Ecology and Evolution

Open Access

Predicting fine-scale distributions of peripheral aquatic species in headwater streams

Christopher R. DeRolph[1], Stacy A. C. Nelson[1], Thomas J. Kwak[2] & Ernie F. Hain[1]

[1]Center for Geospatial Analytics, College of Natural Resources, North Carolina State University, Box 7106, Raleigh, North Carolina 27695
[2]U.S. Geological Survey, North Carolina Cooperative Fish and Wildlife Research Unit, North Carolina State University, Box 7617, Raleigh, North Carolina 27695

Keywords
Conservation planning, habitat modeling, headwater streams, landscape variables, peripheral populations, species distributions, topographic gradient, wild trout.

Correspondence
Current address for Christopher R. DeRolph: Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831.
Tel: 919-621-2903; Fax: 865-576-3989;
E-mail: chrisderolph@gmail.com

Funding Information
This research was funded by Federal Aid in Sport Fish Restoration Funds.

Received: 21 April 2014; Revised: 10 September 2014; Accepted: 7 November 2014

***Ecology and Evolution* 2015; 5(1): 152–163**

doi: 10.1002/ece3.1331

Abstract

Headwater species and peripheral populations that occupy habitat at the edge of a species range may hold an increased conservation value to managers due to their potential to maximize intraspecies diversity and species' adaptive capabilities in the context of rapid environmental change. The southern Appalachian Mountains are the southern extent of the geographic range of native *Salvelinus fontinalis* and naturalized *Oncorhynchus mykiss* and *Salmo trutta* in eastern North America. We predicted distributions of these peripheral, headwater wild trout populations at a fine scale to serve as a planning and management tool for resource managers to maximize resistance and resilience of these populations in the face of anthropogenic stressors. We developed correlative logistic regression models to predict occurrence of brook trout, rainbow trout, and brown trout for every interconfluence stream reach in the study area. A stream network was generated to capture a more consistent representation of headwater streams. Each of the final models had four significant metrics in common: stream order, fragmentation, precipitation, and land cover. Strahler stream order was found to be the most influential variable in two of the three final models and the second most influential variable in the other model. Greater than 70% presence accuracy was achieved for all three models. The underrepresentation of headwater streams in commonly used hydrography datasets is an important consideration that warrants close examination when forecasting headwater species distributions and range estimates. Additionally, it appears that a relative watershed position metric (e.g., stream order) is an important surrogate variable (even when elevation is included) for biotic interactions across the landscape in areas where headwater species distributions are influenced by topographical gradients.

Introduction

Headwater streams are often underrepresented in commonly used topographic maps in the United States and abroad despite their prevalence on the landscape and the importance of their hydrologic and biological form and function (Meyer et al. 2007; Storey et al. 2011). The biota of headwater streams represent many functional biological groups that are important components of biodiversity in a river network. According to Meyer et al. (2007), fishes that use headwaters can be classified into three general classes: (1) specialists that use headwater streams throughout the year, (2) generalists that use headwaters as one of many habitats, and (3) fish that live in large streams but use small streams for spawning and nursery areas. Headwater species may be vulnerable to extirpation due to the variability in conditions of headwater streams, but may also find refuge in headwater streams from threats in other parts of the stream network.

Peripheral populations that occupy habitat at the edge of a species' geographic range are assumed to occupy marginal habitat and are often more isolated than populations closer to the core of a species range (Sagarin and Gaines 2002; Sagarin et al. 2006; Haak et al. 2010).

As a result of higher degrees of genetic drift and selective pressures, genetic characteristics of peripheral populations are thought to be potentially more distinctive and adaptive compared to larger, more stable core populations (Lesica and Allendorf 1995; Nielsen et al. 2001; Hardie and Hutchings 2010). Due to their potential to maximize intraspecies diversity and species' adaptive capabilities and their importance for species persistence in the context of climate change, peripheral populations may warrant a higher conservation priority to managers (Lesica and Allendorf 1995; Nielsen 1999; Nielsen et al. 2001; Haak et al. 2010; Hardie and Hutchings 2010).

The worldwide distribution and abundance of native salmonid populations have undergone large-scale declines for the previous century or more (Piccolo 2011). Recent research has advocated for implementation of a diverse management portfolio into native trout conservation strategies that aims to increase representation, resilience, and redundancy (Haak and Williams 2012). New strategies focused on improving resistance and resilience of natural systems – such as increasing connectivity, increasing the number and size of habitat reserves, and intensive management of stressed populations – have been proposed to aid in climate change adaptation strategies for fishery managers (Haak and Williams 2012).

Distribution modeling using GIS has made it possible to predict fish distributions across a region using landscape-scale variables, and given the logistical and economical limitations of field sampling, it is perhaps the only realistic way to obtain spatially comprehensive fine-scale distribution information to be applied in resource management decisions at the regional scale. Although in-stream habitat variables cannot be included in this type of modeling, landscape variables can be extracted at multiple scales to allow for inclusion of local and accumulative upstream conditions. Several recent studies have demonstrated the efficacy of predicting freshwater fish distributions from GIS-derived landscape variables (Steen et al. 2006, 2008; Brewer et al. 2007; Dauwalter and Rahel 2008; Hopkins and Burr 2009; McKenna and Johnson 2011; Kristensen et al. 2012; Sindt et al. 2012; Filipe et al. 2013; Maloney et al. 2013; McKenna et al. 2013).

The southern Appalachian Mountains are the southern extent of the geographic range of native brook trout *Salvelinus fontinalis* (Fig. 1) and naturalized rainbow trout and brown trout in eastern North America (Maccrimmon and Campbell 1969; Flebbe 1994). From the 1930s to the 1970s in the Great Smoky Mountains of North Carolina and Tennessee, brook trout populations experienced shifts to upper headwater streams due to success of introduced rainbow trout (Kelly et al. 1980; Larson and Moore 1985) and brown trout at lower elevations.

Figure 1. Brook trout *Salvelinus fontinalis*.

The shift over time of brook trout into headwater streams due to encroachment by rainbow trout and brown trout has resulted in isolation of populations, which limits gene flow and can cause genetic drift and inbreeding (Stoneking et al. 1981; Fausch et al. 2009). Genetic research to evaluate suspected phylogenetic differences between southern Appalachian and hatchery reared, northern-derived brook trout has been a prolific topic in the literature, and there are numerous studies that indicate distinct genotypic differences do exist between the two. Long-term survival of brook trout and reductions in their distribution in the southern Appalachians has been a concern for fisheries resource managers since the early twentieth century (King 1937), and much emphasis has been placed on conservation and restoration of brook trout in recent research and management efforts.

