Predictive modeling and spatial mapping of fish distributions in small streams of the Canadian Rocky Mountain foothills

Richard J. McCleary and Marwan A. Hassan

Abstract: We developed an automated procedure for modeling spatial distribution of fish occurrence using logistic regression models and geographic information system (GIS) tools. Predictors were measured from a digital elevation model (DEM) and stream layers. We evaluated the accuracy of GIS measures of reach slope through a comparison with field measures. Resource selection function models were used to explain presence–absence of bull trout (*Salvelinus confluentus*), rainbow trout, (*Oncorhynchus mykiss*), nonnative brook trout (*Salvelinus fontinalis*), and all fishes. Our models were extrapolated based on low, medium, and high levels of probability to produce reach-scale maps across 12 000 km^2. We attempted to improve models by adding land-use variables; however, the terrain best suited to road building and harvest also contained the habitat selected by rainbow trout, whereas bull trout generally selected terrain too steep for land use. These confounding factors emphasize the need for process-based investigations in addition to correlative approaches to identify habitat requirements. This automated method provides a rapid evaluation of fish habitat across remote areas useful for salmonid conservation and research planning.

Résumé : Nous avons mis au point une procédure automatisée pour modéliser la répartition spatiale de l'occurrence des poissons à l'aide de modèles de régression logistique et d'outils du système d'information géographique (GIS). Les valeurs prédictives ont été mesurées d'après un modèle digital d'altitude (DEM) et la couche des cours d'eau. Nous avons évalué la précision des mesures GIS des pentes des sections par comparaison avec des mesures de terrain. Des modèles de fonction de sélection des ressources ont servi à expliquer la présence–absence de l'ombre à tête plate (*Salvelinus confluentus*), de la truite arc-en-ciel (*Oncorhynchus mykiss*), de l'omble de fontaine (*Salvelinus fontinalis*) non indigène et de l'ensemble des poissons. Nous avons extrapolé nos modèles à des niveaux bas, moyens et élevés de probabilité pour produire des cartes à l'échelle des sections sur 12 000 km^2. Nous avons essayé d'améliorer les modèles en ajoutant des variables d'utilisation des terres; cependant, les surfaces les plus appropriées pour la construction de routes et pour les récoltes contiennent aussi les habitats sélectionnés par la truite arc-en-ciel, alors que l'ombre à tête plate choisit généralement des terrains trop escarpés pour utilisation humaine. Ces facteurs confondants soulignent la nécessité de faire des investigations reliées aux processus en plus d'utiliser les méthodes de corrélation pour identifier les besoins en habitats. Notre méthode automatisée fournit une évaluation rapide des habitats des poissons dans des régions éloignées qui est appropriée pour la conservation des salmonidés et elle permet une planification de la recherche.

[Traduit par la Rédaction]

Introduction

Quantifying fish distributions in streams and extrapolating these occurrence patterns beyond sampled areas can provide knowledge to support fish conservation planning and guide detailed experimental investigations (Rosenfeld 2003). Impediments to extrapolation of occurrence patterns include identifying key parameters that can be accurately measured with remote methods and addressing scaling issues related to the hierarchical nature of stream systems.

Approaches to modeling fish distributions differ between regions and land-use types. Differences in terrain limit model transferability. Models of fish occurrence used in headwater streams of the mountainous regions of western North America include descriptors of stream size, reach slope, and disruptions in stream connectivity created by obstructions to fish migration (Fransen et al. 2006). Geomorphic characteristics that dictate thermal regime also influence habitat selection by salmonids (Baxter and Hauer 2000). Fish habitat and distribution in the Great Plains of

Received 19 November 2006. Accepted 12 November 2007. Published on the NRC Research Press Web site at cjfas.nrc.ca on 30 January 2008.
J19668

R.J. McCleary[1] and M.A. Hassan.[2] Department of Geography, The University of British Columbia, Vancouver, BC V6T 1Z2, Canada.

[1]Present address: Foothills Model Forest, Box 6330, Hinton, AB T7V 1X6, Canada.
[2]Corresponding author (e-mail: mhassan@geog.ubc.ca).

 doi:10.1139/F07-161

central North America have been impacted by agriculture, reservoir construction, and urbanization, and as a result, models of fish distribution include a greater number of variables to describe environmental gradients and land use (Oakes et al. 2005).

Impacts to fish habitat from riparian logging are emphasized in some regions; however, these effects are not relevant in the Rocky Mountain foothills (foothills) where riparian buffer strips have been used since the onset of industrial forest harvest in 1956 (Bott et al. 2003). Likely impacts include sedimentation from logging roads (Spillios 1999), bank erosion, and habitat loss due to obstructions at culverts (Scrimgeour et al. 2003). However, relations between land use, habitat impacts, and fish occurrence have not been studied extensively in the continental rain and snow climate of the foothills.

Habitat studies have historically focused on large streams that provide water, power, and habitat for sport fish; less is known about small streams that represent most of the watercourses on the landscape (Moore and Richardson 2003). In comparison to large streams where fish presence can generally be inferred, only a portion of the small streams on the landscape provide habitat suitable for fish. In small streams, land use has a greater possibility of altering fish habitat features created by streamside vegetation and large woody debris (Hassan et al. 2005). Small streams can also be suitable for low-cost culverts at road crossings; however, where fish are present, these structures may obstruct fish passage (Wolter and Arlinghaus 2003). Fish presence at such locations should be ruled out by conducting electrofishing inventories during different years and seasons and by considering characteristics of the channel network. In the foothills, most small streams are inaccessible and located in a region of growing industrial activity. Environmental assessments for proposed projects ideally include electrofishing inventories. However, information on fish-bearing status may be required on short notice or outside of the summer season during which conditions are suitable for electrofishing. The sensitivities of small streams to a variety of land-use impacts and the increasing rate of industrial development warrants a remote-sensing approach to providing fish occurrence information.

The main objective of this research was to develop an automated method to predicting fish occurrence in small foothills streams. This knowledge is important to support planning of road locations and design of stream-crossing structures. Development of this approach is timely in Alberta because of the rapid expansion of the petroleum exploration and extraction industries. Fish occurrence information is also necessary when designing processed-based geomorphic and aquatic studies. The specific objectives of the research are (*i*) to map distributions for individual fish species and for all fishes regardless of species, (*ii*) to determine the most important variables for explaining occurrence, (*iii*) to compare distributions of native and nonnative fish, and (*iv*) to identify relationships between land use and occurrence of individual species.

We sought to address a number of common pitfalls when predicting distributions of organisms. First, the accuracy of variables measured with a geographic information system (GIS) is often not validated with field-measured data. Second, many researchers report a model's predictive accuracy; however, this measure is systematically affected by the frequency of occurrence of the target organism (Olden et al. 2002). Third, few models are validated with an independent data set (Olden et al. 2002). Fourth, it is difficult to design a sampling program to capture multiple environmental gradients when land-use variables are included. Land-use activities are not ubiquitous across the landscape (Van Sickle et al. 2004), rather their locations are based on criteria that may be similar to those of native fishes adapted to specific niches. In the foothills, harvesting has always been limited to terrain with less than 40% slope because of the limitations of logging equipment (Bott et al. 2003). Exclusionary distributions of land use and fish occurrence create a layer of complexity that confounds the interpretation of model outputs. Finding solutions to these four difficulties will advance the use of GIS-based resource selection function models.

