	0.95 ± 0.07	86.3	1	2	0.01
Villosa ortmanni	0.88 ± 0.06	85.0	1.9	2	0.12

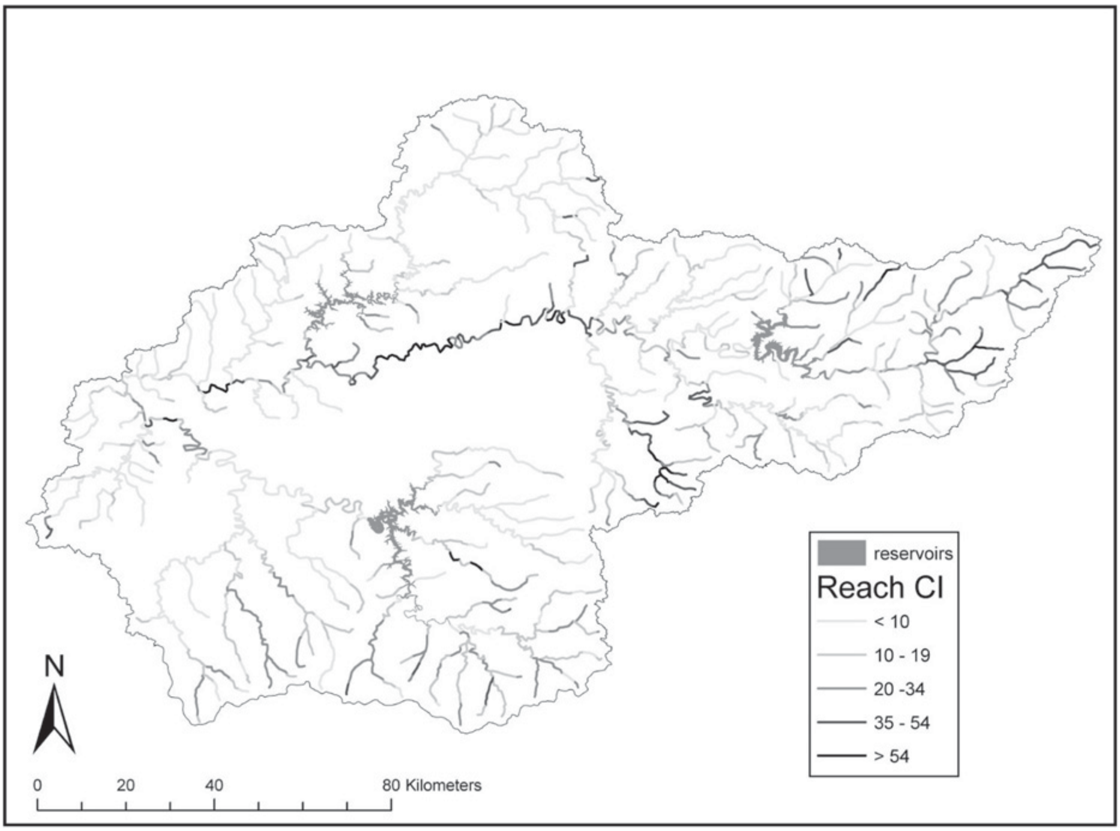

Fig. 2. Conservation index score for each stream reach of the upper Green River catchment. Scores are presented in five classes.