Fisheries managers suspect that the quality of salmonid sport fisheries in the southern Appalachians has declined over time, resulting in the gradual reduction of self-sustaining wild trout populations. Salmonid fisheries are vital economic, cultural, and recreational resources in the southern Appalachian Mountains as well as ecological indicators of stream quality (Habera et al. 2005; NCWRC 2009). As native stocks continue to be extirpated, fish ecologists are tasked with predicting the distribution and abundance of the remaining salmonid stocks to provide decision-support to managers and policymakers responsible for conservation and restoration of threatened or depleted populations (Piccolo 2011). Fish occurrence data for brook trout, rainbow trout, and brown trout exist for only a small portion of the total number of stream reaches that likely support these species. A landscape-level tool to integrate information on stream salmonid distribution would serve a vital need for ecological understanding and fisheries planning and management, given the threats to these peripheral populations from numerous forms of environmental degradation.

Our goal was to provide a tool to assist resource managers in prioritization of conservation efforts for wild trout across the landscape of western North Carolina. Specifically, the objective of this study was to use readily available or easily calculated GIS-derived landscape variables as predictors of occurrence in correlative logistic regression models for each species. To overcome the underrepresentation of headwater streams in more commonly used hydrography datasets like the USGS National Hydrography Dataset, model predictions were applied to a stream network generated specifically for this project that captures potential headwater stream trout habitat (i.e., the potential species range) more consistently across the study area.

Methods

Study area

We used EPA Level IV ecoregion boundaries to determine the study area boundary (Fig. 2). We included all of the westernmost ecoregions in North Carolina that encompass the mountainous terrain. We extended the eastern study area boundary sixteen kilometers to the east, because the ecoregion boundaries are somewhat arbitrary and we wanted to ensure that the entirety of the southern Appalachian Mountains (excluding the foothills) within North Carolina was included. The study area encompassed approximately 23,411 km^2.

Fish survey data

The North Carolina Wildlife Resources Commission (NC WRC) collected trout occurrence data in the field from 1968 to 2009. We sought to model current trout distributions, so we identified a temporal range of empirical data that represented adequate spatial distribution and sample sizes for each species (i.e., sufficient presence and absence points). Field surveys conducted from approximately 2001 to 2009 were primarily focused on private lands in the northeastern portion of the study area, while those conducted from the mid-1990s through 2000 were heavily focused on public lands in the southwestern portion of the study area. We also wanted to use fish data from a time period relatively close to dates at which changeable GIS-derived landscape predictors were derived (e.g., land cover, road density). Based on these criteria, we elected to apply data collected from 1995 to 2009 (Fig. 2) to model trout distributions.

Development of the GIS Framework

The GIS framework implemented for this study was developed using methods proposed by Brenden et al. (2006) for a standardized GIS methodology for stream ecological

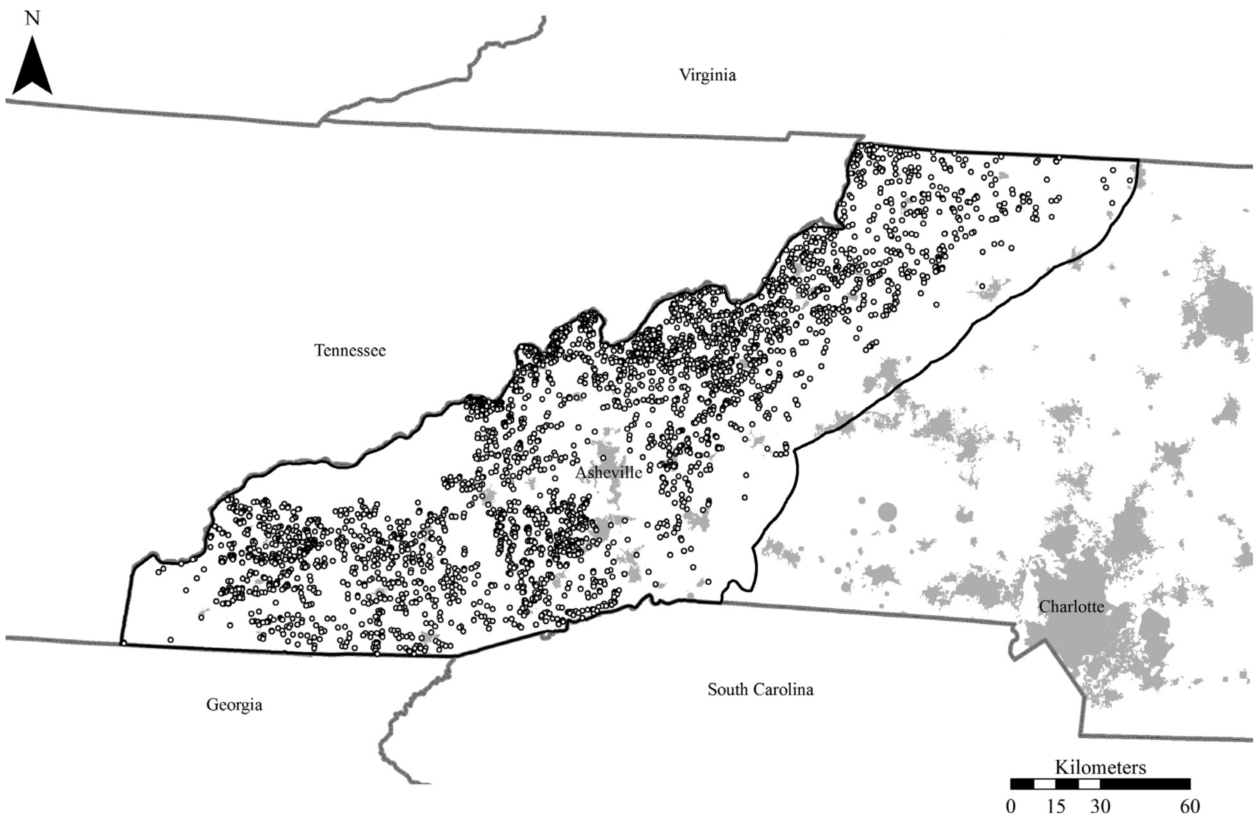

Figure 2. Study area with field data collections points for fish surveys conducted from 1995 to 2009.

research. The framework was intended for research related to the functional linkages between rivers and multiscale landscape variables and has been shown to perform well in predictive modeling with the presence–absence fish data collected on a regional scale (Steen et al. 2006).

