Sea lamprey (*Petromyzon marinus*) habitat and population models in Michigan river networks: Understanding geomorphic context and boundaries

by

Thomas Matthew Neeson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Natural Resources and Environment)
in The University of Michigan
2010

Doctoral Committee:

Associate Research Scientist Sara A. Adlerstein, Co-Chair
Professor Michael J. Wiley, Co-Chair
Professor Steven J. Wright
Associate Research Scientist Rick L. Riolo
Adjunct Associate Research Scientist Edward S. Rutherford

Acknowledgements

I first wish to thank my advisors and co-chairs Drs. Michael Wiley and Sara Adlerstein. They have been mentors, guides, critics and patient editors of my work.

My committee members Drs. Rick Riolo, Edward Rutherford, and Steven Wright have always offered helpful feedback and a broader perspective on my research. Rick Riolo deserves special thanks for guidance on all things computer modeling.

I have been surrounded by a diverse and talented group of fellow graduate students. I particularly wish to thank students who were faithful participants in the Adlerstein/Wiley lab group meetings, the Hooper Seminar series, and the Complex Systems Advanced Academic Workshop series. I presented various pieces of preliminary research at each of these venues, and the commentary from other students was invaluable. I learned equally much from listening to students present their own research.

Beth Sparks-Jackson and Lori Ivan reviewed drafts of Chapters 2 and 3.

Emily Silverman of the USFWS provided important guidance during early stages of this research.

Jeff Slade of the USFWS provided access to and helpful interpretation of both the habitat and larval fish survey data.

Computing resources for Chapters 3 and 4 were provided by the Center for the Study of Complex Systems at the University of Michigan.

This project was partially funded by The Great Lakes Fishery Commission and The University of Michigan Rackham Summer Research Grant. The Paul Welch Fellowship in Limnology allowed me to prioritize research for the first four years of my graduate program. I was supported in my fifth and final year by two graduate student instructor positions from the School of Natural Resources and Environment.

A unique note of thanks must go to my wife Molly. Her support and perspective always kept me moving forward.

Table of Contents

List of Figures

List of Tables

Abstract

Rivers are hierarchical, heterogeneous products of climate and the landscapes through which they flow. This nature motivates two interesting and practical questions in river ecology: how are populations of riverine organisms and their habitat influenced by their geomorphic context, and how best should spatial boundaries for conservation and management be defined?

I address both questions with respect to the sea lamprey (*Petromyzon marinus*), an invasive species in this region with a history of negative impacts on Great Lakes fisheries. High management costs (> $14 million/year) provide a strong practical motivation for developing a deeper understanding of sea lamprey ecology.

In a first analysis, I built a series of regression models to predict larval sea lamprey habitat across the lower peninsula of Michigan. I used the geomorphological concept of process domains as a justification for using regression trees to both stratify the dataset and fit a regression model to each stratum. Trees identified useful partitions of the lower peninsula of Michigan and the importance of geomorphic influences on habitat distribution.

In a second analysis, I developed an individual-based model (IBM) to explicitly test the hypothesis that river network structure influences the distribution of larval sea lamprey. In the model, patterns of larval aggregation varied strongly with network structure across a broad range of parameterizations. My results suggest that larval distribution across a watershed results from a system of emergent feedbacks shaped by the network structure of each watershed and the past migratory and spawning behavior of adults.

In a final analysis, I expanded the IBM model to include heterogeneous distributions of larval habitat, and explored how habitat availability, network structure, and fish migratory behavior interact to shape larval distributions. Using this model I demonstrate how and when larval distribution could be influenced by both network structure and the distribution of larval habitat.

My IBM results suggest that the watershed is an important functional boundary for sea lamprey. I argue for a view of each watershed as a unique place, where larval distribution depends on geomorphic structures that alter feedbacks between larval habitation and adult migratory choices.

Chapter 1

Introduction

Rivers are hierarchical, heterogeneous products of climate and the landscapes through which they flow. They are heterogeneous and dynamic in that they exhibit spatial and temporal variability in flow, chemistry, temperature, and substrate (Giller and Malmqvist 2004). They are hierarchical in that this variability occurs at a series of nested scales, e.g. among faces of substrate particles, among microhabitats, among channel elements such as pools and riffles, among reaches, and among large sections of river networks (Frissell et al. 1986). All of this variability is a complex function of characteristics of the stream's valley (Hynes 1975), including landuse (Milliman et al.1987, Roth et al. 1996, Wang et al. 1997), surficial geology (Whittier et al. 1988), hydrologic processes (Bagnold 1980, Whiting et al. 1999), and human alterations such as dams (Williams and Wolman 1984, Meade and Parker 1985). In glaciated landscapes such as Michigan, each watershed itself occupies a unique physiographic context within a regional landscape of geology and climate.

Rivers' hierarchical, heterogeneous, and context-dependent nature is of fundamental importance for fluvial ecologists. The effects of flow are pervasive on riverine organisms (Hart and Finelli 1999), with the implication that organisms' distribution and population dynamics cannot be understood apart from the physical context of the particular river that they inhabit. Habitat in rivers, whether defined in terms of temperature, chemistry, or substrate, is influenced by landscape and hydrology and results from the interplay of processes at several scales (Wiley et al. 1997, Knighton 1998, Giller and Malmqvist 2004).

In this dissertation I attempt to address two basic questions in fluvial ecology, both of which are motivated by the hierarchical, heterogeneous, and context-dependent nature of rivers. First, I explore how substrate habitat and organisms with a long benthic stage are influenced by geomorphic context. For habitat, this is a question of understanding the linkages between landscape, hydrologic processes, and geomorphic processes, the best scales for defining these linkages, and how these linkages vary over space and time. For organisms, this is a question of understanding how specific facets of a species' life history interface with each of the multiple scales of spatial structure in rivers (Schlosser 1991, 1995).

Second, I address the problem of defining boundaries for conservation and management. Restricting the scope of analysis is prerequisite to all work in conservation biology (Post et al. 2006), yet no clear rules exist for defining the boundaries of geomorphic units (Montgomery et al. 1995), populations (Berryman 2002, Camus and Lima 2002), or ecosystems (Keller and Goley 2000). When multiple potential methods for the delineation of spatial units and their boundaries exist, the choice of boundaries has important operational consequences for sampling design, modeling studies, and the identification of units for conservation and management; Given the hierarchical nature of rivers, best boundaries are likely problem-specific. In the context of predictive habitat models, boundaries would ideally delineate geomorphic process domains (Whiting and Bradley 1993, Montgomery 1999, Brardinoni and Hassan 2006) or hydrologic response units (Becker and Pfutzner 1986, Becker 1995, Maidment and Ximing Cai 1996, Becker and Braun 1999). Spatial and temporal boundaries for population analyses would ideally coincide with units within which individuals strongly interact (Wiens 1989, Post et al. 2006), or be structural boundaries that constrain the migration or genetic exchange of organisms (e.g. watersheds; Wiens 2002, Strayer et al. 2003, Post et al. 2006).

In this dissertation, I address both of the above questions in the context of sea lamprey (*Petromyzon marinus*) as a case study. The sea lamprey is an invasive species in the Laurentian Great Lakes with a long history of negative impacts on salmonines (Smith and Tibbles 1980), particularly lake trout (*Salvelinus namaycush,* Coble et al. 1990). High management costs (> $14 million/year; Slade et al. 2003, Christie et al 2003, Jones 2007) provide a strong practical motivation for developing a deeper understanding of sea lamprey ecology. In the remainder of this Introduction, I give an overview of sea lamprey life history, their history as an invasive species in the Great Lakes, and an overview of the contents of this dissertation.

Sea lamprey life history

The sea lamprey (*Petromyzon marinus*) is a primitive fish with an anadromous, semelparous lifecycle characterized by a stream-resident, filter-feeding larval stage generally three to five years in duration, an ocean- or lake-dwelling juvenile parasitic stage, and a non-feeding sexually mature adult stage during which lamprey migrate from oceans or lakes into rivers to spawn. Adults migrate upstream to spawn in the spring when water temperatures rise to 3 to 4C. Spawning runs reach their peak near 10C, and migration continues until temperatures exceed 18C (Morman et al. 1980).

Adult sea lamprey follow two pheromones during their spawning migration. First, both male and female spawning-phase lamprey are attracted by a pheromone released by stream-resident larvae (Li et al. 1995, Bjerselius et al. 2000, Vrieze and Sorensen 2001, Sorensen and Vrieze 2003). Adult females within the stream also follow a sexual pheromone released by spermiated males who have recently completed the upstream migration (Li et al. 2003). The larval pheromone controls broad-scale migration through the watershed; the sexual pheromone allows females to find nesting males.

Spawning adults eventually construct nests of gravel in areas with steady, unidirectional flow (Morman et al. 1980), and females each deposit several tens of thousands of eggs into a nest. Adults die after spawning. The eventual number of larvae, "ammocoetes", leaving the nest generally represents less than one percent of the number of eggs produced by each female (Applegate 1950).

Ammocoetes remain in the stream for three to seventeen or more years, residing in depositional areas characterized by a soft, burrowable substrate of fine sands, low current velocity, and some amount of organic detritus (Applegate 1950, Malmqvist 1980, Potter et al. 1986). While burrowed, ammocoetes feed primarily on algae, detritus, and bacteria (Moore and Potter 1976, Sutton and Bowen 1994, Yap and Bowen 2003).

After completing the sedentary burrowing stage, ammocoetes undergo a metamorphosis into the parasitic phase of their life history. This metamorphosis typically occurs during late summer and fall, and consists of several distinct morphological changes, including the development of a circular sucking disc, numerous horny teeth, and full development of an eye. After completing metamorphosis, lamprey migrate downstream to begin the parasitic phase of their life. This migration occurs between September and May, and often coincides with peak stream flow periods (Applegate 1950).

During the parasitic phase, lamprey attach themselves to host fishes and feed on the blood and bodily fluids of their host, causing significant mortality. The parasitic phase lasts between 12 and 20 months in freshwater environments (Applegate 1950), but 23 to 28 months in marine environments (Beamish 1980). During the parasitic phase, each lamprey is capable of killing 40 pounds of fish (GLFC 2000). Following the parasitic phase, the juvenile lamprey matures sexually and migrates upstream to spawn.

Sea lamprey in the Great Lakes

The sea lamprey is an anadromous parasitic fish native to the Northern Atlantic and East coast of the United States (Beamish 1980). Christie (1973) argued that sea lamprey are likely native to Lake Ontario as well, while Aron and Smith (1971) suggest that sea lamprey may have gained access to Lake Ontario through the Erie Canal. All researchers agree that sea lamprey are not native to Lake Erie or the Upper Great Lakes.

Invasion of the Great Lakes beyond Lake Ontario occurred after enlargement of the Welland Canal in 1919. The first confirmed specimen was reported in Lake Erie in 1921. Specimens were recorded in Lake Michigan in 1936, Lake Huron in 1937, and Lake Superior in 1938, suggesting that lamprey spread through the Great Lakes relatively rapidly (Smith and Tibbles 1980). Early records documenting the invasion and spread of the sea lamprey are sparse, as the sea lamprey was initially considered a public curiosity and there was little recognition of its potential influence on fisheries.

The effect of sea lamprey predation on Great Lakes fisheries was severe. Lake trout (*Salvelinus namaycush*) were eliminated from Lake Michigan and nearly eliminated in Lakes Huron and Superior. Drastic reductions in the population size of burbot (*Lota lota*), rainbow trout (*Salmo gairdneri*), and lake whitefish (*Coregonus artedii*) are also attributed to sea lamprey (Farmer 1980). Sea lamprey predation had indirect effects on populations of many other species as well. As lake trout and burbot declined, their prey, ciscoes (*Coregonus* spp.), increased dramatically. As sea lamprey predation on ciscoes increased and ciscoe populations declined, alewife (*Alosa pseudoharengus*) populations increased. Alewife's success has had adverse effects on many other species such as yellow perch (*Perca flavescens*), emerald shiner (*Notropis atherinoides*), and rainbow smelt (*Osmerus mordax*) (Smith 1968).

Sea lamprey populations in the Great Lakes are controlled primarily with a chemical lampricide that targets the sedentary larval stage (Slade et al. 2003, Christie et al 2003, Jones 2007). Lampricide control in the Great Lakes has been very successful, and is credited with reducing catches of spawning sea lampreys in Lake Superior to 8% of the precontrol average by 1978 (Smith and Tibbles 1980). Lampricide application is also expensive and time consuming, and a limited budget restricts the number of streams that can be treated annually. Streams are currently ranked for treatment priority using the Empiric Stream Treatment Ranking (ESTR), which considers treatment cost, predicted treatment effectiveness, and the projected number of juvenile sea lamprey that each river system would produce (Christie et al. 2003).

Alternatives to lampricide-based management include several potential strategies that utilize both larval and adult pheromones (Jones 2007). Teeter (1980), one of the first authors to suspect the existence of larval pheromone, noted that pheromone-based management had been successfully used to control a number of insect pests. More recent and thorough reviews of the various potential management uses of larval pheromone are given by Sorensen and Vrieze (2003) and Twohey et al. (2003).

Dissertation Contents

In Chapter 2, I built a series of regression models to predict larval sea lamprey habitat in 43 watersheds across the lower peninsula of Michigan. Given the large spatial extent of the dataset and the nature of sediment transport, I expected that the relationships between landscape covariates and instream habitat would be complex (Langbein and Schumm 1958, Shreve 1979, Madej and Ozaki 1996, Seelbach et al. 1997, Phillips 2003 and 2006). A key methodological question was how best to partition the study area such that the relationship between landscape and riverine habitat within each partition would be simpler and more amenable to regression modeling than the spatially complex global relationship. I first began with a null model without partitioning the dataset, and then developed a second set of models where I partitioned the dataset by watershed. I next used the concept of process domains (Whiting and Bradley 1993, Montgomery 1999, Brardinoni and Hassan 2006) as a motivation for using regression trees to identify partitions of the dataset and fit a unique linear regression model to each domain. The goals of this chapter were to compare several methods for partitioning a regression model between landscape and habitat, and to understand the key geomorphic influences on the distribution of ammocoete habitat, both overall and within identified stratifications.

In Chapter 3, I developed an individual-based computational model (IBM) to directly test the hypothesis that river network structure influences the distribution of larval sea lamprey across a watershed. I used this model to perform a series of computational experiments to explore the relationship between river network structure, model parameterization, and larval distribution over both short and long timescales. I also used a river network based on the Muskegon River, Michigan as a validation test of the general model, and compared model output for this river network with existing larval survey data. Chapter 3 constitutes both a freestanding research contribution and prerequisite material for Chapter 4.

In Chapter 4, I expanded the IBM model developed in Chapter 3 to include spatial heterogeneity in the distribution of larval habitat. Patterns of habitat distribution in this chapter were motivated by observed habitat distributions in Michigan rivers (Chapter 2). I performed a series of computational experiments to examine how habitat, river network structure, and fish behavior interact to determine patterns of larval distribution. Specifically, I measured the correlation between distributions of larvae and habitat, the effect of geomorphic structures on total larval population size, and how the distribution of larval habitat could mediate the population dynamics observed in Chapter 3. I discussed how concepts advanced by the model analysis might supplement a range of management activities. I then used the results of the modeling experiments to draw inferences about the best spatial boundaries for describing larval lamprey distribution and abundance over time.

All three chapters, although diverse, are motivated by a desire to better understand physical and biological processes in rivers, how such processes are influenced by geomorphic structures, and how best to draw boundaries when processes are multi-scale and spatially continuous. In Chapter 2, I sought to understand how the geomorphic structure of the landscape influences instream substrate habitat across process domains, and to identify the boundaries of these process domains in order to partition regression models of larval habitat. In Chapters 3 and 4, I examined how two specific geomorphic structures (habitat distribution and network shape) influence the biological processes that control larval distribution, and used my modeling results to infer best spatial boundaries for describing larval distribution.

References

Applegate, V.C. 1950. Natural history of the sea lamprey (*Petromyzon marinus*) in Michigan. Spec. Sci. Rept. U.S. Fish and Wildlife Serv. 55, 1-237.

Aron, W.I. and Smith, S.H. 1971. Ship canals and aquatic ecosystems. Science 174: 13-20.

Bagnold, R.A. 1980. An empirical correlation of bedload transport rates in flumes and natural rivers. Proceedings of the Royal Society 372A: 453-73.

Beamish, F.W.H. 1980. Biology of the North American anadromous sea lamprey, *Petromyzon marinus.* Can. J. Fish. Aquat. Sci. 37: 1924-1943.

Becker, A. and Pfutzner, B. 1986. Identification and modeling of river flow reductions caused by evapotranspiration losses from shallow groundwater areas. IAHS Publication No. 156.

Becker, A., and Braun, P. 1999. Disaggregation, aggregation, and spatial scaling in hydrological modeling. Journal of Hydrology 217: 239-252.

Berryman, A.A. 2002. Population regulation, emergent properties, and a requiem for density dependence. Oikos 99: 600-606.

Bjerselius, R., Li, W., Teeter, J.H., Seelye, J.G., Johnsen, P.B., Maniak, P.J., Grant, G.C., Polkinghorne, C.N., and Sorensen, P.W. 2000. Direct behavioral evidence that unique bile acids released by larval sea lamprey (*Petromyzon marinus*) function as a migratory pheromone. Canadian Journal of Fisheries and Aquatic Sciences 57: 557 - 569.

Brardinoni, F., and Hassan, M.A. 2006. Glacial erosion, evolution of river long profiles, and the organization of process domains in mountain drainage basins of coastal British Columbia. Journal of Geophysical Research 111: F01013.

Camus, P.A., and Lima, M. 2002. Population, metapopulations, and the open-closed dilemma: The conflict between operational and natural population concepts. Oikos 97: 433-438.

Christie, G.C., Adams, J.V., Steeves, T.B., Slade, J.W., Cuddy, D.W., Fodale, M.F., Young, R.J., Kuc, M., and Jones, M.L. 2003. Selecting Great Lakes streams for lampricide treatment based on larval sea lamprey surveys. J. Great Lakes Res. 29: 152-160.

Christie, W.J. 1973. A review of the changes in fish species composition in Lake Ontario. Great Lakes Fish. Comm. Tech. Rep. 23: 65pp.

Coble D.W., Brueseqitz R.E., Fratt T.W., Scheirer J.W. 1990. Lake trout, sea lampreys, and overfishing in the upper Great Lakes: A review and reanalysis. Transactions of the American Fisheries Society 119: 985–995.

Farmer, G.J. 1980. Biology and physiology of feeding in adult lampreys. Can. J. Fish. Aqua. Sci. 37: 1751-1761.

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

Giller, P.S. and Malmqvist, B. 2004. The biology of streams and rivers. Oxford University Press, New York City, New York.

Hart, D.D., and Finelli, C.M. 1999. Physical-biological coupling in streams: the pervasive effects of flow on benthic organisms. Annual Review of Ecology and Systematics 30: 363-395.

Hynes, H.B. 1975. The stream and its valley. Internationale Vereinigung fir theoretische und angewandte Limnologie 19: 1-15.

Jones, M.L. 2007. Toward improved assessment of sea lamprey population dynamics in support of cost-effective sea lamprey management. J. Great Lakes Res. 33: 35-47.

Keller, D.R. and Golley, F.B. 2000. The philosophy of ecology: from science to synthesis. University of Georgia Press, Athens Georgia.

Knighton, D. 1998. Fluvial forms and processes: a new perspective. John Wiley and Sons, New York, New York.

Langbein, W.B. and Schumm, S.A. 1958. Yield of sediment in relation to mean annual precipitation. Transactions of the American Geophysical Union 39: 1076-1084.

Li, W., Sorensen, P.W., and Gallaher, D.D. 1995. The olfactory system of migratory adult sea lamprey (*Petromyzon marinus*) is specifically and acutely sensitive to unique bile acids released by conspecific larvae. The Journal of General Physiology 105: 569 - 587.

Li., W., Siefkes, M.J., Scott, A.P., and Teeter, J.H. 2003. Sex pheromone communication in the sea lamprey: implications for integrated management. J. Great Lakes Res. 29 (Suppl. 1): 85 - 94.

Li, W., Twohey, M., Jones, M., and Wagner, M. 2007. Research to guide use of pheromones to control sea lamprey. J. Great Lakes Res. 33 (Special Issue 2): 70 - 86.

Madej, M.A. and Ozaki, V. 1996. Channel response to sediment wave propagation and movement, Redwood Creek, California, USA. Earth Surface Processes and Landforms 21: 911-927.

Malmqvist, B. 1980. Habitat selection of larval brook lampreys (*Lampetra planeri*, Bloch) in a south Swedish stream. Oecologia 45:35-38.

Meade, R.H. and Parker, R.S. 1985. Sediment in rivers in the United States. In *National Water Summary 1984,* United States Geological Survey Water-Supply Paper 2275: 49-60.

Milliman, J.D., Yun-Shan, Q., Mei-E, R., and Saito, Y. 1987. Man's influence on the erosion and transport of sediment by Asian rivers: the Yellow River (Huanhe) example. Journal of Geology 95: 751-762.

Montgomery, D.R. 1999. Process domain and river continuum. Journal of the American Water Resources Association 35: 397-410.

Montgomery, D.R., Grant, G.E., and Sullivan, K. 1995. Watershed analysis as a framework for implementing ecosystem management. Water Resources Bulletin, 31: 369 - 386.

Moore, J.W. and Potter I.C. 1976. A laboratory study on the feeding of larvae of the brook lamprey *Lampetra planeri* (Bloch). J. Anim. Ecol. 45: 81-90.

Morman, R.H., Cuddy, D.W., and Ruger, P.C. 1980. Factors influencing the distribution of sea lampreyy (*Petromyzon marinus*) in the Great Lakes. Can. J. Fish. Aquat. Sci. 37: 1811 - 1826.

Phillips, J.D. 2003. Sources of nonlinearity and complexity in geomorphic systems. Progress in Physical Geography 27: 1-23.

Phillips, J.D. 2006. Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change. Hydrol. Earth Syst. Sci. 10: 731-742.

Post, D.M., Doyle, M.W., Sabo, J.L., and Finlay, J.C. 2006. The problem of boundaries in defining ecosystems: A potential landmine for uniting geomorphology and ecology. Geomorphology 89: 111 - 126.

Potter, I.C., Hillard, R.W., Bradley, J.S., and McKay, R.J. 1986. The influence of environmental variables on the density of larval lampreys in different seasons. Oecologia 70: 443-440.

Puth, L.M. and Wilson, K.A. 2001. Boundaries and corridors as a continuum of ecological flow control: lessons from rivers and streams. Conservation Biology 15: 21-30.

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

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

Schlosser, I.J. 1995. Critical landscape attributes that influence fish population dynamics in headwater streams. Hydrobiologia 303: 71 - 81.

Schlosser, I.J. 1995. Dispersal, boundary processes, and trophic-level interactions in streams adjacent to beaver ponds. Ecology 76: 908 - 925.

Seelbach P. W., M. J. Wiley, J. C. Kotanchik, and M. E. Baker. 1997. A landscape-based ecological classification system for river valley segments in lower Michigan (MI-VSEC Version 1.0). Michigan Department of Natural Resources, Fisheries Division, Fisheries Research Report 2036, Lansing.

Shreve, R.L. 1979. Models for prediction in fluvial geomorphology. Mathematical Geology 11: 165-173.

Slade, J.S. Adams, J.V., Christie, G.C., Cuddy, D.W., Fodale, M.F., Heinrich, J.W., Quinlan, H.R., Weise, J.G., Weisser, J.W., and Young, R.J. 2003. Techniques and methods for estimating abundance of larval and metamorphosed sea lampreys in Great Lakes tributaries, 1995 to 2002. J. Great Lakes Res. 29: 137-151.

Smith, B.R. and Tibbles, J.J. 1980. Sea lamprey (*Petromyzon marinus*) in Lakes Huron, Michigan, and Superior: history of invasion and control, 1936-78. Can. J. Fish. Aquat. Sci. 37: 1780-1801.

Smith, S.H. 1968. Species succession and fishery exploitation in the Great Lakes. J. Fish. Res. Board Can. 25: 667-693.

Sorensen, P.W., and Vrieze, L.A. 2003. The chemical ecology and potential application of the sea lamprey migratory pheromone. J. Great Lakes Research. 29 (Suppl. 1): 66 - 84.

Sutton, T.M. and Bowen, S.H. 1994. Significance of organic detritus in the diet of larval lampreys in the Great Lakes basin. Can. J. Fish. Aquat. Sci. 51: 2380-2387.

Teeter, J. 1980. Pheromone communication in sea lampreys (*Petromyzon marinus*): implications for population management. Canadian Journal of Fisheries and Aquatic Sciences 37: 2123-2132.

Twohey, M.B., Sorensen, P.W., and Li, W. 2003. Possible applications of pheromones in an integrated sea lamprey management program. J. Great Lakes Res. 29: 794 - 800.

Vrieze, L.A. and Sorensen, P.W. 2001. Laboratory assessment of the role of a larval pheromone and natural stream odor in spawning stream localization by migratory sea lamprey (*Petromyzon marinus*). Canadian Journal of Fisheries and Aquatic Sciences 58: 2374 - 2385,

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

Whiting, P.J. and Bradley, J.B. 1993. A process-based classification system for headwater streams. Earth Surface Processes and Landforms 18: 603-612.

Whiting, P.J, J.F. Stamm, D.B. Moog, and R.L. Odndorff. 1999. Sediment-transporting flows in headwater streams. Bulletin of the Geological Society of America 111(3): 450-466.

Whittier, T.R., Hughes, M.R., and Larsen, D.P. 1988. Correspondence between ecoregions and spatial patterns in stream ecosystems in Oregon. Canadian Journal of Fisheries and Aquatic Sciences 45: 1264-1278.

Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology 3: 385 - 397.

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

Wiley, M.J., Kohler, S.L., and Seelbach, P.W. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. Freshwater Biology 37: 133-148.

Williams, G.P. and Wolman, M.G. 1984. Downstream effects of dams on alluvial rivers. United States Geological Survey Professional Paper 1286.

Yap, M.R., and Bowen, S.H. 2003. Feeding by northern brook lamprey (*Ichthyomyzon fossor*) on sestonic biofilm fragments: habitat selection results in ingestion of a higher quality diet. J. Great Lakes Res. 29: 15-25.

Chapter 2

Towards a Process Domain-Sensitive Sea Lamprey Habitat Model for Michigan Rivers

Abstract

Habitat mapping is a common and often useful tool in fisheries and other kinds of ecological management in rivers. For species whose preferred habitats are defined primarily by substrate particle size, the complex nature of fluvial geomorphology makes it difficult to create geographically extensive process-informed statistical models that predict habitat. Geomorphic processes are characterized by thresholds, interactions across temporal and spatial scales, temporal and spatial lags between causes and effects, and spatial variation in climate and physiography, all of which are difficult to formulate in a statistical model. An option is to partition the dataset into strata within which the globally complex relationship between geomorphology and habitat can be locally approximated by a simpler model. Such partitions correspond to process domains in geomorphology, areas in which one or several geomorphic processes are the dominant influence on substrate. This partitioning is a preliminary modeling step to identify useful strata and covariates for future studies.

In this study, I used regression trees as a machine learning method for partitioning and identifying useful strata in a larval sea lamprey (*Petromyzon marinus*) habitat model in 43 watersheds in the lower peninsula of Michigan. I used data from over 5000 habitat transects collected by United States Fish and Wildlife Service field agents, and created a geographic database of GIS-derived covariates that represent the principal geomorphic influences on habitat: stream power, low-flow yield, distance to the mouth of the river, distance to the nearest upstream dam or lake, distance to the nearest downstream dam or

lake, and land use and surficial geology in riparian corridor. I created three trees, where tree splits delineated: (1) spatially contiguous units, (2) non-contiguous units defined by values of the covariates, and (3) both contiguous and non-contiguous units. The adjusted R^2 values of the three trees were 0.30, 0.30, and 0.32, respectively, and all three trees outperformed a single model fit to the entire dataset and a set of models fit to each watershed individually. The trees identify useful stratifications of the lower peninsula of Michigan, the important geomorphic influences on habitat, and how the influence of geomorphology on habitat varies across my study region. I discuss the relevance of my findings for sea lamprey management in the Laurentian Great Lakes.

Introduction

Reach-scale maps of riverine species' preferred habitats play a variety of useful roles in conservation and management (Lunetta et al. 1997, Seelbach et al. 1997, Delong and Brusven 2006, Brenden et al. 2008). For some species, key dimensions of their habitat are largely predictable from landscape-scale covariates (e.g. temperature and flow; Zorn et al. 2002, Brenden et al. 2008). For such species it is possible to create statistical models that predict reach-scale habitat from geographic datasets of landscape-scale covariates (Wang et al. 1997, Wehrly et al. 2009), or that implicitly model habitat by directly linking landscape characteristics to fish distribution (Wang et al. 2007, Steen et al. 2008). Species whose preferred habitats depend strongly on site-specific substrate characteristics, however, present a unique challenge.

The distribution of specific types of substrate in a reach is controlled by the local energy distributions and sediment budget (Knighton 1998). Efforts to predict reach-scale sediment distributions across large geographic areas must rely on the tacit assumption that the local energy distributions and sediment budget are correlated with landscape-scale covariates. However, the relationship between energy distributions, sediment budget, and landscape-scale covariates is characterized by strong thresholds (Phillips 2003, 2006), non-additive and higher-order interactions between several ongoing processes (Shreve 1979), and temporal and spatial lags between causes and effects (Madej and Ozaki 1996). Habitat models with a large spatial extent may be further influenced by variation in geography and physiography (Seelbach et al. 1997), latent climate effects (Langbein and Schumm 1958), and non-causal (spurious) correlations among covariates.

Systems with strong thresholds and complicated interactions often exhibit multiple operating regimes (Eikens and Karim 1999). In geomorphology, this idea is referred to as the Process Domain Concept (Montgomery 1999). Each process domain is an area or region within which a limited subset of geomorphic processes dominate fluvial processes and instream habitat (Whiting and Bradley 1993, Brardinoni and Hassan 2006). Channels identified as belonging to the same process domain should therefore behave similarly with respect to the influence of landscape on instream substrate (Whiting and Bradley 1993). A key consequence is that within each process domain, the globally complex relationship between landscape and instream habitat may be approximated by a simpler (e.g. more linear) relationship between habitat and some subset of landscape covariates. The ability to identify process domains, and to partition large datasets accordingly, transforms the task of developing a single complicated model into the task of developing a suite of simpler models, each involving fewer covariates. A similar rationale underlies the delineation of hydrologic response units (HRUs) in hydrological modeling (Becker and Pfutzner 1986, Becker 1995, Maidment and Ximing Cai 1996, Becker and Braun 1999).

The multi-modeling framework provides a useful conceptual strategy for constructing a statistical model of a system hypothesized to span multiple process domains (Eikens and Karim 1999). The model-fitting process consists of two steps: first, the dataset must be partitioned on boundaries of the various process domains; second, a statistical model relating regressors to response is fit to the subset of datapoints in each domain. The resulting model is potentially useful both quantitatively, as a predictive model, and qualitatively, as a method for identifying useful stratifications of a large dataset or geographic region. Even in cases where the resulting domain models have low predictive power, the stratifications identified in the process of partitioning the dataset may still inform future modeling efforts, sample designs, and identification of units for conservation and management.

Identifying the boundaries of process domains in a large and noisy multivariate dataset may be challenging. Previous efforts to identify process domains in fluvial studies have focused on a small geographic area (Brardinoni and Hassan 2006) or type of stream (Whiting and Bradley 1993), where process domains were identified either by field work or reasoning from first principles. Neither method for identifying process domains is possible when constructing a spatially-extensive habitat model spanning physiographic regions. The class of statistical approaches informally known as "machine-learning" methods may be useful for discovering the boundaries of process regimes. In particular, classification and regression trees (CART) provide a method for completing both steps of the multi-modeling procedure outlined by Eikens and Karim (1999): 1) partitioning a dataset on the basis of differences in the relationship between instream habitat and landscape, and 2) finding a best-fitting model for each partition.

In this study, I was interested in exploring the use of CART methods to create a domain-sensitive predictive model of the distribution of larval sea lamprey (*Petromyzon marinus)* habitat in streams in the Lower Peninsula of Michigan. The sea lamprey is an invasive and parasitic fish species in the Laurentian Great Lakes with a long history of decimating Laurentian Great Lakes fish stocks (Applegate 1950, Smith and Tibbles 1980). Models predicting the distribution of larval habitat have application in population models and in guiding field sampling efforts, particularly in evaluating stream reaches upstream of aging low-head dams that are being considered for removal. Larval sea lamprey, termed "ammocoetes", prefer habitat that is characterized by soft, burrowable substrate of fine sands, low current velocity, and some amount of organic detritus (Applegate 1950, Malmqvist 1980, Potter et al. 1986). Organic detritus and fine sands often occur together in "depositional zones" in rivers and so their distribution and abundance is controlled by the geomorphic regime. I explored a dataset that is unique in its spatial extent and geomorphic covariates, and so provided a rich setting for testing hypotheses related to the influence of landscape-scale covariates on instream habitat across process domains. My objective was to understand how relatively large-scale patterns in geomorphic processes (i.e. process domains) might influence the distribution of ammocoete habitat across the Lower Peninsula of Michigan.

