The Role of Physiography in the Relationships Between Land Cover and Stream Fish Assemblages

Jefferson Tyrell Deweber

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science

In

Fisheries and Wildlife Sciences

Emmanuel A. Frimpong
Donald J. Orth
Feng Guo

April 27, 2010
Blacksburg, Virginia

Keywords: stream fish assemblages, land cover, regional framework, watershed classification, physiographic template

The Role of Physiography in the Relationships Between Land Cover and Stream Fish Assemblages

Jefferson Tyrell Deweber

ABSTRACT

Human alteration of the landscape for agricultural and urban land use has been linked to the degradation of streams and stream biota. Natural physical and climatic characteristics, or physiographic template, are important for determining natural land cover and constraining human land use, and are strongly related to stream habitat and stream biotic assemblages. Since the physiographic template differs among watersheds and is an important determinant of the processes being studied, it is important to account for these natural differences among watersheds so that the relationship between land cover and streams can be properly understood. The purpose of this thesis is to develop and assess the utility of a regional framework that classifies watersheds based on physical and climatic predictors of land cover. In Chapter 1, I identified physical and climatic predictors of land cover and classified watersheds into Land cover Distinguished Physiographic Regions (LDPRs) based on these predictors. I was able to identify and create classes based off eight climatic and landform characteristics that determined natural land cover and human land use patterns for both the Eastern and Western U.S. In Chapter 2, I utilized LDPRs to stratify a study region and investigated whether the relationships between land cover and stream fish assemblages varied between these regions. Five commonly used metrics covering trophic, reproductive and taxonomic groupings showed significant variation in their response to agricultural land use across

LDPRs. The results suggest that the physiographic differences among LDPRs can result in different pathways by which land cover alterations impact stream fish communities. Unlike other commonly used regional frameworks, the rationale and methods used to develop LDPRs properly accounts for the causal relationship between physiography and land cover. Therefore, I recommend the use of LDPRs as a tool for stratifying watersheds based on physiography in future investigations so that the processes by which human land use results in stream degradation can be understood.

Acknowledgements

There are a number of people who have made this work possible and I am sure that I will fail to give credit to some who deserve it. First, I would like to express my thanks to my advisor, Emmanuel Frimpong. He has helped guide me throughout my thesis work while allowing me sufficient freedom to learn nearly every task on my own. I am sure that I have become a better scientist, and a better person, because of his leadership. I thank Don Orth and Feng Guo, who have both provided thoughtful comments and revisions that have improved this thesis greatly. This research was funded partially by grants from the Virginia Department of Game and Inland Fisheries and The Edna Bailey Sussman foundation. Tyler Wagner recommended the use of linear mixed models for analysis of Chapter 2. Mike Meador provided auxiliary data from the NAWQA fish sampling that was used in Chapter 2. I would also like to acknowledge the many people who have gone before me to help tell the story of how human actions throughout the landscape result in stream degradation.

Dedication

I dedicate this thesis, along with the hard work and knowledge gained along the way, to my savior Jesus Christ. I suppose that I could have done the work without you, Lord. But without you I certainly would not have had the integrity, peace, joy and stamina that you have given me each and every day. May this work and all future work be pleasing in your sight, and may the man I am becoming along the way bring you praise. Thank you for the new life you have given me, for making me a better person, and for being with me each and every day.

Table of Contents

List of Tables

List of Figures

General Introduction

Streams and rivers are valuable resources that have long intrigued, served, and enlightened mankind through their inimitable qualities. From antiquity, poets have emphasized their beauty, farmers and fishermen have utilized their resources, and merchants have employed them as transportation pathways; in more recent years, ecologists have come to appreciate and study their uniqueness as ecosystems. However, lotic systems are very limited resources that hold only about 0.0003% of the Earth's water (Downes et al. 2002) and are important sources of biodiversity. In North America, a large proportion of the approximately 950 described species of freshwater fishes inhabit lotic systems (Biggs 1986). Human culture is closely tied to rivers and streams, but human populations also threaten the existence of these resources in a number of ways. The human demand for water already approaches an estimated 54% of total accessible runoff and continues to increase, which will necessitate the construction of additional reservoirs in the future (Postel et al. 1996). Postel et al. (1996) also estimate that agricultural activities account for the largest portion of water extracted for human use from freshwater sources, a figure that is sure to rise given the projected increased demand for food supplies. The expansion of cities and intensification of industrial activities have resulted in an array of point and non-point sources of pollution that can seriously impact streams and rivers. However, human perturbations are not limited to water withdrawals, channel alterations or direct inputs of pollutants into surface waters. Human land use can alter the trophic structure of streams, and result in sediments and contaminants from throughout the watershed being carried to the stream channel via indirect pathways. Reductions in point source inputs resulting from regulation revealed that the magnitude of non-point source inputs of pollution from human land use were greater in many areas (Petersen et al. 1987).

Agricultural land use has been highlighted as a major non-point source for increased levels of sediment, nitrogen, phosphorous, and pesticides (Dodds and Oaks 2007; Huryn et al. 2002; Lenat and Crawford 1994; Zhang et al. 2008). Pesticides, in particular, have been shown to directly cause health problems and mortality in invertebrates, fish and humans (Berenzen et al. 2005; Kitamura et al. 2000; Yim et al. 2005). The removal of natural riparian vegetation associated with agriculture results in increased water temperatures and inputs of light; coupled with increasing nutrient runoff, these changes have been associated with increases in autochthonous organic matter (Sponseller et al. 2001). The removal of riparian forest is also associated with reduced inputs of allocthonous organic matter and poor habitat condition (Richards et al. 1996; Roth et al. 1996; Stauffer et al. 2000). Water withdrawal and reduced infiltration can dramatically increase storm flows while reducing base flows (Poff et al. 2006). These various effects can interact in a number of ways to impair stream fish communities. Many studies have associated agricultural activities with lower Index of Biotic Integrity (IBI) scores (Fitzpatrick et al. 2001; Frimpong et al. 2005a; Frimpong et al. 2005b; Wang et al. 1997). Low IBI scores generally reflect decreases in the structure, function, and health of stream fish assemblages (Karr 1981). Higher abundances (Harding et al. 1998), more tolerant species, greater proportions of omnivores, and lower species richness (Horwitz et al. 2008) have been observed in agricultural streams. The direct pathway of influence between agriculture and stream ecosystem health is not often easy to establish, but it is clear that the effects are far reaching.

Urban land use is also a major source of non-point pollution, hydrologic alteration, and degraded fish communities. Many studies have shown that disruption of the natural flow regime is one of the most important alterations caused by urban land use (Arnold et al. 1982; Snyder et al. 2003), as impervious surface area reduces infiltration and typically results in significant increases in storm flow (Poff et al. 2006). Increases in runoff result in the delivery of nutrients, metals, pesticides, and other contaminants that can impair stream health (Lenat and Crawford 1994; Lussier et al. 2008; Sprague and Nowell 2008). Higher stream temperatures are also associated with urbanization due to warming of runoff and more inputs of sunlight due to the deforestation of riparian zones (Paul and Meyer 2001). As with agricultural land use, a number of studies have shown a strong relationship between urban land use and lower fish IBI scores (Roy et al. 2006; Wang et al. 2001). Urbanization has also been associated with the degradation of stream fish communities through the loss of endemics (Scott 2006), lower abundances, removal of intolerant species (Lenat and Crawford 1994), and altered population dynamics (Scott et al. 1986).

Our ability to understand the effects of land use on streams has grown in recent years due to a number of advancements. Although stream ecologists have long recognized the influence of the catchment on the stream (Hynes 1975; Knox 1977), ideas derived from the relatively new field of landscape ecology have increased this perspective greatly (Allan 2004; Johnson and Gage 1997). The importance of advancements in computer technology and the development of Geographic Information Systems (GIS) software cannot be overstated. The coupling of these technological advancements with multivariate and spatial statistics allows us to understand more complex processes at larger extents than was previously feasible (Johnson and Gage 1997). Another asset that has spurred on our capacity for recognizing links between streams and the landscape is the increasing number of publicly available datasets, such as the National Land Cover Dataset (NLCD), the National Hydrography Dataset (NHD), and a great number of other readily available sources that characterize different landscape features (Johnson and Gage 1997).

The importance that these and other developments have had on stream ecology can be seen by making note of the proliferation of land cover studies in the past few years.

In order to practically restore and manage stream ecosystems, we need a clear understanding of the ways in which different types of land cover interact with stream ecosystems. The hierarchical nature of streams and stream processes described by Frissell et al. (1986) provides a useful framework for describing the ways in which streams interact with their landscape. For this reason, a similar hierarchy focused on characterizing the relationships between streams and land use has been developed with three zones representing increasing scale: the reach, the full riparian zone, and the entire catchment upstream of a site. The reach buffer is generally 30 – 200 m in width and up to 2 km in length; the riparian buffer is of similar width to the reach but extends much further longitudinally, often the entire length of the stream network; and the catchment scale is the entire catchment upstream of the sampling site. This hierarchical framework has been used in a number of land use investigations and affords researchers the opportunity of studying three zones for which processes are understood fairly well.

Land use within the reach is responsible for many important processes through localized control of woody debris and litter inputs, stream temperatures, and retention of organic matter and sediments (Ewel et al. 2001; Gregory et al. 1991; Horwitz et al. 2008; Stauffer et al. 2000). At a slightly larger scale, land use within the full riparian zone can be an important determinant of sedimentation, substrate composition, nutrient concentrations, and other abiotic factors that may be transported great distances due to the longitudinal connectivity of streams (Dodds and Oaks 2007; Naiman et al. 1993; Strayer et al. 2003; Vannote et al. 1980). Land use at the catchment scale has been shown to control base flow and determine sediment, nutrient, and contaminant levels (Lenat and Crawford 1994; Stewart et al. 2001; Strayer et al. 2003). Clearly, interactions between land use and streams will differ greatly depending upon proximity to the stream. Thus, identifying the zone within which land use most impacts streams is an important step in understanding the processes that result in degradation.

Despite a large and steadily growing number of multi-scale studies, uncertainty remains regarding the relative importance of human land use near the stream or at a larger scale in determining the health of fish communities. Many studies have concluded that maintenance of a naturally vegetated riparian zone at the reach scale is important for maintaining the health of fish assemblages even if human land use is pervasive at larger scales (Diamond et al. 2002; Frimpong et al. 2005b; Lee et al. 2001). Others have suggested that land cover within the full riparian zone upstream of a site is the best predictor of fish community conditions (Jones et al. 1999; Van Sickle and Johnson 2008; Wilson and Xenopoulos 2008). Alternatively, other investigators have claimed that land cover throughout the catchment is the most important predictor of fish community condition (Long and Schorr 2005; Moerke and Lamberti 2006; Roy et al. 2007; Roy et al. 2006). These contradictory findings show that generalizations about the effects of human land use on stream fishes are not fully possible at this time. The development of effective management strategies for mitigating the impacts of land use hinges upon understanding the processes by which the impacts occur. Bernhardt (2007) discusses some of the methods currently used in the restoration of urban streams, and it is clear that many of these methods are based off of the results of previous studies on the impacts of urban land cover. For example, urban land use is known to greatly impact streams through hydrologic alteration, so urban stream restoration almost always involves altering the way that stormwater reaches the stream by disconnecting impervious surface areas (Bernhardt 2007). It would be an understatement to say that there are many other examples of how watershed management is based upon properly understanding the relationships between land cover and streams.

In an insightful review, Allan (2004) mentions that one important aspect that might perpetuate our inability to better understand stream – land use relationships is covariance of land use with underlying natural characteristics. Several studies have documented the relationship between land cover and landform, soil and geology, which together make the physiographic template (Burgi and Turner 2002; Frimpong et al. 2006; Hietel et al. 2004; Iverson 1988; Reger et al. 2007). Since the physiographic template is also important for determining many properties of streams (Dunne and Black 1970; Horton 1933; Hynes 1975) and shaping biotic assemblages (Burton and Odum 1945; Paller 1994; Wang et al. 2003), it is important to account for differences in physiography among streams. Many investigators have included measurements of various landscape characteristics in their analysis (Fitzpatrick et al. 2001; Frimpong et al. 2005a; Roy et al. 2007; Snyder et al. 2003). However, this method requires spending time and resources, and it is also not a simple matter to select landscape metrics that are ecologically relevant. Other investigators have attempted to account for similarities in watershed characteristics by stratifying the study region through the use of U.S. EPA Ecoregions (Dodds and Oaks 2007; Goldstein et al. 2002; Heitke et al. 2004) or physiographic regions (Baker et al. 2006; Barker et al. 2006). These two commonly used regional frameworks were not designed to control for natural physiographic factors that are related to land cover, which limits their use for investigating the relationships between land cover and streams.

The purpose of this thesis is to describe the development of a regional framework, Physiographically Constrained Land Use/Land Cover Regions (LDPRs), and to demonstrate its utility in studying the relationship between land cover and stream fish assemblages. In Chapter

1, I utilize multivariate regression trees to identify physiographic predictors of land cover in the Eastern and Western U.S. and classify watersheds based on these predictors. I translated these results into a GIS framework to facilitate use in studies. In Chapter 2, I utilize linear mixed models to determine whether relationships between metrics describing aspects stream fish assemblages and land cover types vary between LDPRs. These results are summarized and implications for our understanding of the importance of physiography for stream – land cover interactions, along with related watershed management implications, are discussed.

Chapter 1: Development of a physiographically defined regional framework for investigating interactions between land cover and streams

Abstract – Physiographic characteristics are important determinants of land cover and must be accounted for in investigations aimed at unveiling the relationships between land cover and streams. In this chapter, I describe the development of Land cover Distinguished Physiographic Regions (LDPRs), a regional framework that stratifies watersheds into similar groups based on physiography. I utilized multivariate regression trees to identify physiographic predictors of eight land cover types at multiple extents for the eastern and western U.S. I then named the nodes of the regression tree dendrograms based on physiographic and land cover patterns, and translated the decisions conveyed by the trees into a Geographic Information System (GIS). Eight of the 11 physiographic attributes used were identified as predictors of land cover. Natural land cover was primarily predicted by climatic attributes, with an expected shift from forest to grassland to shrubland as excess precipitation decreased. Agriculture was negatively associated with steep slopes, sandy soils, and warmer temperatures, whereas urban land use showed generally weak relationships with the physiographic template. I discuss the physiographic characteristics and resulting land cover patterns of each of the watershed classes. The relationships between the physiographic template and human activities are complex and likely alter the ways that human land use impacts streams. Unlike other commonly used regional frameworks, the rationale and methods used to develop LDPRs properly accounts for the causal relationship between physiography and land cover. Therefore, I recommend the use of LDPRs as a tool for stratifying watersheds based on physiography in future studies.

Introduction

The negative effects of agricultural and urban land use on streams and stream fish assemblages have been well documented by a large and growing volume of literature (Roy et al. 2007; Snyder et al. 2003; Wilson and Xenopoulos 2008). However, there is little agreement regarding whether land use at the local (i.e. near the stream) or the regional (i.e. at the watershed extent) scale is more important for determining stream health. The development of effective watershed management strategies requires an understanding of the processes by which land cover disturbances at multiple extents impact habitat and biota. Restoration of urbanized and agricultural streams, for example, hinges upon knowing how the impacts occur, attempting to reverse these processes, and then monitoring resulting changes (Bernhardt 2007, Barton 1997). One factor that may lead to inconsistent results across studies is regional variation in the physiographic template, or natural physical and climatic characteristics. The physiographic template is important for determining many properties and processes of streams and also covaries with land cover (Allan 2004; Wang et al. 2004). Thus, investigations must properly account for differences in the physiographic template between watersheds and streams in order to uncover the mechanisms by which land use affects streams and stream biota. In this chapter, I identify physiographic predictors of land cover and develop a regional framework, Land cover Distinguished Physiographic Regions (LDPRs), by classifying watersheds based on these predictors.

Landform and climatic characteristics of a watershed determine many important stream attributes and processes. Steep slopes are often associated with more direct pathways by which storm water reaches the stream and can result in a greater frequency and magnitude of floods (Burt 1992; Ramlal and Baban 2008). Soil and bedrock permeability largely determine groundwater infiltration, controlling stream flow, water chemistry, and many other important hydrologic parameters (Burt 1992; Hynes 1975). The physiographic template also places constraints on different aspects of channel morphology, including number of meanders, wetted width, floodplain size, channel slope and many other important features related to stream ecosystem function and biotic communities (Duval et al. 2004; Frissell et al. 1986; Montgomery 1999; Montgomery and Buffington 1997). These natural differences in processes are also likely to result in different pathways by which land use alters streams (Allan 2004; Brenden et al. 2008; Montgomery 1999; Quinn et al. 2001).

The physiographic template partially determines natural land cover and suitable locations for human land use as well. Studies have found that slope and elevation are positively correlated with forest and grasslands, presumably because areas of high elevation and steep slopes are less suitable for agricultural land use (Frimpong et al. 2006; Hietel et al. 2004; Iverson 1988; Reger et al. 2007). Agricultural land use is also positively associated with various soil characteristics, including soil organic matter content, percentage clay, and moistness (Burgi and Turner 2002; Hietel et al. 2004; Iverson 1988). On the other hand, sandy soils are generally negatively associated with agricultural land use (Burgi and Turner 2002). Although many studies suggest that agricultural land use is strongly constrained by the physiographic setting, I was only able to uncover one study by Iverson (Iverson 1988) that suggests that urbanization is only marginally constrained by water availability.

Due to the importance of the physiographic template for determining land cover patterns, it is important for stream – land cover studies to account for landform and climatic characteristics. Otherwise, it is possible for studies to place too much importance on land cover as a driver of differences in stream biotic assemblages (Allan 2004). Several investigators have heeded this warning and included measurements of various landscape characteristics in their analysis (Fitzpatrick et al. 2001; Frimpong et al. 2005a; Roy et al. 2007; Snyder et al. 2003). A number of different metrics have been used, including watershed area, watershed length, drainage shape, drainage density, stream length, mean channel slope, channel sinuosity, percent slope of the catchment, topographical relief of the catchment, length-slope factor, permeability, and percent sandy surficial deposits (Fitzpatrick et al. 2001; Frimpong et al. 2005a; Roy et al. 2007; Wilson and Xenopoulos 2008). This approach can be very informative and shed light on processes that might otherwise go unnoticed. For example, Snyder et al. (2003) concluded that urban land use had a more disruptive effect on higher gradient streams in their study area. However, this process can require extensive time and resources on measurements and it is rarely a simple matter to select landscape metrics that are ecologically relevant. Furthermore, the same metrics rarely appear in more than a few studies, making comparisons across studies difficult. Such an approach seems largely ineffective if the goal is to understand processes and develop management strategies that transcend a single study area.

A second technique that may be useful for accounting for the physical template is to carry out multi-scale land use investigations within zones of similar physiography. Many studies account for similarities in watershed characteristics by stratifying the study region through the use of U.S. EPA Ecoregions (Dodds and Oaks 2007; Goldstein et al. 2002; Heitke et al. 2004) or physiographic provinces (Baker et al. 2006; Barker et al. 2006). Although the ecoregion and physiographic province can be useful in many cases, neither was developed with land cover-stream studies in mind nor conveys much information regarding stream-watershed interactions. Furthermore, regional land cover was used in the delineation of ecoregions, but inconsistently throughout the country and only in a qualitative and subjective manner (Omernik 1987), which makes the use of ecoregions inappropriate for controlling for underlying physical and climatic characteristics alone. A regional framework that conveys information about stream processes and that is delineated to predict rather than by land cover may prove more informative and useful.

The purpose of this chapter is to determine physiographic predictors of land cover at a broad geographic scale and develop a classification system for selecting watersheds with relatively homogeneous physiographic templates for future investigations through multivariate regression tree analysis. Multivariate regression trees present a unique opportunity to create classes based on multiple response variables that are best predicted by a number of independent variables (De'ath 2002; Everitt and Hothorn 2006). Brenden et al. (2008) used this method to classify river valley segments in Michigan by similarities in fish assemblage attributes as predicted by temperature, catchment area, and channel gradient. In this chapter, I develop a regional framework for classifying watersheds based on physiographic predictors of land cover at both the watershed and riparian extents. I present results of this analysis for both the Eastern and Western United States and discuss the underlying relationships between watershed physiographic characteristics and land cover at multiple extents (30m and 210m riparian buffers and watershed).

Methods

The hydrologic landscape regions (HLR) dataset was developed by the USGS as part of the NAWQA program for the purpose of grouping streams based on characteristics expected to yield similar hydrologic responses (Wolock et al. 2004). In the development of HLRs, Wolock et al. (2004) delineated 43,931 small (~200 km^2) watersheds for the entire U.S. and then used principal components and cluster analysis to assign the watersheds to groups according to similarity in landform, geology and climate. It is important to note that the watersheds delineated by Wolock et al. are truly hydrologic units, which include the watersheds of tributary streams, headwater streams and confluence to confluence reaches. To avoid confusion, I will also refer to these hydrologic units as watersheds throughout this paper. The values of the 13 physiographic variables used in the delineation process were then averaged for each watershed (Wolock et al. 2004) and are included in the publicly available, on-line version of the dataset (USGS 2003). Therefore, the HLR dataset represents a unique, publicly available source of information suited to our purposes. Preliminary analyses using multivariate regression trees for the entire conterminous United States split the HLR dataset roughly into Eastern (Precipitation minus potential evapotranspiration (PMPE) > -7.94 mm/yr) and Western (PMPE < -7.94 mm/yr) halves, with a few exceptions primarily in the Pacific Northwest. Because of the large amount of variation explained by this split, I decided that separate classifications systems for the eastern and western United States were needed in order to have sufficient classification detail.

I then selected two stratified random samples of watersheds for the eastern and western U.S. by randomly selecting 10 percent of the watersheds from each of the original 20 HLR classes. In cases where 10% constituted a sample size smaller than 30, either 30 or all available watersheds for that HLR class were selected in order to include as much of natural variability in the sample as possible. The HLR dataset contains a small number of sliver polygons that are apparently errors resulting from the original watershed delineation process. In order to minimize these errors, watersheds not containing streams were removed from the selected samples and replaced with watersheds containing streams as determined with the USGS National Hydrography Dataset (NHD) of streams (USGS 2000). The final number of watersheds sampled totaled 2200 and 1890 for the eastern and western U.S., respectively.

Land Use/Land Cover Calculations

The 2001 National Land Cover Dataset (NLCD) is available for the conterminous United States at 30 m resolution and was used for all land cover calculations (Homer et al. 2004). The original hierarchical NLCD classes were collapsed into the following eight broad types of land cover: forest, wetland, shrubland, grassland, agriculture, urban, water, and barren. The NHD streams and rivers were clipped from the sampled HLR watersheds and buffered at both 30 and 210 meters. The percent area covered by each of the eight land cover types within the 30 m buffer, 210 m buffer, and full watershed were then calculated. I used ArcGIS 9.2 for all GIS analyses.

Multivariate Regression Trees

Multivariate regression trees (MRTs) analysis was carried out using the mvpart package available for the R statistical environment. This package creates MRTs by successively splitting the original sample based on predictor that maximizes the sums of squared distances between the centroids of the two resulting groups (De'ath 2002). We used the default cross validation method in mvpart, which randomly divides the sample into ten subsets of equal size, then fits models using the other nine subsets and calculates error using the withheld subset. The process is repeated 10 times to give a more accurate representation of cross validation error. Eleven of the 13 physiographic attributes were used as predictors of land cover: watershed area (AREA, square kilometers), average annual precipitation minus potential evapotranspiration (PMPE, mm/yr), average annual temperature (TAVE, °C), average annual precipitation (PPT, mm/yr), an aquifer permeability ranking (AQPERMNEW, ordinal scale of 1-7, lowest - highest), percentage sand (SAND, %), percent total flatland (PFLATTOT, %), percent flatland in upland (PFLATUP, %), percent flatland in lowland (PFLATLOW, percent), watershed relief (RELIEF, m) and average watershed slope (SLOPE, %). Watershed perimeter (PERIMETER, km) was not included as a predictor because it was largely irrelevant for my objectives, and average annual potential evapotranspiration (PET, mm/yr) was not included because of redundancy with other variables.

Regression trees were created for the 30m buffer, 210 m buffer and watershed extents for both the eastern and western U.S. Multivariate regression tree analysis assumes that response variables follow multivariate normal distribution. If all classes of land cover were used in the analysis, the values for each watershed would sum to one, creating a multinomial response and violating the multivariate normal assumption. Therefore, I did not include land cover types that were underrepresented at both the watershed and riparian extents. Shrubland, barren and water were excluded for the eastern, while wetland, barren and water were excluded for the western. Trees were restricted to have a minimum final node size of 50 so that the final classification system would contain a sufficient number of watersheds in each class (an approximate minimum of 500) for future investigations. The MRT regression tree analysis was run 50 times and the simplest tree that fell within 1 standard error of the minimum cross validation error was selected each time (De'ath 2002). For each analysis, the MRT that appeared most often out of the 50 runs was selected as the best and final model.

LDPR Mapping and Comparisons

Before I mapped them in a GIS, Land cover Distinguished Physiographic Regions (LDPRs) were numbered from left to right in order of appearance on the tree and translated into the GIS by following the decisions conveyed by the tree regarding break points in the physiographic attributes. I named classes based on the physiographic template characteristics presented in the trees. The procMeans procedure in SAS 9.2 (SAS Institute Inc. 2007) was used to calculate the mean and range of the independent variables from the sample data for each land cover type used in tree creation. To compare each of the land form and climatic variables between classes, boxplots were created in the R statistical environment. Cross tabulations between the LDPR and HLR classes, and between the watershed and 30m riparian buffer LDPR classes, were carried out using the procFreq procedure in SAS 9.2 (SAS Institute Inc. 2007).

Results

Eastern watershed LDPRs

The selected MRT for watershed land use in the Eastern United States contains 10 final LDPR classes split by 5 landform and climatic variables (Figure 1.1a). The five physiographic variables explain approximately 55.6% of the total variation in the Eastern watershed land cover data, and had a cross validation error of 47.7% (Figure 1.1b). The tree visualizes how watersheds are successively split based on a physiographic attributes that best correlate with land cover, and the amount of variation explained by each split is reflected by the length of the branches for each of the trees. The most important predictor of land cover was precipitation minus potential evapotranspiration (PMPE), explaining 41.4% of the total variance explained by the tree in the first split. This split reveals watersheds with PMPE $\geq$ 233.8 mm/yr go to the left side of the tree and those with less go to the right based on average land cover differences. Successive splits in the tree further divide the groups of watersheds until the final classes are reached at the end nodes. The next split in the tree is defined by SLOPE and explains 22.9% of the explained variance. Four of the variables (PMPE, SLOPE, SAND, and TAVE) appear twice in the tree, while PFLATTOT only appears as a significant predictor in one tree split that explains 4.7% of the explained variance.

The names and characteristics of the 10 classes are given in Table 1.1, and the tree splits show how the physiographic template characteristics constrain these different land cover classes. For example, Eastern class 1.1 (E-1) is characterized by a mix of forest, wetland and agricultural land cover and has a humid climate (PMPE $\geq$ 233.8 mm), flat topography (SLOPE < 0.66%), and sandy soils (SAND $\geq$ 23.5%). Classes E-2 and E-3 have less sandy soils (SAND < 23.5%) and have greater amounts of agricultural land use. E-4 and E-5 are humid with relatively steeper slopes (SLOPE > 0.66%) and are largely dominated by forest land cover. Classes E-6 – E-10 classes are typically less humid (PMPE < 9.2) and generally have a larger proportion of grassland than the first 5 classes, but differ from each other based on other attributes. Some major differences in average land cover between classes result from percent sand, average temperature, and slope. Class E-8 is cooler and has much more agricultural activity than E-7, which is typified by a mix of agriculture, forest, and grassland. A sharp contrast in average land cover can also be seen between E-9 and E-10, higher slopes (SLOPE > 0.68%) are characteristics of watersheds with more grassland than agriculture.

Boxplots that showed differences of a given physiographic predictor between two or more Eastern watershed LDPRs were selected and are shown in Figure 1.2. Since PMPE, TAVE, and SAND appear multiple times in the tree, the variation between classes in each of the corresponding boxplots is not surprising. Minimum elevation (MINELE) does not appear in the model but the boxplot shows a clear trend from typically low elevation watersheds in E-1 to higher elevation watersheds in LDPR E-10. The boxplots also help show differences in physiographic predictors that appear further down and thus only separate two LDPRs in the tree; for example, the corresponding plot shows that E-5 has less total flatland (PFLATTOT) than all other classes, not just E-4. Slope is negatively related to flatland, so it is not surprising to see that LDPR E-5 has a much higher slope than the other classes.

A map of the classification system for the10 Eastern LDPR classes is shown in Figure 1.3 and as expected there are strong spatial patterns for most of the 10 Eastern LDPR classes. Not surprisingly, the wetlands dominated class E-2 appears to be limited almost entirely to areas near the Missisissipi River Delta and the agricultural dominated Midwest is represented well by class E-8. However, some classes (such as E-1 and E-6) do not show such strong spatial correlations but are rather widely distributed. The classification system also closely follows topographical features; for example, the forested classes (E-4 and E-5) are constrained to the regions that we would describe as more mountainous and hilly.

Eastern riparian LDPRs

The 210 meter riparian buffer tree has nine LDPR classes of land cover determined by four physiographic variables which each appear twice in the tree: PMPE, SLOPE, SAND, and TAVE (in order of importance). This tree explained 49.9% of the total variance in land cover data and has a cross validation error of 53.1%. The 210m meter riparian buffer trees are not shown for either the east or west because the 30 meter riparian buffer is probably more representative of actual riparian widths (Lee et al. 2004).

The 30 meter riparian buffer tree for the eastern U.S. (Figure 1.4a) has 10 LDPR classes determined by 5 physiographic characteristics: SLOPE, PMPE, MINELE, SAND, and TAVE. This tree explains 47.5% of the variance and has a cross validation error of 56% (Figure 1.4b). In contrast to the watershed land cover tree, PMPE was second to third in importance in the model and accounted for only 25.2% of the explained variance. Two landform variables, SLOPE and MINELE, explained 38.4% and 25.6% of the variance explained by the model.

TAVE and SAND appeared in the lower branches of the tree and together explained the remaining 10.8%.

The physiographic differences between classes result in land cover signatures in a manner similar to the Eastern watershed classes (Table 1.2). In general, watersheds with SLOPE < 0.69% (ERip-1 – ERip-7) have more agricultural land use in riparian buffers than those with greater slopes. This inverse relationship between agriculture and slope can also be seen by comparing ERip-1 with ERip-2. The model also implies that there is an inverse relationship between temperature and riparian agriculture, which distinguishes ERip-3 from ERip-4 and ERip-6 from ERip-7. Watersheds with sandy soils (ERip-5) also typically have less agricultural land use within their riparian buffers.

A map of the different Eastern 30 m riparian buffer LDPRs (Figure 1.5) and the cross tabulation results (Table 1.3a) comparing the classification of watersheds under the 30 m riparian buffer and watershed LDPRs both show that some of the watershed classes fit nicely into one or a few of the riparian classes, while others vary widely. The strongest relationship was between the watersheds of E-10 and ERip-11, which cross classified 96.3% of the time (Table 1.3a). Most of the watersheds classified into a given watershed LDPR class fall into a number of riparian classes, which reveal how differences between watershed and riparian land use patterns generally prevail under similar physiographic conditions. One clear example is that the largely agriculturally dominated watersheds of class E-8 may either be typically dominated by agriculture (ERip-7), or have more natural land cover in the form of grasslands and forest (ERip-8).

Western watershed LDPRs

The regression tree for watershed land cover in the western U.S. contains 9 LDPR classes split by 5 landform and climatic variables (Figure 1.6a). This tree explains 45.7% of the total variation in watershed land cover, and has a cross validation error of 57.9% (Figure 1.6b). The amount of variation explained by each split in the tree is represented by branch length in Figure 1.6a. PMPE was the most important predictor, explaining 42.2% of the total variance explained by the tree in the first split and appearing two other times. SLOPE was the second most important, explaining 23.7% of the total tree variance, followed by PFLATTOT, and MINELE. In contrast to the Eastern tree, PMPE was the only variable to appear more than once and was the only climatic variable present.

The names and land cover characteristics of the LDPR classes are shown in Table 1.4. The first major split shows that Classes W-1 and W-2 have more humid climates (PMPE >= -7.94) and generally less shrubland in comparison to the other classes. Watersheds of class W-2 are characterized by greater than 70% average forest cover on steep topography, while W-1 watersheds are less steep and have an almost even mix of different land cover types. W-1 also has a representation of urban land cover (6.9%) than all other classes. Watersheds in classes W-7 – W-9 are all relatively more arid (-13.4 mm/yr <=PMPE < -7.9 mm/yr) but have different land cover signatures from each other because of topographical variables. Class W-7 watersheds are lower (MINELE < 1254) and have greater amounts of grassland than do those of class W-8, which have almost even amounts of shrubland and grassland. Class W-9 has large amounts of flatland (PFLATTOT >=63.5%), and is dominated by almost equal amounts of grassland and agricultural land cover.