We generated our own stream network using Arc Hydro Tools Version 1.3 (ESRI 2009) rather than using the National Hydrography Dataset (NHD) because of underrepresentation of headwater streams and inconsistencies in the density of stream networks in the NHD, including significant documented deficiencies in the study area (Meyer et al. 2001; Colson 2006). Omission errors found along USGS topographic map boundaries (Fig. S1), potentially due to different map publication dates or different cartographers working on adjacent quadrangles, result in frequent abrupt termination of streams at the boundary of map sheets. Trout in western North Carolina inhabit headwater streams, so it was important to have a stream network that depicted headwater streams consistently across the study area because they represent the potential species ranges. We used a drainage area threshold of 0.024 km^2 (six acres) as prescribed by the North Carolina Stream Mapping Project for western North Carolina. This threshold, which will reportedly capture approximately 95% of intermittent and perennial stream breakpoints (NCDENR 2005), was applied to one arc second resolution elevation data from the National Elevation Dataset to generate the stream network.

Landscape variables

Candidate predictor variables (Table S1) were chosen based on the literature and professional opinions of NC WRC personnel. The metrics were selected to represent elements that have been linked to fish assemblage structure, and they fall into several broad categories including land cover characteristics, watershed position, surficial geology, climate, terrain, fragmentation, land ownership (i.e., private or public), and hydrologic connectivity to impoundments.

Model development

Multiple logistic regression performs well when fitting models with binary response variables and either categorical or continuous predictor variables (Hosmer and Lemeshow 1989) and is applicable for predictive modeling with the presence–absence fish data collected at a regional scale (Rashleigh et al. 2005; Steen et al. 2006; Dauwalter and Rahel 2008; Sindt et al. 2012).

We partitioned the NC WRC data into calibration and validation subsets (Table 1) using guidelines provided by Huberty (1994) that recommend using 70% of the data for calibration and 30% for validation and Harrell (1997) that recommended that the sample size for the number of absences to be included in the prediction model should be roughly three times the number of presences. To compare model structures (see Appendix S1) and select the final models, we evaluated the AIC score (Akaike 1974), Wald chi-square test statistic (Wald 1943), area under the receiver operating characteristic curve (AUC), and the Hosmer-Lemeshow goodness-of-fit test statistic (Hosmer and Lemeshow 1989). When selecting the final models, we chose the model that had the lowest AIC (Burnham and Anderson 2002) and highest AUC values that also had a significant ($P < 0.05$) Wald chi-square test statistic and a nonsignificant ($P > 0.05$) Hosmer-Lemeshow goodness-of-fit test statistic.

Table 1. Number of presences and absences in the calibration and validation datasets for each trout species.

	Calibration data		Validation data	
Species	Presences	Absences	Presences	Absences
Brook trout	309	927	133	1947
Rainbow trout	520	1560	222	1014
Brown trout	267	801	114	2134

Model validation

We evaluated model accuracy using both statistics derived from an error matrix (Fielding and Bell 1997) and AUC values. We reported three statistics derived from the error matrix: presence accuracy, absence accuracy, and average accuracy (the mean of the presence and absence accuracy values) (Steen et al. 2008). Binary predictions of "present" or "absent" are necessary to evaluate model accuracy using statistics derived from the error matrix. Therefore, continuous model outputs must be converted into binary predictions by setting a threshold probability value above which the species is predicted to be present. For each species model, we elected to use the threshold value that maximized the sum of the presence accuracy and the absence accuracy (Manel et al. 2001; Jimenez-Valverde and Lobo 2007). Following terminology from Steen et al. (2006), we termed this measure "performance." As model accuracy statistics derived from an error matrix are threshold dependent, we also present the AUC values because they are threshold-independent measures of model accuracy.

Results

Model selection

Nine variables were included in the final brook trout model (Table 2, Table S2), which was selected using the

Table 2. Habitat variables used in models for each species.

General influence	Habitat variable (effect scale)	Model(s)
Fragmentation	Number of stream crossings (S)	BKT
	Number of stream crossings (W)	RBT, BNT
	Mean road density (N)	BKT
Land cover	Percent urban land (R)	BKT
	Percent forest (N)	RBT, BNT
Terrain	Mean elevation (N)	BKT, BNT
	Mean slope (N)	BKT, BNT
Watershed position	Strahler stream order	BKT, RBT, BNT
	Shreve stream order	BKT
Climate	Mean annual precipitation (W)	BKT, RBT, BNT
Surficial geology	Percent fine-grained soils (W)	BKT, BNT

W, entire upstream watershed; N, entire upstream riparian corridor; S, local watershed; R, local riparian corridor. Model field indicates which models incorporated the habitat variables. BKT, brook trout; RBT, rainbow trout; BNT, brown trout.

backward selection method with a significance level of $P < 0.10$ for a variable to remain in the model. Four variables were included in the final rainbow trout model (Table 2, Table S3), which was selected using the forward, backward, and forward–backward selection methods with a significance level of $P < 0.01$ for a variable to enter into or to remain in the model (depending on the selection method). Seven variables were included in the final brown trout model (Table 2, Table S4), which was selected using the forward, backward, and forward–backward selection methods with a significance level of $P < 0.05$ for a variable to enter into or to remain in the model (depending on the selection method). Both the manual stepwise and best subsets methods produced reasonable models, but our evaluation criteria ranked them below at least one of the models produced by the automated selection techniques.

Each of the final models had four significant metrics in common: Strahler stream order, number of road crossings (a measure of stream network fragmentation), mean annual precipitation, and percent forest/urban land cover (forest and urban land cover were moderately correlated). We used odds ratios to estimate the effect size of variables selected in the final models (Nielsen et al. 2008) (Tables S2, S3, and S4), with the understanding that we can only interpret the effect of a predictor on the response given the other predictors in the model. Strahler stream order had the highest odds ratio in the final models for rainbow trout and brown trout and the second highest odds ratio in the final brook trout model. Elevation had the highest odds ratio in the brook trout model and the second highest odds ratio in the brown trout model.

Based on our stream network delineation, there are over 95,000 km of streams within the study area. Our models indicate that about 30% (28,000 km) of these streams can support wild trout (Table S5). According to our predictions, approximately 27% of the wild trout streams support allopatric brook trout populations and another 31% support sympatric populations. Approximately 60% of the wild trout streams were predicted to support either allopatric or sympatric populations of rainbow trout, and 46% were predicted to support brown trout.

Model validation

All models provided good levels of discrimination with AUCs ranging from 0.72 to 0.80 (Table 3). Greater than 70% presence accuracy was achieved for all three models. Average accuracy and performance values ranged from 64.86% and 129.71 for rainbow trout to 72.19% and 144.37 for brook trout, respectively.

