Quantitative Recovery Planning:

Understanding how human activities in watersheds can influence population dynamics and genetic structuring of diadromous fishes.

by

© Heather Dawn Bowlby

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Department of Biology

Memorial University of Newfoundland

Submitted September 2016

St. John's, Newfoundland and Labrador

Abstract

Developing quantitative relationships that link human-induced environmental change with changes in population dynamics for species of conservation concern is hindered by: (1) a limited understanding of the cumulative effect (and relative importance) of population regulation, spatial dynamics, and demographic processes, (2) issues with detectability for species-environment interactions owing to data characteristics and (3) the cumulative or confounding nature of multiple threats. Taking a single-species approach based on endangered Atlantic salmon, I have partially addressed these challenges in my four research chapters.

In chapter two, I characterized the conditions under which metapopulation structure would be expected to benefit a population assemblage and found that straying can reduce abundance and heighten extinction risk when productivity is low. For species of conservation concern, I would expect that remediation actions designed to influence demographic rates (e.g. mortality rates) would be more beneficial than actions influencing spatial dynamics. In chapter three, I accounted for the effects of observation and measurement error when quantifying relationships between hydrological variation and survival. Beyond the potential to change our interpretation of ecological relationships, I was able to infer the types of threats affecting juveniles in specific watersheds. In chapter four, I used patterns of effective dispersal to surmise the behavioural mechanism leading to watershed choice among straying adult salmon as well as the relative importance of multiple concurrent threats. My conclusions contradict some current perceptions on threats and suggest new directions for future research. In chapter five, I was able to develop a spatial tool that could inform management decisions or identify priority areas for restoration efforts. However, I was unable to fully characterize how environmental variation influences habitat utilization, distribution patterns, or population-level responses to human activities at multiple spatial extents.

The relationships I describe are among the first to be developed for endangered Atlantic salmon in Nova Scotia at a population level. Several of the analyses represent novel applications to conservation questions and have the potential to be extended or more widely applied. Because freshwater fishes, including diadromous fishes, are collectively one of the most imperiled species groups in the world, such research represents a timely contribution to conservation biology.

Acknowledgements

I would like to thank all of the staff and students that have been involved in field projects, the assessment program, and research projects within Fisheries and Oceans Canada which resulted in the population monitoring data and land use database I used in this research; in particular: Dr. Patrick O'Reilly, Ross Jones, Peter Amiro and Tracy Horsman. I am grateful for the stimulating discussions that my supervisors, Dr. A. Jamie F. Gibson and Dr. Ian A. Fleming, as well as my committee members, Dr. Shawn Leroux and Dr. Yolanda Wiersma, have had with me, and the ideas that I pursued because of them. I would not have even contemplated getting this degree without the encouragement of Dr. Ross Claytor, and I would never have finished it without the support from my family: Nat, Damian and Liam.

"What should we do to bring salmon back?"

Carl Purcell – life-long fisherman and advocate for salmon conservation

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter one: Introduction and overview

The concept of a population is central to our understanding of patterns in ecology, evolutionary biology and conservation biology, and it is a natural unit for which to evaluate the consequences of environmental change. Populations are typically considered to be mid-way along successive levels of aggregation, from an isolated individual to all individuals of a particular species (e.g., individual, aggregation/family, sub-population/deme, population, population aggregation/complex, sub-species, species). Most definitions of a population are largely qualitative and are based on the degree of similarity in either biological attributes or location/interactions in space and time, or on the potential for reproduction among individuals (see Table 1 in Waples & Giogotti 2006). However, organisms that have highly specific habitat requirements for specific life stages tend to be grouped into populations using a landscape boundary. Examples would be diadromous fishes in fresh water, pond-breeding amphibians, or migratory birds and butterflies using specific habitat types such as alpine meadows and old growth forest (Opdam 1991, Marsh & Trenham 2001, Jones 2006). Although these groupings are functionally useful for monitoring, management and policy (Waples & Gaggiotti 2006), it is also important to recognize that dynamics at any level are influenced by variability that exists at both smaller and larger levels of aggregation. It is useful to consider the example of Atlantic salmon (*Salmo salar*) in Atlantic Canada, where populations exhibiting broadly similar life history (Chaput et al. 2006) and genetic characteristics (O'Reilly et al.

2012, Bradbury et al. 2014) have been identified and grouped into regional units for conservation and management (COSEWIC 2010) based on the idea that the dynamics of individual populations would be expected to be similar throughout each grouping. However, measurable differences in life history characteristics, biological attributes (e.g. maturation rates, freshwater residency; Gibson & Bowlby 2013) and genetic structure (O'Reilly et al. 2012) have been described among components of individual populations inhabiting a single river system, where each component would be expected to exhibit somewhat different dynamics. Thus to understand the dynamics of a salmon population within a specific watershed, it is important to place that population within a larger regional context, while at the same time, to consider the extent or potential effect of localized variability within the watershed.

Evaluating how specific environmental factors and intrinsic life history characteristics culminate in changes in population size is central to understanding how populations are regulated (Shaub & Abadi 2011). The most basic description of the abundance of an open population at a given time is: $N_t = N_{t-1} + B - D + I - E$; where current population size (N_t) is jointly determined by population size in the previous year (N_{t-1}) plus births (B), minus deaths (D), plus immigration (I) and minus emigration (E). Changes in the number of individuals born, dying, or moving are typically understood in terms of a rate, called a population vital rate. Variation of vital rates in space and time can result from intrinsic, localized interactions with habitat characteristics in terms of abiotic conditions, prey resources, competitors and predators (Wu & Loucks 1995), or can reflect larger-scale environmental heterogeneity in terms of habitat size, quality, distribution or isolation (Harrison 1991, Hanski & Gilpin 1991). Throughout this thesis, I use an operational definition of habitat (Hall et al. 1997), where habitat is understood to represent any location with resources that enable individuals from a population to survive and/or reproduce. Describing changes in vital rates at a population level in relation to human activities can be thought of hierarchically: it would be based on an understanding of life history strategy, population regulation, behaviour, habitat requirements and spatial habitat structuring; next would be consideration of inter-relationships among these and how spatial and temporal variability may culminate as changes to abundance and distribution; and finally would be the attempt to characterize how anthropogenic activities have influenced variability in ecological processes, leading to population decline.

Demographic vs. spatial perspectives of population dynamics

Ecological inquiry on changes to population dynamics has gone in two general directions: the first being focused on factors affecting demographic processes (reproductive rates, stage-specific mortality and life history strategy), and the second being focused on factors influencing how organisms interact spatially with the landscape (primarily through immigration and emigration; e.g., Sutherland et al. 2013). It is expected that the overall influence of specific factors will vary depending on the system under consideration (Sutherland & Norris 2002, Kinnison & Hairston 2007, Nislow et al. 2011, Kiernan & Moyle 2012). Explicit consideration of life history strategy, population regulation (i.e., compensatory density dependence), habitat requirements and environmental stochasticity is fundamental to understanding population dynamics relative to demographic processes. In terms of life history, theory predicts that characteristics such as iteroparity have evolved as a mechanism to maximize reproductive success in variable environments, given that specific cohorts (i.e., offspring from a specific reproductive event) are exposed to different environmental conditions by being distributed over multiple years (Figge 2004, Moore et al. 2010). Similarly, individual investment in early reproduction relative to growth and delayed maturation varies with the timing of mortality, and reproductive allocation has been linked to the degree of environmental variation, organismal stress, and the availability of resources (references in Winemiller 2005). In terms of population regulation, it is expected that species, populations, or life stages characterized by strong density dependence will partially attenuate sources of density-independent mortality through increases in survival or fecundity (Jonsson et al. 1998, Rose et al. 2001, Nislow et al. 2011). Thus, there has been a substantial debate on the relative importance of density dependent versus stochastic environmental processes on population dynamics at various spatial and temporal scales (Benton et al. 2006, Kiernan & Moyle 2012).

A spatial perspective to population dynamics views distribution and abundance patterns as resulting primarily from the size and geographical distribution of habitat as well as movement of organisms over the landscape (Hanski 1999, Bowler & Benton 2005, Thornton et al. 2011, Kubisch et al. 2014). Natural selection would be expected to favour individuals that prefer, and can access and/or defend high quality habitat for growth and reproduction (Fretwell & Lucas 1970, Bowler & Benton 2005). Although density would be expected to be higher in these habitats, the amount of area used or defended by individual animals is extremely variable (Kubisch et al. 2014). Organisms assess habitat suitability on the basis of characteristics that they can perceive, even though such secondary cues are imperfect and may result in individuals spending considerable time in suboptimal areas (Robertson & Hutto 2006, Johnson 2007). Such variability can obscure relationships between habitat quality and density at larger scales, particularly for generalist species that can profitably exploit a variety of conditions.

Habitat quality and resource availability relative to inter- and intraspecific densities largely determine the relative strength of competition experienced within localized habitats, which in turn, influence emigration rates (Kinnison & Hairston 2007, Pfluger & Balkenhol 2014). In general, animals are more likely to leave habitats of lower quality and areas of high density (Lin & Batzli 2001, Matthysen 2005), but individual dispersal strategies can also be adjusted relative to habitat unpredictability, as well as to characteristics such as sex, age, or morphology (Bowler & Benton 2005, Bonte et al. 2012). Successful immigration depends on the number of individuals reaching a habitat, habitat quality, current occupancy and carrying capacity, as well as the relative fitness of individual dispersers within the new habitat (i.e., reproductive success). Processes such as local adaptation would be expected to reduce effective dispersal (i.e., dispersal that results in successful reproduction; Pfluger & Balkenhol 2014), which is one reason that straying has been seen as partially maladaptive in species such as salmonids (Rieman & Dunham

2000). A comprehensive review of salmonids (Fraser et al. 2011) estimates local adaptation to occur in 55-70% of populations, with local populations having 1.2 times the average fitness advantage relative to foreign populations or relative to their performance in new environments. Fraser et al. (2011) found local adaptation to occur over a variety of spatial scales and to develop relatively quickly, within 6-30 generations. However, the majority of evidence for local adaptation in Atlantic salmon (the best-studied species) is indirect and comes from ecological correlates in fitness-related traits, the failure of many translocations, and the poor performance of domesticated stocks (Garcia de Leaniz et al. 2007). Also relative to successful immigration, it is generally expected that geographically proximate habitats would be encountered more frequently by dispersing individuals and thus the costs associated with locating them would be lower (Bonte et al. 2012). However, such a pattern might not be expected for migratory species where overall dispersal distances are much greater than the average distance between habitats used by a specific life stage (Guillot et al. 2009).

Although a population is a natural unit for which to evaluate the consequences of environmental change, developing quantitative relationships first requires explicit consideration of underlying patterns which determine abundance, density or connectivity at a population level. Relative to changes in demographic rates, the information above suggests that: (1) on observational time scales, average population vital rates (e.g., mortality and maturation rates) will be largely determined by life history strategy, (2) anthropogenic activities would be expected to change environmental conditions in some way, (3) annual variation in vital rates should be related to such environmental variation and thus may arise from anthropogenic activities, (4) density dependent processes occurring within localized habitats would be expected to partially attenuate any effect of environmental change, and (5) dispersal patterns could have an apparent dampening effect on stage-specific mortality rates owing to the re-distribution of individuals. Overall, measured abundance and density patterns are expected to be related to habitat quality, and environmental variation that reduces habitat quality is expected to also reduce localized abundance or density through changes to demographic rates. However, population dynamics would be further affected by the spatial characteristics of habitats, where environmental conditions within habitats could alter the propensity for and success of individual dispersal events. The way in which such processes would culminate in changes in abundance, density or genetic structure is poorly known, yet is expected to be scale-dependent (Benton et al. 2006).

Watersheds in the landscape

The spatial structure and dendritic nature of watersheds regulates or modifies their ecological organization, where the physical and biological processes at small spatial extents are constrained and/or modified by processes occurring at larger spatial extents (Campbell Grant et al. 2007, Johnson et al. 2007). This hierarchical organisation is a result of convergence in the stream network, where water accumulates along the length of the network and transports multiple components from the upper tributaries to its confluence, such as nutrients, sediments, energy, or organisms (Huang et al. 2007). At a catchment scale, precipitation patterns, geology, topography and land cover interact to influence hydrological conditions and landscape structure. At smaller spatial extents, factors such as the distribution and nature of riparian vegetation, or biotic community structure exert further influence on the characteristics of localized habitats as well as their realized variability (Johnson et al. 2007, Ugedal & Finstad 2011). For classification and comparison, this convergence can be represented multiple ways, such as Strahler order classes (Strahler 1957) that describe the relative positions of reaches within a stream network based on upstream confluences, or by using metrics such as upstream catchment area as a more universal descriptor of the relative size of specific reaches (Altermatt 2013). One of the most important conceptual models with which to understand the spatial heterogeneity of streams resulting from convergence (Fausch et al. 2002) is the River Continuum Concept (RCC; Vannote et al. 1980). Averaged over space and time, the RCC predicts relatively gradual changes in geomorphological processes along a stream network resulting from changes in channel morphology as influenced by flow. Thus over large spatial extents, heterogeneity in habitats is associated with downstream decreases in channel slope, riffle development and substrate size, and with increases in stream width, depth and pool size (Schlosser 1991).

Ecosystem function in river networks can depend on the interplay between three main factors: local environmental conditions (e.g., Woodward et al. 2012), the induced directionality (also called anisotropy) in the spread of abiotic or biotic inputs (Rodriguez-Iturbe et al. 2009), and the topological network structure defined by stream reaches and confluences (Benda et al. 2004, Brown & Swan 2010). As general patterns, within-species and community-level diversity are predicted to be higher around confluences and in downstream reaches, largely due to increased habitat complexity (Altermatt 2013) and constraints on the distance and directionality of movement within the network (Kiffery et al. 2006), respectively. However, stream ecosystem dynamics are greatly influenced by disturbance arising from variability in ecosystem processes at multiple spatial and temporal scales (Stanley et al. 2010). One example of such variability is a natural flow regime, which is characterized by a cyclical seasonal pattern, punctuated by daily, seasonal and annual variability in the timing, duration and severity of floods, droughts and intermittent flows (Olden & LeRoy Poff 2003). Annual patterns of population distribution and abundance within the stream network reflect a species' habitat requirements as well as any behavioural response to variability in ecosystem structure and function (Campbell Grant et al. 2007), and are expressed as changes in vital rates and overall population dynamics.

The patchy nature and relative isolation of specific habitat types within a watershed would be expected to lead to multiple groupings of individuals that could exhibit differences in dynamics. Both the extent of isolation as well as the strength of spatial coherence in environmental conditions (i.e., the Moran effect) would partly determine the degree of similarity or synchrony among the different groups (Koenig 2002). The value of this diversity relative to species abundance and persistence is currently understood using economic portfolio theory (Figge 2004), which posits that asynchronous dynamics among a diversified aggregation of distinct groups buffers against temporal variability in an individual group, thereby increasing the overall stability and long-term performance of the aggregation (Moore et al. 2010, Schindler et al. 2010). Ecologically, this translates into lower inter-annual variations in abundance at the level of the population coupled with greater resilience to environmental change and lower risk of global extinction (Earn et al. 2000, Figge 2004).

Considerations related to hypothesis testing

Quantifying changes in population dynamics owing to environmental variability requires some type of statistical model, which can be as simple as a description of central tendency. Any model of population dynamics allows for the organization, conceptualization, and interpretation of ecological data, enabling researchers to form hypotheses and test predictions (Courchamp et al. 2008). By necessity, they are a simplification of the natural world, and conclusions arising from them are critically dependent on both their inherent assumptions as well as the characteristics of the data used in parameterization. There is a wealth of information about environmental processes that might be expected to be related to population dynamics. However, it is not immediately obvious how to aggregate environmental monitoring data into biologically meaningful indices (Mueter et al. 2002). For example, typical environmental monitoring takes place at a fixed location with a fixed sampling frequency (e.g., hourly flow measurements at a hydrological monitoring station), while population-level monitoring data often yields a single annual estimate for the life stages sampled (Levin 1992, Schwarz & Seber 1999). In this instance, the environmental monitoring data would need to be aggregated into some type of annual estimate to be on an appropriate temporal scale for any comparison with changes in population dynamics. Annual variability in a hydrological flow regime can be expressed relative to average conditions, variation in means or medians, timing and duration of extreme conditions, skewness, or a multivariate approach, among others (Olden & LeRoy Poff 2003). Limited population monitoring data, where 20-30 data points are considered to be substantial, plus the large number of possible environmental predictors impose quite strong statistical limitations on hypothesis testing (Graham 2003, Zuur et al. 2010, Fredericksen et al. 2014). Thus, quantitative relationships can be difficult to describe due to issues related to scale and consequently, to detectability (Rose 2000, Fredericksen et al. 2014).

Conservation of fish populations

Freshwater and diadromous fishes are collectively one of the most imperilled species groups in the world (Dudgeon et al. 2006). Globally, they are vulnerable because they provide an abundant and easily accessed food source, occur in areas where human populations are concentrated, and inhabit an ecosystem that is particularly susceptible to damage through direct modification as well as by activities within catchments (Lake et al. 2007, Olden et al. 2010). Human activities that impact upon watersheds often represent an assemblage of changes to fish and fish habitat. For example, infrastructure development such as roads can precipitate multiple alterations to the ecological, demographic or behavioural attributes of populations leading to reduced viability and consequently reduced abundance (reviewed in Gucinski et al. 2001). Alternately, an assemblage of different types of human activities can affect the same ecological process in watersheds. For example, mining operations, urbanization, infrastructure development, agriculture, and forestry have all been linked to changes in hydrological flow patterns (Allan 2004, Broadmeadow & Nisbet 2004, LeRoy Poff et al. 2006), where the overall severity of anthropogenic effects are expected to be exacerbated by changing precipitation patterns due to climate change (Milly et al. 2005).

Given the rate of population declines and species extinction, there is a pressing need to produce quantifiable predictions on the consequences of environmental change to aid in conservation and recovery efforts (Sutherland & Norris 2002, Palmer 2009, Olden et al. 2010). This directly depends on being able to identify the environmental and ecological factors limiting abundance (i.e., the threats) at a population level (Lawler et al. 2002). It would be expected that population decline would be more likely to occur when threats arise that are outside of a population's previous evolutionary experience (i.e., the variable is new, such as species invasions) or when change outpaces adaptation (e.g., increasing severity and frequency of extreme events or catastrophes) (Brook et al. 2008). For species in which the causes of decline are thought to result from multiple simultaneous threats, there is considerable debate on how individual pressures would culminate in changes to abundance. In general, effects on population dynamics might be expected to be non-linear and non-intuitive (Rose 2000), representing some combination of cumulative, antagonistic, and threshold-type responses among agents of environmental change.

Case study to assess practical implications

A recent review of the status of anadromous Atlantic salmon (*Salmo salar*; Linnaeus, 1758) by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) grouped populations into 16 designatable units (DUs), representing evolutionary significant groupings below the species level. Each DU is considered to contain a distinct population assemblage on the basis of genetic data, life history characteristics, the general status of populations, and the geospatial characteristics of watersheds (COSEWIC 2010). Five of the DUs were designated as endangered, one as threatened, four as being of special concern, four as not at risk, one as data deficient and one as extinct (COSEWIC 2010). Four of the five DUs considered to be endangered are located in Atlantic Canada, including: the inner Bay of Fundy (iBoF), outer Bay of Fundy (oBoF), Southern Upland (SU), and Eastern Cape Breton (ECB) regions (Figure 1-1). The fifth was Anticosti Island in Quebec.

Recent population dynamics modeling from monitored populations in the SU region (Gibson & Bowlby 2013) indicated that extinction risk is high, yet could be substantially reduced through relatively small increases to freshwater productivity (e.g., a 20% increase in egg to smolt survival). Therefore, there is strong interest from multiple stakeholders to begin remediation projects to address threats affecting specific watersheds. To develop such remediation priorities requires a thorough understanding of quantitative relationships between anthropogenic activities affecting watersheds and population-level changes in abundance and productivity.

Overall research goals

For many endangered species, there is a pressing need to advance our understanding of species-environment relationships, in order to guide recovery planning and maximize the potential that remediation actions will lead to population increase (Lawler et al. 2002, Scheurell et al. 2006, Driscoll & Lindenmayer 2012). I specifically chose my research questions to partially address three of the main issues related to developing these relationships. First, there is considerable debate on the cumulative effect and relative importance of population regulation, spatial interactions between organisms and the landscape, and localized demographic processes (e.g. Benton et al. 2006, Kiernan & Moyle 2012), yet using one perspective as a basis for recovery planning can lead to very different practical guidance relative to the alternate perspectives (Morales 2011, Driscoll & Lindenmayer 2012). Second, observational data collected at a population level tends to be quite noisy and limited by sample size, which can lead to issues with detectability when trying to describe species' responses to environmental change (Rose 2000, Allan 2004). And lastly, several different types of anthropogenic activities in watershed catchments can lead to cumulative and/or confounding changes to ecological processes in streams (Altermatt 2013, Jeffrey et al. 2015), many of which affect fish survival (LeRoy Poff et al. 2010). Thus at a population level, species responses to environmental change may be largely non-intuitive (Rose 2000).

For each analysis, I aimed to provide a critical evaluation of the applicability of various ecological ideas to conservation questions, to extend current statistical methodologies or propose alternatives, to establish quantitative links between land use or environmental characteristics and population dynamics, and to explicitly consider the practical implications of results for conservation and recovery planning relative to a case study: endangered Atlantic salmon populations. In chapter two, I developed a simulation model based on a simplified diadromous life history that I used to evaluate how spatial structure and demographic variability integrate to determine the dynamics of an assemblage of populations. As such, the analysis is a basis for a critical assessment of whether metapopulation theory should be used to develop remediation priorities for conservation and management for species such as diadromous fishes that exhibit substantial population structuring. In chapter three, I focused on quantifying the influence of environmental change, specifically hydrological variation, on survival in isolation from density-dependent processes using temporal data. By explicitly evaluating sources of uncertainty, I demonstrated how our ecological understanding of environmental change can be directly affected by decisions made *prior* to analyses, and explicitly consider how this would affect recovery planning. In chapter four, I focused on how individuals respond to environmental gradients through movement and use patterns of effective dispersal to infer the relative importance of specific types of threats. This relies on optimally characterizing expected patterns of genetic variability, and I proposed that these need to be understood relative to environmental rather than geographic gradients. In chapter five, I evaluated the support for several ecological hypotheses on habitat utilization, distribution patterns within watersheds, and population-level responses to human activities using juvenile density data and landscape predictors developed at two spatial extents. Although results relative to theoretical predictions are inconclusive due to characteristics associated with data collection, they could be applied in a management context to predict effects on juvenile salmon productivity from proposed industrial development or other anthropogenic changes in watershed catchments. In my final chapter, I synthesized conclusions from my research relative to theoretical expectations, identify limitations and opportunities for further analyses, and discuss practical considerations for moving forward with recovery planning.

Figures

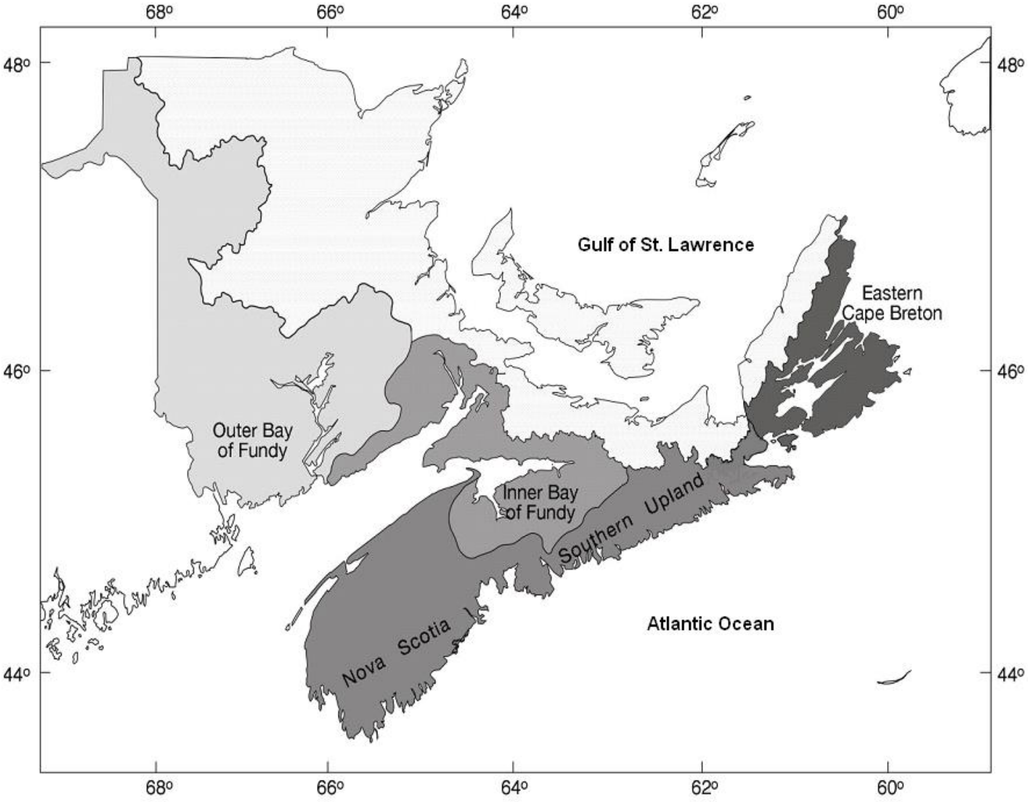

Figure 1-1. Map showing the location of the inner Bay of Fundy (iBoF), outer Bay of Fundy (oBoF), Southern Upland (SU), and Eastern Cape Breton (ECB) regions of Atlantic Canada.

Co-authorship statement

The research in this thesis was undertaken as a component of a dynamic research and assessment program for Atlantic salmon in Nova Scotia, conducted by Fisheries and Oceans Canada. For the manuscripts resulting from my thesis chapters, individuals who contributed greatly to their intellectual development and execution are included as co-authors.

(i) Design and identification of the research proposal.
Many of the general ideas contained in this thesis were previously discussed between me and others; for example the idea that hydrological change or land use would likely be related to juvenile salmon survival. However, I was solely responsible for the design and identification of my research proposal. I did draw on the expertise of Dr. Jamie Gibson to structure how my ideas might best contribute to the goals of the larger salmon assessment program. I consider the main ideas explored in each chapter to have been identified and developed by me. For example, the idea that metapopulation structure may not benefit diadromous populations with low productivity (chapter two); the idea that robust regression might be useful to evaluate data with error (chapter three); the idea to use landscape genetics as a way to evaluate concurrent threats (chapter four); and the idea that spatial distribution patterns operating over multiple spatial extents may influence any understanding of threats (chapter five).

(ii) Practical aspects of the research
I used a wide variety of data sources. Some of these had been collected or developed by others as part of ongoing assessment activities for salmon (e.g. the microsatellite data used in chapter four). For others, I had been involved in collecting and archiving the data (e.g. the electrofishing data used in chapters three and five), or I had been responsible for overseeing their development (e.g. the landscape database used in chapter four). I independently developed the spatial predictors used in chapter five from a pre-existing landscape database. The data source that I made the most use of was the electrofishing data, which represents the culmination of eight years of work by Dr. Jamie Gibson to electronically archive. For chapter four, I offered authorship to Dr. Patrick O'Reilly in recognition of his work in developing the microsatellite data, but he declined. Also, I based the population dynamics model in chapter two on R code that was originally written by Dr. Jamie Gibson for a single population of a diadromous fish. Although I made use of the his novel idea and code development for tracking individuals using multi-dimensional arrays, the simulations presented in chapter two represent substantial modifications to the original model.

(iii) Data analysis

I both identified appropriate analyses and undertook them on my own. I regularly discussed results with my supervisors, and they provided guidance as appropriate on potential ways to think about specific problems. They also provided guidance on resources I could consider to broaden my understanding of specific topics or analysis issues. Chapter four was co-authored by Dr. Ian Fleming, because I drew on his expertise in ecological theory as related to population genetics throughout the development of the analyses.

(iv) Manuscript preparation
I prepared all manuscripts. However, they all went through substantial editing prior to any submission. All co-authors as well as my committee members made substantial contributions to this editing process.

Publications (published or anticipated) and authorship arising from this thesis:

Chapter Two:

Bowlby, HD & AJF Gibson. Does metapopulation structure reduce extinction risk for diadromous species as productivity declines? *Journal still in consideration*.

Chapter Three:

[1]Bowlby HD & AJF Gibson. 2015. Environmental effects on survival rates: robust regression, recovery planning and endangered Atlantic salmon. Ecology and Evolution 5(16):3450-3461.

Chapter Four:

[1]Bowlby, HD, AJF Gibson & IA Fleming. 2016. Applying landscape genetics to evaluate threats affecting endangered Atlantic salmon populations. Conservation Genetics DOI 10.1007/s10592-016-0824-7.

Chapter Five:

Bowlby HD & AJF Gibson. Evaluating the spatial extent and strength of relationships between juvenile Atlantic salmon density and environmental metrics. *Journal still in consideration*.

---

[1] Published open source, under the Creative Commons license

Chapter two: Does metapopulation structure reduce extinction risk for diadromous species as productivity declines?

Abstract

Metapopulation structure (i.e., an aggregation of partially isolated populations interconnected by non-negligible dispersal rates) is typically thought to increase regional abundance, promote population persistence, and help in the re-establishment of extirpated populations. The effectiveness of recovery strategies for species of conservation concern can depend on the validity of these perceptions if finite resources become focused on population connectivity in exclusion of alternate actions. Given that the theoretical research supporting the benefits of metapopulations typically assumed high productivity (i.e. that individual populations could easily increase in size) when evaluating spatial dynamics, metapopulation structure may not be as beneficial when productivity is low. Here, I use a simulation model to compare the effect of among-population dispersal relative to within-population dynamics on the abundance trajectory and extinction probability of a population assemblage under five different productivity scenarios. My analyses demonstrate maximum abundance in the population assemblage at moderate levels of productivity. When productivity is extremely low, straying does not ensure persistence of non-viable populations or enable population re-establishment, yet it does increase extinction risk in viable populations. Combined, these effects lead to faster

declines in abundance in a population assemblage connected by straying (i.e. a metapopulation), relative to one composed of the same number of isolated populations. This result was robust to a wide range of alternate spatial and life history parameterizations of the simulation model. Thus, for endangered species characterized by low or very low productivity, changes to demographic rates appear to have a much larger effect on overall extinction risk relative to spatial dynamics and likely should be the focus of recovery planning.

Introduction

Applying ecological theory to practical conservation issues relies first on describing general patterns that can be used to predict the response of specific systems, and then using these predictions as the basis for management recommendations and applied remediation actions (Driscoll & Lindenmayer 2012). Metapopulation theory has been widely proposed as a basis from which to understand species' responses to human-dominated landscapes, specifically in terms of abundance and distribution patterns to assist in conservation (Hanski 2011, Eaton et al. 2014). For many species, maintaining metapopulation structure has been proposed (e.g., Morales 2011) or implemented (e.g., Gusset et al. 2008) as a way to promote conservation in situations of severe population decline. For example, recovery planning for diadromous fishes typically considers both abundance as well as distribution targets when defining conservation objectives (e.g., Ruckelshaus et al. 2002, Bowlby et al. 2014). For endangered Atlantic salmon (*Salmo salar*), justification for the distribution target was partially based on the perception that individuals from extant watersheds would stray into surrounding watersheds and thus promote persistence or population re-establishment (Bowlby et al. 2014). However, these perceived benefits of straying could be disproportionate to the potential realized benefit, and could reduce effectiveness of recovery actions if limited resources become focused on maintaining or promoting population connectivity in exclusion of alternate actions. Thus perceived benefits of metapopulation structure become important when trying to qualitatively rank various remediation options for the conservation of a suite of populations.

The idea of a metapopulation is strongly related to the theory of island biogeography put forward by MacArthur and Wilson (1967) to explain patterns of biodiversity, but shares similarities with any system that is considered to be a mosaic of islands or patches. Among aggregations of a single species, metapopulation theory begins from the premise that habitat area and isolation are the key determinants of abundance and distribution patterns at a regional level (Hanski 1999, Driscoll 2007) and the main ecological process assumed to govern the dynamics of metapopulations is dispersal (Clinchy et al. 2002, Harrison 1991). Thus, metapopulation structure is commonly understood to refer to an aggregation of partially isolated populations interconnected by non-negligible dispersal rates, while metapopulation dynamics refers to changes in abundance or distribution owing to the spatial configuration of habitat, habitat size and movement of individuals among habitat patches (Pellet et al. 2007). Relative to species conservation, metapopulation structure is perceived to confer resilience to environmental variability or catastrophic habitat loss (e.g., Wilcox et al. 2006, Vuilleumier et al. 2007), thereby reducing overall extinction risk (Gyllenberg & Hanski 1997, Simberloff 1998). Furthermore, immigration into small habitat patches from large ones (called source-sink metapopulations (Harrison 1991)), is thought to maintain individual populations that may otherwise extirpate or help in the re-establishment of extirpated populations, thereby increasing overall persistence. This is commonly referred to as the rescue effect (Brown and Kodric-Brown 1977) or source-sink dynamics (Harrison 1991), even though it is movement from (rather than higher productivity in) the source populations that reduce extinction risk in the smaller sink populations.

The majority of theoretical metapopulation models share the common characteristic that space is organized into discrete patches separated by an inhospitable matrix (Marsh & Trenham 2001, Bender & Fahrig 2005). Such models become spatially explicit when immigration rates are modeled as a function of distance among habitat patches (e.g., Moilanen 2004), incorporating the ecological assumption that there are energetic and evolutionary costs associated with movement (Bonte et al. 2012). Extinction probabilities or population size are typically modeled as a function of habitat area (e.g., Gyllenberg & Hanski 1992), reflecting the ecological assumptions that large habitats can support more individuals than small ones (i.e., have a higher carrying capacity; Kubisch et al. 2014) and that extinction probability is primarily a function of population size (Simberloff 1998). Because within patch dynamics are ignored in the majority of theoretical models (Baguette and Schtickzelle 2003, Pellet et al. 2007), life history processes such as births and deaths are not explicitly accounted for. By excluding consideration of population dynamics, individual populations inhabiting individual patches are assumed to exhibit no life history variation (i.e., population vital rates are identical; Dennis et al. 2003). In addition, using patch area as a proxy for population size means that populations are assumed to be able to grow in size upon reaching suitable habitat, quickly (in lagged models; e.g., Ellner and Fussmann 2003) or instantly (e.g., Gyllenberg and Hanski 1997) reaching the carrying capacity of each habitat patch. Furthermore, because populations in these models increase quickly or instantly to carrying capacity, emigration is assumed to have no negative effect on populations in large habitat patches (i.e., on source populations; Gyllenberg & Hanski 1992).

Although these assumptions are implicit in model structure, they are typically not stated as caveats of the results. Thus, the impression that metapopulation structure promotes regional persistence through source-sink dynamics or the rescue effect could be limited to situations in which population-level productivity is high. For species of conservation concern, population productivity may be substantially reduced from historical values. Deterministically, species or populations with a negative maximum per capita growth rate tend to decline over time, regardless of starting population size or habitat area (Lande et al. 2003, Hutchings et al. 2012a). Even if the maximum per capita growth rate is not predicted to lead to extinction, populations may have very little capacity to increase in size following stochastic declines (i.e., low resiliency). Often, endangered species are those that have undergone large declines in abundance over time

(Gibson et al. 2011, Gibson et al. 2015) and/or have not responded to the abatement of specific threats (Hutchings et al. 2012b). The former is consistent with reduced population productivity and the latter with low resiliency. Both characteristics suggest that demographic rates may have strong influence on population growth and extinction probabilities in a manner independent from habitat area and isolation (i.e., independent of metapopulation dynamics). Thus, metapopulation theory may be a relatively poor basis from which to describe populations with low productivity (i.e., many endangered species), and could even hinder conservation if the perceptions regarding metapopulation dynamics are false when productivity is low.

In nature, considerable population structuring would be expected in species that have highly specific habitat requirements for specific stages, and it is not surprising that they have been described as having metapopulation structure (Marsh & Trenham 2001). Many consider diadromous fishes to be the quintessential example (Rieman & Dunham 2000, Jones 2006, Schtickzelle & Quinn 2007) given that populations inhabit specific watersheds for reproduction and juvenile rearing (Thorstad et al. 2011), yet are not entirely isolated due to relatively low straying rates of adults to non-natal watersheds (Keefer and Caudill 2014). Such emigration would be expected to be a relatively constant percentage of population size, given that straying behaviour is thought to result largely from imprecise homing (Keefer and Caudill 2014) as related to environmental effects on imprinting (e.g., Westley et al. 2015). There is also some evidence that straying behaviour becomes more frequent when population sizes are small (Westley et al. 2015), in that individuals from larger groups tend to be able to respond better to environmental signals, termed the Collective Navigation Hypothesis (Berdahl et al. 2014). Both imprecise homing and the potential for collective navigation suggest that metapopulation structure for diadromous fishes would not be expected to disappear, and may even become more prominent, as abundance declines. Thus, metapopulation structure has the potential to have greater or lesser influence on persistence of diadromous fishes at different levels of population productivity.

This chapter assessed whether the perceived benefits of metapopulation structure for diadromous fishes hold when population dynamics are explicitly incorporated into a spatial model, particularly as population productivity declines. More specifically, the following two questions are addressed: (1) Does metapopulation structure always confer resilience to extinction or maintain greater regional abundance provided individual populations have some ability to grow in size? and (2) Does the rescue effect (movement from source populations to sinks) always reduce extinction risk as productivity declines? To answer these questions, I developed a life history-based simulation model in which the dynamics of individual populations were connected by dispersal. By simulating populations at progressively decreasing levels of productivity while keeping spatial structuring constant, I was able to evaluate the consequences of dispersal on overall abundance and extinction risk. By incorporating a contrast in habitat area among populations, I could evaluate the effect of source-sink dynamics on persistence. Results would have immediate practical application in terms of how recovery planning may be approached for diadromous fishes of conservation concern; specifically, with regard to how much focus should be given to actions that maintain or facilitate population connectivity.

Methods

I extended a simulation model for a single population to consider multiple populations distributed in space, connected by emigration and immigration. For the individual populations inhabiting distinct habitat patches, I modeled local dynamics explicitly in terms of maturity schedules, mortality rates, and environmental variability (detailed in Appendix 2-1). Given that the model approximates a diadromous life history, the habitat patches can be thought of as individual watersheds. These simulations considered 50 watersheds, with large ones (N = 10) containing 200 dimensionless habitat units and small ones (N = 40) containing 40 units (80% smaller). I incorporated this contrast to both approximate the range of sizes that can exist among natural watersheds in a region (e.g., Bowlby et al. 2013) as well as to permit an evaluation of the rescue effect for the simulated metapopulations. I assumed that the individual populations within watersheds were either open with a 10% stray rate (i.e., 10% of adult abundance from a specific population emigrates per year; termed a metapopulation for the remainder of this chapter) or closed (i.e., an annual stray rate of 0; termed isolated populations for the remainder of this chapter). These are the two types of population assemblages (also called population aggregations) that I considered in the simulations. Modeling isolated populations can be thought of as the null hypothesis against which I assessed the effects of straying. Large populations had higher carrying capacity and therefore had larger population sizes than small populations at the same level of productivity (see below). Given the spatial component of the model (i.e., the linear distribution of watersheds and dispersal following IBD), and in the absence of environmental variability, individual large populations would lose more individuals through emigration in a specific year than they would gain through straying (i.e. net loss through emigration); the opposite would be true for small populations (i.e. net gain through immigration), given that small populations outnumber large in these simulations. Thus, large populations were defined as sources and small populations as sinks in the projections, to represent a source-sink metapopulation.

Diadromous fishes are thought to experience population regulation (i.e., compensatory density dependence) within freshwater environments, so I modeled freshwater production as a Beverton-Holt process (Gibson & Myers 2003, Gibson 2006, Bowlby & Gibson 2011) during the first year for each population. There can be considerable variability among populations or species in the time diadromous fishes spend in freshwater (e.g., less than one year for American shad (Hasselman et al. 2013) and one to four years for Atlantic salmon (e.g., Chaput et al. 2006)). Here I have assumed that individuals emigrate to the marine environment after one year, largely to limit the complexity of the simulation model. However, if I had incorporated variability in freshwater residency, I would have expected lower overall extinction probabilities in each specific productivity scenario, given that different components of a cohort would be exposed to different annual conditions in fresh water (Figge 2004, Moore et al. 2010). Immature fish returned to spawn after one or two years at sea and were able to repeat-spawn up to 3 times. Again, there is considerable variability among diadromous populations or species in the time spent in the marine environment prior to spawning as well as in the number of times an individual will spawn (e.g., Taylor 1991). Thus, my simulation model captures the unique characteristics of a diadromous life history while not being focused specifically on a particular species or population aggregation. More generally, this model structure would characterize any migratory life history in which: (1) individuals move between two distinct habitats over ontogeny, (2) density dependence affects the early life stages, (3) individuals vary in the timing of maturation, and (4) the species is iteroparous.

Since dispersal is the main process leading to metapopulation dynamics, the spatial component of this simulation describes immigration as a function of the distance among watersheds. I assumed that individuals immigrated into non-natal watersheds following an Isolation by Distance (IBD) model, which is a pattern thought to characterize population connectivity in multiple taxa (Guillot et al. 2009), including diadromous fishes (Palkovacs et al. 2013). This was accomplished through parameterizing a two-dimensional matrix of probabilities that an individual from a given population would enter any other population, where the probabilities are based on the relative distance among watersheds. To approximate a coastline, I assumed each watershed to be distributed along a line and separated by one unit of distance. Populations

1 and 2 are one unit of distance apart while populations 1 and 50 are 49 units apart. Note that the relative distances among rivers would be preserved regardless of the units that distance is given (e.g. m, km), so the probability of straying into a given watershed does not depend on knowledge of the distance units.

The implicit assumptions about population dynamics in most metapopulation models (i.e., habitat area is a proxy for population size; populations increase quickly to carrying capacity; source populations are not negatively affected by emigration) are all related to a population's ability to grow in size, its productivity. A useful metric to describe productivity or population growth over a species' life cycle is the maximum lifetime reproductive rate (MLR; Myers et al. 1999): the maximum rate at which spawners can produce spawners at low population size. The MLR is related to the idea of the maximum per capita growth rate ($r_{max;}$ Lande et al. 2003) but accounts for age structure in the calculation. For closed populations, it equates to the slope at the origin of a spawner-recruit relationship (α: the maximum rate at which spawners produce recruits; units: recruit/spawner) multiplied by the rate at which recruits produce spawners over their lifetime (the spawner-per-recruit relationship; SPR; units: spawner/recruit) (Myers et al. 1999, Gibson & Myers 2003). As such, the MLR is dimensionless and applies to any life stage (e.g., lifetime egg production per egg; number of spawners produced per spawner throughout its life). Here, smolts are defined as recruits given that compensatory density dependence takes place in fresh water (Gibson 2006, Jonsson et al. 1998). For open populations (i.e., a metapopulation), the SPR relationship for a particular river would be affected by straying, in that the intercept could be shifted away from zero, representing a situation in which spawners could be present even if there were no natal recruits. Differences in the SPR among populations could occur provided at least one of the following differed: (1) stray rates, (2) immigration rates or (3) life history parameters. Even if none of these varied, changes to the intercept of the SPR relationship for individual open populations would be time-varying because the relative contribution of strays to each population would vary over time. However, for the entire metapopulation assemblage, the MLR would not vary over time and would equal that for the isolated populations (Table A2-1-2, Appendix 2-1), given that no mortality was associated with straying in the simulation. Thus, the productivity scenarios for metapopulations and isolated populations were comparable at the level of the population assemblage, but not at the level of the individual populations in the simulations.

I considered five productivity scenarios: High, Medium, Low, Extreme Low and Not Viable, representing progressive declines in populations' ability to grow in size (the MLR) as well as in the equilibrium size of adult populations in small and large watersheds (Table 2-1). Deterministically, the High scenario corresponds to a situation in which the population assemblage can easily grow in size, parameterized such that one spawner can produce a maximum of seven spawners throughout its lifetime at low population size (MLR = 7 spawners/spawner). In contrast, the Not Viable scenario corresponds to a situation in which the population assemblage has no ability to increase in size, given that one spawner can produce a maximum of one spawner throughout its lifetime (MLR = 1 spawner/spawner). The Medium, Low, and Extreme Low scenarios have progressively lower MLR values between these two endpoints (Table 2-1). To aid in the evaluation of source-sink dynamics, I incorporated a quasi-extinction threshold into the simulation model to ensure that small populations acted as extreme sinks in the Not Viable scenario and to prevent biologically unrealistic population sizes (e.g., fractions of an individual) persisting in the projections. If total abundance in any given year was less than 20 spawners, I assumed future production from the cohort was zero; thus the Not Viable scenario leads to deterministic extinction for small isolated populations given that their equilibirum size is lower than the assumed quasi-extinction threshold. The graphical representation of each of these scenarios in the absence of environmental variation (Figure 2-1) shows the deterministic population sizes for small and large populations and how they decline with changes in productivity.

To evaluate if metapopulation structure benefits a population assemblage as productivity declines, I summarized simulations in terms of median abundance (and 25th and 75th quantiles), extinction probabilities, as well as percent difference among trajectories. At the level of the population assemblage, abundance becomes the sum of individual median population sizes in a given year and extinction probability is the probability that all populations are extinct in a given year. In addition, I quantified differences in the overall abundance trajectories of metapopulations relative to isolated populations by calculating a percent difference for each scenario. I subtracted the isolated population trajectory from the metapopulation trajectory, resulting in a 100 year-long vector of differences, before taking the maximum value from this vector, dividing by median abundance in the metapopulation, and multiplying by 100 to get a percentage (%). This standardization gives a comparable metric for each level of productivity, one that is not affected by the rate of population decline or by differences in absolute abundance. Positive values represent situations in which abundance was higher in metapopulations than in isolated populations, and vice versa for negative values.

To determine which populations were contributing to abundance or extinction trajectories of the population assemblage, I calculated median abundance at the level of individual populations, as well as the proportion of individual populations persisting (the inverse of extinction probability). Looking at the dynamics of individual watersheds was useful to determine which populations were ultimately contributing to the overall patterns described for each assemblage in terms of abundance or extinction, as well as to evaluate the potential for the rescue effect. I summarized individual watershed trajectories one of two ways, either by size (small or large) or by “fate” (extinct, recolonized or persisting). Separately for small and large, I calculated the proportion of individual populations that were persisting (the number of non-extinct populations at a given time step, divided by the original number of populations of that size: P) at decadal intervals for each scenario. I used a log linear model (Dick 2004) to calculate a rate of change in the proportion of populations of a particular size persisting per year. I used the log-transformed proportion as the response and the year of the simulation as the predictor. The slope coefficient gives the strength and direction of change in the proportion persisting per year. To evaluate if the rescue effect was likely to facilitate persistence as productivity declined, I compared the abundance trajectories of populations that had different fates in the projections, standardized relative to their starting equilibrium size. If the trajectory remained centered around one, populations fluctuated around their equilibrium, both above and below. If the trajectory was substantially lower than one, populations had reduced capacity to grow in size following stochastic declines. This low resiliency would be related to how strongly individual populations would act as sinks in the projections.

To ensure that my conclusions were not overly sensitive to the parameterization of the simulation model, I evaluated multiple alternatives and compared results to those from the parameterization described above. I have presented background information on eight alternative hypotheses as well as specifics on the resulting 11 alternate parameterizations of the simulation model in Appendix 2-2. In brief, I considered: (1) higher stray rates, (2) two different spatial models of immigration, (3) increased or decreased contrast in watershed size, (4) changes in the relative geographic position of watersheds, (5) less temporal autocorrelation in environmental variability, (6) changes to productivity in the freshwater or marine environment only, (7) a lower quasi-extinction threshold, and (8) a low mortality rate on strays. Again, I calculated the percent difference between metapopulations and isolated populations to quantify any benefit of metapopulation structure. One thing to note is that the alternate parameterizations had the potential to change the range of productivities over which metapopulation structure was beneficial, which would be reflected by zero values for the higher productivity scenarios in particular.

Results

Does metapopulation structure benefit a population assemblage as productivity declines?

Comparing metapopulations and isolated populations in terms of total median abundance (i.e., the sum of median abundances in individual populations at each time-step) suggests that straying has a differential effect on population trajectories as productivity declines. For the High scenario, there was little difference (a maximum of 9%) between the abundance trajectories (Figure 2-2), indicating that metapopulation structure was unlikely to substantially increase abundance when populations were very productive. However, when productivity declined by essentially half, total abundance was progressively greater in a metapopulation than it was for isolated populations (*c.f.* the High, Medium, and Low productivity scenarios; Figure 2-2). However, further declines in productivity erode this benefit, where metapopulations and isolated populations had a more similar abundance trajectory in the Extreme Low scenario, and isolated populations remained larger than metapopulations in the Not Viable scenario (Figure 2-2). Comparing the percent difference between the two trajectories as productivity declines results in a dome-shaped pattern (Figure 2-3), showing the increasing positive effect of metapopulation structure on abundance from the High to Low scenarios (maximum differences of 9%, 18% and 32%, respectively), the partial erosion of this benefit in the Extreme Low scenario (27%), and the reversal in the Not Viable scenario (-26%). Intuitively, this pattern suggests that the benefit to small populations from immigration counter-balances the negative effect of strays from source populations only at higher levels of population productivity. However, total extinction probability (i.e., the probability that all populations are extinct at a specific time) does not show any benefit from metapopulation structure, in that probabilities are essentially equal for the High and Medium scenarios, and are lower for isolated populations in the Low, Extreme Low, and Not Viable scenarios (Figure 2-4). Calculating the maximum difference in extinction probability between metapopulations and isolated populations results in progressively more negative values; -3%, -10%, -18%, and -19% from the High to Extreme Low scenarios, respectively. Although this may seem to contradict the patterns reported above for total abundance, it likely relates to which specific populations persisted in the simulations over time.

The rate at which individual small populations go extinct in the projections was measurably higher for isolated populations than metapopulations when productivity declined from the Medium to the Not Viable scenario (Figure 2-5), as was also shown by the slopes coefficients from the log linear model (*c.f.* -0.001 and 0; -0.004 and -0.002; -0.014 and -0.009 year^{-1} respectively, Table 2-2). For this to be true, the number of individuals immigrating had to have been greater than the 10% emigrating, leading to larger sizes as compared to isolated small populations. However, the reverse is true for large populations, in that annual emigration from a specific large population (10% of population size) is not balanced by immigration from the other 49 populations, leading to smaller population sizes as compared to isolated populations. Therefore, extinction risk is lower for small populations and higher for large populations in the metapopulation at any level of productivity (Figure 2-5; Table 2-2). Given that the differences in total extinction probability are progressively more negative, persistence of the population assemblage must be primarily linked to the persistence of large populations. If total extinction risk for the population assemblage was primarily determined by the persistence of small populations, metapopulations would have had a lower overall extinction probability at some or all levels of productivity. The fact that all differences are negative also demonstrates that straying would always be expected to have a measurable negative effect on large (source) populations.

By increasing the persistence of small populations, metapopulation structure can lead to substantial increases in total abundance as compared to isolated populations when productivity is moderately high (up to 32% in the Low scenario). However, as productivity drops, small populations tend to go extinct quickly in the simulations for both isolated populations and metapopulations (Extreme Low and Not Viable scenarios; Figure 2-5). In the Extreme Low scenario, this reduces the beneficial effect of metapopulation structure on total abundance and makes the trajectories more similar (maximum difference of 27%). In the Not Viable scenario, the difference between the abundance trajectories becomes strongly negative (-26%). Numerically this would mean:

(1) that straying into small populations is not sufficient to ensure that abundance remains above the quasi-extinction threshold and that small populations are lost at a similar rate to isolated ones, and (2) large populations are losing slightly less than the percentage of the population that emigrates in a given year by acting as sources, leading to faster declines relative to isolated populations.

Is the rescue effect likely to facilitate persistence as productivity declines?

Individual populations have one of two fates when isolated: they either persist in the projections or go extinct. Populations that persist in the projections tend to fluctuate around their equilibrium size for the High to Extreme Low productivity scenarios, while populations that go extinct do so relatively rapidly, within a median timeframe of 50 or fewer years (all scenarios; Figure 2-6). All populations experience autocorrelated variation in vital rates which would lead to annual variability in productivity relative to the deterministic MLR of the population aggregation (Appendix 2-1). Therefore, in the Not Viable scenario, populations can persist if they experience favorable environmental conditions (i.e., negative deviates on at-sea mortality rates coupled with positive deviates around the spawner-recruit relationship) enabling them to fluctuate above their equilibrium size. In the metapopulation simulations, the individual trajectories that persist or go extinct are very similar those of isolated populations (*c.f.* the red and black lines in the High to Extreme Low scenarios, Figure 2-6, Figure 2-7). However, looking at the trajectories of recolonized populations only (i.e., trajectories that have declined to zero and subsequently increased over the duration of the simulation) shows that they tend to persist at progressively lower sizes relative to their equilibrium as productivity declines (*c.f.* the blue and black lines among scenarios in Figure 2-7). This demonstrates that populations have progressively less ability to increase in size; in other words, they have progressively lower resiliency to environmental variability. Once this capacity to increase in size reaches a critical value, which would be at a productivity between the Extreme Low and Not Viable scenarios considered here (i.e., around the zero-intercept in the top panel of Figure 2-3), any benefit to metapopulation structure in terms of total abundance is eliminated.