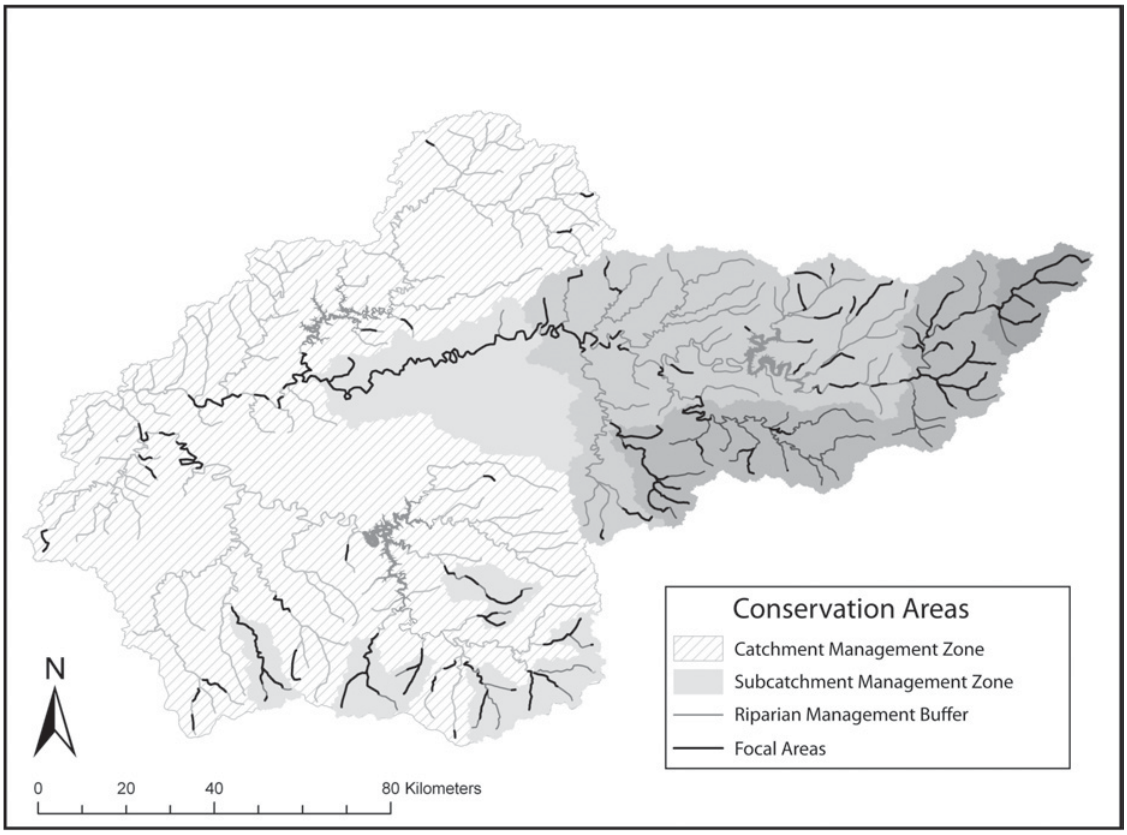

Fig. 3. Suggested conservation areas for the upper Green River catchment identified by spatial scale. Shown are (1) focal areas, (2) riparian management buffers, (3) subcatchment management zones and (4) the catchment management zone. Darker shaded areas are more spatially nested.

designated 15% of the reaches as focal areas. Reducing the cut-off value increased the likelihood of selecting contiguous stream reaches as focal areas (Filipe *et al*., 2004) and conformed to conservation goals of conservation managers to set aside 10–20% of the upper Green River catchment for conservation. The widespread distribution of focal areas warranted identifying the entire catchment as a catchment management zone (Fig. 3). To provide a higher-resolution map of conservation areas, both subcatchment management zones and riparian management buffers were selected by identifying clusters (≥ 3) of focal areas where the CI exceeded the inflection point value of 37 (Fig. 3). This allowed us to further accentuate regions of the watershed to target for conservation. Moreover, these spatial scales closely align with those used in the statistical modeling process and forge a link between predictive models, conservation units and management approaches.

Species–environment relationships

Richness of focal species at the 304 sample sites varied from 0 to 8, with a mean of 3.6 ± 1.8. Nine environmental variables explained the majority (71.1%) of variation in species richness. Soil composition in the subcatchment was important in modeling the richness of focal species: the composition of loamy-skeletal soils, clayey soils and coarse-silty soils contributed a total relative influence of 53.1% in the species richness model. Species richness was positively correlated with percent composition of each of these soil types. Soil composition was positively spatially autocorrelated, thus conferring vital information about stream size, spatial position and topography in the watershed. However, we were interested in the influence of LULC (which is subject to management strategies) on patterns of focal species richness. At the scale of reach buffer, the percent forest and patch densities of developed/exposed land cover (pd_dex) were the two most influential LULC variables on focal species richness, contributing 15.3 and 10.2% to the model, respectively. Interaction plots (Fig. 4) show species richness declining with increasing pd_dex and decreasing forest cover. Species richness increased notably when pd_dex was less than 2 and forest cover exceeded 55%. Similarly, species richness increased when forest cover exceeded 60% and the mean patch size of forest was greater than 15 ha in the riparian buffer (Fig. 4). These variables had a relative influence of 4.9 and 3.3% in the focal species richness model. At the subcatchment scale, the total patch density (a measure of landscape fragmentation) and mean patch size of forest were the most influential LULC variables on focal species riches, contributing 2.9 and 2.1% to the model, respectively. When the landscape fragmentation was minimal