Materials and methods

Study area

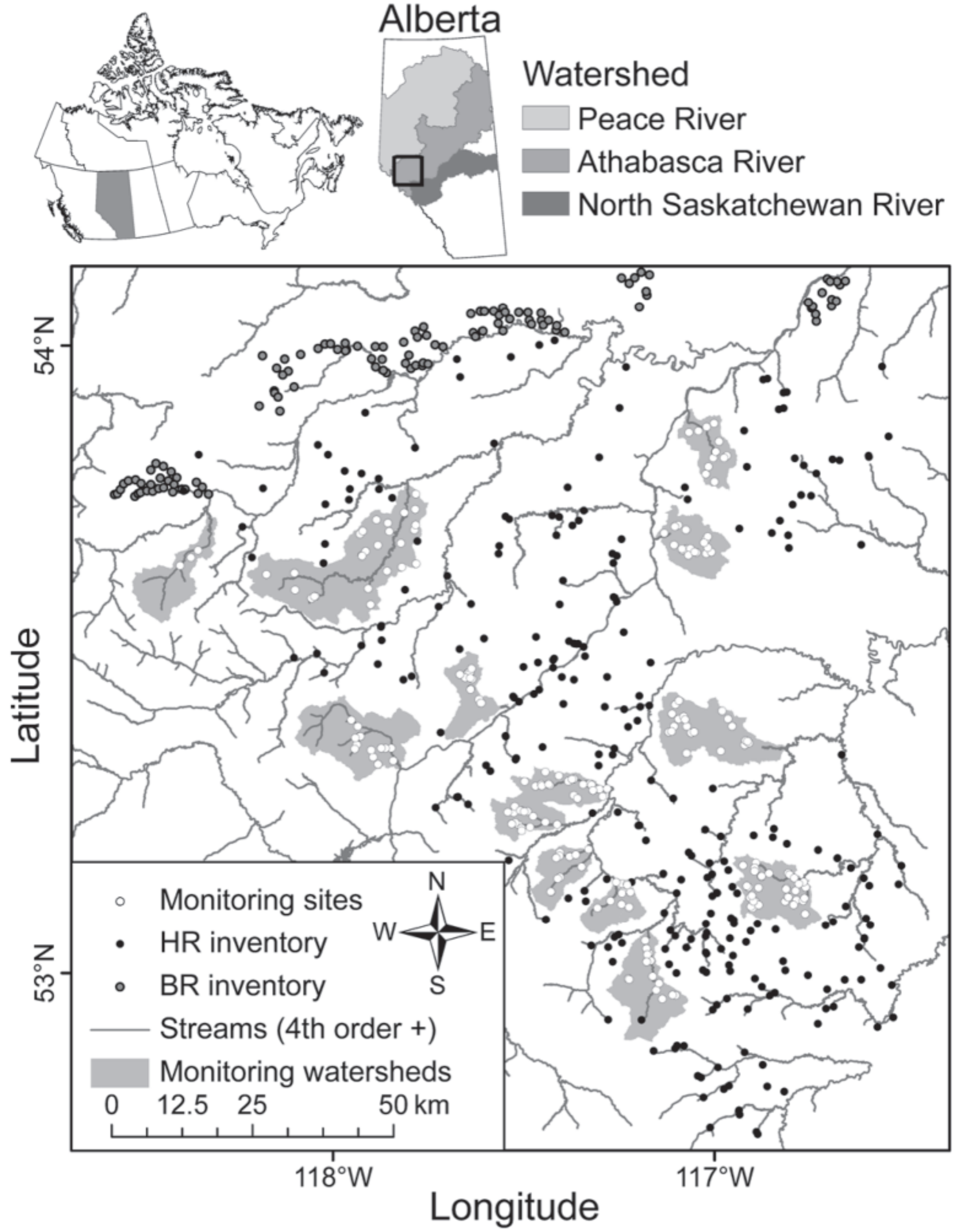
The study area covers a 12 000 km^2 portion of the foothills in west-central Alberta (Fig. 1). The foothills, underlain with sandstone and siltstone, represent a transition zone between the interior boreal plains to the east and the carbonate-dominated front ranges of the Rocky Mountains to the west. The combination of highly erodible bedrock and extensive glaciation has created a rolling topography. Elevations range from 1000 to 1600 m. Lodgepole pine (*Pinus contorta*) and white spruce (*Picea glauca*) forests cover uplands, and either black spruce (*Picea mariana*) – tamarack (*Larix laricina*) forests or shrubs dominate wet areas. Approximately 1% of the land base (120 km^2) is harvested annually, and the first harvest rotation will be completed by 2040.

Fish species

This study focused on bull trout (*Salvelinus confluentus*), rainbow trout (*Oncorhynchus mykiss*), and introduced brook trout (*Salvelinus fontinalis*). Bull trout are designated as a species of special concern in Alberta. The Athabasca River watershed (Fig. 1) is one of three basins east of the continental divide that support native rainbow trout (Taylor et al. 2006). Both nonnative and native rainbow trout inhabit the study area, but small streams provide important habitat for genetically pure native populations (Taylor et al. 2006). Brook trout were first introduced in Alberta in 1922, and widespread introductions in the 1960s included headwater lakes in the Athabasca River watershed (Donald et al. 1980). Bull trout have a very high sensitivity to angling, road-related sediment, road-related migration barriers, and increases in water temperature (Rieman and McIntyre 1993), whereas rainbow trout have a high sensitivity (Ford et al. 1995*b*) and brook trout have a moderate to high sensitivity to these impacts (Ford et al. 1995*a*). Brook trout present a threat to the long-term conservation of bull trout through hybridization (Leary et al. 1993) and juvenile competition (Selong et al. 2001). Knowledge of spatial distributions of these native and nonnative fishes may assist conservation efforts.

We used fish occurrence data from three different projects (Fig. 1): we collected information within the monitoring watersheds (MW) specifically for this study and obtained

Fig. 1. Map of the fish sampling locations by project, including monitoring watersheds (model training data set), Hinton region (HR) inventory (model testing data set), and Berland region (BR) inventory (model testing data set).



data from Hinton region (HR) and Berland River region (BR) inventories for the purposes of model validation. In the HR inventory, forest planners selected sites where fish presence–absence information was required to conserve fish passage at existing and proposed road–stream crossings. The BR inventory, completed in 1995 and 1997, was a reconnaissance survey of a largely pristine area scheduled for forest harvest.

The study area includes the subalpine, upper foothill, and lower foothill natural subregions (Beckingham et al. 1996). For detailed analysis, we selected 15 watersheds representing the range of biophysical and land-use characteristics for the study area (Fig. 1). For the MW and HR inventories, technicians used a backpack electrofisher (model 12A; Smith-Root Inc., Vancouver, Wash.) to sample 300 m long reaches. For the BR inventory, a Coffelt Mark 10 (Coffelt Manufacturing, Inc., Flagstaff, Arizona) was used, and reaches in first- and second-order streams were reduced to 150 m. Electrofishing was deferred when elevated flow or turbidity levels could have affected capture. A single-pass survey with no block nets was used to determine fish presence–absence. We randomly selected sample sites from the various stream-order and stream-slope combinations within each of the 15 monitoring watersheds, and these sites were sampled between 1999 and 2002. Additional fish presence–absence data within the 15 watersheds were obtained for a small number of sites that were monitored between 1996 and 2002.

An automated procedure to measure basin and reach characteristics

To map predicted probabilities of target fish species in the study area, we needed a system to extrapolate findings from specific sample points across the study area. This approach required a set of predictor variables that could be calculated through GIS analyses of existing data. An automated procedure to measure these variables included two stages. Before completing these stages, we prepared a single-line stream network and hydrologically corrected the digital elevation model (DEM) (Duhaime 2003).

In the first stage (reach-break creation), the stream network was divided into a series of distinct sections based on topologic (stream network) and topographic (slope) criteria. Reach breaks were placed at all confluences and topological features such as waterfalls and lake boundaries. We then used a two-step process for the creation of topologic breaks. In the first step, preliminary reach breaks were placed every 100 m in an upstream direction from each topologic reach break. In the second step, preliminary topologic reaches were amalgamated wherever the change in slope from a downstream segment to the next segment upstream failed to exceed a given tolerance. To achieve an average reach length of 300 m for all streams in the study area, the tolerances changed with downstream reach slope as follows: slope ≤ 1%, tolerance = 0.5%; 1% < slope ≤ 2%, tolerance = 1.0%; 2% < slope ≤ 4%, tolerance = 1.5%; 4% < slope ≤ 6%, tolerance = 2.0%; 6% < slope ≤ 10%, tolerance = 3.0%; 10% < slope ≤ 20%, tolerance = 5.0%; 20% < slope ≤ 40%, tolerance = 10%; slope ≤ 40%, tolerance = 20%. The process proceeded in an upstream direction from each stream confluence. Reach breaks were retained at topologic features and stream confluences.

The second stage consisted of the creation of watershed boundaries representing a drainage-area polygon originating at the downstream end of each reach. The polygons were created using a nested approach to ensure consistent boundaries between adjacent watersheds and to facilitate editing. The drainage-area boundary for each reach was saved to permit calculation of basin-scale variables for each reach.

Accuracy of stream slope measures generated from a DEM

To evaluate the accuracy of measures of stream slope derived from a 25 m DEM, we measured true stream slope with a total station (Lecia model TPS700; Lecia Geosystems AG, St. Gallen, Switzerland) and prism pole at 23 stream reaches. Surveys extended for a distance equivalent to 50–100 channel widths and points were captured at all bed-feature and meander transitions. We found a meandering channel pattern in most reaches that extended into headwater positions. Stream banks were well vegetated and root structures promoted meander development. For each survey, line length was the sum of the horizontal distances between successive points. Straight-line distance was calculated using the start and end points. Using GIS (ArcMap Version 9.2; ESRI Corporation, Redlands, California), we isolated the mapped stream segments that corresponded to our field-surveyed reaches. Then we calculated the change in line length by dividing the total length of the mapped stream reach by the total length of the surveyed stream reach (McMaster 1986). Stream reaches with a drainage area greater than 20 km^2 had small changes in line length, indicating reasonable discernment of the channel pattern during aerial photograph interpretation. Below this threshold, the change in line length averaged 78%, indicating greater generalization and poor representation of channel meanders. After applying this channel length correction factor to all reaches with a drainage area < 20 km^2 (corrected slope = slope × 0.78), there was no longer any evidence of differences between field-surveyed slope and GIS slope for the corresponding segment (paired t test, $P_{\text{before}} = 0.002$, $P_{\text{after}} = 0.954$) or between field-surveyed slope and GIS reach slope from the automated procedure ($P_{\text{before}} = 0.017$, $P_{\text{after}} = 0.645$).

Deriving predictive models of fish occurrence

The strategy that we used to explore relationships between fish presence–absence and habitat descriptors has been used in related studies (e.g., Ripley et al. 2005) and is consistent with the information–theoretic approach to model development and selection. This approach differs from traditional hypothesis testing in three ways (Burnham and Anderson 2002). First, the researchers identify a set of candidate models from patterns described in the literature and their observations. Second, the best model is identified using the Akaike's information criterion that considers both parsimony and the fit of each candidate model to the phenomenon or process of interest. Third, this approach does not use hypothesis tests or p values.

We calculated the small-sample version of Akaike's information criterion (AIC_c) and included a preliminary test for overdispersion (Burnham and Anderson 2002). Following the calculation of AIC_c, we estimated an evidence ratio (Δ_i) and Akaike weights (w_i) to help select the most parsimonious model and further provide rankings of model importance and uncertainty. The model with the lowest Δ_i is the best-fit model out of the candidate set of models, and w_i provides an approximate representation for the probability of a particular model fitting the data compared with the candidate set of models. To accomplish our objectives, two separate model-selection exercises were undertaken.

