Large-scale drivers of fish biodiversity differ across an environmentally variable Great Plains watershed

by

Richard J. Lehrter II

B.S., University of Oklahoma, 2014

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Martha Mather

Copyright

Abstract

Understanding the empirical relationships between biotic diversity and components of the environment is crucial for effective research and management, particularly in highly disturbed watersheds. The Smoky Hill River is a semi-arid prairie stream with a historic native fish community that is adapted to the extreme and highly variable climatic and hydrological conditions characteristic of the Great Plains streams. Following a literature review on environmental variables, diversity responses, and analysis methods, I evaluated the importance of land use, flow, discontinuities (dams, confluences), and stream type (mainstem-tributary) variables in explaining fish richness using AIC_c model selection with multiple linear, Poisson and negative-binomial regressions. I then compared these results from 48 sites across three watershed regions to those from a long-term monitoring dataset (ST) using the same candidate variables. Finally, I examined phylogenetic patterns of the fish community using ordination analyses. Patterns and drivers of biodiversity differed with watershed region, land use, stream type, and flow. Fish species richness in the Smoky Hill watershed was negatively correlated with percent developed land in the Lower region of the watershed, but positively correlated with percent herbaceous grassland, the reference prairie condition, in the Upper region of the watershed. Summer mean flow was consistently and positively related to species richness in the Middle and Upper regions of the watershed where flow was limited. In the Lower region of the watershed, species richness was higher in the more flow-moderate tributaries relative to high-flow mainstem sites. In the Middle and Upper flow-limited regions, species richness was lower in the low-flow tributaries than main stem sites. Families of fish species were also related to region and stream type (mainstem vs. tributary). A comparison of two databases showed how

different goals, questions, and methods result in different insights, emphasizing the need for establishing *a priori* goals before sampling.

Table of Contents

List of Figures

List of Tables

Chapter 1 - Large-Scale Drivers of Fish Biodiversity Differ Across an Environmentally Variable Great Plains Watershed

Introduction

Understanding the empirical relationships between biotic diversity and environmental correlates is crucial for effective research and management, particularly in highly disturbed watersheds. Specifically, determining which environmental factors alter biodiversity across a watershed is both important and complex.

While components of rivers have long been shown to follow a natural pattern from headwaters to mouth (Vannote et al. 1980, Davies and Walker 2013), spatial and temporal discontinuities alter this gradient (Poole 2002, Burchsted et al. 2014). Natural discontinuities such as confluences can be significant disruptions in the river continuum, altering localized abiotic components of streams such as depth, width, sediment quality, and water quality (Rice et al. 2008). Recent studies have noticed an alteration of fish species assemblages in close proximity to confluences (Thornbrugh and Gido 2009; Czegledi et al. 2016). Relative to large rivers, tributaries often differ in their species richness and composition due to differences in water discharge, as has been shown using species-discharge relationships (McGarvey and Hughes 2008).

Additionally, impacts to watersheds can often be linked to anthropogenic effects. Humans drastically alter landscapes, and the alteration of land use and composition is considered to be a major environmental issue (Foley et al. 2005). Nutrient-rich runoff from agriculture and livestock has led to decreases in water quality through eutrophication in many reservoirs and other surface waters in the central United States (Carpenter et al. 1998). Developed land is often accompanied by impervious surfaces and the physical alteration of streams, which can increase runoff, decrease in-stream habitat diversity, and negatively impact characteristics of water flow and subsequently localized fish diversity (Jungwirth et al. 1993; Brabec et al. 2002).

Water quantity is greatly impacted by human use through damming and withdrawal for municipalities and agriculture. Water withdrawal for human use generally reduces flow rates, which help to shape important characteristics of aquatic habitat including water quality, substrate, and physical habitat features (Matthews 1998). Additionally, global climate change predicts altered flow regimes and water availability, which may cause an even greater strain on aquatic biodiversity in particularly water-limited systems. Understanding which natural and anthropogenic components of the environment impact stream ecosystems across a watershed is therefore important for sustaining water security for freshwater biodiversity (Vörösmarty et al. 2010).

To determine appropriate methods of environmental and biotic variable quantification as well as commonly used data analyses, I first conducted a standardized literature review using relevant key terms. To test relationships among fish species richness and the environmental variables identified in the literature review, I addressed four research issues by sampling 48 sites across the Smoky Hill watershed. First, I evaluated the importance of land use, flow, discontinuities (dams and confluences), and physical variables (mainstem vs tributary) for three *a priori* selected regions of the watershed (Upper, Middle, and Lower) using AIC selection on multiple linear regression models. Second, I compared results of the above-described variable selection to that which resulted from Poisson and negative binomial models. I refer to this as the KSU dataset. Third, I repeated this variable selection process on a second Smoky Hill long-term fish monitoring dataset collected by the Kansas Department of Wildlife Parks and Tourism (KDWPT). I refer to this as the Stream Team (ST) dataset. I then compared the results of the

KSU and ST analyses. Lastly, I used CCA analyses on the KSU dataset in order to determine whether there were any consistent differences in fish species and families between mainstem and tributary sites.

Methods

Literature Review

Prior to collecting and analyzing data, I wanted to first perform a preliminary screening of similar studies in order to develop appropriate methods of variable quantification and analysis. To identify which environmental variables, responses, and statistical analyses have been examined, I conducted a Web of Science search using the keywords (1) “fish” and (2) “correlates” or “predict” or “driver,” and (3) “ecology” or “environment,” and (4) “stream” or “river” or “lotic” for 14 journals in four categories: (1) *Interdisciplinary* (Science of the Total Environment, Ecological Indicators, PLoS One), (2) *Ecology* (Freshwater Biology, Ecology of Freshwater Fish, Canadian Journal of Fisheries and Aquatic Sciences, Journal of Animal Ecology, Annual Review of Ecology and Systematics), (3) *Ecological Applications* (Biological Conservation, Ecological Applications, River Research and Applications, Aquatic Conservation-Marine and Freshwater Ecosystems), and (4**)** *Global Change* (Global Ecology and Biogeography, Global Change Biology).

Environmental variables were examined in all 60 peer-reviewed papers plus nine highly cited reviews. Each paper was categorized as being either a (1) watershed study or a (2) non-watershed study on the basis of whether the study attempted to obtain a representative sample of (1) the majority of streams located within a specific drainage area or (2) one or only a few streams within a specific drainage area. Land use / land cover and flow dynamics were the dominant environmental variables linked to fish biodiversity in both watershed (Fig. 1A) and non-watershed studies (Fig. 1B). However, the importance of these variables differed between studies that took a watershed and a non-watershed approach. For watershed-approach studies, the most common environmental variables examined were land use (35% of all studies) and flow (32% of studies) (Fig. 1A). In watershed studies, other environmental variables (e.g. temperature, elevation, water chemistry, river kilometer, gradient, and discontinuities) were examined in < 10% of all studies. In non-watershed studies, flow was the most common variable linked to fish biodiversity (32% of all studies), temperature was a focus in 32% of the studies, land use was a focus in 19% of all studies, and other variables were examined in less than 12% of this literature (Fig. 1B).

Species richness (number of species) and guilds (functional richness) were the two most common biotic response metrics in both watershed (Fig. 2A) and non-watershed studies (Fig. 2B). These two responses were used in approximately one-third of all studies reviewed. Abundance was more common in non-watershed studies (26%; Fig. 2B) than in watershed studies (11%; Fig. 2A). Multi-species diversity indices were more common in watershed studies (15%; Fig. 2A) than non-watershed studies (4%; Fig. 2B). Less common biotic metrics included trophic level/isotopes, life history characteristics, and diversity indices (<10%; Fig. 2).

The most common analytical method across all studies was multiple regression, used in 75% of the watershed scale papers (Fig. 3A) and 52% of the non-watershed studies (Fig. 3A). Ordination was used more often in watershed studies (39%; Fig. 3A) than in non-watershed studies (13%; Fig. 3B). Less common analytical tools included categorical and regression trees (10-25% of all studies that linked environmental variables to fish biodiversity).

Flow was the most common environmental variable linked to biotic diversity in the literature; in all 14 studies metrics of flow magnitude and variation had a positive relationship with biotic diversity (Fig. 4; e.g. He et al. 2010, Espinoza et al. 2013, Bond et al. 2015). Stream size including stream type (mainstem vs. tributary) was the second most common variable that affected fish diversity (e.g. Osborne and Wiley 1992, Meyer et al. 2007, Radinger et al. 2015). Of the seven papers that found a relationship between biotic diversity and land use, four found a negative relationship between diversity and developed land (e.g. Lammert and Allan 1999, Doledec et al. 2006, Esselman and Allen 2010, Macedo et al. 2014), two found a negative relationship between diversity and cultivated land, and one found a positive relationship with grassland (Fig. 4). Temperature was linked to fish diversity in seven studies (e.g. Crozier 2010, Braun and Reynolds 2014, Alexiades and Fischer 2015, Letcher et al. 2015), but all of these studies examined abundance and reproduction in temperature-dependent salmonid fishes (Fig. 4). Four papers found relationships between biotic diversity and a spatial environmental gradient (e.g. González-Bergonzoni et al. 2012, Rypel 2012) while only two inferred a relationship between biota and connectivity (Crook et al. 2010). Based on these results from the literature, I prioritized the following as important environmental variables: region (longitudinal area), flow, land use (developed land and grassland), type (mainstem and tributary to reflect physical conditions and size), dams, and confluences.

Study System

General location

The Smoky Hill River watershed is a 900-km long semi-arid prairie stream located in the North American Great Plains eco-region (Summerfelt 1967) that originates in the High Plains of

Eastern Colorado and flows eastward through Kansas (Fig. 5A, B). The headwaters of the river are composed of two streams, the Smoky Hill River and the North Fork Smoky Hill River, which begin in Cheyenne County, Colorado and join together in Logan County, Kansas. From Logan County the river flows through central Kansas to Russell County where it joins the confluence of Big Creek, then to Salina where it joins the Saline and Solomon Rivers, and finally further east to Junction City where it joins the Republican River to form the Kansas River (Fig. 6A, B). The basin drains an area of approximately 50,000 square kilometers (Summerfelt 1967). My study area (650-km long, approximately 30,000 km^2 drainage area) extends from Cedar Bluff Reservoir and Big Creek to the confluence of the Smoky Hill and Republican Rivers (Fig. 6A).

Regions

I divided my study area within the Smoky Hill River Basin (Fig. 6A) into three distinct geographic regions [Upper (U), Middle (M), and Lower (L)] of approximately equal lengths (U=97 km, M=97 km, U=106 km) that are separated by distinct landmarks (Fig. 6B). The Upper region is between Cedar Bluff reservoir and the confluence of the Smoky Hill River and Big Creek (Fig. 6B). The Middle region is between the Smoky Hill-Big Creek confluence and Kanopolis Reservoir (Fig. 6B). The Lower region extends from below Kanopolis Reservoir to the confluence of the Smoky Hill and Republican Rivers (Fig. 6B).

Land use

Agricultural production in the Smoky Hill watershed consists primarily of winter wheat, sorghum, and corn, which are frequently irrigated in the western part of the watershed (Gao et al. 2017). Irrigated land use removes water directly from the stream as well as the underlying aquifer, and can be accompanied by stream fragmentation and hydrologic alteration by dams, all of which have been negatively associated with native fish presence and diversity (Xenopolous et al. 2005; Perkin et al. 2014; Perkin et al. 2017). Rangeland (labeled on Fig. 7A as “other”) is composed of a mix of native and non-native herbaceous grassland, is most abundant in the Upper region of my Smoky Hill study area, and is frequently used for cattle grazing. In addition to altering vegetation, cattle grazing can reduce flow via water diversion and use in stock ponds (Belsky et al. 1999). Historical land cover was predominantly mixed-grass prairie that has been gradually converted to agriculture or is grazed by cattle (Matthews 1988). Human population density in the majority of the Smoky Hill basin averages <5 people per square kilometer with the exception of several moderately large towns including Hays, Ellsworth, and Salina (Dykstra 1983; Fig. 7B). Municipalities are the main source of developed land in the Smoky Hill basin (NLCD 2011). The majority of developed land occurs in the Lower region of our Smoky Hill study area with additional significant development near Hays in the Upper region of the watershed. Urbanization often reduces the quantity of water, affects water quality through pollution, and physically alters instream and riparian habitats (Jurajda 1995, Winter et al. 1998, Cumming 2004).

Flow

The Smoky Hill River flows along a strong precipitation gradient with annual precipitation in the headwaters of the basin averaging roughly half of the amount of rainfall received in the farthest eastern (downstream) portion of the watershed (Gunal and Ransom 2006; Fig. 7C). Thus, precipitation is increasingly limited from downstream to upstream which can make the role of instream flow more important in the Middle and Upper regions. Instream water discharge varies significantly between the three regions of the Smoky Hill basin. Mean annual discharge in the Lower region of the Smoky Hill River is 95.4 m3 s-1, with an observed minimum of 1.84 m3 s-1 and observed maximum of 648.46 m3 s-1 (Fig. 8A- USGS 06877600 at

Enterprise, KS). Mean annual discharge in the Middle region of the Smoky Hill River is 13.37 m3 s-1, with an observed minimum of 0.17 m3 s-1 and observed maximum of 365.29 m3 s-1 (Fig. 8B- USGS 06864500 at Ellsworth, KS). Mean annual discharge in the Upper Smoky Hill Basin is 4.3 m3 s-1 with an observed minimum of 0 m3 s-1 and an observed maximum of 202 m3 s-1 (Fig. 8C- USGS 06861000 at Arnold, KS). The Upper Smoky Hill River has significantly lower flow than the other regions making water limitation more severe, and stream desiccation more likely in this western region. Water is also limited in the Middle region but summer stream desiccation is often restricted to tributaries. Flow is rarely limiting in the Lower region of the Smoky Hill River, where lower flow tributaries can provide more heterogeneous fish habitat than the higher flow in the mainstem (Fig. 6B).

Dams

Two Corps of Engineers (CoE) reservoirs (Cedar Bluff Reservoir and Kanopolis Reservoir) constructed in the late 1940s impound the Smoky Hill River. Both reservoirs were constructed to serve as sources of flood control and water conservation (Sowards et al. 2016; Fig. 6B). Cedar Bluff was also constructed to serve as a source of water for irrigation in response to severe drought in Western Kansas during the 1930s. Both serve secondary purposes as sources of recreation and contain similar populations of fish species including Crappie (*Pomoxis spp.*), Catfish (*Ictalurus spp.*), Black Bass (*Micropterus spp.*), True Bass (*Morone americana & Morone americana x saxatilis*), Walleye (*Stizostedion vitreum),* and Gizzard Shad (*Dorosoma cepedianum*). These CoE reservoirs are likely to affect the three regions differently. The Upper region includes Cedar Bluff reservoir, which may significantly impact fish by habitat and flow alteration and could add reservoir fish to the Smoky Hill watershed downstream of the reservoir. The Middle and Lower regions are divided by Kanopolis reservoir. Consequently, the Upper and

Middle region sites will likely be most affected by distance to the upstream Cedar Bluff Reservoir, while the Lower region sites will be most affected by distance to the upstream Kanopolis Reservoir. Since there is no reservoir downstream of our Lower region study area, the Lower region will not be affected by any downstream CoE dam. The Lower region is farthest from a CoE dam, and CoE dam effects will likely be least for this region. The majority of intact low-head dams in the Smoky Hill Basin occur relatively high on tributaries (National Inventory of Dams). Although low-head dams are important elsewhere for other sustainability questions, none of the sites that I sampled were adjacent to low head dams.

Site type: mainstem vs. tributaries

The majority of the mainstem Smoky Hill River is a large, sand-bed river. Tributaries feeding into it are higher gradient streams that may have intermittent flow (especially in the Middle and Upper region of the watershed) even during years of normal precipitation (Dodds et al. 2004). All three sampled regions of the basin contain both mainstem and tributary sites, but their relative conditions vary. In the Lower region, mainstem sites are considerably larger (7th order stream) with greater discharge relative to the Middle (6th order stream) and Upper (5th and 6th order stream) regions. First and second order tributaries in the Middle region often have low or intermittent flow. Tributaries in the Upper region have irregular flow and are intermittent with isolated pools, especially in the later summer. Thus, the aquatic habitat suitability of mainstem and tributary sites varies across regions.

Confluences

The size and complexity of confluences, or river junctions, varies throughout the basin and is highly dependent on stream order and magnitude of flow. Confluences serve as sources of physical and biological heterogeneity as well as increased productivity in lotic systems (Rice et al. 2006, Kiffney et al. 2006). Throughout the watershed and in all regions, confluences exist where numerous tributaries join each other and merge into the mainstem. Hence, confluences are a ubiquitous discontinuity in my study area. What differs across regions is the size of the streams that merge, which is dependent on stream order and magnitude of flow.

Summary of regional conditions

In summary, the Lower region has the greatest amount of developed land, has the least amount of grassland, has the most precipitation, has the highest flow (especially in the mainstem), has tributary sites with intermediate physical conditions compared to the high-flow habitat in the mainstem, is only affected by the upstream Kanopolis CoE dam, and has a variety of confluences. The Middle region has an intermediate to low amount of developed land, has limited but variable amounts of grassland, has an intermediate level of precipitation, has limited instream flow especially in the summer, has tributary sites with low flow and variable physical conditions compared to the more stable mainstem sites, is affected by both Kanopolis and Cedar Bluff CoE dams, and has a variety of confluences. The Upper region has the lowest amount of developed land, has a substantial amount of grassland, has the lowest level of precipitation, has very limited instream flow, has tributary sites with no or intermittent flow that are physically harsh compared to mainstem conditions, is only affected by the Cedar Bluff CoE dam, and has a variety of confluences.

Research Design

To capture and quantify watershed-scale fish biodiversity in the Smoky Hill River that may be linked to human impacts, I collected environmental and fish community data at 48 sample sites within the three regions of the Smoky Hill River (Fig. 6B). The sample sites were chosen using several criteria. First, many of my sample sites (hereafter called KSU sites and samples) were near sites that were previously sampled for fish by the Kansas Department of Wildlife, Parks and Tourism Stream Team long-term monitoring program (hereafter called ST sites and samples). Second, most KSU mainstem sample sites were paired with two sites on an adjacent tributary- a lower and upper tributary site. Third, all geographic areas of the river were sampled such that, in general, KSU samples were at regular intervals along the river (~15-30 km between mainstem samples). Lastly, near equal sample sites were selected in each of the three river regions (Upper N=17, Middle N=20, Lower N=11). Because I was interested in linking biodiversity to a full range of natural and anthropogenic conditions throughout the watershed (not just differences in existing fish community), the above-described, predetermined KSU sites were always included in the data analysis regardless of fish presence or absence.

KSU Fish Sampling

I used standard fish sampling methods at all 48 study sites between the months of June and September, 2015. At each of the above-described 48 study sites in the Smoky Hill watershed, fish were sampled at three 100-meter transects that I combined to estimate species richness and abundance for a 300-meter longitudinal section of stream. To make samples at all sites comparable throughout the watershed, I chose to sample the dominant habitat: mid-channel pool. At several sites where there were isolated pools, three 100-meter transects of stream channel were sampled beginning at the first substantial pool at that site. All sites were sampled with a crew of 2 to 3 people.

To sample fish, I used a mini-Missouri trawl (Herzog et al. 2009) as a hand trawl (without trawl doors). The Mini-Missouri trawl is constructed with lead weights on the bottom of the net that allow it to remain on the streambed at all times. In this way, it functions similarly to a beach seine. The mini-Missouri trawl that I used was a 2.4-meter wide net with 35 mm outer mesh and 3.2 mm inner mesh. Operators of the mini-Missouri hand trawl can adjust the width (wider or narrower) to fit ambient and changing stream size. Like a beach seine, the mini-Missouri trawl does not sample heterogeneous bottoms well (logjams and brush piles). However, these bottom structures were rarely present or encountered in the Smoky Hill River and were not part of the dominant mid-channel pool habitat. Unlike a seine, the double bag of the mini-Missouri kept rocks separate from fish and thus limits fish tissue damage and mortality. The mini-Missouri trawl is a commonly used fish sampling gear, especially in larger rivers and lentic habitats (Phelps et al. 2010, Fischer and Quist 2014). Characteristics and performance of the mini-Missouri trawl, fished as a hand trawl in similar-sized, geographically adjacent rivers in Kansas, were evaluated elsewhere (Appendix 1). In that quantitative gear assessment, this gear fished as I used it here) captured as many fish species as other common gear in mid channel habitats.

In summary, the hand-pulled mini-Missouri trawl was chosen because it provided a standardized sample across all sites, depths, and substrates in the most common stream habitat (mid-channel pool) throughout the watershed. The mini-Missouri trawl did minimal damage to fish, effectively sampled the most common habitat, always remained on the bottom, and could be adjusted to any stream width as needed. These strengths were especially important for this study in which a primary goal was to compare fish across sites in three regions of the Smoky Hill watershed. After sampling, all fish were identified to species and returned to the river unharmed.

Environmental Correlate Data Collection

Choice of environmental variables

Based on the literature review from the introduction, I targeted data collection from four categories of environmental variables at the 48 sites at which I sampled fish species richness. These categories of environmental variables included (1) Physical characteristics: river kilometer, region, elevation, stream order, type (mainstem/tributary); (2) Discontinuities: distance to confluence, distance to upstream CoE dams, distance to downstream CoE dams; distance to any CoE dam; (3) Land use in basin: developed land (%); forest (%), grassland (%), cultivated (%), rangeland (%); and (4) Flow: mean, minimum, maximum (overall, spring, summer), standard error, number of high and low peaks, reversals, number of zero flow days, 30-day minimum flow, and 30-day maximum flow, calculated using IHA (Indicators of Hydrologic Alteration; Richter et al. 1996) on SWAT (Soil and Water Assessment Tool) and gage flow measurements.

From these data, based on exploratory analysis (correlations among variables), I narrowed my candidate environmental variables to seven core variables that I used throughout the analysis: region, developed land (%), herbaceous grassland (%), mean summer flow (m3 s-1), distance to upstream CoE dam (km), type (mainstem/tributary), and distance to nearest 2nd order confluence. This method of candidate variable selection provided us with variables that represent broad classes of potential drivers. How I calculated the above metrics for my 48 study sites is described below.

Land use

Percent land use was calculated for all KSU study sites using the NLCD 2011 land cover database in ArcMAP version 10.3.1. Land cover was summarized for a 0.5 km radius circular buffer around each site (Fig. 9A).

Flow

I determined flow at each of my 48 sites using one of three sources of data: (1) measured data from a USGS gage, (2) modeled data from a SWAT site, or (3) bootstrapped data from a combination of USGS gages and SWAT sites. For all three sources of flow data, at all sites, daily 15-year flow records were obtained. Daily 15-year flow data were processed through the IHA program to obtain flow metrics. Summer mean flow (June-August) was used in all analyses because this time period is when flow is the most limited. Relative to the first source of flow data, 7 of the 48 sites were <10 km in proximity to a USGS gage with no major inflows between the gage and sample site. In this first source, my sample sites used flow measurements from associated gages. Relative to the second source of flow data, 10 of the 48 sites were <10 km in proximity to a SWAT modeled flow site with no major inflows between the SWAT location and the fish sample site. In this second source of flow data, my sample sites were paired with flow estimates from associated gages.

The remaining 31 sites were not close to a unique gage or SWAT site and required flow estimates from a third source. For these, I combined SWAT and gage flow estimates for region (Upper, Middle or Lower) and stream type (mainstem or tributary) to create six scenarios- Upper Mainstem, Upper Tributary, Middle Mainstem, Middle Tributary, Lower Mainstem, and Lower Tributary. For these, I calculated mean and standard deviation of yearly flow from SWAT and USGS gage estimates for each scenario, created a normal distribution, then randomly sampled

100 times to impute unique flow estimates to each site within this third that had no unique adjacent gage or SWAT site.

I used this flow estimation protocol for the following reasons. First, I judged that multiple temporal estimates across days, months, and years was essential. Hence, limited empirical estimates of velocity, as are often used in local studies, were not adequate. Second, USGS gages or unique SWAT flow estimates were not available adjacent to all fish sample sites. Third, sites used for estimates were selectively matched to similar types and sizes of streams that were nearby, an approach that is a component of much hydrological modeling. Finally, temporally detailed, site-specific data were not available for these sites through other methods.

CoE dams

The distance to upstream Corps of Engineers (CoE) Dams in the Smoky Hill watershed was calculated for all sample sites. This was done with the measure tool in ArcGIS using a digital shapefile obtained from KDWPT for all Corps of Engineers reservoirs in the state of Kansas (Fig. 9C). When measuring distance to dam for a sample site upstream of a CoE dam, the distance was measured from each site to the river inlet of the main part of the lake. If there was no dam upstream of a given sample site, the value for that variable was set at the empirical observed maximum distance from a dam among all sample sites. Although the effects of upstream and downstream dams are ecologically different, preliminary analyses indicated that distance to upstream or downstream dam did not have significant effects on species richness with our dataset and at the scale of our analysis. However, because we wanted to include a distance to dam variable in our variable selection analysis and because we had more sample sites located immediately below a reservoir dam than we did immediately above a reservoir, we chose to include only distance to upstream dam in our model selection analyses. For parsimony, I chose distance to upstream dam for use in further analyses.