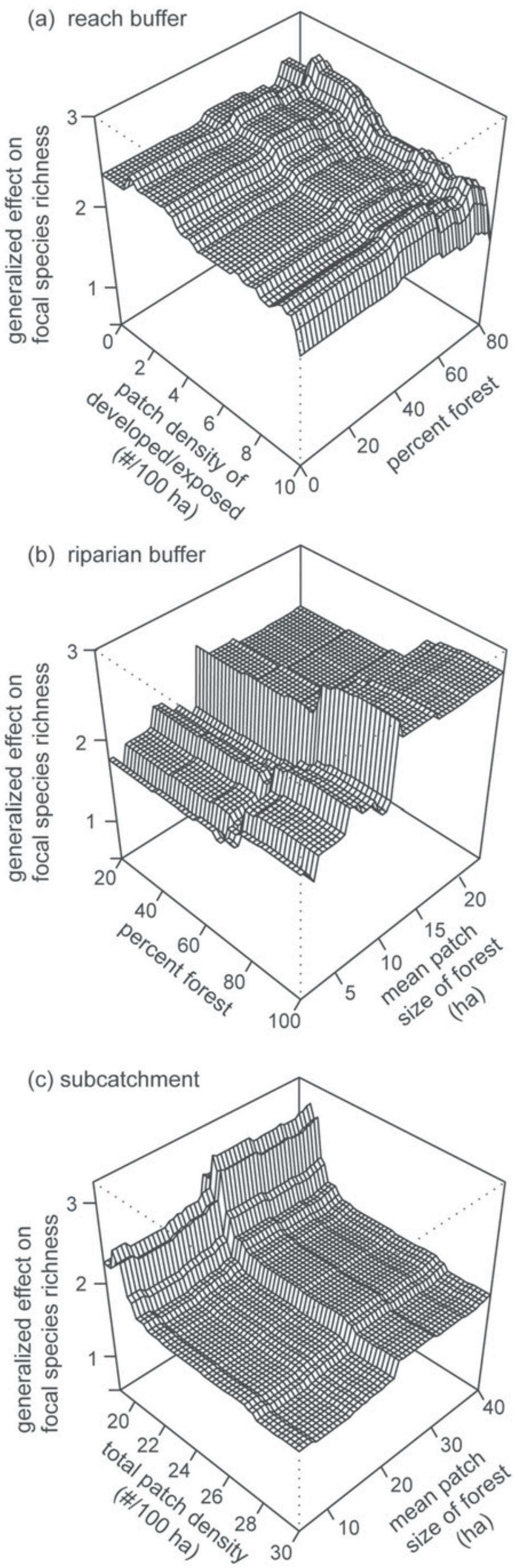

Fig. 4. Pairwise-interaction partial-dependence plots showing relationships of environmental variables and generalized effect on focal species richness. Shown are the two most influential land use/land cover (LULC) environmental variables for each spatial scale analyzed.

(total patch density less than 20) and the mean forest patch size exceeded 30 ha, a marked positive response in species richness was observed (Fig. 4).

Discussion

In this paper, we present a multiscale approach for the selection of management areas for fish and mussel conservation using a case study of the upper Green River catchment, USA. Priority conservation areas were located primarily in the eastern (upper Green River) and southern (upper Barren River) portions of the catchment. And, designating 15% of stream reaches for conservation would protect a very large proportion of aquatic biodiversity and cover the range of all seven endemic species studied. We found that focal species richness is explained most by soil composition in the subcatchment. Although subcatchment soil composition certainly affects stream hydrology and geomorphology, we surmise that the high degree of spatial autocorrelation of the soil types analyzed also confers explicit information about spatial position in the watershed, as well as watershed area. Thus, the relationship between soil composition and focal species distributions may be indirect. Nested within the subcatchment scale focal species richness responded positively to percent forest and negatively to patch density of developed/exposed land in the reach buffer. For both the reach and riparian buffers, retaining forested tracts of land and limiting the level of development and fragmentation would benefit the focal species.

Our approach is anchored in SDMs created using multiscale landscape environmental data, including LULC data. The conservation planning process is made more robust by aligning the spatial scales and environmental variables used at each stage, from species distribution modeling to ranking of stream reaches for conservation, to selection and management of conservation areas. In the end, we identified focal areas, riparian management buffers and subcatchment management zones to target for conservation planning. The congruity between scales used for modeling and those aimed for conservation, as well as the use of LULC data to generate SDMs allow us to provide more constructive management recommendations for conservation areas.