Sensitivity analyses

The sensitivity analyses included simulations parameterized with: (1) higher stray rates, (2) two different spatial models of immigration, (3) increased or decreased contrast in watershed size, (4) changes in the relative geographic position of watersheds, (5) less temporal autocorrelation in environmental variability, (6) changes to productivity in the freshwater or marine environment only, (7) a lower quasi-extinction threshold, and (8) a low mortality rate on strays (Appendix 2-2). Again, I summarized the results relative to the maximum percent difference between the overall abundance trajectories for isolated vs. metapopulations; where I took deviation from a dome-shaped pattern as evidence that these results were sensitive to the parameterization of the simulation. Because several of the alternate parameterizations changed the range of productivities over which metapopulation structure was beneficial, some of the positive deviations peaked in different scenarios (Figure 2-8). For example, when stray rates were 30%, the simulations revealed essentially no benefit to metapopulation structure until the maximum lifetime reproductive rate was 3.6 (Low scenario). Then, the benefits changed rapidly, peaking in the Extreme Low scenario and becoming strongly negative in the Not Viable scenario (Figure 2-8). Negative deviations were not found before the Not Viable scenario in any of the sensitivity analyses, consistent with the results from the main simulation.

Discussion

By explicitly accounting for population dynamics in a spatial metapopulation model, I could characterize the conditions under which metapopulation structure would be expected to benefit a population assemblage. Overall, my analyses suggest that the effect of straying on abundance or extinction probability is highly dependent on underlying productivity. Furthermore, the impression that metapopulation structure, source-sink dynamics, and the rescue effect will always reduce extinction risk for a population assemblage is not supported by my analyses, where such structure becomes increasingly detrimental as the potential for populations to increase in size approaches zero (i.e., when the maximum lifetime reproductive rate approaches one). Relative to the assumption that populations increase quickly to carrying capacity when recolonized, my results suggest that this becomes more severely violated as productivity declines. Source-sink dynamics are unlikely to facilitate the persistence of the population assemblage when productivity is extremely low. Thus recovery plans that prioritize actions to maintain metapopulation structure at the expense of other objectives appear to be a poor basis for conservation of diadromous fishes. These simulations suggest that changes to population-specific vital rates have a much larger impact on overall extinction risk relative to spatial dynamics. Given that these conclusions are robust to a range of alternate ecological hypotheses, they might be expected to apply to other migratory species characterized by low productivity.

Consistent with expectations from previous research on metapopulations (e.g., Hanski & Gilpin 1991, Fronhofer et al. 2012, Smedbol et al. 2002), source-sink dynamics led to substantially higher total abundance when small populations exhibited two characteristics: (1) a non-negligible probability of extinction and (2) the ability to increase to sizes approaching carrying capacity. Extinction probabilities for small populations progressively increased in the High to Low scenarios and the majority of small populations were recolonized at some point in the time series. However, these populations still had sufficient ability to increase in size, leading to substantially greater abundance in the population assemblage. In contrast, overall extinction risk was tied directly to the dynamics of the large source populations. In these same productivity scenarios (High to Low), straying reduced population size and made large populations in metapopulations slightly more vulnerable to extinction. As productivity declined further (the Extreme Low scenario), the numerical benefit from recolonization in small populations was reduced by approximately 20% (*c.f.* 32% difference vs. 27% difference); source-sink dynamics became less beneficial to the population assemblage overall. In the Not Viable scenario, populations had extremely low resiliency to environmental stochasticity and essentially could not recover if abundance declined. For the metapopulation, straying from sources to sinks became doubly detrimental in that it simultaneously reduced population size in sources and did not result in viable populations in sinks, leading to swift declines to extinction.

A prevalent theme in ecology is that higher diversity (at the level of the community, species, population or individual) is expected to contribute substantially to long-term persistence (e.g., Whittaker et al. 2005, Bolnick et al. 2011, Schindler et al. 2010). Metapopulations are typically considered synonymous with diversity at the population level, so at first glance these results may appear contradictory to other well-established ecological theories. A good example is the portfolio effect, where theoretical simulation studies have shown that the complementary (i.e., partially non-synchronous) dynamics of multiple populations can reduce inter-annual variability in total abundance relative to a single large population (summarized in Figge 2004). For actual populations, this reduced variability is expected to lead to greater sustainability of exploited populations (Schindler et al. 2010) as well as to reduced extinction risk overall (Figge 2004, Moore et al. 2010). Recent research on threatened Atlantic salmon shows that declining diversity is positively correlated with population synchrony (Moore et al. 2010), which would be consistent with theoretical predictions that the portfolio effect gets weaker when components of the population aggregation are lost (Schindler et al. 2010). Although my simulations demonstrate how diversity may decline as population productivity is reduced, these results do not directly relate to the portfolio effect because the level of synchrony among populations as well as population diversity is the same in each scenario. In other words, these results do not suggest that metapopulation structure is disadvantageous because of its diversity; rather that straying becomes disadvantageous among diverse populations as productivity declines. In relation to population synchrony, it would be possible to vary synchrony in the simulation model by reducing σ (currently set at 0.75) in the freshwater deviates as abundance declines. My expectation from doing this would be that metapopulation structure (i.e., dispersal from sources to sinks) would become detrimental to persistence at higher levels of productivity (e.g., when MLR > 1 spawners/spawner), because extinction risk increases with synchrony (Moore et al. 2010, Mustin et al. 2013). However, the conclusion that declining productivity leads to a reduction in diversity (i.e., that a higher proportion of populations go extinct as productivity declines) would remain unchanged.

Metapopulation theory as related to extinction risk

If metapopulation theory were to be used as a basis for understanding extinction risk, the idea of population productivity must remain separate from the idea of population size. Although population size is recognized as being correlated with extinction risk within a given timeframe, the majority of ecological processes leading to extinction are not linked to population size *per se*, but rather to changes in population productivity or population vital rates with population size (Melbourne & Hastings 2008). Small or newly established populations do not have large sizes that limit inbreeding, any Allee effects or mal-adaptation to the new environment, or buffer against catastrophic events (Simberloff 1998, Drake & Lodge 2006, Kinnison & Hairston 2007). Of these processes, only specific types of catastrophic events could realistically influence population size in isolation from demographic rates; for example, poaching would reduce abundance without changing environmental characteristics linked to mortality or maturity rates. Demographic Allee effects (also called depensation or depensatory density-dependence in the fisheries literature; e.g., Myers et al. 1995, Courchamp et al. 2008) are commonly understood to be a decline in per capita population growth rates or in individual fitness as population sizes decline (i.e., reduced MLR). Similarly, genetic effects such as inbreeding or mal-adaptation are defined in terms of changes to fitness, where fitness is generally understood to describe the ability of an individual, population, or species to survive and produce offspring (Oor 2009). Although such processes would be expected to have a proportionately larger influence on fitness when populations are small as when they are large, it does not necessarily follow that productivity is high when population sizes are larger or that productive small populations have an inherently higher extinction risk than large populations (Lande 1993). Simulated population sizes in excess of 80 times carrying capacity went extinct when productivity was extremely low (Not Viable scenario), while the extinction probability for individual populations within a metapopulation was essentially zero when productivity was high for these same watersheds. If metapopulation theory was derived relative to population productivity (e.g., Baguette & Schtickzelle 2003) rather than the geographical size of habitat patches (Driscoll 2007, Pellet et al. 2007), it is likely that it would more accurately characterize extinction dynamics and thus be much more useful when applied to conservation questions.

It is important to note that a higher probability of persistence of individual populations in the metapopulation projections does not indicate that extinction risk for a specific population is lower. Extinction and recolonization of specific watersheds occurred frequently in the metapopulation simulations, leading to a variable group of watersheds persisting in any given year. Consistent with the majority of theoretical metapopulation models, I treated both habitats as well as populations as interchangeable in these simulations, even though real diadromous species can exhibit substantial life history variation among populations within a region (Garcia de Leaniz et al. 2007, Palkovacs et al. 2013). Actual survival and spawning success (i.e., fitness) can be a function of the difference between populations, where individuals are expected to have higher mean fitness in watersheds similar to their natal watershed owing to local adaptation (Fraser et al. 2011, Pfluger & Balkenhol 2014). Accounting for such heterogeneity among watersheds would be expected to reduce median abundance as well as the probability of persistence of small populations, because individuals in recolonized populations would have progressively lower mean fitness relative to that in populations that have never gone extinct. In other words, metapopulation structure and the rescue effect would be less beneficial to sink populations than shown here, even at moderate levels of productivity. On the other hand, straying is also expected to maintain or increase genetic variability and thus adaptive potential, leading to populations that may be better able to respond to environmental change (e.g., Burke & Arnold 2001). Such hybridization would act in opposition to any mal-adaptation to the recipient environment. Furthermore, a recent study by Fraser et al. (2014) indirectly suggests that productivity may increase in small populations relative to large owing to increased environmental variability maintaining greater genetic polymorphism in these small populations. Both of these effects would increase the overall viability of sink populations, and the rescue effect could be beneficial at lower levels of productivity (e.g., potentially in the Extreme Low scenario). Unfortunately, any increase to productivity in small populations relative to large, or any fitness benefits from straying, would have to be quite sizable to compensate for numerical declines in the population assemblage and to reduce extinction risk when productivity is low.

Practical implications

To put these results in context, it is useful to consider their implications for conservation of an actual population assemblage. Atlantic salmon (*Salmo salar*) belonging to the Southern Upland (SU) region of Nova Scotia are an identified population assemblage containing 72 individual populations. Collectively, they have undergone extremely large declines in abundance over the last 40 years, and are presently considered to be endangered (Gibson et al. 2011). Population dynamics modeling from two populations suggests that maximum lifetime reproductive rates have declined from approximately 2.5 in the 1980s (similar to the Extreme Low scenario in this chapter) to essentially 1 in the 2000s (similar to the Not Viable scenario in this chapter), indicating extremely low and declining productivity (Gibson & Bowlby 2013). From assessment monitoring and historical tagging (Bowlby et al. 2013), straying has been observed among populations, so the region would be expected to contain a metapopulation. Based on the simulations in this chapter, such metapopulation structure would have been expected to have benefitted individual populations by increasing persistence during the 1980s for the smaller watersheds within the SU. Today, this same metapopulation structure could be acting as an additional source of mortality for the remaining populations (thought to be those inhabiting the larger watersheds), hastening overall population decline. In terms of recovery actions, maintaining a more widespread distribution throughout the SU would have been expected to promote the rescue effect in the 1980s, so concurrent small improvements to multiple watersheds (i.e., actions leading to small gains in productivity for many watersheds) were warranted. Today, these simulations suggest that it would be much more beneficial to focus on substantial increases in productivity to a limited number of populations (up to and including a single large watershed) in order to ensure that productivity remains high enough to prevent total extinction of the SU metapopulation. In addition, the propensity to stray could now be considered largely synonymous with mortality, because individuals that stray have little to no subsequent reproductive output. Our perception of harm from activities or threats that increase straying rates, such as chemical contaminants reducing olfactory imprinting and homing success (Lurling & Scheffer 2007) may need to be higher. In addition, actions to expand distributions, such as ensuring fish passage at barrier structures or remediating watershed fragmentation, could be given lower priority in that they would not be expected to benefit the remaining individuals as much as increasing productivity in the habitat currently occupied. Adult abundance is currently so low that any density dependent constraints on freshwater productivity would be expected to be minimal. However, it is a practical reality that remediating the threats linked to declines in productivity (e.g., land use practices; Allan 2004, Foster et al. 2003) is much more difficult than those that limit distribution (e.g., fragmentation from road crossings) because their effects on the ecological structure and functioning of rivers are gradual, largely indirect and can be cumulative (Lake et al. 2007, Rose 2000).

The main debate surrounding metapopulation theory as applied to single species has centered on the question of whether or not metapopulation structure exists in nature and can be empirically described (e.g., Fronhofer et al. 2012). There has been comparatively little debate related to the validity of theoretical assumptions as applied to real species (although see Dennis et al. 2003, Clinchy et al. 2002, and Hanski 2004), leading to population connectivity being perceived as unilaterally beneficial by conservation biologists and wildlife managers (Eaton et al. 2014). By using a spatially-explicit population dynamics model of a commonplace life history strategy (diadromous fishes), I demonstrate how metapopulation structure would be expected to first benefit and then harm population assemblages as productivity declines. By comparing the simulation results to the characteristics of an actual population assemblage, I demonstrate why our current approach to recovery planning may need to be re-evaluated. Although the conclusions from my analyses follow very logically from an understanding of population dynamics, they have substantial implications for how we perceive population connectivity, source-sink processes, and the rescue effect for species of conservation concern.

Tables

Table 2-1. Description of the scenarios considered in the simulations for metapopulations or isolated population assemblages, detailing the number of populations (*N*), starting population size (for small and large populations), the slope of the origin of the spawner-recruit relationship (α), mortality rates for immature (M^{Sea}) and mature (M^{Adult}) adults, the stray rate (*s*), the spawner-per-recruit relationship (SPR), and the maximum lifetime reproductive rate in spawners/spawner (MLR = α *SPR) corresponding to each. Scenarios were assessed from 500 simulations of 50 populations (40 small, 10 large) over 100 years.

Scenario	Productivity	Type	N	small	large	α	M^{Sea}	M^{Adult}	s	SPR	MLR
1	High	metapop	50	483	2413	10	0.5	0.5	0.1	0.703	7.0
2	High	isolated	50	483	2413	10	0.5	0.5	0	0.703	7.0
3	Medium	metapop	50	366	1831	9	0.55	0.55	0.1	0.569	5.0
4	Medium	isolated	50	366	1831	9	0.55	0.55	0	0.569	5.0
5	Low	metapop	50	264	1319	8	0.6	0.6	0.1	0.455	3.6
6	Low	isolated	50	264	1319	8	0.6	0.6	0	0.455	3.6
7	Extreme Low	metapop	50	172	861	7	0.65	0.65	0.1	0.358	2.5
8	Extreme Low	isolated	50	172	861	7	0.65	0.65	0	0.358	2.5
9	Not Viable	metapop	50	6	30	5	0.75	0.75	0.1	0.208	1.0
10	Not Viable	isolated	50	6	30	5	0.75	0.75	0	0.208	1.0

Table 2-2. A comparison of the mean extinction percentages (%) as productivity declines for large (N = 10) or small (N = 40) watersheds at decadal intervals, as well as the slope of a log-linear regression representing the change in the proportion of populations persisting per year (P/year) (data plotted in Figure 2-5).

			Year of the simulation										Log-linear slope
Size	Productivity	Type	10	20	30	40	50	60	70	80	90	100	
Large	High	Metapop	0	0	0	0.4	0.4	0.4	0.6	0.6	0.6	0.6	0.000
Large	High	Isolated	0	0	0	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.000
Large	Medium	Metapop	0	0.2	0.2	0.6	0.8	1	1.2	1.4	1.4	1.6	0.000
Large	Medium	Isolated	0	0.2	0.2	0.6	0.6	0.8	1	1	1	1.2	0.000
Large	Low	Metapop	0	1.6	3.2	5.8	6.4	8.2	9.8	10.2	10.4	10.7	-0.001
Large	Low	Isolated	0	1.2	2.6	4.2	4.8	5.4	6.6	7.6	8	8.2	-0.001
Large	Extreme Low	Metapop	0	5.8	10.6	15.6	19.8	25	29.8	33.2	37.6	41.2	-0.006
Large	Extreme Low	Isolated	0	4.8	9.6	13.8	17	21.8	25.8	29.6	34	36.4	-0.005
Large	Not Viable	Metapop	24.2	55.6	74.2	82.6	88.4	93.6	96.8	97.8	98.8	99.6	-0.052
Large	Not Viable	Isolated	21.6	49.8	66.6	77.4	85.8	91.8	95	96.8	98.2	98.6	-0.046
Small	High	Metapop	0	0.2	0.2	0.6	0.4	0.6	0.6	0.6	0.6	0.8	0.000
Small	High	Isolated	0	0.2	0.2	0.6	0.8	1.0	1.2	1.2	1.2	1.4	0.000

Size	Productivity	Type	Year of the simulation										Log-linear slope
			10	20	30	40	50	60	70	80	90	100	
Small	Medium	Metapop	0	1.0	1.4	2.0	1.8	2.6	2.6	2.6	2.4	2.4	0.000
Small	Medium	Isolated	0	1.2	2.2	3.2	3.6	4.4	5.6	6.4	7.0	7.2	-0.001
Small	Low	Metapop	0	4.2	7.8	10.4	11.4	12.6	15.2	15.8	16.4	16.6	-0.002
Small	Low	Isolated	0	4.6	8.8	12.6	15.8	19	23.2	26.4	30.8	32.6	-0.004
Small	Extreme Low	Metapop	1.2	12.2	19.6	24.6	30.6	37.6	41.6	48.0	50.8	55.2	-0.009
Small	Extreme Low	Isolated	1.4	12.8	22.2	31.0	38.8	48.4	55.4	61.4	66.2	72.2	-0.014
Small	Not Viable	Metapop	97.8	93.4	94.4	96.8	97.8	98.4	99.4	100	100	100	NA
Small	Not Viable	Isolated	100	100	100	100	100	100	100	100	100	100	NA

Figures

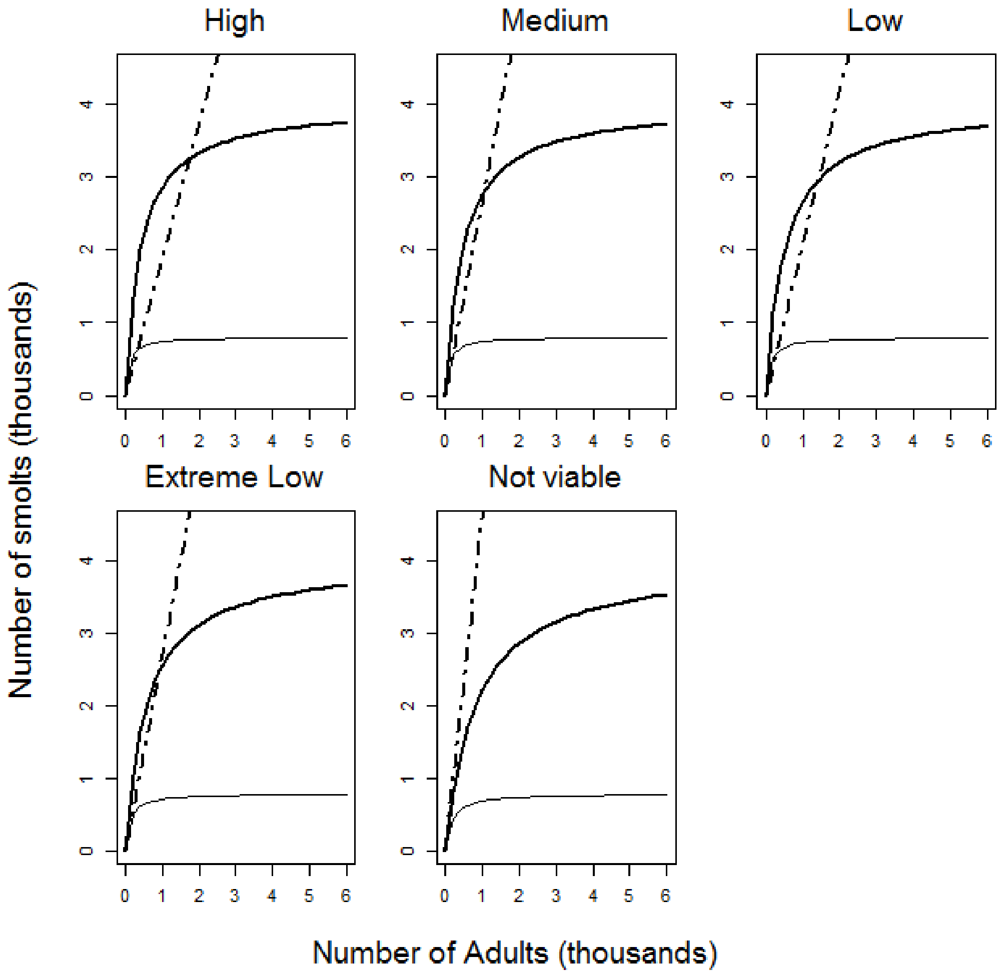

Figure 2-1. Schematic of the productivity scenarios considered in the metapopulation simulations. For each panel, freshwater production for large (thick solid line) and small (thin solid line) isolated populations is shown relative to the replacement line that represents marine survival (dashed line). The point at which the lines intersect represents the deterministic population size (number of smolts or number of adults) for large and small populations; note that it is essentially zero in the last scenario. See Table 2-1 for a description of scenarios.

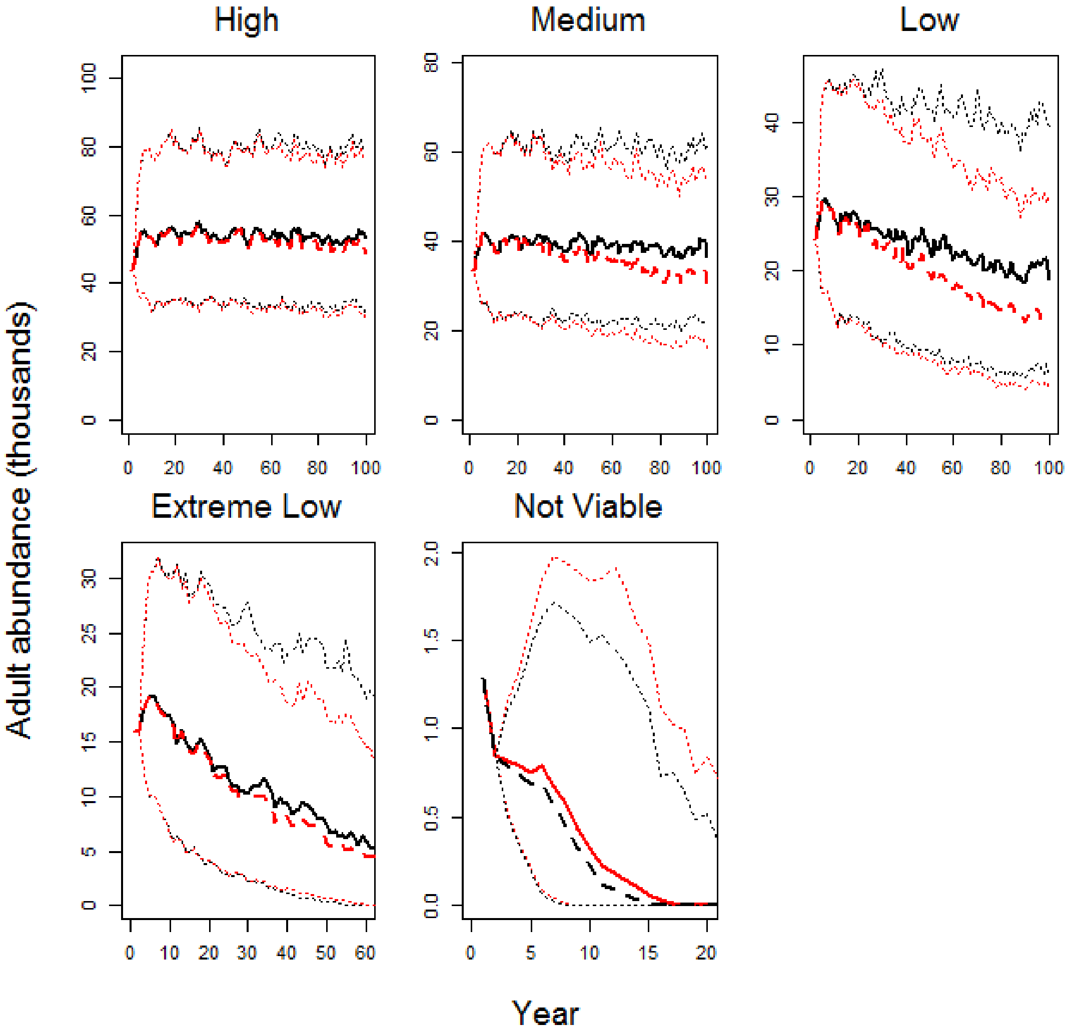

Figure 2-2. Median spawner abundance of the population assemblage (i.e., sum of individual population abundances) assuming metapopulation structure (solid black line) with 25th and 75th quantiles (dotted black lines) compared to isolated populations (dashed red line) with 25th and 75th quantiles (dotted red lines) for five productivity scenarios. Each scenario was assessed from 500 simulations of 50 populations over 100 years. Note the differences in the x- and y-axes among scenarios.

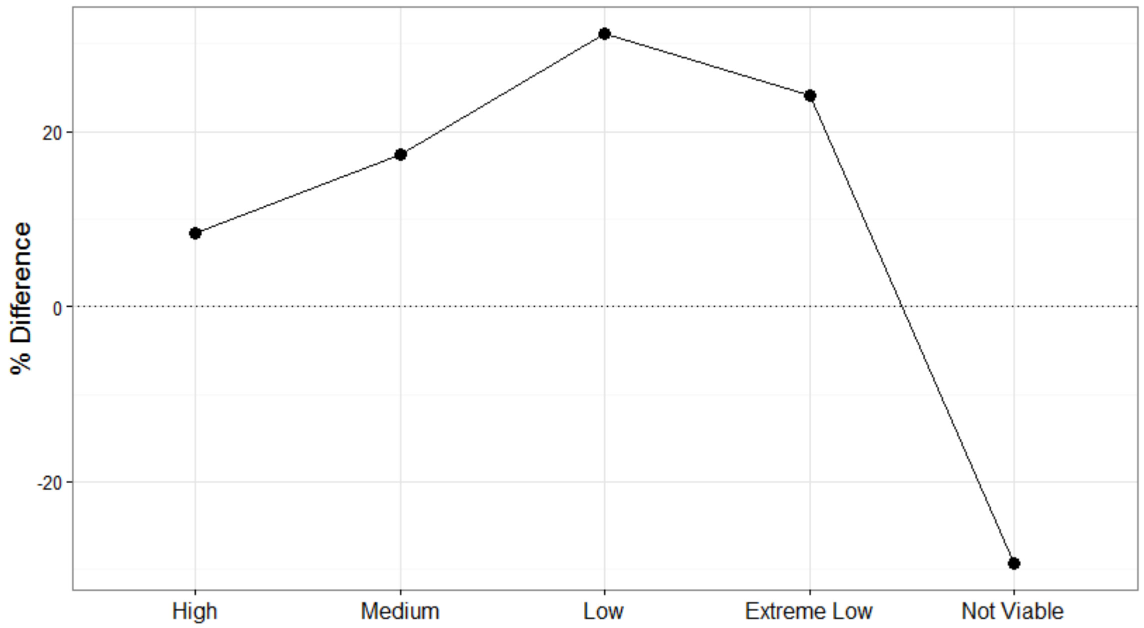

Figure 2-3. Maximum percent difference between total spawner abundance in a metapopulation versus in isolated populations for five productivity scenarios (described in Table 2-1). Metapopulation structure led to higher abundance relative to isolated populations when values were positive, which occurred in all scenarios except Not Viable for all parameterizations.

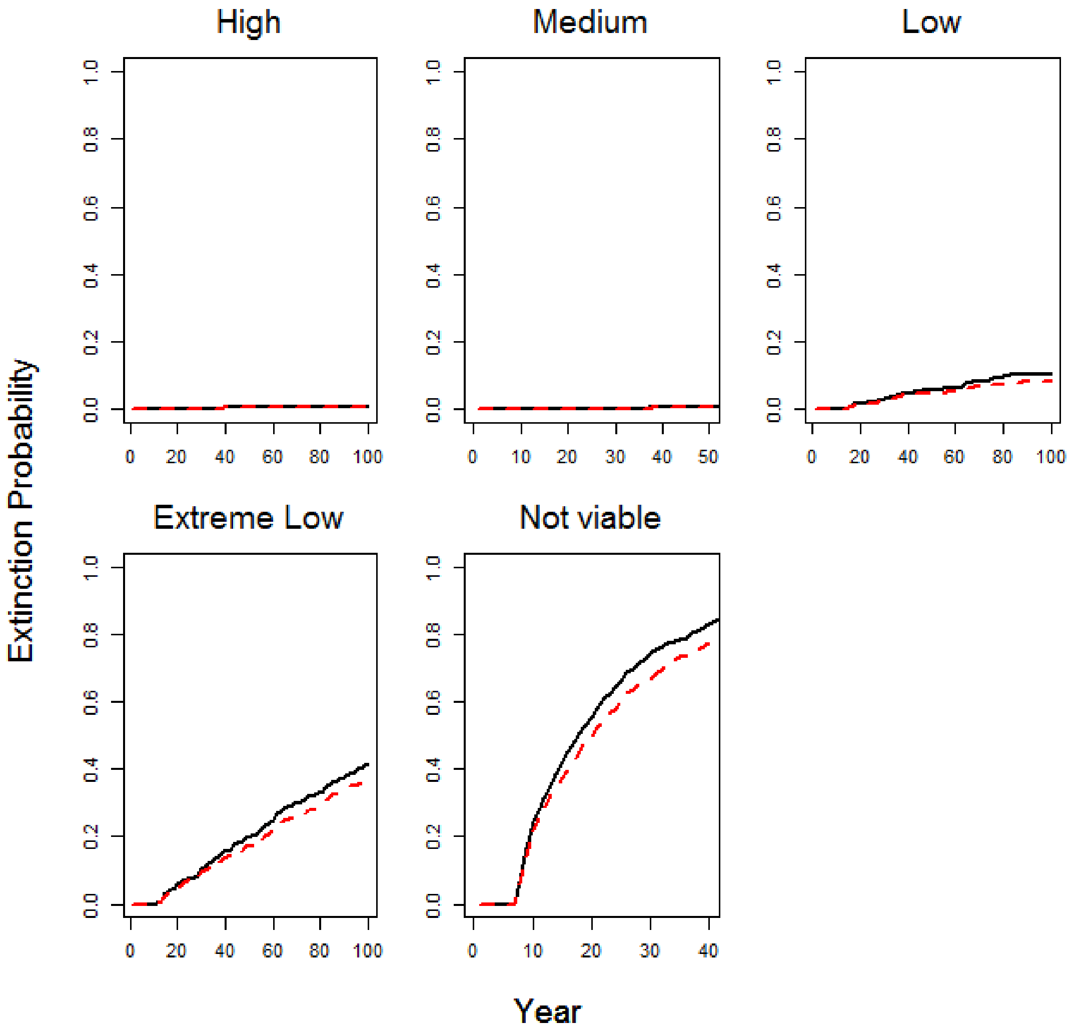

Figure 2-4. Probability of extinction for the population assemblage (i.e., all individual populations are extinct) per year as productivity declines, assuming metapopulation structure (black line) or isolated populations (red dashed line). Each scenario was assessed from 500 simulations of 50 populations (*N* = 25,000) over 100 years.

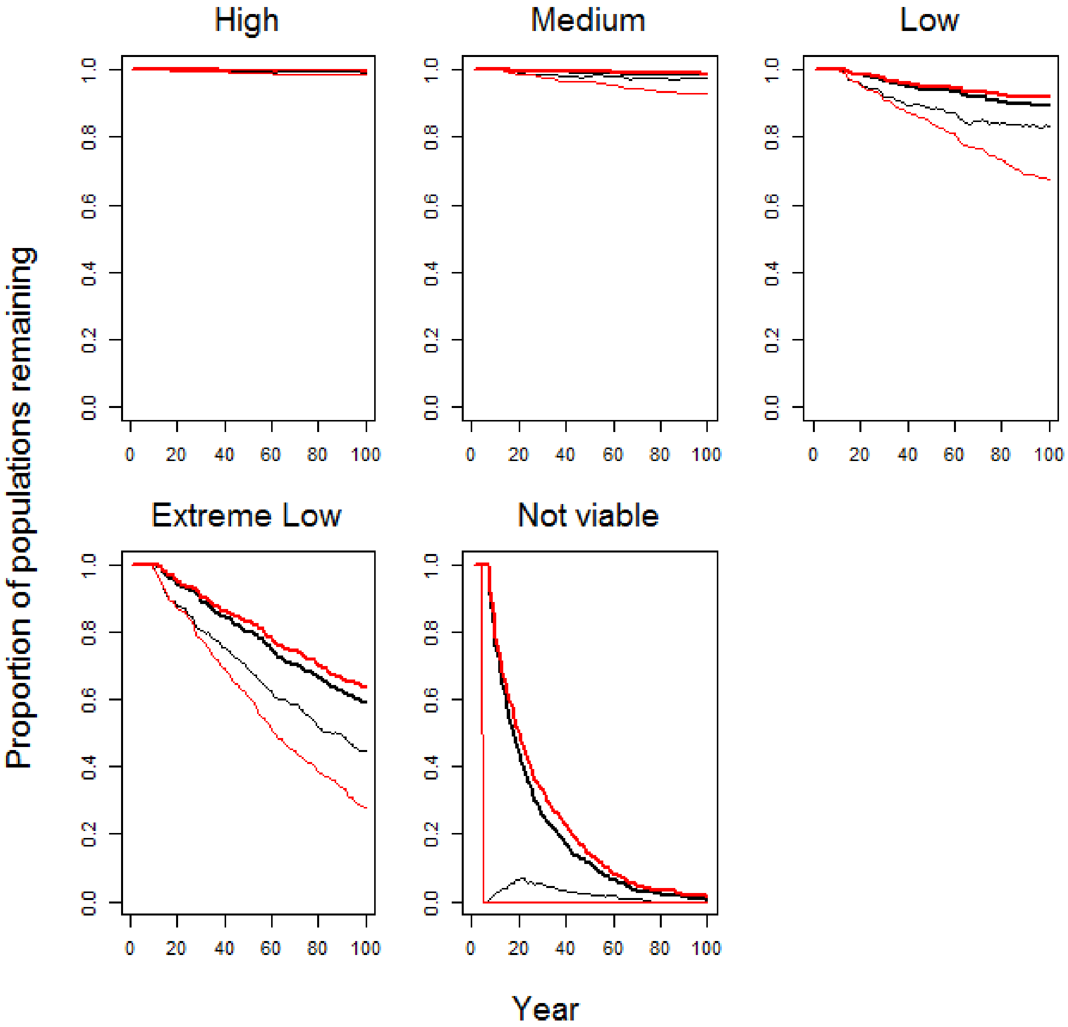

Figure 2-5. Proportion of populations persisting per year in the simulations by size, for large (thick black line) and small (thin black line) populations in a metapopulation, as well as large (thick red line) and small (thin red line) isolated populations. Each scenario was assessed from 500 simulations of 50 populations (*N* = 25,000) over 100 years. Proportions were calculated as the number of populations with non-zero abundance in a given year divided by the original number of populations of that size.

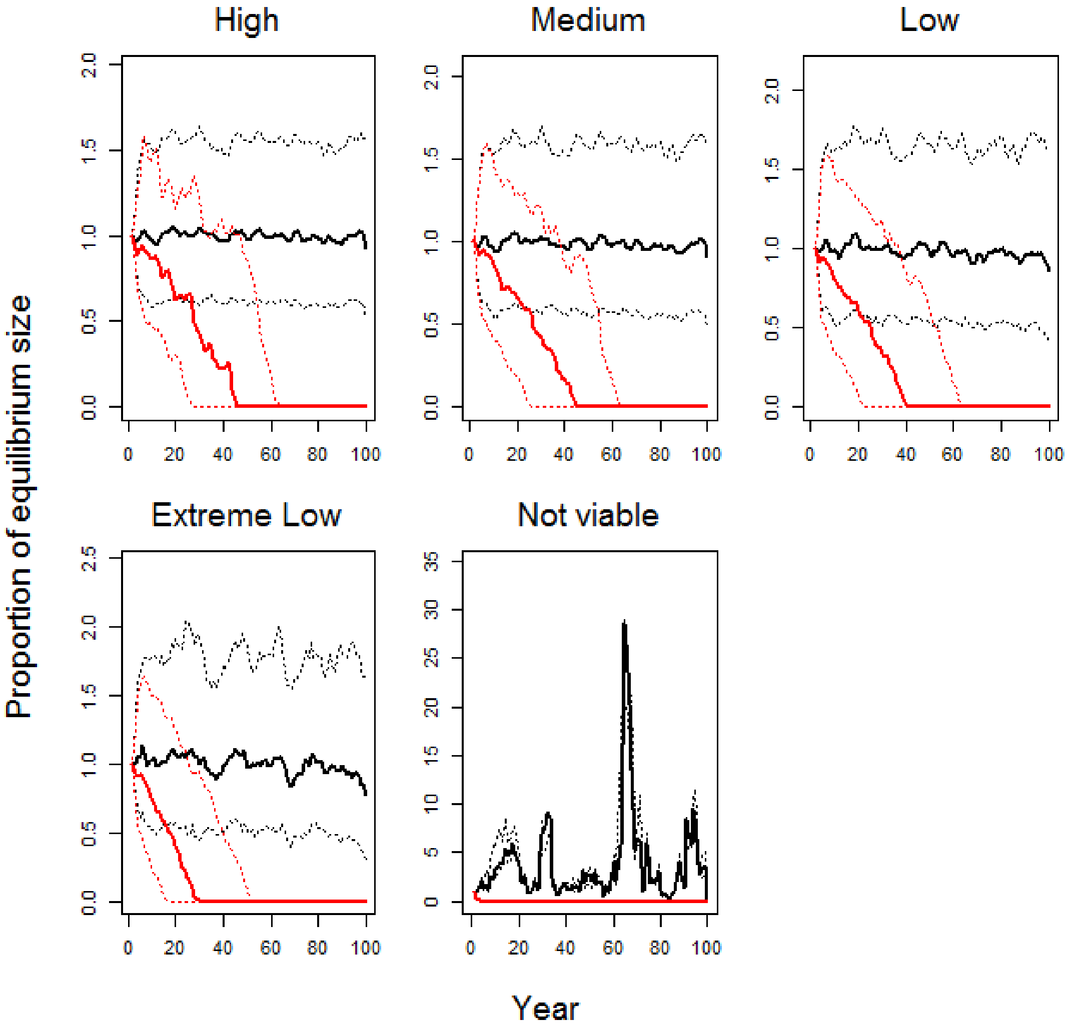

Figure 2-6. Median abundance of spawners as a proportion of equilibrium size (plus 25th and 75th quantiles) for individual isolated populations that had one of two fates: extinction (solid plus dotted red lines) or persistence (solid plus dotted black lines). Note that populations persisting in the Not Viable scenario would have done so because of environmental stochasticity rather than input from straying.

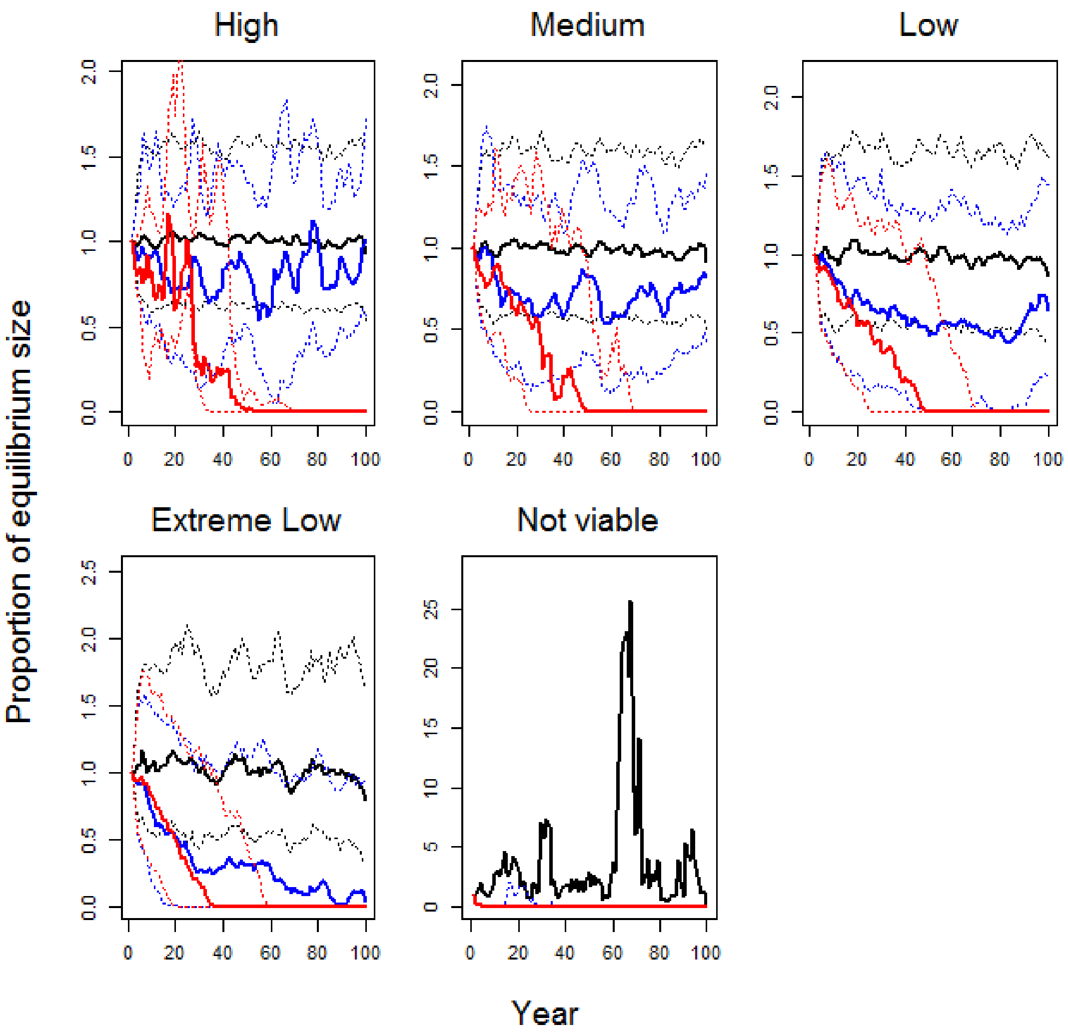

Figure 2-7. Median spawner abundance as a proportion of equilibrium size (plus 25th and 75th quantiles) for populations that had one of three fates in a simulated metapopulation: extinction (solid plus dotted red lines), recolonization (solid plus dotted blue lines) or persistence (solid plus dotted black lines). Note that the combined effects of environmental variation and straying would contribute to sizes in excess of the equilibrium for the populations persisting in the Not Viable scenario.

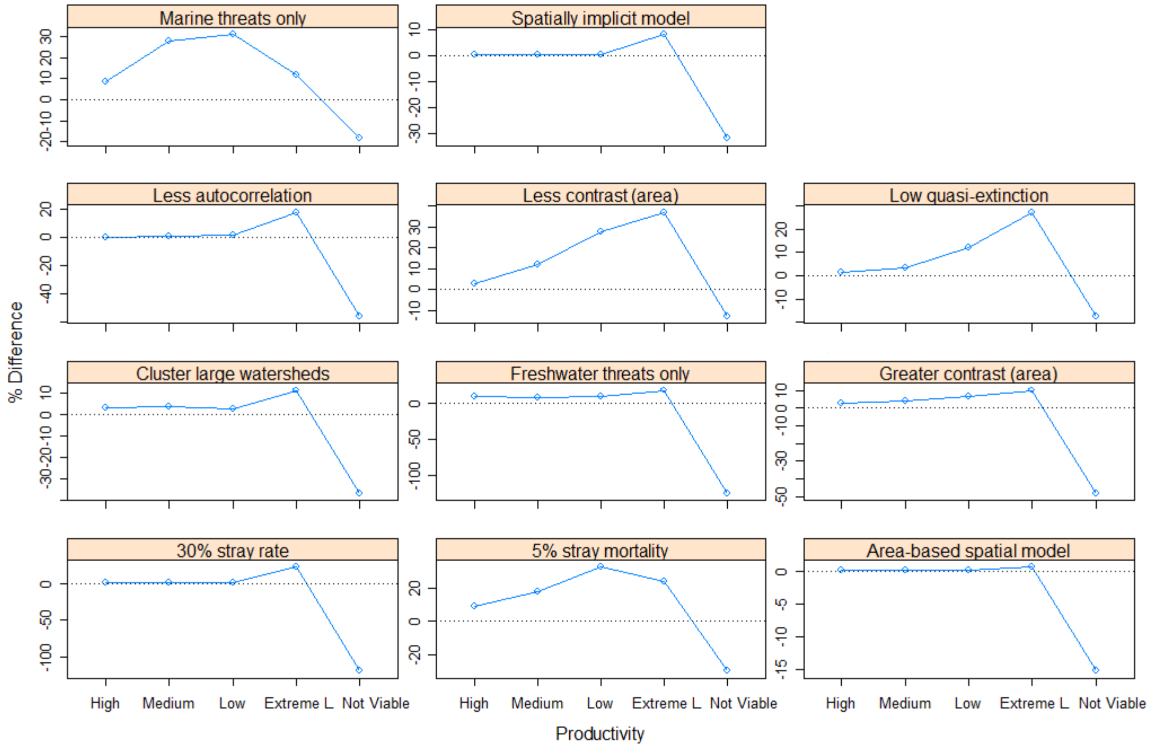

Figure 2-8. Maximum percent difference between total spawner abundance in a metapopulation and isolated populations for five productivity scenarios using 11 alternate parameterizations for the simulation model (described in Appendix 2-2). Metapopulation structure led to higher abundance relative to isolated populations when values are positive. When values remain on zero, there is no benefit to straying at that level of productivity, indicating that the parameterization changes the range of productivities over which metapopulation structure leads to higher abundance.

Chapter three: Environmental effects on survival rates: robust regression, recovery planning and endangered Atlantic salmon

Abstract

Describing how population-level survival rates are influenced by environmental change becomes necessary during recovery planning to identify threats that should be the focus for future remediation efforts. However, the ways in which data are analyzed have the potential to change our ecological understanding and thus subsequent recommendations for remedial actions to address threats. In regression, distributional assumptions underlying short time series of survival estimates cannot be investigated *a priori* and data likely contain points that do not follow the general trend (outliers) as well as contain additional variation relative to an assumed distribution (overdispersion). Using juvenile survival data from three endangered Atlantic salmon *Salmo salar* L. populations in response to hydrological variation, I compared four distributions for the response in lognormal and Generalized Linear Models (GLM). I investigated the influence of outliers as well as overdispersion by comparing conclusions from robust regressions with these lognormal models and GLMs. My analyses strongly supported the use of a lognormal distribution for survival estimates (i.e. modeling the instantaneous rate of mortality as the response), and would have led to ambiguity in the identification of significant hydrological predictors as well as low overall confidence in the importance of the

relationships if I had only considered GLMs. However, using robust regression to evaluate the effect of additional variation and outliers in the data relative to regression assumptions resulted in a better understanding of relationships between hydrological variables and survival that could be used for population-specific recovery planning. This chapter highlights how a systematic analysis that explicitly considers what monitoring data represents and where variation is likely to come from is required in order to draw meaningful conclusions when analyzing changes in survival relative to environmental variation to aid in recovery planning.

Introduction

Effective conservation of endangered species and the development of successful recovery plans rely on the identification of environmental and ecological factors limiting population abundance. Small-scale, mechanistic experiments are typically used to identify environmental variables that have significant influence on individual characteristics such as growth, habitat use, or physiology (e.g., Nislow et al. 2004, Kiernan & Moyle 2012); characteristics that are related to survival rates or population productivity. Subsequently, analyses of an observed time series of abundance data relative to the identified environmental factors is typically used to understand how these mechanisms culminate in changes in survival rates at a population level (Webster 2003, Lawson et al. 2004). However, analyses of temporal trends in data can lack statistical power and give conflicting, or non-significant results relative to theoretical predictions (Shenk et al. 1998), resulting in the impression that a specific environmental factor is not meaningfully related to population size (and thus should not be the focus of recovery efforts). Therefore, analyses should strive to maximize ecological relevance (in terms of choosing variables for analysis) and to appropriately characterize uncertainty or sources of error to minimize the possibility that significant environmental variation remains undetected (Zuur et al. 2010, Fredericksen et al. 2014). Although this is self-evident for any sound scientific inquiry, how one achieves it when describing species-environment relationships at a population level is equivocal at best (e.g., Hilborn & Walters 1992, Ver Hoef & Boveng 2007).

The validity of conclusions from regression analyses depends in part on appropriately characterizing the distributional form of the response, given the parameters, because biased estimates can result from misspecification (i.e., modeling data arising from one distribution with alternate distributions), as is well-described in the theoretical literature (e.g., Dick 2004). Survival values arise from a binomial process: the sum of a sequence of Bernoulli trials where an individual is either alive or dead (Collett 2003). However, the measurement and process errors contributing to estimates of annual abundance also influence the distribution of the relative survival estimates, making it unknown how closely the mean-variance relationship characterizing relative survival matches the expectation from a binomial process. In many cases, ecological data exhibits overdispersion relative to a binomial or poisson process (Lande et al. 2003, Ver Hoef & Boveng 2007, Zuur et al. 2009). A common way to deal with overdispersion in regression models is to use Generalized Linear Models (GLM) and either the quasi-likelihood or negative binomial family of distributions to estimate the regression parameters and the dispersion parameter affecting the variance jointly (e.g., Ver Hoef & Boveng 2007). However, this only accounts for situations in which the variance increases or decreases more quickly than the mean. An alternative would be to use a regression method that estimates the functional relationship between the predictors and response in situations where the underlying assumptions are violated to some extent (i.e., either the predictor or response contains outliers). Termed robust regression (e.g., Hampel et al. 1986, Heritier et al. 2009), these methods address a very specific problem; namely, to remain unbiased given uncertainty in sampling data and slight misspecification in the explanatory model relative to the data-generating process (Hampel et al. 1986). In doing so, they offer several distinct advantages over more commonly used regression techniques, including an increased ability to detect a subtle signal in noisy data as well as the ability to produce unbiased estimates of variance around a fitted relationship for overdispersed data (Cantoni & Ronchetti 2001). Starting from a parametric model (i.e., a particular model form as in a GLM), robust regression builds in protection against outlying behaviour in the data during the estimation process, by reducing the influence of outliers on the objective function (Hampel et al. 1986). As such, the robust counterpart to a GLM should not be considered a competing model *per se*, but rather a method by which to: (1) identify atypical values in a dataset (relative to what is assumed *a priori* by the model) and (2) to reduce bias in the estimated coefficients (particularly the variance) that result from these values. However, the true power of robust regression is that all data are included in fitting and no subjective decisions are needed relative to the quality of specific data points.

My primary goal in this chapter was to quantify changes in survival relative to environmental variation for use in recovery planning. In doing so, I explored the implications of common assumptions underlying Ordinary Least Squares (OLS) and Generalized Linear Models (GLMs) when attempting to describe relationships. I demonstrated how our understanding partially depends on the statistical technique and assumed distributional form of the response chosen prior to the analytical process. Using juvenile survival data from three populations relative to variation in hydrological flows, I found that population-level survival estimates did not match the expectation from a binomial process. I also demonstrated how restricting the analyses to Generalized Linear Models would have led to the identification of multiple significant hydrological predictors, yet low overall confidence in the importance of the relationships. However, using robust regression to evaluate the effect of outliers in the data relative to regression assumptions affirmed the selection of the lognormal model and resulted in a better understanding of the relationships between hydrological variables and survival; one that could be used for population-specific recovery planning. Although the conclusions were framed relative to a specific application, the methods are directly applicable to recovery planning for multiple species in which observational time series of abundances are available to estimate survival rates.

Case study

For endangered Atlantic salmon populations, there is considerable interest from multiple non-government organizations, academics and government departments to implement remedial actions at a watershed scale to promote population increase. Many of the actions related to habitat enhancement (e.g., bank stabilization, digger logs, changing channel morphology) are proposed because of their influence on hydrological flows, with the assumption being that such changes will increase the productive capacity of freshwater environments for Atlantic salmon (Roni et al. 2002). Hydrological variation is thought to be a key factor controlling the population dynamics of freshwater fishes, in that flow levels influence the majority of physical factors (e.g., current velocity, water depth, temperature regime) and ecological interactions (e.g., competition, predation) experienced by fish in freshwater environments (Bunn & Arthington 2002, Kiernan & Moyle 2012). Five major components of flow are considered to be ecologically important across a diverse range of riverine ecosystems: extreme low flows, low flows, high flow pulses, small floods and large floods (Mathews & Richter 2007, LeRoy Poff et al. 2010). Under this categorization, low flows represent typical flow conditions which determine the amount and characteristics (e.g., temperature, connectivity, velocity) of aquatic habitat available for the majority of the year. The other flow categories are thought of as discrete events that typically trigger a behavioural response (Mathews & Richter 2007) and thus might be correlated with survival rates. Extreme low flows describe drought conditions, which are characterized by a decrease in surface area and water volume causing extreme values of several physical and chemical water quality variables, such as temperature, flow velocity, oxygen concentration, or dissolved mineral content (Magoulick & Kobza 2003, Rolls et al. 2012). For aquatic species, droughts induce stress responses and typically increase mortality due to a reduction in habitat connectivity, availability and suitability (Lake 2003). Thus it might be expected that increased frequency or severity of drought conditions experienced by juvenile Atlantic salmon would result in measurable declines in survival rates at a population level. Conversely, high flow pulses (up to bankfull) are thought to recharge river systems by reducing water temperatures, flushing wastes, increasing oxygen availability, and delivering organic matter (Mathews & Richter 2007) and thus would be expected to be positively correlated with survival. However, large floods or quick changes in water level are considered to be less directly beneficial for individuals given that they can move significant amounts of sediment and large woody debris, transport organisms downstream, and alter the direction of the main channel. However, in the long term they also form new habitats and refresh water quality conditions in stagnant portions of the stream (Allan 2004). Given that a decline in survival related to flood conditions is predicted to come from sediment transport and displacement (Caissie 2006), the rise rate of the river could have a more direct influence on survival than flood conditions per se. Understanding how low or high water conditions influence juvenile survival in specific populations would be a first step towards identifying whether or not hydrological change should be a focus of recovery efforts, as well as which specific components of the flow regime should be targeted in specific watersheds.

