Graduate Theses, Dissertations, and Problem Reports

2008

Predictive modeling of freshwater mussels (Unionidae) in the Appalachians

Alison R. Mynsberge
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation

Predictive Modeling of Freshwater Mussels (Unionidae) in the Appalachians

Alison R. Mynsberge

Thesis submitted to the
Davis College of Agriculture, Forestry, and Consumer Sciences
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Wildlife & Fisheries Resources

Patricia M. Mazik, Ph.D., Co-Chair
Michael P. Strager, Ph.D., Co-Chair
Jacquelyn M. Strager, M.S.

Department of Wildlife & Fisheries Resources

Morgantown, West Virginia
2008

Keywords: freshwater mussels, predictive modeling, species occurrences, watershed, Appalachians

ABSTRACT

Predictive Modeling of Freshwater Mussels (Unionidae)
in the Appalachians

Alison R. Mynsberge

Freshwater mussels are in decline, particularly in the Appalachian region of North America. This region contains the world's greatest diversity of freshwater mussels, but many species are now threatened or endangered. Little is known of the basic ecology and distributions of species of freshwater mussels relative to other freshwater organisms. The goal of this study was to use predictive modeling to predict distributions of freshwater mussels in the Appalachians and identify correlated factors using a watershed framework. Models were developed in the upper Mid-Atlantic and Ohio drainage regions using subwatersheds and separately in the Tennessee region using catchments. Models developed at this scale had low predictive ability because few surveys of freshwater mussels are available at the subwatershed scale and a regional extent. Independent data were unavailable to evaluate catchment-based models. Additional mussel surveys are necessary to expand the potential for developing robust predictive models of most freshwater mussel species.

Dedication

This work is dedicated to my husband, Shawn Crimmins, for his extraordinary patience, assistance, and encouragement during my pursuit of a graduate education. I also owe my family thanks for their financial and emotional support. Without their undying encouragement, I would not be here today.

I would also like to thank many people at West Virginia University for helping me through this process. Pat Mazik has helped me work out professional questions and personal concerns throughout my graduate career, Mike and Jackie Strager have been wonderful examples of patient, hard-working, and ethical researchers, and Petra Wood, Matt Shumar, Becky Nestor, Steve Harmon, Sarah Woodruff, and others have offered crucial support for this project. My fellow graduate students, including Zina Hense, Angie Burns, Jen Zimmerman, Kerry Wixted, Kelley Flaherty, Lisa Tager, and others also helped me concentrate on the importance of friendship and cooperation when schoolwork became overwhelming.

Finally, I am deeply indebted to the researchers who have studied freshwater mussels for decades and who graciously consented to sharing their wealth of knowledge with me and with this project. Mussels are fascinating organisms that benefit from the dedication of those who study them, and I hope that more attention will be paid to them so that we can be confident in the conservation of the diversity of unique species we have in North America.

Acknowledgments

Table of Contents

APPENDICES

Chapter 1: Literature Review

Introduction

Freshwater mussels contribute to the functioning of aquatic ecosystems in many ways. Mussels also rely on components of freshwater ecosystems for food, habitat, and reproduction. Because freshwater mussels are an important part of the ecosystems in which they occur, it is necessary to describe the state of these systems to understand the ecology of mussel communities and to discuss implications for their conservation and management. This literature review focuses on freshwater mussels and their ecosystems in the eastern United States, along with techniques for modeling distributions of freshwater mussels and other aquatic species.

Freshwater ecosystems

Importance

Diverse roles of freshwater ecosystems in society and in nature necessitate multiple uses for most freshwater systems. Intact freshwater systems are ecologically necessary to conserve native species, which are valuable for their intrinsic worth and for human use (Naiman and Turner 2000). Healthy freshwater ecosystems provide habitats for terrestrial and aquatic organisms and have ecological and aesthetic value; they include native aquatic communities and provide suitable water sources and even larval habitats for terrestrial organisms. These systems are declining due to human uses, which can directly degrade water and habitat quality or lead to ecosystem changes by disrupting the water cycle (Jackson et al. 2001). Balancing multiple uses across the landscape is therefore an important concern for modern management of freshwater ecosystems (Naiman and Turner 2000).

Management concerns

The downstream nature of flow in stream networks complicates management needs (Dudgeon et al. 2006). The effects of channel alteration, pollution, and land-use changes to the watershed and riparian area can influence not only the immediate channel, but larger downstream systems (Gregory et al. 1991; Roth et al. 1996; Johnson and Gage 1997; Malmqvist and Rundle 2002; Allan 2004). Understanding the hydrologic connectivity that relates immediate stream

condition to upstream and downstream conditions is critical to conserving freshwater ecosystems (Frissell et al. 1986; Hitt and Angermeier 2008), but can be difficult to incorporate into management actions (Pringle 2001). In the past, humans have altered the landscape with little consideration for the effects of changes across a broad extent (Schlosser 1991). Consequently, anthropogenic changes to the environment have affected freshwater ecosystems in a number of ways, including degradation of water quality and habitat for freshwater organisms (Naiman and Turner 2000).

Freshwater ecosystems are threatened in the United States (U.S.) and throughout the world (Nilsson et al. 2005). Impoundments alone have fragmented over 75% of the main channel of three large river systems of the U.S., including the Columbia, the Connecticut, and the Mobile river systems, and a total of 26 large river systems in North America are strongly affected by fragmentation and flow regulation (Dynesius and Nilsson 1994; Nilsson et al. 2005). Ensuring the sustainability of freshwater ecosystems is therefore a difficult task (Dudgeon et al. 2006). In the U.S., federal and state regulations require monitoring and maintenance of water quality and of species listed under the Endangered Species Act, yet research indicates that 123 freshwater animal species have been recorded as extinct in North America since 1900, and projected freshwater extinction rates are similar to the estimated rates of loss for tropical rainforest communities (Ricciardi and Rasmussen 1999). Many experts believe that exotic species, contaminants, agricultural non-point source pollution, habitat removal, and habitat alteration limit the continued survival or recovery of more than 10% of freshwater faunal species in the U.S. (Richter et al. 1997). Conserving functioning freshwater communities requires addressing these threats to aquatic systems.

General threats

Anthropogenic activities directly and indirectly threaten aquatic ecosystems. The composition of aquatic communities is affected by chemical pollutants as well as physical habitat and community changes. Point and non-point source pollution leading to sub-lethal or lethal levels of contaminants have decreased water quality and changed species richness and abundance in freshwater ecosystems (see Fleeger et al. 2003 for review). Physical changes to watersheds, streams, and lakes have contributed to increased siltation and physical habitat degradation, which can affect the short- and long-term survival of aquatic organisms (Waters 1995). The addition of

invasive species and extirpation of endemic species contribute to significantly altered aquatic communities, which can affect the normal functioning of freshwater ecosystems (Miller et al. 1989; Gido and Brown 1999; Hakenkamp and Palmer 1999; Kelly et al. 2006). A variety of human activities is responsible for these threats to aquatic ecosystems.

Concentrated discharges of chemical pollutants from industrial sources have caused massive fish and mussel kills (Sheehan et al. 1989), which are the most obvious detrimental effects of water quality on aquatic fauna. Although water quality may return to normal shortly after an industrial spill, aquatic species may take months or years to recover (Sietman et al. 2001). Lethal pollutants from point sources such as factories or accidents are relatively easy to identify and correct, but less obvious effects on freshwater ecosystems can be caused by sub-lethal pollutants. Pollution may not always cause immediate reductions in the health or fecundity of freshwater organisms, but may lead to chronic deleterious effects (Hanazato 2001). The discharge of sub-lethal levels of chemical pollutants may impair stream quality by reducing the health or fecundity of organisms downstream of the chemical release (Gagne et al. 2002). Because specific chemicals responsible for sub-lethal effects in aquatic ecosystems can be difficult to identify across large extents, analyses of chemical loads in indicator species have been used to determine the types and potential sources of pollutants affecting a system (Flessas et al. 2000; McKinney et al. 2002; Brown et al. 2005; Henderson 2007). Acid mine drainage is a point source that is more difficult to address, and it is a common factor contributing to acidity and high metal concentrations in streams of the Appalachians because of mining activities in the Appalachian Basin coal field stretching from Pennsylvania to Alabama (Averitt 1975).

Non-point source pollution, or pollution originating from runoff or atmospheric deposition, can also affect freshwater systems. Examples of non-point source pollution include acid rain and increased levels of nitrogen and phosphorus from agricultural fertilizers, factors that can affect community composition and that can lead to eutrophication and the creation of an oxygen-limited environment for fish and macroinvertebrates (Miltner and Rankin 1998; Chambers et al. 2006). Acidic streams resulting from acid mine drainage or acid rain in watersheds with low acid-neutralizing capacity are common in the Appalachians (Herlihy et al. 1989) and have different assemblages of fish and macroinvertebrates and lower densities of benthic macroinvertebrates than healthy streams (Haines 1981), even following stream restoration (McClurg et al. 2007). Reducing non-point source pollution is complicated by the

difficulty of regulating pollutants originating from diverse and potentially widespread sources (Carpenter et al. 1998).

Physical changes to the environment can also affect freshwater ecosystems. Anthropogenic perturbations to ecosystems have affected stream habitats and general stream health through increased sedimentation and altered stream flow dynamics at every scale (Petts 1984). Increased sedimentation from erosion in the watershed due to roads, timber harvest, and farming activities can impact species by impairing habitat, altering the composition of aquatic communities, or changing food sources used by organisms (Ney and Van Hassel 1983; Aldridge et al. 1987; Berkman and Rabeni 1987; Newcombe and Macdonald 1991; Waters 1995; Sutherland et al. 2002). Sediment contamination decreases species richness of benthic macroinvertebrates in some systems (de Lange et al. 2004), and siltation can decrease populations of some fish guilds (Rabeni and Smale 1995). Dam construction, channelization, and some forms of dredging have altered habitat availability in freshwater ecosystems by changing the flow regime and sediment loads, influencing stream morphology and substrates and therefore affecting fish communities (Trautman and Gartman 1974; Allen and Hardy 1980; Petts 1984; Poff et al. 1997). Dams can also create barriers to dispersal that interfere with movements of diadromous species to and from spawning grounds and can disrupt gene flow between populations (Pringle 2001; Wissmar and Craig 2004; Katano et al. 2006). Water released from dams and power plants at temperatures differing from an unregulated system can be sources of thermal pollution, causing changes in water temperatures that affect the fish and insect species living downstream, sometimes changing community composition (Petts 1984; Lessard and Hayes 2003).

Changes to fish and invertebrate habitats due to physical and chemical alteration of ecosystems can not only affect the relative abundance of native species, but can also contribute to invasions by exotic or introduced species (Holden and Stalnaker 1975; Miller et al. 1989; Gido and Brown 1999; Jones III 1999). The release of cold water from the hypolimnion of dams can lead to the extirpation of some native species (Edwards 1978) or change warm-water fisheries to cold-water communities, and in some cases, non-native species are deliberately introduced to these altered systems (Swink 1983). Extirpation of endemic species as a result of chemical and physical alteration to the environment or invasion by non-native species can affect community

composition, which may ultimately affect the functioning of freshwater ecosystems (Simon and Townsend 2003; Kelly et al. 2006).

All of these perturbations to the aquatic environment threaten freshwater ecosystems that have inherent importance as functioning biological systems with diverse aquatic flora and fauna. Minimizing anthropogenic threats to freshwater ecosystems and restoring impaired systems are necessary steps to preserve aquatic resources and to maintain and expand populations of threatened and endangered species (Allan and Flecker 1993; Lydeard and Mayden 1995; Saunders et al. 2002). Increasing our knowledge of freshwater ecosystems and the effects of local, riparian, watershed, and upstream threats on organisms and their ecological roles in those systems are important steps toward developing conservation plans for aquatic organisms of concern (Moss 2000; Saunders et al. 2002; Moilanen et al. 2008).

Freshwater mussels

Life history and ecological importance

Freshwater mussels can be useful indicators of freshwater ecosystem health for many reasons (Williams et al. 1993). Mussels can be useful for studies of pollutants affecting freshwater ecosystems because their tissues can store chemicals present in the aquatic environment (e.g., Gagne et al. 2002; McKinney et al. 2002; O'Rourke 2004; Brown et al. 2005). Mussels perform an important function as filterers in freshwater ecosystems, as they can consume algae, zooplankton, rotifers, and detritus (Vaughn et al. 2008). They are also a food source for a variety of terrestrial and aquatic species (Strayer et al. 1999; Vaughn and Hakenkamp 2001; Vaughn et al. 2008). Although few studies have quantitatively examined the contributions of native freshwater mussels to their ecosystems, research by Vaughn et al. (2004) indicates that mussels can have large impacts on the ecosystem by removing algae and excreting nutrients and organic material at rates correlated with mussel biomass. Because mussels can comprise the majority of benthic biomass in some freshwater ecosystems, living mussels and mussels that are consumed or that decompose can be substantial sources of nutrients and organic matter in an ecosystem (Negus 1966).

Diverse freshwater mussel communities are generally limited to streams, rivers, and lakes that have been minimally altered by anthropogenic activities (Vaughn and Taylor 1999). Many

freshwater mussel species have declined as freshwater ecosystems have become impaired. North America is home to one of the most diverse assemblages of freshwater mussels in the world, making the conservation status of mussel species in the U.S. a reason for concern (Bogan 1993; Williams et al. 1993). The U.S. and Canada contain 297 recognized native freshwater mussel species in the Margaritiferidae and Unionidae families (Turgeon et al. 1988), and the U.S. Endangered Species Act currently lists 72 mussel species as endangered or threatened (US FWS 2006). Nearly half of freshwater mussel species native to the U.S. can be found in the Appalachian region (Ortmann 1918), and many of these species are official species of concern by state or federal standards (US FWS2006).

Although mussels are a critical part of freshwater ecosystems, they have generally been studied much less than both terrestrial organisms and most other components of freshwater ecosystems (Lydeard et al. 2004). Because basic research on freshwater mussels is still in its early stages, specific knowledge of the functional roles, habitat requirements, and limitations to population growth for individual species is less understood than the ecology of mussels in general (Strayer et al. 1999; Vaughn and Hakenkamp 2001). Reproductive ecology of unionids has been explored in some detail, although life history information on individual species requires additional study (Haag and Staton 2003). Nearly all freshwater mussels in the family Unionidae require hosts, usually fish, to complete their reproductive cycle (Watters 1998). Glochidia, which are larval unionid mussels, are parasitic on the gills or fins of fish until they develop to the juvenile stage (Howard and Anson 1922). Specific fish and amphibian species capable of hosting glochidia have been identified for some mussel species (Watters 1998). However, suitable fish hosts vary in their ability to transform large numbers of juvenile mussels, and laboratory studies frequently used to evaluate potential hosts may not identify the best host for a species (Khym and Layzer 2000; Rogers and Dimock 2003; Watters et al. 2005). Since the absence of suitable host species could limit the development of juvenile freshwater mussels (Coker et al. 1921), the identification of host species for all mussel species is important for mussel conservation (Vaughn and Taylor 2000).

Freshwater mussels are long-lived compared to most freshwater invertebrates, with most mussel species capable of living several decades (Bogan 1993). Some evidence exists for reproductive senescence in older adults (Downing et al. 1993; Haag and Staton 2003), and some populations show no recent evidence of successful reproduction (Heinricher and Layzer 1999;

Bogan 1993; Strayer et al. 2004). An extinction debt may exist for mussels because species that may appear to be stable based on the numbers of individuals can be limited by habitat destruction, habitat fragmentation, and loss of glochidial host species to isolated or non-reproducing populations that may nonetheless be maintained for up to a century due to the longevity of individual mussels (Tilman et al. 1994; Heinricher and Layzer 1999; Strayer et al. 2004). The numbers or genetic diversity of these populations may be critical to the long-term viability of the species, but without successful reproduction, the extinction of the species will become apparent decades from now when existing individuals die (Tilman et al. 1994; Neves et al. 1997).

Distribution and occurrences

Knowledge of the locations of mussel populations is critical to the conservation of imperiled freshwater mussel species, and the National Native Mussel Conservation Committee has recognized the lack of information regarding mussel distribution as a barrier to understanding and conserving freshwater mussels (National Native Mussel Conservation Committee 1998). Current data on the distributions of freshwater mussel species are collected primarily through localized research studies and larger Natural Heritage Surveys that use a variety of sampling protocols, including timed searches and bridge surveys (Andersen 2002; Strayer and Smith 2003), or that rely on incidental observations of species (Rushton et al. 2004). Individual states cooperate with The Nature Conservancy to collect and synthesize data on many species, including mussels, for the Natural Heritage Survey (Groves et al. 1995). Most state natural heritage surveys record mussel occurrences, but some states only note occurrences of federal and state species of concern, and few states conduct regular scientific surveys for all mussels across the state. In the Appalachian region, regular surveys for some freshwater mussels are conducted by state agencies or in cooperation with state agencies in Virginia, North Carolina, Georgia, Kentucky, Tennessee, and Maryland (Table 1). State agencies in West Virginia, Ohio, Pennsylvania, and New Jersey also regularly monitor freshwater mussels, although the focus of surveys in these states has not normally been the collection of distribution data, and state agencies in New York conduct some monitoring activities (Table 1). Thus, there is wide variation in the types and quality of mussel data collected by states, organizations, and researchers.

Several researchers have concentrated their studies on different regions within the Appalachians. Watters has extensively researched the molluscan fauna of the Ohio drainage region (1995), and recent surveys have also been conducted by Hoggarth (1992) and Clarke and Clarke (1998). Strayer's work has focused on the Mid-Atlantic drainage region, including recent studies of mussels in the Hudson and Delaware regions (1993; 1996; 1999b; Strayer and Fetterman 1999). Harman (1970) has also published research based in the Mid-Atlantic. Ortmann's work included extensive studies of Pennsylvania, sampling both the Mid-Atlantic and Ohio drainage regions, and Ortmann also conducted research in the Tennessee Drainage (1913; 1924; 1925). Historic works of the Tennessee drainage region include Van der Schalie (1938; 1939) and Wilson (1914). There are many studies in the Tennessee drainage region, including those conducted by the Tennessee Valley Authority (TVA) and others related to TVA impoundments (Isom and Yokley 1968; Krenkel et al. 1979). S.A. Ahlstedt has conducted studies since the 1970s concentrated on the Tennessee River system, including research in the Elk River and the Clinch River and its tributaries (e.g., Ahlstedt 1983, 1991). Other researchers who have conducted studies in this region include Stansbery (1972), Isom (1973), Houp (1980), Neves (1982), Dennis (1984), and more recent students of malacology, primarily under Richard J. Neves at Virginia Polytechnic Institute and State University (e.g., Church 1997; Ostby 2005). A comprehensive bibliography of mollusk studies, including mussel research in the Appalachians, is available from the Freshwater Mollusk Conservation Society (2008). Existing knowledge of mussel distributions has been summarized in texts focused on the fauna of specific states or regions. Maryland, North Carolina, Ohio, Pennsylvania, and Tennessee have books on the distributions of mussels for the state (Starnes and Bogan 1988; Bogan and Proch 1995; Watters 1995; Parmalee and Bogan 1998; Bogan 2002). Other states, including West Virginia and New Jersey, have outlined intentions of developing statewide guides to mussels and their distributions in state Comprehensive Wildlife Conservation Plans (Table 1), but these guides are not yet available throughout the Appalachians.

Historic data on the distributions of mussels are necessary for the development of historic and potential ranges and extents, but these data are also limited in scope and accuracy. Museum collections exist for numerous species of mussels; however, the collection date and location are not available for many records, and the geographic distribution of collections is not likely to represent the geographic range of the species because historic data are largely derived from

naturalists whose research may have focused on particular rivers or regions rather than on sampling efforts designed to determine the extent of a species' range (e.g., Neel and Allen 1964). Furthermore, the taxonomy of freshwater mussels changes frequently (National Native Mussel Conservation Committee 1998) and species delineations are still under dispute (Hoeh and Gordon 1996), complicating efforts to incorporate historic data in modern analyses.

Habitat requirements

More information is necessary to determine specific habitat requirements for freshwater mussels (Brim-Box et al. 2002). However, general relationships between freshwater mussel occurrences and aspects of the environment have been identified (Coker et al. 1921; Matteson 1955). Brim-Box et al. (2002) outlined three major factors influencing mussel communities: the distribution and availability of host fish, basin characteristics, and microhabitat variables. Brim-Box et al. (2002) reviewed numerous studies that have attempted to specify habitat requirements by comparing these factors to freshwater mussel communities with variable success, and Hornbach (2001) provided a review of studies that compared variables explaining freshwater mussel distributions at various scales.

Several researchers have studied relationships between basin-scale variables and freshwater mussel communities. Vaughn (1997) found evidence that processes occurring at the drainage basin scale affected freshwater mussels and hypothesized that host-fish interactions helped determine freshwater mussel communities. Haag and Warren (1998) examined the influence of fish communities and microhabitat factors on mussel communities. They found that at a drainage basin scale, the relationship between mussels and host fishes were most important in structuring mussel communities and that the influences of host fishes overwhelmed the effects of microhabitat variables. They identified three groups of mussels that varied in their use of available habitat based on their mussel-fish host relationship: host-generalists (species with larvae that can develop on a variety of fishes), displaying host-specialists (species with larvae that require a specific host fish or fishes and rely on lures to attract hosts), and nondisplaying host-specialists (species with larvae that require a specific host fish or fishes but do not have mechanisms to attract the hosts). There was a positive correlation between the abundance of nondisplaying host-specialists and the abundance of their fish hosts, and whereas host-generalists and displaying host-specialists are able to inhabit both small streams and larger streams,

nondisplaying host-specialists were limited to large streams (Haag and Warren 1998). Additional evidence for a strong effect of fish species richness on the number of mussel species in an area was provided by Watters (1992) in the Ohio drainage region and Vaughn and Taylor (2000) in the Red River drainage, Kentucky.

Fish communities are not the only factors that affect freshwater mussel distributions. Strayer (1983) hypothesized that factors within the catchment are partially responsible for determining mussel distributions. DiMaio and Corkum (1995) described mussel communities specific to drainage basins with high hydrological variability and identified differences from communities occurring in hydrologically stable drainage basins. Their findings confirmed that hydrologic stability could predict mussel distributions, as suggested by Strayer (1983, 1993). Strayer (1983) found strong significant relationships for several mussel species to both surface geology and stream size in Michigan, concluding that hydrology is an important predictor of the distribution of many freshwater mussel species. However, these variables explained less variation in mussel distributions in the northern Atlantic Slope region (Strayer 1993). Morris and Corkum (1996) found differences in mussel communities between streams with forested riparian zones and streams with grassy riparian zones. They hypothesized that temperature and chemical differences in these streams contributed to the differences in mussel communities. At a macrohabitat scale (1-10 km), Strayer (1993) found that hydrological variability, physiographic province, calcium concentration, and stream gradient were weakly correlated with distributions of freshwater mussel species, but stream size was a strong predictor of mussel distributions.

On a microhabitat scale (1 m^2), patches of freshwater mussels (mussel beds) in streams and rivers may be linked to flow refugia, where flows change little during floods (Strayer and Ralley 1993; Strayer 1999b). Other stream bed characteristics such as streambed substrate may also affect freshwater mussel distributions and have been used extensively to describe mussel habitat (e.g., Strayer 1981; Salmon and Green 1983; Strayer and Ralley 1993). Mussel distributions may be determined to a lesser extent by river depth (Salmon and Green 1983) and by the number of aquatic plants in an area (Strayer and Ralley 1993).

Threats specific to freshwater mussels

Many factors impair freshwater ecosystem health, but an understanding of the mechanisms through which mussels are affected by specific threats is necessary to understand

the steps that must be taken to restore the freshwater mussel fauna of the U.S. Nearly half of freshwater mussel species in North America are now considered extinct, endangered, threatened, or species of concern, and reduced habitat quality has contributed to many of these declines (Bogan 1993). Bogan (1993) determined that causes for extinction and local extirpation of mussels and other freshwater bivalves included damming, habitat destruction from siltation, and pollution such as acid mine runoff, and a recent review of the literature reiterates the influence of these factors on mussel populations (Strayer et al. 2004). Anthropogenic effects such as harvesting, pollution, landscape alteration, and habitat destruction have been responsible for decreases in mussel diversity worldwide (Sheehan et al. 1989; Neves et al. 1997), but the southeastern region of the U.S. has experienced larger declines in populations and species richness than any other region over the last 150 years (Bogan 1993; Neves et al. 1997). Experts believe that agriculture, exotic species, and altered nutrient impacts from hydrological dams and agricultural runoff are currently the most severe threats to freshwater mussels, but point source pollution was largely responsible for past extirpations of mussel fauna in the eastern U.S. (Richter et al. 1997), and some researchers believe that harvesting continues to contribute to population declines (Anthony and Downing 2001).

Although mussel kills from point source pollution are a less severe problem now than in the past (Richter et al. 1997), both point and non-point source pollution may continue to contribute to the decline of some mussel populations by affecting recruitment and reducing viability. Environmental conditions that limit the attachment and transformation of glochidia on host fishes or reduce survival of juvenile mussels into adulthood can limit the ultimate survival of a population (Heinricher and Layzer 1999). Chemical pollutants are more harmful to juvenile freshwater mussels than adult freshwater mussels (Dimock and Wright 1993), so levels that are acceptable to adult mussels may be sufficient to prevent the persistence of the population. Juvenile freshwater mussels are vulnerable to acidity, a common threat in Appalachian streams, with a pH of 4.5 leading to mortality of 50% of juveniles after 96 hours (Dimock and Wright 1993). Freshwater mussels exposed to treated wastewater showed increased levels of heavy metals in their soft tissues and signs of stress from heavy metal exposure (Gagne et al. 2002). Populations of freshwater mussels downstream of effluent outlets also have unnaturally high proportions of mussels displaying primarily female characteristics, therefore reproduction may

be altered in populations exposed to estrogenic compounds and other pollutants (Blaise et al. 2003).

Alteration of the flow regime, hypolimnetic releases from dams, and changes in landcover within the riparian corridor from land-use practices such as logging, agriculture, and urban development can change the natural temperature of rivers and streams (Roberts and Barnhart 1999; Bigham 2002), which can affect availability of host fishes as well as mussel transformation success and juvenile survival (Matteson 1955; Dimock and Wright 1993; Bigham 2002). Heinricher and Layzer (1999) transferred adult freshwater mussels in the Cumberland River system from a population with no recruitment found in habitat below a dam to a lake with known reproducing populations. They found evidence of reproduction by translocated individuals within two years and suggested that because suitable hosts were available for larval transformation at both sites, water temperature was the most likely factor inhibiting recruitment (Heinricher and Layzer 1999).

Dams and other barriers affect freshwater mussel dispersal and survival both directly and indirectly (Watters 1996; Vaughn and Taylor 1999; Strayer 2006). Habitat destruction from dredging and impoundments have reduced mussel species richness in affected stream segments (Ortmann 1924; Allen and Hardy 1980; Williams et al. 1992). In addition to altering water temperatures, dams can change mussel communities by separating mussel species from compatible host fish (Layzer et al. 1993; Watters 1996). Smaller physical barriers such as culverts at stream crossings can have the same effect by interfering with movements of hosts (Warren and Pardew 1998). Streams with lethal levels of pollutants or temperatures can also act as barriers if water quality and aquatic communities are in better health upstream and downstream of the polluted area by inhibiting fish from traversing the polluted area (Wishart and Davies 2003; Maes et al. 2008), therefore limiting potential transport of glochidia across the chemical or thermal barrier.

Dredging, channelization, and impoundments are some factors affecting freshwater mussel habitat, but alteration of in-stream habitat can occur through many mechanisms (see Watters 1999 for review). Siltation is a known threat to freshwater mussel communities (Matteson 1955; Houp 1993); increased silt and sediment can occur upstream of impoundments or as a result of land use changes within the watershed (Houp 1993). Agriculture, forestry

operations, and other anthropogenic land uses can contribute to habitat alteration through increased erosion in the watershed or riparian corridor, and agricultural land use can also affect stream conditions by increasing nutrient inputs from fertilized croplands and livestock operations. Eutrophication of freshwater ecosystems from increased nutrient inputs impacts some freshwater mussels (Arter 1989; Strayer 1993), and anoxic conditions that may develop in the benthic zone negatively affect the survival of juvenile mussels (Dimock and Wright 1993; Sparks and Strayer 1998).

Exotic species are a continuing problem for freshwater mussels. Invasive exotic species such as zebra mussels (*Dreissena polymorpha*) and the Asian clam (*Corbicula fluminea*) continue to spread into large streams, rivers and lakes and contribute to declines in native mussel populations (Ricciardi et al. 1998; Strayer 1999a). Zebra mussels can directly kill native mussels by attaching to their shells in large numbers, and the high densities of zebra mussels in many rivers may also indirectly affect native freshwater mussels by limiting food availability (Strayer and Smith 1996; Ricciardi et al. 1998). Exotic fish may directly affect freshwater mussel populations through predation or may affect recruitment of native mussel species by displacing native fish hosts (Strayer 1999a). An invasive molluscivore, the round goby, has invaded the Great Lakes and is becoming established in freshwater systems, where it could have an effect on native freshwater mollusks (French and Jude 2001; Carman et al. 2006). Researchers speculate that introduced plants, animals, and diseases may indirectly affect freshwater mussel populations, but more research is necessary to explore the potential for indirect effects from exotic and alien species (Strayer 1999a; White et al. 2006).

Conservation efforts

A lack of basic knowledge limits the ability to identify priority areas for mussel conservation and restoration, as our knowledge of mussel distributions and the factors limiting their distributions is far behind that of terrestrial species (Lydeard et al. 2004; Strayer 2006). Some conservation plans focus on areas with high conservation value, which are areas containing high concentrations of endemic species threatened by anthropogenic activity, and often these areas have been selected based on the variety of terrestrial taxa without regard for aquatic diversity (e.g., Myers et al. 2000). Studies have shown that these plans fail to include most areas with diverse freshwater mollusk communities (Lawler et al. 2003; Lydeard et al. 2004). Recent

research has attempted to identify areas important to aquatic species (e.g., FitzHugh 2005), but these efforts have highlighted differences in existing knowledge between freshwater fish and freshwater invertebrates (Sowa et al. 2005). Conservation efforts directed toward freshwater mussels require consideration of upstream and watershed conditions, as do efforts directed at conserving any freshwater organism (Dudgeon et al. 2006). Long-term conservation plans are also necessary to ensure the recovery and maintenance of freshwater mussel populations because mussels are generally long-lived and some species have experienced declines over several decades (Bogan 1993; Strayer et al. 2004). Extensive field surveys to determine species distributions and the composition of communities and studies of ecosystem functions will be necessary before conservation biology can fully accommodate nonmarine mollusks, including freshwater mussels (Lydeard et al. 2004).

Outlines of the problems preventing the development of effective freshwater mussel conservation plans and guidelines toward conservation planning for mollusks and specifically for mussels have been published by Lydeard et al. (2004) and the National Native Mussel Conservation Committee (1998). The National Native Mussel Conservation Committee (1998) indicated that establishing historic and current ranges of freshwater mussel species is a necessary step toward implementation of conservation efforts. If conducted at an appropriate scale, this work can help identify range declines and allow researchers to relate distributions and declines to environmental factors that may affect freshwater mussels.

Comprehensive wildlife conservation strategies (CWCS) required for the State Wildlife Grants Program and Wildlife Conservation and Restoration Program and designed and implemented at the state level have addressed monitoring and maintaining populations of freshwater mussels. A review of CWCS for states in the eastern U.S. indicates that monitoring approaches are not consistent among states and that states differ in their assessments of the types of research necessary for conserving freshwater mussel diversity (Table 1). Indiana's CWCS anticipates implementing a rotating monitoring schedule that will resample streams for mussels every 10-12 years until national standards are enacted, while North Carolina has initiated cooperative programs with other agencies and universities to address diverse needs including fish host identification, ecology, systematics, and more detailed information on species distributions. Although a few states anticipate collecting quantitative information during surveys or repeating surveys at frequent intervals to address population trends, few state CWCS provided detailed

information regarding survey protocols and some states suggested the need to consult statisticians to develop sound study designs. Annual surveys are the most frequent surveys discussed in CWCS, although more frequent surveys would be necessary to evaluate seasonal variation in detection probabilities for species. Although the lack of information on freshwater mussels was addressed by most state CWCS, variable coordination across species ranges and differences among survey designs limit the potential effectiveness of CWCS research efforts in supporting the long-term conservation of freshwater mussels.

Predictive modeling

Species distribution models

Models of species distributions can be used to fill in gaps in knowledge regarding the geographic range and habitats of a species. Models can provide information that is too costly to collect, yet is valuable for evaluating the impacts of management decisions on species. Before the widespread use of complex statistical models coupled with Geographic Information Systems (GIS), Habitat Suitability Indices were developed to determine the probability of species occurrence in various habitats on a landscape. Calculation of a Habitat Suitability Index (HSI) relies on existing knowledge such as basic life-history information, the extent of known occurrences for the species, or the relationships between the species and components of its environment (US FWS 1981). Habitat Suitability Indices have been developed for many species using U.S. Fish and Wildlife Service guidelines due to the ease with which they can be developed (US FWS 1981; Roloff and Kernohan 1999). A summary of verified HSI models indicated that although models are simple to develop, details of implementation may affect model quality (Roloff and Kernohan 1999). Habitat Suitability Indices can have relatively high predictive ability when the species of interest has strong, clearly understood relationships to readily definable habitat characteristics, when the relationships used to define the model are accurate within the study area, and when the model incorporates habitats ranging from suitable to unsuitable (Roloff and Kernohan 1999). However, variability in input variables represented in a GIS, uncertainty in the application of variable ranks, difficulty in choosing the correct scale for analysis, and difficulty in obtaining validation data reduces the accuracy and applicability of HSIs (Bender et al. 1996; Roloff and Kernohan 1999; Johnson and Gillingham 2005).

With advances in technology, more techniques have become available to predict the probability that a species occurs in a specific area (Rushton et al. 2004; Phillips et al. 2006). Novel statistical techniques may be more complex than many earlier methods (Efron and Tibshirani 1991), allowing for various types and qualities of dependent and independent variables and more flexibility in the relationships between the independent and dependent variables (e.g., De'ath 2002). The increasing computational power of personal computers gives more researchers access to complex modeling techniques and increases the potential to develop models for many species (Efron and Tibshirani 1991; Boone and Krohn 2002; Rushton et al. 2004; Phillips et al. 2006). Modeling techniques rely on a variety of input data, including abundance, presence/absence, and presence-only species data and categorical, ordinal, or continuous independent variables. A review of available techniques is available in Guisan and Zimmermann (2000).

Niche analyses, such as environmental niche factor analysis, (Hirzel et al. 2002), ecological distances (e.g., Clark et al. 1993), and maximum entropy (Phillips et al. 2006), rely on presence data to identify areas that potentially contain suitable habitat for the species of interest. Niche analyses correctly predict occupied species distributions when species locations are distributed throughout the range of the species, when the species occupies all of the available suitable habitat, and when the environment is represented by the variables used in the analysis (Browning et al. 2005). Niche analyses usually overpredict species distributions when these conditions are not met, making these models unsuitable for species in decline (Engler et al. 2004). When a species is spreading into new areas, niche analyses may perform better than general linear models, such as stepwise linear regression (Hirzel et al. 2001).

Presence-only models are preferred by some researchers because of the uncertainty associated with absence data (Anderson et al. 2002; Hirzel et al. 2002). Assuming correct species identification and accurate recording of locations, researchers can have high confidence in the accuracy of species presences (Tyre et al. 2003; Brotons et al. 2004). Models relying on both presence and absence data can be affected by false absences, which occur when a recorded absence of a species is actually a failure to detect a true occurrence of the species (Tyre et al. 2003), or when true absences are unrelated to the environmental variables that describe habitat suitability in the model (Hirzel et al. 2001). The probability of detecting a species over repeated surveys can be used to evaluate confidence in absence data or can be directly incorporated into

some presence-absence models to improve predictions (Tyre et al. 2003). Methods to develop or improve absence data are the subject of many analyses (e.g., Engler et al. 2004; Lutolf et al. 2006) because models incorporating absence or pseudo-absence data with presence data have been demonstrated to be more accurate than presence-only models (e.g., Brotons et al. 2004), although this is likely true only when observed absences are correlated with reduced habitat suitability (Hirzel et al. 2001).