Methods

The study region is the Lower Peninsula of Michigan (Fig. 2.1). I used data from 43 watersheds that vary in size from 5 km^2 to 16,176 km^2, and exhibit a variety of land use, surficial geology, and hydrology (Table 2.1). From these data I created linear models relating the proportion of habitat transect width composed of preferred larval sea lamprey habitat to a set of geographic information system (GIS)-derived covariates of hydrology, landuse, and surficial geology. Both larval habitat data and geomorphic covariates are described in detail in a later section. First, I created two sets of models by fitting a multiple linear regression to 1) the entire dataset, and 2) to each of the 43 tributaries individually. I then used regression trees as a method for partitioning a multiple linear regression model by creating three trees, with splits based on: 3) latitude and longitude, 4) values of the GIS-derived covariates, and 5) both latitude, longitude, and values of the covariates. I compared the explanatory power of all models, and examined covariates and strata in the dataset identified by the regression tree models.

Modeling Strategy

My analysis proceeded in five stages. For convenience, I will use the notation “s1” through “s5” to refer to these five stages for the remainder of this paper.

As a first and null model (“s1”), I assumed that the relationship between larval sea lamprey habitat and geomorphic covariates would be adequately represented by a single, best fit linear model across the entire spatial extent and range of covariates of my dataset. To test this hypothesis, I fit to the entire dataset one multiple linear regression model relating preferred larval sea lamprey habitat and my GIS-derived covariates.

In the second stage (“s2”), I hypothesized that the relationship between larval sea lamprey habitat and my covariates varied by watershed. Under this hypothesis I assumed that each watershed corresponds to a unique process domain that most strongly determines instream habitat. I again assumed that the relationship between habitat and covariates might be represented by linear models, but also that the relationship between habitat and covariates differs (i.e. has different regression coefficients) among watersheds. To test this set of hypotheses, I fit a multiple linear regression model to each of the 43 selected watersheds.

In the third stage models (“s3”), I again hypothesized that there is spatial variability in the relationship between larval sea lamprey habitat and covariates, but that this variability may be best characterized by spatial units other than the watershed. Under this hypothesis I assumed that the boundaries of process domains are primarily spatial, e.g. between upstream and downstream reaches (Gomi et al. 2002). I created a regression tree as a machine-learning method for identifying these units, where partitions in the tree were based on latitude and longitude (referred to as the “spatial tree”). Each terminal node in the regression tree contained a spatially contiguous subset of datapoints described by a unique multiple linear regression relating habitat to covariates. If this hypothesis were correct, I expected that the set of regressions based on the regression tree partitions would outperform both models 1 and 2.

In the fourth stage (“s4”), I hypothesized that the relationship between habitat and covariates was characterized by functional complexity, i.e. by nonlinearities, thresholds, and interactions between covariates. Under this hypothesis I assumed that the boundaries of process domains were not necessarily spatial but instead primarily covariate-based, analogous to arranging domains on a slope-area plot (Montgomery and Foufoula-Georgiou 1993, Beardinoni and Hassan 2006). One option for modeling such a system is to partition the dataset based on values of the covariates such that datapoints within each partition may be adequately represented by a linear model. To test this hypothesis, I created a regression tree where partitions in the tree were based only on values of the covariates (referred to as the “covariates-only tree”).

In a fifth and final modeling stage (“s5”), hypothesizing that the relationship between habitat and covariates may be characterized by both geographic variability and functional complexity, I created a fifth model which allowed both: 1) spatial stratification of the dataset, and 2) partitions in the dataset based on values of the covariates. Under this hypothesis I assumed that process domains within my study area are at times best identified by their spatial boundaries, and at other times best identified by covariate-based boundaries. I created a regression tree as a machine-learning method for identifying these partitions, with partitions in the tree based on both latitude and longitude, and on values of the covariates (model 5; referred to as the “hybrid tree”).

Statistical Approach and Analysis

For all of the models, my response variable was the natural logarithm of the mean proportion of transect length that was preferred larval habitat in transects measured by USFWS agents (Type I habitat in Slade et al. 2003). Covariates were a set of seven GIS-derived variables chosen to represent key geomorphic influences on habitat. Both larval habitat data and geomorphic covariates are described in detail in a later section.

To fit one model (s1) to my entire dataset, I selected covariates using forward and backward stepwise regression. To fit a unique multiple linear regression model to each watershed in my dataset (s2), I partitioned the dataset by watershed and selected covariates for each watershed's model using forward and backward stepwise regression. I calculated a single R^2 value for the entire dataset by calculating the sum of squares regression (SSR) for all datapoints across all watersheds, and dividing by the total sum of squares (SST) for all datapoints across all watersheds. I examined the performance of the watershed models in more detail with regression diagnostics on the two best-fitting and two worst-fitting watersheds.

For models 3-5 (s3 - s5), I used the GUIDE software package (Loh 2002) to build regression trees where each terminal tree node (or leaf) contains a multiple linear regression model relating the response (preferred larval habitat) to some subset of my seven GIS-derived covariates. In GUIDE, the trees were constructed by iteratively splitting the data on values of "classification variables". The classification variable used for each split is the one that minimizes the p-value of a chi-squared test for curvature in the relationship between habitat and covariates. For each terminal tree node, GUIDE fits a multiple linear regression model using forward and backward stepwise regression.

I measured the explanatory power of each tree by calculating the adjusted R^2 value using all points in the dataset. I also examined the performance of the trees by performing a detailed diagnostic analysis on the models fitted to the two terminal nodes with the highest R^2_{adj} values and the two terminal nodes with the lowest R^2_{adj}. I report the overall importance of each covariate as the number of terminal node regression relationships in which it appeared. I also report the consistency of each covariate, or the frequency with which the regression coefficient of the covariate was consistently positive or consistently negative.

Larval Sea Lamprey Habitat Data

I used an extensive georeferenced database of larval sea lamprey habitat collected during Quantitative Assessment Sampling surveys (QAS; Slade et al. 2003) conducted by United Stated Fish and Wildlife Service (USFWS) personnel between 1998 and 2003. During the surveys, substrate was classified visually into three habitat types primarily on the basis of substrate particle size (Slade et al. 2003): Type I is the habitat most preferred by larvae and characterized by sands and deposits of particulate organic matter, Type II is acceptable but less preferred and generally consists of coarser or shifting sands or gravel, and Type III is unburrowable and therefore unacceptable. Although this is a subjective classification, there is typically high agreement among observers (Mullett and Bergstedt 2003). Tributaries in the QAS system are classified into "reaches" based on measured or perceived differences in larval densities or habitat abundance, characteristics of the stream network and presence of dams or other structures affecting larval distribution (GLFC 2006). USFWS agents measured habitat along a pair of latitudinal transects at a minimum of six randomly selected access sites per reach and measured the length of each habitat type along each transect. The location of one of the transects was recorded using a global positioning system (GPS) unit. The second transect was located up or downstream a distance of twice the mean stream width. I incorporated 5,027 of these waypoints into my GIS database, and corrected waypoint position when supplementary metadata (e.g. sample date, reach and access site) contradicted the GPS coordinates.

Geomorphic Covariates

I derived geomorphic covariates using GIS: stream power, low-flow yield, distance to mouth of the river, distance to the nearest upstream lake or dam, distance to the nearest downstream lake or dam, the percentage of the upstream riparian corridor with forested land cover, and the percentage of the upstream riparian corridor with surficial geology characterized as lacustrine deposits. River segments included in my analysis were those inhabited by and therefore accessible to sea lamprey, i.e. no headwater reaches were included as they are, in Michigan rivers, almost always upstream of migratory barriers. Nevertheless, mean watershed values of the seven GIS-derived covariates varied considerably among watersheds (Table 2.1). Stream power, distance to the mouth of the river, distance to the nearest upstream lake or dam, and distance to the nearest downstream lake or dam each exhibited approximately two orders of magnitude variation among watersheds. Low-flow yield differed slightly more than one order of magnitude among watersheds. The highest correlation between any two covariates was between lacustrine deposits and the distance to the mouth of the watershed (correlation = -0.303); summary statistics for the response, covariates, and correlation among the covariates are reported in Table 2.2.

I calculated stream power, distance to the mouth of the river, and distance to the nearest upstream and downstream dam or lake using stream path files from the U.S. Geological Survey's National Hydrography Dataset (NHD; USGS 2006) updated for greater accuracy for Michigan Rivers (Brenden et al. 2006). I obtained modeled low-flow yield, forested land cover, and lacustrine surficial geology from a high-resolution GIS database constructed through a combined effort of the Great Lakes Aquatic Gap Analysis Program (GLGAP; GLSC 2006) and the Classification and Impairment Assessment of Upper Midwest Rivers (CIAUMT; Brenden et al. 2006; UM 2006). The database contains 31,817 Michigan stream reaches, and I assigned data to my waypoints by using the data values associated with the stream reach containing each waypoint.

I calculated covariates as follows:

Stream power **-** I calculated stream power based on Whiting et al. (1999) as

$$\text{Power} = g * p * \text{Flow} * \text{Slope}$$

where g is the gravitational constant (9.8 m/s^2), p is the density of water, Flow is the 2-year flood flow (m^3/s), and Slope is the stream channel slope (m/m) at the site. Discharge data for estimating the 2-year flood flow were not available for all sites in my analysis, so I estimated 2-year flood flow using drainage area (Personal communication, A.M. Gorman, Ohio Dept. of Natural Resources) as follows:

$$\text{Flow} = a * \text{Area}^b$$

where $a = 6 * 10^{-6}$ and $b = 0.8613$ are regional regression coefficients derived by Gorman for northeastern Ohio from a regional relationship between 2-year flood flow and drainage area. To better meet assumptions of normality, I selected the natural log of stream power as my covariate.

Low-flow yield - Low-flow yield is the 90% exceedance flow divided by upstream drainage area. These data were available in an existing database of Michigan rivers, where the 90% exceedance flow was generated from hydrologic models (Brenden et al. 2006).

Distance to the mouth of the river - Distance to the mouth of the river was calculated for each of my sample waypoints using the ArcView GIS 3.x extensions Network Analyst (ESRI 1998) and Multiple Closest Facilities, which generated a measurement of the distance along the NHD stream path shapefile between each sample waypoint and the river mouth.

Distance to the nearest upstream (downstream) lake or dam - I combined a Michigan Dept. of Natural Resources (MDNR) dataset of dams and a MDNR dataset of inland lakes to create a single data file of the location of 311 dams and 974 inland lakes in my study region. I used the ArcView extensions Network Analyst and Multiple Closest Facilities to calculate the distance from each waypoint in my habitat database to the nearest upstream and nearest downstream dam or lake. To better meet assumptions of normality, I transformed both covariates according to the formula:
$\text{distance} = \log_e(1 / (\text{distance} + 100))$

Land use and geology - Landuse and surficial geology are categorical variables that each exhibit a sum-to-one constraint (Barringer et al. 1990) and collinearity between categories, so I used as covariates only one category of landuse and one category of surficial geology. From an existing dataset (Brenden et al. 2006), I chose the percentage of the riparian corridor that was forested (landuse) and the percentage of the riparian corridor with lacustrine deposits (surficial geology), the two categories that had the highest correlation with lamprey habitat.

Results

Model Performance

The single linear regression fit to the entire dataset (null model, s1) performed poorly ($R^2 = 0.03$). Regression diagnostics indicated the poor fit was due to high variability in the response variable in the pooled dataset that could not be explained by the model covariates. A plot of predicted values against residuals for the entire model exhibited minimal heteroscedasticity with the exception of artifacts due to the 0,1 constraint on the response variable. However, more than 99% of the predicted values fell between 0.10 and 0.25, suggesting that while the 0,1 constraint violated the assumption of constant variance for some points on or near the extremes it did not inappropriately constrain model fitting or predictions. Plots of residuals against each of the covariates in the model did not indicate that further covariate transformations would improve the fit.

In the s2 analysis, multiple linear regression models fit to each of the 43 watersheds individually outperformed the null model, and had an overall R^2 value of 0.24. Detailed analysis of regression diagnostics of the two best and two worst fitting watersheds suggested that the poor fit was not due to correctable patterns in the residuals.

The regression tree created using latitude and longitude as classification variables (analysis s3; the spatial tree) outperformed models fit to each watershed individually. It had an overall R^2 of 0.30 and contained 23 terminal nodes (Fig. 2.2). The regression tree created using the covariates as classification variables (analysis s4; the covariates-only tree) performed similarly ($R^2 = 0.30$), and contained 21 terminal nodes (Fig. 2.3). The tree created using both latitude and longitude and the covariates as classification variables (analysis s5; the hybrid tree) performed best (overall $R^2 = 0.32$) and contained 27 terminal nodes (Fig. 2.4).

Habitat Patterns in Tree Models

All three of the tree models described similar coarse patterns in the distribution of preferred ammocoete habitat. For example, all three described a high proportion of preferred habitat in several of the smaller watersheds in the Northeast of the Lower Peninsula, (Wycamp Lake Outlet, Elliott Creek, and Greene Creek; watershed IDs 24, 28, and 29 in Fig. 2.1; mean = 28.4%). Likewise, the south branch of the Saginaw River (watershed ID 43 in Fig. 2.1) was predicted to have low proportions of preferred ammocoete habitat (mean = 2.44%). In other watersheds, the proportion of preferred habitat varied among streams within that watershed. In the Muskegon River (watershed ID 9 in Fig. 2.1), Bigelow Creek (the Northeast-most tributary) had consistently high proportions of preferred habitat (mean = 22.4%), while Brooks Creek (the middle North tributary) has consistently mid-range proportions of preferred habitat (mean = 13.6%).

There were also finer-scale differences in the predictions generated by each tree. For example, the covariates-only tree and the hybrid tree predicted a high proportion of preferred habitat in the upstream reaches of the Rifle River (watershed ID 42 in Fig. 2.1), while the spatial tree predicted little or no habitat. Variation in explanatory power within and among the trees is reflected in the R^2 values of the individual nodes. In the spatial tree, the overall R^2 value was 0.30, but the R^2 of individual nodes varied between 0.05 and 0.45. Similarly, the R^2 of individual nodes ranged from 0.04 to 0.51 in the covariates-only tree, and from 0.07 to 0.50 in the hybrid tree. R^2 values and fitted regression equations for each terminal node are in Appendix 1.

I found no obvious patterns between the R^2 value of a terminal node and the number of data points or the particular covariates selected. More detailed regression diagnostics of the two best and two worst fitting nodes in the hybrid tree (as measured by R^2_{adj}) showed that the two worst fitting nodes were not hampered by a need for further transformations or other remedial adjustments. I also saw no obvious relationship between the variance of the response and the fit of the model.

Geomorphic Influences on Habitat

Covariates differed in their importance (the number of terminal node regression relationships in which they appeared) and their consistency (the frequency at which the coefficient of the covariate was consistently positive or consistently negative) (Table 2.3). Distance to the mouth of the river and distance to the nearest upstream dam or lake were the most important covariates (occurring in 32 and 31 nodes overall, respectively). Distance to the mouth of the river was negatively related to preferred habitat in 19 of 32 nodes, while distance to the nearest upstream dam or lake was negatively related to preferred habitat in 16 of 31 nodes. Stream power was the third-most important covariate, occurring in 29 nodes, and had the most consistent relationship with preferred larval habitat; there was a negative relationship between stream power and the proportion of preferred habitat in 24 of 29 nodes. The least important covariate, the distance to the nearest downstream dam or lake, occurred in 22 nodes, and was positively related to preferred habitat in 12 of 22 cases. For all covariates, their importance and consistency was approximately similar across all three regression trees.

Dataset Partitions

Covariates differed in the frequency with which they were used as the basis for a split within a regression tree (Table 2.4). Low-flow yield was the most frequently-used covariate, and served as the basis for 8 splits among the three trees, followed by distance to the mouth of the river (7 splits). Stream power, distance to the nearest upstream dam, and riparian corridor forest were each used for 5 splits. Low-flow yield, stream power, and riparian corridor forest were more commonly used in the covariates tree, while distance to the mouth of the river and to the upstream dam were more commonly used in the hybrid tree. Riparian corridor lacustrine deposits and the distance to the nearest downstream dam were the least frequently used at 2 splits each.

In the spatial tree, the first tree split divided the dataset along a North-South line closely corresponding to a divide between Lake Huron and Lake Michigan drainages. Second-tier splits further divided the Western half of the dataset along an East-West line, and the Eastern half of the state along a North-South line. Covariates differed in their importance among these four partitions (Fig. 2.2c). Datapoints in the Northwest portion of the lower peninsula (green in Fig. 2.2) never included the distance to the nearest upstream or downstream dam or lake as a covariate. Datapoints in the Easternmost portion of the dataset (red in Fig. 2.2) never included distance to the nearest downstream dam or lake as a covariate.

In the covariates-only tree, the top-most split was based on stream power. Second-tier splits were based on distance to the mouth of the river and low-flow yield. I again saw differences in the importance of covariates between the resulting four subsets of the dataset (Fig. 2.3c). Datapoints with low stream power and near the mouth of the river (blue in Fig. 2.3) were never modeled with distance to the mouth of the river as a covariate, and only once used distance to the nearest downstream lake or dam.

The top-most split in the hybrid tree also approximated the divide between Lake Huron and Lake Michigan drainages, and second-tier splits were made on distance to the mouth of the river and latitude. Datapoints in red terminal nodes (approximately those points draining to Lake Huron) were never modeled using distance to the nearest downstream lake or dam (Fig. 2.4c). Datapoints in orange terminal nodes (approximately those points draining to Lake Michigan and having a high distance to the mouth of the river) were most frequently modeled using the distance to the nearest upstream dam or lake.

Discussion

Performance of the Models

The null linear model fit to the entire dataset (s1) performed poorly. Treating each watershed as a separate domain (s2) had better explanatory power, suggesting that there is significant spatial variability in the underlying relationship between habitat and geomorphic covariates, and that some of this variability occurred by watershed. All three regression trees (s3, s4, s5) outperformed the watershed-based models, suggesting that there are within-basin differences in the functional relationship between my landscape covariates and lamprey habitat. The performance of the covariates-only tree (model 4) suggests that partitioning the dataset by values of the covariates may allow linear models to approximate portions of a more complex relationship between habitat and covariates. The moderate increase in fit in the hybrid tree (model 5) suggests that a combination of geographic factors (i.e. latitude and longitude) and physiographic factors work best to delimit process domains.

The modest explanatory power of the three best models (R^2 about 0.30) may limit their overall use as predictive models in fisheries management. However, my models explained over 50% of the variability in the response (i.e. $R^2 > 0.5$) in some domains. Identifying these domains is an important step towards understanding where habitat is most predictable from landscape-scale covariates.

Information gained from the results of these models may also serve as the basis for future modeling efforts or studies. Specifically, the regression trees allowed us to 1) identify both spatial and covariate-based stratifications of the lower peninsula of Michigan, 2) identify the most important geomorphic influences on habitat, and 3) describe how the influence of covariates on habitat differs among strata in the dataset.

Inference from Tree Splits

Stratifications identified by the regression trees represent statistically significant differences in the coefficients of the regression relationships between habitat and covariates. For the spatial tree and the latitude and longitude splits in the hybrid tree, the tree splits delineate spatially contiguous units among which the relationship between habitat and covariates differs. For the covariates-only tree and the covariate-based splits in the hybrid tree, the tree splits represent process-based partitions in the dataset within which the complex relationship between habitat and covariates can be better approximated by a linear model.

The partitions identified by my regression tree analysis may inform stratification strategy for future studies and modeling efforts. For example, the top-most split in the spatial and hybrid trees is a split between the East and West sides of Michigan, approximately through the center of the Lower Peninsula. This split may be due to differences in physiography, or to latent climate variation in the form of differences in precipitation regime between the East and West sides of Michigan (Albert 1995). This hypothesis could be tested with further study, and incorporated into future habitat models explicitly (by introducing a covariate describing climate) or implicitly (by stratifying studies or models by an East-West partition.) Second-tier spatial partitions divide the Western half of the Lower Peninsula into North/South units, and divide the Eastern half of the Lower Peninsula into East/West units.

The size of my dataset and the number of tree nodes limited my ability to understand the basis of covariate-based tree splits. In a CART analysis of desert tortoise habitat with a smaller dataset and geographic area, for example, Andersen et al. (2000) were able to use tree splits to identify interactions between covariates. However, covariates that served as key classification variables in my dataset can still be useful for stratifying future studies and models. The top-level splits in the covariates-only tree (stream power, distance to the mouth of the watershed, and 90% exceedance flow) and in the hybrid tree (distance to the mouth of the watershed) likely represent most important partitioning of the dataset for management or future modeling efforts, while partitions further down the tree could be included if costs allow.

Relationship Between Habitat and Covariates

The regression coefficients of some covariates were consistent in sign across tree nodes and reflect previously known relationships between geomorphology and habitat. Preferred ammocoete habitat consists of relatively easily-transported and deposited medium sands and organic detritus, so that I expect a negative correlation between the amount of preferred ammocoete habitat and the sediment deposition rate. Stream power had a negative regression coefficient in 24 of 29 terminal nodes, corresponding to empirical studies that demonstrate that stream power is positively correlated with sediment transport rate (Bagnold 1980) and median substrate particle size (Whiting et al. 1999). The importance of stream power in my models compares well with Zorn et al. (2002), who found that catchment area (a component of stream power) was a primary driver of fish community composition. The amount of forest in the riparian corridor had a positive regression coefficient in 17 of 24 terminal nodes, which corresponds to empirical studies describing the negative correlation between forested landscapes and sediment transport rates (Milliman et al.1987).

Rivers in glaciated landscapes such as Michigan flow through regions of varying landuse and geology, and the relative importance of geomorphic drivers of habitat may differ between such regions. Differences in the importance of covariates between strata reflect this variability in the influence of geomorphology on habitat. For example, habitat in the spatial tree in the Northwest quarter of the Lower Peninsula was never influenced by the distance to the nearest upstream or downstream dam or lake (green points in Fig. 2.2). Habitat in the spatial tree in the Eastern quarter of the Lower Peninsula was never influenced by the distance to the nearest downstream dam or lake (red points in Fig. 2.2). These patterns of variability in the importance of covariates may be characterized at any of several spatial scales (i.e. tree depths), and may be combined with expert knowledge of regional hydrology to inform future studies and models of habitat.

For other covariates, such as low-flow yield and the distance to the nearest upstream dam or lake, the number of nodes with positive and negative regression coefficients was approximately equal. This may reflect changes in the relationship between geomorphology and habitat across Michigan's glaciated landscapes. Alternatively, the sign of the relationships may be due to a mismatch between the scales at which I measured covariates and the scales at which they influence habitat. For example, a dam may limit downstream sediment supply (Williams and Wolman 1984, Meade and Parker 1985), but this effect diminishes with distance and the supply of sediment from additional tributaries downstream. For waypoints with a measured distance to the nearest dam greater than several kilometers, the covariate may be serving as a general description of geographic position. Distance from the river mouth was chosen to incorporate the gradient of increasing substrate particle size from mouth of the rivers to headwaters (Knighton 1980, Brierley and Hickin 1985), but this gradient is rarely smooth and local convex-upwards shapes and scatter may explain the 13 positive regression coefficients in 32 terminal nodes.

Management Applications

A general management objective for sea lamprey control in the Great Lakes is to increase the number of ammocoetes killed per dollar spent (a cost/kill ratio; Slade et al. 2003, Christie et al 2003), and proposed modifications to the sampling protocol used to rank streams for treatment are typically evaluated in terms of their cost/kill impact. Until 2008, estimates of both larval density and total amount of larval habitat were necessary for estimating the expected production of parasitic-phase lamprey in each reach, and so improving the estimates of either would be beneficial for the accuracy of stream rankings. However, simply increasing the amount of field sampling performed is not a feasible way to improve accuracy. Regarding larval density estimates, Hansen et al. (2003) calculate that field sampling efforts "would need to be increased 52% to increase precision of the larval density to 50% of the mean, and by more than 240% to increase precision of the larval density estimate to 33% of the mean." The calculations would be different for habitat, but Hansen's calculations for larvae demonstrate the general point that large increases in field sampling efforts are needed to increase the precision of measurements.

Although the QAS protocol has recently been replaced by a "rapid assessment protocol" that uses historical data in place of additional field sampling of habitat (GLFC 2008), accurate estimates of the distribution and amount of preferred ammocoete habitat are still desirable. The switch from the QAS protocol to the rapid assessment protocol was made on the basis of a cost/benefit analysis, but good predictive habitat models that complement and guide field sampling efforts could significantly change the costs associated with measuring habitat.

Predictive larval habitat models may also be useful for predicting distribution and amount of habitat upstream of dams being considered for removal. Adult sea lampreys are constrained below most dams during their spawning runs, and dam removal could open large areas to infestation. Dam removal is being considered for many rivers to improve ecological integrity (Lavis et al. 2003), and is a pressing management concern given the more than 2000 existing dams in the lower peninsula of Michigan (MDNR data). Predictive habitat models that complement field sampling efforts provide a cost-effective way to estimate the amount of preferred ammocoete habitat that will be opened to infestation.

Given the modest predictive power of my models (R^2 approximately 0.30 for all trees), they may best be used as the basis for a stratified sampling regime (Guisan and Zimmerman 2000, Cole et al. 2001) that guides, rather than replaces, existing field sampling efforts. Specifically, the trees identify domains where the explanatory power of landscape variables is relatively good ($R^2 > 0.5$) as well as domains where the explanatory power is relatively low ($R^2 < 0.1$). These two extremes will require different allocation of field-sampling efforts to arrive at the same degree of confidence in estimates of larval habitat in a reach. Tree-identified domains may therefore serve as the basis for efficiently allocating field survey efforts.

Despite differences in their predictive power, each step in the analysis contributed to a better understanding of the processes and spatial units influencing habitat. The spatial and hybrid trees identify geographic areas that could serve as the basis for management units or stratifications for partitioning future habitat models. The presence or absence of covariates in the regression relationship that describes each of these nodes may highlight the critical influences on larval habitat within that management unit. In addition, the trees identify key geomorphic influences on habitat and key stratifying geomorphic variables, both of which may be used to understand the processes influencing habitat and to inform future modeling efforts.

References

Albert, Dennis A. 1995. Regional landscape ecosystems of Michigan, Minnesota, and Wisconsin: a working map and classification. Gen. Tech. Rep. NC-178. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 250 pp.

Allan, J.D., Erickson, D.L, and Fay, J. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37: 149-161.

Andersen, M.C., Watts, J.M., Freilich, J.E., Yool, S.R., Wakefield, G.I., McCauley, J.F., and Fahnestock, P.B. 2000. Regression tree modeling of desert tortoise habitat in the central Mojave desert. Ecological Applications 10: 890-900.

Angermeier PL, Krueger KL, and Dolloff CA. 2001. Discontinuity in stream-fish distributions: Implications for assessing and predicting species occurrence. Pages 519–527. in Scott JM, Heglund PJ, Morrison ML, eds Predicting Species Occurrences: Issues of Accuracy and Scale. Covelo (CA): Island Press.

Applegate, V.C. 1950. Natural history of the sea lamprey (*Petromyzon marinus*) in Michigan. U.S. Fish and Wildlife Service, Special Scientific Report. Fisheries 55.

Bagnold, R.A. 1980. An empirical correlation of bedload transport rates in flumes and natural rivers. Proceedings of the Royal Society 372A: 453-73.

Barringer, T.H., Dennis Dunn, Battaglin, W.A., and Vowinkel, E.F. 1990. Problems and methods involved in relating land use to ground-water quality. Water Resources Bulletin 26: 1 - 9.

Becker, A. and Pfutzner, B. 1986. Identification and modeling of river flow reductions caused by evapotranspiration losses from shallow groundwater areas. IAHS Publication No. 156.

Becker, A., and Braun, P. 1999. Disaggregation, aggregation, and spatial scaling in hydrological modeling. Journal of Hydrologyy 217: 239-252.

Brardinoni, F., and Hassan, M.A. 2006. Glacial erosion, evolution of river long profiles, and the organization of process domains in mountain drainage basins of coastal British Columbia. Journal of Geophysical Research 111: F01013.

Brieman, L., J.H. Friedman, R.A. Olsen, and C.J. Stone. 1984. Classification and Regression Trees. Wadsworth, Belmont.

Brierley, G.J., Hickin, E.J., 1985. The downstream gradation of particle sizes in the Squamish River, British Columbia: Earth Surface Processes and Landforms, v. 10, p. 597-606.

Brenden, T.O., R.D. Clark, A.R. Cooper, P.W. Seelbach, L. Wang, S. Aichele, E.G. Bissell, and J.S. Stewart. 2006. A GIS framework for collecting, managing, and analyzing multiscale landscape variables across large regions for river conservation and management. Pages 49-74 *in* R.M. Hughes, L. Wang, and P.W. Seelbach, editors. Landscape influences on stream habitat and biological assemblages. American Fisheries Society, Symposium 48, Bethesda, Maryland.

Brenden, T.O., Wang, L, and Seelbach, P.W. 2008. A river valley segment classification of Michigan streams based on fish and physical attributes. Transactions of the American Fisheries Society 2008: 1621 - 1636.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. 1984. Classification and Regression Trees. Wadsworth, Belmont.

Brune, G.M. 1953. The trap efficiency of reservoirs. Transactions of the American Geophysical Union 34: 407-418.

Christie, G.C., Adams, J.V., Steeves, T.B., Slade, J.W., Cuddy, D.W., Fodale, M.F., Young, R.J., Kuc, M., and Jones, M.L. 2003. Selecting Great Lakes streams for lampricide treatment based on larval sea lamprey surveys. J. Great Lakes Res. 29 (Suppl. 1): 152 -160.

Cole, R.G., Healy, T.R., Wood, M.L., and Foster, D.M. 2001. Statistical analysis of spatial pattern: a comparison of grid and hierarchical sampling approaches. Environmental Monitoring and Assessment 69: 85-99.

Cooper, M.J., Uzarski, D.G., and Burton, T.M. 2007. Macroinvertebrate community composition in relation to anthropogenic disturbance, vegetation, and organic sediment depth in four Lake Michigan drowned river-mouth wetlands. Wetlands, 27: 894-903.

Cunjak, R.A. 1996. Winter habitat of selected stream fishes and potential impacts from land-use activity. Can. J. Fish. Aquat. Sci. 53(Suppl. 1): 267-282.

De'ath, G. 2002. Multivariate regression trees: A new technique for modeling species-environment relationships. Ecology 83: 1105-1117.

De'ath, G., and Fabricius, K.E. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology 81: 3178-3192.

Delong, M.D., and Brusven, M.A. 2006. Classification and spatial mapping of riparian habitat with applications toward management of streams impacted by nonpoint source pollution. Environmental Management 15: 565 - 571.

Eikens, B. and Karim, M.N. 1999. Process identification with multiple neural network models. Int. J. Control. 72: 576 - 590.

ESRI, Inc. 1998. *ArcView Network Analyst*, an ESRI White Paper.

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

GLSC (Great Lakes Science Center). 2006. The Great Lakes Aquatic GAP Project. Available: www.glfc.usgs.gov.

Gomi, T., Sidle, R.C., and Richardson, J.S. 2002. Understanding processes and downstream linkages of headwater systems. Bioscience 52: 905 - 916.

Gregory, S.V., Swanson, F.J., McKee, W.A., and Cummins, K.W. 1991. An ecosystem perspective on riparian zones. Bioscience 41: 540-551.

Guisan, A. and Zimmerman, N.E. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135: 147-186.

Hutchinson, G.E. 1957. Concluding remarks. *Cold Spring Harbour Symposium on Quantitative Biology* 22: 415-427.

Knighton, A.D., 1980, Longitudinal changes in size and sorting of stream bed material in four English rivers: Geological Society of America Bulletin, v. 91, p. 55-62.

Knighton, D. 1998. Fluvial forms and processes: a new perspective. John Wiley and Sons, New York, New York. 383 pp.

Kondolf, G.M. 1997. Hungry water: Effects of dams and gravel mining on river channels. Environmental Management 21: 533-551.

Kondolf, G.M., and Wolman, M.G. 1993. The sizes of salmonid spawning gravels. Water Resour. Res. 29: 2275-2285.

Langbein, W.B. and Schumm, S.A. 1958. Yield of sediment in relation to mean annual precipitation. Transactions of the American Geophysical Union 39: 1076-1084.

Lavis, D.S., Hallet. A., Koon, E.M., and McAuley, T.C. 2003. History of and advances in barriers as an alternative method to suppress sea lampreys in the Great Lakes. J. Great Lakes Res. 29 (Suppl. 1): 362-372.

Loh, W-Y. 2002. Regression trees with unbiased variable selection and interaction detection. Statistica Sinica 12: 361-386.

Loh, W.Y. 2006. Regression by parts: Fitting visually interpretable models with GUIDE. Handbook of Computational Statistics, Vol. III, Springer.

Lunetta, R.S., Cosentino, B.L., Montgomery, D.R., Beamer, E.M., and Beechie, T.J. 1997. GIS-based evaluation of salmon habitat in the pacific northwest. Photogrammetric Engineering and Remote Sensing 63: 1219 - 1229.

Macey, D.J. and Potter, I.C. 1978. Lethal temperatures of ammocoetes of the Southern Hemisphere lamprey, *Geotria australis* Gray. Environ. Biol. Fish. 32: 241 - 243.