Methods

Data sources

Time series of juvenile density estimates from the west branch of the St. Mary’s River, the LaHave River above Morgans Falls (both in Nova Scotia, Canada) as well as the Nashwaak River in New Brunswick, Canada (Figure 3-1) were used in this chapter. Of the Atlantic salmon populations considered to be endangered by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC 2010), these rivers are the only three in the Maritime Provinces that: (1) have long-term monitoring programs which enumerate all freshwater life stages for at least a portion of the watershed, and (2) have hydrological monitoring stations gauging daily water flows in a location near to that for the population monitoring data. The annual egg deposition and juvenile density estimates used for analyses are the same as in recent assessments (Gibson & Bowlby 2013 (St. Mary’s and LaHave); and Gibson et al. 2016 (Nashwaak)). Annual egg depositions were estimated from the number and characteristics of adult spawners, and age 0 densities (N/100m^2) were estimated from electrofishing surveys. Age 0 salmon were those sampled in the year of hatching as juveniles between June and September. A Poisson GLM (incorporating site and year effects) was used to estimate age 0 density values for all potential sites (from a random-stratified survey design) prior to calculating the annual mean densities. This standardization was done to reduce annual estimation error and account for any directional biases related to changes in site selection (Gibson et al. 2009). For these populations, it has been shown to produce density estimates that are more consistent with data available for other life stages when analyzed in age- and stage-structured population dynamics models (e.g., Gibson and Bowlby 2012). In other words, estimates can be considered to be representative of juvenile abundance at the population level when multiplied by a scalar representing the habitat area in each watershed. All data were analyzed relative to a specific egg cohort, with age 0 density lagged by one year relative to egg deposition for calculating survival. To make the population-level estimates for the two life stages comparable, age 0 density was multiplied by the habitat scalar to calculate population-level abundance in the watershed. Thus, egg deposition estimates in a given year and age 0 abundance estimates in the following year are integer values, and their ratio is an estimate of survival from the egg to the age 0 life stages for each egg cohort.

Environment Canada maintains hydrological gauging stations on the St. Mary’s River at Stillwater (45°10’27”N 61°58’47”W), on the LaHave River at West Northfield (44°26'50"N 64°35'28"W) and on the Nashwaak River at Durham Bridge (46°07’33”N 66°36’40”W) (Figure 3-1). These stations have been in operation continuously from 1915 on the St. Mary’s and LaHave Rivers and 1961 on the Nashwaak; historical flow data can be downloaded from the Water Survey of Canada’s HYDAT database of archived hydrometric data (http://www.ec.gc.ca/rhc-wsc/default.asp?lang=En&n=9018B5EC-1,

Accessed May 2013). As detailed in the metadata that accompanies the HYDAT database, daily discharge (in cubic meters per second) corresponds to averages of hourly flow recordings and values were estimated for days during which the station was not operational. From the discharge values identified as estimated rather than as measured, I calculated that 1.6%, 0.7% and 1.8% of the time series was estimated for the St. Mary's, LaHave and Nashwaak Rivers, respectively; considering all years up to 2010.

Hydrological variables describing flow conditions were calculated from the Environmental Flow Components module of the Indicators of Hydrologic Alternation software (Mathews & Richter 2007). This module categorizes daily flows into the five ecologically important components identified earlier (extreme low flows, low flows, high flows, small floods and large floods) based on user-defined thresholds. From these categories four annual variables were calculated: (1) the minimum flow value (i.e., the lowest flow value recorded), (2) the frequency of extreme low flows (i.e., the number of days categorized as having extreme low flows), (3) the timing of extreme low flows (i.e., the median ordinal date of all the days classified as extreme lows) and (4) the rise rate (i.e., the median of all positive differences between two daily flow measurements). This value represents how quickly water levels increase following precipitation events or snow-melt, and does not depend on the initial water conditions in the river (i.e., the classification of each flow measurement as extreme low, low, or high flows; small or large floods). The lowest 20% percentile (of all flow measurements regardless of year) was used as the cut-off between extreme low flows and low flows to ensure that a value could be calculated for all variables for all years. For this to be true for high water conditions (i.e., to ensure a value could be calculated for all variables for all years) would have required combining all flows categorized as high flows, small floods and large floods, even though these would be expected to have opposing relationships with survival (e.g., positive for high flows and negative for large floods). Instead, the rise rate was used as an indicator of the flashiness of the river system and the potential for bedload transport (Caissie 2006), with the expectation that faster rise rates would negatively affect survival. Based on Kendall’s tau, correlations among predictor variables were less than 0.6. To ensure that the hydrological conditions corresponded to the time period between autumn egg deposition and juvenile sampling the subsequent summer, a year was considered to begin on November 1 and end at the start date of the summer electrofishing survey (July to September, depending on the year).

Regression analyses

Survival and mortality can be thought of multiple ways, leading to different response variables and model structures for regression analyses. Here, hydrological relationships with age 0 juveniles could be modeled directly, assuming a Poisson distribution for age 0 abundance (thought to be appropriate for count data), a log link, and including an offset for starting population size (egg deposition) in a GLM (McCullagh & Nelder 1989). A second alternative would be to model a binomial process using a logit link in a GLM (McCullagh & Nelder 1989) using the number of successes and number of failures (i.e. a two-part vector of $age\ 0, (age\ 0 - egg)$) as the input for the response variable. A third option would be to model the instantaneous mortality rate assuming a normal error distribution as in a linear regression. Survival is related to mortality by $S = e^{-Zt}$, so the instantaneous mortality rate (*Z*) is: $Z = -\ln(S)$ (Ricker 1975). It is important to note that the lognormal model would be expected to have a slope estimate opposite in sign as compared to the other regressions. Starting from these three models (count data with Poisson errors, an odds ratio with binomial errors, and the instantaneous mortality rate with lognormal errors), I used two different methods to account for potential overdispersion in the GLMs. One was to substitute the quasibinomial and quasipoisson family into the GLMs described above, which estimates a dispersion parameter for the variance. The second was to assume a negative binomial distribution when modeling age 0 abundance (with an offset for the previous year's egg deposition) in a GLM (Ver Hoef & Boveng 2007). The results in this chapter are presented for the lognormal, quasibinomial, quasipoisson and negative binomial models; as detailed in Table 3-1. For the two years in which estimated survival was > 1 on the St. Mary's River, survival was set at 1 in order to be able to fit the quasibinomial, quasipoisson, and negative binomial models.

Atypical values (outliers) and points with high leverage are known to bias parameter estimation using maximum likelihood as in GLMs (Graham 2003). Additionally, the specific data points contributing to such biases as well as the magnitude and direction of the bias cannot be assessed statistically from the output of GLM regressions (Richards 2008, Zuur et al. 2010), although several ad hoc methods of identifying outliers exist (e.g., visual examination of residual plots). Robust regression provides a statistical framework from which to both identify and limit the influence of extreme values or leverage points on parameter estimation. Depending on the specific method used, up to half of the data can take atypical values and still have limited influence on coefficient estimates (e.g., Yohai 1987, Hampel et al. 1986). Therefore, I used robust regression more as an extension of the traditional linear and GLMs, to both evaluate the presence of outliers and to obtain less biased estimates of regression coefficients for use in recovery planning.

Robust regression uses Mallows or Huber-type robust estimators (typically called M-estimators; Jajo 2005, Cantoni & Ronchetti 2001) to estimate model parameters. Although the postulated model (i.e., the assumed distribution of the response and associated linear predictor) used in robust analyses is analogous to that used in traditional regressions or GLMs (Table 3-1), estimation of the β parameters proceeds in a different manner. As a simple example, it is useful to compare the familiar least-squares estimator with an M-estimator to appreciate the main differences among the two techniques. For a linear model, the least-squares estimator for β minimizes the objective function: $\sum_{i=1}^{n} r_i^2$, where each residual (r_i) is: $r_i = y_i - \beta_0 - \sum_{j=1}^{p} x_{i,j}\beta_j$ for each value of the response variable, y_i, and each value of p hydrological predictors ($x_{i,j}$). An M-estimate of β

minimizes the objective function: $\sum_{i=1}^{n} \rho(r_i / s)$, where s is an estimated scale parameter and ρ is called the psi-function. Note that if the weighting function $\rho(r_i / s)$ is equivalent to r_i^2, the parameter estimates will be the same as from ordinary least squares (Jajo 2005). The influence of individual residuals on model fitting is controlled by the derivative of the psi-function: ψ. Multiple functions can be chosen for ψ, but each has the common characteristic of limiting the contribution of data points that deviate substantially from the fitted relationship. The scale parameter can be thought of as a multiplier on the error term, representing deviation from the assumed error distribution.

A simple M-estimate (as above) was not appropriate for this application given that the levels of the predictor were not fixed *a priori* (Maronna et al. 2006). Here, I used MM-estimation in the lmrob function for fitting a robust lognormal model (Yohai 1987) and the Mqle method in the glmrob function for fitting robust binomial and robust Poisson models (Cantoni and Ronchetti 2001), as implemented in the readily available R package ‘robustbase’ (Rousseeuw et al. 2013). I followed recommendations for the tuning constants from Koller & Machler (2013) for the robust lognormal and robust binomial models (k = 4.685 for the redescending ψ used in lmrob; k = 1.345 for the Huber ψ in glmrob). I increased the tuning constant used for the robust Poisson model slightly (k = 1.8) for both rivers. As in GLMs, the Mqle method uses the iteratively re-weighted least squares algorithm during fitting (Cantoni & Ronchetti 2001). Given that there are no robust counterparts to the quasi-family GLMs, I employed the newly available ‘glmrob.nb’ function (Aeberhard et al. 2014) to allow for overdispersion in the response for a robust model. Here I used the redescending Tukey’s biweight function for ψ in the M-estimates of the regression parameters and the same tuning constant as above (e.g., k = 4.685).

In total, I evaluated 4 potential hydrological predictors using the 4 parametric models (Table 3-1) as well as the two different regression types (robust and traditional), and present the results from eight models (lognormal, quasibinomial, quasipoisson, negative binomial, robust lognormal, robust binomial, robust Poisson and robust negative binomial). Model selection proceeded in three general steps: (1) simplification of the initial multivariate model using traditional linear regression or GLMs, (2) evaluation of regression assumptions from diagnostic plots (all traditional models) and estimated overdispersion parameters (quasi- family models), and (3) evaluation of the effect of atypical values on the estimated coefficients using robust regression. Hydrological predictors were both sequentially added and dropped from each regression based on a comparison of nested models using ANOVA (for lognormal models), Likelihood ratios (for GLM models; Zuur et al. 2009) and the Robust Wald test (for robust lognormal or robust GLMs; Sommer & Huggins 1996). Each of these methods is appropriate for model selection among nested candidate models of the specific type for which they were applied. Although there are other methods that are appropriate, there is no single method that was possible for model selection across all model forms (i.e. lognormal, GLMs and robust models). This is discussed in more detail in the following paragraph. In all regressions, the final model included a single predictor. Diagnostic plots of the residuals, quantiles and fitted vs. observed values were examined visually for each model to assess the appropriateness of model assumptions (Zuur et al. 2009, Zuur et al. 2010). If autocorrelation was detected in the residuals on the basis of diagnostic plots, I used AIC to compare the fit from a Generalized Least Squares (GLS) model with a residual first-order correlation structure (ar1) to the GLM fit (Zuur et al. 2009) and repeated model selection.

For the results presented, model diagnostics were similar (i.e., there was no compelling reason to reject individual models based on diagnostic plots), so the assumptions underlying each model appeared to be appropriate. This was not true for some preliminary model formulations that were evaluated yet not included in this Chapter. For example, modeling the relative survival rate as a Gaussian GLM with a log link had extremely strong patterns in the residual plots, indicating that the model structure was inappropriate.

Traditionally, model selection for regression analyses uses an information theoretic approach such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC); both of which assess fit from maximum likelihood scores that are penalized for model complexity (Johnson & Omland 2004). This presents a problem when attempting to compare among the GLMs presented here (i.e., to compare the best-supported models for each distributional form of the response after variable reduction) because the quasi- family is characterized by a mean and variance but not a specified distributional form, which means that the log-likelihood is not defined (Ver Hoef & Boveng 2007). Therefore, it is not possible to use a statistical criterion such as AIC to evaluate model fits from all 4 traditional regressions (lognormal, quasibinomial, quasipoisson, and negative binomial). However, the lognormal and negative binomial models could be directly compared with the Akaike Information Criterion for small samples (AICc), and the quasi- family models could be compared using a quasi-AIC for small samples (QAICc) (e.g., Young et al. 2009), even though both comparisons were between models with different response variables (Johnson & Omland 2004). Further to this, it is possible to assess the appropriateness of the quasi- family models via the variance inflation factor (model deviance divided by residual degrees of freedom) or the dispersion parameter (Collett 2003), where values are expected to be less than ~ 4 when the data structure is well-specified (Anderson et al. 1994). In relation to the robust models, the estimated coefficients would be essentially identical to those estimated from traditional regressions if model assumptions were met (i.e., provided that variation in the response conformed exactly to the assumed distribution and the predictors did not contain outliers). However, as compared to GLMs assuming the same distribution of the response, robust regressions have greater statistical efficiency (reduced variance) and can produce unbiased estimates of coefficients if assumptions are violated to some extent (Jajo 2005, Hampel et al. 1986, Heritier et al. 2009). Therefore, I considered the robust regressions to be way in which I could reduce the potential for non-detection of a subtle relationship with hydrological change and obtain better estimates of coefficients for use in recovery planning, relative to the equivalent traditional model (*c.f.* lognormal with robust lognormal; binomial with robust binomial, Poisson with robust Poisson, and negative binomial with robust negative binomial).

Results

For the St. Mary’s River, the lognormal and GLM regressions did not consistently simplify to the same hydrological predictor. Survival was found to be negatively associated with the frequency of extreme low water events (xlow.freq) from the lognormal and negative binomial models, while the quasipoisson model identified a positive relationship with the timing of extreme low water events (dist.low) and the quasibinomial model retained no predictors (Table 3-2). The lognormal model had much more substantial support on the basis of AIC as compared to the negative binomial (AICc = 64 and 147, respectively) using xlow.freq as the predictor. For the quasipoisson model of survival relative to dist.low, the estimated dispersion parameter (12.05) was substantially greater than 4, indicating that this model was not an adequate characterization of the data (Anderson et al. 1994), even though the predictor was retained in the optimal model. For a visual comparison among the different model forms, it was necessary to plot them relative to a standardized response variable. Here, I chose to use the relative survival rate of each egg cohort, a prevalence ratio between annual estimates of population-level age 0 abundance and the previous year’s egg deposition. This is standard output from the ‘predict.glm’ function in R for the quasibinomial, quasipoisson and negative binomial models. Re-calculating the response to be a survival rate and comparing the fits of the 4 traditional regressions with xlow.freq revealed only minor deviations in the estimated mean slope (Figure 3-2). As above, the estimated dispersion parameters for the quasipoisson and quasibinomial models of survival relative to xlow.freq were unacceptably high (16.16 and 12.05, respectively). Overall, the lognormal model was considered to be the best model structure with which to describe the relationship between hydrological change and survival for the St. Mary's River. When the data were re-examined in the robust analyses, all four robust models found the response to be negatively associated with the frequency of extreme low water events (Table 3-2). This suggests that the weak relationship with dist.low from the quasipoisson GLM was an artifact of outliers or points with high leverage in the data. The lognormal and robust lognormal models had identical slope estimates (*c.f.* 0.064 and 0.064; Table 3-2), indicating that atypical values had no influence on this parameter estimate. However, the 95% confidence intervals (based on the normal approximation) are much smaller for the robust model, particularly at lower survival values (e.g., compare the lognormal and robust lognormal fits; Figure 3-2). The robust lognormal model identified three of 21 data points that were contributing substantially to this difference, in that they were given a weighting (a robustness weighting that corresponds to the $\rho(r_i / s)$ function of the residual divided by the residual) of less than 0.7.

For the Nashwaak River, the different model forms were more consistent in that they all identified a negative relationship between the response and the rise rate of the river (Table 3-2). Slope estimates from the quasibinomial, quasipoisson and negative binomial GLMs were very similar (-0.299, -0.247 and -0.256, respectively; Table 3-2). I recognize that the slope estimate from the quasibinomial GLM is not strictly comparable to the quasipoisson and negative binomial because the response is a log odds ratio. As was the case for the St. Mary's River, the estimated dispersion parameters for the quasibinomial and quasipoisson models were quite high (8.10 and 7.68, respectively), again indicating that these models do not adequately describe the data. Based on AICc, there was substantially more support for the lognormal model as compared to the negative binomial GLM (AICc = 74 and AICc = 298, respectively) and was considered to be the most informative model structure with which to describe these data. Similar to the St. Mary's, the robust weightings indicate that the data are approximately lognormal, given that only three of 39 points are down-weighted by more than 0.7 (Cantoni & Ronchetti 2001). However, these outliers have a greater influence on the estimated coefficients in that the robust lognormal model resulted in a more negative slope relative to the lognormal model (Table 3-2; Figure 3-3).

For the LaHave River, initial fits from the lognormal and GLM models found survival to be related to the frequency of extreme low water events (xlow.freq), but had strongly autocorrelated residuals at a lag of 1. Reanalysis in a Generalized Least Squares model with an ar1 residual correlation structure substantially reduced model AIC, but xlow.freq was no longer retained relative to an intercept-only model. Therefore, a relationship between egg to age 0 survival and hydrological change could not be described for the LaHave River population from these analyses.

Discussion

Linkages between methodology and ecological inference

Investigating changes in survival relative to environmental variation requires a systematic analysis that explicitly considers what the monitoring data represent and where variation is likely to come from in order to draw meaningful conclusions. It is particularly important in cases where the distributional assumptions underlying the methods cannot be investigated *a priori* (Ver Hoef & Boveng 2007, Zuur et al. 2010), as well as in situations where both predictors and response could contain variation that is unaccounted for with a particular model structure (Richards 2008). The relative popularity of GLMs stems from their ability to account for alternate mean-variance relationships and errors arising from certain types of biological processes (McCullagh & Nelder 1989). At first glance, GLMs may have been expected to be the most appropriate method for analyzing juvenile salmon survival relative to hydrological variation given that the observational time series derive from count data and survival is an inherently binomial process. However, my analyses suggest that the appropriate error distribution for the survival estimates deviates from theoretical expectations, likely due to the combined observation and measurement error associated with the population monitoring data.

Simulation studies have demonstrated that violating regression assumptions can produce spurious correlations or can mask significant correlations when data contain additional errors in the predictors or response (Graham 2003). Both of these potential biases were demonstrated by the GLM models for the St. Mary’s River: with the quasipoisson model showing a seemingly spurious correlation with the timing of extreme low flows, and the relationship between survival and the frequency of extreme low water events being masked in the quasi- family models. The most obvious outliers (i.e., the survival estimates > 1) occurred when the median timing of extreme low flows was later in the year (i.e., at higher values of dist.low) and would be expected to have high leverage on model fits. It is likely that dist.low was only retained as a significant predictor by the quasipoisson model because of characteristics of the estimation process. For example, quasipoisson regression gives greater weight to larger counts in the fit from iteratively weighted least squares as compared to alternatives such as the negative binomial (Ver Hoef & Boveng 2007) and the age 0 densities contributing to the survival estimates above one were an order of magnitude larger than the majority of the other values. Although the response values and offsets that would give relative survival values greater than one could have been removed because they were not biologically plausible, this would have been the equivalent of preferentially excluding data points when survival would be expected to be high. As an alternative, robust methods are a powerful way to analyze data that is subject to measurement and process error in that they do not require any *a priori* assessment of data quality (i.e., removal of biologically implausible values or other outliers). Because of the down-weighting imposed by the influence function during estimation, the response values or offsets that give relative survival estimates greater than one would have little influence on model fit. Therefore, the robust regressions should have also identified the relationship between survival and the timing of low water events (dist.low) if it was unrelated to leverage points in the data. Similarly, the impact of outliers was found to be relatively small on the slope estimates for the St. Mary's River (*c.f.* traditional and robust parameter estimates; Table 3-2), but larger on the standard deviation. This influences the significance of parameters in the model, and is likely why the quasibinomial and quasipoisson GLMs did not retain xlow.freq as a predictor.

Practical implications for species of conservation concern

The practical consequences of such statistical considerations can be quite large for this type of a research question. Restricting these analyses to GLMs (e.g., the quasibinomial, quasipoisson and negative binomial models) would have led to ambiguous results among candidate models for the St. Mary's River as well as to slight confidence in the fitted relationships on the Nashwaak River. Furthermore, it would not be immediately obvious whether the assumed distribution was inappropriate, the underlying relationships were weak (i.e., not ecologically important), or if variability in the data (i.e., violations of assumptions) was adversely affecting parameter estimation. Taking this one step further for recovery planning, the GLMs would not form as convincing a basis to argue that remediation actions to alter hydrological flows should be included in a remediation strategy. Extending the analyses using both the lognormal model as well as robust regression enabled me to address all of these uncertainties, and to both identify the hydrological predictor best supported by the data (xlow.freq on the St. Mary's River and rise.rate on the Nashwaak), as well as to reduce biases in the estimated coefficients. The latter was particularly important on the Nashwaak River, given that the estimated slope increased from approximately 0.3 to 0.4, indicating closer to a four-fold rather than a three-fold change in the instantaneous mortality rate over the range of observed rise rates (Figure 3-3).

The identification of population-level changes in survival in relation to hydrological variation gives indirect evidence for the specific threats that have resulted in population decline in these three rivers as well as the expected population response to recovery actions. Furthermore, it would be expected that changes in hydrological conditions that have resulted in increased contrast in the data (i.e., anthropogenic activities that cause more extreme flow values) would enhance our ability to detect relationships with flow (Fredericksen et al. 2014). The primary anthropogenic activities that have been linked to changes in hydrological flow patterns are related to land clearing (Allan 2004, Broadmeadow & Nisbet 2004, LeRoy Poff et al. 2006), which can result from mining operations, urbanization, agriculture or forestry, and the effects of which can be exacerbated by changing precipitation patterns due to climate change (Milly et al. 2005). Extreme low water conditions can arise from a reduced capacity of the watershed to retain run-off owing to the removal of vegetation (Broadmeadow & Nisbet 2004) as well as to water extraction from surface water or aquifers (Allan 2004). Recovery plans that identify the specific location, extent and severity of such activities as well as remedial actions designed to alleviate these threats would be expected to have a positive influence on egg to age 0 survival on the St. Mary’s River. The speed at which water levels increase is related to geology and vegetation patterns which determine the capacity of a drainage area to absorb run-off (Allan 2004, Jewett et al. 1995), as well as to channel morphology, where straighter, deeper streambeds enable faster water flow (Paul & Myer 2001). Given that approximately 90% of the Nashwaak River watershed was clear-cut in 1978-1979 (Jewett et al. 1995), our ability to detect the negative relationship between egg to age 0 survival and hydrological rise rate may represent the effect of a land-use legacy (Greenwood et al. 2012), by increasing the contrast in the data for this Atlantic salmon population. Remediation focused on riparian planting, minimizing erosion and sources of sedimentation, as well as increasing channel complexity would be expected to result in increased egg to age 0 survival in the Nashwaak River. For the LaHave River, autocorrelation in the residuals was the strongest signal found in the data, indicating a decline in egg to age 0 survival over the duration of monitoring that was not related to hydrology. These results do not preclude the possibility that a relationship between egg to age 0 survival and alternate hydrological predictors exist, nor that additional data collection and a longer time series would enable a relatively weak relationship to be described. However, in terms of guiding recovery planning, alternate threats that are not as strongly linked to hydrology, such as the effects of invasive smallmouth bass and chain pickerel (Wathen et al. 2011) or changes to water quality (Paul & Myer 2001) should be investigated.

This chapter provides only one example of how our interpretation of ecological data changes as a result of the assumptions made during the analytical process, and highlights the implications that these assumptions can have for future recovery planning.

Given the declining trends in a large number of freshwater fish species (e.g., Dudgeon et al. 2006) as well as the limited time and resources available for remediation, the efficient identification of priorities for recovery planning is a pressing ecological issue.

Tables

Table 3-1. Description of the lognormal and Generalized Linear Model (GLM) forms considered for analyzing egg to age 0 survival data for Atlantic salmon from three populations, detailing the response variable, response distribution, parametric model and variance estimator. Terms used are: hydrological predictors ($X_{n,i}$), mean value (μ_i), probability of being alive (π_i), age 0 density (λ_i), egg density (n_i), α and β_n are the regression coefficients, θ is an overdispersion parameter, and κ is the scale parameter from a gamma distribution.

Model	Dependent variable	Response Distribution	Parametric model	Variance
Lognormal	$-\ln(S)$	$Y_i \sim N(\mu_i, \sigma_i)$	$\alpha + \beta_1 X_{1,i} + \ldots.\beta_n X_{n,i}$	σ^2
Quasi-binomial (with offset)	$\dfrac{\lambda}{(\lambda - n)}$	$Y_i \sim B(n_i, \pi_i)$	$n_i \dfrac{\exp(\beta_0 + \beta_1 X_{1,i} + \ldots.\beta_n X_{n,i})}{1 + \exp(\beta_0 + \beta_1 X_{1,i} + \ldots.\beta_n X_{n,i})}$	$\theta n_i \pi_i (1 - \pi_i)$
Quasi-Poisson (with offset)	λ	$Y_i \sim P(\mu_i)$	$\ln(\lambda_i) - \ln(n_i) = \beta_0 + \beta_1 X_{1,i} + \ldots\beta_n X_{n,i}$	$\theta\mu_i$
Negative binomial (with offset)	λ	$Y_i \sim NB(\mu_i, \kappa)$	$\ln(\lambda_i) - \ln(n_i) = \beta_0 + \beta_1 X_{1,i} + \ldots\beta_n X_{n,i}$	$\mu_i(1 - \mu_i\kappa)$

Table 3-2. Comparisons of coefficients from eight regression model forms describing egg to age 0 survival relative to the frequency of extreme low water conditions (xlow.freq) for the St. Mary’s River and the median rise rate (rise.rate) for the Nashwaak River. Coefficients from a model that retained an alternate hydrological predictor (the timing of extreme low water events; dist.low) for the St. Mary’s River are also shown. The annual survival rate is input into the quasibinomial regression as a two-variable vector of # alive, # dead. Note that the slope estimates for the models of mortality rates would be expected to be opposite in sign to those of survival rates or age 0 density. The slopes are interpreted as the rate of change in the response variable per unit change in the predictor variable; as such they are not directly comparable among all model forms. Results from the LaHave River are not included because no significant predictors were identified.

River	Model	Dependent Variable	Independent Variable	Value	S.E.	P
St. Mary's	Lognormal	Instantaneous mortality rate	xlow.freq	0.064	0.027	0.030
St. Mary's	Quasibinomial	Annual survival rate	xlow.freq	-0.046	0.038	0.240
St. Mary's	Quasipoisson	Age 0 abundance (offset pop size)	xlow.freq	-0.039	0.031	0.235
St. Mary's	Negative binomial	Age 0 abundance (offset pop size)	xlow.freq	-0.054	0.021	0.010
St. Mary's	Robust lognormal	Instantaneous mortality rate	xlow.freq	0.064	0.020	0.005
St. Mary's	Robust binomial	Annual survival rate	xlow.freq	-0.055	0.007	<< 0.001
St. Mary's	Robust Poisson	Age 0 abundance (offset pop size)	xlow.freq	-0.048	0.003	<< 0.001
St. Mary's	Robust negative binomial	Age 0 abundance (offset pop size)	xlow.freq	-0.059	0.007	<<0.001
St. Mary's	Quasipoisson	Age 0 abundance (offset pop size)	dist.low	0.014	0.006	0.038

River	Model	Dependent Variable	Independent Variable	Value	S.E.	P
Nashwaak	Lognormal	Instantaneous mortality rate	rise.rate	0.326	0.114	0.007
Nashwaak	Quasibinomial	Annual survival rate	rise.rate	-0.299	0.138	0.037
Nashwaak	Quasipoisson	Age 0 abundance (offset pop size)	rise.rate	-0.247	0.119	0.045
Nashwaak	Negative binomial	Age 0 abundance (offset pop size)	rise.rate	-0.256	0.108	0.018
Nashwaak	Robust lognormal	Instantaneous mortality rate	rise.rate	0.401	0.137	0.006
Nashwaak	Robust binomial	Annual survival rate	rise.rate	-0.420	0.052	<< 0.001
Nashwaak	Robust Poisson	Age 0 abundance (offset pop size)	rise.rate	-0.294	0.046	<< 0.001
Nashwaak	Robust negative binomial	Age 0 abundance (offset pop size)	rise.rate	-0.460	0.109	<<0.001

Figures

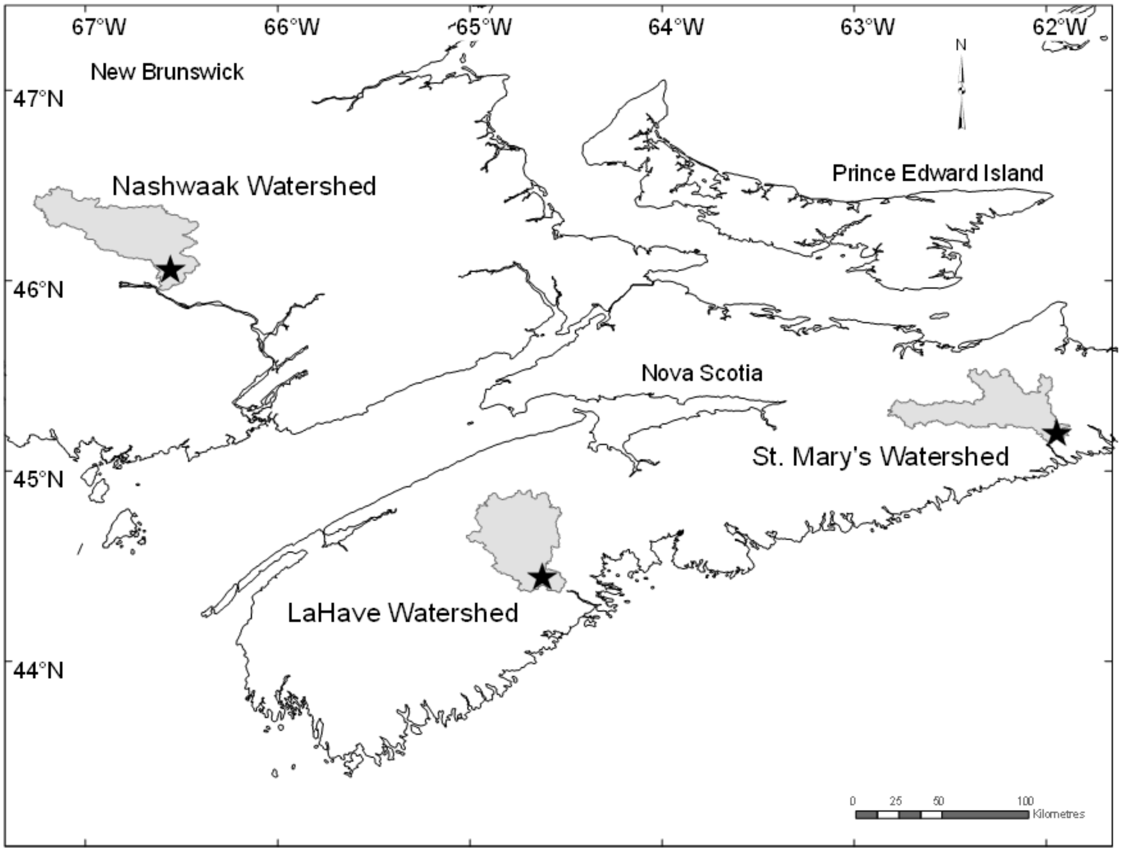

Figure 3-1. Location of the study area in Atlantic Canada showing the boundaries of the St. Mary's, LaHave and Nashwaak watersheds as well as the locations of the hydrological monitoring stations (stars).

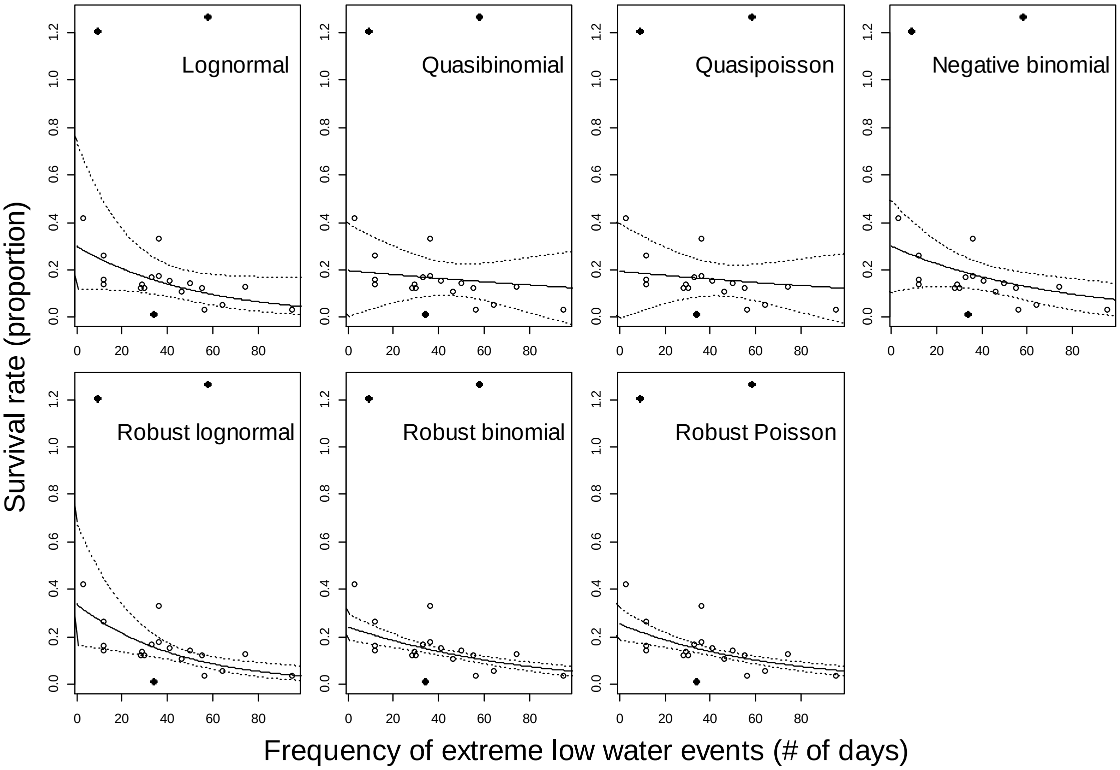

Figure 3-2. A comparison of the fits of seven different regression models to egg to age 0 survival data (egg cohorts: 1989-2009) relative to the frequency of extreme low water events from the St. Mary's River, showing the observed values (points), the fitted model (lines) and 2 s.e. (dashed lines). The response variable was standardized to be an annual survival rate to facilitate comparison. Although the preferred quasipoisson model retained an alternate predictor as significant (Table 3-2) and the quasibinomial model retained no significant predictors, the non-significant relationship with the frequency of extreme low flows is shown here. Observations that were identified as outliers by the robust lognormal model are identified by the filled points. To date, a predict function has not been developed for the newly available robust negative binomial model, which is why the results are not plotted (although see Table 3-2).

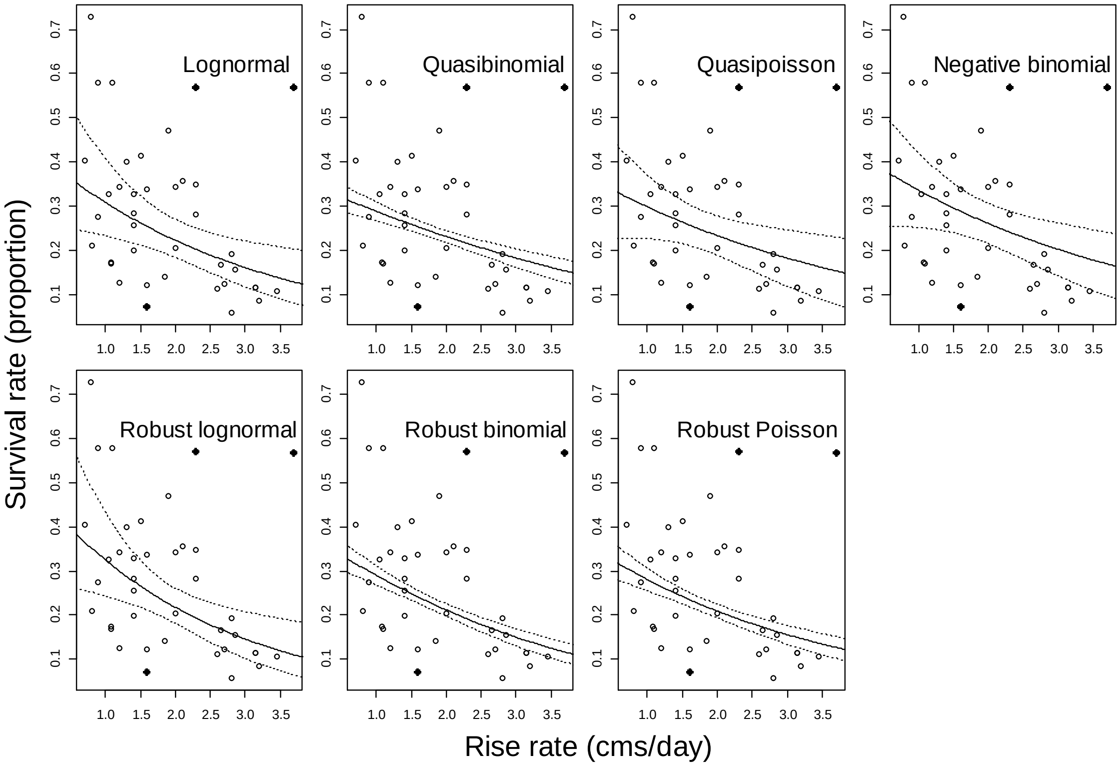

Figure 3-3. A comparison of the fits of seven different regression models to egg to age 0 survival data (egg cohorts: 1970-2009) relative to the rise rate (cms/day) from the Nashwaak River, showing the observed values (points), the fitted model (lines), and 2 s.e. (dashed lines). The response variable was standardized to be an annual survival rate to facilitate comparison. Observations that were identified as outliers by the robust lognormal model are identified by the filled points. To date, a predict function has not been developed for the newly available robust negative binomial model, which is why the results are not plotted (although see Table 3-2).

Chapter four: Applying landscape genetics to evaluate threats affecting endangered Atlantic salmon populations

Abstract

Landscape genetics affords a potential analysis framework to evaluate the effect of contemporary land use on endangered species at a population level. However, historical patterns of population connectivity need to be accounted for prior to testing for the contemporary effect of threats. I had two objectives for this chapter: (1) to optimally describe historical patterns in population connectivity for a diadromous fish species before (2) evaluating whether residual genetic variation was correlated with ecological changes arising from several types of land use. Using endangered Atlantic salmon populations as a case study, I evaluated whether historical patterns in population connectivity were more likely to result from dispersal limitation (Isolation by Distance; IBD) relative to habitat choice and reproductive success (Isolation by Environment; IBE). Second, I used Reciprocal Causal Modeling to identify the types of land use contributing to three threat indices, and subsequently Multiple Regression on Distance Matrices to evaluate the relative severity of each. Pairwise genetic distances were most highly correlated the Water Quality and Sedimentation indices; mean annual pH and abandoned mine density contributed equally to the negative correlation with Water Quality, and road density contributed to the positive correlation with Sedimentation. These results suggest

that straying Atlantic salmon avoid watersheds with reduced water quality (resulting from acidification and abandoned mines) and higher road density, yet are not responding to watershed fragmentation (from road-river crossings and dams) at a population level. This research is among the first to explicitly compare alternate behavioural hypotheses leading to dispersal patterns for diadromous fishes and to quantitatively assess freshwater threats for Atlantic salmon at a population level using landscape genetics.

Introduction

Anthropogenic changes to the environment, particularly in terms of land use, are one of the major causes of species decline worldwide (Foley et al. 2005, Dudgeon et al. 2006, Venter et al. 2006). Effective recovery planning to address these declines relies on understanding the relative importance and severity of multiple threats in order to prioritize among potential recovery actions (Norris 2004). However, in many situations the quantitative links between population-level productivity and the specific land use activities identified as threats are not known (Lawler et al. 2002, Roni et al. 2002). For example, forestry, urbanization, agriculture, roads, industrial corridors or mining activities have all been shown to be related to processes governing sedimentation rates, hydrological flows, temperature regimes and other environmental characteristics of watersheds (Allan 2004, LeRoy Poff et al. 2010). The relationships between such environmental processes (e.g., sedimentation) and changes in freshwater fish survival or other vital rates are often understood quantitatively (e.g., Bunn and Arthington 2002), yet typically cannot be linked back quantitatively to the underlying threat (e.g., extent of forestry activity; Lawler et al. 2002). This means that while it is possible to understand the relative severity of a single threat among watersheds, it is not currently possible to assess the relative severity of multiple threats affecting a single population within a watershed. Even when abundance time series are available, land use patterns and other types of spatial data on threats are typically aggregated over multiple years (due to the time and labour-intensive nature of data collection) thus preventing time series analyses with population abundance. Additionally, different threats can be measured in different units (e.g., proportions, counts, or densities), complicating direct comparison. For these reasons, alternate ways of evaluating multiple concurrent threats and their influence on populations are needed.

One possibility is to use landscape genetics to investigate whether or not specific threats influence population genetic structuring. Measured variation at neutral markers such as microsatellites largely arises from within-population genetic drift causing divergence, tempered by the homogenizing effects of gene flow among populations (Allendorf et al. 2013). Threats that have fragmented populations or caused substantial reductions in abundance would be expected to reduce the rate of gene flow (fewer available migrants) while concurrently increasing the rate of genetic drift (increased isolation among small or declining populations) in a manner potentially dependent on the extent of the threat. One well-studied example is fragmentation of populations in response to roads, where roads have been found to increase genetic isolation of adjacent populations by acting as a barrier to movement (e.g., Epps et al. 2005, Arens et al. 2007, Clark et al. 2010). Such population-level responses to contemporary threats are expected to be relatively rapid. Simulation studies have found that contemporary landscape changes leading to reduced population connectivity are detectable from microsatellite variation in as little as five generations (longer if dispersal rates are extremely low), provided dispersal distances are sufficiently large (Landguth et al. 2010) or population size is low (Dileo et al. 2013). This suggests that the influence of recent threats on populations would be detectable even if the threats had been acting for a relatively short period of time.

Identifying contemporary threats relies on accurately characterizing historical patterns of gene flow (i.e., those existing prior to the advent of the threat), in order to prevent biased inference (Singh et al. 2013). Therefore, understanding how landscape elements influence the dispersal process becomes paramount to understanding expected patterns of population connectivity prior to the advent of threats. The majority of research questions in landscape genetics focus on the importance of landscape elements occurring between the locations or populations of interest (Pfluger and Balkenhol 2014). Resistance surfaces that describe the tendency or ability of individuals to move through this landscape matrix are compared to genetic variability when evaluating spatial genetic structure (Spear et al. 2010, Sawyer et al. 2011, Zeller et al. 2012). More geographically proximate habitats or populations are expected to be more genetically similar owing to a reduction in the number of migrants with increasing distance (Guillot et al. 2009). Ecologically, this Isolation by Distance (IBD) pattern is thought to result from cumulative fitness costs associated with movement or dispersal limitation, thereby making long-distance dispersal events rare (Bowler and Benton 2005, Bonte et al. 2012, Sexton et al. 2013). For non-migratory species, where individual dispersal distances are small relative to the area occupied by the population, it might be expected that IBD would be a good approximation of historical population connectivity (Guillot et al. 2009). However, migratory species such as diadromous fishes are unlikely to be regulated by cumulative fitness costs associated with movement among spawning habitat, given that lifetime dispersal distances are orders of magnitude greater than the area occupied by breeding or reproducing individuals. An alternative model for historical patterns in gene flow could be an Isolation by Environment or IBE model, which posits that gene flow is greatest among similar environments owing to either non-random mating or local adaptation (Sexton et al. 2013, Wang et al. 2013). In this instance, genetic variability among populations would be expected to be related to the degree of similarity in recipient environments (promoting local adaptation) or possibly to the degree of difference (counter-gradient gene flow against local adaptation) (Sexton et al. 2013). To maximize reproductive fitness, dispersing individuals would be expected to respond to the environmental characteristics of recipient habitats through habitat selection (Kubisch et al. 2014) rather than to the characteristics of the environment between habitats.

The overall goals of this chapter were to use a landscape genetics approach to: (1) evaluate the IBD and IBE hypotheses relative to population structuring in a diadromous fish species and (2) evaluate whether residual genetic variation could be explained relative to the magnitude or importance of multiple contemporary threats. To achieve this, I used a case study of genetic variation among 11 endangered Atlantic salmon (*Salmo salar*) populations, where the results could be used directly to facilitate recovery planning. These analyses are structured into four main parts. The first part describes the methods used to obtain and summarize the genetic and landscape data sources, which were transformed into genetic and environmental distance matrices for subsequent analyses. The second component compares the relative support for the IBD and IBE hypotheses of gene flow. The third part explores whether residual genetic variation could be explained by anthropogenic threats, which relied on developing optimal indicators of three general categories of anthropogenic effects (those related to sedimentation, fragmentation and water quality). The fourth component evaluated the relative magnitude and significance of any correlation between residual genetic variation and each threat category. From a theoretical standpoint, the results potentially provide insight into the behavioural processes that may be leading to gene flow among watersheds for Atlantic salmon. From a practical standpoint, they also provide one of the first quantitative comparisons of the relative magnitude of threats among these populations, and can be used to facilitate recovery planning.

Methods

This chapter uses data from 11 Atlantic salmon populations in the Southern Upland (SU) region (Figure 4-1). These tend to inhabit the larger river systems and are well distributed throughout the region (i.e., they are not clustered geographically). As such, they are thought to largely represent the range of variability that exists among populations in the SU. A total of 709 genetic samples from juveniles (fin clips from young-of-the-year, age 1 or age 2) were collected opportunistically during electrofishing surveys, with the timing of each collection and the number of samples varying slightly among rivers (Table 4-1). During data collection, individual samples were identified by the river of origin instead of by individual. Because the geographical coordinates representing individual sampling locations were unknown, the landscape genetics analysis described below had to be done at a population level. Older juveniles were sampled preferentially to minimize the likelihood of multiple sampling of individual families. Microsatellite amplification was carried out at 17 individual loci, where the methods used are detailed in O'Reilly et al. (2012).

From the microsatellite dataset, the potential for null alleles (non-amplified alleles that result in apparent homozygotes), allele dropout (preferential amplification of short alleles) and scoring errors due to stuttering (failure to discriminate adjacent alleles) was assessed using Micro-checker version 2.2.3 (van Oosterhout et al. 2004). I used 8000 Monte-Carlo randomizations of the observed data to calculate the p-values for the χ^2 test of observed vs. expected homozygote frequencies and heterozygote size differences. I also tested for Hardy-Weinburg (HW) equilibrium over loci in all sampled populations using the R package 'adegenet' (Jombart 2008). At each locus within each population, I used 8000 Monte-Carlo simulations to calculate p-values for the χ^2test of observed vs. expected gene frequencies. After Bonferroni correction, I assessed significant departures and any patterns of departures over loci or within populations (Table A4-1-2, Appendix 4-1). Finally, I assessed the potential for linkage disequilibrium (non-random association of alleles at multiple loci) using LinkDos (Garnier-Gere & Dillmann 1992) from GenePop 4.2 (Raymond & Rousset 1995; Rousset 2008) (Table A4-1-3, Appendix 4-1). After the above, I calculated five within-population measures of genetic variability, averaged over loci, to better understand the characteristics of individual populations: allelic richness (i.e., the number of alleles), the effective number of alleles (allelic richness standardized to account for sample size), expected heterozygosity (the expected probability that an individual will be heterozygous at a given locus), observed heterozygosity, and Wright's inbreeding coefficient (F_{is}) within the R package 'gstudio' (Dyer 2012).

Watershed characterization

These analyses required metrics from which the environmental characteristics of watersheds could be described (to parameterize an IBE model), and from which the extent of human activities within watersheds could be characterized (for the evaluation of threats). Freely available data sources describing the physical, geological and land use characteristics of the watersheds sampled for these analyses were combined and analyzed using ArcGIS (ESRI® ArcGIS 10.0 software service pack 3). A detailed description of the GIS methods and data layers is provided in Bowlby et al. (2014). These were used to calculate: (1) the proportion of catchment area composed of multiple geological features within a watershed (e.g., surficial geology types), (2) the proportion of watershed area affected by multiple land use types (e.g., forestry) or (3) the prevalence (i.e., a count) of a specific type of threat (e.g., road/river crossings) in each watershed. In the case of counts, values used in these analyses were expressed as a density, scaled either by the area of the watershed or by the length of the stream network. The sources of geographic data that were used are provided in Table A4-2-1 of Appendix 4-2, while the calculated landscape variables describing the watersheds are provided in Appendix 4-2, Tables A4-2-2 to A4-2-6 inclusive, and the land use activities (i.e., threats) are in Table A4-2-7.

It was necessary to determine the environmental metrics that would be most likely to influence habitat quality for Atlantic salmon for inclusion in the IBE model. Given the hierarchical nature of watersheds, where landscape characteristics affect processes at successively smaller scales (Allan 2004), it is thought that the evolution of life history differences as well as population dynamics of salmon are primarily governed by large-scale characteristics of watersheds (Ugedal and Finstad 2011). Therefore, the data types used in the IBE model included: area metrics (e.g., watershed area, length of stream network), surficial and bedrock geology, natural forest disturbance regimes, and topography (Appendix 4-2; Tables A4-2-2 to A4-2-6, respectively). To avoid weighting subsequent analyses due to different numbers of variables contributing to each information type (e.g., 4 variables describe natural disturbance regime while 14 describe surficial geology), each of the five data types were scaled to ensure comparability and then independently transformed into two non-metric multidimensional scaling (NMDS) axes (Bocard et al. 2011) using the metaMDS function of the R package ‘vegan’ (Oksanen et al. 2013). The Euclidean distance matrix representing landscape variation among watersheds was calculated from these 10 NMDS axes (i.e., the sum of 2 axes representing each of five data types), as described below.

Historical patterns of gene flow

If dispersal limitation was the primary determinant of gene flow among populations (as predicted by the IBD hypothesis), it would be expected that the geographic distance among watersheds would be positively correlated with the degree of genetic difference among populations (Guillot et al. 2009). Alternately, if habitat selection and local adaptation to recipient environments was the primary determinant of gene flow (as predicted by the IBE hypothesis), a stronger correlation may exist between environmental similarity and genetic distance (Sexton et al. 2013). It would also be expected that the better model would have significant explanatory power after any correlation with the other hypothesis was accounted for. To evaluate these hypotheses, I first developed genetic, geographical and environmental distance matrices and compared those using Mantel tests and two-step Reciprocal Causal Modeling (Wasserman et al. 2010). The Mantel tests evaluate whether a single model can explain a significant amount of genetic variation. The Reciprocal modeling (described below) is based on a comparison of partial Mantel tests and evaluates whether IBD or IBE can be considered significantly better than the other model. Looking forward, there are many exciting statistical techniques that are being proposed as more appropriate and rigorous alternatives to Mantel and partial Mantel tests for evaluating IBD relative to IBE as well as other questions in landscape genetics. These include Generalized Dissimilarity

Modelling of distance matrices (Ferrier et al. 2007, Fitzpatrick et al. 2015), spatial regression methods based on conditional or simultaneous autoregressive models of individual-based data (Wagner & Fortin 2016), and Bayesian methods such as BEDASSLE that model the covariance in allele frequencies from SNPs as functions of geographic or environmental data (Bradburd et al. 2013), among others. Unfortunately, none of these could have been applied here as they require a much more comprehensive data set than was available.