Many statistical techniques use both presence and absence data to create predictive models. Both parametric and nonparametric techniques are available to create predictive models of species distributions using presence and absence data. Most techniques predict species presence and absence with greater accuracy when there are approximately equal numbers of presence and absence locations, although class weights can be applied to adjust prediction error bias (Guisan and Zimmermann 2000). Principal components analysis (PCA) creates factors from correlated and independent environmental variables and uses these factors in models predicting probability of occurrence (McGarigal et al. 2000). However, PCA is not robust to violations of distributional assumptions and does not work well with independent data; therefore it is primarily used as a method of data exploration (McGarigal et al. 2000). Multiple logistic regression predicts the probability of occurrence and can be used with categorical or continuous independent variables and is robust to non-normality (Hosmer and Lemeshow 2000; McGarigal et al. 2000). Novel machine-learning methods such as artificial neural networks (ANN) and classification and regression trees (CART) have been used for exploratory data analysis, as these are nonparametric techniques that can be used to classify datasets (Olden and Jackson 2002). Classification and regression trees (CART) are nonparametric analyses that divide species occurrences into groups through successive dichotomous splits based on specific values for each environmental variable (Breiman et al. 1984). Regression trees (RT) are used to predict species richness or probability of species occurrence, while classification trees (CT) are used to predict presence or absence (Breiman et al. 1984). Classification and regression trees differ from most other modeling techniques because categorical, ordered, and continuous variables are compatible with this technique (Breiman et al. 1984). Classification and regression trees are valuable for modeling ecological datasets (Vayssieres et al. 2000; De'ath 2002) because they make no assumptions about the distributions of the variables and they have been shown to be appropriate for modeling spatially autocorrelated data (Cablk et al. 2002).

Modeling aquatic species occurrences in stream networks

Models of aquatic species occurrences have focused primarily on freshwater fishes. Sampling protocols for fishes have been established that have enabled researchers to collect data with known variance (Angermeier et al. 2002), and the accuracy of these datasets has proven suitable for model development. Bayley and Peterson (2001) demonstrated the use of Bayesian methods to incorporate estimates of detection into models, improving predictions of presence and absence of species at study sites, although models addressing probability of detection for aquatic species remain rare. Accurate datasets are especially important for riverine species because factors related to the distribution of these species introduce additional sources of error rarely encountered in models of terrestrial organisms. Dispersal by freshwater aquatic organisms generally occurs along the stream network, which reduces the ability of organisms to occupy all available useable habitats (Angermeier et al. 2002; Hitt and Angermeier 2008). If the species does not occupy all appropriate habitat in the study area, suitable habitat will be more difficult to define using environmental variables and the interpretation of the model results will be adversely affected (Hirzel et al. 2001). Choosing appropriate scales can reduce the effect of this problem (Rieman and McIntyre 1995), and incorporating variables from a range of scales can also improve the understanding of relationships between aquatic species and their environments (Dunham and Vinyard 1997; Angermeier et al. 2002).

Few studies of freshwater organisms have relied on presence-only modeling techniques to predict species occurrences. Recent uses of presence-only data for modeling aquatic species include uses of the genetic algorithm for rule-set prediction (GARP), which creates pseudo-absence data using random points (McNyset 2005; Dominguez-Dominguez et al. 2006). Dominguez-Dominguez et al. (2006) used topographic and climate data at a grid size of approximately 1km^2 to predict distributions of freshwater fishes in central Mexico, but found strong evidence that the models over-predicted the current and historical distributions of most species.

Many more studies of freshwater organisms have relied on presence and absence data. For example, Fransen et al. (2006) used logistic regression to determine the farthest upstream occurrence of fish in streams of western Washington based on variables derived from GIS layers, including variables of stream gradient upstream and downstream of each stream point, upstream

basin area, elevation, and mean annual precipitation. The models did not include information on barriers to fish movements, which the authors determined was responsible for most of the variation in the ability of their model to predict the occurrence of fish in some areas (Fransen et al. 2006).

In addition to traditional regression models, nonparametric techniques such as artificial neural networks (ANN), multivariate adaptive regression splines (MARS), and classification trees (CT) have been used in recent predictive models of aquatic species because they require fewer assumptions about the distributions of the data and the relationships between predictor and dependent variables, making them potentially well suited to ecological studies (De'ath and Fabricius 2000; Vayssieres et al. 2000). Mastrorillo et al. (1997) used ANN and discriminant factorial analysis to predict occurrences of three small stream fish in rivers in France. They incorporated distance to bank, instream cover, proportion of substrate types, depth, and water velocity into their models, which were created at the microhabitat scale (1 m^2) using data from electrofishing points in four regions of a river. Model performance on a random subset of the data withheld from analysis indicated that ANN results were an improvement over DFA results, although the models had similar prediction accuracy (Mastrorillo et al. 1997).

At a larger study scale, or grain, researchers have created models of fish occurrence and fish and crayfish communities based on variables at catchment, segment, and reach scales variables using ANN, CT, and discriminant function analysis (DFA) (Joy and De'ath 2004; Oakes et al. 2005). Joy and De'ath (2004) achieved 84% similarity between observed and modeled species assemblages with ANN. Artificial neural networks had lower error rates than CT and DFA, and models of generalist species had lower overall accuracy than models created for all other fish in the Great Plains (Oakes et al. 2005). Leathwick et al. (2005) used another machine-learning method, MARS, to predict distributions of diadromous fish at a grain of stream segments and found that distance to the sea, stream size, summer temperature, and catchment-scale variables that likely influence variation in stream flow were important predictors of occurrence for most species; riparian shade and measures of downstream slope were also incorporated into models for some species. The variables important to fish at this broader scale differ from analyses focused on microhabitats, and factors determining occurrences of diadromous fish likely differ from factors affecting aquatic species with more restricted ranges.

Researchers have also created models to predict presence or absence by watershed rather than stream reach. Watersheds and catchments have been identified as suitable scales for analyses of freshwater systems because the landscape over which water flows can affect freshwater ecosystems (Hynes 1975; Moss 2000). Studies on bull trout (*Salvelinus confluentus*) and brook trout (*Salvelinus fontinalis*) used CT analysis to predict watersheds where fish are likely to occur by modeling classifications of the status of trout populations (Rieman et al. 1997; Thieling 2006). These trout studies modeled historic and present ranges of the species by using environmental data to predict classifications of present, absent, or reduced trout populations at a subwatershed scale, then used model results to predict distributions and create maps of those distributions as a guide to management activities (Rieman et al. 1997; Thieling 2006). There are limitations to creating models at the scale of subwatersheds related to species data and analysis, as data are generally aggregated from several scales to the subwatershed scale and absences are difficult to assign to a subwatershed without extensive sampling or expert knowledge. The studies by Thieling (2006) and Rieman et al. (1997) should be interpreted with caution because the models were not tested on completely independent datasets, but rather on a subset of the data.

Modeling mussel species occurrences

There is disagreement in the literature regarding the ability to predict mussel occurrence from specific environmental variables, as similar studies have had varying levels of success predicting freshwater mussel occurrences (Hornbach 2001; Brim-Box et al. 2002). Zimmerman (2003) concluded that the lack of data and information available for freshwater mussels and their habitat requirements severely limits the ability to create HSI models and advocated the development of an ecological risk assessment to evaluate relative habitat suitability. Other researchers have correlated environmental variables at small and large scales, ranging from 1 m^2 areas of the stream bed to catchments and watersheds, with distributions of freshwater mussel species (e.g., Strayer 1983, 1993, 1999b; Strayer and Ralley 1993).

At a grain of 1 m^2, current speed, spatial variation in current speed, percentage of medium sand, depth, and number of macrophytes were significant but weak predictors of mussel occurrence in a stream (Strayer and Ralley 1993). Strayer and Ralley (1993) also suggested that the microhabitat scale may be inappropriate for predicting distributions of freshwater mussels, suggesting that factors within the catchment may influence occurrences of freshwater mussels.

At the larger scale of stream reaches, environmental characteristics, including riparian composition, water chemistry, stream shading, stream substrate, and stream width and depth have been used to model presence or absence of specific mussel species within a given reach (Poole and Downing 2004). Bambarger (2006) combined variables representing the microhabitat, sub-segment, and riparian buffer scales and found that a factor related to sediment at the microhabitat scale, geology at the sub-segment scale, and land use at both the sub-segment and riparian buffer spatial scales was highly correlated with freshwater mussel abundance and species richness. Studies of microhabitats or factors at the scale of stream segments over small extents define areas that are suitable for mussel populations and identify factors that may influence the persistence of mussel beds.

Although few studies have found strong significant correlations between environmental variables and freshwater mussel distributions at a local scale, analyses conducted for a large extent across the range of a species may reveal environmental variables related to the distribution of the species. As part of one of the most comprehensive efforts to model freshwater taxa at a regional extent, Sowa et al. (2005) georeferenced occurrence data and developed models of 67 freshwater mussel species within the state of Missouri. Their models were based on 1:100,000 stream segments within 8-digit hydrologic units (HUs) as the scale of analysis and were limited to variables representing stream size, gradient, temperature, flow, and size difference in their models. Although Sowa et al. (2005) developed what is currently the most comprehensive set of models for freshwater mussels by predicting distributions of every species known to occur Missouri, the model results are limited by several factors. They created separate models for different regions within the state, therefore the relationships between occurrences and environmental variables are not likely accurate in other regions. They also assumed that if a species had ever occurred in an 8-digit HU, the species had potential to occur in any suitable stream segment within the 8-digit HU, which may have led to errors of commission. Most importantly, the models represent the potential distributions of mussel species rather than the current distributions because the researchers did not address threats and focused only on environmental variables describing stream segment conditions. Spatially explicit research into mussel distributions and environmental and anthropogenic effects across a large extent may help to identify threats responsible for the broad decline in the ranges and number of species of freshwater mussels, as indirect evidence suggests that processes operating on a regional scale,

such as distributions and movements of host fishes within drainage basins may explain freshwater mussel distributions in North America (Vaughn 1997).

To the best of the author's knowledge, few studies have attempted to predict the current distributions of freshwater mussels using factors at a watershed extent over a large region, although researchers suggest that anthropogenic threats to mussels are important to examine at large spatial scales (Strayer et al. 2004). Recent mussel research has been conducted on the spatial relationships between freshwater mussels and the surrounding landscape, but these studies have generally been limited to relatively small areas, with state boundaries comprising the maximum extent of the study area (Andersen 2002; Arbuckle and Downing 2002; Poole and Downing 2004; Bambarger 2006). Anderson (2002) considered riparian, upstream, and local environmental conditions and found weak but significant correlations between sampled mussel communities and nearby urban or non-row crop agriculture land use. However, her research was restricted to survey points within five watersheds of one drainage subbasin, therefore her results may not represent general trends (Andersen 2002). Research conducted by Arbuckle and Downing (2002) across the state of Iowa considered variables on a watershed scale and identified variables related to sedimentation and pollution from agriculture as having a significant effect on mussel populations. Although this information increases the body of knowledge regarding current threats to freshwater mussels, findings from local studies may not apply in other regions: for example, although agriculture is a major land use in Iowa, it is not common throughout the ranges of all freshwater mussels. Furthermore, the response to perturbations by species that are present in Iowa may differ from responses of species in the eastern U.S., as differences in geology, land use, and other environmental and anthropogenic factors between regions could contribute to different patterns in freshwater mussel occurrences.

Because the Appalachian region of the United States has a history of forest clearing, acid mine drainage, and acid rain, among other anthropogenic changes to the environment (Herlihy et al. 1990, Clarkston 1993), variables representing forestry practices, acid deposition, and other environmental concerns common to the region may be better landscape-level predictors of the suitability of streams for freshwater mussels (Sheehan et al. 1989). Creating accurate predictive models of freshwater mussels in the Appalachians may elucidate factors affecting freshwater mussels in the region and will provide researchers and managers with information on distributions without incurring the costs associated with surveying all potential mussel habitats.

A successful model can also provide a coarse map of areas that can be surveyed to confirm the presence or absence of the mussel species and may allow managers to prioritize areas for restoration and conservation that are most likely to support any freshwater mussels.

Freshwater mussel data

Successful models of freshwater mussels require data suitable to model development. Sampling protocols for freshwater mussels vary considerably among agencies and individuals. Strayer and Smith (2003) have reviewed sampling techniques available for freshwater mussels and discussed in detail the techniques most suitable for the goals of a particular project. Strayer and Smith (2003) included information on determining occurrences of freshwater mussels in a sampling guide designed to be accessible to researchers and managers who are interested in determining the range of a species. They emphasized the importance of defining the spatial grain of a species dataset and the minimum density of mussels that must be present for mussels to be considered present at a site. For mussel surveys, Strayer and Smith (2003) suggest defining an absence as when a species is not found during a timed search within the period of time determined from the minimum encounter rate, although they also suggest that catch-per-unit-effort can be extremely variable. Studies of other fauna may provide alternatives or improvements on these methods to determine presence or absence of a species, including estimates of detection probabilities for species collected using particular sampling methods (Bayley and Peterson 2001).

A panel of experts on freshwater mussels were gathered at a meeting in May 2007 and assigned the task of suggesting mussels most suitable for the development of regional modeling efforts focused in the Appalachians (Figure 1, see Chapter 2 for model development)). The experts identified *Alasmidonta undulata, Alasmidonta varicosa, Alasmidonta viridis, Elliptio complanata, Elliptio dilatata, Epioblasma triquetra, Lasmigona subviridis, Plethobasus cyphyus, Pleurobema clava,* and *Villosa fabalis* as species that represented the variation of freshwater mussel occurrences, ranging from limited to widespread in their habitat preferences, and including species considered common, rare, or declining as of May 2007 (see Table 1 of Appendix B for a summary of these species). Data requests submitted to researchers, state agencies, organizations, and museums indicated that freshwater mussel data for the majority of these species are inadequate for regional modeling efforts (see Table 2 of Appendix B for

available mussel data). *Elliptio dilatata* (spike), which is a widespread species common to the Ohio drainage region, and *Elliptio complanata* (Eastern elliptio), a widespread species common to the Mid-Atlantic drainage, were best represented by recent data readily incorporated into a GIS.

Elliptio dilatata is an abundant species across its range and is actually increasing in some regions of the Appalachians (Houp 1993). This species occurs throughout the Mississippi River drainage; in the Appalachians its range stretches from the St. Lawrence River south to Kentucky and west from West Virginia and central Pennsylvania and New York (Parmalee and Bogan 1998). Strayer (1983) indicated that the species resides in streams and rivers with drainages ranging from 50-30,000 km^2. This species is found primarily in streams on outwash plans and moraines in southeastern Michigan; the species never occurred in streams that drained only lake plain geologies. Huehner (1987) found *E. dilatata* primarily in gravel substrate and never in sand. He determined through laboratory tests of substrate preference and field collections that current velocity is likely more important than substrate in determining local occurrences of *E. dilatata* (Huehner 1987). Houp (1993) suggested that *E. dilatata* is tolerant to sediment, as despite increasing sedimentation, populations nearly doubled in number from 1980-1991 in the North Fork Red River, Kentucky. Hosts established for *E. dilatata* include gizzard shad (*Dorosoma cepedianum*), flathead catfish (*Pylodictis olivaris*), white crappie (*Pomoxis annularis*), and yellow perch (*Perca flavescens*) (Fuller 1974). *E. dilatata* may be host-limited in some habitats (Lyons et al. 2007).

Elliptio complanata is a ubiquitous species. It is the most frequently occurring species in the Neuse River Basin of North Carolina (Andersen 2002) and in the northern Atlantic Slope, including the Susquehanna, Delaware, and Hudson River basins of Pennsylvania and New York (Strayer 1993), and its range stretches south into Virginia. Strayer (1993) described the distribution of *E. complanata* in the northern Atlantic Slope using stream size and physiography in a stepwise discriminant analysis model. Anderson (2002) observed a decline in *E. complanata* within the first 50 m downstream of bridges, indicating that the species may be sensitive to anthropogenic changes to the environment. Hosts for *E. complanata* include bluegill (*Lepomis macrochirus*), pumpkinseed (*Lepomis gibbosus*), green sunfish (*Lepomis cyanellus*), orangespotted sunfish (*Lepomis humilis*), largemouth bass (*Micropterus salmoides*), white crappie (*Pomoxis annularis*), yellow perch (*Perca flavescens*), and banded killfish (*Fundulus*

diaphanus) (Matteson 1948; Watters 1994; Watters et al. 2005). Amyot and Downing (1991) found that a population of *E. complanata* in a lake in Canada undergoes a seasonal vertical migration, with more than 60% of individuals, particularly those with smaller shells, migrating into the sediments by late fall. They cautioned that sampling lake or streambed surfaces could therefore drastically underestimate the number of individuals occurring in an area, particularly when sampling is conducted during the late summer, fall, or winter; they further hypothesized that vertical migration is responsible for lower estimates of juvenile mussels in samples collected by hand than in samples collected with dredges and grabs (Amyot and Downing 1991).

Summary

Freshwater mussels pose a challenge to modeling efforts due to their dependence on more mobile species for recruitment and dispersal and their occurrence within stream networks. However, the large number of freshwater mussel species in the U.S. prohibits inventorying every species throughout its range. The development of successful models of freshwater mussels and a detailed outline of the data required to develop models for mussel species would therefore give managers and scientists the ability to maximize the information available from limited mussel survey efforts. Because of the value of models of freshwater mussel distributions and the factors affecting mussel species, studies of the feasibility of models at various scales are an important component of freshwater conservation efforts.

References

Ahlstedt, S.A., 1983. The molluscan fauna of the Elk River in Tennessee and Alabama. American Malacological Bulletin 1, 43-50.

Ahlstedt, S.A., 1991. Cumberlandian mollusk conservation program: mussel surveys in six Tennessee Valley streams. Walkerana 5, 123-160.

Aldridge, D.W., Payne, B.S., Miller, A.C., 1987. The effects of intermittent exposure to suspended solids and turbulence on three species of freshwater mussels. Environmental Pollution 45, 17-28.

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

Allan, J.D., Flecker, A.S., 1993. Biodiversity conservation in running waters. Bioscience 43, 32-43.

Allen, K., Hardy, J., 1980. Impacts of navigational dredging on fish and wildlife: a literature review. U.S. Fish and Wildlife Service, Biological Services Program, Washington, DC.

Amyot, J.-P., Downing, J.A., 1991. Endo- and epibenthic distribution of the of the unionid mollusc *Elliptio complanata*. Journal of the North American Benthological Society 10, 280-285.

Andersen, E.F., 2002. Effects of land use and land cover on freshwater mussel populations in the upper Neuse River basin, North Carolina: A GIS approach. M.S. thesis, North Carolina State University, Raleigh, NC.

Anderson, R.P., Gomez-Laverde, M., Peterson, A.T., 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. Global Ecology and Biogeography 11, 131-141.

Angermeier, P.L., Krueger, K.L., Dolloff, C.A., 2002. Discontinuity in stream-fish distributions: implications for assessing and predicting species occurrence. In: Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B. (Eds.), Predicting Species Occurrences: Issues of Accuracy and Scale, Island Press, Washington, DC.

Anthony, J.L., Downing, J.A., 2001. Exploitation trajectory of a declining fauna: a century of freshwater mussel fisheries in North America. Canadian Journal of Fisheries and Aquatic Sciences 58, 2071.

Arbuckle, K.E., Downing, J.A., 2002. Freshwater mussel abundance and species richness: GIS relationships with watershed land use and geology. Canadian Journal of Fisheries and Aquatic Sciences 59, 310-316.

Arter, H.E., 1989. Effect of eutrophication on species composition and growth of freshwater mussels (Mollusca: Unionidae) in Lake Hallwil (Aargau, Switzerland). Aquatic Sciences 51, 87-99.

Averitt, P., 1975. Coal Resources of the United States, January 1, 1974. U.S. Geological Survey Bulletin 1412, U.S. Government Printing Office, Washington, DC.

Bambarger, R., 2006. Freshwater mussel communities of the Florida Parishes, Louisiana: the importance of spatial scale. M.S. thesis, Louisiana State University and Agricultural and Mechanical College, Baton Rouge, LA.

Bayley, P.B., Peterson, J.T., 2001. An approach to estimate probability of presences and richness of fish species. Transactions of the American Fisheries Society 130, 620-633.

Bender, L.C., Roloff, G.J., Haufler, J.B., 1996. Evaluating confidence intervals for habitat suitability models. Wildlife Society Bulletin 24, 347.

Berkman, H.E., Rabeni, C.F., 1987. Effect of siltation on stream fish communities. Environmental Biology of Fishes 18, 285-294.

Bigham, S.E., 2002. Host specificity of freshwater mussels - a critical factor in conservation. M.S. thesis, Southwest Missouri State University, Springfield, MO.

Blaise, C., Gagne, F., Salazar, M., Salazar, S., Trottier, S., Hansen, P.-D., 2003. Experimentally-induced feminisation of freshwater mussels after long-term exposure to a municipal effluent. Fresenius Environmental Bulletin 12, 865-870.

Bogan, A.E., 1993. Freshwater bivalve extinctions (Mollusca: Unionoida): a search for causes. American Zoologist 33, 599-609.

Bogan, A.E., 2002. Workbook and key to the freshwater bivalves of North Carolina. North Carolina Museum of Natural Sciences, Raleigh, NC.

Bogan, A.E., Proch, T., 1995. Manual of the Freshwater Bivalves of Maryland. Privately printed.

Boone, R.B., Krohn, W.B., 2002. Modeling tools and accuracy assessment. In: Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B. (Eds.), Predicting Species Occurrences: Issues of Accuracy and Scale, Island Press, Washington, DC.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and regression trees. Chapman & Hall, Boca Raton, FL.

Brim-Box, J., Dorazio, R.M., Liddell, W.D., 2002. Relationships between streambed substrate characteristics and freshwater mussels (Bivalvia: Unionidae) in coastal plain streams. Journal of the North American Benthological Society 21, 253-260.

Brotons, L., Thuiller, W., Araujo, M.B., Hirzel, A.H., 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27, 437-448.

Brown, M.E., Kowalewski, M., Neves, R.J., Cherry, D.S., Schreiber, M.E., 2005. Freshwater mussel shells as environmental chronicles: geochemical and taphonomic signatures of mercury-related extirpations in the North Fork Holston River, Virginia. Environmental Science & Technology 39, 1455.

Browning, D.M., Beaupré, S.J., Duncan, L., 2005. Using partitioned Mahalanobis $D^2(K)$ to formulate a GIS-based model of timber rattlesnake hibernacula. Journal of Wildlife Management 69, 33-44.

Cablk, M., White, D., Kiester, A.R., 2002. Assessment of spatial autocorrelation in empirical models in ecology. In: Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B. (Eds.), Predicting Species Occurrences: Issues of Accuracy and Scale, Island Press, Washington, DC.

Carman, S.M., Janssen, J., Jude, D.J., Berg, M.B., 2006. Diel interactions between prey behaviour and feeding in an invasive fish, the round goby, in a North American river. Freshwater Biology 51, 742-755.

Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8, 559-568.

Chambers, P.A., Meissner, R., Wrona, F.J., Rupp, H., Guhr, H., Seeger, J., Culp, J.M., Brua, R.B., 2006. Changes in nutrient loading in an agricultural watershed and its effects on water quality and stream biota. Hydrobiologia 556, 399-415.

Church, G.W., 1997. Macrohabitat factors affecting distribution patterns of freshwater mussels in the Clinch River (Virginia, Tennessee). Ph.D. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Clark, J.D., Dunn, J.E., Smith, K.G., 1993. A multivariate model of female black bear habitat use for a geographic information system. The Journal of Wildlife Management 57, 519-526.

Clarke, A.H., Clarke, J.J., 1998. Mussels and the mussel beds of the lower Ohio River (ORM 438.0 to 981.0). Malacology Data Net 4, 1-122.

Clarkston, R.B., 1993. Destruction of the upland forest by lumbering and fire. In: Stephenson, S.L. (Ed.), Upland Forests of West Virginia, McClain Printing Co., Parsons, WV.

Coker, R.E., Shira, A.F., Clark, H.W., Howard, A.D., 1921. Natural history and propagation of fresh-water mussels. Bulletin of the United States Bureau of Fisheries 37: 75-182.

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

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

de Lange, H.J., de Jonge, J., den Besten, P.J., Oosterbaan, J., Peeters, E.T.H.M., 2004. Sediment pollution and predation affect structure and production of benthic macroinvertebrate communities in the Rhine–Meuse delta, The Netherlands. Journal of the North American Benthological Society 23, 557-579.

Dennis, S.D., 1984. Distributional analysis of the freshwater mussel fauna of the Tennessee River System, with special reference to possible limiting effects of siltation (assimilation, radioisotopes, community structure). Ph.D. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Di Maio, J., Corkum, L., 1995. Relationship between the spatial distribution of freshwater mussels (Bivalvia: Unionidae) and the hydrological variability of rivers. Canadian Journal of Zoology 73, 663-671.

Dimock, R.V., Jr., Wright, A.H., 1993. Sensitivity of juvenile freshwater mussels to hypoxic, thermal and acid stress. Journal of the Elisha Mitchell Scientific Society 109, 183-192.

Dominguez-Dominguez, O., Martinez-Meyer, E., Zambrano, L., De Leon, G.P.-P., 2006. Using ecological-niche modeling as a conservation tool for freshwater species: live-bearing fishes in central Mexico. Conservation Biology 20, 1730-1739.

Downing, J.A., Rochon, Y., Perusse, M., Harvey, H., 1993. Spatial aggregation, body size, and reproductive success in the freshwater mussel *Elliptio complanata*. Journal of the North American Benthological Society 12, 148-156.

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

Dunham, J.B., Vinyard, G.L., 1997. Incorporating stream level variability into analyses of site level fish habitat relationships: some cautionary examples. Transactions of the American Fisheries Society 126, 323-329.

Dynesius, M., Nilsson, C., 1994. Fragmentation and flow regulation of river systems in the northern third of the world. Science 266, 753-762.

Edwards, R.J., 1978. The effect of hypolimnion reservoir releases on fish distribution and species diversity. Transactions of the American Fisheries Society 107, 71-77.

Efron, B., Tibshirani, R., 1991. Statistical data analysis in the computer age. Science 253, 390-395.

Engler, R., Guisan, A., Rechsteiner, L., 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of Applied Ecology 41, 263-274.

Fleeger, J.W., Carman, K.R., Nisbet, R.M., 2003. Indirect effects of contaminants in aquatic ecosystems. The Science of the Total Environment 317, 207-233.

Flessas, C., Couillard, Y., Pinel-Alloul, B., St-Cyr, L., Campbell, P.G.C., 2000. Metal concentrations in two freshwater gastropods (Mollusca) in the St. Lawrence River and relationships with environmental contamination. Canadian Journal of Fisheries and Aquatic Sciences 57, 126-137.

Fransen, B.R., Duke, S.D., McWethy, L.G., Walter, J.K., Bilby, R.E., 2006. A logistic regression model for predicting the upstream extent of fish occurrence based on geographical information systems data. North American Journal of Fisheries Management 26, 960-975.

French, J.R.P., Jude, D.J., 2001. Diets and diet overlap of nonindigenous gobies and small benthic native fishes co-inhabiting the St. Clair River, Michigan. Journal of Great Lakes Research 27, 300-311.

Freshwater Mollusk Conservation Society. Mollusk bibliography database. 2008. Unpublished. Available from: <http://ellipse.inhs.uiuc.edu:591/mollusk/biblio.html>. Last revised 9 November 2007.

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

Fuller, S.L.H., 1974. Clams and mussels (Mollusca: Bivalvia). In: Hart, C.W., Fuller, S.L.H. (Eds.), Pollution Ecology of Freshwater Invertebrates. Academic Press, New York, NY.

Gagne, F., Blaise, C., Aoyama, I., Luo, R., Gagnon, C., Couillard, Y., Campbell, P., Salazar, M., 2002. Biomarker study of a municipal effluent dispersion plume in two species of freshwater mussels. Environmental Toxicology 17, 149-159.

Gido, K.B., Brown, J.H., 1999. Invasion of North American drainages by alien fish species. Freshwater Biology 42, 387-399.

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

Groves, C.R., Klein, M.L., Breden, T.F., 1995. Natural heritage programs: public-private partnerships for biodiversity conservation. Wildlife Society Bulletin 23, 784-790.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135, 147.

Haag, W.R., Staton, J.L., 2003. Variation in fecundity and other reproductive traits in freshwater mussels. Freshwater Biology 48, 2118-2130.

Haag, W.R., Warren, M.L., 1998. Role of ecological factors and reproductive strategies in structuring freshwater mussel communities. Canadian Journal of Fisheries and Aquatic Sciences 55, 297-306.

Haines, T.A., 1981. Acidic precipitation and its consequences for aquatic ecosystems: a review. Transactions of the American Fisheries Society 110, 669-707.

Hakenkamp, C.C., Palmer, M.A., 1999. Introduced bivalves in freshwater ecosystems: the impact of Corbicula on organic matter dynamics in a sandy stream. Oecologia 119, 445-451.

Hanazato, T., 2001. Pesticide effects on freshwater zooplankton: an ecological perspective. Environmental Pollution 112, 1-10.

Harman, W.N., 1970. New distribution records and ecological notes on central New York Unionacea. American Midland Naturalist 84, 46-58.

Heinricher, J.R., Layzer, J.B., 1999. Reproduction by individuals of a nonreproducing population of *Megalonaias nervosa* (Mollusca: Unionidae) following translocation. American Midland Naturalist 141, 140-148.

Henderson, H.D., 2007. The effects of contaminants on sperm quality and intersex condition of smallmouth bass (*Micropterus dolomieu*) in the Potomac River. M.S. thesis, West Virginia University, Morgantown, WV.

Herlihy, A.T., Kaufman, P.R., Mitch, M.E., 1990. Regional estimates of acid mine drainage impact on streams in the Mid-Atlantic and Southeastern United States. Water, Air, and Soil Pollution 50, 91-107.

Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N., 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83, 2027-2036.

Hirzel, A.H., Helfer, V., Metral, F., 2001. Assessing habitat-suitability models with a virtual species. Ecological Modelling 145, 111-121.

Hitt, N.P., Angermeier, P.L., 2008. Evidence for fish dispersal from spatial analysis of stream network topology. Journal of the North American Benthological Society 27, 304-320.

Hoeh, W.R., Gordon, M.E., 1996. Criteria for the determination of taxonomic boundaries in freshwater unionoids (Bivalvia: Unoinoida): comments on Stiven and Alderman (1992). Malacologia 38, 223-227.

Hoggarth, M.A., 1992. The Unionidae and Corbiculidae of the Little Miami River system in southwestern Ohio. Walkerana 6, 247-293.

Holden, P.B., Stalnaker, C.B., 1975. Distribution and abundance of mainstream fishes in the middle and upper Colorado River basins, 1967-1973. Transactions of the American Fisheries Society 104, 217-231.

Hornbach, D.J., 2001. Macrohabitat factors influencing the distribution of naiads in the St. Croix River, Minnesota and Wisconsin, USA. In: Bauer, G., Wachtler, K. (Eds.), Ecology and evolution of the freshwater mussels Unionoida. Springer-Verlag, Berlin.

Hosmer, D., Lemeshow, S., 2000. Applied Logistic Regression, 2nd edition. John Wiley and Sons, New York, NY.

Houp, R.E., 1980. A survey of the mussels of the Red River (wild river segment) in eastern Kentucky. Transactions of the Kentucky Academy of Science 41, 55-56.

Houp, R.E., 1993. Observations on long-term effects of sedimentation on freshwater mussels (Mollusca: Unionidae) in the North Fork of Red River, Kentucky. Transactions of the Kentucky Academy of Science 54, 93-97.

Howard, A.D., Anson, B.J., 1922. Phases in the parasitism of the Unionidae. The Journal of Parasitology 9, 68-82.

Huehner, M.K., 1987. Field and laboratory determination of substrate preferences of unionid mussels. Ohio Journal of Science 87, 29-32.

Hynes, H.B., 1975. The stream and its valley. Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie 19, 1-15.

Isom, B.G., Yokley Jr, P., Gooch, C.H., 1973. Mussels of the Elk River Basin in Alabama and Tennessee-1965-1967. American Midland Naturalist 89, 437-442.

Isom, B.G., Yokley, P., 1968. The mussel fauna of the Duck River in Tennessee, 1965. American Midland Naturalist 80, 34-42.

Johnson, C.J., Gillingham, M.P., 2005. An evaluation of mapped species distribution models used for conservation planning. Environmental Conservation 32, 117.

Johnson, L., Gage, S., 1997. Landscape approaches to the analysis of aquatic ecosystems. Freshwater Biology 37, 113-132.

Jones III, E.B.D., 1999. Effects of riparian forest removal on fish assemblages in southern Appalachian streams. Conservation Biology 13, 1454-1465.

Joy, M.K., De'ath, R.G., 2004. Predictive modelling and spatial mapping of freshwater fish and decapod assemblages using GIS and neural networks. Freshwater Biology 49, 1036–1052.

Katano, O., Nakamura, T., Abe, S., Yamamoto, S., Baba, Y., 2006. Comparison of fish communities between above- and below-dam sections of small streams; barrier effect to diadromous fishes. Journal of Fish Biology 68, 767-782.

Kelly, D.W., Bailey, R.J., MacNeil, C., Dick, J.T.A., McDonald, R.A., 2006. Invasion by the amphipod *Gammarus pulex* alters community composition of native freshwater macroinvertebrates. Diversity and Distributions 12, 525-534.

Khym, J.R., Layzer, J.B., 2000. Host fish suitability for glochidia of *Ligumia recta*. American Midland Naturalist 143, 178-184.

Krenkel, P.A., Lee, G.F., Jones, R.A., 1979. Effects of TVA impoundments on downstream water quality and biota. In: Ward, J.V., Stanford, J.A. (Eds.), The Ecology of Regulated Streams. Plenum Press, New York, NY.

Lawler, J.J., White, D., Sifneos, J.C., Master, L.L., 2003. Rare species and the use of indicator groups for conservation planning. Conservation Biology 17, 875-882.

Layzer, J.B., Gordon, M.E., Anderson, R.M., 1993. Mussels: the forgotten fauna of regulated rivers. A case study of the Caney Fork River. Regulated Rivers 8, 63-71.

Leathwick, J.R., Rowe, D., Richardson, J., Elith, J., Hastie, T., 2005. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. Freshwater Biology 50, 2034-2052.

Lessard, J.L., Hayes, D.B., 2003. Effects of elevated water temperature on fish and macroinvertebrate communities below small dams. River Research and Applications 19, 721-732.

Lutolf, M., Kienast, F., Guisan, A., 2006. The ghost of past species occurrence: improving species distribution models for presence-only data. Journal of Applied Ecology 43, 802-815.

Lydeard, C., Clark, S.A., Perez, K.E., Cowie, R.H., Ponder, W.F., Bogan, A.E., Bouchet, P., Gargominy, O., Cummings, K.S., Frest, T.J., Herbert, D.G., Hershler, R., Roth, B., Seddon, M., Strong, E.E., Thompson, F.G., 2004. The global decline of nonmarine mollusks. Bioscience 54, 321-330.

Lydeard, C., Mayden, R.L., 1995. A diverse and endangered aquatic ecosystem of the southeast United States. Conservation Biology 9, 800-805.

Lyons, M.S., Krebs, R.A., Holt, J.P., Rundo, L.J., Zawiski, W., 2007. Assessing causes of change in the freshwater mussels (Bivalvia: Unionidae) in the Black River, Ohio. American Midland Naturalist 158, 1-15.

Maes, J., Stevens, M., Breine, J., 2008. Poor water quality constrains the distribution and movements of twaite shad *Alosa fallax fallax* (Lacépède, 1803) in the watershed of river Scheldt. Hydrobiologia 602, 129-143.

Malmqvist, B., Rundle, S., 2002. Threats to the running water ecosystems of the world. Environmental Conservation 29, 134-153.