Madej, M.A. and Ozaki, V. 1996. Channel response to sediment wave propagation and movement, Redwood Creek, California, USA. Earth Surfact Processes and Landforms 21: 911-927.

Malmqvist, B. 1980. Habitat selection of larval brook lampreys (*Lampetra planeri,* Bloch) in a south Swedish stream. Oecologia 45: 35-38.

Manion, P.J.. and McLain, A.L. 1971. Biology of larval sea lamprey (*Petromyzon marinus)* of the 1960 year class, isolated in the Big Garlic River, Michigan, 1960-65. Great Lakes Fish. Comm. Tech. Rep. 16: 35 pp.

Marcus, M.D. et al. 1990. Salmonid-Habitat Relationships in the Western United States: A Review and Indexed Bibliography. General Technical Report No. RM-188. U.S. Department of Agriculture, Forest Service. February.

Meade, R.H. and Parker, R.S. 1985. Sediment in rivers in the United States. In *National Water Summary 1984,* United States Geological Survey Water-Supply Paper 2275: 49-60.

Mertens, M., Nestler, I, and Huwe, B. 2001. GIS-based regionalization of soil profiles with Classification and Regression Trees (CART). J. Plant. Nutr. Soil Sci. 165: 39-43.

Michaelsen, J., Schimel, D.S., Friedl, M.A., Davis, F.W., and Dubayah, R.C. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. Journal of Vegetation Science 5: 673-686.

Miller, H. and Han, J. Geographic Data Mining and Knowledge Discovery. Taylor and Francis, London, U.K., 2001.

Milliman, J.D., Yun-Shan, Q., Mei-E, R., and Saito, Y. 1987. Man's influence on the erosion and transport of sediment by Asian rivers: the Yellow River (Huanhe) example. Journal of Geology 95: 751-762.

Montgomery, D.R. 1999. Process domain and river continuum. Journal of the American Water Resources Association 35: 397-410.

Montgomery, D.R., and Foufoula-Georgiou, E. 1993. Channel network source representation using digital elevation models. Water Resources Research 29: 3925-3934.

Mullett, K.M., and Bergstedt, R.A. 2003. Agreement among observers classifying larval sea lamprey (*Petromyzon marinus*) habitat. J. Great Lakes Res. 29 (Suppl. 1): 183-189.

Parsley, M.J., Beckman, L.G., and McCabe Jr., G.T. 1993. Spawning and rearing habitat use by white sturgeons in the Columbia River downstream from McNary Dam. Transactions of the American Fisheries Society 122: 217-227.

Phillips, J.D. 2003. Sources of nonlinearity and complexity in geomorphic systems. Progress in Physical Geography 27: 1-23.

Phillips, J.D. 2006. Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change. Hydrol. Earth Syst. Sci. 10: 731-742.

Piavis, G. W. 1971. Embryology, p. 361-400. *In* M.W. Hardisty and I.C. Potter [ed.] The biology of lampreys. Vol. 1. Academic Press, London and New York.

Potter, I.C., and Beamish, F.W.H. 1975. Lethal temperatures in ammocoetes of four species of lampreys. Acta Zool. 56: 85 - 91.

Potter, I.C., Hill, B.J., and Gentleman, S. 1970. Survival and behavior of ammocoetes at low Oxygen tensions. J. Exp. Biol. 53: 59-73.

Potter, I.C., Hilliard, R.W., Bradley, J.S., and McKay, R.J. 1986. The influence of environmental variables on the density of larval lampreys in different seasons. Oecologia 70: 433-440.

Rejwan, C., Collins, N.C., Brunner, L.J., Shuter, B.J., and Ridgway, M.S. 1999. Tree regression analysis of the nesting habitat of smallmouth bass. Ecology 80: 341-348.

Seelbach P. W., M. J. Wiley, J. C. Kotanchik, and M. E. Baker. 1997. A landscape-based ecological classification system for river valley segments in lower Michigan (MI-VSEC Version 1.0). Michigan Department of Natural Resources, Fisheries Division, Fisheries Research Report 2036, Lansing.

Shreve, R.L. 1979. Models for prediction in fluvial geomorphology. Mathematical Geology 11: 165-173.

Slade, J.W., Adams, J.V., Christie, G.C., Cuddy, D.W., Fodale, M.F., Heinrich, J.W, Quinlan, H.R., Weise, J.G., Weisser, J.W., and Young, R.J. 2003. Techniques and methods for estimating abundance of larval and metamorphosed sea lampreys in Great Lakes tributaries, 1995 to 2002. J. Great Lakes Res. 29 (Suppl. 1): 137-151.

Smith, B.R. and Tibbles, J.J. 1980. Sea lamprey (*Petromyzon marinus*) in Lakes Huron, Michigan, and Superior: history of invasion and control, 1936-1978. Can. J. Fish. Aquat. Sci. 37: 1780-1801.

Steen, P.J., T.G. Zorn, P.W. Seelbach, and J.S. Schaeffer. 2008. Classification tree models for predictions distributions of Michigan stream fish from landscape variables. Transacations of the American Fisheries Society 137: 976-996.

UM (University of Michigan). 2006. Ecological classification of rivers for environmental assessment. Available: sitemaker.umich.edu/riverclassproject.

USGS (U.S. Geological Survey). 2006. National hydrography dataset. Available: nhd.usgs.gov.

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

Wehrly, K.E., Brenden, T.O., and Wang, L. 2009. A comparison of statistical approaches for predicting stream temperatures across heterogeneous landscapes. Journal of the American Water Resources Association 45: 986-997.

Whiting, P.J. and Bradley, J.B. 1993. A process-based classification system for headwater streams. Earth Surface Processes and Landforms 18: 603-612.

Whiting, P.J, J.F. Stamm, D.B. Moog, and R.L. Odndorff. 1999. Sediment-transporting flows in headwater streams. Bulletin of the Geological Societty of America. 111(3): 450-466.

Williams, G.P. and Wolman, M.G. 1984. Downstream effects of dams on alluvial rivers. United States Geological Survey Professional Paper 1286.

Zorn, T.G., P.W. Seelbach, and M.J. Wiley. 1998. Patterns in the distributions of stream fishes in Michigan's Lower Peninsula. Michigan Department of Natural Resources, Fisheries Research Report 2035, Ann Arbor.

Zorn, T.G., P.W. Seelbach, and M.J. Wiley. 2002. Distributions of stream fishes and their relationship to stream size and hydrology in Michigan's lower peninsula. Transactions of the American Fisheries Society 131: 70-85.

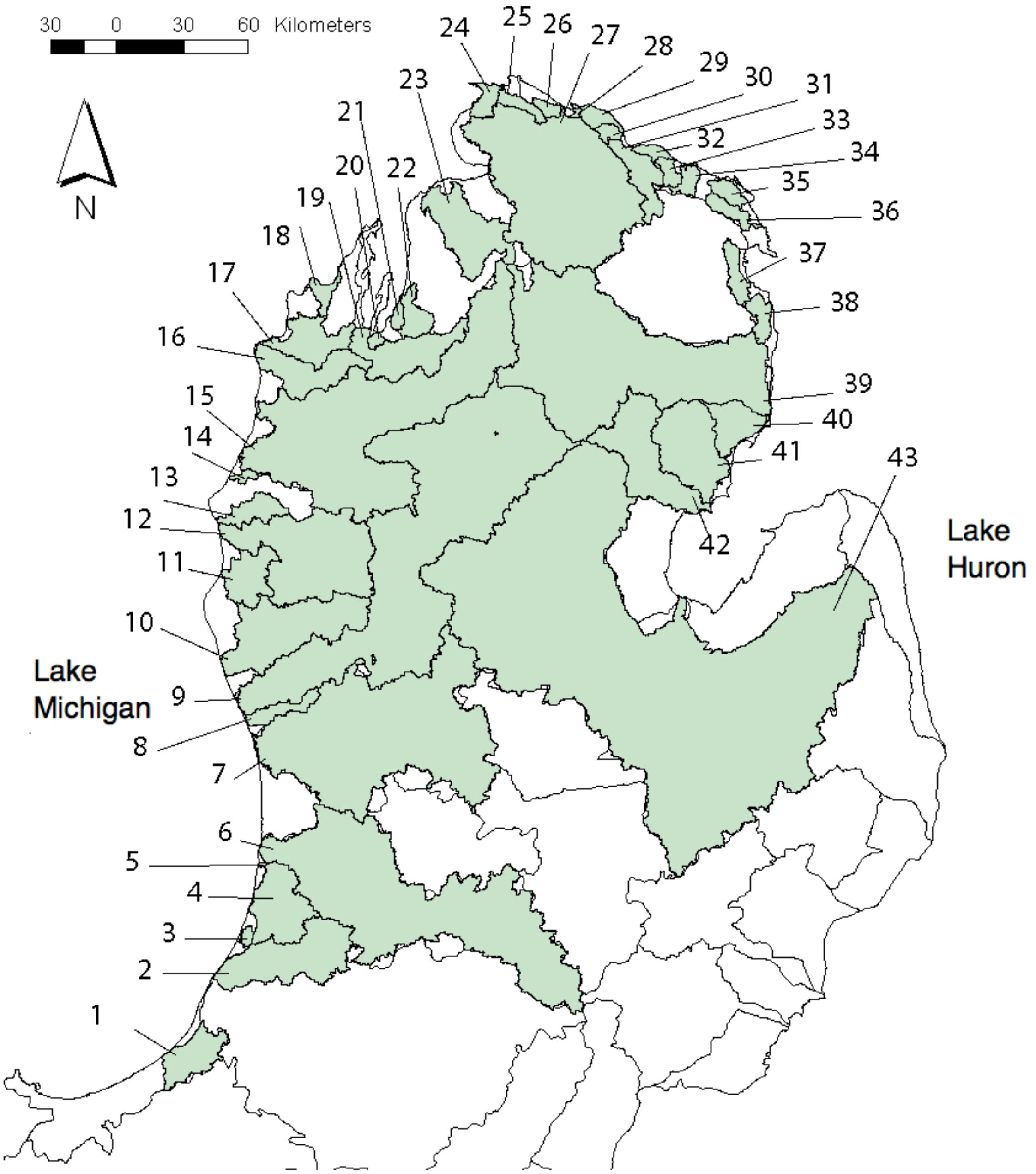
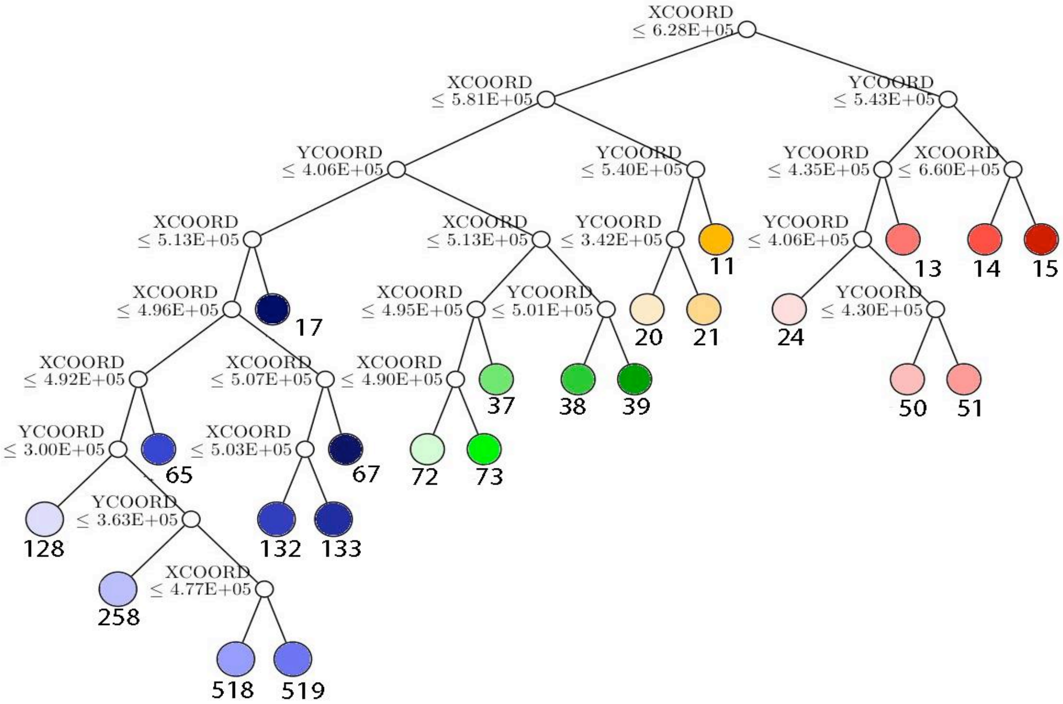
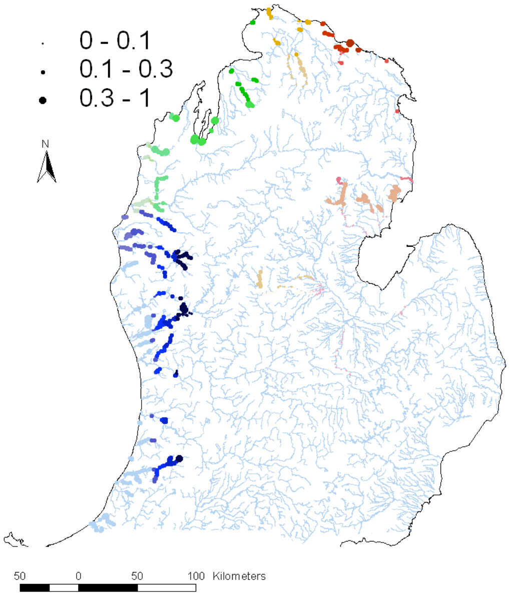


Figure 2.1: The study area is composed of 43 watersheds in the Lower Peninsula of Michigan. Watersheds included in the analysis are colored grey, and identification numbers correspond with ID numbers in Table 2.1.


XCOORD ≤ 6.28E+05
XCOORD ≤ 5.81E+05
YCOORD ≤ 5.43E+05
YCOORD ≤ 4.06E+05
YCOORD ≤ 5.40E+05
YCOORD ≤ 4.35E+05
XCOORD ≤ 6.60E+05
XCOORD ≤ 5.13E+05
XCOORD ≤ 5.13E+05
YCOORD ≤ 3.42E+05
11
YCOORD ≤ 4.06E+05
13
14
15
XCOORD ≤ 4.96E+05
17
XCOORD ≤ 4.95E+05
YCOORD ≤ 5.01E+05
20
21
24
YCOORD ≤ 4.30E+05
XCOORD ≤ 4.92E+05
XCOORD ≤ 5.07E+05
XCOORD ≤ 4.90E+05
37
38
39
50
51
YCOORD ≤ 3.00E+05
65
XCOORD ≤ 5.03E+05
67
72
73
128
YCOORD ≤ 3.63E+05
132
133
258
XCOORD ≤ 4.77E+05
518
519


0 - 0.1
0.1 - 0.3
0.3 - 1
N
50
0
50
100
Kilometers

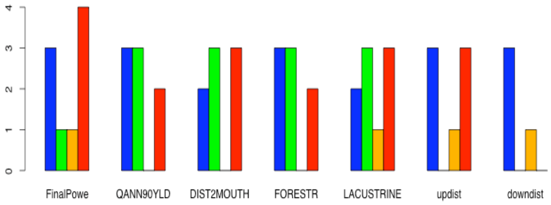
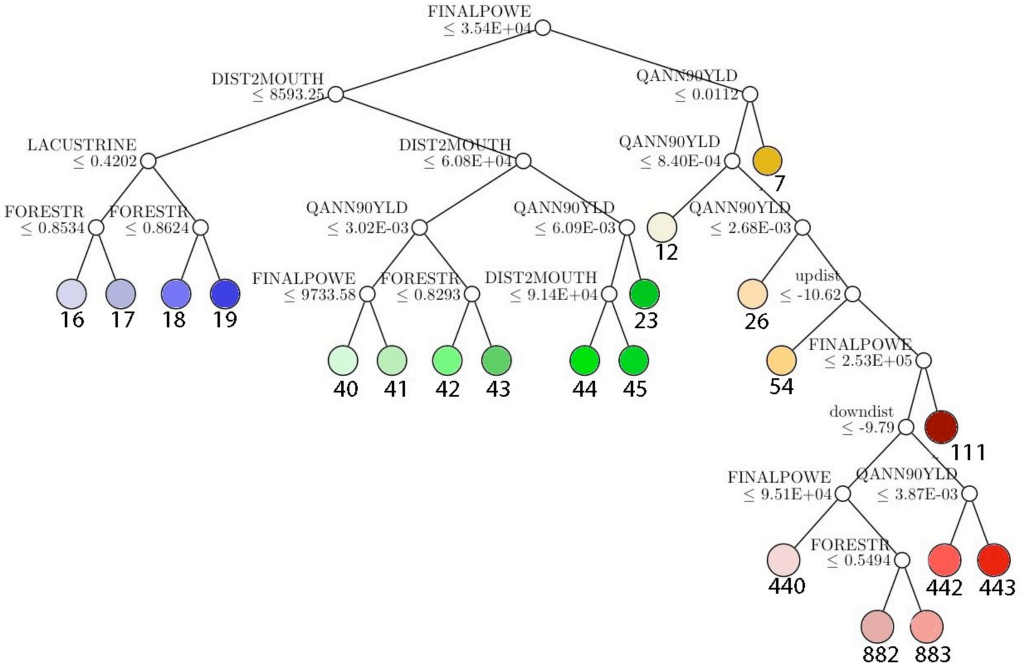
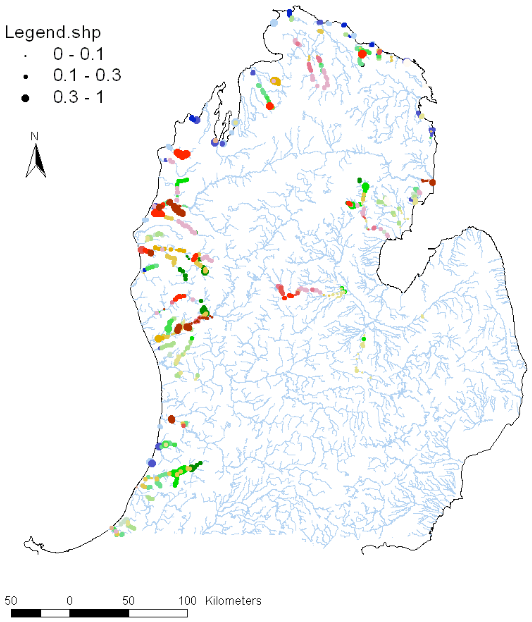

0
1
2
3
4
FinalPowe
QANN90YLD
DIST2MOUTH
FORESTR
LACUSTRINE
updist
downdist

Figure 2.2: a) Fitted regression tree relating preferred larval sea lamprey habitat to GIS-derived geomorphic covariates, where splits in the tree are based on latitude and longitude (“spatial tree”, analysis s3; R^2= 0.30). Each of the tree splits is labeled with the value of the covariate that determined its split: latitude (YCOORD) or longitude (XCOORD). Numbers below each terminal nodes report the node ID, which corresponds to individual regression equations in Appendix Table A1. The color of each node corresponds to waypoints in (b), a map of n=5027 waypoints of habitat transect data used to fit the regression tree. The size of each waypoint reflects the predicted proportion of habitat transect length that is preferred larval habitat. Color “groups” (i.e. blue, green, orange, red) in the tree and map correspond to (c), barplot showing the number of times that each covariate was used in a terminal node regression relationship for all terminal nodes in that color group.


FINALPOWE
≤ 3.54E+04
DIST2MOUTH
≤ 8593.25
QANN90YLD
≤ 0.0112
LACUSTRINE
≤ 0.4202
DIST2MOUTH
≤ 6.08E+04
QANN90YLD
≤ 8.40E-04
7
FORESTR
≤ 0.8534
FORESTR
≤ 0.8624
QANN90YLD
≤ 3.02E-03
QANN90YLD
≤ 6.09E-03
12
QANN90YLD
≤ 2.68E-03
FINALPOWE
≤ 9733.58
FORESTR
≤ 0.8293
DIST2MOUTH
≤ 9.14E+04
23
26
updist
≤ -10.62
16
17
18
19
40
41
42
43
44
45
54
FINALPOWE
≤ 2.53E+05
downdist
≤ -9.79
111
FINALPOWE
≤ 9.51E+04
QANN90YLD
≤ 3.87E-03
FORESTR
≤ 0.5494
440
442
443
882
883


Legend.shp
0 - 0.1
0.1 - 0.3
0.3 - 1
N
50
0
50
100
Kilometers

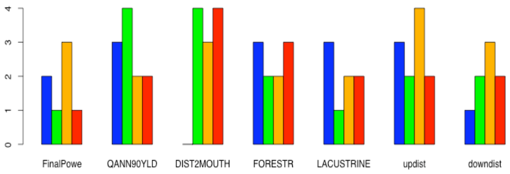
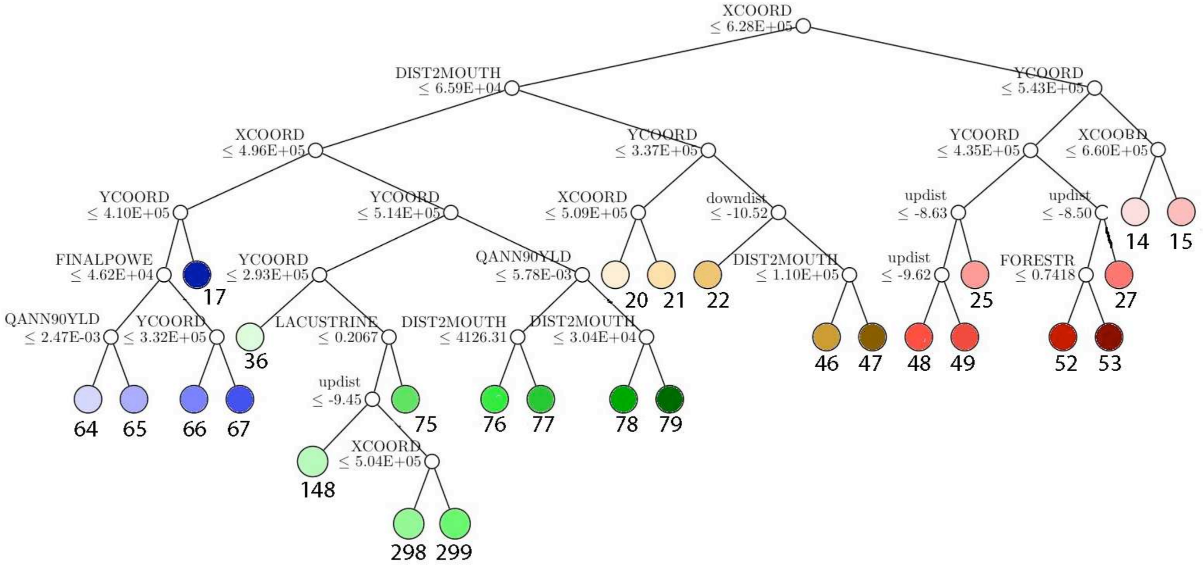
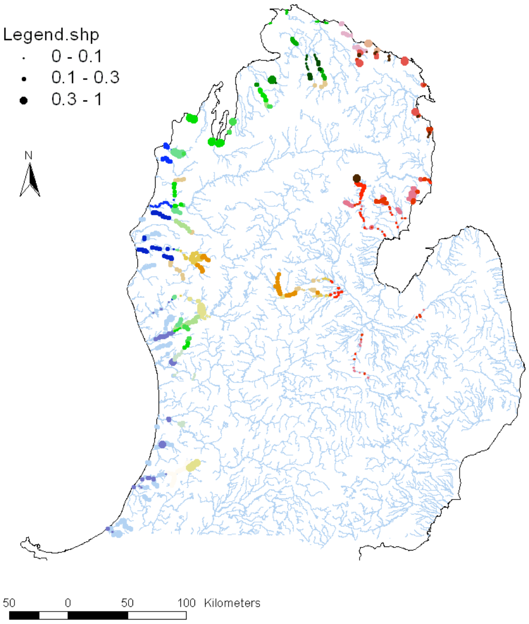

0
1
2
3
4
FinalPowe
QANN90YLD
DIST2MOUTH
FORESTR
LACUSTRINE
updist
downdist

Figure 2.3: a) Fitted regression tree relating preferred larval sea lamprey habitat to GIS-derived geomorphic covariates, where splits in the tree are based on values of the covariates (“covariates-only tree”, analysis s4; $R^2 = 0.30$). Each of the tree splits is labeled with the value of the covariate that determined its split. Numbers below each terminal nodes report the node ID, which corresponds to individual regression equations in Appendix Table A2. The color of each node corresponds to waypoints in (b), a map of n=5027 waypoints of habitat transect data used to fit the regression tree. The size of each waypoint reflects the predicted proportion of habitat transect length that is preferred larval habitat. Color “groups” (i.e. blue, green, orange, red) in the tree and map correspond to (c), barplot showing the number of times that each covariate was used in a terminal node regression relationship for all terminal nodes in that color group.


XCOORD ≤ 6.28E+05
DIST2MOUTH ≤ 6.59E+04
YCOORD ≤ 5.43E+05
XCOORD ≤ 4.96E+05
YCOORD ≤ 3.37E+05
YCOORD ≤ 4.35E+05
XCOORD ≤ 6.60E+05
YCOORD ≤ 4.10E+05
YCOORD ≤ 5.14E+05
XCOORD ≤ 5.09E+05
downdist ≤ -10.52
updist ≤ -8.63
updist ≤ -8.50
14
15
FINALPOWE ≤ 4.62E+04
17
YCOORD ≤ 2.93E+05
QANN90YLD ≤ 5.78E-03
20
21
22
DIST2MOUTH ≤ 1.10E+05
updist ≤ -9.62
25
FORESTR ≤ 0.7418
27
QANN90YLD ≤ 2.47E-03
YCOORD ≤ 3.32E+05
36
LACUSTRINE ≤ 0.2067
DIST2MOUTH ≤ 4126.31
DIST2MOUTH ≤ 3.04E+04
46
47
48
49
52
53
64
65
66
67
updist ≤ -9.45
75
76
77
78
79
148
XCOORD ≤ 5.04E+05
298
299


Legend.shp
0 - 0.1
0.1 - 0.3
0.3 - 1
N
50
0
50
100
Kilometers

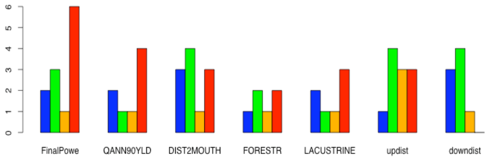

0
1
2
3
4
5
6
FinalPowe
QANN90YLD
DIST2MOUTH
FORESTR
LACUSTRINE
updist
downdist

Figure 2.4: a) Fitted regression tree relating preferred larval sea lamprey habitat to GIS-derived geomorphic covariates, where splits in the tree are based on latitude, longitude, and values of the covariates (“hybrid tree”, analysis s5; $R^2= 0.32$). Each of the tree splits is labeled with the value of the covariate that determined its split. Numbers below each terminal nodes report the node ID, which corresponds to individual regression equations in Appendix Table A3. The color of each node corresponds to waypoints in (b), a map of n=5027 waypoints of habitat transect data used to fit the regression tree. The size of each waypoint reflects the predicted proportion of habitat transect length that is preferred larval habitat. Color “groups” (i.e. blue, green, orange, red) in the tree and map correspond to (c), barplot showing the number of times that each covariate was used in a terminal node regression relationship for all terminal nodes in that color group.

Table 2.1: Summary statistics for each of the 43 watersheds in the study area. Table columns give: the Laurentian Great Lake into which each watershed drains (L. Michigan = 2, L. Huron = 3); ID number, corresponding to Fig. 2.1; DA, the drainage area; n, the number of larval habitat datapoints; the mean proportion of transect width classified as Type 1 habitat in QAS surveys; the standard deviation of measurements of transect width classified as Type 1 habitat. The remaining columns give mean watershed values for each of the covariates used as regressors in my models: FinalPowe, stream power; QANN90YLD, the low-flow yield; Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; downDist, the distance to the nearest downstream dam or lake; FOREST_R, the percentage of the riparian corridor classified as forested landcover; and LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology.

Stream Name	Lake	ID	DA km^2	n	Type I (mean) Proportion	Type I (sd)	FinalPowe N/s	QANN90YLD	Dist2Mouth m	upDist m	downDist m	FORESTR %	LACUSTRINE_R %
Galien River	2	1	454.13	182	22.8	22.5	18862	0.00318116	23747	3329	6754	55.5	37.0
St. Joseph	2	2	1155.61	246	23.7	20.1	33951	0.00466708	62864	8823	48670	71.2	14.5
Brandywine Creek	2	3	41.48	18	30.3	27.7	18075	0.00311000	1598	4618	NA	66.4	65.0
Black River	2	4	742.55	30	22.8	16.9	14385	0.00503388	17978	12800	7412	79.0	100.0
Allegan 4 Creek	2	5	5.40	24	19.6	17.5	3061	0.00108700	458	NA	NA	43.5	27.9
Kalamazoo River	2	6	5039.06	153	11.4	10.1	546366	0.00334988	29634	1463	18059	78.7	92.5
Grand River	2	7	4944.01	138	16.1	19.0	172140	0.00143214	39659	15356	14630	70.9	62.4
Black Creek	2	8	194.31	36	21.3	11.4	12936	0.00169000	15219	13562	10764	66.6	100.0
Muskegon River	2	9	7052.16	256	15.6	15.2	515644	0.00497642	45207	20114	13889	69.4	25.2
White River	2	10	1392.82	80	15.0	13.8	86281	0.00916162	65668	18187	7859	82.0	0.0
Pentwater River	2	11	442.07	42	23.4	13.4	11291	0.00581930	16448	37848	NA	79.4	27.2
Pere Marquette River	2	12	1954.64	443	18.0	15.4	75446	0.00935625	85136	14084	28821	90.8	4.1
Lincoln River	2	13	262.00	46	23.2	12.6	16845	0.00171840	20435	20510	18204	84.6	80.5
Gurney Creek	2	14	48.45	48	15.2	11.4	10600	0.00234124	1061	6759	NA	93.0	30.4
Big Manistee River	2	15	5046.13	389	14.6	12.6	291623	0.00932900	31216	19526	21754	66.6	59.4
Betsie River	2	16	627.38	18	13.3	13.3	75463	0.00368000	12208	4914	NA	82.9	34.0
Platte River	2	17	498.38	204	24.5	17.5	29988	0.00453530	8891	4202	3233	71.1	11.7
Good Harbor Creek	2	18	117.83	66	50.7	32.0	4767	0.00644557	4317	6695	NA	86.5	81.2
Boardman River	2	19	734.47	34	31.7	19.5	76827	0.00845000	2604	2981	NA	32.2	100.0
Mitchell Creek	2	20	40.53	18	37.4	23.9	17554	0.00449201	1293	3388	NA	57.5	42.6
Yuba Creek	2	21	58.76	18	19.3	25.7	6536	0.00554350	719	1279	91	77.1	0.0
Elk Lake Outlet	2	22	217.21	33	11.6	15.1	118609	0.00262760	198	453	NA	14.2	5.8
Jordan River	2	23	860.02	191	18.6	17.4	28005	0.00940700	25455	11176	8563	71.2	71.7
Wycamp Lake Outlet	2	24	138.32	12	21.6	20.9	24624	0.00203778	2227	3473	NA	89.2	0.6
Carp Lake River	2	25	119.68	48	16.4	12.0	28449	0.00166000	3747	11803	NA	88.3	100.0
Little Black River	3	26	67.93	12	10.5	15.0	13516	0.00180338	1072	2004	NA	40.2	80.2
Cheboygan River	3	27	3867.70	168	12.8	13.4	160882	0.00698000	47623	31190	45141	76.0	46.1
Elliott Creek	3	28	14.46	78	22.8	12.7	3339	0.00231322	3193	NA	NA	61.2	41.0
Greene Creek	3	29	127.53	60	22.4	27.2	2905	0.00082718	778	NA	NA	84.0	52.8
Black Mallard Creek	3	30	70.87	60	6.8	9.7	23560	0.00253000	4098	NA	NA	86.9	50.0
Ocqueoc River	3	31	382.65	68	18.8	14.8	51058	0.00465722	22269	3030	11901	84.1	66.1
Schmidt Creek	3	32	81.21	136	21.4	25.7	14879	0.00246032	2082	NA	NA	78.6	51.5
Trout River	3	33	95.22	166	17.2	23.2	35363	0.00326381	3298	2945	1764	84.2	49.7
Swan River	3	34	113.42	32	5.9	7.5	50176	0.00307678	6470	NA	5654	85.1	100.0
Grand Lake Outlet	3	35	124.74	16	72.9	32.1	38809	0.00017776	3331	328	NA	78.8	100.0
Long Lake Creek	3	36	147.51	31	1.3	4.6	61169	0.00030870	1488	5339	NA	73.8	87.7
Devil's River	3	37	185.09	54	11.7	10.2	25089	0.00243093	3479	3284	NA	72.5	46.1
Black River	3	38	161.14	96	14.8	21.6	20392	0.00524000	8895	12507	NA	74.2	76.3
Au Sable River	3	39	5197.00	88	8.9	10.0	325189	0.00587000	8877	10312	NA	39.1	93.9
Tawas Lake Outlet	3	40	403.51	164	18.4	14.5	29631	0.00232465	9471	3437	1325	73.1	60.4
Au Gres River	3	41	1012.30	80	16.5	14.7	39543	0.00068721	39541	12113	4695	74.8	48.6
Rifle River	3	42	1144.29	297	11.7	11.8	42281	0.00350659	73434	8255	20823	72.3	48.3
Saginaw River	3	43	16176.03	554	9.4	10.4	105446	0.00221350	120620	10391	18501	66.1	65.6

Table 2.2: Summary of covariates, transformations, and correlations among covariates used in my models. The first column gives covariates used in my models: prop_type1, the mean proportion of transect width classified as Type 1 habitat in QAS surveys; latitude and longitude of each survey waypoint; FinalPowe, stream power; QANN90YLD, the low-flow yield; Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; downDist, the distance to the nearest downstream dam or lake; FOREST_R, the percentage of the riparian corridor classified as forested landcover; and LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology. Remaining columns give: variable type, where each covariate was used as a response variable ("Resp."), classification variable used as the basis for regression tree splits ("Class."), or regressor variable for model fitting ("Reg."); variable transformations; and a correlation table between all possible pairs of covariates used as regressors in the model-fitting process.