In the first exercise, the candidate models were limited to those containing predictor variables describing fish habitat at each stream reach. These variables, which we call ecological variables, represented key habitat dimensions including size, energy, and climate (Bozek and Hubert 1992) and were measured from readily available digital information sources. The best model developed from the ecological variables provided distribution information exclusive of land use and was well suited to extrapolation. The second set of models included combinations of the best ecological model and three land-use variables. The land-use variables were intended to represent management activities that could have affected the productive capacity of fish habitat in small streams.

We identified six ecological variables and three land-use variables (Table 1). All candidate ecological models included area because of its importance in most fish occurrence models (Rosenfeld 2003). The complete set of candidate ecological models included all combinations of area and noncorrelated variables. The set of candidate ecological – land-use models included the variables from the best ecolog-

Table 1. Reach-scale variables grouped by category.

Variable code	Predictor variable, measurement methods, and fish habitat considerations
Ecological parameters	
Area	Drainage area above downstream end of reach (km^2); common variable in fish occurrence modeling (Rosenfeld 2003)
Basin slope	Mean basin slope (%); calculated as the mean slope percentage from all DEM cells in the catchment of each reach; common variable in fish occurrence modeling; related to annual sediment yield in region (McPherson 1975)
Basin elevation	The average elevation (m) from all DEM cells in the catchment of each reach; ecosystem productivity decreases and annual precipitation increases with elevation within study area (Beckingham et al. 1996)
Percent wetlands	Percent of drainage area occupied by wetlands on forest inventory map; muskegs are common in region and typically form along lower side slopes and depressional areas (Dumanski et al. 1972); wetlands include black spruce (*Picea mariana*) – tamarack (*Larix laricina*) and shrub–herb communities; we observed low fish occurrence and organic substrate within streams that drain wetlands
Reach elevation	Elevation at downstream end of reach (m)
Reach slope	Common variable in fish occurrence modeling (Rosenfeld 2003)
Land-use parameters	
Harvest	Percent of basin harvested since 1956 measured from timber inventory maps; percent harvest was associated with a lower probability of bull trout occurrence within region (Ripley et al. 2005)
Road density	Length of permanent roads within each basin ($km \cdot km^{-2}$) measured on road maps; impact water quality and peak flow (MacDonald et al. 1990); related to decreased bull trout redd numbers (Baxter et al. 1999)
Downstream barrier	All reaches with a potential barrier located in downstream areas were coded yes (1) and others were coded no (0); culverts that retain an uninterrupted natural bottom conserve fish migration and all others may present a barrier depending on hang height, water velocity, fish species, and life stage (Jackson 2003); therefore, barriers included all culverts, except those with substrate retained within the culvert and without any hanging outfall; fish migration barriers at road crossings represent challenge for fish population conservation in some foothills streams of Alberta (Scrimgeour et al. 2003)

Note: Ecological variables were measured at all reaches across the study area. Land-use variables were measured for all reaches within the monitoring watersheds. Bull trout, *Salvelinus confluentus*.

ical model with all combinations of the three land-use variables.

Before model development, colinearity between all continuous ecological and land-use variables was assessed using Pearson correlation tests. To avoid colinearity, we excluded one of the variables when correlations were greater than |0.6|. We used a data-splitting type of cross validation (Olden et al. 2002), with 80% of the MW samples randomly selected for model training and the rest used for model testing. A nonlinear quadratic relationship was expected for ecological variables because of the adaptation of most organisms to a specific niche. Before model development, we used a two-step process to identify such patterns. For each predictor variable, we determined the average percent occurrence across 10 evenly sized classes. Next, we plotted relations for percent occurrence by predictor variable class to identify nonlinear relations. When these patterns were apparent, we included a quadratic term within model structures (i.e., basin slope + basin slope^2). Logistic regression was used to compare fish species present sites (1) with absent sites (0) based on ecological covariates of interest and candidate models tested.

Model interpretation and validation

Interpretation of regression coefficients was complicated by different measurement units among variables. To ensure that the value of regression coefficients corresponded to the magnitude of the effect on probability, all continuous predictors were scaled to standardized scores with mean value of 0 and variance of 1 (Klienbaum et al. 1988). Odds ratios (ratio of probability that event will occur to probability that it will not occur) and associated confidence intervals were also calculated. When a regression coefficient equals 0, its odds ratio equals 1. Therefore, when a confidence interval includes 1, it cannot be concluded that the variable is associated with a change in odds.

We assessed model performance based on receiver–operator characteristic (ROC) and cross-validation. For the best ecological models, we used four data sets (MW training and testing and HR and BR inventories). Land-use data were not available for the HR and BR inventories. The ROC evaluation included two components. First, we reported the area under the ROC curve where 0.5–0.7 represented low, 0.7–0.9 represented intermediate, and >0.9 represented high model accuracy (Manel et al. 2001). Second, we established high and medium probability cutoffs. Logistic regression is well suited to studies on rare organisms because the probability cutoff for establishing presence–absence varies with prevalence. The high probability cutoff corresponded to the point where the model sensitivity (probability of correctly predicting a positive outcome) and model specificity (probability of correctly predicting a negative outcome) were both maximized. We set the medium probability cutoff at one-half the value of the high probability cutoff. Probabilities less than the medium cutoff were assigned a low probability. Based on the high and medium cutoffs, we reported sensitivity (true positives correctly classified), specificity (true negatives correctly classified), and overall accuracy. The Canada

Fisheries Act provides protection of all fish habitat, including stream reaches with infrequent use; therefore correct identification of positive outcomes (sensitivity) is more important than correct identification of negative outcomes (specificity). For the sensitivity evaluation at both cutoffs, low, medium, and high performances were represented by values <70, 70–90, and >90, respectively.

We included Cohen's kappa as a fourth validation statistic, because unlike sensitivity, specificity, and overall accuracy, it provides an assessment of correct predictions outside of chance expectations (Manel et al. 2001). Low, medium, and high model performances were represented by kappa values <0.4, 0.4–0.6, and 0.6–0.8, respectively.

Extrapolating and mapping occurrence models

We used a simple procedure to extrapolate the best AIC models and produce distribution maps by species for the study area. First, we exported a table from the GIS with unique reach numbers and values for all ecological variables for the 40 000+ reaches within the study area. This table was appended to the table containing the model training data and the combined table was imported into the modeling program (SPSS version 10.0.5; SPSS Inc., Chicago, Illinois). We ran the best model for each species and used the function to save predicted probability values for each reach. The predicted probabilities for all reaches were linked back to the reach table in the GIS. Probability of occurrence was mapped using three categories (high, medium, and low) based on the ROC-determined cutoff value for each model.

Sources of error

Errors in fish occurrence can result when the minimum number of fish present in a reach is less than the minimum detectable population for the capture technique. Paul and Post (2001) used similar capture techniques (single-pass electrofishing, 300 m reach, and no block nets) in the foothills and determined a minimum detectable density of 8 individuals·(300 m)$^{-1}$. In recognition of this limitation for determining occurrence, we used three levels of probability of occurrence (low, medium, and high) rather than using a definitive yes–no cutoff.

Errors in the ecological variables (drainage area, basin slope, reach slope, and reach elevation) are related to the resolution of the digital stream layer and accuracy of the DEM. A digital elevation model is a grid with elevations that represent the earth's surface. Raw DEM data typically contain erroneous features that prevent the downhill movement of water across the surface. A drainage enforcement function uses GIS smoothing algorithms to correct these erroneous cells by filling depressions and cropping barriers. The use of a 10 m drainage-enforced DEM improves the accuracy of watershed boundary identification over a 30 m DEM lacking drainage enforcement (Clarke and Burnett 2003). We used a 25 m DEM with drainage enforced.

Land-use variables were measured from a variety of sources. Percent harvest was measured from digital maps of cutover areas. We confirmed this information using digital orthophotographs. Percent harvest was intended to provide a surrogate indicator for changes in the hydrologic regime. However, hydrologic changes are not permanent and recovery occurs with forest regrowth. Road density was calculated using existing digital road information that was confirmed using digital orthophotos. Sediment inputs at a road crossing are related to characteristics of the road surface and ditches (Brooks et al. 2006); however, such factors were not considered in our models. Our criteria for identifying potential migration barriers are consistent with policy recommendations for maintaining fish passage under the Canada Fisheries Act, and a portion of the potential barriers may be passable to certain species and life stages. Beaver dams, logjams, head cuts, and waterfalls are common natural migration barriers that are not mapped and thus were not included in our models. Despite these limitations, this information provides a valuable context for interpreting fish occurrence patterns.

Results

Sampling summary

Our study was limited to small streams with drainage areas of <20 km^2. Bankfull width for streams with drainage areas between 15 and 20 km^2 averaged 4.9 ± 1.5 m. The channel length correction factor (slope × 0.78) was applied to all GIS measures of reach slope for these streams. Based on this drainage area cutoff, sample sizes by project area were as follows: MW, 284; HR inventory, 542; BR inventory, 66. Prevalence of brook trout among these three project areas was as follows: MW, 7%; HR, 12%; BR, 0% (mean for brook trout present areas = 10%). Prevalence of bull trout was as follows: MW, 25%; HR, 8%; BR, 12% (mean = 15%). Prevalence of rainbow trout was as follows: MW, 50%; HR, 29%; BR, 32% (mean = 37%). Prevalence of one or more of any species was as follows: MW, 66%; HR, 43%; BR, 33% (mean = 47%). In addition to the three target species, 14 additional species were encountered as incidental catches and were also included in the "all species" model along with the three target species.