Selected boxplots show differences in the mean and range of physiographic characteristics among western watershed LDPRs (Figure 1.7). PMPE appears as an important classifying variable three times in the tree and it is not surprising to see these differences in the corresponding plot. The RELIEF plot also primarily shows the relationship depicted in the tree, as classes W-6 – W-9 have less relief than W-3 – W-5. Precipitation (PPT) is directly related to PMPE and varies between classes in similar ways, while TAVE is inversely related as shown by the watersheds of classes W-4 and W-6. The inverse relationship between SLOPE and PFLATOT is clear and it is interesting how much these two characteristics vary between many classes even though each only appears as a predictor once in the tree. The plot for RELIEF primarily provides visualization for the relationships shown in the tree, as classes W-6 – W-9 have less relief than others.

A map of the Western watershed LDPRs (Figure 1.8) that shows that there is some spatial aggregation of watersheds with similar physiographic templates and land cover characteristics, but that there is more variability than in the East. Even with this variability, LDPR classes generally follow well defined topographical features or climatic regions. The forest dominated watersheds of W-2 closely follow mountainous regions, and it is apparent that the slightly less steep W-1 watersheds are spatially associated with class W-2. The low precipitation watersheds of W-3 are well distributed, while the more arid class W-4 which receives more precipitation but has less excess moisture (see PMPE boxplot in Figure 1.7) are primarily constrained to the Southwest. The flat, high elevation watersheds of W-8 are spread out over wide areas and show very little spatial aggregation.

Western riparian LDPRs

The western tree for land cover within 210m riparian buffers explained 48.6 % of the variance and had a cross validation error of 55.2%. The model had 9 classes of watersheds with different land cover signatures split by 5 physiographic predictors. As for the east, the 210m trees of the west are not shown or used in subsequent analyses.

The 30m riparian buffer tree (Figure 1.9a) has 11 classes of watersheds with different land cover signatures separated by 4 physiographic predictors. This model explained 51.8% of the total variance and had a cross validation error of 52.3% (Figure 1.9b). PMPE was the most important variable and accounted for 40.1% of the explained variance in the first split and an additional 21.8% in two other branches. SLOPE, the only landform variable, accounted for 26.3% of the explained variance in three splits, while PPT and TAVE explained the remaining 11.7%.

A map of the Western LDPR classes for land cover within the 30m riparian buffer is shown in Figure 1.10. A visual comparison between this classification system and the watershed LDPRs shown in Figure 1.8 suggests that land cover within riparian buffers is more spatially aggregated despite the additional classification resolution, as there are fewer isolated watersheds of a given riparian class. The watersheds of classes WRip-6 and WRip-7, for example, are located in fairly continuous blocks. Table 1.3b shows the results of the cross tabulation between the watershed and 30m riparian buffer LDPRs for the Western U.S. The cross tabulation results are similar to those from the Eastern U.S. and show that there are both differences and similarities between the watershed and riparian classification systems.

Discussion

The multivariate regression tree analysis identified relationships between characteristics of the physiographic template and typical land cover patterns. The topographic and geologic predictors of land cover in my analyses generally agree with the results of other studies. In this study, watersheds with low slope are typified by large amounts of agricultural land use, which agrees with a number of studies carried out across different spatial extents (Frimpong et al. 2006; Hietel et al. 2004; Iverson 1988; Reger et al. 2007). Sandy soils have been shown in previous studies to constrain agriculture (Burgi and Turner 2002); this is clearly shown in LDPRs E-1 and E-6, which have relatively sandy soils and less agricultural land use. While agriculture differs greatly, urban land use differs only moderately between LDPRs, which agrees with the results of Iverson (Iverson 1988). Urban land use at the watershed extent for the Eastern LDPRs ranged from an average of 5.2% in class E-2 to 11.3% in class E-6. In the western U.S., average urban land use at the watershed extent ranged from 0.8% in W-3 to 6.9% in W-1. Landform characteristics are also important in explaining the signatures of natural land cover across classes, especially in the western U.S. where the primary human land use of our analysis (agriculture) is less prevalent. For example, watersheds with lower minimum elevation are typically dominated by grassland (e.g. W-7), while those with higher minimum elevation typically have more shrubland (e.g. W-8).

My analysis covers a large extent across which climate varies widely, so it is not surprising that climatic characteristics appear as important predictors of both natural land cover and human land use. Average annual precipitation minus potential evapotranspiration was the most important predictor of land cover overall in this study and appears to primarily determine natural land cover at regional extents. This variable effectively split watersheds into those that were characterized by forest, wetland, grassland or shrubland land cover at different extents and only within 30 meter riparian buffers in the eastern U.S. was PMPE not the most important predictor. Average annual temperature and precipitation were also important predictors of land cover in my analysis. Of particular note is that warm temperatures are negatively associated with agricultural land use in the eastern U.S., which can be seen by comparing classes E-7 with E-8 and ERip-3 with ERip-4. Previous investigations interested in uncovering predictors of land cover have primarily focused on topographical and soil characteristics within an extent that would not vary appreciably in climatic factors and have thus not identified these climatic variables.

Several of the physiographic predictors of land cover identified by the analyses are known to be important for determining natural stream characteristics and processes, including the processes by which human activities in the watershed impact streams. Slope is of particular importance because of the natural differences between high and low gradient streams, and the ways in which these different types might respond to human impacts (Frimpong et al. 2005a; Snyder et al. 2003). Percent sand in soil is closely related to soil permeability and is related to other geologic properties, which have been related to stream fish assemblage responses in land cover investigations (Moerke and Lamberti 2006; Weigel et al. 2004). Streams in watersheds with highly permeable soils might be more impacted by land cover alterations throughout the watershed due to bypassing of streamside vegetation. Differences in temperature and precipitation can also presumably be important parts of the stage upon which land cover interacts with streams through a number of pathways, such as frequency and magnitude of floods, droughts and overland runoff.

Land cover Distinguished Physiographic Regions (LDPRs) also closely follow real regions of land cover that we can easily identify in the U.S., such as the largely agricultural Midwest or the various mountainous regions dominated by forest. The degree to which the classes mirror these regions is somewhat surprising because the amount of variation in land cover explained by even the best models is fairly low. While the low amount explained variance might at first seem like a drawback, it is an important characteristic because different types of land cover need to vary within a region in order for observational studies to identify the impacts of land cover alterations. Thus it is important that the watersheds within a class have different land cover attributes that vary around the average due to factors unaccounted for in our models. As I mentioned, this will enable investigators interested in using LDPRs for stratification in stream-land cover investigations to find study sites with a wide range of a given land use type within a single class. Then the relationship between a stream response variable of interest and land cover predictors can be identified for sites with a similar physiographic template, and even compared among sites with different templates. Investigations may be able to utilize such an approach in the future, as it is easier to implement than measuring and accounting for physiographic variables on a study-by-study basis and may give more meaningful and interpretable results than stratifying watersheds through the use of existing regional frameworks.

My primary aim in performing this analysis was to create a regional framework for classifying watersheds that could be used to account for natural differences in the physiographic template that need to be considered in land use investigations. A number of river classification systems and related regional frameworks aimed at creating classes based on ecological similarity have been developed in recent years (Brenden et al. 2008; Quinn et al. 2001; Snelder and Biggs 2002), but none of these meet this specific need. Multivariate regression trees turned out to be a powerful tool for identifying physical and climatic variables that covary with land cover. The clear dendrogram format was easily translated into a map format through relatively simple GIS techniques. I recommend this technique for the development of regional frameworks for various purposes, and for similar watershed classification systems at smaller regional extents with higher classification accuracy provided that data of sufficient resolution are available. Land cover Distinguished Physiographic Regions (LDPRs) may be a tool that will help us to better understand and mitigate the impacts of human activities on streams and their biota. This regional framework may also prove useful for identifying unimpaired reference streams that share a common physiographic template, since every LDPR has watersheds with minimal human land use.

Conclusions

This analysis has effectively identified landform and climatic predictors of land cover and classified watersheds based on these variables at multiple extents. Multivariate regression tree analysis proved to be a good method for this analysis, and I was easily able to translate the output into a GIS and classify watersheds into Land cover Distinguished Physiographic Regions (LDPRs) of the Eastern and Western U.S. The most important predictor of land cover was average annual precipitation minus potential evapotranspiration and appears to primarily determine natural land cover at regional extents. Other important physiographic predictors included average annual temperature, slope, minimum elevation, relief, percent total flatland, and percent sand in soil. Agricultural land use was shown to be strongly related to several of these physiographic attributes and was a key land cover type for differentiating LDPRs. Urban land use was less prevalent at all extents and was not strongly related to physiography, although there are differences among some LDPRs. Several of these physical and climatic characteristics that predicted land cover can also be important for determining the structure and function of stream ecosystems. The relationships between underlying natural characteristics, land cover, and stream ecosystems can make it difficult to identify the impacts of human land use on streams and stream biota. The processes by which land cover alteration impacts streams are also complex and likely differ between regions with underlying physiographic differences. I propose that LDPRs be used to stratify regions based on physiography and to tease apart the underlying relationships between the physiographic template and land cover. More specifically, LDPRs are a useful tool for identifying minimally impacted reference streams with a similar physiographic template, selecting streams for bioassessment and biomonitoring programs, and stratifying watersheds in land cover investigations. I also propose that the spatial and classification resolution of LDPRs be improved upon in the future as higher resolution data become available in the future.

Table 1.1 Eastern watershed LDPR names and land cover characteristics.
Mean and range ( ) of the percent of total area covered by 8 land cover types represent typical land cover signatures associated with physiographic attributes. Shrubland, barren and water were not used in the creation of classes. Abbreviations in names: 'ag'= agriculture, 'grass' = grassland.

Code	LDPR Name	Forest	Wetland	Grassland	Agriculture	Urban	Shrubland	Water	Barren
E-1	Humid, sandy forest-ag-wetland watersheds	34.9	17.5	3.7	32.4	9.9	3.9	4.7	0.9
		(0, 90.8)	(0, 80.2)	(0, 27.2)	(0, 87.2)	(0, 94.4)	(0, 45.9)	(0, 99.8)	(0, 5.2)
E-2	Humid, warm, wetland-ag watersheds	6.5	42.4	1.8	30.7	5.2	6.5	10.1	0.2
		(0, 61.2)	(0, 91.0)	(0, 32.1)	(0, 96.9)	(0, 36.3)	(0, 29.0)	(0, 97.1)	(0, 4.6)
E-3	Humid, ag dominated watersheds	19.8	7.4	0.9	58.5	8.8	2.2	2.3	0.1
		(0, 95.0)	(0, 47.8)	(0, 9.6)	(0, 94.6)	(0, 68.8)	(0, 21.8)	(0, 21.3)	(0, 3.1)
E-4	Humid, forest-ag watersheds	55.7	5.7	2.8	20.7	8.3	3.7	2.6	0.2
		(6.2, 97.2)	(0, 36.9)	(0, 20.3)	(0, 86.8)	(0, 84.9)	(0, 22.6)	(0, 32.2)	(0, 6.0)
E-5	Humid, forest dominated watersheds	74.3	2.2	1.5	12.5	5.9	1.7	1.7	0.3
		(19.4, 99.8)	(0, 26.7)	(0, 19.8)	(0, 58.6)	(0, 44.4)	(0, 20.9)	(0, 32.1)	(0, 6.0)
E-6	Semi-humid, warm, very sandy,forest-ag-wetland watersheds	33.4	20	5.3	23.6	11.3	1.9	4.1	0.4
		(0, 88.8)	(0, 100)	(0, 33.6)	(0, 78.1)	(0, 93.0)	(0, 19.5)	(0, 30.2)	(0, 11.4)
E-7	Semi-humid, warm, ag-forest-grassland watersheds	25.8	7.3	18.9	33.9	7.2	2.6	4.1	0.1
		(0.5, 69.3)	(0, 89.8)	(0, 75.4)	(0.1, 71.4)	(0.6, 86.8)	(0, 18.3)	(0, 36.2)	(0, 1.7)
E-8	Semi-humid, ag dominated watersheds	14.3	3.2	4.3	66.7	8.6	0.3	2.5	0
		(0, 85.8)	(0, 77.9)	(0, 85.5)	(1.3, 96.9)	(0.3, 88.0)	(0, 6.7)	(0, 78.4)	(0, 1.8)
E-9	Semi-arid, ag-grassland watersheds	4.1	3.5	21.9	58.5	7.6	1.2	3.0	0.1
		(0, 46.3)	(0, 21.0)	(0, 96.8)	(0, 93.9)	(0, 80.3)	(0, 20.9)	(0, 59.2)	(0, 5.8)
E-10	Semi-arid, grassland-ag watersheds	7.5	1.9	51	31.2	5.3	1.2	1.7	0.2
		(0, 47.6)	(0, 7.1)	(1.8, 95.5)	(0, 91.5)	(0, 91.4)	(0, 45.8)	(0, 21.1)	(0, 6.3)

Table 1.2 Eastern riparian LDPR names and land cover characteristics.
Mean and range ( ) of the percent of total area of 30m riparian buffers covered by 8 land cover types represent typical land cover signatures associated with physiographic attributes. Shrubland, barren and water were not used in the creation of classes. Abbreviations in names: 'ag'= agriculture, 'grass' = grassland.

Code	LDPR Name	Forest	Wetland	Grassland	Agriculture	Urban	Shrubland	Water	Barren
ERip-1	Very flat, wetland-ag riparian	7.3 (0, 59.6)	41.9 (0, 100)	1.8 (0, 27.9)	27.8 (0, 95.7)	6.9 (0, 78.9)	1.1 (0, 12.6)	12.9 (0, 100)	0.4 (0, 6.0)
ERip-2	Flat, wetland-forest riparian	25.8 (0, 85.6)	45.8 (0, 100)	2.1 (0, 14.1)	8.3 (0, 51.1)	5.4 (0, 100)	2.7 (0, 24.2)	9.6 (0, 98.6)	0.4 (0, 9.8)
ERip-3	Semi-humid ag-grass-forest-wetland riparian	13.3 (0, 54.5)	12.2 (0, 67.4)	31.7 (0, 74.7)	28.3 (0, 88.7)	5.4 (0, 46.8)	0.5 (0, 13.4)	8.4 (0, 77.3)	0.3 (0,16.7)
ERip-4	Cool, semi-humid, ag-grass-wetland riparian	2.9 (0, 33.0)	11.4 (0, 52.1)	15.9 (0, 82.1)	58.7 (9.4, 100)	4.6 (0, 18.7)	0.2 (0, 4.9)	6.2 (0, 46.1)	0.0 (0, 1.9)
ERip-5	Very sandy, wetland-forest-ag riparian	26.4 (0, 80.0)	37.4 (1.1, 100)	2.9 (0, 12.0)	18.1 (0, 79.2)	4.9 (0, 47.2)	0.7 (0, 7.3)	9.5 (0, 59.1)	0.1 (0, 1.1)
ERip-6	Warm, humid, forest-ag-wetland riparian	37.4 (0, 100)	17.2 (0, 65.7)	3.1 (0, 53.3)	29.6 (0, 91.7)	4.3 (0, 66.3)	1.9 (0, 16.8)	6.3 (0, 70.6)	0.1 (0, 4.2)
ERip-7	Cool, humid, ag-forest-wetland riparian	19.8 (0, 67.7)	10.6 (0, 72.3)	3.9 (0, 24.0)	52.3 (0, 93.6)	7.7 (0.5, 63.5)	0.4 (0, 15.8)	5.2 (0, 62.7)	0.0 (0, 2.2)
ERip-8	Humid, forest-ag-wetland riparian	38.1 (0.1, 100)	12.1 (0, 76.1)	4.5 (0, 64.4)	31.3 (0, 87.8)	5.0 (0, 57.9)	0.7 (0, 8.1)	8.2 (0, 48.3)	0.1 (0, 3.3)
ERip-9	Very humid, low elevation, forest-wetland riparian	47.3 (0, 86.0)	26.5 (0, 85.0)	0.9 (0, 12.0)	8.2 (0, 61.3)	4.7 (0, 51.3)	3.1 (0, 14.4)	9.1 (0, 100)	0.2 (0, 2.7)
ERip-10	Very humid, forest dominated riparian	62.2 (12.4, 99.8)	7.7 (0, 79.2)	1.9 (0, 23.2)	13.7 (0, 81.3)	6.5 (0, 76.4)	1.4 (0, 18.5)	6.4 (0, 58.6)	0.2 (0, 16.5)
ERip-11	Semi-humid, gras-ag-forest riparian	16.0 (0, 71.6)	9.4 (0, 50.7)	42.4 (0, 96.5)	20.1 (0, 98.2)	3.1 (0, 65.8)	1.0 (0, 28.7)	7.9 (0, 99.9)	0.2 (0, 5.5)

Table 1.3 a-b Cross tabulation for watershed and riparian LDPRs.
The percentage of watersheds in a given Watershed LDPR class (columns) that fall into the different Riparian LDPR classes (rows) are shown for the Eastern (a.) and Western (b.) U.S. Rows sum to 100.

a.

Code	ERip-1	ERip-2	ERip-3	ERip-4	ERip-5	ERip-6	ERip-7	ERip-8	ERip-9	ERip-10	ERip-11
E-1	9.5	58.0	-	-	13.8	11.0	7.8	-	-	-	-
E-2	78.8	21.2	-	-	-	-	-	-	-	-	-
E-3	11.2	14.0	-	-	-	53.1	21.7	-	-	-	-
E-4	-	2.0	-	-	0.2	1.6	0.7	0.7	34.5	60.2	-
E-5	-	-	-	-	-	-	-	0.2	9.4	90.4	-
E-6	9.1	24.7	-	1.3	44.8	-	-	19.5	-	-	0.6
E-7	9.9	6.2	1.2	-	-	37.0	-	37.0	-	-	8.6
E-8	-	-	2.8	2.8	-	10.1	42.4	39.6	-	-	2.4
E-9	2.0	5.2	53.6	39.2	-	-	-	-	-	-	-
E-10	-	-	3.8	-	-	-	-	-	-	-	96.3

b.

Code	WRip-1	WRip-2	WRip-3	WRip-4	WRip-5	WRip-6	WRip-7	WRip-8	WRip-9	WRip-10	WRip-11
W-1	32.8	67.2	-	-	-	-	-	-	-	-	-
W-2	5.3	9.8	84.9	-	-	-	-	-	-	-	-
W-3	-	-	-	-	2.2	-	22.1	66.7	1.1	-	8.0
W-4	-	-	-	-	-	-	-	-	58.2	2.5	39.2
W-5	-	0.5	0.5	0.5	0.9	5.5	59.5	-	3.6	1.4	27.7
W-6	-	-	-	-	-	-	-	23.6	54.0	19.5	2.9
W-7	-	0.4	-	7.1	10.2	64.2	11.9	-	0.9	3.5	1.8
W-8	-	-	-	-	25.4	46.3	14.9	4.5	-	1.5	7.5
W-9	-	-	-	24.1	11.4	53.8	-	-	3.8	7.0	-

Table 1.4 Western watershed LDPR names and land cover characteristics.
Mean and range ( ) of the precent of total area covered by 8 land cover types represent typical land cover signatures associated with physiographic attributes. Wetland, barren and water were not used in the creation of classes. Abbreviations in names: 'ag'= agriculture, 'grass' = grassland, 'shrub' = shrubland.

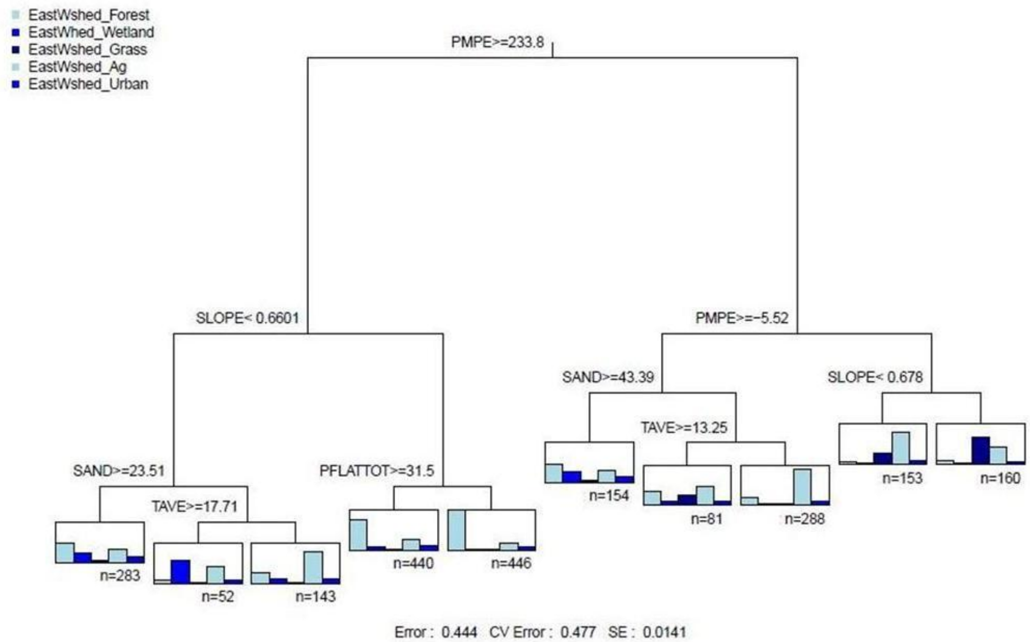
Code	Name	Forest	Shrubland	Grassland	Agriculture	Urban	Wetland	Water	Barren
W-1	Humid, forest-shrub-grass-ag watersheds	26.2	22.4	19.9	17.5	6.9	3.6	3.1	0.4
		(0, 93.1)	(0, 100)	(0, 94.9)	(0, 91.8)	(0, 88.6)	(0, 93.1)	(0, 100)	(0, 5.8)
W-2	Humid, steep, forest dominated watersheds	62.6	19.3	10.0	2.6	1.7	1.0	1.1	1.7
		(0, 96.4)	(0, 94.1)	(0, 75.6)	(0, 85.4)	(0, 42.9)	(0, 12.6)	(0, 26.8)	(0, 49.3)
W-3	Arid, very low precipitation, shrub dominated watersheds	10.2	72.2	8.4	3.3	0.8	0.6	0.4	4.1
		(0, 79.2)	(0.9, 100)	(0, 64.2)	(0, 68.8)	(0, 14.8)	(0, 12.1)	(0, 22.5)	(0, 98.6)
W-4	Arid, shrub dominated watersheds	6.6	70.6	11.2	1.9	4.9	0.4	0.2	4.3
		(0, 58.4)	(2.7, 99)	(0, 71.8)	(0, 47.8)	(0, 94.4)	(0, 10.5)	(0, 3.8)	(0, 79.8)
W-5	Semi-arid, shrub-forest-grass watersheds	27.7	39.8	21.4	5.4	3.1	0.9	0.6	1.1
		(0, 93.7)	(0, 99.3)	(0, 94.2)	(0, 92.8)	(0, 66.3)	(0, 24.5)	(0, 39.9)	(0, 50.5)
W-6	Semi-arid, shrub-ag-grass watersheds	1.8	55.6	16.9	18.9	2.5	1.2	0.7	2.4
		(0, 52.4)	(0, 100)	(0, 98.4)	(0, 93.7)	(0, 22)	(0, 15.5)	(0, 24.9)	(0, 92.7)
W-7	Semi-humid, low elevation, grassland dominated watersheds	5.4	14.6	59.9	15.3	1.5	1.6	0.9	0.7
		(0, 55.1)	(0, 92.9)	(0, 99.3)	(0, 89.4)	(0, 18.6)	(0, 48.3)	(0, 25.1)	(0, 45.7)
W-8	Semi-humid, high elevation, shrub-grass watersheds	1.6	46.6	39.9	7.8	1.4	1.5	0.2	1.1
		(0, 35.9)	(0, 100)	(0, 99.2)	(0, 77.5)	(0, 10.9)	(0, 11.5)	(0, 2.9)	(0, 16.6)
W-9	Semi-humid, flat, grass-ag watersheds	3.1	12.8	41.9	36.1	3.7	1.3	0.8	0.3
		(0, 26.1)	(0, 95.2)	(1.1, 96.1)	(0, 94.5)	(0, 47.9)	(0, 13.9)	(0, 14.5)	(0, 25.5)

Table 1.5 Western riparian LDPR names and land cover characteristics.
Mean and range ( ) of the percent of total area covered by 8 land cover types represent typical land cover signatures associated with physiographic attributes. Wetland, barren and water were not used in the creation of classes. Abbreviations in names: 'ag'= agriculture, 'grass' = grassland, 'shrub' = shrubland.

Code	LDPR Name	Forest	Shrubland	Grassland	Agriculture	Urban	Wetland	Water	Barren
WRip-1	Humid, forest-wetland-ag riparian	43.2 (0, 88.4)	9.8 (0, 45.5)	6.1 (0, 36.7)	11.8 (0, 74.3)	6.8 (0, 51.5)	15.1 (0, 55.3)	6.1 (0, 65.3)	1.0 (0, 29.8)
WRip-2	Semi-humid, shrub-grass-forest-ag riparian	18.7 (0, 93.2)	26.5 (0, 100)	21.9 (0, 95.0)	14.0 (0, 90.7)	4.0 (0, 78.1)	10.1 (0, 100)	4.5 (0, 100)	0.3 (0, 11.9)
WRip-3	Steep, forest dominated riparian	65.1 (0, 100)	14.8 (0, 73.4)	6.7 (0, 77.7)	3.1 (0, 91.2)	1.7 (0, 27.9)	4.9 (0, 51.7)	2.8 (0, 50.7)	0.8 (0, 44.2)
WRip-4	Warm, flat, forest-shrub-grass riparian	26.5 (0, 62.5)	26.0 (0, 59.9)	18.4 (0, 69.8)	9.0 (0, 47.9)	2.5 (0, 31.4)	11.4 (0, 91.4)	5.8 (0, 69.0)	0.4 (0, 12.1)
WRip-5	Arid, flat, cool, shrub-grass-ag riparian	0.4 (0, 6.7)	35.9 (0, 100)	32.9 (0, 91.3)	18.2 (0, 83.9)	1.6 (0, 22.7)	7.6 (0, 80.7)	3.1 (0, 30.6)	0.3 (0, 3.2)
WRip-6	Semi-arid, cool, flat, grassland dominated riparian	3.3 (0, 55.3)	10.0 (0, 87.1)	60.0 (0, 100)	14.6 (0, 94.3)	1.3 (0, 10.9)	7.1 (0, 84.0)	2.9 (0, 49.7)	0.8 (0, 39.8)
WRip-7	Slightly steep, shrub-forest-grass riparian	24.0 (0, 95.2)	44.1 (0.1, 100)	16.8 (0, 80.6)	5.4 (0, 53.0)	1.9 (0, 64.2)	6.7 (0, 57.3)	0.8 (0, 14.6)	0.4 (0, 17.4)
WRip-8	Arid, shrubland dominated riparian	5.2 (0, 92)	74.1 (0, 100)	6.3 (0, 59.2)	4.8 (0, 90.4)	0.9 (0, 14.2)	2.8 (0, 38.6)	1.5 (0, 54.8)	4.4 (0, 100)
WRip-9	Semi-arid, warm, shrubland dominated riparian	4.9 (0, 54.3)	67.5 (0.5, 99.6)	7.5 (0, 61.2)	5.4 (0, 73.2)	4.8 (0, 82.7)	6.7 (0, 53.3)	1.3 (0, 29.6)	1.9 (0, 84.9)
WRip-10	Flat, semi-arid, grass-shrub-ag riparian	1.0 (0, 21.2)	25.0 (0, 100)	52.5 (0, 99.4)	12.9 (0, 79.4)	2.1 (0, 14.8)	2.6 (0, 24.0)	2.6 (0, 72.3)	1.3 (0, 36.8)
WRip-11	Slightly steep, semi-arid, shrub-grass-forest riparian	14.6 (0, 100)	54.1 (0, 98.6)	18.9 (0, 93.3)	4.2 (0, 91.7)	2.2 (0, 18.9)	2.4 (0, 31.2)	2.1 (0, 42.6)	1.6 (0, 18.9)

Figure 1.1 a – b Eastern watershed tree and cross validation error plot.
The dendrogram shows the relationship between physiographic predictor and percent cover for five land cover types at the watershed extent. The original sample of 2200 watersheds is split successively by physiographic variables and the branches represent the amount of variance explained by each split. The number of watersheds (n) in each class is shown below each histogram. The cross validation error plot (b.) shows cross validation error (y-axis) as a function of tree size (x-axis). The horizontal line represents the minimum cross validation error plus one standard error, and the tree with the fewest classes below this line is selected.

a.



b.

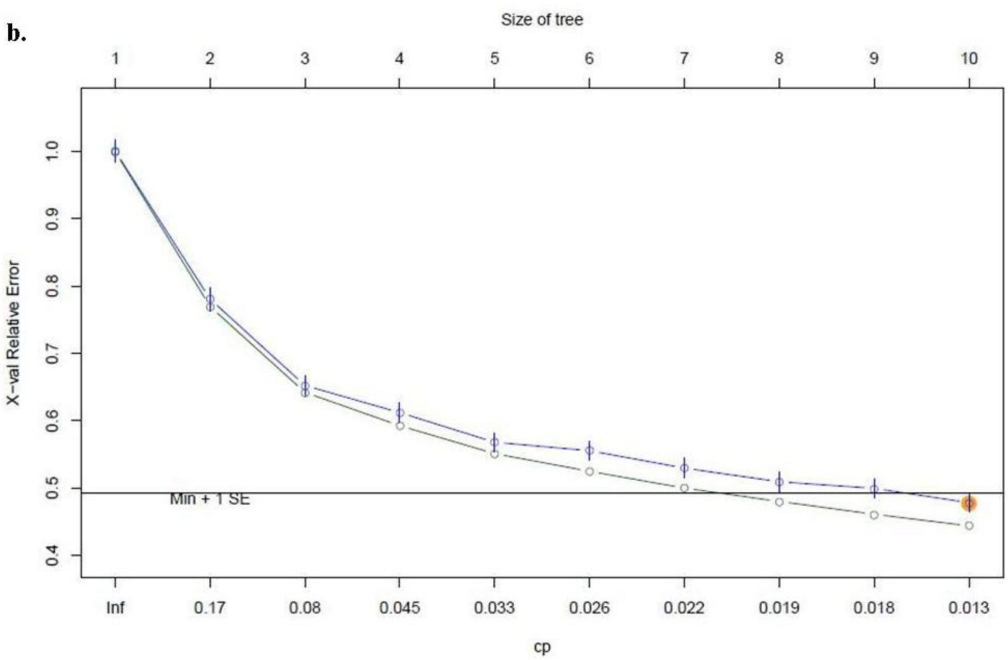
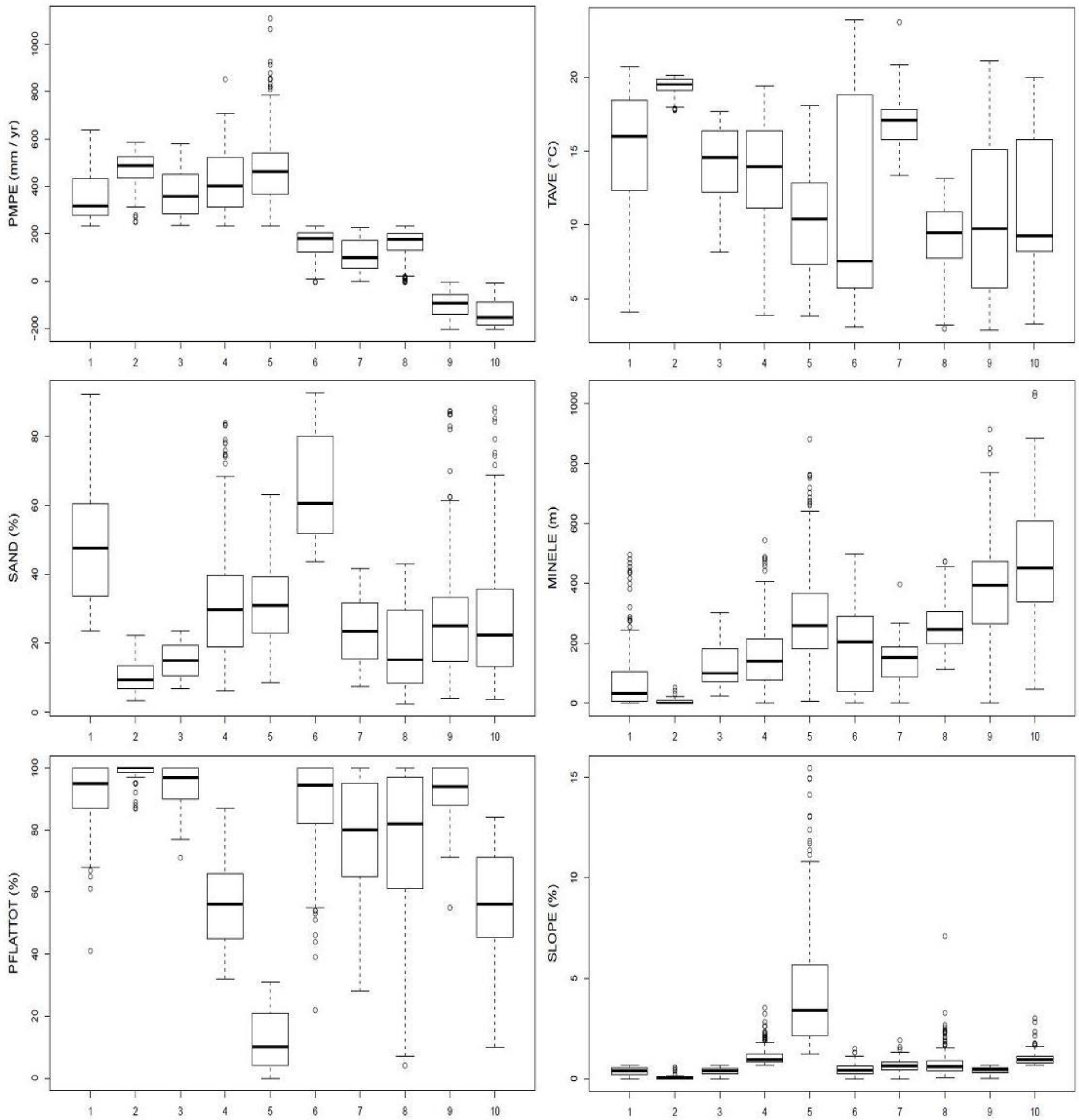


Figure 1.2 Eastern watershed LDPR boxplots.
Selected boxplots compare physiographic characteristics between the 10 eastern watershed LDPR Classes.