To evaluate the strength of the correlation between genetic distance and either the IBD or IBE models relative to a null distribution of no association, I used Mantel tests calculated from 10,000 permutations of the data, where significance was assessed relative to the 95% quantile of the distribution of permutations (i.e., the test statistic would be significant only if it was greater than 95% of the permuted values). The genetic distance matrix was calculated from Nei's pairwise F_{st} (Nei 1978) using the R package 'gstudio' (Dyer 2012). This distance is given by:

$$\widehat{D} = -ln\left[\sqrt[\widehat{G_{xy}}]{\widehat{G_x}\widehat{G_y}}\right]$$

where $\widehat{G_x}$, $\widehat{G_y}$ and $\widehat{G_{xy}}$ represent the bias-corrected averages of allele frequencies over the *r* loci sampled (Nei 1978). The correction prevents overestimating distances based on observed gene frequencies when sample sizes are small. For this calculation, differentiation among populations is assumed to arise through constant mutation rates and genetic drift. The Euclidean matrix representing pairwise geographical distances (IBD model) was calculated from the coastal straight-line distance between each pair of rivers, which represents a minimum distance over water; beginning and ending on the latitude and longitude between each pair of river mouths. Euclidean distance is merely a straight-line distance between two points: $d(p_i, q_i) = \sqrt{(q_i - p_i)^2}$, where p_i and q_i represent the geographic coordinates (i.e., *x* and *y* co-ordinates) for each river (Oksanen et al. 2013). The Euclidean matrix representing pairwise environmental distances (IBE model) was calculated from the centroids in multivariate space using the NMDS axes described above. This Isolation by Environment (IBE) model explicitly did not include any positional information representing the geographical location of watersheds. The IBD and IBE distance matrices were highly correlated (Spearman correlation coefficient = 0.92), which was expected because environmental gradients tend to be spatially autocorrelated (Legendre 1993). Values for each distance matrix are given in Appendix 4-3.

Second, I used the two-step Reciprocal Causal Modeling framework proposed by Wasserman et al. (2010), in which competing models (IBD or IBE) are compared directly to each-other using partial Mantel tests based on the same distance matrices as above. Reciprocal Causal Modeling has been advocated as a method for formal significance testing among highly correlated distance matrices (Cushman et al. 2006, Wasserman et al. 2010). Previous evaluation of the technique using simulated data has suggested that it is efficient in identifying the correct model from a range of alternatives in individual-based landscape genetic analyses (Cushman et al. 2013). However, it suffers from elevated Type I error rates, as do most analyses on highly correlated data (Balkenhol et al. 2009, Legendre 1993). The magnitude of Type I error can be related to the degree of second-order autocorrelation (i.e., the covariance function between observations) in the matrices being compared (Guillot and Rousset 2013). Simulation testing by Cushman et al. (2013) suggests that inference on the optimal model relative to a range of highly correlated alternatives remains correct when based on a comparison of the relative magnitudes of partial Mantel *r* values, even when the potential for Type I error is high. Therefore, as advocated by Cushman et al. (2013), I have also used the relative magnitudes of the Mantel *r* values for assessing the comparative support among models (e.g., Castillo et al. 2014), rather than the significance of individual models in isolation. Reciprocal Causal Modeling works by comparing a partial Mantel test of a given hypothesis after parcelling out the influence of a competing hypothesis and vice versa (e.g., comparing the *r* statistic of Gene~Model1|Model2 with Gene~Model2|Model1). The null hypothesis for these partial Mantel tests is that there is no additional variation captured in the second distance matrix relative to the first, indicating greater explanatory power for Model1. The expectation would be that a better model would have a large and significant *r* value relative to the alternate hypotheses in the first instance, while the *r* value would be small and non-significant in the second instance (although see comments related to Type I error above). In other words, competing models would not show a significant correlation with genetic structure after the influence of the chosen model was accounted for and vice versa. On the basis of the comparisons between p-values, the chosen model can be taken as significantly better than the competing models (Wasserman et al. 2010, Cushman et al.

2013). In an evaluation of Reciprocal Causal Modeling (RCM) using simulated landscape resistance hypotheses, Cushman et al. (2013) suggested that for a resistance model to be formally affirmed as the only hypothesis supported (i.e. significantly better than all competing hypotheses), the following pattern would be observed. First, all of the partial Mantel comparisons between the chosen model while controlling for the effect of each alternative model would be significant (p-values < 0.05). Second, all partial Mantel comparisons between the alternate models relative to the chosen had to be non-significant (p-values > 0.05). All Mantel tests and partial Mantel tests were carried out using the R package ‘vegan’ (Oksanen et al. 2013).

Identification of variables contributing to threats indices

In general, the land use activities identified as threats to Atlantic salmon have been linked to changes in the environmental processes governing watersheds, rather than to changes in population dynamics (Allan 2004). In addition, different types of threats are expected to have the same overall effect on environmental processes, as detailed in the Introduction. Therefore, to quantitatively assess the influence of threats on genetic connectivity among populations, it becomes necessary to first determine which environmental process a threat is most likely to contribute to, and then to assess the relative magnitude of effect from multiple environmental processes. Here I consider a suite of eight land use variables as potential threats and evaluate their independent contributions to three more general environmental processes: sedimentation (S), fragmentation (F) and water quality (WQ) (hereafter called the three threats indices;

Table 4-2; Table A4-2-7, Appendix 4-2). This was done both as a method for variable reduction and as a way to group activities that would be expected to influence populations in similar ways. Albeit indirectly, retention of a specific land use variable in a threat index provides support for both the hypotheses on how specific land use types influence populations as well as whether or not these activities can be considered to be substantial threats for these specific populations. A literature review was used to identify the manner in which specific land use types are thought to influence Atlantic salmon populations or their habitat characteristics. For example, agricultural activities, industrial sites (such as gravel quarries) and roads are considered to be chronic sources of sediments (Gilvear et al. 2002), where increased sedimentation has been linked to reduced habitat quality for salmonids as well as to acute mortality of eggs or juveniles (Soulsby et al. 2001, Julien and Bergeron 2006), particularly during storm events (Lisle 1989). Furthermore, the bedrock geology of the Southern Upland region has little buffering capacity (Watt et al. 1983, Korman et al. 1994), which makes Southern Upland salmon particularly vulnerable acid deposition from either precipitation or land-based sources (Farmer et al. 1980, Watt 1987). Atmospheric deposition in the form of acid rain has substantially reduced pH in many rivers throughout the Southern Upland (Watt 1987). Lastly, salmonids depend on unobstructed movement in a watershed to access spawning and rearing areas, avoid predators, and respond to changing environmental conditions such as temperature, flow, or inter- and intra-species competition (Poplar-Jeffers et al. 2009). Threats that fragment watersheds, such as road/river crossings containing culverts or impassable dams, would be expected to reduce habitat productivity as well as accessibility for Atlantic salmon. Although other types of land use would be expected to influence freshwater environments (e.g., urbanization), the incidence of such threats in the Southern Upland region was very low (e.g., < 1% of watershed area affected) and quite similar among watersheds (data not shown), and so were not included. In terms of duration, each of the land use variables considered here would have become significant prior to the 1970s (Bowlby et al. 2014), which corresponds to more than five generations prior to genetics sampling for these populations.

The preferred model of each threat index was identified using the two-step Reciprocal Causal Modeling framework proposed by Wasserman et al. (2010), introduced above. A separate Euclidean distance matrix was calculated for each potential combination of the land use variables hypothesized to be contributing to each threat index (Table 4-2). Because land use variables were similarly scaled prior to calculating Euclidean distance, included land use types contribute equally in the model. As in Shirk et al. (2010), the model that had the highest correlation with genetic distance was initially identified using Mantel tests based on Spearman correlations (e.g., Gene~Model) and partial Mantel tests (Smouse et al. 1986) controlling for the influence of underlying genetic structure (e.g., Gene~Model|IBE). Here I used the IBE model only in order to avoid overestimating the effect of land use on gene flow, given that spatially correlated landscape features can mask or confound the independent contribution of contemporary land use effects on genetic distance (Dileo et al. 2013). However, I recognize that essentially equivalent results could have been obtained in this study by controlling for IBD. Next, the relative support for the chosen model (which I will now call ‘Chosen’) was assessed using reciprocal comparisons with each competing threat model (which I will call an ‘Alternate’) following: Gene~Chosen|Alternate vs. Gene~Alternate|Chosen. Again, the expectation would be that a better model would have a large and significant *r* value relative to all of the other hypotheses in the third test, while the *r* value would be small and non-significant in the fourth test. Using Reciprocal Causal Modeling to evaluate the contribution of land use variables to each of three threats indices necessitated two assumptions: (1) that the effects of land use were additive and (2) that specific types of land use contributed to one main threat index. By this I mean that road density (for example) would be mainly related to sedimentation, rather than having nearly equivalent contributions to the fragmentation and sedimentation indices. This assumption was evaluated *a posteri* by calculating the Spearman correlation coefficient among the threat indices as well as by substituting alternate land use variables into each threat index (e.g., considering road density rather than road crossings in the fragmentation index).

I have detailed the computational sequence followed in this section relative to a specific example to aid in interpretation (Appendix 4-4).

Influence of anthropogenic land use

To evaluate if any of the threats indices explained significant residual genetic variation (over and above that explained by IBE), I used Multiple Regression on Distance Matrices (MRDM; Legendre et al. 1994, Lichstein 2007). The response was the genetic distance matrix and the possible predictors were the three threat indices. MRDM works by transforming each distance matrix into a vector, calculating the linear regression coefficients, and then estimating significance by holding the predictors constant while randomly permuting the response (Lichstein 2007). For the retained predictors, a positive regression coefficient would indicate increased genetic isolation relative to that expected historically and would provide evidence that the variables contributing to the threat index have had measureable effects on population dynamics. Here, the relative support for various models was assessed using the Akaike Information Criterion for small samples (AICc; Burnham and Anderson 2002). Although individual data points in a distance matrix are not independent, this non-independence is the same in each model and thus does not affect model ranking, making information theoretic approaches (Johnson and Omland 2004) valid for model selection (Engler et al. 2014).

Results

Population genetic diversity

Meaningful comparisons of genetic diversity at neutral markers require that data are not biased by genotyping errors (e.g., null alleles), that all populations have been sampled representatively, and that the tested loci are not under selection (i.e., that they are in Hardy-Weinburg equilibrium (HWE); Allendorf et al. 2013). Round Hill River and Salmon River (Guysborough County) were unlikely to have been sampled representatively, given that juvenile salmon were only captured at one of the two sites electrofished on either river. Clustering among genetically similar individuals would lead to a non-representative sample at the population level (Schoville et al. 2012) and would be expected given the population structuring exhibited by Atlantic salmon and the tendency to home to specific places within a watershed (Keefer & Caudill 2014). The low genetic diversity in Round Hill River relative to the other 10 population sampled (Table 4-1) supports this conclusion, as does the prediction that one quarter of tested loci were out of HWE for Salmon River (Guysborough County) (Table A4-1-2, Appendix 4-1). Therefore, these rivers were removed from further analyses. The remaining rivers exhibited no systematic departures from HWE either within loci or populations. The tested loci and alleles appear to sort independently (i.e., do not exhibit linkage disequilibrium), given that correlations among alleles were extremely low (< 0.1) and significant results were not systematically distributed among samples (Table A4-1-3, Appendix 4-1). There was no evidence for scoring error or allele dropout, yet null alleles were found at locus 9 in five populations (Country Harbour, LaHave, Gold, Moser, and St. Mary's) and at locus 7 in Salmon River (Digby County) (Table 4-1-1, Appendix 4-1). Although accounting for null alleles would slightly change the allele frequencies, the relative differences tend to be preserved (i.e., the frequencies all decline, rather than some increasing and some, decreasing). Therefore, null alleles would not be expected to bias analyses based on pairwise distances, an expectation that I confirmed by both including and excluding locus 9 in the analyses. The results I present are based on the unadjusted allele frequencies at all loci. Measures of within-population genetic variation were similar

(Table 4-1) and indicated relatively high levels of genetic diversity. Expected and observed heterozygosity were higher than have been previously reported for anadromous fishes (DeWoody and Avise 2000), but were consistent with estimates derived from microsatellite data from other salmon species (Jia et al. 2012, Khristaleva et al. 2014).

Population genetic structure

The Mantel test reveals a higher correlation (and one much less likely to have occurred by chance) between Nei's pairwise F_{st} and the landscape distance matrix (r = 0.392, p-value = 0.007) as compared to coastal geographic distance (r = 0.373, p-value = 0.020) (Table 4-3). Comparing the IBD and IBE hypotheses using Reciprocal Causal Modeling suggests that the IBE model may be better supported by the data, given that the partial Mantel r correlation is substantially higher between genetic isolation and the landscape distance matrix after controlling for any correlation with geographic distance (Gene ~ IBE|IBD; r = 0.176), as compared to the reverse (Gene ~ IBD|IBE; r = <-0.001) although neither p-value is significant at a level of 0.05 (Table 4-3). However, relatively little (15.2%) of the permuted null distribution is larger than 0.176 in the Gene ~ IBE|IBD comparison, while half of the permuted null distribution is larger than the r value in the Gene ~ IBD|IBE comparison.

Identification of variables contributing to threats indices

Mantel and partial Mantel tests between genetic distance and a fragmentation index, F, (where each possible combination of road crossings (Rc), dams (D) and pH (pH) was assessed independently) showed the largest correlation with the RcD model (road crossings and dams; Table 4-4). Comparisons of this model ('Chosen') to genetic distance while controlling for each of the other candidate models ('Alternate') had Mantel *r* values ranging from 0.229 to 0.283; although each of the p-values were greater than 0.05 (Table 4-4). Conversely, none of the other models retained any significant correlation with genetic distance once the effect of the RcD model was accounted for, and the absolute values of the Mantel *r* correlations were substantially lower (< 0.135; Table 4-4). Although these results do not suggest that the RcD model can be considered significantly better than each of the other candidate models, the relatively large Mantel *r* correlations when the competing models are accounted for (and vice versa) suggests that this model is the best representation of the fragmentation index.

The Mantel and partial Mantel tests relative to the water quality index (WQ) were unexpected, in that the correlations tended to be strongly negative, suggesting greater population connectivity among more dissimilar watersheds related to mine density, forestry or mean annual pH. The largest negative Mantel *r* correlation was with the MP model (mines and pH). Relative to the competing models, the MP model retained a large negative correlation with genetic distance (*r* values < -0.33; Table 4-4), while the competing models had smaller correlations with genetic distance when the effect of the MP model was accounted for. Based on a comparison of the p-values between Chosen|Alternate (expected to be significant) and Alternate|Chosen (expected to be non-significant), the MP model was significantly better than all of the competing models. It was chosen as the best representation of the water quality index.

Unlike the water quality and fragmentation indices, there was no single model that clearly had the largest correlation with genetic distance for the sedimentation index. With the exception of the agriculture only model (A), the partial Mantel tests demonstrated correlations above 0.48 for each of the other candidate models (Table 4-4). The model that included industry, road density and agriculture (IRdA) had among the highest *r* values for both the Mantel and partial Mantel tests, and so was chosen as the best representation of the sedimentation index (S) for Reciprocal Causal Modeling. However, sensitivity to this choice was also assessed by evaluating the IRd, ARd, IA and Rd models using Reciprocal Causal Modeling (data not shown). The IRdA model had no additional explanatory power relative to the IRd model (*c.f.* $r = 0.146$ and 0.148; Table 4-4), the ARd model (*c.f.* $r = 0.191$ and 0.126; Table 4-4) and the Rd model (*c.f.* $r = 0.174$ and 0.177; Table 4-4); there are essentially equivalent correlations with residual genetic variation in the reciprocal tests. In addition, there is still a correlation for the IRdA model after accounting for the influence of the IA model ($r = 0.194$) while there is not for the IA model after accounting for IRdA ($r = 0.008$; Table 4-4). Similar patterns are found when the IRd, ARd, IA and Rd models are evaluated. All models excluding Rd have essentially no additional explanatory power, while models including one or both of I and A (in addition to Rd) have essentially equivalent explanatory power in the reciprocal tests. Taken together, these results suggest to me that road density is the primary land use variable contributing to the sedimentation index. Industry and agriculture have comparatively little contribution. Therefore, following the principle of parsimony, I chose the Rd model as the best representation of the sedimentation index, although I cannot discount the possibility that industry and agriculture have a lesser contribution.

Influence of anthropogenic land use

The three threats indices were relatively uncorrelated, with Spearman correlation coefficients of < 0.4, and thus unlikely to violate assumptions regarding collinearity for inclusion in multiple regressions (Zuur et al. 2010). In isolation, the IBE model explained very little of the variation in genetic distances (R-square = 0.164, p-value = 0.005; Table 4-5), even though the variable was significant in the regression. Results were very similar between the full model including all of the indices (IBE+S+F+WQ) and the model that excluded the fragmentation index (IBE+S+WQ), although the latter was selected as the preferred model on the basis of AICc; *c.f.* R-square = 0.547 and 0.536, p-values = 0.001 and <0.001, AICc = -108.71 and -110.01, respectively (Table 4-5). As expected, the fragmentation index was not a significant predictor in the full model (slope = 0.11; p-value = 0.529). All other candidate models had R-square values of less than 0.4 (Table 4-5). It is important to remember that R-square values from MRDM regression (as well as *r* statistics from Mantel tests) cannot be interpreted as the proportion of variance in the response explained by the predictors, but only as a measure of the fit of a linear model to the paired sets of distances (Legendre and Fortin 2010). Given that each of the Euclidean distance matrices included in the MRDM models were similarly scaled, the slope estimates from each regression model give a direct comparison of the relative magnitude of each threat index on genetic distance among watersheds. Absolute values of the slope estimates from the sedimentation (slope = 0.44, p-value = 0.012) and water quality (slope = -0.39, p-value = 0.013) indices were similar to those from the landscape distance matrix (slope = 0.41, p-value = 0.007).

Discussion

My analyses do not provide significant support for the hypothesis that environmental variation is the primary determinant of gene flow in Atlantic salmon at a regional scale, as opposed to a pattern of movement determined by geographic isolation, but they do suggest that it is possible. In addition, the behavioural mechanism proposed here to lead to an Isolation by Environment (IBE) model is consistent with the interpretation of genetic variation arising from contemporary threats, while the IBD model is not. In relation to population-level responses to threats, findings are consistent with multiple experimentally-predicted behavioural or physiological changes, and suggest that straying behaviour and subsequent reproductive success are affected by anthropogenic land use in watersheds. Based on this result, landscape genetics could become a powerful basis from which to develop future research priorities and remediation strategies to address population declines in Atlantic salmon and potentially other endangered species.

Population genetic structure

There is a strong theoretical basis by which to argue that Isolation by Distance may be insufficient to describe straying behaviour, effective dispersal and historical patterns of population connectivity in Atlantic salmon, and potentially other migratory species. This contrasts most research on population structuring in diadromous fishes, where IBD is taken to be the expectation or null model of gene flow (e.g., Palstra et al. 2007, King et al. 2001, Bradbury et al. 2014). However, the theoretical concept of Isolation by Distance was developed to represent a specific type of dispersal process, one limited by the cumulative fitness costs associated with movement (Slatkin 1993, Guillot et al. 2009). Such a pattern is unlikely to describe movement among watersheds in Atlantic salmon and other diadromous fishes, given that oceanic migrations are vast relative to any separation among river mouths (Hansen and Quinn 1998, Ritter 1989). Alternatively, the evolution of homing behaviour as well as substantial life history variation among populations provides indirect evidence of local adaptation (Garcia de Leaniz et al. 2007, Fraser et al. 2011), which implies that individuals have higher reproductive success in environments similar to their natal environment (Pfluger and Balkenhol 2014) and could be motivated to seek out similar habitats (Bonte et al. 2012). Thus, among returning adults, habitat selection would be a more likely behavioural motivation for watershed choice while straying, where subsequent reproduction (i.e., effective dispersal) would also be expected to be higher in more similar recipient watersheds. Proximate cues to assess habitat quality could be olfactory (e.g., dissolved mineral content, oxygen saturation) or environmental (e.g., flow rate, temperature); factors that are linked to landscape characteristics and environmental variation among watersheds (reviewed in Allan 2004). If straying adults are using such proximate cues, gene flow among watersheds should correlate with the environmental characteristics of watersheds. Although a marginally stronger correlation existed with IBE as IBD, my IBE model did not explain significantly more genetic variation. Variability in this relatively limited data set coupled with the low power of Mantel tests (Legendre et al. 2015) would hinder the description of any underlying relationship, as could strong spatial autocorrelation between environmental variation and geographic distance (here estimated at 0.92), making IBE closely approximate IBD. This latter pattern would be expected to break down under weak spatial autocorrelation, and could explain why IBD is both supported (e.g., Dionne et al. 2009, King et al. 2001) and not (e.g., Bradbury et al. 2014) in the literature for Atlantic salmon. Given the theoretical basis for the underlying assumptions, I propose that future research on gene flow among populations (i.e., among watersheds) in diadromous species should develop an IBE model in addition to IBD. The potential exists for environmental variation to affect the dispersal process directly, rather than acting exclusively on within population processes such as genetic drift or adaptation.

The negative Mantel *r* correlation between genetic distance and the water quality index can only be explained if returning adults are actively assessing the characteristics of the watersheds they encounter (implicit in the IBE model) as opposed to incurring cumulative fitness costs associated with movement behaviour among watersheds (implicit in the IBD model). Most landscape genetics questions dealing with a species' response to threats test for the effect of barriers or some type of land use that increases resistance to movement (through degrading habitat quality) and thus reduces the ability or tendency to move through areas more heavily affected by the threat (e.g., Sawyer et al. 2011, Zeller et al. 2012). From these examples, IBD is predicted to become stronger than what would be expected historically and any Mantel correlation between genetic distance and the magnitude of the threat would be inherently positive (i.e., high levels of the threat would cause a greater reduction in movement among populations). However, if individuals have the ability to avoid areas more heavily affected by a threat (e.g., particular watersheds) while moving freely through the marine environment, negative correlations become possible, such as the one seen here between the water quality index and genetic distance.

The negative correlation suggests *increased* gene flow among environmentally *dissimilar* watersheds (based on mines+pH), which would be consistent with individuals preferentially entering or experiencing higher spawning success in rivers that are less acidified or contain fewer contaminants relative to their natal watershed. One hypothesis is that observed patterns of genetic divergence could result from a combination of: (1) a reduced capacity to imprint using olfactory cues, leading to an increase in the proportion of a population that strays; and (2) relatively lower survival in watersheds with poor water quality, leading to higher survival of individuals straying to watersheds with good water quality. If chemical contaminants are in the watershed, it seems unlikely that they would influence behaviour in isolation of physiology and vice versa. Besides the potential for acute mortality caused by low pH (e.g., Lacroix and Townsend 1987), acidification is known to interfere with chemosensory functions related to the detection and response to chemical signals (Leduc et al. 2010). One chemosensory function would be the olfactory imprinting that occurs during emigration to the marine environment and enables individuals to return to natal rivers (McCormick et al. 1998). Similarly, a range of chemical compounds (including insecticides used in forestry practices in Nova Scotia; Fairchild et al. 1999) as well as heavy metals (such as those typically present in acid rock drainage from old mine sites; Akcil and Koldas 2006) can have comparable effects on chemosensory function (reviewed in Lurling and Scheffer 2007). Preliminary tag-based research on sub-lethal exposure to an insecticide corroborates this hypothesis, in that homing success was lower for exposed Pacific salmon relative to controls (Scholz et al. 2000). Similarly, higher straying rates among localized stream reaches have been observed when conditions at the natal site were less favorable (Dittman et al. 2010, Cram et al. 2013). A recent review of homing behaviour in anadromous salmonids indicates that environmental conditions influence attraction to non-natal sites and that any chemicals interfering with olfactory imprinting or sensory development would be expected to increase the incidence of straying as well as to influence adult habitat choice (Keefer and Caudill 2014).

Practical implications

During recovery planning for endangered species, threats that tend to be well-understood are often focused on for remediation at the expense of those that are not

(Lawler et al. 2002, Norris 2004). The specific threats identified as contributing to the sedimentation, fragmentation and water quality indices contradict some current perceptions on the optimal focus of recovery planning. It is generally accepted that pH has an extremely large influence on Atlantic salmon in the Southern Upland, where increases in acid precipitation from the 1950s have been implicated in population extirpations and severe population decline (Watt et al. 1983, Lacroix 1989). However, the influence of past mining activities on current water quality has never been identified as a pressing issue for mitigation, although it has been identified as a potential threat (e.g., Bowlby et al. 2014). The present analysis suggests that mining activities affect population connectivity (movement and spawning success) in the same manner and with the same severity as acidification. Similarly, road crossings and fragmentation have been identified as extremely important determinants of freshwater habitat accessibility for salmon, and many user groups are interested in culvert and dam removal or remediation as a method for improving habitat accessibility and presumably increasing population size for Atlantic salmon (Langill and Zamora 2002, Gibson et al. 2005). Here, the larger concern appears to be road density itself, given that the fragmentation index had no significant effect on population connectivity while the sedimentation index did. The presence of roads has been linked to a suite of ecological changes in watersheds beyond sedimentation: including changes in thermal regimes, constriction of channel movement, and changes in channel morphology, contaminants, as well as the spread of invasives in watersheds (Trombulak and Frissell 2000). It has been argued for a wide variety of taxa that the effects of roads on natural populations are one of the most pressing current conservation issues (Forman and Alexander 1998, Clark et al. 2010). It is interesting that my analysis suggests that the presence of roads causes changes in population connectivity of a similar magnitude relative to changes in water quality (i.e., the slope estimate for the S index is comparable to that of the WQ index). However, I have not explicitly linked these results to a population dynamics model used to assess status, so it is unknown how the opposing effects of the two threats culminate in changes in abundance.

Landscape genetics as a method to evaluate threats

It is well-known that the influence of specific environmental factors on systems or populations can vary over different scales (e.g., Schneider 2001), which has implications for how well this population-level analysis can fully characterize the relative magnitude and influence of this suite of threats affecting adult Atlantic salmon populations, particularly the fragmentation index. In general, adult habitat choice can be considered to result from processes at two scales: first, the selection of a watershed to enter from the marine environment (large scale), and second, the selection of a spawning location within that watershed (small scale). Starting from the premise that adult salmon assess watershed characteristics when choosing a watershed for spawning (as suggested by IBE), only threats that influence the characteristics of the water being discharged from the river could be perceived by individuals at a large scale and used as proxies for habitat quality. Of the variables considered here, such threats would be contributing to the sedimentation and water quality indices; the two threat indices identified as significant predictors of population structuring in the regression analyses. At smaller scales, the manner in which adults would be expected to interact with their environment changes, in that movement behaviour is now influenced by the spatial arrangement of suitable spawning substrate as well as the relative connectivity among stream reaches (Waples et al. 2008, Sabo et al. 2010). Stated another way, gene flow within a population would be partially determined by an individual’s ability or motivation to move through the habitat mosaic, as well as by habitat suitability. It may be more appropriate to characterize threats such as full or partial barriers (e.g., road-river crossings) as well as localized habitat degradation caused by land use (e.g., an agricultural field adjacent to a particular stream reach) using resistance surfaces developed for a dendritic landscape (i.e., one in which animals are constrained to move along the stream network; Fagan 2002, Hopken et al. 2013).

Detectability, sublethal effects and cumulative effects (among others) all complicate attempts to quantify relationships between abundance and environmental variability or anthropogenic effects on fish populations (Rose 2000), undermining effective conservation and hindering timely recovery planning (Lawler et al. 2002, Norris 2004). Using genetic variation (as opposed to abundance time series) to investigate population-level responses to threats addresses the latter two issues: multiple correlated hypotheses can be directly compared (assessing cumulative effects) and population-level responses can give indirect evidence of physiological change (sublethal effects), as seen here in relation to the water quality index. Secondary benefits are that these analyses can be done at the level most appropriate for recovery planning (i.e., for populations), can be completed relatively rapidly in that they do not rely on the collection of time series data. Different threats measured in different units (e.g., proportions, counts, or densities) can be directly compared, with the caveat that relationships between genetic distance and environmental distance are assumed to be linear. Although landscape genetics would not be an appropriate method in situations where population decline was extremely rapid (given that there would not be adequate time for evolutionary effects to become detectable; Landguth et al. 2010, Dileo et al. 2013), it would be appropriate in the majority of situations where declines are observed for many years prior to their identification as a conservation concern, provided the same suite of threats are influencing populations over the entire time period. Although these analyses only provide a way to compare the influence and magnitude of threats relative to each-other among a group of watersheds (rather than developing predictive relationships between specific populations and a suite of threats), they appear to be a powerful basis from which to develop future research priorities and remediation strategies to address population declines.

Tables

Table 4-1. River of origin (River), sample size (*N*), year of collection (Year), expected heterozygosity or genetic diversity (H_{exp}), observed heterozygosity (H_{obs}), Wright's inbreeding coefficient (F_{is}), the effective number of alleles (A_e) and allelic richness standardized to a sample size of 26 individuals (*A*). Rivers are listed relative to geographic location along the coast of Nova Scotia (Refer to Figure 4-1).

River	Year	N	H_{exp}	H_{obs}	F_{is}	A_e	A
Round Hill	2000	28	0.664	0.725	-0.090	3.677	7.392
Salmon River (Digby County)	2000	44	0.791	0.811	-0.021	7.346	11.511
Tusket	1999	60	0.831	0.857	-0.034	7.846	11.619
Medway	2001	83	0.824	0.828	-0.005	9.433	13.044
LaHave	2000	49	0.814	0.809	0.009	7.543	11.971
Gold	2001	84	0.827	0.804	0.033	8.549	12.042
Musquodoboit	2000	53	0.834	0.859	-0.029	7.972	11.632
Moser	2000	58	0.801	0.822	-0.027	7.332	11.899
St. Mary's	2000	78	0.830	0.812	0.025	9.552	13.187
Country Harbour	2000	42	0.834	0.828	0.011	8.407	12.092
Salmon River (Guysborough County)	2009	30	0.817	0.796	0.018	8.881	13.323

Table 4-2. Variable definitions for the Reciprocal Causal Modeling used to assess the contribution of various land use activities to three threat indices: Fragmentation, Sedimentation and Water quality. Values for each variable were scaled between zero and one, summed, and transformed into a single Euclidean distance matrix for analyses.

Threat	Variables	Description
Fragmentation[2]	RcPD	Road crossings+pH+dams
Fragmentation	RcP	Road crossings + pH
Fragmentation	RcD	Road crossings + dams
Fragmentation	DP	Dams + pH
Fragmentation	Rc	Road crossings (# road crossings per 10 km of stream length)
Fragmentation	P	pH (categorized mean annual value)
Fragmentation	D	Dams (proportion of stream length above an impassable dam)
Sedimentation	IRdA	Industry + road density + agriculture
Sedimentation	IRd	Industry + road density
Sedimentation	IA	Industry + agriculture
Sedimentation	ARd	Agriculture + road density
Sedimentation	I	Industry (proportion of watershed used for industry)
Sedimentation	Rd	Road density (km of road/km^2 of watershed)
Sedimentation	A	Agriculture (proportion of watershed used for agriculture)
Water quality[3]	FMP	Forestry + mines + pH

---

[2] pH can vary substantially among tributaries within a watershed and can act as a barrier to movement as individuals will not enter nor will they experience high survival in low pH tributaries (Bowlby et al. 2014).

Threat	Variables	Description
Water quality	FM	Forestry + mines
Water quality	FP	Forestry + pH
Water quality	MP	Mines + pH
Water quality	F	Forestry (proportion of watershed used for forestry)
Water quality	M	Mines (# abandoned mine openings per km^2)
Water quality	P	pH (mean annual pH)

---

[3] Industry was originally included in the water quality index, but was removed because it had a positive Mantel correlation while the rest were negative.

Table 4-3. A Mantel test assuming Spearman correlations between pairwise genetic distances (Nei's F^{st}) and two isolation models, as well as an assessment of the two isolation models relative to each other using Reciprocal Causal Modeling. The Isolation by Distance (IBD) model represents a distance matrix based on the coastal length (km) between river mouths (Euclidean), and the Isolation by Environment (IBE) model represents a distance matrix based on the landscape characteristics of the watersheds (Euclidean). Significance was assessed relative to the 95% quantile of the distribution of the permuted data.

		Mantel		IBE|IBD		IBD|IBE	
Isolation model	Genetic metric	*r*	p-value	*r*	p-value	*r*	p-value
IBE	Pairwise F^{st}	0.392	0.007	0.176	0.152		
IBD	Pairwise F^{st}	0.373	0.020			<-0.001	0.504

Table 4-4. Reciprocal Causal Modeling to assess the land use variables contributing to each of three threats categories: Fragmentation, Sedimentation, and Water quality. The best-supported model from the Mantel and partial Mantel tests is shown in bold (representing ‘Chosen’). Chosen|Alternate indicates the Mantel correlation between the best-supported model and genetic distance, parcelling out the influence of each of the other models. Alternate|Chosen indicates the Mantel correlation between each of the alternate models after parcelling out the influence of the best-supported model.

		Mantel		Partial Mantel		Chosen|Alternate		Alternate|Chosen	
Threat	Variables	R	p-value	R	p-value	R	p-value	R	p-value
Fragmentation	RcPD	0.202	0.229	0.091	0.411	0.258	0.169	0.135	0.336
Fragmentation	RcP	0.114	0.314	0.190	0.231	0.229	0.193	0.058	0.408
Fragmentation	Rc	0.165	0.183	0.204	0.140	0.233	0.180	0.114	0.261
Fragmentation	P	0.100	0.331	-0.079	0.640	0.283	0.144	0.100	0.346
Fragmentation	D	-0.085	0.636	-0.052	0.635	0.280	0.156	0.012	0.567
Fragmentation	**RcD**	**0.205**	**0.199**	**0.348**	**0.094**				
Fragmentation	DP	-0.025	0.541	-0.117	0.674	0.282	0.150	0.039	0.493
Sedimentation	**IRdA**	**0.425**	**0.028**	**0.558**	**0.007**				
Sedimentation	IRd	0.438	0.043	0.528	0.019	0.146	0.310	0.148	0.312
Sedimentation	I	0.402	0.060	0.485	0.034	0.257	0.135	0.190	0.259
Sedimentation	Rd	0.482	0.017	0.488	0.015	0.174	0.282	0.177	0.276
Sedimentation	A	0.388	0.074	0.358	0.128	0.420	0.051	0.085	0.427

		Mantel		Partial Mantel		Chosen|Alternate		Alternate|Chosen	
Threat	Variables	R	p-value	R	p-value	R	p-value	R	p-value
Sedimentation	IA	0.378	0.076	0.560	0.008	0.195	0.227	0.008	0.481
Sedimentation	ARd	0.438	0.027	0.523	0.008	0.191	0.265	0.126	0.332
Water quality	FMP	-0.090	0.264	-0.019	0.477	-0.405	0.020	0.080	0.714
Water quality	FM	0.118	0.787	0.063	0.676	-0.402	0.014	0.094	0.746
Water quality	FP	-0.066	0.401	-0.182	0.231	-0.415	0.015	-0.180	0.244
Water quality	**MP**	**-0.383**	**0.015**	**-0.435**	**0.011**				
Water quality	F	-0.011	0.522	-0.348	0.106	-0.400	0.023	-0.084	0.391
Water quality	M	-0.290	0.095	-0.202	0.174	-0.330	0.046	-0.069	0.379
Water quality	P	0.100	0.680	-0.097	0.349	-0.429	0.014	0.161	0.756

Table 4-5. Comparison of the results from the MRDM regressions between pairwise genetic distances (Nei's F_{st}) and all combinations of predictors (expressed as Euclidean distance) to assess any relationship between population connectivity and threats. Distance matrices are: IBE = Isolation by Environment model, F = Fragmentation, WQ = Water quality, S = Sedimentation (refer back to Table 4-4). The chosen model is shown in bold.

Model	R-square	p-value	AIC_c
Full	0.547	0.001	-108.71
IBE+WQ+S	**0.536**	**<0.001**	**-110.01**
IBE+WQ+F	0.377	0.011	-104.85
IBE+S+F	0.399	0.009	-103.48
IBE+S	0.387	0.003	-104.59
IBE+WQ	0.316	0.005	-102.36
IBE+F	0.232	0.050	-99.95
IBE	0.164	0.006	-97.34

Figures

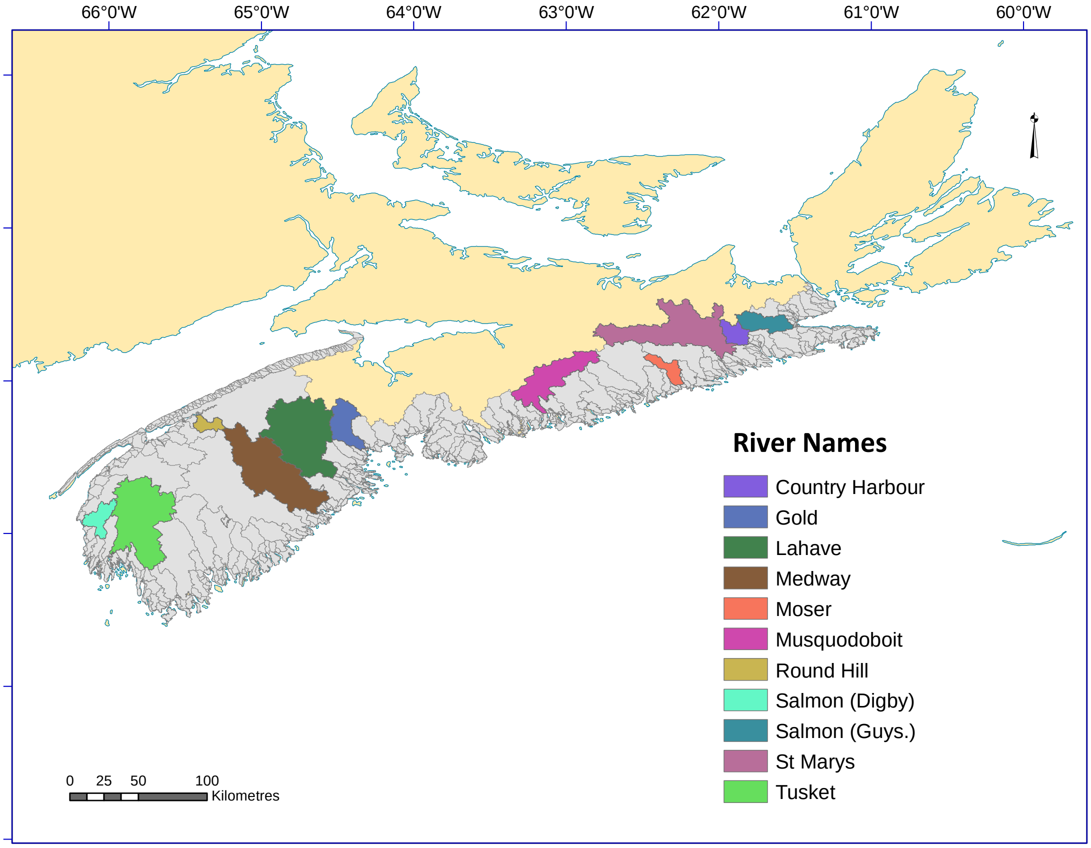

Figure 4-1. Geographical boundaries of the watersheds (coloured) containing the Atlantic salmon populations and rivers sampled for this study as well as other watersheds (grey) within the Southern Upland region of Nova Scotia. There is a Salmon River in Digby County and a second one in Guysborough County.

Chapter five: How is juvenile Atlantic salmon density influenced by the landscape characteristics of watersheds?

Abstract

Remediating negative impacts from anthropogenic activities in watershed catchments first requires quantifying how abundance and distribution of species relate to spatial and temporal variability of landscape characteristics. For Atlantic salmon, questions remain about how the spatial organization of rivers influence ecological processes, how species respond behaviourally to variability, and how multiple types of land use interact to change the environmental conditions experienced in particular locations. Based on mixed effects modeling, I demonstrate how juvenile distribution patterns are consistent with predictions that geomorphological processes occurring over large spatial extents set the context for habitat selection at small spatial extents, and thus largely determine abundance and distribution patterns within rivers. However, I did not find a similar pattern in relation to threats, suggesting that the previous result could be confounded by differences in the efficacy of electrofishing surveys along the stream network. Nevertheless, the influence of human activities on juvenile salmon varied depending on the suite of coincident land use and land cover types. When the proportion of natural forest cover was high, populations appeared to be able to benefit from low levels of anthropogenic land use, declining with increasing human activity only once the average proportion of natural forest cover was low. For management and recovery

planning, I proposed that this separation in land use and land cover characteristics relative to population response could be used to develop a simple quantitative index to help evaluate the effects of future human development in the vicinity of rivers or to identify areas requiring remediation.

Introduction

Freshwater fishes inhabit an ecosystem that is particularly susceptible to damage through direct modification as well as by anthropogenic land use within catchments (Lake et al. 2007, Olden et al. 2010). Such land use directly affects the structure and function of stream ecosystems through modification of the geological, topographical, and/or hydrological characteristics of watersheds (Altermatt 2013, Jeffrey et al. 2015). Given that human modification is typically extensive and severe, expectations are that it reduces habitat quality or quantity for fishes (LeRoy Poff et al. 2010, Dudgeon et al. 2006) leading to reduced survival at a population level and declines in abundance. Developing models to quantify any functional relationship between survival and land use change is critical for recovery planning for endangered wildlife (Venter et al. 2006), particularly for freshwater fish (Lawler et al. 2002, Lake et al. 2007). Yet this relies on understanding how individual responses to spatial and temporal variability of in-stream habitat characteristics culminate in changes to vital rates and population-level abundance.

It is relatively uncertain which landscape patterns reflect biologically meaningful environmental variability, particularly within riverine environments. Salmonid species have been extensively studied relative to habitat requirements (e.g., Gibson 1993, Heggenes 1990, McCormick et al. 1998) and responses to proximate habitat manipulation (e.g., Roni et al. 2001, Bash and Ryan 2002), which can be used to identify environmental variables that potentially influence survival or abundance at a population level. However, the influence of spatial pattern on ecological processes is particularly strong in riverine environments (Wiens 2002). Their spatial network structure and dendritic nature (i.e., characterized by unidirectional flow) imposes a hierarchical structure to their ecological organization, where the physical and biological processes at small spatial extents are constrained and/or modified by processes occurring over larger extents (Campbell Grant et al. 2007, Johnson et al. 2007). The River Continuum Concept (RCC; Vannote et al. 1980) is arguably one of the most important conceptual landscape models of stream networks, one that contextualizes this spatial heterogeneity (Fausch et al. 2002). The RCC predicts relatively gradual changes in ecosystem processes along a stream network (averaged over space and time) resulting from changes in channel morphology as influenced by flow. Thus at a large spatial extent, heterogeneity in habitats is associated with downstream decreases in channel slope, riffle development and substrate size, yet increases in width, depth and pool size (Schlosser 1991). Relative to salmon, the RCC predicts a longitudinal gradient in juvenile habitat suitability from the tributaries to the main stem of rivers, given that suitable flow and substrate characteristics for salmon would be predominantly found in headwaters (Heggenes 1990, Gibson 1993). Irrespective of spatial heterogeneity, relative position in the stream network could influence the ability of individuals to change their behaviour in response to localized changes in habitat suitability; for example, by limiting the potential for movement into refuge areas, particularly in an upstream direction (Brown & Swan 2010). Overall, fish distribution patterns within a watershed would be expected to be related to the amount, distribution and suitability of habitat types, where these are expected to be primarily determined by stream geometry and flow characteristics.

It is also relatively unknown how populations may compensate for spatial and temporal variability in environmental characteristics. Species living in variable environments might be expected to be similarly dynamic and able to respond effectively to environmental stochasticity through changes in growth rates, reproductive strategy, habitat use, or distribution (Pearman et al. 2008, Kubisch et al. 2014). Relative to threats, mobile organisms such as fishes would be expected to redistribute in response to local habitat degradation and thus localized abundance may reflect processes occurring within reaches, sub-catchments, or catchments, as opposed to localized variation (Semokorowski & Pratt 2007, Johnson et al. 2007). Also, population-level changes in mortality rates may not become measureable until substantial losses in productivity are realized. At the same time, compensatory behaviour could lead to threshold-type or non-linear responses to threats (Rose 2000, Benton et al. 2006). Potential synergistic and antagonistic interactions among multiple individual stressors would complicate any quantitative relationship between land use and population dynamics, leading to cumulative effects on populations that may be quite different from the sum of their independent effects (Darling & Cote 2008, McCluney et al. 2014). Furthermore, ecological changes resulting from human activities within watersheds have the potential to culminate along the length of the stream network and have substantial impact on locations distant from their source (Fausch et al. 2002, McCluney et al. 2014). In order to characterize how populations respond to localized land use, it is thus necessary to consider not only the relative magnitude of and potential interactions among anthropogenic activities in watersheds, but also the extent over which they occur.

Questions and predictions

This chapter represents a first attempt to describe functional relationships between landscape characteristics of watershed catchments (particularly human activities) and endangered juvenile Atlantic salmon (*Salmo salar*) populations from the Southern Upland region of Nova Scotia, Canada. Given substantial uncertainty in how the spatial organization of rivers influence ecological processes, how species respond behaviourally to variability, and how multiple types of land use interact to change the environmental conditions experienced in particular locations, it became difficult to determine what patterns may exist as well as to develop optimal environmental predictors *a priori*. Therefore, I am presenting this analysis as a series of questions that may be used to inform future research. Throughout the analysis, I controlled for changes in population size over time, given that Southern Upland salmon populations have undergone extensive declines in abundance since the 1980s (Gibson et al. 2011). The five questions and some associated predictions that I evaluated in this chapter are:

1. *Do densities change longitudinally along the stream network?* If juvenile habitat suitability varies longitudinally in relation to water flow accumulation (as predicted by RCC), juvenile densities might be expected to be negatively related to relative position in the watershed; given flow, depth and substrate preferences (Heggenes 1990).

2. *Is habitat use related to flow characteristics?* The physical properties of habitat types as well as their distribution within a stream network are largely determined by the flow regime (Altermatt 2013), which is closely related to the topographical characteristics of watersheds (LeRoy Poff et al. 2010). From this, I would predict a functional relationship between stream gradient and juvenile density.
3. *How might the hierarchical structure of rivers influence relationships?* If mobile organisms such as fishes primarily respond to environmental processes or changes in such processes occurring in the catchment, sub-catchment or reach (Johnson et al. 2007, Hitt & Angermeier 2008), juvenile densities might be more strongly related to metrics describing the characteristics of tributaries or sub-catchments, rather than those describing specific sampling sites.
4. *How do populations respond to anthropogenic land use?* The general perception is that essentially any human activity occurring within watershed catchments, be it forestry, agriculture, industry, or urbanization, negatively affects freshwater productivity and should be considered a threat to populations (LeRoy Poff et al. 2010, Dudgeon et al. 2006, Venter et al. 2006). Thus, juvenile densities would be predicted to decline as the total proportion of human land use increases, potentially at or after a threshold value.
5. *What is the form of the population response to threats?* Spatial heterogeneity caused by variation in the relative proportions of specific types of land use is expected to lead to complex and non-intuitive changes to the localized abiotic characteristics of rivers (Rose 2000) through synergistic and antagonistic

interactions among individual threats (Darling & Cote 2008). Thus, I would predict that changes in juvenile density from a specific type of land use would depend on the suite of coincident land use or land cover types.

Methods

Juvenile salmon density estimation

For diadromous fishes, juvenile densities are often estimated at multiple sites within watersheds, offering the potential to use changes in density as a proxy for changes in mortality rates leading to changes in abundance when evaluating population-level responses to land use. Using density data as a proxy for abundance assumes that there is a predictable relation between the numbers of fish and the spatial dimensions of the sampled unit. In many salmonid species, relationships are positive, but have substantial variability due to factors other than habitat size, including inter-species interactions or food availability (Torgersen et al. 2012) as well as intra-species characteristics such as increasing territory size with body size (Grant et al. 1998). There is the potential that such factors may obscure any relationship between landscape characteristics and juvenile density; or that abundance, biomass, or a measure of habitat saturation (rather than density) becomes the appropriate metric for analysis (Torgersen et al. 2012). Given the variability in the electrofishing site characteristics included in this study (see below), I have kept density as a proxy for abundance, while recognizing the limitations of the approach.

Information on localized juvenile Atlantic salmon densities came from an extensive population monitoring data set of 12 rivers in the Southern Upland region of Nova Scotia collected by electrofishing (Figure 5-1). For comparability with the environmental data on land use, only sites (N = 134) sampled in the years 1995 to 2005 were included, resulting in 452 sampling events (Table 5-1). During this timeframe, electrofishing survey methods were not standardized; single-pass surveys, depletion surveys and mark-recapture surveys were all used (Table A5-1-1, Appendix 5-1). Of these, single pass surveys as well as mark-recapture surveys had open sampling sites, where physical barriers were not used to prevent movement into or out of the site while sampling. Multi-pass surveys typically employed barrier nets. Multiple habitat types were sampled in each watershed as well as within the majority of individual sites (e.g., a single site could include riffle, run and pool habitat types). A typical survey sampled individual reaches (10^1 m) or stream segments (10^2 m; as defined in Frissell et al. 1986), which averaged over any preferential use of specific habitat features by age class (e.g., older juveniles that prefer faster water; Blanchet et al. 2006), or any variation in juvenile densities as related to specific habitat types (e.g. preferential use of riffles relative to pools; Gibson 1993). To calculate total juvenile densities (number/100m^2), I summed catches of fry and parr during the first pass in each survey, multiplied by survey catchability, and scaled by site area (Bowlby & Gibson 2012). I estimated median capture efficiency (q) for open sites and depletion sites independently from the entire data set. The capture efficiency is defined as the proportion of total abundance captured on the first pass, where q = 0.34 for open sites and q = 0.56 for depletion sites (Appendix 5-1). Over

95% of the data derived from open-site electrofishing, so the standardization relative to catchability had very little effect on the analyses. A more substantial limitation of the data was the need to combine catches of fry and parr, even though the life stages would be expected to prefer slightly different habitat types (Heggenes 1990, Gibson 1993, Gibson et al. 2008). I did this to reduce the proportion of zeros in the data, but it was also necessary to be able to estimate the random effects distribution, which typically requires 10 or more categories (Zuur et al. 2009).

For the majority of sites (58%), the area sampled (in m^2) as well as shocking time (in seconds) had been recorded in the data; however, there were instances where only the shocking time was recorded (16%) and where no effort measure was given (27%). To approximate sampling area when only shocking time was recorded, I developed a zero-intercept linear regression, with area as the dependent variable and shocking time as the independent variable, and used the estimated slope coefficient to predict areas (in m^2). For sites with no measure of effort, I applied the median area from all sites sampled on the particular river in which area was given or had been predicted from the linear model above.

GIS analyses

To develop land use predictors, I intersected and/or summarized 6 main data layers in ESRI ArcGIS® 10.2: (1) the electrofishing site locations, (2) a hydrologically-connected flow network for each watershed, (3) watershed boundaries, (4) a Digital Elevation Model (DEM) for Nova Scotia, (5) site-specific and sub-catchment-scale buffers, and (6) a layer detailing land use and land cover types for Nova Scotia (details on

1 and 5 to follow; details on 2,3,4 and 6 provided in Appendix 5-2). For the electrofishing sites, the precision and projection of the geographical coordinates recorded in the data varied considerably, with some listing only grid references (3-digit UTM coordinates corresponding to 100 m resolution from the NAD27 datum for Zone 20 North), some giving full UTM coordinates (but referenced to the NAD83 datum for Zone 20 North), and some listed in latitude and longitude (also using the NAD83 datum). I gave grid referenced sites more exact coordinates by locating them in Google Earth (Google Inc. 2005), often based on qualitative site descriptions recorded in the original data. Once imported into ArcGIS, I projected all data (site locations as well as all other data layers) into UTM (NAD 83 datum for Zone 20 North) to ensure consistency and the ability to calculate distance and area metrics. As is common in data derived from different sources, most of the geographical locations of the electrofishing sites did not fall exactly on the stream network. Thus, I snapped all electrofishing site locations to the most proximate segment of the flow network; 91% of sites moved < 50 meters (median = 14.5 m; range = 0.07, 735 m). I verified snapped locations *a posteri* with the site descriptions recorded in the electrofishing database to ensure that they were on the correct stream segment. Locations that were electrofished in multiple years often had slightly different coordinates corresponding to the same site name (typically within 50 m of each other). In these instances, I used the coordinates recorded from the most recent survey.

I developed two types of environmental predictors for comparison with the juvenile densities. The first was site-specific, corresponding to land use characteristics in the vicinity of the stream reaches sampled for juvenile density. To develop these, site-specific buffers were delineated assuming a 500 m radius around each point location for an electrofishing site, which gave a polygon area of 785,397 m^2. I made the site-specific buffers so large in order to reduce the effect of uncertainty in the site location; given the qualitative descriptions of some of the sampling locations, the need to snap sites to the stream network, the fact that upwards of 100 m of stream length was fished at some sites, and that coordinates could have been taken at either the upstream or downstream end.