In the Pearson correlation test, mean basin slope was correlated with mean basin elevation and percent wetlands. The variables retained included area (A), mean basin slope + mean basin slope^2 (B), reach slope + reach slope^2 (S), and elevation (E). The eight models tested included A, A+B, A+S, A+E, A+B+S, A+B+E, A+S+E, and A+B+S+E.

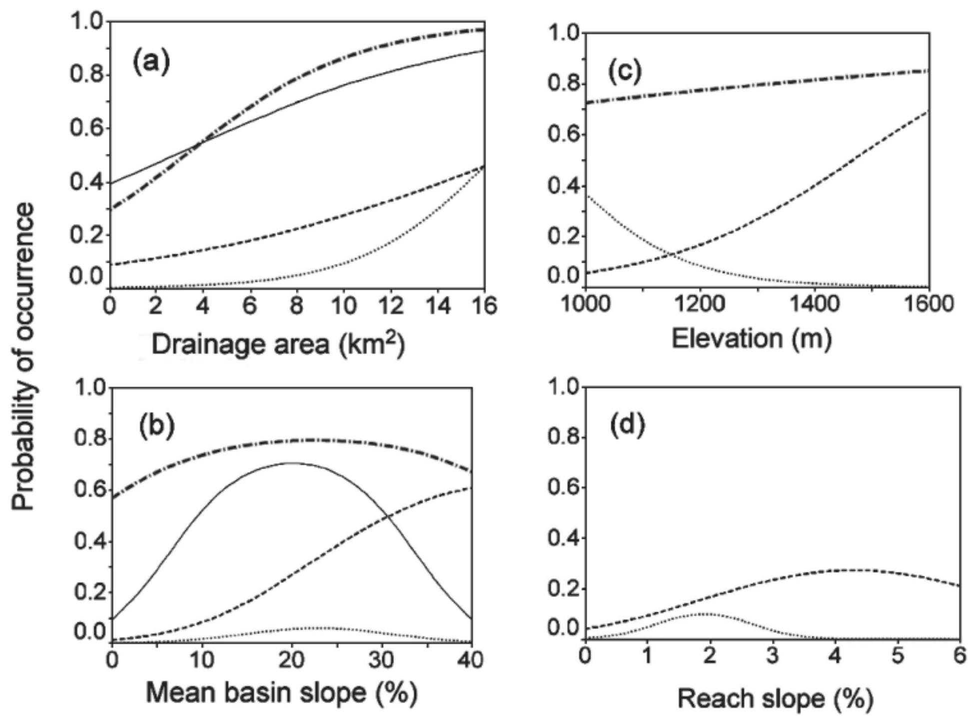
Three of the four independent ecological variables shared similar means and standard deviations among the three projects. Drainage area averaged 7.4 km^2 (MW, 8.0 ± 5.0 km^2; HR, 6.2 ± 5.1 km^2; BR, 7.9 ± 5.8 km^2), reach slope averaged 2.7% (MW, 2.8% ± 1.8%; HR, 3.2% ± 2.8%; BR, 2.2% ± 1.6%), and elevation averaged 1250 m above sea level, ASL (MW, 1270 ± 150 m ASL; HR, 1290 ± 190 m ASL; BR, 1190 ± 230 m ASL). Although basin slope had averaged 13.2% between the projects, the mean values were from a broader range (MW, 18.2% ± 11.4%; HR, 12.6% ± 7.6%; BR, 8.9% ± 8.5%). Land-use variables were only measured for the monitoring watersheds project. Percent harvest averaged 20.1% (0.0%–97.6%), and road density averaged 0.34 km·km^{-2} (0.0–2.54 km·km^{-2}). We inspected 302 crossings that intersected mapped streams, including those in intermittent watercourses. Average hang height of the culverts rated as potential barriers was 0.27 m. We identified downstream barriers at 25% of electrofished reaches.

Table 2. Logistic regression models of ecological variables predicting the probability of occurrence of brook trout (*Salvelinus fontinalis*), bull trout (*Salvelinus confluentus*), rainbow trout (*Oncorhynchus mykiss*), and all species present.

Species	Model name	K	AIC_c	Δ_i	w_i	$p\hat{C}$
Brook trout	Area + basin slope + elevation + reach slope	7	100.34	0.00	0.45	0.97
	Area + elevation + reach slope	5	100.78	0.44	0.36	0.63
	Area + basin slope + elevation	5	102.72	2.38	0.14	0.88
Bull trout	Area + basin slope + elevation + reach slope	7	192.54	0.00	0.62	0.32
	Area + basin slope + elevation	5	193.57	1.03	0.37	0.10
Rainbow trout	Area + basin slope	4	267.45	0.00	0.55	0.69
	Area + basin slope + elevation	5	269.43	1.98	0.20	0.76
	Area + basin slope + reach slope	6	269.66	2.21	0.18	0.02
All species	Area + basin slope	4	246.86	0.00	0.35	0.08
	Area + basin slope + elevation	5	247.84	0.98	0.21	0.21
	Area + basin slope + reach slope	5	248.91	2.06	0.13	0.10
	Area + elevation	3	249.27	2.41	0.10	0.02

Note: K, the number of model parameters; AIC_c, the small sample version of Akaike's information criteria (AIC); Δ_i, AIC difference between a given model and highest ranked model; w_i, Akaike model selection weights; and $p\hat{C}$, the Hosmer and Lemeshow goodness-of-fit chi-square.

Fig. 2. Predicted probability of occurrence by habitat variables for brook trout (*Salvelinus fontinalis*) (dotted line), bull trout (*Salvelinus confluentus*) (broken line), rainbow trout (*Oncorhynchus mykiss*) (solid line), and all fishes (dashed–dotted line): (*a*) drainage area, (*b*) mean basin slope, (*c*) elevation (m), and (*d*) reach slope. The range of the *x* axis represents the approximate observed range of each variable. For predictions, all values for other variables were standardized to mean values.



Best ecological models

Based on the AIC_c weights (w_i), global models for both brook trout and bull trout occurrence had the highest probability as the best model (Table 2). For the rainbow trout and "all species" models, there was less evidence for one best model. According to Burnham and Anderson (2002), an evidence ratio (Δ_i) of less than 2.0 between models does not show strong support for any one model and they advise on choosing based on ecological–management sense. For rainbow trout, we selected the model with the highest w_i as the best model. This model was the most parsimonious, and based on $p\hat{C}$, the model also fit the data well (Table 2). For all species, we selected the model with the second highest w_i (area + basin slope + elevation) as the best ecological model. There was some evidence based on $p\hat{C}$ that the model with the highest w_i did not fit the data, whereas there was no evidence that the second-ranked model did not fit the data (Table 2).

Under average conditions (i.e., basin slope and elevation), rainbow trout were more prevalent in smaller streams than bull trout and brook trout. Streams with a drainage area greater than 3 km^2 had a high probability of rainbow trout (cutoff = 0.51) (Fig. 2). The high probability threshold for both brook trout (cutoff = 0.08) and bull trout (cutoff = 0.25)

Table 3. Classification analysis for high and medium predicted occurrence probability categories for best ecological models by species (brook trout (*Salvelinus fontinalis*), bull trout (*Salvelinus confluentus*), rainbow trout (*Oncorhynchus mykiss*), and all species).

Species	Cutoff	Data set	*n*	Prevalence	AUC	PCC (1)	PCC (0)	PCC	Kappa
Brook trout	High (0.08)	Training	238	9	0.92	81	82	82	0.36
		HR	540	13	0.63	29	88	80	0.16
		Mean			0.77	55	85	81	0.26
	Medium (0.04)	Training	238	9		95	70	72	0.27
		HR	540	13		33	82	75	0.12
		Mean				64	76	74	0.19
Bull trout	High (0.27)	Training	238	25	0.88	77	79	78	0.49
		Testing	46	22	0.69	60	72	70	0.27
		HR	540	8	0.68	40	85	81	0.16
		BR	66	12	0.96	63	97	92	0.62
		Mean			0.80	60	83	80	0.39
	Medium (0.13)	Training	238	25		92	60	68	0.38
		Testing	46	22		80	58	63	0.26
		HR	540	8		53	71	70	0.11
		BR	66	12		75	93	91	0.62
		Mean				75	71	73	0.34
Rainbow trout	High (0.51)	Training	238	50	0.79	71	73	72	0.45
		Testing	46	52	0.78	67	73	70	0.39
		HR	540	29	0.62	47	77	68	0.24
		BR	66	32	0.81	57	87	77	0.46
		Mean			0.75	61	77	72	0.38
	Medium (0.25)	Training	238	50		93	39	66	0.32
		Testing	46	52		92	32	63	0.24
		HR	540	29		86	24	42	0.07
		BR	66	32		90	51	64	0.33
		Mean				90	36	59	0.24
All species	High (0.63)	Training	238	68	0.81	76	74	76	0.48
		Testing	46	54	0.74	76	57	67	0.34
		HR	540	43	0.69	58	75	68	0.34
		BR	66	33	0.77	82	68	73	0.45
		Mean			0.75	73	69	71	0.40
	Medium (0.31)	Training	238	68		99	17	73	0.21
		Testing	46	54		100	14	61	0.15
		HR	540	43		92	15	49	0.07
		BR	66	33		91	41	58	0.25
		Mean				96	22	60	0.17

Note: High cutoff, the threshold of optimal sensitivity and specificity from ROC curve; medium cutoff, 0.5 × the high cutoff value; AUC, area under ROC curve; PCC (1), model sensitivity defined as percent of fish occurrence sites correctly classified; PCC (0), model specificity defined as percent of no fish occurrence sites correctly classified; PCC, overall prediction success.