			Correlation					
Variable	Variable Type	Transformation	FINAL_POWE	QANN90_YLD	DIST2MOUTH	FOREST_R	LACUSTRINE_R	updist
prop_type1	Resp.	log(prop_type1 + 0.05)						
Latitude	Class.	none						
Longitude	Class.	none						
FINAL_POWE	Class., Reg.	log(FINAL_POWE)	*					
QANN90_YLD	Class., Reg.	none	0.08	*				
DIST2MOUTH	Class., Reg.	none	0.005	0.07	*			
FOREST_R	Class., Reg.	none	-0.09	0.054	0.091	*		
LACUSTRINE_R	Class., Reg.	none	0.039	-0.203	-0.303	-0.248	*	
updist	Class., Reg.	log(1 / (updist + 100))	0.073	0.006	-0.06	-0.096	-0.015	*
downdist	Class., Reg.	log(1 / (downdist + 100))	0.022	0.072	0.191	0.005	0.018	0.02

Table 2.3: Summary of the importance (the number of terminal node regression models in which the covariate appeared) and consistency (the frequency at which the coefficient of the covariate was consistently positive or consistently negative) for covariates in each of the three trees, and in total. The positive and negative signs denote the number of terminal node regression models in which the covariate had a positive and negative coefficient, respectively. Covariates are: Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; FinalPowe, stream power; QANN90YLD, the low-flow yield; FOREST_R, the percentage of the riparian corridor classified as forested landcover; LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology; and downDist, the distance to the nearest downstream dam or lake.

	Spatial Tree			Regressors Tree			Both Tree			Total		
	Importance	-	+	Importance	-	+	Importance	-	+	Importance	-	+
Dist2mouth	9	6	3	11	6	5	12	7	5	32	19	13
Updist	8	5	3	12	5	7	11	6	5	31	16	15
Finalpowe	9	8	1	8	6	2	12	10	2	29	24	5
Qann90Yld	8	4	4	11	5	6	6	3	3	25	12	13
Forestr	8	3	5	10	3	7	6	1	5	24	7	17
Lacustrine	9	5	4	8	6	2	7	5	2	24	16	8
Downdist	5	2	3	9	5	4	8	3	5	22	10	12

Table 2.4: Summary of the number of regression tree partitions based on each classification variable, for each of the three trees and in total. Covariates are: Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; FinalPowe, stream power; QANN90YLD, the low-flow yield; FOREST_R, the percentage of the riparian corridor classified as forested landcover; LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology; and downDist, the distance to the nearest downstream dam or lake.

	Spatial Tree	Regressors Tree	Hybrid Tree	Total
XCOORD	12	-	5	17
YCOORD	10	-	6	16
Qann90Yld	-	6	2	8
Dist2mouth	-	3	4	7
Finalpowe	-	4	1	5
Forestr	-	4	1	5
Updist	-	1	4	5
Downdist	-	1	1	2
Lacustrine	-	1	1	2

Appendix A

Table A1: Details of the fitted regression relationships for each of the terminal tree nodes in model 3, “spatial tree”. Columns give: Node, the ID number of each terminal tree node, corresponding to Fig. 2.2a; n, the number of waypoints assigned to that terminal node; R^2 for the multiple linear regression model fit to data assigned to that terminal node; and the fitted regression equation itself. Covariates are: Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; FinalPowe, stream power; QANN90YLD, the low-flow yield; FOREST_R, the percentage of the riparian corridor classified as forested landcover; LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology; and downDist, the distance to the nearest downstream dam or lake.

Node	n	R^2	Fitted Regression Equation
128	344	0.0532	-0.944 - 0.978 * Forestr
258	95	0.3498	4.17 - 2.82E-5 * Dist2mouth - 0.362 * Lacustrine + 0.452 * Downdist
518	98	0.2045	-1.69
519	91	0.1823	-0.658 - 56.8 * Qann90yld -1.17E-5 * Dist2mouth
65	110	0.2201	-2.21 - 127 * Qann90yld + 1.69 * Forestr
132	270	0.1407	-2.93 - 2.31E-7 * Finalpowe-0.316 * Lacustrine - 0.123 * Updist
133	101	0.1598	-1.79
67	159	0.1913	-2.29 - 3.69E-7 * Finalpowe + 0.0687 * Updist - 0.155 * Downdist
17	187	0.4271	-7.44 - 6.26E-8 * Finalpowe + 61.8 * Qann90yld + 1.47 * Forestr - 0.119 * updist - 0.297 * Downdist
72	133	0.4543	-2.569 - 3.93E-7 * Finalpowe + 1.72 * Forestr - 0.430 * Lacustrine
73	117	0.0471	-2.047
37	201	0.3968	2.44 - 123 * Qann90yld -1.41E-5 * Dist2mouth + 0.778 * Lacustrine + 0.199 * Updist + 0.121 * Downdist
38	131	0.2794	-0.429 - 183 * Qann90yld + 2.77E-5 * Dist2mouth - 0.564 * Forestr + 0.933 * Lacustrine
39	165	0.1384	-2.75 + 105 * Qann90yld - 2.58E-5 * Dist2mouth + 1.13 * Forestr
20	167	0.0897	-1.96
21	207	0.1725	0.193 - 3.23E-6 * Finalpowe - 0.524 * Lacustrine + 0.175 * Downdist
11	160	0.1725	-3.45 - 0.181 * Updist
24	230	0.0617	-3.50 + 8.00E-7 * Finalpowe - 9.99E-2 * Updist
50	308	0.0978	-1.23 - 4.46E-6 * Finalpowe - 3.30E-6 * Dist2mouth - 0.315 * Lacustrine
51	102	0.1690	-3.85 + 208 * Qann90yld + 1.30 * Forestr
13	277	0.2648	0.656 - 7.77E-6 * Finalpowe + 85.6 * Qann90yld + 1.23E-5 * Dist2mouth - 1.39 * Forestr + 0.798 * Lacustrine + 0.243 * Updist
14	150	0.2620	-2.65 + 1.14 * Lacustrine
15	243	0.3868	-2.88 - 1.18E-5 * Finalpowe + 4.07E-4 * Dist2mouth - 0.0550 * Updist

Table A2: Details of the fitted regression relationships for each of the terminal tree nodes in model 4, “covariates-only tree”. Columns give: Node, the ID number of each terminal tree node, corresponding to Fig. 2.3a; n, the number of waypoints assigned to that terminal node; R^2 for the multiple linear regression model fit to data assigned to that terminal node; and the fitted regression equation itself. Covariates are: Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; FinalPowe, stream power; QANN90YLD, the low-flow yield; FOREST_R, the percentage of the riparian corridor classified as forested landcover; LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology; and downDist, the distance to the nearest downstream dam or lake.

Node	n	R^2	Fitted Regression Equation
16	317	0.2653	-0.525 + 3.36E-5 * Finalpowe - 1.153 * Forestr + 2.33 * Lacustrine + 0.0956 * Updist
17	93	0.3723	1.85 - 0.0439 * Qann90yld + 0.298 * Updist
18	369	0.3549	-2.48 - 3.84E-5 * Finalpowe + 76.8 * Qann90yld + 1.15 * Forestr + 1.43 * Lacustrine + 0.0818 * Updist
19	111	0.2858	-22.5 - 355 * Qann90yld + 15.5 * Forestr - 6.76 * Lacustrine - 1.28 * Downdist
40	101	0.1199	-1.45
41	255	0.1541	0.306 - 2.88E-5 * Finalpowe - 160 * Qann90yld - 8.28E-6 * Dist2mouth + 0.0921 * Updist
42	385	0.1486	-2.63 * -1.19E-5 * Dist2mouth + 0.567 * Forestr - 0.441 * Lacustrine - 0.118 * Updist
43	131	0.2644	-0.302 - 1.98E-5 * Finalpowe - 0.185 * Updist + 0.282 * Downdist
44	110	0.2639	-5.00 + 140 * Qann90yld + 2.51E-5 * Dist2mouth - 0.0921 * Downdist
45	152	0.1903	-6.38 + 115 * Qann90yld + 9.32E-6 * Dist2mouth - 0.307 * Downdist
23	101	0.2498	-2.78 - 212 * Qann90yld + 3.52 * Forestr
12	116	0.5115	2.88 + 4.37E-6 * Finalpowe + 176 * Qann90yld + 1.38E-5 * Dist2mouth - 0.581 * Lacustrine + 0.669 * Updist
26	301	0.1770	1.02 - 2.21E-7 * Finalpowe - 9.50E-6 * Dist2mouth + 0.0942 * Updist + 0.166 * Downdist
54	397	0.1899	-8.13 + 45.3 * Qann90yld - 0.841 * Forestr - 0.610 * Lacustrine - 0.697 * Updist + 0.0799 * Downdist
440	235	0.1264	-3.67 + 33.3 * Qann90yld - 0.177 * Updist
882	100	0.1480	2.29 - 1.02 * Lacustrine + 0.350 * Updist
883	89	0.0455	-2.46 + 1.70E-6 * Dist2mouth
442	113	0.2161	3.13 - 6.34E-6 * Dist2mouth - 2.13 * Forestr - 1.60 * Lacustrine + 0.231 * Downdist
443	246	0.2916	-3.17 - 3.63E-6 * Finalpowe - 4.20E-6 * Dist2mouth + 1.14 * Forestr - 0.144 * Downdist
111	244	0.1896	-2.32 - 42.7 * Qann90yld - 5.81E-6 * Dist2mouth + 1.21 * Forestr
7	107	0.5135	-27.6 - 1.81E-6 * Finalpowe + 2.87E-5 * Dist2mouth + 1.60 * Forestr - 0.723 * Updist - 1.63 * Downdist

Table A3: Details of the fitted regression relationships for each of the terminal tree nodes in model 5, "hybrid tree". Columns give: Node, the ID number of each terminal tree node, corresponding to Fig. 2.4a; n, the number of waypoints assigned to that terminal node; R^2 for the multiple linear regression model fit to data assigned to that terminal node; and the fitted regression equation itself. Covariates are: Dist2Mouth, the distance to the mouth of the watershed; upDist, the distance to the nearest upstream dam or lake; FinalPowe, stream power; QANN90YLD, the low-flow yield; FOREST_R, the percentage of the riparian corridor classified as forested landcover; LACUSTRINE_R, the percentage of the riparian corridor classified as lacustrine surficial geology; and downDist, the distance to the nearest downstream dam or lake.

Node	n	R^2	Fitted Regression Equation
64	233	0.0754	0.711 - 5.03 * Qann90yld + 0.138 * Updist
65	253	0.1665	0.185 - 2.03E-5 * Finalpowe - 1.29E-5 * Dist2mouth - 0.486 * Lacustrine + 0.0879 * Downdist
66	118	0.2247	2.45 - 1.28 * Forestr + 0.326 * Downdist
67	118	0.3694	-2.83 - 8.87E-7 * Finalpowe - 2.30E-5 * Dist2mouth - 0.520 * Lacustrine - 0.226 * Downdist
17	249	0.1104	-1.42 - 54.5 * Qann90yld - 7.01E-6 * Dist2mouth
36	169	0.1274	-2.70 - 2.58E-7 * Finalpowe - 2.02E-5 * Dist2mouth - 0.161 * Downdist
148	118	0.1128	-1.80
298	108	0.5323	1.57 - 2.58E-5 * Dist2mouth + 0.167 * Updist + 0.0826 * Downdist
299	87	0.3395	-2.27 + 18.5 * Lacustrine - 0.171 * Updist + 0.0771 * Downdist
75	212	0.3398	2.89 - 5.74E-7 * Finalpowe + 1.011 * Forestr + 0.418 * Updist + 0.0954 * Downdist
76	112	0.1489	-2.08 + 2.01E-4 * Dist2mouth
77	87	0.2378	-1.71
78	96	0.4877	-16.2 + 319 * Qann90yld + 2.12 * Forestr - 1.19 * Updist
79	89	0.1593	-2.59 - 4.56E-6 * Finalpowe + 1.86E-5 * Dist2mouth
20	105	0.0757	-1.37
21	131	0.5005	-5.28 - 7.51E-7 * Finalpowe + 149 * Qann90yld - 0.297 * Updist
22	104	0.1401	-1.67 + 1.83 * Forestr + 0.161 * Updist
46	183	0.4039	-7.29 + 2.66E-5 * Dist2mouth - 1.12 * Lacustrine - 0.335 * Downdist
47	163	0.0705	-2.99 - 0.128 * Updist
48	239	0.1084	-1.87 - 7.65E-7 * Finalpowe - 0.360 * Lacustrine
49	171	0.2377	-1.64 * 1.35E-6 * Finalpowe - 9.51E-6 * Dist2mouth + 0.544 * Forestr - 0.830 * Lacustrine
25	230	0.2561	-2.43 - 3.27E-6 * Finalpowe + 65.9 * Qann90yld - 6.45E-6 * Dist2mouth - 0.116 * Updist
52	101	0.2856	-2.52 - 3.20E-6 * Finalpowe + 156 * Qann90yld
53	81	0.2789	-3.02 + 1.02E-5 * Finalpowe + 1.96E-5 * Dist2mouth
27	95	0.2880	-0.561 - 177 * Qann90yld + 1.63 * Forestr + 0.241 * Updist
14	150	0.2620	-2.65 + 1.14 * Lacustrine
15	243	0.3868	-2.88 - 1.18E-5 * Finalpowe + 4.07E-4 * Dist2mouth - 0.0550 * Updist

Chapter 3

River Network Structure Affects Sea Lamprey Distribution

Abstract

I investigated the influence of river network structure on the distribution of larvae of sea lamprey (*Petromyzon marinus*), an anadromous, semelparous parasitic fish that is invasive in the Laurentian Great Lakes. Adult lamprey migrating upstream to spawn follow a pheromone released by instream larvae. The size (i.e. flow) of a tributary dilutes the concentration of larval pheromone. The downstream propagation pattern of larval pheromone through a watershed thus depends on patterns in the relative sizes of confluent tributaries. I hypothesized that river network structure by itself must therefore influence the migration decisions of adult lamprey following the larval pheromone. To test this hypothesis, I formulated a simple behavioral rule for how adults might choose between confluent tributaries on the basis of pheromone concentration and tributary size. I then created an individual-based computational model of this system, and performed two sets of computational experiments. First, I initialized the model using randomly-generated river networks, and quantified how river network shape influenced the distribution of sea lamprey in the model systems. I found a strong positive relationship between network diameter (a measure of network shape) and the skewness of the distribution of larvae across reaches in the model-generated networks. Larvae in the model aggregate over time, and the degree and rate of this aggregation depends on network diameter. I observed these patterns over several orders of magnitude of the model parameter describing the sensitivity of adults to larval pheromone, indicating that these patterns are robust to changes in model parameterization. Second, I initialized the model using a river network based on the Muskegon River, Michigan, and compared larval distribution by reach generated by the model to available survey data from the Muskegon River. I found a significant correlation between model-generated larval abundance and the field-measured larval densities (rho = 0.46; $p < 0.005$). In addition, I observed an inverse relationship between the drainage area of each subwatershed of the Muskegon River and the skewness of the distribution of larvae across all reaches, which demonstrates the magnitude of the effect of path-dependent feedbacks influencing larval abundance. In the model, patterns of larval distribution across a watershed are a result of a complex system of emergent feedbacks contingent on each unique network structure and the past migratory and spawning behavior of adults. The processes in the model are likely complimentary to models in which the local-scale distribution of sea lamprey is determined by the distribution of preferred larval and adult spawning habitat. My results reinforce the importance of considering rivers as "riverscapes", wherein rivers are continuous, hierarchical, and heterogeneous, and spatial pattern at several scales influences the population dynamics of resident fishes.

Introduction

One of the most obvious spatial patterns in a watershed is the river's own network structure, i.e. its branching pattern of tributaries. Each river network is unique and inherently large-scale; consequently, if network structures shape system-wide processes in rivers, there are important implications for both spatial uniqueness and the minimum sufficient spatial scale for ecological studies and management. This issue motivates much recent work that demonstrates how river network structure influences the distribution of substrates (Benda et al. 2004a, b; Kiffney et al. 2006; Rice et al. 2006), fish diversity (Osborne and Wiley 1992), population size of salamanders (Lowe and Bolger 2002), disturbance propagation through a network (Jones et al. 2000) and species' recovery from disturbances (Fagan 2002, Fagan et al. 2002), movement rate of individuals (Johnson et al. 1995), and rates of predator-prey interaction (Cuddington and Yodzis 2002).

In this study, I examine the effect of river network structure on the distribution and relative abundance of sea lamprey larvae (*Petromyzon marinus*) in a watershed. The sea lamprey is a primitive fish with an anadromous, semelparous lifecycle characterized by a stream-resident, filter-feeding larval stage generally three to five years in duration, an ocean- or lake-dwelling juvenile parasitic stage, and a non-feeding sexually mature adult stage during which lamprey migrate from oceans or lakes into rivers to spawn (Beamish 1980). The sea lamprey is an invasive species in the Laurentian Great Lakes with a long history of negative impacts on salmonines (Smith and Tibbles 1980), particularly lake trout (*Salvelinus namaycush,* Coble et al. 1990). Lamprey populations are controlled primarily with a chemical lampricide that targets larvae and is applied to the most heavily populated streams every four years. High management costs (> $14 million/year; Slade et al. 2003, Christie et al 2003, Jones 2007) provide a strong practical motivation for understanding as much as possible about sea lamprey ecology.

Adult sea lamprey migrate upstream to spawn by following two pheromones. First, both male and female spawning-phase lamprey migrating upstream are attracted by a pheromone released by stream-resident larvae (Li et al. 1995, Bjerselius et al. 2000, Vrieze and Sorensen 2001, Sorensen and Vrieze 2003). Adult females within the stream also follow a sexual pheromone released by spermiated males who have recently completed the upstream migration (Li et al. 2003). The larval pheromone controls broad-scale migration through the watershed; the sexual pheromone allows females to find nesting males. I consider only the effect of larval pheromone in this paper.

Adult lamprey migrate upstream through a series of confluences en route to spawning sites, and the relationship between network structure and the directional choices adults make at confluences likely has important consequences for where they ultimately spawn. The size (i.e. flow) of a tributary dilutes the concentration of larval pheromone. The downstream propagation pattern of larval pheromone through a watershed thus depends on patterns in the relative sizes of confluent tributaries. I hypothesize that river network structure by itself must therefore influence the migration decisions of adult lamprey, and in turn their ultimate spawning location and the distribution of eggs and larvae. Furthermore, each cohort alters the pheromone landscape for subsequent generations. This implies that the distribution of larvae in a watershed must result from a system of emergent feedbacks related to the behavior of individual adults as they migrate through the river network. Patterns of larval densities across a watershed should therefore be contingent on both fish behavior and network structure (i.e. contingent on location and history, Malanson 1999), and arise from the interaction of processes that span several spatial and temporal scales.

In this paper I develop an individual-based model and perform a series of computational experiments to explore how river network structure could influence the distribution of sea lamprey in a watershed. First, I used information from published literature and from a state-wide dataset to formulate behavioral rules describing how adult sea lamprey might choose between confluent tributaries as they migrate upstream to spawn and how larval sea lamprey drift downstream. I created a model in which individual sea lamprey use those rules in an abstract representation of a river network. Second, I carried out computational experiments in which the model was initialized with randomly-generated river networks, and quantified how river network structure influenced the distribution of sea lamprey in the model systems. Third, I tested the model with a river network structure based on the Muskegon River, a tributary to Lake Michigan, and compared model output with historical data of larval density available from surveys in that watershed.

Methods

The pattern of interest in this study is the distribution of sea lamprey larvae across a watershed at the resolution of a reach, defined here as a section of river between two confluences. I hypothesized that larval distribution at the resolution of the reach emerges from the interaction of river network structure, larval pheromone, and the behavioral choices of adults migrating upstream to spawn. My model consists of an abstract representation of a river network and discrete time biological submodels that describe the upstream migration of adult sea lamprey, and the downstream movement of larvae and larval pheromone. An overview of the model follows.

Computational Model

The model runs at a daily timestep. The basic spatial unit in the model is the reach, defined here as the section of river between two confluences. Each reach is characterized by 1) an index of size (upstream drainage area), 2) pheromone concentration, 3) number of sea lamprey larvae of each year class, 4) number of adult sea lamprey, and 5) links to the one downstream and two (or zero) upstream reaches. The river network consists of a hierarchical arrangement of *n* such reaches.

Each year, female adult sea lamprey begin the spawning migration at the mouth of the river and proceed upstream until they spawn, after which they die (Fig. 3.1). Adult females are modeled as individuals, and are characterized by their location (i.e. the reach they occupy at timestep *t*) and a Boolean variable describing whether they have spawned. For the sake of simplicity I assumed that all females find a nest and spawn.

Larvae are modeled as identical members of a year class, and for each reach I model the number of larvae of age zero to four. Larvae in natural populations drift downstream over time at a seasonally-variable yet unknown rate (Applegate 1950, Manion and McLain 1971). In my model, larvae drift downstream at a constant rate, where 30% of the larvae move downstream one reach each year. Larvae in the model are otherwise inactive (i.e. any movement occurs below the spatial resolution of the model), and serve mainly as a pheromone source to which adults are attracted upstream. At age four, larvae mature and leave the system.

Decision Rules

When migrating upstream and encountering a confluence, adult sea lamprey must decide which tributary to follow. As extensive exploration of each upstream reach is likely energetically too costly, lamprey must make their decision based on information available at the confluence. I formulated a basic decision rule governing upstream migration of spawning adults by considering two biological observations. First, spawning-phase lamprey migrating upstream are attracted by a pheromone released by stream-resident larvae (Li et al. 1995, Bjerselius et al. 2000), and laboratory experiments demonstrate that adults preferentially swim towards higher pheromone concentrations (Vrieze and Sorensen 2001). Second, lamprey spread readily through watersheds recently treated with lampricide, where chemical treatment results in a watershed largely without pheromone signals. Under these conditions I assume that adults must be biased towards the larger tributary at each confluence; if this were not the case, highest larval densities in all watersheds would occur in small tributaries near the river mouth, a pattern not observed in the USFWS dataset described later in this chapter. Thus, the decision rule for choosing between two confluent tributaries *a* and *b* is a rule where the probability of selecting a given path is a linear function of both pheromone concentration and tributary size, written as:

$$P(a) = \frac{[Ph]_a + \alpha S_a}{[Ph]_a + [Ph]_b + \alpha S_a + \alpha S_b} \tag{1}$$

where P(*a*) is the probability of choosing tributary *a*, $[Ph]_a$ is the pheromone concentration in tributary *a*, S_a is the size of tributary *a* (i.e. flow or discharge), and α is a model parameter for weighting the importance of tributary size relative to pheromone concentration. Discharge scales linearly with upstream drainage area (Dunne and Leopold 1978); in the model, the drainage area of each reach (km^2) is a proxy for discharge that is derivable solely from river network structure.

The equation (1) itself can be parameterized to represent a range of possible choice rules. When $\alpha = 0$, adult lamprey base their migration decisions exclusively on pheromone concentration. When α is very large, adult lamprey effectively ignore pheromone and base their migration decisions nearly on the relative sizes of confluent tributaries. Intermediate values of α represent the range of conditions in which adults consider both pheromone and tributary size in their upstream migration.

In a natural system, the selection of spawning depends on acceptable habitat, which consists of a substrate of medium gravel with steady, unidirectional stream flow between the ranges of 1.1 to 1.3 m/s in the most heavily used sites (Applegate 1950, Morman et al. 1980). Adults also migrate upstream considerable distances before spawning, such that they must bypass otherwise acceptable spawning substrates en route to their eventual spawning destination. This suggests that the selection of spawning must also be related to distance traveled from the river mouth or the size (i.e. flow) of the river segment. In the model, the probability of spawning is set as the inverse of the size of the reach (upstream drainage area in km^2), so that the probability of spawning is highest for small upstream reaches. Spawning always occurs if an adult has migrated to a reach with no further upstream reaches.

The distribution of larval pheromone is recalculated after spawning during each day in which spawning occurs. Each larva in the model creates one 'unit' of pheromone, and each reach has unit drainage area and contributes one 'unit' of discharge to all downstream reaches. For each 1st-order reach (i.e. a reach with no further upstream reaches), pheromone concentration is calculated as:

$$[Ph] = \frac{n * r}{Q}$$

where n is the number of resident larvae, r is the rate of pheromone production per larvae (units of volume of pheromone per time), and Q is the discharge (i.e. flow) of the river in that reach (units of volume of water per time). I assume that the pheromone load propagates downstream with no losses, such that the pheromone concentration for each 2nd-or greater order reach is calculated as the total number of resident and upstream larvae divided by the discharge (i.e. the number of upstream reaches).

Experimental Framework

I conducted two sets of computational experiments. First, I ran the model using randomly-generated river networks of differing shape. My goal in this set of experiments was to quantify the relationship between river network structure and larval sea lamprey distribution. Towards this goal, I also sought to understand the general behavior of the model, and the relationships between biological submodels, river network structure, and model parameterization. In a second set of experiments, I ran the model using a single river network based on the structure of Muskegon River, MI, and compared larval abundance generated by the model to available survey data. My goal in this set of experiments was to use the Muskegon River as a case study of the ability of the model to reproduce empirical data.

The basic output of the model is the number of larvae in each reach during each timestep. Model behavior is stochastic and depends on the probabilistic migration decisions of adults, such that each model run is unique. For some experiments, I focus on and describe the variability in larval distribution across model runs resulting from this stochasticity. For other experiments I pooled results from multiple model runs to create aggregate measures describing the distribution of larvae in the river networks.

Parameterization

I used USFWS reports and other literature on larval abundance, fecundity, and egg survival to produce an estimate of the number of female spawners entering the Muskegon River, MI each year, and an estimate of the number of larvae produced by each female. I assumed an average year class of 1,459,564 larvae in the system (based on Schleen and Klar 2000, Klar and Young 2004). I estimated that 0.75% of eggs laid eventually mature (Applegate 1950), and that females produce an average of 65,000 eggs (Applegate 1950, Manion and McLain 1971, Manion 1972), with the result that each adult female ultimately produces on average 488 larvae. I divided average year class size by average female production of larvae (488) to estimate that on average 2990 females enter the Muskegon each year.

For the random network experiments, I assumed that the number of females entering a river system each year would be proportional to the total drainage area of the system. Accordingly, I multiplied the number of females entering the Muskegon each year (2990) by the ratio of the drainage area of each random network to the drainage area of the Muskegon (46.72) to parameterize the model with 64 females entering a river network each year.

Random River Network Experiments

I generated random river networks with 49 reaches of equal length, and measured river network shape as the diameter, or maximum link distance, of each network. Network diameter is here defined as the number of reach segments in the path between the mouth of the network and the most distant first-order tributary (Fig. 3.2). Very short and compact or “branchy” river networks have a low diameter, while elongated river networks have a high diameter.

Network diameter — a measure commonly used in network topology studies (Jackson 2008) — differs from the bifurcation ratio more commonly used to describe channel network shape by geomorphologists and fluvial ecologists (Knighton 1998). Although network diameter is positively correlated with bifurcation ratio, this relationship is not one-to-one (i.e. two networks of the same diameter could have different bifurcation ratios).

My algorithm for generating random river networks began with an initial downstream reach (the "mouth" of the river) and its two immediate upstream tributaries, one of which was arbitrarily defined to be the mainstem of the river. The non-mainstem branch of the network was branched into two tributaries with probability p, or left as a 1st-order tributary with probability 1-p. The mainstem branch of the network is branched into two tributaries with probability 1, and one of these two branches is again arbitrarily labeled the mainstem. The process was repeated at subsequent network depths (where depth is the length of the path from the river mouth) until the total number of reaches reached 49. The process is stochastic, and a variety of network configurations are possible for each value of p. To create river networks of further varying shape, I varied the parameter p between 0 (which led to unbranchy networks with the maximum diameter possible, 25) and 1 (which led to branchy river networks with the minimum diameter possible, 6).

I performed a total of six experiments with randomly-generated river networks (Table 3.1). For the first three experiments (exp. 1 - 3), I analyzed the short-term behavior of the model by running the model for four years and measuring larval distribution at the end of each four-year model run. Understanding larval distribution after four years is most relevant to sea lamprey management via chemical lampricide, where the most heavily populated reaches are treated every four years and prioritized by larval abundance. For these three model runs, I was interested in the degree of aggregation of larvae within the river network, i.e. the extent to which a small number of reaches contained high numbers of larvae. I measured this aggregation by calculating the skewness of the distribution of larvae across all reaches in the network. Model runs with high positive skewness had one or several reaches with high larval densities and many other reaches with low or absent densities, while model runs with lower (but positive) skewness had smaller differences in larval density between reaches.

For exp. 1, I examined the effect of the model parameter α on the mean relationship between network diameter and larval distribution by calculating the mean skewness of larval distribution at each combination of α and network diameter. For exp. 2, I examined stochasticity in the relationship between network diameter and larval distribution by fixing $\alpha = 1$ and running the model 30 times at each network diameter.

This stochasticity results from both probabilistic adult behavior and the fact that there are multiple possible network configurations for a given diameter. For exp. 3, I isolated the effects of stochasticity resulting from probabilistic adult behavior by running the model 1500 times with $\alpha = 1$ and a fixed river network of diameter 8 (Fig. 3.2c).

For the next three experiments (exp. 4 - 6), I ran the model for longer time periods and described the changes in larval distribution over time. This is an unlikely scenario in the Laurentian Great Lakes, where the most heavily-populated reaches are treated with chemical lampricide every four years. However, model runs longer than four years provide important information about both un-managed lamprey populations and the dynamics of the model. For these three model runs I were again interested in the degree of aggregation of larvae within the network; I measured the degree of aggregation as the total number of reaches inhabited by larvae during each timestep.

For exp. 4, I examined the mean effect of network diameter on the distribution of larvae over time. For each of network diameters 6,7,8, and 25, I fixed $\alpha = 1$ and measured over 50 model runs the mean number of reaches inhabited each year over 50 years. For exp. 5, I examined stochasticity in the change in larval distribution over time by examining three individual model runs each of length 100 years. For exp. 6, I examined the effect of the model parameter α on the change in the number of reaches inhabited over time for a fixed network of diameter 8 (Fig. 3.2c).

Muskegon River Experiments

I ran the model using a river network based on the structure of the Muskegon River, and compared the distribution predicted by the model to densities of sea lamprey larvae in samples collected during USFWS surveys in the Muskegon River. The Muskegon River is the largest Lake Michigan tributary in the State of Michigan, with a length of 341 km and a drainage area of over 5,900 km^2 (O'Neal 1997). The modeled section of the river was the mainstem between the mouth of the system near Muskegon, MI, and Croton Dam (a migratory barrier for adults), and included all tributaries to that section of the mainstem (n = 167 reaches; Fig. 3.3). I excluded from the model five tributaries to Muskegon Lake classified as Areas of Concern and therefore uninhabitable for sea lamprey (M.J. Wiley, personal communication), and represented the braided wetlands area of the lower Muskegon as a single channel.

Measurements of network diameter are relative to the total number of reaches in a network, with the consequence that there is no straightforward relationship between the structure of the Muskegon River and the diameters of the randomly-generated networks. Among rivers of its size, the Muskegon is of approximately moderate branchiness (Fig. 3.4). The Muskegon River below Croton Dam (i.e. treating the mainstem at Croton as a 1st order stream) has a bifurcation ratio of 4.15.

I used georeferenced data describing the distribution and density of larval sea lamprey collected between 1998 and 2003 during Quantitative Assessment Surveys (QAS; Slade et al. 2003) conducted by United States Fish and Wildlife Service (USFWS) personnel. During the QAS surveys, larval sea lamprey were collected by electrofishing 15 m^2 plots using the methodology described by Slade et al. (2003). I described each reach in the Muskegon River by the mean larval density in plots of preferred habitat (i.e. “Type 1” habitat in Slade et al. 2003) from all samples in that reach, scaled by the number of years elapsed since lampricide treatment.