Eastern watershed LDPR classes E-1 thru E-10

Figure 1.3 Eastern watershed LDPR map.
Map created by translating the decisions of the multivariate regression tree analysis into a Geographic Information System (GIS). The map shows the distribution of watersheds having similar physiographic characteristics that coincide with land cover at the watershed extent for the eastern U.S.

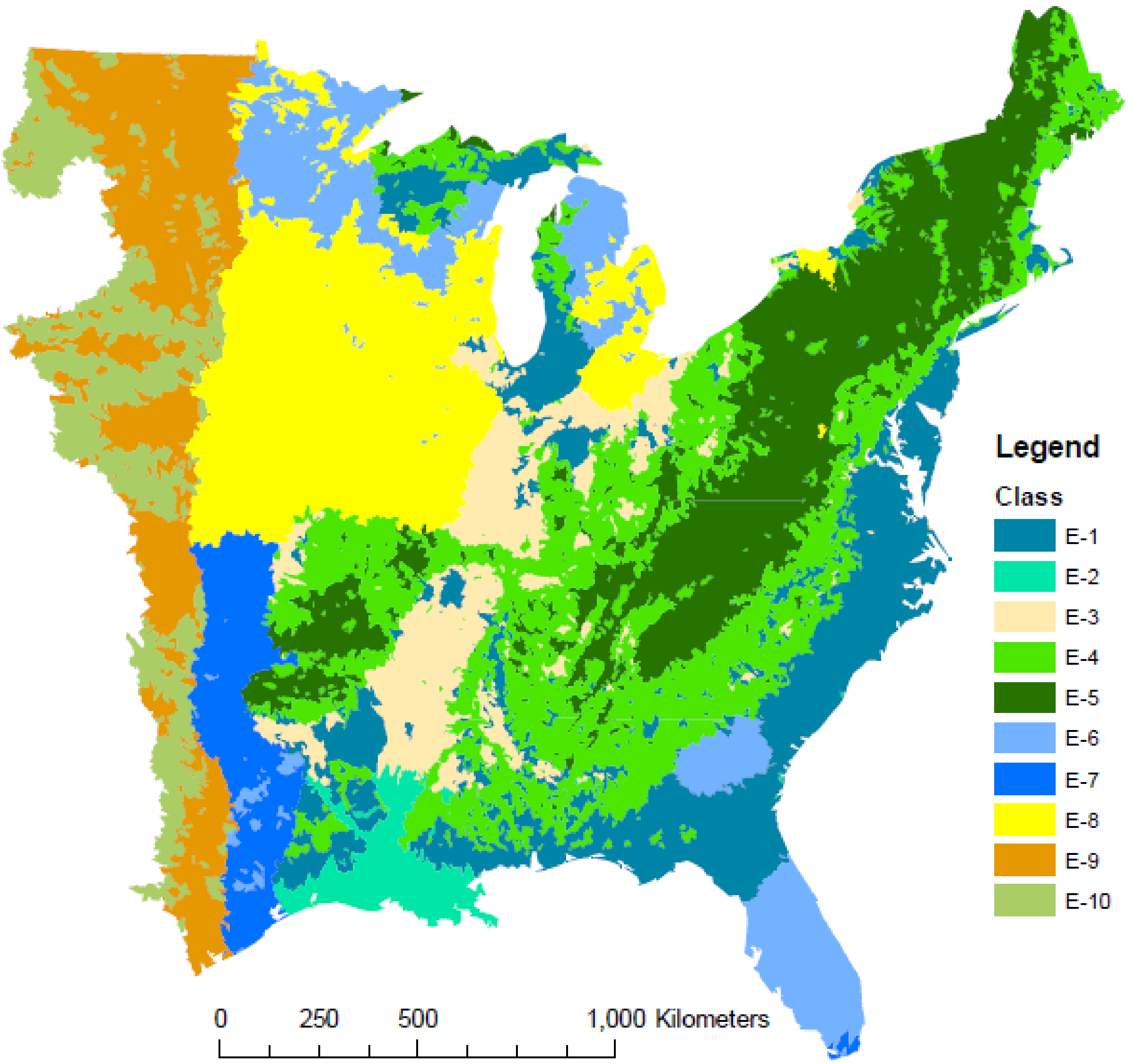
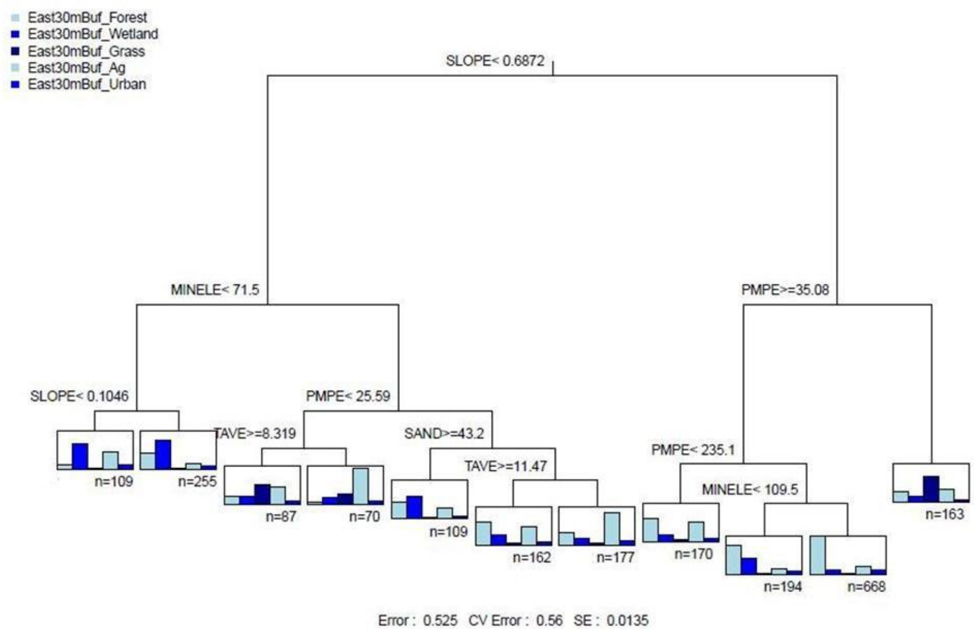


Figure 1.4 a – b Eastern riparian tree and cross validation error plot.
The dendrogram shows the relationship between physiographic predictor and percent cover for five land cover types within a 30 meter riparian buffer. The original sample of 2200 watersheds is split successively by physiographic variables and the branches represent the amount of variance explained by each split. The number of watersheds (n) in each class is shown below each histogram. The cross validation error plot (b.) shows cross validation error (y-axis) as a function of tree size (x-axis). The horizontal line represents the minimum cross validation error plus one standard error, and the tree with the fewest classes below this line is selected.

a.



b.

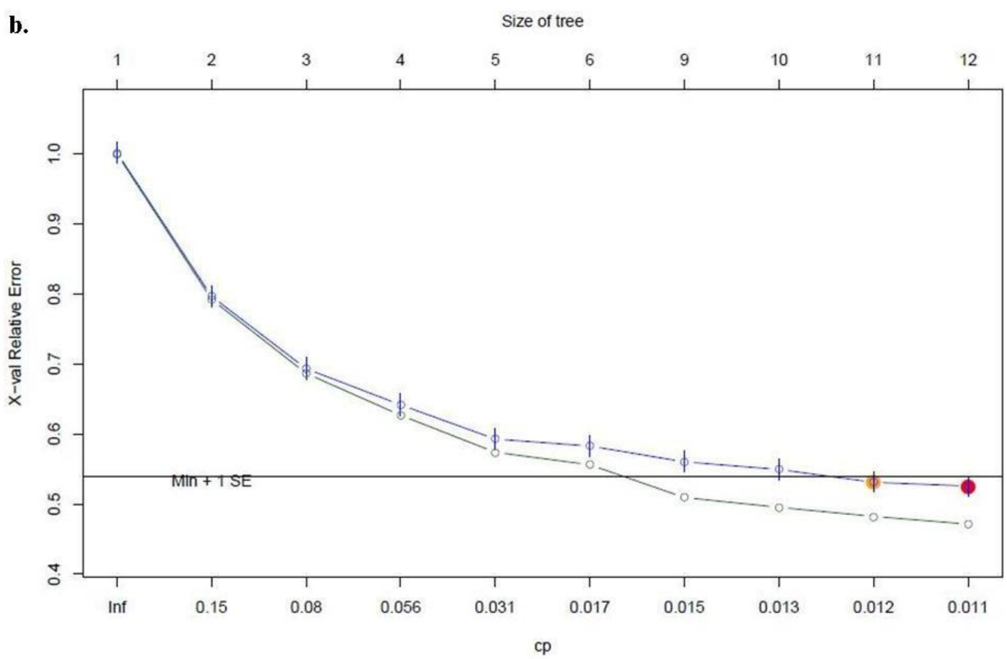


Figure 1.5 Eastern riparian LDPR map.
Map created by translating the decisions of the multivariate regression tree analysis into a Geographic Information System (GIS). The map shows the distribution of watersheds having similar physiographic characteristics that coincide with land cover within 30 meter riparian buffers for the eastern U.S.

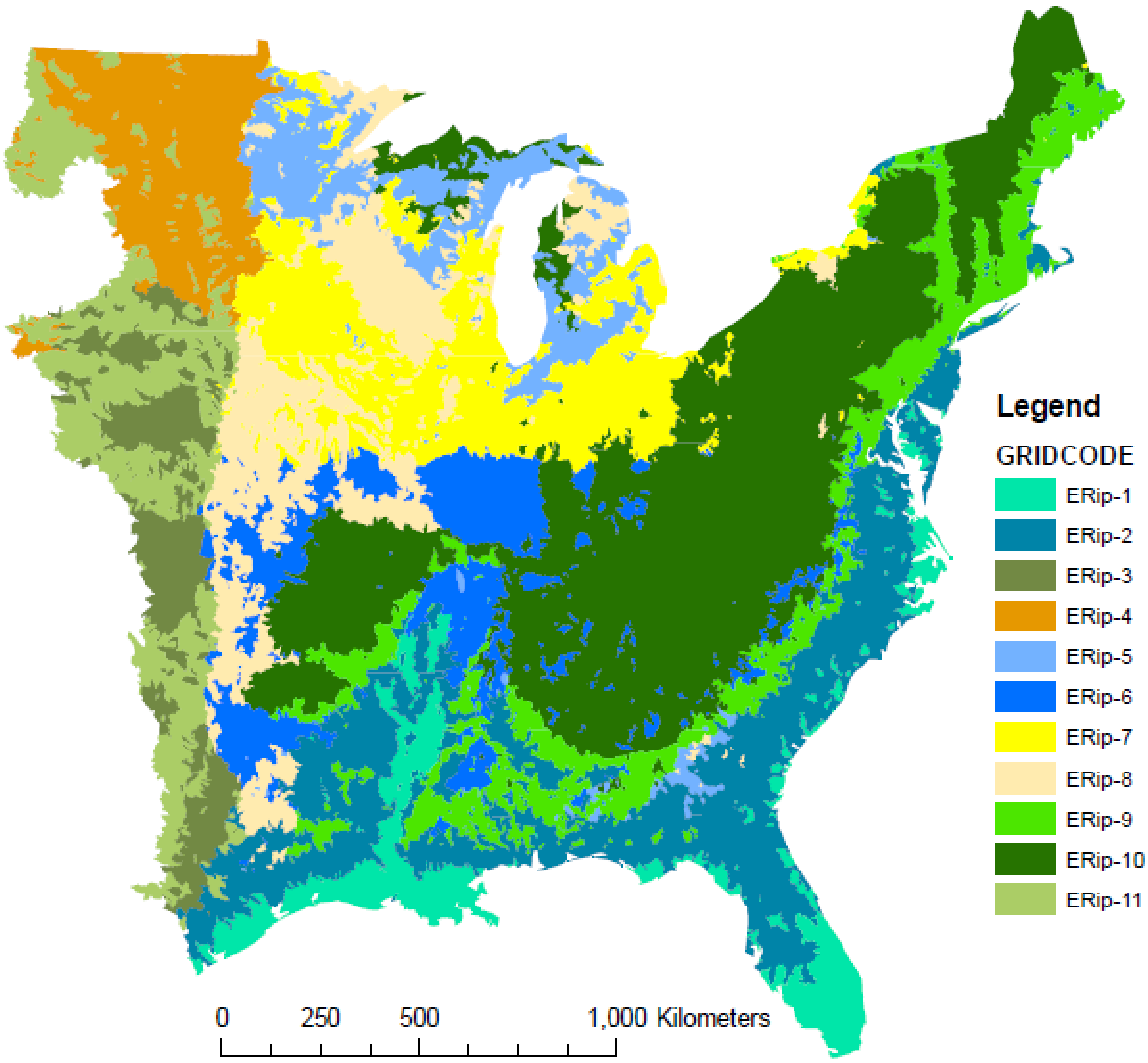
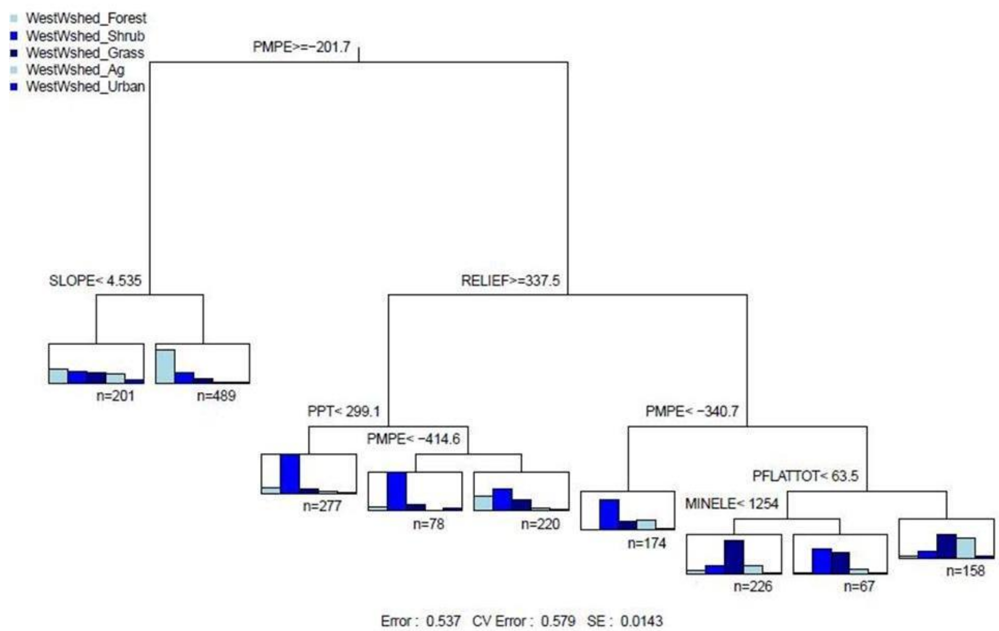


Figure 1.6 a – b Western watershed tree and cross validation error plot.
The dendrogram show relationships between physical and climatic predictors and the percent cover of five types of land cover at the watershed extent. The original sample of 1890 watersheds is split successively by physiographic variables and the branches represent the amount of variance explained by each split. The number of watersheds (n) is each class is shown below each histogram. The cross validation error plot shows the cross validation error (y-axis) as a function of tree size (x-axis). The horizontal line represents the minimum cross validation error plus one standard error, and tree with the least number of nodes below this line is selected.

a.



b.

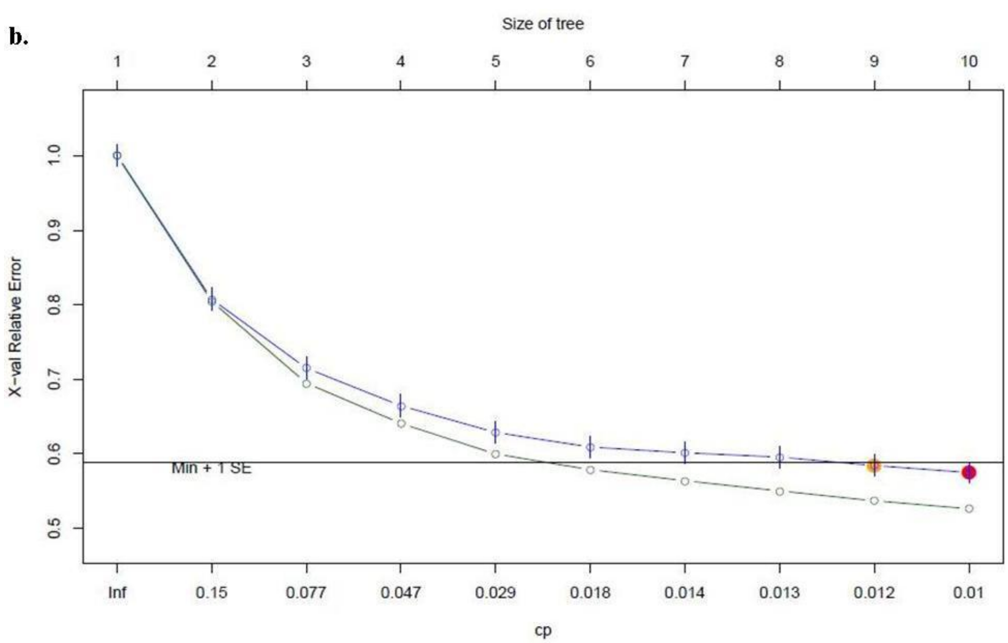


Figure 1.7 Western watershed LDPR boxplots.
Selected boxplots compare physiographic characteristics between nine western watershed LDPR classes.

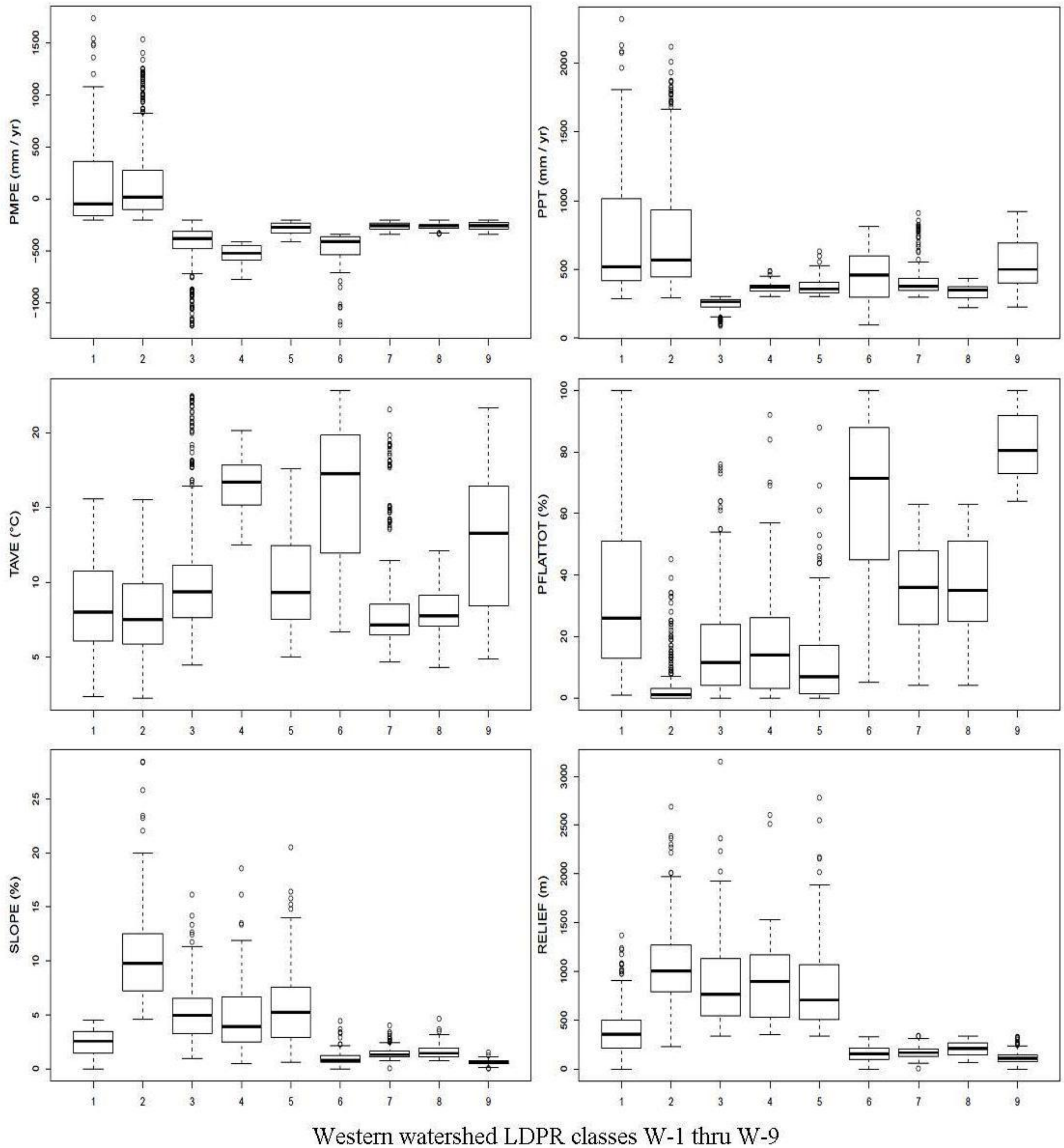


Figure 1.8 Western watershed LDPR map.
Map created by translating the decisions of the multivariate regression tree analysis into a Geographic Information System (GIS). The map shows the distribution of watersheds having similar physiographic characteristics that coincide with land cover at the watershed extent for the western U.S.

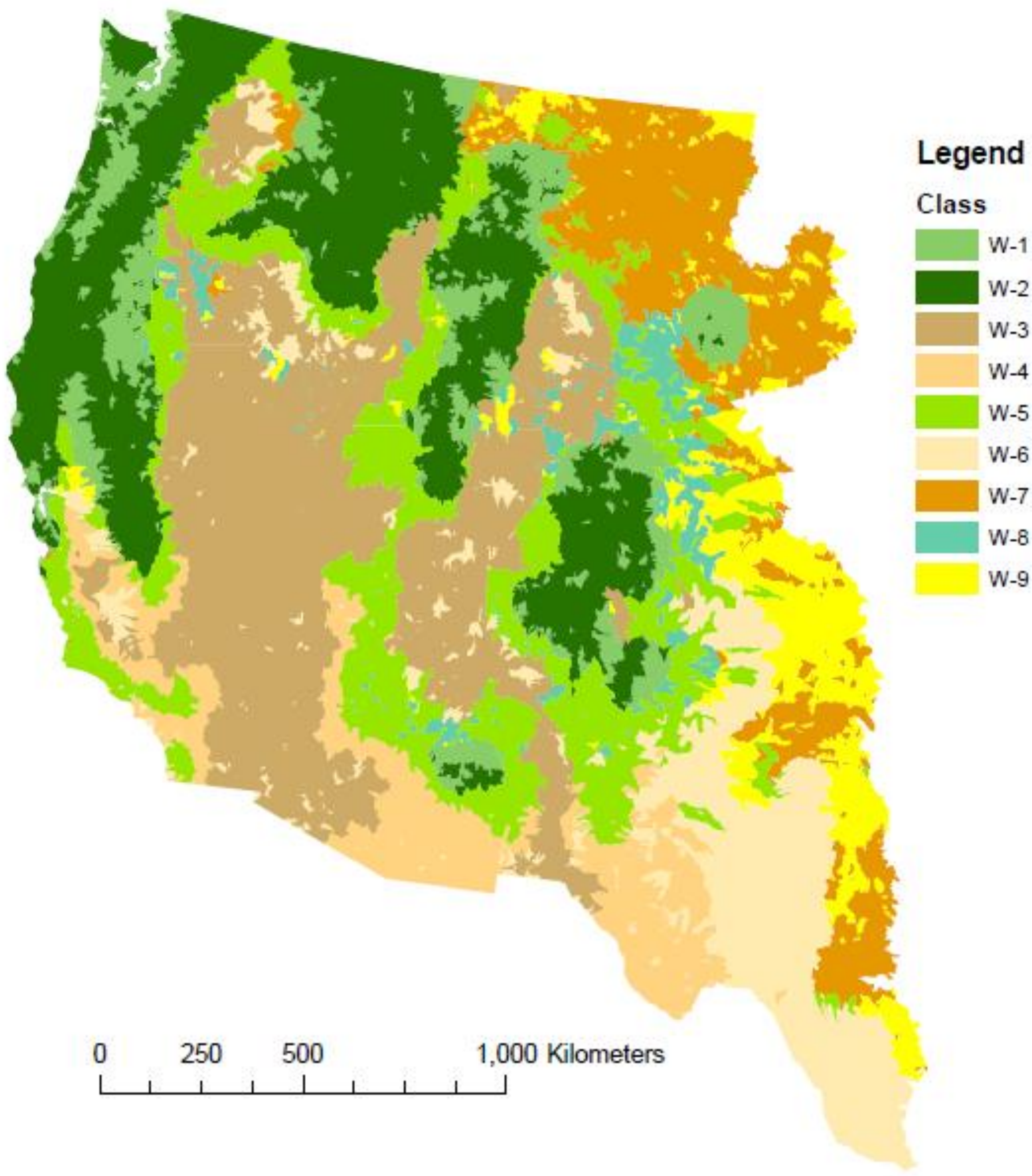
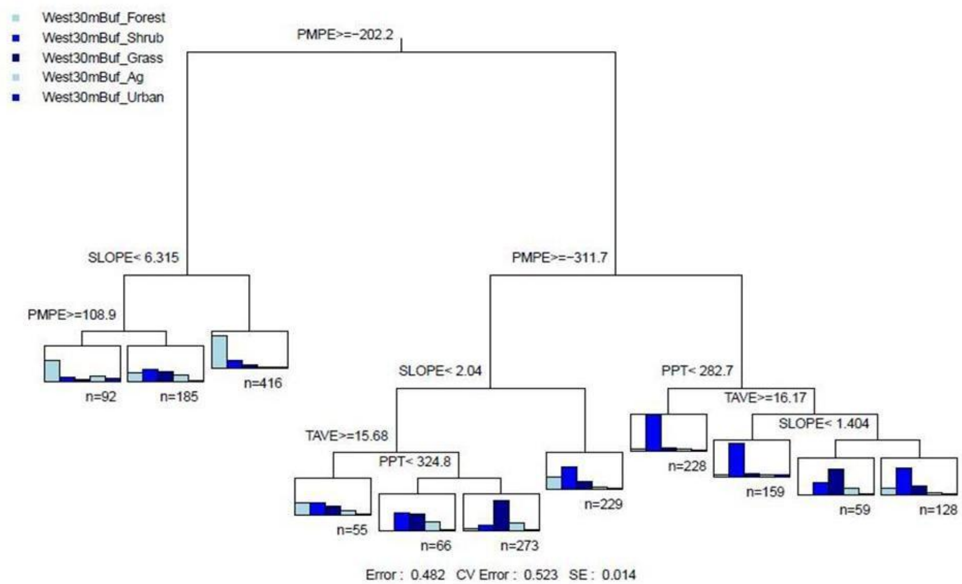


Figure 1.9 a – b Western riparian tree and cross validation error plot.
The dendrogram show relationships between physical and climatic predictors and the percent cover of five types of land cover within 30 meter riparian buffers. The original sample of 1890 watersheds is split successively by physiographic variables and the branches represent the amount of variance explained by each split. The number of watersheds (n) is each class is shown below each histogram. The cross validation error plot shows the cross validation error (y-axis) as a function of tree size (x-axis). The horizontal line represents the minimum cross validation error plus one standard error, and tree with the least number of nodes below this line is selected.

a.



b.