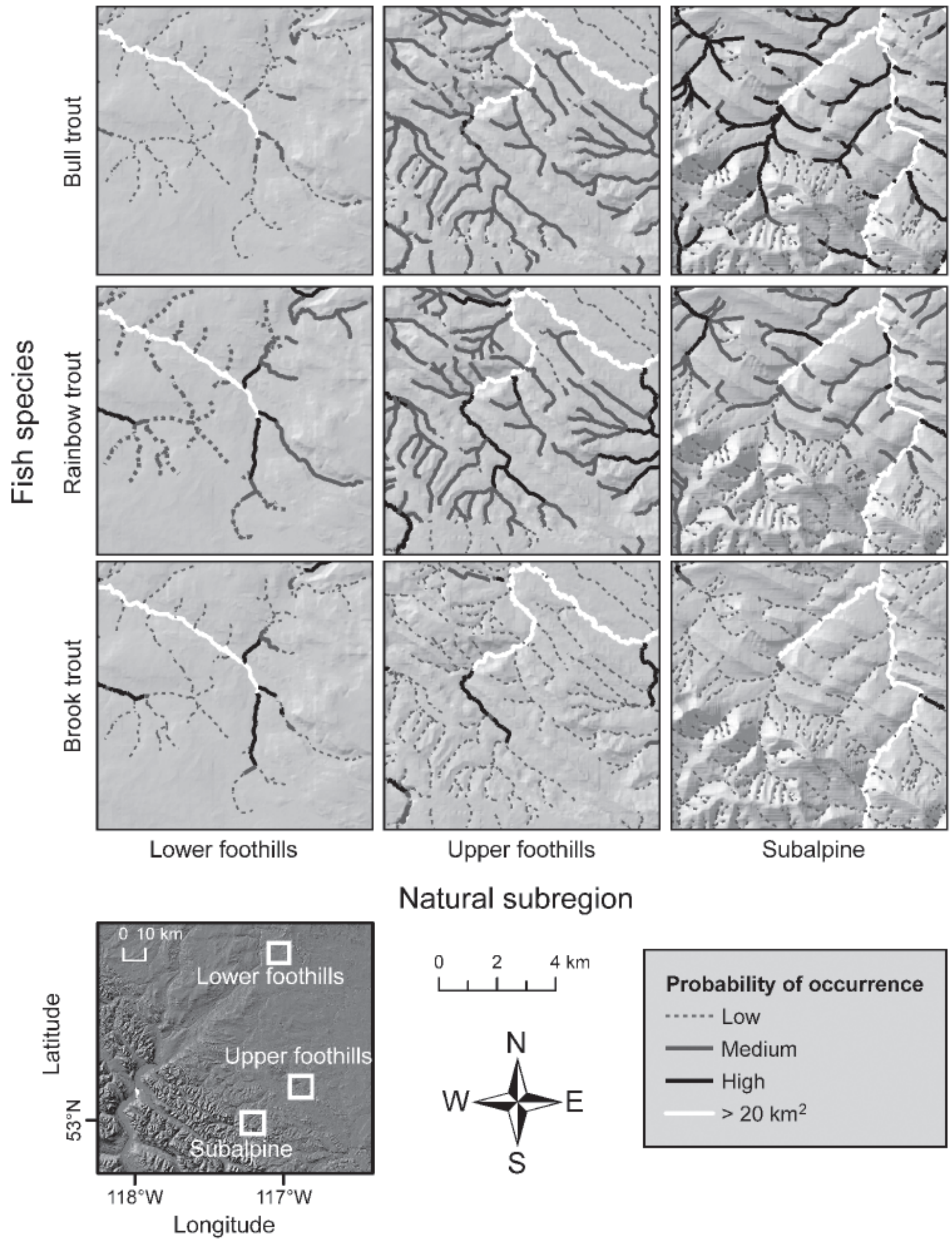
was near 10 km^2. Different ecological niches among the three species were apparent from the predicted probabilities for basin slope and reach slope (Fig. 2). Maximum probability for brook trout, bull trout, and rainbow trout occurred in basins with average slopes of 16%–28%, 30%–40%, and 12%–28%, respectively. Bull trout probability increased with elevation, whereas brook trout displayed the opposite trend.

Model performance varied among species and data sets (Table 3). Based on average area under the ROC curve, all models had moderate performance at high cutoff values. Model sensitivity improved with lower cutoff values for all models. For rainbow trout and all species, sensitivity improved to high with reduced cutoff values. The use of the lower cutoff value resulted in reduced specificity, overall accuracy, and kappa. Based on kappa values, all models performed moderately at the high cutoff value, except for the brook trout model, which had a low performance rating. Performance of all models (based on area under ROC curve, sensitivity, and kappa) decreased with the HR inventory. Sixty percent of the incorrectly classified bull trout sites were located within 2 km of a stream with a drainage area >20 km^2.

Predicted spatial distributions

The different niches for the three target species were also apparent from fish species distribution maps (Fig. 3). Brook trout had the greatest predicted extent within the streams originating in the lower foothills, whereas rainbow trout and bull trout had the greatest predicted extent within streams originating in the upper foothills and in the subalpine natural subregion, respectively.

Fig. 3. Maps of predicted occurrence probability for bull trout (*Salvelinus confluentus*), rainbow trout (*Oncorhynchus mykiss*), and brook trout (*Salvelinus fontinalis*) within representative areas of the lower foothills, upper foothills, and subalpine natural subregions.



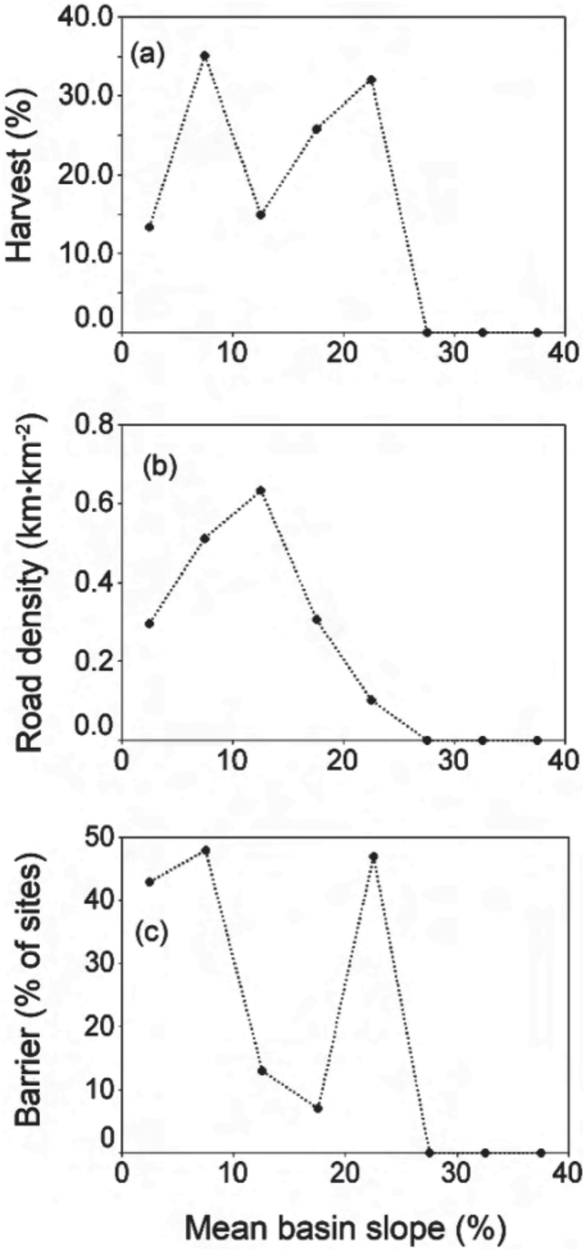
Best ecological – land-use models

Underlying relationships between land-use variables and basin slope present a confounding factor to consider when interpreting correlations between land-use variables and fish occurrence (Fig. 4). All three land-use activities including harvest (H), road density (R), and downstream barrier (D) were limited to reaches with basin slopes <25%. The eight models tested included the best ecological model (E), E+H, E+R, E+D, E+H+R, E+H+D, E+R+D, and E+H+R+D.

The five top-ranked brook trout models shared a $\Delta_i < 2.0$, indicating little support for one best model (Table 4). In the highest ranked model, area had the greatest odds ratio, followed by harvest. The regression parameter values indicate a greater probability of capturing brook trout as harvest increases, but a lower probability as roads increase (Table 5). Percent harvest also had a low standard error. Road density had a high standard error and also included the value 1 within the odds ratio confidence interval to indicate "no effect" is within the interval. The kappa value indicates medium performance of the best ecological – land-use model, whereas the ecological model had a lower rating (Table 6). All other performance measures indicate similar trends.