The second type of environmental predictor represented sub-catchment variability, corresponding to land use characteristics along all tributaries upstream of each electrofishing site (example in Figure 5-2). I delineated one sub-catchment buffer for each electrofishing site individually and then saved them as unique feature layers (median buffer size = 15,198,221 m^2; range = 1,495,113 to 583,214,860 m^2). The length of the stream network contained in these buffers represented stream (10^2 m) or segment (10^3 m) systems (Frissell et al. 1986). To identify all segments upstream of a specific electrofishing site from the hydrologically-connected flow network previously developed for Southern Upland watersheds (Bowlby et al. 2014), I used the Trace Upstream tool from Network analyst toolbox in ArcGIS® 10.2. These segments were buffered at a distance of 500 m on either side of the stream and then dissolved into a single upstream polygon. For a small number of sites (N = 9), inappropriate upstream buffers revealed issues with connectivity in the flow network, so I selected the upstream segments at these sites manually. Because the tool calculates buffers around each stream segment individually, it was the possibility that the buffers for upstream sub-catchments would include area downstream of the actual electrofishing site (i.e., the buffer would include the entire river segment that the electrofishing site was located on). *Post-hoc* evaluation revealed that this was a prevalent issue and in some cases nearly doubled the amount of area considered to be upstream of a given site, particularly for sites on smaller tributaries. Therefore, I clipped each sub-catchment buffer polygon at approximately a 90° angle through the electrofishing site and deleted any downstream area. All spatial analyses using the sub-catchment buffers were based on these clipped polygons. In some instances, unique electrofishing sites were very close together and had overlapping site-specific buffers. This lead to identical sub-catchment buffers being delineated for these specific sites. There was not enough data to be able to add site as an additional random effect in the model (Table A5-1-1, Appendix 5-1), which would have been done to account for repeated measurements of juvenile density in the same year relative to identical environmental predictors. Instead, I assumed that these sites were the same and averaged their site-specific environmental characteristics and juvenile densities.

Once I had delineated buffers, I used them as boundaries within which I summarized the topographic characteristics of the stream network, as well as the proportions of land use and land cover types surrounding the stream network. This means that the variables describing land use composition had the same spatial grain yet differed in spatial extent, where grain represents the spatial resolution of the data (e.g. pixel size in a raster layer) and extent refers to the overall size of the area being studied (Turner 1989, Gustafson 1998). I compiled a digital elevation model (DEM) for Nova Scotia from 20 m horizontal resolution and 1 m vertical resolution raster data available from GeoBase (Table A4-2-1, Appendix 4-2), before calculating the slope (in degrees) of each pixel using the Spatial Analyst extension. Iterating over unique electrofishing sites, I used Zonal Statistics Analysis to calculate the mean and range of slopes for the site-specific and sub-catchment-specific buffers, also in ArcGIS® 10.2. Zonal Statistics Analysis takes values from a raster data layer (here slope values) and calculates statistics relative to a user-specified zone. For the sub-catchment buffers, I used the stream segment contained within them as a zone. For the site-specific buffers, I used 100 m of stream length upstream and downstream of each electrofishing point location (200 m total length) as the zone.

To develop variables describing land use composition, I intersected both the site-scale and sub-catchment buffers with land use and land cover data from the Forest Inventory Cycle 2&3 (Nova Scotia Department of Natural Resources, downloaded November 2011; Table A4-2-1, Appendix 4-2). These data are based primarily on aerial photography from 1995 to present and digitized from 1:10000 scale images (1 mm on the image represents 10 m on the ground) which ensured that the resolution was meaningful relative to my chosen buffer sizes. The spatial grain was of a similar magnitude between environmental metrics describing stream topography (20 m horizontal resolution) and ones representing landscape composition (10 m horizontal resolution). Finally, I calculated the amount of area (m^2) corresponding to each type of land use (represented by a FORNON code) within each buffer and appended these into site-specific and sub-catchment-specific data tables. To ensure repeatability and consistency, I programmed each of the analysis steps described above into the Model Builder component of ArcGIS® 10.2. In the original data, FORNON represents a field name rather than an acronym.

There were multiple unique FORNON codes representing different human activity types (e.g., treated stands vs. Christmas trees for silviculture; hay fields and blueberries for agriculture), as well as stand characteristics (e.g., brush with 25%, 50% or 75% crown closure of alders). To reduce the number of variables considered, I re-categorized all areas from codes corresponding to similar types of land use or land cover and then calculated proportions for each category (Table A5-2-1, Appendix 5-2).

Specifically related to the evaluation of human activities in watershed catchments, potential predictors included the proportions of clearcutting, forestry activity, industry, urbanization, corridors or agriculture in the site-scale or sub-catchment buffers (Table 5-2). However, there were relatively strong correlations among many of these variables (Spearman correlations > 0.6). To reduce the potential for retention of spurious predictors arising from collinearity in multiple regressions (Dormann et al. 2013, Zuur et al. 2009) and to prevent focusing the analyses on threats perceived to be the most important (Norris et al. 2004), I summed the proportions of individual types of anthropogenic land use (e.g., industry, clearcutting, agriculture, etc.) at the site-scale and in sub-catchments when evaluating anthropogenic land use. However, I evaluated the relative contribution of each individual threat type *a posteri* using a Quadratic Discriminant Analysis (QDA; Borcard et al. 2011), based on the results of the multiple regression described below.

Statistical analyses

On the basis of diagnostic plots, I excluded points that took extreme values (i.e., falling orders of magnitude outside the range of the rest of the data) for salmon density or the environmental predictors from further analyses (N = 7 points, distributed over three watersheds). For the rest of the data, I evaluated potential relationships between juvenile salmon density and the environmental predictors using a series of generalized linear mixed models (GLMM) implemented in the R package ‘glmmADMB’ (Fournier et al. 2012, Skaug et al. 2012). For all models, I used a negative binomial distribution (log link) for the response, given that initial diagnostics suggested that the data were overdispersed relative to a Poisson distribution. A Poisson distribution in count data would only be expected to arise in situations where there is a constant probability of appearance of the target species in a given sampling event, and where all individuals present are observed and correctly identified. Variability in counts arising from differences in detection efficiency or a non-constant probability of appearance among sampling events are two common causes of overdispersion, and are better modelled using a negative binomial distribution (Linden & Mantyniemi 2011).

Similarities among density estimates might be expected due to characteristics of the watersheds themselves; for example, more productive rivers or those with larger adult populations would be expected to have proportionately higher juvenile densities relative to other watersheds. In each model, I incorporated ‘watershed’ as a random effect to account for the lack of independence among data points sampled in the same watershed (Zuur et al. 2009). Although model selection was not done using frequentist methods (see below), allowing for a constant spatial correlation structure among data from the same watershed reduces the overall degrees of freedom in the model, which affects the critical values from the F-distribution against which parameter significance may be assessed (i.e. guards against Type 1 error). As fixed effects, I considered 12 candidate variables related to my five original questions (Table 5-3), which were scaled between zero and one to ensure comparability. I followed the general approach of Albanese et al. (2009) to reduce the number of models considered in the analyses: starting from a global model incorporating all predictors, I constructed a subset of 13 candidate models, beginning with predictors that I considered to be important based on previous knowledge of these systems. This meant that the questions I originally outlined were not evaluated in isolation (i.e., one model formulation per question), but in conjunction. I used an information-theoretic approach for model selection (Johnson & Omland 2004) based on the Akaike Information Criterion (AIC) as well as Akaike weights representing the relative likelihood of each model (Albanese et al. 2009). The weights can be thought of as the probability that a specific model is the best given the data as well as the set of candidate models, expressed as a ratio (Wagenmakers & Farrell 2004).

I also performed a *post hoc* evaluation of parameter variability and stability of the chosen model using Markov-Chain Monte-Carlo (MCMC) methods, based on 10,000 MCMC iterations as implemented by the ‘glmmadmb’ function (Skaug et al. 2012), discarding the first 1500 iterations as the burn-in and retaining 8500 for inference. Here, convergence in the MCMC chains is suggestive of model stability, and credible intervals for parameters that do not span zero are indicative of reliability. I assessed convergence to the posterior distribution visually from traceplots (e.g. Gibson et al. 2015) as well as by the Geweke diagnostic using the R package ‘coda’ (Plummer et al. 2006). The Geweke calculation compares the standardized Z-score (i.e., the difference between sample means divided by the estimated standard error) of the first 10% and last 50% of the chain.

Absolute values less than 2 are often considered representative of convergence (Su et al. 2001). For visualization of the fixed effects retained in the chosen model, I plotted partial regression fits, varying the predictor of interest while keeping others constant at median values.

When the site-specific proportion of human use (s.phuman) was below approximately 0.4, the regression analyses suggested a different relationship with juvenile density than when it was above 0.4 (see Results). To visualize individual variables potentially contributing to this grouping, I compared boxplots of the site-specific proportions of natural forest cover and each threat type for two categories of sites, ‘low’ (s.phuman ≤ 0.4) and ‘high’ (s.phuman > 0.4). It is important to note that the natural forest variable is not the inverse of the proportion of human use, but excludes other natural cover types such as wetlands (cover types defined as natural in Table A5-2-1; Appendix 5-2). The site-specific threats that I included were the proportions of clearcutting, forestry (silviculture), agriculture, industrial corridors, other industry and urban area. In addition to the individual comparisons, I used Quadratic Discriminant Analyses (QDA) to evaluate the combined suite of environmental characteristics for sites of each category. Here, the ‘low’ category would represent environmental conditions that would be expected to be beneficial to juvenile salmon populations, while the ‘high’ category would represent conditions that would be expected to lead to population decline (see Results). Although QDA is most commonly used to evaluate metrics contributing to differences between distinct ecological types (e.g., species presence/absence, Slater & Michael 2012; age classes, Fukuda et al. 2013), it has also been used in situations where the categorization is related to the environmental metrics that were subsequently analyzed (e.g., niche overlap, Broennimann et al. 2012; habitat quality, Gordon et al. 2016). QDA works by projecting the input variables to a linear or quadratic axis that maximizes the separation between categories (Bocard et al. 2011). Note that this axis is dimensionless. Unlike Linear Discriminant Analyses, the within-group covariance matrix of the explanatory variables does not need to be homogeneous, which makes it more appropriate for these data. To calculate the misclassification rate that arose from the fitted discriminant axis, I used a jackknife (i.e., leave-one-out cross validation; Bocard et al. 2011), iterating over sites.

Results

There was relatively strong support for the NB12 model on the basis of AIC and Akaike weights, with differences > 3.4 in AIC when compared to the majority of other models, and an overall evidence ratio of 0.594 (Table 5-4). As fixed effects, this model incorporated year, site position along the stream network (site.pos), average slope within the site buffer (s.degslope), the proportion of human use within the sub-catchment (phuman) and a quadratic term for site-specific human use (s.phuman+s.phuman$^{\wedge 2}$) (see Table 5-2 for definitions; Figure 5-3). The next best model (NB13) had a second quadratic term for human use in the sub-catchment (phuman+phuman$^{\wedge 2}$) and an evidence ratio of 0.239. However, the addition of the quadratic term did not change the log-likelihood of the NB13 model relative to the NB12 model, although it did change the AIC value (Table 5-4). This suggests that the extra parameter did not explain any additional variation in juvenile densities and that the change in the AIC was entirely a result of the penalty imposed for the extra parameter (Supplement C accompanying Evans & Davis 2011). If the NB13 model was removed from consideration when calculating Akaike weights, the evidence ratio for the NB12 model increased to 0.78. All other models contributed less than 0.12 in Akaike weight. Even with combining age classes, there was some indication that the data were zero-inflated (10.3% of the juvenile density estimates were zero), as reflected by the estimated dispersion parameter for the negative binomial distribution being < 1 (theta = 0.908, s.d. = 0.07). However, re-analysis using a zero-inflated model (as implemented by the glmmadmb function) had little effect on the estimated fixed or random coefficients other than increasing the estimate of theta, suggesting that zero-inflation did not bias parameter estimation.

Based on maximum likelihood, the estimated variance for the random effects in the NB12 model was 0.381 (s.d. = 0.617). Based on the MCMC, there was good convergence for the posterior distributions, with the absolute values of the Geweke Z statistic being smaller than 2 for all parameters (range = -1.73, 1.49). For most fixed effects, 95% Bayesian credible intervals derived from MCMC did not include zero. The estimated slope for the fixed effect of year was negative (slope = -0.63, s.e. = 0.19; Table 5-5), suggesting that average densities declined from 1995 to 2005 (Figure 5-3A). The model included a stronger negative relationship relative to position (slope = -1.90, s.e. = 0.42; Table 5-5), indicating that juvenile densities were higher in habitats located in the tributaries relative to the main stem (Figure 5-3C). Model comparison indicated that the relative position of a sampling site (site.pos) was a better predictor than the area upstream of the site (total.area) of this relationship (*c.f.* NB2 and NB3, AIC = 3327 and 3332, respectively; Table 5-4).

Relative to localized habitat characteristics, the NB12 model retained the median slope of the site-specific buffer (s.degslope) as a fixed effect (Figure 5-3B). Juvenile density slightly increased with the median slope of the site (slope = 0.07, s.e. = 0.04; Table 5-5). The quadratic function for the effect of human activity at the site level (s.phuman + s.phuman$^{\wedge 2}$) demonstrated modest increases in mean density until the proportion reached approximately 0.4; thereafter, mean density tended to decline, albeit with a high level of variability (Figure 5-3D). Relative to characteristics of sub-catchments, the NB12 model retained the proportion of human use (phuman) as a significant fixed effect, yet the relationship was strongly positive (slope = 2.54, s.e. = 0.78; Table 5-5). Similarly unexpected were the nearly identical slope estimates for site-specific and sub-catchment-specific proportions of human use (s.phuman and phuman; *c.f.* 2.76 and 2.54, respectively; Table 5-5), suggesting increases in average juvenile density until the proportion of human use in a watershed reaches 0.4 (*c.f.* Figure 5-3D and Figure 5-3E). It is also worth noting that the model with the second-highest evidence ratio (NB13) included a quadratic term for the proportion of human use in a sub-catchment (phuman), similar to that for the site-specific proportion of human use (s.phuman) in the top model.

The proportion of natural forest cover was substantially different for sites categorized as having low levels of human use (s.phuman ≤ 0.4) relative to high, but the distributions of the individual threats variables overlapped substantially (Figure 5-4). As expected, the threats exhibited less variability for sites belonging to the ‘low’ category as well as consistently lower medians relative to the ‘high’ category, except for ‘Industry’ (Figure 5-4). However, proportions of industry at the site-scale were extremely low relative to the other types of threats (< 6% of site-specific area for 93% of the data). In general, sites with high proportions of human use were characterized by relatively little natural forest cover combined with substantially higher median proportions of clearcutting, forestry (silviculture), agriculture, roads and urban area (Figure 5-4). Looking at these individual variables in multivariate space suggested a strong separation of sites categorized as ‘low’ relative to ‘high’ along a quadratic discriminant axis (Figure 5-5). Based on a jackknife (i.e., leave-one-out cross-validation) of the QDA, the misclassification rate was 2% for all sites (2 misclassifications relative to 107 data points for ‘low’ and 8 misclassifications relative to 334 data points for ‘high’).

Discussion

The density of juvenile Atlantic salmon (*Salmo salar*) from endangered Southern Upland populations was functionally related to multiple characteristics of the landscape. Results conformed well to hypotheses that habitat suitability for juvenile salmon varies longitudinally in the stream network (Q1; Table 5-3). The identified relationship between juvenile density and position implies that suitable habitats for juvenile growth and rearing are concentrated in the tributaries of Southern Upland watersheds. Over smaller spatial extents, juvenile density was related to the mean gradient of electrofishing sites (Q2; Table 5-3), with higher densities occurring at higher gradients. It is possible that the hierarchical structure of rivers may influence these patterns (Q3; Table 5-3). Expected juvenile densities decreased approximately 3-fold moving from the upper tributaries (site.pos = 0) to the main stem (site.pos = 1). Increases in expected juvenile density were essentially half this magnitude when the mean gradient of sites increased from zero to approximately seven degrees. These results are consistent with previous research on salmonids that suggested individuals concentrate primarily relative to geomorphological features and then select specific habitats within reaches based on localized characteristics (e.g., Baxter & Hauer 2000, Torgersen et al. 1999). If geomorphology sets the context for habitat selection at small spatial extents (Fausch et al. 2002), ecological processes occurring at larger spatial extents principally control habitat characteristics for fish (and consequently abundance and distribution) within rivers (e.g., Fausch et al. 2002, Thornton 2011). Future research in the Southern Upland could focus on developing explicit links between geomorphological patterns and adult distribution or juvenile productivity, similar to those developed for populations in the Gaspe region of Quebec (e.g., Davey & Lapointe 2007, Kim & Lapointe 2011). In order to accomplish this, it would be necessary to re-design the electrofishing surveys to employ either a nested or high-frequency sampling design that would enable evaluation of habitat or landscape relationships at varying spatial extents (e.g. Bult et al. 1998, Deschenes & Rodriguez 2007, Le Pichon et al. 2016).

The relationships between juvenile density and human land use were somewhat unexpected in that they did not seem to change at different spatial extents. In situations where human activities affected up to 40% of buffer area, the slope estimates for the sub-catchment and site-specific predictors were essentially identical. This is directly counter to expectations that land cover within catchments (i.e., at larger spatial extents) primarily determines site-specific abiotic characteristics and ecological processes (Fagan 2002, Semokorowski & Pratt 2007, Johnson et al. 2007) or that fishes can largely compensate for site-specific habitat degradation by moving (Brown & Swan 2010, Hitt & Angermeier 2008). Although it is possible that the site-specific buffers were too large to adequately characterize conditions at smaller spatial extents, this should have been true relative to distribution patterns as well (Q2; Table 5-3). If ecological changes resulting from human land use occurred primarily at larger spatial extents and culminated throughout a sub-catchment, changes in juvenile densities as related to sub-catchment characteristics would have been expected to be proportionately greater, potentially showing a threshold response at lower values or having a steeper slope estimate than the site-specific response. Values of the phuman predictor represent much larger geographic areas (e.g., 40% of 1.5 million m^2 relative to 40% of 780,000 m^2) with much greater potential for culmination, or synergistic effects (Darling & Cote 2008). However, the coefficients estimated for the site-specific and sub-catchment-scale predictors of human activity imply that low levels (ranging from 0 to 40% of area affected) have a positive effect on salmon density regardless of spatial extent and that culmination is minimal. This conclusion is unlikely to be an artifact due to correlation between the site-scale and sub-catchment predictors (Kendall’s *tau* = 0.38). It is also possible that a quadratic relationship with human activity at the sub-catchment scale could have been found if these data contained a larger range of values. Here, there were only 3 observations larger than 0.5; however, 80% (37 out of 47 values) of the residuals for the fitted NB12 model were negative when phuman was greater than 0.4, which suggests that a quadratic relationship is possible and might be detectable if the data contained a larger range of values.

Why would the response of juvenile salmon be consistent over multiple spatial extents in relation to anthropogenic land use, while it appears to differ with spatial extent in relation to distribution patterns? Unfortunately, the simplest explanation is likely related to limitations of the sampling methodology. The same variation in flow conditions predicted by RCC along the length of the stream network could markedly influence the efficiency of electrofishing surveys. Backpack electrofishing is most effective in shallow constrained areas (Speas et al. 2004), and catchability of salmon would be expected to be systematically lower at deeper and wider sites (Speas et al. 2004, McInerny & Cross 2000), particularly when barrier nets are not used. This is indirectly supported by the overall lack of electrofishing sites with a relative position > 0.5 in Southern Upland rivers (10 unique sites out of 134). Although site-specific catchabilities were estimated for some of the locations electrofished in a given year (i.e., the mark-recapture and depletion surveys), these estimates were systematically distributed relative to position in the watershed because they tended to take place where the initial number of juveniles captured was high. Thus, catchability of juvenile salmon is relatively unknown for main stem locations from these data. The direction of bias would be indistinguishable from hypothesized changes in juvenile distribution owing to variation in ecological processes with spatial extent. Any future work on developing quantitative linkages between large-scale geomorphological patterns and juvenile productivity would need to explicitly account for changes in catchability if it were to be based on electrofishing surveys.

Relationships between juvenile density and anthropogenic land use

Systematic trends in catchability would not be expected relative to human land use predictors, so these analyses are still informative on the functional form and relative magnitude of population-level change resulting from human activity as well as the manner in which individual land use types interact to cause such change (Q4 & Q5; Table 5-3). These analyses strongly support the prediction that the spatial arrangement of human activity exerts varying influence on the ecological processes in streams, leading to non-linear changes in population dynamics (Rose 2000, McCluney et al. 2014). In other words, specific land cover types do not have an intrinsic quality for juvenile salmon (being always good or always bad), and there appears to be no universal contrast between human-modified and natural landscapes when understanding change (Haila 2002, Fisher and Lindenmayer 2007). Results are consistent with evidence from regional and global studies in landscape ecology that large areas of native vegetation tend to benefit native species (reviewed in Fischer & Lindenmayer 2007). It is striking that the median proportion of natural forest cover at a site is more than 2.5 times higher and has an overall magnitude of 72% when localized human activity is positively associated with changes in juvenile salmon density (i.e., sites categorized as low in the QDA) than when it is negatively associated (i.e., sites categorized as high in the QDA). This difference is unlikely to be an artifact of the distribution of high and low sites among rivers, in that similar differences exist among high and low sites from the same river (data not shown). Previous research has found natural forest cover to be highly beneficial to various fish communities (Tanentzap et al. 2014, Allan 2004). The distinction here is that activities typically considered harmful to freshwater fishes were associated with *increases* in juvenile salmon density when natural forest was the predominant land cover type. While it is likely that a high proportion of natural forest cover in watershed catchments confers resilience to perturbation from human activities (McCluney et al. 2014), these results also suggest that it can enable populations to benefit from low levels of anthropogenic land use. This non-intuitive result would be similar to the inverse patterns in abundance expected from a trophic cascade (Pace et al. 1999), where removal or substitution of a top predator in an ecosystem alters the resource base available to another species at a lower trophic level, leading to population increases rather than declines. Some level of eutrophication from fertilizer application would be expected from land use activities such as silviculture or agriculture (Broadmeadow & Nisbet 2004) and would be expected to increase primary production. Fertilization has been used in aquaculture operations to increase fish productivity (Chislock et al. 2013) and in natural watersheds to enhance sockeye salmon production (Hyatt et al. 2004). Similarly, low levels of riparian clearing would open the canopy, potentially increasing light penetration and stream temperature in a manner beneficial to primary production or to juvenile growth (Allan 2004, Nislow 2005). Conversely, when juvenile density declines with increasing human activity (sites categorized as high in the QDA), the median proportion of natural forest is low (27%), while median proportions of clearcutting and other forestry activity approximately doubles (6% vs. 3% and 6% vs. 2.4%, respectively) and agricultural activity becomes substantial (24% vs. 0%). This suggests that there is a point above which changes in canopy cover, sediment and organic inputs, as well as hydrology lead to habitat deterioration (Gilvear et al. 2002, Allan 2004) and become detrimental to juvenile salmon.

Practical application

A landscape approach to recovery planning is important because it enables the ecological consequences of human activities to be identified and predicted at scales relevant to resource managers (Fausch et al. 2002). Although this research did not develop quantitative relationships with specific land use types for juvenile Atlantic salmon, it could be useful in a practical context for prediction; particularly in Southern Upland rivers where Atlantic salmon populations are not assessed. Using the discriminant axis generated by the QDA, the landscape characteristics of an unknown site could be used to classify it as low or high with a relatively high degree of certainty (2% misclassification rate). This means that from readily accessible data on land use and land cover (i.e., the variables in Table 5-2-1, Appendix 5-2) managers would have a way to evalate whether potential development in the vicinity of rivers would be detrimental to juvenile salmon, or could identify areas requiring remediation. This would be a simple, quantitative way to use changes in landscape area as a basis for decision-making during recovery planning (Semokorowski & Pratt 2007). Relative to the status quo, having any data-based criterion is preferable in that risk to populations from new development can be reduced and limited financial resources can be more meaningfully allocated.

Tables

Table 5-1. Summary of the number of sites electrofished for each river and year from which Atlantic salmon juvenile density was estimated.

River	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Row total
East (Chester)	13	12	2		1	2						30
Gold River	1	1				1	5			10	11	29
LaHave River	30		15	12	14	17	16	17	12	16	16	165
Liscomb			6	1	1	1						9
Medway River						3		5		5	3	16
Middle (Chester)	3	3				2						8
Musquodoboit		9			8	2						19
Petite						5						5
Salmon (Guysborough Co.)		8					9					17
Salmon (Port Dufferin)						2			5	1		8
St. Mary's River	15	8	12	12	12	10	11	12	11	11	10	124
West (Sheet Harbour)	4	4	5	4	3	2						22
Column total	66	45	40	29	39	47	41	34	28	43	40	452

Table 5-2. Description of the types of data used in these analyses. Characteristics of the electrofishing surveys were specific to the site and year sampled (site-specific). The suite of possible environmental predictors developed from the spatial land use data at small (site) and large (sub-catchment) spatial extents are given. Variables that were considered in the regression analyses are identified relative to their type and others are left blank. Measures of area (area.est and total.area) as well as year were standardized to a unit interval [0,1] for analyses.

Type	Extent	Name	Variable type	Description
site-specific		area.est		the amount of area that was electrofished at the site (m^2)
site-specific		sal		the number of juvenile salmon caught (all ages combined)
site-specific		sal.den	Dependent	the estimated density of juvenile salmon (#/100m^2)
site-specific		siteID		the site name
site-specific		watershed	Random effect	the watershed name
site-specific		year	Independent	the year(s) that the site was sampled
site-specific		site.pos	Independent	position of the electrofishing site along the stream network (length upstream/total length)
buffer	site	s.natfor		the proportion of natural forest in the site buffer
buffer	site	s.clearcut		the proportion of the site buffer that was clearcut
buffer	site	s.forest		the proportion of the site buffer used for silviculture, including treated stands
buffer	site	s.agri		the proportion of the site buffer used for agriculture

Type	Extent	Name	Variable type	Description
buffer	site	s.urban		the proportion of the site buffer that is urbanized
buffer	site	s.industry		the proportion of the site buffer used for industry
buffer	site	s.corridor		the proportion of the site buffer used for industrial corridors
buffer	site	s.natural		the proportion of the site buffer with natural cover types
buffer	site	s.phuman	Independent	the proportion of the site buffer that has been modified by human activities (sum of s.clearcut, s.forest, s.agri etc.)
buffer	site	s.degslope	Independent	the average slope of the site buffer
buffer	site	s.rangeslope	Independent	the maximum difference in slope (max - min) in the site buffer
buffer	sub-catchment	total.area	Independent	the size of the sub-catchment buffer (m^2); represents the relative drainage area upstream
buffer	sub-catchment	natfor		the proportion of natural forest cover in the sub-catchment buffer
buffer	sub-catchment	clearcut		the proportion of the sub-catchment buffer that was clearcut
buffer	sub-catchment	forest		the proportion of the sub-catchment buffer used for silviculture, including treated stands
buffer	sub-catchment	agri		the proportion of the sub-catchment buffer used for agriculture
buffer	sub-catchment	urban		the proportion of the sub-catchment buffer that is urbanized
buffer	sub-catchment	industry		the proportion of the sub-catchment buffer used for industry
buffer	sub-catchment	corridor		the proportion of the sub-catchment buffer used for industrial corridors

Type	Extent	Name	Variable type	Description
buffer	sub-catchment	pnatural		the proportion of the sub-catchment buffer with natural land cover types
buffer	sub-catchment	phuman	Independent	the proportion of the sub-catchment buffer that has been modified by human activities (sum of clearcut, forest, agri etc.)
buffer	sub-catchment	deg.slope	Independent	the average slope of the sub-catchment buffer
buffer	sub-catchment	range.slope	Independent	the maximum difference in slope (max - min) in the sub-catchment buffer

Table 5-3. Summary of the ecological questions considered and the specific landscape predictors used in their evaluation. Predictors are defined in Table 5-2. Grey shading was added to aid in interpretability.

	Description of ecological question	Predictors (site)	Predictors (sub-catchment)
Q1	Do densities change longitudinally along the stream network?	site.pos	total.area
Q2	Is habitat use related to flow characteristics?	s.degslope; s.rangeslope	deg.slope; range.slope
Q3	How might the hierarchical structure of rivers influence relationships?	s.degslope; s.rangeslope; s.phuman	deg.slope; range.slope; phuman
Q4	How do populations respond to anthropogenic land use?	s.phuman	phuman
Q5	What is the form of the population response to threats?	s.phuman+s.phuman^2	phuman+phuman^2

Table 5-4. Candidate models and model selection based on AIC and Akaike weights (w_i(AIC)) relative to hypotheses on how juvenile Atlantic salmon distribute in watersheds and are affected by threats (see Table 5-2 for parameter definitions and Table 5-3 for questions). Each model represents a regression fit assuming a negative binomial distribution of juvenile salmon density (#/100m^2). The number of predictors (*k*), log-likelihood (Loglik) and delta AIC are also shown. The top model is identified in bold.

	Predictors	*k*	Loglik	AIC	ΔAIC	w_i(AIC)
NB.full	year,site.pos,s.degslope,s.rangeslope, deg.slope,range.slope,s.phuman,s.phuman$^{\wedge 2}$,phuman,phuman$^{\wedge 2}$	10	-1677	3357	46.3	<0.001
NB1	Year	1	-1667	3342	31.6	<0.001
NB2	year, site.pos	2	-1659	3327	16.4	<0.001
NB3	year, total.area	2	-1661	3332	21.7	<0.001
NB4	year,site.pos, s.degslope	3	-1656	3323	12.3	0.001
NB5	year,site.pos, s.rangeslope	3	-1658	3328	17.1	<0.001
NB6	year,site.pos, deg.slope	3	-1659	3329	18.4	<0.001
NB7	year,site.pos, range.slope	3	-1656	3323	12.6	0.001
NB8	year,site.pos,s.degslope, s.phuman	4	-1655	3325	13.9	<0.001
NB9	year,site.pos,s.degslope, s.phuman,s.phuman$^{\wedge 2}$	5	-1652	3319	8.2	0.009
NB10	year,site.pos, s.degslope, phuman	4	-1650	3314	3.4	0.111
NB11	year,site.pos,s.degslope, phuman,phuman$^{\wedge 2}$	5	-1650	3316	5.2	0.043
NB12	**year,site.pos,s.degslope, s.phuman,s.phuman$^{\wedge 2}$,phuman**	**6**	**-1646**	**3311**	**0**	**0.594**
NB13	year,site.pos,s.degslope,area.est, s.phuman,s.phuman$^{\wedge 2}$,phuman, phuman$^{\wedge 2}$	8	-1646	3313	1.8	0.239

Table 5-5. Parameter type (fixed or random), slope estimates, standard error (s.e), and estimated significance based on Maximum Likelihood estimation (P), as well as the median and 95% credible interval based on MCMC for the NB12 model describing changes in juvenile salmon densities.

	type	slope	s.e.	P	2.5%	median	97.5%
(Intercept)	F	2.07	0.30	<<0.001	1.46	2.08	2.70
year	F	-0.63	0.19	<0.001	-0.94	-0.62	-0.13
s.degslope	F	0.07	0.04	0.043	-0.01	0.08	0.19
site.pos	F	-1.90	0.42	<<0.001	-2.68	-1.92	-0.79
s.phuman	F	2.76	1.02	0.007	0.77	2.69	4.91
s.phuman$^{\wedge 2}$	F	-3.36	1.23	0.007	-6.06	-3.27	-0.82
phuman	F	2.54	0.78	0.001	0.88	2.52	4.11
East River (Chester)	R				-1.63	-0.71	-0.02
Gold River	R				-0.71	-0.11	0.56
LaHave River	R				-0.43	0.26	1.12
Liscomb River	R				-0.89	-0.34	0.21
Medway River	R				-0.91	-0.35	0.26
Middle River (Chester)	R				-2.11	-1.57	-0.92
Musquodoboit River	R				0.31	1.03	1.79
Petite Riviere	R				0.48	1.21	1.93
Salmon River (Guys. Co.)	R				0.55	1.26	1.98
Salmon River (P.D.)	R				-0.82	-0.34	0.17
St. Mary's River	R				-0.71	0.23	1.07
West River (S.H.)	R				-1.33	-0.57	0.02

Figures

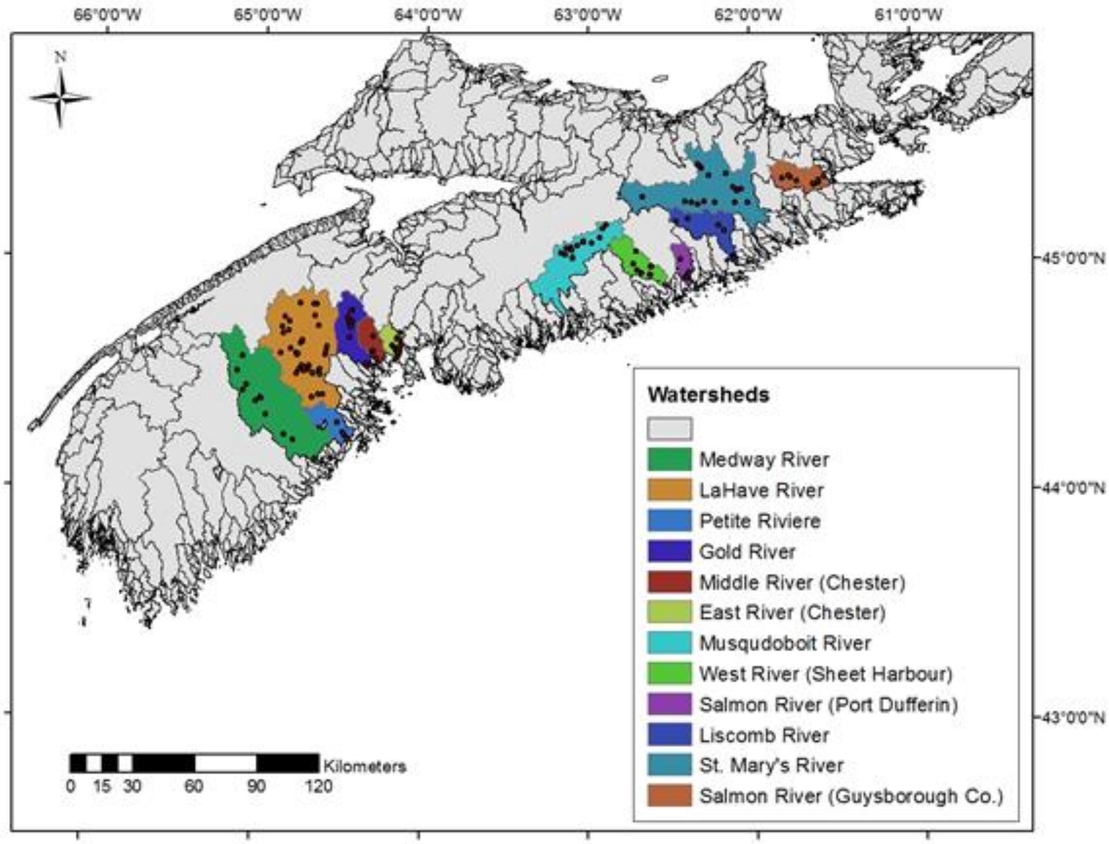

Figure 5-1. Map of the study area in Nova Scotia, Canada, showing the catchment areas of the sampled watersheds (coloured and labeled with the river name) as well as the electrofishing site locations (dots). All other watersheds are shown in grey with black outlines.

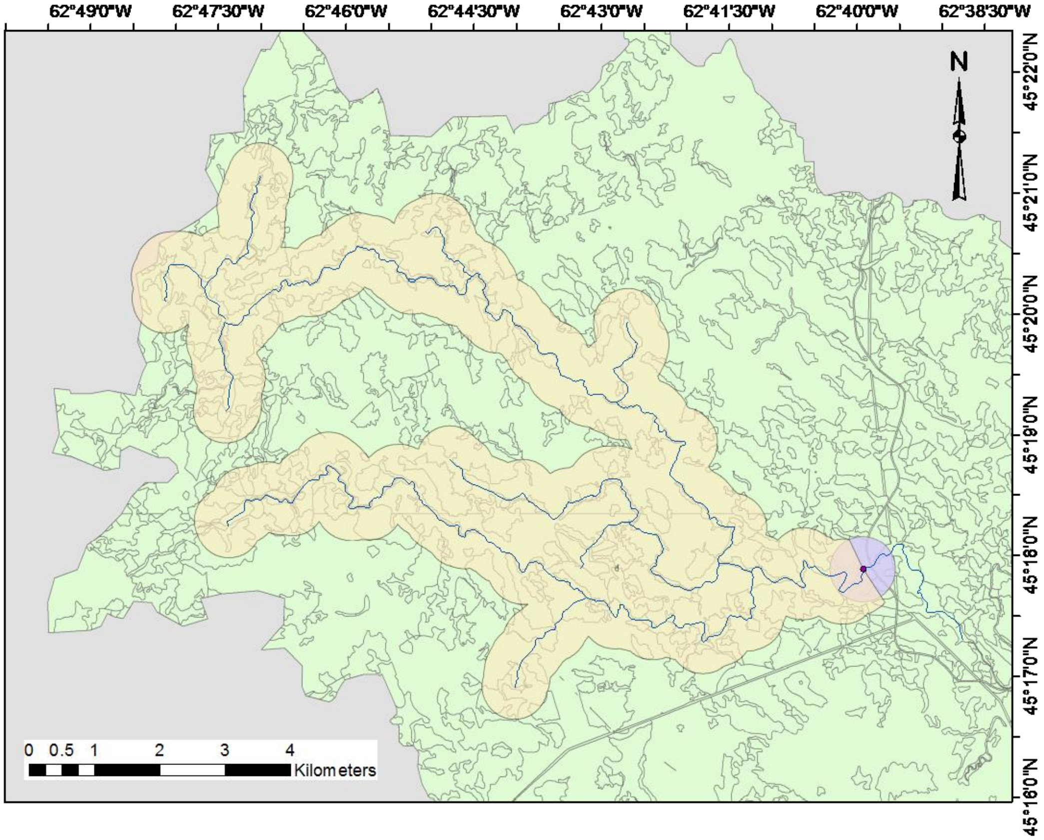

Figure 5-2. An example of the site-scale (purple circle) and sub-catchment (orange polygon) buffers for an electrofishing site (red point) on the Nelson tributary of the St. Mary's River watershed (green). The land use and land cover polygons from the Forest Inventory data are shown as grey lines.

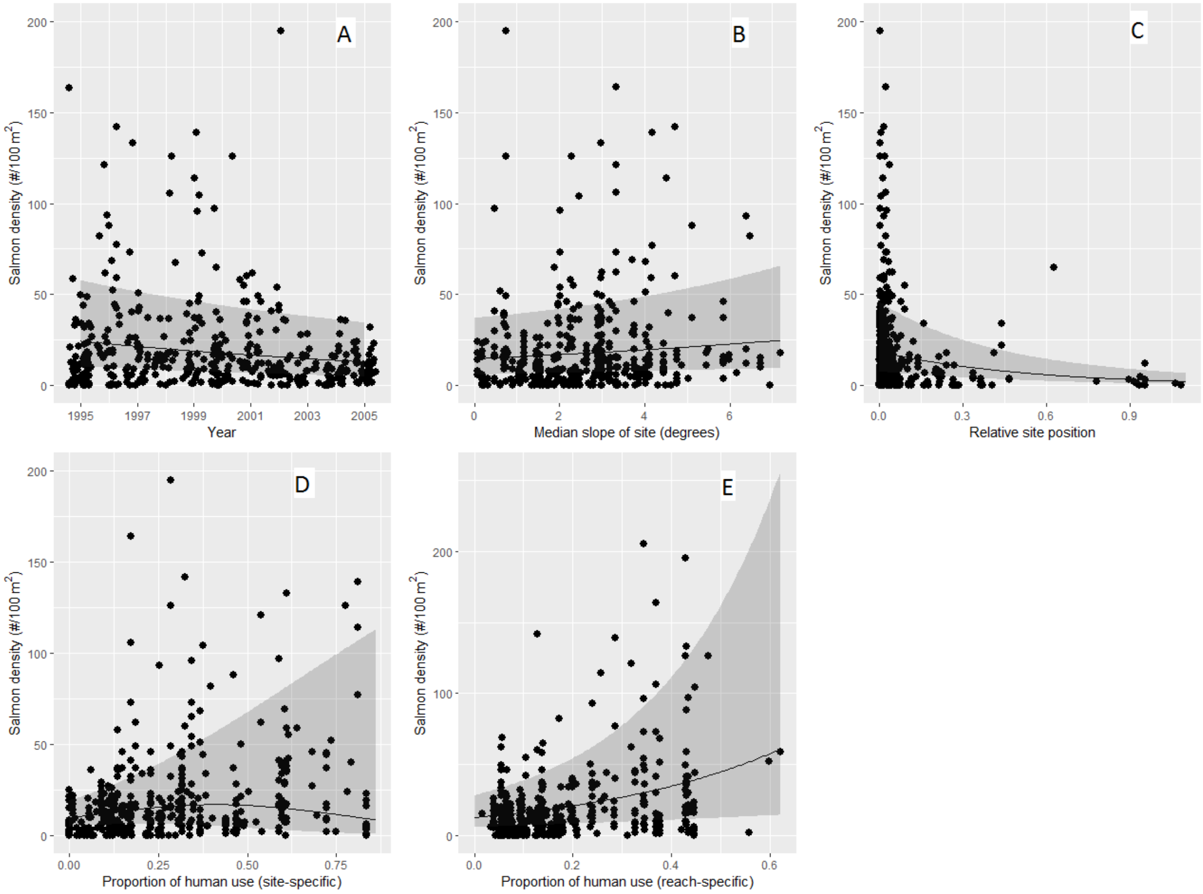

Figure 5-3. Partial regression fits (lines) plus 95% confidence interval (grey shading) relative to the juvenile density data (points) for each of the fixed effects retained in the NB12 model. X-axes are as follows. Panel A: year of electrofishing survey (year), panel B: median slope of a site (s.degslope), panel C: relative site position (site.pos), panel D: site-specific proportion of human use (s.phuman+s.phuman$^{\wedge 2}$), and panel E: proportion of human use in the sub-catchment (phuman). Multiple observations from the same year are off-set slightly along the x-axis in panel A in order to see all of the individual density estimates. See Table 5-2 for parameter definitions.

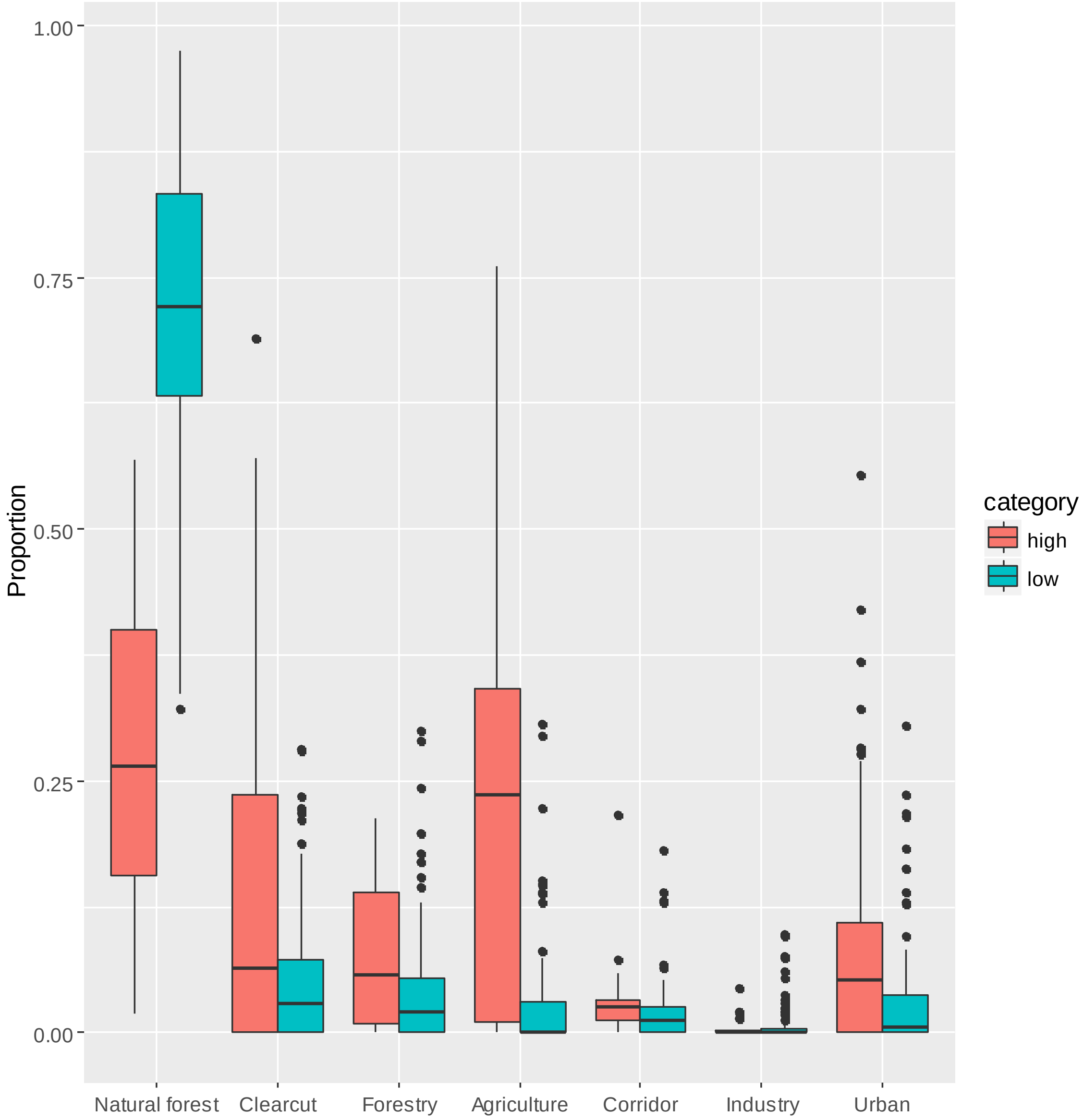

Figure 5-4. Comparison of the distributions of site-scale land use variables between sites categorized as having low levels (blue bars) or high levels (red bars) of human impact. Horizontal lines represent the median and the lower and upper boundaries of the box are drawn at the 25th and 75th percentiles of the distribution, respectively.

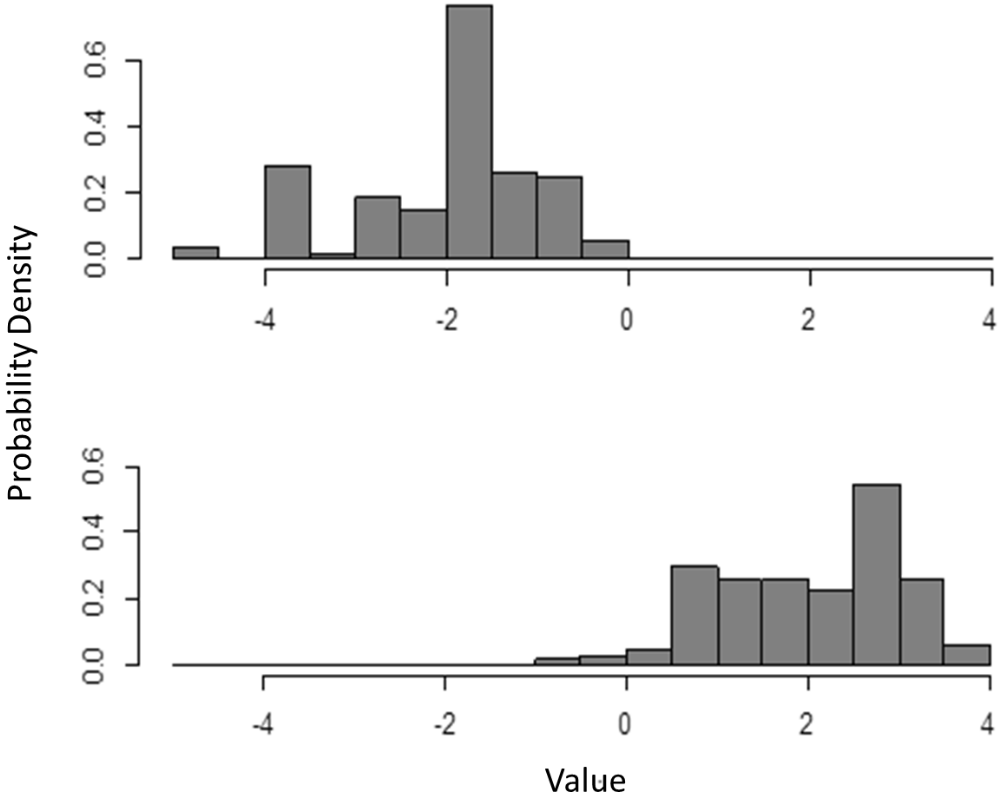

Figure 5-5. Distribution of observations along the discriminant axis separating sites with high proportions of human use (s.phuman > 0.4; top panel) from those with low (s.phuman ≤ 0.4; bottom panel), estimated using Quadratic Discriminant Analysis.

Chapter six: Conclusions and Synthesis

Recovery planning and the effective management of endangered species requires understanding: (1) how ecosystem processes interact across spatial scales to create and maintain the habitats needed by species to complete their life history; (2) how demographic processes vary across space and through time in response to this habitat heterogeneity (Labbe & Fausch 2000); and (3) how spatially variable processes interact to determine the viability of a larger demographic unit. As such, what I have called quantitative recovery planning requires the integration of several fields of theoretical ecology with diverse analytical methodologies appropriate for observational data. It is similar in complexity to fields of inquiry such as landscape genetics, which integrate diverse ideas originating from landscape ecology and population genetics, coupled with analytical methods that merge quantitative genetics with spatial statistics (Balkenhol et al. 2009, Schoville et al. 2012). From my perspective, there is no straightforward or simple approach to follow, which is reflected in the diversity of ideas and analytical methods included in my four research chapters. Environmental variability could be expected to be strongly related to both spatial and demographic factors influencing population-level abundance and distribution (Pearman et al. 2008, Schurr et al. 2012), particularly for species with a high degree of habitat specificity within patchy network environments (Baguette & Schtickzelle 2003). At the same time, populations might be expected to respond to environmental variability through changes in survival, growth, reproduction or behaviour (Rose et al. 2001).

Population responses to environmental variation

Linking changes in population-level abundance and distribution directly to environmental metrics integrates over other life history and evolutionary responses (Sih et al. 2011), which has implications for how we understand population responses to threats. For example, species might respond to environmental variation through changes in growth rates, reproductive strategy, habitat use or distribution (Pearman et al. 2008). Thus, behavioural responses to environmental variation, in addition to the environmental variation itself, would be expected to change the potential for resource acquisition as well as inter- and intra-species interactions experienced by an individual (Bowler & Benton 2005, Kubisch et al. 2014). A wealth of research suggests that individual behavioural strategies are context-dependent and vary with the environmental conditions encountered by individuals (Sih et al. 2011, Kubisch et al. 2014). For example, factors such as the strength of intraspecific or interspecific competition, and local habitat characteristics such as patch size or relative quality have been shown to influence the propensity of an individual to disperse (Bowler & Benton 2005, Poethke et al. 2010). Such context-dependent behavioural strategies would contribute to the potential for resource acquisition or the relative fitness of individuals, and would be expressed at a population level as changes vital rates or dynamics (e.g., Rose et al. 2001), in addition to changes in distribution (Bowler & Benton 2005). If these responses promote a numerical increase at low density and result in reduced population growth rates at high density, they are compensatory and will dampen inter-annual variations in population-level abundance (Rose et al. 2001, Lande et al. 2003). As such, they would help to maintain large population size (Simberloff 1998), and promote population increase following stochastic declines (i.e., improve population resilience, Hutchings et al. 2012b), thus reducing extinction risk.

As I discussed in chapter three, any compensatory response would weaken relationships with environmental change (i.e., threats); reducing detectability (Rose 2000), and having the potential to bias our ecological understanding of species-environment relationships. As I suggested in chapters four and five, this would be similar to ignoring the influence of spatial variation or pattern on the life history dynamics of riverine fishes, which would be expected to bias our understanding of abundance and distribution patterns relative to landscape-level change. I did not directly evaluate the extent to which compensatory processes buffer environmental change in spatially and temporally variable environments, but the idea that these processes would affect both the strength and nature of population-level responses to threats is a re-occurring theme throughout my thesis. In addition, the correlations among threats and genetic distance from chapter four suggest that behavioural responses to environmental variation can be relatively strong, and thus could be a major contributor to demographic change or population dynamics.

Individual behavior depends on the immediate conditions encountered, as constrained by life history characteristics (Kubisch et al. 2014), which means that behaviour is one step removed from the consequences of that behaviour in terms of survival and reproductive success (Robertson & Hutto 2006). To guard against false inference, I evaluated quantitative relationships with threats in the context of other expected ecological patterns arising from changes in behaviour in chapters four and five. Using site-specific density as a proxy for habitat suitability, quality, or productive potential in chapter five (Kubisch et al. 2014), as well as changes in genetic characteristics as a proxy for population connectivity in chapter four, meant I was implicitly assuming that individuals are drawn to suitable habitat types, and that density patterns would elucidate the landscape characteristics necessary to maintain population productivity. However, it is well-known that human-perceived landscape patterns often do not correspond to functionally suitable habitat for specific species or groups of species, particularly when landscapes have been heavily modified (Fischer & Lindenmayer 2007). Individual behaviour can become mal-adaptive if human-induced environmental change uncouples the cues that individuals use to assess habitat quality from the true quality of the environment, making poor quality habitat relatively attractive (termed ecological or evolutionary traps; Schlaepfer et al. 2002, Gilroy & Sutherland 2007). Paradoxically, ecological restoration efforts themselves have been identified as a frequent cause of ecological traps, mainly because of their focus on a single life stage of a single species coupled with an imperfect mechanistic understanding of the ecological relationships being manipulated (Jeffres & Moyle 2012, Robertson et al. 2013). Relative to recovery planning, the unknown existence of ecological traps due to habitat modification or land use in watersheds would be expected to influence our ability to identify threats as well as to quantify population response.