Predicted distribution maps

There are over 400,000 interconfluence stream reaches in the stream network generated for the study area, and occurrence for each species was predicted for each of these stream reaches. We produced a series of 1:50,000 scale distribution maps (see Fig. S2 for a sample) for use

Table 3. Model validation results.

Species	Threshold of occurrence (0–1)	Presence accuracy (%)	Absence accuracy (%)	Average accuracy (%)	Performance (presence + absence)	AUC value (0–1)
Brook trout	0.28	70.7	73.7	72.2	144.4	0.802
Rainbow trout	0.21	75.7	54.0	64.9	129.7	0.716
Brown trout	0.25	71.9	67.0	69.5	139.0	0.792

Threshold of occurrence values represent probability values above which the species is predicted to be present. The presence and absence accuracies were calculated by applying the validation datasets in a error matrix. Average accuracy represents the average of the presence and absence accuracies. Performance represents the sum of the presence and absence accuracies. AUC values represent the area under the receiver operating characteristic curve.

in resource management efforts that depict our stream network with occurrence predictions (please contact the corresponding author for access to these maps).

We also generated generalized distribution maps for each species that show predicted distributions across the entire study area and a density surface created from the predicted probabilities for each stream reach. It appears that the areas with the highest probability of brook trout occurrence (Fig. 3) are the highlands in the Great Smoky Mountains National Park and Nantahala National Forest in the southwestern portion of the study area and privately owned lands located at higher elevations in the northwestern portion of the study area. Areas with the highest probability of rainbow trout occurrence (Fig. 4) are the public lands in the southwestern portion of the study area. Probabilities of occurrence for brown trout (Fig. 5) were generally low across the study area, with the exception of a portion of the private lands in the northwestern portion of the study area. The spatial distribution patterns shown on our predicted probability maps appear to closely agree visually with patterns found by Flebbe (1994) in a study of the distributions of brook trout, rainbow trout, and brown trout and their relation to latitudinal and elevational gradients in Virginia and North Carolina.

Discussion

We applied a multiple logistic regression approach to a GIS to predict the distributions of three salmonids whose ranges include headwater streams. Overall, about 7 of 10 predictions were accurate for the occurrence models, which suggests that landscape-scale variables alone can be applied to predict occurrence of salmonids in our study area when comprehensive local-scale data are unavailable.

Implications for headwater stream ecology

Headwater streams provide many unique functions across a landscape and are important terrestrial-aquatic and headwater-downstream linkages (Meyer et al. 2001, 2007; Lowe and Likens 2005; Wipfli et al. 2007; Clarke et al. 2008; Storey et al. 2011). Strahler stream order was identified as the most influential variable in two of the three final models and the second most influential variable in the other model, which supports our decision to generate

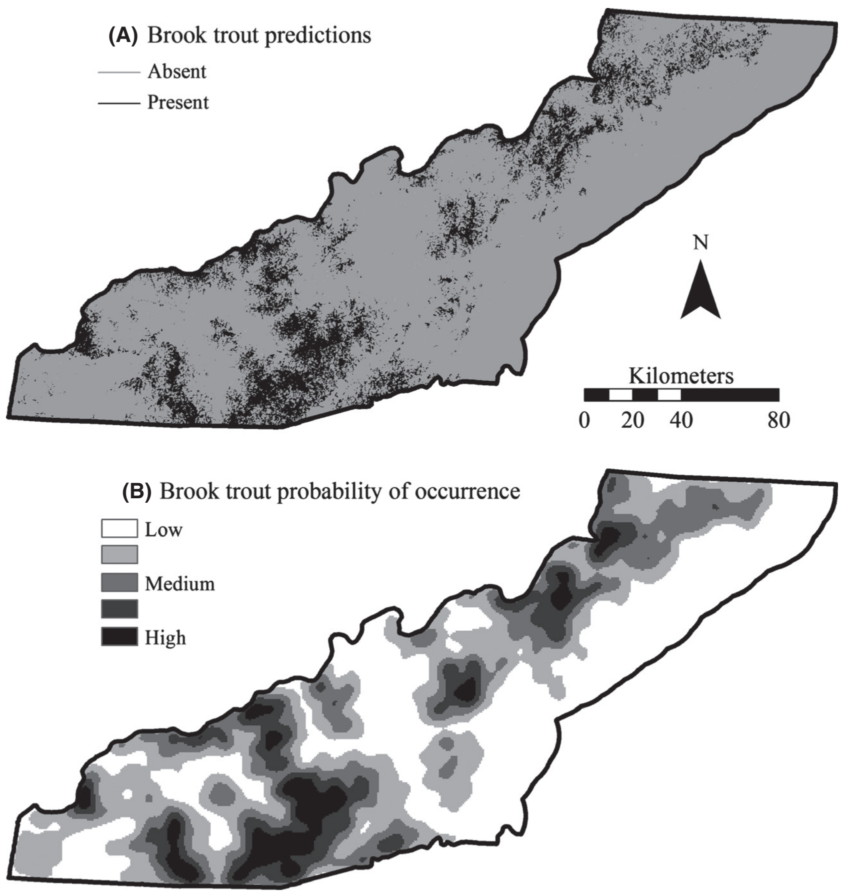

Figure 3. Generalized brook trout distribution maps showing presence/absence predictions and probabilities of occurrence across the study area. Probability surface created by calculating the density of predicted probabilities for each stream reach. Probabilities separated into five classes using Jenks natural breaks method.