Mastrorillo, S., Lek, S., Dauba, F., Belaud, A., 1997. The use of artificial neural networks to predict the presence of small-bodied fish in a river. Freshwater Biology 38, 237-246.

Matteson, M.R., 1948. Life history of *Elliptio complanatus* (Dillwyn, 1817). American Midland Naturalist 40, 690-723.

Matteson, M.R., 1955. Studies on the natural history of the Unionidae. American Midland Naturalist 53, 126-145.

McClurg, S.E., Petty, J.T., Mazik, P.M., Clayton, J.L., 2007. Stream ecosystem response to limestone treatment in acid impacted watersheds of the Allegheny Plateau. Ecological Applications 17, 1087-1104.

McGarigal, K., Cushman, S., Stafford, S., 2000. Multivariate Statistics for Wildlife and Ecology Research. Springer Science and Business Media, Inc., New York, NY.

McKinney, R.A., Lake, J.L., Charpentier, M.A., Ryba, S., 2002. Using mussel isotope ratios to assess anthropogenic nitrogen inputs to freshwater ecosystems. Environmental Monitoring and Assessment 74, 167-192.

McNyset, K.M., 2005. Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas. Ecology of Freshwater Fish 14, 243-255.

Miller, R.R., Williams, J.D., Williams, J.E., 1989. Extinctions of North American fishes during the past century. Fisheries 14, 22-38.

Miltner, R.J., Rankin, E.T., 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshwater Biology 40, 145-158.

Moilanen, A., Leathwick, J., Elith, J., 2008. A method for spatial freshwater conservation prioritization. Freshwater Biology 53, 577-592.

Morris, T.J., Corkum, L.D., 1996. Assemblage structure of freshwater mussels (Bivalvia: Unionidae) in river with grassy and forested riparian zones. Journal of the North American Benthological Society 15, 576-586.

Moss, B., 2000. Biodiversity in fresh waters: an issue of species preservation or system functioning? Environmental Conservation 27, 1-4.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., de Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853-858.

Naiman, R.J., Turner, M.G., 2000. A future perspective on North America's freshwater ecosystems. Ecological Applications 10, 958-970.

National Native Mussel Conservation Committee, 1998. National strategy for the conservation of native freshwater mussels. Journal of Shellfish Research 17, 1419-1428.

Neel, J.K., Allen, W.R., 1964. The mussel fauna of the upper Cumberlandian Basin before its impoundment. Malacologia 1, 427-459.

Negus, C.L., 1966. A quantitative study of growth and production of unionid mussels in the River Thames at Reading. The Journal of Animal Ecology 35, 513-532.

Neves, R.J., Bogan, A.E., Williams, J.D., Ahlstedt, S.A., Hartfield, P.W., 1997. Status of aquatic mollusks in the southeastern United States: a downward spiral of diversity. In: Benz, G.W., Collins, D.E. (Eds.), Aquatic Fauna in Peril: The Southeastern Perspective. Southeast Aquatic Research Institute, Lenz Design and Communications, Decatur, GA.

Neves, R.J., Zale, A.V., 1982. Freshwater mussels (Unionidae) of Big Moccasin Creek, southwestern Virginia. Nautilus 96, 52-54.

Newcombe, C.P., Macdonald, D.D., 1991. Effects of suspended sediments on aquatic ecosystems. North American Journal of Fisheries Management 11, 72-82.

Ney, J.J., Van Hassel, J.H., 1983. Sources of variability in accumulation of heavy metals by fishes in a roadside stream. Archives of environmental contamination and toxicology 12, 701-706.

Nilsson, C., Reidy, C.A., Dynesius, M., Revenga, C., 2005. Fragmentation and flow regulation of the world's large river systems. Science 308, 405-408.

O'Rourke, S.M., 2004. Quantitative biomonitoring of polychlorinated biphenyls in the Detroit River using the freshwater mussel *Elliptio complanata*. M.S. thesis, University of Windsor, Windsor, Ontario.

Oakes, R.M., Gido, K.B., Falke, J.A., Olden, J.D., Brock, B.L., 2005. Modelling of stream fishes in the Great Plains, USA. Ecology of Freshwater Fish 14, 361-374.

Olden, J.D., Jackson, D.A., 2002. A comparison of statistical approaches for modelling fish species distributions. Freshwater Biology 47, 1976-1995.

Ortmann, A.E., 1913. The Alleghenian Divide, and its influence upon the freshwater fauna. Proceedings of the American Philosophical Society 52, 287-390.

Ortmann, A.E., 1918. The nayades (freshwater mussels) of the upper Tennessee drainage, with notes on synonymy and distribution. Proceedings of the American Philosophical Society 57, 521-626.

Ortmann, A.E., 1924. Mussel shoals. Science 60, 565-566.

Ortmann, A.E., 1925. The naiad-fauna of the Tennessee River system below Walden Gorge. American Midland Naturalist 9, 321-372.

Ostby, B.J.K., 2005. Characterization of suitable habitats for freshwater mussels in the Clinch River, Virginia and Tennessee. M.S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Parmalee, P.W., Bogan, A.E., 1998. The Freshwater Mussels of Tennessee. University of Tennessee Press, Knoxville, TN.

Petts, G., 1984. Impounded Rivers: Perspectives for Ecological Management. John Wiley and Sons, Chichester, England.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231-259.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The natural flow regime. Bioscience 47, 769-784.

Poole, K.E., Downing, J.A., 2004. Relationship of declining mussel biodiversity to stream-reach and watershed characteristics in an agricultural landscape. Journal of the North American Benthological Society 23, 114-125.

Pringle, C.M., 2001. Hydrologic connectivity and the management of biological reserves: a global perspective. Ecological Applications 11, 981-998.

Rabeni, C.F., Smale, M.A., 1995. Effects of siltation on stream fishes and the potential mitigating role of the buffering riparian zone. Hydrobiologia 303, 211-219.

Ricciardi, A., Neves, R.J., Rasmussen, J.B., 1998. Impending extinctions of North American freshwater mussels (Unionoida) following the zebra mussel (*Dreissena polymorpha*) invasion. The Journal of Animal Ecology 67, 613-619.

Ricciardi, A., Rasmussen, J.B., 1999. Extinction rates of North American freshwater fauna. Conservation Biology 13, 1220-1222.

Richter, B.D., Braun, D.P., Mendelson, M.A., Master, L.L., 1997. Threats to imperiled freshwater fauna. Conservation Biology 11, 1081-1093.

Rieman, B.E., Lee, D.C., Thurow, R.F., 1997. Distribution, status, and likely future trends of bull trout within the Columbia River and Klamath River basins. North American Journal of Fisheries Management 17, 1111-1125.

Rieman, B.E., McIntyre, J.D., 1995. Occurrence of bull trout in naturally fragmented habitat patches of varied size. Transactions of the American Fisheries Society 124, 285-296.

Roberts, A.D., Barnhart, M.C., 1999. Effects of temperature, pH, and CO_2 on transformation of the glochidia of *Anodonta suborbiculata* on fish hosts and in vitro. Journal of the North American Benthological Society 18, 477-487.

Rogers, C.L., Dimock, R.V., Jr., 2003. Acquired resistance of bluegill sunfish *Lepomis macrochirus* to glochidia larvae of the freshwater mussel *Utterbackia imbecillis* (Bivalvia: Unionidae) after multiple infections. Journal of parasitology 89, 51-56.

Roloff, G.J., Kernohan, B.J., 1999. Evaluating reliability of habitat suitability index models. Wildlife Society Bulletin 27, 973-985.

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

Rushton, S.P., Ormerod, S.J., Kerby, G., 2004. New paradigms for modelling species distributions? Journal of Applied Ecology 41, 193-200.

Salmon, A., Green, R.H., 1983. Environmental determinants of unionid clam distribution in the Middle Thames River, Ontario. Canadian Journal of Zoology 61, 832-838.

Saunders, D.L., Meeuwig, J.J., Vincent, A.C.J., 2002. Freshwater protected areas: strategies for conservation. Conservation Biology 16, 30-41.

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

Sheehan, R.J., Neves, R.J., Kitchel, H.E., 1989. Fate of freshwater mussels transplanted to formerly polluted reaches of the Clinch and North Fork Holston Rivers, Virginia, USA. Journal of Freshwater Ecology 5, 139-150.

Sietman, B.E., Whitney, S.D., Kelner, D.E., Blodgett, K.D., Dunn, H.L., 2001. Post-extirpation recovery of the freshwater mussel (Bivalvia: Unionidae) fauna in the upper Illinois River. Journal of Freshwater Ecology 16, 273-281.

Simon, K.S., Townsend, C.R., 2003. Impacts of freshwater invaders at different levels of ecological organisation, with emphasis on salmonids and ecosystem consequences. Freshwater Biology 48, 982-994.

Sowa, S.P., Diamond, D.D., Abbitt, R., Annis, G., Gordon, T., Morey, M.E., Sorensen, G.R., True, D., 2005. A gap analysis for riverine ecosystems of Missouri. Final report submitted to the USGS National Gap Analysis Program.

Sparks, B.L., Strayer, D.L., 1998. Effects of low dissolved oxygen on juvenile *Elliptio complanata* (Bivalvia:Unionidae). Journal of the North American Benthological Society 17, 129-134.

Stansbery, D.H., 1972. The mollusk fauna of the North Fork Holston River at Saltville, Virginia. Bulletin of the American Malacological Union, Inc. 1971, 45-46.

Starnes, L.B., Bogan, A.E., 1988. The mussels (Mollusca: Bivalvia: Unionidae) of Tennessee. The American Malacological Bulletin 6, 19-37.

Strayer, D.L., 1981. Notes in the microhabitats of unionid mussels in some Michigan streams. American Midland Naturalist 106, 411.

Strayer, D.L., 1983. The effects of surface geology and stream size on freshwater mussel (Bivalvia, Unionidae) distribution in southeastern Michigan, USA. Freshwater Biology 13, 253-264.

Strayer, D.L., 1993. Macrohabitats of freshwater mussels (Bivalvia: Unionacea) in streams of the northern Atlantic Slope. Journal of the North American Benthological Society 12, 236-246.

Strayer, D.L., 1999a. Effects of alien species on freshwater mollusks in North America. Journal of the North American Benthological Society 18, 74-98.

Strayer, D.L., 1999b. Use of flow refuges by unionid mussels in rivers. Journal of the North American Benthological Society 18, 468-476.

Strayer, D.L., 2006. Challenges for freshwater invertebrate conservation. Journal of the North American Benthological Society 25, 271-287.

Strayer, D.L., Caraco, N.F., Cole, J.J., Findlay, S., Pace, M.L., 1999. Transformation of freshwater ecosystems by bivalves: a case study of zebra mussels in the Hudson River. Bioscience 49, 19-27.

Strayer, D.L., Downing, J.A., Haag, W.R., King, T.L., Layzer, J.B., Newton, T.J., Nichols, S.J., 2004. Changing perspectives on pearly mussels, North America's most imperiled animals. Bioscience 54, 429-439.

Strayer, D.L., Fetterman, A.R., 1999. Changes in the distribution of freshwater mussels (Unionidae) in the Upper Susquehanna River Basin. American Midland Naturalist 142, 328-339.

Strayer, D.L., Ralley, J., 1993. Microhabitat use by an assemblage of stream-dwelling Unionaceans (Bivalvia), including two rare species of *Alasmidonta*. Journal of the North American Benthological Society 12, 247-258.

Strayer, D.L., Smith, D.R., 2003. A Guide to Sampling Freshwater Mussel Populations. American Fisheries Society, Bethesda, Maryland.

Strayer, D.L., Smith, L.C., 1996. Relationships between zebra mussels (*Dreissena polymorpha*) and unionid clams during the early stages of the zebra mussel invasion of the Hudson River. Freshwater Biology 36, 771-779.

Sutherland, A.B., Meyer, J.L., Gardiner, E.P., 2002. Effects of land cover on sediment regime and fish assemblage structure in four southern Appalachian streams. Freshwater Biology 47, 1791-1805.

Swink, W.D., 1983. Survey of stocking policies for tailwater trout fisheries in the southern United States. Progressive Fish-Culturalist 45, 67-71.

Thieling, T.M., 2006. Assessment and predictive model for brook trout (*Salvelinus fontinalis*) population status in the eastern United States. M.S. thesis, James Madison University, Harrisonburg, VA.

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

Trautman, M.B., Gartman, D.K., 1974. Re-evaluation of the effects of man-made modification on Gordon Creek between 1887 and 1973 and especially as regards fish fauna. Ohio Journal of Science 74, 162-173.

Turgeon, D.D., Bogan, A.E., Coan, E.V., Emerson, W.K., Lyons, W.G., Pratt, W.L., Roper, E.F.E., Scheltema, A., Thompson, F.G., Williams, J.D., 1988. Common and Scientific Names of Aquatic Invertebrates from the United States and Canada: Mollusks. American Fisheries Society, Bethesda, MD.

Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K., Possingham, H.P., 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecological Applications 13, 1790-1801.

US FWS (U.S. Fish and Wildlife Service), 1981. Standards for the development of habitat suitability index models. Ecological Services Manual 103. Division of Ecological Services, U.S. Fish and Wildlife Service, Dept. of the Interior, Washington, D.C.

US FWS (U.S. Fish and Wildlife Service), 2006. Species Information: Threatened and Endangered Animals and Plants. Available from: <http://www.fws.gov/endangered/wildlife.html>

van der Schalie, H., 1938. The naiades (fresh-water mussels) of the Cahaba River in northern Alabama. Occasional Papers of the Museum of Zoology, University of Michigan 392, 1-29.

van der Schalie, H., 1939. Additional notes on the naiades (freshwater mussels) of the lower Tennessee River. American Midland Naturalist 22, 452-457.

Vaughn, C.C., 1997. Regional patterns of mussel species distributions in North American rivers. Ecography 20, 107-115.

Vaughn, C.C., Gido, K.B., Spooner, D.E., 2004. Ecosystem processes performed by unionid mussels in stream mesocosms: species roles and effects of abundance. Hydrobiologia 527, 35.

Vaughn, C.C., Hakenkamp, C.C., 2001. The functional role of burrowing bivalves in freshwater ecosystems. Freshwater Biology 46, 1431-1446.

Vaughn, C.C., Nichols, S.J., Spooner, D.E., 2008. Community and foodweb ecology of freshwater mussels. Journal of the North American Benthological Society 27, 409-423.

Vaughn, C.C., Taylor, C.M., 1999. Impoundments and the decline of freshwater mussels: a case study of an extinction gradient. Conservation Biology 13, 912-920.

Vaughn, C.C., Taylor, C.M., 2000. Macroecology of a host-parasite relationship. Ecography 23, 11.

Vayssieres, M.P., Plant, R.E., Allen-Diaz, B.H., 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. Journal of Vegetation Science 11, 679.

Warren, M.L., Pardew, M.G., 1998. Road crossings as barriers to small-stream fish movement. Transactions of the American Fisheries Society 127, 637-644.

Waters, T.F., 1995. Sediment in Streams: Sources, Biological Effects, and Control. American Fisheries Society, Bethesda, MD.

Watters, G.T., 1992. Unionids, fishes, and the species-area curve. Journal of Biogeography 19, 481-490.

Watters, G.T., 1994. An annotated bibliography of the reproduction and propagation of the Unionoidea (primarily of North America). Ohio Biological Survey Miscellaneous Contributions 1, 1-158.

Watters, G.T., 1995. A Guide to the Freshwater Mussels of Ohio. Division of Wildlife, Ohio Department of Natural Resources.

Watters, G.T., 1996. Small dams as barriers to freshwater mussels (Bivalvia, Unionoida) and their hosts. Biological Conservation 75, 79-85.

Watters, G.T., 1998. Metamorphosis of freshwater mussel glochidia (Bivalvia: Unionidae) on amphibians and exotic fishes. American Midland Naturalist 139, 49-57.

Watters, G.T., 1999. Freshwater mollusks and water quality: a review of the effects of hydrologic and instream habitat alterations, In Proceedings of the First Freshwater Mollusk Conservation Society Symposium, pp. 261-274.

Watters, G.T., Menker, T., Thomas, S., Kuehnl, K., 2005. Host identifications or confirmations. Ellipsaria 7, 11-12.

White, E.M., Wilson, J.C., Clarke, A.R., 2006. Biotic indirect effects: a neglected concept in invasion biology. Diversity and Distributions 12, 443-455.

Williams, J.D., Fuller, S.L.H., Grace, R., 1992. Effects of impoundments on freshwater mussels (Mollusca: Bivalvia: Unionidae) in the main channel of the Black Warrior and Tombigbee Rivers in Western Alabama. Bulletin of the Alabama Museum of Natural History 13, 1-10.

Williams, J.D., Warren, M.L., Cummings, K.S., Harris, J.L., Neves, R.J., 1993. Conservation status of freshwater mussels of the United States and Canada. Fisheries 18, 6-22.

Wilson, C.B., Clark, H.W., 1914. The mussels of the Cumberland River and its tributaries. Report of the Commissioner of Fisheries for 1912. Document No. 781. U.S. Bureau of Fisheries, Washington, DC.

Wishart, M.J., Davies, B.R., 2003. Beyond catchment considerations in the conservation of lotic biodiversity. Aquatic Conservation: Marine Freshwater Ecosystems 13, 429-437.

Wissmar, R.C., Craig, S.D., 2004. Factors affecting habitat selection by a small spawning charr population, bull trout, *Salvelinus confluentus*: implications for recovery of an endangered species. Fisheries Management and Ecology 11, 23-31.

Zimmerman, L.L., 2003. Propagation of juvenile freshwater mussels (Bivalvia: Unionidae) and assessment of habitat suitability for restoration of mussels in the Clinch River, Virginia. M.S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Tables

Table 1. Links to comprehensive wildlife conservation plans for states within the Appalachian region of the eastern U.S.

State	Web Link (May 2008)
Alabama	http://www.wildlifeactionplans.org/pdfs/action_plans/al_action_plan.pdf
Georgia	http://www1.gadnr.org/cwcs/Documents/strategy.html
Indiana	http://www.djcase.com/incws/manuscript/FINAL_CWS_MANUSCRIPT.pdf
Kentucky	http://fw.ky.gov/kfwis/stwg/
Maryland	http://www.wildlifeactionplans.org/pdfs/action_plans/md_action_plan.pdf
Mississippi	http://www.mdwfp.com/Level2/cwcs/Final.asp
New Jersey	http://www.state.nj.us/dep/fgw/ensp/wap/pdf/wap_draft.pdf
New York	http://www.wildlifeactionplans.org/pdfs/action_plans/ny_action_plan.pdf
North Carolina	http://www.ncwildlife.org/pg07_wildlifespeciescon/WAP_complete.pdf
Ohio	http://www.fws.gov/midwest/FederalAid/state_plans.html
Pennsylvania	http://www.wildlifeactionplans.org/pdfs/action_plans/pa_action_plan.pdf
Tennessee	http://www.state.tn.us/twra/cwcs/tncwcs2005.pdf
Virginia	http://bewildvirginia.org/wildlifeplan/plan.asp
West Virginia	http://www.wvdnr.gov/Wildlife/PDFFiles/wvwcap.pdf
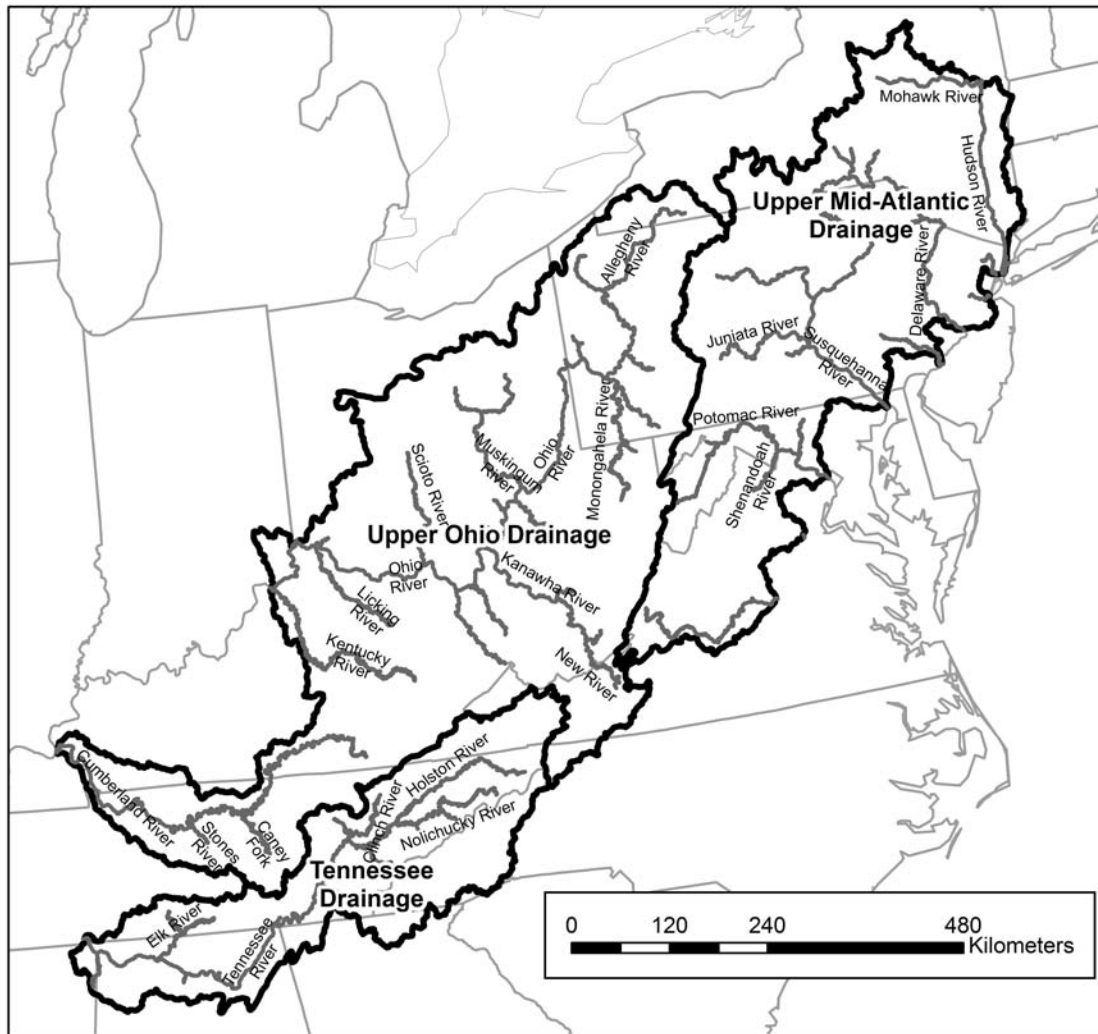
Figures

Figure 1. Focus area for the development of predictive models for freshwater mussels in the Appalachians.

Chapter 2: Predictive Models for the Freshwater Mussel Species (Mollusca: Unionidae) *Elliptio complanata* and *Elliptio dilatata* in the Appalachians

Abstract

Eastern North America contains the greatest diversity of freshwater mussels in the world. Additional information on threats and on habitat requirements and distributions of freshwater mussels is necessary to preserve diverse freshwater mussel communities, as many species are in decline. Models of freshwater mussels can predict species distributions by determining environmental and anthropogenic factors within the watershed or riparian area or factors upstream that may determine occurrences. Subwatershed-based models developed for *Elliptio complanata* and *E. dilatata* in the Mid-Atlantic and Ohio drainage regions using existing survey data performed well on training datasets, but did not accurately predict independent species occurrences Additional studies are necessary to evaluate the quality of existing data, the utility of subwatershed-based models, and the feasibility of modeling freshwater mussel distributions across large extents.

Introduction

Research into the conservation and management of freshwater mussels is in its infancy compared to research on charismatic or commercially important species, such as many fish and mammals. Past designs for conservation efforts directed at preserving ecosystems have focused on terrestrial species (e.g., Kerr 1997; Myers et al. 2000). Although newer approaches have focused on aquatic species (e.g., Sowa et al. 2005), these plans either identify gaps in available data for freshwater mussels or fail to ensure the conservation of freshwater invertebrates (Lydeard et al. 2004; Strayer 2006).

There are many barriers to specifically addressing the conservation of freshwater mussels. Unlike terrestrial ecosystems, streams and rivers are networked to each other and connected to the terrestrial landscape (Frissell et al. 1986), therefore freshwater ecosystems cannot be isolated from activities and land-use changes in the watershed. Protecting freshwater species from extinction is therefore more complicated than terrestrial conservation efforts (Moss 2000). Only in recent years has there been national recognition of the need to conserve freshwater mussels, which historically were most diverse in the Appalachian region of the United States (National Native Mussel Conservation Committee 1998). Researchers have made recent advances in the science of propagating and managing freshwater mussels at a local scale, and recent studies have begun to evaluate anthropogenic impacts on mussels at a regional scale

(Gatenby et al. 2006), but additional research is necessary to address gaps in knowledge of freshwater mussel distributions, ecology, life-history characteristics, and threats (Vaughn et al. 2008).

Conserving freshwater ecosystems requires aquatic ecologists to consider terrestrial impacts within the watershed and riparian corridor and upstream water quality as well as conditions immediately within a particular stream reach (Gregory et al. 1991; Allan & Johnson 1997; Wang et al. 2003; Allan 2004; Schomberg et al. 2005). Maintaining management-defined riparian zones can prevent or reduce impairment of freshwater ecosystems (Kiffney et al. 2003; Northington & Hershey 2006), and researchers have found that riparian zone conditions can affect freshwater mussel communities at a local scale (Morris & Corkum 1996). Overland water flow and connectivity from headwater streams to large rivers also contribute to the conditions affecting freshwater ecosystems (Vannote et al. 1980; Hitt & Angermeier 2008). Increasing anthropogenic influences upstream can negatively affect downstream aquatic communities by decreasing water quality or creating barriers to dispersal (Watters 1996; Freeman et al. 2007). Although techniques to incorporate upstream and riparian variables into analyses have been developed (FitzHugh 2005), few studies have incorporated measures of riparian area or cumulative anthropogenic impacts into analyses specific to freshwater mussels (see Strayer 2008 for a review), particularly over a large extent (e.g., Arbuckle & Downing 2002; Poole & Downing 2004). Previous studies have found that factors related to regional anthropogenic activities can partially explain the local distribution of freshwater mussel species in Iowa (Arbuckle & Downing 2002). This suggests that modeling the regional distributions of declining mussel species may highlight regional or widespread factors that limit species occurrence or determine habitat suitability in the Appalachians, providing useful information for managers and scientists in areas with important molluscan fauna.

Occurrences of freshwater mussels are more difficult to model than plants or vertebrate fauna for several reasons. Modeling freshwater aquatic species requires consideration of habitat networks, which is a factor rarely considered in models of terrestrial species. Freshwater networks differ from terrestrial habitat corridors because the linear and unidirectional nature of stream flow limits species movements and because upstream conditions may influence the system downstream (Hitt & Angermeier 2008). Despite these difficulties, several models of freshwater fishes have been developed that adequately predict species occurrences (e.g.,

Mastrorillo et al. 1997; Olden & Jackson 2002; Leathwick et al. 2005; Dominguez-Dominguez et al. 2006). However, freshwater mussels are additionally difficult to model because although sampling techniques for freshwater fish are firmly established with tested accuracy and error for many species (Angermeier et al. 2002), fewer researchers have extensively tested sampling regimes necessary to determine occurrences of mussels (Strayer & Smith 2003). Knowledge of the accuracy of research surveys can be incorporated into models to improve predictions for freshwater fishes (Bayley & Peterson 2001), but the paucity of mussel survey data and the need for research on the accuracy and error of existing mussel surveys may limit the development of predictive models for freshwater mussels.

Models of freshwater mussels would ideally identify factors that limit the occurrences of rare species, as these species are most in need of conservation planning (Lydeard et al. 2004). Models of rare terrestrial species with specific habitat requirements have been created that define the extent of potentially suitable habitat (Hernandez et al. 2006). However, rare freshwater species may violate modeling assumptions about the dependent variable because they are confined to a stream network (Hirzel et al. 2001). Rare species are not likely distributed across all suitable habitats within a study region, and rare mussels are further limited from moving into suitable habitat by the linear movements of host fishes within a stream network. When suitable habitat remains unoccupied by a species, the ability of the model to define factors determining suitable habitat is negatively affected (Hirzel et al. 2001). Additionally, the probability of detecting a rare species is normally lower than the probability of detecting common species, so false absences are potentially more likely for rare species. False absences may adversely affect model results by identifying areas where species actually occur as unsuitable (Hirzel et al. 2001). Data on common species are more likely to conform to model expectations and are easier to obtain than data for federally listed endangered and threatened species. Because models of common mussel species should be affected by fewer potential limitations than rare freshwater mussels, common species were chosen for model development.

By focusing on common, widespread species of freshwater mussels in the Appalachians, a region with a diversity of mussels, this research can highlight methods and potential outcomes associated with the development of models in an area where model development may be important for conserving species of concern. The objectives of this research are to: 1) create and test models of freshwater mussel species distributions using watershed, riparian, and upstream

attributes, 2) predict the presence or absence of freshwater mussel populations in areas where data are lacking, and 3) identify the types of environmental variables and quality of species data suitable for the development of predictive models. A successful model of common, widespread freshwater mussel species in the Appalachian study area will highlight the feasibility of this modeling technique for freshwater mussel research, allow managers to prioritize areas for restoration and conservation that are most likely to support any freshwater mussels, and provide a template for modeling additional mussel species. A successful model will also provide a coarse map of areas that can be surveyed to confirm the presence or absence of modeled mussel species.

Methods

Study area

The study area for this project is a watershed-based region that encompasses headwater and mid-order streams within the Appalachian Mountains (Figure 1). The study area includes the upper portions of the Mid-Atlantic and Ohio drainages and stretches from northern Alabama and Georgia in the south to New York in the north. Because hydrologic boundaries limit mussel distributions and drainage regions approximate mussel faunal provinces containing different communities of freshwater mussels, models were developed separately for each drainage region (Parmalee & Bogan 1998). The study area was designed to facilitate comparisons of mussel distributions in the Appalachian region with conservation plans and studies for other species of concern (Link & Sauer 2002; Nuttle et al. 2003; Sauer et al. 2003; Buehler et al. 2006).

The boundaries used to define relevant hydrological units for the majority of the study area were 11-, 12-, and 14-digit hydrological unit (HU) watershed data boundaries (hereafter subwatersheds) prepared for each state by the U.S. Geological Survey and the Natural Resources Conservation Service (Seaber et al. 1987). Boundaries at the desired scale are unavailable for New York, so the most detailed watersheds available in those areas, which are 10-digit HUs, were substituted. The boundaries of these watersheds were edited to more approximately match the structure of 12-digit subwatersheds, which enclose rather than border streams and rivers.

Drainage regions within the study area differ in size and level of hydrological modification. The area of each drainage region included in the study area varies from 172,000

km^2 to 266,000 km^2. The Mid-Atlantic region consists of 1,714 subwatersheds averaging 100.11 km^2 in size and drains to the Atlantic. Large rivers in the Mid-Atlantic include the Hudson, Susquehanna, Delaware, Potomac, and Shenandoah. The Hudson, Susquehanna, and Potomac Rivers are affected by damming over more than 25% of their length (Dynesius & Nilsson 1994). The Ohio region contains 3,168 subwatersheds averaging 84 km^2 in size and includes the Ohio, Allegheny, Cumberland, and New Rivers. Rivers within the Ohio drainage region vary considerably in conservation status. The status of large rivers in the Ohio varies from the New River, which is a National River managed by the National Park Service, to the Ohio River, which is used extensively for navigation.

Mussel data

Mussel data were obtained for two widespread species, *Elliptio dilatata* (spike) and *E. complanata* (eastern elliptio) (Turgeon et al. 1988), which are common throughout their ranges (Houp 1993; Strayer 1993; Andersen 2002).   For each species, available data were compiled into either a presence-absence training dataset for model development or an independent presence-only test dataset for model evaluation. The data contained in the presence-only datasets were derived from separate data sources to ensure independence from the training dataset.

Regional mussel experts and state agencies provided freshwater mussel survey data (presences and absences) for sites across the study area (Table 1). Survey data are differentiated from occurrence (presence-only) data because areas where researchers searched for mussels, but did not find them, are recorded. Recent survey data were used to develop training datasets containing presence and absence locations for model development in each drainage region. Presence by subwatershed for the training dataset was derived from occurrences recorded in recent surveys, which are defined by river lengths, points, and subwatersheds. Surveys with river miles were defined in a geographic information system (GIS) as point or line locations and were converted to the Albers Conical Equal Area projection used for environmental data. Presence data from survey lines that overlapped more than one subwatershed were omitted to avoid falsely assigning a presence to an unsurveyed area. Absences were assigned to subwatersheds along recent survey routes when all surveys covering the subwatershed failed to locate the species. The average nearest neighbor distance was calculated to test for clustering of survey data for each region (Table 3).

Recent mussel data were defined as any recorded surveys or occurrences since 1985 (Table 1, 2). Choosing a cutoff date creates a defined distinction between historic occurrences and recent occurrences. Recent surveys would have a high likelihood of detecting individuals that have recolonized areas restored since the Clean Water Act of 1972 (Strayer et al. 2004) and earlier surveys were difficult to obtain. A more recent cutoff date was not chosen because few mussel surveys have been conducted that span several subwatersheds.

The training dataset for the Mid-Atlantic drainage region contained 143 subwatersheds in which *E. complanata* had been detected by surveys and 114 subwatersheds where *E. complanata* had not been located by available surveys (Table 2). These data were derived from surveys covering the state of Pennsylvania and most of Maryland and were used to develop models for *E. complanata* in the Mid-Atlantic. The training dataset for model development in the Ohio drainage region included 47 presences and 126 absences of *E. dilatata* by subwatershed derived from surveys in Pennsylvania and along the Ohio River (Table 2).

Museums and state natural heritage surveys provided occurrence (presence-only) data for the species of interest (Table 1). The presence-only datasets were developed separately from the survey (presence/absence) datasets to create an independent test dataset for each species. The test dataset for each species contained mussel occurrences not included in model creation and was developed to enable an independent evaluation of model predictions (Hernandez et al. 2006). Presence in a subwatershed was assigned based on species occurrences since 1985 recorded in museum and state natural heritage survey records (Table 1). Because most state natural heritage surveys limit records to occurrences, these data were appropriate for the test (presence-only) dataset. The Mid-Atlantic test dataset contained 40 presences of *E. complanata* by subwatershed based on presence-only data (Table 2). The independent test dataset for the Ohio drainage region contained 205 presence-only subwatersheds for *E. dilatata* derived from state agency and museum records (Table 2). The determination of species presence in the test dataset did not always correspond to the training (presence/absence) dataset; however, occurrences within a subwatershed recorded from presence-only data were not used to change an absence assigned to the training dataset based on survey data so that independence of the datasets could be preserved.

Although the training and test datasets were independently developed, geographic overlap occurred between the datasets for both species. Some subwatersheds were therefore represented in both datasets (Table 2). The test (presence-only) dataset for *E. dilatata* in the Mid-Atlantic contained three subwatersheds that overlapped the training dataset. Surveys located *E. dilatata* two of these subwatersheds, but based on recent surveys, the species was classified as absent for the third subwatershed in the training dataset. The presence-only dataset for the Ohio overlapped 14 subwatersheds where surveys recorded *E. complanata* and 17 subwatersheds where *E. complanata* was not recorded by surveyors (Table 2).

Independent variables

Anthropogenic and environmental variables representing potential threats or limiting factors to freshwater mussels were selected. Only variables at a resolution and extent that could be summarized for each subwatershed within a drainage region were considered for inclusion in model development. These variables were derived from existing data available in a GIS (Table 4). Metrics were calculated by subwatershed and by riparian corridor within the subwatershed following Hudy (2005). The riparian corridor was defined as a 100 m buffer from streams, rivers, and lakes represented by the flowline, stream area, and waterbody layers in the improved National Hydrography Dataset (NHDPlus) (Horizon Systems 2006). This riparian zone width is commonly used in mussel research and other aquatic studies (Richards et al. 1996; Andersen 2002; McKinney et al. 2002; Thieling 2006) and is larger than the 30 m resolution of the remotely sensed environmental variables. Although some legal stream buffer zones are smaller than 100 m (Willson & Dorcas 2003), a larger stream corridor is more suitable for the grain of the analyses in this study.