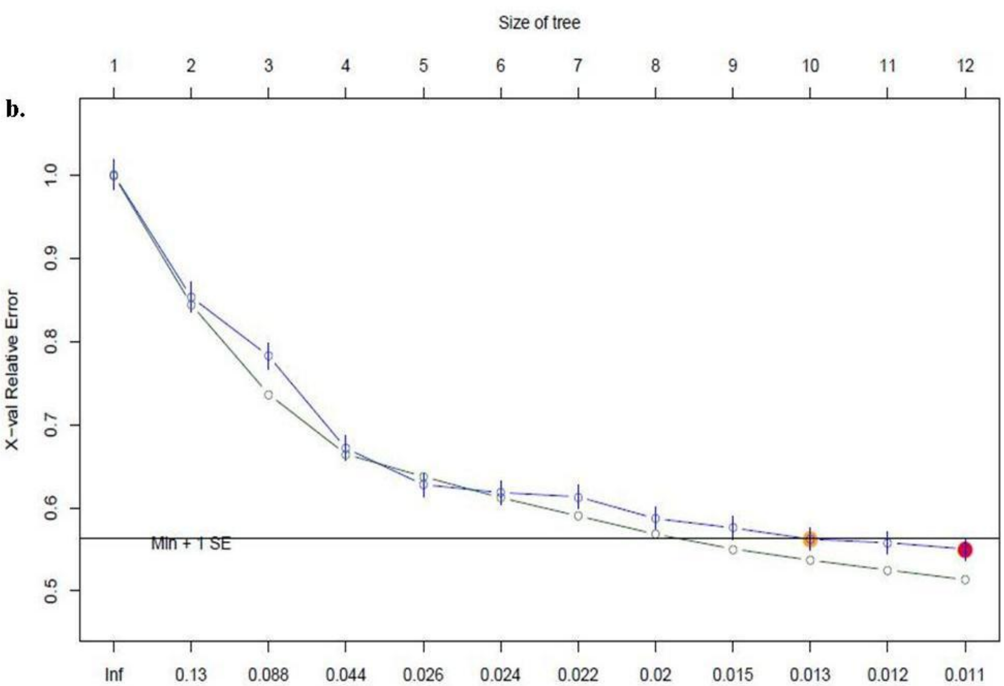
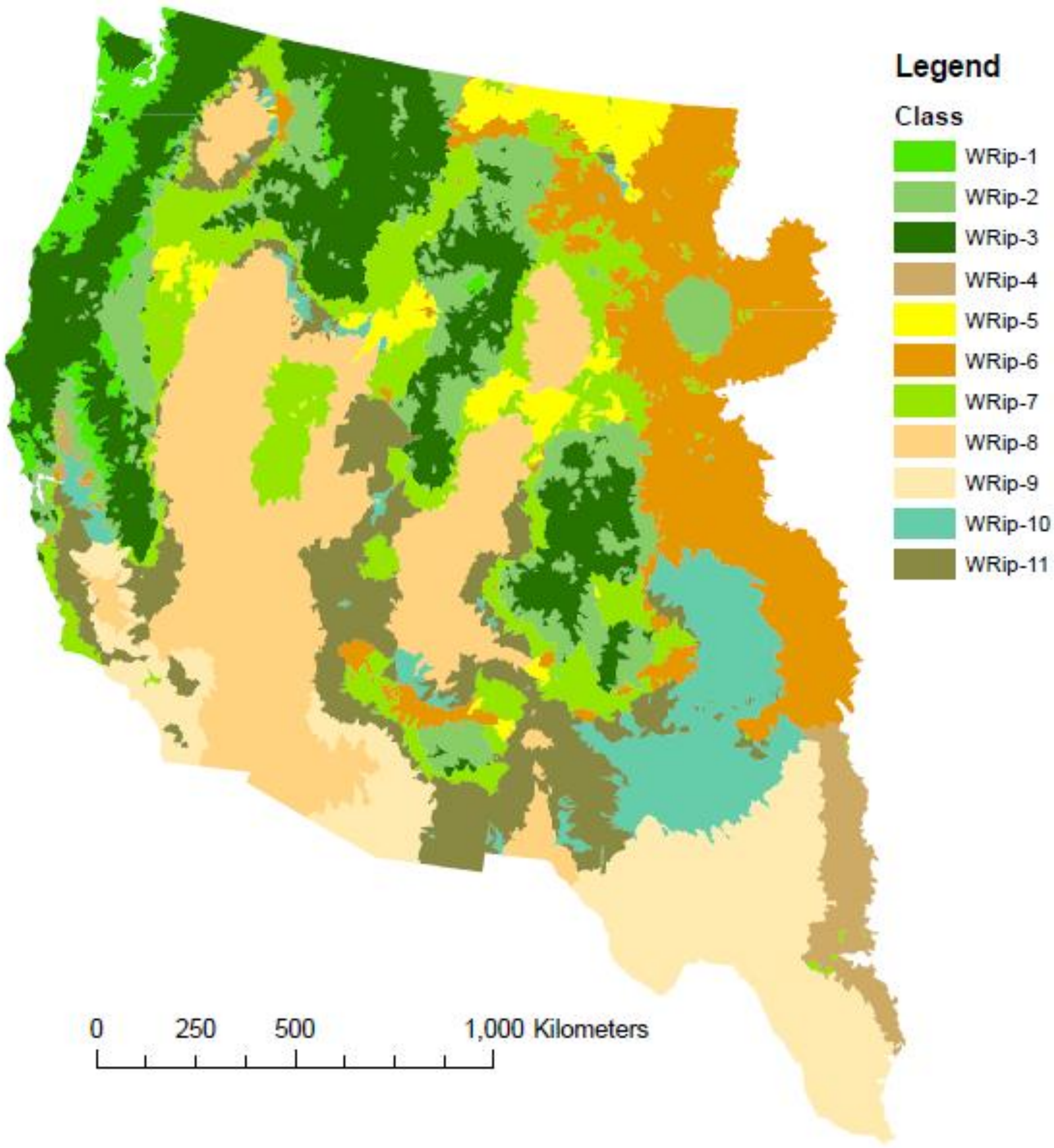


Figure 1.10 Western riparian LDPR map.
Map created by translating the decisions of the multivariate regression tree analysis into a Geographic Information System (GIS). The map shows the distribution of watersheds having similar physiographic characteristics that coincide with land cover within 30 meter riparian buffers for the western U.S.



Chapter 2: Regional variation in the relationships between land cover and stream fish assemblages

Abstract – The physiographic template is an important determinant of stream fish assemblage characteristics and land cover characteristics, and must be accounted for in investigations of the relationships between land cover and fish assemblages. In Chapter 1, I developed Land cover Distinguished Physiographic Regions (LDPRs) as a watershed classification scheme to account for differences in physiography. In this chapter, I use LDPRs to stratify the watersheds of 161 streams in the Upper Mississippi River Basin. I then used linear mixed models analysis to determine whether the relationships between fish metrics commonly used in bioassessment and land cover varied across LDPRs, as determined by the presence of significant random effects slopes and/or intercepts. The proportions of forest, natural, agriculture and urban land cover were calculated at both the watershed and riparian (30 m x 1000 m) extents and used as predictors for five fish metrics: proportion of benthic invertivore species, proportion of centrarchid species, proportion of darter, madtom and sculpin species, proportion of omnivore species, and proportion of substrate generalist species. All metrics had models with significant random intercepts, and the proportion of centrarchid species, proportion of darter, madtom and sculpin species, and proportion of omnivore species all had models with significant random slopes as well. The effects of agricultural land use at the watershed extent on these three metrics showed the most variation across LDPRs, whereas the effects of urban land use did not vary. These results suggest that there are regional differences in the processes by which land cover interacts with streams and reiterates the importance of accounting for underlying physical and climatic features in further studies. I suggest that LDPRs or a similarly delineated watershed

classification system be used in future studies in order to better understand the impacts of land use on streams within a given region and compare these impacts across regions.

Introduction

The relationships between streams, stream biota and land cover have been well studied and are reviewed thoroughly by Allan (2004) and Wang et al. (2004). Despite the large body of research, there is little agreement regarding the relative importance of land cover at multiple spatial extents from the stream reach to the watershed. It has also been suggested that the relationships between land cover and streams may depend upon regional properties such as the predominant regional land cover type (Wang et al. 2004). Development of proper watershed management strategies depends upon mechanistic understandings of important processes, and the impact of human land use on streams is certainly of great importance to the health of stream ecosystems. Bernhardt and Palmer (2007) discuss many examples of how understanding the impacts of urban land use on streams has resulted in the development of effective restoration strategies, such as disconnecting impervious surface areas and replanting riparian zones. Thus, a lack of consensus regarding how land cover at different extents impacts streams limits our ability to effectively protect and restore streams. One reason for the lack of agreement is that the physiographic template, or physical and climatic characteristics, varies between watersheds and regions. The physiographic template also covaries with land cover and sets the stage on which stream-land cover interactions occur, which necessitates accounting for natural physiographic differences between streams in studies (Allan 2004; Wang et al. 2004). In the previous chapter, I described the development of Land cover Distinguished Physiographic Regions (LDPRs). In this chapter, I analyze the relationships between stream fish metrics and land cover types using linear mixed models to determine whether the relationships vary between LDPRs.

The physiographic template partially determines stream and river characteristics (Hynes 1975; Montogmery 1999), and as a result is related to the associated biotic communities. Steep slopes can be directly related to floods, soil and bedrock permeability determine groundwater infiltration which can control baseflow, and regional geology places constraints on channel morphology (Burt 1992; Frissell et al. 1986; Hynes 1975; Montgomery 1999). These natural differences in processes are also likely to result in different pathways by which streams interact with their watersheds (Montgomery 1999). A large number of studies have shown that fish assemblages are closely tied to gradient, watershed area, and temperature (Burton and Odum 1945; Paller 1994). In a more recent study, Wang et al. (2003) found that the key environmental features shaping fish communities in an undegraded ecoregion were stream gradient, temperature, groundwater inputs, and geology.

Many of these same physiographic characteristics also determine natural land cover and suitable locations for agriculture and urban development. Studies have found that slope and elevation are positively correlated with forest and grasslands, presumably because areas of high elevation and steep slopes are less suitable for agricultural land use (Frimpong et al. 2006; Hietel et al. 2004; Iverson 1988; Reger et al. 2007). Agricultural land use is positively associated with organic matter content, percentage clay, and moistness of soil (Burgi and Turner 2002; Hietel et al. 2004; Iverson 1988), and negatively associated with sandy soils (Burgi and Turner 2002). Although there are few studies available regarding constraints on urban land use, Iverson (1988) suggests that urbanization has fewer constraints aside from water availability.

Since physiographic characteristics are strongly related to many stream properties, stream fish assemblage characteristics, and land cover signatures, it is important for studies to account for these underlying characteristics (Allan 2004). Several investigators have heeded this warning and included a number of measurements of landscape characteristics in their analysis (Fitzpatrick et al. 2001; Frimpong et al. 2005b; Roy et al. 2007). Snyder and others (2003) measured stream gradient and were able to show that urban land use had a more disruptive effect on higher gradient streams in their study area. However, the process of selecting and measuring physiographic variables can be difficult require extensive time and resources on measurements and the same variables rarely overlap across studies, making comparisons difficult. Many studies account for similarities in watershed characteristics by stratifying the study region through the use of U.S. EPA ecoregions (Dodds and Oaks 2007; Goldstein et al. 2002; Heitke et al. 2004) or physiographic regions (Baker et al. 2006; Barker et al. 2006). Although the ecoregion and physiographic regions can be useful in some cases, neither was developed with land cover-stream studies in mind nor conveys much information regarding stream-watershed interactions. Furthermore, the delineation of ecoregions involved consideration of regional land cover properties (Bailey 2004; Omernik 1987), so that the underlying relationship between physiography and land cover is not accounted for.

In the previous chapter, I developed LDPRs as a watershed classification scheme that groups watersheds based on physiographic template characteristics that are predictors of land cover signatures at the riparian and watershed extents for the Eastern and Western United States. The purpose of this chapter is to assess the utility of LDPRs as a regional framework for stream-land cover investigations by determining whether the relationships between stream fish assemblages and land cover differ across LDPRs. Linear mixed models have been recommended as a way to accurately describe the hierarchical nature of fisheries data in statistical analyses (Wagner et al. 2006). In this chapter, I will use linear mixed models to contrast the responses of commonly used fish metrics to land cover types at the watershed and riparian (30 m x 1 Km upstream) across LDPRs. I present the results of this analysis with sites grouped by LDPRs developed to predict watershed (Classes) and riparian (RipClasses) land cover to determine whether the definition used to classify watersheds has any bearing on the results.

Methods

Fish Data

I acquired data from samples of fish communities that were collected as part of the United States Geological Survey (USGS) National Water Quality Assessment (NAWQA) program for the years 1993-2002 (Gilliom et al. 1995; USGS 2010) and the Environmental Protection Agency (EPA) Regional Environmental Monitoring and Assessment Program (REMAP) data for the MidAtlantic Highlands sampled between the years 1993-1998 (EPA 1993; EPA 2009). In order to ensure that fish assemblages in our samples were similar while still spanning a wide geographic range, I selected sites within the freshwater zoogeographic province for the Upper Mississippi River basin (region A2A; Maxwell et al. 1995). I chose this province because it encompasses the greatest number of sampling sites and spans a large area covering multiple LDPRs. Figure 2.1 shows that the NAWQA sites had a fairly even spread across zoogeographic region A2A, while the REMAP sites were predominantly located in the highlands region and belonged to Class E-5 and RipClass E-10. Taken together, the fish sampling sites are spread across a rather diverse area and have a relatively even spread across region A2A. Unlike the REMAP data, the NAWQA dataset included samples from multiple years at a given site; therefore, I selected the sample from each NAWQA site from the year nearest to 2001 so that sampling was nearest the year of land cover imagery collection.

Watersheds were delineated for REMAP sites using preprocessed Digital Elevation Models acquired with the National Hydrography Dataset Plus (NHDPlus) after snapping the sites to NHDPlus streams (EPA and USGS 2009). For the NAWQA data, I was able to acquire delineated watersheds and I used these for my analyses (Mike Meador, USGS Water Resources Division, personal communication). I visually checked to ensure that sites had correctly delineated watersheds and that each was located near a stream. Sites that did not meet these criteria were removed. The initial number of fish sites within region A2A numbered 201, with 166 sites from the NAWQA data and 35 from the REMAP data. In order to help minimize the effect of watershed area on the results, I only included sites with watersheds between 10 and 10,000 km^2, which reduced the total number of sites used in this analysis to 161 fish sites.

Land Cover and LDPR Calculations

The 2001 National Land Cover Dataset (NLCD) at 30 m resolution was used for all land cover calculations (Homer et al. 2004). The original hierarchical NLCD classes were collapsed into the following eight broad types of land cover: forest, wetland, shrubland, grassland, agriculture, urban, water, and barren. NHDPlus streams and rivers were buffered with a 30 m (each side of the stream) by 1 km buffer upstream of each site. The proportion of the total areas covered by each of the eight land cover types within the riparian buffer and watershed extents was then calculated. I also calculated natural land cover to represent an undisturbed condition in watersheds with multiple natural land cover types by summing the proportions of forest, wetland, shrubland and grassland.

Since the watersheds often span across multiple LDPRs, I developed a definition to assign sites to a given Class or RipClass while removing those that did not clearly belong to a single. A site was assigned to a watershed LDPR (Class) or riparian LDPR (RipClass) whenever both the location of the fish sampling site and the majority of the watershed belonged to the same Class or RipClass. Sites that did not meet these criteria for either Class or RipClass were not assigned to any LDPR and were excluded from the respective analysis. All GIS analyses were carried out in ArcGIS 9.2.

Fish Metrics

A relatively large number of fish metrics (29) expected to respond to land cover was selected based on previously published studies, hypothesized relationships and availability of information. FishTraits (Frimpong and Angermeier 2009) is a freely available source of information for a number of species traits and was my primary source of trait information for classifying fish into guilds, functional groups and taxonomic groups. A ranking of resilience, or capacity to withstand exploitation, was also obtained from the Fish Base website for each species (Frose and Pauly 2010). All proportional metrics were calculated as both the proportion of species richness and the proportion of total abundance in a guild or taxon. For clarity, I use a notation for all proportional metrics throughout for which each metric ends in PS (e.g. BenInv_PS) or PA (e.g. BenInv_PA) for proportion of species richness or proportion of abundance, respectively. In Table 2.1, I give a description of each metric, including the specific definition used to calculate the metrics derived from the FishTraits dataset or from FishBase information. Microsoft Excel was used for the calculation of all metrics.

Statistical Analysis

Land cover types that were underrepresented at the watershed and/or riparian extents were not used in analyses. The land cover types used were forest, natural, agriculture and urban land cover. All proportional fish metrics and land cover proportions were arcsine (square root (proportion)) transformed and then rescaled from 0 – 1 by multiplying by $2/\pi$. Although not often used in regression, the arcsine square root transformation of a proportional response variable has been shown to help achieve linearize relationships (Ray et al. 1996). Since the transformation moves values that are near the extremes of 0 and 1 to more central locations, it also helped visualization of the relationships. Natural and forest land cover were then linearly inverted by subtracting from 1 so that all analyses were run with the intercept representing a least disturbed condition. I also removed sites from the analysis belonging to greatly underrepresented LDPR Classes and RipClasses to help reduce the inequality in sample sizes. Prior to fitting final statistical models, I used a correlation matrix, visually analyzed scatterplots and fit preliminary models to select several fish metrics that potentially exhibited varying relationships with different land cover types between LDPR Classes. Due to the large amount of time required, I did not do preliminary analyses contrasting responses of fish metrics to land cover types with sites grouped by RipClasses but assumed that the selected metrics would show trends similar to the Class analyses. I selected the following five metrics showed the potential for having significant random intercept and slope terms with sites grouped by Class for linear mixed models analysis: BenInv_PS, Centrar_PS, DrtMdSc_PS, Omni_PS, and SubGen_PS (Table 2.1). Preliminary analysis for all other metrics (see Table 2.1) either did not suggest strong relationships with land cover or did not have strong contrasting relationships with land cover across LDPR Classes, and I did not use them in further analysis.

For the selected metrics, I fit hierarchical models using proc Mixed in Statistical Analysis System (SAS) 9.2 (SAS Institute Inc. 2007) to determine whether relationships with land cover varied between LDPRs as determined by the presence of random intercepts and/or random slopes. Wagner et al. (2006) provides a rather thorough discussion of the benefits and appropriate uses of mixed models in fisheries research. I also found the overview of implementing proc Mixed in SAS by Der and Everitt (2002) to be very helpful and used this resource throughout analysis. For each model, I only included one land cover type as a predictor in order to simplify interpretation. I also included log transformed watershed area as a predictor, and the multiplicative interaction between watershed area and the land cover type. A linear regression model of this form with only fixed effects and random errors, ε_i, is expressed by the formula

$$\text{Equation 1:} \quad y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + \varepsilon_i \,,$$

where y_i represents a proportional response of a given fish metric observed at site i, β_0 represents the intercept, β_1 represents the slope coefficient of the relationship between the response and a land cover predictor, x_1 represents the proportion of a given land cover, β_2 represents the slope coefficient of the relationship between the response and watershed area, x_2 represents watershed area, and β_3 represents the coefficient for the multiplicative interaction between x_1 and x_2. An extension of the above model that contains terms for random slopes and intercepts for each group of sites can be expressed by the formula

$$\text{Equation 2:} \quad y_{ij} = \beta_0 + \beta_1 x_{1i} + \alpha_j + \gamma_j x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} + \varepsilon_{ij} \,,$$

where α_j represents a random intercept for each group j, and γ_j represents the random slope coefficient of the relationship between the response and a given land cover for group j. The models are hierarchical, with the multiple regression model with only fixed effects in Equation 1 nested within the linear mixed model of equation 2.

Each of these models represents a potential relationship between land cover and stream fish assemblage response with sites grouped by Class or RipClass. A model with only fixed effects will be interpreted as meaning that there is not sufficient evidence to claim different intercept values of a predictor or different predictor – fish metric relationships among Classes or

RipClasses groups. A model with random intercepts and fixed slope means that there is a common relationship between a predictor and fish metric but differences in the value of the metric whenever the predictor is at intercept. For models with forest and natural land cover as predictors, random intercepts represent natural differences among Classes or RipClasses, as human disturbance is minimal. For models with agriculture and urban land use, random intercepts do not represent natural differences, as a watershed with no agriculture could be entirely urban. A model with a fixed intercept and random slopes points to a common intercept but varying land cover - fish metric responses across Classes or RipClasses. Finally, a model with random intercepts and slopes would point to intercept differences and also different effects among Classes or RipClasses.

I compared models having all of the fixed and random terms in Equation 2 above to determine which offered the best fit for the relationship between each fish metric and land cover variable. A model was selected as the best representative for each relationship if the appropriate tests for the fixed and random effects were significant at $\alpha = 0.05$ and Akaike's Information Criterion (AIC) was minimized. An unstructured covariance matrix was assumed for models having more than one random effect, since the random components may be correlated with one another and likely do not share a common variance. I recorded information for the best model for each relationship, including AIC and p-values for each effect, in Microsoft Excel for tabular presentation. In order to compare the presence of random intercept and slopes between the two LDPR classification schemes, each relationship between a fish metric and land cover type was run with sites grouped by LDPR Classes and RipClasses. The recorded AIC values can be used to compare the best predictive models with different land cover types at the two extents when sites are stratified by the same grouping factor, either Class or RipClass. However, the AIC values cannot be used to compare the fit of Class and RipClass models because a different number of sites were used for the two analyses.

Presentation of Results

I present information from the best models for each relationship in tabular form through the use of Microsoft Excel. Boxplots were created in the R statistical environment to contrast the characteristics of the sites grouped within LDPR Classes and RipClasses. All regression models except those with random intercepts and fixed slopes were plotted using proc Gplot in SAS. Models containing random intercepts with a fixed slope were graphed by representing these models in an analysis of covariance (ANCOVA) design using proc GLM in SAS. These graphs were used because ANCOVA can fairly accurately represent these models and because of the difficulty of obtaining the appropriate plots via other procedures in SAS. Relationships containing significant watershed area and watershed area interaction terms were not graphed due to the difficulty of interpreting three-dimensional graphs, especially whenever random effects are significant.

Results

LDPR Class Characteristics

The number of fish sampling sites that were included in mixed models analysis was 149 sites in six LDPR Classes (E-1, E-3, E-4, E-5, E-6, and E-8) and 135 in three RipClasses (ERip-5, ERip-7 and ERip-10). The number of sites in each Class was 13 in E-1, 28 in E-3, 11 in E-4, 44 in E-5, 12 in E-6, and 41 in E-8; and in each RipClass was 11 in ERip-5, 70 in ERip-7, and 54 in ERip-10. The sites not included in the analyses were either members of underrepresented Classes or RipClasses, or did not meet the site location and watershed majority agreement criteria for assignment to a LDPR.

Box plots for untransformed watershed area and the percent watershed and riparian cover of the common land cover types that were used in mixed models (forest, natural, agriculture, and urban) are shown in order to contrast characteristics of the sites grouped within Classes (Figures 2.2 and 2.3) and RipClasses (Figures 2.4 and 2.5). The watershed area box plots (Figures 2.2a and 2.4a) also display the number of sites within each Class at the top of the plots. Shrubland, grassland, wetland, and water are not shown in the box plots because these land cover types were underrepresented for most sites and were not used in statistical models as predictors of fish assemblage metrics. Watershed area may differ slightly between LDPRs, but these differences do not appear to be great and should not confound my results, especially since area is analyzed as a potential effect (Figures 2.2a and 2.4a). The box plots reveal that the sites have different watershed land cover patterns when grouped by Class or RipClass; even so, the ranges generally overlap considerably (Figures 2.2b-e and 2.4b-e). For example, watershed agricultural land use in every Class has a range from 0 percent to nearly 60 percent, although medians differ greatly (Figures 2.2d). However, for watershed percent forest and natural land covers (Figures 2.2b-c and 2.4b-c), there are differences and only marginal overlap among some groups. The sites in E-5 and ERip-10, for instance, have considerably more forested land in their watersheds and overlap with other Classes and RipClasses is rather low (Figures 2.2b and 2.4b). Riparian land cover signatures are more similar across Classes and RipClasses, with sites having widely overlapping ranges of forest, natural, agricultural and urban (Figures 2.3a-d and 2.5a-d).

The number of individual fish in the samples used in this analysis was 98,959, with the number of species collected ranging from 2 to 40 species per site. Analysis of scatter plots revealed that a number of metrics representing fish assemblage characteristics varied across classes and had relationships with land cover types. I selected the following five metrics that appeared to show the most contrasting trends across Classes for linear mixed models analysis: BenInv_PS, Centrar_PS, DrtMdSc_PS, Omni_PS, and SubGen_PS. The proportion of darter, madtom and sculpin species was moderately correlated with the proportion of benthic invertivore species since most darter, madtom and sculpin species are benthic invertivores. However, there are a number of fish species of the Cyprinidae family that are also benthic invertivores so I decided to include both of these metrics in my analysis. Models for Class and RipClass were fit between each of these metrics and 1- watershed forest (Wshd_For), 1 - watershed natural (Wshd_Nat), watershed agriculture (Wshd_Ag), watershed urban (Wshd_Urb), 1 - riparian forest (Buf_For), 1 - riparian natural (Buf_Nat), riparian agriculture (Buf_Ag), and riparian urban (Buf_Urb) land covers.

Benthic Invertivores

Summary information for the best relationships between the eight land cover predictors and the proportion of benthic invertivore species (BenInv_PS) with sites grouped within both Class and RipClass are shown in Table 2.2. Each row shows the best representative model for the relationship between the land cover variable in that row and fish metric being analyzed. The best overall model for this metric that had Class as the grouping factor contained a random intercept and fixed Wshd_Urb slope (AIC= -262.3, Figure 2.6); an equivalent model with Wshd_Ag had the second best fit (AIC = -250.8, Figure 2.7). Figure 2.6 shows that Wshd_Urb negatively impacts percent benthic invertivore species in a similar way for all classes, and that LDPR Classes E-1, E-3, E-4 and E-5 had generally higher intercepts than Classes E-6 and E-8. For this and other random intercept models that have fixed slopes, differences in intercept among Classes represent different estimates of the level of BenInv_PS whenever the given predictor is minimized; keep in mind that riparian and natural land cover predictors are represented from highest proportion to lowest. Figure 2.7 shows that Wshd_Ag had a positive relationship with BenInv_PS and that Class E-5 had the highest intercept, followed by E-4, E-3, E-1, E-6 and E-8. The best models for wshd_for, wshd_nat, and all four riparian land cover types did not have better fits than a model with only random intercepts (AIC = -246.1).

The models predicting BenInv_PS with sites grouped by RipClass did not include any random terms. The single best predictor for the RipClass models was also Wshd_Urb, with the best model for this relationship containing a fixed slope and a watershed area term, both having negative relationships (AIC = -222.7). The second best model had negative Wshd_Ag * area interaction term with both individual slopes positive (AIC = -212.7). Fixed slopes for Wshd_For, Wshd_For, and Buf_For were significant and provided better fits (AIC = -211.6, AIC = -210.1, and -210.9, respectively) than the null, intercept only model. Figures 2.8 – 2.10 show that the relationships between Wshd_For, Wshd_Nat and Buf_For are almost identical, with BenInv_PS decreasing as natural and forest land cover decreases (1 – Wshd_For, 1 - Wshd_Nat or 1 - Buf_For increases) from left to right. The intercept-only RipClass model provided a better fit than RipClass models that included Buf_Nat, Buf_Ag and Buf_Urb.

Centrarchids

The best models for each of the relationships between the eight types of land cover and the proportion of centrarchid species (Centrar_PS) with both Class and RipClass as site grouping factors are summarized in Table 2.3. The best overall Class model for predicting Centrar_PS had random intercepts and random Wshd_Ag slopes (AIC=-256.8), and the plot shows that both intercepts and slopes differ but that all Classes converge near the same point when watershed agricultural land use is maximized (Figure 2.11). Listed in order of highest to lowest intercept, the Classes with negative slopes include E-8, E-3, E-4 and E-1, while Class E-6 had almost no slope and E-5 had a fairly strong positive slope (Figure 2.11). A model with only a fixed Wshd_Urb slope was a close second (AIC=-256.3). The plot of the fixed Wshd_Urb slope model in Figure 2.12 shows that Centrar_PS increased sharply with Wshd_Urb, and suggests that the relationship may be better represented by a logarithmic curve with a threshold around 0.2. Models for Wshd_For (AIC = -253.3) and Wshd_Nat (AIC = -251.9) had significant, negative watershed area – fixed slope interactions with the individual slopes positive.

The Class models with riparian land cover predictors generally had poorer fits than those with watershed land cover. Class models with random intercepts and random slopes for Buf_For (AIC=-243.6) and Buf_Nat (AIC=-242.6) are shown in Figures 2.13 and 2.14, respectively. The plot in Figure 2.13 shows that Classes E-1, E-3, E-4, E-6, and E-8 have very similar intercepts and that Class E-8 had the strongest negative trend as riparian forest decreases left to right; Class E-5 had a much lower intercept and a strong positive increase in Centrar_PS as Buf_For increased. The Wshd_Nat plot in Figure 14 shows that Classes E-4, E-3, and E-1 have nearly identical intercepts and slightly negative slopes. Classes E-5 and E-6 have the lowest intercepts and strong positive slopes, while E-8 has an intermediate intercept and a negative slope (Figure 2.14). A random intercepts with a fixed Buf_Urb slope model (AIC = -243.6) is plotted in Figure 2.15. The plot reveals an overall increasing trend with Class E-5 having the lowest intercept and Class E-1 having the highest. Despite agriculture being a significant predictor at the watershed scale for Class models, riparian agriculture (Buf_Ag) did not offer improved predictions beyond the random intercept only model (AIC = -246.1).

Again, the RipClass models did not include any random terms, and the best had a fixed Wshd_Urb slope without random terms, similar to the corresponding Class model (AIC = -211.4, Figure 2.12). The best models for Wshd_For and Wshd_Nat had significant, negative watershed area – fixed slope interactions with both of the individual slopes positive (AIC = -210.5 and -209.4, respectively); the corresponding model for Wshd_Ag provided only a marginally improved fit (AIC =198.0) over the intercept only model (AIC = -197.2). The null, intercept-only model (AIC = -197.2) was not improved upon by RipClass models containing any of the riparian land cover variables.

Darters, Madtoms, and Sculpins

The best models for the relationships between each land cover type and the proportion of Darters, Madtom and Sculpin species (DrtMdSc_PS) are shown in Table 2.4. The best overall Class model predicting DrtMdSc_PS had random intercepts with a positive watershed area - fixed Wshd_Urb slope interaction with both individual slopes negative (AIC= -295.8). The second best model included random intercepts, random slopes, and a fixed positive Wshd_Ag slope (AIC = -275.2). A plot showing random intercepts and random slopes for this model reveals that Class E-5 had a high intercept and negative slope, Class E-6 had an intermediate intercept and a very slight negative slope, and all other Classes had lower intercepts and positive slopes (Figure 2.16). A model with fixed, negative Wshd_For slope (AIC = -260.6, Figure 2.17) was the only other model to offer better predictions of DrtMdSc_PS than one with a random intercepts only (AIC = -259.8).

The RipClass model with a fixed Wshd_Urb slope and watershed area term, both negative, had the best fit (AIC = -265.9) of all RipClass models. The two models for Wshd_For and Wshd_Nat both had fixed slopes and similar fits (AIC = -240.1 and -238.7), and the plots show very similar negative relationships (Figures 2.17 and 2.18, respectively). The RipClass model with a watershed land cover predictor had random intercepts and a negative watershed area – fixed Wshd_Ag slope interaction with both individual slopes positive (AIC = -239.5).

Riparian land cover again provided poorer predictions than watershed land cover, with Buf_Nat and Buf_Ag not offering improvements beyond a model with only random intercepts (AIC = -231.7). A model with a fixed, slightly negative Buf_For slope (AIC = -232.4, Figure 19) and a model with random intercepts and a watershed area - fixed Buf_Urb slope interaction term (AIC = -232.9) were only marginally better than the random intercept only model. Analysis of the random intercept only RipClass model showed that Class ERip-10 had the highest intercept values of DrtMdSc_PS, followed by ERip-7 and ERip-5.

Omnivores

The best models for the proportion of omnivore species (Omni_PS) all contained either watershed area as a predictor or a watershed area – fixed land cover predictor interaction (Table 2.5). All Class models with watershed land cover predictors had very similar fits, although the models had different terms. The best had random intercepts and a negative watershed area - fixed Wshd_Urb slope interaction with both individual slopes positive (AIC = -341.7). A close second was comprised of a fixed, positive Wshd_For slope and a negative area term without random effects (AIC = -341.6). Class models with Wshd_Nat and Wshd_Ag had random intercepts, random slopes and negative watershed area terms (AIC = -341.5 and -341.2, respectively). No riparian land cover types improved the fit of a model with random intercepts and watershed area as a predictor (AIC = -338.5).

The RipClass models did not have any significant random effects, and the best had a negative watershed area – fixed Wshd_Urb slope interaction with both individual slopes positive (AIC = -284.2). Three equivalent RipClass models that included a positive fixed slope and negative area term were fit for Wshd_For (AIC = -284.1), Wshd_Nat (AIC = -283.0), and Buf_For (AIC = -276.3). The best RipClass models including Buf_Nat (AIC = -274.3) and

Buf_Urb (AIC = -274.9) as predictors had significant, negative watershed area – fixed slope interactions with both individual slopes positive. Agricultural land use within both the watershed and riparian extents did not have better fits than a model that included a negative relationship with watershed (AIC = -272.1).

Substrate Generalists

The models providing the best fit for the proportion of substrate generalist species (SubGen_PS) for each land cover type are summarized in Table 2.6, and reveals that all contain at least one random intercept or slope effect. The best overall model that had Class as the grouping factor was comprised of random intercepts, a fixed, positive Wshd_Urb slope and a positive watershed area term (AIC = -220.3). A Class model with a negative watershed area – fixed Wshd_For slope interaction with the individual slopes positive had the second best fit (AIC = -211.2). The best Class model with Wshd_Nat as a predictor had a positive fixed slope and positive watershed area term (AIC = -206.3). A Class model with a fixed, negative Wshd_Ag slope and positive watershed area term provided the poorest fit of all watershed land cover predictors (AIC = -205.2). The best Class models for Buf_Nat (AIC = -204.9) and Buf_For (AIC = -203.7) had random intercepts and a negative watershed area – fixed slope interactions with all individual slopes positive. The last Class model with a significant land cover predictor had random intercepts with a fixed, positive Buf_Urb slope and a positive watershed area term (AIC = -205.7). Riparian agriculture (Buf_Ag) was not a significant predictor in any model offering better fit than a model with random intercepts and a watershed area term (AIC = -203.0).

The four RipClass models that had significant land cover terms had the same terms in each model as the respective Class models. The best RipClass model had random intercepts, a fixed positive Wshd_Urb slope and a positive watershed area term (AIC = -185.1), similar to the best Class model. A model with a negative watershed area – fixed Wshd_For interaction with both individual terms positive provided the second best fit (AIC = -181.9), followed by a model with a fixed positive Wshd_Nat slope and a positive watershed area term (AIC = -178.6). The Buf_Nat model with random intercepts and a negative watershed area – fixed slope interaction with both individual terms positive was the only other RipClass model to have a better fit (AIC = -165.6) than the model with random intercepts and a watershed area term (AIC = -165.6).