Confluences

Distance to nearest 2nd order or larger confluence was calculated using ArcGIS 10.3.1. Distance to first order streams was not included in confluence measurements due to their frequent intermittency in the Smoky Hill Basin. For my calculations, first, the drainage basin was delineated using a Digital Elevation Model (DEM) acquired from the State of Kansas Data Access and Support Center (DASC). The DEM file provides coverage of Kansas in 7.5 by 7.5 minute blocks. SWAT (Soil and Water Assessment Tool) was used to determine the locations of all streams within the Smoky Hill River watershed using the DEM data. Once this was done, stream order was calculated for all streams within the Smoky Hill River watershed. Using the measure tool, the distance from each fish sample site to 2nd order or larger streams both upstream and downstream was calculated (Fig. 9B).

Regression Analysis Methods

I used information theoretic model selection on multiple linear regression (MLR) for small datasets (AICc) to determine which environmental variables [developed land, herbaceous grassland, summer mean flow, distance to upstream CoE dam, distance to confluence, and site type (mainstream or tributary)] influenced species richness in each of the three regions of the Smoky Hill watershed. Limiting the number and types of variables to those that test meaningful ecological concepts is crucial for maintaining parsimony, as model selection with too many parameters can lead to over-fitting (Burnham and Anderson 2004). Although multiple regression can be used for many purposes, here, my primary goal was variable selection (i.e., to identify variables that were most influential in biotic patterns).

Starting with an a priori analysis of the variables recognized to be important for biodiversity in lotic systems in the ecological literature, I ran a set of analyses to empirically identify variable importance for species diversity in the three regions of the Smoky Hill watershed. In each analysis the dataset was run as a whole and by region (Upper, Middle, Lower). Species richness was log transformed in all analyses in order to meet the assumptions of normality and homogeneity of variance. All models were tested for normality and homogeneity of variance using residual and QQ plots and met these assumptions.

Regressors included in the AICc analysis included proportion of developed land (of total land use), proportion of grassland (of total land use), 15-year summer mean flow (m3 s-1), distance to upstream Corps of Engineers (CoE) dam (km), stream type (mainstem or tributary), and distance to 2nd order confluence (km). Stream type was categorized for each site and coded as a factor in MLR analyses (mainstem = 0, tributary = 1). Candidate variables were standardized for ease of comparison between variables of different units. For each analysis, multiple regression models were compared using AICc, a model selection tool for small sample sizes (Burnham and Anderson 2002) using the MuMIn package in program R. Prior to model selection, variance inflation factors (VIF; Function vif in Program R; Core Team 2014) were calculated in order to detect multicollinearity. Variables with VIF >10 were discarded. Models within Δ AICc <4 were retained as top models. Although Δ AICc <2 are also shown, I chose the larger range (Δ AICc <4) to capture the full range of information in the dataset.

For these top models, I calculated model averaged regression coefficients (β), standard errors (SE), and variable importance (MuMIn in program R). A high variable importance value shows a higher relative importance of that variable in the model selection process and thus a higher probability that the variable is in the best model (Burnham and Anderson 2002). A larger slope indicates a stronger relationship between the regressor and fish species richness. Model weights and adjusted R2 (Adj. R2) were also calculated.

To test whether the type of statistical models changed these results, I also ran the same AICc model selection protocol, described above, using Poisson and negative-binomial distributions. These generalized linear models were run in Program R using the MASS package version 7.3-47.

Fish Community Analysis Methods

A Canonical Correspondence Analysis (CCA) was performed on the log-transformed fish abundance data set as a whole and by region (Program R using the vegan package version 2.4-3) in order to capture potential relationships between fish community response and environmental variables. The same variables used in our regression analyses were used in the CCAs. Sample sites and fish species were included in the ordinations. Each fish species detected was colored by phylogenetic order. Details of fish species distribution is discussed elsewhere (Appendix 2).

AIC$_c$ Analysis- ST Dataset

I ran the same multiple linear regression analysis, described above, on a fish community dataset collected by the Kansas Department of Wildlife, Parks and Tourism (KDWPT) Stream Team from 1996-2013 as part of a long- term fish monitoring program. The Stream Team collected 33 useable samples in the Smoky Hill River between 1996 and 2013 using a combination of seining and electrofishing. For the locations of these samples, I calculated the same environmental variables as described above with one exception: land use calculations were matched to the year of collection with the temporally most similar NLCD raster file (i.e., NLCD land cover data exists only for the years 1992, 2001, 2006, and 2011). I followed AICc analysis with scatterplots and histograms of environmental and fish variables (including sample quartiles for environmental variables) to compare sample site selection and results for KSU and ST datasets.

Results

Overview

Below, I review several types of results. First, I summarize the relationships between all explanatory variables and fish species richness in 1) the full KSU dataset, and 2) the KSU by region datasets (Q1). I then summarize the results of Poisson and negative-binomial models for the same datasets (Q2). Third, I provide a summary of the individual fish species and fish guilds patterns using a canonical correspondence analysis ordination on the KSU fish abundance dataset (Q3). Finally, I provide the results of a multiple linear regression analysis with $\mathrm{AIC_c}$ selection for the ST dataset followed by plots to help identify why differences existed between the KSU and ST datasets (Q4).

Variable Selection Using Multiple Linear Regression- KSU Dataset (Q1)

Full dataset

The three most important variables for species richness in the full KSU dataset multiple linear regression AICc were developed land (%), summer mean flow (m3 s-1), and stream type (mainstem vs tributary) (Table 1). Developed land had a variable importance of 1.0, meaning it was included in all models with Δ AICc <2. Developed land was included (< 0.05) in four of the seven top models and this variable had a model averaged slope of -0.31, indicating a decrease in species richness with increasing developed land. Summer mean flow also had a variable importance of 1.0 (Table 1) and was significant in six of seven top models with a Δ AICc <4. This variable had a model averaged slope of 0.36, indicating an increase in species richness with an increase in summer mean flow. Stream type (mainstem or tributary) was in all top models and had a variable importance of 1.0. Stream type had a model averaged slope of -0.90, indicating higher species richness in the mainstem Smoky Hill River relative to its tributaries. All top models had highly significant p-values.

Lower region

In the Lower region of the Smoky Hill River, two regressors were present in the three top models (Table 2A; Lower Region models #1-3; Δ AICc 0-2.92; Adj. R2 0.33-0.52; p <0.05). All models had p-values <0.05 (Table 2A). Specifically, percent developed land and stream type (mainstem vs tributary) were present in two top models, respectively. Developed land had a variable importance of 0.82, and was significant in one top model. Developed land had a Δ AICc <4 model averaged slope of -0.28, indicating that this environmental variable was negatively associated with species richness (Fig. 10A). Stream type had a variable importance of 0.34, was in one top model with a Δ AICc <4, and had a model averaged slope of 0.39, indicating that tributaries had higher species richness relative to mainstem sites in this region (Fig. 11D).

Middle region

In the Middle region of the Smoky Hill River, six regressors were present in the seven top models (Table 2B; Middle Region - models #1-7; Δ AICc 0-3.69; Adj R2 0.72-0.76; p <0.00005). All models had highly significant p-values (p <0.00005). Percent developed land was in two top models but the slope of this variable was not significant (Fig. 10B). Herbaceous grassland was present in two models, but the slope of this variable also was not significant (Fig. 10E). Summer mean flow was present and significant in all seven top models with variable importance of 1.0. Summer flow had a Δ AICc <4 model averaged slope of 0.54 (Table 2B; Fig. 10H), indicating that species richness increased with summer mean flow. Distance to upstream dam and distance to confluence were also in top models but were not significant, indicating that they did not significantly explain the variation in species richness in this region (Fig. 11B & 11H). Stream type was present and significant in all seven top models, and had a variable importance of 1.0. Stream type had a Δ AICc <4 model averaged slope of -1.27, indicating higher richness in the mainstem sites relative to tributaries in this region (Table 2B, Fig. 11E).

Upper region

In the Upper region of the Smoky Hill River, five regressors were present in the four top models (Table 2C; Upper Region models 1-4; Δ AIC$_c$ 0-3.31; Adj R^2 0.70-0.78; p <0.0005). All models were highly significant with *p*-values <0.0005 (Table 2C). Percent developed land was in one top model but was not significant (Table 2C). Herbaceous grassland was significant in one top model with a variable importance of 0.36. Herbaceous grassland had a Δ AIC$_c$ <4 model averaged slope of 0.24, indicating that fish species richness increased with proportion of grassland (Table 2C, Fig. 10F). Summer mean flow was significant in all four top models with a variable importance of 1.0, and a Δ AIC$_c$ <4 model averaged slope of 0.41, indicating that species richness increased with summer mean flow (Fig. 10J). Stream type was also significant in all four top models with a variable importance of 1.0. A Δ AIC$_c$ <4 model averaged slope of -1.56 indicated that higher species richness occurred in mainstem sites compared to tributaries in this region (Fig. 11F). Distance to confluence was in three of four top models with a model averaged slope of 0.38 and variable importance of 0.90 (Table 2C). The sites that were farthest from a tributary were also the sites highest in the watershed (Fig. 11J).

Variable Selection Using Poisson and Negative Binomial Models- KSU Dataset (Q2)

Variables identified in multiple linear regressions for both full and by region datasets were the same as those identified by Poisson and negative-binomial models (Table 3 & 4).

CCA Ordinations (Q3)

While there is some interspersion of mainstem and tributary sites, two distinct groupings of mainstem and tributary sites occurred in ordination space (Fig. 12). Several species were grouped around mainstem in the ordination space. These include several large bodied species such as freshwater drum (*Aplodinotus grunniens*), channel catfish (*Ictalurus punctatus*), river carpsucker (*Carpiodes carpio*), but also several smaller bodied species that were predominantly caught at mainstem sites such as northern plains killifish (*Fundulus kansae*) and western mosquitofish (*Gambusia affinis*). Another group of fishes was also associated distinctly with tributaries in the ordination space. This included some smaller bodied fish species such as orangethroat darter (*Etheostoma spectabile*), central stoneroller (*Campostoma anomalum*), bluntnose minnow (*Pimephales notatus*), and members of the order perciformes including several sunfish species (*Lepomis spp.*). The rest of the species were grouped around 0, indicating no particular relationship with either mainstem or tributary sites (Fig. 12).

CCA trends, by region, generally reflected similar patterns (Fig. 13). Lower Smoky Hill River sites were grouped around the mainstem factor with several small-bodied fish species including red shiner (*Cyprinella lutrensis*), sand shiner (*Notropis stramineus*), western mosquitofish, and northern plains killifish. The tributary factor in the Lower region was grouped with a number perciformes fishes, mainly centrarchid species including longear sunfish (*Lepomis megalotis*), green sunfish (*Lepomis cyanellus*), orangespotted sunfish (*Lepomis humilis*), and bluegill (*Lepomis macrochirus*) as well as the orangethroat darter. The central stoneroller was also primarily associated with tributary sites.

The Middle region CCA showed distinct groupings of mainstem and tributary sites (Fig. 14). Fish species that grouped around the mainstem in the ordination space included northern plains killifish as well as several larger bodied fish species such as freshwater drum, channel catfish, and river carpsucker. Tributary sites were composed mainly of fish of the order perciformes, including several sunfish species. Black bullheads (*Ameiurus melas*) were also sampled mainly in tributary sites.

The Upper region CCA showed less distinct patterns than the Middle and Lower regions (Fig. 15). While there was some grouping of perciformes fishes around the tributary factor in ordination space, both mainstem and tributary sites had a considerable amount of interspersion.

Multiple Linear Regression- ST Dataset (Q4)

Full dataset

Relationships between environmental variables and fish richness for the ST dataset were weak. Few ST results matched KSU trends, as variation in sampling year and season was high in the ST long-term monitoring dataset. For the entire watershed, no environmental variables had a significant relationship with species richness for the ST dataset. ST models had high, non-significant p-values, and R2 values were low [Table 6; ωi = 0.02 - 0.13, Adj. R2 = -0.03 - 0.07, p-value = 0.14 – 0.61].

Lower region

In the Lower region of the Smoky Hill River, for the ST dataset, only three regressors in two models had significant *p*-values (Table 7A). In a single model, summer mean flow was significant, but variable importance was low (0.25). In a single model, the distance to upstream CoE dam was significant, but again, variable importance was low (0.18). The slope of the regressors in the ST analysis did not match the direction of the relationships in the KSU data. However, in the ST dataset, the relationship between species richness and flow (Fig. 16G), and species richness and distance to nearest CoE dam (Fig. 17A) were highly variable and possibly influenced by the two lowest points. In the Lower region of the Smoky Hill River, for the ST dataset, stream type (mainstem vs tributary) affected fish species richness.

Middle region

In the Middle region of the Smoky Hill River, stream type was significant in both top models with a variable importance of 1.0 (Table 7B; Fig. 17E). As in the KSU dataset, mainstem sites had higher species richness than tributary sites (ST model averaged slope of -0.25).

Upper region

In the Upper region of the Smoky Hill River, for the ST dataset, 15 models had $AIC_c < 4$. However, these "top" models had high *p*-values (Table 7C; $p = 0.18-0.89$), were not significant, and R^2 values were low (Table 7C; Adj. $R^2 = -0.08-0.11$). In these weak models, no environmental variables were significantly related to fish species richness.

KSU-ST sample site comparison

Both KSU and Stream Team datasets were collected over the same geographic area (Fig. 19). However, the timing of sampling was quite different. KSU samples were collected over one summer season during 2015 (June-August), while Stream Team samples were collected in different seasons over a 17-year period (1996-2013; May-September). In addition to temporal variation, the choice of sampling sites between the KSU and ST datasets varied in important ways that reflected the different purposes for which the two datasets were collected. Furthermore, there was significant variation in the number of species captured within the study region by KSU and ST. KSU sampled a mean species richness of 4.6, while ST sampled a mean species richness of 11.4.

Land use

For land use, two of top three KSU models in the Lower region analysis showed a negative relationship between species richness and developed land, while none of the ST models in the Lower region showed this relationship. As an explanation, KSU sample sites were chosen to include a greater amount of variation in developed land (Quartile 1-3: 0.09-0.17) than ST sites (< 0.10; Fig. 20).

Herbaceous grassland in the Upper region showed a positive relationship with species richness in the KSU dataset, while there was no relationship between herbaceous grassland and fish species richness in the ST dataset. As an explanation, the KSU sites were chosen to cover a wider range of herbaceous grassland values than the ST dataset (including many more sites in the Lower range, 0-0.2; Fig. 21).

Flow

In the Lower region, for the ST dates, summer mean flow was negatively related to richness in a single model (Table 7A) whereas no relationship for this variable in this region was significant for the KSU dataset. This trend may be driven by a single low-flow point in the ST dataset (Fig. 22). In the Middle and Upper regions, for the KSU dataset, a positive relationship existed between flow and fish species richness. This trend was not observed for this variable in this region in the ST dataset. As an explanation for differences, KSU sample sites had a greater amount of variation in summer mean flow than ST sites (including sites with low flow; Fig. 23-24).

CoE dams

In the ST dataset, but not in the KSU dataset, the distance to upstream dam was significantly related to fish species richness in the Lower region. Again, this difference in results may be related to differences in the choice of sample sites. The KSU dataset had more sites closer to (5 vs. 1 <20 km) and farther away from (2 vs. 0 >100) CoE dams than the ST dataset (Fig. 25).

Confluences

For the Upper region, the KSU dataset showed a positive, significant relationship between species richness and distance to confluence while the ST dataset did not show this trend. Because the KSU dataset included many points in the Upper region, the KSU dataset had a wider range of confluence values including a few more that were very close to a confluence (< 0.5 km) and many more that were very far from a confluence (> 5 km) (Fig. 26).

Discussion

Role of Region for Watershed Patterns

Across a watershed, environmental variables that affect species richness can vary regionally. In a river such as the Smoky Hill that spans a large geographic area that parallels environmental gradients (i.e., precipitation), components of environmental variables can vary significantly in the Upper, Middle, and Lower watershed. A whole system analysis can mask complex regional trends. Thus, it is essential to consider regional patterns. For example, in my analysis, whole watershed trends did not detect a number of important regional patterns.

Human conversion of land cover for municipalities and agriculture is rarely debated in its importance to biotic communities and impacts on in-stream habitat quality. The Great Plains of the central United States is an area of intense agriculture and ranching, both of which pose significant threats to stream health (Hall et al. 1999). In the Smoky Hill watershed, developed land was negatively associated with species richness. The adverse effects that developed land can have on instream habitat and species richness through impacts such as channelization, water withdrawal, and altered connectivity has been widely reported elsewhere (Jurajda 1995, Winter et al. 1998, Cumming 2004). A study within a sub-basin of the Kansas River found increased nutrient concentrations in tributaries located in the lower-reaches of the basin, reflecting an increased density of row-crop agriculture (Dodds and Oakes 2006). In the Lower, most populated region of the Smoky Hill River, the negative effect of developed land on species richness was more pronounced than in the less populated Middle and Upper regions.

In contrast, a positive relationship between species richness and herbaceous grassland was observed. This trend was strongest in the Upper region, weak in the Middle region, and nonexistent in the Lower region. When comparing the Lower, Middle, and Upper regions of the

Smoky Hill watershed, the proportion of herbaceous grassland in the Upper watershed is greater. The positive relationship that I noted between richness and herbaceous grassland in the Middle and Upper regions is likely due to the positive effect of a greater proportion of herbaceous grassland at sites in these regions. This response could be reflective of fish in the Middle and Upper regions responding positively to the natural land use "reference condition," which in a semi-arid prairie stream like the Smoky Hill is herbaceous grassland (Stoddard et. al 2006).

Flow alteration and water withdrawal are obvious major threats to aquatic organisms. Water withdrawal for human consumption generally reduces flow rates, which help to shape important characteristics of aquatic habitat (Matthews 1998). In the Smoky Hill watershed, higher summer mean flow was generally associated with higher species richness. In the KSU dataset, summer mean flow had no effect on species richness in the Lower region, but had a high importance and consistent significance in the Middle and Upper regions. This is reflective of the major decrease in flow between the Lower region and the Middle and Upper regions. In our Middle and Upper region sample sites in the Smoky Hill, flow is often considerably lower, even intermittent, in headwater and tributary streams. This is a common characteristic of prairie streams (Williams 1988). However, global climate change literature predicts altered flow regimes and water availability, which may place an even greater strain on aquatic organisms in particularly water limited systems (Covich et al. 1997). Since humans are highly dependent on aquatic resources and thus closely linked to them, it is important to diagnose and limit threats to water quality and quantity at the source in order to provide water security for humans as well as for freshwater diversity (Spinazola et al. 1985; Vörösmarty et al. 2010).

Stream type follows a similar pattern, with tributaries in the Lower region having higher species richness than mainstem sites and tributaries in the Middle and Upper regions having lower species richness than mainstem sites. This reflects different habitat availability between mainstem and tributaries in the different regions. When flow is high as in the Lower region, tributaries may represent more moderate and diverse habitats relative to the large, fast-flowing mainstem. Additionally, tributaries in dendritic systems serve as natural spatial discontinuities, altering components of stream size, discharge, substrate, and water chemistry- and have subsequently been shown to increase local biotic diversity (Kiffney et al. 2006).

Although distance to upstream dam and distance to confluence were the least influential variables across all regions in our analysis, this is not necessarily an indication of a lack of an effect of these environmental components on biota or fish response. The negative effects of dams on aquatic ecosystems, in particular, have been well documented elsewhere in the scientific community (Poff et al. 1997; Olden and Naiman 2009; Dugan et al. 2010). One specific study in the Kansas River Basin observed highly localized effects of reservoirs on upstream fish assemblages (Falke and Gido 2006). However, the sampling design used in this study included a considerably more detailed sampling of streams at sites closer to reservoirs and in tributaries directly connected to reservoirs. So, while large dams can serve as major discontinuities in the river continuum, their effects on fish communities were not seen using our watershed-scale sampling design. In other whole river studies, the effects of dams on fish can also be masked (Welch et al. 2008; Rechisky et al. 2013).

Fish Response

Fish species richness provided insights about biotic patterns. One of the simplest and most common ways that ecologists have measured species diversity is by species richness (Humphries et al. 1995). Even with the development of different ways of quantifying biological diversity, maximizing species richness has remained an explicit or implicit goal of conservation studies (Gotelli and Colwell 2001).

CCA ordinations of the KSU dataset further support the idea of differences in fish communities between mainstem and tributary sites, possibly related to differences in flow, habitat type, and heterogeneity. Tributary sites were dominated by a diversity of perciformes fishes, while mainstem sites were composed mainly of cypriniformes and other smaller-bodied species. Similar differences in fish richness and assemblage structure were noticed in another study of the Kansas River basin, with abrupt changes in fish assemblages between mainstem river sites and tributaries above confluences (Thornbrugh and Gido 2009).

Furthermore, while our analyses provided some useful information as to which environmental variables affect species richness and fish communities, looking in more detail at which fish are considered to be native and non-native may provide some interesting insights. I categorized all species collected in 2015 KSU samples using the USGS Nonindigenous Aquatic Species (NAS) database, which suggested that 33% of all fish collected in my samples were either native transplants or exotic species (Appendix 3). Further research looking at how non-native species negatively impact lotic ecosystems as well as the benefits of restoring native biodiversity is needed.

Choice of Sites, Timing of Sampling, and Sampling Gear Choice Affect Results

The results of the same data analysis procedure differed for the KSU and ST datasets. These differences are not surprising because the two datasets were collected for completely different reasons that motivated very different sampling designs and choice of sample sites. The KSU dataset was designed to collect environmental and fish data across the entire watershed in a narrow time period in order to minimize confounding temporal variation caused by year and month. The KSU dataset intentionally included sites spread throughout the three regions that were equally divided between mainstem and tributaries. The KSU dataset also included sites with a range of values for each environmental variable. Because the KSU research sought to link anthropogenic and climate impacts to fish biodiversity, data were collected at a predetermined set of sites, not just where fish were abundant and diverse in previous samples. One major implication of this is that some pre-determined KSU sample sites had little to no flow, and thus few or no fish. This contributed to a decreased mean richness at KSU sites relative to ST sites. Additionally, this likely also contributed to the importance of summer mean flow in our model selection analysis.

On the other hand, the ST dataset was designed as a long-term fish sampling and intentionally focused on sites where fish had been collected in the past rather than at sites chosen for research design reasons a priori. In addition, the ST sites intentionally sampled over a wide range of seasons and years. For example, the ST dataset (N=33) was collected over a 17-year period between 1996 and 2013 with a maximum of 10 samples collected in one year (2010). The KSU dataset (N=48) was sampled entirely during the summer of 2015, representing the conditions during that single season. Since tributaries and headwaters in prairie streams are known to vary considerably in their hydrology from year to year, the ST decision yielded both generality and variability. Because the KSU dataset was collected under less variable conditions, this dataset provided a more cohesive relationship between environmental variables and fish diversity.

Another important difference between the ST and KSU sampling methods was choice of sampling gear. The KSU data collection sought to maintain standardization between sites by using the same sampling distance with the same gear type (hand trawl) that best sampled the most common habitat type- mid-channel pool. Because of different goals and protocols, the Stream Team applied more effort at each site, sampled in more habitats at each site, and used a variety of gear (seining and electro-fishing) at each site. This difference in sampling protocol resulted in higher mean richness in ST samples, likely because this protocol caught more uncommon and rare species in the Smoky Hill basin.

Although neither the KSU nor ST sample design is better or worse than the other, my analysis indicates that choice of sampling sites and design affects the outcome. Consequently, thoughtful decisions about goals are essential before sampling and analysis occurs.

Future Considerations

Below, I discuss potential implications that this research may have for future study and monitoring of Great Plains Streams. First, it is critical to consider the scale at which environmental variables may affect biota. I intentionally conducted my study at a large scale in order to see which variables may be affecting biota throughout the watershed. While this did provide information on relationships between components of the environment and biota, we obtained more information by dividing our study area into regions. Finer scale trends, examined through microhabitat variables, may yield different patterns still. So, a conscious choice on scale is essential.

Lastly, when sampling across such a large spatial extent, intensive single year sampling may be more useful for examining ecological relationships than multiple year sampling aimed towards monitoring fish species presence and absence. This may be particularly applicable in a highly variable system like the Smoky Hill River, where a flowing tributary from which fish were collected one year may be dry the next. In future studies it is important to clearly state research goals and develop sampling methods that fit those goals.