Most past efforts at conservation planning for aquatic systems have focused exclusively on the catchment or subcatchment scales (Collares-Pereira and Cowx, 2004; Moilanen *et al*., 2008) and the implementation of best management practices (Linke *et al*., 2007). These broad scales are important, and the widespread distribution of focal areas in the upper Green River catchment merits designation of the entire catchment as a management zone (Master *et al*., 1998). Accordingly, the KDFWR (2005) has targeted the entire upper Green River catchment for conservation planning and many programs (*e.g*., Green River Bioreserve Program) are already in place to encourage aquatic biodiversity conservation. However, there are detectable "hotspots" we identified within the watershed where conservation efforts can be further focused. Moreover, the issue of identifying best management practices crops up for each spatial scale considered. With the inclusion of LULC data in the modeling process, scientists can provide some level of guidance for resource managers regarding these issues.

A cautionary note, we are not suggesting our ranking methodology for selecting conservation areas supplant more sophisticated reserve-selection approaches. We use this approach to simplify the case study at hand. Being straightforward and easy to implement, our approach may appeal to some resource managers and be sufficient for various scenarios; however, it has serious shortcomings. The conservation index and selection method used do not invoke any criteria for conserving overall species diversity and may result in a number of species being excluded from conservation areas. In this study, at least a portion of the range of every species was included in selected focal areas; however, we recognize that this may not always be the case. Using a complementarity-based approach, which takes into account criteria such as site similarities and differences, could produce contrasting results. Moilanen *et al*. (2008) present an excellent case study demonstrating the use of such an approach in a freshwater ecosystem. In our study, however, meta-population dynamics and the data needed to set parameters for species responses to connectivity are lacking for most of the taxa examined. Certainly, as the spatial scale of a conservation planning project increases, the need to address such matters becomes essential.

Management considerations

Once freshwater conservation areas have been selected, the question of how to manage these resources remains (Moyle and Yoshiyama, 1994; Saunders *et al*., 2002; Abell *et al*., 2007). As has been observed for many aquatic systems in the eastern US (Jones *et al*., 1999; Scott *et al*., 2002; Singkran and Meixler, 2008), forest cover greatly influenced the aquatic fauna at all spatial scales. Within focal areas (*e.g*., reach buffer scale), our models indicate that conservation managers should aim at retaining at least 55% forest cover and highly limit agricultural and urban land uses. Comparably, our models suggest a target of $>60\%$ forest cover with limited fragmentation at the scale of riparian management buffers (riparian buffer scale). Similarly, within the subcatchment management zones (subcatchment scale), large contiguous tracts of forest cover with reduced fragmentation appear very important in maintaining stream habitat and species diversity (Jones *et al*., 1999).

Increased forest cover in riparian zones and subcatchments has been shown to decrease sediment delivery to streams, create more stabile hydrologic regimes and lower nutrient loadings (Allan, 2004). Each of these hydrologic characteristics has direct impact on stream habitat quality, which influences species assemblages. Fish assemblage diversity often peaks when hydrologic variability is low to moderate (Poff and Allan, 1995), likely due to the creation of more heterogeneous habitat. Similarly, McRae *et al.* (2004) found that mussel diversity is positively correlated with flows of moderate to high stability, which shape substrate composition and other microhabitat features.

Despite the recognition that conservation of reach and riparian buffers exerts disproportionately large benefits in relation to the area targeted (Allan, 2004), the subcatchment cannot be ignored. Roth *et al*. (1996) and Wang *et al.* (1997) found that streams draining watersheds consisting of $<50\%$ forest remained relatively degraded regardless of forest cover in the riparian zone. It is likely that processes at the larger subcatchment scale were simply overwhelming more local-scale processes (Jones *et al*., 1999). Thus, a multiscale approach is most appropriate to ensure adequate management of conservation areas.

Notably, at each of the scales, we found that "pristine" environmental conditions were not necessary to realize benefits for the focal species we examined. It is important to recognize that conservation plans for freshwater ecosystems should acknowledge human presence and activity within the conservation areas (Dudgeon *et al*., 2006; Abell *et al*., 2007). Our results suggest that a total preservation approach (*e.g*., no human activity or modification of the watershed) is not always necessary to ensure perpetuation of target species, and this supports some prior research. For example, Poole and Downing (2004) found that retaining 50% forest cover in the reach buffer resulted in no losses or increases in species richness of freshwater mussels in Iowa. Relationships and thresholds are expected to vary with species and specific systems, but these results are encouraging given the reality that humans have modified the majority of watersheds and receiving streams on the planet.