Discussion

The results of this study reiterate the impacts of agricultural and urban land use on stream fish assemblages, which have been reviewed thoughtfully by Allan (2004). Urban land use was by far the most important predictor of the five metrics analyzed in this study, and other studies comparing the impacts of urban land use to agricultural land use have reached similar conclusions (Snyder et al. 2003; Wang et al. 2001). The impacts of urbanization were predicted to be similar across all Classes (i.e. no random slopes) regardless of differences in the physiographic template, although sites did differ in the baseline proportions of metrics (i.e. random intercepts).

The models also suggest that watershed land cover has a stronger impact on fish assemblages than land cover within 30 m x 1 km reach riparian zone. There is little agreement in the literature regarding the relative importance of land cover at the reach, riparian and watershed extents. A number of studies have pointed to watershed land cover as the most important predictor of fish assemblage health (Roy et al. 2007; Snyder et al. 2003) while other studies have pointed out the importance of reach land cover (Frimpong et al. 2005b; Van Sickle and Johnson 2008). Allan (2004) discussed a number of reasons why studies might reach different conclusions, including study design and differences in the dominant regional land cover type.

This study was designed to cover a large extent with many differences in the physiographic template and there was more regional variation in watershed land cover, as intact riparian zones are somewhat common even in regions with agriculturally dominated watersheds. Thus, while this study supports the hypothesis that watershed land use is a more important determinant of fish assemblage health, it is possible that another study specifically designed to compare the impacts of land cover across multiple extents might reach a different conclusion. Watershed area was also identified as an important predictor in this study, especially for percent omnivore species and percent substrate generalist species. These results are not surprising, as there was a large range of watershed area in this study and the relationships between watershed area and stream fish assemblages are well known (Burton and Odum 1945; Paller 1994; Schlosser 1982).

Many of the models with sites grouped by Land cover Distinguished Physiographic Regions (LDPR) Classes had significant random intercepts and slopes, which was true for only a few with sites grouped by RipClass. The fact that random terms are present in Class models but not RipClass models suggests that the definitions used to define regional boundaries are important, and that the regions defined by physiographic predictors of land cover at the watershed extent offer more explanatory power in this analysis. If models for both groupings had significant random terms, one could conclude that the results were due to the spatial proximity of sites within the groups and not the definitions. Other researchers have used random groupings (Brenden et al. 2008), arbitrary regions (Wolock et al. 2004) or site proximity models (Pyne et al. 2007) to assess the strength of classification provided by regional framework. Such an assessment is unnecessary given the differences between results for the two LDPR watershed stratification schemes.

The models that had sites grouped by LDPR Class predicted differences in natural fish assemblage characteristics, as predicted by different intercepts. The sites assigned to Class E-5 were primarily located in more forested, mountainous areas and typically have higher gradient, cooler streams. Not surprisingly, the samples from these sites were characterized by higher proportions of specialized benthic invertivores and darters, madtom and sculpin, and lower proportions of centrarchids whenever watershed human disturbance was minimal. Class E-8 sites were located on a very different physiographic templates and tended to have agriculturally dominated watersheds. In contrast to Class E-5, the fish assemblages of the sites in LDPR E-8 were predicted to have lower proportions of specialized benthic invertivores and darters, madtom and sculpin, and higher proportions of centrarchids at low levels of human disturbance. Scott and Helfman (2001) discuss similar trends for streams of the southeastern U.S., with highland streams having diverse assemblages with a number of darter species and lowland streams having more cosmopolitan species, such as many centrarchids. Class E-6 watersheds are characterized by sandy soils and thus have land cover signatures similar to E-5 with less agriculture and more natural land cover. However, fish assemblages of E-6 were generally predicted by the models to be more similar to E-8 when undisturbed by human land use. Despite differences in slope and soil properties, Classes E-1, E-3 and E-4 were predicted to have naturally similar fish assemblages as represented by the proportions of centrarchid, specialized benthic invertivore, and darters, madtom and sculpin species.

This analysis was not designed to highlight natural differences in fish assemblages, but it is clear from the results that there is significant covariance between human land use and fish assemblage characteristics due to physiography. Temperature, geology, stream gradient, and groundwater inputs have been identified as three important characteristics determining differences in fish assemblages (Brenden et al. 2008; Burton and Odum 1945; Paller 1994; Wang et al. 2003). The most important physiographic template characteristics used to distinguish LDPRs include temperature, percent sand in soil, slope, relief, percent total flatland and precipitation. Since the factors identified as important determinants of fish assemblage structure are closely related to those distinguishing LDPRs, it should not be surprising that natural fish assemblage differences exist between Classes. Because watershed area was an important predictor for the proportion of omnivores and substrate generalists and this study includes a wide range of watershed sizes, it is not possible to speculate on natural differences among Classes for these metrics. A future study contrasting these metrics within similar sized watersheds across LDPRs could be used to determine how these metrics differ among regions.

Not only were differences evident in the natural proportions of the selected fish metrics among classes, but also in the way some of the metrics responded to land cover alterations as shown by the models with random slopes. The three metrics that included at least one model with random slopes were percent centrarchid species, percent darter, madtom and sculpin species, and percent omnivore species. As mentioned previously, it is not possible to assess how the relationship between proportion omnivore and land cover differs among Classes due to the importance of watershed area as a predictor for this metric. The relationships between the proportion of centrarchid species and watershed agriculture and riparian forest suggested that the different slopes among Classes resulted in the proportions becoming more similar as human alteration increased, regardless of the intercepts. Class E-5 in particular tended to have the lowest natural proportion of Centrar_PS and human alteration in the riparian zone or watershed agriculture resulted in increased proportions, while E-8 generally had the strongest negative trends. These two LDPRs differ from one another primarily based on slope and precipitation, and as a result the streams and rivers have different physical properties and biotic communities. The higher gradient streams of Class E-5 have cooler temperatures and less pool habitat, and thus do not support high proportions of centrarchid species. Agricultural land use can result in more lacustrine habitat, and this alteration represents one pathway by which Class E-5 streams may be altered to the benefit of deep bodied centrarchids (Roy et al. 2007). The strong negative relationship in Class E-8 may represent the impacts of pollution from very high levels of agriculture present in this region, which would remove the less tolerant members of this family (Karr 1981; Schleiger 2000). Others have found that homogenization of fish communities is related to human land use through sedimentation, which would benefit many centrarchids (Scott and Helfman 2001; Scott 2006; Walters 2002).

The proportion of darter, madtom and sculpin species generally responded to watershed agricultural land use in a somewhat similar manner for all LDPRs with the exceptions of Class E-5, which had the highest predicted natural proportions and the only negative relationship and Class E-6, which had little or no slope. All of the other Classes were predicted to have much lower proportions and had similar, positive slopes. Due to the mountainous terrain, it is likely that the streams and rivers in Class E-5 support a diverse assemblage of daters, madtom and sculpin dependent upon cobble/gravel substrate as mentioned previously. Agricultural land use in the watershed would likely result in the removal species that are less tolerant to warming temperatures, siltation, and stressor inputs (Karr 1981; Schleiger 2000; Scott and Helfman 2001; Waite and Carpenter 2000). The positive relationships of this metric with agricultural land use in other LDPR Classes may largely represent a gradient from urbanization to agriculture, as land cover types are calculated as proportions of a given area and are not independent of each other (King et al. 2005). One piece of evidence supporting this hypothesis is that Class E-4 typically has the highest levels of watershed urbanization (see boxplot in Figure 2.2e) and also the strongest positive relationship with watershed agriculture. Class E-6, on the other hand, had much more forested and natural land cover and had very little relationship with agricultural land use. It is also possible that the increase represents introductions of nonnative, relatively tolerant species that are not present in the more pristine streams of these regions.

These potential scenarios are only examples of the way that physiographic differences can be used to explain differences in land cover-stream fish assemblage relationships, and there are many other plausible explanations. Regardless of the mechanism, these results suggest that responses of at least some fish assemblage to land cover alterations differ between regions and that these differences can be explained by physiography. Since investigations have typically been carried out in regions believed to be homogenous, there is not a large body of evidence to support these results. Utz (2009) found that responses of stream macroinvertebrate assemblages to urban land use differed between Piedmont and Coastal Plain regions in Maryland. Other researchers have identified important physical attributes that affect the response of fish assemblages, including gradient for urban land use (Snyder et al. 2003) and length slope factor for agricultural (Frimpong et al. 2005a). These results also show that fish communities become increasingly homogeneous as human land use increases. The replacement of endemic fishes native to highland streams by more cosmopolitan fishes from lowland fishes as a result of habitat alteration by humans at a smaller scale in the southeastern U.S. has been described (Scott and Helfman 2001; Scott 2006). In this study, the highland streams of Class E-5 show similar trends with the replacement of the more endemic darter, madtom and sculpin taxa by more cosmopolitan centrarchids as human land use increases. The models in this study also suggest that the proportions of darter, madtom and sculpin species increase while the proportions of centrarchids decrease in the lowland Classes, especially E-8, as human land use increases. These results suggest that homogenization of fish communities is occurring across large regions that differ physiographically and have naturally different stream fish assemblages.

This investigation has provided further support for the need to consider physiographic template differences when investigating the impacts of land use on streams, and shows that LDPRs can be used to successfully stratify a study region based on physiography. In attempting to understand the relationship between two cause and effect factors (e.g. land cover and stream fish assemblages), it is important to control for underlying factors (e.g. physiography) that are acting at different scales to determine many of the properties of the interactions being studied (Schumm and Lichty 1965). The characteristics of stream biotic communities largely depend upon stream habitat characteristics, which in turn depend upon regional physiographic features (Frissell et al. 1986). Natural land cover and human land use are strongly related to several of these same physiographic properties, which necessitates the need to account for physiographic differences in land cover-stream investigations (Allan 2004).

A large number of land use investigations have carried out regions within or across ecoregions (Bailey 2004; Omernik 1987) and physiographic regions such as those delineated by Fenneman (1928). However, there are some drawbacks to both of these approaches. Although ecoregions have a number of relevant uses, land cover patterns were one variable used to determine regional boundaries (Bailey 2004). More specifically, this means that a boundary between two ecoregions could be based off of different land cover properties even when the physiographic template is similar. It is also possible that regions with different physiographic templates could be joined in one ecoregion because they have similar land cover properties. If ecoregions are used as a stratification tool, it can be difficult to determine which differences in the stream are due to the impacts of land use and which are due to underlying physiographic differences. The delineation of LDPRs differs from this approach by identifying properties of the physiographic template that determine land cover and then delineating regions based on the identified predictors. This approach properly treats land cover as dependent upon physiography and does not define regional boundaries based off of land cover. While it is true that each LDPR is distinguished by land cover patterns, each land cover type varies widely around the regional average and upon a similar physiographic template. Thus, LDPRs can be used to identify streams with similar physiography that range from unimpacted to greatly impacted.

Physiographic regions avoid the pitfall of including land cover characteristics in classification and thus provide a better stratification option than ecoregions. However, physiographic regions do not provide detailed information regarding the differences between regions that could be useful for explaining why land cover-stream interactions might differ. Furthermore, the regional boundaries are determined by geomorphology (Fenneman 1928), which is only one of the determining factors of stream properties. Thus, physiographic regions leave researchers with the need to measure and analyze the impacts of a number of underlying variables, which consumes resources and can leave one doubting whether the selected variables are ecologically relevant. The use of LDPRs to stratify regions improves upon this approach by including a range of physical, climatic, and geologic characteristics, while also giving objective explanations for how regions differ.

Land cover Distinguished Physiographic Regions (LDPRs) were developed to classify watersheds in a way that accounts for differences in the physiographic template that are important determinants of land cover properties. This study has shown that there are both differences among LDPRs in natural fish assemblage characteristics and the way these assemblages respond to land cover alterations. There are two main advantages to be gained by utilizing LDPRs instead of existing regional frameworks to classify watersheds in land cover-stream investigations. First, the LDPR dataset not only classifies watersheds but also includes the original physiographic measurements. Thus, LDPRs can be used to not only identify regional differences but to explain these differences based on the physiographic attributes either subjectively or through statistical procedures. Second, only physiographic characteristics that were shown to predict land cover patterns were used in the development of LDPRs. This is important because natural land cover and human land use is partially determined by a number of physiographic attributes that can also determine stream structure and function. One shortcoming of LDPRs as a stratification tool is that they were delineated through the use of rather coarse, readily available one kilometer data that limits their utility at smaller extents. However, the delineation process was rather easy to perform and with the advent of higher resolution data, LDPRs or a similar classification scheme can be improved upon in the future.

For these reasons, I recommend the utilization of LDPRs as a watershed stratification tool that can be used in a number of ways. First, control streams that are unimpacted by human land use can be identified because there are watersheds within every LDPR that have minimal human land use, even in agriculturally dominated areas. Second, environmental monitoring programs aimed at characterize the streams of a region and identify major human impacts can stratify watersheds though the use of LDPRs. I noticed that the watershed and reach land cover for the NAWQA and REMAP sampling sites used in this analysis had higher proportions of human land use than the average for the LDPRs that they belonged to. For example, the watersheds for the sampling sites in Class 4 had an average of nearly 20% urban land cover (Figure 2.2e), which is much higher than the average of 8.3% for this LDPR (see Table 1.1). This is no small difference, as a number of studies have suggested that there is a threshold for urban land use around 10-15% above which streams are almost without exception impaired (e.g. Roy et al. 2007, Wang et al. 2001). Similar discrepancies between regional averages in land cover patterns and the averages for the sampling sites used in this analysis are present for other LDPRs as well. Since these programs are used for regional environmental and water quality monitoring and to develop regional standards, the streams sampled need to be representative of the respective region. It is likely that the streams have higher proportions of human land use due to the difficulty of accessing and sampling streams in undeveloped areas. It is also possible that the high levels of human land use reflect the use of watershed stratification schemes that do not properly tease apart the relationship between land cover and physiography. The use of LDPRs would enable streams that are truly representative of a region with a common physiographic template to be selected, which would strengthen inferences about human impacts.

Another use of LDPRs for investigating relationships between land cover and streams because a number of streams within a single LDPR that have a range of a given land cover types can be sampled. Since the physiographic template within a single LDPR is relatively similar, a study designed in this way would not have to measure individual physiographic variables for each stream. However, one way to further assess the strength of LDPRs as a stratification tool would be to compare the amount of variation explained by individual physiographic variables with the amount explained by the LDPR classification. The final use of LDPRs that I will discuss is for better understanding the role of physiography in the interactions between land cover and streams, especially as relates to the importance of land cover at different extents. This can be studied by selecting a similar number of streams with a similar range of land cover at multiple extents from more than one LDPR, then contrasting the relationships across LDPRs.

Such an approach could identify whether there are differences in the relative importance of land cover at different spatial extents among LDPRs. Since physiography is important for determining the links between streams and their valleys, it seems possible that the importance of riparian land cover could be the best predictor of a given stream response variable in one LDPR, while watershed land cover could be the best predictor in another. Physiographic properties that differ between the two LDPRs could then be used to help explain why these relationships differ, which would greatly enhance our understanding in this area. I recommend the use of LDPRs as a watershed stratification tool for these purposes and others that need to account for physiographic differences throughout a region.

Conclusions

This study was designed to assess the utility of a watershed classification system, Land cover Distinguished Physiographic Regions (LDPRs), for identifying regional differences in fish assemblages and their responses to land cover alterations. Mixed linear models analysis revealed that the responses of benthic invertivores, omnivores, substrate generalists, centrarchids, and Darter, madtom and sculpin species had varying responses across LDPR Classes, but not RipClasses. The responses of several metrics to the conversion of forested land to agriculture and urban land use suggest that homogenization of fish assemblages, which supports the findings of Scott and Helfman (2001). These metrics describe fish assemblages based on reproductive, trophic, and taxonomic guilds and have been commonly used in other investigations. Thus, it seems likely that the significant differences seen in these metrics reflect actual differences in the ways that fish assemblages and potentially other taxa respond to land cover alterations across regions. Presently there is a tendency to expect that there is one watershed management strategy that will help mitigate the impacts of human land use on streams for all regions, since land use impacts on streams are assumed equal. These results do not support this view and rather suggest that physiography must be considered in the development of watershed management strategies. I recommend that further studies investigating stream-land cover interactions utilize LDPRs as a stratification tool to control for regional variation and to further assess the validity of this new regional framework. I also recommend that a similar classification system, or an update of LDPRs, be developed as higher resolution data become available in the future for the entire United States or smaller regions.

Table 2.1 Description of fish metrics.
The table shows the full list of metrics used in preliminary analysis, along with a full description and the formula used to define the metric. The external source where the trait or taxonomic information was acquired is shown as either from FishTraits (Frimpong and Angermeier 2009) and Fish Base (Frose and Pauly 2010).

Metric Abbreviation	Metric Description	Source	Formula
N_Species	Species richness	n/a	n/a
Abundance	Total number of individuals	n/a	n/a
Simpson	Simpson's Diversity Index	n/a	1 - ∑(ni/Nt)2, where ni is the abundance of species i and N is the total abundance.
Resil_PS and Resil_PA	Resilience	Fish Base	Defined as a species described by Fish Base as highly resilient
BenInv_PS and BenInv_PA	Specialized Benthic Invertivore	FishTraits	Defined as a species that is a specialized benthic feeder and a specialized invertivore, one consuming small invertebrates and larval fish, but not macrophytes, detritus, algae, phytoplankton, fish, crabs, crayfish, or blood. By this definition, a species may consume eggs or other unidentifiable matter. FishTraits formula: BENTHIC = 1, SURWCOL=0, "INVLVFSH"=1, all other food items equal 0 except "EGGS" and "OTHER".
Omni_PS and Omni_PA	Omnivore	FishTraits	Defined as a species for which phytoplankton, algae, macrophytes, and detritus were described as consisting of between 25% and 75% of the diet FishTraits formula: 0.25 < [("ALGPHYTO" + "MACVASCU" + "DETRITUS")/ ("ALGPHYTO" + "MACVASCU" + "DETRITUS" + "INVLVFISH" + "FSHCRCRB" + "BLOOD" + "EGGS" + "OTHER")] >0.75
Gen_PS and Gen_PA	Trophic Generalist	FishTraits	Defined as a species whose diet commonly consisted of four or more food items. FishTraits formula: "ALGPHYTO" + "MACVASCU" + "DETRITUS" + "INVLVFISH" + "FSHCRCRB" + "BLOOD" + "EGGS" + "OTHER" ≥ 4
OmniGen_PS and OmniGen_PA	Omnivore and Trophic Generalist	FishTraits	Defined as a species that is either an Omnivore or Trophic Generalist.
HerbDet_PS and HerbDet_PA	Herbivore and Detritivore	FishTraits	Defined as a species whose diet commonly consisted of one or more of phytoplankton, algae, macrophytes, and detritus, but did not include any other food items. FishTraits formula: "ALGPHYTO" + "MACVASCU" + "DETRITUS") > 1, "INVLVFISH" + "FSHCRCRB" + "BLOOD" + "EGGS" + "OTHER" < 1
SubIn_PS and SubIn_PA	Substrate Indifferent	FishTraits	Defined as a species that is described as either pelogophilic or a bearer. FishTraits formula: A_1_1 or C_1_3_4_C2_4 = 1
SubGen_PS and SubGen_PA	Substrate Generalist	FishTraits	Defined as a species described as a phyto-lithophil, polyphil, ariadnophil, or speleophil not specific to rock cavities. FishTraits formula: A_1_4, A_2_4C, B_2_2, B_2_4, B_2_7B, or B_2_7C = 1
SubInGe_PS and SubInGe_PA	Substrate Indifferent and/or Generalist	FishTraits	Defined as either a substrate indifferent or a substrate generalist species.
Litho_PS and Litho_PA	Simple Lithophil	FishTraits	Defined as a species described as a non-guarding lithophil associated with rock, gravel or sand substrate. FishTraits: A_1_3A or A_1_3B = 1
Speleo_PS and Speleo_PA	Specialized Speleophils	FishTraits	Defined as a species described as either a non-guarding or guarding rock cavity speleophil. FishTraits formula: A_2_4A or B_2_7A = 1
CatMox_PS and CatMox_PA	Catostomid or Moxostomid	FishTraits	Defined as a species belonging to either the genus Catostomus or Moxostomus. FishTraits formula: GID = "Catostom" or "Moxostom")
Centrar_PS and Centrar_PA	Centrarchid	FishTraits	Defined as a species belonging to the family Centrarchidae. FishTraits formula: FID = "Centrarc"
DrtMdSc_PS and DrtMdSc_PA	Darter, Madtom, and Sculpin	FishTraits	Defined as a species that is a member of one of the following genera: Etheostoma, Percina, Noturus, or Cottus. FishTraits formula: GID = "Etheosto", "Percina", "Noturus", or "Cottus"

Table 2.2 Benthic Invertivore Models.

Each row in the table shows the best model for the relationship between the proportion of benthic invertivore species (BenInv_PS) and the land cover variable in that row (Land Cover column) with sites grouped by either Class or RipClass (Group column). The definitions of land cover types are as follows: 1 – Wshd_For = 1 – proportion of watershed forest, 1 – Wshd_Nat = 1 – proportion of watershed natural, Wshd_Ag = proportion of watershed agriculture, Wshd_Urb = proportion of watershed urban,1 - Buf_For = 1 – proportion of riparian buffer forest, 1 – Buf_Nat = 1 – proportion of riparian buffer natural, Buf_Ag = proportion riparian buffer agriculture, Buf_Urb = proportion riparian buffer urban. The Best Model column lists the terms that are present for the best model for each relationship (see text) and the AIC scores are shown as an assessment of model fit. The columns Fixed, Null Test, Area*Slope, and Area give p-values for the appropriate significance tests for the fixed land cover slope, random effects, watershed area – fixed slope interaction and watershed area, respectively. The Plot column gives the figure where the appropriate plot, if applicable, can be found. Note that AIC scores can be used to compare models with a common grouping factor, but cannot be used to compare models between Class and RipClass groupings since sample sizes are not equal.

Metric	Group	Land Cover	Best Model	AIC	Fixed	Null Test	Area * Slope	Area	Plot
BenInv_PS	Class	1 - Wshd_For	Random intercept only	-246.1	n/a	< 0.0001	n/a	n/a	n/a
BenInv_PS	Class	1- Wshd_Nat	Random intercept only	-246.1	n/a	< 0.0001	n/a	n/a	n/a
BenInv_PS	Class	Wshd_Ag	Random Intercept w/ fixed Wshd_Ag	-250.8	0.0081	< 0.0001	n/a	n/a	Fig. 7
BenInv_PS	Class	Wshd_Urb	Random Intercept w/ fixed Wshd_Urb	-262.3	< 0.0001	< 0.0001	n/a	n/a	Fig. 6
BenInv_PS	Class	1 - Buf_For	Random intercept only	-246.1	n/a	< 0.0001	n/a	n/a	n/a
BenInv_PS	Class	1 - Buf_Nat	Random intercept only	-246.1	n/a	< 0.0001	n/a	n/a	n/a
BenInv_PS	Class	Buf_Ag	Random intercept only	-246.1	n/a	< 0.0001	n/a	n/a	n/a
BenInv_PS	Class	Buf_Urb	Random intercept only	-246.1	n/a	< 0.0001	n/a	n/a	n/a
BenInv_PS	RipClass	1 - Wshd_For	Fixed Wshd_For	-211.6	0.0087	n/a	n/a	n/a	Fig. 8
BenInv_PS	RipClass	1 - Wshd_Nat	Fixed Wshd_Nat	-210.1	0.0197	n/a	n/a	n/a	Fig. 9
BenInv_PS	RipClass	Wshd_Ag	Fixed Wshd_Ag * area interaction	-212.7	n/a	n/a	0.0069	n/a	n/a
BenInv_PS	RipClass	Wshd_Urb	Fixed Wshd_Urb w/ area	-222.7	< 0.0001	n/a	n/a	0.0048	n/a
BenInv_PS	RipClass	1 - Buf_For	Fixed Buf_For	-210.9	0.0130	n/a	n/a	n/a	Fig. 10
BenInv_PS	RipClass	1 - Buf_Nat	Intercept only	-206.7	n/a	n/a	n/a	n/a	n/a
BenInv_PS	RipClass	Buf_Ag	Intercept only	-206.7	n/a	n/a	n/a	n/a	n/a
BenInv_PS	RipClass	Buf_Urb	Intercept only	-206.7	n/a	n/a	n/a	n/a	n/a

Table 2.3 Centrarchid Models.

Each row in the table shows the best model for the relationship between the proportion of centrarchid species (Centrar_PS) and the land cover variable in that row (Land Cover column) with sites grouped by either Class or RipClass (Group column). The definitions of land cover types are as follows: 1 – Wshd_For = 1 – proportion of watershed forest, 1 – Wshd_Nat = 1 – proportion of watershed natural, Wshd_Ag = proportion of watershed agriculture, Wshd_Urb = proportion of watershed urban,1 - Buf_For = 1 – proportion of riparian buffer forest, 1 – Buf_Nat = 1 – proportion of riparian buffer natural, Buf_Ag = proportion riparian buffer agriculture, Buf_Urb = proportion riparian buffer urban. The Best Model column lists the terms that are present for the best model for each relationship (see text) and the AIC scores are shown as an assessment of model fit. The columns Fixed, Null Test, Area*Slope, and Area give p-values for the appropriate significance tests for the fixed land cover slope, random effects, and watershed area – fixed slope interaction, respectively. The Plot column gives the figure where the appropriate plot, if applicable, can be found. Note that AIC scores can be used to compare models with a common grouping factor, but cannot be used to compare models between Class and RipClass groupings since sample sizes are not equal.

Metric	Group	Land Cover	Best Model	AIC	Fixed	Null Test	Area * Slope	Plot
Centrar_PS	Class	1 - Wshd_For	Fixed Wshd_For w/ area interaction	-253.3	< 0.0001	n/a	< 0.0001	n/a
Centrar_PS	Class	1- Wshd_Nat	Fixed Wshd_Nat w/ area interaction	-251.9	< 0.0001	n/a	< 0.0001	n/a
Centrar_PS	Class	Wshd_Ag	Random intercept and Wshd_Ag	-256.8	n/a	< 0.0001	n/a	Fig. 11
Centrar_PS	Class	Wshd_Urb	Fixed Wshd_Urb	-256.3	< 0.0001	n/a	n/a	Fig. 12
Centrar_PS	Class	1 - Buf_For	Random intercept and Buf_For	-243.6	n/a	0.0006	n/a	Fig. 13
Centrar_PS	Class	1 - Buf_Nat	Random intercept and Buf_Nat	-242.6	n/a	0.001	n/a	Fig. 14
Centrar_PS	Class	Buf_Ag	Random intercept only	-241.5	n/a	0.0008	n/a	n/a
Centrar_PS	Class	Buf_Urb	Random intercept w/ fixed Buf_Urb	-243.6	0.0428	0.001	n/a	Fig. 15
Centrar_PS	RipClass	1 - Wshd_For	Fixed Wshd_For w/ area interaction	-210.5	n/a	0.0046	0.0006	n/a
Centrar_PS	RipClass	1 - Wshd_Nat	Fixed Wshd_Nat w/ area interaction	-209.4	n/a	n/a	0.0008	n/a
Centrar_PS	RipClass	Wshd_Ag	Fixed Wshd_Ag w/ area interaction	-198.0	n/a	n/a	0.0115	n/a
Centrar_PS	RipClass	Wshd_Urb	Fixed Wshd_Urb	-211.4	< 0.0001	n/a	n/a	Fig. 12
Centrar_PS	RipClass	1 - Buf_For	Intercept only	-197.2	n/a	n/a	n/a	n/a
Centrar_PS	RipClass	1 - Buf_Nat	Intercept only	-197.2	n/a	n/a	n/a	n/a
Centrar_PS	RipClass	Buf_Ag	Intercept only	-197.2	n/a	n/a	n/a	n/a
Centrar_PS	RipClass	Buf_Urb	Intercept only	-197.2	n/a	n/a	n/a	n/a

Table 2.4 Darter, Madtom and Sculpin Models.

Each row in the table shows the best model for the relationship between the proportion of darter, madtom and sculpin species (DrtMdSc_PS) and the land cover variable in that row (Land Cover column) with sites grouped by either Class or RipClass (Group column). The definitions of land cover types are as follows: 1 – Wshd_For = 1 – proportion of watershed forest, 1 – Wshd_Nat = 1 – proportion of watershed natural, Wshd_Ag = proportion of watershed agriculture, Wshd_Urb = proportion of watershed urban,1 - Buf_For = 1 – proportion of riparian buffer forest, 1 – Buf_Nat = 1 – proportion of riparian buffer natural, Buf_Ag = proportion riparian buffer agriculture, Buf_Urb = proportion riparian buffer urban. The Best Model column lists the terms that are present for the best model for each relationship (see text) and the AIC scores are shown as an assessment of model fit. The columns Fixed, Null Test, Area*Slope, and Area give p-values for the appropriate significance tests for the fixed land cover slope, random effects, watershed area – fixed slope interaction and watershed area, respectively. The Plot column gives the figure where the appropriate plot, if applicable, can be found. Note that AIC scores can be used to compare models with a common grouping factor, but cannot be used to compare models between Class and RipClass groupings since sample sizes are not equal.

Metric	Group	Land Cover	Best Model	AIC	Fixed	Null Test	Area * Slope	Area	Plot
DrtMdSc_PS	Class	1 - Wshd_For	Fixed Wshd_For	-260.6	0.0001	n/a	n/a	n/a	Fig. 17
DrtMdSc_PS	Class	1 - Wshd_Nat	Random intercept only	-259.8	n/a	0.0002	n/a	n/a	n/a
DrtMdSc_PS	Class	Wshd_Ag	Random intercept and Wshd_Ag w/ fixed Wshd_Ag	-275.2	0.0452	< 0.0001	n/a	n/a	Fig. 16
DrtMdSc_PS	Class	Wshd_Urb	Random intercept w/ fixed Wshd_Urb * area interaction	-295.8	n/a	0.0121	0.0352	n/a	n/a
DrtMdSc_PS	Class	1 - Buf_For	Random intercept Only	-259.8	n/a	0.0002	n/a	n/a	n/a
DrtMdSc_PS	Class	1 - Buf_Nat	Random intercept Only	-259.8	n/a	0.0002	n/a	n/a	n/a
DrtMdSc_PS	Class	Buf_Ag	Random intercept Only	-259.8	n/a	0.0002	n/a	n/a	n/a
DrtMdSc_PS	Class	Buf_Urb	Random intercept Only	-259.8	n/a	0.0002	n/a	n/a	n/a
DrtMdSc_PS	RipClass	1 - Wshd_For	Fixed Wshd_For	-240.1	0.0002	n/a	n/a	n/a	Fig. 17
DrtMdSc_PS	RipClass	1 - Wshd_Nat	Fixed Wshd_Nat	-238.7	0.0004	n/a	n/a	n/a	Fig. 18
DrtMdSc_PS	RipClass	Wshd_Ag	Random intercept w/ fixed Wshd_Ag * area interaction	-239.5	n/a	0.0007	0.0142	n/a	n/a
DrtMdSc_PS	RipClass	Wshd_Urb	Fixed Wshd_Urb w/ area	-265.9	< 0.0001	n/a	n/a	0.0101	n/a
DrtMdSc_PS	RipClass	1 - Buf_For	Fixed Buf_For	-232.4	0.0134	n/a	n/a	n/a	Fig. 19
DrtMdSc_PS	RipClass	1 - Buf_Nat	Random intercept only	-231.7	n/a	0.0195	n/a	n/a	n/a
DrtMdSc_PS	RipClass	Buf_Ag	Random intercept only	-231.7	n/a	0.0195	n/a	n/a	n/a
DrtMdSc_PS	RipClass	Buf_Urb	Random Intercept w/ fixed Buf_Urb * area interaction	-232.9	n/a	0.0436	0.0499	n/a	n/a

Table 2.5 Omnivore Models.
Each row in the table shows the best model for the relationship between the proportion of omnivore species (Omni_PS) and the land cover variable in that row (Land Cover column) with sites grouped by either Class or RipClass (Group column). The definitions of land cover types are as follows: 1 – Wshd_For = 1 – proportion of watershed forest, 1 – Wshd_Nat = 1 – proportion of watershed natural, Wshd_Ag = proportion of watershed agriculture, Wshd_Urb = proportion of watershed urban,1 - Buf_For = 1 – proportion of riparian buffer forest, 1 – Buf_Nat = 1 – proportion of riparian buffer natural, Buf_Ag = proportion riparian buffer agriculture, Buf_Urb = proportion riparian buffer urban. The Best Model column lists the terms that are present for the best model for each relationship (see text) and the AIC scores are shown as an assessment of model fit. The columns Fixed, Null Test, Area*Slope, and Area give p-values for the appropriate significance tests for the fixed land cover slope, random effects, watershed area – fixed slope interaction and watershed area, respectively. The Plot column gives the figure where the appropriate plot, if applicable, can be found. Note that AIC scores can be used to compare models with a common grouping factor, but cannot be used to compare models between Class and RipClass groupings since sample sizes are not equal.