The five top-ranked bull trout models shared a $\Delta_i < 2.0$, indicating little support for one best model (Table 4). The highest ranked model included downstream barriers as the only land-use variable. The ranking from largest odds ratio in this model was basin slope, elevation, and then drainage area (Table 5). One parameter in each quadratic term (i.e.,

Fig. 4. Correlations between mean basin slope and land-use variables using bin-averaged values of each land-use variable with bins distinguished by a 5% step in mean basin slope: (*a*) percent of basin harvested, (*b*) road density, and (*c*) percent of sites with downstream barrier present.



basin slope^2 and reach slope) had high standard errors and included 1 within the confidence interval. The presence of a downstream barrier was associated with a reduction in odds by a factor of 0.274, and the odds ratio confidence interval did not include 1. The performance evaluation of the highest ranked model indicated improved performance over the ecological model (Table 6).

There was little evidence ($\Delta_i < 2.0$) for one best rainbow trout model (Table 4). Based on $p\hat{C}$, the second-ranked model, which included harvest, fit the data well, whereas there was evidence that the model with the highest w_i did not. Therefore we selected the model that included harvest as the only land-use variable as the best model. Harvest was associated with a negligible increase in odds, and the confidence interval did include 1 to indicate the potential of no effect (Table 5). Overall model performance decreased with the addition of the only plausible land-use variable into the best ecological model for rainbow trout (Table 6).

Discussion

Occurrence patterns of nonnative and native salmonids

Brook trout had higher probability of occurrence in larger, low-elevation channels. These findings were similar to those of another study in the Rocky Mountains where brook trout tended to persist in the main channels at lower elevations after dispersing from upstream stocking locations (Paul and Post 2001). In our study, brook trout occurrence increased in basins with higher harvest levels, and although other researchers have attributed these patterns to habitat degradation resulting from land-use activities, such relationships cannot be inferred within this study area. Timber harvest within riparian areas can cause habitat changes, including increased water temperature (MacDonald et al. 1990), and these changes may favor brook trout over bull trout (Selong et al. 2001) and native cutthroat trout (*Oncorhynchus clarki clarki*) (Shepard 2004). However, timber-harvesting practices in the study area have always included retention of streamside buffers. Furthermore, land-use activities were most prevalent in the intermediate relief basins that brook trout commonly inhabited and largely absent in higher-relief areas where brook trout were less frequently encountered. Alternate explanations include brook trout stocking at stream crossings rather than habitat modification.

The best ecological model for bull trout performed moderately. A logistic regression model for bull trout occurrence from a nearby foothills watershed had a similar ROC value (0.80); however, it included field measures (stream width and percent fines) combined with GIS measures of stream slope and harvest (Ripley et al. 2005). They found increased bull trout occurrence with increased stream width and reduced occurrence as percent fines, stream slope, and harvest levels increased. In contrast, we found a nonlinear relationship with stream slope and little evidence for including percent harvest in our best overall model. Our model also indicated a nonlinear relationship with basin slope tending towards steeper terrain and an increase in occurrence with increased elevation. Our best model included the downstream barrier as a land-use variable, and this contributed to an improvement in the ROC score from 0.78 to 0.81. All of the sample reaches with barriers had basin slopes <25%, which is below the optimal relief for bull trout. Although there may be some potential for a confounding effect from the exclusive occurrences of bull trout and this land-use variable, we did capture bull trout immediately below an impassible railway crossing in Pinto Creek, but not in any of the upstream tributaries. Similar patterns were observed in Japan where extinction rates of white-spotted char (*Salvelinus leucomaenis*) upstream from dams increased at a greater rate in headwater streams than in downstream areas (Morita and Yamamoto 2002).

Table 4. Logistic regression models of ecological and land-use variables predicting the probability of occurrence of brook trout (*Salvelinus fontinalis*), bull trout (*Salvelinus confluentus*), and rainbow trout (*Oncorhynchus mykiss*).

Species	Model name	K	AIC_c	Δ_i	w_i	$p\hat{C}$
Brook trout	Ecological, harvest, roads	9	93.29	0.00	0.37	0.18
	Ecological, harvest	8	93.51	0.22	0.33	0.51
	Ecological, harvest, barrier	9	95.37	2.08	0.13	0.79
	Global	10	95.45	2.16	0.13	0.19
Bull trout	Ecological, barrier	8	188.57	0.00	0.26	0.82
	Ecological, harvest, barrier	9	189.20	0.63	0.19	0.91
	Ecological, roads, barrier	9	189.79	1.22	0.14	0.99
	Ecological, harvest	8	189.89	1.32	0.14	0.21
	Global	10	190.36	1.79	0.11	0.93
	Ecological, harvest, roads	9	190.74	2.17	0.09	0.38
Rainbow trout	Ecological, harvest, roads	6	265.85	0.00	0.25	0.00
	Ecological, harvest	5	266.29	0.43	0.20	0.56
	Ecological, roads	5	266.78	0.93	0.16	0.34
	Ecological	4	267.45	1.60	0.11	0.69
	Global	7	267.97	2.12	0.09	0.04
	Ecological, roads, barrier	6	268.28	2.43	0.07	0.37
	Ecological, harvest, barrier	6	268.34	2.48	0.07	0.36

Note: K, the number of model parameters; AIC_c, the small sample version of Akaike's information criteria (AIC); Δ_i, AIC difference between a given model and highest ranked model; w_i, Akaike model selection weights; and $p\hat{C}$, the Hosmer and Lemeshow goodness-of-fit chi-square.

Table 5. Logistic regression parameter estimates from the best ecological – land-use models for each of three fish species (brook trout (*Salvelinus fontinalis*), bull trout (*Salvelinus confluentus*), rainbow trout (*Oncorhynchus mykiss*)) and all species present, including standard error of the model coefficient (SE), odds ratio, and associated lower and upper 95% confidence intervals (CIs).

					95% CIs	
Model	Parameter	Estimate	SE	Odds ratio	Lower	Upper
Brook trout	Area	1.763	0.421	5.827	2.554	13.294
	Basin slope	0.606	0.807	1.833	0.377	8.920
	Basin slope^2	–1.234	0.813	0.291	0.059	1.434
	Reach slope	–2.522	1.712	0.080	0.003	2.301
	Reach slope^2	–2.924	1.533	0.054	0.003	1.083
	Elevation	–2.088	0.713	0.124	0.031	0.501
	Percent harvest	1.419	0.501	4.132	1.548	11.031
	Road density	–0.949	0.707	0.387	0.097	1.547
	Constant	–3.557	0.782	0.029		
Bull trout	Area	0.703	0.260	2.019	1.212	3.363
	Basin slope	1.442	0.474	4.230	1.672	10.701
	Basin slope^2	–0.326	0.186	0.722	0.502	1.039
	Reach slope	0.647	0.380	1.911	0.907	4.025
	Reach slope^2	–0.385	0.182	0.680	0.476	0.972
	Elevation	0.908	0.296	2.480	1.388	4.432
	Barrier downstream	–1.295	0.570	0.274	0.090	0.836
	Constant	–0.906	0.309	0.404		
Rainbow trout	Area	0.751	0.173	2.120	1.511	2.974
	Basin slope	0.434	0.228	1.544	0.988	2.413
	Basin slope^2	–0.989	0.192	0.372	0.255	0.542
	Percent harvest	0.011	0.006	1.011	0.999	1.023
	Constant	0.621	0.241	1.861		

Note: To assist in interpretation of regression coefficients, all independent variables, except the categorical variable "barrier downstream", were scaled to standardized scores.

Table 6. Classification analysis for high occurrence category for best ecological – land-use models by species (brook trout (*Salvelinus fontinalis*), bull trout (*Salvelinus confluentus*), rainbow trout (*Oncorhynchus mykiss*)).

Species	Model	Cutoff	Data set	*n*	Prevalence	AUC	PCC (1)	PCC (0)	PCC	Kappa
Brook trout	Ecological	0.08	Training	238	9	0.922	81	82	82	0.359
	Ecological, harvest, roads	0.09	Training	238	9	0.94	86	86	86	0.455
Bull trout	Ecological	0.27	Training	238	25	0.883	77	79	78	0.488
			Testing	46	22	0.685	60	72	70	0.265
			Mean			0.784	68	75	74	0.377
	Ecological, barriers	0.31	Training	238	25	0.889	82	84	83	0.595
			Testing	46	22	0.732	60	69	67	0.235
			Mean			0.811	71	77	75	0.415
Rainbow trout	Ecological	0.51	Training	238	50	0.793	71	73	72	0.445
			Testing	46	52	0.777	67	73	70	0.392
			Mean			0.785	69	73	71	0.419
	Ecological, harvest	0.51	Training	238	50	0.795	71	72	72	0.437
			Testing	46	52	0.714	67	59	63	0.258
			Mean			0.756	69	66	67	0.348

Note: Cutoff, the threshold of optimal sensitivity and specificity from ROC curve; AUC, area under ROC curve; PCC (1), model sensitivity defined as percent of fish occurrence sites correctly classified; PCC (0), model specificity defined as percent of no fish occurrence sites correctly classified; PCC, overall prediction success.

The high odds ratio for basin slope illustrates a strong relation with bull trout occurrence. Basin slope was positively correlated with mean basin elevation, a climatic descriptor, and negatively correlated with percent wetlands, a terrain indicator. This variable has not been included as a candidate variable in other watershed-scale models of bull trout distribution (Paul and Post 2001; Ripley et al. 2005), but a similar variable was incorporated into other watershed-scale models of fish occurrence (Porter et al. 2000).