A custom automated GIS routine in ArcView 3.2 (ESRI 2000) was used to define all subwatersheds upstream of each subwatershed. Additional variables were derived using the cumulative values of environmental and anthropogenic variables by including data upstream of each subwatershed for variables that could logically be calculated cumulatively. ArcGIS 9.1 and 9.2 (ESRI 2002) was used for all other development and summarization of variables (Figure 2). More detailed information on methods and sources for environmental variables is available in Appendix A, and a complete listing of variables, summary statistics, and data sources are available in Appendix C.

Model development

Models were developed using statistical techniques commonly used for presence and absence data. All of the available mussel survey data aggregated by subwatershed for each drainage region were used as the training datasets for model development. The training datasets for *E. complanata* (Mid-Atlantic) and *E. dilatata* (Ohio) were entered into classification tree (CT) models (mvpart package, De'ath 2007) and multiple logistic regression models predicting the categories of presence or absence for each species (Breiman et al. 1984; Hosmer & Lemeshow 2000) in the R programming language (R Core Development Team 2007).

Multiple logistic regression models and CT models differ in several aspects that can affect model outcomes. Multiple logistic regression predicts the probability of occurrence and can be used with categorical or continuous independent variables (Hosmer & Lemeshow 2000; McGarigal et al. 2000). Logistic regression requires the choice of a cutoff to predict the categories of species presence or absence from probability of occurrence, which can affect the outcome of the model (Fielding & Bell 1997). Classification tree analysis is a nonparametric technique that has been used for exploratory data analysis because it is robust to large numbers of independent variables relative to the number of samples and it can use interactions between variables to classify datasets (Olden & Jackson 2002). This technique summarizes results in an easily interpretable dichotomous key predicting class (such as presence or absence) based on values for the input variables (Breiman et al. 1984). Classification trees differ from most other modeling techniques because categorical, ordered, and continuous variables are compatible with this technique (Breiman et al. 1984). This method is valuable for modeling ecological datasets (Vayssieres et al. 2000; De'ath 2002) because it makes no assumptions about the distributions of the variables and it has been shown to be appropriate for modeling spatially autocorrelated data (Cablk et al. 2002).

All variables available for a drainage region were entered into the CT model for the species in that drainage region. The 154 variables in the Mid-Atlantic and 149 variables in the Ohio entered into the CT models included variables related to stream size (maximum and minimum Strahler and link order within each subwatershed, Strahler and maximum link order difference in the subwatershed, cumulative drainage area, maximum, minimum, and maximum difference in mean annual discharge within a subwatershed), land cover (proportion of

agriculture, forest, and other land cover types in the subwatershed, riparian area, and cumulative watershed), elevation (minimum, maximum and average elevation in the subwatershed, riparian area, and cumulative watershed, minimum, maximum, and average slope in the subwatershed, riparian area, cumulative watershed, and along streams), and other environmental and anthropogenic variables related to water quality represented in a GIS (Tables 4, 5). Because multiple logistic regression can be affected by correlation among independent variables, subsets of the variables entered into CT models for each region were developed for logistic regression using the caret package (Kuhn et al. 2008) in R (2007) by omitting highly correlated variables until the largest correlation in the dataset for each region was 0.5 (Table 5). A value of 0.5 was used as the cutoff for classifying subwatersheds as present or absent from logistic regression output, which ranges from zero to one (Hosmer & Lemeshow 2000). The CT model automatically outputs classifications of presence or absence (Breiman et al. 1984) that can be evaluated against test data. The presence-only dataset was withheld from model development as an independent test dataset for model evaluation (after Hernandez et al. 2006).

Resulting models for each species were applied to subwatersheds across the modeled drainage region to create predictions of presence and absence for each species. These predictions were compared against the presence-only test data obtained from museums and natural heritage surveys to evaluate the predictive ability of the models (Hernandez et al. 2006). Accuracy, errors of omission (where a species is known to occur but predicted absent, ranging from 0-100%) and commission (where a species is considered absent but predicted present, ranging from 0-100%), and kappa scores (ranging from negative one to one, with larger values indicating greater agreement between predicted and observed values) were used to evaluate the performance of logistic regression and CT models for both species on the training data. Errors of omission were used to compare predictions with the presence-only test data. Variables used in either model were tested for differences between subwatersheds with surveyed presences and subwatersheds with absences using the nonparametric Wilcoxon rank sum test (Zar 1998).

Results

Modeled variables

The logistic regression and classification tree models incorporated different variables to predict occurrences of the species, with classification trees retaining fewer variables in the models (Tables 5, 6). The Wilcoxon rank sum test indicated that only the number of downstream dams, maximum nitrate deposition in the subwatershed, density of dams in the subwatershed, and maximum slope in the subwatershed were significantly different between surveyed subwatersheds classified as present and absent for *E. complanata*. The number of downstream dams was the only predictor retained in the classification tree model for *E. complanata*; this model performed well upon cross-validation (Table 7). The Wilcoxon test indicated that the proportion of wetlands and the proportion of grass landcover in the subwatershed were the only variables that differed significantly between presence and absence subwatersheds for *E. dilatata*. Three variables were included in the classification tree model for *E. dilatata*, but only the cumulative proportion of shrub landcover was significantly different between absent and present surveyed subwatersheds. This model performed poorly upon cross-validation (Table 7).

Model predictions

The logistic regression and the classification tree models performed well on the training datasets for both species, with accuracy ranging from 68.9 - 81.5% (Table 8). None of the models accurately predicted the majority of the independent presence-only data for both species, as 62.4 - 97.5% of presences were incorrectly classified as absences by the models (Table 8). The logistic regression model for *E. dilatata* performed best on the independent dataset with 37.6% of presence-only subwatersheds correctly classified as present.

Discussion

Models developed for *E. complanata* and *E. dilatata* include variables that research suggests may affect distributions of freshwater mussels (Tables 6, 7). Dams affect freshwater mussels by destroying habitat and by affecting distributions of host fishes, which can limit dispersal and recruitment of freshwater mussels by reducing the chance that mussel larvae, or glochidia, will successfully attach to a host (Ortmann 1924; Williams et al. 1992; Layzer et al.

1993; Watters 1996; Heinricher & Layzer 1999). Municipal effluents are also detrimental to freshwater mussels (Gagne et al. 2001; Blaise et al. 2003), and population density may be an effective estimate of municipal effluents in an area. Although the proportion of grassland within the subwatershed is not known to predict occurrences of freshwater mussels, stream segments with grassy riparian areas have different mussel communities than segments with forested corridors (Morris & Corkum 1996). Although species in the genus *Elliptio* were not included in that study, these species may be more likely to occur in communities such as those associated with grassy riparian areas.

Model limitations

The classification tree and logistic regression models developed for *E. complanata* and *E. dilatata* in the Mid-Atlantic and Ohio drainage regions generally performed well on the training datasets, but the models performed poorly on independent presence-only test data (Table 8). The inability of models for either species to predict known species occurrences from an independent dataset indicates that model results are largely spurious. Several factors likely contributed to the poor performance of the models on the independent data, including the predictive ability of variables available in a GIS, variable selection and model development techniques employed, and the quality of the data used for analysis at this scale.

The variables developed for this modeling effort were restricted to those that could be calculated in a GIS from existing datasets at the scale of subwatersheds. Variables were also summarized by riparian corridors and for cumulative area for each subwatershed in an attempt to examine factors upstream of each subwatershed that may determine mussel occurrences. The scale for analysis was chosen because factors occurring within a watershed have been hypothesized to affect freshwater mussels, and variables entered into the models were considered potentially relevant to freshwater mussels. However, the potential for these variables to predict freshwater mussels could have been affected by variable accuracy and scale. Most datasets available in a GIS across a broad scale have a maximum resolution of 30 m and vary in temporal and positional accuracy and precision, affecting the accuracy of the variables derived from those datasets and their potential for predicting freshwater mussel distributions. It is also possible that distributions of these mussel species are affected by the same variables used in this study, but that the variables are important at scales different from that used in this study. Finally, mussel

distributions may depend on factors inadequately represented in a GIS. If greater accuracy, different scales, or different variables are important to modeling freshwater mussels, these models would be expected to perform poorly.

The models chosen for the logistic regression models were determined by removing variables most responsible for correlations in the dataset. This automated process did not account for logical relationships between the explanatory variables and mussel occurrences, therefore models created from these variable subsets may lack valuable information that could have been derived from easily interpretable variables. Because all available variables were entered into the classification tree models, these models selected the variables most responsible for distinguishing subwatersheds where species were present from those where species were not found in surveys. Although the explanatory variables included in predictive models ideally have a logical relationship to the dependent variable, the variables selected as important by classification tree models for both the Mid-Atlantic and Ohio regions may not have direct effects on the occurrence of these species at the subwatershed scale. Although variables relevant to distinguishing presences from absences may have been masked due to correlations with the variables preserved in the final classification tree models, the poor performance of the models on the independent test dataset indicate that other factors contributed to poor model performance.

Modeling techniques may also affect the accuracy of predictions. The classification tree procedure in the mvpart package (De'ath 2007) in R is a data-partitioning technique capable of identifying variables that best distinguish between presence and absence data. Cross-validation of this technique should reduce overfitting so that models are not specific to the training dataset and are therefore able to predict new data accurately. However, cross-validation is only effective at reducing overfitting if the data used in model development accurately represent the dependent variable of interest. Because the training datasets for each species were highly clustered (Table 3), it is unlikely that these datasets accurately represented the range of conditions in which these species occur across the study area.

Several modeling considerations that may be responsible for poor performance on the test data have been explored, but the dependent variable may also contribute to poor model development. Because prediction accuracy on the training dataset did not suggest overfitting, but model predictions remained more accurate on the training than the test dataset (Table 8), the

training dataset and the test dataset likely differ in their ability to capture true species occurrences. Models created to explain presence and absence data work best when there is high confidence in both the presence and absence data. However, the presence and absence data used in model development may have been limited in spatial distribution, precision, and accuracy (Table 3).

Surveys for the species were biased geographically, and very few surveys were available for the Mid-Atlantic and Ohio drainage regions. Survey data that were available for model development were limited to the central regions of each drainage and were not distributed randomly across the areas (Table 3). The scale at which the data were analyzed and the scale at which data were collected further complicate the dependent variable considered in this study. Freshwater mussels are difficult to sample, and although a presence is almost certainly an actual presence, failure to detect a species in a stream does not necessitate absence. The survey data available for the Mid-Atlantic region were biased toward large rivers, meaning that although a large river within the subwatershed may have been surveyed for mussels, tributaries were not necessarily searched. It is unlikely that any surveys searched all potential mussel habitat in any subwatershed because these surveys were not conducted with subwatershed-scale analysis in mind. This suggests that absences identified by surveys in the study areas may not correlate with absence of a species throughout the subwatershed. It was not possible to develop a measure of confidence in the absence data because the available data indicated that only two of the surveyed subwatersheds were independently sampled more than once. The size of the subwatersheds used in this study, although comparable with successful studies of other aquatic species (Rieman et al. 1997; Thieling 2006), may be prohibitively large for thorough searches of the stream segments.

At the subwatershed scale, the paucity of survey data from which models were created limited the ability to confidently apply model predictions to the entire drainage region, therefore areas where presence or absence of the species were predicted should be considered to be of unknown classification status until field surveys can further evaluate the accuracy of the predictions. When absence data is not available, modelers have experimented with pseudo-absence data created using random points (Engler et al. 2004), but model performance is inconsistent (Lutolf et al. 2006; Pearce & Boyce 2006; Tsoar et al. 2007). Independent absence data were not available for either species and detection estimates are not known for the survey

data used to create the model, so it is not possible to reasonably assess the ability of modeling at this scale to predict absences of freshwater mussels.

Absence data were difficult to confirm for both species, but the presence data for *E. complanata* may have also contained errors. The models for *E. complanata,* none of which accurately predicted independent presence data, may have been complicated by taxonomic problems with the species. Although experts suggested that identification was relatively simple in the Mid-Atlantic region and difficult only in the Atlantic Slope, *E. complanata* has been difficult for some researchers to identify.

Previous studies that successfully modeled distributions of aquatic species using subwatersheds focused on species of fish in the salmonid family (Rieman 1997, Thieling 2006). These studies avoided many of the problems encountered in this study by focusing on species that are easily identified and whose habitat requirements are more commonly known. Although individuals of these species can travel great distances and may reside in both small streams and large rivers during the course of their lives, knowledge of the timing and direction of movements make these life-history characteristics a benefit to distributional studies. Because of the mobile nature of these fish, determining presence or absence at the scale of a subwatershed could be accomplished with a high degree of confidence by sampling a few stream segments in each subwatershed during appropriate times of the year. This differs dramatically from mussels, which are difficult to identify during the mobile larval stage, and which move very little as juveniles and adults. Sampling efficiencies of fish survey methods are therefore much higher for juvenile and adult fish than for juvenile and adult mussels. The same degree of confidence in presence or absence data for mussels would require sampling a much larger proportion of stream segments in each subwatershed than is necessary for fishes. Freshwater mussel research would benefit from studies of how sampling gear (hand sampling or grab sampling, for instance), sampling effort, and the timing of sampling events can affect the probability of detecting a species (Strayer & Smith 2003).

Modeling directions

The benefit of the subwatershed approach was the potential to examine the effects of cumulative and riparian variables on species occurrence, as landscape factors may affect the aquatic environment (Frissell et al. 1986). However, this study demonstrated the difficulties of

modeling freshwater mussels at this scale. Because absence and presence data at the subwatershed scale are difficult to obtain, presence-only modeling using smaller watershed boundaries may be best suited to future broad-scale studies of freshwater mussels.

Defining presence by individual stream segments within a watershed framework would set a smaller sampling area for mussel data, environmental variables, and anthropogenic impacts. Segment-level watersheds (catchments) are available at the 1:100,000 scale for the U.S. through NHDPlus (Horizon Systems 2006) (Figure 2). Because individual stream segments are much shorter stream lengths than the sum of the stream and river segments within a subwatershed, this approach to analysis would allow thorough, repeated sampling, as some researchers already do for monitoring studies, without sacrificing the ability to readily calculate upstream and riparian variables. Sampling all potential mussel habitat within a stream segment to record presence or absence of a species rather than aggregating data collected at various scales to the subwatershed scale increases the probability of detecting a species in a catchment compared to the larger subwatershed. Probabilistic sampling of freshwater mussels in stream segments can also aid in the calculation of sampling efficiencies and detection probabilities for some species (Bayley & Peterson 2001), which could benefit general freshwater mussel research. In the absence of funds for additional sampling, studies could incorporate historic occurrence data with precise location information into the explanatory variable of presence-only models to determine environmental variables that affect the distribution of a species.

Catchments would be an appropriate modeling alternative for future studies of freshwater mussel distributions for several reasons. Unlike grid cells or points, catchments preserve the ability to easily incorporate cumulative, riparian, and watershed variables into the analysis. Although finer-grained analyses using catchments at a regional extent is currently difficult using typical personal computers, the rapid advancement of technology should make these analyses possible in the near future. Catchment analysis could provide better results not only by analyzing environmental variables at a scale more appropriate to freshwater mussels, but by supporting more precise mussel data associated with those environmental variables.

Management implications

Freshwater mussels are long-lived organisms that have declined due to short-term, local changes in the environment and have also been affected by diffuse perturbations (Gagne et al.

2001). The models developed in this study rely on recent occurrences of freshwater mussels in a subwatershed, but may not accurately identify presence or absence of a species. These models do not distinguish populations that are reduced or not recruiting from populations that are healthy or even expanding, and they do not incorporate the effects of local environmental changes on local freshwater mussel populations. Because model predictions did not correspond to known presences of the species, the models and their results should be interpreted with caution. These models and others created at this scale could be valuable for identifying variables that may be correlated with the occurrence of freshwater mussels at the scale of small watersheds in the Appalachians, but the concerns addressed above suggest that additional research is necessary to evaluate whether robust models of freshwater mussel distributions and factors affecting those distributions can be developed.

A national or international georeferenced database of mussel occurrences and known absences and cooperative research-based survey efforts can improve the quality and availability of freshwater mussel data for research based on small sampling units compared over large drainage regions, including the development of current maps of freshwater mussel distributions. Current efforts are focused on developing a national atlas for mussels in the U.S. (Arthur Bogan, unpublished data) and on creating georeferenced databases of survey information to promote access to existing data (e.g., West Virginia Department of Natural Resources (WV DNR) Comprehensive Wildlife Conservation Strategy). Although agencies and researchers developing mussel databases intend to provide data in a format useful for ecological and environment research, outlines of sampling and database plans for individual states suggest that databases will likely continue to differ among states, as earlier efforts have been limited by political boundaries or have aggregated data to different scales. Databases that specify location, accuracy, and precision of sampled sites, define the sampling period (e.g., a single date or over the course of a year), and describe the purpose and type of sampling (e.g., abundance estimates of a single species or inventories of all species at a site) will be most valuable for modeling efforts, as combining data collected for different reasons with different techniques can limit the accuracy of models and model predictions. For example, the usefulness of the WV DNR's intended database identifying the length of stream occupied by mussel species at each recent occurrence will depend on whether the database describes the minimum length between mussel beds necessary to distinguish one occurrence from another and whether the database identifies areas where

particular species were absent. The proposed development of an Atlas of Freshwater Mussels of the United States has the most potential to provide useful data for regional-scale modeling efforts because the researchers plan to georeference available species records for the entire United States and confirm identifications of mussel specimens in collections (Arthur Bogan, unpublished data). However, if absence data are not included in this database, modeling applications based on this database would be limited to those techniques that do not require absence data.

Although mussel researchers may choose to concentrate their survey efforts on areas known to contain suitable mussel habitat, models will benefit from the inclusion of unsuitable areas from which all or certain species of mussels are known to be absent. Efforts to develop a database of areas where extirpations have occurred containing information on how long species remain absent from specific sites would promote development of models that can distinguish anthropogenic factors affecting mussel occurrences. Finally, relatively frequent surveys for particular species will be particularly useful in evaluating how seasonal variation, weather patterns, or other factors affect sampling efficiency, or the probability that researchers detect species when they occur.

Increasing the availability of recorded presences and absences for mussel species relevant to the segment scale can increase the potential to model rare or locally distributed mussel species. Sowa et al. (2005) used 1,157 community records containing 12,304 mussel occurrences to successfully model potential habitat for 67 species at the 8-digit HU scale within a state; this research attempted to model species at the finer subwatershed scale throughout the Appalachians using fewer mussel occurrences (Appendix B). Previous researchers have developed methods to determine when a watershed containing historic data remains undersampled (Sowa et al. 2005), and following efforts to assign coordinates to historic mussel collections, these methods can be adapted to analyzing the availability of mussel data at a regional extent. Sampling can then be focused to address under-sampled regions and to ensure that data are available in a subset of catchments representing the range of environmental conditions and mussel communities across the drainage region. These sampling efforts would require cooperation from universities, consultants, non-profits, and state and federal agencies, and could be incorporated into Comprehensive Wildlife Conservation Plans.

Cooperative sampling efforts and studies of freshwater mussels at a variety of scales using proven sampling techniques can address the concerns of the National Native Mussel Conservation Committee (1998) by increasing knowledge of freshwater mussel distributions, ecology, habitat requirements, and limitations while developing cooperative partnerships for mussel research and conservation. By establishing sampling protocols for specific mussel species at a regional level, a freshwater mussel dataset can be developed that is valuable for estimating probability of detection and deriving predictive models that incorporate both local- and landscape-level effects on species of interest. Novel field research could lead to findings related to minimum sampling efforts required to detect a species, sampling efficiencies of different sampling gear and methods, and temporal variation in sampling success. Modeled distributions of species across their geographic ranges, if successful, could provide valuable information on species locations. This information can provide a basis for future studies by establishing minimum sampling effort for species detections, defining relationships between season and sampling success, and by identifying factors that may affect species occurrences.

The use of survey data to develop accurate, current distributional maps and to identify factors affecting freshwater mussels over a large extent at a scale relevant to scientists and managers would be an important first step toward addressing the many problems affecting our ability to conserve and manage freshwater mussels (National Native Mussel Conservation Committee 1998). Although maps cannot replace sampling efforts for identifying local mussel populations, models and mapped distributions can be valuable for identifying priority areas for sampling and for summarizing species occurrences over large areas. Agency efforts and funding opportunities promoting additional research on freshwater mussel distributions should improve our understanding of the diversity of native mussels in the Appalachians and throughout the United States.

Acknowledgments

References

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

———— & Johnson, L. (1997) Catchment-scale analysis of aquatic ecosystems. *Freshwater Biology* **37**(1): 107-111.

Andersen, E. F. (2002) Effects of land use and land cover on freshwater mussel populations in the upper Neuse River basin, North Carolina: A GIS approach. M.S. thesis, Raleigh, North Carolina: North Carolina State University.

Angermeier, P. L., Krueger, K. L. & Dolloff, C. A. (2002) Discontinuity in stream-fish distributions: implications for assessing and predicting species occurrence. In: *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall & F. B. Samson, Washington, DC: Island Press.

Arbuckle, K. E. & Downing, J. A. (2002) Freshwater mussel abundance and species richness: GIS relationships with watershed land use and geology. *Canadian Journal of Fisheries and Aquatic Sciences* **59**(2): 310-316.

Bayley, P. B. & Peterson, J. T. (2001) An approach to estimate probability of presences and richness of fish species. *Transactions of the American Fisheries Society* **130**: 620-633.

Blaise, C., Gagne, F., Salazar, M., Salazar, S., Trottier, S. & Hansen, P.-D. (2003) Experimentally-induced feminisation of freshwater mussels after long-term exposure to a municipal effluent. *Fresenius Environmental Bulletin* **12**(8): 865-870.

Breiman, L., Friedman, J., Olshen, R. & Stone, C. (1984) *Classification and Regression Trees*. Boca Raton, Florida: Chapman & Hall.

Buehler, D. A., Welton, M. J. & Beachy, T. A. (2006) Predicting cerulean warbler habitat use in the Cumberland Mountains of Tennessee. *Journal of Wildlife Management* **70**(6): 1763-1769.

De'ath, G. (2007). Mvpart: Multivariate partitioning. R package version 1.2-6.

———— . & Fabricius, K. E. (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* **81**(11): 3178-3192.

Dominguez-Dominguez, O., Martinez-Meyer, E., Zambrano, L. & De Leon, G. P.-P. (2006) Using ecological-niche modeling as a conservation tool for freshwater species: live-bearing fishes in central Mexico. *Conservation Biology* **20**(6): 1730-1739.

Dynesius, M. & Nilsson, C. (1994) Fragmentation and flow regulation of river systems in the northern third of the world. *Science* **266**: 753-762.

Engler, R., Guisan, A. & Rechsteiner, L. (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology* **41**(2): 263-274.

ESRI (Environmental Systems Research Institute) (2000) ArcView v. 3.2. Redlands, California.

ESRI (Environmental Systems Research Institute) (2002) ArcView v. 9.2. Redlands, California.

ESRI Data & Maps (2002) Redlands, California: Environmental Systems Research Institute.

Fielding, A. H. & Bell, J. F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24**(1): 38-49.

FitzHugh, T. W. (2005) GIS tools for freshwater biodiversity conservation planning. *Transactions in GIS* **9**(2): 247-263.

Freeman, M. C., Pringle, C. M. & Jackson, C. R. (2007) Hydrologic connectivity and the contribution of stream headwaters to ecological integrity at regional scales. *Journal of the American Water Resources Association* **43**(1): 5-14.

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

Gagne, F., Blaise, C., Salazar, M., Salazar, S. & Hansen, P.-D. (2001) Evaluation of estrogenic effects of municipal effluents to the freshwater mussel *Elliptio complanata. Comparative Biochemistry and Physiology Part C* **128**: 213-225.

Gatenby, C. M., Patterson, M. A., Devers, J. L. & Kreeger, D. A. (2006) Condition of freshwater mussels held in refugia at the White Sulphur Springs National Fish Hatchery, West Virginia. *Journal of Shellfish Research* 25(2): 729.

Gregory, S. V., Swanson, F. J., McKee, W. A. & Cummins, K. W. (1991) An ecosystem perspective of riparian zones. *BioScience* **41**(8): 540-551.

Heinricher, J. R. & Layzer, J. B. (1999) Reproduction by individuals of a nonreproducing population of *Megalonaias nervosa* (Mollusca: Unionidae) following translocation. *American Midland Naturalist* **141**: 140-148.

Hernandez, P. A., Graham, C. H., Master, L. L. & Albert, D. L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* **29**: 773-785.

Hirzel, A. H., Helfer, V. & Metral, F. (2001) Assessing habitat-suitability models with a virtual species. *Ecological Modelling* **145**: 111-121.

Hitt, N. P. & Angermeier, P. L. (2008) Evidence for fish dispersal from spatial analysis of stream network topology. *Journal of the North American Benthological Society* **27**(2): 304-320.

Homer, C., Dewitz, J., Fry, J., Coan, M.., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N. & Wickham, J. (2007) Completion of the 2001 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering and Remote Sensing* **73(**4): 337-341.

Horizon Systems (2006) NHDPlus. URL http://www.horizon-systems.com/nhdplus/

Hosmer, D. & Lemeshow, S. (2000) *Applied Logistic Regression*. New York: John Wiley and Sons.

Houp, R. E. (1993) Observations on long-term effects of sedimentation on freshwater mussels (Mollusca: Unionidae) in the North Fork of Red River, Kentucky. *Transactions of the Kentucky Academy of Science* **54**(3-4): 93-97.

Hudy, M., Thieling, T. M., Gillespie, N. & Smith, E. P. (2005) Distribution, status, and perturbations to brook trout within the eastern United States. Final report to the steering committee of Eastern Brook Trout Joint Venture.

Kerr, J. T. (1997) Species richness, endemism, and the choice of areas for conservation. *Conservation Biology* **11**(5): 1094-1100.

Kiffney, P. M., Richardson, J. S. & Bull, J. P. (2003) Responses of periphyton and insects to experimental manipulation of riparian buffer width along forest streams. *Journal of Applied Ecology* **40**(6): 1060-1076.

Kuhn, M., Wing, J., Weston, S. & Williams, A. (2008) Caret: classification and regression training. R package version 3.13.

Layzer, J. B., Gordon, M. E. & Anderson, R. M. (1993) Mussels: the forgotten fauna of regulated rivers. A case study of the Caney Fork River. *Regulated Rivers* **8**: 63-71.

Leathwick, J. R., Rowe, D., Richardson, J., Elith, J. & Hastie, T. (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology* **50**(12): 2034-2052.

Link, W. A. & Sauer, J. R. (2002) A hierarchical analysis of population change with application to cerulean warblers. *Ecology* **83**(10): 2832-2840.

Lutolf, M., Kienast, F. & Guisan, A. (2006) The ghost of past species occurrence: improving species distribution models for presence-only data. *Journal of Applied Ecology* **43**(4): 802-815.

Lydeard, C., Clark, S. A., Perez, K. E., Cowie, R. H., Ponder, W. F., Bogan, A. E., Bouchet, P., Gargominy, O., Cummings, K. S., Frest, T. J., Herbert, D. G., Hershler, R., Roth, B., Seddon, M., Strong, E. E. & Thompson, F. G. (2004) The global decline of nonmarine mollusks. *Bioscience* **54**(4): 321-330.

Mastrorillo, S., Lek, S., Dauba, F. & Belaud, A. (1997) The use of artificial neural networks to predict the presence of small-bodied fish in a river. *Freshwater Biology* **38**(2): 237-246.

McKinney, R. A., Lake, J. L., Charpentier, M. A. & Ryba, S. (2002) Using mussel isotope ratios to assess anthropogenic nitrogen inputs to freshwater ecosystems. *Environmental Monitoring and Assessment* **74**(2): 167-192.

Miller, D. A. & White, R. A. (1998) A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions* **2**(2): 1-26.

Morris, T. J. & Corkum, L. D. (1996) Assemblage structure of freshwater mussels (Bivalvia: Unionidae) in river with grassy and forested riparian zones. *Journal of the North American Benthological Society* **15**(4): 576-586.

Moss, B. (2000) Biodiversity in fresh waters: an issue of species preservation or system functioning? *Environmental Conservation* **27**(1): 1-4.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., de Fonseca, G. A. B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature* **403**: 853-858.

NADP (National Atmospheric Deposition Program NRSP-3) (2007) NADP Program Office, Illinois State Water Survey. URL http://nadp.sws.uiuc.edu/

NAS (Nonindigenous Aquatic Species) Program (2007) Zebra Mussel (*Dreissena polymorpha*) URL http://nas.er.usgs.gov/taxgroup/mollusks/zebramussel/

National Atlas (2007) National Atlas of the United States. URL http://nationalatlas.gov/atlasftp.html/

National Native Mussel Conservation Committee (1998) National strategy for the conservation of native freshwater mussels. *Journal of Shellfish Research* **17**(5): 1419-1428.

Northington, R. M. & Hershey, A. E. (2006) Effects of stream restoration and wastewater treatment plant effluent on fish communities in urban streams. *Freshwater Biology* **51**(10): 1959-1973.

Nuttle, T., Leidolf, A. & Burger, L. W., Jr. (2003) Assessing conservation value of bird communities with Partners In Flight-based ranks. *The Auk* **120**(2): 541-549.

Olden, J. D. & Jackson, D. A. (2002) A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology* **47**(10): 1976-1995.

Ortmann, A. E. (1924) Mussel shoals. *Science* **60**: 565-566.

Parmalee, P. W. & Bogan, A. E. (1998) The Freshwater Mussels of Tennessee. Knoxville, TN: University of Tennessee Press.

Pearce, J. L. & Boyce, M. S. (2006) Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology* **43**(3): 405-412.

Poole, K. E. & Downing, J. A. (2004) Relationship of declining mussel biodiversity to stream-reach and watershed characteristics in an agricultural landscape. *Journal of the North American Benthological Society* **23**(1): 114-125.

R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Research Triangle Institute (2001) The National water pollution control assessment model (NWPCAM) version 2, Draft report prepared for the U.S. Environmental Protection Agency, Research Triangle Park, North Carolina: Research Triangle Institute.

Richards, C., Johnson, L. B. & Host, G. E. (1996) Landscape-scale influences on stream habitats and biota. *Canadian Journal of Fisheries and Aquatic Sciences* **53**(Suppl. 1): 295-311.

Rieman, B. E., Lee, D. C. & Thurow, R. F. (1997) Distribution, status, and likely future trends of bull trout within the Columbia River and Klamath River basins. *North American Journal of Fisheries Management* **17**: 1111-1125.

Sauer, J. R., Fallon, J. E. & Johnson, R. (2003) Use of North American Breeding Bird Survey data to estimate population change for Bird Conservation Regions. *The Journal of Wildlife Management* **67**(2): 372-389.

Schomberg, J. D., Host, G., Johnson, L. B. & Richards, C. (2005) Evaluating the influence of landform, surficial geology, and land use on streams using hydrologic simulation modeling. *Aquatic Sciences* **67**(4): 528-540.

Seaber, P. R., Kapinos, F. P. & Knapp, G. L. (1987) Hydrologic Unit Maps. U.S. Geological Survey Water-Supply Paper 2294, 63 p.

Sethi, S. A., Selle, A. R., Doyle, M. W., Stanley, E. H. & Kitchel, H. E. (2004) Response of unionid mussels to dam removal in Koshkonong Creek, Wisconsin (USA). *Hydrobiologia* **525**(1-3): 157-165.

Sowa, S. P., Diamond, D. D., Abbitt, R., Annis, G., Gordon, T., Morey, M. E., Sorensen, G. R., True, D. (2005) A gap analysis for riverine ecosystems of Missouri. Final report submitted to the USGS National Gap Analysis Program.

Strayer, D. L. (1993) Macrohabitats of freshwater mussels (Bivalvia: Unionacea) in streams of the northern Atlantic Slope. *Journal of the North American Benthological Society* **12**(3): 236-246.

———— (2006) Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society* **25**(2): 271-287.

———— (2008) *Freshwater Mussel Ecology: A Multifactor Approach to Distribution and Abundance*. Berkeley, California: University of California Press.

————, Downing, J. A., Haag, W. R., King, T. L., Layzer, J. B., Newton, T. J. & Nichols, S. J. (2004) Changing perspectives on pearly mussels, North America's most imperiled animals. *BioScience* **54**(5): 429-439.

———— & Smith, D. R. (2003) *A Guide to Sampling Freshwater Mussel Populations*. Bethesda, Maryland: American Fisheries Society.

Thieling, T. M. (2006) Assessment and predictive model for brook trout (*Salvelinus fontinalis*) population status in the eastern United States. M.S. thesis, Harrisonburg, Virginia: James Madison University.

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. (2007) A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions* **13**(4): 397-405.

Turgeon, D. D., Bogan, A. E., Coan, E. V., Emerson, W. K., Lyons, W. G., Pratt, W. L., Roper, E. F. E., Scheltema, A., Thompson, F. G. & Williams, J. D. (1988) Common and Scientific Names of Aquatic Invertebrates from the United States and Canada: Mollusks. Bethesda, Maryland: American Fisheries Society.

US EPA (U.S. Environmental Protection Agency) (2002) NHD Indexed locations for Section 303(d) Listed Waters. URL http://www.epa.gov/waters/data/downloads.htm/

USGS (U.S. Geological Survey), Eros Data Center (1999) National Elevation Dataset (30m). URL http://gisdata.usgs.net/ned/

USGS (U.S. Geological Survey) (2003) U.S. Geological Survey Minerals Information. URL http://minerals.usgs.gov/minerals/

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

Vaughn, C. C., Nichols, S. J. & Spooner, D. E. (2008) Community and foodweb ecology of freshwater mussels. *Journal of the North American Benthological Society* **27**(2): 409-423.

Wang, L., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., Kanehl, P., Baker, E., Niemela, S. & Stewart, P. M. (2003) Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, U.S.A. *Canadian Journal of Fisheries & Aquatic Sciences* **60**: 491-505.

Watters, G. T. (1996) Small dams as barriers to freshwater mussels (Bivalvia, Unionoida) and their hosts. *Biological Conservation* **75**(1): 79-85.

Williams, J. D., Fuller, S. L. H. & Grace, R. (1992) Effects of impoundments on freshwater mussels (Mollusca: Bivalvia: Unionidae) in the main channel of the Black Warrior and Tombigbee Rivers in Western Alabama. *Bulletin of the Alabama Museum of Natural History* **13**: 1-10.

Willson, J. D. & Dorcas, M. E. (2003) Effects of habitat disturbance on stream salamanders: implications for buffer zones and watershed management. *Conservation Biology* **17**(3): 763-771.

Zar, J. H. (1998) *Biostatistical Analysis.* Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Tables

Table 1. Available data and training and test data for models of *Elliptio complanata* and *E. dilatata*. Total refers to the total number of records provided by each source. The date range is determined from all records. The number of recent records are found in the column ≥1985, the number of records with associated coordinates or area locations are found in the columns Coord or Area, and Both refers to the number of spatially explicit recent records from each source.