Metric	Group	Land Cover	Best Model	AIC	Fixed	Null Test	Area * Slope	Area
Omni_PS	Class	1 - Wshd_For	Fixed Wshd_For w/ area	-341.6	< 0.0001	n/a	n/a	0.0004
Omni_PS	Class	1 - Wshd_Nat	Random intercept and Wshd_Nat w/ area	-341.5	n/a	0.0002	n/a	< 0.0001
Omni_PS	Class	Wshd_Ag	Random intercept and Wshd_Ag w/ area	-341.2	n/a	0.0002	n/a	< 0.0001
Omni_PS	Class	Wshd_Urb	Random intercept w/ Fixed Wshd_Urb * area interaction	-341.7	n/a	0.0101	0.0161	n/a
Omni_PS	Class	1 - Buf_For	Random intercept w/ area	-338.5	n/a	0.0004	n/a	0.0001
Omni_PS	Class	1 - Buf_Nat	Random intercept w/ area	-338.5	n/a	0.0004	n/a	0.0001
Omni_PS	Class	Buf_Ag	Random intercept w/ area	-338.5	n/a	0.0004	n/a	0.0001
Omni_PS	Class	Buf_Urb	Random intercept w/ area	-338.5	n/a	0.0004	n/a	0.0001
Omni_PS	RipClass	1 - Wshd_For	Fixed Wshd_For w/ area	-284.1	0.0002	n/a	n/a	0.0017
Omni_PS	RipClass	1 - Wshd_Nat	Fixed Wshd_Nat w/ area	-283.0	0.0003	n/a	n/a	0.0024
Omni_PS	RipClass	Wshd_Ag	Area only	-272.1	n/a	n/a	n/a	0.0323
Omni_PS	RipClass	Wshd_Urb	Fixed Wshd_Urb *area interaction	-284.2	n/a	n/a	0.0152	n/a
Omni_PS	RipClass	1 - Buf_For	Fixed Buf_For w/ area	-276.3	0.013	n/a	n/a	0.0029
Omni_PS	RipClass	1 - Buf_Nat	Fixed Buf_Nat * area interaction	-274.3	n/a	n/a	0.0228	n/a
Omni_PS	RipClass	Buf_Ag	Area only	-272.1	n/a	n/a	n/a	0.0323
Omni_PS	RipClass	Buf_Urb	Fixed Buf_Urb * area interaction	-274.9	n/a	n/a	0.0252	n/a

Table 2.6 Substrate Generalist Models.

Each row in the table shows the best model for the relationship between the proportion of substrate generalist species (SubGen_PS) and the land cover variable in that row (Land Cover column) with sites grouped by either Class or RipClass (Group column). The definitions of land cover types are as follows: 1 – Wshd_For = 1 – proportion of watershed forest, 1 – Wshd_Nat = 1 – proportion of watershed natural, Wshd_Ag = proportion of watershed agriculture, Wshd_Urb = proportion of watershed urban,1 - Buf_For = 1 – proportion of riparian buffer forest, 1 – Buf_Nat = 1 – proportion of riparian buffer natural, Buf_Ag = proportion riparian buffer agriculture, Buf_Urb = proportion riparian buffer urban. The Best Model column lists the terms that are present for the best model for each relationship (see text) and the AIC scores are shown as an assessment of model fit. The columns Fixed, Null Test, Area*Slope, and Area give p-values for the appropriate significance tests for the fixed land cover slope, random effects, watershed area – fixed slope interaction and watershed area, respectively. The Plot column gives the figure where the appropriate plot, if applicable, can be found. Note that AIC scores can be used to compare models with a common grouping factor, but cannot be used to compare models between Class and RipClass groupings since sample sizes are not equal.

Group	Land Cover	Metric	Best Model	AIC	Fixed	Null Test	Area * Slope	Area
Class	1 - Wshd_For	SubGen_PS	Fixed Wshd_For * area interaction	-211.2	n/a	< 0.0001	0.0221	n/a
Class	1 - Wshd_Nat	SubGen_PS	Fixed Wshd_Nat w/ area	-206.3	< 0.0001	n/a	n/a	0.0033
Class	Wshd_Ag	SubGen_PS	Random intercept w/ fixed Wshd_Ag and area	-205.2	0.0336	< 0.0001	n/a	0.0011
Class	Wshd_Urb	SubGen_PS	Random intercept w/ fixed Wshd_Urb and area	-220.3	< 0.0001	< 0.0001	n/a	< 0.0001
Class	1 - Buf_For	SubGen_PS	Random intercept w/ fixed Buf_For * area interaction	-203.7	n/a	< 0.0001	0.0386	n/a
Class	1 - Buf_Nat	SubGen_PS	Random intercept w/ fixed Buf_Nat * area interaction	-204.9	n/a	< 0.0001	0.029	n/a
Class	Buf_Ag	SubGen_PS	Random intercept w/ area	-203.0	n/a	< 0.0001	n/a	0.0051
Class	Buf_Urb	SubGen_PS	Random intercept w/ fixed Buf_Urb and area	-205.7	0.0317	< 0.0001	n/a	0.0012
RipClass	1 - Wshd_For	SubGen_PS	Fixed Wshd_For * area interaction	-181.9	n/a	n/a	0.0433	n/a
RipClass	1 - Wshd_Nat	SubGen_PS	Fixed Wshd_Nat w/ area	-178.6	< 0.0001	n/a	n/a	0.0079
RipClass	Wshd_Ag	SubGen_PS	Random intercept w/ area	-164.1	n/a	< 0.0001	n/a	0.0103
RipClass	Wshd_Urb	SubGen_PS	Random intercept w/ fixed Wshd_Urb and area	-185.1	< 0.0001	< 0.0001	n/a	< 0.0001
RipClass	1 - Buf_For	SubGen_PS	Random intercept w/ area	-164.1	n/a	< 0.0001	n/a	0.0103
RipClass	1 - Buf_Nat	SubGen_PS	Random intercept w/ fixed Buf_Nat * area interaction	-165.6	n/a	< 0.0001	0.0287	n/a
RipClass	Buf_Ag	SubGen_PS	Random intercept w/ area	-164.1	n/a	< 0.0001	n/a	0.0103
RipClass	Buf_Urb	SubGen_PS	Random intercept w/ area	-164.1	n/a	< 0.0001	n/a	0.0103

Figure 2.1 Map of Study Area.
The map at bottom shows the location of the study area in the U.S. NAWQA Sites (triangles) and REMAP sites (plus symbols) are shown nested within zoogeographic region A2A, the Upper Mississippi River Basin. The sites have a fairly even spread across the study region and cover a range of LDPR Classes, represented by the different colors in the background.

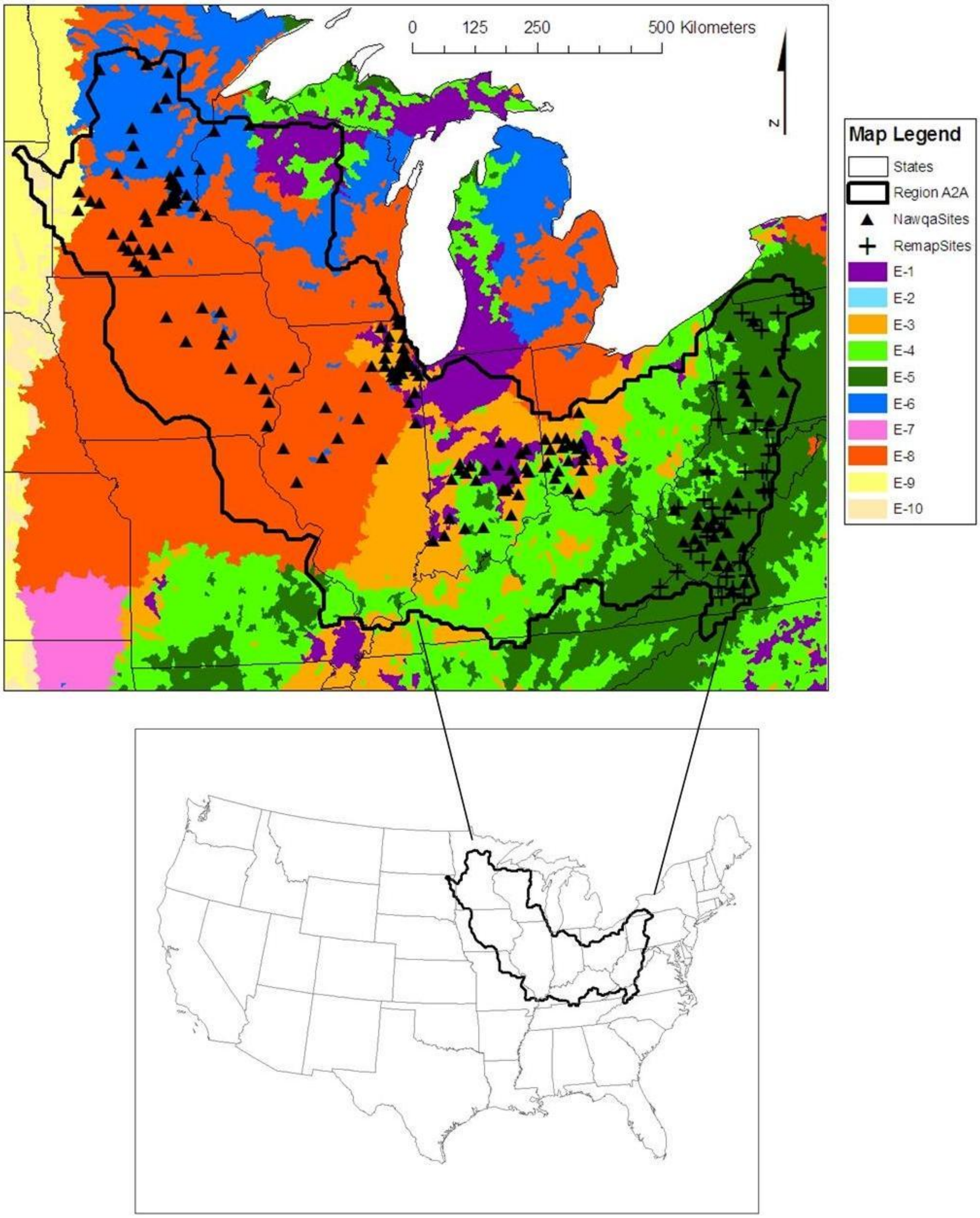


Figure 2.2 a – e Boxplots for watershed characteristics of LDPR Classes.
Boxplots display (a) watershed area, and (b) watershed forest, (c) natural, (d) agriculture and (e) urban land cover characteristics of the 149 fish sites grouped by Eastern LDPR watershed Classes. The numbers of fish sampling sites (n) in each Class are displayed at the top of the watershed area box plot (a).

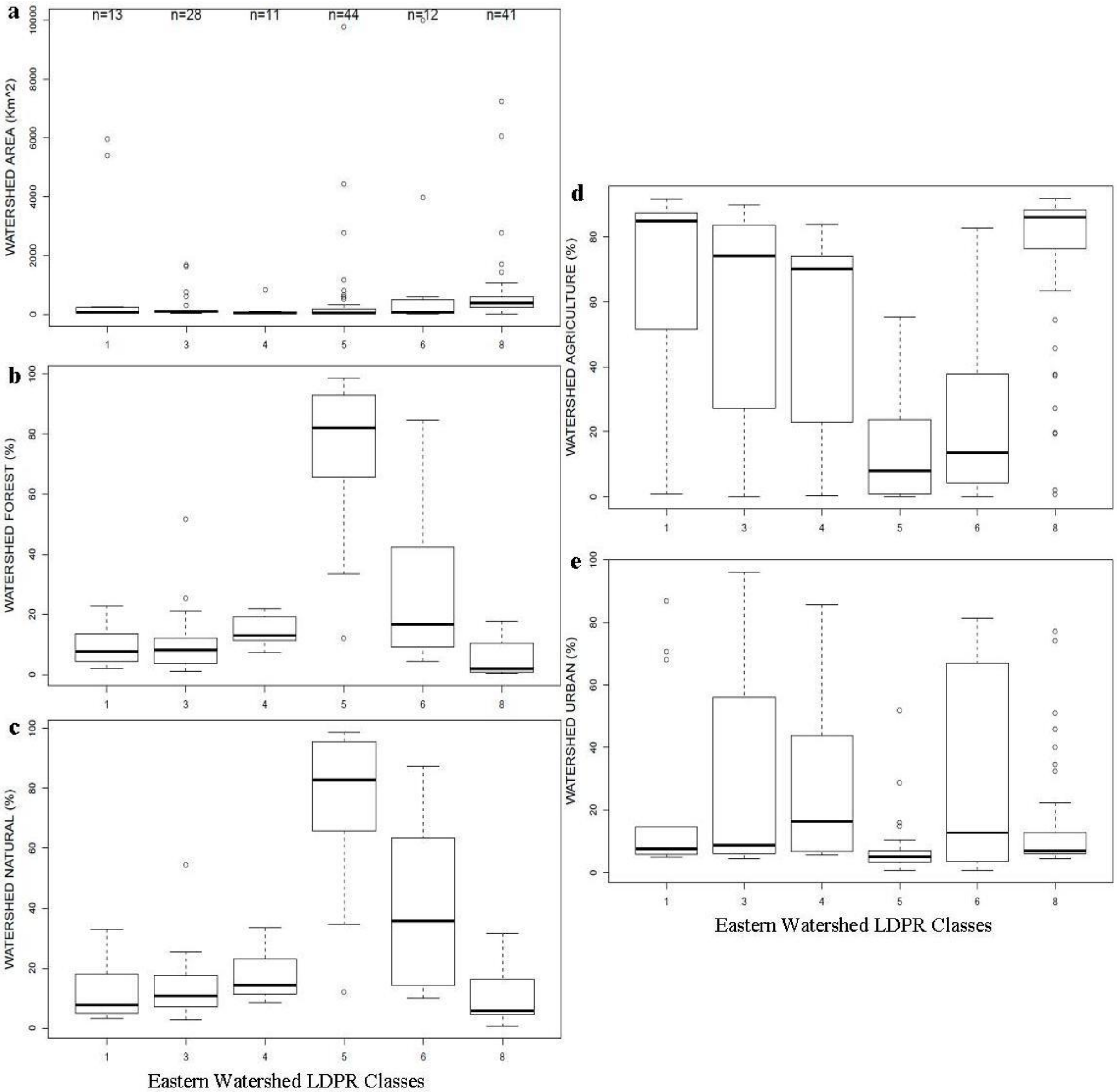


Figure 2.3 a – d Boxplots for riparian land cover of LDPR Classes.
Boxplots display (a) riparian forest, (b) natural, (c) agriculture and (d) urban land cover characteristics of the 149 fish sites grouped by Eastern LDPR watershed Classes.

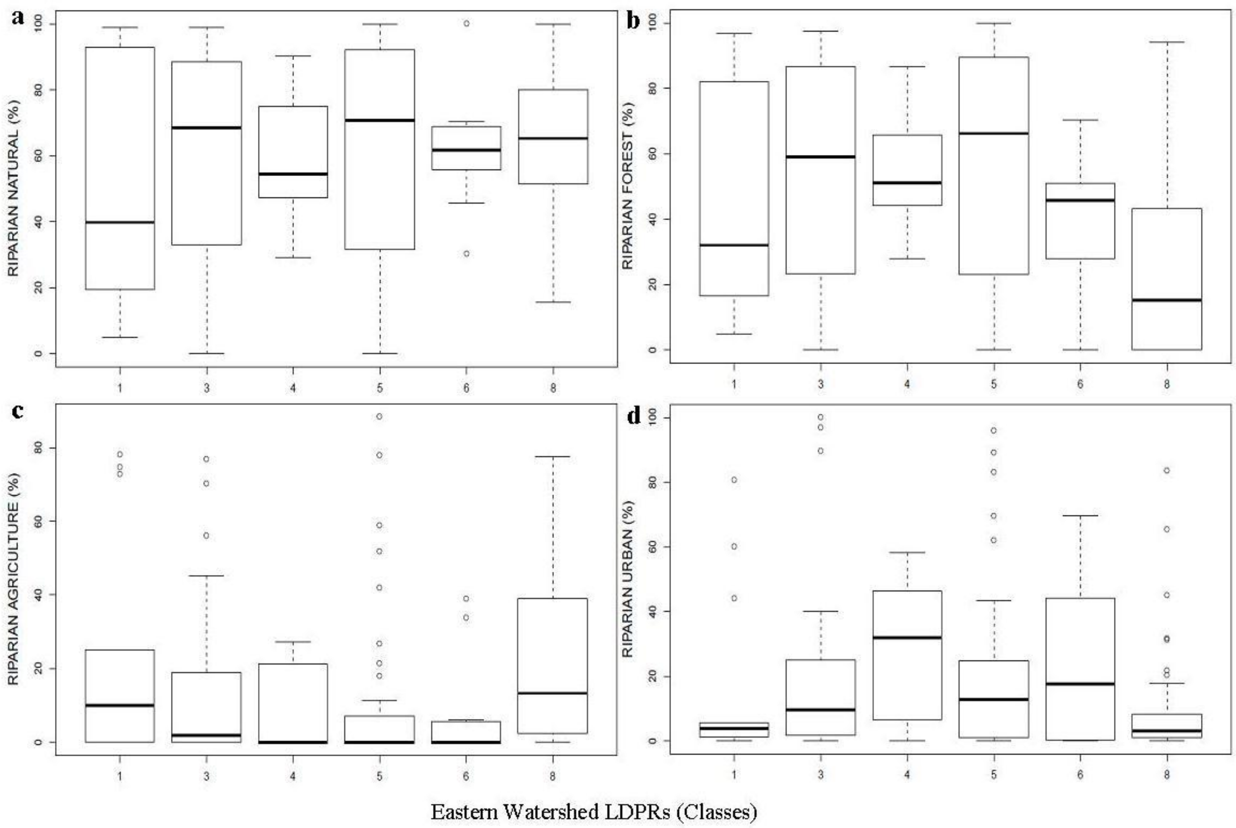


Figure 2.4 a – e Boxplots for watershed characteristics of LDPR RipClasses.
Boxplots display watershed area (a), and watershed forest (b), natural (c), agriculture (d) and urban (e) land cover characteristics of the 149 fish sites grouped by Eastern LDPR riparian RipClasses. The numbers of fish sampling sites (n) in each Class are displayed at the top of the watershed area box plot (a).

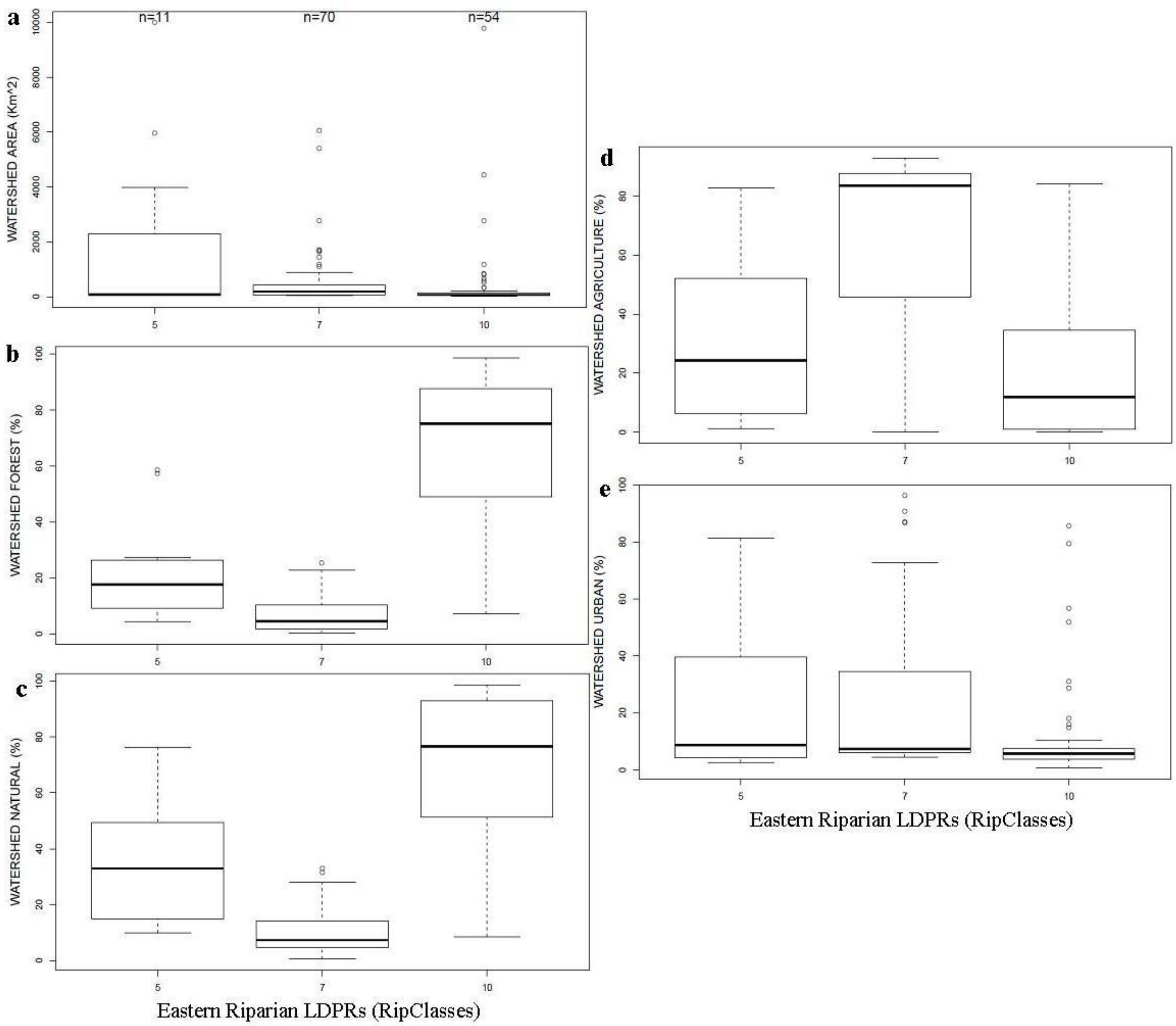


Figure 2.5 a –d Boxplots for riparian land cover of LDPR RipClasses.
Boxplots display (a) riparian forest, (b) natural, (c) agriculture and (d) urban land cover characteristics of the 149 fish sites grouped by Eastern LDPR riparian RipClasses.

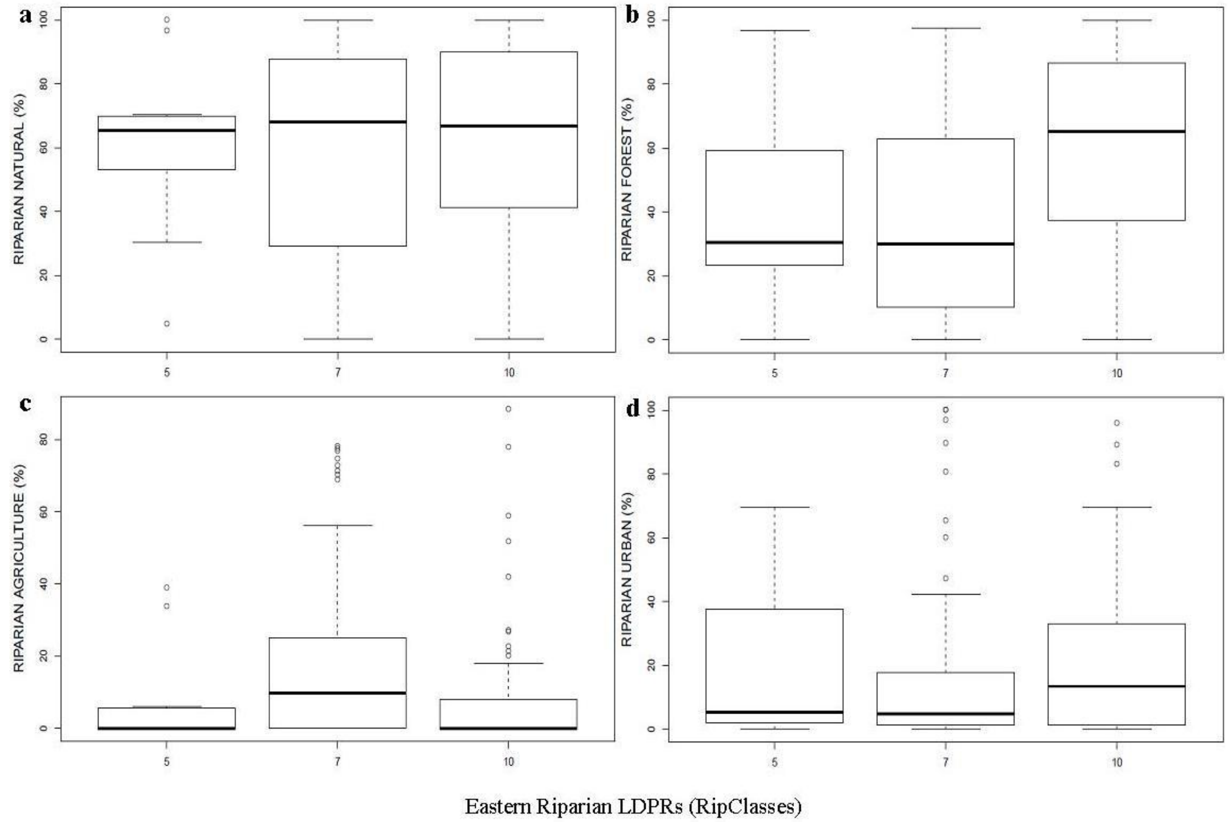


Figure 2.6 Benthic invertivore and watershed urban plot.
Plot of the relationship between proportion benthic invertivore species (BenInv_PS) and proportion watershed urban land use (Wshd_Urb), with different lines for each Class to represent the random intercepts model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

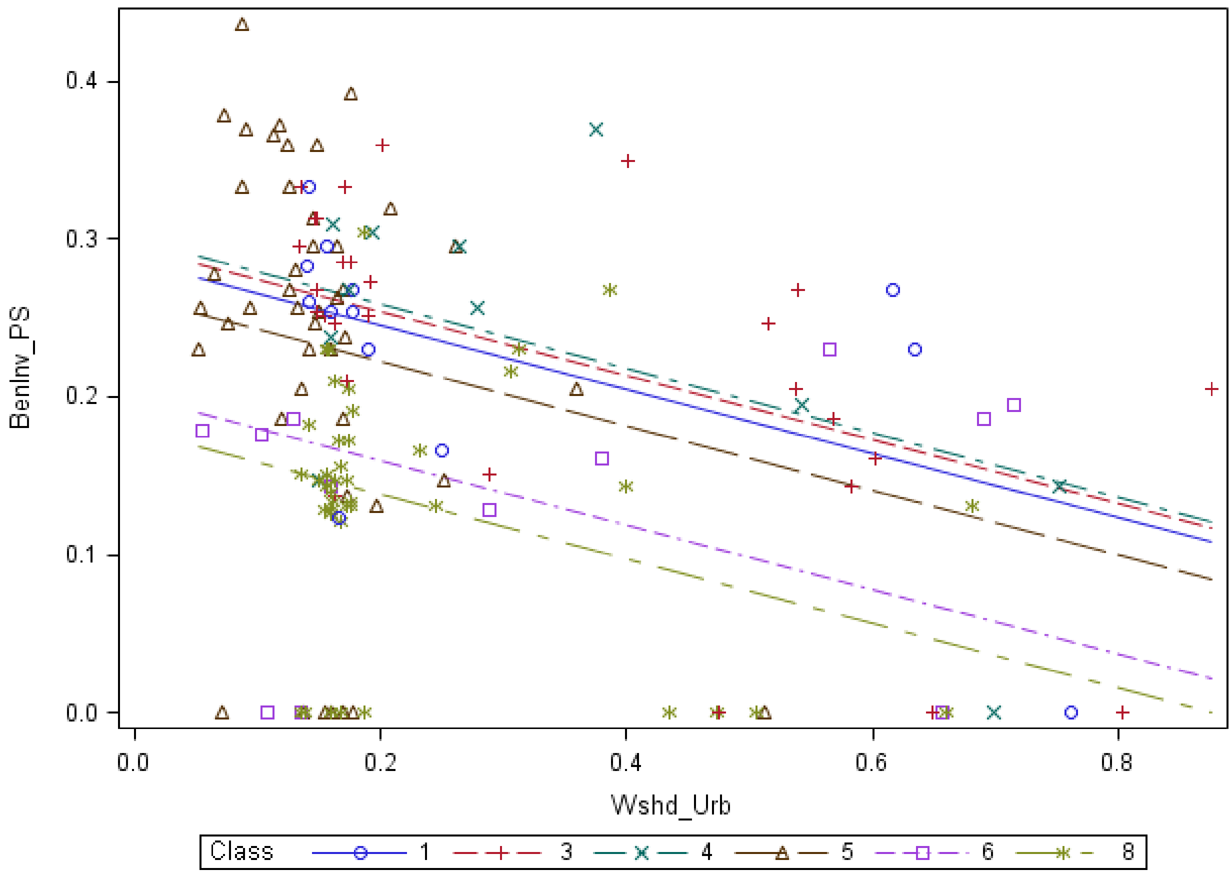


Figure 2.7 Benthic invertivore and watershed agriculture plot.
Plot of the relationship between proportion benthic invertivore species (BenInv_PS) and proportion watershed agriculture land use (Wshd_Ag), with different lines for each Class to represent the random intercepts model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

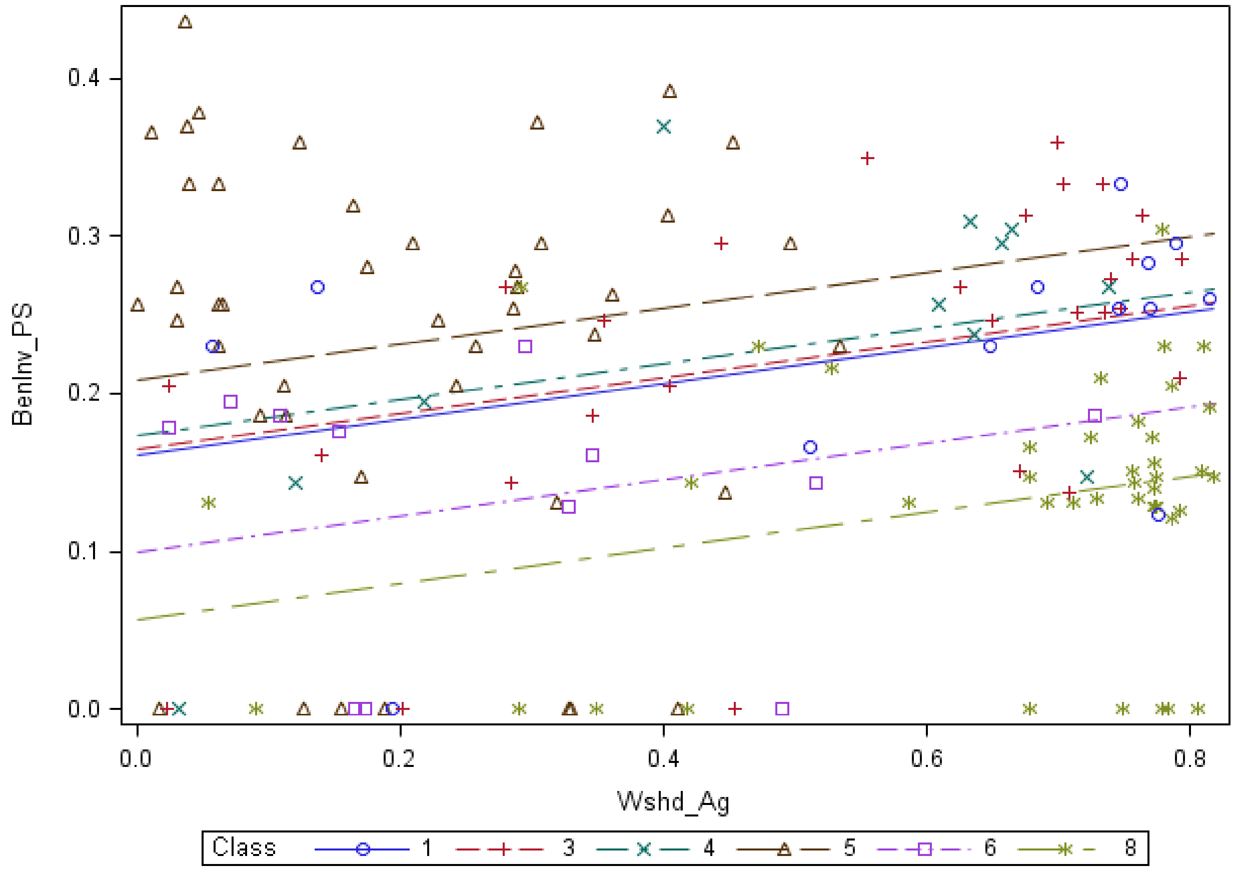


Figure 2.8 Benthic invertivore and watershed forest plot.
Plot of the relationship between proportion benthic invertivore species (BenInv_PS) and 1 – proportion of watershed forest land cover (Wshd_For). Only one line is shown for all RipClasses as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

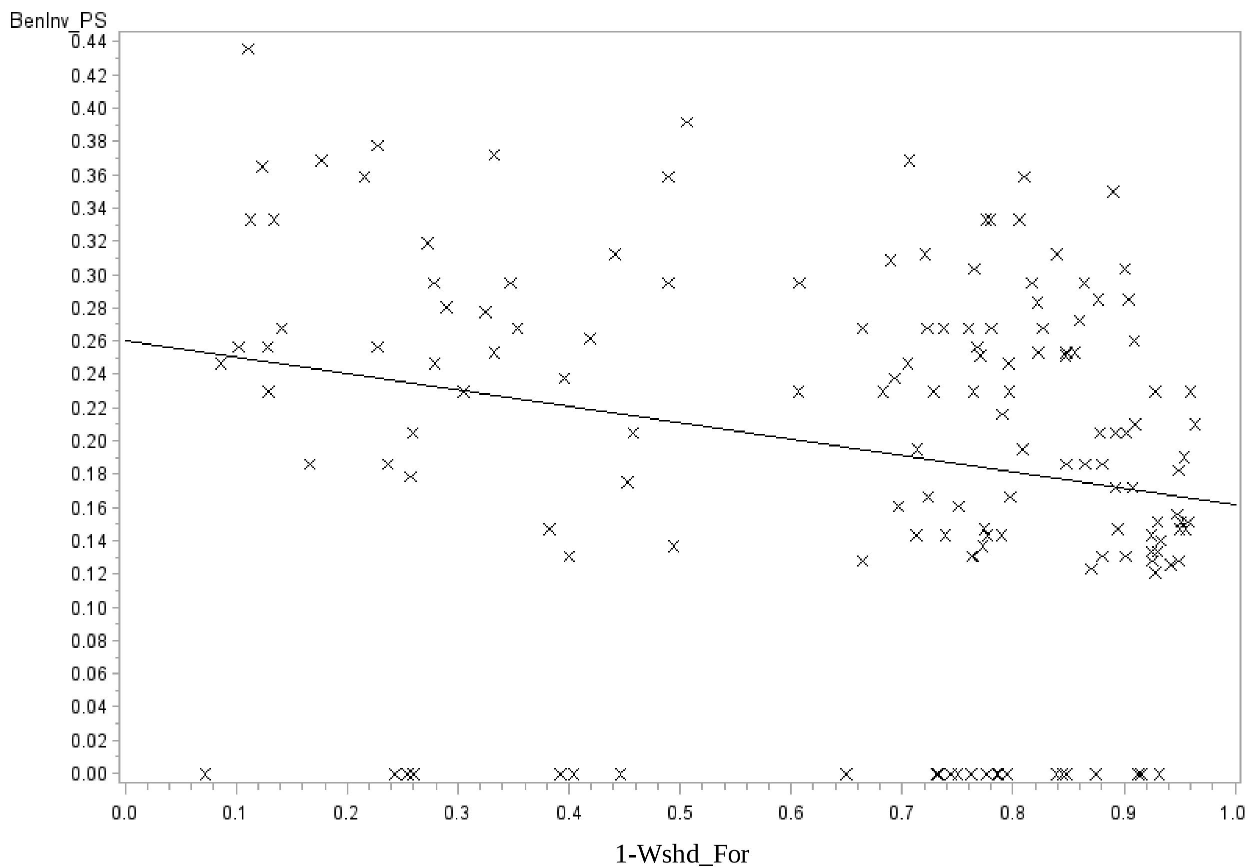


Figure 2.9 Benthic invertivore and watershed natural plot.
Plot of the relationship between proportion benthic invertivore species (BenInv_PS) and 1 – proportion of watershed natural land cover (Wshd_Nat). Only one line is shown for all RipClasses as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by $2/\pi$ to rescale from 0 to 1.