For each model run in the Muskegon River we began with a watershed with no resident larvae and ran the model for the equivalent of four years. For each reach, I calculated a mean larval abundance after four years across 500 models runs. I used Spearman’s rank coefficient to measure correlation between mean larval densities in the USFWS dataset and mean modeled larval abundance in each reach.

Model behavior depends primarily on α, the parameter that weights the importance of tributary size relative to pheromone concentration for migrating adults choosing a tributary. For the Muskegon River model runs, I chose a final value of α = 0.36 by running the model 30 times at each of a range of α values, and for each value of α comparing the correlation between model output and USFWS survey data.

Results

Random Network Experiments

In experiment 1, I observed a strong positive relationship between river network diameter and the skewness of the distribution of larval abundance across all 49 reaches (Fig. 3.5a). I observed this pattern across four orders of magnitude in the parameter α, which weights the relative importance of tributary size and pheromone concentration for migratory adults ($\alpha = 0.01$ to $\alpha = 100$). This relationship between network diameter and larval distribution disappeared at very large values of α (i.e. > 1000), a clearly unrealistic situation in which adult lamprey pay little attention to pheromone concentration during migration.

There was considerable variability across model runs in the skewness of larval distribution for each network diameter (experiment 2; Fig. 3.5b). This variability was due to both differences in network structure between the generated network in each model run (i.e. there are many unique network structures with the same diameter), and the unique history of adult migration during each model run. Repeated model runs with a fixed river network of diameter 8 exhibited a similar range of skewness (experiment 3; Fig. 3.5c). Variability between model runs in this experiment was due solely to the probabilistic behavior of adults during spawning runs.

Larvae in the model aggregated over time, and the degree and rate of aggregation depended on the river network structure (experiment 4; Fig. 3.6a). In the highest diameter network ($D = 25$), larvae aggregated into an average of three reaches in approximately 10 years. In the lowest diameter networks ($D = 6$), larvae eventually populated on average 11 or 12 reaches, and the system took approximately 40 years to reach this state. Although the mean behavior of the model exhibited these ‘stable’ larval spatial distributions, adults during each model run make probabilistic migration decisions that enable them to repopulate uninhabited reaches at any time during the model run (experiment 5; Fig. 3.6b). Some amount of colonization of uninhabited reaches occurred in each of the model runs in experiment 5. Last, in the model runs using a fixed river network of diameter 8, larvae aggregated over time across several orders of magnitude of the parameter α (experiment 6; Fig. 3.6c). Larvae failed to aggregate over time only when adults paid little attention to pheromone concentration during migration (i.e. $\alpha > 1000$).

Muskegon River Experiments

The parameterized model successfully reproduced the general pattern of larval distribution and relative abundance in the Muskegon River. I found a significant correlation between the log of the modeled mean larval abundance and the log of the field-measured mean larval density in the USFWS dataset (Spearman's rho = 0.46, $p < 0.005$; Fig. 3.7). The model reproduced very high larval abundances in the upper mainstem reaches of the Muskegon (Fig. 3.3), along with a range of medium and lower abundances in other locations in the watershed.

Larval abundance in each subwatershed of the Muskegon differed between model runs. Furthermore, subwatersheds differed in the shape of their frequency distribution of larval abundance across model runs. The upper reaches of the mainstem and large subwatersheds (e.g. Brooks Creek; see Fig. 3.3) consistently exhibited high larval abundance and a symmetric frequency distribution (Fig 3.8a, b). Medium-sized subwatersheds (e.g. Sand Creek) exhibited positive skewness in the frequency distribution of larval abundance across model runs (Fig. 3.8c). This pattern was strongest in very small subwatersheds (e.g. Reach 143). Low larval abundance was common in these subwatersheds, but infrequently the subwatershed accumulated moderate larval abundance (Fig. 3.8d).

Discussion

My results demonstrate that river network structure by itself has a strong influence on the distribution of sea lamprey larvae. The general result that network structure influences larval distribution was very robust to changes in model parameterization, and disappeared only when adults paid little attention to larval pheromone. This scenario seems unrealistic given the empirical studies demonstrating the behavioral response of adult sea lamprey to larval pheromone (Li et al. 1995, Bjerselius et al. 2000, Vrieze and Sorensen 2001, Sorensen and Vrieze 2003).

Further evidence that adults do indeed pay attention to both larval pheromone and tributary size is provided by the best-fit α for the Muskegon River ($\alpha = 0.36$), which is neither zero (evidence that adults ignore tributary size) nor infinite (evidence that adults ignore pheromone). Finally, the correlation between model and data for the Muskegon River (rho = 0.46) is encouraging given the relative biological simplicity and coarse spatial resolution of the model.

My model omits many factors known to influence larval distribution, including density-dependent (Purvis 1979, Weine and Pajos 1998) and independent (Applegate 1950) variation in larval growth and survival, variation in the duration of the larval stage (Applegate 1950), and temporal and spatial variation in larval drift rates (Manion and McLain 1971). However, the general result that network structure influences larval distribution rests on only three premises: 1) network structure influences the downstream propagation of larval pheromone, 2) migrating adults are sensitive to the resulting pheromone landscape, and 3) differences among reaches in the number of spawning adults results in corresponding variation in larval abundance. Elements of lamprey ecology omitted from my model would negate the effect of network structure on larval distribution only if they invalidated one of these three premises, e.g. if a paucity of larval habitat constrained larval distribution to the extent that it was entirely independent of patterns of adult migration (see Chapter 4).

Implications for Ecology and Management

In my model, patterns of larval distribution across a watershed are a result of a system of emergent feedbacks contingent on each unique network structure and the past migratory and spawning behavior of adults. These patterns arise when microscopic interactions (larvae interacting with adults through the proxy of pheromone) generate macroscopic patterns (the distribution of larvae in a watershed) via nonlinear feedbacks. The population of sea lamprey inhabiting a watershed is therefore a complex adaptive system (Levin 2002), with important implications for predictability and fisheries management (Mahon et al. 2008). Small subwatersheds in the Muskegon River, for example, are most susceptible to path-dependent feedbacks which place intrinsic limits on the predictability of larval populations in these subwatersheds.

The mechanisms in my model are different than those typically considered to be important for sea lamprey distribution and population dynamics. The preferred habitat for larval sea lamprey (Applegate 1950, Potter et al. 1986) and spawning (Applegate 1950) is well defined, and the local-scale distribution of larvae is typically assumed to be determined by the availability of preferred substrate habitat. In the model, the large-scale distribution of larvae across a watershed is a function of a complex system of emergent feedbacks constrained by river network structure. Although the ability of the model to predict larval densities may be moderate, my results suggest that local-scale habitat models may never be solely adequate as predictive models for sea lamprey. Understanding patterns of river network structure and the ecological processes constrained by those patterns (i.e. larval pheromone and the migratory choices of adult sea lamprey) provides important additional information to that based only on descriptions of available reach habitat.

I hypothesize that local habitat preferences and the migratory behavior of adults are complimentary processes operating at different spatial and temporal scales. In watersheds with no or minimal larval populations and thus no pheromone, adult migration is determined primarily by patterns of the relative sizes of confluent tributaries (i.e. network structure). Adults in these systems may be biased towards the largest tributaries, but still disperse widely. Such behavior likely contributed to the success of the sea lamprey as an invasive species in the Laurentian Great Lakes. Once lamprey have colonized a watershed, good larval habitat may confer increased growth and survival to resident larvae, who in turn create a strong and consistent pheromone signal that draws adults to areas of the watershed with good habitat. Larvae in my model aggregate over time, and it is likely that the distribution of larval and spawning habitat mediates the location and rate of this aggregation. Disturbances that reduce larval populations (e.g. natural physical disturbances or the application of chemical lampricide) interrupt this process by reducing larval pheromone concentration and enabling wider dispersal of adults through a watershed. Larvae in such systems may populate a greater range of marginal habitats (e.g. Type II habitat in Slade et al. 2003), and the distribution of larvae may exhibit less correlation with the distribution of preferred habitat (e.g. Type I habitat in Slade et al. 2003).

The importance of multiple scales of spatial pattern, and the recognition that different biological processes may be constrained by each of these scales of spatial pattern, reinforces views of rivers as “riverscapes” (Fausch et al. 2002) with many interacting spatial and temporal scales of variability (Wiens 2002, Baker and Wiley 2009). Larval distribution in each watershed is contingent on location (each watershed’s unique river network) and history (Malanson 1999).

Model Behavior and Path-Dependent Feedbacks

Both mean modeled larval abundance and the frequency distribution of larval abundance across model runs varied among subwatersheds in the Muskegon River (Fig. 3.8). Tributaries with the largest drainage areas had the highest larval abundance, and a symmetric frequency distribution of a larval abundance across model runs (e.g. Upper Mainstem, Brooks Cr.). Tributaries with intermediate drainage area had lower larval abundance and slight right skewness in the distribution of larvae across model runs (e.g. Sand Cr.). Tributaries with the smallest drainage area (e.g. Reach 143) had the lowest larval abundances and highest right skewness in the distribution of larvae across model runs.

Mean larval abundance across model runs for each tributary is correlated with drainage area of the tributary. This correlation is attributable to the fact that migrating adults in the model are biased towards the larger tributary at each confluence. The general pattern that results is that larger tributaries accumulate greater numbers of larvae.

The skewness of the frequency distribution of larval abundance across model runs is negatively correlated with tributary drainage area. For the first year class of spawners in any model run, the probability of migrating to a tributary is a function solely of patterns in the relative sizes of tributaries and is constant across models runs. For all subsequent year classes of spawners, their migration decisions depend in part on the distribution of larval pheromone that results from the probabilistic migration decisions of previous generations of spawners. The model thus contains path-dependent feedbacks whereby reaches that accumulate larvae early in a model run have a greater likelihood of attracting subsequent year classes of adults.

For each tributary, the relative contributions of larval pheromone and tributary size to attracting subsequent year classes of spawners determines the magnitude of these path-dependent feedbacks and the resulting skewness of the frequency distribution of larval abundance across model runs. Tributaries with large drainage areas (e.g. the Mainstem of the Muskegon R.) are always attractive to spawners because of their large size, and their frequency distribution shows minimal skewness. The smallest tributaries (e.g. Reach 143) are typically unattractive to spawners due to their small size. However, in model runs where a larger-than-average number of adults spawn in such a reach, the pheromone released by the new cohort of larvae appreciably increases the attractiveness of the reach to subsequent year classes of adults, initiating a feedback loop driving larval abundance higher.

Model Applications

In the experiments with randomly generated networks I observed a positive relationship between network diameter and larval aggregation. An appealing application would be the ability to measure the network diameter of a river and then make quantitative predictions of larval distribution. However, network diameter is one of only several possible measures of network structure (Jackson 2008), and furthermore is a meaningful measure of comparing river networks with the same number of reaches. Real watersheds differ in size (drainage area) and drainage density (the number of tributaries per drainage area), so any measure of network structure must be scalable with network size. Scaling with size may be nonlinear for both the measure of network structure and the dynamics of this model, which may complicate the search for relationships between river network structure and lamprey densities in real systems. Understanding the best measures for describing spatial pattern may be the most significant challenge in relating sea lamprey population dynamics to river network structure, but must occur before ecologists can develop criteria to relate pattern (river network structure; lamprey distribution) to its causes and consequences (Levin 1992).

Despite these lingering questions about the details of the relationship between network structure and larval distribution, the Muskegon R. model runs illustrate several patterns that are likely generalizable to other watersheds. Adults in the model are biased towards the larger tributary at each confluence; as a result, larvae accumulate in the largest subwatersheds, particularly where a barrier to further upstream migration results in a concentration of spawners just below that barrier (e.g. Croton Dam). The network context of each reach is also important: reaches of the same size, and containing the same amount of larval and spawning habitat, may accumulate very different larval abundances depending on their location in a river network.

A final consideration is the skewness of the frequency distribution of larval abundance across model runs, which I suggest is due to path-dependent feedbacks in the model. These path-dependent feedbacks will place a fundamental limit on the predictability of larval densities in some reaches, and so better characterizing when and how they occur is an important step towards building predictive models of larval distribution.

References

Applegate, V.C. 1950. Natural history of the sea lamprey (*Petromyzon marinus*) in Michigan. Spec. Sci. Rept. U.S. Fish and Wildlife Serv. 55, 1-237.

Baker, ME and MJ Wiley. 2009. Multiscale controls of flooding in riparian forests of Lower Michigan. *Ecology* 90:145-159.

Beamish, F. W. H. 1980. Biology of the North American anadromous sea lamprey, *Petromyzon marinus*. Can. J. Fish. Aquat. Sci. 37: 1924 - 1943.

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

Benda, L, Andras, K., Miller, D., and Bigelow, P. 2004. Confluence effects in rivers: Interactions of basic scale, network geometry, and disturbance regimes. Water Resources Research. 40, W05402, doi:10.1029/2003WR003583

Bjerselius, R., Li, W., Teeter, J.H., Seelye, J.G., Johnsen, P.B., Maniak, P.J., Grant, G.C., Polkinghorne, C.N., and Sorensen, P.W. 2000. Direct behavioral evidence that unique bile acids released by larval sea lamprey (*Petromyzon marinus*) function as a migratory pheromone. Canadian Journal of Fisheries and Aquatic Sciences. 57, 557 - 569.

Bruns, D.A., Minshall, G.W., Cushing, C.E., Cummins, K.W., Brock, J.T., and Vannote, R.K. 1984. Tributaries as modifiers of the river continuum concept: Analysis by polar ordination and regression models. Arch. Hydrobiol. 99, 208-220.

Christie, G.C., Adams, J.V., Steeves, T.B., Slade, J.W., Cuddy, D.W., Fodale, M.F., Young, R.J., Kuc, M., and Jones, M.L. 2003. Selecting Great Lakes streams for lampricide treatment based on larval sea lamprey surveys. J. Great Lakes Res. 29, 152-160.

Coble D.W., Bruesegitz R.E., Fratt T.W., Scheirer J.W. 1990. Lake trout, sea lampreys, and overfishing in the upper Great Lakes: A review and reanalysis. Transactions of the American Fisheries Society: Vol. 119, No. 6 pp. 985–995

Cuddington, K.M. & Yodzis, P. (2002). Predator–prey dynamics and movement in fractal environments. Am. Nat., 160, 119–134.

Dunne, T., and Leopold, L.B. 1978. Water in environmental planning: New York, W.H. Freeman and Company, 818 p.

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

Fagan, W.F., Unmack, P.J., Burgess, C. & Minckley, W.L. (2002). Rarity, fragmentation and extinction risk in desert fishes. Ecology, 83, 3250–3256.

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

Fisher, S.G. 1997. Creativity, idea generation, and the functional morphology of streams. Journal of the North American Benthological Society. 16, 305-318.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz and G. Huse. 2006. A Standard Protocol for Describing Individual-Based and Agent-Based Models. Ecological Modelling. 198: 115-126.

Jackson, M.O. 2008. Social and economic networks. Princeton University Press, Princeton, New Jersey.

Johnson, A.R., Hatfield, C.A. & Milne, B.T. (1995). Simulated diffusion dynamics in river networks. Ecol. Model., 83, 311–325.

Jones, J.A., Swanson, F.J., Wemple, B.C., and Snyder, K.U. 2000. Effects of roads on hydrology, geomorphology, and disturbance patches in stream networks. Conservation Biology. 14, 76-85.

Jones, M.L. 2007. Toward improved assessment of sea lamprey population dynamics in support of cost-effective sea lamprey management. J. Great Lakes Res. 33, 35-47.

Kiffney, P.M., Greene, C.M., Hall, J.E., and Davies, J.R. 2006. Tributary streams create spatial discontinuities in habitat, biological productivity, and diversity in mainstem rivers. Can. J. Fish. Aquat. Sci. 63: 2518-2530.

Klar, G.T, and Young, R.J. 2005. Integrated management of sea lampreys in the Great Lakes 2004. Annual Report to the Great Lakes Fishery Commission.

Lancaster, J., and Hildrew, A.G. 1993. Flow refugia and the microdistribution of lotic macroinvertebrates. J. N. Am. Benth. Soc. 12: 385-393.

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

Levin, S.A. 2002. Complex adaptive systems: Exploring the known, the unknown, and the unknowable. Bulletin (new series) of the American Mathematical Society. 40, 3 - 19.

Li, W., Sorensen, P.W., and Gallaher, D.D. 1995. The olfactory system of migratory adult sea lamprey (*Petromyzon marinus*) is specifically and acutely sensitive to unique bile acids released by conspecific larvae. The Journal of General Physiology. 105, 569 - 587.

Li., W., Siefkes, M.J., Scott, A.P., and Teeter, J.H. 2003. Sex pheromone communication in the sea lamprey: implications for integrated management. J. Great Lakes Res. 29 (Suppl. 1): 85 - 94.

Li, W., Twohey, M., Jones, M., and Wagner, M. 2007. Research to guide use of pheromones to control sea lamprey. J. Great Lakes Res. 33 (Special Issue 2): 70 - 86.

Lowe, W.H. & Bolger, D.T. (2002). Local and landscape-scale predictors of salamander abundance in New Hampshire headwater streams. Conserv. Biol., 16, 183–193.

Mahon, R., McConney, P., and Roy, R.N. 2008. Governing fisheries as complex adaptive systems. Marine Policy. 32, 104-112.

Malanson, G. P. 1999. Considering complexity. Annals of the Assocation of American Geographers. 89, 746 – 753.

Manion, P.J. 1972. Fecundity of the sea lamprey (*Petromyzon marinus*) in Lake Superior. Transactions of the American Fisheries Society 101: 718-720

Manion, P.J. and McLain, A.L. 1971. Biology of larval sea lamprey *Petromyzon marinus* of the 1960 year class, isolated in the Big Garlic River, Michigan. Great Lakes Fishery Commission, Technical Report 16.

Minshall, G.W., Cummins, K.W., Petersen, R.C., Cushing, C.E., Bruns, D.A., Sedell, J.R., and Vannote, R.L. 1985. Developments in stream ecosystem theory. Canadian Journal of Fisheries and Aquatic Sciences. 42, 1045-1055.

Morman, R.H., Cuddy, D.W., and Ruger, P.C. 1980. Factors influencing the distribution of sea lamprey (*Petromyzon marinus*) in the Great Lakes. Can. J. Fish. Aquat. Sci. 37: 1811 - 1826.

Newbold, J.D., Elwood, J.W., O'Neill, R.V., and Van Winkle, W. 1982. Nutrient spiraling in streams: implications for nutrient limitation and invertebrate activity. American Naturalist. 120: 628-652.

North, M.J., Collier, N.T., and Vos, J.R. 2006. Experiences creating three implementations of the Repast agent modeling toolkit. ACM Transactions on Modeling and Computer Simulation. 16, 1 - 25.

O'Neal, R.P. 1997. Muskegon River Watershed Assessment. Michigan Department of Natural Resources, Fisheries Division, Fisheries Special Report 19. 187 pp.

Osborne, L.L. and Wiley, M.J. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. Canadian Journal of Fisheries and Aquatic Sciences. 49, 671 - 681.

Perry, J.A. and Schaeffer, D.J. 1987. The longitudinal distribution of riverine benthos: A river dis-continuum? Hydrobiologia. 148, 257-268.

Poole, G.C. 2002. Fluvial landscape ecology: addressing uniqueness within the river discontinuum. Freshwater Biology. 47, 641 - 660.

Potter, I.C., Hillard, R.W., Bradley, J.S., and McKay, R.J. 1986. The influence of environmental variables on the density of larval lampreys in different seasons. Oecologia 70: 443-440.

Pringle C.M., Naiman R.J., Bretschko G., Karr J.R., Oswood M.W., Webster J.R., Welcomme R.L. & Winterbourn M.J. (1988) Patch dynamics in lotic systems: the stream as a mosaic. Journal of the North American Benthological Society, 7, 503–524.

Rice, S.P., Ferguson, R.I., and Hoey, T.B. 2006. Tributary control of physical heterogeneity and biological diversity at river confluences. Can. J. Fish. Aquat. Sci. 63: 2553-2566.

Schleen, L.P, and Klar, G.T. 2000. Integrated management of sea lampreys in the Great Lakes 1999. Annual Report to Great Lakes Fishery Commission.

Slade, J.S., Adams, J.V., Christie, G.C., Cuddy, D.W., Fodale, M.F., Heinrich, J.W., Quinlan, H.R., Weise, J.G., Weisser, J.W., and Young, R.J. 2003. Techniques and methods for estimating abundance of larval and metamorphosed sea lampreys in Great Lakes tributaries, 1995 to 2002. J. Great lakes Res. 29: 137-151.

Smith, B.R. and Tibbles, J.J. 1980. Sea lamprey (*Petromyzon marinus*) in Lakes Huron, Michigan, and Superior: history of invasion and control, 1936-78. Can. J. Fish. Aquat. Sci. 37: 1780-1801.

Sorensen, P.W., and Vrieze, L.A. 2003. The chemical ecology and potential application of the sea lamprey migratory pheromone. J. Great Lakes Research. 29 (Suppl. 1): 66 - 84.

Townsend C.R. (1989) The patch dynamics concept of stream community ecology. Journal of the North American Benthological Society, 8, 36–50.

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

Vrieze, L.A. and Sorensen, P.W. 2001. Laboratory assessment of the role of a larval pheromone and natural stream odor in spawning stream localization by migratory sea lamprey (*Petromyzon marinus*). Canadian Journal of Fisheries and Aquatic Sciences. 58, 2374 - 2385.

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

Table 3.1: Summary of model setup for computational experiments with randomly-generated river networks. For each of experiments 1-6, the table gives: the diameters of the river networks; values of α, the model parameter that weights the relative importance of larval pheromone and tributary size for migratory adults; the length of each model run, in years; the total number of model runs performed in each experiment; and the figure showing results from the experiment.

Experiment	River Network Diameters	α	Length (years)	Number of Model Runs	Resulting Figure
1	{6:25}	{0.01, 1, 100, 1000}	4	3000 per value of α	3a
2	{6,7,8,25}	1	4	30 per network diameter	3b
3	8 (fixed)	1	4	1500	3c
4	{6,7,8,25}	1	50	50 per network diameter	4a
5	{6, 25}	1	100	1 per network diameter	4b
6	8 (fixed)	{0.001, 0.1, 1, 10, 1000}	20	100 per value of α	4c

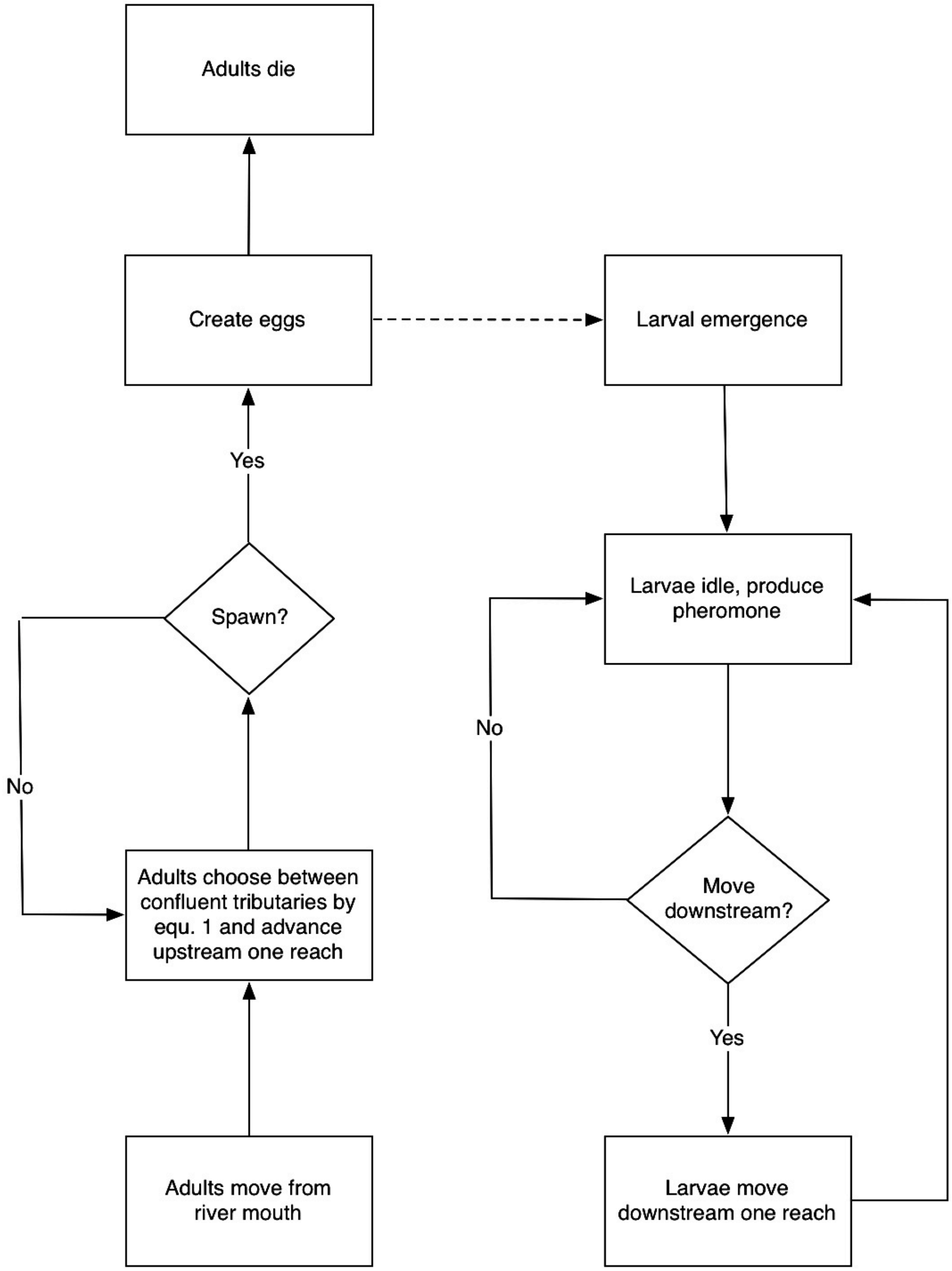


Figure 3.1: Diagram of model sea lamprey adults (a) and larvae (b) describing processes for upstream (adult) and downstream (larval) movement. Adult sea lamprey begin at the river mouth and migrate upstream until spawning. At each confluence, adults choose between confluence tributaries on the basis of tributaries' pheromone concentration and size (equ. 1). The probability of spawning in a reach is the inverse of the reach's drainage area. Larvae are inactive apart from downstream migration and the production of pheromone, and leave the system as parasitic juveniles after four years.

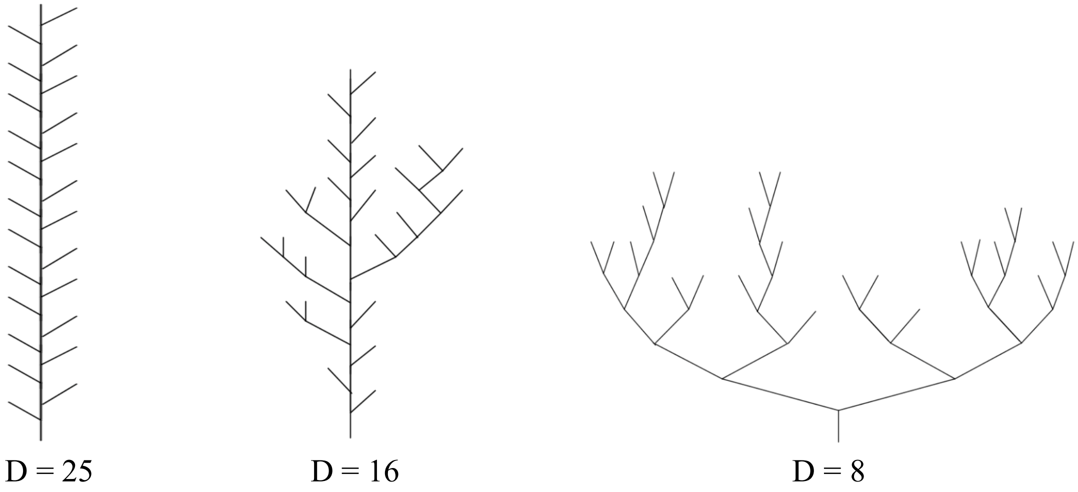


Figure 3.2: Examples of river networks generated by the model. Each network contains 49 reaches. High diameter river networks are elongated and less compact (left-most network; diameter = 25, bifurcation ratio = 25), while low diameter river networks are short, compact, and branchy (right-most network; diameter = 8, bifurcation ratio = 3.04). The center network has a diameter of 16 and a bifurcation ratio of 5. Note that there are many unique network configurations with the same diameter.

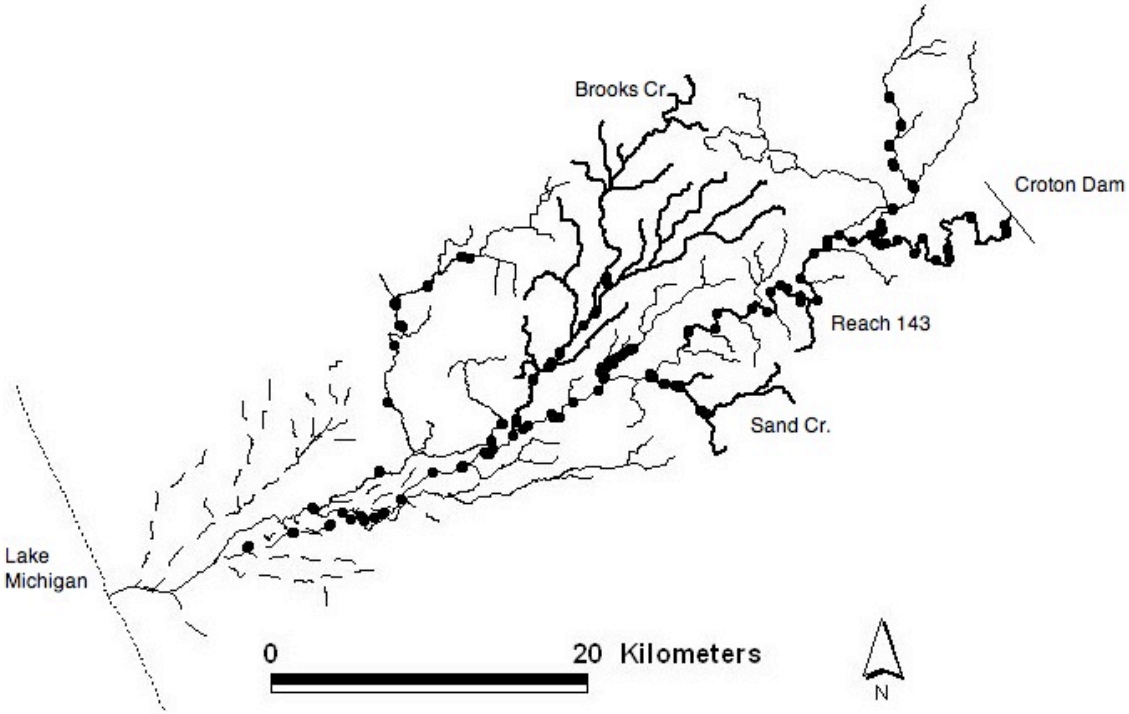


Figure 3.3: Map of the Muskegon River below Croton Dam, showing the outline of the Muskegon River, reaches included in my model (all heavy and solid lines), subwatersheds referenced in this paper (heavy lines, and labeled), uninhabitable Areas of Concern and braided channels removed from the model (dashed lines), and locations of USFWS sea lamprey larval samples (circles). Area upstream of Croton Dam is approximately 6,022 km^2.