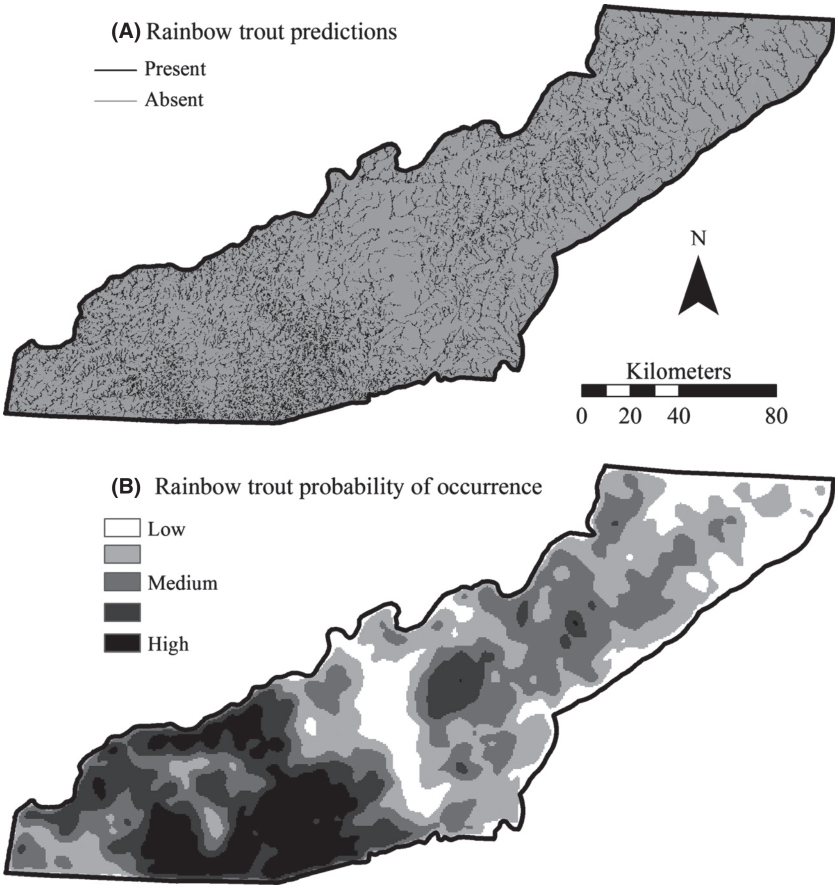

Figure 4. Generalized rainbow trout distribution maps showing presence/absence predictions and probabilities of occurrence across the study area. Probability surface created by calculating the density of predicted probabilities for each stream reach. Probabilities separated into five classes using Jenks natural breaks method.

our own stream network that represents headwater streams and potential species ranges more consistently across the study area. When mapping species potential ranges and predicting species distributions in headwater systems, it is important to capture the potential species range as accurately as possible so that potential habitat is not excluded or underrepresented.

Conservation and management implications

Model results can serve as baseline estimates for wild trout distributions in the study area and be used as a planning tool by resource managers to aid in establishing a diverse management portfolio that aims to maximize representation, resilience, and redundancy of these peripheral wild trout populations. Knight et al. (2009) identified eight "hallmarks of best practice" to promote delivery of effective conservation planning through integration of sound design and application of spatial prioritization techniques. Our methodology and findings partially or entirely satisfy the first six hallmarks: (1) identify who wants the assessment and the products they need, (2) situate spatial prioritization techniques within a broader operational model, (3) involve experts and implementers in the spatial prioritization, (4) identify conservation opportunities not simply conservation priorities, (5) translate prioritization outputs into planning products for end-users, and (6) complement planning products with an implementation strategy.

Most existing conservation lands were not intended primarily for conservation of aquatic species, and thus we anticipated finding gaps between predicted species populations and protected lands. Identifying potential conservation areas with the goal of aquatic species conservation should be more effective than the more traditional ad hoc methods (Frissell et al. 1986; Dudgeon et al. 2006), but deciding where on the landscape to invest conservation resources can be a challenge for managers. One option is to focus on headwater catchments because of their connectivity to downstream ecosystems and their contribution to biodiversity and biological integrity of river networks (Roth et al. 1996; Meyer et al. 2001).

The modeling approach implemented in this study can be used to build "what-if" scenarios that evaluate changes over time in trout distributions or predict future trout distributions based on environmental change pertinent to

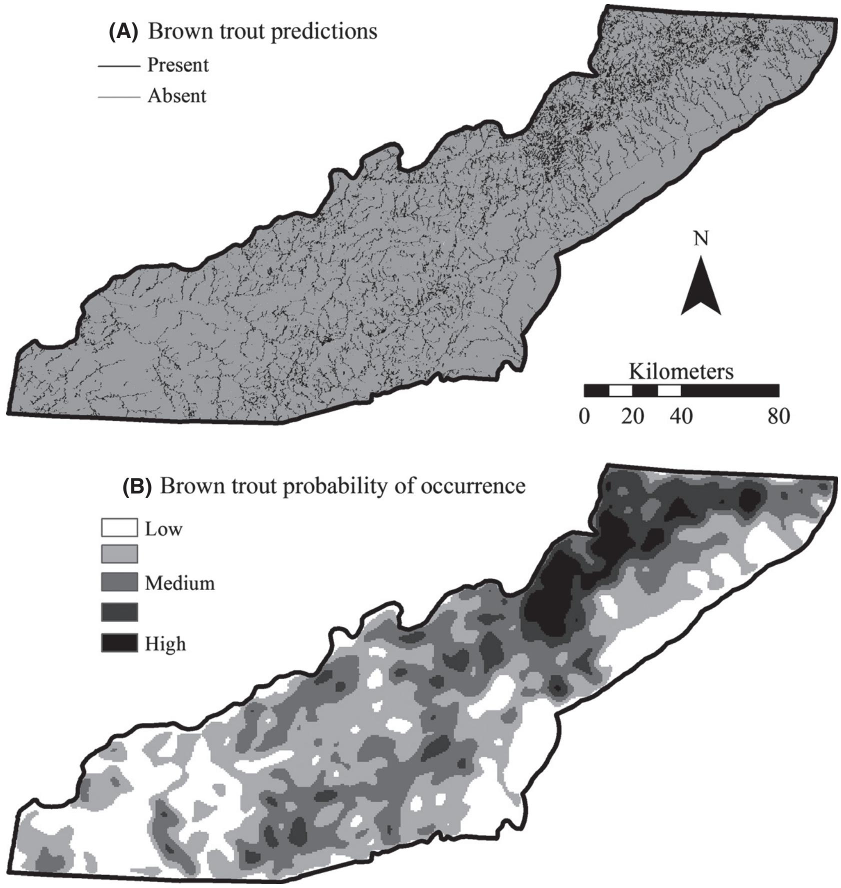

Figure 5. Generalized brown trout distribution maps showing presence/absence predictions and probabilities of occurrence across the study area. Probability surface created by calculating the density of predicted probabilities for each stream reach. Probabilities separated into five classes using Jenks natural breaks method.

the southern Appalachians, such as climate change, urbanization, and hemlock mortality. Each of our final models included climate and land cover variables that could be manipulated to build what-if scenarios. Due to long-term changes in water temperature and flow resulting from climate change, distributions of some salmonids, including brook trout, are expected to become even more constrained and fragmented (Haak et al. 2010; Isaak et al. 2010; Wenger et al. 2011) until only fragile populations remain in small headwater stream thermal refuges (Flebbe et al. 2006; Haak et al. 2010; Argent and Kimmel 2013; Whiteley et al. 2013; Trumbo et al. 2014). Roberts et al. (2009) found that the replacement of hemlock overstory with hardwood species in the Great Smoky Mountains National Park will have a negligible effect on long-term stream conditions, but impacts from the transition period from hemlock to hardwoods are likely to be significant. See Appendix S2 for additional potential management implications of this modeling effort.