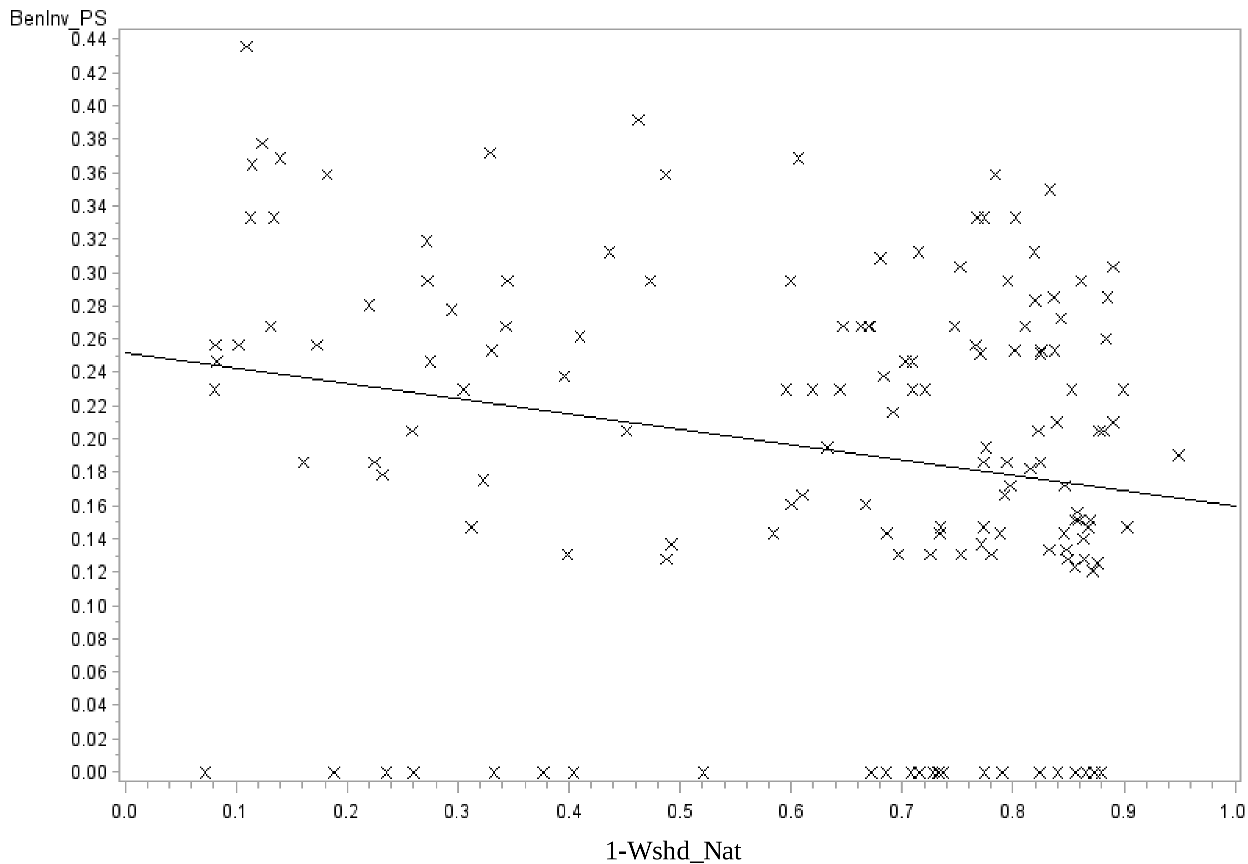


Figure 2.10 Benthic invertivore and riparian forest plot.
Plot of the relationship between proportion benthic invertivore species (BenInv_PS) and 1 – proportion of riparian buffer forest (Buf_For). Only one line is shown for all RipClasses as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

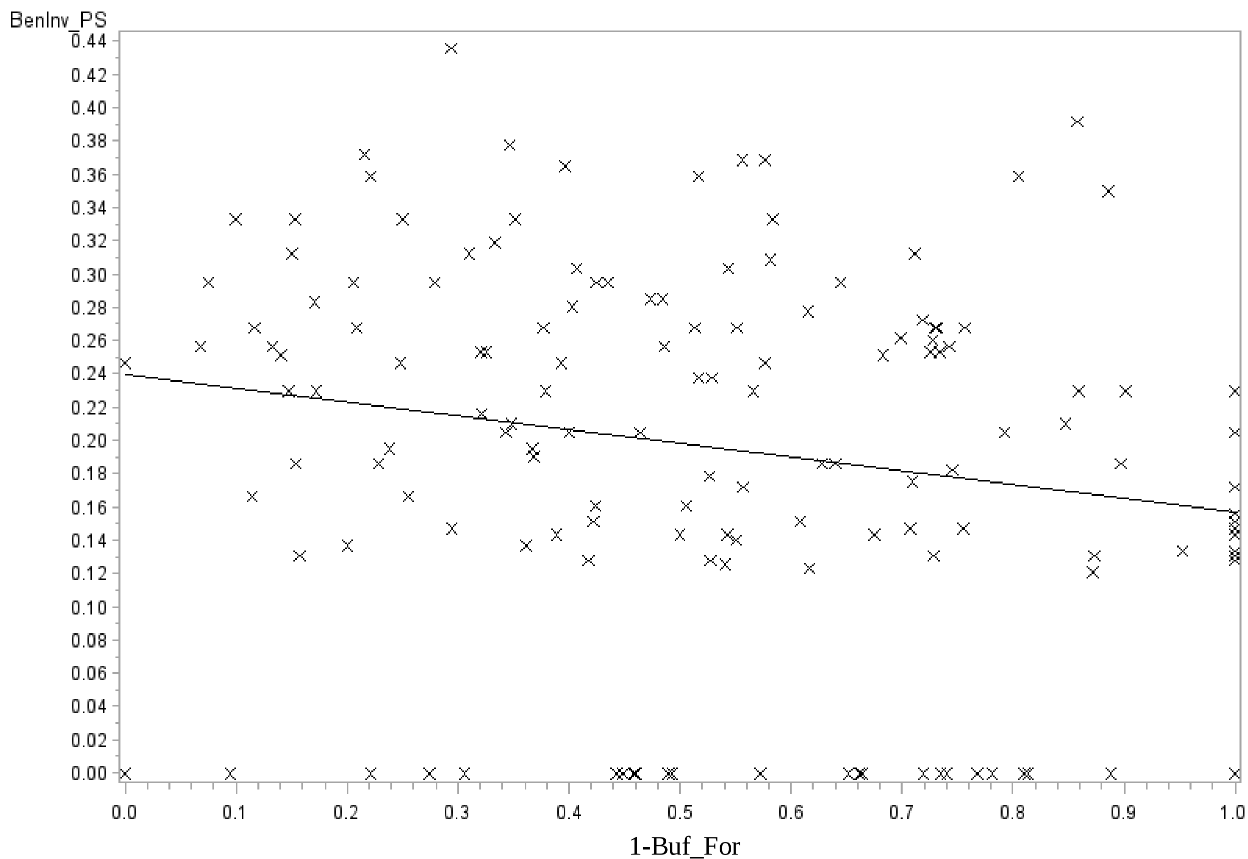


Figure 2.11 Centrarchid and watershed agriculture plot.
Plot of the relationship between proportion Centrarchid species (Centrar_PS) and proportion watershed agriculture land use (Wshd_Ag), with different lines for each Class to represent the random intercepts and slopes model. Both variables have been arcsin(square root) transformed and multiplied by $2/\pi$ to rescale from 0 to 1.

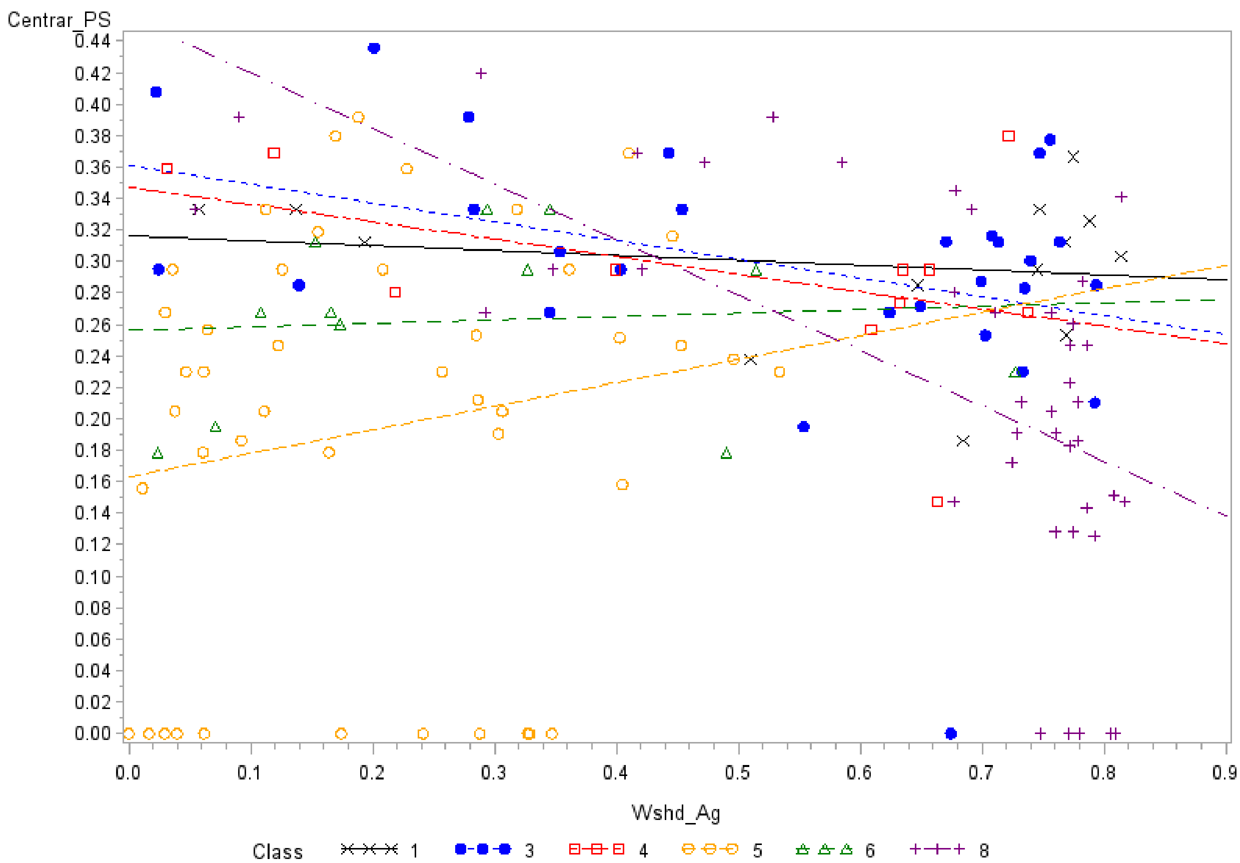


Figure 2.12 Centrarchid and watershed urban plot.

Plot of the relationship between proportion Centrarchid species (Centrar_PS) and proportion watershed urban land use (Wshd_Urb). Only one line is shown for all Classes as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by $2/\pi$ to rescale from 0 to 1.

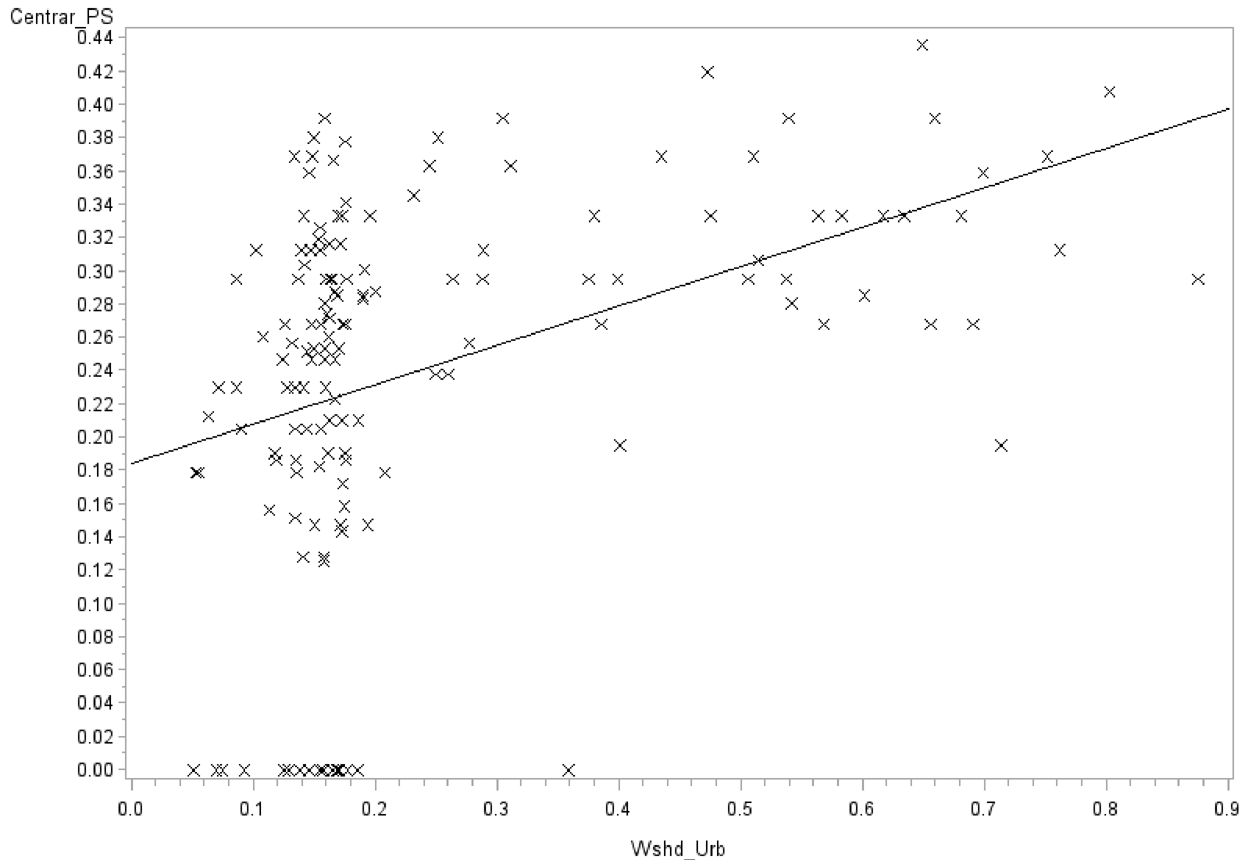


Figure 2.13 Centrarchid and riparian forest plot.
Plot of the relationship between proportion Centrarchid species (Centrar_PS) and 1 - proportion riparian buffer forest land cover (Buf_For), with different lines for each Class to represent the random intercepts and slopes model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

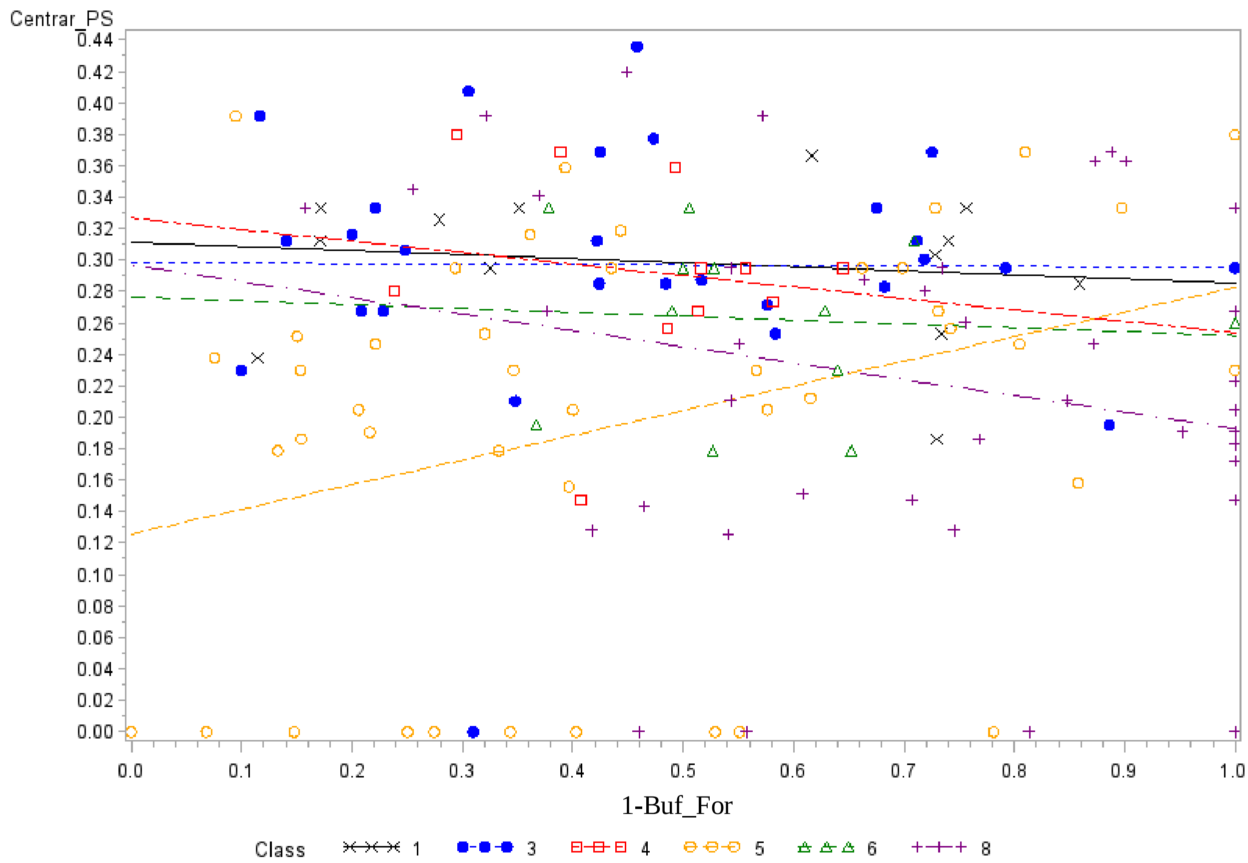


Figure 2.14 Centrarchid and riparian natural plot.

Plot of the relationship between proportion Centrarchid species (Centrar_PS) and 1 - proportion riparian buffer natural land cover (Buf_Nat), with different lines for each Class to represent the random intercepts and slopes model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

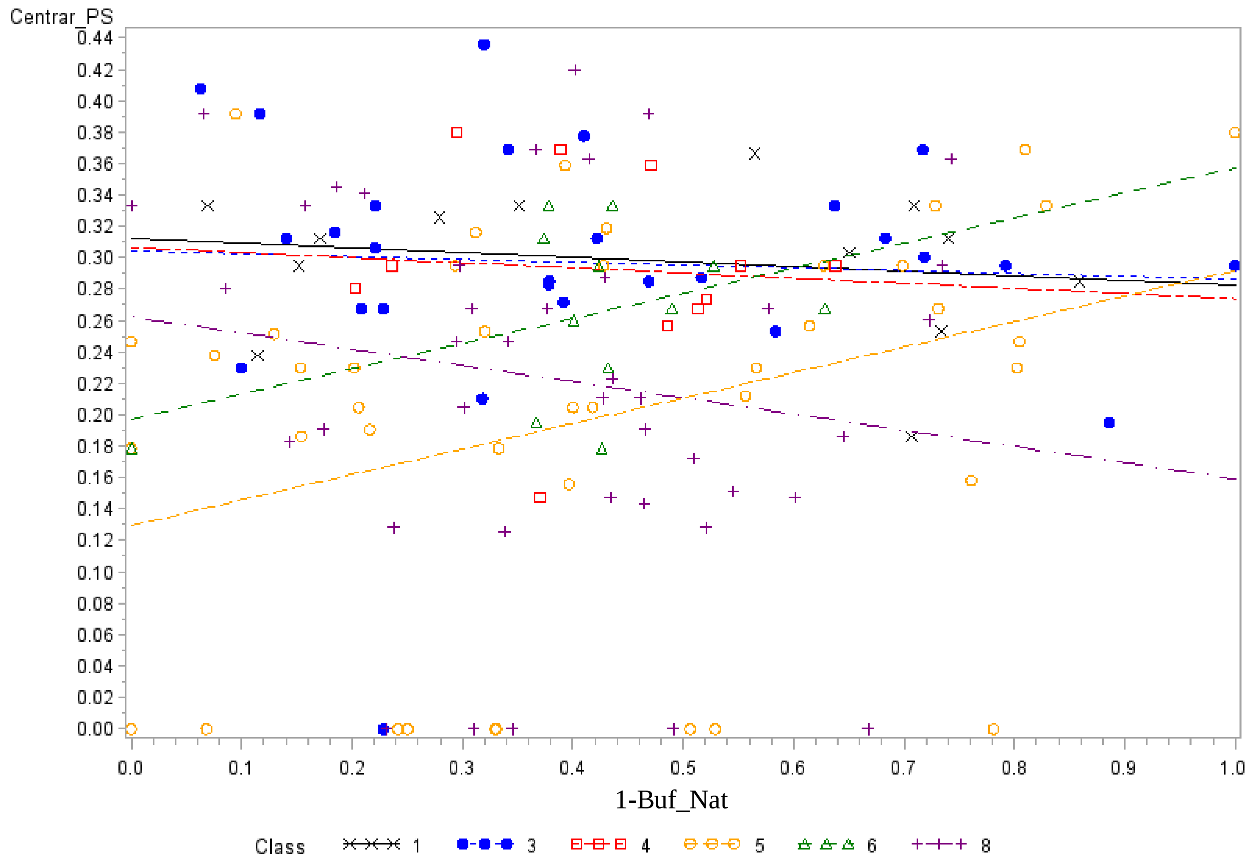


Figure 2.15 Centrarchid and riparian urban plot.

Plot of the relationship between proportion of Centrarchid species (Centrar_PS) and proportion of riparian urban land use (Buf_Urb), with different lines for each Class to represent the random intercepts model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

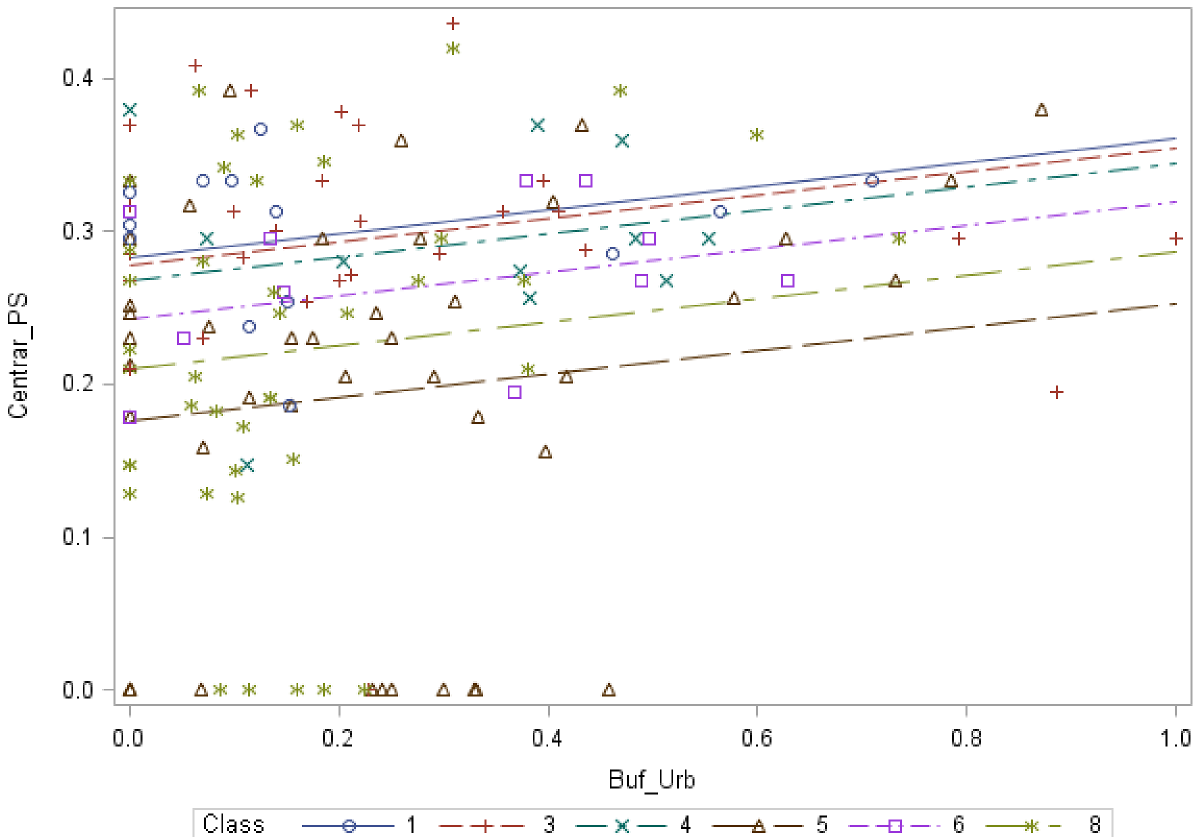


Figure 2.16 Darter, madtom, sculpin and watershed agriculture plot.
Plot of the relationship between proportion darters, madtom, and sculpin species (DrtMdSc_PS) and proportion watershed agriculture land use (Wshd_Ag), with different lines for each Class to represent the random intercepts and slopes model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

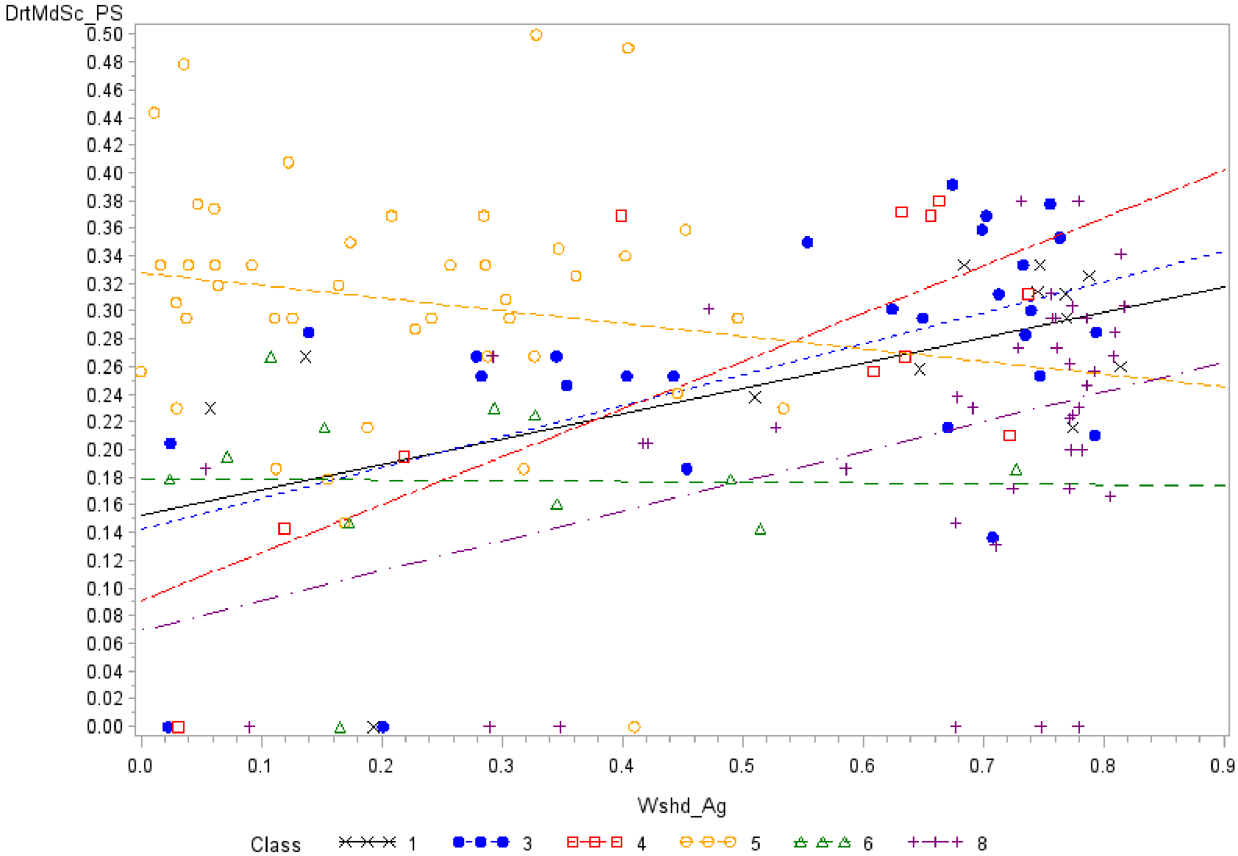


Figure 2.17 Darter, madtom, sculpin and watershed forest plot.
Plot of the relationship between proportion darter, madtom, and sculpin species (BenInv_PS) and 1 – proportion of watershed forest (Wshd_For). Only one line is shown for all Classes as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

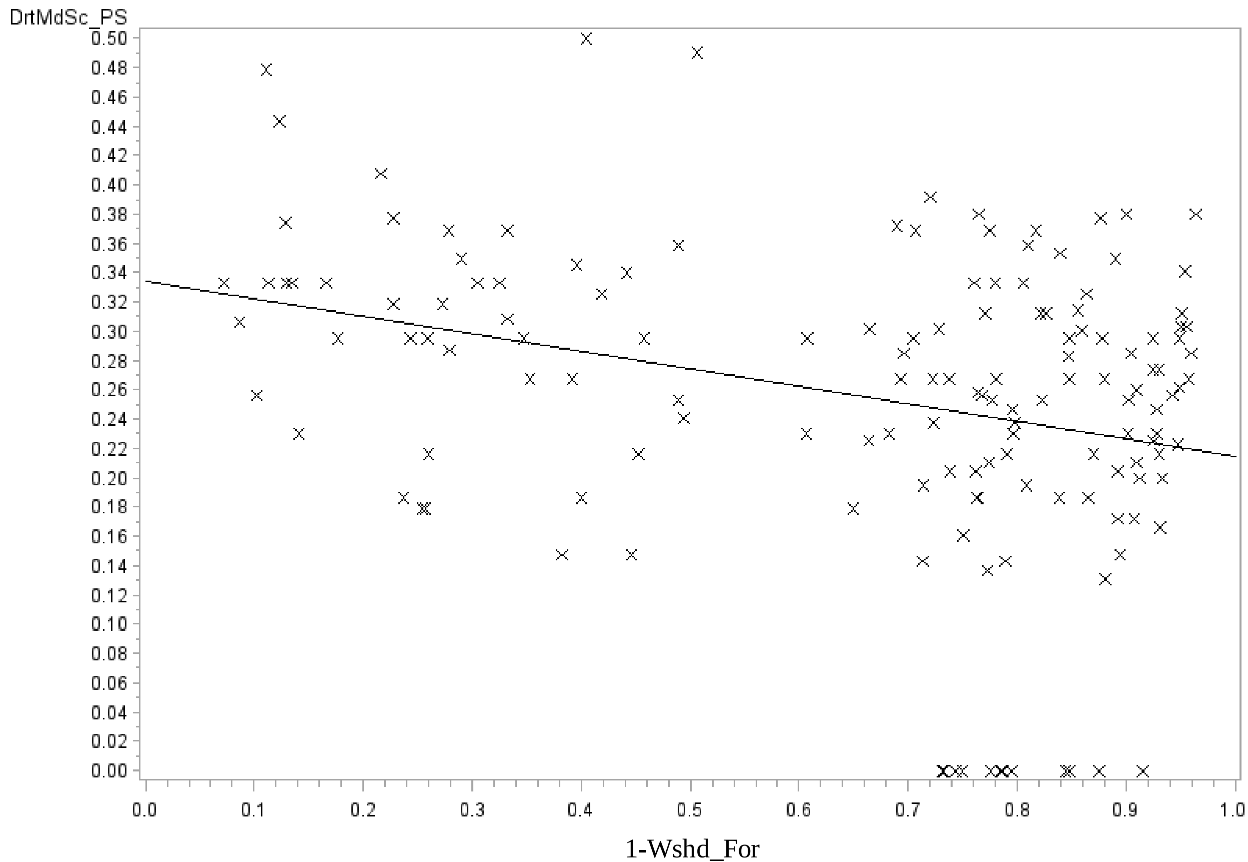


Figure 2.18 Darter, madtom, sculpin and watershed natural.
Plot of the relationship between proportion darter, madtom, and sculpin species (BenInv_PS) and 1 – proportion of watershed natural (Wshd_For). Only one line is shown for all Classes as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by 2/π to rescale from 0 to 1.