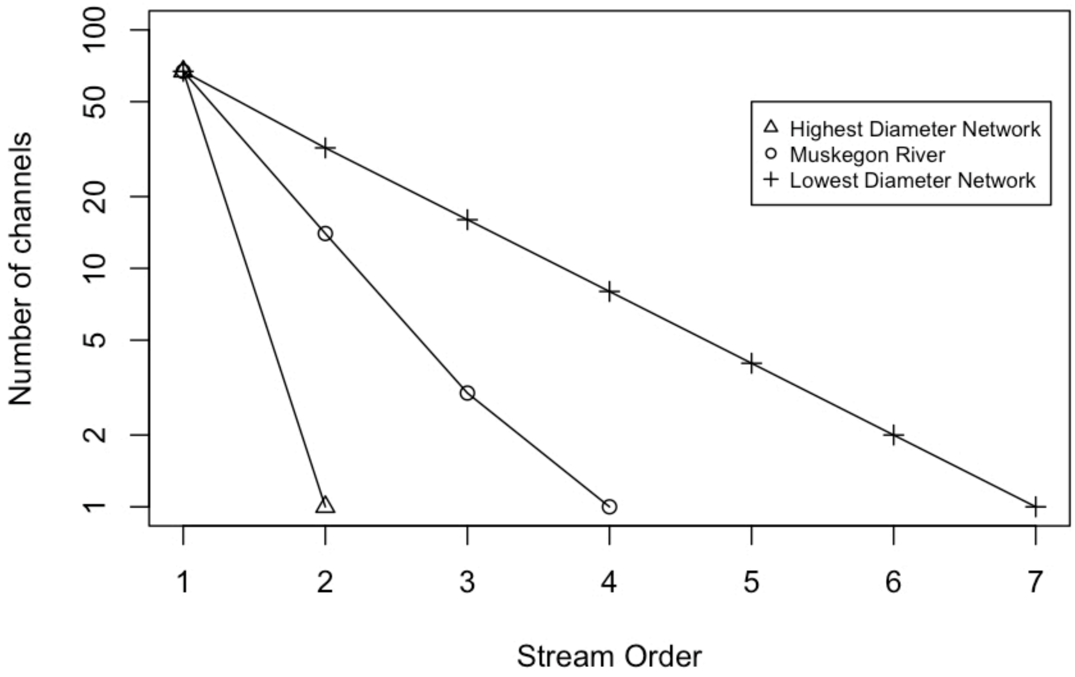
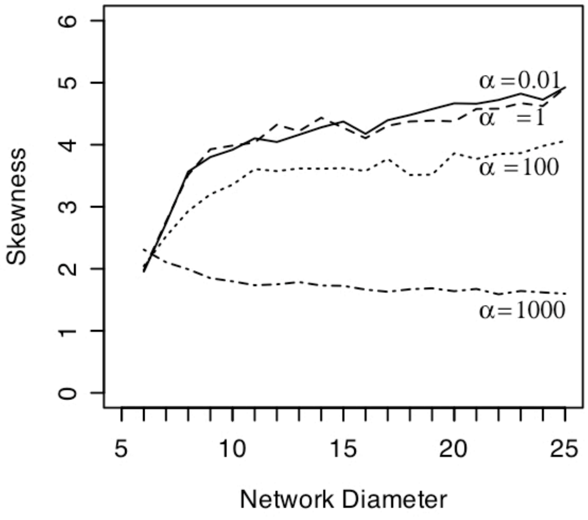
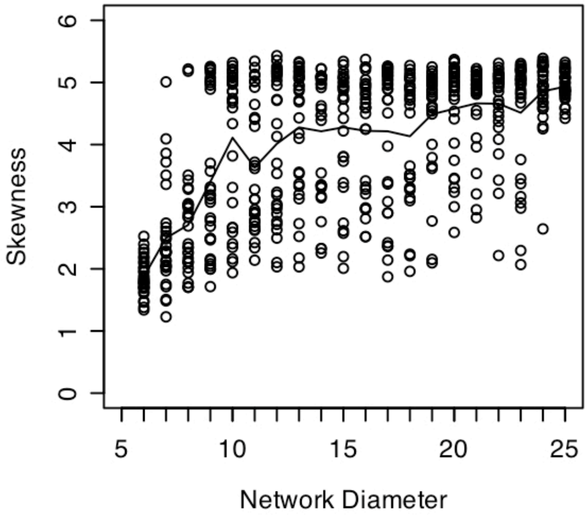


Figure 3.4: Plot of the number of channels of a given stream order, defined using the Horton-Strahler classification, for the Muskegon River below Croton Dam (open circles; see Fig. 3.3) and two hypothetical river networks with the same number of reaches as the Muskegon River (n = 167). The two hypothetical river networks represent, for channel networks with 167 reaches, the network configurations with the maximum possible network diameter (triangles) and the minimum possible network diameter (crosses).

a)


Skewness
0
1
2
3
4
5
6
5
10
15
20
25
Network Diameter
α =0.01
α =1
α =100
α=1000

b)


Skewness
0
1
2
3
4
5
6
5
10
15
20
25
Network Diameter

c)

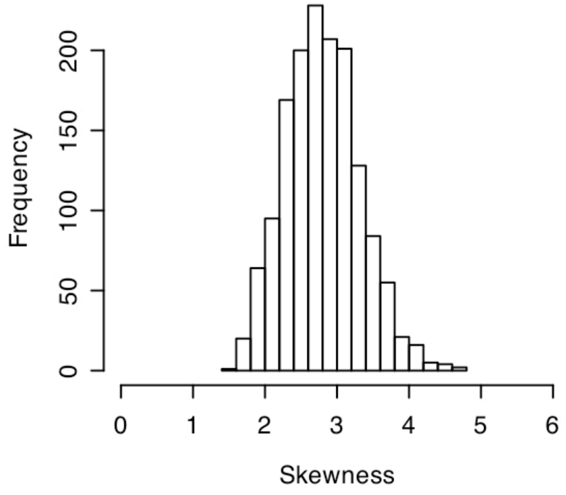
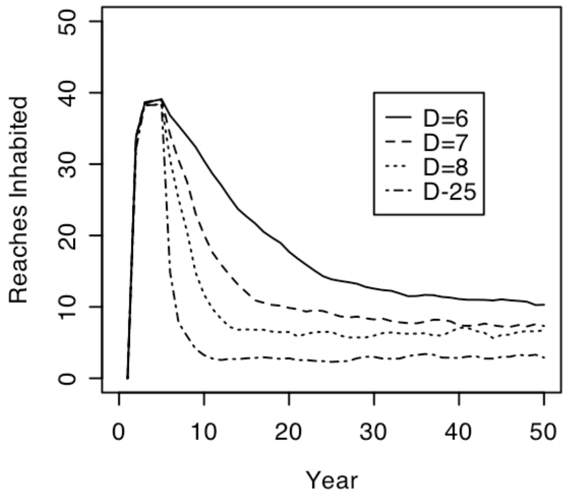

Frequency
0
50
100
150
200
0
1
2
3
4
5
6
Skewness

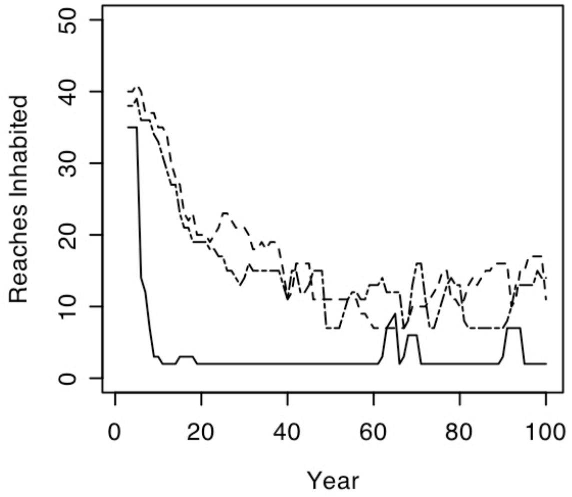
Figure 3.5: Results of the computational experiments using randomly-generated river networks and four-year model runs. Panel A gives the results of experiment 1, the relationship between network diameter and the skewness of the distribution of larvae across reaches in the network for four selected values of α, the parameter that weights the importance of tributary size and pheromone concentration for migratory adults. Results are from 3000 model runs at each α. Lines are drawn through the mean skewness at each diameter. Panel B gives the results of experiment 2, 30 model runs at each network diameter (points) for $\alpha = 1$. The line is drawn through the mean skewness at each diameter. Panel C gives the results of experiment 3, a histogram of measured skewness over 1500 model runs for a fixed river network of diameter 8 (see Fig. 3.2**)**.

a)



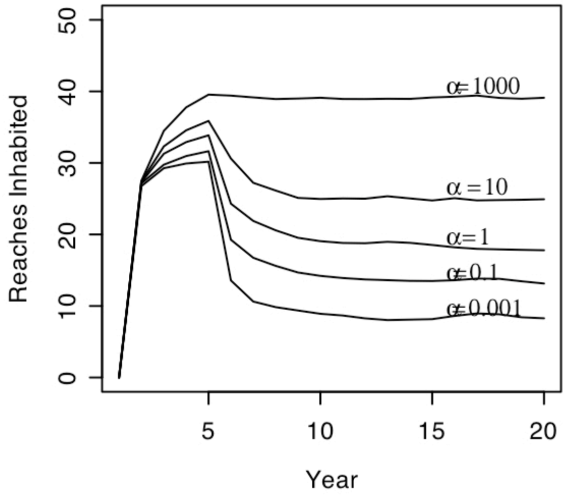
Reaches Inhabited
0
10
20
30
40
50
Year
0
10
20
30
40
50
D=6
D=7
D=8
D-25

b)



Reaches Inhabited
0
10
20
30
40
50
Year
0
20
40
60
80
100

c)



Reaches Inhabited
0
10
20
30
40
50
Year
5
10
15
20
α=1000
α =10
α=1
α=0.1
α=0.001

Figure 3.6: Results of the computational experiments using randomly-generated river networks and long-term model runs. Panel A gives the results of experiment 4, the change in the number of reaches inhabited over the course of 50 years for four network diameters, a measure of network shape. Each line shows the mean number of reaches inhabited over 50 model runs. Panel B gives the results of experiment 5, the change in the number of reaches inhabited over the course of 100 years for three individual model runs of diameter 6 (both dashed lines) and 25 (solid line). Panel C gives the results of experiment 6, the change in the number of reaches inhabited over the course of 20 years for the preset river of diameter 8 (Fig. 3.2c) over 100 model runs at each of five values of α, the parameter that weights the importance of tributary size and pheromone concentration for migratory adults.

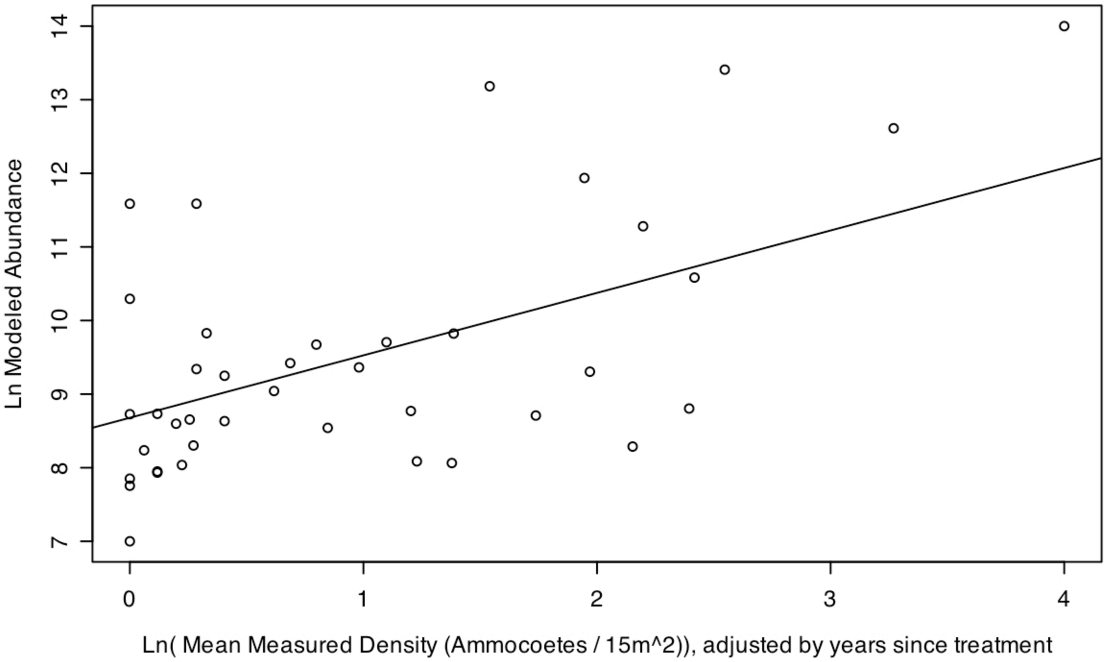


Figure 3.7: Relationship between the natural log of mean larval density in Type 1 habitat in the USFWS database adjusted by years since treatment (horizontal axis) and the natural log of the mean modeled larval abundance over 500 model runs (vertical axis), for reaches in the Muskegon River, Michigan. The relationship is significant at $p < 0.005$ (Spearman's rho = 0.46). The line is a best-fit linear regression.

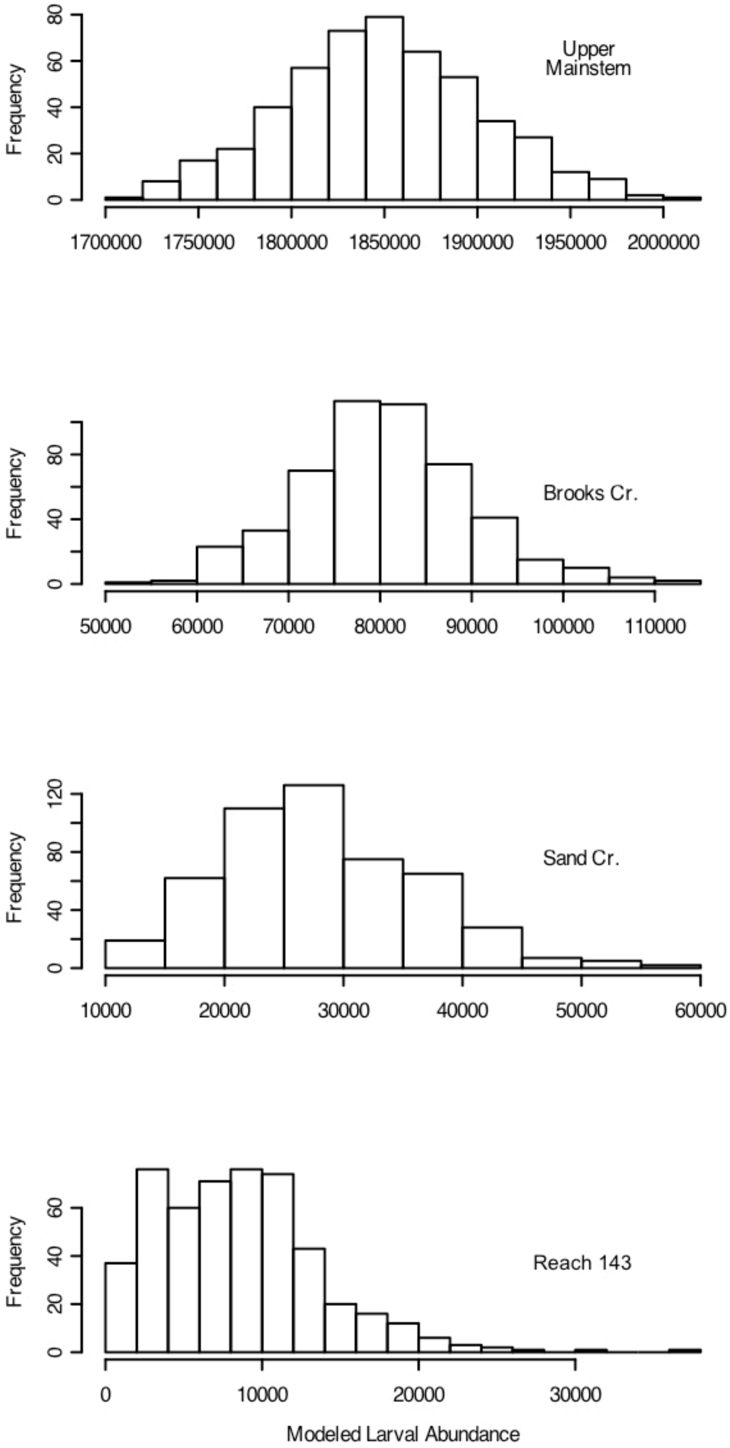


Figure 3.8: Frequency distributions of the total abundance of larval sea lamprey in four subwatersheds of the Muskegon River (see Fig. 3.3) across 500 model runs. Note differences in horizontal axes between plots.

Chapter 4

The watershed as place: Sea lamprey as a case study in defining boundaries for population analyses

Abstract

Restricting the scope of analysis is often prerequisite to work in conservation biology, yet no clear rules exist for defining the boundaries of populations or ecosystems. One option is to identify the spatial and temporal domains within which members of a population strongly interact. A preliminary step is to understand how the spatial structure of the environment might influence the population dynamics of resident species. In this study, I expanded an existing individual-based computational model and performed a series of computational experiments to explore how both river network structure and the distribution of larval habitat might influence the distribution of larvae of sea lamprey *(Petromyzon marinus)*. I observed four main results: 1) Larval distribution depended on both river network structure and the distribution of larval habitat, but the relative importance of these two factors depended on the basin-wide abundance of larval habitat. Larvae were most aggregated in elongated river networks with abundant larval habitat. 2) Larvae in the model aggregated over time, and the degree and rate of this aggregation depended on river network structure and the amount and distribution of larval habitat. Larvae aggregated most quickly, and eventually into the smallest number of reaches, in elongated river networks with abundant larval habitat. 3) The correlation between larval abundance and amount of larval habitat in a reach depended on the overall abundance of larval habitat; and 4) Total larval population size in a watershed depended on factors that control larval aggregation, i.e. river network structure and the amount and distribution of larval habitat. Based on these modeling results, I argue that lamprey within a watershed

constitute an interlinked population whose dynamics cannot be understood at a spatial scale finer than that of the watershed, and that the watershed therefore constitutes an important functional boundary for population analyses of sea lamprey. I discuss implications of these modeling results for sea lamprey management in the Great Lakes.

Introduction

Restricting the scope of analysis is often prerequisite to work in conservation biology (Post et al. 2006), yet no clear rules exist for defining the boundaries of populations (Berryman 2002, Camus and Lima 2002), or ecosystems (Keller and Goley 2000). Spatial and temporal boundaries for population analyses should ideally coincide with units over which population dynamics are "closed" (*sensu* Wiens 1989) or at least "well-bounded" (*sensu* Post et al. 2006). Identifying the spatial and temporal domains within which members of a population strongly interact is therefore an important first step in developing appropriate boundaries for conservation and management.

A central tenet of landscape ecology is that the spatial structure of the environment influences the population dynamics of resident species (Dunning et al. 1992, Wiens et al. 1993, Fahrig and Merriam 1994, Pickett and Cadenasso 1995, Wiens 1999, Fausch et al. 2002). Many environments exhibit heterogeneity at multiple scales, including rivers (Leopold et al. 1964), forests (Picket and Cadenasso 1995), and intertidal zones (Kotliar and Wiens 1990). Populations in any of these environments are influenced by processes operating at multiple scales of spatial structure (e.g. Fausch et al. 2002, Baker and Wiley 2009), particularly when resource requirements and animal behavior differ between life history stages (Dunning et al. 1992). Yet to merely say that processes at multiple spatial scales influence organisms is to "mouth a platitude" (Kareiva 1990) and makes no contribution towards defining operational boundaries for population analyses. The challenge for ecologists is to establish quantitative links between specific aspects of organisms' behavior and specific aspects of spatial structure in the environment (Schlosser 1991, 1995). Such links are essential for understanding how population dynamics at any one location are constrained by the structure of the environment.

In this paper, I consider how geomorphic structure at two spatial scales (local substrate composition and basin-wide network shape) could influence the distribution of sea lamprey (*Petromyzon marinus*), an anadromous fish whose adults are sensitive to a pheromone released by stream-resident benthic larvae. This paper is a case study for organisms with similar life histories: many insects (Carde and Minks 1998) and fish (Sorensen et al. 1998) communicate via pheromone, and their population dynamics may generally be determined by the distribution of life-stage specific resources, spatial structures that alter the diffusion of pheromone, and pheromone-driven feedbacks between life history stages. Species that communicate across life history stages via pheromone are a special case of a broader class of organisms, so that this paper is also a case study of how the environment constrains population dynamics when resource requirements differ between life history stages.

More specifically, the sea lamprey is a primitive fish with an anadromous, semelparous lifecycle characterized by a stream-resident, filter-feeding and burrowing larval stage generally three to five years in duration, an ocean- or lake-dwelling juvenile parasitic stage, and a non-feeding sexually mature adult stage during which lamprey migrate from oceans or lakes into rivers to spawn (Beamish 1980). The sea lamprey is an invasive species in the Laurentian Great Lakes with a long history of negative impacts on salmonines (Smith and Tibbles 1980), particularly lake trout (*Salvelinus namaycush,* Coble et al. 1990). Lamprey populations in the Great Lakes are controlled primarily with a chemical lampricide that targets larvae and is applied to the most heavily populated streams every four years. High management costs (> $14 million/year; Slade et al. 2003, Christie et al 2003, Jones 2007) provide a strong practical motivation for developing a deeper understanding of sea lamprey ecology.

On a local scale, the distribution of sea lamprey larvae is correlated with the distribution of their habitat (Applegate 1950, Malmqvist 1980, Potter et al. 1986). Larvae primarily occur in "depositional zones", characterized by a soft, burrowable substrate of fine sands, low current velocity, and some amount of organic detritus. Larval habitat can be further subdivided into preferred habitat and suitable habitat based on grain size and abundance of fine organic detritus (i.e. Types I and II, Slade et al. 2003). Larvae are unable to burrow into hard substrates (Slade et al. 2003, Quintella et al. 2007), with the implication that larvae cannot colonize these areas. Each stream in its entirety thus consists of a mosaic of patches of ideal, acceptable, and unacceptable larval habitat, as well as patches of adult spawning habitat.

The distribution of sea lamprey larvae is also necessarily related to the branching pattern of reaches in each watershed, i.e. river network structure (Chpt. 3, this dissertation). Adult lamprey migrate upstream through a series of confluences en route to spawning sites, and the relationship between river network structure and the directional choices adults make at confluences has important consequences for where they ultimately spawn. Larval sea lamprey release a pheromone to which adults home during their upstream spawning migration (Li et al. 1995, Bjerselius et al. 2000, Vrieze and Sorensen 2001, Sorensen and Vrieze 2003). The size (i.e. flow) of a tributary dilutes the concentration of this larval pheromone. The downstream propagation pattern of larval pheromone through a watershed thus depends on patterns in the relative size of confluent tributaries. River network structure by itself must therefore influence the migration decisions of spawning-phase lamprey, and in turn their ultimate spawning location and the subsequent distribution of young of the year larvae. Furthermore, larvae created by each year class of adults alter the pheromone distribution for subsequent generations. The distribution of larvae in a watershed is thus a result of a complex system of emergent feedbacks contingent on both past fish behavior and network structure.

In this paper I perform a series of computational experiments to explore how the distribution of larval habitat could mediate the watershed-scale population dynamics that are contingent on network structure and adult migration. The overarching goal was to understand the temporal and spatial scales at which lamprey population dynamics might be considered “well-bounded”. To meet this goal, I sought to identify how and when network structure, habitat, and fish behavior influence larval distribution.

Methods

I began with an existing computational model developed in a previous study to test the hypothesis that river network structure influences larval sea lamprey distribution. This model was previously validated against historical survey data from the Muskegon River, Michigan; Chapter 3 of this dissertation contains that validation and a detailed analysis of the relationships between network structure, larval distribution, and model parameterization, over both short and long timescales. Below, I expand this existing model by adding larval habitat. I give a brief overview of the expanded new model, discuss my formulation of habitat and how it interacts with river network structure, and explain my experimental framework and analyses.

Computational Model

The response variable in this study is the distribution of sea lamprey larvae across a watershed at the resolution of a reach, defined here as a section of river between two confluences. I hypothesized that larval distribution at the resolution of the reach emerges from the interaction of river network structure, the distribution of larval habitat, the production of larval pheromone, and the behavioral choices of adults migrating upstream to spawn. The model therefore consists of an abstract representation of a river network and its habitats linked to discrete time biological submodels that describe the upstream migration of adults, and the downstream movement of larvae and of larval pheromone. An overview of the model follows.

The model runs at a daily timestep. The basic spatial unit in the model is the reach, and the river network consists of a hierarchical arrangement of n such reaches. Each reach is characterized by 1) an index of size (upstream drainage area, represented by link number), 2) pheromone concentration, 3) number of sea lamprey larvae of each year class, 4) number of adult sea lamprey, 5) the identity of the one downstream and two (or zero) upstream reaches, and 6) the amount of larval habitat.

Adult sea lamprey

In my model, female adult sea lamprey begin the spawning migration each year at the mouth of the river and proceed upstream until they reach a 1st-order tributary and spawn. Adult females are modeled as individuals, and are characterized by their location (i.e. the reach they occupy at timestep t) and a Boolean variable describing whether they have spawned. For the sake of simplicity I allow all females to find a nest and spawn. The decision rule for where females spawn is a simplified version of the rule used in

Chapter 3, where the probability of spawning in a reach is a function of the inverse of the reach link order.

When migrating upstream and encountering a confluence, adult sea lamprey must decide which of the two tributaries to follow. The decision rule for choosing between two confluent tributaries is a rule where the probability of selecting a given path is a linear function of both pheromone concentration and tributary size, written as:

$$P(a) = \frac{[Ph]_a + \alpha S_a}{[Ph]_a + [Ph]_b + \alpha S_a + \alpha S_b} \qquad (1)$$

where P(a) is the probability of choosing tributary a, $[Ph]_a$ is the pheromone concentration in tributary a, S_a is the size of tributary a (upstream drainage area in km^2), and α is a model parameter for weighting the importance of tributary size relative to pheromone concentration. Chapter 3 contains the rationale and derivation of this equation.

The equation (1) itself can be parameterized to represent a range of possible choice rules. When $\alpha = 0$, adult lamprey base their migration decisions exclusively on pheromone concentration. When α is very large, adult lamprey ignore pheromone and base their migration decisions nearly exclusively on the relative sizes of confluent tributaries. Intermediate values of α represent the range of conditions in which adults consider both pheromone and tributary size in their upstream migration. In this chapter, I fixed $\alpha = 1$ for all experiments. Chapter 3 contains an analysis of model behavior for a range of values of α between 0.01 and 1000.

Larval Sea Lamprey

Larvae are modeled as identical members of a year class, and for each reach I model the number of larvae of each age zero to four. Larvae drift downstream at a constant rate, where 30% of the larvae move downstream one reach each year. Larvae in the model are otherwise inactive (i.e. any movement occurs below the spatial resolution of the model), and serve mainly as a pheromone source to which adults are attracted upstream. At age four, larvae mature and leave the system.

The maximum potential larval abundance in each reach is modeled as a linear function of the amount of larval habitat in that reach. Larvae migrating from upstream reaches colonize a reach successfully if the resident larval population is less than this maximum. The fate of larvae unable to colonize a reach depends on their age (i.e. size): young of the year recently created by spawning adults perish if they are unable to colonize their immediate reach. Larvae one year or older drift downstream one reach at a time and re-attempt the colonization process, and continue this process until they either successfully colonize a reach or are flushed out of the river system.

The distribution of larval pheromone is recalculated after spawning during each day in which spawning occurs. Each larva in the model creates one 'unit' of pheromone, and each reach has unit drainage area and contributes one 'unit' of discharge to all downstream reaches. The pheromone concentration in each 1st-order reach (i.e. a reach with no further upstream reaches) is calculated as:

$$[Ph] = \frac{n * r}{Q}$$

where n is the number of resident larvae, r is the rate of pheromone production (in units of volume of pheromone per time) per larvae, and Q is the discharge (i.e. flow) of the river in that reach (in units of volume of water per time). I assume that the pheromone load propagates downstream with no losses, such that the pheromone concentration for each 2nd-or greater order reach is calculated as the total number of resident and upstream larvae divided by the discharge (i.e. the number of upstream reaches).

Model Networks and Habitat

Each river network consists of a hierarchical arrangement of 49 unit reaches. I measure river network shape as the diameter of the network, defined as the maximum link distance between the mouth of the river and the most distal tributary. Networks generated by my model differ in their branchiness, or the degree to which a network is short and compact (low diameter) rather than elongated (high diameter; Fig. 4.1).

I sought to distribute larval habitat across the model networks in ways that would be consistent with the way natural hydrologic processes might generate patterns of larval habitat in nature. Key dimensions of larval habitat (e.g. substrate particle size and fine organic detritus) are correlates of flow. However, the exact distribution of habitat in any river is a complicated function of local stream energy and sediment budget (Knighton 1998). Rather than duplicate this variability, I parameterized the model to recreate three general trends in the distribution of larval habitat found in a database of larval sea lamprey habitat surveys in the Lower Peninsula of Michigan (Chapter 2, this dissertation). In some streams, habitat was concentrated upstream (as in Good Harbor Creek, watershed 18 in Chpt. 2). To recreate this distribution, I modeled habitat as a function of the inverse of link number, such that small (link 1) tributaries have the largest amounts of larval habitat and the largest mainstem reaches have the smallest amounts of habitat. As a second distribution, I modeled the same amount of habitat in all reaches regardless of their size (as in the Big Manistee River, watershed 15 in Chapter 2). As a third distribution, I considered streams where habitat was concentrated downstream (as in the Betsie River, watershed 16 in Chapter 2). To recreate this distribution, I modeled habitat as a positive function of link number, with the result that smallest tributaries have the smallest amounts of larval habitat.

The branching pattern of reaches in each watershed (i.e. river network structure) determines the spatial pattern of flow across each watershed (represented by link numbers in Fig. 4.2a). In my model, the amount of larval habitat in each reach is a function of link number; as a result, the spatial patterns in link number resulting from network structure are mirrored in the distribution of larval habitat (Fig. 4.2b). An important consequence is that the spatial proximity of larval habitat to locations of adult spawning (i.e. habitat complementation, Dunning et al. 1992) is affected by both network structure (Fig. 4.2b) and the longitudinal distribution of larval habitat (Fig. 4.2c). These patterns in the distribution of larval habitat will have important consequences for larvae drifting downstream from locations of adult spawning.

To explore the influence of habitat abundance, I included a model parameter that allowed scaling of the overall amount of larval habitat in a watershed by some linear constant independently of habitat distribution parameters. The amount of habitat in each reach is calculated as a function of link number, and then multiplied by this scaling parameter (Fig. 4.2d). Throughout this paper, I reference four habitat abundance parameterizations: habitat sparse (1x), minimally abundant (10x), moderately abundant (100x), and very abundant (10000x).

Experimental Framework

I conducted a series of computational experiments to answer three research questions: 1) How do the availability of larval habitat alter the relationships between network structure and larval distribution established in Chapter 3? 2) How does the correlation between the number of larvae and the amount of habitat in a reach change with network structure and habitat availability? and 3) How does the total larval population size change with network structure and the amount and distribution of habitat? Table 4.1 contains summary details for all experiments.

Question 1: How does the availability of larval habitat influence previously-established relationships between river network structure and larval distribution?

I previously found that in short-duration model runs (i.e. after four years) river network diameter was positively correlated with the aggregation of larvae within a river network (i.e. high diameter networks often contain a few reaches with high larval abundance; exp. 1 and Fig. 3.4a in Chapter 3). In this study, I hypothesized that the availability of larval habitat might at times constrain the distribution of larvae and thus alter the previously-observed relationship between network structure and larval distribution. To test this hypothesis, I performed an experiment with a series of three treatments using networks of varying diameter, where habitat was extremely abundant (10000x, trt. 1a), moderately abundant (100x, trt. 1b), and sparse (1x, trt. 1c; Table 4.1). Within each treatment, I also ran the model with habitat concentrated upstream, constant in all reaches, and concentrated downstream. For each combination of habitat amount, habitat distribution, and network diameter, I calculated the mean skewness of the distribution of larvae across all reaches in the network. Model runs with high positive skewness had one or several reaches with high larval densities and many other reaches with low or absent densities, while model runs with lower (but positive) skewness had smaller differences in larval density between reaches.

Over the long term (i.e. over 20 or more years), I previously found that larvae in a river network became aggregated over time, and that the degree and rate of this aggregation depends on river network structure (exp. 4 and Fig. 3.5a in Chapter 3). Larvae in high-diameter networks converged more quickly, and eventually populated a smaller number of reaches, than did larvae in low-diameter networks. Similarly to the four-year case, I hypothesized that the availability of larval habitat might at times constrain the distribution of larvae and thus alter the previously-observed relationship between network structure and larval aggregation. To test this hypothesis, I performed a second experiment with three treatments, where larval habitat was extremely abundant (trt. 2a), moderately abundant (trt. 2b), and sparse (trt. 2c). Habitat in this experiment was always weighted upstream. For each treatment, I ran the model with network diameters of D = 6, 7, 8, and 25, and calculated the mean number of reaches inhabited each year over 50 model runs. These network diameters (6,7,8, and 25) were network shapes that in Chapter 3 generated least larval aggregation (D = 6), intermediate larval aggregation (D = 7, 8) and the greatest amount of larval aggregation (D = 25).

Question 2: How do network structure and the availability of larval habitat influence the correlation between larval abundance and amount of habitat in a reach?

If larval habitat is so abundant that it does not constrain larval abundance in any reach, then model patterns of larval distribution across each river network should depend only on river network structure and patterns of adult migration. In this case, differences in the amount of habitat among reaches should be unrelated to differences in larval abundance among those reaches. Consequently, the correlation between the number of larvae and the amount of habitat in each reach should be low. Conversely, in a river network where larval habitat is rare, habitat might function as a bottleneck and saturate with larvae in all reaches regardless of network-dependent patterns of adult migration. In this case, differences in the amount of habitat among reaches should be exactly mirrored by differences in larval abundance among those reaches, and the correlation between the amount of habitat and the number of larvae in a reach might be high. With respect to question 2, then, I hypothesized that the correlation between the number of larvae and the amount of habitat in a reach would be related to the overall availablity of larval habitat.

I tested this hypothesis by performing an experiment with two treatments related to the correlation between larval abundance and the amount of habitat in a reach (summarized in Table 4.1). As a first treatment (3a), I explored how both the amount and type of distribution of larval habitat in the river network influenced the correlation between the number of larvae and the amount of larval habitat in a reach. I used a fixed river network of diameter 8 (Fig. 4.2b), each of habitat weighted upstream and downstream, and varied the amount of habitat from sparse (1x) to very abundant (10000x). For each combination of amount and distribution of habitat, I measured the correlation between larval abundance and reach habitat for all reaches in the network.