Point Records	*Elliptio dilatata*					*Ellipto complanata*					No Mussels			
	Date Range	≥ 1985	Coord	Both	Total	Date Range	≥ 1985	Coord	Both	Total	≥ 1985	Coord	Both	Total
Carnegie Museum	1884-2006	30	42	15	365	1889-2005	7	6	4	196				N/A
DE Museum	1898-2001	4	0	0	165	1856-1996	3	0	0	296				N/A
IL Natural History Museum	1878-2006	550	87	79	733	1907-1998	50	14	9	107				N/A
MD Biological Stream Survey	N/A	0	0	0	0	1995-1997	72	72	72	72	778	778	778	778
NC Natural Heritage	1991-2006	94	69	69	94	2000-2002	2	2	2	2				N/A
OH State U. Museum	1825-2001	322	1313	283	1740	1850-2002	53	46	37	284				N/A
Smithsonian	1933-1974	0	0	0	17	1949-1991	20	19	0	60				N/A
U. of MI Museum of Zoology	1863-1997	0	0	0	1126	1866-1979	0	0	0	1061				N/A
WV Natural Heritage	1919-2001	113	581	113	581	1919-2000	55	141	55	141				N/A
Subwatershed Records	Date Range	≥ 1985	Area	Both	Total	Date Range	≥ 1985	Area	Both	Total	≥ 1985	Area	Both	Total
Western PA Conservancy	N/A	39	39	39	39	N/A	125	125	125	125	156	156	156	156
Stream Segment Records by Subwatershed	Date Range	≥ 1985	Area	Both	Total	Date Range	≥ 1985	Area	Both	Total	≥ 1985	Area	Both	Total
Ecological Specialists, Inc.	1969-1999	9	9	9	9		0	0	0	0	16	16	16	16

Table 2. Number of subwatersheds in the study area with data for *Elliptio complanata* (Mid-Atlantic region) and *E. dilatata* (Ohio region). Presence and absence subwatersheds were determined from survey data. The test dataset (presence-only) contains data derived from occurrence records. The overlap between the two datasets is indicated by the number of subwatersheds with presence or absence data in the training dataset that are also represented in the test dataset by occurrence-derived presences.

	Training Datasets		Test Datasets	Test and Training Data Overlap	
	Presence	Absence	Presence	Surveyed Present	Surveyed Absent
E. complanata	143	114	40	2	1
E. dilatata	47	126	205	14	17

Table 3. Nearest neighbor index (NNI) classification (clustered vs. dispersed) and scores measuring degree of clustering for the training (presence/absence) datasets for *Elliptio complanata* in the Mid-Atlantic and *E. dilatata* in the Ohio drainage regions. P-values less than 0.05 indicate significant clustering.

	Mid-Atlantic	Ohio
NNI	clustered	clustered
z-score	-2.27	-6.59
p-value	<0.05	<0.01

Table 4. Data sources for variables included in model development.

Variable Categories	Data Source	Scale or Resolution	Relevant Time Period	Source Details
Presence/absence data	survey data	subwatershed	1985-2007	Mussel data were compiled from the Maryland Biological Stream Survey, Ecological Specialists, Inc, and a database compiled by the Western Pennsylvania Conservancy
Presence-only data	occurrences	subwatershed	1985-2007	Mussel data were compiled from West Virginia DNR, Carnegie Museum of Natural History, The Ohio State University Museum, and Illinois Natural History Museum
Zebra mussel occurrences	NAS	1:2,000,000	1988-2007	Point locations for zebra mussels compiled by the Nonindigenous Aquatic Species (NAS) Program were downloaded from the National Atlas
Dams	NID	1:2,000,000	2002-2004	A shapefile containing locations of dams compiled by the National Inventory of Dams was downloaded from the National Atlas
Roads and population data	ESRI	1:100,000	2000	U.S. Detailed Streets and U.S. Census Block Centroid Populations were obtained from Environmental Systems Research Institute, Inc. (ESRI) data CDs provided with ArcGIS software
Metal and mining facilities	USGS	1:2,000,000	2003	Shapefiles of locations of metal processing plants and mines compiled by the USGS were downloaded from the National Atlas
Polluted streams	USEPA	1:100,000	2002	303-D listed streams were obtained from the U.S. Environmental Protection Agency (USEPA)
Stream data	NHDPlus	1:100,000	2005	National Hydrography Dataset Plus (NHDPlus) data were downloaded from Horizon Systems, Inc.
Elevation and slope	NED	30m	varies	The National Elevation Dataset (NED) was obtained directly from the USGS
Sulfate and nitrate deposition	NADP	2.5km	2005	Deposition data were obtained from the National Atmospheric Deposition Program (NADP)
Soil acidity	CONUS	1km	1998	Soil acidity was determined from upper soil layers in CONUS-SOIL data prepared by Miller and White (1998).
Land use and land cover	NLCD	30m	1999-2001	National Land Cover Database (NLCD) land use grids were downloaded from the Multi-Resolution Land Characteristics Consortium (MRLC) by mapping zone (Homer et al. 2007.

Table 5. Reduced variable subsets entered into logistic models of *E. complanata* and *E. dilatata.* These subsets of variables were selected by removing the most highly correlated variables (r>0.5) from the full variable set. NLCD values refer to land use/land cover classes for the 2001 National Land Cover Database.

E. complanata		*E. dilatata*	
Variables	Meaning	Variables	Meaning
NUMDSDAMS	number of downstream dams	WET	proportion of woody wetlands (NLCD 90) in the subwatershed
NO3_MAX	maximum NO$_3$ deposition in subwatershed	DISTUSDAM	upstream distance to nearest dam
DAMDENAREA	density of dams in the subwatershed	MAX_STMSLP	maximum streamslope from NHDPlus streamflow data
CU_HAY_RC	cumulative proportion of pasture/hay in the riparian corridor	GRS	proportion of grass (NLCD 71) in the subwatershed
EMH_RC	proportion of emergent herbaceous (NLCD 95) in the riparian corridor	CU_HDV	cumulative proportion of high-intensity development in the subwatershed
SLOPE_MAX	maximum slope in subwatershed	NUMDSDAMS	number of downstream dams
CURIPAREA	cumulative area of the riparian corridor of the subwatershed		
WET	proportion of woody wetlands (NLCD 90) in the subwatershed		
CUMIDENAR	cumulative density of mining facilities in the subwatershed		
MAFLOU_MIN	minimum mean annual flow (unit runoff method, Research Triangle Institute 2001) from NHDPlus streamflow data		
CRP_RC	proportion of cultivated crops (NLCD 82) in the riparian corridor		

Table 6. Variables and coefficients (BETA) for logistic regression models for freshwater mussel species E. complanata (ECOM) and E. dilatata (EDIL). The level of parameter significance is indicated by the p-value. Significance of Wilcoxon tests for differences in the variables between presence and absence subwatersheds are given under WILC.

Species	Variables	Beta	P-value	WILC
ECOM	NUMDSDAMS	0.1886	<0.0001	0.0004
	NO3_MAX	0.9115	<0.0001	0.0142
	DAMDENAREA	-55.0468	0.0009	0.0064
	CU_HAY_RC	-1.1873	0.0022	0.1564
	EMH_RC	-46.2130	0.0123	0.3173
	SLOPE_MAX	0.0229	0.0052	0.0074
	CURIPAREA	-0.00030	0.0157	0.1886
	WET	56.0705	0.0130	0.3819
	CUMIDENAR	-68.3266	0.0338	0.0688
	MAFLOU_MIN	-57.2486	0.0226	0.3329
	CRP_RC	4.8671	0.0449	0.2091
EDIL	WET	36.4601	0.0002	0.0032
	DISTUSDAM	-1.67E-7	0.0047	0.4477
	MAX_STMSLP	0.3237	0.0083	0.6456
	GRS	51.2941	0.0248	0.0057
	CU_HDV	-691369	0.0266	0.6414
	NUMDSDAMS	-0.0669	0.0308	0.5140

Table 7. Variables and cross-validation error (CV Error) for classification tree models for *Elliptio complanata* (ECOM) and *E. dilatata* (EDIL). CV Error > 1 indicates that the model did not perform well on the training dataset. The significance of differences between values in presence subwatersheds and values in absence subwatersheds calculated using the Wilcoxon test is given under WILC for each variable. Significant WILC scores ($p<0.05$) indicates that subwatersheds with surveyed presences differ significantly from subwatersheds with surveyed absences for that variable.

Species	CV Error	Variables	WILC
ECOM	0.716	NUMDSDAMS	0.0004
EDIL	1.25	CU_SHB	<0.0001
		LO	0.2469
		POPDENAREA	0.7708

Table 8. Summary of logistic regression (LOG) and classification tree (CT) models predicting training and independent test data for *E. complanata* (ECOM) and *E. dilatata* (EDIL). Measures of model evaluation given for training data include accuracy (ACC), errors of omission (OMIT) and commission (COM) (%) for presences, and kappa; errors of omission equal accuracy for evaluation on the test data.

| Species | Model | Presence/Absence Training Data | | | | Presence-Only Test Data |
		ACC	OMIT	COM	KAPPA	OMIT
ECOM	CT	68.87	34.21	35.96	0.371	87.50
	LOG	72.80	24.80	30.40	0.541	97.50
EDIL	CT	81.50	59.57	8.51	0.444	91.22
	LOG	76.30	20.30	40.00	0.367	62.44

Figures
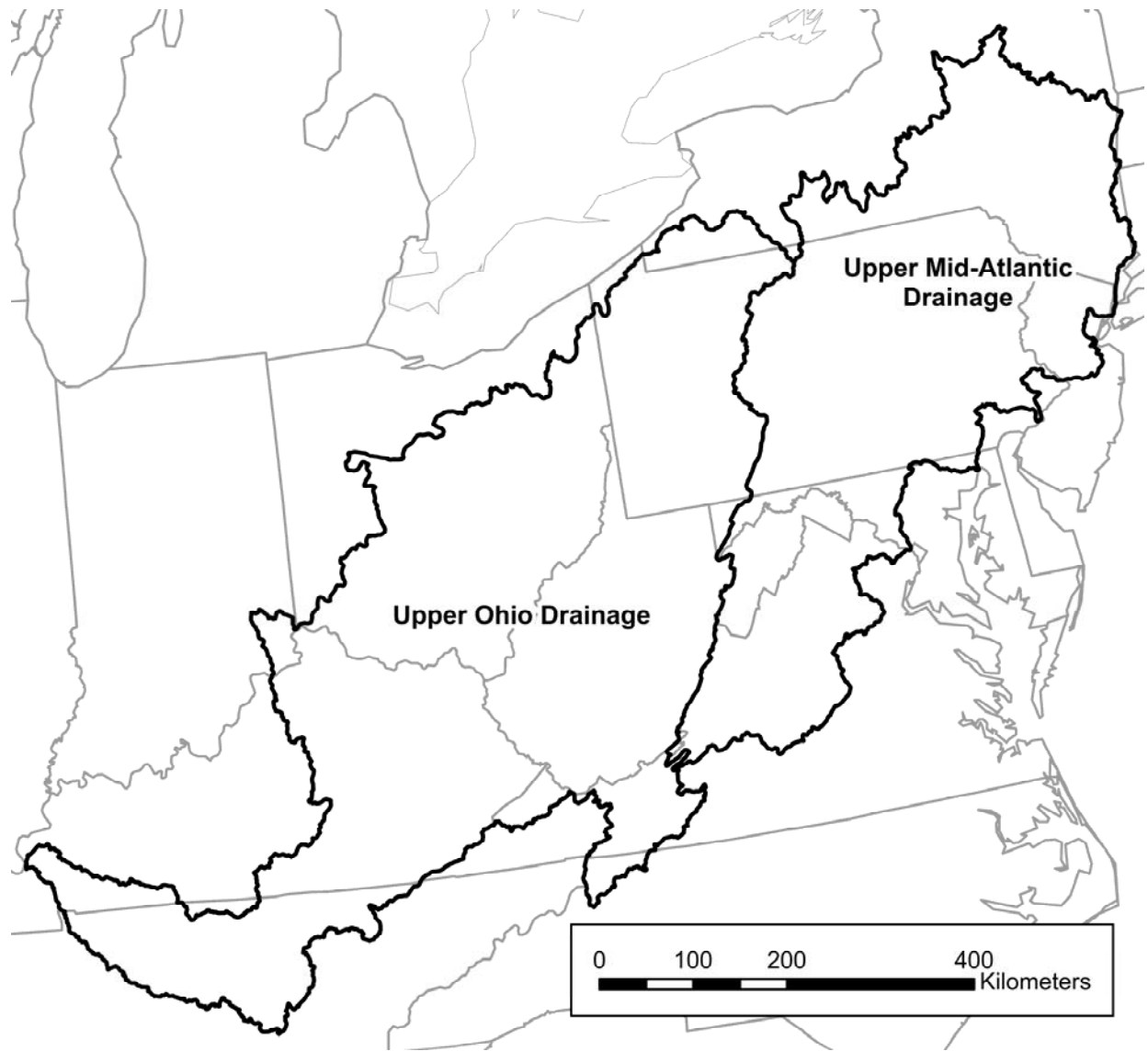
Figure 1. Study area for modeling mussel distributions in the Appalachians.
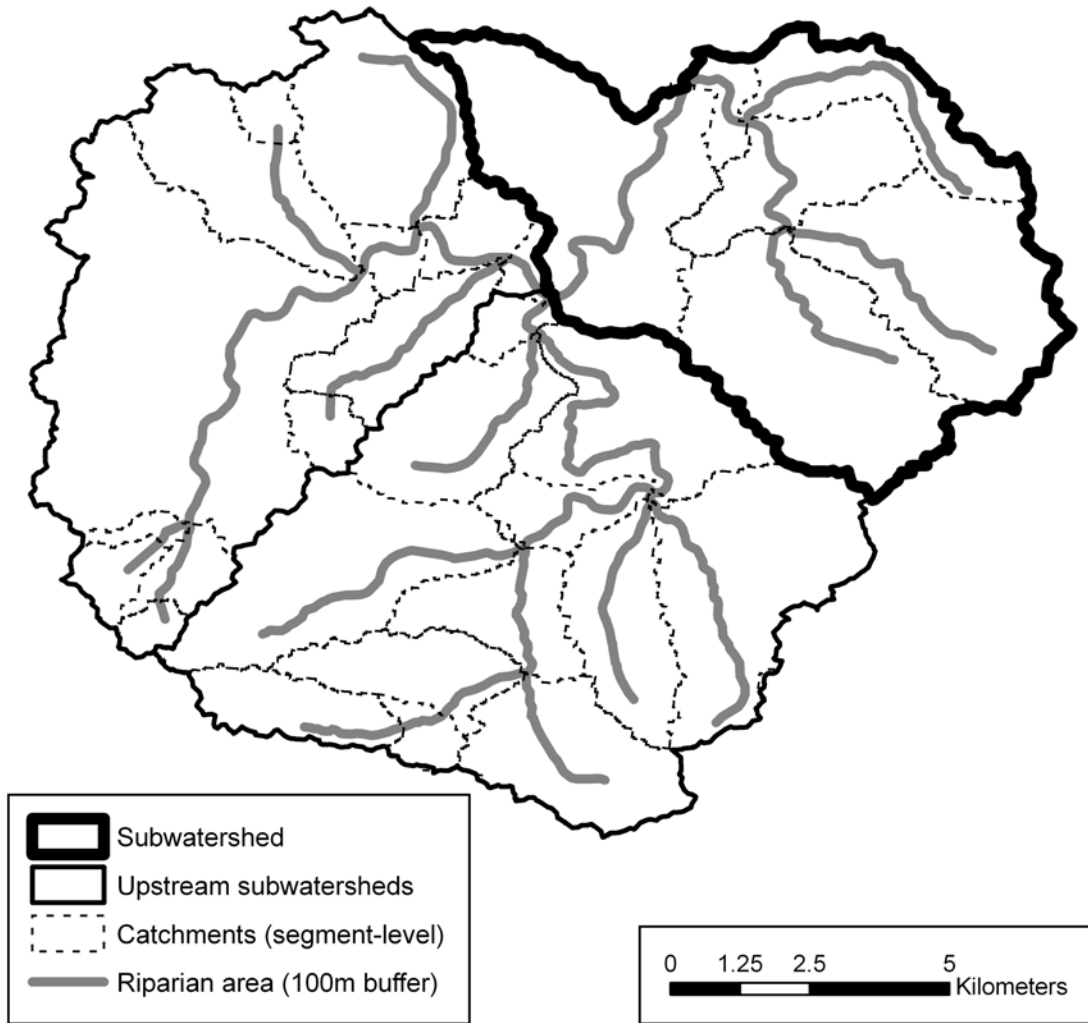
Figure 2. Variables are calculated by subwatershed, cumulative area (including upstream subwatersheds), riparian area, and cumulative riparian area (including upstream riparian areas). Stream segments are contained by catchments, which are subdivisions of subwatersheds.

Chapter 3: Catchment-Scale Predictive Models for Freshwater Mussel Species (Mollusca: Unionidae) in the North Fork Holston River, VA

Abstract

Freshwater mussels are declining in species richness and abundance, particularly in the Appalachians in the eastern United States. Developing effective ways to predict occurrences of freshwater mussels and to discern environmental factors associated with freshwater mussels or their absences are required to effectively manage the large diversity of mussels in the region. Models were developed from available mussel survey data and readily available digital environmental datasets to predict occurrences of freshwater mussels. Significant models could not be developed for all species due to issues with small sample sizes, but results suggest that developing predictive models of mussel occurrences by catchments may be useful in future research.

Introduction

North America contains some of the most diverse assemblages of freshwater mussels in the world, with 297 recognized freshwater mussel species in the *Margaritiferidae* and *Unionidae* families (Turgeon et al. 1988). This diversity of mussels is threatened in the United States, with 72 mussel species federally listed as endangered or threatened (US FWS 2006). Nearly half of freshwater mussel species native to the United States (U.S.) can be found in the Appalachian region, and many of these species are official species of concern by state or federal standards. The loss of this diversity deserves attention because freshwater mussels are an important component of freshwater ecosystems in North America, comprising a significant amount of the benthic biomass and contributing to water filtration and nutrient cycling in many systems (Negus 1966).

Although mussels are a critical part of freshwater ecosystems, they have generally been studied much less than both terrestrial organisms and most other components of freshwater ecosystems (Lydeard et al. 2004). Because basic research on freshwater mussels is still in its early stages, complete information is unavailable regarding the ecology of many mussel species, including their functional roles, habitat requirements, and limitations to population growth (Strayer et al. 1999; Vaughn and Hakenkamp 2001), although researchers are beginning to address these information gaps (Vaughn et al. 2004). The available ecological literature on freshwater mussels focuses primarily on the reproductive cycle and issues of stream quality, such

as invasive species or stream fragmentation, that affect species and communities (e.g. Diamond and Serveiss 2001; Hallac and Marsden 2000; Schneider et al. 1998; Watters 1996; Watters et al. 2005). More information is needed to determine specific habitat requirements for freshwater mussels.

There is disagreement in the literature regarding the ability to predict mussel occurrence from specific environmental variables (Brim-Box et al. 2002; Hornbach 2001). However, general relationships between freshwater mussel occurrence and aspects of the environment have been identified. Many freshwater mussel studies have examined relationships between mussels and microhabitat variables or local water quality measures (e.g., Strayer 1981; Strayer and Ralley 1993). Other studies have examined factors at a broader scale. Brim-Box et al. (2002) determined that basin characteristics such as distributions of host fishes are one of three major factors influencing mussel communities, and Morris and Corkum (1996) found differences in mussel communities between streams with forested riparian zones and streams with grassy riparian zones. Strayer (1983) found a significant relationship between mussels and both surface geology and a measure of stream size, concluding that hydrology is an important predictor of the distribution of freshwater mussels. DiMaio and Corkum (1995) described mussel communities specific to drainage basins with high hydrological variability and different communities specific to hydrologically stable drainage basins.

Basin-scale research on freshwater mussels and research suggesting that factors related to the riparian area and the watershed of a stream segment can affect components of freshwater ecosystems within a given stream reach (Vannote et al. 1980) indicate that landscape components are important to include in studies of freshwater mussels. For example, land use types within a watershed may alter the variability of water quality parameters, therefore research incorporating landscape factors in analyses of mussel occurrence may be more robust than analyses limited to highly variable stream-reach conditions assessed during one sampling effort. Because freshwater mussels occur in the network of freshwater ecosystems, assessing impacts of the terrestrial landscape on their occurrence can be difficult. By summarizing components of the landscape comprising the watershed in which species occur or by calculating landscape factors upstream of sampled sites, both terrestrial and aquatic variables with local and cumulative impacts can be related to occurrences of freshwater organisms (Strayer et al. 2003).

Only recently have ecological studies of freshwater mussels incorporated broad landscape factors affected by human influences (Arbuckle and Downing 2002; Bambarger 2006; Poole and Downing 2004). Bambarger (2006) used a hierarchical approach that combined variables representing the microhabitat, sub-segment, and riparian buffer scales and found that sediment measures at the microhabitat scale, geology at the sub-segment scale, and land use at both the sub-segment and riparian buffer spatial scales was highly correlated with freshwater mussel abundance and species richness. Poole and Downing found evidence for watershed-level impacts of agriculture on mussel populations (Arbuckle and Downing 2002; Poole and Downing 2004). Although few studies have successfully correlated landscape or cumulative variables with metrics of mussel populations, additional research is necessary to fully explore these issues.

This study will identify variables at catchment, upstream, and riparian scales that affect occurrences of freshwater mussels within stream reaches using a classification tree approach. Classification trees are appropriate for ecological studies because the analysis method is non-parametric and outputs an easily interpretable tree, in the form of a dichotomous key, which uses values of the explanatory variables to split the data into separate classes (Breiman et al. 1984; De'ath and Fabricius 2000). Results of the analyses of mussel occurrence at the catchment scale in the North Fork Holston River, VA, will be compared to findings from a similar study conducted at the subwatershed scale for mussel species in the Mid-Atlantic and Ohio regions. The results of this exploratory project can provide additional information on the importance of upstream, riparian corridor, and catchment landscape conditions to the occurrence of freshwater mussels, a list of variables that may affect mussel occurrences at the catchment scale, and the degree of spatial autocorrelation affecting analyses at this scale. This project suggests an alternative to traditional studies of freshwater mussels at the microhabitat scale and identifies benefits and costs associated with attempts to broaden the focus of freshwater mussel research to include aspects of the terrestrial landscape.

Methods

Study area

The upper North Fork Holston River is located in western Virginia and flows southwest into the Holston and ultimately the Tennessee Rivers. The uppermost reaches of the North Fork Holston River and its tributaries were the focus of this study, draining an area of approximately

560 km^2 (Figure 1). Tributaries of the North Fork Holston River in the region of analysis include Laurel Creek and Lick Creek. Both the Holston and the nearby Clinch River systems within the Tennessee River System had historically diverse freshwater mussel communities, and these rivers have been impacted by point-source pollutants (Brown 2004). Researchers have extensively studied the molluscan fauna of this region (Brown et al. 2005; Jones et al. 2001; Neves and Widlak 1988), therefore recent detailed surveys are available for the North Fork Holston River.

Data

Abundance data for fifteen species of freshwater mussels at sampling points along the upper North Fork Holston River and its tributaries in Virginia were obtained from the Department of Wildlife and Fisheries at Virginia Polytechnic Institute and State University (Table 1). Data were aggregated to catchments (Figure 1) developed at the 1:100,000 scale for NHDPlus (Horizon Systems 2006) to facilitate landscape analysis in a geographic information system (GIS), and abundance data were converted to presence and absence for ease of analysis and interpretation (Table 2). Presence was assigned to catchments containing at least one point with a recorded abundance greater than zero. Absence was assigned when the species was not located at any survey point within the catchment. Environmental data layers were converted to a common projection in a GIS. Variables were calculated for the catchment, a 100-m riparian buffer around the stream segment within the catchment, and the catchments upstream of the stream segment comprising the cumulative area for the catchment using ArcView 3.2 and 9.1 (ESRI 2000, 2002) (Figure 2, Table 2). Mussel data were available for thirty catchments averaging 3.89 km^2 in area and ranging from a minimum of 0.16 to 12.22 km^2. The 1:100,000 scale stream segments making up these catchments varied in length from 0.27 to 7.06 km in length, with an average length of 3.06 km. A complete listing of data sources and relevant GIS analysis methods used to derive environmental variables may be found in Appendix A.

Statistical analyses

The degree of clustering of catchments containing mussel data was determined using the Nearest Neighbor Index because clustering indicates that mussel data are not distributed randomly across the region of interest. Correlations among variables were analyzed in the R programming language (2007) and a subset of variables were identified omitting correlations

greater than 0.8 (Thieling 2006). Forward logistic regression was used to analyze presence and absence of mussel species using the reduced subset of independent variables with a p-value of less than 0.05 to enter the model. Logistic regression results provide an estimate of the probability of occurrence. This was converted to presence or absence for each species by classifying the species as present in catchments with a probability of species occurrence greater than 0.5 and absent for smaller probabilities of occurrence. A nonparametric Wilcoxon rank sum test was used to compare values in present and absent catchments for variables included in species models. The exact Wilcoxon test was used to adjust for a small sample size (Zar 1998).

Additional models of freshwater mussel occurrences were based on cross-validated classification trees (CT). Classification trees are non-parametric models with fewer assumptions than traditional statistical techniques; these models also predict a class such as presence or absence directly rather than calculating probability of occurrence (Breiman et al. 1984). Trees perform well with categorical and continuous variables and are easily interpretable, therefore they have become valuable for the exploration of environmental datasets (Breiman et al. 1984; De'ath and Fabricius 2000). The optimal tree can be selected from models created using a cross validation procedure, in which data are repeatedly split into k partitions, the model is created on k-1 partitions, and the model is tested on the kth partition (Breiman et al. 1984). Cross-validated tree models were created using 10 partitions and 50 iterations for predictions of presence and absence of each mussel species at the catchment scale using all environmental variables. Cross-validation error estimates (CV error), which are calculated for trees by averaging error rates over the total number of partitions, were used to pick the optimal model for each species (within one standard error of the minimum CV error) and to determine if models were robust to cross-validation (CV error < 1). Variables that differed between logistic regression and CT models were evaluated for correlation. Model evaluation statistics for logistic regression models and cross-validated trees were also calculated in R.

Moran's I was used to measure spatial autocorrelation for variables used in splits for cross-validated trees. The Getis-Ord General G statistic was computed for spatially autocorrelated variables to identify whether high or low values clustered. All spatial analyses were conducted in ArcGIS 9.2 (ESRI 2002).

Results

Available mollusk data were significantly clustered based on the Nearest Neighbor Index (Table 3). Forward logistic regression using the variable subsets failed to converge because of complete separation of the classes, although parameter estimates were calculated for thirteen of the fifteen models (Table 4). Logistic regression and CT models failed for *Elliptio dilatata* and *Strophitus undulatus* because occurrence data were skewed to one class, either presence or absence (Tables 4 and 5). Wilcoxon tests indicated that for some models, the forward-selection process incorporated variables that are not independently significantly different between classes (Table 4). Catchments in which the species were present differed significantly from absence catchments for all the variables included in the models for *Fusconaia barnesiana, Lasmigona costata, Pleurobema oviforme, Ptychobranchus fasciolaris,* and *Villosa iris* (Table 4). When considered independently, differences were insignificant for at least one variable in the logistic regression models for *Villosa v. vanuxemensis, Ptychobranchus subtentum, Medionidus conradicus, Lexingtonia dolabelloides, Lampsilis ovata,* and *Fucsonaia cor* (Table 4). Many variables used in analyses were significantly clustered (Table 6).

Classification trees created for the mussels *Fusconaia barnesiana*, *Lexingtonia dolabelloides*, *Lampsilis ovata, Ptychobranchus subtentum*, and *Lasmigona costata* performed poorly on cross-validation (CV error > 1) (Table 5). Trees for the mussels *Ptychobranchus fasciolaris*, *Medionidus conradicus*, and *Actinonaias pectorosa* identified cumulative area as the only variable differentiating between predictions of presence and absence Table 5. Presences and absences of *Pleurobema oviforme* were best separated by the maximum slope in the riparian corridor, *Villosa v. vanuxemensis* was best separated by the cumulative proportion of shrubland, and *Fusconaia cor* and *Lampsilis fasciola* were separated by minimum elevation in the catchment, and *Villosa iris* was split by the cumulative density of the population (Table 5). Many of the variables used in splits in classification trees were significantly clustered (Table 6).

Discussion

Model interpretation and limitations

Independent data are not available to test these species models. However, the overall predictive ability of the models can be inferred from the results. Robust models were identified

as those with CV error < 1, but the cross-validation procedure itself is robust only when the data adequately represent the variable of interest. In this case, the sampled catchments are a small subset of the ranges of the mussel species. Models created with more environmental variables than dependent observations are prone to over-fitting the data with spurious relationships, and this is particularly likely given the geographic restriction of the catchments. Classification tree models with CV error > 1 were created for eight of thirteen mussels we intended to model, but in all cases the best model contained only one split, and all of the variables defining model splits were significantly spatially autocorrelated. These models should therefore be interpreted with caution until independent data are available to evaluate these findings.

Although the results of these models may largely be spurious, some of the variables identified should be considered in future studies. For instance, the proportion of grassland in the catchment was identified in the logistic regression model for *Villosa iris*. A previous study by Morris and Corkum (1996) found differences in mussel communities between streams with forested riparian zones and streams with grassy riparian zones, indicating that changes in the terrestrial landscape related to greater amounts of grassy areas may affect freshwater mollusks. The frequent occurrence of cumulative area in models confirms previous findings that mussel communities change as stream size and basin area increase (Watters 1992; Myers-Kinzie et al. 2001). Although the lack of independent data limits confidence in the specific models created in this study, the inclusion in our models of variables identified by previous studies indicates that catchment-scale and cumulative land cover variables may be important in predicting occurrences of freshwater mollusks.

Future study considerations

Although the results from these models are not conclusive, models developed at a catchment scale aimed at predicting occurrences of freshwater mussels may be more appropriate than models based on larger or smaller sampling areas for several reasons. Catchment analyses easily incorporate data obtained at the scale of stream segments, which can be easily searched for existing mussel populations. Predictive studies based on point or quadrat sampling may be explored for modeling freshwater mussels because studies based on smaller sampling units would allow more precise incorporation of immediate habitat conditions, but several factors contribute to decreased suitability and utility of models developed for sampling units smaller

than the watershed-based catchments used as sampling units in this study. Researchers studying freshwater mussels have questioned the importance of microhabitat variables for prediction occurrences (Strayer and Ralley 1993), incorporating cumulative, riparian, and watershed variables into models for smaller sampling units is computationally intensive, and the potential for increasing the resolution of mussel distributions is likely outweighed by the need for more precise species and environmental data.

Analyses based on larger, hydrologically relevant sampling units, such as subwatersheds (see Chapter 2) and 8-digit hydrologic units (HUs) (Sowa et al. 2005), have also been conducted for freshwater mussels. Freshwater mussel surveys or occurrences are normally conducted at a scale finer than subwatersheds, which can incorporate stream segments from more than one stream or river. Aggregating survey or occurrence data to this scale of analysis limits the researcher's ability to clearly and correctly associate mussel data with the environmental variables influencing its site of occurrence. For instance, mussels located in a tributary near its mouth at a large river would be assigned to a subwatershed containing the large river; therefore, the occurrence would erroneously be associated with all of the variables directly associated with the subwatershed and cumulative variables associated with upstream tributaries of the large river (see Chapter 2).

Because of the difficulties associated with sampling environmental conditions and species at smaller sampling units or of summarizing data from small or variable sampling units to larger sampling units, additional research should explore the suitability of catchment analyses for predicting occurrences of freshwater mussels. Sampling for mussels in all potential habitats by segment is relatively easy, as the average segment length in this region of 3 km is much shorter than the total stream length with potential mussel habitat in a subwatershed. Sampling at the segment scale would provide researchers with data appropriate for analyzing distributions at the catchment scale in a geographic information system (GIS). Sampling performed in a subset of catchments representing the range of environmental and anthropogenic characteristics of a drainage region could be initiated with cooperation from universities, consultants, non-profits, and state and federal agencies. By establishing uniform sampling protocols for specific mussel species at a regional level, a freshwater mussel dataset can be developed that is valuable for estimating probability of detection and deriving predictive models that incorporate both local- and landscape-level effects on species of interest. Information on minimum sampling efforts

required to detect individuals of a species, temporal variation in sampling success, and distributions of species across their geographic ranges can provide a basis for future studies by establishing minimum sampling effort for species detections at the segment scale, defining relationships between season and sampling success, and by identifying factors that may affect species occurrences.

References

Arbuckle, K.E., Downing, J.A., 2002. Freshwater mussel abundance and species richness: GIS relationships with watershed land use and geology. Canadian Journal of Fisheries and Aquatic Sciences 59, 310-316.

Bambarger, R., 2006. Freshwater mussel communities of the Florida Parishes, Louisiana: the importance of spatial scale, In Biological Sciences. p. 38. Louisiana State University and Agricultural and Mechanical College, Baton Rouge, LA.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and regression trees. Chapman & Hall, Boca Raton, FL.

Brim-Box, J., Dorazio, R.M., Liddell, W.D., 2002. Relationships between streambed substrate characteristics and freshwater mussels (Bivalvia: Unionidae) in coastal plain streams. Journal of the North American Benthological Society 21, 253-260.

Brown, M., 2004. Geochemical and Taphonomic Signatures of Freshwater Mussel Shells as Evidence of Mercury-Related Extirpations in the North Fork Holston River, Virginia, In Department of Geosciences. p. 70. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Brown, M.E., Kowalewski, M., Neves, R.J., Cherry, D.S., Schreiber, M.E., 2005. Freshwater mussel shells as environmental chronicles: geochemical and taphonomic signatures of mercury-related extirpations in the North Fork Holston River, Virginia. Environmental Science & Technology 39, 1455.

De'ath, G., 2007. Mvpart: Multivariate partitioning. R package version 1.2-6.

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

Di Maio, J., Corkum, L., 1995. Relationship between the spatial distribution of freshwater mussels (Bivalvia: Unionidae) and the hydrological variability of rivers. Canadian Journal of Zoology 73, 663-671.

Diamond, J.M., Serveiss, V.B., 2001. Identifying sources of stress to native aquatic fauna using a watershed ecological risk assessment framework. Environmental Science and Technology 35, 4711-4718.

Hallac, D.E., Marsden, J.E., 2000. Differences in tolerance to and recovery from zebra mussel (*Dreissena polymorpha*) fouling by *Elliptio complanata* and *Lampsilis radiata*. Canadian Journal of Zoology 78, 161-166.

Horizon Systems, 2006. NHDPlus. URL *http://www.horizon-systems.com/nhdplus/*

Hornbach, D.J., 2001. Macrohabitat factors influencing the distribution of naiads in the St. Croix River, Minnesota and Wisconsin, USA, In Ecology and evolution of the freshwater mussels Unionoida. eds G. Bauer, K. Wachtler, pp. 213-230. Springer-Verlag, Berlin, Germany.

Jobson, H.E., 1996. Prediction of traveltime and longitudinal dispersion in rivers and streams. U.S. Geological Survey Water Resources Investigations Report 96-4013, U.S. Geological Survey, Washington, DC.

Jones, J.W., Neves, R.J., Patterson, M.A., Good, C.R., DiVittorio, A., 2001. A status survey of freshwater mussel populations in the Upper Clinch River, Tazewell County, Virginia. Banisteria 17, 20-30.

Lydeard, C., Clark, S.A., Perez, K.E., Cowie, R.H., Ponder, W.F., Bogan, A.E., Bouchet, P., Gargominy, O., Cummings, K.S., Frest, T.J., Herbert, D.G., Hershler, R., Roth, B., Seddon, M., Strong, E.E., Thompson, F.G., 2004. The global decline of nonmarine mollusks. Bioscience 54, 321-330.

Morris, T.J., Corkum, L.D., 1996. Assemblage structure of freshwater mussels (Bivalvia: Unionidae) in river with grassy and forested riparian zones. Journal of the North American Benthological Society 15, 576-586.

Myers-Kinzie, M.L., Wente, S.P., Spacie, A., 2001. Occurrence and distribution of freshwater mussels in small streams of Tippecanoe County, Indiana. Proceedings of the Indiana Academy of Science 110, 141-150.

Negus, C.L., 1966. A quantitative study of growth and production of unionid mussels in the River Thames at Reading. The Journal of Animal Ecology 35, 513-532.

Neves, R.J., Widlak, J.C., 1988. Occurrence of glochidia in stream drift and on fishes of the upper North Fork Holston River, Virginia. American Midland Naturalist 119, 111.

Poole, K.E., Downing, J.A., 2004. Relationship of declining mussel biodiversity to stream-reach and watershed characteristics in an agricultural landscape. Journal of the North American Benthological Society 23, 114-125.

R Development Core Team, 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Research Triangle Institute, 2001. The National water pollution control assessment model (NWPCAM) version 2, Draft report prepared for the U.S. Environmental Protection Agency, Research Triangle Institute, Research Triangle Park, North Carolina.