Presently, the upper Green River catchment and riparian buffer is about 65% forested overall. Many sections of riparian buffer, particularly areas along the main stem of the Green River and Barren River, are greater than 70% forested; thus, achieving the goal of retaining about 60% forestation seems reasonable for most areas of the catchment. Moreover, 40 500 ha of land within the catchment have been enrolled into the Conservation Reserve Enhancement Program, which provides incentives to landowners to reforest cleared lands. The catchment is also the location of Mammoth Cave National Park, several state parks and two wildlife management areas, which all strive to maintain lands in natural condition. Thus, many programs would be in place to increase the efficacy of any conservation plan.

In summary, our approach enables conservation planners to develop precise suggestions for landscape management of spatial units comprising conservation areas. Such precise recommendations are necessary for those involved in conservation planning and perhaps will increase the likelihood of success when working with stakeholders and landowners (Collares-Pereira and Cowx, 2004). Another important aspect of such a landscape-based approach is the potential for integration with terrestrial conservation plans (Abell *et al*., 2007;

Moilanen *et al*., 2008). Approaches integrating terrestrial and aquatic systems need to become the standard for conservation efforts as streams and the landscapes they drain are irrevocably linked.

Acknowledgements. We would like to thank two anonymous reviewers for their suggestions; their input greatly improved the quality of this manuscript. This work was supported by a grant from the KDFWR. Partial funding was also provided by the Sigma Xi Grants-in-Aid of Research program.

References

Abell R., 2002. Conservation biology for the biodiversity crisis: a freshwater follow-up. *Conserv. Biol.*, *16*, 1435–1437.

Abell R., Allan J.D. and Lehner B., 2007. Unlocking the potential of protected areas for freshwater. *Biol. Conserv.*, *134*, 48–63.

Allan J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annu. Rev. Ecol. Evol. Syst.*, *35*, 257–284.

Brotons L., Thuiller W., Arau'jo M.B. and Hirzel A.H., 2004. Presence-absence versus presence-only methods for predicting bird habitat suitability. *Ecography*, *27*, 437–448.

Collares-Pereira M.J. and Cowx I.G., 2004. The role of catchment scale environmental management in freshwater fish conservation. *Fisheries Manag. Ecol.*, *11*, 303–312.

Dudgeon D., Arthington A.H., Gessner M.O., Kawabat Z.I., Knowler D.J., Lévêque C., Naiman R.J., Prieur-Richard A.-H., Soto D., Stiassny M.L.J. and Sullivan C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, *81*, 163–182.

Elith J., Graham C.H., Anderson R.P., Dudík M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J., Peterson A.T., Phillips S.J., Richardson K.S., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S. and Zimmermann N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*, 129–151.

Elith J., Leathwick J.R. and Hastie T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.*, *77*, 802–813.

Environmental System Research Institute, 2006. ArcGIS/ArcInfo for Windows XP Version 9.2.0, Environmental System Research Institute, Redlands, CA, USA.

Filipe A.F., Marques T.A., Seabra S., Tiago P., Ribeiro F., Moreria Da Costa L., Cowx I.G. and Collares-Pereira M.J., 2004. Selection of priority areas for fish conservation in Guadiana River basin, Iberian Peninsula. *Conserv. Biol.*, *18*, 189–200.

Guisan A. and Thuiller W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.*, *8*, 993–1009.

Hirzel A. and Guisan A., 2002. Which is the optimal sampling strategy for habitat suitability modeling. *Ecol. Model.*, *157*, 331–341.

Hopkins R.L. II, 2009. Use of landscape pattern metrics and multiple spatial scales in aquatic species distribution models. *Landscape Ecol.*, *24*, 943–955.

Hopkins R.L. II and Burr B.M., 2009. Modeling freshwater fish distributions using multiscale landscape data: a case study of six narrow range endemics. *Ecol. Model.*, *220*, 2024–2034.

Hopkins R.L. II, Burns M.D., Burr B.M. and Hopman L.J., 2009. Building a centralized database for Kentucky fishes: progress and future applications. *J. Kentucky Acad. Sci.*, *69*, 73–78.

Jones E.B.D. III, Helfman G.S., Harper J.O. and Boltstad P.V., 1999. Effects of riparian forest removal on fish assemblages in southern Appalachian streams. *Conserv. Biol.*, *13*, 1454–1465.