In addition to individual behaviour, it is increasingly being recognized that evolutionary processes can influence population responses to environmental change over contemporary time scales (De Meester & Pantel 2014). Such contemporary evolution would be expected to increase the potential for populations to persist even at low abundance by maximizing the productive potential of the remaining individuals relative to novel environmental conditions and reducing the likelihood of stochastic extinctions (Kinnison & Hairston 2007, Cameron et al. 2013). Fitness is partially determined by the suite of adaptive traits possessed by an individual (Bolnick et al. 2011); where variation among individuals in phenotypic traits would be ultimately expressed as population productivity (Sih et al. 2011). Previous research has found that populations with relative high levels of genetic variation are less sensitive to environmental variation and can maintain more consistent population sizes because they maintain the ability to evolve (reviewed in De Meester & Pantel 2014). In addition, population assemblages characterized by higher rates of gene flow may be better able to maintain the ability for contemporary evolution, given the potential for immigrants to positively contribute to the local phenotypic distribution (Kinnison & Hairston 2007). Experimental research has demonstrated that adaptive responses to environmental change can reverse population declines, and thus have the potential for evolutionary rescue of endangered species (Cameron et al. 2013). However, despite theoretical advances, the degree to which eco-evolutionary processes influence the dynamics of natural populations is relatively unknown, as are the specific conditions in which contemporary evolution would be expected to lead to population persistence (De Meester & Pantel 2014).

Atlantic salmon populations in the Southern Upland still have considerable genetic variation (chapter four) and are linked through dispersal, so it is possible that contemporary evolution might be occurring in response to environmental change. The simulations in chapter two did not allow for contemporary evolution, which may be expected to increase the productivity of individual populations over time. However, population dynamics modeling for salmon in the Southern Upland suggests that productivity is extremely low, such that the few remaining populations are close to being deterministically extinct (Gibson & Bowlby 2013). The observation that populations still persist in SU watersheds may reflect an extinction debt, where deterministically extinct populations may persist for a considerable amount of time in the landscape prior disappearing and in the absence of further environmental change (Tilman et al. 1994, Kuussaari et al. 2009). Critically endangered populations are often thought to be affected by mutually-enforcing detrimental changes to demographic, genetic, or behavioural processes resulting in population decline (Fagan & Holmes 2006), which can be exacerbated by synergistic interactions between multiple threats (Darling & Cote 2008). By studying the population dynamics of species that have actually gone extinct, Fagan & Holmes (2006) found that year-to-year rates of decline as well as year-to-year variability increased as the time to extinction decreased, two characteristics that are predicted by theory (Lande et al. 2003). However, without knowing that a population is actually destined for extinction, it is difficult to use such characteristics to approximate extinction risk, or to determine the ultimate fate of a population. Although Southern Upland Atlantic salmon populations are at critically low abundance (Gibson et al. 2011) and are presently characterized by extremely low productivity (Gibson & Bowlby 2013), there is still the potential to implement recovery actions for the remaining populations. Ideally, these would slow or reverse population declines and maximize the potential for evolutionary rescue.

Practical application for SU Atlantic Salmon

Ecological theory developed in relation to conservation questions is very useful to contextualize the scope for population responses, yet it can have little immediate practical relevance to recovery planning for specific populations of endangered species (Driscoll & Lindenmayer 2012). When theoretical predictions are evaluated using data from natural populations, invariably there is evidence that supports as well as contradicts the hypothesis being tested; some examples are related to: ecological synergies (Brook et al. 2008, Darling & Cote 2008), eco-evolution (Kinnison & Hairston 2007), species-area relationships and spatial pattern (Thornton et al. 2011), metapopulations (Fronhofer et al. 2012), or ecological correlates with biodiversity patterns (Chase & Knight 2013). Typically and understandably, authors conclude that different processes are important in different systems, and often call for further refinement of the relative contribution of demographic, evolutionary, spatial or ecological factors affecting populations in order to better understand dynamics, demographic change or the extinction process. While this has led to a much better understanding of the range of potential population responses to environmental variability, it does not translate into specific practical guidance for conservation, unless the guidance is that recovery planning is case-specific. Furthermore, assuming that a particular theory describes a specific species or system *a priori* can constrain the focus of research programs to the exclusion of alternate hypotheses and can lead to inefficient or ineffective management or recovery strategies being applied (as demonstrated in chapter two; Haila 2002, Driscoll & Lindenmayer 2012, Jeffres & Moyle 2012, Stoll et al. 2016).

From my perspective, the key questions for recovery planning for Atlantic salmon in the Southern Upland region include: (1) should a large or small number of watersheds be targeted for recovery simultaneously? (2) Which remediation actions would be expected to have the largest effect on productivity for each specific river? And (3) is our understanding of the relative severity of different types of threats accurate? Relative to question (1), the metapopulation simulations in chapter two suggest that recovery planning should focus on increasing productivity in the remaining populations in large watersheds. If there was a choice between distributing a low level of remediation effort among many watersheds relative to a higher level of effort in a smaller number of watersheds, the simulations in chapter two support the latter. Working in smaller watersheds might be more tractable in terms of the relative amount of the stream network and surrounding landscape that could be affected by remediation, yet would be expected to have little influence on regional persistence. Within these larger watersheds, the spatial landscape database and the quadratic discriminant axis detailed in chapter five could be used to identify areas that may be priorities for restoration, based on the proportion of human activity in the vicinity of the stream network. Although the hydrology analyses in chapter three suggests specific remediation actions that would be expected to lead to population increase for the St. Mary’s and Nashwaak Rivers, I was not able to quantify population responses to the magnitude of multiple threats at a regional level (question 2 above). As such, these analyses do not provide substantial information on the specific amount of a stream network or the specific remediation actions that would be expected to produce the greatest population response.

Relative to (3), results from chapters two, four and five challenge some of the current thinking on the importance or impact of specific types of threats affecting Southern Upland Atlantic salmon. If straying becomes detrimental to population persistence as productivity declines (as suggested by chapter two), threats that contribute to increased stray rates should have a higher perception of harm. This would not necessarily have been true when population declines started, in that productivity was substantially higher in the 1980s (Gibson & Bowlby 2013) and thus source populations would have been expected to better compensate for emigration. Such switches in the relative importance of specific threats over time are consistent with the general expectation that the original causes of population decline can become separated from the factors keeping populations at low abundance (Fisher & Lindenmayer 2007). Also counter to current perceptions, my landscape genetics analysis in chapter four implies that several previously unrecognized threats, such as historical mining activity, influence effective dispersal in a manner similar to well-described and substantial threats such as acidification (Watt 1987, Korman et al. 1994). However, I was not able to evaluate how attraction or avoidance behaviour by emigrating adults relative to watershed characteristics would culminate in changes to population dynamics in that analysis. Further evaluation or at least the consideration of relatively unstudied threats would be warranted during recovery planning; particularly because the analysis in chapter five suggests that all types of human activity in watersheds cause population decline once they collectively reach a certain threshold. The analyses did not support the general perception that a specific practice such as clearcutting would exert the greatest influence on habitat distribution and quality in stream networks for juvenile salmon, relative to other types land use (Gilvear et al. 2002, Allan 2004). Unfortunately, this complicates recovery planning in that it becomes difficult to prioritize among remediation actions designed to address different types of threats, unless spatially-explicit restoration plans are developed for individual rivers.

Methodological considerations

Pattern does not equal process (Clinchy et al. 2002) and using observational data for ecological inquiry is inherently difficult. Similar time series can arise through multiple ecological mechanisms (Clinchy et al. 2002, Benton et al. 2006). Apparently strong and ecologically plausible environmental correlations often break down with the addition of new data (Myers 1998, Hilborn & Walters 1992). The sampling scheme itself or other study design issues have been found to be correlated with our general understanding of specific ecological relationships, such as species-area relationships (e.g., Drakare et al. 2006, Mortelliti et al. 2010) or biodiversity patterns (Chase & Knight 2013), as well as questions of spatial scale and patterning in landscape ecology (Thornton et al. 2011). Throughout this thesis, I attempted to explicitly consider potential pitfalls in the data I had available, in the indices or environmental metrics evaluated, as well as in the analytical methods I used. For example, I included robust regression in chapter three to explicitly account for sampling errors and extreme values in the hydrological data that could lead to spurious relationships or could mask significant ones, particularly for such short time series of egg to age 0 survival. When it was not possible to control for specific characteristics, such as the behaviour of electrofishing crews affecting the juvenile density data used in chapter five, I evaluated spatial grain relative to other apparent ecological patterns in the broader analysis. In all chapters, I looked for overall consistency as related to theory, which is why I proposed that Isolation by Environment (IBE) was more likely to characterize genetic patterns for Atlantic salmon relative to Isolation by Distance (IBD) in chapter four; the threats relationships could not be explained unless this was true.

From my perspective, the largest potential pitfalls affecting the landscape genetics analysis in chapter four as well as the analysis of the electrofishing data in chapter five relate to (1) a potential temporal mis-match between data sources and (2) time lags between environmental change and population-level response. Although the land use data and electrofishing data used in chapter five were collected during similar timeframes, I was not able to ensure that both sources of data represented the same year for particular sites. Relative to population responses, landscape modification causes both acute and delayed changes to the spatial distribution and relative quality of stream habitats (Nislow 2005, McCluney et al. 2014). Therefore, even if both data sources represented the same year, there could still be a time lag between changes in land cover and population response. Similarly, in order to evaluate threats using landscape genetics, a sufficient number of generations need to have passed since the threat originated for any effect to become measureable (Landguth et al. 2010). I was not able to directly test whether my results were sensitive to these temporal constraints due to data availability, although it is likely that they would have had some impact on the relationships described (Graham 2003). If these analyses were to be extended, there would be the potential to substitute the Nova Scotia Change Atlas: 1990 – 2007 (available from Global Forest Watch Canada; Cheng & Lee 2009) as an alternate characterisation of land cover. Although this would likely rely on using electrofishing data or genetic samples collected after 2007 (as opposed to those from 1995 to 2005 as used here), it might be one way to incorporate a temporal aspect into these analyses.

Summary

In my thesis, I have taken a single-species approach to understanding population responses to environmental change and to recovery planning. This was largely because of the inherent complexity in developing quantitative relationships between landscape-level variation and *population-level* response. The relationships I describe are among the first to be developed for endangered Atlantic salmon populations in Nova Scotia (although see Watt 1987 relative to acidification).

By developing a simulation model for a simplified diadromous life history in chapter two, I was able to evaluate how spatial structure and demographic variability integrate to determine regional abundance and extinction risk of a population assemblage. My results suggest that maintaining metapopulation structure should not be the primary focus for conservation and management of diadromous fishes when population productivity is low; demographic processes seem to be more important determinants of persistence relative to spatial processes. In chapter three, I demonstrated how our understanding of survival-hydrology relationships can be directly affected by decisions made *prior* to analyses. Responses to hydrological change were population-specific and would necessitate population-specific remediation plans. In chapter four, I proposed that population connectivity in diadromous fishes should be evaluated relative to environmental patterns rather than geographic distance, particularly when using landscape genetic analyses to evaluate contemporary threats. Results suggested that straying Atlantic salmon avoid watersheds with reduced water quality and higher road density, yet were not responding to watershed fragmentation at a population level. In chapter five, I was able to develop a relatively simple tool that could inform management decisions or identify priority areas for restoration efforts, even though I was not able to fully evaluate theoretical predictions on the influence of environmental variation on habitat utilization, distribution patterns, or population-level responses to human activities at multiple spatial scales.

This research was motivated by a single question from a long-time angler and conservation advocate: *“What should we do to bring salmon back?”* In trying to answer this question, I hope that this research, as well as the expertise I have gained, will substantially contribute to future recovery planning and remediation efforts for endangered species, particularly Atlantic salmon populations in the Southern Upland region of Nova Scotia, Canada.

References

Aeberhard, W.H., E. Cantoni & S. Heritier. 2014. Robust inference in the negative binomial regression model with an application to falls data. Biometrics 70: 920-931.

Akcil, A. & S. Koldas. 2006. Acid Mine Drainage (AMD): Causes, treatment and case studies. Journal of Cleaner Production 14:1139-1145.

Albanese, B., P.L. Angermeier & J.T. Peterson. 2009. Does mobility explain variation in colonization and population recovery among stream fishes? Freshwater Biology 54:1444-1460.

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

Allendorf, F.W., G. Luikart & S.N. Aitken. 2013. Conservation and the Genetics of Populations, Second Edition. Blackwell Publishing Ltd. UK

Altermatt, F. 2013. Diversity in riverine metacommunities: a network perspective. Aquatic Ecology 47:365-377.

Amiro, P.G., A.J.F. Gibson & H.D. Bowlby. 2006. Atlantic salmon (*Salmo salar*) overview for eastern Cape Breton, Eastern Shore, Southwest Nova Scotia and inner Bay of Fundy rivers (SFA 19 to 22) in 2005. DFO Canadian Science Advisory Secretariat Research Document 2006/024.

Anderson, D.R., K.P. Burnham & G.C. White. 1994. AIC model selection in overdispersed capture-recapture data. Ecology 75:1780-1793.

Arens, P., T. van der Sluis, W.P.C. van't Westende, B. Vosman, C.C. Vos & M.J.M. Smulders. 2007. Genetic population differentiation and connectivity among fragmented Moor frog (*Rana arvalis*) population in the Netherlands. Landscape Ecology 22:1489–1500.

Baguette, M. & N. Schtickzelle. 2003. Local population dynamics are important to the conservation of metapopulations in highly fragmented landscapes. Journal of Applied Ecology 40:404-412.

Balkenhol, N., L.P. Waits & R.J. Dezzani. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography 32:818–830.

Barton, P.S., S.A. Cunningham, A.D. Manning, H. Gibb, D.B. Lindenmayer & R.K. Didham. 2013. The spatial scaling of beta diversity. Global Ecology and Biogeography 22:639-647.

Bash. J.S. & C.M. Ryan. 2002. Stream restoration and enhancement projects: Is anyone monitoring? Environmental Monitoring 29:877-85.

Baxter, C.V. & F.R. Hauer. 2000. Geomorphology, hyporheic exchange, and selection of spawning habitat by bull trout (*Salvelinus confluentus*). Canadian Journal of Fisheries and Aquatic Sciences 57:1470-1481.

Beckmann, J.P. 2011. Understanding the role of behavioral mechanisms and impacts of habitat fragmentation on structuring small declining populations. Animal Conservation 14:229-230.

Benda, L, N. LeRoy Poff, D. Miller, T. Dunne, G. Reeves, G. Pess & M. Pollock. 2004. The network dynamics hypothesis: how channel networks structure riverine habitats. BioScience 54:413-427.

Bender, D.J. & L. Fahrig. 2005. Matrix structure obscures the relationship between interpatch movement and patch size and isolation. Ecology 86:1023-1033.

Benton, T.G., S.J. Plaistow & T.N. Coulson. 2006. Complex population dynamics and complex causation: devils, details and demography. Proceedings of the Royal Society B 273:1173-1181.

Berk, R.A. 1990. A primer on robust regression. Modern Methods of Data Analysis [eds.] J. Fox and J.S. Long. pp. 293-324. Sage Publications. London.

Blackburn, T.M. & K.J. Gaston. 1996. Abundance-body size relationships: the area you census tells you more. Oikos 75:303-309.

Blanchet, S., J.J. Dodson & S. Brosse. 2006. Influence of habitat structure and fish density on Atlantic salmon *Salmo salar* L. territorial behaviour. Journal of Fish Biology 68:951-957.

Bocard, D., F.Gillet & P. Legendre. 2011. Numerical Ecology with R. Springer Science and Business Media LLC. New York, NY.

Bolnick, D.I., P. Amarasekare, M.S. Araujo, R. Burger & others. 2011. Why intraspecific trait variation matters in community ecology. Trends in Ecology and Evolution 26:183-192.

Bonte, D., H. van Dyck, J.M. Bullock, A. Coulon & others. 2012. Costs of dispersal. Biological Reviews 87:290-312.

Bowlby, H.D. & A.J.F. Gibson. 2011. Reduction in fitness limits the useful duration of supplementary rearing in an endangered salmon population. Ecological Applications 21:3032-3048.

Bowlby, H.D. & A.J.F. Gibson. 2012. Inferring adult status and trends from juvenile density data for Atlantic salmon. North American Journal of Fisheries Management 32:1225-1236.

Bowlby, H.D., A.J.F. Gibson & A. Levy. 2013. Recovery Potential Assessment for Southern Upland Atlantic salmon: Status, past and present abundance, life history and trends. DFO Canadian Science Advisory Secretariat Research Document 2013/005.

Bowlby, H.D., T. Horsman, S.C. Mitchell & A.J.F. Gibson. 2014. Recovery Potential Assessment for Southern Upland Atlantic Salmon: Habitat Requirements and Availability, Threats to Populations, and Feasibility of Habitat Restoration. DFO Canadian Science Advisory Secretariat Research Document 2013/006.

Bowler, D.E. & T.G. Benton. 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. Biological Reviews of the Cambridge Philosophical Society 80:205-225.

Bradburd, G., P. Ralph & G.M. Coop. 2013. Disentangling the effects of geographic and ecological isolation on genetic differentiation. Evolution. doi:10.1111/evo.12193.

Bradbury, I.R., L.C. Hamilton, M.J. Robertson, C.E. Bourgeois, A. Mansour & J.B. Dempson. 2014. Landscape structure and climatic variation determine Atlantic salmon genetic connectivity in the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences 71:246-258.

Brett, M.T. 2004. When is a correlation between non-independent variables “spurious”? Oikos 105:647-656.

Broadmeadow, S. & T.R. Nisbet. 2004. The effects of riparian forest management on the freshwater environment: a literature review of best management practice. Hydrology and Earth System Sciences 8:286-305.

Broennimann, O., M.C. Fitzpatrick, P.B. Pearman, B. Petitpierre, L. Pellissier, & others. Measuring ecological niche overlap from occurrence and spatial environmental data. Global Ecology and Biogeography 21:481-497.

Brook, B.W., N.S. Sodhi & C.J.A. Bradshaw. 2008. Synergies among extinction drivers under global change. Trends in Ecology and Evolution 23:453-460.

Brown, B.L. & C.M. Swan. 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. Journal of Animal Ecology 79:571-580.

Brown, J.H. & A. Kodric-Brown. 1977. Turnover rates in insular biogeography: effect of immigration on extinction. Ecology 58:445-449.

Bult, T.P., R.L. Haedrich & D.C. Schneider. 1998. New technique describing spatial scaling and habitat selection in riverine habitats. Regulated Rivers: Resource Management 14:107-118.

Bunn, S.E. & A.H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management 30:492-507.

Burdahl, A., P.A.H. Westley, S.A. Levin, I.D. Couzin and T.P. Quinn. 2014. A collective navigation hypothesis for homeward migration in anadromous salmonids. Fish and Fisheries. doi: 10.1111/faf.12084.

Burke, J.M. & M.L. Arnold. 2001. Genetics and the fitness of hybrids. Annual Review of Genetics 35:31-52.

Burnham, K.P. & D.R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer, NY.

Caissie, D. 2006. The thermal regime of rivers: a review. Freshwater Biology 51:1389-1406.

Cameron, T.C., D. O'Sullivan, A. Reynolds, S.B. Piertney & T.G. Benton. 2013. Eco-evolutionary dynamics in response to selection on life-history. Ecology Letters 16:754-763.

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

Cantoni, E. & E. Ronchetti. 2001. Robust inference for Generalized Linear Models. Journal of the American Statistical Association 96:1022-1030.

Castillo, J.A., C.W. Epps, A.R. Davis & S.A. Cushman. 2014. Landscape effects on gene flow for a climate-sensitive montane species, the American pika. Molecular Ecology 23:843-856.

Chaput, G., J.B. Dempson, F. Caron, R. Jones & J. Gibson. 2006. A synthesis of life history characteristics and stock grouping of Atlantic salmon (*Salmo salar* L.) in eastern Canada. DFO Canadian Science Advisory Secretariat Research Document 2006/015.

Chase, J.M & T.M. Knight. 2013. Scale-dependent effect sizes of ecological drivers on biodiversity: why standardized sampling is not enough. Ecology Letters 16:17-26.

Cheng, R. & P.G. Lee. 2009. Recent (1990-2007) anthropogenic change within the forest landscapes of Nova Scotia. Edmonton, AB: Global Forest Watch Canada. 62 pp.

Chislock, M.F., E. Doster, R.A. Zitomer & A.E. Wilson. 2013. Eutrophication: causes, consequences, and controls in aquatic ecosystems. Nature Education Knowledge 4:10. 8pp.

Clark, R.W., W.S. Brown, R. Stechert & K.R. Zamudio. 2010. Roads, interrupted dispersal, and genetic diversity in timber rattlesnakes. Conservation Biology 24:1059-1069.

Clinchy, M., D.T. Haydon & A.T. Smith. 2002. Pattern does not equal process: what does patch occupancy really tell us about metapopulation dynamics? The America Naturalist 159:351-362.

Collett, D. 2003. Modelling binary data, Second edition. Texts in Statistical Science Series, Chapman & Hall/CRC, Boca Raton.

COSEWIC. (2010) COSEWIC assessment and status report on the Atlantic salmon *Salmo salar* Nunavik population, Labrador population, Northeast Newfoundland population, South Newfoundland population, Southwest Newfoundland population, Northwest Newfoundland population, Quebec Eastern North Shore population, Quebec Western North Shore population, Anticosti Island population, Inner St. Lawrence population, Lake Ontario population, Gaspe-Southern Gulf of St. Lawrence population, Eastern Cape Breton population, Nova Scotia Southern Upland population, Inner Bay of Fundy population Outer Bay of Fundy population in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. (www.sararegistry.gc.ca/status/status_e.cfm).

Courchamp, F., L. Berec & J. Gascoigne. 2008. Allee effects in ecology and conservation. Oxford University Press, Oxford, UK.

Cram, J.M., C.E. Torgersen, R.S. Klett, G.R. Pess, D. May, T.N. Pearsons & A.H. Dittman. 2013. Tradeoffs between homing and habitat quality for spawning site selection by hatchery origin Chinook salmon. Environmental Biology of Fish 96:109–122.

Cushman, S.A., K.S. McKelvey, J. Hayden & M.K. Schwartz. 2006. Gene-flow in complex landscapes: testing multiple models with causal modeling. American Naturalist 168:486–499.

Cushman, S.A., T.N. Wasserman, E.L. Landguth & A.J. Shirk. 2013. Re-evaluating causal modeling with Mantel tests in landscape genetics. Diversity 5:51-72.

Darling, E.S. & I.M. Cote. 2008. Quantifying the evidence for ecological synergies. Ecology Letters 11:1278-1286.

Datry, T., H. Pella, C. Leigh, N. Bonada & B. Hugueny. 2016. A landscape approach to advance intermittent river ecology. Freshwater Biology. doi:10.1111/fwb.12645.

Davey, C. & M. LaPointe. 2007. Sedimentary links and the spatial organization of Atlantic salmon (*Salmo salar*) spawning habitat in a Canadian Shield river. Geomorphology 83:82-96.

De Meester, L. & J. Pantel. 2014. Eco-evolutionary dynamics in freshwater systems. Journal of Limnology 73:193-200.

Dennis, R.L.H., T.G. Shreeve & H. van Dyck. 2003. Towards a functional resource-based concept for habitat: a butterfly biology viewpoint. Oikos 102:417-426.

Deschenes, J. & M.A. Rodriguez. 2007. Hierarchical analysis of relationships between brook trout (*Salvelinus fontinalis*) density and stream habitat features. Canadian Journal of Fisheries and Aquatic Sciences 64:777-785.

DeWoody, J.A. & J.C. Avise. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. Journal of Fish Biology 56: 461-473.

Dick, E.J. 2004. Beyond ‘lognormal versus gamma’: discrimination among error distributions for generalized linear models. Fisheries Research 70:351-366.

Dileo, M.F., J.D. Rouse, F.A. Davila & S.C. Lougheed. 2013. The influence of landscape on gene flow in the eastern massasauga rattlesnake (*Sistrurus c. catenatus*): insight from computer simulations. Molecular Ecology 22:4483-4498.

Dionne, M., F. Caron, J. Dodson & L. Bernatchez. 2009. Comparative survey of within-river genetic structure in Atlantic salmon: relevance for management and conservation. Conservation Genetics 10:869–879.

Dittman, A.H., D. May, D.A. Larsen, M.L. Moser, M. Johnston & D. Fast. 2010. Homing and spawning site selection by supplemented hatchery- and natural-origin Yakima River spring Chinook salmon. Transactions of the American Fisheries Society 139:1014–1028.

Dormann, C.F., J. Elith, S. Bacher, C. Buchmann, G. Carl & others. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27-46.

Drakare, S. J.J. Lennon & H. Hillebrand. 2006. The imprint of the geographical, evolutionary and ecological context on species-area relationships. Ecology Letters 9:215-227.

Drake, J.M. & D.M. Lodge. 2006. Allee effects, propagule pressure and the probability of establishment: risk analysis for biological invasions. Biological Invasions 8:365-375.

Driscoll, D.A. 2007. How to find a metapopulation. Canadian Journal of Zoology 85:1031-1048.

Driscoll, D.A. & D.B. Lindenmayer. 2012. Framework to improve the application of theory in ecology and conservation. Ecological Monographs 82:129-147.

Dudgeon, D., A.H. Arthington, M.O. Gessner & others. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews 81:163-182.

Dulvy, N.K., J.R. Ellis, N.B. Goodwin, A. Grant, J.D. Reynolds & S. Jennings. 2004. Methods of assessing extinction risk in marine fishes. Fish and Fisheries 5:255-276.

Dyer, R.J. 2012. The gstudio package. Virginia Commonwealth University, Virginia.

Earn, D.J.D., S.A. Levin & P. Rohani. 2000. Coherence and conservation. Science 290:1360-1364.

Eaton, M.J., P.T. Hughes, J.E. Hines & J.D. Nichols. 2014. Testing metapopulation concepts: effects of patch characteristics and neighborhood occupancy on the dynamics of an endangered lagomorph. Oikos 123:662-676.

Ellner, S.P. & G. Fussmann. 2003. Effects of successional dynamics on metapopulation persistence. Ecology 84:882-889.

Engler, J.O., N. Balkenhol, K.J. Filz, J.C. Habel & D. Rödder. 2014. Comparative landscape genetics of three closely related sympatric Hesperid butterflies with diverging ecological traits. PLoS ONE 9:e106526

Epps, C.W., P.J. Palsboll, J.D. Wehausen, G.K. Roderick, R.R. Ramey & D.R. McCullough. 2005. Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecology Letters 8:1029–1038.

Evans, J.A. & A.S. Davis. 2011. Consequences of parameterization and structure of applied demographic models: a comment on Pardini et al. (2009). Ecological Applications 21:608-613.

Fagan, W.F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. Ecology 83:3243-3249.

Fagan, W.F. & E.E. Holmes. 2006. Quantifying the extinction vortex. Ecology Letters 9:51-60.

Fairchild, W.L., E.O Swansburg, J.T. Arsenault & S.B. Brown. 1999. Does an association between pesticide use and subsequent declines in catch of Atlantic salmon (*Salmo salar*) represent a case of endocrine disruption? Environmental Health Perspectives 107:349-357.

Farmer, G.J., T.R. Goff, D. Ashfield & H.S. Samant. 1980. Some effects of the acidification of Atlantic salmon rivers in Nova Scotia. Canadian Technical Report of Fisheries and Aquatic Sciences 972.

Fausch, K.D., C.E. Torgersen, C.V. Baxter & H.W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52:483-498.

Ferrier, S., G. Manion, J. Elith & K. Richardson. 2007. Using generalized dissimilarity modelling to analyze and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions 13:252-264.

Figge, F. 2004. Bio-folio: applying portfolio theory to biodiversity. Biodiversity and Conservation 13:827-849.

Fischer, J. & D.B. Lindenmayer. 2007. Landscape modification and habitat fragmentation: a synthesis. Global Ecology and Biogeography 16:265-280.

Fitzpartick, M.C. & S.R. Keller. 2015. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. Ecology Letters 18:1-16.

Foley, J.A., R. DeFries, G.P. Asner, C. Barford & others. 2005. Global consequences of land use. Science 309:570-574.

Forman, R.T.T. & L.E. Alexander. 1998. Roads and their major ecological effects. Annual Review of Ecology and Systematics 29:207–257.

Foster, D., F. Swanson, J. Aber, I. Burke, N. Brokaw, D. Tilman & A. Knapp. 2003. The importance of land-use legacies to ecology and conservation. BioScience 53:77-88.

Fournier, D.A., H.J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. Maunder, A. Nielsen & J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods Software 27:233-249.

Fraser, D.J., L.K. Weir, L. Bernatchez, M.M. Hansen & E.B. Taylor. 2011. Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. Heredity 106:404-420.

Fraser, D.J., P.V. Debes, L. Bernatchez & J.A. Hutchings. 2014. Population size, habitat fragmentation, and the nature of adaptive variation in a stream fish. Proceedings of the Royal Society B: Biological Sciences 281: 20140370.

Frederiksen, M., J.-D. Lebreton, R. Pradel, R. Choquet & O. Gimenez. 2014. Identifying links between vital rates and environment: a toolbox for the applied ecologist. Journal of Applied Ecology 51:71-81.

Fretwell, S.D. & H.L. Lucas. 1970. On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheoretica 19:136-156.

Frissell, C.A., W.J. Liss, C.E. Warrren & M.C. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10:199-214.

Fronhofer, E.A., A. Kubisch, F.M. Hilker, T. Hovestadt & H.J. Poethke. 2012. Why are metapopulations so rare? Ecology 93:1967-1978.

Fukuda, N., J. Aoyama, M.J. Miller, Y. Suzuki & K. Tsukamoto. 2013. Discriminant analysis of age to enable rapid estimation of 0-age Japanese eels *Anguilla japonica*. Coastal Marine Science 36:19-25.

Gahagan, B.I., J.C. Vokoun, G.W. Whitledge & E.T. Schultz. 2012. Evaluation of otolith microchemistry for identifying natal origin of anadromous river herring in Connecticut. Marine and Coastal Fisheries: Dynamics Management and Ecosystem Science 4:358-372.

Garcia de Leaniz C, I.A. Fleming, S. Einum, E. Verspoor, & others. 2007. A critical review of inherited adaptive variation in Atlantic salmon. Biological Reviews 82:173-211.

Garnier-Gere, P. & C. Dillmann. 1992. A computer program for testing pairwise linkage disequilibrium in subdivided populations. Journal of Heredity 83:239.

Gibson, R.J. 1993. The Atlantic salmon in fresh water: spawning, rearing and production. Reviews in Fish Biology 3:39-73.

Gibson, A.J.F. 2006. Population regulation in Eastern Canadian Atlantic salmon (*Salmo salar*) populations. DFO Canadian Science Advisory Secretariat Research Document, 2006/016.

Gibson, A.J.F. & R.A. Myers. 2003. A meta-analysis of the habitat carrying capacity and maximum reproductive rate of anadromous alewife in Eastern North America. American Fisheries Society Symposium 35:211-221.

Gibson, A.J.F., H.D. Bowlby & P.G. Amiro. 2008. Are wild populations ideally distributed? Variations in density-dependent habitat use by age class in juvenile Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences 65:1667-1680.

Gibson, A.J.F., R.A. Jones & H.D. Bowlby. 2009. Equilibrium analyses of a population's response to recovery activities: a case study with Atlantic salmon. North American Journal of Fisheries Management 29:958-974.

Gibson, A.J.F., H.D. Bowlby, D. Hardie & P. O'Reilly. 2011. Populations on the brink: Atlantic salmon (*Salmo salar*) in the Southern Upland Region of Nova Scotia, Canada. North American Journal of Fisheries Management 31:733–741.

Gibson, A.J.F. & H.D. Bowlby. 2013. Recovery potential assessment for Southern Upland Atlantic Salmon: population dynamics and viability. DFO Canadian Science Advisory Secretariat Research Document, 2012/142.

Gibson, A.J.F., H.D. Bowlby& A.L. Levy. 2015. Dynamics of endangered eastern Cape Breton Atlantic Salmon populations. North American Journal of Fisheries Management 35:372-387.

Gibson, A.J.F., R.A. Jones & G.J. MacAskill. 2016. Recovery potential assessment for outer Bay of Fundy Atlantic Salmon (*Salmo salar*): population dynamics and viability. DFO Canadian Science Advisory Secretariat Research Document, 2016/032.

Gibson, A.J.F., H.D. Bowlby & A.L. Levy. 2015. Dynamics of endangered Eastern Cape Breton Atlantic salmon populations. North American Journal of Fisheries Management. 35:372-387.

Gibson, R.J. 1993. The Atlantic salmon in fresh water: spawning, rearing and production. Reviews in Fish Biology and Fisheries 3:39-73.

Gibson, R.J., R.L. Haedrich & C.M. Wernerheim. 2005. Loss of fish habitat as a consequence of inappropriately constructed stream crossings. Fisheries 30:10-17.

Gilroy, J.J. & W.J. Sutherland. 2007. Beyond ecological traps: perceptual errors and undervalued resources. Trends in Ecology and Evolution 22:351-356.

Gilvear, D.J., K.V. Heal & A. Stephen. 2002. Hydrology and ecological quality of Scottish river ecosystems. Science of the Total Environment 294:131-159.

Google Inc. (2005). Google Earth. Mountain View, CA

Gordon, B., P.E. Rothrock & P. Labus. 2016. Testing the use of best professional judgement to create biological benchmarks for habitat assessment of wetlands and oak savannas in northwestern Indiana. Ecological Indicators 60:410-419.

Goslee, S.C. & D.L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22:1–19.

Graham, M.H. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809-2815.

Grant, J.W.A., S.O. Steingrimsson, E.R. Keeley & R.A. Cunjak. 1998. Implications of territory size for the measurement and prediction of salmonid abundance in streams. Canadian Journal of Fisheries and Aquatic Sciences. 55(Suppl.1):181-190.

Greenwood, M.J., J.S. Harding, D.K. Niyogi & A.R. McIntosh. 2012. Improving the effectiveness of riparian management for aquatic invertebrates in a degraded agricultural landscape: stream size and land-use legacies. Journal of Applied Ecology 49:213-222.

Gucinski, H., M.J. Furniss, R.R. Ziemer & M.H. Brookes. 2001. Forest roads: A synthesis of scientific information. General Technical Report. PNWGTR-509. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR, USA.

Guillot, G. & F. Rousset. 2013. Dismantling the Mantel tests. Methods in Ecology and Evolution. 4:336-344.

Guillot, G., R. LeBlois, A Coulon & A.C. Frantz. 2009. Statistical methods in spatial genetics. Molecular Ecology 18:4734-4756.

Gusset, M., S.J. Ryan, M. Hofmeyr, G. van Dyk & others. 2008. Efforts going to the dogs? Evaluating attempts to re-introduce endangered wild dogs in South Africa. Journal of Applied Ecology 45:100-108.

Gustafson, E.J. 1998. Quantifying landscape spatial pattern: what is the state of the art? Ecosystems 1:143-156.

Gyllenberg, M. & I. Hanski. 1992. Single-species metapopulation dynamics: a structured model. Theoretical Population Biology 42:35-61.

Gyllenberg, M. & I. Hanski. 1997. Habitat deterioration, habitat destruction, and metapopulation persistence in a heterogeneous landscape. Theoretical Population Biology 52:198-215.

Haila, Y. 2002. A conceptual genealogy of fragmentation research: from island biogeography to landscape ecology. Ecological Applications 12:321-334.

Hall, L.S., P.R. Krausman & M.L. Morrison. 1997. The habitat concept and a plea for standard terminology. Wildlife Society Bulletin (1973-2006) 25:173-182.

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw & W.A. Stahel. 1986. Robust Statistics: The Approach Based on Influence Functions. Wiley Series in Probability and Statistics, John Wiley and Sons, New York.

Hansen, L.P. & T.P. Quinn. 1998. The marine phase of the Atlantic salmon (*Salmo salar*) life cycle, with comparisons to Pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences 55(Suppl. 1):104-118.

Hanski, I. 1999. Metapopulation dynamics. Nature 396:41-49.

Hanski, I. 2004. Metapopulation theory, its use and misuse. Basic and Applied Ecology 5:225-229.

Hanski, I. 2011. Habitat loss, the dynamics of biodiversity, and a perspective on conservation. Ambio 40:248-255.

Hanski, I. & M. Gilpin. 1991. Metapopulation dynamics: brief history and conceptual domain. Biological Journal of the Linnean Society 42:3-16.

Harrison, S. 1991. Local extinction in a metapopulation context: an empirical evaluation. Biological Journal of the Linnean Society 42:73-88

Hasselman, D.J., D. Ricard & P. Bentzen. 2013. Genetic diversity and differentiation in a wide ranging anadromous fish, American Shad (*Alosa sapidissima*), is correlated with latitude. Molecular Ecology 22:1558-1573.

Heggenes, J. 1990. Habitat utilization and preferences in juvenile Atlantic salmon (*Salmo salar*) in streams. Regulated Rivers: Research & Management 5:341-354.

Heritier, S., E. Cantoni, S. Copt & M.-P. Victoria-Feser. 2009. Robust methods in biostatistics. Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., UK.

Hilborn, R. 2001. Risk analysis for salmon spawning reference levels. In: E. Prevost & G. Chaput [eds.] Stock, recruitment and reference points, assessment and management of Atlantic salmon. French National Institute for Agricultural Research (INRA), Paris, France. pp. 177–193.

Hilborn, R. & C.J. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Reviews in Fish Biology and Fisheries, Kluwer Academic Publishers, Netherlands.

Hitt, N.P. & P.L. Angermeier. 2008. River-stream connectivity affects fish bioassessment performance. Environmental Management 42:132-150.

Hopken. M.W., M.R. Douglas & M.E. Douglas. 2013. Stream hierarchy defines riverscape genetics of a North American desert fish. Molecular Ecology 22:956-971.

Huang, S.-L., W.W. Budd, S.-L. Chan & Y.-C. Lin. 2007. Stream order, hierarchy, and energy convergence of land use. Ecological modeling 205:255-264.

Hutchings, J.A., R.A. Myers, V.B. Garcia, L.O Lucifora & A. Kuparinen. 2012a. Life history correlates of extinction risk and recovery potential. Ecological Applications 22:1061-1067.

Hutchings, J.A., S.H.M. Butchart, B. Collen, M.K. Schwartz & R.S. Waples. 2012b. Red flags: correlates of impaired species recovery. Trends in Ecology and Evolution 27:542-546.

Hyatt, K.D., D.J.McQueen, K.S. Shortreed & D.P. Rankin. Sockeye salmon (*Oncorhynchus nerka*) nursery lake fertilization: review and summary of results. Environmental reviews 12:133-162.

Jajo, N.K. 2005. A review of robust regression and diagnostic procedures in linear regression. Acta Mathematicae Applicatae Sinica 21:209-224.

Jeffres, C. & P. Moyle. 2012. When good fish make bad decisions: coho salmon in an ecological trap. North American Journal of Fisheries Management 32:87-92.

Jeffrey, J.D., C.T. Hasler, J.M. Chapman, S.J. Cooke & C.D. Suski. 2015. Linking landscape-scale disturbances to stress and condition of fish: implications for restoration and conservation. Integrative and Comparative Biology 22:1-13.

Jewett, K., D. Daugharty, H.H. Krause & P.A. Arp. 1995. Watershed responses to clear-cutting: Effects on soil solutions and stream water discharge in central New Brunswick. Canadian Journal of Soil Science 75:475-490.

Jia, Z., Y. Zhang, S. Chen & L. Shi. 2012. Genetic diversity and differentiation of masu salmon (*Oncorhynchus masou masou*) between and within cultured populations inferred from microsatellite DNA analysis. Zoological Research 33(E3-4): E33-E39.

Johnson, J.B. & K.S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19:101-108.

Johnson, M.D. 2007. Measuring habitat quality: a review. The Condor 109:489-504.

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

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jones, C.M. 2006. Estuarine and diadromous fish metapopulations. In: J.P. Kritzer & P.F. Sale [eds.]. Marine Metapopulations. Elsevier Academic Press, San Diego, CA.

Jonsson, N., B. Jonsson & L.P. Hansen. 1998. The relative role of density-dependent and density-independent survival in the life cycle of Atlantic salmon *Salmo salar*. Journal of Animal Ecology 67:751-762

Julien, H.P. & N.E. Bergeron. 2006. Effect of fine sediment infiltration during the incubation period on Atlantic salmon (*Salmo salar*) embryo survival. Hydrobiologia. 563:61-71.

Keefer, M.L. & C.C. Caudill. 2014. Homing and straying by anadromous salmonids: a review of mechanisms and rates. Reviews in Fish Biology and Fisheries 24:333-368.

Khrustaleva, A.M., M.T. Limborg, J.E. Seeb. 2014. Genetic variation among major sockeye salmon populations in Kamchatka Peninsula inferred from SNP and microsatellite DNA analyses. North Pacific Anadromous Fish Commission 1519. 17p.

Kiernan, J.D. & P.B. Moyle. 2012. Flows, droughts, and aliens: factors affecting the fish assemblage in a Sierra Nevada, California, stream. Ecological Applications 22:1146-1161.

Kiffney, P.M., C.M. Greene, J.E. Hall & J.R. Davies. 2006. Tributary streams create spatial discontinuities in habitat, biological productivity, and diversity in mainstem rivers. Canadian Journal of Fisheries and Aquatic Sciences 63:2518-2530.

Kim, M. & M. LaPointe. 2011. Regional variability in Atlantic salmon (*Salmo salar*) riverscapes: a simple landscape ecology model explaining the large variability in size of salmon runs across Gaspe watersheds, Canada. Ecology of Freshwater Fish 20:144-156.

King, T.L., S.T. Kalinowski, W.B. Schill, A.P. Spidle & B.A. Lubinski. 2001. Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. Molecular Ecology 10:807–821.

Kinnison, M.T. & N.G. Hairston Jr. 2007. Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence. Functional Ecology 21:444-454.

Koenig. W.D. 1999. Spatial autocorrelation of ecological phenomena. Trends in Ecology and Evolution 14: 22-26.

Koenig, W.D. 2002. Global patterns of environmental synchrony and the Moran effect. Ecography 25:283-288.

Koller, M. & M. Machler. 2013. Definitions of the psi-functions available in robustbase. robustbase : Basic robust statistics. R. package version 0.9-10. http://CRAN.R-project.org/package=robustbase

Korman, J., D.R. Marmorek, G.L Lacroix, P.G. Amiro, J.A. Ritter, W.D. Watt, R.E. Cutting & D.C.E. Robinson. 1994. Development and evaluation of a biological model to assess regional-scale effects of acidification on Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences 51:662-680.

Kubisch, A., R.D. Holt, H-J. Poethke & E.A. Fronhofer. 2014. Where am I and why? Synthesizing range biology and the eco-evolutionary dynamics of dispersal. Oikos 123:5-22.

Kuussaari, M., R. Bommarco, R.K. Heikkinen, A. Helm & others. 2009. Extinction debt: a challenge for biodiversity conservation. Trends in Ecology and Evolution 24:564-571.

Labbe, T.R. & K.D. Fausch. 2000. Dynamics of intermittent stream habitat regulate persistence of a threatened fish at multiple scales. Ecological Applications 10:1774-1791.

Lacroix, G.L. 1989. Production of juvenile Atlantic salmon (*Salmo salar*) in two acidic rivers of Nova Scotia. Canadian Journal of Fisheries and Aquatic Sciences 46:2003-2018.

Lacroix, G.L. & D.R. Townsend. 1987. Responses of juvenile Atlantic salmon (*Salmo salar*) to episodic increases in acidity of Nova Scotia rivers. Canadian Journal of Fisheries and Aquatic Sciences 44:1475-1484.

Lake, P.S. 2003. Ecological effects of perturbation by drought in flowing waters. Freshwater Biology 48:1161-1172.

Lake, P.S., N. Bond & P. Reich. 2007. Linking ecological theory with stream restoration. Freshwater Biology 52:597-615.

Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. The American Naturalist 142:911-927.

Lande, R., S. Engen & B.-E. Saether. 2003. Stochastic population dynamics in ecology and conservation. Oxford Series in Ecology and Evolution. Oxford University Press, NY.

Landguth, E.L., S.A. Cushman, M.K. Schwartz, K.S. McKelvey, M. Murphy & G. Luikart. 2010. Quantifying the lag time to detect barriers in landscape genetics. Molecular Ecology 19:4179–4191.

Langill, D.A. & P.J. Zamora. 2002. An audit of small culvert installations in Nova Scotia: habitat loss and habitat fragmentation. Canadian Technical Report of Fisheries and Aquatic Sciences, 2422.

Lawler, J.J., S.P. Campbell, A.D. Guerry, M.B. Kolozsvary,R.J. O'Connor & L.C. Seward. 2002. The scope and treatment of threats in endangered species recovery plans. Ecological Applications 12:663-667.

Lawson, P.W., E.A. Logerwell, N.J. Mantua, R.C. Francis & V.N. Agostini. 2004. Environmental factors influencing freshwater survival and smolt production in Pacific Northwest coho salmon (*Oncorhynchus kisutch*). Canadian Journal of Fisheries and Aquatic Sciences 61:360-373.

Leduc, A.O.H.C., E. Roh, C.J. Macnaughton, F. Benz, J. Rosenfeld & G.E. Brown. 2011. Ambient pH and the response to chemical alarm cues in juvenile Atlantic salmon: mechanisms of reduced behavioral responses. Transactions of the American Fisheries Society 139:117-128.

Legault, C.M. 2005. Population viability analysis of Atlantic salmon in Maine, USA. Transactions of the American Fisheries Society 134:549-562.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659-1673.

Legendre, P., F.-J. Lapointe & P. Casgrain. 1994. Modeling brain evolution from behavior: a permutational regression approach. Evolution 48:1487-1499.

Legendre, P. & M.-J. Fortin. 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. Molecular Ecology Resources 10:831–844.

Legendre, P., M.-J. Fortin & D. Borcard. 2015. Should the Mantel test be used in spatial analysis? Method in Ecology and Evolution 6:1239-1247.

Le Pichon, C., E. Tales, J. Belliard & C.E. Torgersen. 2016. Spatially intensive sampling by electrofishing for assessing longitudinal discontinuities in fish distribution in a headwater stream. Fisheries Research 185:90-101.

LeRoy Poff, N., B.P. Bledsoe & C.O. Cuhaciyan. 2006. Hydrologic variation with land use across the contiguous United States: geomorphic and ecological consequences for stream ecosystems. Geomorphology 79:264-285.

LeRoy Poff, N., B.D. Richter, A.H. Arthington, S.E. Bunn & others. 2010. The ecological limits of hydrologic alteration (ELOHA): a new framework for developing regional environmental flow standards. Freshwater Biology 55:147-170.

Levin, S.A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.

Lichstein, J.W. 2007. Multiple regression on distance matrices: a multivariate spatial analysis tool. Plant Ecology 188:117-131.

Lin, Y.-T.K. & G.O. Batzli. 2001. The influence of habitat quality on dispersal, demography and population dynamics of voles. Ecological Monographs 71:245-275.

Linden, A. & S. Mantyniemi. 2011. Using the negative binomial distribution to model overdispersion in ecological count data. Ecology 92:1414-1421.

Lisle, T.E. 1989. Sediment transport and resulting deposition in spawning gravels, north coastal California. Water Resources Research 25:1303-1319.

Lurling, M. & M. Scheffer. 2007. Info-disruption: pollution and the transfer of chemical information between organisms. Trends in Ecology and Evolution 22: 374-379.

Magoulick, D.D. & R.M. Kobza. 2003. The role of refugia for fishes during drought: a review and synthesis. Freshwater Biology 48:1186-1198.

Maronna, R., D. Martin & V. Yohai. 2006. Robust statistics, theory and methods. Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., West Sussex.

Marsh, D.M. & P.C. Trenham. 2001. Metapopulation dynamics and amphibian conservation. Conservation Biology 15:40-49.

Mathews, R. & B. Richter. 2007. Application of the Indicators of Hydrologic Alternation software in environmental flow-setting. Journal of the American Water Resources Association 43:1-14.

Matthysen, E. 2005. Density-dependent dispersal in birds and mammals. Ecography 28:416-416.

McCluney, K.E., N. LeRoy Poff, M.A. Palmer, J.H. Thorpe & others. 2014. Riverine macrosystems ecology: sensitivity, resistance, and resilience of whole river basins with human alterations. Frontiers in Ecology and the Environment 12:48-58.

McCormick, S.D., L.P. Hansen, T.P. Quinn & R.L. Saunders. 1998. Movement, migration and smolting of Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences 55(Suppl. 1):77-92.

McCullagh, P. & J.A. Nelder. 1989. Generalized Linear Models: second edition. Monographs on Statistics and Applied Probability, 37, Chapman & Hall, London.

McInerny, M.C. & T.K. Cross. 2000. Effects of sampling time, intraspecific density, and environmental variables on electrofishing catch per effort of largemouth bass in Minnesota Lakes. North American Journal of Fisheries Management 20:328-336.

Melbourne, B.A. & A. Hastings. 2008. Extinction risk depends strongly on factors contributing to stochasticity. Nature 454:100-103.

Melvin, G.D., M.J. Dadswell & J.D. Martin. 1985. Fidelity of American shad, *Alosa sapidissima* (Clupeidae), to its river of previous spawning. Canadian Journal of Fisheries and Aquatic Sciences 43:640-646.

Milly, P.C.D., K.A. Dunne & A.V. Vecchia. 2005. Global pattern of trends in streamflow and water availability in a changing climate. Nature 438:347-350.

Moilanen, A. 2004. SPOMSIM: software for stochastic patch occupancy models of metapopulation dynamics. Ecological Modelling 179:533-550.

Moore, J.W., M. McClure, L.A. Rogers & D.E. Schindler. 2010. Synchronization and portfolio performance of threatened salmon. Conservation Letters 3:340-348.

Morales, J.M. 2011. The metapopulation that was not? Animal Conservation 14:227-228.

Morris, W.F. & D.F. Doak. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer, Sunderland, Massachusetts, USA.

Mortelliti, A., G. Amori & L. Boitani. 2010. The role of habitat quality in fragmented landscapes: a conceptual overview and prospectus for future research. Oecologia 163:535-547.

Mueter, F.J., R.M. Peterman & B.J. Pyper. 2002. Opposite effects of ocean temperature on survival rates of 120 stocks of Pacific salmon (*Oncorhynchus* spp.) in northern and southern areas. Canadian Journal of Fisheries and Aquatic Sciences 59:456-463.

Myers, R.A. 1998. When do environment–recruitment correlations work? Reviews in Fish Biology and Fisheries 8:285-305.

Myers, R.A., N.J. Barrowman, J.A. Hutchings & A.A. Rosenberg. 1995. Population dynamics of exploited fish stocks at low population levels. Science 269:1106-1108.

Myers, R.A., G. Mertz & J. Bridson. 1997. Spatial scales of interannual recruitment variations of marine, anadromous, and freshwater fish. Canadian Journal of Fisheries and Aquatic Sciences 54:1400-1470.

Myers, R.A., K.G. Bowen & N.J. Barrowman. 1999. Maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences 56: 2404-2419.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Nislow, K.H. 2005. Forest change and stream fish habitat: lessons from ‘Olde’ and New England. Journal of Fish Biology 67(Suppl. B):186-204.

Nislow, K.H., A.J. Sepulveda & C.L. Folt. 2004. Mechanistic linkage of hydrologic regime to summer growth of age-0 Atlantic salmon. Transactions of the American Fisheries Society 133:79-88.

Nislow, K.H., M. Hudy, B.H. Letcher & E.P. Smith. 2011. Variation in local abundance and species richness of stream fishes in relation to dispersal barriers: implications for management and conservation. Freshwater Biology 56:2135-2144.

Norris, K. 2004. Managing threatened species – the ecological toolbox, evolutionary theory and the declining-population paradigm. Journal of Applied Ecology 41:413-426.

O’Reilly, P., S. Rafferty & J. Gibson. 2012. Within- and among-population genetic variation in the Southern Upland Designatable Unit of Maritime Atlantic Salmon (*Salmo salar* L.). DFO Canadian Stock Assessment Secretariat Research Document, 2012/077.

Oksanen, J., F. Guillaume Blanchet, R. Kindt, P. Legendre & others. 2013. vegan: Community Ecology Package. R package version 2.0-10. http://CRAN.R-project.org/package=vegan.

Olden, J.D. & N. LeRoy Poff. 2003. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. River Research and Applications 19:101-121.