Habitat variable selection

All predictor variables used in this study were GIS-derived landscape-scale variables, some of which were calculated at multiple scales. The integration of multiscale data capable of capturing different attributes of species biogeography is crucial to developing strong species distribution models (Hopkins and Burr 2009). In general (see Appendix S3 for details), the final models included the "natural" metrics of watershed position, terrain, climate, and surficial geology and the "anthropogenic" metrics of land cover and fragmentation.

Model limitations and sources of error

There are several types of error inherent in this correlative modeling approach. The stream lengths presented for our entire stream network and for the species predicted distributions are sensitive to the application of the 0.024 km^2 km drainage area threshold, and thus these length estimates may be best used for relative comparisons of species ranges and potential niche occupation than for true channel length estimates. When using large existing datasets for regional studies, data quality, and resolution are a valid concern. We used the highest resolution dataset we could locate for each landscape metric we included in our analysis and, given that a few of our landscape metrics (e.g., climate and soil data) were obtained from relatively coarse resolution maps, they lack the desired scale and accuracy that was attained for our higher resolution stream network.

A disconnect may exist between the spatial scale at which fish data were collected in the field and the scale at which our ecological models operate. This disparity of scales can affect the apparent importance of a habitat variable. The perceived importance of a particular habitat feature can also be affected and confounded by incorporating datasets with varying resolution (Lammert and Allan 1999; Brewer et al. 2007). Annual precipitation was included in all three of our final models and low-resolution soil characteristics were included in the models for brook trout and brown trout, so it is possible that issues of scale with these datasets contributed to model error (Steen et al. 2008).

Temporal variation in species distributions was not included in our models because the fish survey data came from single samples of individual stream reaches (Wiley et al. 1997; Steen et al. 2008). The fish data were collected over a 15-year period by different biologists for varying purposes, and it was impossible to determine which samples may have included fish misidentifications or to estimate detection probability for the three trout species. Species absences can arise either from a genuine lack of species presence or from the inability to detect a species at a particular site if it was present (Oakes et al. 2005), and our random selection of absence sites from the field survey data possibly included some segments where trout were present.

There appears to be inadequate representation of larger, higher order streams in the sampling data, perhaps due to wadeability constraints. The majority (85%) of the sampling was carried out in second, third, and fourth Strahler order streams. The spatial bias in sampling may reflect an absence of wild trout in higher order streams, but we were unable to incorporate such absences into our models without explicit data. Using flawed training data to build models can reduce model quality, but the errors in the training data are manifested in model accuracy measurements so they are reflected in our results (for an in-depth look at addressing deficiencies in field collection of training data see Vaughan and Ormerod (2003)).

While our approach does implicitly incorporate ecological processes driven by spatial variation in species traits related to landscape characteristics, mechanistic relationships between species functional traits and their environment are not explicitly included (Filipe et al. 2013). However, correlative modeling approaches have practical advantages over mechanistic methods due to their relatively flexible and simplistic data requirements and their capacity to incorporate both biotic and abiotic interactions (Kearney et al. 2010).

Research has shown that species interactions between brook trout, rainbow trout, and brown trout can be very important drivers of species distributions (Nyman 1970; Fausch and White 1981; Larson and Moore 1985; Dewald and Wilzbach 1992; Weaver and Kwak 2013). While some recent species distribution modeling research (Wenger et al. 2011) has included biotic interactions directly as candidate predictors, other research (Wisz et al. 2013) has suggested that surrogate variables can be used as proxies to capture spatial turnover or gradients in the distribution of biotic interactions across the landscape. Although our models do not include biotic interactions as candidate predictors, the models include landscape characteristics like stream order and elevation and there is significant research that suggests salmonids in the study area generally follow an elevational gradient.

Conclusion

The purpose of this study was to use GIS-derived landscape variables as predictors of occurrence in statistical models for brook trout, rainbow trout, and brown trout. Based on the aforementioned sources of error inherent in this type of modeling, we did not expect model accuracies to exceed those that we obtained. Our validation results indicate that all species were modeled with acceptable error and that using landscape data in a logistic regression framework is an appropriate method for predicting wild trout distributions in our study area. The underrepresentation of headwater streams in more commonly used hydrography datasets is an important consideration that warrants close examination when forecasting headwater species distributions and range estimates. Research on impacts to fish assemblages along upstream–downstream gradients has projected that headwater streams will suffer the most from the effects of climate change (Buisson and Grenouillet 2009). Additionally, it appears that a relative watershed position metric (e.g., stream order) is an important surrogate variable (even when elevation is included) for biotic interactions across the landscape in areas where headwater species distributions are influenced by topographical gradients.

Acknowledgments

This research was funded by Federal Aid in Sport Fish Restoration Funds. We would like to thank Dr. Montserrat Fuentes for providing statistical guidance and feedback and Adam Terando for comments on an early manuscript draft. We would also like to thank Doug Besler and Mallory Martin from the North Carolina Wildlife Resources Commission for providing valuable guidance and insight into the trout fisheries of western North Carolina. The

North Carolina Cooperative Fish and Wildlife Research Unit is jointly supported by North Carolina State University, North Carolina Wildlife Resources Commission, U.S. Geological Survey, U.S. Fish and Wildlife Service, and Wildlife Management Institute. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Conflict of Interest

None declared.

References

Akaike, H. 1974. New look at statistical-model identification. IEEE Trans. Autom. Control 19:716–723.

Argent, D., and W. Kimmel. 2013. Potential impacts of climate change on brook trout (*Salvelinus fontinalis*) populations in streams draining the Laurel Hill in Pennsylvania. J. Freshw. Ecol. 28:489–502.

Brenden, T. O., R. D. Clark, A. R. Cooper, P. W. Seelbach, L. Wang, S. S. Aichele, et al. 2006. A GIS framework for collecting, managing and analyzing multiscale landscape variables across large regions for river conservation and management. Pp. 49–74 *in* R. M. Hughes, L. Wang, P. W. Seelbach, eds Landscape influences on stream habitats and biological assemblages. American Fisheries Society Symposium 48, Bethesda, Maryland.

Brewer, S. K., C. F. Rabeni, S. P. Sowa, and G. Annis. 2007. Natural landscape and stream segment attributes influencing the distribution and relative abundance of riverine smallmouth bass in Missouri. North Am. J. Fish. Manag. 27:326–341.