Figures and Tables

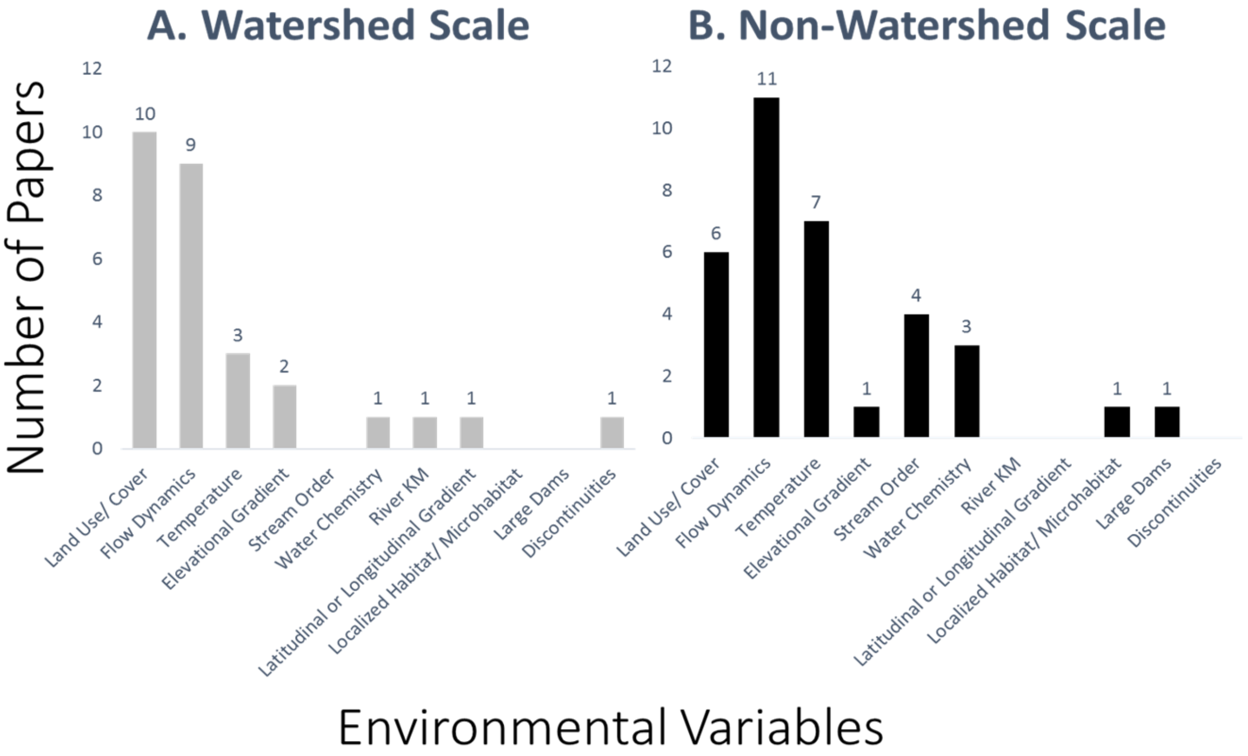


Figure 1. Bar plots numerically representing the environmental variables present in each study in the literature at the A) watershed scale, and B) non-watershed scale. Since not every study in the search empirically analyzed an environmental variable, not all studies are represented in this plot. The Y-axis represents the number of papers in the literature search that analyzed a given variable, while the X-axis represents each variable that was analyzed.

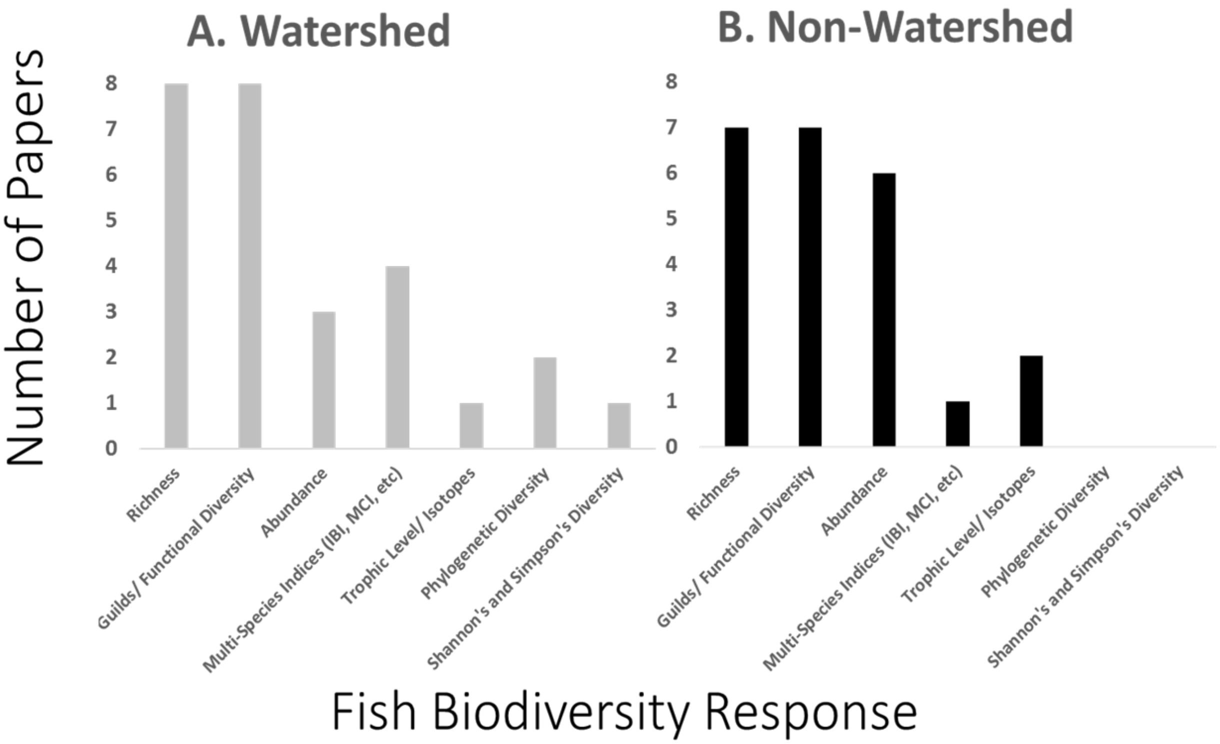


Figure 2. Bar plots numerically representing the biodiversity response variables present in each study in the literature search at the A) watershed scale, and B) non-watershed scale. Since not every study in the search empirically analyzed a biodiversity response variable, not all studies are represented in this plot. The Y-axis represents the number of papers in the literature search that analyzed a given variable, while the X-axis represents each variable that was analyzed.

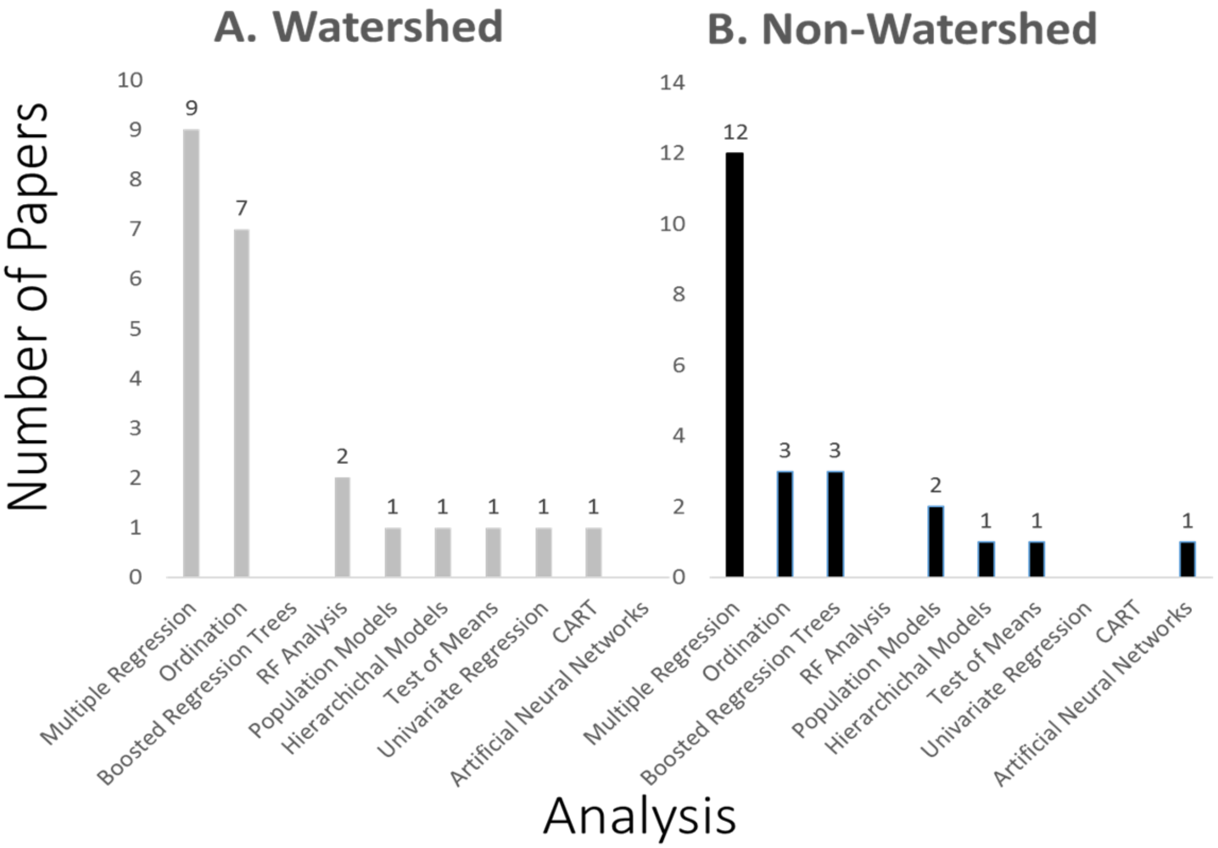


Figure 3. Bar plots numerically representing the statistical analysis used in each study in the literature search at the A) watershed scale, and B) non-watershed scale. Since not every study in the search used statistical analyses, not every study is represented in this plot. The Y-axis represents the number of papers in the literature search that used a specific type of statistical analysis, while the X-axis represents each type of statistical analysis that was used.

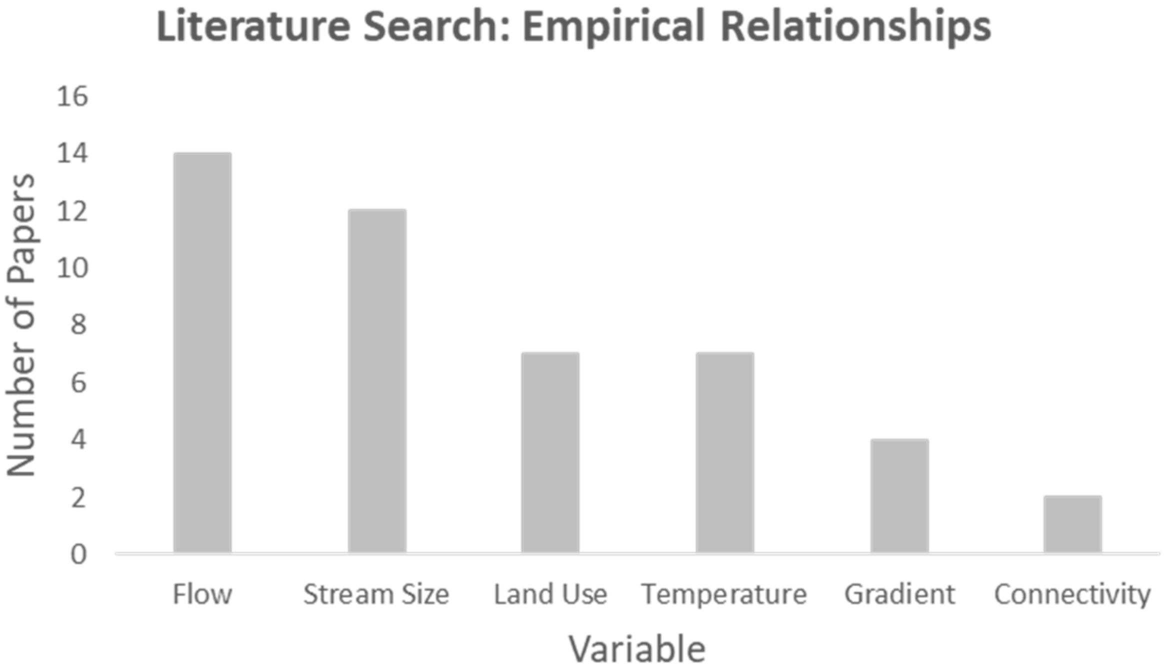


Figure 4. Bar plot indicating relationships inferred between environmental variables and a measure of biotic diversity in the papers included in my literature search. Not all papers included in the literature review are included here, as not all papers produced empirical results inferring relationships between biota and environmental metrics.

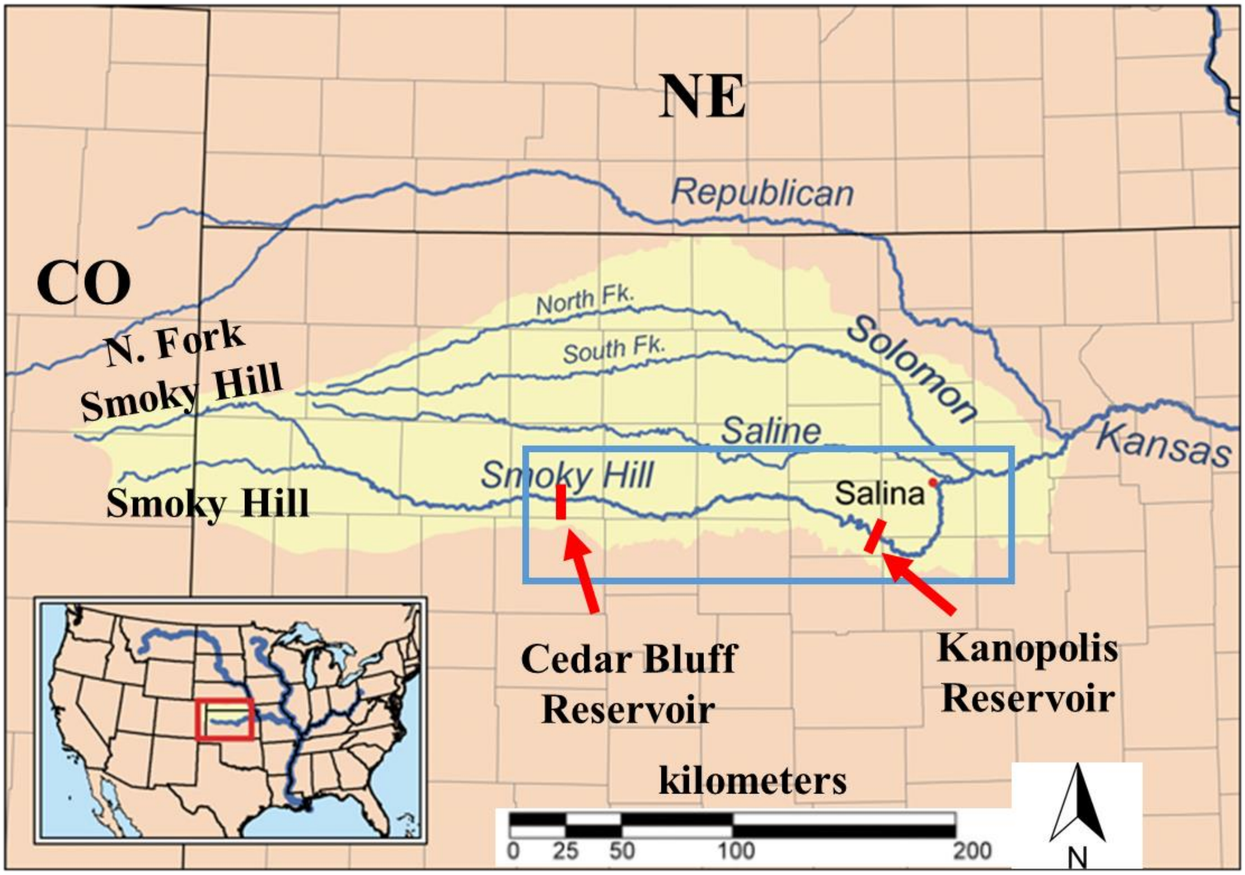


Figure 5. Map showing the origin of the Smoky Hill River and the North Smoky Hill Rivers in eastern Colorado and their junction in Logan County, KS. The river then flows through central Kansas where it joins the Saline and Solomon rivers near Salina, KS, then further east to Junction City where it joins the Republican River to form the Kansas River. The study area for this project is indicated the light blue box. Corps of Engineers dams are indicated by thick red lines. Figure credit- USGS.

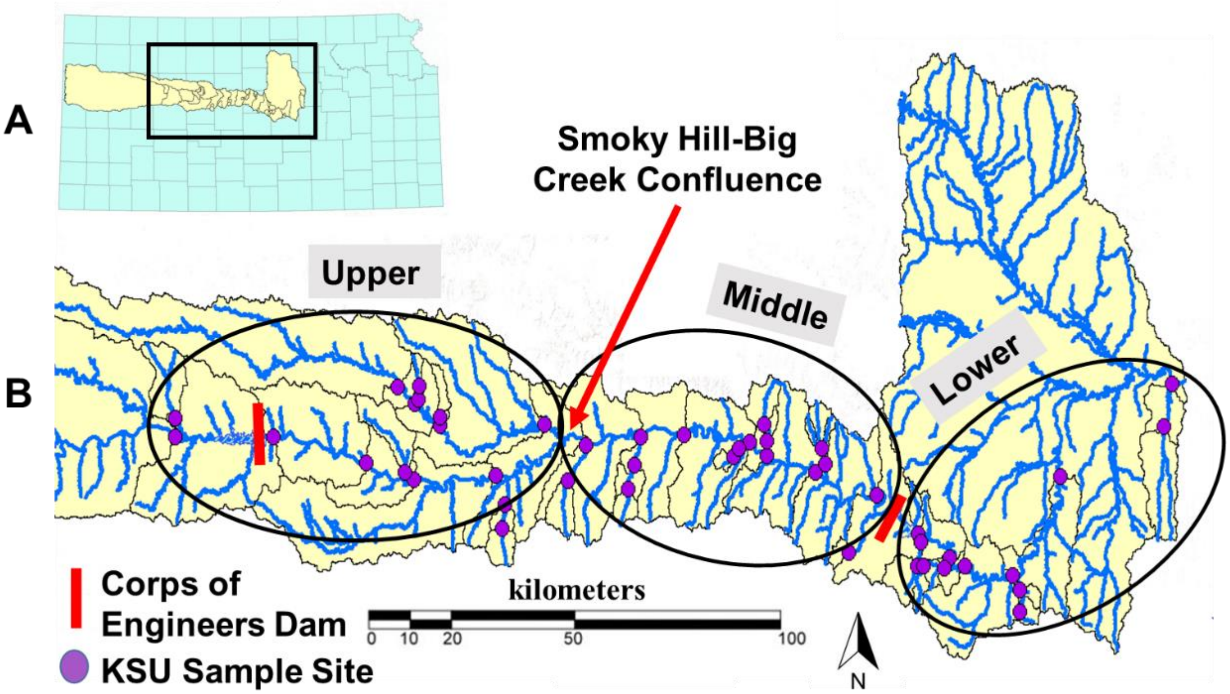


Figure 6. Map of study area within the Smoky Hill River-Big Creek watershed, KS. Included in the map of the watershed are the divisions for the three regions of the reservoir (Upper region, Middle region, and Lower region), locations of Corps of Engineers Dams, and the location of the confluence of the Smoky Hill River and Big Creek. Regions are used to describe trends across a large geographic area. The division between Lower and Middle regions is Kanopolis reservoir, while the division between the Middle and Upper regions is the confluence of Big Creek.

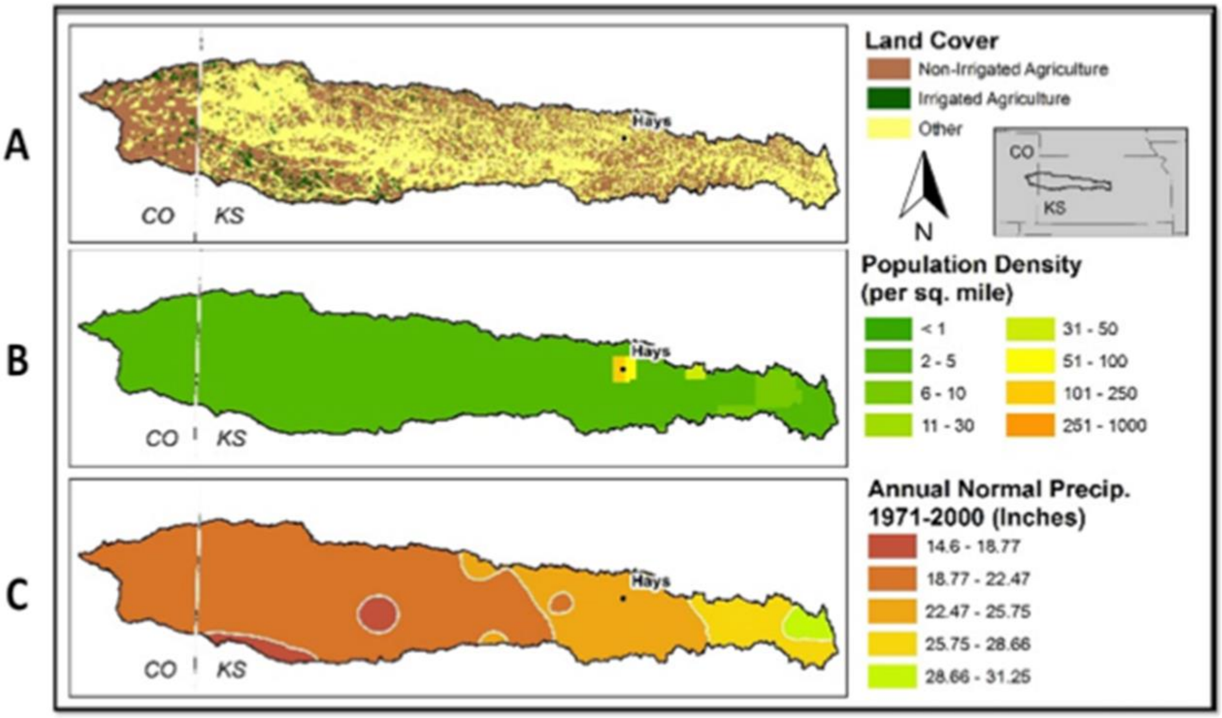


Figure 7. Maps showing the delineation of the Smoky Hill-Big Creek watershed within KS and patterns of (A) land cover, (B) population density, and (C) annual normal precipitation. Figure credit- Melinda Daniels.

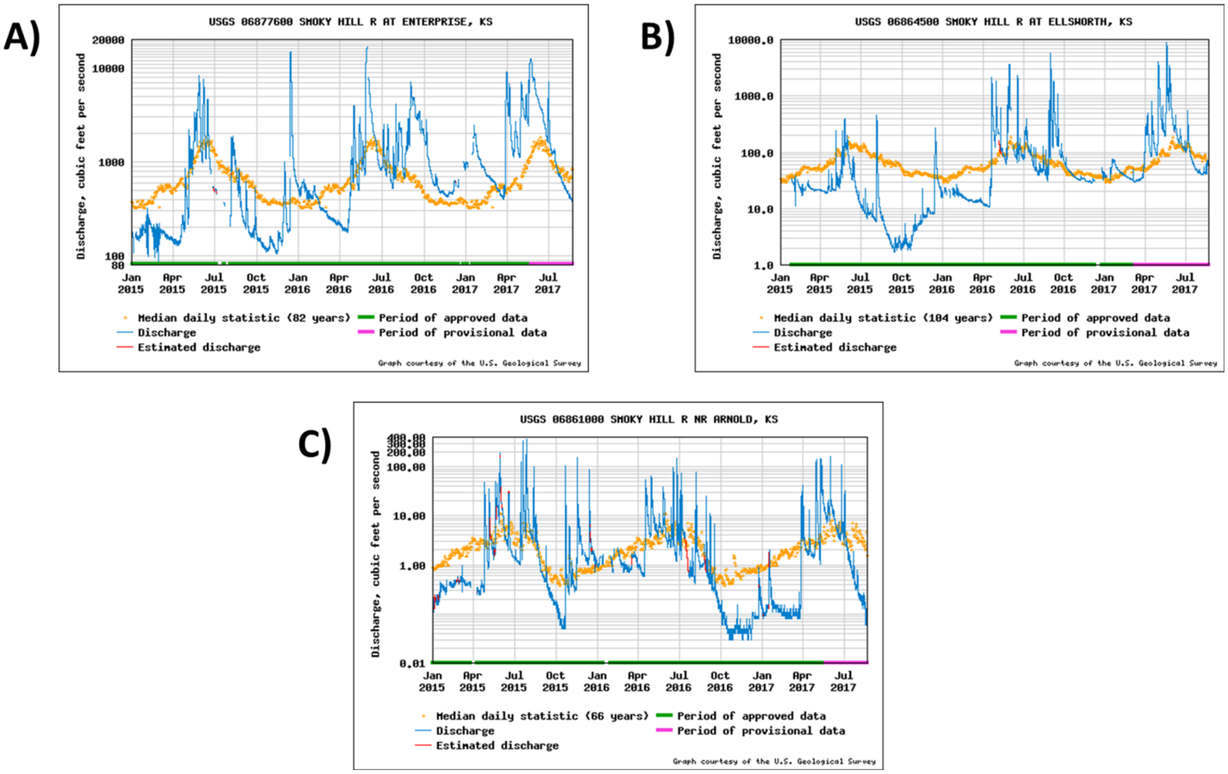


Figure 8. Discharge in cubic feet per second for USGS gauges on the Smoky Hill River in three regions of the basin: A) Smoky Hill River at Enterprise, KS (Lower Region), B) Smoky Hill River at Ellsworth, KS (Middle Region), and C) Smoky Hill River near Arnold, KS (Upper Region). Graphs and data courtesy of the United States Geological Survey (USGS).