In a second treatment (3b), I examined how both river network structure and the overall availability of larval habitat influenced the correlation between the number of larvae and the amount of larval habitat in a reach. For each of network diameters 6 and 25, I varied the amount of larval habitat from sparse (1x) to very abundant (10000x). For each combination of network diameter and habitat, I measured the correlation between larval abundance and reach habitat for all reaches in the network.

Question 3: How do network structure and the availability of larval habitat influence the size of the larval population in a river?

Density-dependent effects in my model are implemented as a simple threshold, where larvae successfully colonize any reach where larval abundance is below the maximum that habitat can sustain, but never colonize a reach saturated with larvae. Larval mortality occurs when larvae are unable to colonize any reaches and drift downstream past the mouth of the river. I therefore expected a positive relationship between the amount of larval habitat in a network and total larval population size. However, when larvae are not uniformly distributed across a network, some reaches saturate with larvae more quickly, implying that some mortality might be related to the degree of aggregation of larvae within a river network. I hypothesized that total larval population size should therefore be sensitive to factors that control larval aggregation, e.g. river network structure and the distribution of habitat.

In this experiment, I used a series of three treatments, where larval habitat in the river network was weighted upstream (trt. 4a), equal in all reaches (trt. 4b), and weighted downstream (trt. 4c). Each treatment consisted of running the model with each of network diameters D = 6, 7, 8, and 25, and with increasing amounts of habitat from sparse (1x) to very abundant (10000x). For each combination of habitat distribution and network diameter, I examined how the total larval population size changed with the total amount of larval habitat in the watershed. These network diameters (6,7,8, and 25) were network shapes that in Chapter 3 generated lowest skewness (D = 6), intermediate skewness (D = 7, 8), and highest skewness (D = 25) of the distribution of larvae across all reaches in the network.

Results

Question 1: How does the availability of larval habitat influence previously-established relationships between river network structure and larval distribution?

When habitat was very abundant (trt. 1a), I observed a strong positive relationship between river network diameter and the skewness of the distribution of larval abundance across all 49 reaches (Fig. 4.3a). This relationship did not depend on the longitudinal distribution of larval habitat across the river network: when larval habitat was so abundant in the watershed that it never limited larval population size in any reach, the distribution of habitat had no influence on the distribution of larvae. This result is essentially the same as experiment 1 in Chapter 3.

When larval habitat was moderately abundant (trt. 1b), the relationship between network diameter and larval skewness depended on the longitudinal distribution of larval habitat (Fig. 4.3b). When larval habitat was concentrated upstream and complemented locations of adult spawning, I observed a strong positive relationship between river network diameter and the skewness of the distribution of larval abundance. When larval habitat was equal in all reaches, river network diameter had less influence on the skewness of the distribution of larvae. When larval habitat was concentrated downstream, river network diameter had no influence on the skewness of the distribution of larvae.

When larval habitat was sparse (trt. 1c), I observed little relationship between network diameter and larval distribution (Fig. 4.3c). Regardless of network diameter, most reaches in the network were saturated with larvae. I also saw little effect of the distribution of habitat on the skewness of the distribution of larval abundance.

Larvae in the model aggregated over time, and the degree and rate of this aggregation depended on both river network structure and the amount of larval habitat (exp. 2; Fig. 4.4). Larvae aggregated more quickly, and eventually into a smaller number of reaches, with increasing network diameter and with increasing amounts of larval habitat. When larval habitat was most abundant (trt. 2a, Fig. 4.4a), results corresponded to experiment 4 in Chpt. 3. Larvae did not aggregate only in model runs with both the lowest diameter river network (D = 6) and smallest amounts of larval habitat (trt 2c; Fig. 4.4c).

Question 2: How do network structure and the availability of larval habitat influence the correlation between larval abundance and amount of habitat in a reach?

The correlation between the number of larvae and the amount of habitat in a reach depended in a complex way on both river network structure and the amount and distribution of larval habitat across the river network. When habitat was concentrated upstream in the river, the correlation between larvae and habitat in a reach decreased with increasing amounts of habitat. When habitat was concentrated downstream, the correlation between larvae and habitat was highest at intermediate amounts of habitat (trt. 3a; Fig 4.5a).

When larval habitat was rare, river network structure had a large influence on the correlation between the number of larvae and the amount of habitat in a reach (trt. 3b; Fig. 4.6b). But, the influence of river network structure decreased as habitat became more abundant. The highest correlation between larvae and habitat occurred in low-diameter networks when habitat was sparse; the lowest correlation occurred in high-diameter networks when habitat was very abundant.

Question 3: How do network structure and the availability of larval habitat influence the size of the larval population in a river?

Total larval population size depended on both river network structure and the amount and distribution of habitat. When larval habitat was sparse, larval population size was always low (exp. 3; Fig. 4.6). Larval population size always increased with increasing amounts of larval habitat, but this relationship was dependent upon river network structure and habitat distribution. At intermediate abundances of habitat, larval population size was highest in low diameter rivers with habitat concentrated upstream. Larval population size was conversely lower in high-diameter rivers, or when habitat was weighted downstream. Network structure had little influence on total larval population size when habitat was very abundant, except when habitat was downstream and distant from locations of adult spawning.

Discussion

Spatial structures that influence larval distribution

Previous work related larval distribution to the distribution of larval habitat (Applegate 1950, Malmqvist 1980, Potter et al. 1986) and to network-dependent patterns of adult migration (Chpt. 3, this dissertation). In this study, I demonstrated how and when larval distribution could be influenced by both the distribution of larval habitat, which varies across reaches, and river network structure, which varies among watersheds.

The relative importance of habitat distributions and network structure depends on the overall abundance of larval habitat (Fig. 4.7). When habitat is very abundant and never constrains larval abundance, patterns of larval distribution across a watershed do not depend on habitat distribution and are instead a result of a system of emergent feedbacks contingent on each unique network structure and the past migratory and spawning behavior of adults (Fig. 4.7a). When habitat is only moderately abundant, habitat constrains larval abundance in some reaches, and the network-scale feedbacks between larval distribution and adult migration are mediated by the distribution of habitat (Fig. 4.7b). When habitat is extremely rare, all reaches saturate with larvae regardless of patterns of adult migration. Larval distribution no longer depends on network structure and feedbacks between larval distribution and adult migration, and instead depends only on the distribution of larval habitat (Fig. 4.7c).

I consider the two most extreme scenarios (habitat extremely abundant and habitat extremely rare) unlikely in nature. Habitat likely constrains larval abundance in some reaches, and so patterns of larval distribution within each river depend on river network structure, the amount and distribution of habitat, and the migratory behavior of past generations of adults.

Spatial and temporal boundaries

The structure of the entire river network and the distribution of all larvae in that network determine the downstream propagation pattern of larval pheromone, i.e. the pheromone landscape. Adults migrating upstream to spawn base their migration decisions on this pheromone landscape, such that the behavior of each adult lamprey is influenced by all larvae in the watershed. Larvae created by each adult then alter the pheromone landscape for subsequent generations. Lamprey within a watershed therefore constitute an interlinked population whose dynamics cannot be understood at a spatial scale finer than that of the watershed.

My modeling results suggest that the trajectory of sea lamprey populations may be particularly susceptible to stochastic, historic events. Patterns of larval distribution across a watershed result from a system of path-dependent feedbacks between larvae and adults. The probabilistic migration behavior of each adult potentially alters these feedbacks and in turn the population trajectory. Feedbacks between larvae and adults are also sensitive to disturbances that reduce larval abundance. For example, a localized physical disturbance (e.g. scouring event) that results in high larval mortality directly decreases the pheromone concentration in all downstream reaches and simultaneously increases the relative attractiveness of all other reaches in the river network to migrating adults. Localized events such as disturbances therefore have network-wide effects that alter in turn the phermone landscape, adult migratory behavior, and the distribution of future cohorts of larvae.

Watershed as place

I reiterate three points: 1) Patterns in the distribution of larvae over time depend fundamentally on a watershed's unique geomorphic context, i.e. river network structure and the distribution of larval habitat; 2) The watershed is an important functional unit within which lamprey form a strongly-interacting population; and 3) Localized, historic events have strong system-wide effects that impact the feedbacks between larvae and adults. Based on these three points, I argue for a view of each watershed as a unique place, where larval distribution depends on river network structure, the distribution of habitat, and stochastic, historic events that have altered feedbacks between larvae and adults.

My modeling results suggest that it may be impossible to predict larval densities in a stream reach solely on the basis of a reach's intrinsic properties (i.e. habitat) or on the basis of a spatially averaged stock-recruitment relationship. Rather than arguing for holism over reductionism, however, I reiterate Malanson's (1999) suggestion to consider places as emergent phenomena. This is a subtle yet important shift in viewpoint. I advocate considering the watershed as the functional unit for describing lamprey populations not simply because the watershed is larger than the reach, but because the watershed is the spatial unit within which lamprey constitute a strongly-interacting population whose dynamics reflect both historic events and the spatial structure of the environment.

The watershed has long been recognized as an important functional unit in geomorphology (Chorley 1969, Montgomery 1999, Post et al. 2006). As a result of its functional role, the watershed provides a convenient structure-based boundary (Strayer et al. 2003) for research and management in both geomorphology and ecology (Montgomery 1995). However, I argue for the watershed as a functional unit for studying and managing lamprey not simply for convenience, but because it constitutes a functional boundary (*sensu* Post et al. 2006) for lamprey populations, determining the spatial extent over which a population of lamprey strongly interact.

I identified the watershed as the important functional unit for studying and describing lamprey by first hypothesizing how specific fish life history processes might interact with specific spatial patterns in the environment. I then worked to understand the scale at which lamprey population dynamics emerged from the interaction of fish life history processes and the spatial structure of the environment. This process is an "operational and non-arbitrary" (Wiens 1989) method for defining scale based on an organism-centered view of the landscape (Wiens et al. 1993, With et al. 1997, Wiegan et al. 1999).

Contributions to sea lamprey management

Larval assessment My model helps to explain how the migratory behavior of adults might influence the distribution of larvae across reaches within a river system. Previous authors suggested that the absence of larvae in some streams might diminish their relative attractiveness to lake-dwelling adults just beginning the migratory journey (Morman et al. 1980, Moore and Schleen 1980, Sorensen and Vrieze 2003). It thus appears well established that adult history leads to differences in larval abundance among watersheds. However, a recent review of management-related sea lamprey research needs included no discussion of the consequences of adult behavior for larval abundance among reaches within one watershed (Jones 2007). My modeling results suggest that migratory adult behavior may have importance consequences for not only the distribution of larvae among watersheds, but also for the distribution of larvae among reaches within a watershed.

Management of sea lamprey in the Laurentian Great Lakes occurs primarily via a chemical lampricide that targets the larval stage (Christie et al. 2003, Slade et al. 2003, Jones 2007). Accurate estimates of larval distribution are critical to cost-effective application of lampricide (Jones 2007). My results have several implications for future efforts to build predictive models of larval distribution. First, larval distribution depends on network-wide population dynamics that depend on larval pheromone and adult behavior. It is therefore unlikely that larval densities could ever be predicted solely from the intrinsic properties of a reach (i.e. habitat). Second, network-wide feedbacks lead to aggregations of larvae and localized density-dependent effects, with the result that larval population size is likely not a linear function of the amount of habitat in a watershed. Network-wide feedbacks leading to aggregations of larvae also alter the correlation between the number of larvae and the amount of habitat in a reach. Density-dependent effects and variability in the correlation between larvae and habitat further complicate the task of using habitat to predict larval abundance.

Stock-recruitment models My model is a spatially-explicit tool for understanding how the strength of density-dependent effects might vary among reaches within a watershed. Stock-recruitment relationships for sea lamprey in the Great Lakes are poorly understood

(Jones 2007), although models have been developed for specific watersheds (Haeseker 2003). Nonlinearity in stock-recruitment relationships is driven by density-dependent effects (Myers 2001), which have been observed in sea lamprey (Purvis 1979, Weine and Pajos 1998, Jones et al. 2003). Larval pheromone and the migratory behavior of spawning adults determine locations within the watershed with the largest aggregations of larvae, and so the greatest density-dependent effects. The downstream drift of larvae is only moderate (Morman et al. 1980, Klar et al. 1997), which suggests that aggregations of larvae created by patterns of adult migration are not diffused by larval behavior. Spatially-explicit models of adult migration such as mine could identify both reaches with low larval densities (few density dependent effects) and reaches with high larval densities (large density dependent effects) and thus refine watershed-wide estimates of stock-recruitment relationships.

References

Applegate, V.C. 1950. Natural history of the sea lamprey (*Petromyzon marinus*) in Michigan. Spec. Sci. Rept. U.S. Fish and Wildlife Serv. 55, 1-237.

Baker, ME and MJ Wiley. 2009. Multiscale controls of flooding in riparian forests of Lower Michigan. Ecology 90: 145-159.

Beamish, F. W. H. 1980. Biology of the North American anadromous sea lamprey, *Petromyzon marinus*. Can. J. Fish. Aquat. Sci. 37: 1924 - 1943.

Bergstedt, R.A., and Seelye, J.G. 1995. Evidence for lack of homing by sea lampreys. Trans. Am. Fish. Soc. 124: 235 - 239.

Berryman, A.A. 2002. Population regulation, emergent properties, and a requiem for density dependence. Oikos 99: 600-606.

Bjerselius, R., Li, W., Teeter, J.H., Seelye, J.G., Johnsen, P.B., Maniak, P.J., Grant, G.C., Polkinghorne, C.N., and Sorensen, P.W. 2000. Direct behavioral evidence that unique bile acids released by larval sea lamprey (*Petromyzon marinus*) function as a migratory pheromone. Canadian Journal of Fisheries and Aquatic Sciences 57: 557 - 569.

Camus, P.A., and Lima, M. 2002. Population, metapopulations, and the open-closed dilemma: The conflict between operational and natural population concepts. Oikos 97: 433-438.

Carde, R.T., and Minks, A.K. 1997. *Pheromone Research: New Directions*. New York: Chapman and Hall.

Christie, G.C., Adams, J.V., Steeves, T.B., Slade, J.W., Cuddy, D.W., Fodale, M.F., Young, R.J., Kuc, M., and Jones, M.L. 2003. Selecting Great Lakes streams for lampricide treatment based on larval sea lamprey surveys. J. Great Lakes Res. 29: 152-160.

Chorley, R.J. 1969. The drainage basin as the fundamental geomorphic unit, in Water, Earth, and Man, ed. by R. Chorley, Methuen & Co, Ltd. 1969.

Coble D.W., Brueseqitz R.E., Fratt T.W., Scheirer J.W. 1990. Lake trout, sea lampreys, and overfishing in the upper Great Lakes: A review and reanalysis. Transactions of the American Fisheries Society: Vol. 119, No. 6 pp. 985–995

Dunning, J.B., Danielson, B.J., and Pulliam, H.R. 1992. Ecological processes that affect populations in complex landscapes. Oikos 65: 169 - 175.

Fahrig, L., and Merriam, G. 1994. Conservation of fragmented populations. Conservation Biology 8: 50 - 59.

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

Haeseker, S.L., Jones, M.L., and Bence, J.R. 2003. Estimating uncertainty in the stock-recruitment relationship for St. Marys River sea lampreys. J. Great Lakes Res. 29: 728 - 741.

Jones, M.L. 2007. Toward improved assessment of sea lamprey population dynamics in support of cost-effective sea lamprey management. J. Great Lakes Res. 33: 35-47.

Kareiva, P. 1990. Population dynamics in spatially complex environments: theory and data. Phil. Trans. R. Soc. Lond. B 330, 175 - 190.

Keller, D.R. and Golley, F.B. 2000. The philosophy of ecology: from science to synthesis. University of Georgia Press, Athens Georgia.

Kotliar, N.B., and Wiens, J.A. 1990. Multiple scales of patchiness and patch structure: a hierarchical framework for the study of heterogeneity. Oikos 59: 253-260.

Li, W., Sorensen, P.W., and Gallaher, D.D. 1995. The olfactory system of migratory adult sea lamprey (*Petromyzon marinus*) is specifically and acutely sensitive to unique bile acids released by conspecific larvae. The Journal of General Physiology 105: 569 - 587.

Li., W., Siefkes, M.J., Scott, A.P., and Teeter, J.H. 2003. Sex pheromone communication in the sea lamprey: implications for integrates management. J. Great Lakes Res. 29 (Suppl. 1): 85 - 94.

Li, W., Twohey, M., Jones, M., and Wagner, M. 2007. Research to guide use of pheromones to control sea lamprey. J. Great Lakes Res. 33 (Special Issue 2): 70 - 86.

Malanson, G. P. 1999. Considering complexity. Annals of the Association of American Geographers 89: 746 – 753.

Montgomery, D.R., Grant, G.E., and Sullivan, K. 1995. Watershed analysis as a framework for implementing ecosystem management. Water Resources Bulletin 31: 369 - 386.

Moore, H.H. and Schleen, L.P. 1980. Changes in spawning runs of sea lamprey (*Petromyzon marinus*) in selected streams of Lake Superior after chemical control. Can. J. Fish. Aquat. Sci. 37: 1851-1860.

Morman, R.H., Cuddy, D.W., and Ruger, P.C. 1980. Factors influencing the distribution of sea lampreyy (*Petromyzon marinus*) in the Great Lakes. Can. J. Fish. Aquat. Sci. 37: 1811 - 1826.

Myers, R.A. 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES Journal of Marine Science 58: 937-951.

Pickett, S.T.A. and Cadenasso, M.L. 1995. Landscape ecology: spatial heterogeneity in ecological systems. Science 269: 331-334.

Post, D.M., Doyle, M.W., Sabo, J.L., and Finlay, J.C. 2006. The problem of boundaries in defining ecosystems: A potential landmine for uniting geomorphology and ecology. Geomorphology 89: 111 - 126.

Potter, I.C., Hillard, R.W., Bradley, J.S., and McKay, R.J. 1986. The influence of environmental variables on the density of larval lampreys in different seasons. Oecologia 70: 443-440.

Quintella, B.R., Andrade, N.O, Dias, N.M., and Almeida, P.R. 2007. Laboratory assessment of sea lamprey larvae burrowing performance. Ecology of Freshwater Fish 16: 177-182.

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

Schlosser, I.J. 1995. Critical landscape attributes that influence fish population dynamics in headwater streams. Hydrobiologia 303: 71 - 81.

Schlosser, I.J. 1995. Dispersal, boundary processes, and trophic-level interactions in streams adjacent to beaver ponds. Ecology 76: 908 - 925.

Slade, J.S., Adams, J.V., Christie, G.C., Cuddy, D.W., Fodale, M.F., Heinrich, J.W., Quinlan, H.R., Weise, J.G., Weisser, J.W., and Young, R.J. 2003. Techniques and methods for estimating abundance of larval and metamorphosed sea lampreys in Great Lakes tributaries, 1995 to 2002. J. Great lakes Res. 29: 137-151.

Smith, B.R. and Tibbles, J.J. 1980. Sea lamprey (*Petromyzon marinus*) in Lakes Huron, Michigan, and Superior: history of invasion and control, 1936-78. Can. J. Fish. Aquat. Sci. 37: 1780-1801.

Sorensen, P.W., and Vrieze, L.A. 2003. The chemical ecology and potential application of the sea lamprey migratory pheromone. J. Great Lakes Research. 29 (Suppl. 1): 66 - 84.

Sorensen, P.W., Christensen, T.A., and Stacey, N.E. 1998. Discrimination of pheromonal cues in fish: emerging parallels with insects. Curr. Opin. Neurobiol. 8: 458 - 467.

Twohey, M.B., Sorensen, P.W., and Li, W. 2003. Possible applications of pheromones in an integrated sea lamprey management program. J. Great Lakes Res. 29: 794 - 800.

Vrieze, L.A. and Sorensen, P.W. 2001. Laboratory assessment of the role of a larval pheromone and natural stream odor in spawning stream localization by migratory sea lamprey (*Petromyzon marinus*). Canadian Journal of Fisheries and Aquatic Sciences. 58, 2374 - 2385,

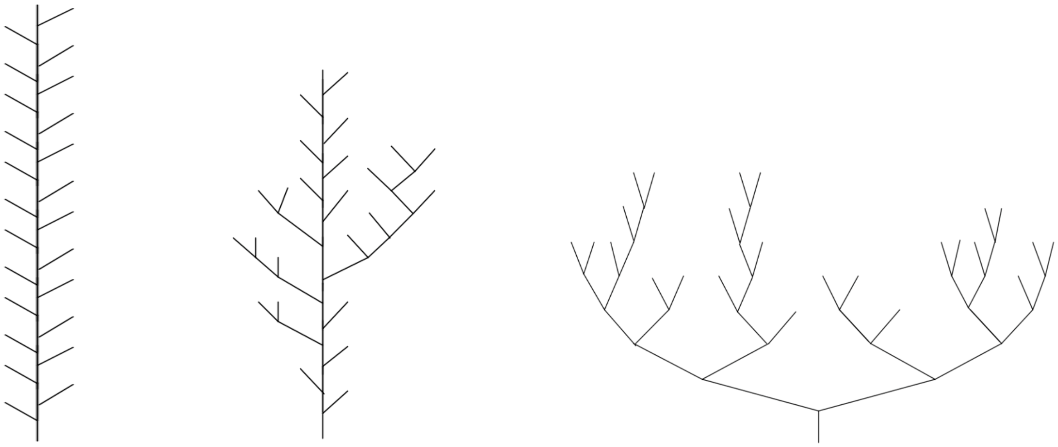
Waldrop, M. Mitchell Waldrop. 1994. Complexity: the emerging science at the edge of order and chaos. Harmondsworth, England. Penguin

Wiegan, T., Moloney, K.A., Naves, J., and Knauer, F. 1999. Finding the missing link between landscape structure and population dynamics: a spatially explicit perspective. American Naturalist 154: 605 - 627.

Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology 3: 385 - 397.

Wiens, J.A., Stenseth, N.C., Van Horne, B, and Ims, R.A. 1993. Ecological mechanisms and landscape ecology. Conservation Biology 11: 1069 - 1080.

With, K.A., Gardner, R.H., and Turner, M.G. 1997. Landscape connectivity and population distributions in heterogeneous landscapes Oikos 78: 151 - 169.



D = 25 D = 16 D = 8

Figure 4.1: Examples of river networks generated by my model. Each network contains 49 reaches. High diameter river networks are elongated and less compact (left-most network; diameter = 25), while low diameter river networks are short, compact, and branchy (right-most network; diameter = 8). The center network has a diameter of 16. Note that there are many possible unique network configurations with the same diameter.

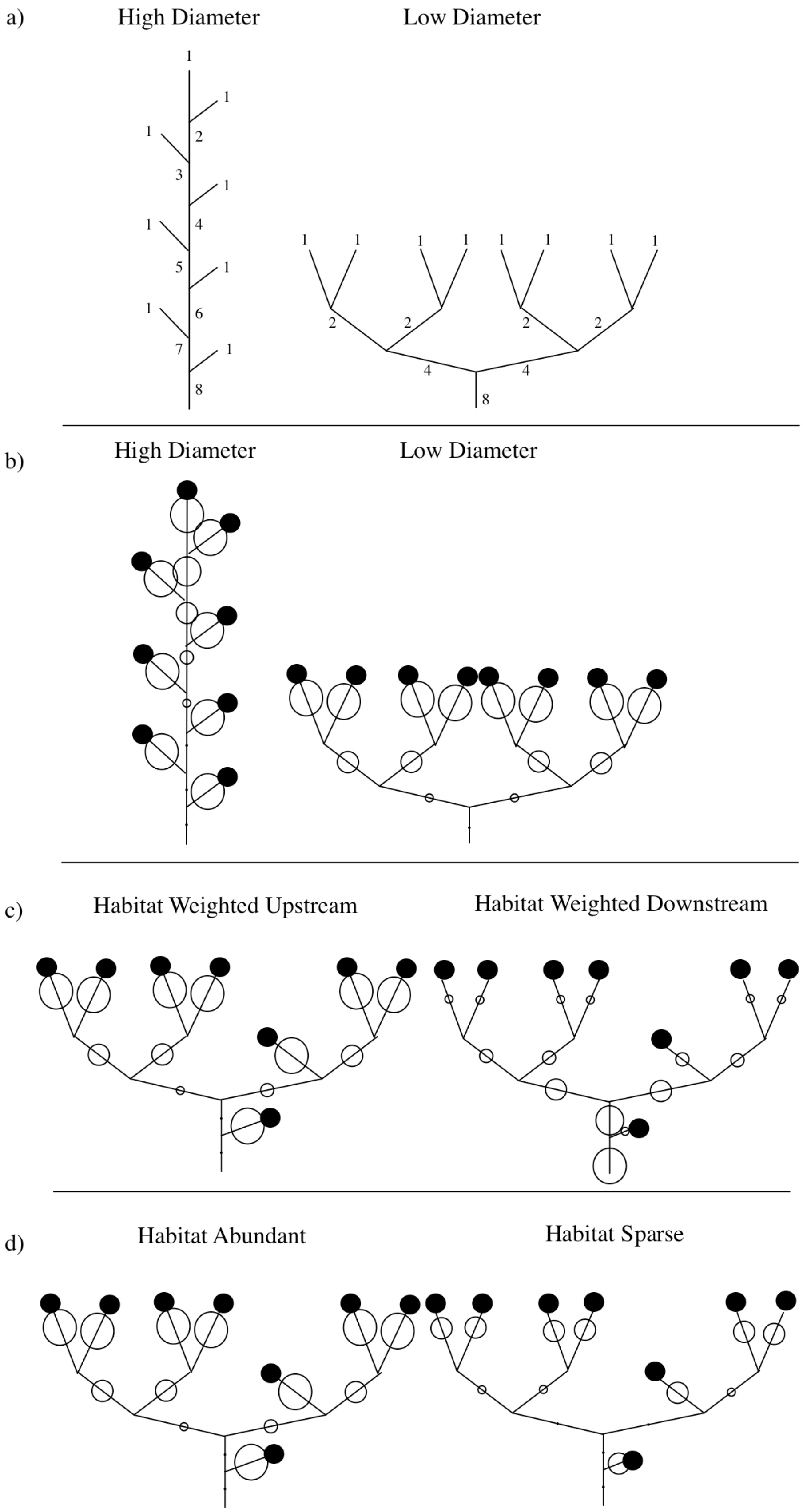
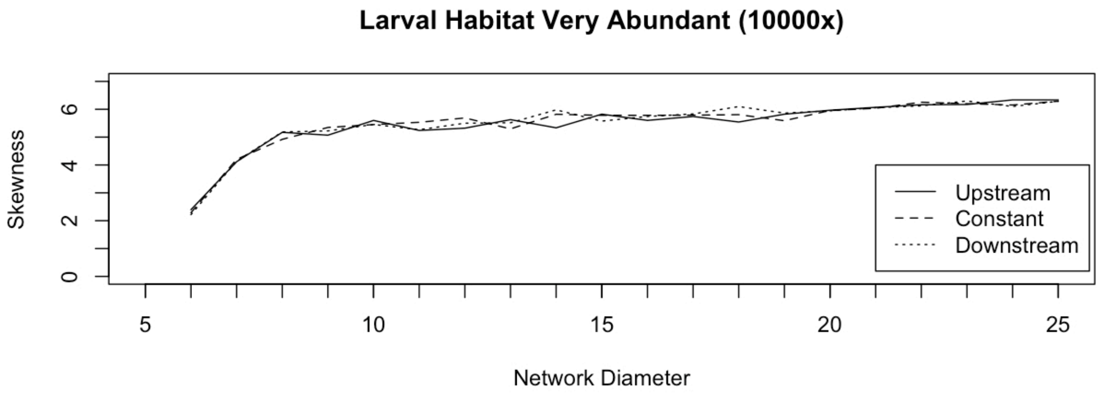
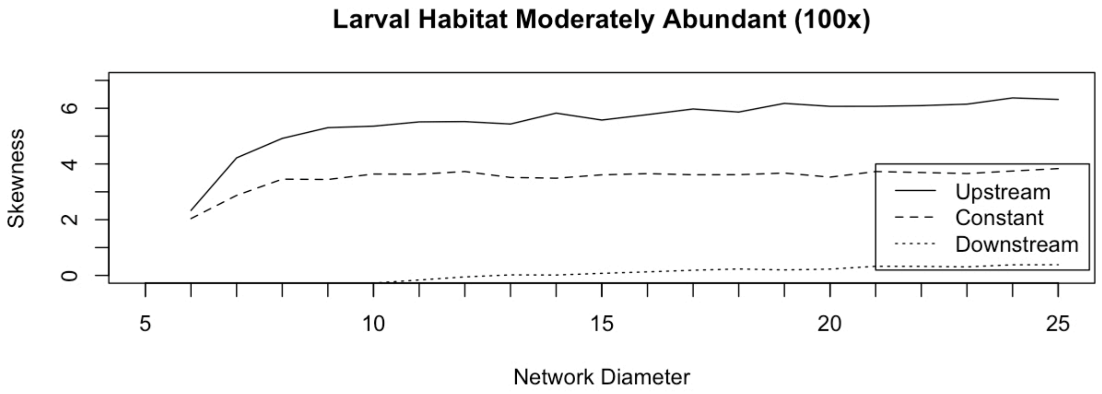

a)
High Diameter
Low Diameter
1
1
1
2
3
1
1
4
5
1
1
6
7
1
8
1
1
1
1
1
1
1
1
2
2
2
2
4
4
8
b)
High Diameter
Low Diameter
c)
Habitat Weighted Upstream
Habitat Weighted Downstream
d)
Habitat Abundant
Habitat Sparse

Figure 4.2: Diagrams showing the relationships between river network structure, link number, habitat distribution, and habitat abundance, using as examples river networks with n = 15 reaches. Networks generated by my model contain n = 49 reaches, as in Fig. 4.1. Panel A demonstrates differences in the spatial pattern of link number (labeled on each reach) between a high diameter river network and a low diameter river network, where link number is defined as the total number of upstream first-order reaches. Link number increases with drainage area and so is strongly correlated with flow. Key dimensions of larval habitat are correlates of flow. I used this relationship as the basis for setting the amount of habitat in a reach as a function of link number. Panel B demonstrates resulting differences in the spatial arrangement of larval habitat (open circles) relative to spawning locations (closed circles), between a high diameter river network and a low diameter river network. Here, the amount of larval habitat in a reach is a function of the inverse of link number, and is represented by circle size. Panel C demonstrates differences in the spatial arrangement of larval habitat and spawning locations between habitat skewed upstream (i.e. habitat is a function of the inverse of link number and negatively correlated with flow) and habitat skewed downstream (i.e. habitat is a function of link number and positively correlated with flow), for the same river network. Panel D demonstrates the effect of altering the total amount of larval habitat in the network by scaling the amount of habitat in each reach by some constant factor *c.*

a)


Larval Habitat Very Abundant (10000x)
Skewness
0
2
4
6
5
10
15
20
25
Network Diameter
Upstream
Constant
Downstream

b)


Larval Habitat Moderately Abundant (100x)
Skewness
0
2
4
6
5
10
15
20
25
Network Diameter
Upstream
Constant
Downstream

c)

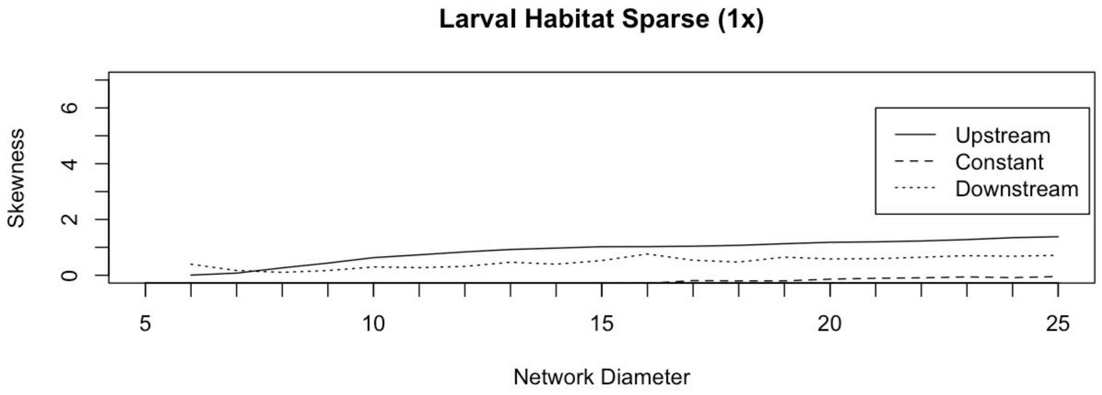
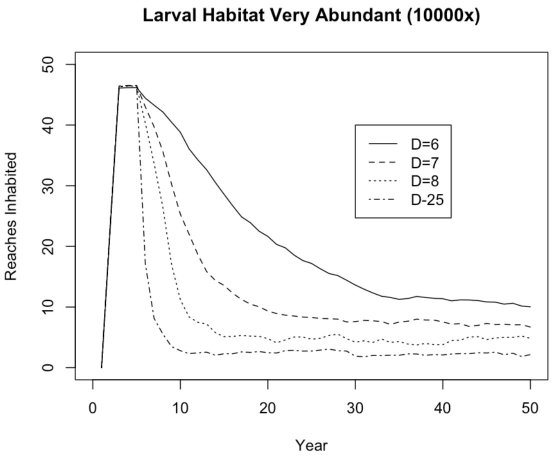
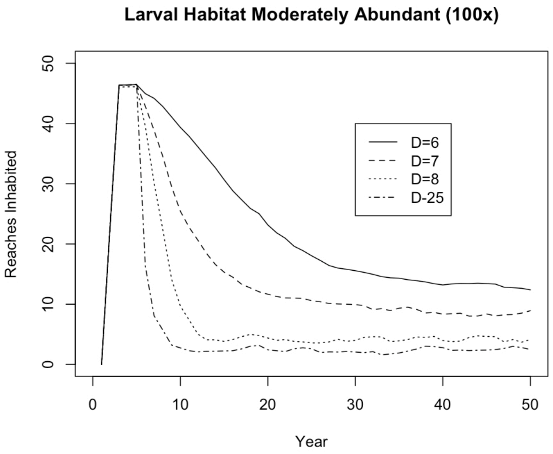