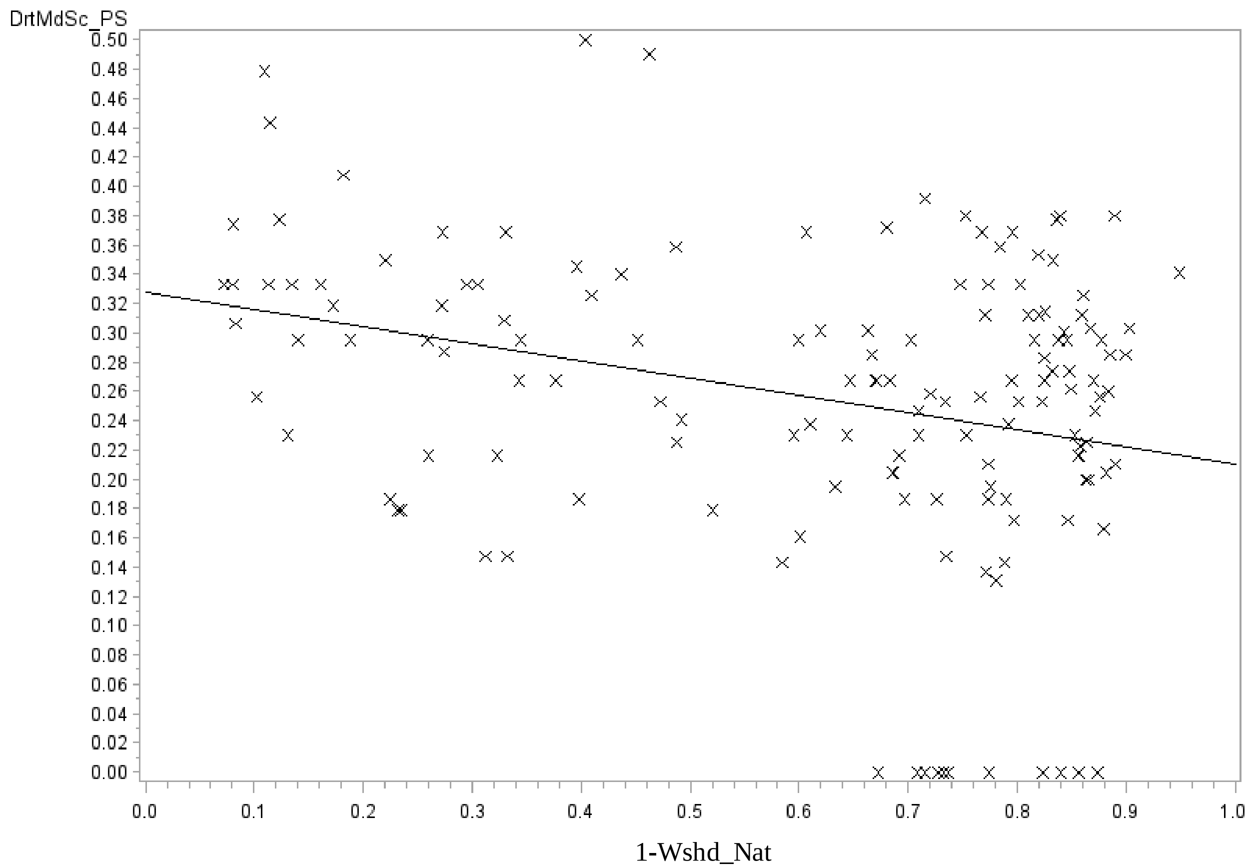
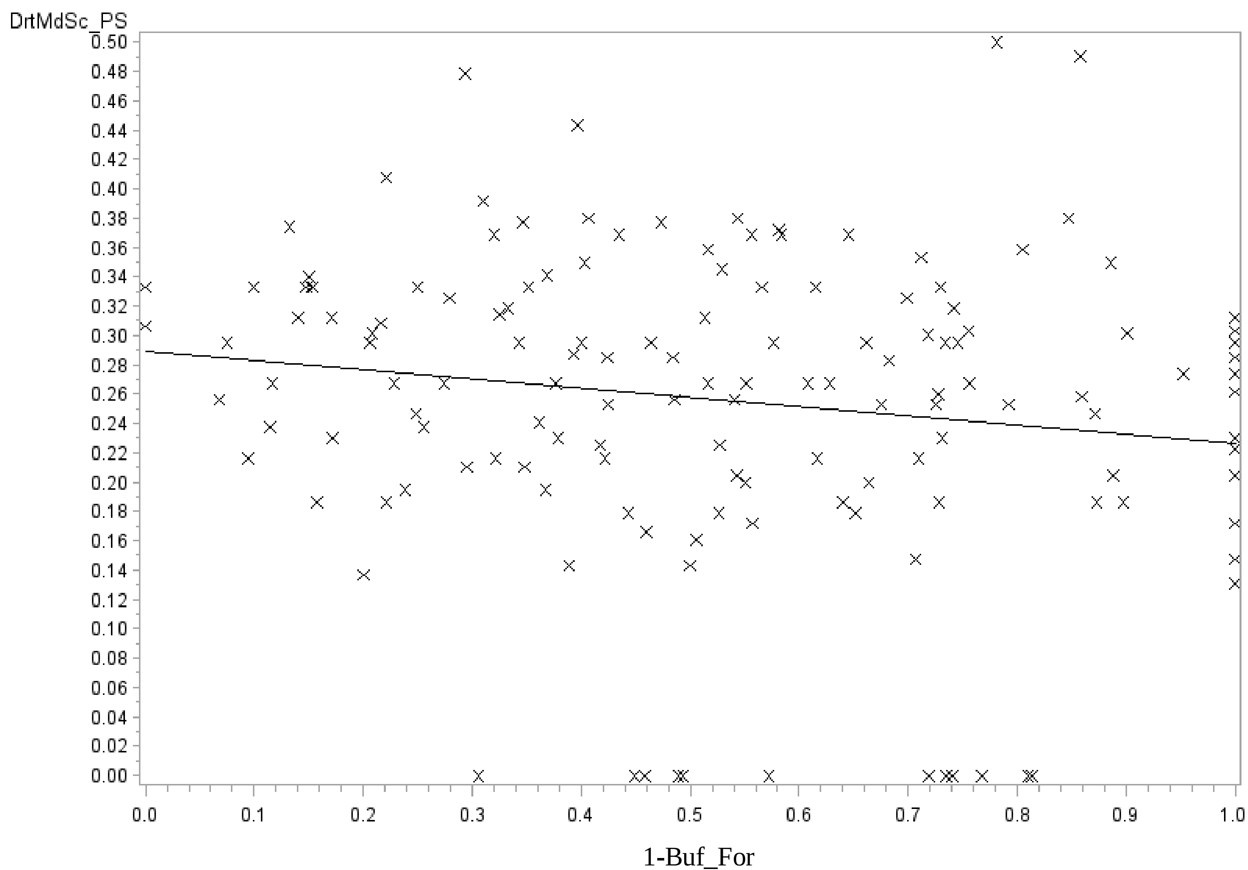


Figure 2.19 Darter, madtom, sculpin and riparian forest plot.
Plot of the relationship between proportion darter, madtom, and sculpin species (BenInv_PS) and 1 – proportion of riparian buffer forest (Buf_For). Only one line is shown for all Classes as there were no significant random effects in the best model. Both variables have been arcsin(square root) transformed and multiplied by $2/\pi$ to rescale from 0 to 1.



Summary and Conclusions

The physiographic predictors of land cover that were identified in Chapter 1 generally agree with the results of previous studies and certainly seem sensible. The most important predictor of land cover was average annual precipitation minus potential evapotranspiration and appears to primarily determine natural land cover at regional extents. Other important physiographic predictors included average annual temperature, slope, minimum elevation, relief, percent total flatland, and percent sand in soil. Agriculture was less prevalent in watersheds with high minimum elevations, steep slopes, less flatland, and sandy soils; the natural land cover types, especially forest, were positively related to these characteristics. While agriculture land use was strongly related to several of these physiographic attributes, urban land use was poorly predicted and had only marginal differences between LDPRs. When translated into a GIS and visually analyzed, the LDPR classes tend to follow regional land use patterns surprisingly well. The LDPRs predicted to have more forest are generally located in the mountainous regions of the Eastern and Western U.S. where forested land is indeed abundant, and the LDPRs with the greatest amounts of agricultural land use were located in the Midwest. Through the use of multivariate regression trees, I was able to effectively identify predictors of land cover at multiple extents and develop a regional framework that classifies watersheds throughout the U.S. based on these predictors. I recommend this method because the dendrogram output is easy to understand and translate into a GIS framework. One limitation to LDPRs that I must reiterate is that the 1 Km spatial resolution limits the utility of this framework to relatively large regions. However, higher resolution data will likely be available in the near future, which will enable the development of higher resolution LDPRs or a similar watershed classification system. Physiographically Constrained Land Use/Land Cover Regions are a unique regional framework that classifies watersheds based solely on physiographic predictors of land cover, explains how regions differ, and provides measurements of physiographic variables for each watershed.

In Chapter 2, I assessed the utility of LDPRs by identifying whether there were differences in fish assemblages and their responses to land cover alterations across these regions. Mixed linear models analysis revealed that the responses of benthic invertivores, omnivores, substrate generalists, Centrarchids, and Darter, madtom and sculpin species differed in their relationships with land cover across LDPRs. These metrics describe fish assemblages based on reproductive, trophic, and taxonomic groupings and have been commonly used in other investigations. Thus, it seems likely that the significant differences seen in these metrics reflect actual differences in the ways that fish assemblages and potentially other taxa respond to land cover alterations across regions. These results are significant because physiographic differences among regions are not often considered in the development of management strategies aimed at mitigating the impacts of agricultural and urban land use on streams. If the results shown in this study are accurate, it is apparent that one single watershed management strategy cannot be applied across all regions because the relationships between watershed and streams are different. I recommend the use of LDPRs as a stratification tool in further studies so that our understanding of the processes by which human alteration of the landscape results in stream degradation can be better understood.

References

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

Arnold, C. L., P. J. Boison, and P. C. Patton. 1982. Sawmill Brook: An example of rapid geomorphic change related to urbanization. Journal of Geology 90:155-166.

Bailey, R. G. 2004. Identifying ecoregion boundaries. Environmental Management 34:S14-S26.

Baker, M. E., D. E. Weller, and T. E. Jordan. 2006. Improved methods for quantifying potential nutrient interception by riparian buffers. Landscape Ecology 21:1327-1345.

Barker, L. S., G. K. Felton, and E. Russek-Cohen. 2006. Use of Maryland biological stream survey data to determine effects of agricultural riparian buffers on measures of biological stream health. Environmental Monitoring and Assessment 117:1-19.

Barton, D. R., and M. E. D. Farmer. 1997. The effects of conservation tillage practices on benthic invertebrate communities in headwater streams in southwestern Ontario, Canada. Environmental Pollution 96:207-215.

Berenzen, N., T. Kumke, H. K. Schulz, and R. Schulz. 2005. Macroinvertebrate community structure in agricultural streams: Impact of runoff-related pesticide contamination. Ecotoxicology and Environmental Safety 60:37-46.

Bernhardt, E. S., and M. A. Palmer. 2007. Restoring streams in an urbanizing world. Freshwater Biology 52:738-751.

Biggs, J. C. 1986. Introduction to the zoogeography of North American freshwater fishes. Pages 1-16 *in* C. H. Hocutt, and E. O. Wiley, editors. The zoogeography of North American freshwater fishes. Wiley Interscience, New York.

Brenden, T. O., L. Z. Wang, and P. W. Seelbach. 2008. A river valley segment classification of Michigan streams based on fish and physical attributes. Transactions of the American Fisheries Society 137:1621-1636.

Burgi, M., and M. G. Turner. 2002. Factors and processes shaping land cover and land cover changes along the Wisconsin River. Ecosystems 5:184-201.

Burt, T. P. 1992. The hydrology of headwater catchments. Pages 3-28 *in* P. Calow, and G. E. Petts, editors. The rivers handbook: Hydrological and ecological principles, volume 1. Blackwell Scientific Publications, Oxford.

Burton, G. W., and E. P. Odum. 1945. The distribution of stream fish in the vicinity of Mountain Lake, Virginia. Ecology 26:182-194.

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

Der, G., and B. S. Everitt. 2002. A handbook of stastical analyses using SAS, 2nd edition. Chapman & Hall/CRC, Boca Raton, FL.

Diamond, J. M., D. W. Bressler, and V. B. Serveiss. 2002. Assessing relationships between human land uses and the decline of native mussels, fish, and macroinvertebrates in the Clinch and Powell River watersheds, USA. Environmental Toxicology and Chemistry 21:1147-1155.

Dodds, W. K., and R. M. Oaks. 2007. Headwater influences on downstream water quality. Environmental Management 41:367-377.

Downes, B. J., L. A. Barmuta, P. G. Fairweather, D. P. Faith, M. J. Keough, P. S. Lake, B. D. Mapstone, and G.P. Quinn. 2002. Monitoring ecological impact: Concepts and practice in flowing waters. Cambridge University Press, Cambridge, UK.

Dunne, T., and R. D. Black. 1970. Partial area contributions to storm runoff in a small New England watershed. Water Resources Research 6:1296-1311.

Duval, A., E. Kirby, and D. Burbank. 2004. Tectonic and lithologic controls on bedrock channel profiles and processes in coastal California. Journal of Geophysical Research-Earth Surface 109:18.

EPA (U.S. Environmental Protection Agency). 1993. Regional environmental monitoring and assessment program. EPA, Report 625/R-93/012, Washington D.C.

EPA (U.S. Environmental Protection Agency). 2009. Regional environmental monitoring and assessment program: Data, Available: http://www.epa.gov/emap/remap/html/three/data/index.html. (April 3, 2006).

EPA (U.S. Environmental Protection Agency) and USGS (U.S. Geological Survey). 2009. NHDPlus User Guide. The NHDPlus Team. Available: ftp://ftp.horizon-systems.com/NHDPlus/documentation/NHDPLUS_UserGuide.pdf. (December 2009).

Everitt, B. S., and T. Hothorn. 2006. A handbook of statistical analyses using R. Chapman & Hall/CRC, Boca Raton, Florida.

Ewel, K. C., C. Cressa, R. T. Kneib, P. S. Lake, L. A. Levin, M. A. Palmer, P. Snelgrove, and D. H. Wall. 2001. Managing critical transition zones. Ecosystems 4:452-460.

Fenneman, N. M. 1928. Physiographic divisions of the United States. Annals of the Association of American Geographers 18:261-353.

Fitzpatrick, F. A., B. C. Scudder, B. N. Lenz, and D. J. Sullivan. 2001. Effects of multi-scale environmental characteristics on agricultural stream biota in eastern Wisconsin. Journal of the American Water Resources Association 37:1489-1507.

Frimpong, E., T. M. Sutton, B. A. Engel, and T. P. Simon. 2005a. Spatial-scale effects on relative importance of physical habitat predictors of stream health. Environmental Management 36:899-917.

Frimpong, E. A., and P. L. Angermeier. 2009. FishTraits: A database of ecological and life-history traits of freshwater fishes of the United States. Fisheries 34:487-495.

Frimpong, E. A., A. L. Ross-Davis, J. G. Lee, and S. R. Broussard. 2006. Biophysical and socioeconomic factors explaining the extent of forest cover on private ownerships in a midwestern (USA) agrarian landscape. Landscape Ecology 21:763-776.

Frimpong, E. A., T. M. Sutton, K. J. Lim, P. J. Hrodey, B. A. Engel, T. P. Simon, J. G. Lee, and D. C. Le Master. 2005b. Determination of optimal riparian forest buffer dimensions for stream biota-landscape association models using multimetric and multivariate responses. Canadian Journal of Fisheries and Aquatic Sciences 62:1-6.

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

Froese, R. and D. Pauly, editors. 2010. FishBase. Available: www.fishbase.org. (February, 2010).

Gilliom, R. J., Alley, W. M., and M. E. Gurtz. 1995. Design of the National Water-Quality Assesment Program: Occurrence and distribution of water quality conditions. U.S. Geological Survey, Circular 1112, Reston, Virginia.

Goldstein, R. M., L. X. Wang, T. P. Simon, and P. M. Stewart. 2002. Development of a stream habitat index for the Northern Lakes and Forests Ecoregion. North American Journal of Fisheries Management 22:452-464.

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

Harding, J. S., E. F. Benfield, P. V. Bolstad, G. S. Helfman, and E. B. D. Jones. 1998. Stream biodiversity: The ghost of land use past. Proceedings of the National Academy of Sciences of the United States of America 95:14843-14847.

Heitke, J. D., C. L. Pierce, G. T. Gelwicks, G. A. Simmons, and G. L. Siegwarth. 2004. Habitat, land use, and fish assemblage relationships in Iowa streams: Preliminary assessment in an agricultural landscape. Pages 287-303 *in* R. M. Hughes, L. Wang, and P. W. Seelbach, editors. Symposium on Influences of Landscape on Stream Habitat and Biological Communities. American Fisheries Society, Madison, WI.

Hietel, E., R. Waldhardt, and A. Otte. 2004. Analysing land-cover changes in relation to environmental variables in Hesse, Germany. Landscape Ecology 19:473-489.

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

Horton, R. E. 1933. The role of infiltration in the hydrologic cycle. Transactions of the American Geophysical Union 14:446-460.

Horwitz, R. J., T. E. Johnson, P. F. Overbeck, T. K. O'Donnell, W. C. Hession, and B. W. Sweeney. 2008. Effects of riparian vegetation and watershed urbanization on fishes in streams of the mid-Atlantic piedmont (USA). Journal of the American Water Resources Association 44:724-741.

Huryn, A. D., V. M. B. Huryn, C. J. Arbuckle, and L. Tsomides. 2002. Catchment land-use, macroinvertebrates and detritus processing in headwater streams: Taxonomic richness versus function. Freshwater Biology 47:401-415.

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

Iverson, L. R. 1988. Land-use changes in Illinois, USA: The influence of landscape attributes on current and historic land use. Landscape Ecology 2:45-61.

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

Jones, E. B. D., G. S. Helfman, J. O. Harper, and P. V. Bolstad. 1999. Effects of riparian forest removal on fish assemblages in southern Appalachian streams. Conservation Biology 13:1454-1465.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

King, R. S., M. E. Baker, D. F. Whigham, D. E. Weller, T. E. Jordan, P. F. Kazyak, and coauthors. 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. Ecological Applications 15:137-153.

Kitamura, S., T. Kadota, M. Yoshida, N. Jinno, and S. Ohta. 2000. Whole-body metabolism of the organophosphorus pesticide, fenthion, in goldfish, *Carassius auratus*. Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology 126:259-266.

Knox, J. C. 1977. Human impacts on Wisconsin stream channels. Annals of the Association of American Geographers 67:323-342.

Lee, K. E., R. M. Goldstein, and P. E. Hanson. 2001. Relation between fish communities and riparian zone conditions at two spatial scales. Journal of the American Water Resources Association 37:1465-1473.

Lee, P., C. Smyth, and S. Boutin. 2004. Quantitative review of riparian buffer width guidelines from Canada and the United States. Journal of Environmental Management 70:165-180.

Legendre, P., and L. Legendre. 1998. Numerical Ecology, 2nd edition. Elsevier, Amsterdam.

Lenat, D. R., and J. K. Crawford. 1994. Effects of land-use on water-quality and aquatic biota of 3 North Carolina piedmont streams. Hydrobiologia 294:185-199.

Long, J., and M. S. Schorr. 2005. Effects of watershed urban land use on environmental conditions and fish assemblages in Chattanooga area streams (Tennessee-Georgia). Journal of Freshwater Ecology 20:527-537.

Lussier, S. M., S. N. da Silva, M. Charpentier, J. F. Heltshe, S. M. Cormier, D. J. Klemm, M. Chintala, and S. Jayaraman. 2008. The influence of suburban land use on habitat and biotic integrity of coastal Rhode Island streams. Environmental Monitoring and Assessment 139:119-136.

Maxwell, J. R., C. J. Edwards, M. E. Jensen, S. J. Paustian, H. Parrott, and D. M. Hill. 1995. A hierarchical framework of aquatic ecological units in North America (Nearctic Zone). U.S. Department of Agriculture Forest Service, General Technical Report NC – 176, Washington D.C.

Moerke, A. H., and G. A. Lamberti. 2006. Scale-dependent influences on water quality, habitat, and fish communities in streams of the Kalamazoo River Basin, Michigan (USA). Aquatic Sciences 68:193-205.

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

Montgomery, D. R., and J. M. Buffington. 1997. Channel-reach morphology in mountain drainage basins. Geological Society of America Bulletin 109:596-611.

Naiman, R. J., H. Decamps, and M. Pollock. 1993. The role of riparian corridors in maintaining regional biodiversity. Ecological Applications 3:209-212.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Paller, M. H. 1994. Relationships between fish assemblage structure and stream order in South Carolina coastal plain streams. Transactions of the American Fisheries Society 123:150-161.

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

Petersen, R. C., B. L. Madsen, M. A. Wilzbach, C. H. D. Magadza, A. Paarlberg, A. Kullberg, and K. W. Cummins. 1987. Stream management: Emerging global similarities. Ambio 16:166-179.

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

Postel, S. L., G. C. Daily, and P. R. Ehrlich. 1996. Human appropriation of renewable fresh water. Science 271:785-788.

Pyne, M. I., R. B. Rader, and W. F. Christensen. 2007. Predicting local biological characteristics in streams: A comparison of landscape classifications. Freshwater Biology 52:1302-1321.

Quinn, J. M., P. M. Brown, W. Boyce, S. Mackay, A. Taylor, and T. Fenton. 2001. Riparian zone classification for management of stream water quality and ecosystem health. Journal of the American Water Resources Association 37:1509-1515.

Ramlal, B., and S. M. J. Baban. 2008. Developing a gis based integrated approach to flood management in Trinidad, West Indies. Journal of Environmental Management 88:1131-1140.

Ray, W., Y. Chien, and C. Chen. 1996. Comparison of probit analysis versus arcsine square root transformation on LC50 estimation. Aquacultural Engineering 15:193-207.

Reger, B., A. Otte, and R. Waldhardt. 2007. Identifying patterns of land-cover change and their physical attributes in a marginal European landscape. Landscape and Urban Planning 81:104-113.

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

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

Roy, A. H., B. J. Freeman, and M. C. Freeman. 2007. Riparian influences on stream fish assemblage structure in urbanizing streams. Landscape Ecology 22:385-402.

Roy, A. H., M. C. Freeman, B. J. Freeman, S. J. Wenger, J. L. Meyer, and W. E. Ensign. 2006. Importance of riparian forests in urban catchments contingent on sediment and hydrologic regimes. Environmental Management 37:523-539.

SAS Institute Inc. 2007. SAS/STAT user’s guide. SAS Institute Inc., Cary, NC.

Schleiger, S. L. 2000. Use of an index of biotic integrity to detect effects of land uses on stream fish communities in west-central Georgia. Transactions of the American Fisheries Society 129:1118-1133.

Schlosser, I. J. 1982. Fish community structure and function along two habitat gradients in a headwater stream. Ecological Monographs 52:395-414.

Schumm, S. A., and R. W. Lichty. 1965. Time, space, and causality in geomorphology. American Journal of Science 263:110-119.

Scott, J. B., C. R. Steward, and Q. J. Stober. 1986. Effects of urban-development on fish population-dynamics in Kelsey Creek, Washington. Transactions of the American Fisheries Society 115:555-567.

Scott, M. C., and G. S. Helfman. 2001. Native invasions, homogenization, and the mismeasure of integrity of fish assemblages. Fisheries 26:6-15.

Scott, M. C. 2006. Winners and losers among stream fishes in relation to land use legacies and urban development in the southeastern U.S. Biological Conservation 127:301-309.

Snelder, T. H., and B. J. F. Biggs. 2002. Multiscale river environment classification for water resources management. Journal of the American Water Resources Association 38:1225-1239.

Snyder, C. D., J. A. Young, R. Villella, and D. P. Lemarie. 2003. Influences of upland and riparian land use patterns on stream biotic integrity. Landscape Ecology 18:647-664.

Sponseller, R. A., E. F. Benfield, and H. M. Valett. 2001. Relationships between land use, spatial scale and stream macroinvertebrate communities. Freshwater Biology 46:1409-1424.

Sprague, L. A., and L. H. Nowell. 2008. Comparison of pesticide concentrations in streams at low flow in six metropolitan areas of the United States. Environmental Toxicology and Chemistry 27:288-298.

Stauffer, J. C., R. M. Goldstein, and R. M. Newman. 2000. Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural streams. Canadian Journal of Fisheries and Aquatic Sciences 57:307-316.

Stewart, J. S., L. Z. Wang, J. Lyons, J. A. Horwatich, and R. Bannerman. 2001. Influences of watershed, riparian-corridor, and reach-scale characteristics on aquatic biota in agricultural watersheds. Journal of the American Water Resources Association 37:1475-1487.

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

USGS (U.S. Geological Survey). 2000. National Hydrography Dataset. USGS. Available: http://nhd.usgs.gov/index.html. (December 2, 2006).

USGS (U.S. Geological Survey). 2003. Hydrologic landscape regions of the United States. USGS water resources NSDI node. Available: http://water.usgs.gov/GIS/metadata/usgswrd/XML/hlrus.xml. (April 3, 2006).

USGS (U.S. Geological Survey). 2010. Nawqa Data Export: Biological Community. USGS. Available: http://infotrek.er.usgs.gov/nawqa_queries/biomaster/index.jsp. (April 3, 2006).

Utz, R. M., R. H. Hilderbrand, and D. M. Boward. 2009. Identifying regional differences in threshold responses of aquatic invertebrates to land cover gradients. Ecological Indicators 9:556-567.

Van Sickle, J., and C. B. Johnson. 2008. Parametric distance weighting of landscape influence on streams. Landscape Ecology 23:427-438.

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

Wagner, T., D. B. Hayes, and M. T. Bremigan. 2006. Accounting for multilevel data structures in fisheries data using mixed models. Fisheries 31:180-187.

Waite, I. R., and K. D. Carpenter. 2000. Associations among fish assemblage structure and environmental variables in Willamette Basin streams, Oregon. Transactions of the American Fisheries Society 129:754-770.

Walters, D. M., D. S. Leigh, and A. B. Bearden. 2002. Urbanization, sedimentation, and the homogenization of fish assemblages in the Etowah River Basin, USA. Pages 5-10 *in* 9th International Symposium on the Interactions between Sediments and Water, Banff, Canada.

Wang, L. Z., J. Lyons, and P. Kanehl. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28:255-266.

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

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

Wang, L. Z., P. W. Seelbach, and R. M. Hughes. 2004. Introduction to landscape influences on stream habitats and biological assemblages. Pages 1-23 *in* R. M. W. L. S. P. W. Hughes, editor Symposium on Influences of Landscape on Stream Habitat and Biological Communities, Madison, WI.

Weigel, B. M., J. Lyons, P. W. Rasmussen, and L. Z. Wang. 2004. Relative influence of environmental variables at multiple spatial scales on fishes in Wisconsin's warmwater nonwadeable rivers. Pages 493-511 *in* R. M. Hughes, Wang L.Z., Seelbach P. W., editor Symposium on Influences of Landscape on Stream Habitat and Biological Communities. American Fisheries Society, Madison, WI.

Wilson, H. F., and M. A. Xenopoulos. 2008. Landscape influences on stream fish assemblages across spatial scales in a northern Great Plains ecoregion. Canadian Journal of Fisheries and Aquatic Sciences 65:245-257.

Wolock, D. M., T. C. Winter, and G. McMahon. 2004. Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. Environmental Management 34:S71-S88.

Yim, U. H., S. H. Hong, W. J. Shim, and J. R. Oh. 2005. Levels of persistent organochlorine contaminants in fish from korea and their potential health risk. Archives of Environmental Contamination and Toxicology 48:358-366.

Zhang, X. Y., X. M. Liu, Y. Z. Luo, and M. H. Zhang. 2008. Evaluation of water quality in an agricultural watershed as affected by almond pest management practices. Water Research 42:3685-3696.

Appendix A: Table of Acronyms

Table 1. Table of Acronyms

The table below gives the full name of each acronym that is commonly used throughout the text. Further description of the acronym can be found in the text where the acronym is first introduced.

Acronym	Full Name
EPA	United States Environmental Protection Agency
HLR	Hydrologic Landscape Regions
LDPR	Land cover Destinguished Physiographic Regions
NAWQA	National Water Quality Assessment program of the USGS
NHD	National Hydrography Dataset
NLCD	National Land Cover Dataset
REMAP	Regional Environmental Monitoring and Assessment Program of the EPA
USGS	United States Geological Survey