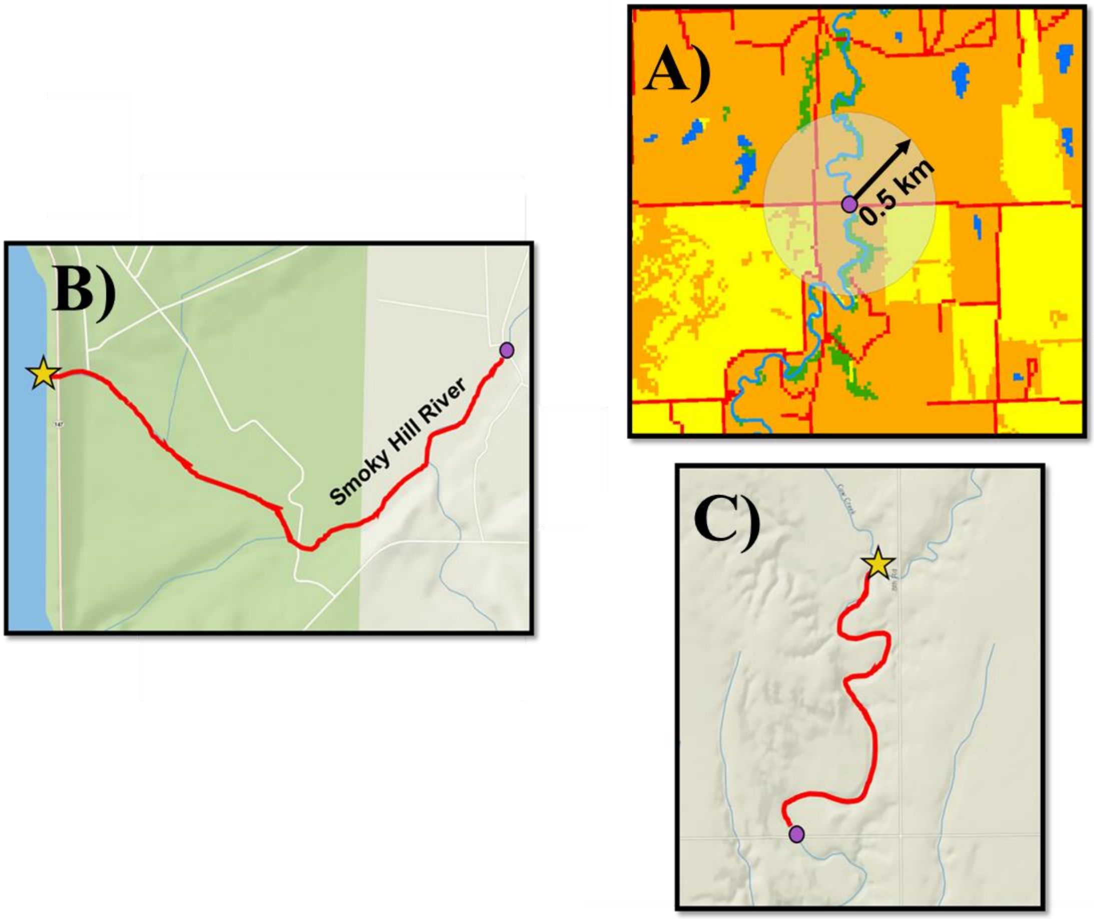


Figure 9. Example figure showing methods used to calculate A) buffer land use with 0.5 km radius buffer surrounding each sample site, B) distance to upstream dam using the measure tool in ArcGIS, and C) distance to nearest confluence using the measure tool in ArcGIS. KSU sample points are indicated by purple dot, while the end location measured to for distance measurements is indicated by a gold star.

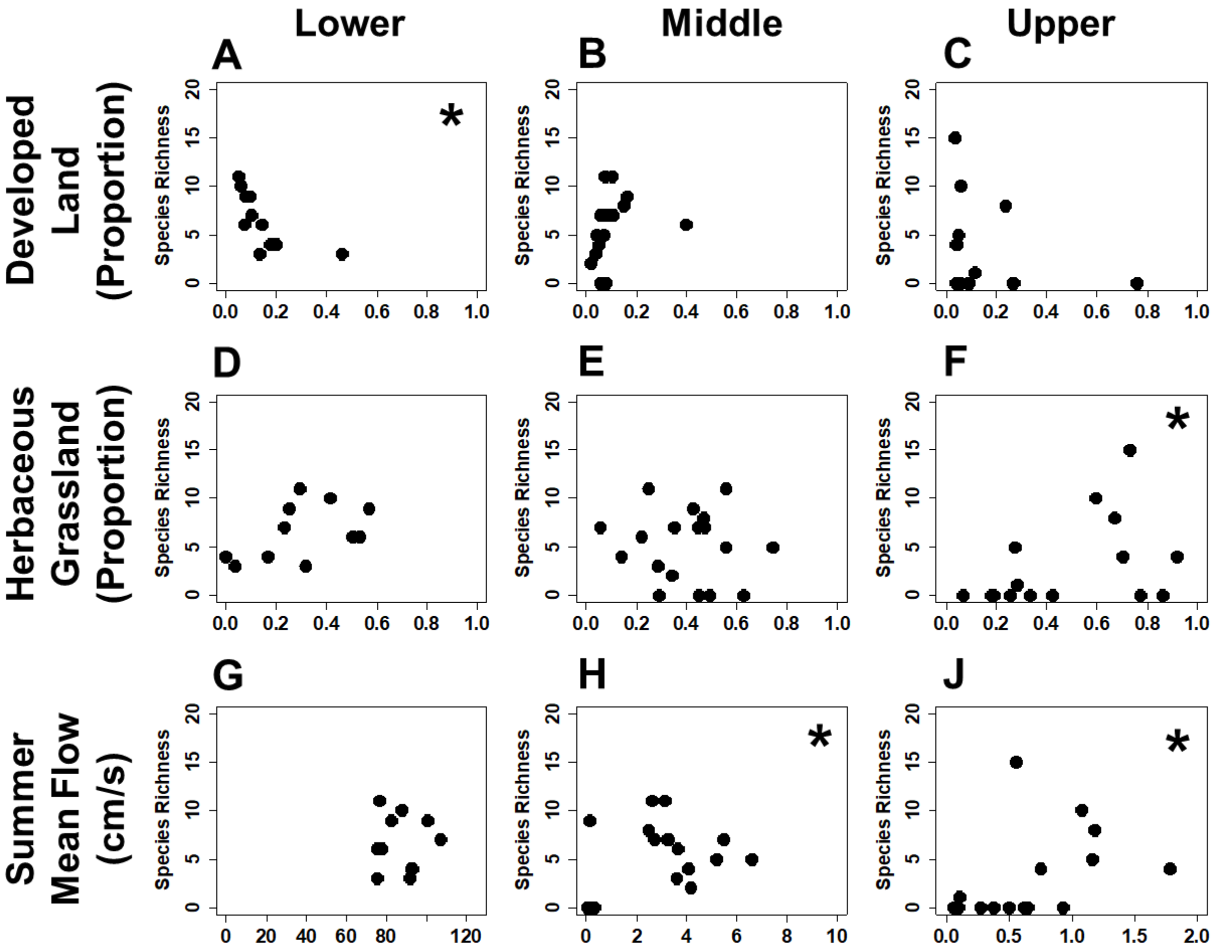


Figure 10. KSU data scatterplot matrix of first three environmental variables vs. species richness for all three regions. Significant variables are indicated by an asterisk in the top right corner of the scatterplot.

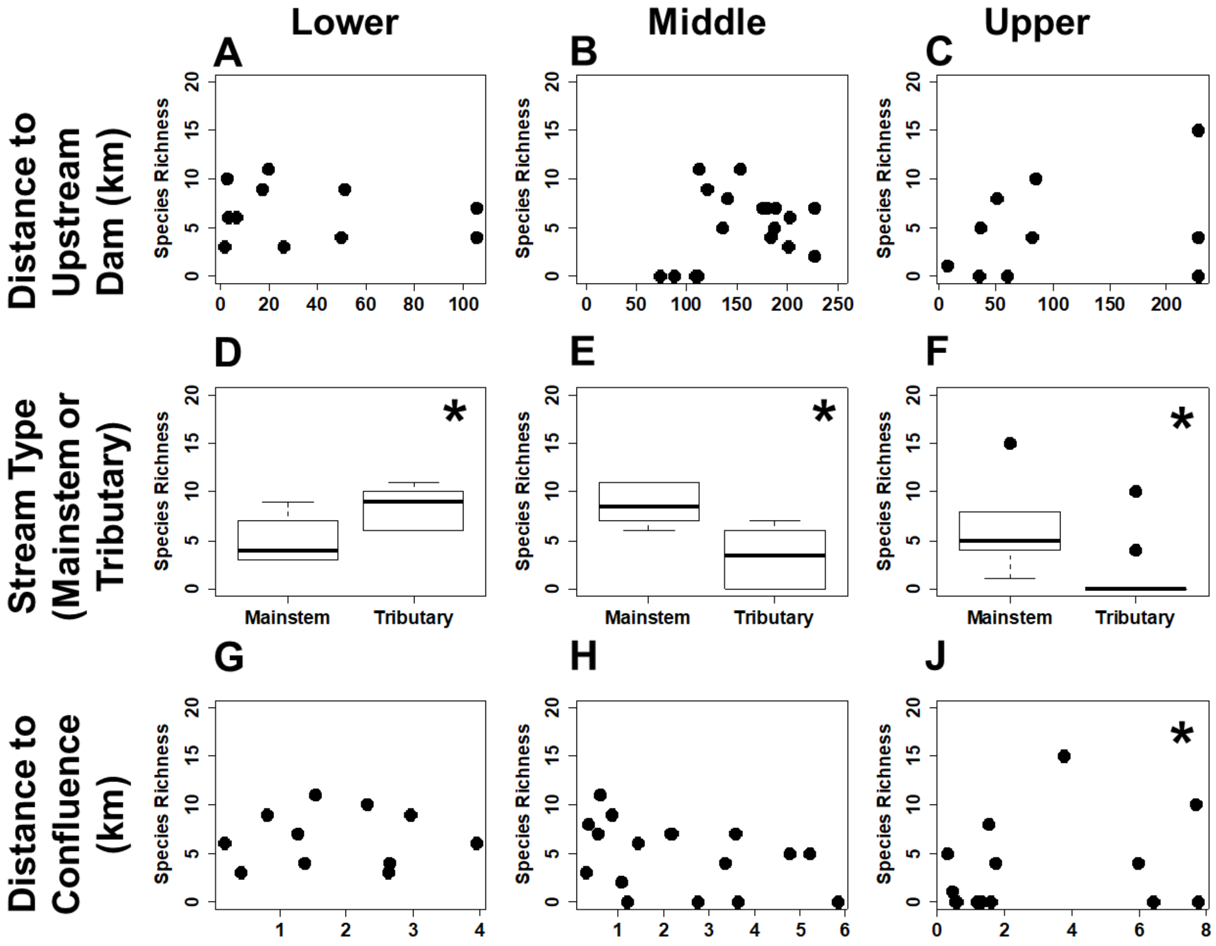


Figure 11. KSU data scatterplot matrix of second three environmental variables vs. species richness for all three regions. Significant variables are indicated by an asterisk in the top right corner of the scatterplot.

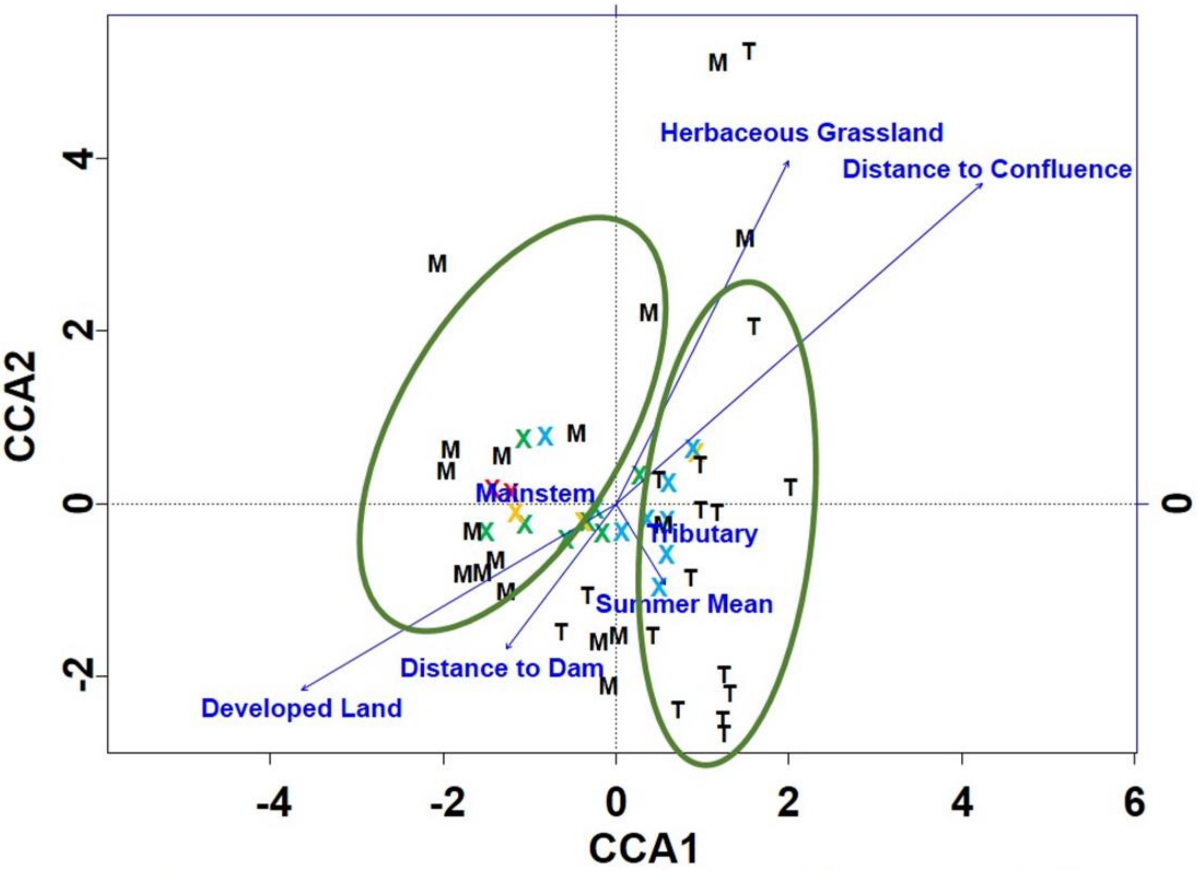


Figure 12. Canonical correspondence analysis plot for full KSU dataset with green circles showing distinct groupings of tributary sites (indicated by "T" before a set of numbers) and mainstem sites (indicated by "SH" before a set of numbers). Fishes comprising four orders were collected across all sample sites, with each species represented by a colored X on the plot. Light blue represents perciformes, green represents cypriniformes, orange represents siluriformes, and red represents cyprinodontiformes. Environmental factors potentially affecting fish presence are in blue.

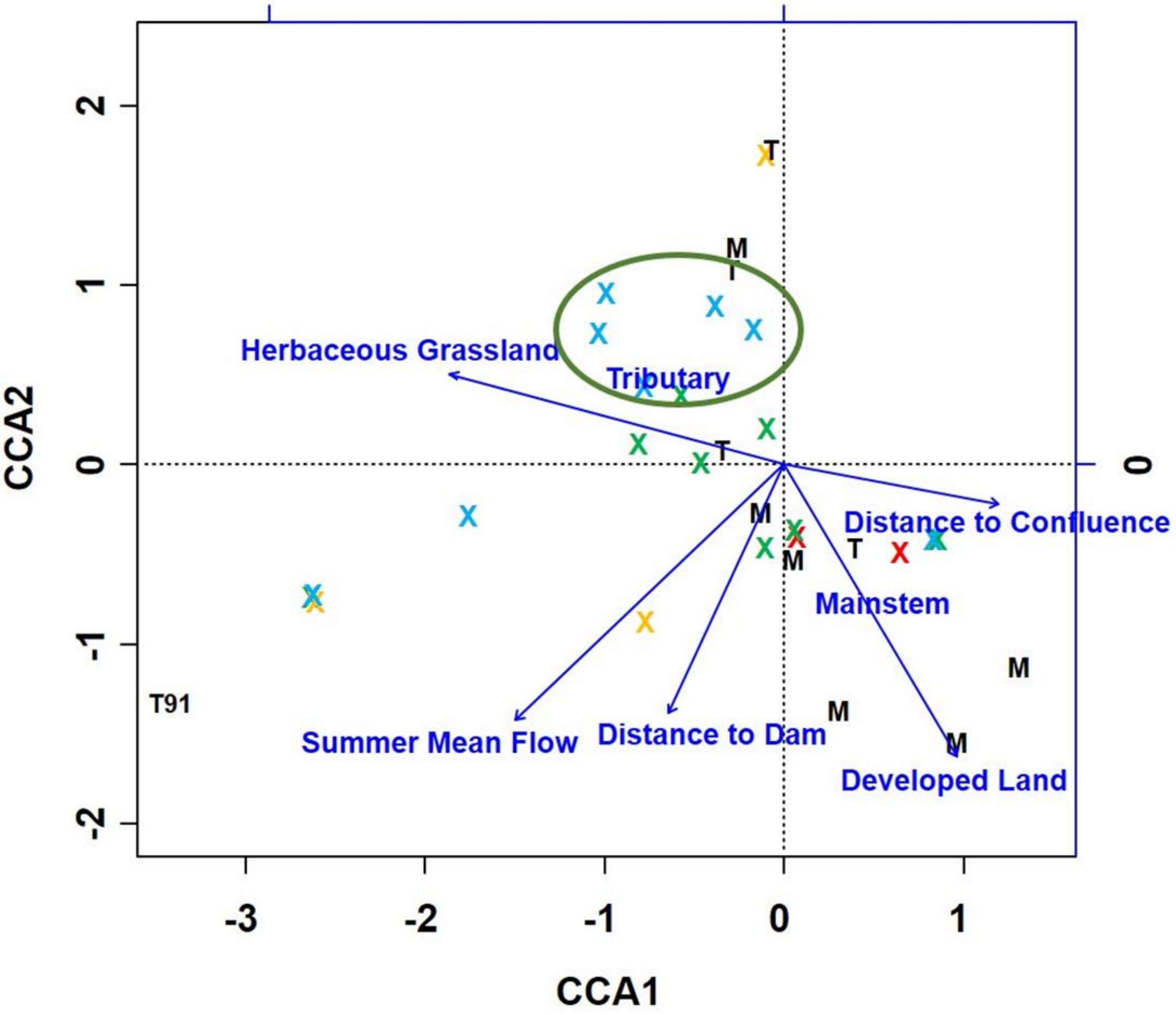


Figure 13. Canonical correspondence analysis plot for Lower region KSU dataset with green circles showing distinct groupings of tributary sites (indicated by “T” before a set of numbers) and mainstem sites (indicated by “SH” before a set of numbers). Fishes comprising four orders were collected across all sample sites, with each species represented by a colored X on the plot. Light blue represents perciformes, green represents cypriniformes, orange represents siluriformes, and red represents cyprinodontiformes. Environmental factors potentially affecting fish presence are in blue.

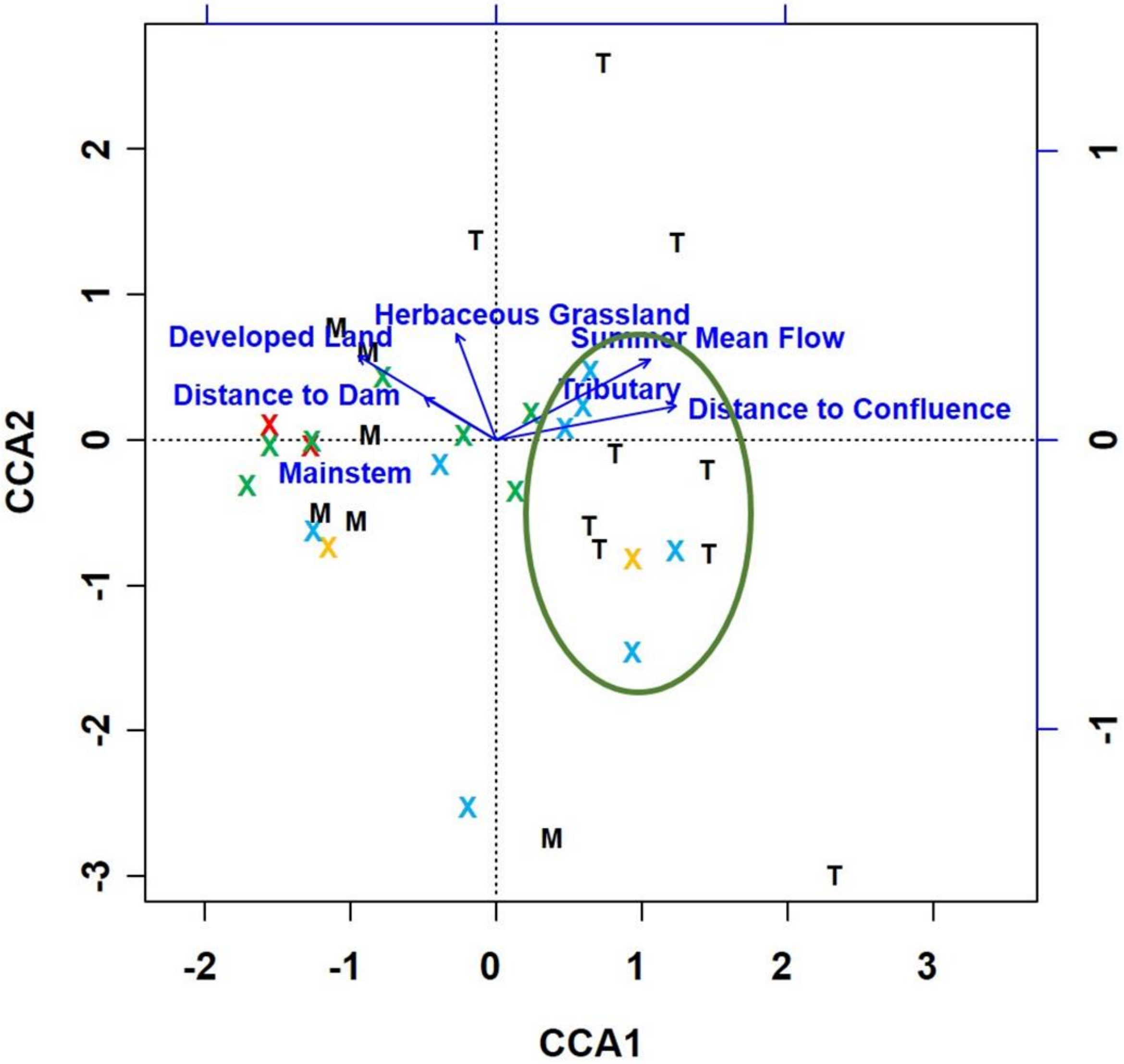


Figure 14. Canonical correspondence analysis plot for Middle region KSU dataset with green circles showing distinct groupings of tributary sites (indicated by "T" before a set of numbers) and mainstem sites (indicated by "SH" before a set of numbers). Fishes comprising four orders were collected across all sample sites, with each species represented by a colored X on the plot. Light blue represents perciformes, green represents cypriniformes, orange represents siluriformes, and red represents cyprinodontiformes. Environmental factors potentially affecting fish presence are in blue.