Kentucky Department of Fish and Wildlife Resources, 2005. Kentucky's Comprehensive Wildlife Conservation Strategy. Accessed online 2 May 2010, http://fw.ky.gov/kfwis/stwg.

Linke S., Pressey R.L., Bailey R.C. and Norris R.H., 2007. Management options for river conservation planning: condition and conservation re-visited. *Freshwater Biol.*, *52*, 918–938.

Master L.L., Flack S.R. and Stein B.A., 1998. Rivers of life: critical watersheds for protecting freshwater biodiversity, The Nature Conservancy, Arlington, MA, 71 p.

McGarigal K. and Marks B.J., 1995. FRAGSTATS. Spatial analysis program for quantifying landscape structure, USDA Forest Service General Technical Report PNW-GTR-351.

McRae S.E., Allan J.D. and Burch J.B., 2004. Reach and catchment-scale determinants of the distribution of freshwater mussels (Bivalvia: Unionidae) in southeastern Michigan (U.S.A.). *Freshwater Biol.*, *49*, 127–142.

Moilanen A., Leathwick J. and Elith J., 2008. A method for spatial freshwater conservation prioritization. *Freshwater Biol.*, *53*, 577–592.

Moyle P.B. and Yoshiyama R.M., 1994. Protection of aquatic biodiversity in California: a five- tiered approach. *Fisheries*, *19*, 6–17.

Naeem S. and Wright J.P., 2003. Disentangling biodiversity effects on ecosystem functioning: deriving solutions to seemingly insurmountable problems. *Ecol. Lett.*, *6*, 567–579.

Neves R.J., Bogan A.E., Williams J.D., Ahlstedt S.A. and Hartfield P.W., 1997. Status of aquatic mollusks in the southeastern United States: a downward spiral of diversity. *In*: Benz G.W. and Collins D.E. (eds.), Aquatic fauna in peril: the southeastern perspective, Special Publication 1, Southeast Aquatic Research Institute, Lenz Design and Communications, Decatur, 43–85.

Poff N.L. and Allan J.D., 1995. Functional organization of stream fish assemblages in relation to hydrologic variability. *Ecology*, *76*, 606–627.

Poole K.E. and Downing J.A., 2004. Relationship of declining mussel biodiversity to stream-reach and watershed characteristics in an agricultural landscape. *J. N. Am. Benthol. Soc.*, *23*, 114–125.

R Development Core Team, 2006. R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna.

Rahel F.J., 2000. Homogenization of fish faunas across the United States. *Science*, *288*, 854–856.

Ricciardi A. and Rasmussen J.B., 1999. Extinction rates of North American freshwater fauna. *Conserv. Biol.*, *13*, 1220–1222.

Ridgeway G., 2006. Generalized boosted regression models. Accessed online 16 May 2009, ftp://ftp.auckland.ac.nz/pub/software/CRAN/doc/packages/gbm.pdf.

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

Saunders D.L., Meeuwig J.J. and Vincent A.C.J., 2002. Freshwater protected areas: strategies for conservation. *Conserv. Biol.*, *16*, 30–41.

Scott M.C., Helfman G.S., McTammany T.E., Benfield E.F. and Boltstad P.V., 2002. Multiscale influences on physical and chemical stream conditions across Blue Ridge landscape. *J. Am. Water Resour. Assoc.*, *38*, 1379–1392.

Singkran N. and Meixler M.S., 2008. Influences of habitat and land cover on fish distributions along a tributary to Lake Ontario, New York. *Landscape Ecol.*, *23*, 539–551.

Srivastava D.S. and Vellend M., 2005. Biodiversity-ecosystem function research: is it relevant to conservation? *Annu. Rev. Ecol. Evol. Syst.*, *36*, 267–294.

Strayer D.L., 2006. Challenges for freshwater invertebrate conservation. *J. N. Am. Benthol. Soc.*, *25*, 271–287.

Tilman D., 1999. The ecological consequences of changes in biodiversity: a search for general principles. *Ecology*, *80*, 1455–1474.

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

Warren M.L. Jr, Angermeier P.L., Burr B.M. and Haag W.R., 1997. Decline of a diverse fish fauna: patterns of imperilment and protection in the south-eastern United States. *In*: Benz G.W. and Collins D.E. (eds.), Aquatic fauna in peril: the southeastern perspective, Special Publication 1, Southeast Aquatic Research Institute, Lenz Design and Communications, Decatur, 105–164.