Buisson, L., and G. Grenouillet. 2009. Contrasted impacts of climate change on stream fish assemblages along an environmental gradient. Divers. Distrib. 15:613–626.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical-theoretical approach. Springer, New York, NY.

Clarke, A., R. Mac Nally, N. Bond, and P. Lake. 2008. Macroinvertebrate diversity in headwater streams: a review. Freshw. Biol. 53:1707–1721.

Colson, T. P. 2006. Stream network delineation from high-resolution digital elevation models. Doctoral dissertation. North Carolina State University, Raleigh, North Carolina.

Dauwalter, D. C., and F. J. Rahel. 2008. Distribution modelling to guide stream fish conservation: an example using the mountain sucker in the Black Hills National Forest, USA. Aquat. Conserv. Mar. Freshw. Ecosys. 18:1263–1276.

Dewald, L., and M. A. Wilzbach. 1992. Interactions between Native Brook Trout and Hatchery Brown Trout - Effects on Habitat Use, Feeding, and Growth. Trans. Am. Fish. Soc. 121:287–296.

Dudgeon, D., A. Arthington, M. Gessner, Z. Kawabata, D. Knowler, C. Leveque, et al. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biol. Rev. 81:163–182.

ESRI. 2009. Arc Hydro Tools version 1.3 Final for ArcGIS 9.2.

Fausch, K. D., and R. J. White. 1981. Competition between Brook Trout (*Salvelinus fontinalis*) and Brown Trout (*Salmo trutta*) for positions in a michigan stream. Can. J. Fish Aquat. Sci. 38:1220–1227.

Fausch, K. D., Rieman B. E., Dunham J. B., Young M. K., and D. P. Peterson. 2009. Invasion versus isolation: trade-offs in managing native salmonids with barriers to upstream movement. Conservation Biology 23:859–870.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24:38–49.

Filipe, A., D. Markovic, F. Pletterbauer, C. Tisseuil, A. De Wever, S. Schmutz, et al. 2013. Forecasting fish distribution along stream networks: brown trout (*Salmo trutta*) in Europe. Divers. Distrib. 19:1059–1071.

Flebbe, P. A. 1994. A regional view of the margin: salmonid abundance and distribution in the Southern Appalachian Mountains of North Carolina and Virginia. Trans. Am. Fish. Soc. 123:657–667.

Flebbe, P. A., L. D. Roghair, and J. L. Bruggink. 2006. Spatial modeling to project Southern Appalachian trout distribution in a warmer climate. Trans. Am. Fish. Soc. 135:1371–1382.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification - viewing streams in a watershed context. Environ. Manage. 10:199–214.

Haak, A. L., and J. E. Williams. 2012. Spreading the risk: native trout management in a warmer and less-certain future. North Am. J. Fish. Manag. 32:387–401.

Haak, A. L., J. E. Williams, H. M. Neville, D. C. Dauwalter, and W. T. Colyer. 2010. Conserving peripheral trout populations: the values and risks of life on the edge. Fisheries 35:530–549.

Habera, J., S. Moore, and S. D. A. F. S. Trout. 2005. Managing southern Appalachian brook trout: a position statement. Fisheries 30:10–20.

Hardie, D., and J. Hutchings. 2010. Evolutionary ecology at the extremes of species' ranges. Environ. Rev. 18:1–20.

Harrell, F. E. 1997. Predicting outcomes: applied survival analysis and logistic regression. University of Virginia School Medicine, Charlottesville, Virginia.

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

Hosmer, D., and S. Lemeshow. 1989. Applied logistic regression. Wiley & Sons, New York, NY.

Huberty, C. J. 1994. Applied discriminant analysis. Wiley Interscience, New York, NY.

Isaak, D. J., C. H. Luce, B. E. Rieman, D. E. Nagel, E. E. Peterson, D. L. Horan, et al. 2010. Effects of climate change and wildfire on stream temperatures and salmonid thermal habitat in a mountain river network. Ecol. Appl. 20:1350–1371.

Jimenez-Valverde, A., and J. Lobo. 2007. Threshold criteria for conversion of probability of species presence to either-or presence-absence. Acta Oecol. Int. J. Ecol. 31:361–369.

Kearney, M., B. Wintle, and W. Porter. 2010. Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. Conserv. Lett. 3:203–213.

Kelly, G. A., J. S. Griffith, and R. D. Jones 1980. Changes in the distribution of trout in Great Smoky Mountains National Park, 1900–1977. US Fish and Wildlife Service Technical Paper 102.

King, W. 1937. Notes on the distribution of native speckled and rainbow trout in the streams of Great Smoky Mountains National Park. J. Tenn. Acad. Sci. 12:351–361.

Knight, A. T., R. M. Cowling, H. P. Possingham, and K. A. Wilson. 2009. From theory to practice: designing and situating spatial prioritization approaches to better implement conservation action. Pp. 249–259 *in* A. Moilanen, K. A. Wilson, H. P. Possingham, eds. Spatial conservation prioritization: quantitative methods and computational tools. Oxford University Press Inc., New York, NY.

Kristensen, E., A. Baattrup-Pedersen, and H. Andersen. 2012. Prediction of stream fish assemblages from land use characteristics: implications for cost-effective design of monitoring programmes. Environ. Monit. Assess. 184:1435–1448.

Lammert, M., and J. Allan. 1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environ. Manage. 23:257–270.

Larson, G., and S. Moore. 1985. Encroachment of exotic rainbow trout into stream populations of native brook trout in the Southern Appalachian Mountains. Trans. Am. Fish. Soc. 114:195–203.

Lesica, P., and F. W. Allendorf. 1995. When are peripheral populations valuable for conservation? Conserv. Biol. 9:753–760.

Lowe, W., and G. Likens. 2005. Moving headwater streams to the head of the class. Bioscience 55:196–197.

Maccrimmon, H. R., and J. S. Campbell. 1969. World distribution of brook trout, *Salvelinus fontinalis*. J. Fish. Res. Board Can. 26:1699.

Maloney, K., D. Weller, D. Michaelson, and P. Ciccotto. 2013. Species distribution models of freshwater stream fishes in Maryland and their implications for management. Environ. Model. Assess. 18:1–12.

Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. J. Appl. Ecol. 38:921–931.

McKenna, J., and J. Johnson. 2011. Landscape models of brook trout abundance and distribution in lotic habitat with field validation. North Am. J. Fish. Manag. 31:742–756.

McKenna, J., D. Carlson, and M. Payne-Wynne. 2013. Predicting locations of rare aquatic species' habitat with a combination of species-specific and assemblage-based models. Divers. Distrib. 19:503–517.