Schneider, D.W., Ellis, C.D., Cummings, K.S., 1998. A Transportation Model Assessment of the Risk to Native Mussel Communities from Zebra Mussel Spread. Conservation Biology 12, 788-800.

Strayer, D.L., 1981. Notes in the microhabitats of unionid mussels in some Michigan streams. American Midland Naturalist 106, 411.

Strayer, D.L., 1983. The effects of surface geology and stream size on freshwater mussel (Bivalvia, Unionidae) distribution in southeastern Michigan, USA. Freshwater Biology 13, 253-264.

Strayer, D.L., Beighley, R.E., Thompson, L.C., Brooks, S., Nilsson, C., Pinay, G., Naiman, R.J., 2003. Effects of land cover on stream ecosystems: Roles of empirical models and scaling issues. Ecosystems 6, 407-423.

Strayer, D.L., Caraco, N.F., Cole, J.J., Findlay, S., Pace, M.L., 1999. Transformation of freshwater ecosystems by bivalves: a case study of zebra mussels in the Hudson River. Bioscience 49, 19-27.

Strayer, D.L., Ralley, J., 1993. Microhabitat use by an assemblage of stream-dwelling Unionaceans (Bivalvia), including two rare species of *Alasmidonta*. Journal of the North American Benthological Society 12, 247-258.

Thieling, T.M., 2006. Assessment and predictive model for brook trout (*Salvelinus fontinalis*) population status in the eastern United States, In Biology. p. 59. James Madison University, Harrisonburg, Virginia.

Turgeon, D.D., Bogan, A.E., Coan, E.V., Emerson, W.K., Lyons, W.G., Pratt, W.L., Roper, E.F.E., Scheltema, A., Thompson, F.G., Williams, J.D., 1988. Common and Scientific Names of

Aquatic Invertebrates from the United States and Canada: Mollusks. American Fisheries Society, Bethesda, Maryland.

US FWS (U.S. Fish and Wildlife Service), 2006. Species Information: Threatened and Endangered Animals and Plants. http://www.fws.gov/endangered/wildlife.html (December 2006).

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

Vaughn, C.C., Gido, K.B., Spooner, D.E., 2004. Ecosystem processes performed by unionid mussels in stream mesocosms: species roles and effects of abundance. Hydrobiologia 527, 35.

Vaughn, C.C., Hakenkamp, C.C., 2001. The functional role of burrowing bivalves in freshwater ecosystems. Freshwater Biology 46, 1431-1446.

Watters, G.T., 1992. Unionids, fishes, and the species-area curve. Journal of Biogeography 19, 481-490.

Watters, G.T., 1996. Small dams as barriers to freshwater mussels (Bivalvia, Unionoida) and their hosts. Biological Conservation 75, 79-85.

Watters, G.T., Menker, T., Thomas, S., Kuehnl, K., 2005. Host identifications or confirmations. Ellipsaria 7, 11-12.

Zar, J.H., 1998. Biostatistical Analysis, 4th edn. Prentice-Hall, Inc., Upper Saddle River, New Jersey.

Tables

Table 1. Freshwater mussel species considered for model development.

Abbreviation	Scientific Name	Common Name
APEC	*Actinonaias pectorosa*	pheasantshell
EDIL	*Elliptio dilatata*	spike
FBARN	*Fusconaia barnesiana*	Tennessee pigtoe
FCOR	*Fusconaia cor*	shiny pigtoe
LFAS	*Lampsilis fasciola*	wavyrayed lampmussel
LOVA	*Lampsilis ovata*	pocketbook
LDOL	*Lexingtonia dolabelloides*	slabside pearlymussel
LCOST	*Lasmigona costata*	flutedshell
MCON	*Medionidus conradicus*	Cumberland moccasinshell
POVI	*Pleurobema oviforme*	Tennessee clubshell
PFAS	*Ptychobranchus fasciolaris*	kidneyshell
PSUB	*Ptychobranchus subtentum*	fluted kidneyshell
SUND	*Strophitus undulatus*	creeper
VIRS	*Villosa iris*	rainbow
VVAN	*Villosa v. vanuxemensis*	mountain creekshell

Table 2. Dependent and independent variables considered in models.

Variable	Definition	Scale or Resolution	Data Source*	Relevant Time Period
VIRS	Presence (23) and absence (7) of *Villosa iris* by catchment	stream reach	Survey data	1998-2006
VVAN	Presence (24) and absence (6) of *V. v. vanuxemensis* by catchment	stream reach	Survey data	1998-2006
SUND	Presence (1) and absence (29) of *Strophitus undulatus* by catchment	stream reach	Survey data	1998-2006
MCON	Presence (17) and absence (13) of *Medionidus conradicus*	stream reach	Survey data	1998-2006
LOVA	Presence (4) and absence (26) of *Lampsilis ovata* by catchment	stream reach	Survey data	1998-2006
LFAS	Presence (11) and absence (19) of *L. fasciola* by catchment	stream reach	Survey data	1998-2006
FBARN	Presence (9) and absence (21) of *Fusconaia barnesiana* by catchment	stream reach	Survey data	1998-2006
FCOR	Presence (6) and absence (24) of *F. cor* by catchment	stream reach	Survey data	1998-2006
LDOL	Presence (10) and absence (20) of *Lexingtonia dolabelloides* by catchment	stream reach	Survey data	1998-2006
LCOST	Presence (4) and absence (26) of *Lasmigona costata* by catchment	stream reach	Survey data	1998-2006
PFAS	Presence (19) and absence (11) of *Ptychobranchus fasciolaris* by catchment	stream reach	Survey data	1998-2006
PSUB	Presence (5) and absence (25) of *P. subtentum* by catchment	stream reach	Survey data	1998-2006
POVI	Presence (14) and absence (16) of *Pleurobema oviforme* by catchment	stream reach	Survey data	1998-2006
EDIL	Presence (2) and absence (28) of *Elliptio dilatata* by catchment	stream reach	Survey data	1998-2006
APEC	Presence (12) and absence (18) of *Actinonaias pectorosa* by catchment	stream reach	Survey data	1998-2006
AREASQKM	Area of the catchment (sq km)	1:100,000	NLCD 2001	1999-2001
L_FLOW	Length of stream centerlines (km)	1:100,000	NLCD 2001	1999-2001
DENPOP	Population density (people*km^{-2})	1:100,000	ESRI	2000
DENRD	Density of roads in the catchment (rds*km^{-2})	1:100,000	ESRI	2000
DENINT	Density of road-stream intersections (intersections*km^{-2})	1:100,000	ESRI	2000
PROPINT	Proportion of road-stream intersections (rd*stream km^{-1})	1:100,000	ESRI	2000
DEN303D	Density of 303-D listed streams (km*km^{-2})	1:100,000	US EPA	2002
PROP303	Proportion of 303-D streams (km*stream km^{-1})	1:100,000	US EPA	2002
DENFLOW	Density of streams (stream km*km^{-2})	1:100,000	NHDPlus	2005
WET	Proportion of wetlands in the catchment (%)	30m	NLCD 2001	1999-2001
SHB	Proportion of shrublands in the catchment (%)	30m	NLCD 2001	1999-2001
ODV	Proportion of open development in the catchment (%)	30m	NLCD 2001	1999-2001
MXF	Proportion of mixed forest in the catchment (%)	30m	NLCD 2001	1999-2001
MH	Proportion of medium- and high-intensity development in the catchment (%)	30m	NLCD 2001	1999-2001
MDV	Proportion of medium-intensity in the catchment (%)	30m	NLCD 2001	1999-2001
LO	Proportion of open- and low-intensity development in the catchment (%)	30m	NLCD 2001	1999-2001

Variable	Definition	Scale or Resolution	Data Source*	Relevant Time Period
LDV	Proportion of low-intensity development in the catchment (%)	30m	NLCD 2001	1999-2001
HUM	Proportion of human land use in the catchment (%)	30m	NLCD 2001	1999-2001
HDV	Proportion of high-intensity development in the catchment (%)	30m	NLCD 2001	1999-2001
HAY	Proportion of hay in the catchment (%)	30m	NLCD 2001	1999-2001
H2O	Proportion of water in the catchment (%)	30m	NLCD 2001	1999-2001
GRS	Proportion of grasslands in the catchment (%)	30m	NLCD 2001	1999-2001
FOR	Proportion of forest in the catchment (%)	30m	NLCD 2001	1999-2001
EVG	Proportion of evergreen forest in the catchment (%)	30m	NLCD 2001	1999-2001
DEV	Proportion of developed lands in the catchment (%)	30m	NLCD 2001	1999-2001
DEC	Proportion of deciduous forest in the catchment (%)	30m	NLCD 2001	1999-2001
CRP	Proportion of croplands in the catchment (%)	30m	NLCD 2001	1999-2001
BAR	Proportion of barren lands in the catchment (%)	30m	NLCD 2001	1999-2001
AG	Proportion of agriculture in the catchment (%)	30m	NLCD 2001	1999-2001
ELEV_MIN	Minimum elevation in the catchment (m)	30m	NED	Varies
ELEV_MAX	Maximum elevation in the catchment (m)	30m	NED	Varies
ELEV_AVG	Average elevation in the catchment (m)	30m	NED	Varies
SLOPE_MIN	Minimum slope in the catchment (%)	30m	NED	varies
SLOPE_MAX	Maximum slope in the catchment (%)	30m	NED	varies
SLOPE_AVG	Average slope in the catchment (%)	30m	NED	varies
CUMDRAINAG	Cumulative drainage area (km^2)	1:100,000	NHDPlus	2005
MAFLOWU	Mean annual flow calculated using the unit runoff method (Rsch. Tri. Inst. 1996) (cfs)	1:100,000	NHDPlus	2005
MAVELU	Mean annual velocity calculated using the Jobson method (Jobson 1996) (fps)	1:100,000	NHDPlus	2005
MAXELEVSMO	Maximum elevation along stream centerline (smoothed) (m)	1:100,000	NHDPlus	2005
MINELEVSMO	Minimum elevation along stream centerline (smoothed) (m)	1:100,000	NHDPlus	2005
SLOPE	Slope along stream centerline (%)	1:100,000	NHDPlus	2005
STREAMLEVE	Stream level	1:100,000	NHDPlus	2005
STREAMORDE	Strahler stream order	1:100,000	NHDPlus	2005
CUAREASQK	Cumulative area (upstream of the catchment) (sq km)	1:100,000	NHDPlus	2005
CUDENPOP	Cumulative population density (people*km^{-2})	1:100,000	ESRI	2000
CUDENRD	Cumulative density of roads (rd*km^{-2})	1:100,000	ESRI	2000
CUDENINTAR	Cumulative density of road-stream intersections (intersections*km^{-2})	1:100,000	ESRI	2000
CUDEN303	Cumulative density of 303-D listed streams (km*km^{-2})	1:100,000	US EPA	2002
CUWET	Cumulative proportion of wetlands (%)	30m	NLCD 2001	1999-2001

Variable	Definition	Scale or Resolution	Data Source*	Relevant Time Period
CUSHB	Cumulative proportion of shrublands (%)	30m	NLCD 2001	1999-2001
CUWET	Cumulative proportion of wetlands (%)	30m	NLCD 2001	1999-2001
CUSHB	Cumulative proportion of shrublands (%)	30m	NLCD 2001	1999-2001
CUODV	Cumulative proportion of open development (%)	30m	NLCD 2001	1999-2001
CUMXF	Cumulative proportion of mixed forest (%)	30m	NLCD 2001	1999-2001
CUMH	Cumulative proportion of medium- and high-intensity development (%)	30m	NLCD 2001	1999-2001
CUMDV	Cumulative proportion of medium-intensity (%)	30m	NLCD 2001	1999-2001
CULO	Cumulative proportion of open- and low-intensity development (%)	30m	NLCD 2001	1999-2001
CULDV	Cumulative proportion of low-intensity development (%)	30m	NLCD 2001	1999-2001
CUHUM	Cumulative proportion of human land use (%)	30m	NLCD 2001	1999-2001
CUHDV	Cumulative proportion of high-intensity development (%)	30m	NLCD 2001	1999-2001
CUHAY	Cumulative proportion of hay (%)	30m	NLCD 2001	1999-2001
CUH2O	Cumulative proportion of water (%)	30m	NLCD 2001	1999-2001
CUGRS	Cumulative proportion of grasslands (%)	30m	NLCD 2001	1999-2001
CUFOR	Cumulative proportion of forest (%)	30m	NLCD 2001	1999-2001
CUEVG	Cumulative proportion of evergreen forest (%)	30m	NLCD 2001	1999-2001
CUDEV	Cumulative proportion of developed lands (%)	30m	NLCD 2001	1999-2001
CUDEC	Cumulative proportion of deciduous forest (%)	30m	NLCD 2001	1999-2001
CUCRP	Cumulative proportion of croplands (%)	30m	NLCD 2001	1999-2001
CUBAR	Cumulative proportion of barren lands (%)	30m	NLCD 2001	1999-2001
CUAG	Cumulative proportion of agriculture (%)	30m	NLCD 2001	1999-2001
WET_100	Proportion of wetlands in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
SHB100	Proportion of shrublands in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
ODV100	Proportion of open development in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
MXF100	Proportion of mixed forest in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
MH100	Proportion of medium- and high-intensity development in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
MDV100	Proportion of medium-intensity in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
LO100	Proportion of open- and low-intensity development in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
LDV100	Proportion of low-intensity development in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
HDV100	Proportion of human land use in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
HAY100	Proportion of high-intensity development in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
H2O100	Proportion of hay in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
GRS100	Proportion of water in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
FOR100	Proportion of grasslands in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001

Variable	Definition	Scale or Resolution	Data Source*	Relevant Time Period
EVG100	Proportion of evergreen forest in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
DEV100	Proportion of developed lands in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
DEC100	Proportion of deciduous forest in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
BAR100	Proportion of barren lands in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
AG100	Proportion of agriculture in a 100-m riparian buffer (%)	30m	NLCD 2001	1999-2001
MINSLP100	Minimum slope in a 100-m riparian buffer (%)	30m	NED	varies
MAXSLP100	Maximum slope in a 100-m riparian buffer (%)	30m	NED	varies
AVGSLP100	Average slope in a 100-m riparian buffer (%)	30m	NED	varies
MINELEV100	Minimum elevation in a 100-m riparian buffer (m)	30m	NED	varies
MAXELEV100	Maximum elevation in a 100-m riparian buffer (m)	30m	NED	varies
AVGELEV100	Average elevation in a 100-m riparian buffer (m)	30m	NED	varies

*The mussel dataset was compiled by Brett Ostby of Virginia Polytechnic Institute and State University from surveys conducted by himself, Jones and Neves (2001), and the Virginia Department of Game and Inland Fisheries. The National Elevation Dataset (NED) for the eastern United States was obtained directly from the USGS. National Land Cover Database (NLCD) land use grids were downloaded from the Multi-Resolution Land Characteristics Consortium (MRLC) by mapping zone. National Hydrography Dataset Plus (NHDPlus) data were downloaded from Horizon Systems, Inc. U.S. Detailed Streets and U.S. Census Block Centroid Populations were obtained from Environmental Systems Research Institute, Inc. (ESRI) data CDs provided with ArcGIS software.

Table 3. Nearest neighbor index classification and scores for sampled catchments (N=30).

	Catchments
NNI	clustered
z-score	-3.3
p-value	<0.01

Table 4. Variables and coefficients (BETA) for logistic regression models for freshwater mussel species. P-VALUE indicates level of parameter significance. Significance of Wilcoxon exact tests are given under WILC. Measures of model evaluation include percent accuracy (ACC), sensitivity (SENS), and specificity (SPEC) and the false-positive (F_POS) and false-negative (F_NEG) rates.

SPECIES	VARIABLES	BETA	P-VALUE	WILC	ACC	SENS	SPEC	F_POS	F_NEG
APEC	LDV	2790.3	<0.0001	0.0117	90	83.3	94.4	9.1	10.5
	CUAREASQK	0.2046	<0.0001	<0.0001					
EDIL	none								
FBARN	CUBAR	-1034.3	0.04709	0.078	60	11.1	81	80	32
FCOR	CUCRP	14387.3	0.0002	<0.0001	96.7	100	95.8	14.3	0
	PROP303	-14.9632	0.0075	0.5569					
LFAS	LDV	-12.0986	0.0002	0.27	83.3	78.9	90.9	6.3	28.6
	CUAREASQK	0.1145	0.0013	<0.0001					
	SLOPE	311.1	0.0008	0.0305					
LOVA	CUAREASQK	0.1968	0.0007	0.0028	83.3	78.9	90.9	6.3	28.6
	DENPOP	-0.2711	0.0092	0.4185					
LCOST	H2O	33727.7	<0.0001	0.0138	90	50	96.2	33.3	7
	BAR	1196	0.0013	0.0189					
LDOL	H2O	-3058.9	0.0011	1	86.7	80	90	20	10
	ELEV_AVG	-0.3097	0.0177	0.0004					
	STREAMORDE	29.1841	0.0172	0.0005					
	PROP303	-33.827	0.0014	0.272					
MCON	SLOPE_MIN	21.55	<0.0001	0.492	93.3	88.2	100	0	13.3
	CUAREASQK	0.094	0.0001	<0.0001					
POVI	STREAMLEVE	-7.2815	0.0077	0.0092	93.3	100	87.5	12.5	0
	LDV	-14327.1	0.0245	0.0189					
	STREAMORDE	2.2944	0.0041	0.0233					
	CUAREASQK	0.0889	0.0063	0.0275					
PFAS	CUAREASQK	0.2104	<0.0001	<0.0001	90	81.8	94.7	10	10
	LDV	3007.7	<0.0001	0.0021					
PSUB	WET	4211.9	0.0002	0.134	90	60	96	25	7.7
	BAR	4382	0.0028	0.0969					
SUND	none								
VIRS	CUSHB	-828.8	0.0003	0.0033	90	91.3	85.7	4.5	25
	ELEV_AVG	-0.3185	0.0055	0.0186					
	GRS	1254.6	0.0055	0.0017					
VVAN	LDV	-1738.3	0.0001	0.8522	90	91.7	83.3	4.3	28.6
	CU_SHB	-1223.3	<0.0001	<0.0001					

Table 5. Variables and cutoffs used to predict presence (PRES NODE) from classification trees developed for freshwater mussel species. P-VALUE indicates level of parameter significance. Cross-validation error (CV ERROR) indicates how well the model performed under cross-validation. Measures of model evaluation include kappa, percent accuracy (ACC), sensitivity (SENS), and specificity (SPEC) and the false-positive (F_POS) and false-negative (F_NEG) rates.

SPECIES	VARIABLE	PRES NODE	CV ERROR	ACC	SENS	SPEC	F_POS	F_NEG	KAPPA
APEC	CUAREASQK	>=243.2	0.203	96.67	91.7	100	0	5.26	92.96
EDIL	NO SPLITS POSSIBLE			93.33	0	NA	N/A	6.67	0
FBARN	LO100	<0.0955	1.61	96.67	88.9	100	0	4.55	91.8
	SHB100	<0.02615							
	HUM	>=0.494							
FCOR	ELEV_MIN	<540.5	0.207	96.67	100	95.83	14.29	0	90.2
LFAS	ELEV_MIN	<601.5	0.838	90	89.5	90.91	5.56	16.67	78.87
LOVA	CUAREASQK	>=337.4	1.53	100	100	100	0	0	100
	ODV	<0.0682							
LCOST	CRP	>=0.03425	1.95	93.33	75	96.15	25	3.85	71.15
LDOL	GRS	>=0.004	1.19	93.33	100	90	16.67	0	85.71
MCON	CUAREASQK	>=119.2	0.185	93.33	88.2	100	0	13.33	86.67
POVI	MAXSLP100	>=100.5	0.786	83.33	64.3	100	0	23.81	65.75
PFAS	CUAREASQK	>=257.4	0.182	96.67	90.9	100	0	5	92.68
PSUB	BAR	>=0.00955	1.18	93.33	60	100	0	7.41	71.43
SUND	NO SPLITS POSSIBLE			96.67	0	100	N/A	3.33	0
VIRS	CUDENPOP	>=2.515	0.64	96.67	100	85.71	4.17	0	90.2
VVAN	CUSHB	<0.06065	0.673	96.67	100	83.33	4	0	88.89

Table 6. Spatial autocorrelation and cluster analysis of variables used in classification trees. Clusters of high or low values are significant at the α=0.05 level.

	Moran's I	z-score	Pattern	General G z-score	Pattern
BAR	0.03	2.5	clustered	4.8	highs cluster
CRP	-0.04	0.1	random	N/A	
DENPOP	0.04	1.9	random	N/A	
ELEV_AVG	0.36	8.1	clustered	-3.7	lows cluster
ELEV_MIN	0.39	9.0	clustered	-3.5	lows cluster
GRS	0.26	5.8	clustered	6.2	highs cluster
H2O	-0.05	-0.6	random	N/A	
HUM	0.03	1.4	random	N/A	
LDV	0.09	3.8	clustered	2.6	highs cluster
ODV	-0.04	0	random	N/A	
PROP303	0.02	1.3	random	N/A	
SLOPE	-0.09	-2	random	N/A	
SLOPE_MIN	-0.07	-0.8	random	N/A	
STREAMORDE	-0.08	2.4	clustered	2.2	highs cluster
STREAMLEVE	0.01	1	random	N/A	
WET	0.08	2.6	clustered	2.8	highs cluster
CUAREASQK	0.26	5.3	clustered	4.7	highs cluster
CUBAR	-0.01	0.5	random	N/A	
CUCRP	0.2	4.9	clustered	2	highs cluster
CUDENPOP	0.13	3.6	clustered	2.3	highs cluster
CUGRS	0.34	7.1	clustered	4.5	highs cluster
CUSHB	0.04	1.6	random	N/A	
LO100	-0.06	-0.5	random	N/A	
MAXSLP100	-0.06	-0.4	random	N/A	
SHB100	-0.03	0.1	random	N/A	
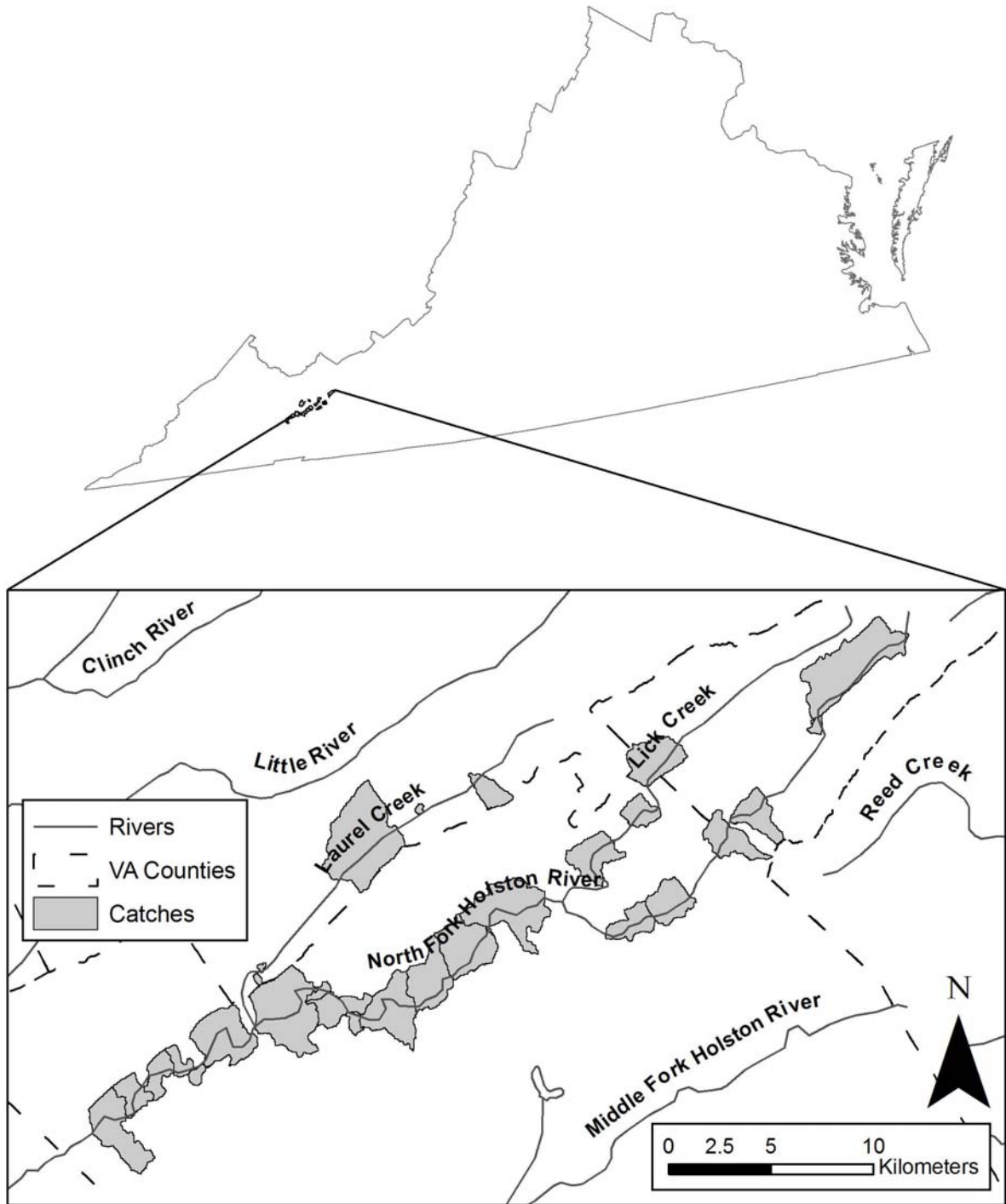
Figure 1. The study area includes 30 catchments (catches) along the North Fork Holston River and its major tributaries, Laurel Creek and Lick Creek, in western Virginia.
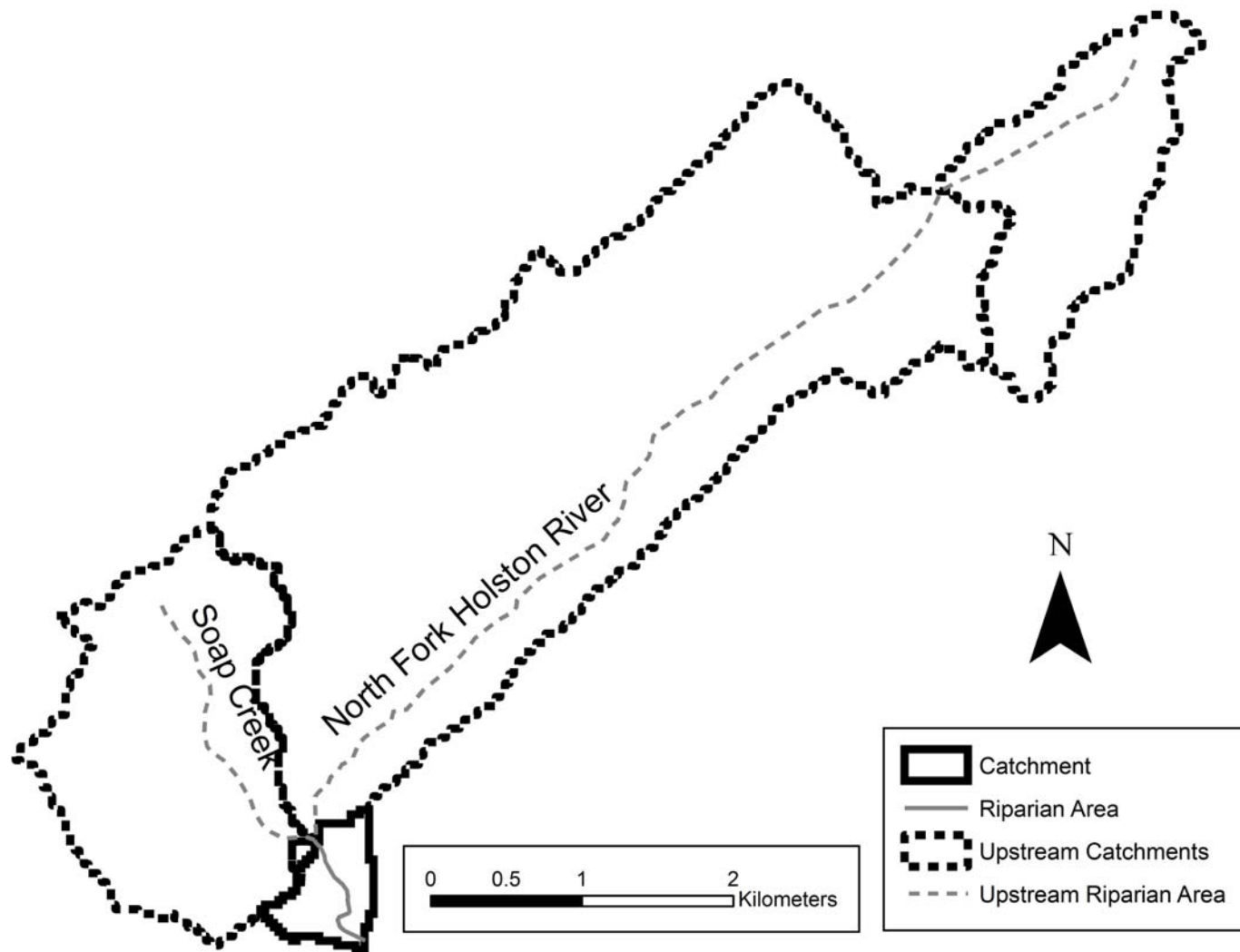
Figure 2. Upstream calculations and riparian area calculations for catchments within the study area.

Appendices

A. Data Layers

Projection and Datum

All vector data layers were clipped to the study area boundary and converted to the North American Datum of 1983. Raster data layers were clipped to a buffered study area boundary. All data layers were displayed in the Albers Conical Equal Area projection for data analysis in order to minimize distortion of area and obtain relatively accurate measures of area across the study region. The minimum unit map scale for the project is equivalent to 1:100,000.

Variable Development

Anthropogenic and environmental variables were selected that represent potential threats or limiting factors to freshwater mussels and could be represented in a GIS (Appendix C). Only variables that were at a resolution and extent that allowed values to be summarized for each subwatershed throughout at least one upper drainage regions within the study area were developed. Metrics were calculated by subwatershed and by riparian corridor within the subwatershed following Theiling (2006). The riparian corridor was defined as being within a distance of 100m from streams, rivers, and lakes represented by the stream area and waterbody layers in the improved National Hydrography Dataset (NHDPlus) (Horizon Systems 2006). This distance is commonly used (Richards et al. 1996; Andersen 2002; McKinney et al. 2002; Thieling 2006) and is larger than the 30m resolution of the remotely sensed environmental variables. Although some legal stream buffer zones are smaller than 100m (Willson & Dorcas 2003), a larger stream corridor is more suitable for the grain of the analyses in this study.

A custom automated GIS routine was used to define all subwatersheds upstream of each subwatershed in the study area. Additional variables were based on the cumulative values of environmental and anthropogenic variables upstream of each subwatershed for variables that could logically be calculated cumulatively (Table 4). ArcGIS 9.1 (ESRI, 2002) was used to develop and summarize all other variables.

Mussel Surveys

Regional mussel experts and state agencies provided freshwater mussel survey data for many sites across the study area (see Chapter 2, Table 1). Additional occurrence data for the species of interest were obtained from museums (Table 2). Surveys with river miles were converted to point or line locations using a script in ArcView 3.2 (ESRI 2000), displayed in ArcMap 9.1 and 9. 2 (ESRI 2002), and exported to shapefiles for conversion into a common projection. All shapefile occurrence or survey data were converted to a common projection to facilitate aggregation into a single dataset.

Presence of a mussel species was defined for model development as any occurrence recorded in surveys conducted since 1985 within 11-, 12-, and 14-level subwatersheds. Choosing a cutoff date creates a defined distinction between historic occurrences and recent occurrences. This year was chosen because surveys conducted after this date would have a high likelihood of detecting individuals that have recolonized areas restored since the Clean Water Act and because earlier surveys proved more

difficult to obtain. A more recent cutoff date was not selected because few mussel surveys that span more than one subwatershed have been conducted across the study area in the recent past. Absences were assigned to subwatersheds along recent survey routes when all surveys covering the subwatershed failed to locate the species. Survey lines that overlapped more than one subwatershed were omitted from analysis.

Presence-only data provided independent test data for model predictions. Presence was assigned to a subwatershed for the test dataset based on any species occurrence recorded since 1985, excluding survey data included in the presence-absence dataset. Presence-only data were assigned to surveyed subwatersheds when available, although the determination of species presence did not always correspond to the survey data.

Zebra Mussels

Zebra mussels are an invasive species that can affect native freshwater mussel communities. Presence and absence of zebra mussel by subwatershed was calculated using a shapefile of zebra mussel occurrences developed by the Nonindigenous Aquatic Species (NAS) Program and downloaded from the National Atlas. This dataset contained occurrences noted from 1988-2007 by several agencies rather than results of a formal survey effort for zebra mussels.

Elevation and Derivatives

Mean, maximum, and minimum elevation and slope were calculated within each watershed from 30m Digital Elevation Models developed by the USGS as part of the National Elevation Dataset. These values were also calculated along flowlines (1:100K stream centerlines) obtained from the NHDPlus dataset using smoothed elevation values (Horizon Systems 2006).

Basin Area

The area of each subwatershed was calculated in ArcGIS 9.1. The area of all upstream subwatersheds was also determined to create measures of the cumulative drainage area and local drainage area for each subwatershed.

Riparian Area

The area of the riparian corridor was calculated in ArcGIS 9.1 using a 100m perimeter around streams, rivers, and lakes represented by the stream area and waterbody layers in the improved National Hydrography Dataset (NHDPlus) (Horizon Systems 2006). The cumulative riparian corridor area was also derived for each subwatershed.

Stream Flow and Discharge

Enhanced NHDPlus grids developed from the National Hydrography Dataset and the National Elevation Dataset by the USGS, US EPA, and government contractors were used to extract min, max, and range of link and Strahler stream order, mean annual flow calculated with the unit runoff method (Research Triangle Institute 2001), and mean annual velocity calculated using the Jobson method (Jobson 1996) by subwatershed (Horizon Systems 2006).

Roads

Roads were represented by ESRI's 1:50,000 map scale U.S. Detailed Streets layer created using 2000 data (ESRI 2002). Road density in kilometers of road per square kilometers of land area was calculated within the riparian corridor and across each subwatershed (Thieling 2006). Point intersections of the road and flowline layers were derived to represent stream crossings and summarized as the density of stream crossings in the subwatershed and in the riparian corridor. Cumulative sums of road-stream intersections and riparian zone intersections for each subwatershed were calculated and cumulative density was determined by dividing by cumulative subwatershed and riparian corridor area measures (Thieling 2006).

Dams

Locations of impoundments, which can alter water temperature and serve as barriers to mussel dispersal, were obtained from the U.S. Army Corps of Engineers National Inventory of Dams (2005). The number of dams per kilometer of land within each subwatershed and the distance to the nearest downstream dam was calculated from this layer. The cumulative proportion of dams per area was also calculated upstream of each subwatershed.

Mining and Metal Operations

Locations of metal and mineral processing plants and mining operations available via the U.S. Geological Survey Minerals Information web page were downloaded from the National Atlas. Variables were calculated based on the density of metal operations and the density of mining operations in the subwatershed, in the riparian corridor, and in the cumulative area including upstream subwatersheds and the cumulative area including the upstream riparian corridor.

Land Use

Land use variables were derived from the 2001 National Land Cover Database (NLCD), which is based on Landsat 7 imagery (USGS 2006). The NLCD land use coverages include 29 land cover classes and originally tested with a single-pixel land cover accuracy of 73-77% (Homer et al. 2004) with an average accuracy of 83.9% estimated using cross-validation techniques (Homer et al. 2007). Individual and combined classes represent variables of interest. Land cover classes were calculated as percentages within each subwatershed, upstream subwatersheds, riparian corridor, and upstream riparian corridor.