Olden, J.D., M.J. Kennard, F. Leprieur, P.A. Tedesco, K.O. Winemiller & E. Garcia-Berthou. 2010. Conservation biogeography of freshwater fishes: recent progress and future challenges. Diversity and Distributions 16:496-513.

Oor, H.A. 2009. Fitness and its role in evolutionary genetics. Nature Reviews: Genetics 10:531-539.

Opdam, P. 1991. Metapopulation theory and habitat fragmentation: a review of Holarctic breeding bird studies. Landscape Ecology 5:93-106.

Pace, M.L., J.J. Cole, S.R. Carpenter & J.F. Kitchell. 1999. Trophic cascades revealed in diverse ecosystems. Trends in Ecology and Evolution 14:483-488.

Palkovacs, E.P., D.J. Hasselman, E.E. Argo, S.R. Gephard & others. 2013. Combining genetic and demographic information to prioritize conservation efforts for anadromous alewife and blueback herring. Evolutionary Applications 7:212-226.

Palmer, M.A. 2009. Reforming watershed restoration: science in need of application and applications in need of science. Estuaries and Coasts 32:1-17.

Palstra F.P., M.F. O’Connell & D.E. Ruzzante. 2007. Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. Molecular Ecology 16:4504-4522.

Paul, M.J. & J.L. Meyer. 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32:333-365.

Pearman, P.B., A. Guisan, O. Broennimann & C.F. Randin. 2008. Niche dynamics in space and time. Trends in Ecology and Evolution 23:149-158.

Pearson, K. 1896. Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. Proceedings of the Royal Society of London 60:489-498.

Pellet, J., E. Fleishman, D.S. Dobkin, A. Gander & D.D. Murphy. 2007. An empirical evaluation of the area and isolation paradigm of metapopulation dynamics. Biological Conservation 136:483-495.

Perry, G.L.W. & N.J. Enright. 2007. Contrasting outcomes of spatially implicit and spatially explicit models of vegetation dynamics in a forest-shrubland mosaic. Ecological Modelling 207:327-338.

Pfluger, F.J. & N. Balkenhol. 2014. A plea for simultaneously considering matrix quality and local environmental conditions when analyzing landscape impacts on effective dispersal. Molecular Ecology 23:2146-2156.

Plummer, M., N. Best, K. Cowles & K. Vines. 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC, R News 6:7-11.

Poethke, H.J., W.W. Weisser & T. Hovestadt. 2010. Predator-induced dispersal and the evolution of conditional dispersal in correlated environments. The American Naturalist 175:577-586.

Pompanon, F., A. Bonin, E. Bellemain & P. Taberlet. 2005. Genotyping errors: causes, consequences and solutions. Nature Reviews 6:847-859.

Poplar-Jeffers, I.O., J.T. Petty, J.T. Anderson, S.J. Kite, M.P. Stragler & R.H. Fortney. 2009. Culvert replacement and stream habitat restoration: Implications from brook trout management in an Appalachian watershed, USA. Restoration Ecology 17:404-413.

Prairie, Y.T. & D.F. Bird. 1989. Some misconceptions about the spurious correlation problem in the ecological literature. Oecologia 81:285-288.

Pyper, B.J., F.J. Mueter & R.M Peterman. 2005. Across-species comparisons of spatial scales of environmental effects on survival rates of Northeast Pacific salmon. Transactions of the American Fisheries Society 134:86-104.

Raymond, M. & F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249.

Richards, S.A. 2008. Dealing with overdispersed count data in applied ecology. Journal of Applied Ecology 45:218-227.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada, 191.

Rieman, B.E. & J.B. Dunham. 2000. Metapopulations and salmonids: a synthesis of life history patterns and empirical observations. Ecology of Freshwater Fish 9:51-64.

Ritter, J.A. 1989. Marine migration and natural mortality of North American Atlantic salmon (*Salmo salar* L.). Canadian Manuscript Report of Fisheries and Aquatic Sciences, 2041.

Robertson, B.A. & R.L. Hutto. 2006. A framework for understanding ecological traps and an evaluation of existing evidence. Ecology 87:1075-1085.

Robertson, B.A., J.S. Rehage & A. Sih. 2013. Ecological novelty and the emergence of evolutionary traps. Trends in Ecology and Evolution 28:552-560.

Rodriguez-Iturbe, I., R. Muneepeerakul, E. Bertuzzo, S.A. Levin & A. Rinaldo. 2009. River networks as ecological corridors: a complex systems perspective for integrating hydrologic, geomorphologic, and ecologic dynamics. Water Resources Research 45:W01413. doi:10.1029/2008WR007124.

Rolls, R.J., C. Leigh & F. Sheldon. 2012. Mechanistic effects of low-flow hydrology on riverine ecosystems: ecological principles and consequences of alteration. Freshwater Science 31:1163-1186.

Roni, P., T.J. Beechie, R.E. Bilby, F.E. Leonetti, M.M. Pollock & G.R. Pess. 2002. A review of stream restoration techniques and a hierarchical strategy for prioritizing restoration in Pacific Northwest watersheds. North American Journal of Fisheries Management 22:1-20.

Rose, K.A. 2000. Why are quantitative relationships between environmental quality and fish populations so elusive? Ecological Applications 10:367-385.

Rose, K.A., J.H. Cowan Jr., K.O. Winemiller, R.A. Myers & R. Hilborn. 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. Fish and Fisheries 2:293-327

Rousseeuw, P., C. Croux, V. Todorov, A. Ruckstuhl, M. Salibian-Barrera, T. Verbeke, M. Koller & M. Maechler. 2013. robustbase: Basic robust statistics. R. package version 0.9-10. http://CRAN.R-project.org/package=robustbase

Rousset, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106.

Ruckelshaus, M.H., P. Levin, J.B. Johnson & P.M. Kareiva. 2002. The Pacific salmon wars: what science brings to the challenge of recovering species. Annual Review of Ecology and Systematics 33:665-706.

Ruetz, C.R., J.C. Trexler, F. Jordan, W.F. Loftus & S.A. Perry. 2005. Population dynamics of wetland fishes: spatio-temporal patterns synchronized by hydrological disturbance? Journal of Animal Ecology 74:322-332.

Sabo, J.L., J.C. Finlay, T. Kennedy & D.M. Post. 2010. The role of discharge variation in scaling of drainage area and food chain length in rivers. Science 330:965–967.

Sawyer, S.C., C.W. Epps CW & J.S. Brashares. 2011. Placing linkages among fragmented habitats: do least-cost models reflect how animals use landscapes? Journal of Applied Ecology 48:668–678.

Schaub, M. & F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology 152(Suppl. 1):S227-S237.

Scheuerell, M.D., R. Hilborn, M.H. Ruckelshaus, K K. Bartz, K.M. Lagueux, A.D. Haas & K. Rawson. 2006. The Shiraz model: a tool for incorporating anthropogenic effects and fish-habitat relationships in conservation planning. Canadian Journal of Fisheries and Aquatic Sciences 63:1596-1607.

Schindler, D.E., R. Hilborn, B. Chasco, C.P. Boatright, T.P. Quinn, L.A. Rogers & M.S. Webster. 2010. Population diversity and the portfolio effect in an exploited species. Nature 465:609-613.

Schlaepfer, M.A., M.C. Runge & P.W. Sherman. 2002. Ecological and evolutionary traps. Trends in Ecology and Evolution 17:474-480.

Schlosser, I.J. 1991. Stream fish ecology: a landscape perspective. BioScience 41:704-712.

Schneider, D.C. 2001. The rise of the concept of scale in ecology. BioScience 51:545-553.

Scholz, N.L., N.K. Truelove, B.L. French, B.A. Berejikian, T.P. Quinn, E. Casillas & T.K. Collier. 2000. Diazinon disrupts antipredator and homing behaviors in Chinook salmon (*Oncorhynchus tshawytscha*). Canadian Journal of Fisheries and Aquatic Sciences 57:1911–1918.

Schoville, S.D., A. Bonin, O. Francois, S. Lobeaux, C. Melodelima & S. Manel. 2012. Adaptive genetic variation on the landscape: methods and cases. Annual Review of Ecology, Evolution, and Systematics 43:23-43.

Schtickzelle, N. & T.P. Quinn. 2007. A metapopulation perspective for salmon and other anadromous fish. Fish and Fisheries 8:297-314.

Schurr, F.M., J. Pagel, J. Sarmento Cabral, J. Groeneveld, & others. 2012. How to understand species' niches and range dynamics: a demographic research agenda for biogeography. Journal of Biogeography 39:2146-2162.

Schwarz, C.J. & G.A. Seber. 1999. Estimating animal abundance: review III. Statistical Science 14:427-456.

Sexton, J.P., S.B. Hangartner & A.A. Hoffman. 2013. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution 68:1-15.

Shenk, T.M., G.C. White & K.P. Burnham. 1998. Sampling-variance effects on detecting density dependence from temporal trends in natural populations. Ecological Monographs 68:445-460.

Shirk, A., D.O. Wallin, S.A. Cushman, R.C. Rice & C. Warheit. 2010. Inferring landscape effects on gene flow: a new multi-scale model selection framework. Molecular Ecology 19:3603–1619.

Sih, A., M.C.O. Ferrari & D.J. Harris. 2011. Evolution and behavioural responses to human-induced rapid environmental change. Evolutionary Applications 4:367-387.

Simberloff, D. 1998. Small and Declining Populations. In: W.J. Sutherland [ed.] Conservation Science and Action. Blackwell Publishing Ltd., Oxford, UK

Singh, N.D., J.D. Jensen, A.G. Clark & C.F. Aquadro. 2013. Inferences of demography and selection in an African population of *Drosophila melanogaster*. Genetics 193:215-228.

Skaug, H., D. Fournier, B. Bolker, A. Magnusson & A. Nielsen. 2014. Generalized Linear Mixed Models using AD Model Builder. R package version 0.8.0.

Slater, H. & E. Michael. 2012. Predicting the current and future potential distributions of lymphatic filariasis in Africa using maximum entropy ecological niche modeling. Plos ONE 7: e32202. doi:10.1371/journal.pone.0032202.

Slatkin, M. 1993. Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47:264–279.

Smedbol, R.K., A. McPherson, M.M. Hansen & E. Kenchington. 2002. Myths and moderation in marine 'metapopulations'? Fish and Fisheries 3:20-35.

Smokorowski, K.E. & T.C. Pratt. 2007. Effect of a change in physical structure and cover on fish and fish habitat in freshwater ecosystems – a review and meta-analysis. Environmental Reviews 15:15-41.

Smouse, P.E., J.C. Long & R.R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology 35:627–632.

Sommer, S. & R.M. Huggins. 1996. Variables selection using the Wald Test and a robust Cp. Applied Statistics 45:15-29.

Soulsby, C., A.F. Youngson, H.J. Moir & I.A. Malcolm. 2001. Fine sediment influence on salmonid spawning habitat in a lowland agricultural stream: a preliminary assessment. Science of the Total Environment 265:295-307.

Spear, S.F., N. Balkenhol, M.-J. Fortin & B.H. McRae. 2010. Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Molecular Ecology 19:3576-3591.

Speas, D.W., C.J. Walters, D.L. Ward & R. Scott Rogers. 2004. Effects of intraspecific density and environmental variables on electrofishing catchability of brown and rainbow trout in the Colorado River. North American Journal of Fisheries Management 24:586-596.

Stanley, E.H., S.M. Powers & N.R. Lottig. 2010. The evolving legacy of disturbance in stream ecology: concepts, contributions, and coming challenges. Journal of the North American Benthological Society 29:67-83.

Stoll, S., P. Breyer, J.D. Tonkin, D. Fruh & P. Haase. 2016. Scale-dependent effects of river habitat quality on benthic invertebrate communities – Implications for stream restoration practice. Science of the Total Environment 553:495-503.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. Transactions, American Geophysical Union 38:913-920.

Su, Z., M.D. Adkison & B.W. Van Allen. 2001. A hierarchical Bayesian model for estimating historical salmon escapement and escapement timing. Canadian Journal of Fisheries and Aquatic Sciences 58:1648-1662.

Sutherland, W.J. & K. Norris. 2002. Behavioural models of population growth rates: implications for conservation and prediction. Philosophical Transactions of the Royal Society of London B 357:1273-1284.

Sutherland, W.J., R.P Freckleton, H.C.J. Godfray, S.R. Beissinger & others. 2013. Identification of 100 fundamental ecological questions. Journal of Ecology 101:58-67.

Tanentzap, A.J., E.J. Szkokan-Emilson, B.W. Kielstra, M.T. Arts, N.D. Yan & J.M. Gunn. 2014. Forest fuel fish growth in freshwater deltas. Nature Communications. 5:4077. doi: 10.1038/ncomms5077.

Taylor, E.B. 1991. A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture 98:185-207.

Thornton, D.H., L.C. Branch & M.E. Sunquist. 2011. The influence of landscape, patch, and within-patch factors on species presence and abundance: a review of focal patch studies. Landscape Ecology 26:7-18.

Thorstad, E.B., F. Whoriskey, A.H. Rikardsen & K. Aarestrup. 2011. Aquatic nomads: the life and migrations of the Atlantic salmon. In: Aas O, Einum S, Klemetsen A, Skurdal J [eds.] Atlantic Salmon Ecology. Blackwell, London. pp. 1-32.

Tilman, D., R.M. May, C.L. Lehman & M.A. Nowak. 1994. Habitat destruction and the extinction debt. Nature 371:65-66.

Torgersen, C.E., D.M. Price, H.W. Li & B.A. McIntosh 1999. Multiscale thermal refuge and stream habitat associations of chinook salmon in northeastern Oregon. Ecological Applications 9:301-319.

Torgersen, C.E., C.V. Baxter, J.L. Ebersole & R.E. Gresswell. 2012. Incorporating spatial context into the analysis of salmonid-habitat relationships. *In* M. Church, P.M. Biron, and A. Roy [eds.]. Gravel-bed Rivers: Processes, Tools, Environments. John Wiley & Sons Limited, Chichester, UK . pp. 216-224.

Traill, L.W., C.J.A. Bradshaw & B.W. Brook. 2007. Minimum viable population size: a meta-analysis of 30 years of published estimates. Biological Conservation 139:159-166.

Trombulak, S.C. & C.A. Frissell. 2000. Review of ecological effects of roads on terrestrial and aquatic communities. Conservation Biology 14:18-30.

Trzcinski, M.K., A.J.F. Gibson, P.G. Amiro & R.G. Randall. 2004. Inner Bay of Fundy Atlantic salmon (*Salmo salar*) critical habitat case study. DFO Canadian Science Advisory Secretariat Research Document 2004/114.

Turner, M.G. 1989. Landscape ecology: the effect of pattern on process. Annual Review of Ecology and Systematics 20:171-197.

Ugedal, O. & A.G. Finstad. 2011. Landscape and land use effects on Atlantic salmon. In: O. Aas, S. Einum, A, Klemetsen, and J. Skurdal [eds.] Atlantic Salmon Ecology. Blackwell Publishing Limited, Oxford, UK. pp. 333-349.

van Oosterhout, C., W.F. Hutchinson, D.P.M. Wills & P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

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

Venter, O., N.N. Brodeur, L. Nemiroff, B. Belland, I.J. Dolinsek & J.W.A. Grant. 2006. Threats to endangered species in Canada. BioScience 56:903-910.

Ver Hoef, J.M. & P.L. Boveng. 2007. Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data? Ecology 88:2766-2772.

Vuilleumier, S., C. Wilcox, B.J. Cairns & H.P. Possingham. 2007. How patch configuration affects the impact of disturbances on metapopulation persistence. Theoretical Population Biology 72:77-85.

Wagenmakers, E.-J. & S. Farrell. 2004. AIC model selection using Akaike weights. Psychonomic Bulletin and Review 11:192-196.

Wagner, H.H. & M.-J. Fortin. 2016. Basics of Spatial Data Analysis: Linking Landscape and Genetic Data for Landscape Genetic Studies. In: N Balkenhol, SA Cushman, AT Storfer, LP Waits [eds.] Landscape genetics: Concepts, methods, applications. John Wiley & Sons Ltd, UK.

Wang, I.J., R.E. Glor & J.B. Losos. 2013. Quantifying the roles of ecology and geography in spatial genetic divergence. Ecology Letters 16:175-182

Waples, R.S. & O.R. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419-1439.

Waples, R.S., R.W. Aable, M.D. Scheuerell & B.L. Sanderson. 2008. Evolutionary responses by native species to major anthropogenic changes to their ecosystems: Pacific salmon in the Columbia River hydropower system. Molecular Ecology 17:84–96.

Waples, R.S., T. Beechie & G.R. Pess. 2009. Evolutionary history, habitat disturbance regimes, and anthropogenic changes: What do these mean for resilience of Pacific salmon populations? Ecology and Society 14:3. http://www.ecologyandsociety.org/vol14/iss1/art3/

Wasserman, T.N., S.A. Cushman, M.K. Schwartz & D.O. Wallin. 2010. Spatial scaling and multi-model inference in landscape genetics: *Martes americana* in northern Idaho. Landscape Ecology 25:1601–1612.

Wathen, G., S.M. Coghlan, J. Zydlewski & J.G. Trial. 2011. Habitat selection and overlap of Atlantic salmon and smallmouth bass juveniles in nursery streams. Transactions of the American Fisheries Society 140:1145-1157.

Watt, W.D. 1987. A summary of the impact of acid rain on Atlantic salmon (*Salmo salar*) in Canada. Water, Air, and Soil Pollution 35:27-35.

Watt, W.D., C.D. Scott & W.J. White. 1983. Evidence of acidification of some Nova Scotia rivers and its impact on Atlantic salmon, *Salmo salar*. Canadian Journal of Fisheries and Aquatic Sciences 40:462-473.

Webster, M.S. 2003. Temporal density dependence and population regulation in a marine fish. Ecology 84:623-628.

Westley, P.A.H., A.H. Dittman, E.J. Ward & T.P. Quinn. 2015. Signals of climate, conspecific density, and watershed features in patterns of homing and dispersal by Pacific salmon. Ecology. doi/pdf/10.1890/14-1630.1.

Whittaker, R.J., M.B. Araujo, P. Jepson, R.J. Ladle, J.E.M. Watson & K.J. Willis. 2005. Conservation Biogeography: assessment and prospect. Diversity and Distributions 11:3-23.

Wiens, J.A. 2002. Riverine landscapes: taking landscape ecology into the water. Freshwater Biology 47:501-515.

Wilcox, C., B.J. Cairns & H. Possingham. 2006. The role of habitat disturbance and recovery in metapopulation persistence. Ecology 87:855-863.

Winemiller, K.O. 2005. Life history strategies, population regulation, and implications for fisheries management. Canadian Journal of Fisheries and Aquatic Sciences 62:872-885.

Woodward, G., M.O. Gessner, P.S. Giller, V. Gulis & others. 2012. Continental-scale effects of nutrient pollution on stream ecosystem functioning. Science 336:1438-1440.

Wu, J. & O.L. Loucks. 1995. From balance of nature to hierarchical patch dynamics: a paradigm shift in ecology. The Quarterly Review of Biology 70:439-466.

Yeakel, J.D., J.W. Moore, P.R. Guimaraes Jr. & M.A.M. de-Aguiar. 2014. Synchronization and stability in river metapopulation networks. Ecology Letters 17:273-283.

Yohai, V.J. 1987. High breakdown-point and high efficiency robust estimates for regression. The Annals of Statistics 15:642-656.

Young, K.V., J.F. Dower & P. Pepin. 2009. A hierarchical analysis of the spatial distribution of larval fish prey. Journal of Plankton Research 31:687-700.

Zeller, K.A., K. McGarigal & A.R. Whiteley. 2012. Estimating landscape resistance to movement: a review. Landscape Ecology 27:777–797.

Zuur, A.F., E.N. Ieno, N.J. Walker, A.A. Saveliev & G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Statistics for Biology and Health, Springer Science and Business Media, LLC, New York.

Zuur, A.F., E.N. Ieno & C.S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. Methods in Ecology Evolution 1:3-14.

Appendices

Appendix 2-1: Details on the simulation model

Projection Model

The core of the simulation was a life history-based population viability analysis (PVA; e.g., Morris & Doak 2002), which modeled the dynamics of 50 populations over 100 years concurrently, assuming total isolation (i.e., no immigration or emigration) or metapopulation structure (i.e., populations were partially connected by straying). For each population, the initial number of one-sea-winter (1SW) and two-sea-winter (2SW) adults (Table A2-1-1) was calculated from smolt production at equilibrium, decremented by mortality in the marine environment prior to spawning (equations 1 and 2, Table A2-1-2). Using equilibrium smolt production (equations 15 and 16, Table A2-1-2) ensured that isolated populations as well as metapopulations were initialized at the same starting population sizes. The contribution from multi-sea-winter (MSW) repeat-spawning adults becomes the number of first-time spawners decremented by mortality that occurs between spawning events (equations 3 and 4, Table A2-1-2). For a metapopulation, the number of animals emigrating from each population in each year is a function of the total number of adults as well as the survival rate of strays (equation 6, Table A2-1-2). Immigrants (in the same year) are distributed among non-natal populations according to an assumed stray matrix (equation 7, Table A2-1-2). This is the spatial component of the model in that immigration probabilities were a function of the relative geographical position of populations (i.e., an isolation by distance (IBD) model; Guillot et al. 2009). The number of virgin 1SW and 2SW spawners in *t+2* and *t+3* years, respectively, becomes a density-dependent function of the total number of returns in year *t* (representing freshwater production) and the probability of maturing as 1SW as well as survival in the marine environment in years *t+1*, or *t+1* and *t+2*, respectively (equations 9 and 10, Table A2-1-2). In order to prevent a situation in which unrealistically small population sizes persist in the projection (e.g., spawner abundance of less than 1), abundance in year *t* was assessed relative to a quasi-extinction threshold (q = 20) and future production in years *t+2* (1SW) and *t+3* (2SW) was set to zero if population size was lower than the threshold. To account for individuals that spawn multiple times from a given cohort, the number of MSW repeats was calculated analogously to the initialization (equation 11, Table A2-1-2).

Model parameterization

Life history parameter values for input into the PVA were chosen to realistically approximate the population dynamics of a diadromous fish (Table A2-1-1). For flexibility, the model was set up so that all parameters could be population-specific; however, most of them were kept constant in order focus the analysis and limit the total number of scenarios considered. One extension to these analyses could be to vary stray rates among populations, given that significant inter-population variation in dispersal has been observed for captively-reared individuals (Westley et al. 2015). I assumed that 10% of a population strayed (emigrated) in a year, given that empirical estimates of stray rates from mark-recapture experiments are in the range of 10% for many salmonid species, including Atlantic salmon (Keefer and Caudill 2014). Estimates from other diadromous species can to be lower (e.g., Melvin et al. 1985 for American Shad) or higher (see discussion in Gahagan et al. 2012 for river herring). However, many of the higher estimates rely on discriminant analysis of microchemical or genetic signatures rather than actual positive identification of individual fish. Note that for the simulations of isolated populations, the stray rate was zero.

Each habitat unit within a watershed was assumed to be able to produce 20 juveniles (R_p^{asy} = 20) and the maximum rate at which spawners produced smolts (α) was 10. Both of these values are realistic relative to observed freshwater production of juvenile Atlantic salmon in the Southern Upland (Gibson & Bowlby 2013). Areas and distances were all parameterized as relative values (e.g., 40 as compared to 200 habitat units) rather than geographical measurements (e.g., 40 km^2 as compared to 200 km^2). Half of the smolts produced matured as 1SW and the starting values for the annual at-sea mortality rates of immature and adult fish were 0.5. If more fish were assumed to mature as 1SW, the overall productivity of the populations would increase (because fewer individuals would forgo reproduction after one year in the marine environment). It is worth pointing out that the productivity scenarios considered in this chapter could be achieved multiple ways (i.e., by changing the input values for a combination of parameters leading to the same estimated maximum lifetime reproductive rate (MLR)). It is also worth noting that if the simulations were to be parameterized for a specific set of rivers, it would be simple to change any of the starting values to reflect measured conditions. Simulations used 500 iterations of each scenario and were projected over 100 years.

I purposefully did not explicitly define the units for some of the parameters in the simulation model. This is because the results would not depend on the spatial or temporal scale of specific parameters, provided the combination of parameter values resulted in the same MLR productivity scenario. As an example related to time, I could have defined t = 2 years rather than t = 1. This would double the temporal scale of freshwater residency and projections comparable to these results would have been made over 200 rather than 100 years. However, the abundance in specific populations and the relative difference between the metapopulation and isolated population trajectories would remain the same, leading to the same conclusions when different levels of productivity were compared. As an example related to space, asymptotic juvenile density (R_p^{asy}) is in numbers per habitat unit (as detailed in Table A2-1-1), where watershed size (h) or the number of habitat units is not explicitly defined. It is possible to calculate juvenile abundance at carrying capacity in a river as $R_p^{asy}*h$. If this value was 10,000; R_p^{asy} could be 20 and h could be 500. In a real river that has 10,000 km^2 of habitat area, it would make the units for R_p^{asy} smolts/km^2 and h km^2. Alternately, if our real river had 10,000 m^2 of habitat, nothing would change in the parameterization but the units. Without making the parameterization of the model explicit to a real-world group of rivers, the units are not invariant.

Incorporating environmental variability

For greater biological realism in the population dynamics models, I incorporated autocorrelated variability into annual survival rates, following the general approach of Hilborn (2001). I assumed a log-normal distribution for the deviates around freshwater production ($dev_{t,p}^{fw}$) and a uniform distribution for marine mortality (dev_t^{mar}) (equations 13 and 14, Table A2-1-2), constrained to be between zero and one. Instantaneous mortality in year t depended on the degree of autocorrelation (d) as well as the standard deviation of the residuals of the parameter value (σ_w). Random deviation around the mean mortality rates at sea (M_t^{Sea} or M_t^{Aqult}) came from a set of random draws from a uniform distribution. Populations were assumed to mix while at sea, which is the same as assuming that the deviation in mean mortality rates was the same for all populations in a given year. In contrast, individual watersheds would be expected to have weakly synchronous dynamics because in-river conditions would be influenced by large-scale climactic patterns as well as localized processes (Ruetz et al. 2005). Therefore, the deviates describing fresh water varied over time as described above, yet individual populations differed from each yearly value according to a normal distribution with a sigma of 0.75 (equation 13, Table A2-1-2). Therefore, the mean correlation among population abundances over time (i.e., strength of population synchrony) in each scenario was set at 0.25, a value which matches observed population-level correlations in an exploited Pacific salmonid species (Schindler et al. 2010).

Table A2-1-1. Description of the parameters describing population vital rates (and starting values), indexing subscripts (with range), spatial model (stray.matrix), environmental variation (deviates), and population sizes (for smolts and adults) used in the metapopulation simulation. The manner in which the parameter is treated in the equations is given as the parameter type. The starting values describe the High productivity scenario and were progressively decreased in the remaining scenarios (see methods and Table 2-1). The spatial model considered was an Isolation by Distance (IBD) model in which the probability of an individual entering a specific watershed was a function of the distance between watersheds.

Model parameter	Description	Parameter type	Starting value
a	Previous spawnings	Indexing variable	0-3
t	Time	Indexing variable	1-100
p	Population	Indexing variable	1-50
q	Quasi-extinction threshold	Constant	20
$N_{a,t,p}$	Abundance of adults for a given previous spawning history in a specific year and population	Population-specific; Time-varying;	Equilibrium (E_p^{Adult})
α_p	Maximum survival from spawner to smolt	Population-specific	10
R_p^{asy}	Asymptotic juvenile density (# per habitat unit)	Population-specific	20
h_p	Watershed size (# habitat units; dimensionless)	Population-specific	200 (large) 40 (small)
M_t^{Sea}	Mortality rate of immature salmon at sea	Time-varying	0.5
m_1	Probability of maturing as 1SW	Constant	0.5
M_p^{Adult}	Mortality rate of repeat spawners	Population-specific	0.5

Model parameter	Description	Parameter type	Starting value
s_p	Stray rate (proportion)	Population-specific	0.1
stray.matrix	Spatial model of immigration	Hypothesized	IBD
d_{fw}	Temporal autocorrelation in freshwater	Constant	0.6
d_{mar}	Temporal autocorrelation at sea	Constant	0.6
σ_{fw}	Variation in freshwater survival	Constant	0.3
σ_{mar}	Variation in at-sea survival	Constant	0.5
$d_{t,p}^{fw}$	Deviate for freshwater environment	Population-specific; Time-varying	Random
dev_t^{mar}	Deviate for marine environment	Time-varying	Random
E_p^{Smolt}	Smolt abundance at equilibrium	Constant	Calculated
E_p^{Adult}	Adult abundance at equilibrium	Constant	Calculated

Table A2-1-2. Equations used in the metapopulation simulation, describing the model initialization (equations 1-4), straying among populations in a given year (equations 5-9), the projection of abundance in future years (equations 10-12), the manner in which random variability was added to freshwater and marine survival (equations 13 and 14), and equilibrium calculations for smolts and adults (equations 15-17). Parameter values, indexing subscripts and definitions are in Table A2-1-1.

Description		Equation
Model initialization		
Abundance in year 1 of 1SW spawners	1	$N_{0,1,p} = E_p^{Smolt} m_1 \left(1 - M^{Sea}\right)$
Abundance in year 2 of 2SW spawners	2	$N_{0,2,p} = E_p^{Smolt} \left(1 - m_1\right)\left(1 - M^{Sea}\right)^2$
Repeat spawners in year 1	3	$N_{a+1,1,p} = N_{a,1,p} \left(1 - M^{Adult}\right)$
Repeat spawners in year 2	4	$N_{a+1,2,p} = N_{a,2,p} \left(1 - M^{Adult}\right)$
Straying among populations		
Returns to natal river	5	$N_{a,t,p}^{Homing} = N_{a,t,p} \left(1 - s_p\right)$
Number of emigrants (strays)	6	$N_{a,t,p}^{Em} = N_{a,t,p} \left(s_p\right)$
Total number of immigrants (combined over *a*)	7	$N_{t,p}^{Imm} = \sum_a \sum_p N_{a,t,p}^{Em} \left(stray.matrix\right)$
Number of 1SW immigrants	8	$N_{0,t,p}^{Imm} = N_{t,p}^{Imm} / \left(1 + \left(1 - M^{Adult}\right) + \left(1 - M^{Adult}\right)^2 + \left(1 - M^{Adult}\right)^3\right)$

Description		Equation
Returns plus immigrants	9	$N^{Return}_{a,t,p} = \left\{ \begin{array}{ll} N^{Imm}_{0,t,p} + N^{Homing}_{0,t,p} & \text{if } a = 0 \\ N_{a,t,p} & \text{if } a > 0 \end{array} \right\}$
<u>*Projection Model*</u>		
1SW virgin spawners	10	$N_{0,t+2,p} = N_{0,t+2,p} + d^{fw}_{t,p} \dfrac{\alpha_p N^{Return}_{a,t,p}}{\left(1 + \dfrac{\alpha_p N^{Return}_{a,t,p}}{R^{asy}_p h_p}\right)} m_1 \left(1 - M^{Sea}_{t+1}\right)$
2SW virgin spawners	11	$N_{0,t+3,p} = d^{fw}_{t,p} \dfrac{\alpha_p N^{Return}_{a,t,p}}{\left(1 + \dfrac{\alpha_p N^{Return}_{a,t,p}}{R^{asy}_p h_p}\right)} \left(1 - m_1\right)\left(1 - M^{Sea}_{t+1}\right)\left(1 - M^{Sea}_{t+2}\right)$
Repeat spawners	12	$N_{a+1,t+1,p} = N_{a,t,p} \left(1 - M^{Adult}_{p}\right)$
<u>*Incorporating random variability*</u>		
Autocorrelated annual mortality rate for fresh water	13	$w_t = w_{t-1} d_{fw} + w^*_t \sigma_w \quad w^*_t \sim N(0,1)$ $d^{fw}_{t,p} = \exp\left(\left(w_t + dev_p\right) - \sigma^2_d / 2\right) \quad dev_p \sim N(0,0.75)$
Autocorrelated annual marine mortality	14	$dev^{mar}_t = w_{t-1} d_{mar} + w^*_t \sigma_w \quad w^*_t \sim Unif(-1,1)$ $M^{Sea}_t = \overline{M} + dev^{mar}_t$

Description		Equation
Equilibrium calculations		
Spawner-per-recruit (SPR)	15	$SPR = N_{1SW} + N_{2SW} \quad \text{where:}$ $N_{1SW} = \left(1 - M^{Sea}\right) m_1 \sum_a \left(1 - M^{adult}\right)^a$ $N_{2SW} = \left(1 - M^{Sea}\right)^2 (1 - m_1) \sum_a \left(1 - M^{adult}\right)^a$
Deterministic equilibrium for adults	16	$E^{Adult} = \left(\frac{\alpha\left(SPR - 1\right) R^{asy} h}{\alpha} \right)$
Deterministic equilibrium for smolts	17	$E^{Smolt} = \left(\frac{\alpha E^{Adult}}{\frac{1 + \alpha E^{Adult}}{R^{asy} h}} \right)$

Appendix 2-2. Analysis of sensitivity to the initial parameterization

Changes to the simulation model that substantially reduced extinction probabilities of small populations resulted in little or no benefit to metapopulation structure. Thus, the simulation results often showed small percent differences between the trajectories of metapopulations and isolated populations at higher levels of productivity; mainly in the High, Medium and Low scenarios (Figure 2-8). However, in all cases, there was a distinct positive peak in percent difference followed by a reversal in the Not Viable scenario, consistent with the main simulations.

Assumption1: Stray rate

Metapopulations are characterized by intermediate stray rates; those low enough to prevent synchrony among populations, yet high enough to prevent effective population isolation (Driscoll 2007). Estimates of stray rates for diadromous fishes vary depending on monitoring intensity, location, species, and monitoring method. Tagging estimates tend to be in the range of 5-10% (Keefer & Caudill 2014, Melvin et al. 1985) while estimates based on molecular techniques or elemental signatures tend to be higher (e.g., Gahagan et al. 2012). Although I used a value similar to tagging estimates for Atlantic salmon for the main simulations (10%), I used a rate that is 3x higher (30%) here. Median adult abundance over time as well as extinction rates for the metapopulation assemblage were very similar to scenarios assuming a 10% stray rate, in that the apparent benefit of metapopulation structure disappeared in the Not Viable as compared to the Medium to Extreme Low scenarios (30% stray rate; Figure 2-8). However, the conclusions related to large populations in the Not Viable scenario become notably stronger, in that extinction risk increased substantially for a metapopulation as compared to an isolated group of populations, as indicated by the increased negative difference between the isolated and metapopulation trajectories.

At a zero stray rate, populations are isolated by definition. Here I have effectively profiled over increasing stray rates (0, 10% and 30%) and found that high levels of straying result in lower extinction risk to small populations at high levels of productivity, indicated by the extremely small positive percent differences in the High to Low scenarios (30% stray rate; Figure 2-8). However, the reversal in the Not Viable scenario becomes notably stronger as stray rates increase. Therefore, I did not show a sensitivity run to a lower stray rate or to stray rates that either increase or decrease with population size. My expectation would be that there would be smaller maximum differences between the isolated vs. the metapopulation trajectories (i.e., abundances would be more similar) when stray rates are lower as compared to the main simulation, corresponding to a slight shift in the peak of the function from the Low to Extreme Low scenarios. However, the negative difference in the Not Viable scenario should remain the same.

Assumption 2: Immigration hypotheses

Several hypotheses have been proposed for what attracts individuals to enter non-natal watersheds. An isolation-by-distance (IBD) model is based on the idea that proximity to natal watersheds determines the probability of immigration, where proximity is typically measured by geographic distance (Guillot et al. 2009). An alternate hypothesis is that returning adults follow large aggregations of individuals or preferentially return to large watersheds (Keefer & Caudill 2014), where the probability of immigration would be related to population size or habitat area. To test the effect of this alternate hypothesis on these analyses, I developed a stray matrix in which the probability of entering a specific watershed was proportional to the recipient watershed’s size. Using this matrix essentially eliminated the benefit of metapopulation structure in terms of total abundance, although the values were still marginally positive from the High to Extreme Low scenarios and became negative in the Not Viable scenario (area-based spatial model; Figure 2-8).

Assumption 3: Geographic position of watersheds

Using an IBD model to parameterize the stray matrix makes the simulation spatially explicit (Perry & Enright 2007), and the geographic position of watersheds relative to each-other becomes important. This could affect the simulations in one major way, in that results could become sensitive to the specific vector used to parameterize habitat area among populations. Here I used a repeating motif of ‘large’,’small’,’small’,’small’,’small’(x10) to describe the habitat area of specific watersheds. In a spatially explicit model, this means that large populations are distributed equally (i.e., distributed with maximum distance from one another) throughout the region, since populations were assumed to positioned along a line. However, there also is the possibility that results are sensitive to how clumped or dispersed populations of different sizes are. This possibility was evaluated in two ways, first by making the stray matrix spatially implicit (Perry & Enright 2007) and then by changing the distribution of large populations. A spatially implicit stray matrix is one in which the probability of immigration into any watershed is the same. If conclusions were a direct result of the geographic position of watersheds, the spatially implicit model removes any consideration of geographic position. Substituting this matrix into the simulations reduced the range of productivities over which metapopulations were beneficial relative to the simulations assuming IBD (spatially-implicit model; Figure 2-8), in a way that was virtually identical to increasing the stray rate. The second way of evaluating the sensitivity of results to the geographic position of watersheds was to change their distribution in the vector describing habitat area. Maximum clumping of populations of similar size would be achieved by making populations 1 to10 large and populations 11 to 50 small. Again, this had no qualitative effect on the functional form of the percent differences as productivity declined (cluster large watersheds; Figure 2-8), yet the benefits of metapopulation structure were lower than in the main simulations.

Assumption 4: Relative watershed size

The specific vector used to describe watershed size gave the area of large populations as 200 habitat units and small populations as 40 habitat units (80% lower). This was thought to be a sufficient level of contrast in order to demonstrate differences in dynamics relative to watershed size and to evaluate the rescue effect. Typical coastlines exhibit orders of magnitude difference among individual watershed areas, with larger systems (e.g., first or second order streams) interspersed with smaller, coastal systems (Bowlby et al. 2014). An alternative model parameterization would be to fully randomize the starting vector for watershed size (i.e., for each iteration, the size of specific watersheds becomes a random draw from a distribution characterized by a specific mean and variance). However, in order to (1) evaluate the effect of environmental variability in isolation of variability in watershed size, as well as (2) to ensure a high level of contrast among watersheds to maximize the potential to detect a difference between the dynamics related to size (while keeping the number of simulations manageable), I did not do this. I did test the robustness of conclusions under situations in which the contrast between sizes of watersheds was greater (2 orders of magnitude; 1000 habitat units or 10 habitat units) and lesser (50% lower; 200 habitat units or 100 habitat units). As expected, when the contrast among watershed sizes is greater, the trajectories become dominated by the dynamics of large populations, resulting in small difference between total adult abundance in the High to Extreme Low scenarios (greater contrast (area); Figure 2-8). When the contrast in watersheds was smaller, the benefit of metapopulation structure to individual populations is maintained for longer as productivity declines due to increased persistence of small populations in the projections, as shown by the larger positive value in the Extreme Low scenario (less contrast (area); Figure 2-8).

Assumption 5: Environmental autocorrelation

Two types of environmental autocorrelation were incorporated into these analyses: (1) temporal autocorrelation in freshwater and marine environments and (2) partial synchrony among populations in freshwater. Relative to (1), increasing temporal autocorrelation describes a situation in which good years are increasingly likely to be followed by good years (and bad by bad). Such long-term cyclical patterns in climatic variables are well-described (Legendre 1993) and would be expected to influence fish mortality rates. Here I evaluated the influence of weaker temporal autocorrelation (d = 0.3) in freshwater and marine environments on a metapopulation relative to a group of isolated populations. As expected, this reduced extinction probabilities in small populations and reduced the magnitude of the positive differences in the High to Low scenarios (less autocorrelation; Figure 2-8). The reversal in the Not Viable scenario was still apparent.

Assumption 6: Environment-specific changes in productivity

In the original simulations, productivity among scenarios changed by adding a multiplier on both the parameters that describe freshwater production and at-sea mortality (i.e., both the number of animals produced in freshwater and then the number that return to spawn from the marine environment are reduced). However, it is possible that changes to productivity could result from threats to populations in either the marine or freshwater environment in isolation. For example, Atlantic salmon populations in rivers draining into the inner Bay of Fundy are characterized by extremely high at-sea mortality rates, estimated to be in excess of 95%. Population dynamics modeling predicts non-viable maximum lifetime reproductive rates and deterministic equilibrium sizes of zero related to these changes in at-sea mortality (Trzcinski et al. 2004). Alternately, threats such as impassable dams or acidification would be expected to substantially reduce the capacity of freshwater environments to produce juveniles (Bowlby et al. 2014), independently of any conditions in the marine environment.

Here, I used the equilibrium calculations to determine the specific values of alpha (for freshwater) and at-sea mortality and mortality on repeat-spawners (for marine) that would lead to similar maximum lifetime reproductive rates as in the simulations presented. This is effectively a comparison of specific types of threats (either those affecting juveniles in freshwater or those affecting adults at sea) and whether or not metapopulation structure becomes more or less beneficial as related to the timing of changes in productivity. When threats affect populations exclusively in the marine environment, the qualitative pattern is essentially identical as in the main simulations (marine threats only; Figure 2-8). When threats only influence the capacity of spawners to produce recruits (i.e., alpha), overall extinction probabilities for small populations are reduced and any benefit of metapopulation structure to total abundance is very slight in the High to Extreme Low scenarios (freshwater threats only; Figure 2-8). This result would be related to equilibrium adult abundance for populations in each of these two sensitivity runs. For example: when only marine mortality increases, equilibrium abundance for adults in the Medium scenario remains comparable to the original Medium scenario (*c.f.* 1748 and 350 for large and small populations, respectively, with 1831 and 366 from Table 1). When only freshwater production declines, equilibrium abundance for adults remains very high: 2410 and 482 for large and small populations, respectively. Biologically this suggests that straying becomes more influential to population persistence when threats are affecting life stages characterized by density-independent survival.

Assumption 7: Quasi-extinction threshold

For natural populations, there is extensive debate on how true extinction is measured or characterized (e.g., Dulvy et al. 2004). The idea of an extinction threshold (the minimum amount of habitat area, connectivity and quality required for a species or population to persist; Kuussaari et al. 2009) would be closely related to the idea of a minimum viable population size (typically defined as a percentage of the original population size (e.g., 80%) persisting for a given number of years (e.g., 20 years); Traill et al. 2007). Below such thresholds, populations are thought to be functionally extinct (i.e., will decline to zero), although there may be a time lag between reaching these thresholds and actual extinction (Tilman et al. 1994). Empirical estimation is extremely uncertain and is context-dependent, both on the species and on localized environmental processes (Traill et al. 2007). Estimating the minimum number of individuals required in a population to ensure reproductive output would be a similar problem to the thresholds identified above, in that it would likely be extremely uncertain and context-dependent. I originally used a quasi-extinction threshold of 20 individuals to represent a population size below which spawning would not take place. Ideally this would represent 10 males and 10 females in the population, although demographic stochasticity could be significant at such small population sizes and could markedly skew the sex ratio (Simberloff 1998). Here, I have compared those results to ones using a quasi-extinction threshold of 2.

Biologically, only two individuals are required to produce offspring (one male and one female), although the chances of them successfully spawning in a large river system may be low. Changing the quasi-extinction threshold reduced the overall extinction risk of individual populations (similar to the autocorrelation sensitivity analysis above), and meant that metapopulation structure had maximum benefit in the Extreme Low scenario (low quasi-extinction; Figure 2-8). However, abundance in isolated populations was still higher than in a metapopulation in the Not Viable scenario.

Assumption 8: Survival of strays

To ensure that the structure of the simulation models for metapopulations and isolated populations were the same, reproductive output of strays and homing individuals was the same and no additional mortality affected individuals as they strayed. Given that there is the potential that some individuals leave the metapopulation due to imprecise homing (Keefer & Caudill 2014), that the additional energy strays expend during migration causes a low level of mortality (Bowler & Benton 2005, Bonte et al. 2012), or that individuals have slightly lower fitness in non-natal habitat (Rieman & Dunham 2000, Pfluger & Balkenhol 2014), I applied a low mortality rate (5%) to strays. The results from this sensitivity analysis very closely mirror the main simulation results in chapter two, although the actual values for the differences at each level of productivity are smaller; less positive or less negative as compared to the equivalent scenario in the chapter text (5% mortality; Figure 2-8).

Appendix 4-1: Evaluation of the microsatellite data relative to assumptions

I used Micro-checker version 2.2.3 to test for null alleles, scoring error and allele drop-out (van Oosterhout et al. 2004). Null alleles are those that do not amplify at a specific locus, producing false homozygotes; scoring error can arise from PCR amplifications that produce stutter products which differ from the original allele by multiples of the repeated microsatellite sequence; allele drop-out relates to preferred amplification of small alleles (van Oosterhout et al. 2004, Pompanon et al. 2005). Any of these errors can cause apparent deviations from Hardy-Weinburg equilibrium (see below). At each locus and for each population, micro-checker did not find evidence for scoring error due to stuttering or large allele drop-out. Table A4-1-1 summarizes the evidence for null alleles at individual loci for each population analyzed. The rescaled gene frequencies, using the Brookfield 1 method to account for null alleles (van Oosterhout et al. 2004), show relatively small differences as compared to the un-corrected gene frequencies. Using the rescaled frequencies to calculate F_{st} had no measureable effect on the resulting distance matrix relative to using the observed data. Therefore, the observed data were used for the analyses presented in chapter four.

I tested for Hardy-Weinburg (HW) equilibrium over loci in all sampled populations using the R package ‘adegenet’ (Jombart 2008). Hardy-Weinburg equilibrium stipulates that the amount of genetic variation in a population should remain constant in the absence of selection, which means that the expected gene frequencies for a locus with two alleles equates to: $p^2 + 2pq + q^2 = 1$. At each locus within each population, 8000 Monte-Carlo simulations were used to calculate p-values for the chi-square test of observed vs. predicted gene frequencies. Significance was assessed relative to a Bonferroni-corrected critical p-value (0.05/17). Results are given in Table A4-1-2.

Linkage disequilibrium occurs if there is non-random association of alleles at multiple loci. I tested for potential linkage disequilibrium using LinkDos (Garnier-Gere & Dillmann 1992) from GenePop 4.2 (Raymond & Rousset 1995; Rousset 2008). From the 17 microsatellite loci, there are 136 unique combinations of loci pairs per population, or 1496 comparisons in total from the 11 populations. At an alpha of 0.05, 75 comparisons might be expected to be significant due to chance. There were only 43 significant comparisons, which are given in Table A4-1-3.

Table A4-1-1. Loci and populations in which evidence for null alleles were found. Each instance is highlighted in yellow.

	Round Hill	Salmon River (Digby County)	Tusket	Medway	LaHave	Gold	Musquodoboit	Moser	St. Mary's	Country Harbour	Salmon River (Guys. County)
Locus 1	no	no	no	no	no	no	no	no	no	no	no
Locus 2	no	no	no	no	no	no	no	no	no	no	no
Locus 3	no	no	no	no	no	no	no	no	no	no	no
Locus 4	no	no	no	no	no	no	no	no	no	no	no
Locus 5	no	no	no	no	no	no	no	no	no	no	no
Locus 6	no	no	no	no	no	no	no	no	no	no	no
Locus 7	no	yes	no	no	no	no	no	no	no	no	no
Locus 8	no	no	no	no	no	no	no	no	no	no	no
Locus 9	no	no	no	no	yes	yes	no	yes	yes	yes	no
Locus 10	no	no	no	no	no	no	no	no	no	no	no
Locus 11	no	no	no	no	no	no	no	no	no	no	no
Locus	no	no	no	no	no	no	no	no	no	no	no

	Round Hill	Salmon River (Digby County)	Tusket	Medway	LaHave	Gold	Musquodoboit	Moser	St. Mary's	Country Harbour	Salmon River (Guys. County)
12											
Locus 13	no	no	no	no	no	no	no	no	no	no	no
Locus 14	no	no	no	no	no	no	no	no	no	no	no
Locus 15	no	no	no	no	no	no	no	no	no	no	no
Locus 16	no	no	no	no	no	no	no	no	no	no	no
Locus 17	no	no	no	no	no	no	no	no	no	no	no

Table A4-1-2. Loci that showed significant deviation from Hardy-Weinburg equilibrium (p-values highlighted in yellow) based on a permutation test.

River	L 1	L 2	L 3	L 4	L 5	L 6	L 7*	L 8	L 9**	L 10	L 11	L 12	L 13	L 14	L 15	L 16	L 17
Round Hill	0.965	0.012	1	0.71	0.883	1	0.583	0.358	0.905	0.574	0.004	0.012	0.921	0.49	1	0.028	**0**
Salmon River (Digby Co.)	0.815	0.175	1	0.886	1	0.961	0.006	0.869	**0**	0.056	**0**	0.278	**0**	0.632	0.118	0.935	0.41
Tusket	0.498	0.921	0.305	0.615	1	0.782	0.895	0.932	0.191	0.999	0.531	0.639	0.671	0.884	0.479	0.214	1
Medway	0.983	1	0.473	1	0.407	0.999	0.461	0.952	1	0.986	0.655	0.972	0.995	0.547	**0**	0.869	0.924
LaHave	1	**0**	1	1	1	0.847	0.954	0.998	0.023	0.915	0.867	0.075	0.093	0.946	0.523	0.075	0.242
Gold	0.833	0.514	0.792	0.999	0.991	1	0.632	1	**0**	0.464	0.967	0.988	0.2	0.963	0.804	0.834	0.251
Musquodoboit	0.015	**0**	0.432	0.129	0.98	0.816	0.987	1	0.736	1	0.832	0.726	0.873	**0**	0.301	0.902	1
Moser	0.397	0.105	0.116	0.457	0.695	0.908	0.999	1	**0**	0.974	0.091	**0**	**0**	0.967	0.999	0.524	0.01
St. Mary's	1	1	0.094	1	0.912	0.998	0.915	0.995	0.328	1	0.943	0.029	1	0.496	0.857	0.561	0.978
Country Harbour	0.314	1	0.953	1	0.154	1	1	0.074	**0.001**	1	0.999	0.17	0.182	0.97	0.389	0.973	**0**
Salmon River (Guys. Co.)	0.985	**0**	1	0.612	0.06	1	0.986	0.998	0.989	1	0.027	**0**	**0**	0.825	1	0.03	**0**

* Salmon River (Digby County) also had null alleles at Locus 7

** Country Harbour, St. Mary’s, Moser, Gold and LaHave Rivers had null alleles at Locus 9

Table A4-1-3. Non-random associations of alleles after Bonferroni correction (i.e., only significant results) from the 17 microsatellite loci for each of the 11 populations analyzed using LinkDos.

Population	Comparison		P-Value	S.E.
Round Hill	Loc5	Loc12	0.000015	0.000015
Round Hill	Loc10	Loc12	0.000055	0.000055
Round Hill	Loc9	Loc13	0	0
Salmon River (Digby)	Loc1	Loc2	0	0
Salmon River (Digby)	Loc4	Loc5	0	0
Salmon River (Digby)	Loc5	Loc6	0	0
Salmon River (Digby)	Loc6	Loc7	0	0
Salmon River (Digby)	Loc4	Loc10	0	0
Salmon River (Digby)	Loc5	Loc10	0	0
Salmon River (Digby)	Loc6	Loc11	0.00009	0.000066
Salmon River (Digby)	Loc1	Loc13	0	0
Salmon River (Digby)	Loc2	Loc13	0	0
Salmon River (Digby)	Loc10	Loc13	0	0
Salmon River (Digby)	Loc11	Loc16	0.000014	0.000014
Salmon River (Digby)	Loc11	Loc17	0	0
Tusket	Loc4	Loc5	0	0
Tusket	Loc1	Loc11	0	0
Tusket	Loc11	Loc12	0	0
Tusket	Loc11	Loc14	0.000263	0.000263
Tusket	Loc11	Loc15	0.000306	0.000142
Tusket	Loc8	Loc17	0	0
Medway	Loc14	Loc16	0.000156	0.000102
Gold	Loc8	Loc12	0.000241	0.000241
Gold	Loc10	Loc13	0	0
Musquodoboit	Loc2	Loc10	0	0
Musquodoboit	Loc2	Loc12	0	0
Moser	Loc2	Loc5	0	0
Moser	Loc4	Loc5	0.000011	0.000011
Moser	Loc1	Loc6	0.000024	0.000024
Moser	Loc5	Loc6	0	0
Moser	Loc5	Loc7	0	0
Moser	Loc2	Loc10	0	0
Moser	Loc6	Loc10	0.000264	0.000201
Moser	Loc7	Loc10	0.000126	0.00009
Moser	Loc4	Loc11	0.000145	0.000145
Moser	Loc6	Loc11	0	0

Population	Comparison		P-Value	S.E.
Moser	Loc10	Loc12	0.00016	0.00016
Moser	Loc8	Loc17	0	0
St. Mary's	Loc4	Loc12	0.000363	0.000248
Country Harbour	Loc2	Loc16	0	0
Country Harbour	Loc7	Loc16	0	0
Salmon River (Guysborough)	Loc11	Loc12	0	0
Salmon River (Guysborough)	Loc1	Loc14	0.000339	0.000268

Appendix 4-2: Spatial data and watershed characteristics

The spatial and geographic data sources as well as data processing steps are fully described in the Appendices that accompany Bowlby et al. (2014). All geographic measurements made of spatial data used Universal Transverse Mercator projection, NAD83 datum for Zone 20 North. Tabular queries to aggregate information and generate basic statistical information (e.g., sum, mean) were carried out primarily using Microsoft® Access 2002 software (service pack 3). Important characteristics of the spatial data include: (1) calculated stream length includes inferred flow segments through all waterbodies large enough to be represented as polygons, (2) the bedrock geology types were aggregated into five general categories to give proportionately less weight to minor formations, while the surficial geology types were not aggregated, (3) the natural disturbance regimes were based on a wide variety of environmental data related to forest structure and composition, and are thus partially redundant with other data types included in the analyses (4) several of the land use codes in the Forest Inventory Data used to evaluate land use were reclassified to represent larger groupings of human activity types, and (5) data used to characterize dams as well as historic mining activities were known to be incomplete. The data sources are given in Table A4-2-1. River-specific values and variables for the five types of landscape data describing watersheds used in the NMDS analyses are given in Tables A4-2-2 to A4-2-6, inclusive. Values for each of the threats categories are given in Table A4-2-7.