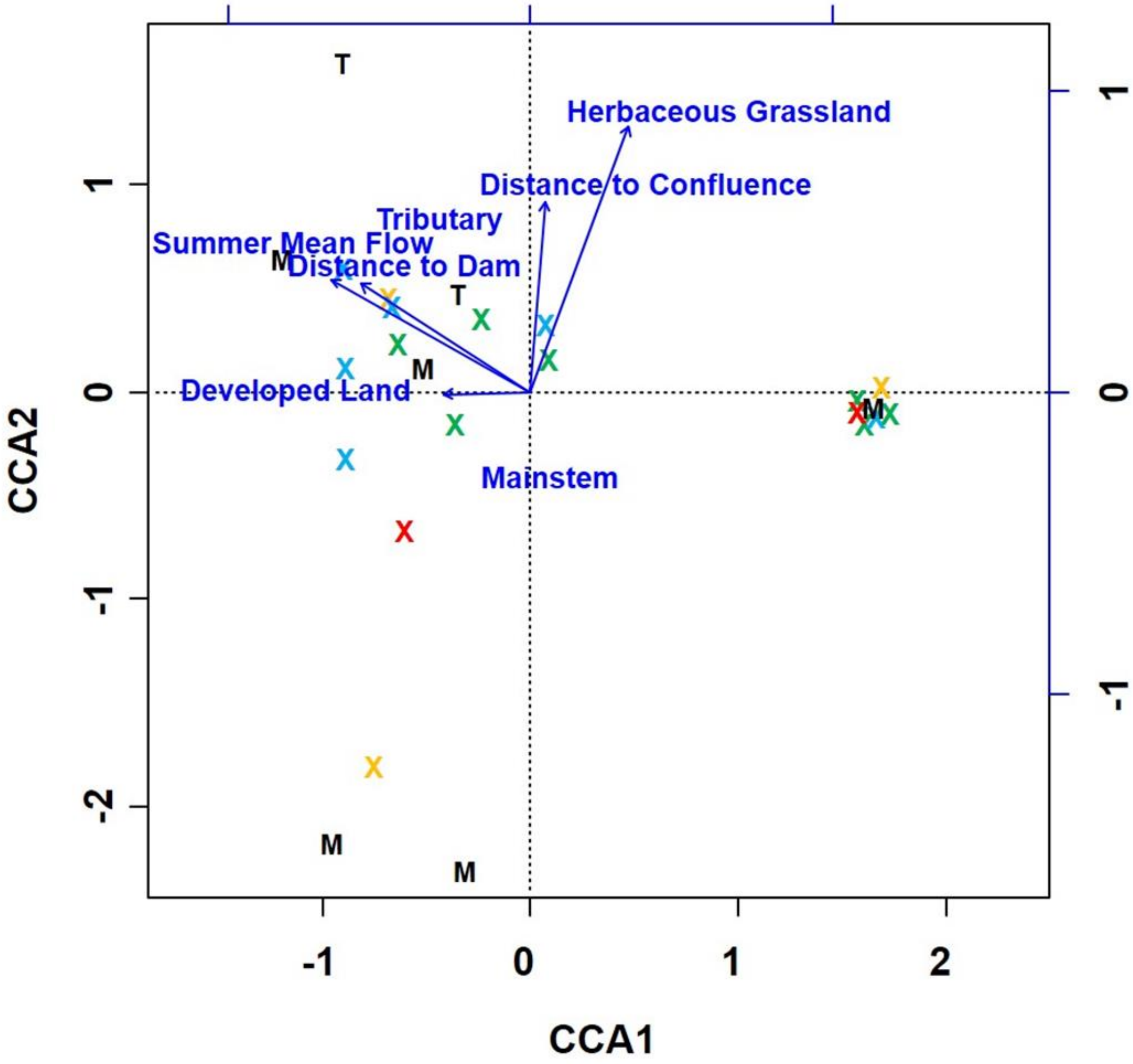


Figure 15. Canonical correspondence analysis plot for Upper region KSU dataset with green circles showing distinct groupings of tributary sites (indicated by "T" before a set of numbers) and mainstem sites (indicated by "SH" before a set of numbers). Fishes comprising four orders were collected across all sample sites, with each species represented by a colored X on the plot. Light blue represents perciformes, green represents cypriniformes, orange represents siluriformes, and red represents cyprinodontiformes. Environmental factors potentially affecting fish presence are in blue.

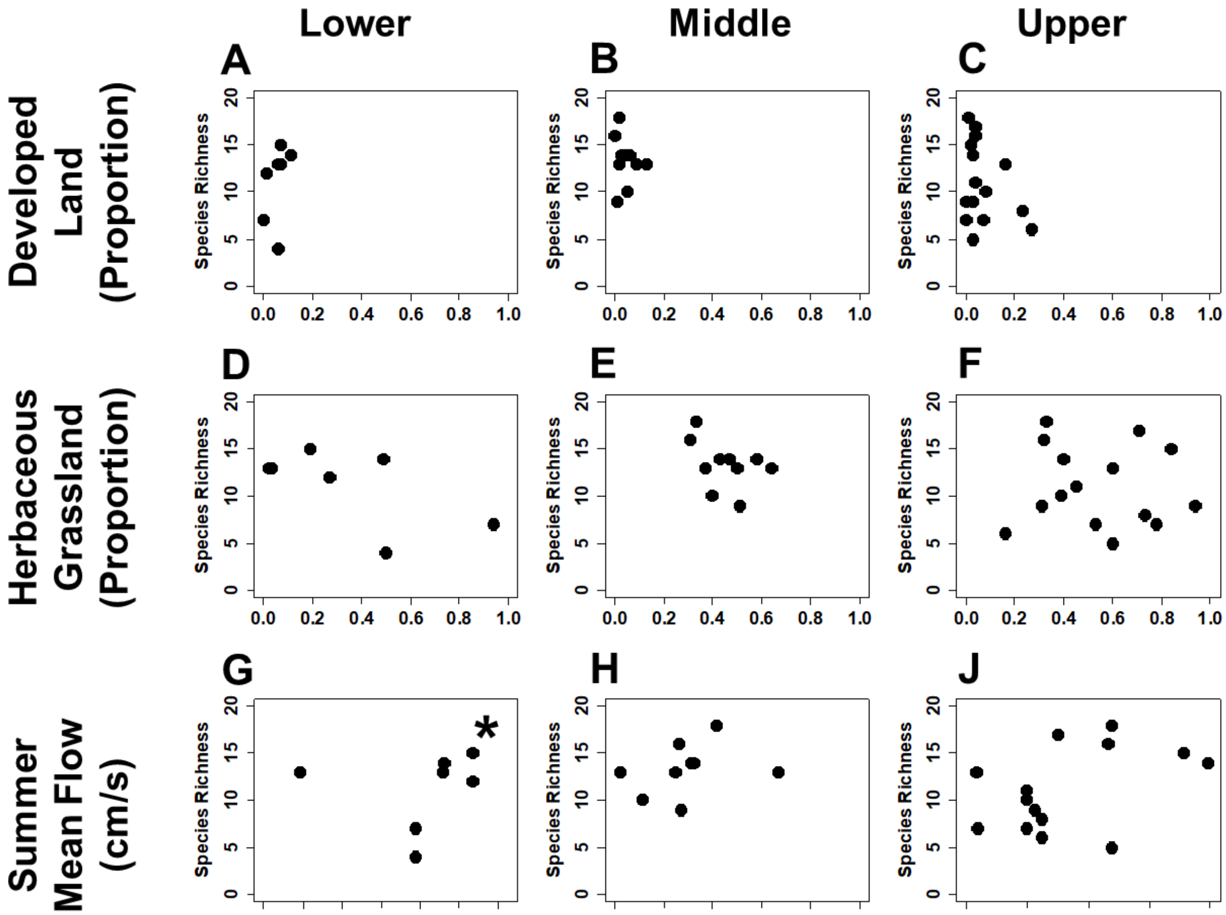


Figure 16. Stream Team data scatterplot matrix of first three environmental variables vs. species richness for all three regions. Significant variables are indicated by an asterisk in the top right corner of the scatterplot.

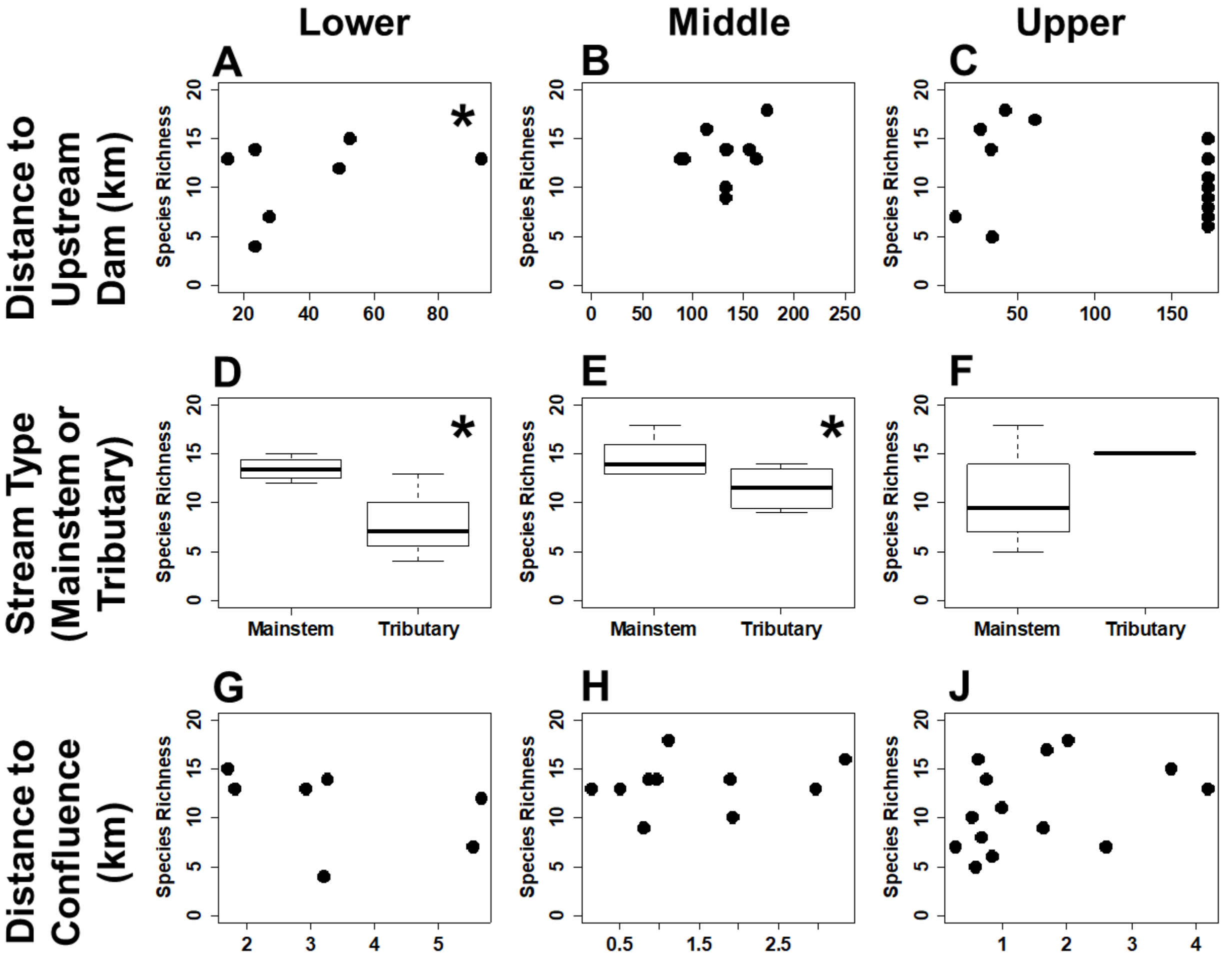


Figure 17. Stream Team data scatterplot matrix of second three environmental variables vs. species richness for all three regions. Significant variables are indicated by an asterisk in the top right corner of the scatterplot.

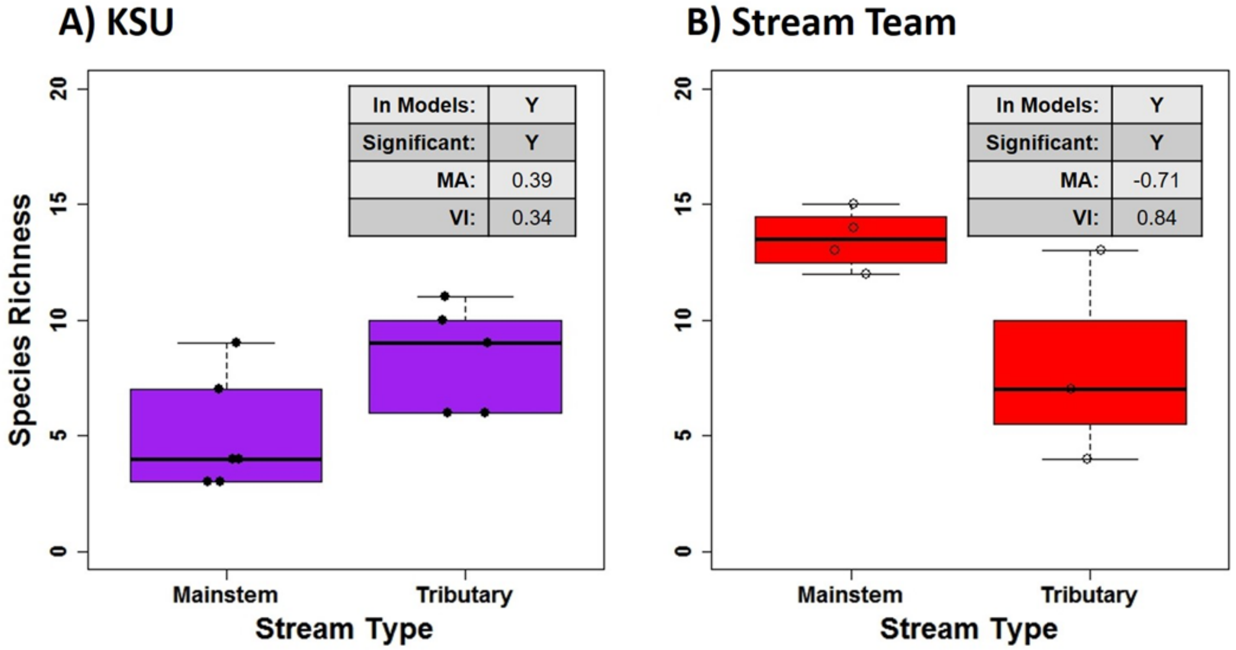


Figure 18. Stream type vs. species richness plots for A) KSU Lower region dataset, and B) Stream Team Lower region dataset. A box on each plot indicates 1) whether this variable was in top models Δ AIC$_c$ <4, 2) if the variable was significant in any top models, 3) Δ AIC$_c$ <4 model average (MA), 4) and Δ AIC$_c$ <4 variable importance (VI).

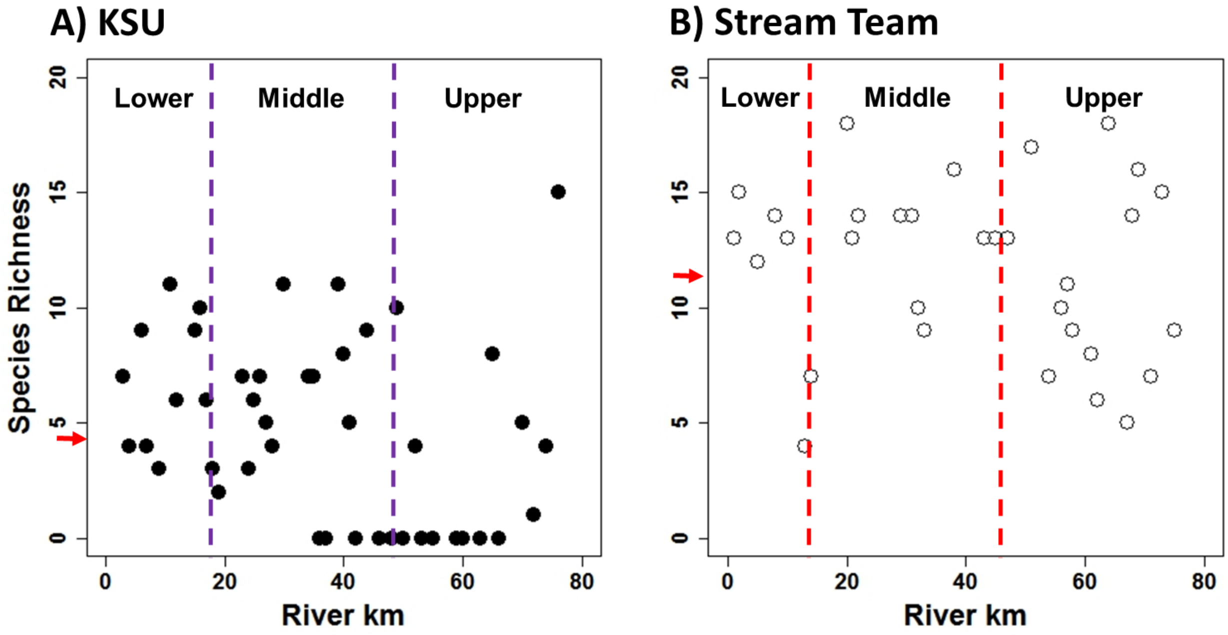


Figure 19. River kilometer vs. species richness plots for A) KSU full dataset, and B) Stream Team full dataset. Mean richness is indicated by a red arrow on the y-axis of each plot. Lower, Middle, and Upper regions are divided on each plot by a vertical dashed line.

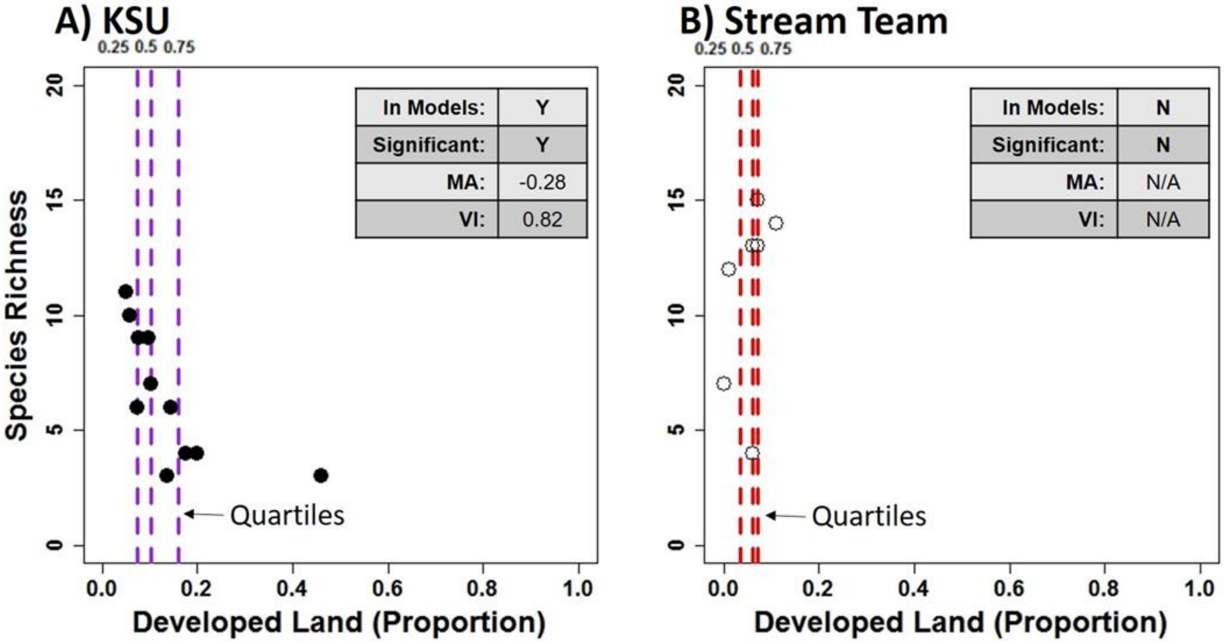


Figure 20. Proportion developed land vs. species richness plots for A) KSU Lower region dataset, and B) Stream Team Lower region dataset. 0.25, 0.5, and 0.75 quartiles for the explanatory variables are marked indicated by vertical dashed lines. A box on each plot indicates 1) whether this variable was in top models Δ AIC$_c$ <4, 2) if the variable was significant in any top models, 3) Δ AIC$_c$ <4 model average (MA), 4) and Δ AIC$_c$ <4 variable importance (VI).

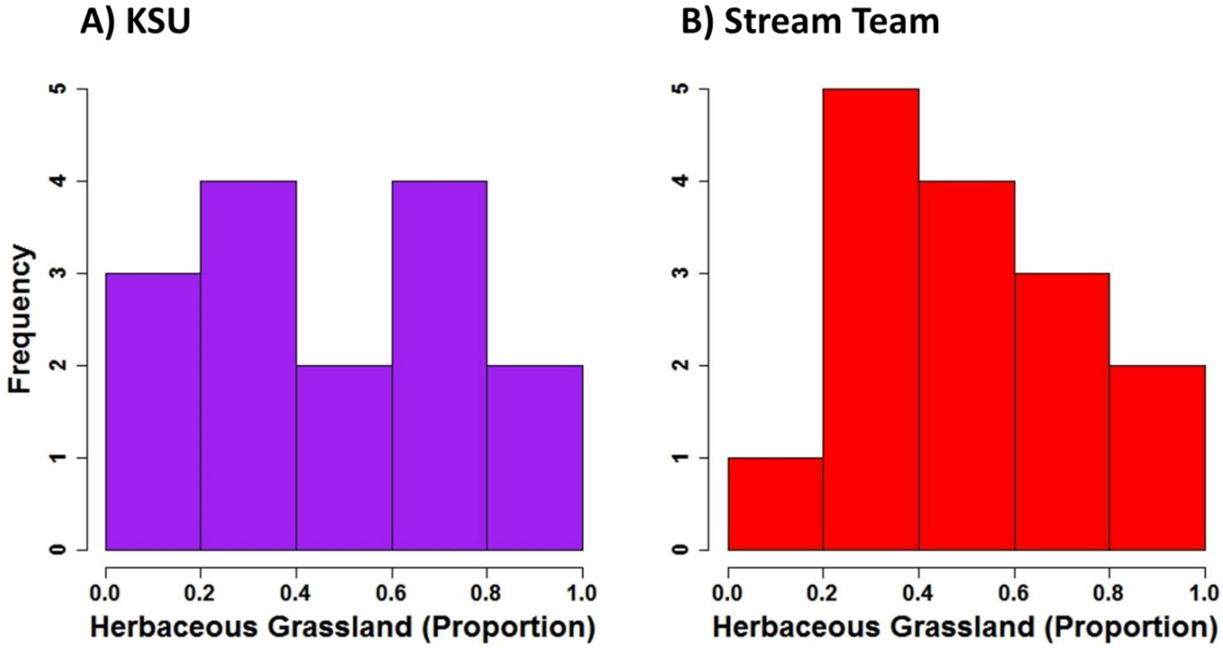


Figure 21. Histogram of Upper region proportion herbaceous grassland vs. frequency for A) KSU and B) Stream Team data.

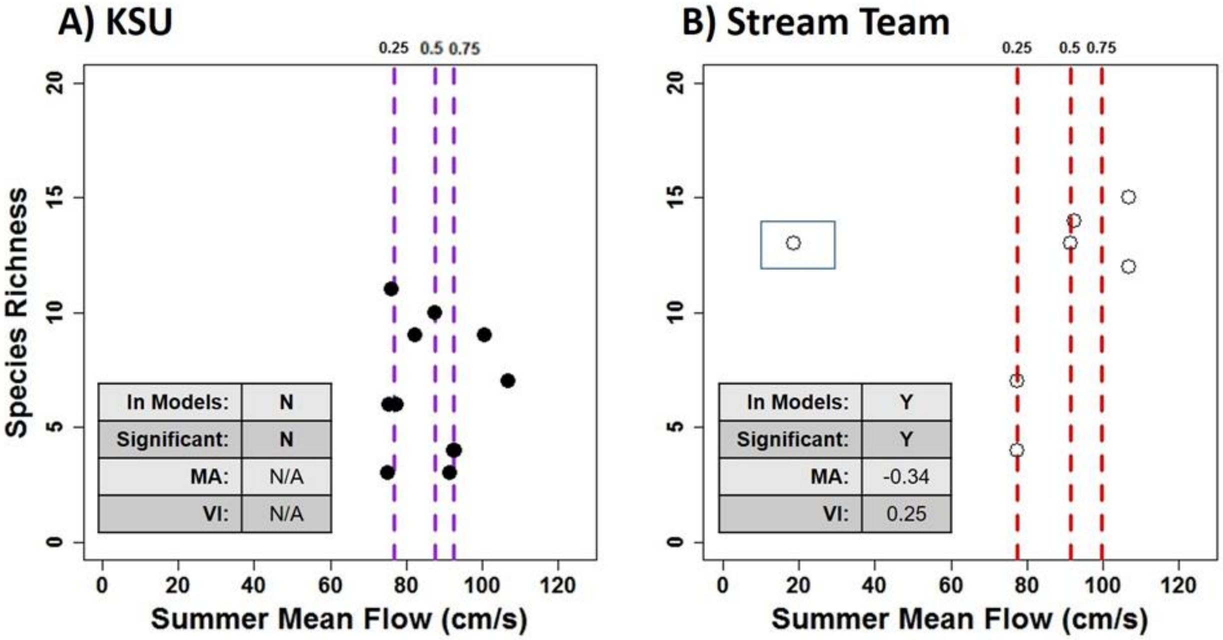


Figure 22. Summer mean flow (cm/s) vs. species richness plots for A) KSU Lower region dataset, and B) Stream Team Lower region dataset. 0.25, 0.5, and 0.75 quantiles for the explanatory variables are marked indicated by vertical dashed lines. A box on each plot indicates 1) whether this variable was in top models Δ AIC$_c$ <4, 2) if the variable was significant in any top models, 3) Δ AIC$_c$ <4 model average (MA), 4) and Δ AIC$_c$ <4 variable importance (VI).