Polluted Streams

The proportion of impaired, 303(d)-listed stream kilometers within each subwatershed for the upper Mid-Atlantic and Tennessee drainages was determined. This metric was omitted from models in the Ohio drainage where 303(d) stream miles cannot be determined by subwatershed due to inconsistent data for the state of Ohio (US EPA 2002). Although 303(d) listed streams are only impaired in relation to their designated use, these were the best available data representing impaired water quality at the extent of drainage basins.

Human Population

The number of people per square kilometer (population density) in each subwatershed and within the riparian corridor was calculated using block centroids derived from 2000 Census data (US Census Bureau 2000) compiled by ESRI (2002). Population density and riparian corridor population density upstream of the watershed were also computed.

Acid Deposition

Mean wet deposition of sulfate and nitrate in kilograms per hectare was calculated using interpolated data available in a 2.5 km cell resolution grid from the National Atmospheric Deposition Program (2005).

Soil Buffering Capacity

An estimate of soil buffering capacity was developed following the method of Thieling (2006). The pH soil dataset from the 1km-resolution CONUS-SOIL database (Miller & White 1998) was used to calculate the percentage of soil within the riparian corridor with a pH greater than or equal to 5.0. These data were created from the 1992 State Soil Geographic Data Base (STATSGO), which has a maximum map scale of 1:250,000 (Miller & White 1998).

References

Andersen, E. F. (2002) Effects of land use and land cover on freshwater mussel populations in the upper Neuse River basin, North Carolina: A GIS approach. M.S. thesis, North Carolina State University, Raleigh, NC.

ESRI (Environmental Systems Research Institute) (2000) ArcView v. 3.2. Redlands, CA.

ESRI (Environmental Systems Research Institute) (2002) ArcView v. 9.2. Redlands, CA.

ESRI Data & Maps (2002) Redlands, CA: Environmental Systems Research Institute.

Homer, C., Huang, C., Yang, L., Wylie, B., Coan, M. (2004) Development of a 2001 national landcover database for the United States. *Photogrammetric Engineering and Remote Sensing* **70**(7): 829-840.

Homer, C., Dewitz, J., Fry, J., Coan, M.., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., and Wickham, J. (2007) Completion of the 2001 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering and Remote Sensing* **73(**4): 337-341.

Horizon Systems (2006) NHDPlus. *http://www.horizon-systems.com/nhdplus/* (December 2006).

Jobson, H.E. (1996) Prediction of traveltime and longitudinal dispersion in rivers and streams. U.S. Geological Survey Water Resources Investigations Report 96-4013, U.S. Geological Survey, Washington, DC.

McKinney, R. A., Lake, J. L., Charpentier, M. A. & Ryba, S. (2002) Using mussel isotope ratios to assess anthropogenic nitrogen inputs to freshwater ecosystems. *Environmental Monitoring and Assessment* **74**(2): 167-192.

Miller, D.A. & White, R.A. (1998) A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interactions* **2**(2): 1-26.

NADP (National Atmospheric Deposition Program NRSP-3) (2007) NADP Program Office, Illinois State Water Survey. *http://nadp.sws.uiuc.edu/*

US EPA (US Environmental Protection Agency) (2002) NHD Indexed locations for Section 303(d) Listed Waters. *http://www.epa.gov/waters/data/downloads.htm*

US Census Bureau (2000) United States Census 2000. *http://www.census.gov/main/www/cen2000.html*

NAS (Nonindigenous Aquatic Species) Program (2007) Zebra Mussel (*Dreissena polymorpha*) *http://nas.er.usgs.gov/taxgroup/mollusks/zebramussel/*

National Atlas (2007) National Atlas of the United States *http://nationalatlas.gov/atlasftp.html/*

Research Triangle Institute (2001) The National water pollution control assessment model (NWPCAM) version 2, Draft report prepared for the U.S. Environmental Protection Agency, Research Triangle Park, North Carolina: Research Triangle Institute.

Richards, C., Johnson, L. B. & Host, G. E. (1996) Landscape-scale influences on stream habitats and biota. *Canadian Journal of Fisheries and Aquatic Sciences* **53**(Suppl. 1): 295-311.

Thieling, T. M. (2006) Assessment and predictive model for brook trout (*Salvelinus fontinalis*) population status in the eastern United States. M.S. thesis, James Madison University, Harrisonburg, VA.

USGS (U.S. Geological Survey), Eros Data Center (1999) National Elevation Dataset (30m). *http://gisdata.usgs.net/ned/*

USGS (U.S. Geological Survey) (2003) U.S. Geological Survey Minerals Information. *http://minerals.usgs.gov/minerals/*

Willson, J. D. & Dorcas, M. E. (2003) Effects of habitat disturbance on stream salamanders: implications for buffer zones and watershed management. *Conservation Biology* **17**(3): 763-771.

B. Species Information and Data

Table 1. Scientific and common names and information on distribution characteristics of species considered for model development.

Scientific Name	Common Name	Occurrence	Drainage Regions
Alasmidonta undulata	Triangle Floater	Declining	Mid-Atlantic
Alasmidonta varicosa	Brook Floater	Declining	Mid-Atlantic
Alasmidonta viridis	Slippershell Mussel	Common	Ohio/Tennessee
Elliptio complanata	Eastern Elliptio	Common	Mid-Atlantic/Tennessee
Elliptio dilatata	Spike	Common	Ohio/Tennessee
Epioblasma triquetra	Snuffbox	Declining	Ohio/Tennessee
Lasmigona subviridis	Green Floater	Declining	Mid-Atlantic/Ohio/Tennessee
Plethobasus cyphyus	Sheepnose	Declining	Ohio/Tennessee
Pleurobema clava	Clubshell	Rare	Ohio/Tennessee
Villosa fabalis	Rayed Bean	Declining	Ohio/Tennessee

Table 2. Available point data for species suggested by expert panel. Both refers to the number of spatially explicit recent records from each source. Records ≥1985 with coordinates are records with collection dates no earlier than 1985 that contain associated spatial reference data. All recent records are identified as ≥1985. All records with associated coordinates are identified under the coordinates heading. The total number of records includes all recent and historic records, with or without coordinates.

	Carnegie Museum	DE Museum	GA Nat Her Atl Slope	GA Nat Her TN Basin	IL Nat His Museum	MD MBSS	NC Nat Her	OSU Museum	Smith-sonian	U. of MI Museum	Total
≥1985 with coordinates											
A. undulata	1	0	0	0	1	0	0	9	1	0	12
A. varicosa	0	0	17	0	0	0	0	1	0	0	18
A. viridis	0	0	0	0	49	0	87	152	0	0	288
E. complanata	4	0	0	0	9	72	2	37	0	0	124
E. dilatata	15	0	0	0	79	0	69	283	0	0	446
E. triquetra	4	0	0	0	20	0	0	109	0	0	133
L. subviridis	0	0	0	0	0	0	8	9	0	0	17
P. cyphyus	0	0	0	0	14	0	0	19	0	0	33
P. clava	11	0	0	0	28	0	0	166	0	0	205
V. fabalis	5	0	0	0	1	0	0	66	0	0	72
Absence	0	0	11	22	0	778	0	0	0	0	811
≥1985											
A. undulata	3	6	0	0	14	0	0	11	7	0	41
A. varicosa	0	1	17	0	2	0	0	2	0	0	22
A. viridis	0	0	0	0	253	0	111	158	0	0	522
E. complanata	7	3	0	0	50	72	2	53	20	0	207
E. dilatata	30	4	0	0	550	0	94	322	0	0	1000
E. triquetra	13	0	0	0	113	0	0	113	0	0	239
L. subviridis	1	0	0	0	10	0	8	10	5	0	34
P. cyphyus	0	0	0	0	78	0	0	26	0	0	104
P. clava	18	2	0	0	176	0	0	189	0	0	385
V. fabalis	5	1	0	0	33	0	0	70	0	0	109
Absence	0	0	11	22	0	778	0	0	0	0	811

Table 2 (cont). Available point data for species suggested by expert panel. Both refers to the number of spatially explicit recent records from each source. Records ≥1985 with coordinates are records with collection dates no earlier than 1985 that contain associated spatial reference data. All recent records are identified as ≥1985. All records with associated coordinates are identified under the coordinates heading. The total number of records includes all recent and historic records, with or without coordinates.

	Carnegie Museum	DE Museum	GA Nat Her Atl Slope	GA Nat Her TN Basin	IL Nat His Museum	MD MBSS	NC Nat Her	OSU Museum	Smith-sonian	U. of MI Museum	Total
Coordinates											
A. undulata	1	0	0	0	1	0	0	10	16	0	28
A. varicosa	0	0	17	0	0	0	0	1	8	0	26
A. viridis	15	0	0	0	49	0	87	506	1	0	658
E. complanata	6	0	0	0	14	72	2	46	19	0	159
E. dilatata	42	0	0	0	87	0	69	1313	0	0	1511
E. triquetra	9	0	0	0	20	0	0	476	1	0	506
L. subviridis	0	0	0	0	0	0	8	13	12	0	33
P. cyphyus	13	0	0	0	14	0	0	180	1	0	208
P. clava	24	0	0	0	28	0	0	395	0	0	447
V. fabalis	11	0	0	0	1	0	0	229	0	0	241
Absence	0	0	13	22	0	778	0	0	0	0	813
Total Records											
A. undulata	54	55	0	0	32	0	0	56	42	109	348
A. varicosa	39	9	17	0	7	0	0	24	16	26	138
A. viridis	56	23	0	0	304	0	111	580	1	601	1676
E. complanata	196	296	0	0	107	72	2	284	60	1061	2078
E. dilatata	365	165	0	0	733	0	94	1740	17	1126	4240
E. triquetra	127	38	0	0	145	0	0	580	34	182	1106
L. subviridis	37	7	0	0	14	0	8	74	51	31	222
P. cyphyus	74	1	0	0	126	0	0	251	36	103	591
P. clava	95	25	0	0	197	0	0	484	90	89	980
V. fabalis	53	11	0	0	46	0	0	274	2	86	472
Absence	0	0	13	22	0	778	0	0	0	0	813

C. Model Variables

Table 1. Variables considered in the development of models for *Elliptio dilatata* and *Elliptio complanata* in the Ohio (OH) and Mid-Atlantic (M-A) drainage regions in the Appalachians. NLCD derivatives are based on general classes (e.g., 2x includes 21, 22, 23, and 24). Source is described in Table 2.

Source	Variable	Meaning	Units	OH	M-A
[1]	PA_EDIL	presence or absence of *E. dilatata*	N/A	x	
[2]	P_EDIL	presence of *E. dilatata*	N/A	x	
[1]	PA_ECOM	presence or absence of *E. complanata*	N/A		x
[2]	P_ECOM	presence of *E. complanata*	N/A		x
[3]	ZEBRAPA	presence or absence of zebra mussels in the subwatershed	N/A	x	x
[4]	AREA_KM2	area of the subwatershed	km^2	x	x
[4]	AREARC_KM2	area of the riparian corridor within the subwatershed	km^2	x	x
[4]	CUAREA_KM2	cumulative area of the subwatershed	km^2	x	x
[4]	CUAREARC	cumulative area of the riparian corridor of the subwatershed	km^2	x	x
[5]	DAMDENAREA	density of dams in the subwatershed	$dams*km^{-2}$	x	x
[5]	DAMDEN_RC	density of dams in the riparian corridor	$dams*km^{-2}$	x	x
[5]	CUDAMDENAR	cumulative density of dams in the subwatershed	$dams*km^{-2}$	x	x
[5]	CUDAMDENRC	cumulative density of dams in the riparian corridor	$dams*km^{-2}$	x	x
[5]	DISTDSDAMS	downstream distance to nearest dam	km	x	x
[5]	NUMDSDAMS	number of downstream dams	dams	x	x
[6]	RDDENAR	density of roads in the subwatershed	km of roads$*km^{-2}$	x	x
[6]	INTDENAREA	density of road-stream intersections in subwatershed	intersections$*km^{-2}$	x	x
[6]	INTDENSTRM	density of road-stream intersections per length of stream	intersections$*$stream km^{-1}	x	x
[6]	INTDENRC	density of road-stream intersections in the riparian corridor	intersections$*km^{-2}$	x	x
[6]	CUINTDENAR	cumulative density of road-stream intersections in subwatershed	intersections$*km^{-2}$	x	x
[6]	CUINTDENRC	cumulative density of road-stream intersections in riparian corridor	intersections$*km^{-2}$	x	x
[6]	CURDDEN_STM	cumulative density of roads per length of stream	km of roads$*$stream km^{-1}	x	x
[6]	CURDDENAR	cumulative density of roads in the subwatershed	km of roads$*km^{-2}$	x	x

Source	Variable	Meaning	Units	OH	M-A
[6]	CUINTDEN_SM	cumulative density of road-stream intersections per length of stream	intersections*km^{-2}	x	x
[6]	POPDENAREA	density of the human population in the subwatershed	people*km^{-2}	x	x
[6]	POPDEN_RC	density of the human population in the riparian corridor	people*km^{-2}	x	x
[6]	CUPOPDENAR	cumulative density of the human population in the subwatershed	people*km^{-2}	x	x
[6]	CUPOPDENRC	cumulative density of the human population in the riparian corridor	people*km^{-2}	x	x
[7]	MINEDENAR	density of mining facilities in the subwatershed	mines*km^{-2}	x	x
[7]	MINEDENRC	density of mining facilities in the riparian corridor	mines*km^{-2}	x	x
[7]	METALDENAR	density of metal facilities in the subwatershed	metal facilities*km^{-2}	x	x
[7]	METALDENRC	density of metal facilities in the riparian corridor	metal facilities*km^{-2}	x	x
[7]	CUMIDENRC	cumulative density of mining facilities in the riparian corridor	mines*km^{-2}	x	x
[7]	CUMIDENAR	cumulative density of mining facilities in the subwatershed	mines*km^{-2}	x	x
[7]	CUMEDENRC	cumulative density of metal facilities in the riparian corridor	metal facilities*km^{-2}	x	x
[7]	CUMEDENAR	cumulative density of metal facilities in the subwatershed	metal facilities*km^{-2}	x	x
[8]	303DDEN_AR	density of 303d streams in the subwatershed	km of 303d streams*km^{-2}		x
[8]	303DDEN_RC	density of 303d streams in the riparian corridor	km of 303d streams*km^{-2}		x
[8]	CU303DENAR	cumulative proportion of 303d streams in the subwatershed	km of 303d streams*km^{-2}		x
[8]	CU303DENSM	cumulative proportion of 303d streams per cumulative kilometer of streams	km of 303d streams*stream km^{-1}		x
[8]	CU303DENRC	cumulative proportion of 303d streams in the riparian corridor	km of 303d streams*km^{-2}		x
[8]	PROP_303D	proportion of 303-D streams in subwatershed	%		x
[9]	MIN_STMSLP	minimum streamslope from NHDPlus streamflow data	%	x	x
[9]	MAX_STMSLP	maximum streamslope from NHDPlus streamflow data	%	x	x
[9]	MAFLOU_MAX	maximum mean annual flow (unit runoff method) from NHDPlus streamflow data	m^3*s^{-1}	x	x
[9]	MAFLOU_MIN	minimum mean annual flow (unit runoff method) from NHDPlus streamflow data	m^3*s^{-1}	x	x
[9]	MAFLOU_DIF	difference between maximum and minimum mean annual flow	m^3*s^{-1}	x	x
[9]	MAVELU_MAX	maximum mean annual velocity (Jobson method) from NHDPlus streamflow data	m*s^{-1}	x	x
[9]	MAVELU_MIN	minimum mean annual velocity (Jobson method) from NHDPlus streamflow data	m*s^{-1}	x	x
[9]	MAVELU_DIF	difference between maximum and minimum mean annual velocity	m*s^{-1}	x	x
[9]	LINK_MAX	maximum link order within the subwatershed	N/A	x	x

Source	Variable	Meaning	Units	OH	M-A
[9]	LINK_MIN	minimum link order within the subwatershed	N/A	x	x
[9]	LINK_DIF	difference between maximum and minimum link order within the subwatershed	N/A	x	x
[9]	STRAHL_MAX	maximum Strahler stream order within the subwatershed	N/A	x	x
[9]	STRAHL_MIN	minimum Strahler stream order within the subwatershed	N/A	x	x
[9]	STRAHL_DIF	difference between maximum and minimum Strahler stream order within the subwatershed	N/A	x	x
[9]	MAXELEVSMO	maximum of smoothed maximum elevation from NHDPlus streamflow data	m	x	x
[9]	MINELEVSMO	minimum of smoothed minimum elevation of flowlines in the subwatershed from NHDPlus	m	x	x
[10]	ELEV_MIN	minimum elevation in subwatershed	m	x	x
[10]	ELEV_MAX	maximum elevation in subwatershed	m	x	x
[10]	ELEV_MEAN	mean elevation in subwatershed	m	x	x
[10]	RC_EL_MIN	minimum elevation in riparian corridor	m	x	x
[10]	RC_EL_AVG	mean elevation in riparian corridor	m	x	x
[10]	RC_EL_MAX	maximum elevation in riparian corridor	m	x	x
[10]	SLP_MIN	minimum slope in subwatershed	%	x	x
[10]	SLP_MAX	maximum slope in subwatershed	%	x	x
[10]	SLP_MEAN	mean slope in subwatershed	%	x	x
[10]	RC_SLP_MIN	minimum slope in riparian corridor	%	x	x
[10]	RC_SLP_AVG	mean slope in riparian corridor	%	x	x
[10]	RC_SLP_MAX	maximum slope in riparian corridor	%	x	x
[11]	NO3_MIN	minimum NO_3 deposition in subwatershed	$kg*ha^{-1}$	x	x
[11]	NO3_MAX	maximum NO_3 deposition in subwatershed	$kg*ha^{-1}$	x	x
[11]	NO3_MEAN	mean NO_3 deposition in subwatershed	$kg*ha^{-1}$	x	x
[11]	SO4_MIN	minimum SO_4 deposition in subwatershed	$kg*ha^{-1}$	x	x
[11]	SO4_MAX	maximum SO_4 deposition in subwatershed	$kg*ha^{-1}$	x	x
[11]	SO4_MEAN	mean SO_4 deposition in subwatershed	$kg*ha^{-1}$	x	x
[12]	PHSOIL	proportion of subwatershed with soil pH <5.0	%	x	x
[13]	FOR	proportion of all forested land (NLCD 4x) in the subwatershed	%	x	x

Source	Variable	Meaning	Units	OH	M-A
[13]	FOR_RC	proportion of forested land in the riparian corridor	%	x	x
[13]	CU_FOR	cumulative proportion of forested land in the subwatershed	%	x	x
[13]	CU_FOR_RC	cumulative proportion of forested land in the riparian corridor	%	x	x
[13]	H2O	proportion of all water (NLCD 1x) in the subwatershed	%	x	x
[13]	H2O_RC	proportion of water in the riparian corridor	%	x	x
[13]	CU_H2O	cumulative proportion of water in the subwatershed	%	x	x
[13]	CU_H2O_RC	cumulative proportion of water in the riparian corridor	%	x	x
[13]	AG	proportion of all agriculture (NLCD 8x) in the subwatershed	%	x	x
[13]	AG_RC	proportion of agriculture in the riparian corridor	%	x	x
[13]	CU_AG	cumulative proportion of agriculture in the subwatershed	%	x	x
[13]	CU_AG_RC	cumulative proportion of agriculture in the riparian corridor	%	x	x
[13]	DEV	proportion of all developed land (NLCD 2x) in the subwatershed	%	x	x
[13]	DEV_RC	proportion of developed land in the riparian corridor	%	x	x
[13]	CU_DEV	cumulative proportion of developed land in the subwatershed	%	x	x
[13]	CU_DEV_RC	cumulative proportion of developed land in the riparian corridor	%	x	x
[13]	BAR	proportion of all barren land cover (NLCD 3x) in the subwatershed	%	x	x
[13]	BAR_RC	proportion of barren land cover in the riparian corridor	%	x	x
[13]	CU_BAR	cumulative proportion of barren land cover in the subwatershed	%	x	x
[13]	CU_BAR_RC	cumulative proportion of barren land cover in the riparian corridor	%	x	x
[13]	HUM	proportion of all human impacts (NLCD 2x, 3x and 8x) in the subwatershed	%	x	x
[13]	HUM_RC	proportion of human impacts in the riparian corridor	%	x	x
[13]	CU_HUM	cumulative proportion of human impacts in the subwatershed	%	x	x
[13]	CU_HUM_RC	cumulative proportion of human impacts in the riparian corridor	%	x	x
[13]	DEC	proportion of deciduous forest (NLCD 41) in the subwatershed	%	x	x
[13]	DEC_RC	proportion of deciduous forest in the riparian corridor	%	x	x
[13]	CU_DEC	cumulative proportion of deciduous forest in the subwatershed	%	x	x
[13]	CU_DEC_RC	cumulative proportion of deciduous forest in the riparian corridor	%	x	x
[13]	EVG	proportion of evergreen forest (NLCD 42) in the subwatershed	%	x	x

Source	Variable	Meaning	Units	OH	M-A
[13]	EVG_RC	proportion of evergreen forest in the riparian corridor	%	x	x
[13]	CU_EVG	cumulative proportion of evergreen forest in the subwatershed	%	x	x
[13]	CU_EVG_RC	cumulative proportion of evergreen forest in the riparian corridor	%	x	x
[13]	MXF	proportion of mixed forest (NLCD 43) in the subwatershed	%	x	x
[13]	MXF_RC	proportion of mixed forest in the riparian corridor	%	x	x
[13]	CU_MXF	cumulative proportion of mixed forest in the subwatershed	%	x	x
[13]	CU_MXF_RC	cumulative proportion of mixed forest in the riparian corridor	%	x	x
[13]	GRS	proportion of grass (NLCD 71) in the subwatershed	%	x	x
[13]	GRS_RC	proportion of grass in the riparian corridor	%	x	x
[13]	CU_GRS	cumulative proportion of grass in the subwatershed	%	x	x
[13]	CU_GRS_RC	cumulative proportion of grass in the riparian corridor	%	x	x
[13]	HDV	proportion of high-intensity development (NLCD 24) in the subwatershed	%	x	x
[13]	HDV_RC	proportion of high-intensity development in the riparian corridor	%	x	x
[13]	CU_HDV	cumulative proportion of high-intensity development in the subwatershed	%	x	x
[13]	CU_HDV_RC	cumulative proportion of high-intensity development in the riparian corridor	%	x	x
[13]	MDV	proportion of med-intensity development (NLCD 23) in the subwatershed	%	x	x
[13]	MDV_RC	proportion of med-intensity development in the riparian corridor	%	x	x
[13]	CU_MDV	cumulative proportion of med-intensity development in the subwatershed	%	x	x
[13]	CU_MDV_RC	cumulative proportion of med-intensity development in the riparian corridor	%	x	x
[13]	LDV	proportion of low-intensity development (NLCD 22) in the subwatershed	%	x	x
[13]	LDV_RC	proportion of low-intensity development in the riparian corridor	%	x	x
[13]	CU_LDV	cumulative proportion of low-intensity development in the subwatershed	%	x	x
[13]	CU_LDV_RC	cumulative proportion of low-intensity development in the riparian corridor	%	x	x
[13]	ODV	proportion of open development (NLCD 21) in the subwatershed	%	x	x
[13]	ODV_RC	proportion of open development in the riparian corridor	%	x	x
[13]	CU_ODV	cumulative proportion of open development in the subwatershed	%	x	x
[13]	CU_ODV_RC	cumulative proportion of open development in the riparian corridor	%	x	x
[13]	LO	proportion of low-intensity or open development (NLCD 21-22) in the subwatershed	%	x	x

Source	Variable	Meaning	Units	OH	M-A
[13]	LO_RC	proportion of low-intensity or open development in the riparian corridor	%	x	x
[13]	CU_LO	cumulative proportion of low-intensity or open development in the subwatershed	%	x	x
[13]	CU_LO_RC	cumulative proportion of low-intensity or open development in the riparian corridor	%	x	x
[13]	MH	proportion of low-intensity or open development (NLCD 23-24) in the subwatershed	%	x	x
[13]	MH_RC	proportion of low-intensity or open development in the riparian corridor	%	x	x
[13]	CU_MH	cumulative proportion of low-intensity or open development in the subwatershed	%	x	x
[13]	CU_MH_RC	cumulative proportion of low-intensity or open development in the riparian corridor	%	x	x
[13]	WET	proportion of woody wetlands (NLCD 90) in the subwatershed	%	x	x
[13]	WET_RC	proportion of woody wetlands in the riparian corridor	%	x	x
[13]	CU_WET	cumulative proportion of woody wetlands in the subwatershed	%	x	x
[13]	CU_WET_RC	cumulative proportion of woody wetlands in the riparian corridor	%	x	x
[13]	EMH	proportion of emergent herbaceous (NLCD 95) in the subwatershed	%	x	x
[13]	EMH_RC	proportion of emergent herbaceous in the riparian corridor	%	x	x
[13]	CU_EMH	cumulative proportion of emergent herbaceous in the subwatershed	%	x	x
[13]	CU_EMH_RC	cumulative proportion of emergent herbaceous in the riparian corridor	%	x	x
[13]	HAY	proportion of pasture/hay (NLCD 81) in the subwatershed	%	x	x
[13]	HAY_RC	proportion of pasture/hay in the riparian corridor	%	x	x
[13]	CU_HAY	cumulative proportion of pasture/hay in the subwatershed	%	x	x
[13]	CU_HAY_RC	cumulative proportion of pasture/hay in the riparian corridor	%	x	x
[13]	CRP	proportion of cultivated crops (NLCD 82) in the subwatershed	%	x	x
[13]	CRP_RC	proportion of cultivated crops in the riparian corridor	%	x	x
[13]	CU_CRP	cumulative proportion of cultivated crops in the subwatershed	%	x	x
[13]	CU_CRP_RC	cumulative proportion of cultivated crops in the riparian corridor	%	x	x
[13]	SHB	proportion of shrub (NLCD 52) in the subwatershed	%	x	x
[13]	SHB_RC	proportion of shrub in the riparian corridor	%	x	x
[13]	CU_SHB	cumulative proportion of shrub in the subwatershed	%	x	x
[13]	CU_SHB_RC	cumulative proportion of shrub in the riparian corridor	%	x	x

Table 2. Information on data sources for variables described in Table 1.

	Data Source	Scale or Resolution	Relevant Time Period	Source Details
1	survey data	subwatershed	1985-2007	Mussel data were compiled from Western Pennsylvania Conservancy, Maryland Biological Stream Survey, and Ecological Specialists, Inc.
2	occurrences	subwatershed	1985-2007	Mussel data were compiled from West Virginia DNR, Carnegie Museum, Ohio State University Museum, and Illinois Natural History Museum
3	NAS	1:2,000,000	1988-2007	Point locations for zebra mussels compiled by the Nonindigenous Aquatic Species (NAS) Program were downloaded from the National Atlas
4	USGS/NRCS	1:100,000	varies	Subwatershed boundaries were edited from USGS and Natural Resources Conservation Service (NRCS) boundaries
5	NID	1:2,000,000	2002-2004	A shapefile containing locations of dams compiled by the National Inventory of Dams was downloaded from the National Atlas
6	ESRI	1:100,000	2000	U.S. Detailed Streets and U.S. Census Block Centroid Populations were obtained from Environmental Systems Research Institute, Inc. (ESRI) data CDs provided with ArcGIS software
7	USGS	1:2,000,000	2003	Shapefiles of locations of metal processing plants and mines compiled by the U.S. Geological Survey (USGS) were downloaded from the National Atlas
8	USEPA	1:100,000	2002	303-D listed streams were obtained from the U.S. Environmental Protection Agency (USEPA)
9	NHDPlus	1:100,000	2005	National Hydrography Dataset Plus (NHDPlus) data were downloaded from Horizon Systems, Inc.
10	NED	30m	varies	The National Elevation Dataset (NED) was obtained directly from the USGS
11	NADP	2.5km	2005	Deposition data were obtained from the National Atmospheric Deposition Program (NADP)
12	CONUS	1km	1998	Soil acidity was determined from upper soil layers in CONUS-SOIL data prepared by Miller and White (1998).
13	NLCD	30m	1999-2001	National Land Cover Database (NLCD) land use grids were downloaded from the Multi-Resolution Land Characteristics Consortium (MRLC) by mapping zone (Homer et al. 2007).

Table 3. Summary statistics for variables in the Mid-Atlantic (Mid-Atl.) and Ohio drainage regions.

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Mid-Atl.	AREA_KM2	100.11	1.9608	7.3903	1205.2	107.15	2.9499	23.948	502.46	98.867	2.246	7.3903	1205.2
	AREARC_KM2	14.442	0.2765	1.2141	146.26	16.446	0.4762	2.5641	80.189	14.088	0.3133	1.2141	146.26
	CUAREA_KM2	1714.7	156.19	15.745	71185	5508.4	806.56	32.078	71185	1045.5	107.41	15.745	70115
	CUAREARC	246.59	22.571	1.2141	10261	791.96	116.07	2.5641	10261	150.39	15.648	1.2141	10092
	DAMDENAREA	0.0027	0.0002	0	0.0765	0.0024	0.0004	0	0.0528	0.0028	0.0002	0	0.0765
	DAMDEN_RC	0.0196	0.0015	0	0.7324	0.0164	0.003	0	0.3145	0.0202	0.0017	0	0.7324
	CUDAMDENAR	0.0028	0.0002	0	0.0606	0.0022	0.0002	0	0.0348	0.0029	0.0002	0	0.0606
	CUDAMDENRC	0.0207	0.0013	0	0.7324	0.0149	0.0016	0	0.2583	0.0217	0.0015	0	0.7324
	DISTDSDAMS	811028	65956	-71.51	1E+07	2E+06	226571	-71.51	1E+07	679540	65951	-66.44	1E+07
	NUMDSDAMS	9.1581	0.1742	0	27	6.7626	0.3651	0	22	9.5806	0.1924	0	27
	RDDENAR	2033.6	28.239	165.76	11825	2368.8	76.171	879.33	11825	1974.5	30.129	165.76	11034
	INTDENAREA	3.9662	0.0446	0.1705	15.862	4.2702	0.1014	1.3024	10.471	3.9126	0.0492	0.1705	15.862
	INTDENSTRM	0.0008	9E-06	4E-05	0.0032	0.0008	2E-05	0.0003	0.0021	0.0008	1E-05	4E-05	0.0032
	INTDENRC	0.5661	0.0068	0.023	2.1081	0.6509	0.0153	0.1498	1.4992	0.5512	0.0074	0.023	2.1081
	CUINTDENAR	0.5523	0.0061	0.023	2.1081	0.6241	0.0119	0.1879	1.4992	0.5396	0.0068	0.023	2.1081
	CUINTDENRC	3.9556	0.0393	0.1705	15.862	4.2218	0.079	1.4685	10.471	3.9086	0.0439	0.1705	15.862
	CURDDEN_STM	2.8872	0.0456	0.2285	23.607	2.8114	0.0836	1.1828	12.205	2.9006	0.0516	0.2285	23.607
	CURDDENAR	1878.3	22.689	165.76	9446.1	2042.9	51.219	879.33	7466.8	1849.3	25.045	165.76	9446.1
	CUINTDEN_SM	0.0008	8E-06	4E-05	0.0032	0.0008	2E-05	0.0003	0.0021	0.0008	9E-06	4E-05	0.0032
	POPDENAREA	84.592	5.8278	0	5258.7	117.59	16.935	0.6765	3520.7	78.771	6.1609	0	5258.7
	POPDEN_RC	48.469	2.3618	0	1161.6	64.401	6.7454	0	1117.9	45.658	2.5046	0	1161.6
	CUPOPDENAR	61.542	3.1522	0	1574.6	74.669	8.2431	0.9293	1122.7	59.227	3.4088	0	1574.6
	CUPOPDENRC	39.34	1.8897	0	1117.9	44.047	5.7342	0	1117.9	38.509	1.9797	0	942.37
	MINEDENAR	0.0048	0.0003	0	0.1775	0.0071	0.0011	0	0.1775	0.0044	0.0004	0	0.1553
	MINEDENRC	0.0066	0.0049	0	7.6687	0	0	0	0	0.0077	0.0058	0	7.6687
	METALDENAR	3E-05	3E-05	0	0.0566	0.0002	0.0002	0	0.0566	0	0	0	0
	METALDENRC	0.0002	0.0002	0	0.2604	0.001	0.001	0	0.2604	0	0	0	0
	CUMIDENRC	0.0066	0.0049	0	7.6687	0	0	0	0	0.0077	0.0058	0	7.6687
	CUMIDENAR	0.0022	0.0002	0	0.1553	0.0016	0.0005	0	0.0493	0.0023	0.0003	0	0.1553
	CUMEDENRC	0	0	0	0	0	0	0	0	0	0	0	0
	CUMEDENAR	3E-05	2E-05	0	0.0249	4E-05	2E-05	0	0.0042	3E-05	2E-05	0	0.0249
	303DDEN_AR	76.17	3.1	0	811.32	137.6	10.477	0	811.32	65.334	3.0592	0	788.71
	303DDEN_RC	566.89	23.948	0	8244.2	922.3	71.957	0	4985.6	504.2	24.802	0	8244.2

119

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Mid-Atl.	CU303DENAR	63.848	2.4786	0	811.32	91.14	6.8348	0	811.32	59.034	2.6359	0	788.71
	CU303DENSM	0.0967	0.0039	0	1.0469	0.1268	0.0102	0	0.9885	0.0914	0.0042	0	1.0469
	CU303DENRC	484.84	19.612	0	5320.3	640.02	51.459	0	4985.6	457.47	21.138	0	5320.3
	PROP_303D	0.1118	0.0047	0	1.1331	0.1798	0.0139	0	0.9885	0.0999	0.0048	0	1.1331
	MIN_STMSLP	0.001	6E-05	0	0.021	0.0002	6E-05	0	0.0117	0.0011	7E-05	0	0.021
	MAX_STMSLP	0.4682	0.0627	0.0043	56.9	0.4725	0.106	0.0103	20.335	0.4674	0.0713	0.0043	56.9
	MAFLOU_MAX	29.76	2.6396	0	1133	90.199	12.882	0.3084	1133	19.1	1.994	0	1117.9
	MAFLOU_MIN	0.0181	0.0017	0	1.3505	0.0035	0.0004	0	0.0306	0.0206	0.002	0	1.3505
	MAFLOU_DIF	29.742	2.6397	0	1133	90.196	12.882	0.3082	1133	19.079	1.9941	0	1117.9
	MAVELU_MAX	0.4972	0.0048	0	2.47	0.6175	0.0187	0.3115	2.47	0.476	0.0044	0	2.0492
	MAVELU_MIN	0.1119	0.0023	0	0.4054	0.0659	0.005	0	0.2625	0.12	0.0025	0	0.4054
	MAVELU_DIF	0.3853	0.006	0	2.47	0.5516	0.0206	0.1064	2.47	0.356	0.0058	0	1.8824
	LINK_MAX	5.0134	0.0269	2	9	4.786	0.0668	2	8	5.0535	0.0293	2	9
	LINK_MIN	2.6628	0.0239	0	6	2.2374	0.0579	0	4	2.7378	0.0257	0	6
	LINK_DIF	2.3506	0.0207	0	5	2.5486	0.0464	1	5	2.3157	0.0228	0	5
	STRAHL_MAX	3.8226	0.0334	1	8	4.6887	0.0945	1	8	3.6699	0.0341	1	8
	STRAHL_MIN	0.9959	0.0017	0	2	0.9767	0.0094	0	1	0.9993	0.0012	0	2
	STRAHL_DIF	2.8267	0.0336	0	7	3.7121	0.096	0	7	2.6706	0.0342	0	7
	MAXELEVSMO	496.17	5.3689	1.2629	1322.4	393.78	11.947	43.71	988.48	514.23	5.8287	1.2629	1322.4
	MINELEVSMO	157.08	3.6767	0	760.46	78.277	7.1828	0	606.43	170.98	4.0279	0	760.46
	ELEV_MIN	216.79	3.129	-68	775	147.62	6.2176	-68	638	228.99	3.4161	-8	775
	ELEV_MAX	617.95	6.3634	0	1482	505.72	12.837	105	1063	637.75	7.0098	0	1482
	ELEV_MEAN	367.63	4.2769	0	1083.7	270.93	9.153	24.531	846.28	384.69	4.6245	0	1083.7
	RC_EL_MIN	217.74	3.1134	-24	775	148.51	6.1839	-24	639	229.96	3.398	-5	775
	RC_EL_AVG	319.96	3.801	6.7557	995.25	231.35	8.1242	6.7557	791.65	335.59	4.1024	8.0591	995.25
	RC_EL_MAX	525.45	5.5696	53	1348	421.47	12.235	67	1011	543.79	6.0622	53	1348
	SLP_MIN	0	0	0	0	0	0	0	0	0	0	0	0
	SLP_MAX	88.712	0.757	14	470	91.79	1.971	23	203	88.17	0.8193	14	470
	SLP_MEAN	13.965	0.1665	1.2378	40.826	12.286	0.3692	2.1545	34.095	14.261	0.1837	1.2378	40.826
	RC_SLP_MIN	0	0	0	0	0	0	0	0	0	0	0	0
	RC_SLP_AVG	12.675	0.1436	0.922	42.292	11.919	0.3383	1.4529	34.289	12.809	0.1578	0.922	42.292
	RC_SLP_MAX	77.759	0.6957	14	468	84.051	1.8798	17	203	76.649	0.7447	14	468
	NO3_MIN	11.439	0.0362	7.4383	14.34	11.787	0.0943	7.4825	14.279	11.378	0.039	7.4383	14.34
	NO3_MAX	11.718	0.0346	7.5564	14.81	12.042	0.0898	7.5564	14.346	11.661	0.0374	7.5674	14.81