Table A4-2-1. Description and data sources of information used in the geographic analyses of watersheds of the Southern Upland. All on-line data accessed between October 15, 2011 and January 15, 2012, unless otherwise noted.

Description	Data Source / Data Credit
Hydrology – rivers and water bodies	GeoBase's National Hydro Network (NHN), Level 1, Edition 1 / Natural Resources Canada
Secondary Watersheds	Custom Data Product derived from NSTDB[1] obtained from Nova Scotia Department of Environment
Digital Elevation Data (DEM)	GeoBase's Canadian Digital Elevation Data (CDED)
Bedrock Geology, DP ME 43, Version 2, 2006	Nova Scotia Department of Natural Resources – Mineral Resources Branch
Surficial Geology DP ME 36, Version 2, 2006	Nova Scotia Department of Natural Resources – Mineral Resources Branch
Forest Inventory Cycle 2 & 3	Nova Scotia Department of Natural Resources – Forestry Branch
Ecological Land Classification	Nova Scotia Department of Natural Resources – Forestry Branch
Roads	GeoBase's National Road Network, Edition 8.0 / Natural Resources Canada
Dams – NHN	GeoBase's National Hydro Network (NHN), Level 1, Edition 1 / Natural Resources

Description	Data Source / Data Credit
	Canada
Dams – NSE	Nova Scotia Department of Environment and Fisheries and Oceans Canada, Maritimes Region, Habitat Protection and Sustainable Development Division (pers. comm., DFO–HPSD March 2011)
Hydro Power Generating Stations	Nova Scotia Power Inc. ((pers. comm. NSPI, February 6, 2012)
Aquaculture – licensed marine sites in Nova Scotia	Nova Scotia Department of Fisheries and Aquaculture
Nova Scotia Abandoned Mine Openings (AMO) Database	Nova Scotia Department of Natural Resources – Mineral Resources Branch
Fall and spring trout stocking distribution lists for 2010, 2011	Nova Scotia Department of Fisheries and Aquaculture

[1] NSTDB = Nova Scotia Topographic Database.

Table A4-2-2. Area metrics describing watershed characteristics. Perimeter, stream length, inferred flow and total flow are in km, others in km^2. Inferred flow refers to straight-line distances through water features considered to be polygons in ArcGIS. Inland water refers to lake and stillwater area.

River	Perimeter	Area	Stream length	Total Flow	Inferred flow	Inland water
Round Hill	84.37	122.12	50.01	35.59	85.6	3.59
Salmon (Digby)	108.06	234.07	88.2	113.61	201.82	29.13
Tusket	307.15	1456.22	497.2	692.47	1189.67	133.41
Medway	296.01	1519.14	559.84	698.72	1258.56	130.58
LaHave	258.69	1524.16	648.41	515.37	1163.78	85.61
Gold	130.7	386.19	145.65	117.77	263.42	16.89
Musquodoboit	248.55	719.08	734.26	213.38	947.64	24.19
Moser	92.62	177.3	92.87	78.84	171.7	10.23
St. Mary's	365.04	1336.82	1025.32	297.99	1323.31	38.41
Country Harbour	79.91	183.47	142.92	43.88	186.8	6.16
Salmon (Guys.)	111.46	298.71	180.05	95.12	275.16	14.46

Table A4-2-3. Surficial geology types given as proportions of watershed area. Formations were not grouped because even rare ones could have a large impact on watershed characteristics; for example, by providing isolated locations of less acidified water. Columns are: A – Alluvial, B – Bedrock, G – Glaciolac, M – Morraine, K – Kame, Mar– Marine, N – None, O – Organic, Ow – Outwash, R – Residuum, SD – Silty drumlin, ST – Silty till, StD – Stony drumlin, StT – Stony till.

River	A	B	G	M	K	Mar	N	O	Ow	R	SD	ST	StD	StT
Round Hill	0.460	1.924	0.005	6.612	0.472	0.423	0.000	1.264	0.580	0.000	13.254	10.937	0.000	64.069
Salmon (Digby)	0.000	0.000	0.000	15.776	0.655	0.000	0.000	4.191	0.000	0.000	5.094	0.000	9.47	64.815
Tusket	0.081	0.000	0.000	7.859	0.328	0.000	1.459	4.523	0.000	0.000	1.711	0.000	3.99	80.049
Medway	0.053	5.028	0.000	3.307	0.61	0.000	0.641	3.509	0.156	0.000	10.491	2.044	2.157	72.004
LaHave	0.454	11.03	0.049	4.118	0.924	0.000	2.012	2.889	0.008	0.018	15.489	21.372	0.031	41.605
Gold	0.078	9.496	0.000	4.131	0.815	0.000	0.973	2.236	0.164	0.000	7.560	29.658	0.000	44.889
Musquodoboit	5.733	9.154	0.407	3.043	0.305	0.000	0.861	2.773	0.582	0.000	0.575	38.829	0.000	37.737
Moser	0.000	16.248	0.000	0.000	3.929	0.000	1.118	1.195	0.000	0.000	10.439	36.945	0.000	30.125
St. Mary's	2.864	6.538	0.000	1.691	2.497	0.000	0.501	1.217	0.000	0.061	2.828	24.224	0.000	57.578
Country Harbour	0.937	0.620	0.000	0.000	0.464	0.000	0.000	1.418	0.000	0.000	2.722	41.870	0.000	51.969
Salmon (Guys.)	0.403	19.834	0.000	4.714	1.565	0.000	1.858	1.75	0.000	0.000	1.314	37.738	0.000	30.823

Table A4-2-4. Bedrock geology types given as proportions of watershed area, grouped into five major categories. Undivided represents areas which were not divided into specific formations.

River	Granite	Other	Sandstone	Slate	Undivided
Round Hill	0.9774	0	0.0226	0	0
Salmon (Digby)	0	0.0894	0	0.9106	0
Tusket	0.2715	0.0434	0	0.6851	0
Medway	0.2680	0.0202	0	0.7117	0
LaHave	0.4981	0.0524	0	0.4495	0
Gold	0.7704	0.1624	0	0.0671	0
Musquodoboit	0.1036	0.2848	0	0.6116	0
Moser	0	0	0	1	0
St. Mary's	0.0323	0.2624	0	0.1571	0.5482
Country Harbour	0.1160	0.0608	0	0.1735	0.6497
Salmon (Guys.)	0.0412	0.6850	0	0.1061	0.1678

Table A4-2-5. Natural disturbance regime given as proportions of watershed area. Definitions are as follows: Open seral: ecosystems where site conditions restrict or limit tree growth (e.g., flooding, extreme wind exposure); Frequent: ecosystems which result in the rapid mortality of an existing stand and the re-establishment of a stand of even age; Gap: ecosystems where areas are seldom exposed to disturbances, but are characterized by stands with multiple age classes; Infrequent: ecosystems where stand initiating events (e.g., disturbance) characterize the development of the forest, but the interval between such events is long.

River	Frequent	Gap	Infrequent	Open
Round Hill	36.91	8.43	48.12	2.53
Salmon (Digby)	16.11	29.01	36.20	5.58
Tusket	38.10	13.71	32.89	5.63
Medway	34.74	21.10	32.10	3.05
LaHave	22.89	25.07	43.79	2.38
Gold	15.81	25.25	50.99	3.34
Musquodoboit	31.87	29.56	30.67	4.27
Moser	53.59	30.46	8.510	1.65
St. Mary's	39.71	29.47	26.26	1.57
Country Harbour	30.26	18.73	47.18	0.24
Salmon (Guys.)	21.16	56.10	16.90	0.62

Table A4-2-6. Topographical characteristics of each watershed based on a Digital Elevation Model with 20 m horizontal and 1m vertical resolution. Mean slope is in degrees, elevations (maximum, mean and standard deviation) are given in meters, and roughness refers to the mean standard deviation of the slope within a 5x5 cell window (i.e. 100 x 100 m) of a watershed.

River	Mean slope	Max. elevation	Mean elevation	SD elevation	Roughness
Round Hill	3.68	265	169.94	48.05	1.66
Salmon (Digby)	2.15	97	43.97	15.47	1.04
Tusket	2.05	208	80.13	42.57	1.03
Medway	2.25	269	116.00	50.62	1.20
LaHave	2.92	288	156.09	53.27	1.43
Gold	3.13	276	157.50	54.29	1.50
Musquodoboit	3.52	221	92.28	49.59	1.68
Moser	3.12	176	97.25	37.32	1.57
St. Mary's	3.01	311	134.57	52.48	1.43
Country Harbour	4.07	203	96.51	39.75	1.94
Salmon (Guys.)	3.46	240	126.11	33.83	1.69

Table A4-2-7. Summary of the land use variables considered in these analyses, their abbreviations, and the threat indices (F = fragmentation, S = sedimentation, WQ = water quality) that they were evaluated relative to (as in Table 4-4). Impassable dams ins the percentage of stream length that is inaccessible; road density is the length of road (km) per km^2 of watershed area; road crossings are a count of the number per 10 km of stream length; pH categories are defined below; agriculture, forestry and industry are percentages of the watershed area affected; and mines is a count of the total number of abandoned mines per km^2 of watershed area.

	Impassable dams	Road density	Road crossings	pH category[4]	Agriculture	Forestry	Industry	Mines
Abbreviation	D	Rd	Rc	P	A	F	I	M
Index	F	S	F	F, WQ	S	WQ	WQ	WQ
Salmon (Digby)	11.00	0.81	2.87	3	1.54	10.40	0.96	0.000
Tusket	94.50	0.55	0.99	2	0.66	8.45	0.37	0.036
Medway	17.70	0.74	1.69	3	0.90	9.36	0.71	0.284
LaHave	0.00	0.87	3.09	3	2.64	9.09	0.80	0.006
Gold	0.00	0.67	2.58	3	1.12	9.73	0.60	0.339
Musquodoboit	16.20	0.91	3.47	4	7.43	15.35	0.63	0.054
Moser	0.00	0.46	1.22	3	0.00	8.27	0.06	0.000
St Marys	0.00	0.85	3.29	4	1.74	30.25	0.57	0.028
Country Harbour	0.00	1.10	3.75	4	2.42	23.95	0.82	0.229
Salmon (Guys)	0.00	0.59	2.58	4	0.67	15.52	0.79	0.030

---

[4] Estimated mean annual pH: 2 = 4.7-5.0, 3 = 5.1-5.4, 4 = above 5.4. Means were determined from monthly measurements at multiple sites distributed in tributaries throughout each watershed. Note that the expectation is that Atlantic salmon mortality would decrease with increasing pH; therefore, in analyses, the categories were recoded so that higher values represented rivers thought to experience higher mortality (i.e., 4 became 2 and 2 became 4). Also note that ‘4’ is an assumed category for Salmon River (Guysborough) on the basis of known geological patterns in the region as well as the categorization of geographically proximate rivers.

Appendix 4-3. Distance matrices used in the landscape genetic analyses

Genetic distance matrix

The genetic distance matrix was calculated from Nei's pairwise F_{st} (Nei 1978).

	Country Harbour	Gold	LaHave	Medway	Moser	Musquodoboit	Salmon (Digby)	St. Mary's	Tusket
Country Harbour	0.0000								
Gold	0.2617	0.0000							
LaHave	0.2866	0.1943	0.0000						
Medway	0.2744	0.1458	0.1871	0.0000					
Moser	0.2786	0.3230	0.3423	0.2798	0.0000				
Musquodoboit	0.2727	0.2533	0.2406	0.2178	0.3515	0.0000			
Salmon (Digby)	0.3486	0.2307	0.3169	0.1344	0.3373	0.2893	0.0000		
St. Mary's	0.2103	0.1495	0.2190	0.1329	0.2439	0.1905	0.2332	0.0000	
Tusket	0.3214	0.2426	0.2433	0.1628	0.3064	0.2927	0.1891	0.2005	0.0000

Isolation by Distance model

The Euclidean matrix showing pairwise geographical distances (IBD model) was calculated from the coastal straight-line distance between each pair of rivers, where coastal distance represented a minimum distance over water.

	Country Harbour	Gold	LaHave	Medway	Moser	Musquodoboit	Salmon (Digby)	St. Mary's	Tusket
Country Harbour	0.00								
Gold	821.60	0.00							
LaHave	956.88	198.02	0.00						
Medway	1034.06	325.75	147.01	0.00					
Moser	302.82	604.15	760.23	855.61	0.00				
Musquodoboit	523.28	375.30	543.74	651.05	264.23	0.00			
Salmon (Digby)	1406.68	989.20	860.85	744.95	1336.23	1206.74	0.00		
St. Mary's	168.03	717.87	865.17	952.72	153.12	395.25	1385.74	0.00	
Tusket	1318.52	825.99	687.71	567.13	1218.14	1068.70	213.02	1280.75	0.00

Isolation by Environment model

The Euclidean matrix showing pairwise environmental distances (IBE model) was calculated from centroids in multivariate space; where the environmental input variables that defined the centroids were calculated from five types of landscape data (Appendix 4-2; Tables A4-2-2 to A4-2-6, inclusive).

	Country Harbour	Gold	LaHave	Medway	Moser	Musquodoboit	Salmon (Digby)	St. Mary's	Tusket
Country Harbour	0.00								
Gold	35.00	0.00							
LaHave	38.00	3.00	0.00						
Medway	40.00	5.00	2.00	0.00					
Moser	9.01	26.00	29.00	31.00	0.00				
Musquodoboit	20.01	15.01	18.01	20.01	11.01	0.00			
Salmon (Digby)	55.00	20.01	17.01	15.01	46.01	35.00	0.00		
St. Mary's	3.02	32.00	35.00	37.00	6.01	17.00	52.00	0.00	
Tusket	52.00	17.00	14.00	12.00	43.00	32.01	3.07	49.00	0.00

Appendix 4-4. An example of Reciprocal Causal Modeling

I developed this example in relation to the identification of the water quality index described in chapter four. I am also going to present data as well as equations directly in the text to try to clarify the method. The threats variables which are relevant to this example are the river-specific values for pH category, the percentage of each watershed affected by forestry activity, and the density of abandoned mine openings (#/km^2 of stream length). I have taken these from Table A4-2-7 (Appendix 4-2), and they are shown below.

	pH category[5]	Forestry	Mines
Abbreviation	P	F	M
Index	F, WQ	WQ	WQ
Salmon (Digby)	3	10.40	0.000
Tusket	2	8.45	0.036
Medway	3	9.36	0.284
LaHave	3	9.09	0.006
Gold	3	9.73	0.339
Musquodoboit	4	15.35	0.054
Moser	3	8.27	0.000
St Marys	4	30.25	0.028
Country Harbour	4	23.95	0.229
Salmon (Guys)		15.52	0.030

---

[5] Estimated mean annual pH: category 2 = 4.7-5.0, category 3 = 5.1-5.4, category 4 = above 5.4. Means were determined from monthly measurements at multiple sites distributed in tributaries throughout each watershed. Note that the expectation is that Atlantic salmon mortality would decrease with increasing pH; therefore, in analyses, the categories were recoded so that higher values indicated rivers thought to experience higher mortality (i.e., 4 became 2 and 2 became 4). Also note that ‘4’ is an assumed category for Salmon River (Guysborough) on the basis of known geological patterns in the region as well as the categorization of geographically proximate rivers.

There are a total of seven possible tables that can be made from these variables, representing groupings of variables alone or in combination. The first two examples would be: FMP (i.e., all of the variables; the table shown above), FM (i.e., Forestry and Mines exclusively; the table shown below).

	Forestry	Mines
Abbreviation	F	M
Index	WQ	WQ
Salmon (Digby)	10.40	0.000
Tusket	8.45	0.036
Medway	9.36	0.284
LaHave	9.09	0.006
Gold	9.73	0.339
Musquodoboit	15.35	0.054
Moser	8.27	0.000
St Marys	30.25	0.028
Country Harbour	23.95	0.229
Salmon (Guys)	15.52	0.030

For analyses using Reciprocal Causal Modeling, each of the 7 potential threats tables needed to be turned into a distance matrix. To do this, I first scaled the values (x_i) of each variable between zero and one, following the equation: $(x_i - min(x))/(max(x) - min(x))$. As discussed in chapter four, scaling the variables this way ensures that all threat variables contribute equally in the analysis. Although it would have been possible to try multiple alternate ways of weighting each variable (e.g., Cushman et al. 2006), I did not because of the complexity it would have added to the analysis.

Here, I have shown the scaled values for the Forestry variable:

	Forestry
Abbreviation	F
Index	WQ
Salmon (Digby)	0.097
Tusket	0.008
Medway	0.049
LaHave	0.037
Gold	0.066
Musquodoboit	0.322
Moser	0.000
St Marys	1.000
Country Harbour	0.713
Salmon (Guys)	0.330

The above table shows the data which were used to calculate the ‘F’ threat model. From this table, the straight-line distance between each pair of points (p_i, q_i) was transformed into Euclidean distance following:

$$d(p_i, q_i) = \sqrt{(q_i - p_i)^2}$$

As an example, the Euclidean distance between the Medway River and the LaHave River becomes:

$$d(Medway, LaHave) = \sqrt{(0.049 - 0.037)^2} = 0.012$$

The entire distance matrix representing the F model is:

	CH	G	LH	MED	M	MUSQ	SD	SMR	T
Country Harbour (CH)	0								
Gold (G)	0.647	0							
LaHave (LH)	0.676	0.029	0						
Medway (MED)	0.664	0.017	0.012	0					
Moser (M)	0.713	0.066	0.037	0.050	0				
Musquo-doboit (MUSQ)	0.391	0.256	0.285	0.273	0.322	0			
Salmon Digby (SD)	0.616	0.030	0.060	0.047	0.097	0.225	0		
St. Mary's (SMR)	0.287	0.934	0.963	0.950	1.000	0.678	0.903	0	
Tusket (T)	0.705	0.058	0.029	0.041	0.008	0.314	0.089	0.992	0

I used this table, identified as the F model, as one of the candidates for Reciprocal Causal Modeling.

Two-step Reciprocal Causal Modeling as proposed by Wasserman et al. (2010) directly compares highly correlated models to each-other using partial Mantel tests. Such reciprocal comparisons have been applied in situations where the data being used in alternate models (e.g., distance matrices representing threats, environmental characteristics, landscape resistance, or geography) are derived from varying combinations of a specific suite of variables (e.g., Cushman et al. 2006). The method becomes two steps when Mantel and partial Mantel tests are first used to identify a candidate model that is subsequently used in the reciprocal comparisons (Wasserman et al. 2010). This can substantially reduce the complexity of the analyses because reciprocal comparisons of each individual model with all other candidate models may no longer be necessary.

For the first step in the analyses of threats contributing to the Water Quality index for chapter four, I identified a candidate model using Mantel and partial Mantel tests. As detailed in Guillot & Rousset (2013), Mantel tests derive from two distance matrices ($D_{i,j}^{x}$) and ($D_{i,j}^{y}$), and the test statistic *r* becomes: $= \sum D_{i,j}^{x} D_{i,j}^{y}$. To assess significance of the correlation, this test statistic is compared to a distribution of randomly permuted values of one of the matrices, where I calculated: $\tilde{r} = \sum \widetilde{D}_{i,j}^{x} D_{i,j}^{y}$ 10,000 times. The p-value becomes the proportion of permutations that lead to a higher correlation coefficient than the test statistic (Oksanen et al. 2013). Partial Mantel tests control for the effect of a third distance matrix when calculating the test statistic, but the permutation procedure is the same.

For the analysis in chapter four, the threat model with the highest correlation with genetic distances included abandoned mine density and pH category (MP), with *r* = -0.383 (p-value = 0.015) in the Mantel test (i.e., Gene~MP), and *r* = -0.435 (p-value = 0.011) in the partial Mantel test after controlling for any influence of IBE on genetic distance (i.e., Gene~MP|IBE; Table 4-4). The second step of the analysis was based on reciprocal comparisons (i.e., two sets of comparisons) among the threats models themselves. For the first set, I calculated the correlation between genetic distance and the MP distance matrix, after controlling for the effect of each alternate threat distance matrix for the Water Quality index (columns labelled ‘Chosen|Alternate’ in Table 4-4). In the order that they are presented in Table 4-4, the partial Mantel tests became: Gene~MP|FMP, Gene~MP|FM, Gene~MP|FP, Gene~MP|F, Gene~MP|M, and Gene~MP|P. As discussed in the description of the method, the expectation is that a better model would have a larger and significant *r* value in each of these comparisons (Wasserman et al. 2010). For the second set of comparisons, I tested for residual correlations between genetic distance and the alternate threat distance matrices after controlling for the influence of the chosen model (columns labeled ‘Alternate|Chosen’ in Table 4-4). Here, the partial Mantel tests were: Gene~FMP|MP, Gene~FM|MP, Gene~FP|MP, Gene~ F|MP, Gene~M|MP, and Gene~P|MP. As discussed in the description of the method, the expectation is that none of the alternate models will demonstrate a significant correlation with genetic distance after the effect of the chosen model is accounted for (Wasserman et al. 2010). Both of these expectations are met relative to the Water Quality index (Table 4-4), so following Cushman et al. (2013), I considered the MP model to be significantly better than the competing threats models, with the caveat discussed in the methods that such tests can be prone to Type I error. However, even if I did not consider the difference to be significant, the simulations in Cushman et al. (2013) suggest that the relative magnitudes of the *r* values in the reciprocal tests can still be used to argue that the chosen model is the best representation of a threat index.

Appendix 5-1. Electrofishing methodology and data

The electrofishing methods implemented on various rivers between 1995 and 2005 were not standardized, differing relative to survey objectives and rivers. Open sites with an estimated catchability, depletion surveys and mark-recapture surveys were all used to estimate fry (age-0) and parr (age-1+) density. However, depletion surveys tended to be done on a limited number of sites within a watershed prior to 2000 (Table A5-1-1, Appendix 5-1), and the vast majority of the data used either open sites or mark-recapture (Table A5-1-1, Appendix 5-1). For both open site and mark-recapture surveys, sites were fished in an upstream direction and no barrier nets were used. For depletion surveys, barrier nets were often used to block off the lower and upper boundaries of a site, and then sites were fished either in an upstream or downstream direction. Between three and five passes were done at each site, where the last pass was one in which very few (e.g., < 10) or no salmon were captured. Traditionally for salmon assessment, a Peterson estimate would be used to calculate densities at mark-recapture sites, and catches at open sites (from the same river and year) would be multiplied by mean capture efficiency and scaled by site area (e.g., Bowlby & Gibson 2012). The mean capture efficiency used for assessments in a specific year was averaged from that year's and the previous year's estimates at mark-recapture sites (Bowlby & Gibson 2012). Similarly for barrier sites, multiple pass depletion methods (e.g., Schwarz & Seber 1999) have been used to estimate juvenile densities by age directly for each site. Due to differences in survey methods throughout the dataset, and because I was interested in total juvenile density at each site (not partitioned by age class), I did not follow the methods that had been used previously to calculate juvenile densities in annual assessments.

For comparability among years, surveys and rivers, I calculated total juvenile density from the number of salmon captured in the first pass of each survey, multiplied by capture efficiency (*q*) and scaled by site area. Capture efficiency is defined as the proportion of the total abundance of juvenile salmon captured on the first electrofishing pass at a site. Thus it was necessary to account for differences in capture efficiency (*q*) among methods. Here, I used all mark-recapture surveys in all rivers and years to estimate a median capture efficiency ($q = 0.34$) that was applied to all open sites and mark-recapture surveys (Table A5-1-1, Appendix 5-1). Capture efficiency for these two survey methods would be expected to be similar because of the way in which sites would have been sampled (i.e., fishing in an upstream direction without barrier nets). Similarly, I used all depletion surveys in all rivers and years to calculate a median capture efficiency for multiple pass surveys ($q = 0.56$). Because multiple pass sites typically used barrier nets, I expected that capture efficiencies from these surveys would have been higher.

Table A5-1-1. Summary of the electrofishing dataset, showing the rivers, the site numbers, which years they were electrofished, and which method was used (‘O’: open site, ‘MR’: mark-recapture, and ‘M’: multi-pass). Blank cells represent years in which no survey took place at a specific site.

River	Site	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
East (Chester)	EstCh001	O	O									
East (Chester)	EstCh002	O	O									
East (Chester)	EstCh003	O	O									
East (Chester)	EstCh004	O	O									
East (Chester)	EstCh005	O	O									
East (Chester)	EstCh006	O	O									
East (Chester)	EstCh008	O	O									
East (Chester)	EstCh009	M	O									
East (Chester)	EstCh010	O	O									
East (Chester)	EstCh011	O	O									
East (Chester)	EstCh012	M	O									
East (Chester)	EstCh013	M		O			O					
East (Chester)	SU27B	M	M	O		O	O					
Gold River	Gold001						O				O	MR
Gold River	Gold002	M						O			O	M
Gold River	Gold003		M								MR	MR
Gold River	Gold004										MR	MR
Gold River	Gold005							O			MR	MR
Gold River	Gold007							O			MR	M
Gold River	Gold009										MR	MR
Gold River	Gold010							O			O	MR
Gold River	Gold011										O	O
Gold River	Gold012										O	MR
Gold River	Gold014											O

River	Site	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Gold River	Gold021							O				
LaHave River	LHav001					MR						
LaHave River	LHav002	MR		MR		O	O	O	O	O	O	O
LaHave River	LHav005	MR		MR		O						
LaHave River	LHav008	MR		MR	MR	MR	O	MR	MR		MR	MR
LaHave River	LHav009	MR										
LaHave River	LHav010	MR		MR								
LaHave River	LHav014	O										
LaHave River	LHav016	MR										
LaHave River	LHav022	MR										
LaHave River	LHav028	MR										
LaHave River	LHav031	MR				O	O					
LaHave River	LHav032	O										
LaHave River	LHav036	MR										
LaHave River	LHav039	MR										
LaHave River	LHav040	O										
LaHave River	LHav042	MR										
LaHave River	LHav048								MR			
LaHave River	LHav101	O			MR	MR	O	M	MR	MR	MR	MR
LaHave River	LHav102			MR	MR		O	MR	MR	MR	MR	MR
LaHave River	LHav103	MR		MR	MR		O	MR	MR	MR	O	MR
LaHave River	LHav104	MR		MR			O	O	O	O	O	O
LaHave River	LHav105	MR				MR	O	MR	MR	MR	MR	MR
LaHave River	LHav106	MR		MR	MR	MR	O	MR	MR	MR	MR	MR
LaHave River	LHav107	MR		MR	MR		O	MR	MR	MR	MR	MR
LaHave River	LHav108	MR		MR	MR	MR	O	MR	MR	MR	MR	MR
LaHave River	LHav109	MR		MR	MR	MR	O	M	MR	MR	MR	MR

River	Site	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
LaHave River	LHav110	MR		MR	MR	MR	O	MR	MR	MR	O	MR
LaHave River	LHav111	MR		MR	MR	MR	O	O	O	MR	O	O
LaHave River	LHav112	MR			MR	MR	O	M	MR		MR	MR
LaHave River	LHav113	MR		MR			O	M	MR		MR	MR
LaHave River	LHav114	MR		MR	MR	MR	O	M	MR		MR	MR
Liscomb	Lisc001			MR	O	O						
Liscomb	Lisc002			NA								
Liscomb	Lisc005			MR			MR					
Liscomb	Lisc007			O								
Liscomb	Lisc009			NA								
Liscomb	Lisc013			O								
Medway River	Medw101								O		O	
Medway River	Medw104										O	
Medway River	Medw108								MR			
Medway River	Medw123								O			
Medway River	SU20A						O					
Medway River	SU20B						O					
Medway River	SU20C						O		MR			
Middle (Chester)	Midd003	O	O									
Middle (Chester)	Midd004	M	M									
Middle (Chester)	Midd008	O	O									
Middle (Chester)	SU26A						O					
Middle (Chester)	SU26B						O					
Musquodoboit	MQ8811		MR			O						
Musquodoboit	MQ8812		MR			MR						
Musquodoboit	MQ8813.1					O						

River	Site	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Musquodoboit	MQ8815.1		MR									
Musquodoboit	MQ8819.2		MR									
Musquodoboit	MQ882					MR						
Musquodoboit	MQ8820.1		MR									
Musquodoboit	MQ8821.6		MR									
Musquodoboit	MQ883.1					O						
Musquodoboit	MQ884		MR			MR						
Musquodoboit	MQ888		MR			MR	O					
Musquodoboit	MQ8883		MR			MR						
Musquodoboit	SU40A						O					
Petite	SU21A						O					
Petite	SU21B						O					
Petite	SU21C						O					
Petite	SU21D						O					
Petite	SU21E						O					
Salmon (Guys. Co.)	SrGuy005		O					O				
Salmon (Guys. Co.)	SrGuy008		O					O				
Salmon (Guys. Co.)	SrGuy010		O					O				
Salmon (Guys. Co.)	SrGuy011		O					O				
Salmon (Guys. Co.)	SrGuy014		O					O				
Salmon (Guys. Co.)	SrGuy015							M				
Salmon (Guys. Co.)	SrGuy016		O					O				
Salmon (Guys. Co.)	SrGuy017		O					O				
Salmon (Guys. Co.)	SrGuy019		O					O				
Salmon (Port Dufferin)	SalPD001									O		
Salmon (Port	SalPD002									O		

River	Site	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Dufferin)												
Salmon (Port Dufferin)	SalPD004									O		
Salmon (Port Dufferin)	SalPD006						O			O	O	
Salmon (Port Dufferin)	SU50B						O			O		
St. Mary's River	STMR8510.2	MR	MR	O	MR	O	O	MR	O	O	O	
St. Mary's River	STMR8510.8	MR	MR	MR	O	MR	O	MR	MR			O
St. Mary's River	STMR853.2	MR		MR	MR	MR	O					
St. Mary's River	STMR854	MR	MR	MR	MR	MR	O	MR	MR	MR	MR	MR
St. Mary's River	STMR855.1	MR	MR	MR	MR	MR	O		MR	M	MR	MR
St. Mary's River	STMR857.1	MR										
St. Mary's River	STMR858.1	O	O	O	MR	O	O	O	O	O	O	O
St. Mary's River	STMR859.4			MR	MR	MR	O	O	MR	O	O	O
St. Mary's River	STMR863	MR		MR	MR	MR		MR	MR	MR	MR	
St. Mary's River	STMR864	MR	MR									
St. Mary's River	STMR865	MR										
St. Mary's River	STMR867	MR	MR	MR	MR	MR	O	MR	MR	MR	MR	MR
St. Mary's River	STMR922	MR										
St. Mary's River	STMR923	MR		MR				MR	MR	MR	O	MR
St. Mary's River	STMR924	MR		MR	MR	MR	O	MR	MR	MR	O	MR
St. Mary's River	STMR925.1+2			MR	MR	MR		MR	MR	MR	MR	MR
St. Mary's River	STMR928	MR	MR		MR	MR	O	MR	MR	MR	MR	
West (Sheet Harbour)	SU47	O	O	O	MR	O	O					
West (Sheet Harbour)	SU47C	MR			MR		O					

River	Site	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
West (Sheet Harbour)	WestSH002		MR	MR		MR						
West (Sheet Harbour)	WestSH005	O	O	MR	MR	MR						
West (Sheet Harbour)	WestSH007	O		O								
West (Sheet Harbour)	WestSH011		O	O	O							

Appendix 5-2. Details on spatial data layers and analyses

The spatial data layers used in these analyses were originally compiled and analyzed at a watershed scale in Bowlby et al. (2014) as well as in chapter four. All geographic measurements made from these spatial data used Universal Transverse Mercator projection, NAD83 Datum for Zone 20 north.

Watershed boundaries and flow network

Watershed boundaries and the associated flow network formed the basis for my analyses of the electrofishing data relative to the physical and geological characteristics of watersheds, as well as the extent of human impact in freshwater environments. Because my research was an extension of that presented in Bowlby et al. (2014), I decided to use the same spatial data layers to characterize watershed boundaries and the flow network. Thus, I used the Secondary Watershed Layer developed by the Nova Scotia Department of the Environment to describe watershed boundaries as well as a flow network (Bowlby et al. 2014) derived from 1:50,000 scale data (from the National Hydrographic Network) to characterize stream networks. Relative to the original NHN data, the flow network from Bowlby et al. (2014) included: (1) inferred arcs through watershed features that were large enough to be represented as polygons to ensure that these components of the stream network could be identified as parts of specific rivers, and (2) topological connections to represent the direction of water flow within the network to enable analyses requiring upstream or downstream accumulation.

Digital elevation model

I combined the individual raster data files available from GeoBase (20 m horizontal resolution and 1 m vertical resolution) into a Digital Elevation Model (DEM) for all of Nova Scotia using the merge tool in ArcGIS 10.2$^{©}$. For the Zonal statistic analyses of the site-specific or upstream buffers at each electrofishing site, I used the Spatial Analyst extension in ArcGIS® to calculate mean, maximum and minimum slopes (in degrees) from the DEM, as well as the standard deviation of slope within the buffered area to calculate topographic roughness.

Catchment land use and land cover

The Forest Inventory Data (Forest Inventory cycle 2 & 3, downloaded in November, 2011) classifies areas relative to numerical codes (FORNON codes) and an associated description for each type of land use. To reduce the number of categories considered in my analyses, I aggregated those codes into more general categories that would represent similar types of land cover or land use (Table A5-2-1, Appendix 5-2). Total areas of each type of activity in the site-specific or sub-catchment buffers were calculated from these new categories. I considered wetlands, brush, natural forest, rocky and dead to be natural cover types and agriculture, clearcutting, forestry (i.e. silviculture), urban, industry and corridors to be human-impacted areas.

Table A5-2-1. Aggregation of the FORNON code classifications for the land use analyses, giving the original FORNON code and description, as well as the category and variable names corresponding to Table 5-2 and the type (either natural or human-impacted). For the regression analyses, all areas within site-specific or sub-catchment buffers that were given a particular code were summed to calculate proportions (e.g. s.natfor would be the proportion of area classified as FORNON code ‘0’ in the site-specific buffer for a particular electrofishing site).

FORNON code	Description	New category	Type	Variable name
0	Natural stand, not treated, containing trees capable of reaching 3m at maturity	Natural forest	Natural	s.natfor natfor
71	Beaver flowage - an area occupied by beavers.	Wetland	Natural	
72	Open bogs - any area that is wet all year and has less than 25% live tree cover and primarily ericaceous plants or mosses.	Wetland	Natural	
73	Treed bogs - any area with primarily ericaceous plants or mosses with stunted softwood or hardwood species having 25% or more live tree cover.	Wetland	Natural	
74	Coastal habitat area - a wetland that lies in the ocean.	Wetland	Natural	
75	Lake wetland - a wetland that lies within freshwater (lake or river).	Wetland	Natural	
77	Inland waterbodies - lakes, rivers, reservoirs, canals, ponds.	Wetland	Natural	
78	Ocean	Wetland	Natural	
33	Brush - areas containing less than 25% merchantable tree cover and contains at least 25% cover of non-merchantable woody plants	Brush	Natural	
38	Alders - less than 75% crown closure of alders.	Brush	Natural	
39	Alders - 75% or more crown closure of alders.	Brush	Natural	
83	Brush - areas containing less than 25% merchantable tree cover and containing at least 25% cover of non-merchantable woody plants. Code being replaced with forested class.	Brush	Natural	
88	Alders - any area with less than 75% crown closure of alders, dry land only. Being replaced with forested class.	Brush	Natural	
89	Alders - 75% or more crown closure of alders, dry land only. Being replaced with a forested class.	Brush	Natural	

FORNON code	Description	New category	Type	Variable name
76	Cliffs, dunes or coastal rocks - areas of land between the high tide mark and the forest or non-forest stand.	Rocky	Natural	
84	Rock barren - areas covered by at least 50% exposed rock with less than 25% live tree cover.	Rocky	Natural	
85	Barren - any area with less than 25% live tree cover containing ericaceous vegetation with less than 50% rock outcrops and/or boulder cover and less than 50% other woody plant cover.	Rocky	Natural	
94	Beach - land between the normal water live and the forest or non-forest category.	Rocky	Natural	
2	Burn - completely destroyed by fire leaving less than 25% crown closure	Dead	Natural	
6	Wind throw - any stand where trees are pushed more than 45 degrees from vertical by wind action	Dead	Natural	
7	Dead - a stand containing dead trees with less than 25% crown closure of residual live material with dead material standing or laying on ground and no evidence of regeneration	Dead	Natural	
8	Dead - a stand containing dead trees with 25-50% crown closure of residual live material with dead material standing or laying on ground and no evidence of regeneration	Dead	Natural	
9	Dead - a stand containing dead trees with 51-100% crown closure of residual live material with dead material standing or laying on ground and no evidence of regeneration	Dead	Natural	
13	Dead - a stand with 26-50% of equivalent crown closure of dead material with evidence of regeneration. Equivalent crown closure being an estimate of what crown closure would be if the dead material were alive.	Dead	Natural	
86	Agriculture - any hay field, pasture, tilled crop or orchard that contains no merchantable species.	Agriculture	Human	s.agri agri
91	Blueberries - areas that appear to have been or are being used for blueberry production.	Agriculture	Human	s.agri agri
5	Old field - a field with indications of merchantable tree species growing in	Agriculture	Human	s.agri

FORNON code	Description	New category	Type	Variable name
	with less than 25% crown closure and less than 1m in height.			agri
60	Clear cut - a completely cut stand with any residuals making up less than 35% crown closure and little or no evidence of regeneration.	Clearcut	Human	s.clearcut clearcut
1	Treated - treatment not classified but not Christmas trees	Forestry	Human	s.forest forest
3	Christmas tree cultivation	Forestry	Human	s.forest forest
4	Sugar bush cultivation	Forestry	Human	s.forest forest
10	Research stand - treated stand which contains sample plots for the evaluation of response.	Forestry	Human	s.forest forest
11	Seed orchard - stands designated as areas reserved for seed production	Forestry	Human	s.forest forest
12	Treated stand - an area where silviculture activity has occurred and the specific treatment has been identified by field data.	Forestry	Human	s.forest forest
20	Plantation - a group of artificially established trees.	Forestry	Human	s.forest forest
61	Partial depletion - a cut stand where hardwood residuals make up 25% or more of crown closure; verified by photo interpreters or field data.	Forestry	Human	s.forest forest
62	Partial depletion, not verified - a stand identified from satellite imagery as a partial cut.	Forestry	Human	s.forest forest
92	Miscellaneous - non-forest area: old mill site, rifle range, tower site, observation site, lake shore bottom, quarry, mining activity, wharf, pier, causeway, dam, airstrips, etc.	Industry	Human	s.industry industry
93	Sanitary land fill - municipal garbage disposal.	Industry	Human	s.industry industry
95	Gravel pit - an active or non-active area used for extracting gravel.	Industry	Human	s.industry industry
96	Pipeline corridor - a 25 meter buffer around a defined linear feature of a gas	Corridor	Human	s.corridor

FORNON code	Description	New category	Type	Variable name
	or oil pipeline route.			corridor
97	Powerline corridor - a corridor of land with limited use due to powerlines.	Corridor	Human	s.corridor corridor
98	Road corridor - polygons of varying widths for roads.	Corridor	Human	s.corridor corridor
99	Rail corridor - 20 m polygons around active and abandoned rail lines.	Corridor	Human	s.corridor corridor

Appendix 5-3: Does sampling area influence our ecological understanding?

The functional relationship between anthropogenic ecological changes and biodiversity patterns partially depends on the spatial grain (i.e., the size of sampling units) and extent of sampling used in the study (Chase & Knight 2013, Barton et al. 2013). In a single-species context, it is possible that sampling considerations could exert similar effects on our ability to describe changes in density relative to landscape metrics. Thus, the amount of area sampled may be an important variable to control for when evaluating other hypotheses. One way to do this would be to use electrofishing sampling area (area.est) as a predictor in the optimal model presented in chapter five.

Incorporating area into the negative binomial mixed model introduces a mathematical relationship between a component of the response and the added predictor, in that area now appears on both sides of the regression equation (Blackburn & Gaston 1996, Brett 2004). This situation has long been recognized to have the potential to produce spurious correlations (*sensu* Pearson 1896). However, there has been considerable debate in the literature over the issue of spurious correlation as related to appropriate ecological inference (summarized in Brett 2004). Two main issues have been identified: (1) the relative contribution of measurement error to total variability in the shared parameter (e.g. Prairie & Bird 1989), and (2) unequal variance between parameters (e.g. Brett 2004).

In relation to (1): multiple authors suggest correlations between ${}^{Y}\!/_{X}$ and X can still be valid provided the shared measurement error term is likely to be small (Blackburn

& Gaston 1996, Prairie & Bird 1989, Brett 2004). I have no way of directly estimating measurement error in relation to sampling area for the electrofishing surveys. For sites at which sampling area was reported, measurement errors are likely to be small. Measurement errors could be more substantial for sites at which sampling area was not recorded and needed to be estimated, yet it is highly unlikely that measurement error is the main source of variation in sampling area. In relation to (2): simulation of linear regression between Y/X and X suggests that such regressions often produce spurious negative correlations, but also that their overall magnitude is strongly related to differences in variability (expressed as a CV) between the shared and non-shared terms (Brett 2004). In my dataset, the CV = 0.82 for census area and CV = 1.47 for salmon counts (CV calculated as the s.d./mean over all watersheds). This gives a ratio between the CV's of X and Y of 0.56. Overall variability in the counts is substantially higher than variability in the sampling areas. Relative to the analysis done by Brett (2004), this would suggest that any spurious correlation due to having a common term on both sides of the equation would be low.

Here, I have compared the optimal model from chapter five (hereafter called 'optimal') with one that incorporates electrofishing site area (area.est) as a predictor (hereafter called 'alternate'). I have evaluated the relative support for this increase in model complexity using model selection based on AIC (Johnson & Omland 2004) and AIC weights (Albanese et al. 2009) as in chapter five. I have also compared the parameter estimates that result from the alternate model in relation to the optimal model (Table A5-3-1, Appendix 5-3) and have plotted the partial regression fits (Figure A5-3-1, Appendix

5-3). There was very strong support for the alternate model on the basis of AIC and Akaike weights, with a difference of seven in AIC as compared to the optimal model (*c.f.* 3303 and 3303) and an Akaike weight of 0.97.

There was a very strong negative relationship between juvenile density and the stream area that was electrofished at each site in the alternate model (Figure A5-3-1, Appendix 5-3). For these types of data, densities are typically considered to be more indicative of relative abundance as compared to raw counts because the influence of area is standardized among sites (Schwartz & Seber 1999, Chase & Knight 2013). In addition to any potential spurious correlation, a strong negative relationship with area (slope = -1.05, s.e. = 0.34; Table A5-3-1, Appendix 5-3) could exist in situations where the amount of area being electrofished is not independent of the number of salmon being captured. In other words, small areas could have been sampled when juvenile salmon were abundant and larger areas may have only been fished when juveniles became scarcer. This could happen if electrofishing crews stopped fishing sooner if they were catching a lot of fish. Another explanation could be that the amount of area electrofished is largely dependent the position of the site in the watershed (i.e. dependent on a correlation between site.pos and area.est). This seems less likely for these data, given that the position of sites along the stream network is not strongly related to the areas electrofished (Kendall’s tau < 0.3; Figure A5-3-2, Appendix 5-3).

If changes in behaviour on the part of an electrofishing crew contribute to the apparent correlation between juvenile salmon density and sampling area, this could have unintended consequences on other ecological inferences from the analysis in chapter five.

For example, sites that had large areas that were electrofished tended to be those with moderately low proportions (e.g., 0.2-0.4) of many of the threats predictors, including the proportion of human use at the site and sub-catchment scales (s.phuman and phuman, respectively). Incorporating the sampled area (area.est) as a fixed effect in the alternate model moderately decreases the slope estimates for the quadratic relationship with human land use from the optimal model (*c.f.* s.phuman = 3.1 and 2.9; s.phuman$^{\wedge 2}$ = -3.9 and -3.6; Table A5-3-1, Appendix 5-3), and also slightly shifts the peak of the quadratic function (Figure A5-3-1, Appendix 5-3). It is worthwhile noting that this occurs irrespective of whether the correlation between juvenile density and sampling area is spurious. For future recovery planning, the results in this appendix suggest that observational data sources can contain unintended patterns that exert influence on the variables of interest (Prairie & Bird 1989, Rose 2000). New research using data collection methods specifically designed to evaluate population response relative to landscape changes could be a more powerful basis for inference.

Table A5-3-1. A comparison of coefficient estimates of the fixed effects from the optimal negative binomial Generalized Linear Mixed Model reported in chapter five with an alternate model incorporating sampling area of the electrofishing surveys as an additional predictor.

	Optimal		Alternate	
	Slope	s.e.	slope	s.e.
(Intercept)	2.07	0.30	2.07	0.29
year	-0.63	0.19	-0.51	0.19
s.degslope	0.07	0.04	0.10	0.04
site.pos	-1.90	0.42	-1.68	0.42
s.phuman	2.76	1.02	2.90	1.01
s.phuman$^{\wedge 2}$	-3.36	1.23	-3.59	1.23
phuman	2.54	0.78	2.69	0.79
area.est			-1.05	0.34

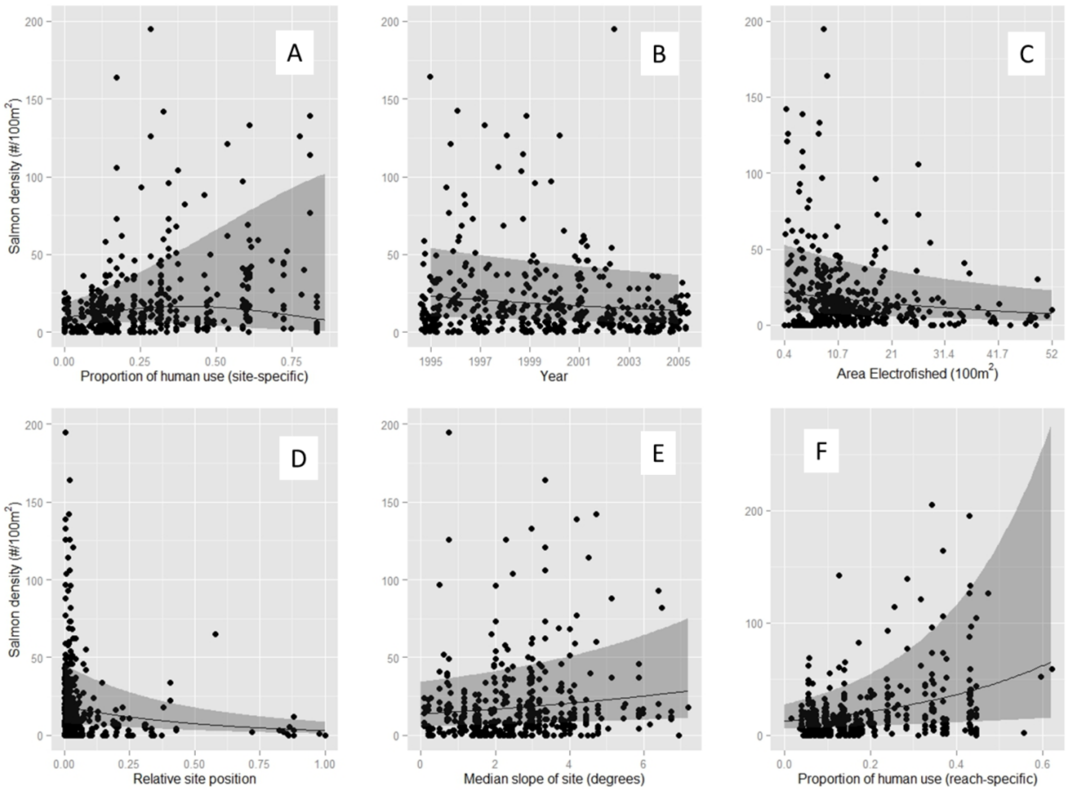

Figure A5-3-1. Partial regression fits (lines) plus 95% confidence interval (grey shading) relative to the juvenile density data (points) for the fixed effects of the negative binomial GLMM incorporating sampling area as a predictor. X-axes are as follows: panel A: site-specific proportion of human use (s.phuman+s.phuman^2), panel B: year of electrofishing survey (year), panel C: area electrofished (area.est), panel D: relative site position (site.pos), panel E: median slope of a site (s.degslope) and panel F: proportion of human use within a sub-catchment (phuman). Multiple observations from the same year are off-set slightly along the x-axis in panel B in order to see all of the individual density estimates. See Table 5-2 for parameter definitions.

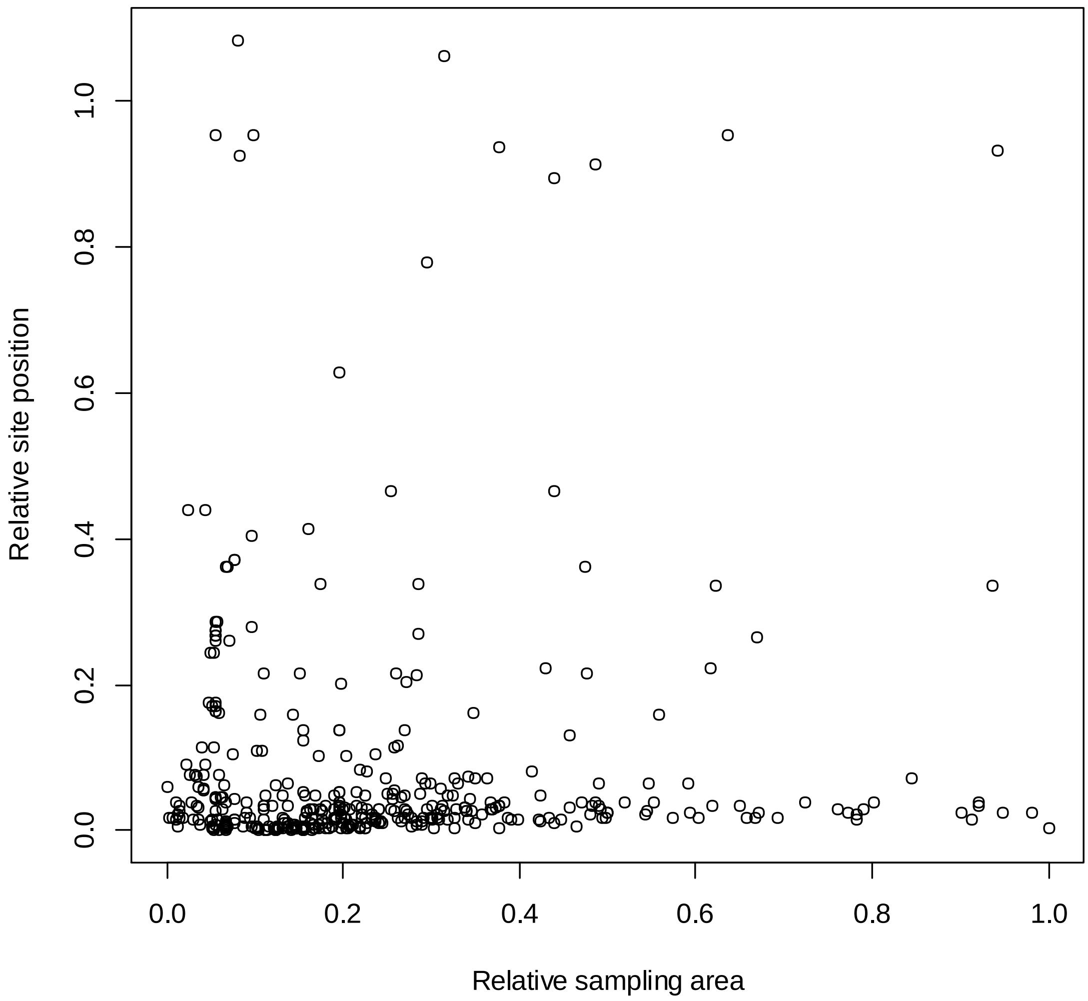

Figure A5-3-2. A plot of site positions in a watershed relative to the amount of area sampled during electrofishing surveys for the data analyzed here and in chapter five.