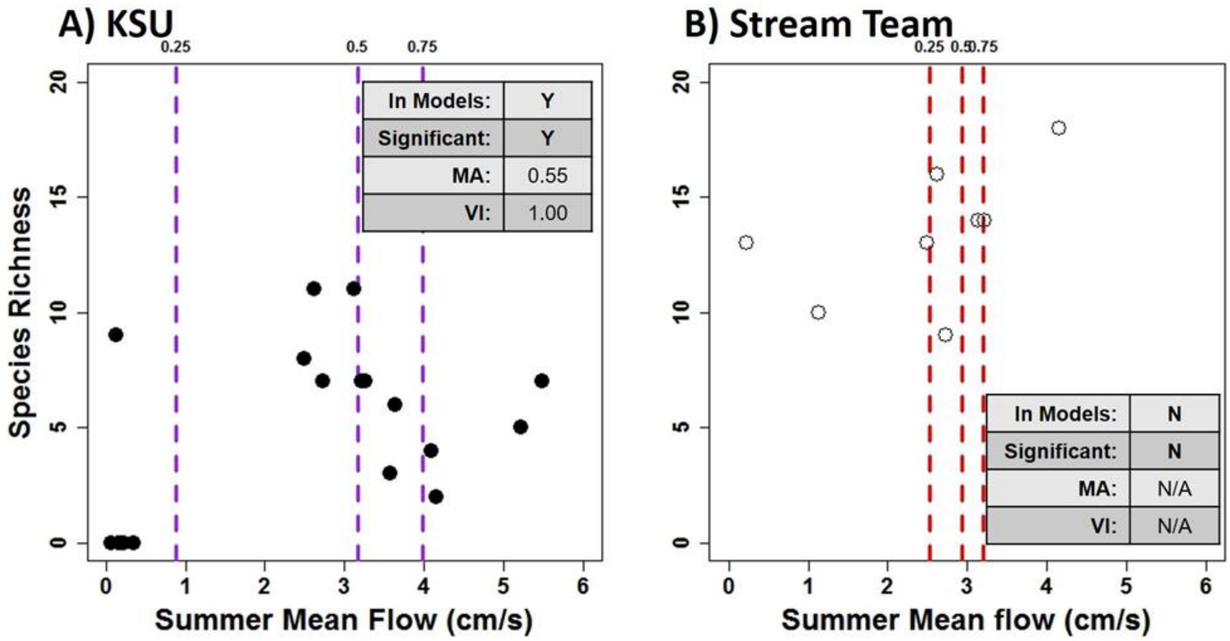


Figure 23. Summer mean flow (cm/s) vs. species richness plots for A) KSU Middle region dataset, and B) Stream Team Middle region dataset. 0.25, 0.5, and 0.75 quantiles for the explanatory variables are marked indicated by vertical dashed lines. A box on each plot indicates 1) whether this variable was in top models Δ AIC$_c$ <4, 2) if the variable was significant in any top models, 3) Δ AIC$_c$ <4 model average (MA), 4) and Δ AIC$_c$ <4 variable importance (VI).

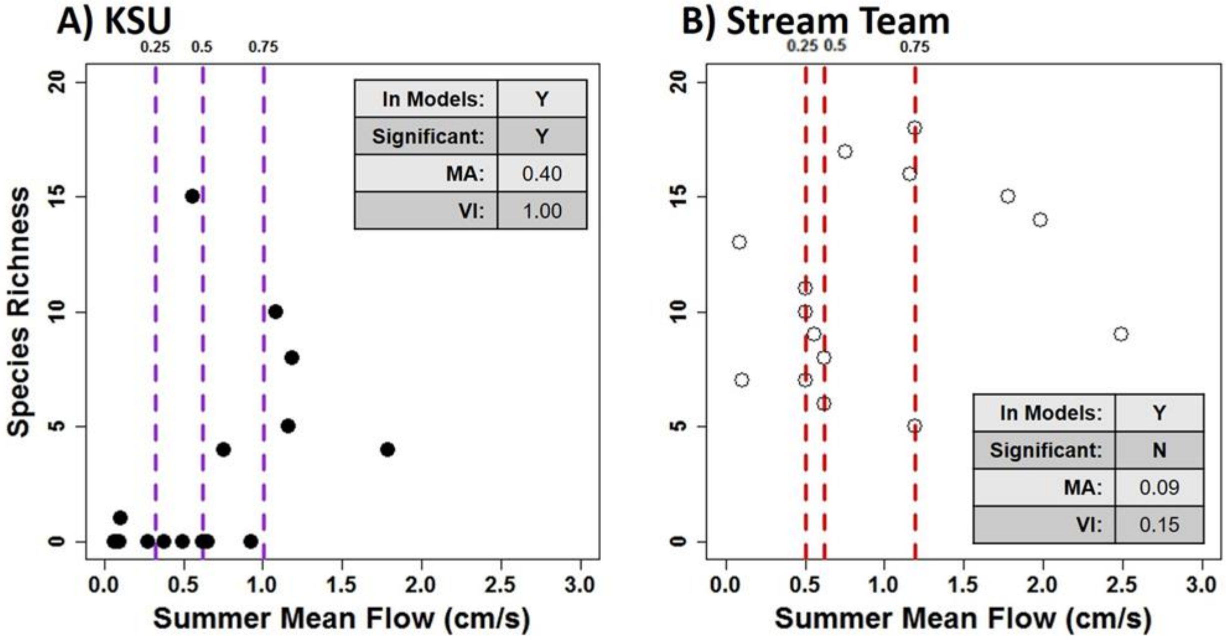


Figure 24. Summer mean flow (cm/s) vs. species richness plots for A) KSU Upper region dataset, and B) Stream Team Upper region dataset. 0.25, 0.5, and 0.75 quantiles for the explanatory variables are marked indicated by vertical dashed lines. A box on each plot indicates 1) whether this variable was in top models Δ AIC$_c$ <4, 2) if the variable was significant in any top models, 3) Δ AIC$_c$ <4 model average (MA), 4) and Δ AIC$_c$ <4 variable importance (VI).

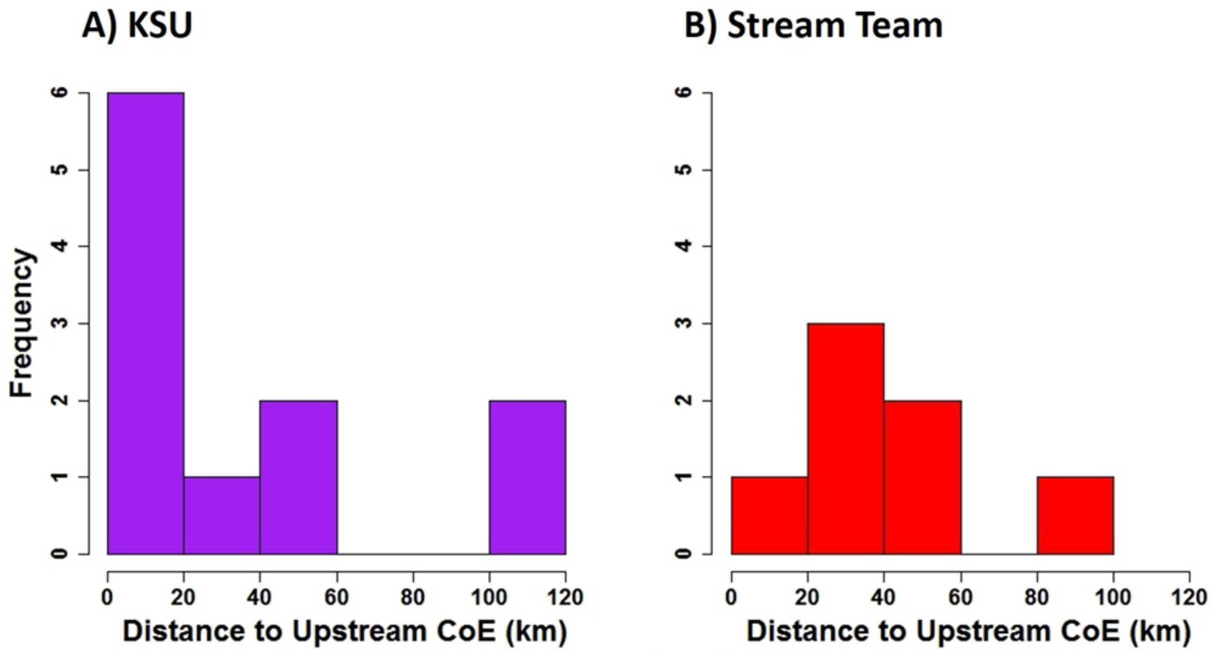


Figure 25. Histogram of Lower region distance to upstream CoE reservoir vs. frequency for A) KSU and B) Stream Team data.

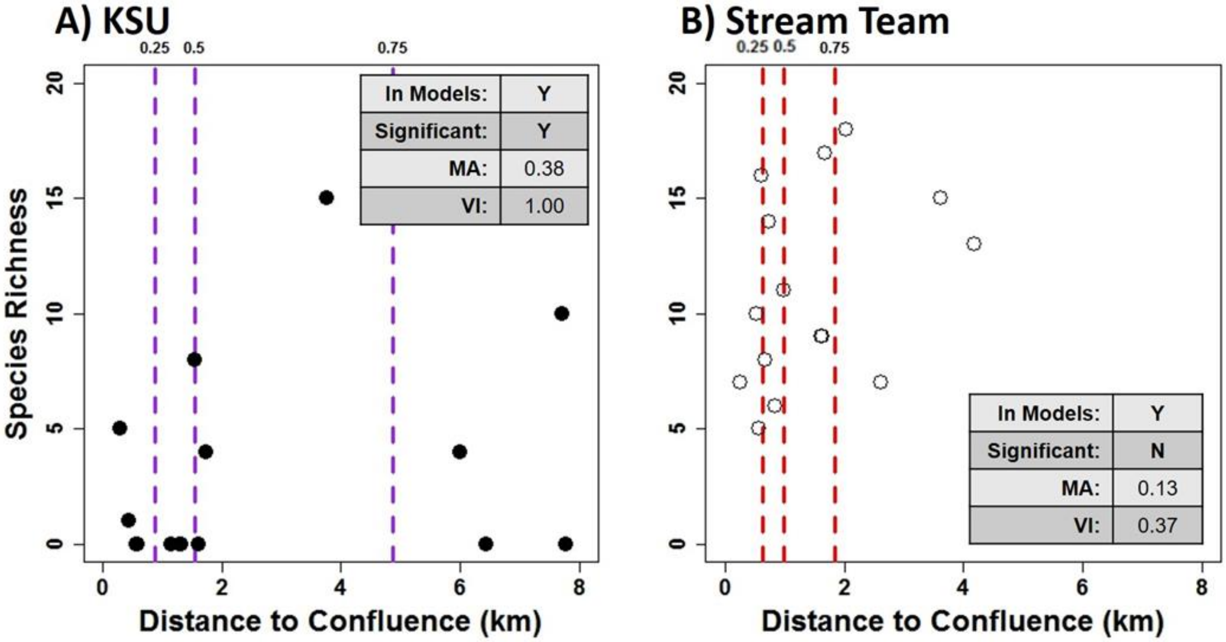


Figure 26. Distance to confluence (km) vs. species richness plots for A) KSU Upper region dataset, and B) Stream Team Upper region dataset. 0.25, 0.5, and 0.75 quantiles for the explanatory variables are marked indicated by vertical dashed lines. A box on each plot indicates 1) whether this variable was in top models Δ AIC$_c$ <4, 2) if the variable was significant in any top models, 3) Δ AIC$_c$ <4 model average (MA), 4) and Δ AIC$_c$ <4 variable importance (VI).

Table 1. KSU multiple linear regression Δ AIC$_c$ <4 model selection table for the full Smoky Hill River dataset. The dataset includes the six final *a priori* selected explanatory variables. The response variable was species richness at each sample site. Species richness was log$_{10}$ transformed. Coefficients (β) and standard errors (SE) are included for all explanatory variables in each model. Model averages (MA) and variable importance (VI) were summarized for each explanatory variable in models both Δ AIC$_c$<2 and Δ AIC$_c$<4. Model evaluation criteria include number of parameters (K), Δ AIC$_c$, Akaike weights (ω_i), model adjusted R^2 (Adj. R^2), and p-values. Variables were considered influential correlates if present and significant in models Δ AIC$_c$<4. Significant variables are bolded and underlined.

	Explanatory Variables						Evaluation Criteria					
	Developed Land (Proportion)	Herbaceous Grassland (Proportion)	Summer Mean Flow (cm s^{-1})	Distance to Upstream Dam (km)	Stream Type (Mainstem or Tributary)	Distance to Confluence (km)						
Model #	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	β_5 (SE)	β_6 (SE)	K	AIC$_c$	Δ AIC$_c$	ω_i	Adj. R^2	p-value
1	**-0.34 (.12)**		**0.41 (.14)**	0.27 (.14)	**-0.93 (.25)**		6	121.92	0.00	0.42	0.36	0.0001
2	**-0.27 (.12)**		**0.27 (.12)**		**-0.85 (.25)**		5	122.99	1.07	0.25	0.32	0.0002
3	**-0.27 (.12)**		**0.28 (.12)**		**-0.89 (.27)**	0.07 (.13)	6	125.29	3.37	0.08	0.31	0.0005
4	-0.25 (.13)	0.04 (.14)	**0.28 (.13)**		**-0.86 (.26)**		6	125.53	3.61	0.07	0.30	0.0005
5		0.29 (.14)	**0.49 (.16)**	0.31 (.16)	**-0.99 (.27)**		6	125.59	3.67	0.07	0.30	0.0005
6	**-0.27 (.12)**				**-0.98 (.26)**		4	125.65	3.73	0.06	0.26	0.0004
7			**0.27 (.13)**		**-0.81 (.26)**		4	125.89	3.97	0.06	0.26	0.0005
<4 MA:	-0.30	0.16	0.35	0.27	-0.90	0.07						
<4 VI:	0.88	0.14	0.94	0.49	1.00	0.08						
<2 MA:	-0.31		0.36	0.27	-0.90							
<2 VI:	1.00		1.00	0.63	1.00							

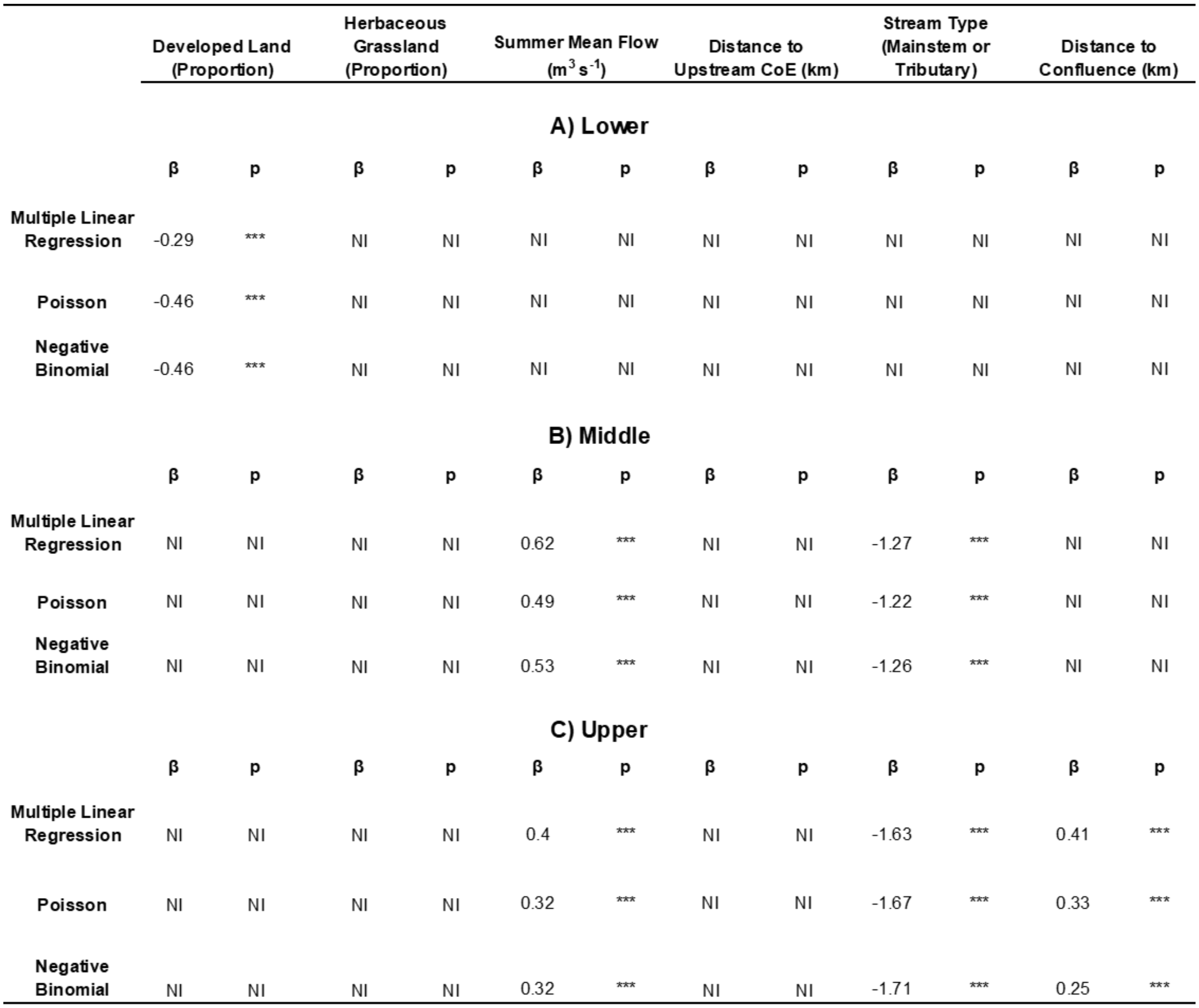
Table 2. KSU multiple linear regression AIC_c <4 model selection table for the A) Lower region, B) Middle region, and C) Upper regions of the Smoky Hill River watershed. The dataset includes the six final *a priori* selected explanatory variables (X_1-X_6). The response variable is species richness at each sample site. Species richness was log transformed. Standardized coefficients (β) and standard errors (SE) are included for all explanatory variables in each model. Model averages (MA) and variable importance (VI) are summarized for each explanatory variable in models both $\Delta\ \text{AIC}_c<2$ and $\Delta\ \text{AIC}_c<4$. Model evaluation criteria includes number of parameters (K), $\Delta\ \text{AIC}_c$, Akaike weights (ω_i), model adjusted R^2 (Adj. R^2), and *p*-values. Significant variables are bolded and underlined.

	Explanatory Variables						Evaluation Criteria					
					A) Lower							
	Developed Land (Proportion)	Herbaceous Grassland (Proportion)	Summer Mean Flow (cm s^{-1})	Distance to Upstream Dam (km)	Stream Type (Mainstem or Tributary)	Distance to Confluence (km)						
Model #	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	β_5 (SE)	β_6 (SE)	K	AIC_c	$\Delta\ \text{AIC}_c$	ω_i	Adj. R^2	p-value
1	**-0.29 (.09)**						3	11.38	0.00	0.66	0.47	0.01
2					**0.49 (.20)**		3	13.98	2.60	0.18	0.33	0.04
3	-0.22 (.10)				-0.22 (.20)		4	14.31	2.92	0.15	0.52	0.02
<4 MA:	-0.28				0.39							
<4 VI:	0.82				0.34							
					B) Middle							
Model #	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	β_5 (SE)	β_6 (SE)	K	AIC_c	$\Delta\ \text{AIC}_c$	ω_i	Adj. R^2	p-value
1			**0.62 (.11)**		**-1.27 (.24)**		4	35.51	0.00	0.37	0.72	0.00001
2			**0.45 (.15)**	0.24 (.15)	**-1.23 (.23)**		5	36.21	0.70	0.26	0.75	0.00001
3		0.18 (.13)	**0.43 (.15)**	0.35 (.17)	**-1.30 (.23)**		6	38.13	2.61	0.10	0.76	0.00003
4	-0.10 (.14)		**0.62 (.11)**		**-1.39 (.30)**		5	38.55	3.04	0.08	0.71	0.00003
5			**0.61 (.12)**		**-1.19 (.31)**	-0.06 (.15)	5	38.91	3.40	0.07	0.71	0.00004
6		0.04 (.13)	**0.63 (.12)**		**-1.29 (.26)**		5	38.99	3.48	0.06	0.71	0.00004
7	-0.13 (.13)		**0.44 (.15)**	0.26 (.15)	**-1.38 (.28)**		6	39.20	3.69	0.06	0.74	0.00004
<4 MA:	-0.11	0.12	0.54	0.27	-1.27	-0.06						
<4 VI:	0.14	0.16	1.00	0.42	1.00	0.07						
<2 MA:			0.55	0.24	-1.25							
<2 VI:			1.00	0.41	1.00							
					C) Upper							
Model #	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)	β_5 (SE)	β_6 (SE)	K	AIC_c	$\Delta\ \text{AIC}_c$	ω_i	Adj. R^2	p-value
1			**0.40 (.13)**		**-1.63 (.29)**	**0.41 (.14)**	5	36.52	0.00	0.50	0.75	0.0001
2		0.22 (.13)	**0.40 (.12)**		**-1.53 (.28)**	**0.33 (.13)**	6	37.80	1.28	0.26	0.78	0.0001
3	-0.19 (.13)		**0.44 (.13)**		**-1.50 (.29)**	**0.36 (.14)**	6	38.96	2.44	0.15	0.77	0.0002
4		**0.32 (.14)**	**0.46 (.14)**		**-1.30 (.31)**		5	39.83	3.31	0.10	0.70	0.0003
<4 MA:	-0.19	0.24	0.41		-1.56	0.38						
<4 VI:	0.15	0.36	1.00		1.00	0.90						
<2 MA:		0.22	0.40		-1.60	0.38						
<2 VI:		0.35	1.00		1.00	1.00						

Table 3. Full KSU dataset model comparison table with the results of the top model for each regression analysis performed (MLR, Poisson, and Negative-Binomial). Coefficients (β) and variable significance (p) are reported for each variable in each model. NI indicates that a variable is not included in the top model. Significance is indicated by three asterisks and non-significance is indicated by a period.

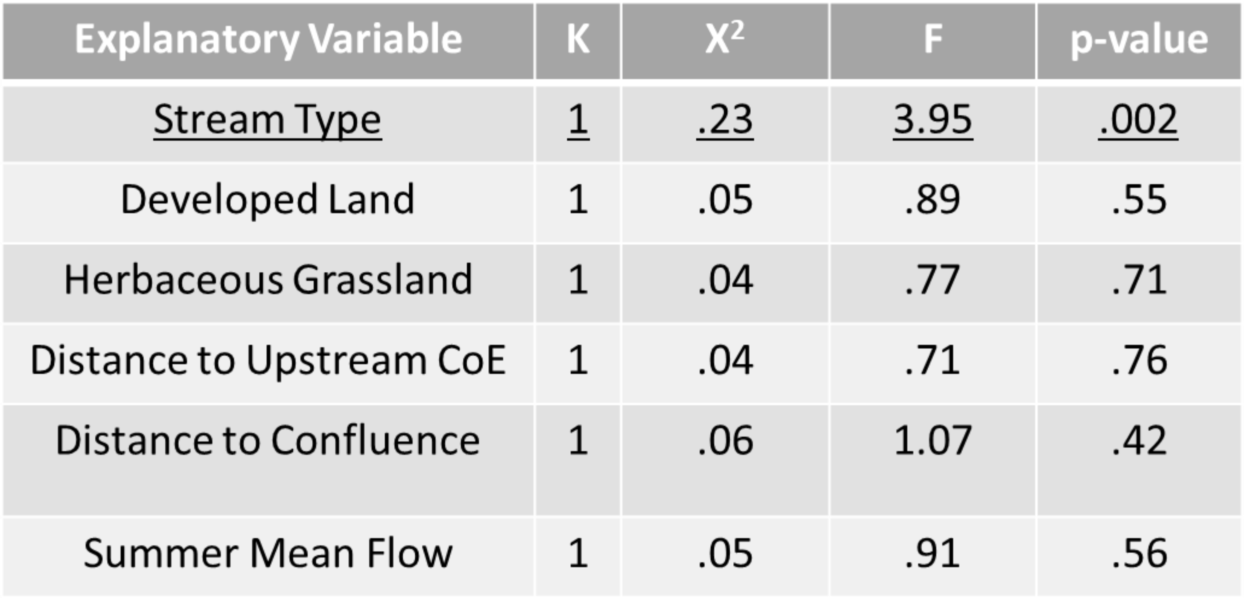
	Developed Land (Proportion)		Herbaceous Grassland (Proportion)		Summer Mean Flow $(m^3 s^{-1})$		Distance to Upstream CoE (km)		Stream Type (Mainstem or Tributary)		Distance to Confluence (km)	
	β	p	β	p	β	p	β	p	β	p	β	p
Multiple Linear Regression	-0.27	***			0.27	***			-0.85	***		
Poisson	-0.59	***			0.27	***	0.18	***	-0.81	***		
Negative Binomial	-0.64	***			0.28	.			-0.99	***		

Table 4. By region KSU dataset model comparison table with the results of the top model for each regression analysis performed (MLR, Poisson, and Negative-Binomial). Coefficients (β) and variable significance (p) are reported for each variable in each model. NI indicates that a variable is not included in the top model. Significance is indicated by three asterisks and non-significance is indicated by a period.