The use of reach elevation as a surrogate for climate proved successful in a bull trout model within the Rocky Mountains (Paul and Post 2001). Improvements to the climatic variable could capture the variations in water temperature due to groundwater inputs in intermediate relief areas. For example, mean summer temperature (June–September) in a study area stream that supports juvenile bull trout averaged 5.4 °C, whereas the mean temperature at a site in a similar adjacent watershed that does not typically support juvenile bull trout averaged 8.1 °C. These differences were attributed to different groundwater inputs between basins (Jablonski 1980). A number of reaches shared medium or high probability of bull trout and brook trout, including an important bull trout spawning stream. Knowledge of such locations where the two species overlap is important for identifying where threats to bull trout persistence occur (Rieman et al. 2006). Therefore, future efforts to model bull trout occurrence should include the locations of groundwater upwelling areas, especially within lower-elevation streams that may otherwise provide marginal bull trout habitat.

The differences between our findings and those of another correlative investigation into bull trout occurrence in the foothills (Ripley et al. 2005) emphasize the need for process-based approaches that consider the confounding relationship between land use and basin relief. Bull trout are vulnerable to a range of human impacts including illegal angler harvest, disruption of migration at stream crossings and dams, loss of habitat productivity due to sedimentation, and increases in water temperatures. To support effective conservation efforts, future investigations should endeavor to rank these impacts.

In comparison to bull trout, rainbow trout occurred more frequently in reaches that also supported brook trout, and rainbow trout were more common in watersheds that supported high levels of land use. Rainbow trout tended towards small streams that were located in intermediate-relief terrain ideal for timber production and were less prevalent in lower-relief areas that support extensive muskegs. We observed that muskeg streams had sinuous channels, but these low-gradient reaches often lacked important rainbow trout habitat attributes including gravel substrate and pools. Given these observations, we were surprised that reach slope was not indicated within the best rainbow trout model. This absence of reach slope from the model has two divergent explanations. One previous study found no indication that trout populations were negatively affected by increases in stream slope (Isaak and Hubert 2000). However, two studies found that channel gradient constrained the upstream extent of coastal cutthroat and rainbow trout (Latterell et al. 2003; Fransen et al. 2006). We suggest the use of higher-resolution DEM data in future studies that explore the relationships between rainbow trout occurrence and stream slope. There are also important considerations when interpreting the suggested positive relationship between rainbow trout occurrence and land use. Our findings were similar to those for winter steelhead (*Oncorhynchus mykiss*) in Oregon where the proportion of watershed converted to shrub and young forest was positively related to winter steelhead redd abundance (Steel et al. 2004). This relation may be confounded by a strong affinity that this fish species may have for streams located within terrain well suited for land use. Also consider that contrary to our findings, one long-term study within our study area found that rainbow trout fry densities decreased after harvest when water yield exceeded a critical flow during the incubation period (Sterling 1992). Additional processed-based investigations are required to identify land-use activities that may negatively decrease survival and growth rates of native rainbow trout across their range.

The “all species” model performed moderately well for predicting fish occurrence among all data sets. These results support the use of the “all species” model as a tool for fish habitat conservation efforts in the study area. Maintaining and restoring connectivity between streams is a cornerstone of most fish conservation strategies and also a legal requirement under the Canada Fisheries Act. Twenty-four percent of the 302 stream crossings were identified as potential migration barriers within medium or high probability streams. Culvert owners could complete a simple GIS exercise to determine the extent of potential fish habitat upstream of each crossing and use this information to identify infrastructure maintenance priorities or identify sites requiring additional inventories.

Limitations and opportunities for modeling fish occurrence using remote methods

We identified a number of considerations for related efforts in the future. First, correlative investigations that quantify fish occurrence from multiple predictor variables are popular approaches used to improve our knowledge of factors affecting fish occurrence; however, they have limitations introduced by nonlinear relationships and nonmeasured correlates (Rosenfeld 2003). Furthermore, unlike mechanistic models and true experiments, these descriptive approaches cannot indicate cause–effect relations. Nonlinear relationships between a species and features of its environment are expected where a species is adapted to a specific ecological niche. These relations have been identified using artificial neural networks (Oakes et al. 2005) or by including a quadratic term within a logistic regression approach (Mattingly and Galat 2002). We selected the latter approach, which proved useful for indicating habitat partitioning among the target species in headwater streams. However, the Pearson correlation test was not well suited for identifying relationships between nonlinear variables. We identified such relationships by plotting predictor variables of interest.

Second, the accuracy of remote-sensed parameters may limit the interpretations of true habitat preferences, particularly for channel slope, which can provide an indication of expected channel unit distribution, including pool extent and frequency. We found that channel sinuosity in small streams is not accurately mapped, and meandering channels extended beyond commonly accepted gradient thresholds of between 2% and 4% (Church 1992). Sinuosity has an important effect on fish habitat because it decreases stream power by reducing stream slope and increasing form roughness within habitat elements, including undercut banks and pools (Leopold et al. 1964). Therefore, fish habitat characteristics are expected to vary between a straight channel with a 3% slope and a meandering channel with 3% slope. In watershed science, there is a trend to use streams automatically generated from high-resolution DEMs rather than air photo interpreted streams (Benda et al. 2007). Generated streams follow the path from one grid cell in the DEM to the lowest adjacent grid cell. However, line length reduction will remain a factor for these generated streams when the size of the watercourse is smaller than the grid-cell size.

Third, the bull trout model could be improved by incorporating additional variables including measures of the size of downstream habitats such as the Strahler order downstream of the next confluence (Oakes et al. 2005), the locations of natural migration barriers in downstream areas (Fransen et al. 2006), and variables that characterize temperature preferenda (Dunham et al. 2003), including indicators of groundwater upwelling areas (Baxter and Hauer 2000).

Fourth, scaling issues are a key component of resource selection function models (Boyce 2006). Future studies on relations between bull trout distribution and land use should address scaling. Ripley et al. (2005) included reaches with areas up to 374 km^2, whereas we limited our study to reaches with a drainage area of ≤20 km^2. We used this approach because we anticipated that correlations with land use would be more pronounced because of close linkages between aquatic habitat and terrestrial environment in small channels. However, study designs that include large streams may be more suited to addressing impacts from angling. Our automated procedure is well suited to scaling up from the reach scale to major watersheds >1000 km^2 where resource management strategies are applied. This ability to align a habitat model with the scale suitable to support resource management decisions can increase the utility of models in conservation and management programs (Boyce 2006).

In conclusion, an automated approach for modelling spatial fish occurrence patterns provides new insight into fish habitat requirements in small foothills streams. The approach is conducive to validation with external data sets. Fundamentally, the maps also provide an opportunity for biologists to visually compare their own inherent spatial models of fish occurrence with those developed mathematically and then to generate ideas on how to further improve our understanding of the factors affecting fish distribution.

Acknowledgements

We thank the Foothills Model Forest funding partners, including the Alberta Conservation Association, Alberta Sustainable Resource Development, the Forest Resources Improvement Program, Jasper National Park, Natural Resources Canada – Canadian Forest Service, Hinton Wood Products (a Division of West Fraser Mills Ltd.), and Trout Unlimited Canada for providing financial assistance for this study. RJM was also supported by Natural Sciences and Engineering Research Council of Canada (NSERC) through an Industrial Postgraduate Scholarship (IPS). John Belanger and Scott Nielsen are acknowledged for their input on the use of logistic regression and the information–theoretic approach. Special thanks must be accorded to Chris Spytz for valuable input at all stages of the project. We also thank two anonymous reviewers for providing helpful suggestions to improve the manuscript.

References

Baxter, C.V., and Hauer, F.R. 2000. Geomorphology, hyporheic exchange, and selection of spawning habitat by bull trout (*Salvelinus confluentus*). Can. J. Fish. Aquat. Sci. **57**(7): 1470–1481.

Baxter, C.V., Frissell, C.A., and Hauer, F.R. 1999. Geomorphology, logging roads, and the distribution of bull trout spawning in a forested river basin: implications for management and conservation. Trans. Am. Fish. Soc. **128**: 854–867.

Beckingham, J.D., Corns, I.G.W., and Archibald, J.H. 1996. Field guide to the ecosites of west-central Alberta. The University of British Columbia Press, Vancouver, British Columbia.

Benda, L., Miller, D., Andras, K., Bigelow, P., Reeves, G.H., and Michael, D. 2007. NetMap: a new tool in support of watershed science and resource management. For. Sci. **53**(2): 206–219.

Bott, R., Murphy, P., and Udell, R. 2003. Learning from the forest: a fifty-year journey towards sustainable forest management. Fifth House Ltd., Calgary, Alta.

Boyce, M.S. 2006. Scale for resource selection models. Divers. Distrib. **12**: 269–276.

Bozek, M.A., and Hubert, W.A. 1992. Segregation of resident trout in streams as predicted by three habitat dimensions. Can. J. Zool. **70**: 886–890.

Brooks, E.S., Boll, J., Elliot, W.J., and Dechert, T. 2006. Global positioning system/GIS-based approach for modeling erosion from large road networks. J. Hydrolog. Eng. **11**(5): 418–426.

Burnham, K.P., and Anderson, D.R. 2002. Model selection and multimodel inference: a practical information–theoretic approach. Springer-Verlag, New York.

Church, M. 1992. Channel morphology and typology. *In* The rivers handbook: hydrological and ecological principles. *Edited by* G.E. Peets. Blackwell Scientific Publications, Oxford, UK. pp. 126–143.

Clarke, S., and Burnett, K.M. 2003. Comparison of digital elevation models for aquatic data development. Photogramm. Eng. Remote Sens. **69**(12): 1367–1375.