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Mid-Atl.	NO3_MEAN	11.579	0.0353	7.4653	14.49	11.914	0.0921	7.5041	14.326	11.52	0.0381	7.4653	14.49
	SO4_MIN	16.503	0.0528	12.173	21.558	17.277	0.1579	12.375	21.267	16.366	0.0548	12.173	21.558
	SO4_MAX	17.026	0.0524	12.361	21.602	17.779	0.1538	12.515	21.581	16.894	0.0546	12.361	21.602
	SO4_MEAN	16.765	0.0524	12.222	21.584	17.523	0.156	12.419	21.453	16.631	0.0545	12.222	21.584
	PHSOIL	0.7637	0.0074	0	1	0.7045	0.0187	0	1	0.7741	0.008	0	1
	FOR	0.6035	0.0056	0.0083	0.9872	0.5	0.016	0.0185	0.953	0.6217	0.0059	0.0083	0.9872
	FOR_RC	0.5825	0.005	0.0085	0.993	0.525	0.0131	0.0126	0.9495	0.5927	0.0053	0.0085	0.993
	CU_FOR	0.6216	0.0052	0.0083	0.9872	0.5627	0.0139	0.0383	0.9307	0.632	0.0056	0.0083	0.9872
	CU_FOR_RC	0.5918	0.0047	0.0085	0.993	0.5645	0.0113	0.0537	0.9267	0.5967	0.0051	0.0085	0.993
	H2O	0.0109	0.0006	0	0.3158	0.021	0.0025	0	0.2816	0.0091	0.0005	0	0.3158
	H2O_RC	0.0182	0.0006	0	0.2801	0.0309	0.0024	0	0.2801	0.0159	0.0006	0	0.2564
	CU_H2O	0.0067	0.0002	0	0.0953	0.0068	0.0005	0	0.0633	0.0067	0.0003	0	0.0953
	CU_H2O_RC	0.0115	0.0003	0	0.1192	0.0133	0.0007	0	0.0612	0.0112	0.0003	0	0.1192
	AG	0.261	0.0048	0	0.8915	0.3488	0.0144	0.0026	0.8915	0.2456	0.005	0	0.8575
	AG_RC	0.2371	0.0043	0	0.8679	0.2836	0.0123	0.0018	0.846	0.2289	0.0046	0	0.8679
	CU_AG	0.2578	0.0045	0	0.8575	0.3249	0.0128	0.0014	0.8234	0.2459	0.0048	0	0.8575
	CU_AG_RC	0.2414	0.0041	0	0.8679	0.283	0.0112	0.0014	0.814	0.2341	0.0044	0	0.8679
	DEV	0.0917	0.0026	0.0002	0.9237	0.1121	0.007	0.0072	0.8547	0.0881	0.0028	0.0002	0.9237
	DEV_RC	0.1073	0.0022	0.0004	0.8351	0.1173	0.0055	0.0034	0.721	0.1055	0.0024	0.0004	0.8351
	CU_DEV	0.0791	0.0021	0.0002	0.8228	0.085	0.0045	0.0084	0.6077	0.0781	0.0023	0.0002	0.8228
	CU_DEV_RC	0.0998	0.0017	0.0004	0.6411	0.0991	0.0032	0.0067	0.4532	0.0999	0.002	0.0004	0.6411
	BAR	0.0034	0.0002	0	0.0809	0.0043	0.0004	0	0.0408	0.0032	0.0002	0	0.0809
	BAR_RC	0.0024	0.0002	0	0.1508	0.0033	0.0004	0	0.0306	0.0022	0.0002	0	0.1508
	CU_BAR	0.0033	0.0002	0	0.0809	0.0043	0.0004	0	0.0447	0.0031	0.0002	0	0.0809
	CU_BAR_RC	0.0023	0.0002	0	0.1508	0.003	0.0003	0	0.0247	0.0021	0.0002	0	0.1508
	HUM	0.3561	0.0056	0.0002	0.9878	0.4652	0.016	0.0171	0.9699	0.3369	0.0058	0.0002	0.9878
	HUM_RC	0.3468	0.0048	0.0004	0.9899	0.4043	0.0125	0.0462	0.9322	0.3367	0.0051	0.0004	0.9899
	CU_HUM	0.3401	0.0052	0.0002	0.9878	0.4141	0.0142	0.0456	0.9565	0.3271	0.0055	0.0002	0.9878
	CU_HUM_RC	0.3435	0.0046	0.0004	0.9899	0.3852	0.0113	0.0496	0.9117	0.3361	0.005	0.0004	0.9899
	DEC	0.4802	0.005	0.0082	0.9501	0.4264	0.0139	0.0182	0.8586	0.4897	0.0053	0.0082	0.9501
	DEC_RC	0.4239	0.0043	0.0043	0.8848	0.4212	0.0109	0.0116	0.8848	0.4244	0.0047	0.0043	0.8775
	CU_DEC	0.5002	0.0046	0.0082	0.9319	0.4798	0.012	0.0346	0.8143	0.5038	0.005	0.0082	0.9319
	CU_DEC_RC	0.4323	0.004	0.0044	0.8775	0.4438	0.0094	0.05	0.8042	0.4303	0.0044	0.0044	0.8775
	EVG	0.0508	0.0011	0	0.3237	0.03	0.0017	0.0003	0.1583	0.0545	0.0012	0	0.3237

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Mid-Atl.	EVG_RC	0.0705	0.0016	0	0.4705	0.049	0.0034	0.0001	0.3218	0.0743	0.0017	0	0.4705
	CU_EVG	0.0513	0.001	2E-05	0.2892	0.0359	0.0016	0.002	0.1583	0.054	0.0011	2E-05	0.2892
	CU_EVG_RC	0.0734	0.0015	0	0.4705	0.0606	0.0031	0.002	0.3218	0.0756	0.0017	0	0.4705
	MXF	0.0725	0.002	0	0.6206	0.0435	0.0048	0	0.508	0.0776	0.0022	0	0.6206
	MXF_RC	0.088	0.0025	0	0.7046	0.0548	0.0054	0	0.5518	0.0939	0.0027	0	0.7046
	CU_MXF	0.0701	0.0018	0	0.6206	0.047	0.0035	0	0.3278	0.0741	0.002	0	0.6206
	CU_MXF_RC	0.0862	0.0023	0	0.7046	0.0601	0.0044	0	0.4157	0.0908	0.0026	0	0.7046
	GRS	0.0034	0.0002	0	0.0837	0.0004	9E-05	0	0.0132	0.004	0.0002	0	0.0837
	GRS_RC	0.0025	0.0002	0	0.0589	0.0002	6E-05	0	0.0089	0.0029	0.0002	0	0.0589
	CU_GRS	0.0038	0.0002	0	0.0626	0.0015	0.0002	0	0.0171	0.0042	0.0002	0	0.0626
	CU_GRS_RC	0.0028	0.0002	0	0.0589	0.0011	0.0002	0	0.0145	0.0031	0.0002	0	0.0589
	HDV	0.0033	0.0003	0	0.3066	0.0051	0.0013	0	0.3066	0.003	0.0003	0	0.1589
	HDV_RC	0.0024	0.0002	0	0.2001	0.0034	0.0009	0	0.2001	0.0022	0.0002	0	0.1579
	CU_HDV	0.0021	0.0002	0	0.0921	0.0022	0.0003	0	0.0433	0.0021	0.0002	0	0.0921
	CU_HDV_RC	0.0016	0.0001	0	0.1579	0.0015	0.0002	0	0.0239	0.0017	0.0002	0	0.1579
	MDV	0.0092	0.0006	0	0.3194	0.0127	0.0017	0	0.29	0.0086	0.0006	0	0.3194
	MDV_RC	0.0076	0.0005	0	0.2853	0.0098	0.0014	0	0.2039	0.0072	0.0005	0	0.2853
	CU_MDV	0.0064	0.0004	0	0.1966	0.007	0.0008	0	0.1106	0.0063	0.0004	0	0.1966
	CU_MDV_RC	0.0056	0.0003	0	0.1461	0.0055	0.0005	0	0.1043	0.0056	0.0003	0	0.1461
	LDV	0.0241	0.001	0	0.4059	0.0342	0.0027	0	0.2694	0.0223	0.0011	0	0.4059
	LDV_RC	0.0226	0.0008	0	0.242	0.0288	0.002	0	0.1652	0.0215	0.0009	0	0.242
	CU_LDV	0.0191	0.0008	0	0.4059	0.0234	0.0018	0.0002	0.2066	0.0183	0.0009	0	0.4059
	CU_LDV_RC	0.0189	0.0006	0	0.242	0.0209	0.0012	0.001	0.1348	0.0185	0.0007	0	0.242
	ODV	0.0551	0.001	0.0002	0.4897	0.0601	0.0024	0.0032	0.3102	0.0542	0.0011	0.0002	0.4897
	ODV_RC	0.0748	0.0011	0.0004	0.4024	0.0753	0.0024	0.0022	0.2052	0.0747	0.0012	0.0004	0.4024
	CU_ODV	0.0515	0.0009	0.0002	0.4897	0.0524	0.0019	0.0068	0.3102	0.0513	0.001	0.0002	0.4897
	CU_ODV_RC	0.0737	0.001	0.0004	0.4024	0.0712	0.0019	0.0046	0.1902	0.0742	0.0011	0.0004	0.4024
	LO	0.0792	0.0019	0.0002	0.6259	0.0943	0.0048	0.0059	0.4966	0.0765	0.0021	0.0002	0.6259
	LO_RC	0.0973	0.0017	0.0004	0.536	0.1041	0.0039	0.003	0.358	0.0961	0.0018	0.0004	0.536
	CU_LO	0.0705	0.0016	0.0002	0.6259	0.0758	0.0035	0.0083	0.4966	0.0696	0.0018	0.0002	0.6259
	CU_LO_RC	0.0926	0.0014	0.0004	0.536	0.0921	0.0027	0.0065	0.325	0.0927	0.0016	0.0004	0.536
	MH	0.0125	0.0009	0	0.5966	0.0178	0.0029	0	0.5966	0.0116	0.0009	0	0.4697
	MH_RC	0.01	0.0007	0	0.4254	0.0132	0.0022	0	0.4039	0.0094	0.0007	0	0.4254
	CU_MH	0.0086	0.0005	0	0.2887	0.0092	0.001	1E-04	0.1539	0.0084	0.0006	0	0.2887

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Mid-Atl.	CU_MH_RC	0.0072	0.0004	0	0.3041	0.007	0.0007	0	0.1282	0.0072	0.0005	0	0.3041
	WET	0.0151	0.0009	0	0.5141	0.0079	0.0008	0	0.1071	0.0164	0.001	0	0.5141
	WET_RC	0.0348	0.0015	0	0.486	0.0262	0.0021	0	0.2325	0.0363	0.0018	0	0.486
	CU_WET	0.0159	0.0008	0	0.3101	0.0086	0.0008	0	0.0911	0.0172	0.0009	0	0.3101
	CU_WET_RC	0.0356	0.0015	0	0.4698	0.0255	0.0019	0	0.182	0.0374	0.0017	0	0.4698
	EMH	0.0022	9E-05	0	0.0382	0.0035	0.0003	0	0.0243	0.002	9E-05	0	0.0382
	EMH_RC	0.0069	0.0003	0	0.0881	0.0113	0.001	0	0.0849	0.0061	0.0003	0	0.0881
	CU_EMH	0.0018	6E-05	0	0.0275	0.0021	0.0001	0	0.0114	0.0017	7E-05	0	0.0275
	CU_EMH_RC	0.0051	0.0002	0	0.0881	0.0062	0.0004	0	0.0458	0.0049	0.0002	0	0.0881
	HAY	0.1807	0.0034	0	0.7066	0.2223	0.0095	0.0014	0.6916	0.1734	0.0037	0	0.7066
	HAY_RC	0.1758	0.0034	0	0.8366	0.1903	0.0083	0.0015	0.6256	0.1732	0.0037	0	0.8366
	CU_HAY	0.1781	0.0031	0	0.7066	0.2081	0.0079	0.0011	0.5562	0.1728	0.0034	0	0.7066
	CU_HAY_RC	0.1811	0.0032	0	0.8366	0.195	0.0073	0.001	0.6256	0.1786	0.0036	0	0.8366
	CRP	0.0803	0.0024	0	0.6655	0.1264	0.0067	0.0004	0.4232	0.0722	0.0025	0	0.6655
	CRP_RC	0.0613	0.0018	0	0.5176	0.0933	0.005	0.0001	0.4404	0.0557	0.0018	0	0.5176
	CU_CRP	0.0797	0.0023	0	0.6655	0.1168	0.006	0.0003	0.4528	0.0731	0.0024	0	0.6655
	CU_CRP_RC	0.0604	0.0017	0	0.5176	0.0881	0.0045	0.0001	0.4182	0.0555	0.0018	0	0.5176
	SHB	0.0088	0.0004	0	0.1769	0.002	0.0004	0	0.0398	0.01	0.0005	0	0.1769
	SHB_RC	0.0083	0.0005	0	0.2103	0.0021	0.0005	0	0.077	0.0094	0.0005	0	0.2103
	CU_SHB	0.0101	0.0005	0	0.1769	0.0042	0.0006	0	0.0486	0.0111	0.0005	0	0.1769
	CU_SHB_RC	0.0097	0.0005	0	0.2103	0.0041	0.0006	0	0.077	0.0106	0.0005	0	0.2103
Ohio	AREA_KM2	84.01	0.7322	8.9306	763.95	104.02	2.8363	36.232	288.45	82.854	0.7517	8.9306	763.95
	AREARC_KM2	13.665	0.1197	1.2024	84.572	16.989	0.4981	3.0852	35.546	13.473	0.1224	1.2024	84.572
	CUAREA_KM2	3570	332.62	14.52	201624	17980	3078.1	43.24	182097	2737.7	296.84	14.52	201624
	CUAREARC	584.06	55.064	2.1132	33412	2938	510.48	5.5728	30351	448.09	49.139	2.1132	33412
	DAMDENAREA	0.0022	0.0001	0	0.1109	0.0036	0.0006	0	0.0684	0.0021	0.0001	0	0.1109
	DAMDEN_RC	0.0135	0.0008	0	0.7168	0.0239	0.0044	0	0.4533	0.0129	0.0008	0	0.7168
	CUDAMDENAR	0.002	0.0001	0	0.1109	0.002	0.0003	0	0.0355	0.002	0.0001	0	0.1109
	CUDAMDENRC	0.0123	0.0007	0	0.7168	0.0121	0.002	0	0.2366	0.0123	0.0007	0	0.7168
	DISTDSDAMS	28484	9457.7	-25	1E+07	38.775	3.1867	-13	176	30127	10003	-25	1E+07
	NUMDSDAMS	18.677	0.2123	0	56	33.353	0.7273	10	52	17.829	0.2104	0	56
	RDDENAR	1861.9	19.042	50.01	15394	2222.6	93.654	108.43	10729	1841.1	19.338	50.01	15394
	INTDENAREA	0.6508	0.0061	0	4.5592	0.6463	0.0219	0.0604	1.7456	0.6511	0.0064	0	4.5592
	INTDENSTRM	0.0008	8E-06	0	0.0093	0.0008	3E-05	4E-05	0.0025	0.0008	8E-06	0	0.0093

123

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Ohio	INTDENRC	3.9627	0.0384	0	46.784	3.917	0.1313	1.0988	12.472	3.9653	0.0399	0	46.784
	CUINTDENAR	0.6601	0.0054	0	2.6883	0.5937	0.0149	0.1418	1.3096	0.6639	0.0056	0	2.6883
	CUINTDENRC	4.0309	0.0323	0	26.061	3.6785	0.0694	1.0988	7.6407	4.0513	0.0339	0	26.061
	CURDDENAR	1737	13.655	50.01	13590	1844.9	35.63	882.62	4432.4	1730.8	14.288	50.01	13590
	CUINTDEN_SM	0.0008	6E-06	0	0.0053	0.0007	1E-05	0.0002	0.0015	0.0008	7E-06	0	0.0053
	POPDENAREA	61.619	2.8254	0	2605.8	84.301	13.876	0.6188	1716.5	60.309	2.8779	0	2605.8
	POPDEN_RC	41.001	1.5314	0	1108.8	45.311	4.659	0	422.35	40.752	1.5974	0	1108.8
	CUPOPDENAR	47.771	1.9149	0	2265	43.005	3.7204	2.9533	490.1	48.047	2.0141	0	2265
	CUPOPDENRC	34.562	1.2019	0	1108.8	30.337	1.8871	0	202.38	34.807	1.2665	0	1108.8
	MINEDENAR	0.0034	0.0002	0	0.1394	0.0075	0.0011	0	0.0739	0.0031	0.0002	0	0.1394
	MINEDENRC	2E-09	3E-10	0	4E-07	4E-09	1E-09	0	1E-07	2E-09	3E-10	0	4E-07
	METALDENAR	3E-05	1E-05	0	0.0183	0.0003	0.0001	0	0.0142	2E-05	1E-05	0	0.0183
	METALDENRC	1E-11	1E-11	0	4E-08	0	0	0	0	1E-11	1E-11	0	4E-08
	CUMIDENRC	2E-09	2E-10	0	2E-07	2E-09	4E-10	0	5E-08	2E-09	2E-10	0	2E-07
	CUMIDENAR	3E-06	1E-06	0	0.0032	3E-05	2E-05	0	0.0032	2E-06	9E-07	0	0.0026
	CUMEDENRC	0	0	0	0	0	0	0	0	0	0	0	0
	CUMEDENAR	0	0	0	0	0	0	0	0	0	0	0	0
	MIN_STMSLP	0.0003	2E-05	0	0.0104	9E-05	5E-05	0	0.007	0.0003	2E-05	0	0.0104
	MAX_STMSLP	0.2067	0.0185	0.0006	30.847	0.3503	0.0789	0.0111	11.995	0.1984	0.019	0.0006	30.847
	MAFLOU_MAX	58.812	5.3085	0.1421	3185.3	302.65	49.86	0.7218	2806.9	44.727	4.6976	0.1421	3185.3
	MAFLOU_MIN	0.0151	0.004	0	12.623	0.0086	0.001	0	0.1325	0.0155	0.0042	0	12.623
	MAFLOU_DIF	58.797	5.3086	0.0511	3185.3	302.64	49.86	0.715	2806.9	44.712	4.6977	0.0511	3185.3
	MAVELU_MAX	0.4716	0.0037	0.2503	3.1759	0.6692	0.0263	0.2677	1.8557	0.4602	0.0035	0.2503	3.1759
	MAVELU_MIN	0.1309	0.0015	0	0.3878	0.097	0.0059	0	0.2606	0.1328	0.0016	0	0.3878
	MAVELU_DIF	0.3407	0.0043	0.0025	2.9387	0.5722	0.0282	0.0901	1.7394	0.3274	0.0041	0.0025	2.9387
	LINK_MAX	6.6506	0.0222	3	13	5.9422	0.0962	4	10	6.6915	0.0226	3	13
	LINK_MIN	4.233	0.0188	2	9	3.474	0.0874	2	7	4.2768	0.0189	2	9
	LINK_DIF	2.4176	0.0136	0	6	2.4682	0.0496	1	5	2.4147	0.0141	0	6
	STRAHL_MAX	3.9612	0.0219	1	8	4.896	0.134	2	8	3.9072	0.0215	1	8
	STRAHL_MIN	1.0032	0.0017	1	5	1	0	1	1	1.0033	0.0018	1	5
	STRAHL_DIF	2.958	0.022	0	7	3.896	0.134	1	7	2.9038	0.0215	0	7
	MAXELEVSMO	435.24	3.9855	116.11	1530.8	481.77	12.668	232.13	878.24	432.55	4.1468	116.11	1530.8
	MINELEVSMO	206.57	2.8164	0	954.69	211.78	13.066	0	728.42	206.27	2.8824	0	954.69
	ELEV_MIN	274.18	2.3838	86	1086	316.44	9.9175	137	744	271.74	2.4485	86	1086

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Ohio	ELEV_MAX	504.7	4.6311	159	1745	537.39	13.937	275	990	502.82	4.8302	159	1745
	ELEV_MEAN	375.23	3.1149	119.43	1274.6	416.81	11.174	193.77	822.07	372.83	3.2259	119.43	1274.6
	RC_EL_MIN	274.51	2.3822	90	1086	316.84	9.8697	144	744	272.06	2.4473	90	1086
	RC_EL_AVG	341.24	2.8818	109.97	1177.1	385.19	10.712	179.78	792.74	338.71	2.9785	109.97	1177.1
	RC_EL_MAX	464.08	4.1401	156	1545	501.95	12.696	259	922	461.9	4.3144	156	1545
	SLP_MIN	0	0	0	0	0	0	0	0	0	0	0	0
	SLP_MAX	77.283	0.5896	5	217	78.647	2.1659	17	154	77.204	0.611	5	217
	SLP_MEAN	17.118	0.2006	0.2377	52.895	12.867	0.4555	0.2377	29.085	17.363	0.2097	0.3439	52.895
	RC_SLP_MIN	0	0	0	0	0	0	0	0	0	0	0	0
	RC_SLP_AVG	16.011	0.1706	0.7094	43.248	12.687	0.4444	1.4574	24.506	16.203	0.178	0.7094	43.248
	RC_SLP_MAX	71.858	0.5501	5	178	72.746	2.0229	17	145	71.807	0.5701	5	178
	NO3_MIN	12.089	0.0329	7.4822	16.775	12.158	0.1496	7.4822	15.684	12.085	0.0338	7.4822	16.775
	NO3_MAX	12.401	0.0336	7.491	16.796	12.727	0.1445	7.501	16.042	12.382	0.0345	7.491	16.796
	NO3_MEAN	12.245	0.0332	7.4833	16.789	12.441	0.1471	7.4845	15.878	12.234	0.0341	7.4833	16.789
	SO4_MIN	18.651	0.0578	12.373	27.665	17.298	0.1945	12.373	24.682	18.729	0.0598	12.373	27.665
	SO4_MAX	19.199	0.059	12.396	27.73	18.004	0.2059	12.4	26.473	19.268	0.061	12.396	27.73
	SO4_MEAN	18.924	0.0584	12.376	27.718	17.647	0.1999	12.376	25.46	18.998	0.0604	12.377	27.718
	PHSOIL	0.8357	0.0056	0	1	0.7011	0.0296	0	1	0.8435	0.0057	0	1
	FOR	0.5923	0.0044	0.0055	0.9963	0.5985	0.0121	0.0095	0.9191	0.592	0.0046	0.0055	0.9963
	FOR_RC	0.5834	0.0035	0.0092	0.9958	0.6059	0.0115	0.133	0.918	0.5821	0.0037	0.0092	0.9958
	CU_FOR	0.0007	5E-06	7E-06	0.0011	0.0007	1E-05	0.0003	0.001	0.0007	5E-06	7E-06	0.0011
	CU_FOR_RC	180.1	21.462	0.045	31028	691.18	124.25	1.7405	8878.8	150.58	21.419	0.045	31028
	H2O	0.0111	0.0006	0	0.9665	0.0309	0.0059	0	0.9665	0.01	0.0005	0	0.3343
	H2O_RC	0.0174	0.0006	0	0.5254	0.0331	0.0037	0	0.5254	0.0165	0.0006	0	0.4037
	CU_H2O	6E-06	2E-07	0	0.0002	2E-05	2E-06	0	0.0002	5E-06	2E-07	0	0.0001
	CU_H2O_RC	0.8388	0.0992	0	140.6	3.384	0.5958	0	41.82	0.6918	0.0985	0	140.6
	AG	0.2622	0.0042	0	0.9492	0.2145	0.0094	0.0021	0.5595	0.2649	0.0044	0	0.9492
	AG_RC	0.2282	0.0036	0	0.9408	0.1589	0.0078	6E-05	0.4023	0.2322	0.0037	0	0.9408
	CU_AG	0.0003	5E-06	0	0.0011	0.0003	9E-06	6E-06	0.0006	0.0003	5E-06	0	0.0011
	CU_AG_RC	10.264	1.1726	0	1674.2	29.905	5.7404	6E-05	409.1	9.1292	1.1922	0	1674.2
	DEV	0.0968	0.0021	0.0009	0.9872	0.1189	0.0091	0.0082	0.7732	0.0955	0.0021	0.0009	0.9872
	DEV_RC	0.1272	0.0019	0.0001	0.9189	0.1423	0.0088	0.0137	0.7336	0.1263	0.002	0.0001	0.9189
	CU_DEV	1E-04	2E-06	1E-06	0.0011	9E-05	4E-06	1E-05	0.0005	1E-04	2E-06	1E-06	0.0011
	CU_DEV_RC	0.13	0.0153	0	22.584	0.4104	0.0788	0	5.5593	0.1138	0.0155	0	22.584

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Ohio	BAR	0.0034	0.0002	0	0.1636	0.0025	0.0004	0	0.0576	0.0034	0.0002	0	0.1636
	BAR_RC	0.0026	0.0001	0	0.1728	0.0017	0.0003	0	0.0317	0.0026	0.0001	0	0.1728
	CU_BAR	4E-06	2E-07	0	0.0001	3E-06	5E-07	0	6E-05	4E-06	2E-07	0	0.0001
	CU_BAR_RC	0	0	0	0	0	0	0	0	0	0	0	0
	HUM	0.3624	0.0046	0.0037	0.9894	0.3359	0.0116	0.0115	0.7771	0.3639	0.0048	0.0037	0.9894
	HUM_RC	0.3579	0.0036	0.0017	0.9821	0.3029	0.0104	0.0138	0.7412	0.3611	0.0038	0.0017	0.9821
	CU_HUM	0.0004	5E-06	4E-06	0.0011	0.0004	1E-05	2E-05	0.0008	0.0004	5E-06	4E-06	0.0011
	CU_HUM_RC	16.459	1.9239	0.0042	2774.7	52.385	9.9317	0.0138	692.47	14.384	1.9462	0.0042	2774.7
	DEC	0.5495	0.0043	0.001	0.9754	0.5485	0.011	0.0065	0.8677	0.5495	0.0045	0.001	0.9754
	DEC_RC	0.5227	0.0033	0.0031	0.9432	0.5225	0.0108	0.098	0.782	0.5227	0.0034	0.0031	0.9432
	CU_DEC	0.0006	5E-06	1E-06	0.0011	0.0006	1E-05	0.0003	0.001	0.0006	5E-06	1E-06	0.0011
	CU_DEC_RC	25.63	3.1062	0.0031	4491.9	97.624	17.928	0.2505	1295.6	21.471	3.1021	0.0031	4491.9
	EVG	0.0214	0.0006	0	0.3217	0.0228	0.002	0.0001	0.1396	0.0214	0.0006	0	0.3217
	EVG_RC	0.0297	0.0009	0	0.4508	0.0441	0.0045	0.0001	0.332	0.0288	0.0009	0	0.4508
	CU_EVG	2E-05	6E-07	0	0.0004	3E-05	2E-06	5E-07	0.0002	2E-05	6E-07	0	0.0004
	CU_EVG_RC	1.5601	0.1835	0	253.28	7.0298	1.1769	0.0001	81.26	1.2441	0.1802	0	253.28
	MXF	0.0214	0.0009	0	0.4807	0.0272	0.0046	0	0.4229	0.021	0.0009	0	0.4807
	MXF_RC	0.0311	0.0013	0	0.7084	0.0393	0.0057	0	0.4777	0.0306	0.0013	0	0.7084
	CU_MXF	3E-05	9E-07	0	0.0005	4E-05	4E-06	0	0.0003	2E-05	9E-07	0	0.0005
	CU_MXF_RC	1.6258	0.1554	0	188.59	5.9264	0.9478	0	64.183	1.3774	0.1538	0	188.59
	GRS	0.0244	0.0006	0	0.4209	0.0101	0.0007	0	0.0433	0.0252	0.0006	0	0.4209
	GRS_RC	0.0254	0.0006	0	0.3168	0.0089	0.0007	0	0.0406	0.0264	0.0007	0	0.3168
	CU_GRS	3E-05	7E-07	0	0.0005	1E-05	7E-07	0	4E-05	3E-05	7E-07	0	0.0005
	CU_GRS_RC	1.0728	0.1175	0	169.67	2.1553	0.499	0	41.798	1.0103	0.1208	0	169.67
	HDV	0.0031	0.0002	0	0.2624	0.0049	0.0011	0	0.1203	0.0029	0.0002	0	0.2624
	HDV_RC	0.0025	0.0002	0	0.2772	0.0053	0.0012	0	0.1303	0.0024	0.0002	0	0.2772
	CU_HDV	2E-06	1E-07	0	0.0002	1E-06	2E-07	0	2E-05	2E-06	1E-07	0	0.0002
	CU_HDV_RC	0.1076	0.0136	0	18.635	0.4535	0.0823	0	5.3543	0.0876	0.0135	0	18.635
	MDV	0.0086	0.0004	0	0.4447	0.0125	0.0022	0	0.2671	0.0083	0.0004	0	0.4447
	MDV_RC	0.0091	0.0004	0	0.3441	0.0132	0.0021	0	0.2285	0.0089	0.0004	0	0.3441
	CU_MDV	7E-06	3E-07	0	0.0003	5E-06	7E-07	0	0.0001	7E-06	3E-07	0	0.0003
	CU_MDV_RC	0.4193	0.054	0	77.454	1.5577	0.2913	0	19.132	0.3536	0.0543	0	77.454
	LDV	0.0237	0.0008	0	0.4653	0.0296	0.0035	0	0.2772	0.0233	0.0009	0	0.4653
	LDV_RC	0.0258	0.0007	0	0.3234	0.0283	0.0029	0	0.2026	0.0257	0.0007	0	0.3234

Region	Variable	All				Surveyed				Unsurveyed			
		Mean	StdErr	Min	Max	Mean	StdErr	Min	Max	Mean	StdErr	Min	Max
Ohio	CU_LDV	2E-05	7E-07	0	0.0005	2E-05	2E-06	9E-08	0.0003	2E-05	8E-07	0	0.0005
	CU_LDV_RC	1.185	0.1466	0	213.06	3.9292	0.7522	0	52.611	1.0265	0.1484	0	213.06
	ODV	0.0615	0.0008	0.0008	0.5078	0.0718	0.0032	0.004	0.2884	0.0609	0.0008	0.0008	0.5078
	ODV_RC	0.0897	0.001	0.0001	0.5394	0.0954	0.0038	0.0137	0.3042	0.0894	0.001	0.0001	0.5394
	CU_ODV	6E-05	7E-07	1E-06	0.0005	7E-05	2E-06	1E-05	0.0002	6E-05	8E-07	1E-06	0.0005
	CU_ODV_RC	4.3533	0.5306	0.0001	768.77	16.129	3.0147	0.0137	213.52	3.673	0.531	0.0001	768.77
	LO	0.0852	0.0015	0.0009	0.8108	0.1014	0.0063	0.0061	0.5655	0.0842	0.0016	0.0009	0.8108
	LO_RC	0.1155	0.0015	0.0001	0.7907	0.1237	0.0061	0.0137	0.5067	0.1151	0.0015	0.0001	0.7907
	CU_LO	9E-05	1E-06	1E-06	0.0009	9E-05	3E-06	1E-05	0.0004	9E-05	1E-06	1E-06	0.0009
	CU_LO_RC	5.5382	0.6768	0.0001	981.83	20.059	3.7634	0.0137	261.16	4.6995	0.6792	0.0001	981.83
	MH	0.0116	0.0006	0	0.7071	0.0174	0.0032	0	0.3874	0.0113	0.0007	0	0.7071
	MH_RC	0.0116	0.0006	0	0.6213	0.0185	0.0033	0	0.3588	0.0112	0.0006	0	0.6213
	CU_MH	9E-06	4E-07	0	0.0005	7E-06	9E-07	0	0.0001	9E-06	5E-07	0	0.0005
	CU_MH_RC	0.5269	0.0675	0	96.089	2.0112	0.3729	0	23.733	0.4412	0.0677	0	96.089
	WET	0.0032	0.0002	0	0.1404	0.0128	0.0018	0	0.1404	0.0026	0.0002	0	0.1345
	WET_RC	0.0092	0.0005	0	0.4028	0.0369	0.005	0	0.3665	0.0076	0.0004	0	0.4028
	CU_WET	3E-06	2E-07	0	0.0001	2E-05	2E-06	0	9E-05	3E-06	1E-07	0	0.0001
	CU_WET_RC	0.5609	0.0663	0	85.725	3.2664	0.4804	0	29.827	0.4046	0.0633	0	85.725
	EMH	0.0005	3E-05	0	0.0677	0.0017	0.0002	0	0.0179	0.0004	3E-05	0	0.0677
	EMH_RC	0.0015	9E-05	0	0.0749	0.0057	0.0008	0	0.071	0.0013	9E-05	0	0.0749
	CU_EMH	5E-07	4E-08	0	8E-05	2E-06	2E-07	0	2E-05	4E-07	4E-08	0	8E-05
	CU_EMH_RC	0.0824	0.01	0	12.835	0.4352	0.0653	0	4.1161	0.0621	0.0098	0	12.835
	HAY	0.1497	0.0025	0	0.8581	0.1392	0.0057	0.0013	0.335	0.1503	0.0026	0	0.8581
	HAY_RC	0.1377	0.0022	0	0.8268	0.1082	0.0057	0	0.3299	0.1394	0.0023	0	0.8268
	CU_HAY	0.0002	3E-06	0	0.001	0.0002	5E-06	3E-06	0.0004	0.0002	3E-06	0	0.001
	CU_HAY_RC	5.9945	0.6811	0	964.8	19.145	3.6147	6E-05	268.31	5.2349	0.6871	0	964.8
	CRP	0.1125	0.0035	0	0.9411	0.0753	0.0055	1E-05	0.3124	0.1147	0.0037	0	0.9411
	CRP_RC	0.0904	0.0029	0	0.9368	0.0507	0.0037	0	0.1904	0.0927	0.0031	0	0.9368
	CU_CRP	0.0001	4E-06	0	0.001	0.0001	6E-06	2E-06	0.0004	0.0001	4E-06	0	0.001
	CU_CRP_RC	4.2692	0.5117	0	709.4	10.76	2.1586	0	171.82	3.8943	0.526	0	709.4
	SHB	0.0062	0.0003	0	0.163	0.01	0.0015	0	0.1137	0.0059	0.0003	0	0.163
	SHB_RC	0.0051	0.0002	0	0.1556	0.0066	0.0009	0	0.0641	0.005	0.0002	0	0.1556
	CU_SHB	7E-06	3E-07	0	0.0002	1E-05	2E-06	0	0.0001	7E-06	3E-07	0	0.0002
	CU_SHB_RC	0.2607	0.0297	0	40.263	1.242	0.1883	0	11.541	0.2041	0.0292	0	40.263