	Developed Land (Proportion)		Herbaceous Grassland (Proportion)		Summer Mean Flow ($m^3 s^{-1}$)		Distance to Upstream CoE (km)		Stream Type (Mainstem or Tributary)		Distance to Confluence (km)	
A) Lower												
	β	p	β	p	β	p	β	p	β	p	β	p
Multiple Linear Regression	-0.29	***	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
Poisson	-0.46	***	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
Negative Binomial	-0.46	***	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
B) Middle												
	β	p	β	p	β	p	β	p	β	p	β	p
Multiple Linear Regression	NI	NI	NI	NI	0.62	***	NI	NI	-1.27	***	NI	NI
Poisson	NI	NI	NI	NI	0.49	***	NI	NI	-1.22	***	NI	NI
Negative Binomial	NI	NI	NI	NI	0.53	***	NI	NI	-1.26	***	NI	NI
C) Upper												
	β	p	β	p	β	p	β	p	β	p	β	p
Multiple Linear Regression	NI	NI	NI	NI	0.4	***	NI	NI	-1.63	***	0.41	***
Poisson	NI	NI	NI	NI	0.32	***	NI	NI	-1.67	***	0.33	***
Negative Binomial	NI	NI	NI	NI	0.32	***	NI	NI	-1.71	***	0.25	***

Table 5. Canonical correspondence analysis diagnostic table for the full KSU dataset including all explanatory variables. Significant variables are indicated by an underline. Diagnostic criteria include the number of parameters (K), chi-square (X^2), f-value (F), and p-value.

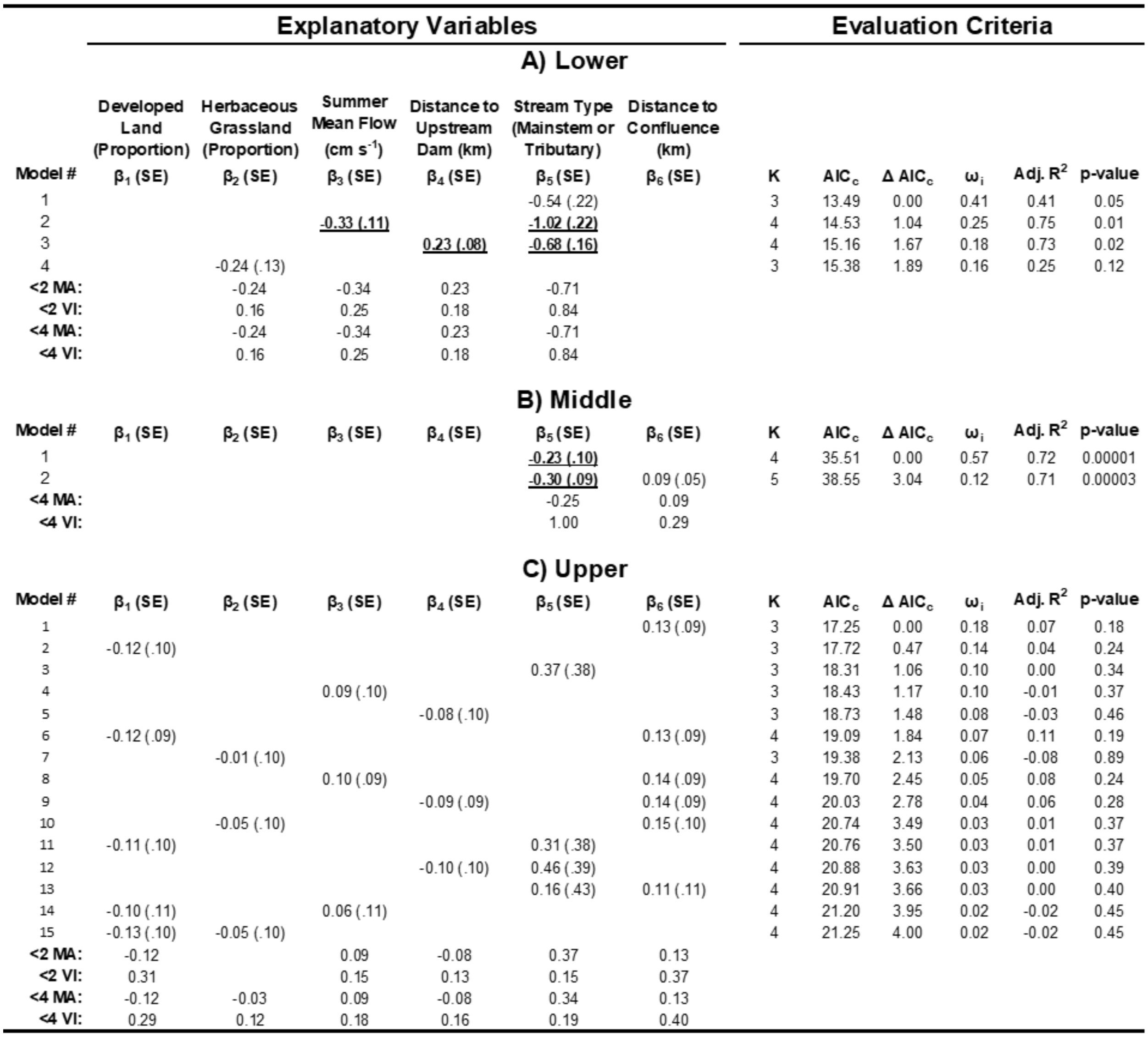


Explanatory Variable	K	X^2	F	p-value
Stream Type	1	.23	3.95	.002
Developed Land	1	.05	.89	.55
Herbaceous Grassland	1	.04	.77	.71
Distance to Upstream CoE	1	.04	.71	.76
Distance to Confluence	1	.06	1.07	.42
Summer Mean Flow	1	.05	.91	.56

Table 6. Stream Team multiple linear regression Δ AICc <4 model selection table for the full Smoky Hill River dataset. The dataset includes the six final *a priori* selected explanatory variables. The response variable was species richness at each sample site. Species richness was $\log_{10}$ transformed. Coefficients (β) and standard errors (SE) are included for all explanatory variables in each model. Model averages (MA) and variable importance (VI) were summarized for each explanatory variable in models both $\Delta \text{AIC}_c<2$ and $\Delta \text{AIC}_c<4$. Model evaluation criteria include number of parameters (K), ΔAIC_c, Akaike weights (ω_i), model adjusted R^2 (Adj. R^2), and p-values. Variables were considered influential correlates if present and significant in models $\Delta \text{AIC}_c<4$. Significant variables are bolded and underlined.

	Explanatory Variables						Evaluation Criteria					
Model #	Developed Land (Proportion) β_1 (SE)	Herbaceous Grassland (Proportion) β_2 (SE)	Summer Mean Flow (cm s^{-1}) β_3 (SE)	Distance to Upstream Dam (km) β_4 (SE)	Stream Type (Mainstem or Tributary) β_5 (SE)	Distance to Confluence (km) β_6 (SE)	K	AIC_c	ΔAIC_c	ω_i	Adj. R^2	p-value
1		-0.09 (.06)					3	25.16	0.00	0.13	0.04	0.14
2	-0.08 (.06)	-0.10 (.06)					4	25.67	0.50	0.10	0.07	0.13
3	-0.07 (.06)						3	26.00	0.83	0.08	0.02	0.23
4					-0.14 (.14)		3	26.47	1.30	0.07	0.00	0.32
5	-0.08 (.06)				-0.16 (.14)		4	27.10	1.94	0.05	0.03	0.25
6		-0.10 (.06)	-0.05 (.06)				4	27.13	1.97	0.05	0.03	0.25
7		-0.08 (.06)			-0.10 (.14)		4	27.22	2.06	0.05	0.02	0.26
8	-0.09 (.06)	-0.11 (.06)	-0.06 (.06)				5	27.42	2.26	0.04	0.07	0.18
9			-0.02 (.06)				3	27.48	2.32	0.04	-0.03	0.79
10		-0.10 (.06)		0.03 (.06)			4	27.54	2.38	0.04	0.01	0.30
11						0.01 (.06)	3	27.54	2.38	0.04	-0.03	0.92
12				0.01 (.06)			3	27.55	2.38	0.04	-0.03	0.92
13	-0.09 (.06)	-0.08 (.06)			-0.12 (.14)		5	27.64	2.48	0.04	0.06	0.19
14		-0.09 (.06)				0.01 (.06)	4	27.70	2.54	0.04	0.01	0.33
15	-0.09 (.06)	-0.11 (.06)		0.05 (.06)			5	27.78	2.61	0.03	0.06	0.20
16	-0.07 (.06)		-0.02 (.06)				4	28.44	3.28	0.02	-0.01	0.46
17	-0.08 (.06)	-0.10 (.06)				0.01 (.06)	5	28.45	3.29	0.02	0.04	0.26
18	-0.08 (.06)			0.02 (.06)			4	28.48	3.32	0.02	-0.01	0.47
19	-0.07 (.06)					-0.01 (.06)	4	28.60	3.43	0.02	-0.02	0.49
20					-0.16 (.15)	0.03 (.06)	4	28.87	3.70	0.02	-0.03	0.56
21		-0.13 (.07)	-0.10 (.08)			0.08 (.08)	5	28.87	3.70	0.02	0.02	0.30
22			-0.01 (.06)		-0.14 (.14)		4	29.00	3.84	0.02	-0.03	0.59
23				0.01 (.06)	-0.14 (.14)		4	29.06	3.89	0.02	-0.03	0.61
<2 MA:	-0.08	-0.10	-0.04	0.02	-0.13	0.02						
<2 VI:	0.44	0.55	0.19	0.16	0.25	0.16						
<4 MA:	-0.08	-0.09	-0.05		-0.15							
<4 VI:	0.49	0.58	0.10		0.24							

Table 7. Stream Team multiple linear regression Δ AICc <4 model selection table for the A) Lower region, B) Middle region, and C) Upper regions of the Smoky Hill River. The dataset includes the six final *a priori* selected explanatory variables. The response variable was species richness at each sample site. Species richness was $\log_{10}$ transformed. Coefficients (β) and standard errors (SE) are included for all explanatory variables in each model. Model averages (MA) and variable importance (VI) were summarized for each explanatory variable in models both Δ $AIC_c<2$ and Δ $AIC_c<4$. Model evaluation criteria include number of parameters (K), Δ AIC_c, Akaike weights (ω_i), model adjusted R^2 (Adj. R^2), and p-values. Variables were considered influential correlates if present and significant in models Δ $AIC_c<4$. Significant variables are bolded and underlined.



	Explanatory Variables						Evaluation Criteria					
	A) Lower											
	Developed Land (Proportion)	Herbaceous Grassland (Proportion)	Summer Mean Flow ($cm\ s^{-1}$)	Distance to Upstream Dam (km)	Stream Type (Mainstem or Tributary)	Distance to Confluence (km)						
Model #	**β_1 (SE)**	**β_2 (SE)**	**β_3 (SE)**	**β_4 (SE)**	**β_5 (SE)**	**β_6 (SE)**	**K**	**AIC_c**	**Δ AIC_c**	**ω_i**	**Adj. R^2**	**p-value**
1					-0.54 (.22)		3	13.49	0.00	0.41	0.41	0.05
2			**-0.33 (.11)**		**-1.02 (.22)**		4	14.53	1.04	0.25	0.75	0.01
3				**0.23 (.08)**	**-0.68 (.16)**		4	15.16	1.67	0.18	0.73	0.02
4		-0.24 (.13)					3	15.38	1.89	0.16	0.25	0.12
<2 MA:		-0.24	-0.34	0.23	-0.71							
<2 VI:		0.16	0.25	0.18	0.84							
<4 MA:		-0.24	-0.34	0.23	-0.71							
<4 VI:		0.16	0.25	0.18	0.84							
	B) Middle											
Model #	**β_1 (SE)**	**β_2 (SE)**	**β_3 (SE)**	**β_4 (SE)**	**β_5 (SE)**	**β_6 (SE)**	**K**	**AIC_c**	**Δ AIC_c**	**ω_i**	**Adj. R^2**	**p-value**
1					**-0.23 (.10)**		4	35.51	0.00	0.57	0.72	0.00001
2					**-0.30 (.09)**	0.09 (.05)	5	38.55	3.04	0.12	0.71	0.00003
<4 MA:					-0.25	0.09						
<4 VI:					1.00	0.29						
	C) Upper											
Model #	**β_1 (SE)**	**β_2 (SE)**	**β_3 (SE)**	**β_4 (SE)**	**β_5 (SE)**	**β_6 (SE)**	**K**	**AIC_c**	**Δ AIC_c**	**ω_i**	**Adj. R^2**	**p-value**
1						0.13 (.09)	3	17.25	0.00	0.18	0.07	0.18
2	-0.12 (.10)						3	17.72	0.47	0.14	0.04	0.24
3					0.37 (.38)		3	18.31	1.06	0.10	0.00	0.34
4			0.09 (.10)				3	18.43	1.17	0.10	-0.01	0.37
5				-0.08 (.10)			3	18.73	1.48	0.08	-0.03	0.46
6	-0.12 (.09)					0.13 (.09)	4	19.09	1.84	0.07	0.11	0.19
7		-0.01 (.10)					3	19.38	2.13	0.06	-0.08	0.89
8			0.10 (.09)			0.14 (.09)	4	19.70	2.45	0.05	0.08	0.24
9				-0.09 (.09)		0.14 (.09)	4	20.03	2.78	0.04	0.06	0.28
10		-0.05 (.10)				0.15 (.10)	4	20.74	3.49	0.03	0.01	0.37
11	-0.11 (.10)				0.31 (.38)		4	20.76	3.50	0.03	0.01	0.37
12				-0.10 (.10)	0.46 (.39)		4	20.88	3.63	0.03	0.00	0.39
13					0.16 (.43)	0.11 (.11)	4	20.91	3.66	0.03	0.00	0.40
14	-0.10 (.11)		0.06 (.11)				4	21.20	3.95	0.02	-0.02	0.45
15	-0.13 (.10)	-0.05 (.10)					4	21.25	4.00	0.02	-0.02	0.45
<2 MA:	-0.12		0.09	-0.08	0.37	0.13						
<2 VI:	0.31		0.15	0.13	0.15	0.37						
<4 MA:	-0.12	-0.03	0.09	-0.08	0.34	0.13						
<4 VI:	0.29	0.12	0.18	0.16	0.19	0.40						

References

Alexiades, A. V., & Fisher, W. L. (2015). Broad-scale habitat classification variables predict maximum local abundance for native but not non-native trout in New York streams. *Aquatic Conservation: Marine and Freshwater Ecosystems, 25*(1), 31-40.

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.

Belsky, A. J., Matzke, A., & Uselman, S. (1999). Survey of livestock influences on stream and riparian ecosystems in the western United States. *Journal of Soil and Water Conservation, 54*(1), 419-431.

Benda, L., Poff, N. L., Miller, D., Dunne, T., Reeves, G., Pess, G., & Pollock, M. (2004). The network dynamics hypothesis: How channel networks structure riverine habitats. *AIBS Bulletin, 54*(5), 413-427.

Bond, N. R., Balcombe, S. R., Crook, D. A., Marshall, J. C., Menke, N., & Lobegeiger, J. S. (2015). Fish population persistence in hydrologically variable landscapes. *Ecological Applications, 25*(4), 901-913.

Box, G. E. (1953). Non-normality and tests on variances. *Biometrics, 40*(3/4), 318-335.

Brabec, E., Schulte, S., & Richards, P. L. (2002). Impervious surfaces and water quality: A review of current literature and its implications for watershed planning. *CPL Bibliography, 16*(4), 499-514.

Branco, P., Segurado, P., Santos, J., Pinheiro, P., & Ferreira, M. (2012). Does longitudinal connectivity loss affect the distribution of freshwater fish? *Ecological Engineering, 48*, 70-78.

Braun, D. C., & Reynolds, J. D. (2014). Life history and environmental influences on population dynamics in sockeye salmon. *Canadian Journal of Fisheries and Aquatic Sciences, 71*(8), 1198-1208.

Burchsted, D., Daniels, M., & Wohl, E. E. (2014). *Introduction to the Special Issue on Discontinuity of Fluvial Systems.*

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261-304.

Callihan, R. A. (2013). *Distribution, Proliferation and Significance of Small Impoundments in Kansas.*

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications, 8*(3), 559-568.

Cote, D., Kehler, D. G., Bourne, C., & Wiersma, Y. F. (2009). A new measure of longitudinal connectivity for stream networks. *Landscape Ecology, 24*(1), 101-113.

Covich, A., Fritz, S., Lamb, P., Marzolf, R., Matthews, W., Poiani, K., . . . Winter, T. (1997). Potential effects of climate change on aquatic ecosystems of the Great Plains of North America. *Hydrological Processes, 11*(8), 993-1021.

Crook, D. A., Lowe, W. H., Allendorf, F. W., Erős, T., Finn, D. S., Gillanders, B. M., . . . Jennings, S. (2015). Human effects on ecological connectivity in aquatic ecosystems: Integrating scientific approaches to support management and mitigation. *Science of the Total Environment, 534,* 52-64.

Cross, F. B., & Collins, J. T. (1995). *Fishes in Kansas.* University of Kansas, Natural History Museum.

Crozier, L. G., Zabel, R. W., Hockersmith, E. E., & Achord, S. (2010). Interacting effects of density and temperature on body size in multiple populations of chinook salmon. *Journal of Animal Ecology, 79*(2), 342-349.

Cumming, G. S. (2004). The impact of low-head dams on fish species richness in Wisconsin, USA. *Ecological Applications, 14*(5), 1495-1506. doi:10.1890/03-5306

Datry, T., Larned, S. T., & Tockner, K. (2014). Intermittent rivers: A challenge for freshwater ecology. *Bioscience, 64*(3), 229-235.

Davies, B. R., & Walker, K. F. (2013). *The ecology of river systems.* Springer Science & Business Media.

Dodds, W. K., Gido, K., Whiles, M. R., Fritz, K. M., & Matthews, W. J. (2004). Life on the edge: The ecology of Great Plains prairie streams. *Bioscience, 54*(3), 205-216.

Dodds, W. K., & Oakes, R. M. (2006). Controls on nutrients across a prairie stream watershed: land use and riparian cover effects. *Environmental Management*, *37*(5), 634-646.

Dolédec, S., Phillips, N., Scarsbrook, M., Riley, R. H., & Townsend, C. R. (2006). Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. *Journal of the North American Benthological Society, 25*(1), 44-60.

Dugan, P. J., Barlow, C., Agostinho, A. A., Baran, E., Cada, G. F., Chen, D., ... & Marmulla, G. (2010). Fish migration, dams, and loss of ecosystem services in the Mekong basin. *AMBIO: A Journal of the Human Environment*, *39*(4), 344-348.

Dykstra, R. R. (1983). *The cattle towns.* U of Nebraska Press.

Espinoza, T., Marshall, S., & McDougall, A. (2013). Spawning of the endangered australian lungfish (neoceratodus forsteri) in a heavily regulated river: A pulse for life. *River Research and Applications*, *29*(10), 1215-1225.

Esselman, P., & Allan, J. (2010). Relative influences of catchment-and reach-scale abiotic factors on freshwater fish communities in rivers of northeastern Mesoamerica. *Ecology of Freshwater Fish, 19*(3), 439-454.

Falke, J. A., & Gido, K. B. (2006). Spatial effects of reservoirs on fish assemblages in Great Plains streams in Kansas, USA. *River Research and Applications, 22*(1), 55-68.

Fausch, K. D., Lyons, J., Karr, J. R., & Angermeier, P. L. (1990). Fish communities as indicators of environmental degradation. Paper presented at the *American Fisheries Society Symposium, 8,* 123-144.

Fencl, J. S., Mather, M. E., Costigan, K. H., & Daniels, M. D. (2015). How big of an effect do small dams have? using geomorphological footprints to quantify spatial impact of low-head dams and identify patterns of across-dam variation. *PloS One, 10*(11), e0141210.

Fischer, J. R., & Quist, M. C. (2014). Characterizing lentic freshwater fish assemblages using multiple sampling methods. *Environmental Monitoring and Assessment, 186*(7), 4461-4474.

Franssen, N. R., Gido, K. B., Guy, C. S., Tripe, J. A., Shrank, S. J., Strakosh, T. R., . . . Paukert, C. P. (2006). Effects of floods on fish assemblages in an intermittent prairie stream. *Freshwater Biology, 51*(11), 2072-2086.

Frimpong, E. A., & Angermeier, P. L. (2009). Fish traits: A database of ecological and life-history traits of freshwater fishes of the United States. *Fisheries, 34*(10), 487-495.

Gao, J., Sheshukov, A. Y., Yen, H., Kastens, J. H., & Peterson, D. L. (2017). Impacts of incorporating dominant crop rotation patterns as primary land use change on hydrologic model performance. *Agriculture, Ecosystems & Environment, 247*, 33-42.

González-Bergonzoni, I., Meerhoff, M., Davidson, T. A., Teixeira-de Mello, F., Baattrup-Pedersen, A., & Jeppesen, E. (2012). Meta-analysis shows a consistent and strong latitudinal pattern in fish omnivory across ecosystems. *Ecosystems, 15*(3), 492-503.

Gorman, O. T. (1986). Assemblage organization of stream fishes: The effect of rivers on adventitious streams. *The American Naturalist, 128*(4), 611-616.

Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters, 4*(4), 379-391.

Gunal, H., & Ransom, M. (2006). Clay illuviation and calcium carbonate accumulation along a precipitation gradient in Kansas. *Catena*, *68*(1), 59-69.

Hall, C. J., Jordaan, A., & Frisk, M. G. (2011). The historic influence of dams on diadromous fish habitat with a focus on river herring and hydrologic longitudinal connectivity. *Landscape Ecology, 26*(1), 95-107.

Hall, R. I., Leavitt, P. R., Quinlan, R., Dixit, A. S., & Smol, J. P. (1999). Effects of agriculture, urbanization, and climate on water quality in the northern Great Plains. *Limnology and Oceanography, 44*(3part2), 739-756.

Hammond, P. C. (1995). Conservation of biodiversity in native prairie communities in the United States. *Journal of the Kansas Entomological Society, 68*(1), 1-6.

He, Y., Wang, J., Lek-Ang, S., & Lek, S. (2010). Predicting assemblages and species richness of endemic fish in the upper Yangtze river. *Science of the Total Environment, 408*(19), 4211-4220.

Herzog, D. P., Ostendorf, D. E., Hrabik, R. A., & Barko, V. A. (2009). The mini-missouri trawl: A useful methodology for sampling small-bodied fishes in small and large river systems. *Journal of Freshwater Ecology, 24*(1), 103-108.

Hitchman, S. M., Mather, M. E., Smith, J. M., & Fencl, J. S. (2017). Identifying keystone habitats with a mosaic approach can improve biodiversity conservation in disturbed ecosystems. *Global Change Biology,*

Humphries, C. J., Williams, P. H., & Wright, R. V. (1995). Measuring biodiversity value for conservation. *Annual Review of Ecology and Systematics, 26*(1), 93-111.

Jungwirth, M., Moog, O., & Muhar, S. (1993). Effects of river bed restructuring on fish and benthos of a fifth order stream, Melk, Austria. *River Research and Applications, 8*(1-2), 195-204.