Larval Habitat Sparse (1x)
Skewness
0
2
4
6
5
10
15
20
25
Network Diameter
Upstream
Constant
Downstream

Figure 4.3: Results of the four-year computational experiments regarding the influence of the amount and distribution of larval habitat on the relationship between river network structure and larval distribution (question 1; exp. 1). The figure shows the results of three experimental treatments, where habitat was first very abundant (trt. 1a, Panel A), then moderately abundant (trt. 2a, Panel B), and finally sparse (trt. 3a, Panel C). Each panel gives the relationship between river network diameter (horizontal axis) and the skewness of the distribution of larvae across all reaches in the river network (vertical axis) for the three sets of model runs conducted in each experiment: habitat weighted upstream (solid line), habitat equal in all reaches (dashed line), and habitat weighted downstream (dotted line).

a)


Larval Habitat Very Abundant (10000x)

b)


Larval Habitat Moderately Abundant (100x)

c)

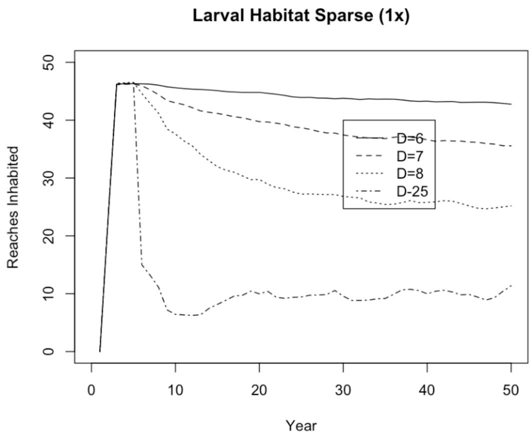
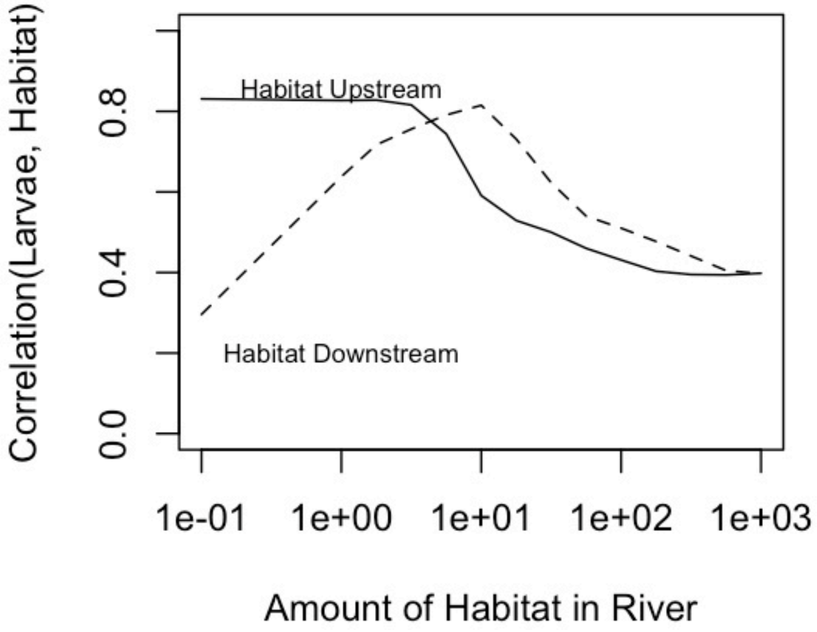

Larval Habitat Sparse (1x)

Figure 4.4: Results of the fifty-year computational experiments regarding the influence of the amount and distribution of larval habitat on the relationship between river network structure and larval distribution (question 1; exp. 2). The figure shows the results of three experimental treatments, where habitat was first moderately abundant (trt. 2a, Panel A), then minimally abundant (trt. 2b, Panel B), and finally sparse (trt. 2c, Panel C). Habitat was weighted upstream in all three experiments. Each panel gives the change in the number of reaches inhabited over the course of fifty years for network diameters D = 6, 7, 8, and 25 (line types, see legend). Each line shows the mean number of reaches inhabited over 30 model runs.

a)


Correlation(Larvae, Habitat)
0.0
0.4
0.8
Habitat Upstream
Habitat Downstream
1e-01
1e+00
1e+01
1e+02
1e+03
Amount of Habitat in River

b)

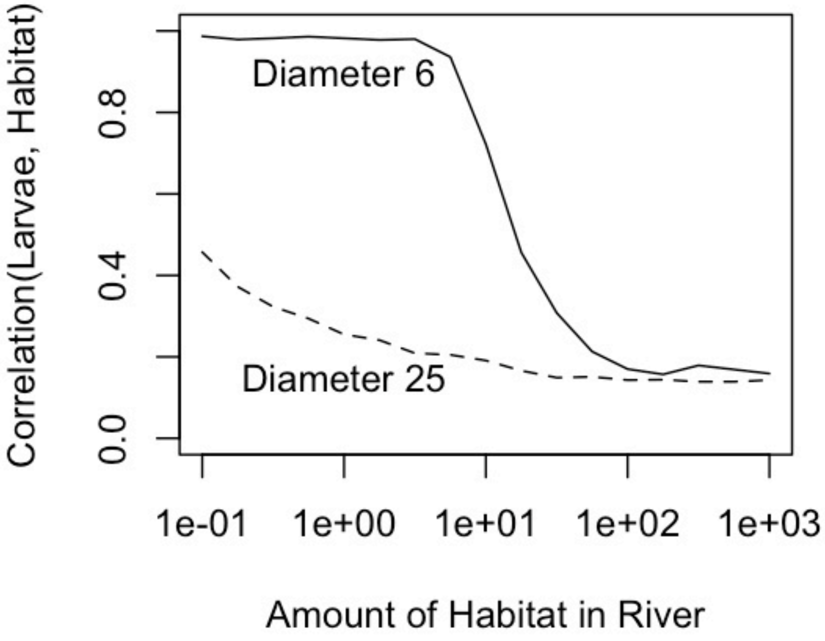
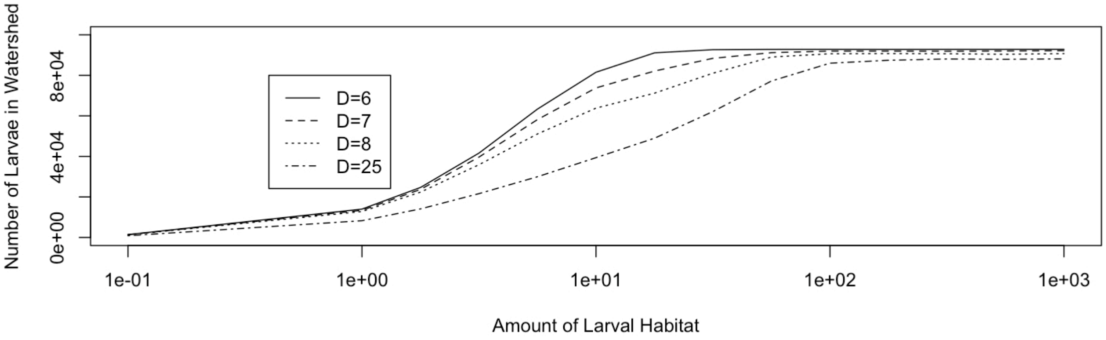
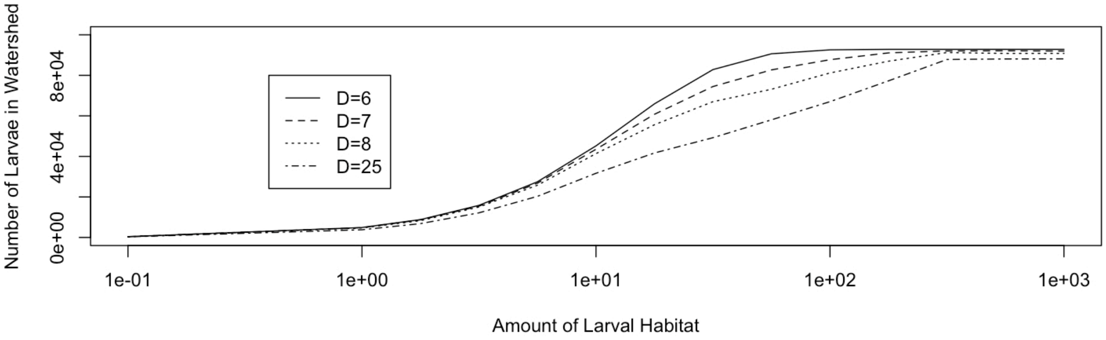

Correlation(Larvae, Habitat)
0.0
0.4
0.8
Diameter 6
Diameter 25
1e-01
1e+00
1e+01
1e+02
1e+03
Amount of Habitat in River

Figure 4.5: Results of the computational experiment regarding the influence of network structure and the amount and distribution of habitat on the correlation between the number of larvae and the amount of habitat in a reach (question 2; exp. 3). The vertical axis in both panels is the correlation between the number of larvae and the amount of habitat in a reach, measured for all reaches in the river network, after four years. Panel A (trt. 3a) shows how this correlation changes with the amount of larval habitat in the river network (horizontal axis) when habitat is weighted upstream (solid line) and when habitat is weighted downstream (dashed line), both for a fixed river network of diameter 8 (shown in Fig. 4.1). Panel B (trt. 3b) shows how this correlation changes with the amount of larval habitat in the river network (horizontal axis), for a river network of diameter 6 (solid line) and a river network of diameter 25 (dashed line), both with larval habitat skewed upstream.

a)


D=6
D=7
D=8
D=25
Number of Larvae in Watershed
0e+00
4e+04
8e+04
1e-01
1e+00
1e+01
1e+02
1e+03
Amount of Larval Habitat

b)


D=6
D=7
D=8
D=25
Number of Larvae in Watershed
0e+00
4e+04
8e+04
1e-01
1e+00
1e+01
1e+02
1e+03
Amount of Larval Habitat

c)

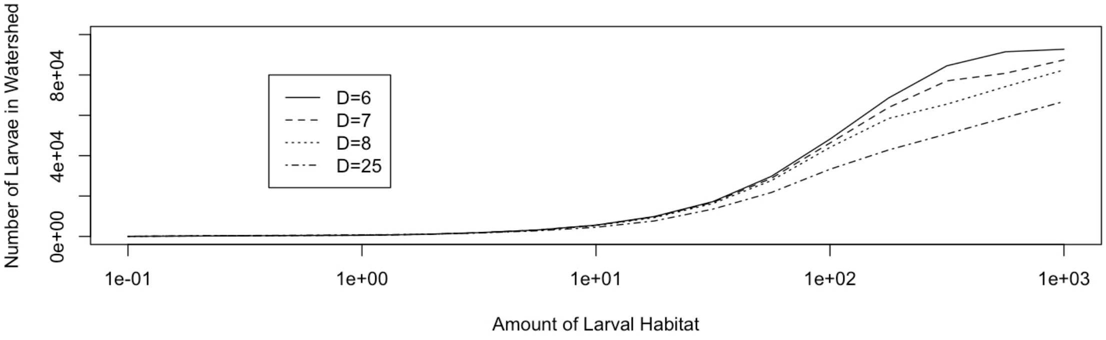
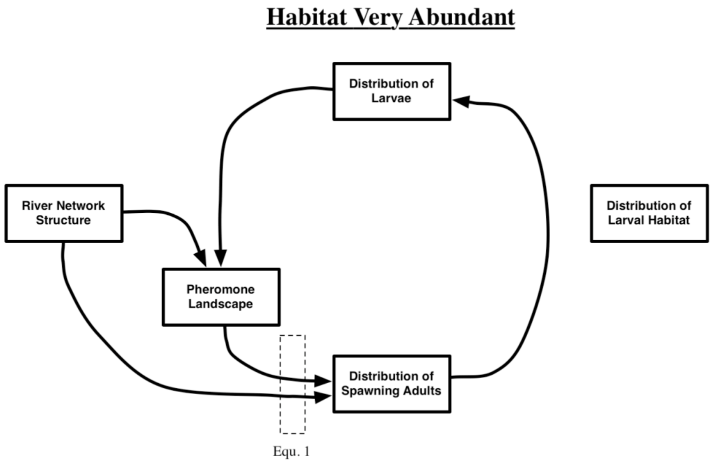
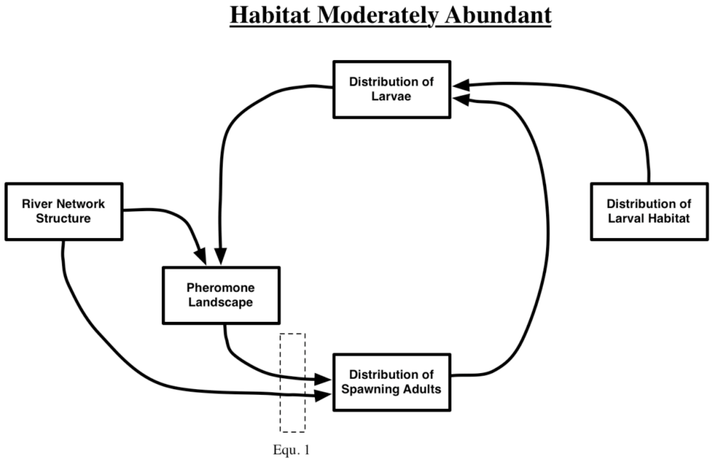

D=6
D=7
D=8
D=25
Number of Larvae in Watershed
0e+00
4e+04
8e+04
1e-01
1e+00
1e+01
1e+02
1e+03
Amount of Larval Habitat

Figure 4.6: Results of the four-year computational experiment regarding the influence of network structure and the amount and distribution of habitat on the total number of larvae in the river network (question 3, exp. 4). The figure shows the results of three experiments, where habitat was first weighted upstream (trt. 4a, Panel A), then equal in all reaches (trt. 4b, Panel B), and finally weighted downstream (trt. 4c, Panel C). Each panel shows how the total number of larvae in the river network (vertical axis) changes with the amount of larval habitat in the river network (horizontal axis) for river networks of diameter D = 6, 7, 8, and 25 (line types, see legend).

a)


Habitat Very Abundant
Distribution of Larvae
River Network Structure
Distribution of Larval Habitat
Pheromone Landscape
Distribution of Spawning Adults
Equ. 1

b)


Habitat Moderately Abundant
Distribution of Larvae
River Network Structure
Distribution of Larval Habitat
Pheromone Landscape
Distribution of Spawning Adults
Equ. 1

c)

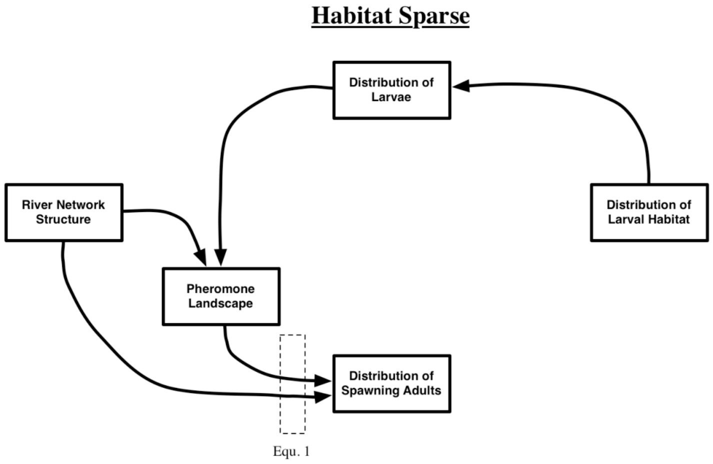

Habitat Sparse
Distribution of Larvae
River Network Structure
Distribution of Larval Habitat
Pheromone Landscape
Distribution of Spawning Adults
Equ. 1

Figure 4.7: Diagrams of feedbacks between the distribution of adults and larvae in my model, when larval habitat is very abundant (Panel A), moderately abundant (Panel B), and sparse (Panel C). In all three scenarios, larval distribution and patterns in the relative sizes of confluent tributaries (i.e. river network structure) determine the downstream propagation pattern of larval pheromone, the pheromone landscape. Adult lamprey begin their spawning run at the mouth of the river and migrate upstream by probabilistically choosing between confluent tributaries on the basis of their relative size and pheromone concentration according to equ. 1. The abundance of larval habitat determines how the distribution of spawning influences the distribution of larvae. When larval habitat is very abundant (Panel A) it never constrains larval distribution, and the distribution of larvae is determined only by the distribution of spawning. Newly-created larvae alter the pheromone landscape for subsequent generations, creating a path-dependent feedback loop between larvae and the migration and spawning of adults. When larval habitat is only moderately abundant (Panel B) and limits larval abundance in some locations, feedbacks between larvae and adults are mediated by the distribution of larval habitat. When larval habitat is extremely rare (Panel C), larval habitat is a bottleneck that saturates with larvae regardless of the distribution of spawning adults, and there are no feedbacks between the distribution of larvae and the distribution of spawning adults.

Table 4.1: Summary of model setup for all computational experiments described in this chapter. The table gives: research question, experiment, and experimental treatment, which correspond to those in Methods and Results; Parameterization, the number of unique combinations of model parameters in each experimental treatment; the duration of the model run, in years; the network diameters of the rivers used; the amount of larval habitat, including sparse (1x), minimally abundant (10x), moderately abundant (100x), and very abundant (10000x); the longitudinal distribution of larval habitat, including concentrated upstream (U), equal in all reaches (C), and concentrated downstream (D); the measure of interest for results, including the skewness of the distribution of larval abundance across all reaches in the river network ("skewness"), the number of reaches inhabited during each timestep, the correlation between the number of larvae and the amount of habitat in each reach ("cor(larvae, habitat)"), and the total number of larvae in the watershed ("population size"); and the resulting Figure.

Question	Experiment	Treatment	Parameterizations	Duration (years)	Network Diameter	Habitat Amount	Habitat Distribution	Response Variable Measured	Figure
1	1	1a	1-60	4	6:25	10000x	U, C, D	skewness	3a
1	1	1b	61-120	4	6:25	100x	U, C, D	skewness	3b
1	1	1c	121-180	4	6:25	1x	U, C, D	skewness	3c
1	2	2a	181-184	50	6,7,8,25	100x	U	number of reaches inhabited	4a
1	2	2b	185-188	50	6,7,8,25	10x	U	number of reaches inhabited	4b
1	2	2c	189-192	50	6,7,8,25	1x	U	number of reaches inhabited	4c
2	3	3a	193-226	4	8	1x:10000x	U, D	cor(larvae, habitat)	5a
2	3	3b	227-294	4	6,7,8,25	1x:10000x	U	cor(larvae, habitat)	5b
3	4	4a	295-350	4	6,7,8,25	1x:10000x	U	population size	6a
3	4	4b	351-407	4	6,7,8,25	1x:10000x	C	population size	6b
3	4	4c	408-463	4	6,7,8,25	1x:10000x	D	population size	6c

Chapter 5

Conclusion

In this final chapter I provide a review of the main findings of my dissertation, and then discuss the implications of my results for sea lamprey management in the Great Lakes. Although earlier chapters included some discussion of management, here I examine how the results of each chapter provide context for interpreting the results of other chapters. Specifically, I consider how patterns in the distribution of larval habitat in Michigan rivers (Chapter 2) provide a point of reference for interpreting my sea lamprey population models (Chapters 3 and 4), and how those population models better define the role of habitat maps in lamprey management.

Principal Findings

Chapter 2: Towards a process domain-sensitive sea lamprey habitat model for Michigan rivers

In Chapter 2, I hypothesized that the relationship between landscape and instream habitat across the lower peninsula of Michigan spanned multiple process domains, or regions among which the relationship between landscape and habitat differed functionally. A key methodological question was how best to partition the study area such that the relationship between landscape and riverine habitat within each partition was simpler and more amenable to regression modeling than the complex global relationship. The null linear regression model fit to the entire study region performed poorly, as expected. Stratifying the study area by watershed provided a large increase in fit relative to the null model, suggesting that there was spatial variability in the relationship between landscape and habitat. Last, I used regression trees as a machine learning method for partitioning the study area and creating a domain-sensitive predictive habitat model. I created three trees, where tree splits delineated: (1) spatially contiguous units, (2) non-contiguous units defined by values of the covariates, and (3) both contiguous and non-contiguous units. The adjusted R^2 values of the three trees were 0.30, 0.30, and 0.32, respectively, and all three trees outperformed the set of models fit to each watershed individually.

The trees identify useful stratifications of the lower peninsula of Michigan, the important geomorphic influences on habitat, and how the influence of geomorphology on habitat varies across the study region. Both the spatial tree and the hybrid tree (models 3 and 5) identified an East-West split through the center of the lower peninsula of Michigan. The spatial and hybrid trees identify geographic areas that could serve as the basis for management units or stratifications for partitioning future habitat models. The distance from the mouth of the river, the distance to the nearest upstream dam or lake, and stream power were the covariates used most frequently in the terminal node regression relationships across all three trees. The importance of covariates also differed among geographic regions. In the spatial tree (s3), for example, sampling locations in the Northwest portion of the lower peninsula (green in Fig. 2.2) never included the distance to the nearest upstream or downstream dam or lake as a covariate. Datapoints in the Easternmost portion of the dataset (red in Fig. 2.2) never included distance to the nearest downstream dam or lake as a covariate.

Chapter 3: River network structure affects sea lamprey distribution

In Chapter 3, I developed a computational model to test the hypothesis that river network structure influences the distribution of larval sea lamprey across a watershed. I performed two sets of experiments with this model. First, I initialized the model using randomly generated river networks, and quantified how river network shape influenced the distribution of sea lamprey in the model systems. I found a strong positive relationship between network diameter (a measure of network shape) and the aggregation of larvae across reaches in the model-generated networks. Second, I initialized the model using a river network based on the Muskegon River, Michigan, and compared larval distribution by reach generated by the model to available survey data. I found a significant correlation between model-generated larval abundance and the field-measured larval densities (rho = 0.46; $p < 0.005$).

In my model, river network structure by itself had a strong influence on the distribution of sea lamprey larvae. The general result that network structure influences larval distribution was very robust to changes in model parameterization, and disappeared only when adults paid little attention to larval pheromone. This scenario seems unrealistic given the empirical studies demonstrating the behavioral response of adult sea lamprey to larval pheromone (Li et al. 1995, Bjerselius et al. 2000, Vrieze and Sorensen 2001, Sorensen and Vrieze 2003).

The mechanisms in my model are different than those typically considered to be important for sea lamprey distribution and population dynamics. The preferred habitat for larval sea lamprey (Applegate 1950, Potter et al. 1986) and spawning (Applegate 1950) is well defined, and the local-scale distribution of larvae is typically assumed to be determined by the availability of preferred substrate habitat. In my model, the large-scale distribution of larvae across a watershed is a function of a complex system of feedbacks constrained by river network structure. My results reinforce the importance of considering rivers as “riverscapes”, wherein rivers are continuous, hierarchical, and heterogeneous, and spatial pattern at several scales influences the population dynamics of resident fishes.

Chapter 4: The watershed as place: Sea lamprey as a case study in defining boundaries for population analyses and conservation biology

In Chapter 4, I expanded the computational model developed in Chapter 3 to include larval habitat. I performed a series of experiments to examine how habitat, river network structure, and fish behavior interact to determine larval distribution, the correlation between larvae and habitat, and total larval population size. In particular, I explored how the amount and distribution of habitat in a river network mediates the population dynamics observed in Chapter 3.

I observed four main results: 1) Larval distribution in my model depended on both river network structure and the distribution of larval habitat, but the relative importance of these two factors was determined by the overall availability of larval habitat. Furthermore, the availability of larval habitat determined whether network-dependent feedbacks existed between adults and larvae; 2) As in Chapter 3, larvae in the model aggregated over time. In Chapter 4, the degree and rate of this aggregation depended on both river network structure and the availability of larval habitat; 3) The correlation between larval abundance and amount of larval habitat in a reach depended on the overall abundance of larval habitat; and 4) Total larval population size in a watershed depended on factors that control larval aggregation, i.e. river network structure and the amount and distribution of larval habitat.

Previous work related larval distribution to the distribution of larval habitat (Applegate 1950, Malmqvist 1980, Potter et al. 1986) and to network-dependent patterns of adult migration (Chapter 3). In Chapter 4, I demonstrated how and when larval distribution could be influenced by both the distribution of larval habitat and river network structure. Based on my modeling results and other published literature, I argued that the watershed constitutes an important functional unit for sea lamprey, where lamprey within a watershed interact yet are largely decoupled from lamprey in adjacent watersheds.

Implications for Sea Lamprey Management in the Great Lakes

Sea lamprey in the Great Lakes are controlled primarily with a chemical lampricide that targets the sedentary larval stage (Slade et al. 2003, Christie et al 2003, Jones 2007). Budget constraints limit the number of streams that may be treated each year (Christie et al. 2003). Cost-effective application of lampricide involves prioritizing stream reaches for lampricide treatment on the basis of expected production of metamorphosing larvae. Broadly, both larval habitat models and population models are useful if they are capable of either directly or indirectly improving the accuracy and precision of estimates of larval year-class abundance.

Ecological systems of interest to conservation biologists and managers are often complex to the degree that no single field survey, population model, or habitat model is likely to suffice as the sole vehicle for understanding that system. Larval sea lamprey in rivers are patchily distributed (Applegate 1950, Malmqvist 1980, Potter et al. 1986) and difficult to sample, with the result that field surveys provide only rough (Hansen et al.

2003, Jones 2007) and expensive ($4 million/year, Jones 2007) estimates of population size. Computational models of larval populations might inform field surveys or be used in a meta-analysis to refine field-based estimates of larval abundance (Arnqvist and Wooster 1995), yet such models incorporate their own sources of uncertainty. Furthermore, my results in Chapter 4 suggest that population models would require river-specific estimates of the distribution of larval habitat in order to generate river-specific estimates of larval distribution and population size. Maps of larval habitat are themselves difficult to create: the linkages between landscape-scale geomorphic covariates and instream habitat are sufficiently complex (Langbein and Schumm 1958, Shreve 1979, Madej and Ozaki 1996, Seelbach et al. 1997, Phillips 2003 and 2006) that statistical models that predict reach-scale instream habitat may never be more than moderately successful.

Field surveys, habitat models, and population models must each play mutually-supporting roles in an integrated attempt to better predict larval abundance. A key question is how these various methods might best complement one another. In the next paragraphs, I attempt to synthesize previous chapters to suggest how the results of my population models in Chapters 3 and 4 better define the management role of habitat models, and how the results of my habitat modeling work in Chapter 2 provide a point of reference for interpreting population models.

My model results in Chapter 4 demonstrate how larval distribution could be a function of both network structure and the availability of larval habitat, with the implication that larval habitat maps might be at best only moderately useful as a proxy for the distribution of larvae. Habitat maps might be most useful as one component of an integrated effort to evaluate how removal of aging low-head dams might allow sea lamprey to spread through river systems. In this application, habitat maps could be combined with computational models that predict adult migration patterns (e.g. Chapter 4) to estimate which reaches upstream of existing dams might attract the highest densities of lamprey. My regression tree models in Chapter 2 furthermore identify domains where larval habitat is moderately predictable from geomorphic covariates ($R^2 > 0.5$) as well as domains where larval habitat is less predictable ($R^2 < 0.1$), and thus identify domains where habitat models might be most (and least) useful for predicting distribution of larval habitat.

Maps of larval habitat across the lower peninsula of Michigan provide a point of reference for interpreting the results of the population model developed in Chapter 4. In my model, the overall abundance of larval habitat determined the relative importance of network structure and larval habitat in structuring larval distribution (Fig. 4.7). Maps of larval habitat therefore provide watershed-specific information about which spatial structures are likely to have the greatest influence on larval distribution. In river systems where larval habitat is abundant (e.g. Good Harbor Creek, mean of 50.7% preferred larval habitat in survey transects; Watershed ID 18 in Fig. 2.1), the distribution of larval habitat may play a secondary role to network structure in determining larval distribution. In river systems where larval habitat is rare (e.g. Swan Creek, mean of 5.9% preferred larval habitat in survey transects; Watershed 34 in Fig. 2.1), the distribution of larval habitat may play a more important role in determining the distribution of larvae. Watersheds where larval habitat is rare may therefore be good candidates for fine-scale mapping of larval habitat that might then serve as a proxy for the distribution of larvae.

References

Applegate, V.C. 1950. Natural history of the sea lamprey (*Petromyzon marinus*) in Michigan. Spec. Sci. Rept. U.S. Fish and Wildlife Serv. 55, 1-237.

Arnqvist, G., and Wooster, D. 1995. Meta-analysis: synthesizing research findings in ecology and evolution. Trends in Ecology and Evolution 10: 236-240.

Bjerselius, R., Li, W., Teeter, J.H., Seelye, J.G., Johnsen, P.B., Maniak, P.J., Grant, G.C., Polkinghorne, C.N., and Sorensen, P.W. 2000. Direct behavioral evidence that unique bile acids released by larval sea lamprey (*Petromyzon marinus*) function as a migratory pheromone. Canadian Journal of Fisheries and Aquatic Sciences 57: 557 - 569.

Christie, G.C., Adams, J.V., Steeves, T.B., Slade, J.W., Cuddy, D.W., Fodale, M.F., Young, R.J., Kuc, M., and Jones, M.L. 2003. Selecting Great Lakes streams for lampricide treatment based on larval sea lamprey surveys. J. Great Lakes Res. 29: 152-160.

Hansen, M.J., Adams, J.V., Cuddy, D.W., Richards, J.M., Fodale, M.F., Larson, G.L, Ollila, D.J., Slade, J.W., Steeves, T.B., Young, R.J., and Zerrenner, A. 2003. Optimizing larval assessment to support sea lamprey control in the Great Lakes. J. Great Lakes Res. 29: 766-782.

Jones, M.L. 2007. Toward improved assessment of sea lamprey population dynamics in support of cost-effective sea lamprey management. J. Great Lakes Res. 33: 35-47.

Langbein, W.B. and Schumm, S.A. 1958. Yield of sediment in relation to mean annual precipitation. Transactions of the American Geophysical Union 39: 1076-1084.

Madej, M.A. and Ozaki, V. 1996. Channel response to sediment wave propagation and movement, Redwood Creek, California, USA. Earth Surface Processes and Landforms 21: 911-927.

Malmqvist, B. 1980. Habitat selection of larval brook lampreys (*Lampetra planeri,* Bloch) in a south Swedish stream. Oecologia 45: 35-38.

Phillips, J.D. 2003. Sources of nonlinearity and complexity in geomorphic systems. Progress in Physical Geography 27: 1-23.

Phillips, J.D. 2006. Evolutionary geomorphology: thresholds and nonlinearity in landform response to environmental change. Hydrol. Earth Syst. Sci. 10: 731-742.

Potter, I.C., Hillard, R.W., Bradley, J.S., and McKay, R.J. 1986. The influence of environmental variables on the density of larval lampreys in different seasons. Oecologia 70: 443-440.

Seelbach P. W., M. J. Wiley, J. C. Kotanchik, and M. E. Baker. 1997. A landscape-based ecological classification system for river valley segments in lower Michigan (MI-VSEC Version 1.0). Michigan Department of Natural Resources, Fisheries Division, Fisheries Research Report 2036, Lansing.

Shreve, R.L. 1979. Models for prediction in fluvial geomorphology. Mathematical Geology 11: 165-173.

Slade, J.S., Adams, J.V., Christie, G.C., Cuddy, D.W., Fodale, M.F., Heinrich, J.W., Quinlan, H.R., Weise, J.G., Weisser, J.W., and Young, R.J. 2003. Techniques and methods for estimating abundance of larval and metamorphosed sea lampreys in Great Lakes tributaries, 1995 to 2002. J. Great Lakes Res. 29: 137-151.

Sorensen, P.W., and Vrieze, L.A. 2003. The chemical ecology and potential application of the sea lamprey migratory pheromone. J. Great Lakes Research. 29 (Suppl. 1): 66 - 84.

Vrieze, L.A. and Sorensen, P.W. 2001. Laboratory assessment of the role of a larval pheromone and natural stream odor in spawning stream localization by migratory sea lamprey (*Petromyzon marinus*). Canadian Journal of Fisheries and Aquatic Sciences 58: 2374 - 2385,