Donald, D.B., Anderson, R.S., and Mayhood, D.W. 1980. Correlations between brook trout growth and environmental variables for mountain lakes in Alberta. Trans. Am. Fish. Soc. **109**: 603–610.

Duhaime, K. 2003. Watershed delineation and stream classification for the Northern East Slopes Region. Facet Decision Systems Inc., Vancouver, B.C. Available from www.fmf.ca/FW/FW_Wc8. pdf [accessed 30 Oct. 2007].

Dumanski, J., Macyk, T.M., Veauvy, C.F., and Lindsay, J.D. 1972. Soil survey and land evaluation of the Hinton–Edson area. Alberta Alberta Institute of Pedology Report No. S-72-31, Soils Division, Research Council of Alberta, Edmonton, Alta.

Dunham, J.B., Rieman, B.E., and Chandler, G. 2003. Influences of temperature and environmental variables on the distribution of bull trout within streams at the southern margin of its range. N. Am. J. Fish. Manag. **23**: 894–904.

Ford, B., Higgins, P., Lewis, A., Cooper, K., Watson, T., Gee, C., Ennis, G., and Sweeting, R. 1995*a*. Brook trout (*Salvelinus fontinalus*). *In* Literature reviews of the life history, habitat requirements and mitigation/compensation strategies for selected fish species in the Peace, Liard and Columbia River drainages of British Columbia. Triton Environmental Consultants for Department of Fisheries and Oceans and the Province of British Columbia, Vancouver, B.C. Sec. 4. pp. 1–22.

Ford, B., Higgins, P., Lewis, A., Cooper, K., Watson, T., Gee, C., Ennis, G., and Sweeting, R. 1995*b*. Rainbow trout (*Oncorhynchus mykiss*). *In* Literature reviews of the life history, habitat requirements and mitigation/compensation strategies for selected fish species in the Peace, Liard and Columbia River drainages of British Columbia. Triton Environmental Consultants for Department of Fisheries and Oceans and the Province of British Columbia, Vancouver, B.C. Sec 2. pp. 1–25.

Fransen, B.R., Duke, S.D., McWethy, G., Walter, J.K., and Bilby, R.E. 2006. A logistic regression model for predicting the upstream extent of fish occurrence based on geographical information systems data. N. Am. J. Fish. Manag. **26**: 960–975.

Hassan, M.A., Church, M., Lisle, T.E., Brardinoni, F., Benda, L., and Grant, G.E. 2005. Sediment transport and channel morphology of small, forested streams. J. Am. Water Resour. Assoc. **41**(4): 853–876.

Isaak, D.J., and Hubert, W.A. 2000. Are trout populations affected by reach-scale stream slope? Can. J. Fish. Aquat. Sci. **57**: 468–477.

Jablonski, P.D. 1980. Pretreatment water quality of the Tri-Creeks experimental watershed. Report No. 7, Alberta Energy and Natural Resources, Forest Land Use Branch, Watershed Management, Edmonton, Alberta.

Jackson, S.D. 2003. Ecological considerations in the design of river and stream crossings. *In* International Conference on Ecology and Transportation, 24–29 August 2003, Lake Placid, N.Y. *Edited by* C.L. Irwin, P. Garrett, and K.P. McDermott. Center for Transportation and the Environment, Raleigh, North Carolina. pp. 20–29.

Klienbaum, D.G., Kupper, L.L., and Muller, K.E. 1988. Applied regression analysis and other multivariate methods. PWS-Kent Publishing Company, Boston, Massachusetts.

Latterell, J.J., Naiman, R.J., Fransen, B.R., and Bisson, P.A. 2003. Physical constraints on trout (*Oncorhynchus* spp.) distribution in the Cascade Mountains: a comparison of logged and unlogged streams. Can. J. Fish. Aquat. Sci. **60**: 1007–1017.

Leary, R.F., Allendorf, F.W., and Forbes, S.H. 1993. Conservation genetics of bull trout in the Columbia and Klamath River drainages. Conserv. Biol. **7**: 856–865.

Leopold, L.B., Wolman, M.G., and Miller, J.P. 1964. Fluvial processes in geomorphology. Freeman and Company, San Fransisco, Calif.

MacDonald, L., Smart, A., and Wissmar, R. 1990. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest. Report No. EPA 910/9-91-001, US Environmental Protection Agency, Region 10, Seattle, Wash.

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

Mattingly, H.T., and Galat, D.L. 2002. Distributional patterns of the threatened Niangua darter, *Etheostoma nianguae*, at three spatial scales with implications for species conservation. Copeia, 2002(3): 573–585.

McMaster, R.B. 1986. A statistical-analysis of mathematical measures for linear simplification. Am. Cartogr. **13**(2): 103–116.

McPherson, H.J. 1975. Sediment yields from intermediate-sized stream basins in southern Alberta. J. Hydrol. **25**: 243–257.

Moore, R.D., and Richardson, J.S. 2003. Progress towards understanding the structure, function, and ecological significance of small stream channels and their riparian zones. Can. J. For. Res. **33**: 1349–1351.

Morita, K., and Yamamoto, S. 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. Conserv. Biol. **16**(5): 1318–1323.

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

Olden, J.D., Jackson, D.A., and Peres-Neto, P.R. 2002. Predicitive models of fish species distributions: a note on proper validation and chance predictions. Trans. Am. Fish. Soc. **131**: 329–336.

Paul, A.J., and Post, J.R. 2001. Spatial distribution of native and nonnative salmonids in streams of the Eastern Slope of the Canadian Rocky Mountains. Trans. Am. Fish. Soc. **130**: 417–430.

Porter, M., Rosenfeld, J., and Parkinson, E. 2000. Predictive models of fish species distribution in the Blackwater drainage, British Columbia. N. Am. J. Fish. Manag. **20**: 349–359.

Rieman, B.E., and McIntyre, J.D. 1993. Demographic and habitat requirements for conservation of bull trout. Gen. Tech. Rep. INT-302, US Department of Agriculture Forest Service, Intermountain Research Station, Ogden, Utah.

Rieman, B.E., Peterson, J.T., and Myers, D.L. 2006. Have brook trout (*Salvelinus fontinalis*) displaced bull trout (*Salvelinus confluentus*) along longitudinal gradients in central Idaho streams? Can. J. Fish. Aquat. Sci. **63**: 63–78.

Ripley, T., Scrimgeour, G.J., and Boyce, M.S. 2005. Bull trout (*Salvelinus confluentus*) occurrence and abundance influenced by cumulative industrial developments in a Canadian boreal forest watershed. Can. J. Fish. Aquat. Sci. **62**(11): 2431–2442.

Rosenfeld, J. 2003. Assessing the habitat requirements of stream fishes: an overview and evaluation of different approaches. Trans. Am. Fish. Soc. **132**: 953–968.

Scrimgeour, G.J., Hvenegaard, P., Tchir, J., Kendall, S., and Wildeman, A. 2003. Cumulative effects of watershed disturbance on stream fish communities in the Kakwa and Simonette River Basins, Alberta. Northern Watershed Project Rep. No. 3, Alberta Conservation Association, Edmonton, Alta. Available from www.ab-conservation.com/about_us/reports_publications/NWP/NWP%20REPORT%203.pdf [accessed 30 Oct. 2007].

Selong, J.H., McMahon, T.E., Zale, A.V., and Barrows, F.T. 2001. Effect of temperature on growth and survival of bull trout, with application of an improved method for determining thermal tolerance in fishes. Trans. Am. Fish. Soc. **130**(6): 1026–1037.

Shepard, B.B. 2004. Factors that may be influencing nonnative brook trout invasion and their displacement of native westslope cutthroat trout in three adjacent southwestern Montana streams. N. Am. J. Fish. Manag. **24**: 1088–1100.

Spillios, L.C. 1999. Sediment intrusion and deposition near road crossings in small foothill streams in west central Alberta. M.Sc. thesis, Department of Renewable Resources, University of Alberta, Edmonton, Alberta.

Steel, E.A., Feist, B.E., Jensen, D.W., Pess, G.R., Sheer, M.B., Brauner, J.B., and Bilby, R.E. 2004. Landscape models to understand steelhead (*Oncorhynchus mykiss*) distribution and help prioritize barrier removals in the Willamette basin, Oregon, U.S.A. Can. J. Fish. Aquat. Sci. **61**: 999–1011.

Sterling, G.S. 1992. Fry emergence survival of rainbow trout, *Oncorhyncus mykiss* (Walbaum), following timber harvest in two foothills streams of west-central Alberta. M.Sc. thesis, Department of Forest Science, University of Alberta, Edmonton, Alta.

Taylor, E.B., Tamkee, P., Sterling, G.S., and Hughson, W. 2006. Microsatellite DNA analysis of rainbow trout (*Oncorhynchus mykiss*) from western Alberta, Canada: native status and evolutionary distinctiveness of “Athabasca” rainbow trout. Conserv. Genet. **8**: 1–15.

Van Sickle, J., Baker, J., Herlihy, A., Bayley, P., Gregory, S.V., Haggerty, P., Ashkenas, L., and Li, J. 2004. Projecting the biological condition of streams under alternative scenarios of human land use. Ecol. Appl. **14**(2): 368–380.

Wolter, C., and Arlinghaus, R. 2003. Navigation impacts on freshwater fish assemblages: the ecological relevance of swimming performance. Rev. Fish Biol. Fish. **13**: 63–89.