Jurajda, P. (1995). Effect of channelization and regulation on fish recruitment in a flood plain river. *Regulated Rivers: Research & Management, 10*(2-4), 207-215. doi:10.1002/rrr.3450100215

Kiffney, P. M., Greene, C. M., Hall, J., & Davies, J. (2006). Tributary streams create spatial discontinuities in habitat, biological productivity, and diversity in mainstem rivers. *Canadian Journal of Fisheries and Aquatic Sciences, 63*(11), 2518-2530.

Koleff, P., Gaston, K. J., & Lennon, J. J. (2003). Measuring beta diversity for presence–absence data. *Journal of Animal Ecology, 72*(3), 367-382.

Lammert, M., & Allan, J. D. (1999). Assessing biotic integrity of streams: Effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. *Environmental Management, 23*(2), 257-270.

Launchbaugh, J. L., & Owensby, C. E. (1978). Kansas rangelands. *Kansas Agricultural Experiment Station Bulletin, 622,* 56.

Letcher, B. H., Schueller, P., Bassar, R. D., Nislow, K. H., Coombs, J. A., Sakrejda, K., . . . O'donnell, M. J. (2015). Robust estimates of environmental effects on population vital rates: An integrated capture–recapture model of seasonal brook trout growth, survival and movement in a stream network. *Journal of Animal Ecology, 84*(2), 337-352.

Macedo, D. R., Hughes, R. M., Ligeiro, R., Ferreira, W. R., Castro, M. A., Junqueira, N. T., . . . Pompeu, P. S. (2014). The relative influence of catchment and site variables on fish and macroinvertebrate richness in cerrado biome streams. *Landscape Ecology, 29*(6), 1001-1016.

Matthews, W. J. (1988). North American prairie streams as systems for ecological study. *Journal of the North American Benthological Society, 7*(4), 387-409.

Matthews, W. J., Cashner, R. C., & Gelwick, F. P. (1988). Stability and persistence of fish faunas and assemblages in three Midwestern streams. *Copeia*, 945-955.

Matthews, W. J., Harvey, B. C., & Power, M. E. (1994). Spatial and temporal patterns in the fish assemblages of individual pools in a Midwestern stream (USA). *Environmental Biology of Fishes*, *39*(4), 381-397.

Matthews, W. J., & Robison, H. W. (1998). Influence of drainage connectivity, drainage area and regional species richness on fishes of the interior highlands in Arkansas. *The American Midland Naturalist*, *139*(1), 1-19.

Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks. *JAWRA Journal of the American Water Resources Association*, *43*(1), 86-103.

Olden, J. D., & Naiman, R. J. (2010). Incorporating thermal regimes into environmental flows assessments: modifying dam operations to restore freshwater ecosystem integrity. *Freshwater Biology*, *55*(1), 86-107.

Osborne, L. L., & Wiley, M. J. (1992). Influence of tributary spatial position on the structure of warmwater fish communities. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*(4), 671-681.

Peet, R. K. (1974). The measurement of species diversity. *Annual Review of Ecology and Systematics*, *5*(1), 285-307.

Perkin, J. S., & Gido, K. B. (2012). Fragmentation alters stream fish community structure in dendritic ecological networks. *Ecological Applications*, *22*(8), 2176-2187.

Perkin, J. S., Gido, K. B., Cooper, A. R., Turner, T. F., Osborne, M. J., Johnson, E. R., & Mayes, K. B. (2015). Fragmentation and dewatering transform Great Plains stream fish communities. *Ecological Monographs*, *85*(1), 73-92.

Phelps, Q. E., Tripp, S. J., Garvey, J. E., Herzog, D. P., Ostendorf, D. E., Ridings, J. W., . . . Hrabik, R. A. (2010). Habitat use during early life history infers recovery needs for shovelnose sturgeon and pallid sturgeon in the middle Mississippi river. *Transactions of the American Fisheries Society, 139*(4), 1060-1068.

Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., ... & Stromberg, J. C. (1997). The natural flow regime. *BioScience, 47*(11), 769-784.

Poole, G. C. (2002). Fluvial landscape ecology: Addressing uniqueness within the river discontinuum. *Freshwater Biology, 47*(4), 641-660.

Radinger, J., Wolter, C., & Kail, J. (2015). Spatial scaling of environmental variables improves species-habitat models of fishes in a small, sand-bed lowland river. *PloS One, 10*(11), e0142813.

Rechisky, E. L., Welch, D. W., Porter, A. D., Jacobs-Scott, M. C., & Winchell, P. M. (2013). Influence of multiple dam passage on survival of juvenile chinook salmon in the Columbia River estuary and coastal ocean. *Proceedings of the National Academy of Sciences*, *110*(17), 6883-6888.

Rice, S. P., Ferguson, R. I., & Hoey, T. B. (2006). Tributary control of physical heterogeneity and biological diversity at river confluences. *Canadian Journal of Fisheries and Aquatic Sciences, 63*(11), 2553-2566.

Richter, B. D., Baumgartner, J. V., Powell, J., & Braun, D. P. (1996). A method for assessing hydrologic alteration within ecosystems. *Conservation Biology*, *10*(4), 1163-1174.

Ridings, J. W., & Schell, S. R. (2010). Evaluation of a substrate-disturbing brail added to the mini-missouri trawl for capture of madtoms, sculpins, and darters. *Journal of Freshwater Ecology, 25*(3), 337-344.

Rypel, A. L. (2012). Meta-analysis of growth rates for a circumpolar fish, the northern pike (esox lucius), with emphasis on effects of continent, climate and latitude. *Ecology of Freshwater Fish*, *21*(4), 521-532.

Samson, F., & Knopf, F. (1994). Prairie conservation in North America. *Bioscience, 44*(6), 418-421.

Schick, R. S., & Lindley, S. T. (2007). Directed connectivity among fish populations in a riverine network. *Journal of Applied Ecology, 44*(6), 1116-1126.

Schiemer, F. (2000). Fish as indicators for the assessment of the ecological integrity of large rivers. *Hydrobiologia, 422*, 271-278.

Sowards, B. J., Pinkall, R. L., Fleming, W. L., Hofmeier, J. R., & Stark, W. J. (2016). Kansas freshwater mussel populations of the upper Saline and Smoky Hill rivers with emphasis on the status of the cylindrical papershell (anodontoides ferussacianus). *Transactions of the Kansas Academy of Science, 119*(3-4), 325-335.

Stoddard, J. L., Larsen, D. P., Hawkins, C. P., Johnson, R. K., & Norris, R. H. (2006). Setting expectations for the ecological condition of streams: The concept of reference condition. *Ecological Applications, 16*(4), 1267-1276.

Summerfelt, R. C. (1967). Fishes of the Smoky Hill river, Kansas. *Transactions of the Kansas Academy of Science (1903-), 70*(1), 102-139.

Thornbrugh, D. J., & Gido, K. B. (2009). Influence of spatial positioning within stream networks on fish assemblage structure in the Kansas river basin, USA. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(1), 143-156.

Tuomisto, H. (2010). A diversity of beta diversities: Straightening up a concept gone awry. Part 1: defining beta diversity as a function of alpha and gamma diversity. *Ecography, 33*(1), 2-22.

Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences, 37*(1), 130-137.

Vellend, M. (2001). Do commonly used indices of β-diversity measure species turnover? *Journal of Vegetation Science, 12*(4), 545-552.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., … Reidy Liermann, C. (2010). Global threats to human water security and river biodiversity. *Nature, 467*(7315), 555.

Welch, D. W., Rechisky, E. L., Melnychuk, M. C., Porter, A. D., Walters, C. J., Clements, S., ... & Schreck, C. (2008). Survival of migrating salmon smolts in large rivers with and without dams. *PLoS Biology*, *6*(10), e265.

Whittaker, R. H. (1977). Evolution of species diversity in land communities [birds and vascular plants]. *Evolutionary Biology.*

Winter, J., & Duthie, H. (1998). Effects of urbanization on water quality, periphyton and invertebrate communities in a southern Ontario stream. *Canadian Water Resources Journal*, *23*(3), 245-257.

Winter, T. C. (1998). Ground water and surface water: A single resource. DIANE Publishing Inc.

Xenopoulos, M. A., Lodge, D. M., Alcamo, J., Märker, M., Schulze, K., & Van Vuuren, D. P. (2005). Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology, 11*(10), 1557-1564.

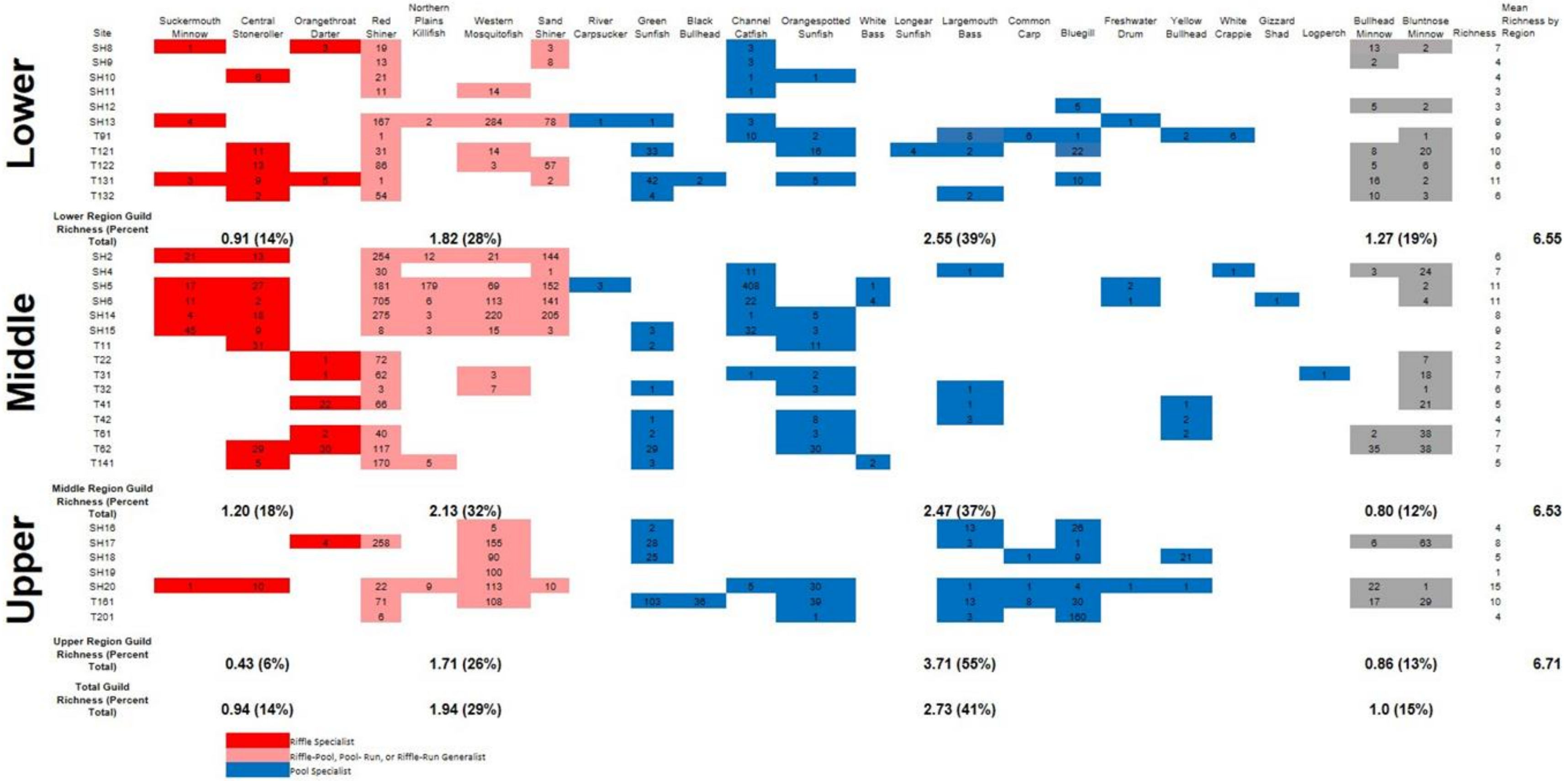
Appendix A- Gear Comparison

In a gear comparison in the adjacent Neosho River Basin, the mini-Missouri trawl (fished as described here) caught as many fish species as a beach seine and more species than backpack electrofishing or nets (Fencl 2015; Mean richness by gear type: Backpack Electrofishing- 4, Hoop Net- 6, Beach Seine- 6, Mini-Missouri Trawl- 7). The Mini-Missouri net, used as a hand trawl that was pulled along the bottom, was an effective gear in the most common mid-channel pool habitat. Use of this gear allowed me to collect comparable samples in all regions of the watershed, which was my research goal. Other advantages of the seine-like Mini-Missouri hand trawl were that (a) the width could be adjusted for different sized streams through operator position, (b) because of the gear construction, the net always stayed on the bottom regardless of operator skill, and the (c) double bag protected fish from rocks, and, as a consequence, improved post-release condition survival.

Appendix B- Fish Heat Maps

I constructed heat maps of fish abundance by species and functional guilds to provide a visual summary of the fish community data. . Fish were classified into six empirically-based functional habitat guilds (Hitchman et al. 2017): riffle specialists, pool specialists, two-habitat generalists (riffle-pool, pool-run, or riffle-run generalists), and generalists. For species not commonly caught in the Neosho River habitat sampling for which the above described empirical classifications were created, my guild classification used habitats described elsewhere (e.g., Frimpong and Angermeier 2009). In the heat maps, riffle specialists were colored red, pool specialists were colored blue, two-habitat generalists were colored pink, and generalists were colored gray. The number and identity of individual species and guilds for these heat maps, mean guild richness, and mean proportion guild richness were summarized by 1) river region (Upper, Middle, Lower: Table B.1) and 2) stream type (mainstem, tributary: Table B.2).

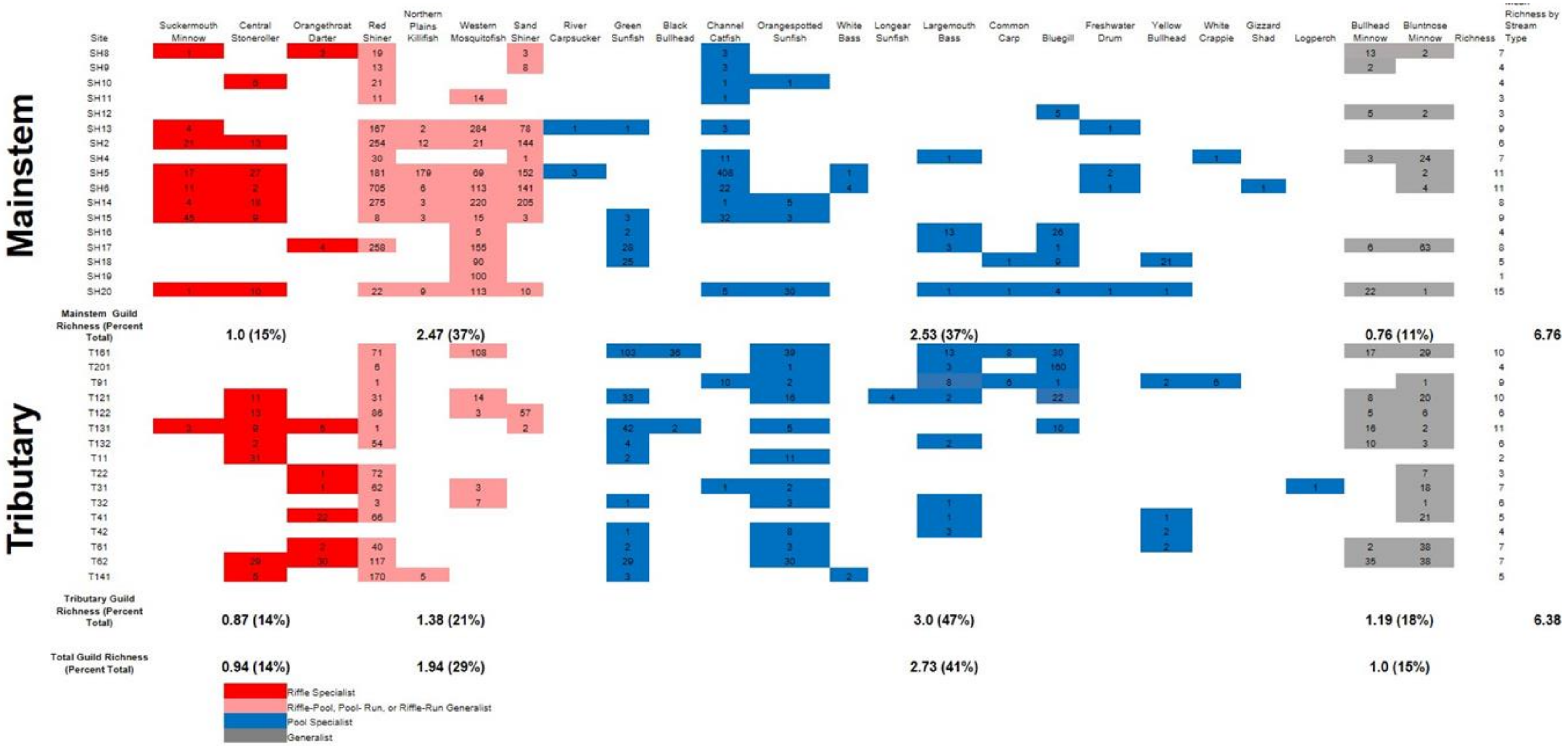
Table B.1. Heat map of fish abundance in the collected by KSU in the Smoky Hill River. In this heat map fish species are organized by river region. Mean total guild richness and percent total mean guild richness are summarized. Guilds include riffle specialists (red), dual-guild generalists (pink), pool specialists (blue), and generalists (grey).



	Site	Suckermouth Minnow	Central Stoneroller	Orangethroat Darter	Red Shiner	Northern Plains Killifish	Western Mosquitofish	Sand Shiner	River Carpsucker	Green Sunfish	Black Bullhead	Channel Catfish	Orangespotted Sunfish	White Bass	Longear Sunfish	Largemouth Bass	Common Carp	Bluegill	Freshwater Drum	Yellow Bullhead	White Crappie	Gizzard Shad	Logperch	Bullhead Minnow	Bluntnose Minnow	Richness	Mean Richness by Region
Lower	SH8	1		3	19			3				3												13	2	7	
	SH9				13			8				3												2		4	
	SH10		8		21							1	1													4	
	SH11				11		14					1														3	
	SH12																	5						5	2	3	
	SH13	4			167	2	284	78	1	1		3							1							9	
	T91				1							10	2			8	6	1		2	6				1	9	
	T121		11		31		14			33			16		4	2		22						8	20	10	
	T122		13		86		3	57																5	6	6	
	T131	3	9	6	1			2		42	2		5					10						16	2	11	
	T132		2		54					4						2								10	3	6	
	Lower Region Guild Richness (Percent Total)		0.91 (14%)				1.82 (28%)									2.55 (39%)								1.27 (19%)			6.55
Middle	SH2	21	13		254	12	21	144																		6	
	SH4				30			1				11				1					1			3	24	7	
	SH5	17	27		181	179	69	152	3			408		1					2						2	11	
	SH6	11	2		705	6	113	141				22		4					1			1			4	11	
	SH14	4	18		275	3	220	205				1	5													8	
	SH15	45	9		8	3	15	3		3		32	3													9	
	T11		31							2			11													2	
	T22			1	72																				7	3	
	T31			1	62		3					1	2										1		18	7	
	T32				3		7			1			3			1									1	6	
	T41			22	66											1				1					21	5	
	T42									1			8			3				2						4	
	T61			2	40					2			3							2				2	38	7	
	T62		29	30	117					29			30											35	38	7	
	T141		5		170	5				3				2												5	
	Middle Region Guild Richness (Percent Total)		1.20 (18%)				2.13 (32%)									2.47 (37%)								0.80 (12%)			6.53
Upper	SH16						5			2						13		26								4	
	SH17			4	258		155			28						3		1						6	63	8	
	SH18						90			25							1	9		21						5	
	SH19						100																			1	
	SH20	1	10		22	9	113	10				5	30			1	1	4	1	1				22	1	15	
	T161				71		108			103	38		39			13	8	30						17	29	10	
	T201				6								1			3		160								4	
	Upper Region Guild Richness (Percent Total)		0.43 (6%)				1.71 (26%)									3.71 (55%)								0.86 (13%)			6.71
	Total Guild Richness (Percent Total)		0.94 (14%)				1.94 (29%)									2.73 (41%)								1.0 (15%)			

🟥 Riffle Specialist
▨ Riffle-Pool, Pool- Run, or Riffle-Run Generalist
🟦 Pool Specialist

Table B.2. Heat map of fish abundance in the collected by KSU in the Smoky Hill River. In this heat map fish species are organized by stream type. Mean total guild richness and percent total mean guild richness are summarized. Guilds include riffle specialists (red), dual-guild generalists (pink), pool specialists (blue), and generalists (grey).



	Site	Suckermouth Minnow	Central Stoneroller	Orangethroat Darter	Red Shiner	Northern Plains Killifish	Western Mosquitofish	Sand Shiner	River Carpsucker	Green Sunfish	Black Bullhead	Channel Catfish	Orangespotted Sunfish	White Bass	Longear Sunfish	Largemouth Bass	Common Carp	Bluegill	Freshwater Drum	Yellow Bullhead	White Crappie	Gizzard Shad	Logperch	Bullhead Minnow	Bluntnose Minnow	Richness	Richness by Stream Type
Mainstem	SH8	1		2	19			3				3												13	2	7	
	SH9				13			8				3												2		4	
	SH10		6		21							1	1													4	
	SH11				11		14					1														3	
	SH12																	5						5	2	3	
	SH13	4			167	2	284	78	1	1		3							1							9	
	SH2	21	13		254	12	21	144																		6	
	SH4				30			1				11				1					1			3	24	7	
	SH5	17	27		181	179	69	152	3			408		1					2						2	11	
	SH6	11	2		705	6	113	141				22		4					1			1			4	11	
	SH14	4	18		275	3	220	205				1	5													8	
	SH15	45	9		8	3	15	3		3		32	3													9	
	SH16						5			2						13		26								4	
	SH17			4	258		155			28						3		1						6	63	8	
	SH18						90			25							1	9		21						5	
	SH19						100																			1	
	SH20	1	10		22	9	113	10				5	30			1	1	4	1	1				22	1	15	
	Mainstem Guild Richness (Percent Total)		1.0 (15%)			2.47 (37%)										2.53 (37%)								0.76 (11%)			6.76
Tributary	T161				71		108			103	36		39			13	8	30						17	29	10	
	T201				6								1			3		160								4	
	T91				1							10	2			8	6	1		2	6				1	9	
	T121		11		31		14			33			16		4	2		22						8	20	10	
	T122		13		86		3	57																5	6	6	
	T131	3	9	6	1			2		42	2		5					10						16	2	11	
	T132		2		54					4						2								10	3	6	
	T11		31							2			11													2	
	T22			1	72																				7	3	
	T31			1	62		3					1	2										1		18	7	
	T32				3		7			1			3			1									1	6	
	T41			22	66											1				1					21	5	
	T42									1			8			3				2						4	
	T61			2	40					2			3							2				2	38	7	
	T62		29	30	117					29			30											35	38	7	
	T141		5		170	5				3				2												5	
	Tributary Guild Richness (Percent Total)		0.87 (14%)			1.38 (21%)										3.0 (47%)								1.19 (18%)			6.38
	Total Guild Richness (Percent Total)		0.94 (14%)			1.94 (29%)										2.73 (41%)								1.0 (15%)			

🟥 Riffle Specialist
🟥 Riffle-Pool, Pool- Run, or Riffle-Run Generalist
🟦 Pool Specialist
⬜ Generalist

Appendix C- Nonindigenous Species

The USGS Nonindigenous Aquatic Species database uses GIS technology and historical samples to create a digital repository of nonindigenous species distributions in the United States. Using this database, I grouped 24 fish species collected during our 2015 fish sampling season into three groups: native, native transplant, and exotic (Table C.1). This was done by using a query specific to the Smoky Hill Basin, KS. 8 of 24 (33.3%) species collected in our samples were listed in database as either a native transplant or exotic to the Smoky Hill Basin (https://nas.er.usgs.gov).

Table C.1. Table including all species collected in 2015 KSU Smoky Hill samples. Species are grouped by native, native transplant, and exotic categories according to their classification in the USGS Nonindigenous Aquatic Species (NAS) database for the Smoky Hill Basin.

Native	Native Transplant	Exotic
Suckermouth Minnow	Bullhead Minnow	Common Carp
Central Stoneroller	Bluegill	
Red Shiner	Largemouth Bass	
Sand Shiner	White Crappie	
River Carpsucker	Western Mosquitofish	
Bluntnose Minnow	Gizzard Shad	
Green Sunfish	White Bass	
Orangespotted Sunfish		
Longear Sunfish		
Freshwater Drum		
Orangethroat Darter		
Logperch		
Northern Plains Killifish		
Channel Catfish		
Black Bullhead		
Yellow Bullhead		