Meyer, J., J. Wallace, M. Press, N. Huntly, and S. Levin. 2001. Lost linkages and lotic ecology: rediscovering small streams. Ecol. Achieve. Challange. 295–317.

Meyer, J., D. Strayer, J. Wallace, S. Eggert, G. Helfman, and N. Leonard. 2007. The contribution of headwater streams to biodiversity in river networks. J. Am. Water Resour. Assoc. 43:86–103.

NCDENR. 2005. Implementation Plan to Improve the Mapping and Digital Representation of Surface Waters in North Carolina.

NCWRC. 2009. The Economic Impact of Mountain Trout Fishing in North Carolina.

Nielsen, J. L. 1999. The evolutionary history of steelhead (*Oncorhynchus mykiss*) along the US Pacific Coast: developing a conservation strategy using genetic diversity. ICES J. Mar. Sci. 56:449–458.

Nielsen, J. L., J. M. Scott, and J. L. Aycrigg. 2001. Endangered species and peripheral populations: cause for conservation. Endange. Species Update 18:194–197.

Nielsen, C., P. Hartvig, and J. Kollmann. 2008. Predicting the distribution of the invasive alien *Heracleum mantegazzianum* at two different spatial scales. Divers. Distrib. 14:307–317.

Nyman, O. L. 1970. Ecological interaction of brown trout (*Salmo Trutta*) L., and brook trout (*Salvelinus Fontinalis*) (Mitchill), in a stream. Can. Field Nat. 84:343–350.

Oakes, R. M., K. B. Gido, J. A. Falke, J. D. Olden, and B. L. Brock. 2005. Modelling of stream fishes in the Great Plains, USA. Ecol. Freshw. Fish 14:361–374.

Piccolo, J. J. 2011. Challenges in the conservation, rehabilitation and recovery of native stream salmonid populations: beyond the 2010 Luarca symposium. Ecol. Freshw. Fish 20:346–351.

Rashleigh, B., R. Parmar, J. Johnston, and M. Barber. 2005. Predictive habitat models for the occurrence of stream fishes in the mid-Atlantic highlands. North Am. J. Fish. Manag. 25:1353–1366.

Roberts, S., R. Tankersley, and K. Orvis. 2009. Assessing the potential impacts to riparian ecosystems resulting from hemlock mortality in great Smoky Mountains national park. Environ. Manage. 44:335–345.

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

Sagarin, R. D., and S. D. Gaines. 2002. The 'abundant centre' distribution: to what extent is it a biogeographical rule? Ecol. Lett. 5:137–147.

Sagarin, R. D., S. D. Gaines, and B. Gaylord. 2006. Moving beyond assumptions to understand abundance distributions across the ranges of species. Trends Ecol. Evol. 21:524–530.

Sindt, A., M. Quist, and C. Pierce. 2012. Habitat associations of fish species of greatest conservation need at multiple spatial scales in wadeable iowa streams. North Am. J. Fish. Manag. 32:1046–1061.

Steen, P. J., D. R. Passino-Reader, and M. J. Wiley. 2006. Modeling brook trout presence and absence from landscape variables using four different analytical methods. Pp. 515–531 *in* R. M. Hughes, L. Wang, P. W. Seelbach, eds. Landscapes influences on stream habitats and biological assemblages. American Fisheries Society Symposium 48, Bethesda, Maryland.

Steen, P. J., T. G. Zorn, P. W. Seelbach, and J. S. Schaeffer. 2008. Classification tree models for predicting distributions of Michigan stream fish from landscape variables. Trans. Am. Fish. Soc. 137:976–996.

Stoneking, M., D. Wagner, and A. Hildebrand. 1981. Genetic evidence suggesting subspecific differences between northern and southern populations of brook trout (Salvelinus fontinalis). Copeia, 810–819.

Storey, R., S. Parkyn, M. Neale, T. Wilding, and G. Croker. 2011. Biodiversity values of small headwater streams in contrasting land uses in the Auckland region. NZ J. Mar. Freshwat. Res. 45:231–248.

Trumbo, B., K. Nislow, J. Stallings, M. Hudy, E. Smith, D. Kim, et al. 2014. Ranking site vulnerability to increasing temperatures in southern appalachian brook trout streams in virginia: an exposure-sensitivity approach. Trans. Am. Fish. Soc. 143:173–187.

Vaughan, I. P., and S. J. Ormerod. 2003. Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. Conserv. Biol. 17:1601–1611.

Wald, A. 1943. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Trans. Am. Math. Soc. 54:426–482.

Weaver, D., and T. Kwak. 2013. Assessing effects of stocked trout on nongame fish assemblages in southern appalachian mountain streams. Trans. Am. Fish. Soc. 142:1495–1507.

Wenger, S., D. Isaak, C. Luce, H. Neville, K. Fausch, J. Dunham, et al. 2011. Flow regime, temperature, and biotic interactions drive differential declines of trout species under climate change. Proc. Natl Acad. Sci. USA 108:14175–14180.

Whiteley, A., J. Coombs, M. Hudy, Z. Robinson, A. Colton, K. Nislow, et al. 2013. Fragmentation and patch size shape genetic structure of brook trout populations. Can. J. Fish Aquat. Sci. 70:678–688.

Wiley, M. J., S. L. Kohler, and P. W. Seelbach. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. Freshw. Biol. 37:133.

Wipfli, M., J. Richardson, and R. Naiman. 2007. Ecological linkages between headwaters and downstream ecosystems: transport of organic matter, invertebrates, and wood down headwater channels. J. Am. Water Resour. Assoc. 43:72–85.

Wisz, M., J. Pottier, W. Kissling, L. Pellissier, J. Lenoir, C. Damgaard, et al. 2013. The role of biotic interactions in shaping distributions and realized assemblages of species: implications for species distribution modelling. Biol. Rev. 88:15–30.

Supporting Information

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

Appendix S1. Summary of model selection.
Appendix S2. Additional management implications.
Appendix S3. Discussion of final models.
Table S1. List of landscape predictor variables used in the creation of the presence-absence models for predicting distributions of wild trout in western North Carolina.
Table S2. Summary of brook trout model.
Table S3. Summary of rainbow trout model.
Table S4. Summary of brown trout model.
Table S5. Prediction results for wild trout occurrence for all possible combinations ofallopatry and sympatry.
Figure S1. Comparison of NHD High Resolution flowlines and 6-acre drainage area stream network.
Figure S2. Example 1:50,000 scale predicted distribution map for brook, rainbow and brown trout.

